



**DGK** Veröffentlichungen der DGK  
Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

---

Reihe C

Dissertationen

Heft Nr. 781

**Rainer Schützle**

**Entwicklung und Evaluierung  
eines formgestützten Location Referencing Verfahrens**

**München 2016**

---

Verlag der Bayerischen Akademie der Wissenschaften

ISSN 0065-5325

ISBN 978-3-7696-5193-5





**DGK** Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

---

Reihe C

Dissertationen

Heft Nr. 781

Entwicklung und Evaluierung  
eines formgestützten Location Referencing Verfahrens

Von der Fakultät Luft- und Raumfahrttechnik und Geodäsie  
der Universität Stuttgart  
zur Erlangung der Würde eines  
Doktors der Ingenieurwissenschaften (Dr.-Ing.)  
genehmigte Abhandlung

Vorgelegt von

**M.Sc. Rainer Schützle**

aus Heilbronn

München 2016

---

Verlag der Bayerischen Akademie der Wissenschaften

**Adresse der DGK:**



Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK)

Alfons-Goppel-Straße 11 • D – 80 539 München

Telefon +49 – 89 – 23 031 1113 • Telefax +49 – 89 – 23 031 - 1283 / - 1100

e-mail post@dgk.badw.de • <http://www.dgk.badw.de>

Hauptberichter: Prof. Dr.-Ing. habil. Volker Schwieger

Mitberichter: Prof. Dr.-Ing. Liqiu Meng

Prof. Dr.-Ing. habil. Dieter Fritsch

Tag der mündlichen Prüfung: 06.06.2016

Diese Dissertation ist auf dem Server der DGK unter <http://dgk.badw.de/>  
sowie auf dem Server der Universität Stuttgart unter <http://elib.uni-stuttgart.de/>  
elektronisch publiziert

---

© 2016 Bayerische Akademie der Wissenschaften, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,  
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen.

*Für meinen Vater.*



# Inhaltsverzeichnis

<b>Abkürzungsverzeichnis</b>	<b>IX</b>
<b>Kurzfassung</b>	<b>XI</b>
<b>Abstract</b>	<b>XIII</b>
<b>1 Einführung</b>	<b>1</b>
1.1 Umfeld	1
1.2 Motivation und Zielsetzung	1
1.3 Begriffsdefinitionen	2
1.4 Aufbau der Arbeit	3
<b>2 Zuordnung von korrespondierenden Kartenelementen</b>	<b>5</b>
2.1 Problemstellung	5
2.2 Digitale Straßenkarten	5
2.2.1 Topologische Grundlagen	5
2.2.2 Repräsentation von Realweltobjekten	6
2.2.3 Kartenunterschiede	7
2.2.4 Kartenformate	8
2.3 Grundlagen der Zuordnungsverfahren	10
2.3.1 Ähnlichkeitsbegriff	10
2.3.2 Kardinalitäten der Zuordnungen	11
2.3.3 Zuordnung mittels Mustererkennung in Straßennetzwerken	12
2.3.4 Interoperabilität von Geoinformationssystemen	13
2.3.5 Interoperabilität beim Location Referencing	14
2.3.6 Einteilung Zuordnungsverfahren	15
2.4 Übersicht bestehender Arbeiten	15
2.4.1 Buffer Growing	15
2.4.2 Iterative Closest Point Algorithm	17
2.4.3 Mehrstufige Zuordnungsverfahren	18
2.4.4 Delimited-Stroke-oriented Matching Algorithm	18
2.4.5 Weitere Verfahren	20
2.4.6 Zusammenfassung	26
2.5 Zuordnung auf Schema-Ebene	27
2.6 Zuordnung auf Objekt-Ebene	28
2.6.1 Geometrische Zuordnungsmaße	28
2.6.2 Topologische Zuordnungsmaße	33
2.6.3 Gemischte geometrische und topologische Zuordnungsmaße	35
2.6.4 Thematische bzw. semantische Zuordnungsmaße	36
2.7 Bewertungsfunktionen	38
2.7.1 Kosten- vs. Leistungsfunktion	39
2.7.2 Gesamt-Ähnlichkeitsmaß	39
2.8 Zusammenfassung	40

<b>3</b>	<b>Location Referencing</b>	<b>41</b>
3.1	Statisches und dynamisches Location Referencing . . . . .	41
3.2	Anwendungen . . . . .	42
3.2.1	Digitale Straßenkarten in ITS Anwendungen . . . . .	42
3.2.2	Kartenupdate bei Service-Providern . . . . .	44
3.2.3	Elektronischer Horizont . . . . .	44
3.2.4	Automatisiertes Fahren . . . . .	46
3.3	Dynamische Location Referencing Verfahren . . . . .	47
3.3.1	ILOC . . . . .	47
3.3.2	SPOT . . . . .	48
3.3.3	GoodLane . . . . .	49
3.3.4	PivotPoint . . . . .	50
3.3.5	AGORA . . . . .	51
3.3.6	MEI-LIN . . . . .	53
3.3.7	TPEG2-ULR . . . . .	53
3.4	OpenLR . . . . .	54
3.4.1	Standard . . . . .	54
3.4.2	Implementierung . . . . .	57
3.4.3	Analyse . . . . .	63
<b>4</b>	<b>Automatische Zuordnung von Kartenelementen zwischen TomTom und OpenStreetMap</b>	<b>71</b>
4.1	Motivation . . . . .	71
4.2	Schema-Matching . . . . .	72
4.2.1	TomTom MultiNet . . . . .	73
4.2.2	OpenStreetMap . . . . .	73
4.2.3	Dynamisches Rating mit Schema-Matching Informationen . . . . .	76
4.3	Ermittlung der Decoding-Kandidaten . . . . .	79
4.3.1	Dynamischer Suchradius . . . . .	79
4.3.2	Bewertung und Vorauswahl der gefundenen Kanten . . . . .	81
4.3.3	Decoding ohne direkte Knoten-Zuordnung . . . . .	81
4.3.4	Zusammensetzen aller möglichen Routenverläufe . . . . .	82
4.4	Bewertung der LRP-Kandidaten . . . . .	85
4.4.1	Richtungswinkel . . . . .	85
4.4.2	Valenz der Knoten . . . . .	85
4.5	Bewertung der Gesamtlocation – Form-Matching . . . . .	86
4.5.1	Anzahl signifikanter Richtungsänderungen . . . . .	87
4.5.2	Kumulierte Richtungsänderungen . . . . .	88
4.6	Berechnung der Leistungsfunktion . . . . .	88
4.6.1	Bewertungsanteile je LRP . . . . .	89
4.6.2	Bewertung der Form für die gesamte Ortsreferenz . . . . .	89
4.6.3	Zusammenführung zur Gesamt-Leistungsfunktion . . . . .	90
4.7	Zuordnung mittels Frechét-Distanz . . . . .	90
4.7.1	Diskrete Frechét-Distanz . . . . .	91
4.7.2	Bewertungsmaß . . . . .	92
4.8	Zusammenfassung . . . . .	93
<b>5</b>	<b>Qualität von Geodaten</b>	<b>95</b>
5.1	Einführung . . . . .	95
5.2	Qualitätsmodelle für Geodaten . . . . .	97
5.2.1	Datenqualität in der Geoinformatik . . . . .	97
5.2.2	Qualitätsmodell für Verkehrsinformationen . . . . .	98
5.2.3	Zusammenfassung . . . . .	100

5.3	Qualitätsmodelle für Kartenzuordnungen . . . . .	101
5.3.1	Einführung . . . . .	101
5.3.2	Relevante Qualitätsmerkmale . . . . .	101
5.3.3	Übersicht existierende Ansätze . . . . .	102
5.3.4	Qualitätsmodell im Projekt ROSATTE . . . . .	104
5.3.5	Klassifikation von Kartenzuordnungen . . . . .	106
5.3.6	Beurteilung von Kartenzuordnungen . . . . .	108
5.3.7	Zusammenfassung . . . . .	110
<b>6</b>	<b>Evaluierungsgrundlagen zur Bewertung von Location Referencing Übertragungen</b>	<b>113</b>
6.1	Definition von SOLL-Ortsreferenzen . . . . .	113
6.1.1	Grundlagen . . . . .	113
6.1.2	Lösungsraum von Kartenzuordnungen . . . . .	114
6.1.3	Realisierung . . . . .	115
6.2	Qualitätsparameter auf Objektebene . . . . .	116
6.2.1	Vollständigkeit . . . . .	117
6.2.2	Geometrische Genauigkeit . . . . .	117
6.2.3	Geometrische Korrektheit . . . . .	118
6.2.4	Topologische Korrektheit . . . . .	119
6.2.5	Gesamt Betrachtung der Zuordnung . . . . .	120
6.3	Qualitätsparameter auf Datensatz-Ebene . . . . .	121
6.3.1	Vollständigkeit . . . . .	121
6.3.2	Geometrische Genauigkeit . . . . .	122
6.3.3	Geometrische Korrektheit . . . . .	122
6.3.4	Topologische Korrektheit . . . . .	123
6.3.5	Gesamt Betrachtung des Datensatzes . . . . .	123
<b>7</b>	<b>Testumgebung, Test Szenarien und Empirische Untersuchungen</b>	<b>125</b>
7.1	Testumgebung zur Analyse von Location Referencing Verfahren . . . . .	125
7.1.1	Batch-Prozessierung . . . . .	125
7.1.2	Anbindung von Kartendatenbanken . . . . .	125
7.1.3	Graphische Benutzeroberfläche . . . . .	125
7.1.4	Reporting . . . . .	126
7.1.5	Visualisierung der Ergebnisse . . . . .	127
7.2	Praktische Durchführung der Untersuchungen . . . . .	128
7.2.1	Verwendete Kartendaten . . . . .	128
7.2.2	Verwendete Testdatensätze / Testszenarien . . . . .	129
7.3	Ergebnisse . . . . .	130
7.3.1	Szenario: Autobahn . . . . .	131
7.3.2	Szenario: Inter-Urban . . . . .	134
7.3.3	Szenario: Urban . . . . .	135
7.3.4	Szenario: Kreuzungen . . . . .	138
7.4	Diskussion der Ergebnisse . . . . .	141
7.4.1	Performanz von OpenLR und Form-Matcher im Vergleich . . . . .	141
7.4.2	Abhängigkeit von der Übertragungsrichtung . . . . .	144
7.4.3	Untersuchung Codierung und Decodierung auf einer Karte . . . . .	145
7.4.4	Signifikanz der OpenLR Zuordnungs-Attribute . . . . .	146
7.4.5	Qualitätsermittlung ohne SOLL-Werte . . . . .	147
<b>8</b>	<b>Zusammenfassung und Ausblick</b>	<b>151</b>
	<b>Literaturverzeichnis</b>	<b>155</b>

<b>Anhang A: OpenStreetMap Datenprozessierung</b>	<b>167</b>
A.1 Datenausschnitt mit osmosis . . . . .	167
A.2 Routingfähiges Netzwerk erzeugen mit osm2po . . . . .	167
A.3 Erzeugen von Shaphefiles mit spezifizierten Attributinformationen mit ogr2ogr . . . . .	168
A.4 Zusammenführen von Geometrie und Attributinformationen . . . . .	168
<b>Anhang B: LocationViewer</b>	<b>169</b>
B.1 Umsetzung in ArcGIS Toolbox . . . . .	169
B.2 Eingangsdaten . . . . .	169
B.3 Programmablauf . . . . .	170
B.4 Ergebnis / Darstellung in ArcGIS . . . . .	171
<b>Danksagung</b>	<b>173</b>
<b>Lebenslauf</b>	<b>175</b>

---

# Abkürzungsverzeichnis

<b>ACC</b>	Adaptive Cruise Control
<b>ADAS</b>	Advanced Driver Assistance System
<b>DAB</b>	Digital Audio Broadcasting
<b>FOW</b>	Form of Way
<b>FRC</b>	Functional Road Class
<b>GDF</b>	Geographic Data Files
<b>GoodLane</b>	Geographical Objects: Dynamical Linking and Navigable Embedding
<b>GUI</b>	Graphical User Interface
<b>ILOC</b>	Intersection Location
<b>INSPIRE</b>	Infrastructure for Spatial Information in the European Community
<b>ITS</b>	Intelligent Transportation Systems
<b>LRP</b>	Location Reference Point
<b>OGC</b>	Open Geospatial Consortium
<b>RDS</b>	Radio Data System
<b>ROSATTE</b>	ROad Safety ATtribute exchange infrastructure in Europe
<b>RTTI</b>	Real Time Traffic Information
<b>SPOT</b>	Spatial Object Tag
<b>TISA</b>	Traveller Information Services Association
<b>TMC</b>	Traffic Message Channel
<b>TN-ITS</b>	Transport Network ITS Spatial Data Deployment Platform
<b>TPEG</b>	Transport Protocol Experts Group
<b>ULR</b>	Universal Location Referencing
<b>WGS84</b>	World Geodetic System 1984



## Kurzfassung

Mit der wachsenden Verfügbarkeit von leistungsfähigen und kostengünstigen mobilen Endgeräten sowie flächendeckenden Datenübertragungsnetzen entstehen immer mehr Anwendungen sowohl in Geräten für Endkunden als auch professionelle Systeme, beispielsweise im Automobil-Bereich, bei denen Informationen mit einer Ortsangabe, der sogenannten Georeferenz, versehen werden. Im geodätischen Sinne besteht eine solche Georeferenz aus einem Koordinatensatz in einem definierten Koordinatensystem und bezeichnet somit einen Ort auf der Erde in eindeutiger Weise. Für praktische Anwendungen ist es jedoch häufig hilfreich, wenn zu dieser koordinatenmäßigen Darstellung auch Bezüge zu Objekten der Realwelt, etwa Gebäuden oder Straßen, hergestellt werden. Ein Katalog solcher Abbildungen von Realweltobjekten samt ihrer Beschreibung kann allgemein als (digitale) Karte bezeichnet werden. Um nun Informationen, welche auf die in einer solchen Karte abgebildeten Objekte referenziert sind, zwischen verschiedenen Endgeräten oder Systemen austauschen zu können, ist eine standardisierte Verallgemeinerung der Georeferenzierung erforderlich. Dies wird üblicherweise als *Location Referencing* bezeichnet.

Location Referencing Verfahren wurden im Rahmen von Forschungsprojekten verschiedentlich entwickelt, implementiert und praktisch angewendet. Dabei wurde beispielsweise im Rahmen des EU-geförderten Forschungsprojektes ROSATTE deutlich, dass diese LR-Verfahren grundsätzlich für die Übertragung von kartenbasierten Informationen zwischen verschiedenen Kartensystemen geeignet sind. Für bestimmte Anwendungsfälle mit hohen Qualitätsanforderungen sind jedoch noch weitere Verbesserungen notwendig (SCHÜTZLE und FRANK 2010).

Eine Location Referencing Übertragung kann technisch als Zuordnungsproblem zwischen zwei digitalen Karten verallgemeinert werden. Dabei wird die auf der Ausgangskarte basierende Georeferenz des zu übertragenden Objektes verallgemeinert. Diese wird als Ortsreferenz bezeichnet. In der als Zielnetz bezeichneten Karte des empfangenden Systems muss zur Zuordnung dann die korrespondierende Ortsreferenz identifiziert werden. Daher werden zunächst Grundlagen der Zuordnungsverfahren aus der Geoinformatik sowie bestehende Arbeiten auf diesem Gebiet dargestellt. Des Weiteren wird zwischen der Zuordnung auf Schema- und Objekt-Ebene unterschieden, was im späteren Verlauf der Arbeit zum Verständnis von Zuordnungsproblemen beiträgt. Die Zusammenführung der vorgestellten Zuordnungsmaße für eine Zuordnungsentscheidung kann mittels Kosten- oder Leistungsfunktionen erfolgen.

Grundsätzlich lassen sich für das Location Referencing zwei Varianten, nämlich das *statische* und das *dynamische* Location Referencing unterscheiden. Aufgrund der starken Dynamik der möglichen LR-Anwendungen erscheint eine statische Referenzierung mit zentral verwalteten Ortstabellen für die heutige und zukünftige praktische Anwendung des Location Referencing wenig sinnvoll. Daher folgt nach der Vorstellung möglicher LR-Anwendungen eine Übersicht bestehender dynamischer LR-Verfahren. Aufgrund seiner schon weiten Verbreitung und freien Zugänglichkeit in Hinsicht auf Lizenz und Implementierung wird OpenLR als Grundlage für die Entwicklungen im Rahmen der vorliegenden Arbeit detailliert vorgestellt und einer detaillierten Analyse unterzogen. Dabei wird der entsprechende Standard und die verfügbare Referenz-Implementierung des Verfahrens untersucht und Verbesserungsmöglichkeiten identifiziert.

Auf der vorangegangenen LR-Analyse aufbauend, wird im Rahmen dieser Arbeit ein neues Location Referencing Verfahren, der sogenannte *Form-Matcher*, entwickelt. Das neue Verfahren baut auf OpenLR auf, verwendet jedoch zusätzliche formbasierte Zuordnungsmerkmale, z.B. die Anzahl signifikanter Richtungsänderungen einer linienhaften Ortsreferenz. Für die Suche der Zuordnungskandidaten

werden alle den vollständigen Verlauf der Ortsreferenz im Zielnetz abdeckenden Routen-Kandidaten identifiziert und dann einer individuellen Bewertung entsprechend der verwendeten Zuordnungsmaße unterzogen. Des Weiteren berücksichtigt das neue Verfahren insbesondere auch die topologischen Eigenschaften der Ortsreferenzen bei der Decodierung im Zielnetz.

Die Beschreibung der Qualität der Kartenzuordnungen und damit auch die Evaluierung der LR-Übertragung soll auf einschlägigen Qualitätsmodellierungen aufbauen. Daher werden zunächst allgemeine Ansätze zur Qualitätsbeschreibung von Geodaten aus der Geoinformatik und auch speziell dem Verkehrsbereich vorgestellt. Daneben werden existierende Ansätze zur Beschreibung der Qualität von Kartenzuordnungen, beispielsweise aus dem ROSATTE-Projekt sowie der häufig verwendete Precision-Recall-Ansatz vorgestellt und auf ihre Tauglichkeit zur Verwendung im Rahmen der vorliegenden Arbeit hin untersucht.

Auf diesen Grundlagen aufbauend, wird ein beschreibendes Qualitätsmodell sowie ein Verfahren zur Ermittlung der Qualität von Kartenzuordnungen für diese Arbeit entwickelt. Grundlage sind die Qualitätsmerkmale *Vollständigkeit*, *Genauigkeit* und *Korrektheit*. Von ihnen ausgehend werden Qualitätsparameter auf Objektebene abgeleitet und mittels Leistungsfunktion zusammengeführt. Zur globalen Beurteilung können diese lokalen Qualitätsparameter auch auf Datensatzebene zusammengeführt werden.

Schließlich wurde das neu entwickelte LR-Verfahren sowie die Qualitätsbeurteilung zur Durchführung praktischer Untersuchungen in der sogenannten LR-Testumgebung implementiert. Aus den verfügbaren digitalen Straßenkarten *TomTom MultiNet* und *OpenStreetMap* wurden Datensätze für vier unterschiedliche Szenarien *Autobahn*, *Landstraße*, *Innenstadt* und *Kreuzungen* erstellt. Die damit durchgeführten empirischen Untersuchungen zeigen eine Qualitätssteigerung der LR-Übertragungen mittels Form-Matcher gegenüber OpenLR bei einer Übermittlung von OpenStreetMap Karten nach TomTom MultiNet. Für diese Übertragungsrichtung konnte der Anteil korrekter Zuordnungen von 63 % auf 75 % gesteigert werden. In umgekehrter Richtung zeigt OpenLR mit einem Anteil an korrekten Zuordnungen von 71 % ein höheres Qualitätsniveau, das der Form-Matcher mit 69 % korrekten Zuordnungen nicht verbessern kann.

## Abstract

The growing availability of powerful and cost-efficient mobile devices as well as area-wide mobile communication networks lead to the development of an increasing number of applications for the consumers and also in the professional sector (e.g. the automotive industry). All these applications imply that a location, the so-called georeference, is assigned to the transmitted information. In a geodetic sense, a georeference consists of a set of coordinates in a defined coordinate system and marks a place on earth in a distinct way. For practical applications however, it is often helpful to link these coordinates to real-world objects, e.g. buildings or roads. A catalogue of such descriptions of real-world objects can be regarded as a (digital) map. In order to exchange information that is referenced to objects depicted in such a map, a standardized generalization of the georeference is essential. This is typically referred to as *Location Referencing*.

Location Referencing techniques have been developed, implemented, and applied in practice within a number of research projects. The EU-funded research project ROSATTE, for example, showed that these LR techniques are basically suitable for the transmission of map-based information between different map systems. Certain applications with high demands on quality, however, require further improvement of the existing LR methods (SCHÜTZLE und FRANK 2010).

A Location Referencing transmission can be generalized from a technical point of view as an allocation problem between two digital maps. In doing so, the georeference of the object to be transmitted from the source network is being generalized. The result thereof is denoted as *Location Reference*. For allocating this location reference, the corresponding location reference has to be identified in the target map of the receiving system. For this reason, the basic mapping techniques used in geoinformatics, as well as already existing publications in this field are firstly presented. Furthermore, a distinction is made between mapping at schema level and mapping at object level. This will lead to a better understanding of some of the matching problems mentioned in the course of this thesis. The presented matching parameters can be combined for a decision on allocation by cost functions or power functions.

Location Referencing can basically be classified in *static* and *dynamic* Location Referencing. Typical LR applications show high dynamics together with high requirements on up-to-dateness and the need of immediate adaption to real-world changes. For this reason, a static referencing with centrally administrated location tables seems to be less reasonable for current and future practical applications of Location Referencing. An overview of existing dynamic LR methods is given after the discussion of possible LR applications. Since it is already widely used and accessible with respect to its license and implementation, OpenLR will be presented and analyzed in detail as a basis for the developments in the present thesis. For this purpose, the corresponding OpenLR Whitepaper as the published standard and the reference implementation which is publicly available from the OpenLR website are investigated in detail and thus possibilities for their improvement are identified.

Based on the preceding analysis, a new Location Referencing method, the so-called *form-matcher*, has been developed in the present thesis. The new method builds on OpenLR, but uses additional, form-based matching parameters, e.g. the number of significant changes of direction of a linear location reference. Searching for possible matching candidates, the complete LR route is considered in the target network instead of evaluating the location reference points individually. Then these identified route candidates are investigated using the selected matching parameters. Furthermore, the new method considers in particular the topological properties of the location references for the decoding in the target network.

The quality specification of the map assignments and thus the evaluation of the LR-transmissions should be based on existing quality models. Therefore, general approaches to a quality specification in geoinformatics as well as in the traffic and transportation domain are presented at first. Besides, existing approaches to quality description and modelling, for example the one derived within the ROSATTE research project, or the widely-used precision and recall approach are presented and examined with respect to their suitability for the use in the present thesis.

Based on these quality principles, a new descriptive quality model for map assignments and a corresponding evaluation method have been developed in the present thesis. The quality model consists of the quality characteristics *completeness*, *accuracy* and *correctness*. Local quality parameters are derived at object level from these characteristics and merged by using a power function. For the global quality assessment, these local quality parameters can be merged at dataset level.

Finally, the newly developed LR method together with the quality assessment methodology were implemented in the so-called LR testbed to enable practical investigations. Four different datasets for the scenarios *motorway*, *inter-urban*, *urban* and *intersections* were created from the *TomTom MultiNet* and *OpenStreetMap* road maps. The empirical investigations show a quality improvement of the LR transmissions via form-matcher from OpenStreetMap to TomTom MultiNet compared to OpenLR. For this direction of transmission, the percentage of correct assignments could be raised from 63 % to 75 %. For the opposite direction, however, OpenLR shows a higher quality level with a portion of 71 % of correct assignments. This portion could not be improved by the form-matcher with its 69 % correctness.

# 1 Einführung

## 1.1 Umfeld

Im Bereich der intelligenten Verkehrssysteme (engl. Intelligent Transportation Systems, ITS) und der Fahrerassistenzsysteme (engl. Advanced Driver Assistance Systems, ADAS) ist in den zurückliegenden Jahren eine deutliche Entwicklung von autarken Systemen, die lediglich auf lokale Sensorik im Fahrzeug verfügen, hin zu kooperativen Ansätzen zu beobachten. Schwerpunkt ist dabei der Austausch von Informationen, die zuvor entweder von Fahrzeugen oder infrastrukturbasierten Systemen detektiert wurden.

Neben der allgemeinen technischen Entwicklung und der damit verbundenen steigenden Verfügbarkeit von mobilen Datendiensten, wird diese Entwicklung hauptsächlich von zwei Hauptfaktoren vorangetrieben, nämlich der Steigerung der Verkehrssicherheit und damit verbundenen Senkung der Opfer- und Verletztanzahlen und den immer weiter ansteigenden Komfort-Anforderungen der Kunden, u.a. auch zur Bewältigung der immer komplexer werdenden Verkehrssituationen.

Verschiedene Dienste-Anbieter verwenden unterschiedliche digitale Straßenkarten als Basiskarten. Darauf werden dann die eigentlichen Sachinformationen referenziert. Die Georeferenzierung besteht also lediglich aus dem Verweis auf ein oder mehrere Straßen-/Linien-Objekte in der Basiskarte. Sollen diese Objekte nun übertragen bzw. ausgetauscht werden, so muss eine für die empfangende Anwendung und die von ihr verwendete Karte verständliche Georeferenzierung bereitgestellt werden. Da aber in dynamischen Anwendungen der Sender nicht weiß, welche Karte (Art/Version) der Empfänger verwendet, muss eine Karten-unabhängige Georeferenzierung generiert werden. Erschwerend kommt hinzu, dass die unterschiedlichen verfügbaren Karten einen unterschiedlichen Erfassungsstand sowie natürlich auch individuelle Erfassungsfehler und Modellierungsunterschiede aufweisen.

Die zu übertragenden Sachinformationen können punkt-, linien- oder flächenhafte Ausprägungen haben. Flächenhafte Objekte können dabei durch Kreise, Rechtecke (sog. Bounding Boxes), frei definierte Polygone, o.ä. definiert sein. Im Rahmen dieser Arbeit soll jedoch insbesondere der Fall linienhafter Ortsreferenzen aufgrund des häufigen Auftretens in der Praxis verbunden mit der Komplexität der Referenzierung betrachtet werden.

## 1.2 Motivation und Zielsetzung

Location Referencing Verfahren wurden in der Praxis bereits implementiert und erprobt (beispielsweise im von der EU geförderten Forschungsprojekt ROSATTE). Dabei hat sich gezeigt, dass die LR-Verfahren grundsätzlich funktionieren, jedoch ist großes Potential zur weiteren Verbesserung besteht. Aufbauend auf diesen Erkenntnissen soll im Rahmen der vorliegenden Arbeit ein verfügbares und in der Praxis auch verbreitetes Location Referencing Verfahren zunächst analytisch und empirisch evaluiert werden. Dazu ist die Entwicklung eines geeigneten Qualitätsmodells und einer Auswerte- und Analysestrategie notwendig. Die daraus gewonnenen Kenntnisse sollen in die Entwicklung eines verbesserten Location Referencing Verfahrens münden. Dabei soll die praktische Umsetzbarkeit, insbesondere die zur Verfügung stehenden Daten betreffend, gewährleistet bleiben. Anregungen zur Weiterentwicklung

von Location Referencing Verfahren können insbesondere auch aus dem Bereich der Kartenzuordnungen, wie sie in der Geoinformatik und Kartographie verwendet werden, gewonnen werden. Diese Zuordnungsverfahren werden häufig für die Zuordnung ganzer Karten verwendet. Es existiert eine Vielzahl unterschiedlicher Ansätze und Methoden. Die Zuordnungsverfahren in der Geoinformatik und dem Location Referencing hatten bislang keine fachlichen Berührungspunkte, daher sollen die entsprechenden Verfahren in dieser Arbeit erstmals ausführlich einander gegenübergestellt werden.

Ziel ist die Entwicklung eines Location Referencing Verfahrens, welches auf den bestehenden und in der Praxis bereits angewendeten Ansätzen aufbaut. Die Performanz der Zuordnung soll im Vergleich zu den existierenden Verfahren durch eine Kombination mit den in der Geoinformatik gebräuchlichen Zuordnungsmethoden gesteigert werden.

Schwerpunktmäßig sollen dabei die Möglichkeiten der geometrisch/topologischen Zuordnung untersucht werden.

### 1.3 Begriffsdefinitionen

Im Rahmen der vorliegenden Arbeit wird die im Folgenden beschriebene Terminologie verwendet, sie basiert neben den in der Geoinformatik gebräuchlichen Bezeichnungen teilweise auf den in den Arbeiten von HACKELÖER et al. (2015) und ZHANG (2009) verwendeten Termini sowie den im OpenLR Whitepaper (TOMTOM 2012) geprägten Begriffen:

- **Straßennetz:** Das Netz einer digitalen Straßenkarte  $M^i$  besteht aus einer Menge von Knoten  $\mathbf{N}^i := \{N_1^i, \dots, N_{n_i}^i\}$ ,  $n_i \in \mathbb{N}$  und Kanten  $\mathbf{E}^i := \{E_1^i, \dots, E_{m_i}^i\}$ ,  $m_i \in \mathbb{N}$ , sodass  $M^i = (\mathbf{N}^i, \mathbf{E}^i)$  mit  $\mathbf{E}^i \subseteq (\mathbf{N}^i \times \mathbf{N}^i)$ .
- **Straßensegment (engl. Road Segment):** Ein Straßensegment besteht aus einer Kante oder einer Folge von Kanten und wird von echten Kreuzungspunkten (Valenz  $> 2$ ) begrenzt. Im Sonderfall einer Sackgasse endet die entsprechende Kante an einer Seite mit einem einwertigen Knoten.
- **Quellkarte / Quellnetz:** In der Quellkarte  $M^S = (\mathbf{N}^S, \mathbf{E}^S)$  werden die Ortsreferenzen definiert und vom *Encoder* in *Location Reference Codes* codiert, die dann in ein empfangendes System (z.B. Fahrerassistenzsystem) übertragen werden.
- **Zielkarte / Zielnetz:** Die Zielkarte  $M^T = (\mathbf{N}^T, \mathbf{E}^T)$  wird vom empfangenden System verwendet und ist dort Basis für alle ortsbasierten Anwendungen. In der Zielkarte müssen alle empfangenen Ortsreferenzen (Location Reference Codes) decodiert und verortet werden, damit sie von den jeweiligen Anwendungen im Zielsystem nutzbar sind.
- **Punkt (engl. Point, Shape Point):** Ein Punkt  $p_i$  wird durch seine Koordinaten in einem anwendungsspezifisch festzulegenden Referenzsystem, definiert. Im Bereich der digitalen Straßenkarten und deren Anwendungen hat sich die Verwendung von Koordinaten im *World Geodetic System 1984 (WGS84)* etabliert. Alle Koordinatenangaben in diesem Anwendungsgebiet werden ohne besondere Nennung immer in diesem System referenziert übertragen. Dann wird ein Punkt  $p_i = (L_i, B_i)$  durch seine geographische Länge  $L_i$  und seine geographische Breite  $B_i$  beschrieben.
- **Knoten (engl. Vertex, Node):** Ein Knoten  $N$  ist ein Punkt, der den Anfang oder das Ende einer Kante definiert.
- **Linie (engl. Line, Segment):** Eine Linie ist definiert durch ihren Start- und Endpunkt  $l = (p_s, p_e)$ .
- **Kante (engl. Edge):** Eine Kante besteht aus einer oder mehreren zusammenhängenden Linien  $\mathbf{E}^i := \{l_j^i, \dots, l_{n_i}^i\}$ ,  $n_i \in \mathbb{N}$ . Der Startpunkt der ersten Linie und der Endpunkt der letzten Linie bilden jeweils den Startknoten  $N_s$  bzw. Endknoten  $N_e$  der Kante.

- **Feature:** Nach der Definition des Open Geospatial Consortium (OGC) repräsentiert ein Feature ein räumliches Realweltobjekt mit seiner Geometrie und weiteren Eigenschaften (Attributen), wie zum Beispiel Straße, Fluss, Gebäude. Ein Straßenobjekt in einer digitalen Straßenkarte besteht i.d.R. aus einer Kante, welche durch zwei Knoten begrenzt wird. Es besitzt weitere Eigenschaften, wie z.B. den Straßennamen, die Straßenklasse etc.
- **Ortsreferenz:** Eine Ortsreferenz (engl.: Location)  $\mathbf{L}_j^i$  beschreibt einen Routenverlauf  $j$  auf dem zugrundeliegenden Straßennetz  $M^i$ . Ihre Geometrie wird daher durch eine oder mehrere zusammenhängende Kanten beschrieben. Beginnt oder endet der Referenzierungsbereich der Ortsreferenz nicht mit dem Start- bzw. Endpunkt der sie beschreibenden Kanten, können optional ein Startoffset  $O_s$  (in OpenLR auch positiver Offset) und ein Endoffset  $O_e$  (in OpenLR auch negativer Offset) in [m] angegeben werden, so dass  $\mathbf{L}_j^i = \{\mathbf{E}_j^i, \mathbf{E}_k^i, \dots, \mathbf{E}_x^i, O_s, O_e\}$ .
- **Decodierte Ortsreferenz:** Beim Decodieren einer Ortsreferenz auf der Zielkarte  $M^T$  wählt der Decoder die am wahrscheinlichsten mit der gesuchten Ortsreferenz übereinstimmende Ortsreferenz aus und erzeugt die decodierte Ortsreferenz  $\hat{\mathbf{L}}_j^T$ .
- **Location Reference Point:** Ein LRP ist der Grundbaustein für eine OpenLR Ortsreferenz. Er wird mindestens am Anfangs- und Endpunkt der Ortsreferenz gesetzt, sowie an Zwischenknoten, an denen die Ortsreferenz vom kürzesten Weg abweicht. Der LRP enthält verschiedene Parameter aus der unterliegenden Karte, entsprechend der OpenLR Spezifikation.
- **Location String:** Ein Location String definiert eine Ortsreferenz für ihre Prozessierung in der LR-Testumgebung. Er umfasst neben dem Ortsreferenz Typ (hier: Line) auch ihren Bezeichner (ID), den Start- und End-Offset in Metern und als Hauptbestandteil eine Liste mit Kanten-Bezeichnern (IDs) auf welche die Ortsreferenz referenziert ist.
- **Location String File:** Ein Location String File ist eine Textdatei, die einen oder in der Regel mehrere Location Strings enthält. Mit dieser Datei wird die Stapelverarbeitung mehrerer Ortsreferenzen in der LR-Testumgebung gesteuert.
- **Location Reference Code:** Ein Location Reference Code ist die in maschinenlesbarer Form encodierte Repräsentation einer OpenLR-Ortsreferenz.
- **Location File:** Ein Location File enthält eine Liste mit Location Reference Codes und dient zur Dokumentation eines Satzes von Ortsreferenzen.

## 1.4 Aufbau der Arbeit

Zunächst werden in Kapitel 2 die Grundlagen für die Zuordnung korrespondierender Kartenelemente aus dem Bereich der Geoinformatik dargelegt. Nach der Beschreibung der realen Welt in digitalen Karten werden die Möglichkeiten zur Modellierung von Ähnlichkeiten geographischer Kartenobjekte zur Zuordnung beschrieben. Nach einer Übersicht über bestehende Zuordnungsverfahren wird die Zuordnung von Kartenelementen auf Schema- und Objekt-Ebene besprochen. Die Zusammenführung verschiedener Zuordnungsmerkmale zur Entscheidungsfindung ist mittels Kosten- oder Leistungsfunktion möglich. Das Kapitel schließt mit einer detaillierteren Darstellung relevanter Zuordnungsverfahren.

Eine Einführung in das Location Referencing folgt in Kapitel 3. Nach einer Motivation mit aktuellen Projekten und Anwendungsmöglichkeiten folgt eine Übersicht über die Entwicklung dynamischer Location Referencing Verfahren. Schließlich wird die dem praktischen Teil der vorliegenden Arbeit zugrunde liegende OpenLR Implementierung näher beschrieben und analysiert.

Das Kapitel 4 beschreibt schließlich die im Rahmen der vorliegenden Arbeit entwickelte Location Referencing Verfahren. Ausgehend von der Schema-Zuordnung für die verwendeten Kartendatensätze wird

die gesamte Prozesskette für eine Location Referencing Zuordnung beschrieben. Diese gliedert sich in Ermittlung der Decoding-Kandidaten, deren Bewertung mittels formbeschreibender Merkmale sowie Entscheidungsfindung mittels Leistungsfunktion. Das Kapitel endet mit einem Ausblick auf weitere Verbesserungsmöglichkeiten für das Location Referencing.

Im darauffolgenden Kapitel 5 werden Grundlagen und verschiedene Möglichkeiten der Qualitätsmodellierung von Geodaten dargelegt. Ausgehend von diesen allgemeinen Ansätzen werden Qualitätsbeschreibungen und Klassifizierungsschemata für Kartenzuordnungen abgeleitet und vorgestellt.

Im Kapitel 6 werden die zuvor allgemein besprochenen Qualitätsgrundlagen auf die konkret für die im Rahmen der vorliegenden Arbeit durchzuführenden Location Referencing Übertragungen angewendet. Zunächst wird die Definition von Referenz-Lösungen sowie der Lösungsraum für Location Referencing Übertragungen diskutiert. Darauf folgend werden schließlich Qualitätsparameter auf Objekt- und Datensatz-Ebene definiert.

Kapitel 7 beschreibt dann die durchgeführten empirischen Untersuchungen. Zunächst wird die Testumgebung, die zur praktischen Durchführung der Location Referencing Übertragungen und Evaluation der verwendeten Verfahren entwickelt wurde, vorgestellt. Nach den verwendeten Kartenausschnitten und den darin definierten Testszenarien und -Datensätzen werden die Ergebnisse der Untersuchungen vorgestellt und diskutiert.

Die Arbeit schließt mit einer Zusammenfassung in Kapitel 8 die auch einen Ausblick auf die kommenden Herausforderungen und die weiteren Entwicklungsmöglichkeiten des Location Referencing enthält.

## 2 Zuordnung von korrespondierenden Kartenelementen

### 2.1 Problemstellung

Kernproblem des Location Referencing ist die Zuordnung von Objekten einer digitalen Straßenkarte zu einem oder mehreren Objekten einer anderen Karte, die dieselben Realweltobjekte repräsentieren. Diese Zuordnung findet jedoch nur mittelbar statt. Es werden lediglich Informationen über ein in der Realwelt existierendes Objekt von dessen Repräsentation in der Quellkarte auf das oder die ihm in der Zielkarte entsprechenden Objekte übertragen. Das Zuordnungsproblem stellt sich also in der Zielkarte: welche Objekte entsprechen hier dem oder denen aus der Quellkarte und können dann mit den zu übertragenen Informationen versehen werden?

Praktisch gesehen wird meist zunächst der Suchraum in der Zielkarte eingeschränkt. Dies kann beispielsweise anhand der durch Koordinaten bestimmten Position des gesuchten Objektes und eines Suchradius erfolgen. Daraus ergibt sich in der Zielkarte ein Satz von *Zuordnungskandidaten*. Aus dem Referenzobjekt der Quellkarte und jeweils einem Objekt aus dem Satz der Zuordnungskandidaten werden *Zuordnungspaare* gebildet. Für jedes Zuordnungspaar werden die beiden Objekte aus Quell- und Zielkarte dann anhand verschiedener Eigenschaften verglichen. Ziel ist die Identifikation einer eindeutigen (*bijektiven*) Zuordnung zwischen zwei *Ortsreferenzen* in den beiden Karten.

### 2.2 Digitale Straßenkarten

#### 2.2.1 Topologische Grundlagen

Eine digitale Karte ist ein planarer Graph bestehend aus Knoten, Kanten und Flächen mit topologischen Beziehungen zwischen den Elementen. In einer anderen, ebenso gebräuchlichen, Terminologie werden dieselben Elemente als 0-Zelle, 1-Zelle und 2-Zelle bezeichnet. Jedes Element des Graphen kann eine beliebige Anzahl an Attributen besitzen (DERETSKY und RODNY 1993).

Eine digitale Straßenkarte repräsentiert das reale Straßennetz mittels dieser Knoten und Kanten. Die Knoten bilden dabei die Endpunkte der Kanten und repräsentieren entweder Kreuzungen, in denen sich zwei oder mehr Straßen treffen oder Straßenenden, an denen eine Straße endet, ohne in eine andere überzugehen (z.B. am Ende einer Sackgasse). Die Kanten repräsentieren die eigentlichen linienhaften Straßenobjekte. Obwohl Kanten theoretisch eine Krümmung aufweisen können, werden gekrümmte Straßenverläufe in der Praxis durch Geradenstücke in digitalen Straßenkarten angenähert. Um eine (routingfähige) Verbindung zweier Kanten zu erhalten, müssen diese in einem gemeinsamen Knoten aufeinandertreffen. Sich schneidende oder überlagernde Kanten besitzen keine topologische Verbindung, wie etwa bei Brücken oder Tunnels. Kanten haben auch eine implizite Richtung, die durch die Reihenfolge der Anfangs-, End- und ggfs. Bruchpunkte definiert wird. Eine solche Kante wird üblicherweise durch eine polygonale Linie oder kurz Polylinie (engl. *Polygonal line*) dargestellt. Polylinien können neben ihren Anfangs- und Endpunkten auch eine beliebige Anzahl von Zwischen- oder Bruchpunkten (engl. *Shape points*) enthalten. Die Bruchpunkte teilen die Polylinie in eine Sequenz

von Liniensegmenten und bilden den geometrischen Verlauf der realen Straße ab bzw. nähern diesen an. Dadurch lassen sich auch längere Streckenverläufe durch ein einziges Linienelement in der Karte darstellen. Enthält eine Polylinie keine weiteren Bruchpunkte, so bildet sie in diesem Sonderfall eine normale Linie. Knoten im Straßennetzwerk haben üblicherweise drei oder mehr abgehende Kanten. Manche Kartenhersteller fügen jedoch auch sogenannte *bivalente*, d.h. zweiwertige Knoten an solchen Stellen in ihr Straßennetzwerk ein, an denen sich einzelne Werte sogenannter *segmatisierter Attribute* der betroffenen Kanten ändern. Eine solche Auftrennung der Kanten ist notwendig, wenn die verwendete Datenbank eine Segmatisierung der Attribute nicht unterstützt, wie beispielsweise bei den hier verwendeten Shape-Files der Fall. In diesen Fällen werden Attribute nur ein mal für jede Kante gespeichert. Deshalb kann eine Kante nicht zwei oder mehr verschiedene Werte für das selbe Attribut annehmen. (angeregt durch GABAY und DOYTSHER (2000), SAFRA et al. (2006), HACKELÖER et al. (2013))

## 2.2.2 Repräsentation von Realweltobjekten

Karten bilden, wie alle geographischen Datenbanken, Objekte oder Entitäten der realen Welt ab. Diese Realwelt ist grundsätzlich nicht unterteilt. Erst durch das Bedürfnis des Menschen, die Realität zu vereinfachen und beschreibbar zu machen, tritt eine Objektbildung auf (VOLZ 2006b).

In manchen Fällen wird ein Realweltobjekt (z.B. eine Straße) durch mehr als ein Objekt in der digitalen Straßenkarte repräsentiert, etwa wenn mehrere Fahrstreifen separat in der Karte erfasst sind. Neben seinen geometrischen und topologischen Eigenschaften kann ein Objekt in der digitalen Straßenkarte auch semantische Attribute besitzen. Diese beschreiben Eigenschaften der Realweltobjekte mit Hilfe von numerischen oder alphanumerischen Attributwerten. Abbildung 2.1 zeigt anschaulich wie durch Wahrnehmung einer realen Situation durch verschiedene Beobachter und deren jeweilige Weiterverarbeitung unterschiedliche Repräsentationen des selben Realweltobjektes entstehen können.

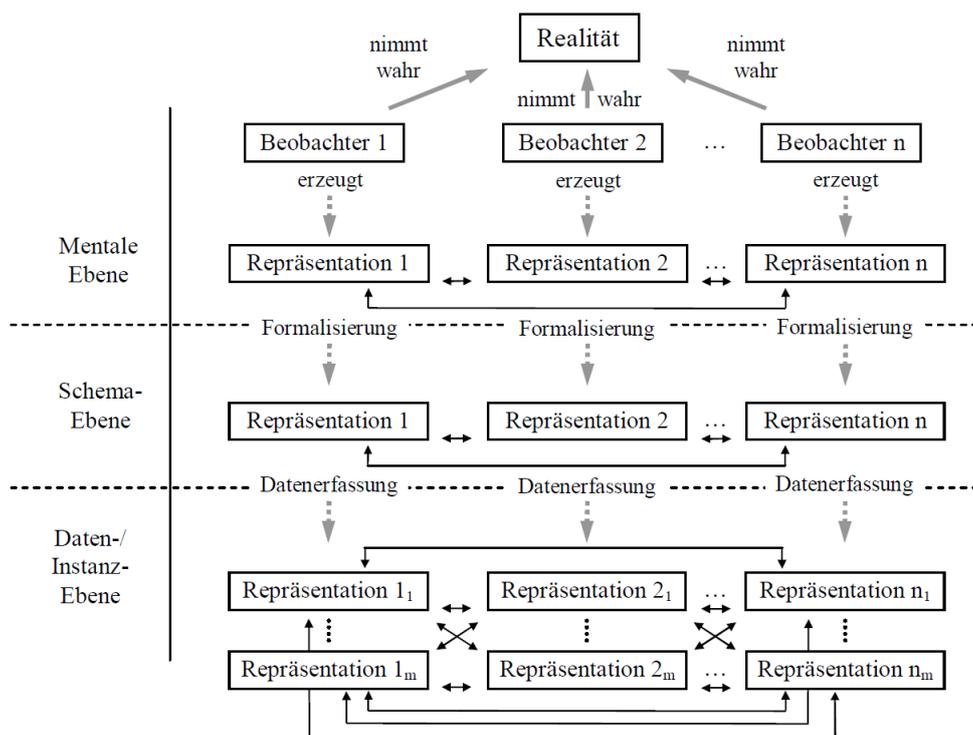


Abbildung 2.1: Ebenen der Repräsentation (aus VOLZ (2006b))

Die Zahl der Datensätze mit Raumbezug, die die gleichen Realwelt-Objekte repräsentieren, wächst stetig (OLTEANU-RAIMOND et al. 2015). Diese Datensätze werden durch verschiedene Organisationen mit unterschiedlichen Mitteln erstellt. Die darin gespeicherten Objekte werden oft unterschiedlich repräsentiert und auch für ganz unterschiedliche Anwendungen bereitgehalten. Daraus ergeben sich teilweise signifikante Unterschiede in den einzelnen Realwelt-Abbildungen der Datensätze bzw. geographischen Datenbanken.

Auch digitale Straßen-Karten von verschiedenen Quellen und Versionen, die das gleiche geografische Gebiet abbilden, weisen zum Teil deutliche Unterschiede in den Repräsentationen derselben Realweltobjekte auf. Gründe dafür sind u.a. Maßstab, Auflösung, Erfassungsstandards, Genauigkeit der Quellen (Datenerfassung), unterschiedliche Zeitpunkte der Datenerfassung oder Fehler bei der Erfassung und Bearbeitung der zugrundeliegenden Daten (WALTER und FRITSCH 1999, EDWARDS und SIMPSON 2002, MUSTIÈRE und DEVOGELE 2008).

### 2.2.3 Kartenunterschiede

Durch die zuvor dargestellten Unterschiede in der Repräsentation von Realweltobjekten entstehen unterschiedliche Darstellungen in den Karten. In BISHR (1998) werden drei Arten der Heterogenität von Daten unterschieden:

1. *Semantische Heterogenität*: Ein Realweltobjekt kann mehr als eine Beschreibung in den Datenbanken besitzen, entsprechend der verschiedenen Fachdisziplinen und ihren Anwendungen. Beispielsweise hat ein Straßennetzwerk, das für Instandhaltungszwecke genutzt wird, andere semantische Beschreibungen, als Daten, die explizit für Navigations-Anwendungen verwendet werden.
2. *Schematische Heterogenität*: Objekte in einer Datenbank werden in einer anderen Datenbank als Eigenschaften modelliert. Oder Objektklassen können unterschiedliche Aggregationen oder Generalisierungshierarchien besitzen, obwohl sie das selbe Realweltobjekt repräsentieren.
3. *Syntaktische Heterogenität*: Die Datenbanken können in unterschiedlichen Datenbank Management Systemen (DBMS) oder mit unterschiedlichen Paradigmen (z.B. objektorientiert oder relational) modelliert sein. Syntaktische Heterogenität ist auch auf die geometrische Repräsentation der Objekte bezogen, z.B. Raster- oder Vektordaten.

Bezogen auf die in den Datenbanken gespeicherten Objekte differenzieren HACKELÖER et al. (2013) drei Arten von Kartenunterschieden:

1. *Topologische Unterschiede*: Straßen können in einer Karte vorhanden sein, aber in der anderen nicht. Aufgrund unterschiedlicher Erfassung und auch Detailliertheit können Straßen in den beiden Karten topologisch unterschiedlich modelliert sein, d.h. mit einer unterschiedlichen Zahl von Knoten und Kanten dargestellt. Darüber hinaus sind komplexe Strukturen wie z.B. große Kreuzungen häufig unterschiedlich modelliert, was vor allem hier auch zu geometrischen Unterschieden in der Lage der einzelnen Kreuzungsknoten führen kann.
2. *Geometrische Unterschiede*: Aufgrund von Genauigkeitsunterschieden beim Erfassen der Geometrien, haben Knoten, die die selben Realweltobjekte repräsentieren, häufig abweichende Positionen in den Karten. Auf der anderen Seite muss eine relative geometrische Nähe zweier Objekte in den beiden Karten nicht auf eine Entsprechung zwischen ihnen hindeuten.
3. *Semantische Unterschiede*: Knoten und Kanten können semantische Informationen, wie beispielsweise Straßennamen tragen. Eine semantische Gleichheit zweier Objekte deutet darauf hin,

dass sie das gleiche Realweltobjekt referenzieren, muss ein festgestellter semantischer Unterschied nicht unbedingt bedeuten, dass zwei verschiedene Objekte betrachtet werden. Die semantischen Attribute unterscheiden sich dafür zu sehr sowohl im Inhalt (z.B. unterschiedliche Schreibweisen und mögliche Abkürzungen) wie auch in der Verfügbarkeit.

Aufgrund dieser Kartenunterschiede sind verschiedene Kardinalitäten bei der Zuordnung von Kartenobjekten möglich, siehe Kapitel 2.3.2.

## 2.2.4 Kartenformate

### 2.2.4.1 GDF - Geographic Data Files

Geographic Data Files (GDF) ist eine von der Automobilindustrie maßgeblich im Rahmen der Entwicklung von Fahrzeugnavigationssystemen eingeführte Spezifikation zum Austausch von digitalen Straßenkarten, die später international standardisiert und inzwischen in Version 5.0 veröffentlicht wurde (ISO 14825 2011). GDF wird hauptsächlich von kommerziellen Anbietern digitaler Straßenkarten (z.B. TomTom und HERE) verwendet.

Das konzeptionelle Datenmodell von GDF beschreibt alle relevanten Objekte der Realwelt (z.B. Straßen) durch *Features*, *Attribute* und *Relationen*. Im Mittelpunkt dieses Datenmodells stehen die *Features*; sie repräsentieren die geographischen Objekte der realen Welt. *Features* werden in drei verschiedenen Ebenen (engl. *Levels*) gebildet: *Level-0*, *Level-1* und *Level-2*. Eigenschaften von *Features* werden als *Attribute* bezeichnet. Beziehungen zwischen *Features* werden durch *Relationen* modelliert.

#### *Level-0: Geometrie*

Die topologischen Primitive Knoten, Kanten und Flächen bilden die Ebene 0 und beschreiben die Geometrie aller Kartenobjekte.

#### *Level-1: Einfache Features*

Auf dieser Ebene werden aus einem oder mehreren Geometrie-Objekten einfache *Features* gebildet, die durch ihre Zusammensetzung aus mehreren Grundelementen topologische Eigenschaften und auch *Attribute* besitzen können. *Simple Features* sind beispielsweise ein Straßenabschnitt ('*Road Element*') als *Linien-Feature* oder Kreuzungen ('*Junction*') als *Punkt-Features*. Alle *Features* werden mit Hilfe eines *Feature Class Codes* kategorisiert. Beispielsweise wird das *Feature Junction* mit Code '4120' oder das *Road Element* mit Code '4110' beschrieben.

#### *Level-2: Komplexe Features*

Komplexe *Features* können aus einem oder mehreren einfachen oder komplexen *Features* gebildet werden. Sie bekommen eigene *Attribute* die nur für dieses zusammengesetzte *Feature* gelten. Dies erlaubt die Modellierung z.B. von komplexen Kreuzungen ('*Intersection*', *Feature Code* 4145) oder Straßen ('*Road*', *Feature Code* 4140), aber auch die Aggregation beispielsweise von Städten zu Landkreisen usw.

### 2.2.4.2 NDS - Navigation Data Standard

Seit geraumer Zeit arbeitet die Fahrzeugindustrie zusammen mit Kartenherstellern und Serviceprovidern an einem standardisierten Kartenformat für Navigationssysteme. Dazu haben sie sich in einer Interessengemeinschaft *Navigation Data Standard (NDS) e.V.* zusammengeschlossen (NDS 2015). NDS ist ein Karten- und Endgerätehersteller unabhängiges physisches Speicherformat für vorkompilierte Navigationskarten. Die in diesem Format vorliegenden Daten können also direkt auf die in den Fahrzeugen

verbauten Endgeräte (z.B. Steuergeräte für Navigations- und Fahrerassistenzsysteme) aufgespielt werden. Das Datenformat enthält neben den klassischen Straßennetzen u.a. auch ein Geländemodell, Kreuzungsdarstellungen, detaillierte Stadtmodelle, Daten für Fahrerassistenzsysteme (ADAS) sowie Daten zur Sprachausgabe. Diese Daten wurden bislang von jedem Fahrzeughersteller für die individuellen Funktionsumfänge der einzelnen Fahrzeugmodelle zusammengestellt und für die verbauten Systeme des jeweiligen Zulieferers individuell kompiliert. Dies führte vor allem unter dem Aspekt der steigenden Anforderungen an die Aktualität und Vollständigkeit der Kartendatenbanken im Fahrzeug zu einem sehr hohen Aufwand, da jedes (in der Regel jährlich veröffentlichte) Kartenupdate vom Fahrzeughersteller für jedes seiner Fahrzeugmodelle unter Berücksichtigung des aktuellen Firmware-Stands und Funktionsumfangs individuell kompiliert werden musste (MÜLLER 2010). Mit NDS genügt nun die einmalige Kompilierung beispielsweise eines Kartenupdates, dieses kann dann in alle Fahrzeuge dieses Herstellers gleichermaßen verteilt werden. Durch den reduzierten Aufwand werden nun erheblich kürzere Update-Zyklen ermöglicht.

Aktuell konzentriert sich die Arbeit an NDS auf die Spezifikation von inkrementellen Updates zur Aktualisierung und Erweiterung der Kartendaten in den Fahrzeugdatenbanken. Gerade begonnen hat die Erweiterung in Bezug auf Connected Services und das hochautomatisierte Fahren (BEHRENS et al. 2015). Die NDS-Spezifikationen stehen ausschließlich den Mitgliedern des Vereins zur Verfügung.

In Bezug auf die im Rahmen der vorliegenden Arbeit betrachtete Übermittlung und Aktualisierung von Kartendaten in Echtzeit ist die Entwicklung von NDS von hoher Relevanz. Allerdings wird NDS hauptsächlich für die einheitliche Bereitstellung und Verbreitung von Aktualisierungen der Fahrzeugsysteme durch den Hersteller entwickelt und eingesetzt. Im Gegensatz dazu finden Location Referencing Verfahren für die systemübergreifende Referenzierung von Informationen und deren direkte Übertragung zwischen den Systemen ihre Anwendung. Von daher wird auf die weitere Betrachtung oder Verwendung von NDS im Rahmen dieser Arbeit verzichtet.

### 2.2.4.3 OpenStreetMap

OpenStreetMap wurde im Jahr 2004 in Großbritannien gegründet, um eine frei zugängliche Karte für das Vereinigte Königreich zu erstellen, denn die bis dato von der dortigen *Ordnance Survey* gepflegten Datenbestände waren nur einem beschränkten Nutzerkreis und auch nicht unentgeltlich zugänglich (OPENSTREETMAP 2015a). 2006 wurde dann die *OpenStreetMap Foundation* gegründet, um das Projekt auch auf andere Länder und Kontinente ausweiten zu können und möglichst allen Menschen freien Zugang zu geographischen Daten zu ermöglichen. Die Erfassung, Verwaltung und Fortführung der Daten erfolgte von Beginn an durch Freiwillige, die sich kostenfrei für das Projekt engagierten. Dazu wurden eigene Programme und Datenmodelle entwickelt. Teilweise stellten auch Organisationen ihre Datenbestände kostenfrei für das Projekt zur Verfügung.

Für diese Art der Erfassung und Fortführung raumbezogener Datensätze durch freie Communities wurde von GOODCHILD (2007) der Begriff *Volunteered Geographic Information (VGI)* geprägt.

Im OpenStreetMap Datenmodell werden Objekte durch *nodes* (Knoten), *ways* (Kanten) und *relations* (Gruppen von Objekten) dargestellt. Die OpenStreetMap Community zählt derzeit ca. 2,2 Millionen Mitglieder, die sich an der Weiterentwicklung und Aktualisierung der Karte beteiligen. Der *world* Datensatz umfasst derzeit etwa 3 Mrd. *nodes* und 300 Mio. *ways* (OPENSTREETMAP 2015d). Die Eigenschaften (oder auch Attribute) aller Kartenobjekte werden durch sogenannte 'tags' beschrieben. Diese OSM tags werden von OSM Mitgliedern mitsamt ihrer Definition und Nutzungsrichtlinien auf der Website des OSM Wiki vorgeschlagen (OPENSTREETMAP 2015c). Sie werden dort innerhalb der OSM Community diskutiert und auch weiterentwickelt.

Ein 'tag' setzt sich immer aus einer key/value Kombination zusammen. Der Schlüssel (engl. 'key') bezeichnet die jeweilige Eigenschaft. Dies können die Objektgruppe (z.B. railway, landuse) oder weitere

Attribute wie z.B. *name* oder *addr:housenumber* sein. Innerhalb einer Objektgruppe werden die einzelnen Objekttypen mit einem Wert (engl. 'value') benannt. So werden Straßen unter dem key *highway* zusammengefasst, während die einzelnen Straßenklassen als value geführt werden, z.B. *motorway* oder *primary*.

## 2.3 Grundlagen der Zuordnungsverfahren

### 2.3.1 Ähnlichkeitsbegriff

Nach OLTEANU-RAIMOND et al. (2015) beruht die Zuordnung von geographischen Daten auf dem Konzept der Ähnlichkeit. Zwei Objekte aus verschiedenen Karten werden einander zugeordnet, wenn sie sich ähnlich sind.

HOLT (1998) beschäftigt sich ganz allgemein mit dem Begriff der *Ähnlichkeit* und dann speziell mit der Ähnlichkeit von räumlichen Objekten und hat entsprechende Ansätze auf dem Gebiet der Kognitionspsychologie, der Philosophie sowie der Informationswissenschaft studiert. Er bedient sich dabei der menschlichen Wahrnehmung als Grundlage seiner Überlegungen und baut den Begriff der Ähnlichkeit auf dem intuitiven (visuellen) Erkennen von Ähnlichkeiten auf. Dabei zeigt er, dass das Menschliche Auge im Zusammenspiel mit dem Sehzentrum im Gehirn visuell wahrgenommene Objekte anhand von *Farbe*, *Form* und *Bewegung* verarbeitet und kategorisiert. Es klassifiziert dabei die Objekte, bildet Konzepte heraus und generalisiert und abstrahiert das Wahrgenommene. Dadurch können Ähnlichkeiten zwischen einem in der Vergangenheit wahrgenommenen Objekt und dem aktuell betrachteten erkannt werden.

Übertragen auf das Arbeiten mit Karten, oder speziell mit digitalen Straßenkarten, können wir feststellen, dass die menschliche Wahrnehmung sehr gut geeignet ist, intuitiv die nicht-übereinstimmenden Elemente beispielsweise zweier Kartenausschnitte desselben Gebiets zu erkennen. Im Anwendungsfall des Location Referencing, also der Suche nach entsprechenden Elementen in zwei oder mehreren digitalen Straßenkarten, kann das Auge sehr schnell erkennen, ob zwei zu vergleichende Elemente einander entsprechen oder nicht. COBB et al. (1998) nennen als Ziel für eine automatische Zuordnung von Kartenobjekten das Nachahmen der Fähigkeit der (menschlichen) Kartographen, die ausgebildet sind, Zuordnungsentscheidungen auf Grundlage sich widersprechender Informationen zu treffen.

Bei der Entwicklung von automatischen Zuordnungsverfahren gilt es also, genau dieses intuitive Erkennen von Ähnlichkeiten des menschlichen Auges mithilfe von Algorithmen nachzubilden. Dabei kann neben der zuvor genannten *Form* bzw. allgemeiner der *Geometrie* eines Objektes auch dessen *Topologie* und *Semantik* zur Ähnlichkeitsbestimmung herangezogen werden, wie in Kapitel 2.6 gezeigt wird.

WALTER (1997) beschreibt das Zuordnungsproblem in technischer Hinsicht wie folgt:

*„Hierzu wird ein Primitiv in einem der Datensätze identifiziert, um anschließend das entsprechende Primitiv im anderen Datensatz zu suchen und zuzuordnen. Dies bedeutet, dass aus der Menge der Primitive des zweiten Datensatzes dasjenige zugeordnet werden soll, welches dem Ausgangselement am ähnlichsten ist, sofern solch ein Primitiv überhaupt existiert. Um dies durchführen zu können, muss die Ähnlichkeit der Primitive mit Hilfe von Ähnlichkeitsmaßen berechnet werden.“*

### 2.3.2 Kardinalitäten der Zuordnungen

WALTER (1997) zeigt die verschiedenen Fälle, die bei einer Zuordnung von von linienhaften Objekten von einem in einen anderen Datensatz auftreten können. Danach kann im Idealfall ein Element aus Datensatz A genau einem Element des Datensatzes B zugeordnet werden (1:1). Aufgrund der unterschiedlichen Modellierungen der Kartendatenbanken ist es jedoch nicht unwahrscheinlich, dass ein Element aus Datensatz A mehreren Elementen aus Datensatz B zugeordnet wird oder dass andersherum mehrere Elemente aus Datensatz A genau einem Element aus Datensatz B zugeordnet werden müssen (1:n) oder (n:1). Daraus folgt, dass auch eine Kombination daraus möglich ist, also mehrere Elemente aus Datensatz A wiederum anderen Elementen aus Datensatz B zugeordnet werden (m:n). Abbildung 2.2 illustriert die genannten Fälle.

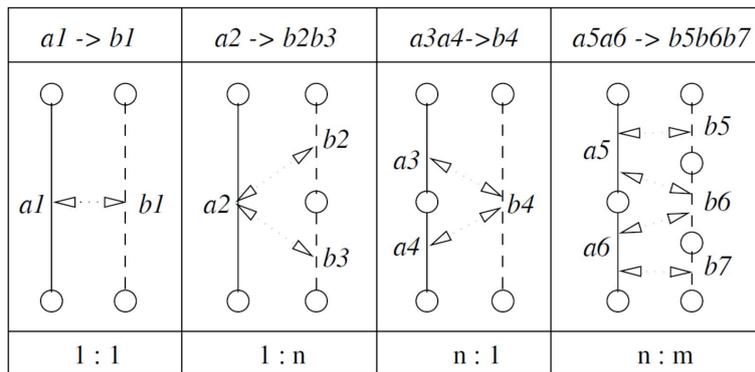


Abbildung 2.2: Kardinalitäten der Zuordnungen (aus WALTER (1997))

In CHEN (2011) werden die Kardinalitäten der Zuordnungen hauptsächlich in Bezug auf die Zuordnung von Kreuzungspunkten (Knoten) dargestellt. Dabei wird sehr schön die Schwierigkeit und teilweise auch Unmöglichkeit der Identifikation einer eindeutigen Zuordnung beschrieben, wie beispielhaft für drei Fälle in Abbildung 2.3 dargestellt.

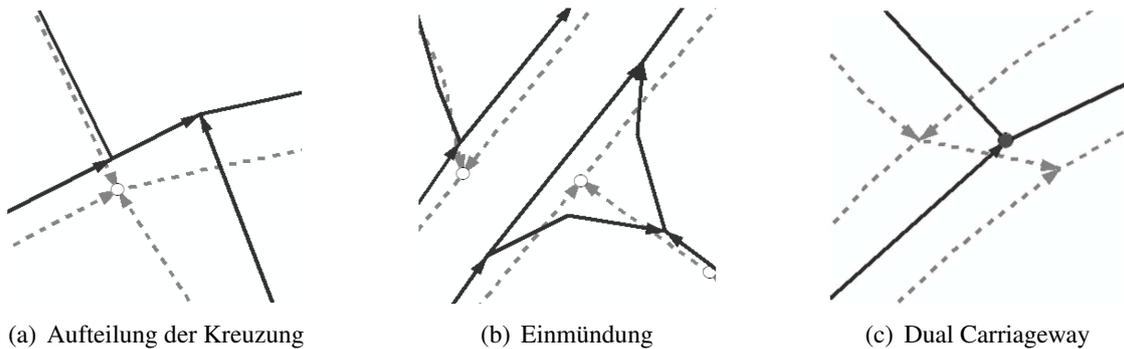


Abbildung 2.3: Kartenunterschiede (aus CHEN (2011))

Zur besseren Beschreibung real auftretender Zuordnungsfälle definiert ZHANG (2009) zwei weitere Arten von Zuordnungen, nämlich die *teilweise Zuordnung* und die *equivalente Zuordnung*. Die teilweise Zuordnung beschreibt eine erfolgte Zuordnung, bei der jedoch das zuzuordnende Objekt (die Kante) aufgrund unterschiedlicher Modellierung in den Karten nur unvollständig zugeordnet werden konnte. Dabei kann weiter zwischen *Erweiterung*, *Enthaltung* und *Verschiebung* unterschieden werden, siehe Abbildung 2.4(a). Die *equivalente Zuordnung* modelliert Fälle, in denen Objekte aus einer Karte durch mehrere gleichwertige Objekte der anderen Karte repräsentiert werden, beispielsweise bei getrennt

erfassten Fahrspuren (dual carriageway) oder einem Kreisverkehr, siehe Abbildung 2.4(b) und 2.4(c).

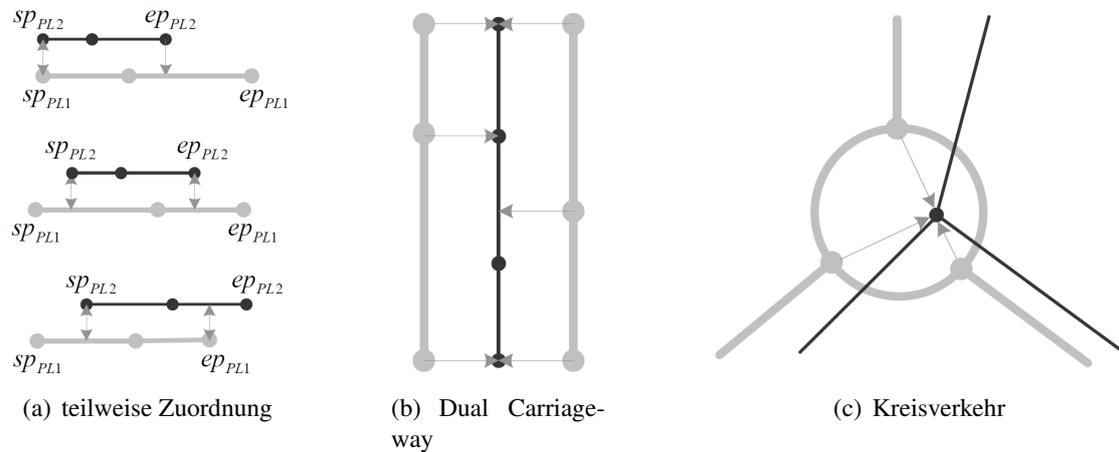


Abbildung 2.4: Teilweise und äquivalente Zuordnung (aus ZHANG (2009))

### 2.3.3 Zuordnung mittels Mustererkennung in Straßennetzwerken

Bei der Zuordnung von Straßennetzwerken können bestimmte Besonderheiten dieser Netzwerke zur Bestimmung von korrespondierenden Elementen herangezogen werden. ZHANG et al. (2010) definieren beispielsweise neben normalen, einspurigen Straßen fünf *Strukturen*, d.h. in Straßennetzwerken wiederkehrende Elemente mit hohem Wiedererkennungswert, die als Ausgangspunkt für eine Zuordnung der beiden Netzwerke dienen können:

- Kreisverkehre
- Mehrspurige Straßen
- Auffahrten an Kleeblatt-Kreuzungen
- Kurze Straßenstücke
- Kurze Sackgassen

Einen ähnlichen Ansatz wählt auch CHEN (2011), der für die formbasierte Zuordnung von Elementen einer Straßenkarte zwischen folgenden Form-Klassen unterscheidet, aus denen auch Kombinationen gebildet werden können:

- Einfache Straße
- Mehrspurige Straße
- Gabelung (Kreuzung)
- Ring
- Punkt

SESTER (1995) unterscheidet grundlegend verschiedene Knotentypen, die in Straßennetzwerken auftreten können. ZHANG et al. (2005) klassifizieren Kreuzungssituationen anhand der Valenz des Kreuzungsknotens und der richtungsmäßigen Verteilung der abgehenden Kanten, siehe Abbildung 2.5.

HEINZLE und ANDERS (2007) und ANDERS (2007) entwickelten ein Verfahren zur Mustererkennung in Straßennetzwerken. Sie unterscheiden dabei zwischen:

- langgestreckte Straßen (engl. strokes)
- Gitter
- Stern
- Ring

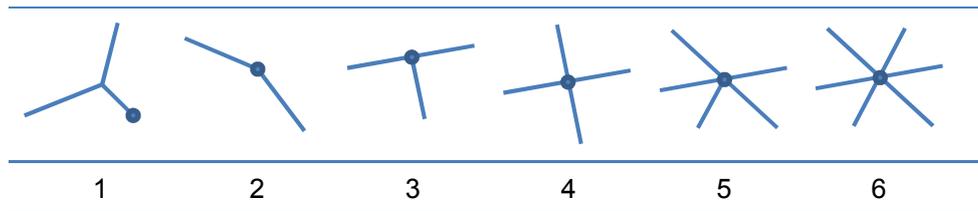


Abbildung 2.5: Klassifikation von Kreuzungen anhand Knoten-Valenz (nach ZHANG et al. (2005))

Da die einzelnen Objekte in digitalen Karten unterschiedlich modelliert und daher schlecht mit den Repräsentationen in einer anderen Karte vergleichbar sind, schlagen COBB et al. (1998) die Vereinigung von Objektteilen zu einem logischen Verbund (engl. composite) vor dem Feature Matching vor.

Zusätzlich zu den individuellen Eigenschaften der einzelnen Objekte einer digitalen Straßenkarte können auch die Beziehungen zwischen den Objekten für die Zuordnung verwendet werden. Das von ZHANG (2009) vorgeschlagene Verfahren beispielsweise ist kontextbasiert (vgl. Kapitel 2.4.4). Unter dem *Geographischen Kontext* versteht man in diesem Zusammenhang die räumlichen Beziehungen zwischen Objekten. Diese können durch Topologie, Entfernungen oder Richtungen beschrieben werden (SAMAL et al. 2004), siehe auch Kapitel 2.6.2. Darauf aufbauend schlagen KIM et al. (2010) zur Zuordnung von Objekten aus verschiedenen Datensätzen die Verwendung sogenannter Orientierungspunkte (engl. landmarks), die einen hohen Wiedererkennungswert im anderen Datensatz besitzen, vor. Übertragen auf die Zuordnung von Straßennetzwerken können beispielsweise Kreuzungen mit Straßen höherer Ordnung (evtl. auch Zusammengesetzt als komplexe Objekte) als solche 'landmarks' dienen.

### 2.3.4 Interoperabilität von Geoinformationssystemen

Die Geoinformatik und mit ihr der Bereich der Geographischen Informationssysteme (GIS) erfährt seit Jahren eine sehr dynamische Entwicklung, beschleunigt durch das schnelle Anwachsen sowohl der Rechenkapazitäten als auch der zu bearbeitenden und zu bewältigenden Menge an Informationen. Darüber hinaus gibt es ein zunehmendes Interesse, diese Informationen zwischen verschiedenen Nutzern weltweit zu teilen (RUIZ et al. 2011) bzw. Daten aus verschiedenen Quellen zu einem bestimmten Anwendungs- oder Auswertezweck zusammenzuführen. Dies wird auch unter dem Begriff *Interoperabilität* verstanden, nämlich der Fähigkeit eines Geo-Informationssystems, heterogene Daten (vgl. Kapitel 2.2.3) zu integrieren und gemeinsam verarbeiten zu können (VOLZ 2006b).

Weil heutzutage gleiche oder zumindest ähnliche geographische Daten (Geodaten) für das gleiche Gebiet häufig von mehr als einem Anbieter erfasst werden, sind die meisten Realwelt-Objekte in einer Vielzahl von geographischen Datenbanken erfasst. Man spricht hierbei auch von sogenannten *Mehrfachrepräsentationen* (VOLZ 2006b).

Die Zuordnung von Objekten eines Datensatzes zu den entsprechenden (homologen) Objekten anderer Datensätze, also die grundsätzliche Aufgabe des Location Referencing, wird auch im Bereich der Geoinformatik bearbeitet. Dort wird das Zuordnungsproblem aber häufig im Zusammenhang mit der Zusammenführung ganzer Datensätze und nur selten für die Zuordnung einzelner Objekte behandelt. Dies geschieht entweder zum Zwecke der Qualitätssteigerung (in Bezug auf Genauigkeit oder Zuverlässigkeit der Information) durch Nutzung von gleichartigen Informationen aus verschiedenen, redundanten Datenquellen oder zur Erhöhung des Informationsgehalts durch das Zusammenführen von unterschiedlichen Informationen aus verschiedenen Datenquellen, die das gleiche Gebiet abdecken. Deshalb ist die Zuordnung von Daten mit Raumbezug (engl. spatial Data Matching) in der Geoinformatik häufig auch unter den Begriffen der Datenintegration (engl. *Integration*) (z.B. SHEEREN et al. (2004), KAMPSHOFF (2005), OLTEANU et al. (2006), THAKKAR et al. (2007), BUTENUTH et al. (2007)) oder der Verschmelzung von Datensätzen (engl. *Conflation*) (z.B. SAALFELD (1988), COBB et al. (1998), YUAN und TAO

(1999), BLASBY et al. (2003), SAMAL et al. (2004), HACKELÖER et al. (2015)) zu finden, wobei beide Begriffe synonym verwendet werden (RUIZ et al. 2011). Der Begriff Conflation kommt vom lateinischen *con flare* und bedeutet soviel wie *Verschmelzen* (LYNCH und SAALFELD 1985). Im Gegensatz zum Location Referencing werden bei der Datenintegration bzw. Conflation die korrespondierenden Objekte nach erfolgreicher Zuordnung zu neuen Objekten zusammengeführt. Dabei müssen ggfs. auftretende geometrische oder semantische Konflikte aufgelöst werden (COBB et al. 1998). WALTER und FRITSCH (1995) und GÖSSELN und SESTER (2003) klassifizieren verschiedene Arten der Datenintegration:

- Typ 1: Datensätze von gleicher Datenquelle,
- Typ 2: Datensätze repräsentiert im gleichen Datenmodell,
- Typ 3: Datensätze gespeichert in ähnlichen aber nicht identischen Datenmodellen,
- Typ 4: Datensätze von heterogenen Quellen.

Alternativ zur Verschmelzung der Daten aus verschiedenen Quellen können auch explizite Beziehungen oder Relationen zwischen den korrespondierenden Objekten der einzelnen Datenquellen abgespeichert werden. Diese Möglichkeit der Modellierung wurde beispielsweise im Projekt *Nexus* am Institut für Photogrammetrie der Universität Stuttgart verwendet, wo aus fest gespeicherten Relationen zwischen Mehrfachrepräsentationen mobile, ortsbezogene Dienste entwickelt wurden (VOLZ 2006b). Einen ähnlichen Ansatz verfolgt auch das geplante Projekt *Integrated Core Net - iCoN* zugrunde (POLLESCH 2011). Dabei sollen korrespondierende Elemente von verschiedenen Datenquellen (Basisdaten wie digitale Straßenkarten sowie Fachdaten von Verwaltung und offenen Communities) durch ein integriertes Kern-Netz miteinander verknüpft werden.

### 2.3.5 Interoperabilität beim Location Referencing

Beim Location Referencing spielt die Interoperabilität zwischen den verschiedenen Kartendatenbanken keine direkte Rolle, da ja gerade durch die Definition eines Karten-unabhängigen Übertragungsstandards die Unabhängigkeit von den spezifischen Modellierungen erreicht werden soll. Allerdings sind bei der Entwicklung und Definition von Location Referencing Methoden/Standards natürlich auch die Aspekte der Interoperabilität insoweit zu berücksichtigen, dass möglichst alle für LR-Anwendungen genutzten Karten auch tatsächlich mit dem durch den LR-Standard definierten Datenmodell/Datenformat kompatibel bzw. interoperabel sind. Eine Interoperabilität der beim Location Referencing verwendeten Karten ist also lediglich mittelbar erforderlich bzw. gegeben.

Anders als in der Geoinformatik ist das Location Referencing nicht primär für die Zuordnung ganzer Karten oder Datensätze gedacht. Ziel des LR ist vielmehr die Möglichkeit einer Referenzierung einzelner Kartenelemente zur Übertragung der darauf referenzierten Informationen (z.B. Geschwindigkeitsbegrenzung, Point of Interest) in ein System, welches ggfs. eine andere Kartengrundlage verwendet. Ein vergleichbares Vorgehen zur Übertragung von Informationen zwischen verschiedenen Karten für Routing und Navigation ist in ZHANG et al. (2012) beschrieben. Dennoch sind die in der Geoinformatik verwendeten Zuordnungsansätze grundsätzlich auch auf die Anwendung beim Location Referencing übertragbar. Man kann sagen, dass das Zuordnungsproblem beim Location Referencing ein Spezialfall der Zuordnung bei der Integration von ganzen Kartendatenbanken darstellt. Deshalb werden im Folgenden einige wichtige Arbeiten aus dem Bereich der Integration von geographische Daten (Geodaten) und insbesondere die darin beschriebenen Zuordnungsansätze vorgestellt.

### 2.3.6 Einteilung Zuordnungsverfahren

Nach YUAN und TAO (1999) wird je nach Anwendungsfall zwischen der *horizontalen* und *vertikalen* Zuordnung unterschieden. BLASBY et al. (2003) nennen neben den zuvor genannten Verfahren noch die *interne* Zuordnung.

**Horizontale Zuordnung:** Die horizontale Zuordnung bezieht sich auf das Eliminieren von Diskrepanzen im Überlappungsbereich angrenzender Datensätze. Dabei können sowohl Unterschiede in der Geometrie als auch in der Semantik auftreten. Diese sind für Objekte, welche in beiden Datensätzen vorhanden sind, zu bereinigen. Daraus entsteht dann ein homogener Datensatz, der das gesamte Gebiet der beiden Ursprungsdatensätze abdeckt.

**Vertikale Zuordnung:** Um unterschiedliche Datensätze, die jedoch das gleiche Gebiet abdecken, einander zuzuordnen, werden vertikale Zuordnungsverfahren verwendet. Die Datensätze können entweder unterschiedliche Versionen der selben Datenquelle enthalten oder unterschiedliche thematische Informationen für das selbe Gebiet enthalten. Für die Zuordnung werden korrespondierende Objekte in beiden Datensätzen gesucht. Nachdem der gegenseitige Bezug hergestellt wurde, können entweder Attribute zwischen einander zugeordneten Objekten ausgetauscht werden oder ganze Objekte von einem Datensatz in den anderen überführt werden. Das in dieser Arbeit behandelte Zuordnungsproblem beim Location Referencing ist vom Typ vertikale Zuordnung.

**Interne Zuordnung:** Interne Zuordnungsverfahren werden verwendet, um einander entsprechende Objekte innerhalb des selben Datensatzes zu identifizieren. Diese finden beispielsweise Anwendung im Rahmen von Datensatzbereinigungen.

WALTER (1997), WALTER und FRITSCH (1999), XIONG (2000) und HACKELÖER et al. (2015) unterscheiden Zuordnungsverfahren auch nach der Integrationsrichtung. Dabei wird von *Bottom-Up* und *Top-Down* Ansätzen gesprochen. Beim Bottom-Up-Matching werden zunächst Korrespondenzen zwischen einfachen geometrischen Elementen (z.B. Knoten oder Kanten) gesucht, um anschließend die zu diesen Geometrieelementen gehörenden Objektstrukturen zuzuordnen. Beim Top-Down-Ansatz hingegen wird umgekehrt von der Zuordnung (komplexer) Objektstrukturen (z.B. Straßenelemente) auf die Korrespondenzen der einzelnen geometrischen Elemente geschlossen.

Grundsätzlich lassen sich Zuordnungsverfahren auch nach der Ebene, auf welcher sie durchgeführt werden, unterscheiden. Es wird zwischen der Zuordnung auf Schema-Ebene und Objekt- bzw. Instanzebene unterschieden (DEVOGELE et al. 1998, DUNKARS 2003, OLTEANU 2007b, RAIMOND und MUSTIÈRE 2008, SHEEREN et al. 2009, CHEN 2011). Eine detailliertere Übersicht hierzu wird in den Kapiteln 2.5 und 2.6 gegeben.

## 2.4 Übersicht bestehender Arbeiten

### 2.4.1 Buffer Growing

Die Zuordnung von Linienelementen zweier Karten kann auch mithilfe von Pufferzonen erfolgen. Einfache Pufferzonen um die zuzuordnenden Segmente in der Quellkarte führen bei 1:1 oder 1:n Zuordnungen zum Ziel. Ist das gesuchte Element in der Zielkarte jedoch länger und liegt somit nicht vollständig im Puffer um das Ausgangselement, so muss der Puffer in der Quellkarte um weitere Elemente erweitert werden. Dieses Verfahren zur Zuordnung von n:1 oder n:m Zuordnungen wird *Buffer Growing* genannt und wurde von WALTER (1997) entwickelt. Nach einer Datenvorverarbeitung zur Eliminierung von globalen Fehlern (Translationen und/oder Rotationen) zwischen den Datensätzen wird das Buffer Growing zur Identifikation der potenziellen Zuordnungskandidaten verwendet. Abbildung 2.6 zeigt eine beispielhafte Darstellung. Das Segment a2 soll einem Element der Karte B zugeordnet werden. Es liegt jedoch

kein Element aus der Karte B vollständig im Puffer um  $a_2$ , mit diesem einfachen Puffer ist somit keine Zuordnung möglich. Der Puffer muss also auf die Elemente  $a_2a_3$  ausgeweitet werden. Damit liegen die gesuchten Elemente  $b_1b_2$  vollständig im Puffer und können den Segmenten  $a_2a_3$  zugeordnet werden. Eine zusätzliche Erweiterung des Puffers um das Element  $a_4$ , ist nicht notwendig, da der Puffer um das Element  $a_4$  vollständig dem Element  $b_3$  zugeordnet werden kann. Danach sorgt eine Suchraumbeschränkung für eine Eliminierung aller unwahrscheinlichen Kandidaten unter Ausnutzung von geometrischen Eigenschaften wie Winkel, Länge, Form oder Entfernung. Thematische Beschränkungen waren ebenfalls angedacht, werden jedoch nicht verwendet. Unter den verbleibenden Kandidaten wird anhand einer Leistungsfunktion zur Berechnung ihrer Ähnlichkeit mit dem Element aus der Quellkarte der am besten passende Kandidat ermittelt. Die Zuordnung wird dazu als Kommunikation betrachtet, in dem Nachrichten von einem Sender zu einem Empfänger über einen Kanal übertragen werden. Die Suche nach dem besten Kanal entspricht dem Zuordnungsproblem zwischen zwei Datensätzen angesehen werden. Die Leistungsfunktion berechnet dabei die *gegenseitige Information* der beiden zu vergleichenden Elemente, ein Begriff aus der Informationstheorie. Dazu werden Form, Winkel, Länge und Position sowie die topologischen Eigenschaften (hier: Valenz) der zu vergleichenden Elemente herangezogen. Derjenige Kandidat mit dem größten Leistungswert wird als der wahrscheinlichste für eine Zuordnung angesehen.

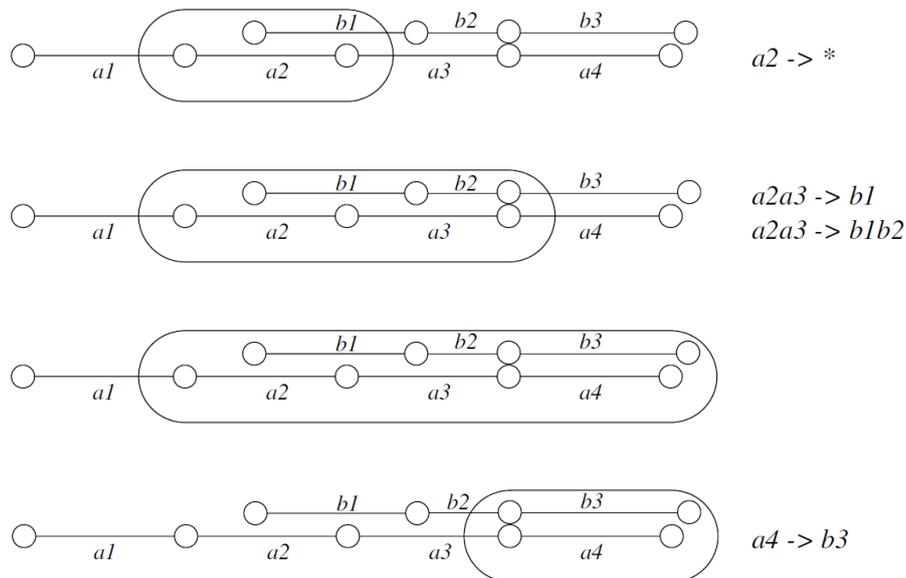


Abbildung 2.6: Buffer Growing (aus WALTER (1997))

Eine erste Implementierung des Verfahrens mit Anwendung auf ATKIS und GDF Daten findet sich in SESTER et al. (1998). WALTER und FRITSCH (1999) haben das Verfahren implementiert und insbesondere die Bestimmung der Parameter der Leistungsfunktionen anhand eines statistischen Verfahrens untersucht. MANTEL und LIPECK (2004) haben das Verfahren innerhalb eines mehrstufigen Zuordnungsverfahrens weiterentwickelt und implementiert. Dem Buffer Growing geht dabei eine Semantische Klassifikation der Ausgangsdatensätze zur Suchraumbeschränkung voraus. Da das Verfahren sowohl merkmalsbasierte als auch relationale Eigenschaften auswertet, ist es ein relationales Zuordnungsverfahren. ZHANG et al. (2005) implementieren ebenfalls ein Buffer Growing Verfahren in ArcGIS zur Verwendung mit BasisDLM Road-Layer und TeleAtlas Straßenobjekten. ZHANG und MENG (2007) entwickeln das Verfahren nochmals weiter und erweitern es um eine *unsymmetrical Buffer Growing* genannte Komponente zur Verwendung von nicht-symmetrischen Puffern.

### 2.4.2 Iterative Closest Point Algorithm

Der *Iterative Closest Point (ICP)* Algorithmus wurde zuerst von BESL und MCKAY (1992) ganz allgemein für dreidimensionale Formen (beschrieben durch Punkte, Linien, implizite oder parametrische Kurven oder Flächen usw.) aufgestellt. Zweck war die Berechnung der Registrierung zweier Formen (z.B. Punktwolken) zueinander, d.h. die Beschreibung von Transformationsparametern, die notwendig sind, um die eine Punktwolke in die andere zu überführen. Der Begriff *Registrierung* stammt dabei aus der Photogrammetrie und wird heutzutage auch im Kontext des Laserscanning häufig gebraucht. Zunächst wird für jeden Punkt der geometrisch nächste Punkt des anderen Datensatzes bestimmt. Aus diesen Zuordnungen werden dann für den gesamten Datensatz Transformationsparameter (die Registrierung) für Rotation und Translation bestimmt und angewendet. Dieses Prozedere wird solange iteriert, bis die gewünschte Registrierungs-Genauigkeit erreicht ist.

Für die Zuordnung von Straßennetzwerken wurde die Identifikation von homologen Knotenpaaren nicht nur auf Grundlage geometrischer Größen (Distanz) durchgeführt, sondern auch um topologische und geometrische Eigenschaften erweitert (z.B. Valenz der Knoten und Richtungswinkelunterschiede der ausgehenden Kanten in ZHANG (2009)). Abbildung 2.7 zeigt in Teil (a) zunächst zwei zuzuordnende Netze, in (b) mit dargestellten Suchradien für die Suche nach korrespondierenden Knoten im grünen Netz, in (c) dann durch Doppelpfeile angedeutet die aus den gefundenen Knoten-Zuordnungen abgeleiteten Kanten-Zuordnungen der ersten Iteration. Für die noch nicht zugeordneten Knoten und Kanten können diese Zuordnungsschritte mit gelockerten Bedingungen (angedeutet in (d) mit größeren Suchradien) iterativ wiederholt werden, bis schließlich die gesamten Netze zugeordnet werden konnten.

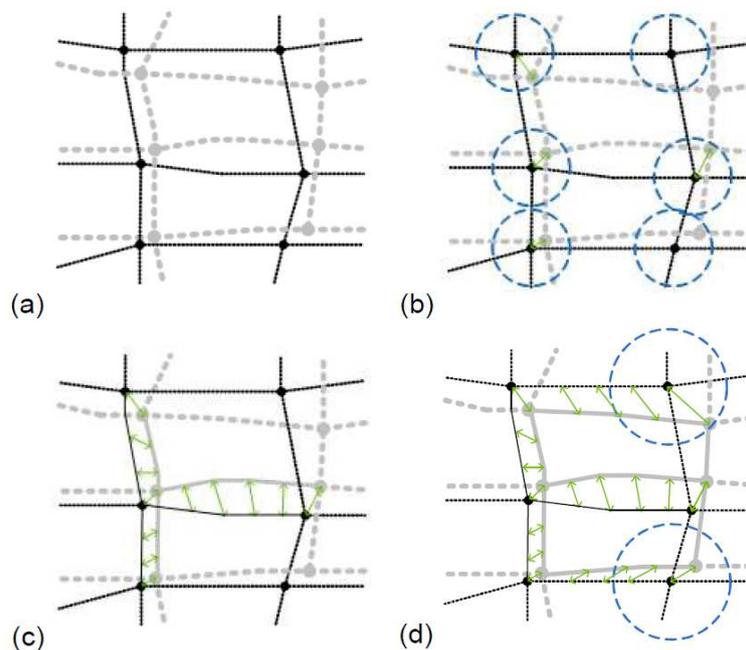


Abbildung 2.7: ICP Verfahren (aus ZHANG (2009))

GÖSSELN und SESTER (2003), GÖSSELN (2005), SESTER et al. (2007) und BUTENUTH et al. (2007) verwenden den ICP Algorithmus im Rahmen von Datenintegrationsverfahren von zweidimensionalen Datensätzen als 4-Parameter Ähnlichkeits- oder Helmert-Transformation (Lage, Maßstab, Orientierung). Die Zuordnung zwischen den Objekten der Datensätze basiert auf semantischen und geometrischen Eigenschaften. ICP wird dabei für die grobe Zusammenführung der einzelnen Objekte eingesetzt und es werden damit Transformationsparameter für ein *best fit* berechnet.

### 2.4.3 Mehrstufige Zuordnungsverfahren

VOLZ (2006a) beschreibt einen iterativen Ansatz zur kombinierten Zuordnung von Knoten und Kanten aus Straßennetzen. Zunächst werden diese einer intensiven Datenvorverarbeitung unterzogen, in der mögliche geometrische Abweichungen zwischen den beiden zuzuordnenden Datensätzen mittels einer Rubber-Sheeting-Transformation beseitigt werden. Die Transformation basiert auf der Identifikation von Passpunkten mit sehr hoher Übereinstimmungswahrscheinlichkeit in den beiden Datensätzen. Dazu dürfen die Knoten nicht weiter als 100 m auseinanderliegen und müssen über mindestens 4 abgehende Kanten verfügen, die jeweils ähnliche Längen- und Richtungswinkelwerte aufweisen. Ähnlich wie das Verfahren von DOYTSHER und FILIN (2000) wird die Suche nach übereinstimmenden Knoten in beiden Richtungen (bi-direktional) durchgeführt, um eine gegenseitige Eindeutigkeit der Zuordnung (engl. *mutual unambiguosness*) der Passpunkte sicherzustellen. In einem zweiten Vorverarbeitungsschritt wird das Problem aufgegriffen, dass mögliche n:m Zuordnungen wesentlich schwieriger zu bestimmen sind, als dies für 1:1 Zuordnungen der Fall ist. Durch eine Aufspaltung der Kanten sollen die n:m Zuordnungen weitestgehend durch 1:1 Zuordnungen ersetzt werden. Die Zuordnung der Kanten erfolgt mit Puffern. Die eigentliche Zuordnung basiert auf der Identifizierung von Startknoten (engl. *seed nodes*) mit sehr hoher Zuordnungswahrscheinlichkeit, von denen ausgehend alle weiteren Kanten und Knoten des Netzes auf mögliche Übereinstimmungen analysiert werden. Der seed-node Ansatz basiert auf den Arbeiten von KRAUT (2003) und XIONG und SPERLING (2004). Die Startknoten werden ähnlich den Passpunkten für die Rubber-Sheeting Transformation mit sehr hohen Anforderungen an die geometrische Nähe sowie die topologische Übereinstimmung der ausgehenden Kanten identifiziert. Ausgehend von diesen Seed Points werden die davon abgehenden Kanten und deren Endknoten anhand topologischer und geometrischer Ähnlichkeitsmaße bewertet. Diese beinhalten Distanz und Valenz für die Knoten sowie Längenunterschied, Richtungswinkelunterschied, Vertex-Hausdorff-Distanz, Adjazenzbeziehungen zwischen den Start- und Endknoten für die Kanten. Die verschiedenen Ähnlichkeitsmaße werden jeweils gewichtet zu einem Gesamt-Ähnlichkeitsmaß zusammengeführt, welches entsprechend einer Fallunterscheidung der Endknoten ausgewertet wird. Dabei wird unterschieden, ob die Endknoten ihrerseits als seed node identifiziert wurden, und ob sie bereits einem Knoten im anderen Netz zugeordnet wurden oder nicht. Für 1:1 Zuordnungen genügt dieser einfache Vergleich, für 1:n Zuordnungen allerdings wird ein Verfahren angelehnt an das *Buffer Growing* von WALTER (1997) und WALTER und FRITSCH (1999) angewendet. Diese Kantenzuordnung wird iterativ solange wiederholt, bis die Netze vollständig zugeordnet wurden. Insgesamt kann das Verfahren auch als Erweiterung des ICP-Algorithmus angesehen werden (ZHANG und MENG 2008, ZHANG 2009).

### 2.4.4 Delimited-Stroke-oriented Matching Algorithm

ZHANG und MENG (2008) stellen mit dem *Delimited Stroke Oriented (DSO)* Algorithmus ein neues Verfahren zur automatischen Zuordnung von Straßennetzwerken vor. Der Ansatz beruht auf der Annahme, dass Buffer Growing und Iterative Closest Point (inkl. dem mehrstufigen Verfahren nach VOLZ (2006a)) zwar gute Ergebnisse liefern, aber diese sind insbesondere in Gebieten mit komplexem geographischem Kontext oder wenn die Datensätze nur wenig oder gar keine semantischen Informationen enthalten nicht zufriedenstellend genug. Daher soll das neu entwickelte Verfahren stärkeren Gebrauch von Kontext-Informationen machen und damit die topologischen Beziehungen zwischen den einzelnen Straßenobjekten besser ausnützen können. Dies soll die geometrische und semantische Zuordnung stützen. Der DSO gliedert sich in vier einzelne Prozesse auf: (a) Erstellung eines Indexes zur Beschreibung der Beziehungen zwischen verbundenen Objekten; (b) Konstruktion der *delimited strokes*; (c) Zuordnung der *delimited strokes*; (d) Weitergehende Zuordnungen ausgehend von den *seed nodes*.

Das Erstellen eines Indexes zur Beschreibung der Beziehungen zwischen verbundenen Objekten im ersten Schritt könnte man auch als Vorverarbeitung oder Aufbereitung der Datensätze für das eigentliche Matching bezeichnen. Es werden alle topologischen Beziehungen (hier als *Kontext Information*

bezeichnet) zwischen den Straßen-Objekten identifiziert und explizit in einer Tabelle abgespeichert. Die Tabelle enthält also für jede Kante die Identifikator/Bezeichner der mit ihrem Start- und Endknoten verbundenen anderen Kanten.

Im darauffolgenden zweiten Schritt werden aus den zuzuordnenden Straßennetzwerken die *delimited strokes* gebildet. Delimited Strokes (DS) sind per Definition Sequenzen von verbundenen Liniensegmenten, die eine hohe Stetigkeit (engl. *good continuity*) aufweisen, d.h. keine signifikanten Richtungsänderungen in ihrem Verlauf beschreiben. Zum Anderen müssen die Liniensequenzen von sog. *Efficient terminating nodes* begrenzt werden, also von Knoten die eine Valenz von entweder 1 oder größer als 3 haben. Dadurch kommen nur solche Knoten als Start-/Endknoten in Frage, die mit relativ hoher Wahrscheinlichkeit im anderen Netz auffindbar sind. Abbildung 2.8 beschreibt diese beiden Grundprinzipien des DSO.

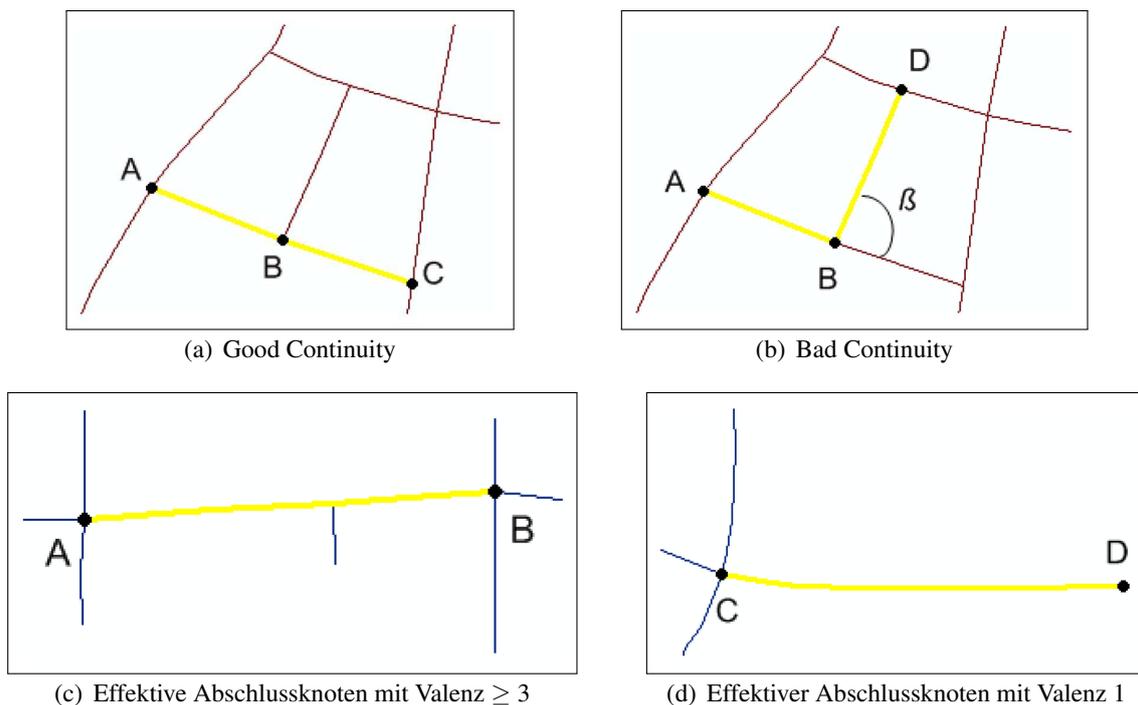


Abbildung 2.8: Grundprinzipien des Delimited Stroke Algorithmus (aus ZHANG (2009))

Der dritte Schritt umfasst schließlich das eigentliche Matching der delimited strokes, für das zunächst ein Ähnlichkeitsmaß zum Vergleich zweier delimited strokes definiert wird. Die verwendeten Ähnlichkeitsmaße beruhen auf den Arbeiten von ZHANG et al. (2005) und ZHANG und MENG (2007) und wurden in ZHANG (2009) zusammengefasst und detailliert veröffentlicht. Sie beinhalten unterschiedliche geometrische sowie topologische Eigenschaften.

Die geometrische Ähnlichkeit der beiden Polylinien wird basierend auf den Distanzen zwischen den Anfangs- und Endpunkten, den Differenzen der Länge, Richtungswinkeln und der beschriebenen Fläche sowie einer Formbeschreibung auf Basis der *Turning Function* (vgl. Abschnitt 2.6.1.4) sowie einer an die Frechét-Distanz angelehnte Distanz-Funktion berechnet. Die topologischen Ähnlichkeiten der Anfangs- und Endpunkte werden basierend auf der Anzahl und den jeweiligen Richtungswinkeln der ausgehenden Kanten bestimmt. Eine Steigerung der Zuordnungs-Güte kann durch Erkennen von Strukturen im Straßennetz (z.B. Kreisverkehre) und durch die Hinzunahme von semantischen Informationen (z.B. Straßename oder Klasse) erreicht werden. Des Weiteren wird durch die Verwendung von räumlichen Indizes der Zugriff auf die Kartendaten optimiert und somit die Berechnungsgeschwindigkeit beschleunigt.

Zu Beginn des Zuordnungsvorgangs wird ausgehend vom Startknoten des delimited strokes der Quellkarte nach Start-/Endknoten von delimited strokes der Zielkarte gesucht, die eine geometrische und topologische Ähnlichkeit aufweisen. Von diesen Knotenkandidaten ausgehend werden die delimited strokes der Zielkarte untersucht, ob sie einen ähnlichen geometrischen Verlauf haben, wie das Referenzobjekt. Im Fall von gravierenden Längenunterschieden werden weitere delimited strokes angehängt, bis die Anfangs- und Endpunkte der Sequenzen eine gewisse räumliche Nähe aufweisen. Dann erst werden die Ähnlichkeitsmaße berechnet und mit einem vorgegebenen Grenzwert verglichen. Durch die dynamische Verkettung von delimited strokes ist das Verfahren in der Lage, nicht nur 1:1 Zuordnungen, sondern auch n:m Zuordnungen zu detektieren. Alle im Zielnetz gefundenen delimited strokes mit einem Ähnlichkeitsmaß oberhalb des Grenzwertes, werden als potenzielle Zuordnungskandidaten betrachtet. Um die Genauigkeit und Zuverlässigkeit einer Zuordnung zu steigern, werden nicht nur die einzelnen delimited strokes einer Ähnlichkeitsbetrachtung und Vergleich mit dem Referenzobjekt unterzogen. Zur Einbeziehung des geographischen Kontextes in die Zuordnung werden die delimited stroke Kandidaten vielmehr mit Hilfe der zu Beginn identifizierten topologischen Beziehungen auf einen größeren Netzbereich ausgeweitet. Für diesen Netzausschnitt wird eine Gesamtähnlichkeit berechnet und derjenige Netz-Kandidat mit der höchsten Ähnlichkeit als wahrscheinlichste Zuordnung angesehen.

Die dadurch gewonnen Knoten-Zuordnungen im Netzausschnitt können dann noch als *seed points* betrachtet und wiederum für eine Ausdehnung der Zuordnung im Netz verwendet werden. Die mit den seed points verbundenen Kanten und Knoten werden wieder einer Ähnlichkeitsbetrachtung unterzogen und im Falle einer großen Übereinstimmung in den Vergleichsparametern eine Zuordnung festgestellt. Das zugeordnete Netz kann also immer weiter ausgeweitet werden, die Zuordnung endet erst an Stellen, an denen die zu vergleichenden Kanten und Knoten eine starke Abweichung im Ähnlichkeitsmaß aufweisen. ZHANG et al. (2010) haben das Verfahren noch für spezielle Strukturen wie z.B. mehrspurige Fahrbahnen oder Kreisverkehre adaptiert und implementiert. ZHANG et al. (2012) haben das automatische delimited stroke Verfahren zur Übertragung von Attributen zwischen verschiedenen Karten um eine manuelle Nachbearbeitung ergänzt und implementiert.

### 2.4.5 Weitere Verfahren

Eine der ersten Veröffentlichungen über eine automatische Zuordnung von Karten ist ROSEN und SAALFELD (1985). Die Arbeit wird motiviert durch die Zusammenführung von Karten des United States Geological Survey und des Bureau of the Census. Bis dato wurden einzelne Karten mithilfe von lokalen Rubber-Sheeting Transformationen interaktiv aligniert. Auf Grundlage der Erfahrungen aus dem manuellen Ausrichten der Karten, wonach Menschen intuitiv Objekte in den Karten zur Deckung bringen können und sich dabei, obwohl natürlich unterschiedliche Ansichten im Detail möglich sind, die Ergebnisse im Großen und Ganzen erstaunlich ähnlich sind, wird untersucht, nach welchen Kriterien diese Zuordnung geschieht und wie dies dann zur automatisierten Umsetzung genutzt werden kann. Das entwickelte Verfahren arbeitet iterativ, wobei zuerst die Knoten den „wahrscheinlichsten“ Entsprechungen in der anderen Karte zugeordnet werden und danach die mit den Knoten verbundenen Linien. Die Zuordnung erfolgte anhand geometrischer und topologischer Kriterien. Dieses Verfahren wurde in SAALFELD (1988) weiterentwickelt. Demnach werden zuerst die stärksten Zuordnungskriterien angewendet, gefolgt von den weniger starken bis hin zu den ganz schwachen zuletzt. Dies führt dazu, dass zunächst die identischen Elemente zugeordnet werden, und erst danach die einander ähnlichen Objekte. Nach jeder Iteration wird die gegenseitige Ausrichtung der Karten mithilfe von Rubber-Sheeting Transformationen neu berechnet und verbessert.

DERETSKY und RODNY (1993) beschreiben ein Verfahren zur Kombination geometrischer und attributiver Daten aus zwei digitalen Straßenkarten zu einer ultimativen Karte. Dazu werden übereinstimmende Knoten, Kanten und Flächen in beiden Karten nach geometrischen und topologischen Kriterien identifiziert. Das dargestellte Verfahren basiert auf Ketten von Kanten, mit denen korrespondierende

Kreuzungen zwischen den Ketten bestimmt werden können. Die Ketten werden mithilfe von geometrischen und semantischen Eigenschaften der Kanten gebildet. Durch die Ketten werden die Karten in kleinere, paarweise zugeordnete Gebiete aufgeteilt. Die gefundenen Kreuzungspunkte werden als Relationen zwischen den Ketten betrachtet und in einer separaten Datenbank abgespeichert.

GABAY und DOYTSHER (1994) beschreiben ein Verfahren zur Bestimmung der gemeinsamen Linien in zwei geometrisch und topologisch ähnlicher aber mit unterschiedlicher Genauigkeit erfasster kartographischer Datensätze. Dies ist das erste Verfahren, bei dem auch solche Linien identifiziert werden können, die kein Gegenstück in der anderen Karte haben. Das Verfahren basiert auf der Bildung von Puffern um die Punkte und Linien der einen Karte, wobei die Radien bzw. Abstände der Puffer den Genauigkeitswerten der Punkte entsprechen. Daher wird diese Methode auch Epsilon-Band-Methode genannt (LEMARIÉ und RAYNAL 1996, RUIZ et al. 2011). Da dieses Verfahren ausschließlich auf der geometrischen Ähnlichkeit einzelner Linien basiert, eignet es sich auch bei Karten, die große topologische Unterschiede aufweisen (GABAY und DOYTSHER 2000). Zur Zuordnung zweier Linien wird zunächst geprüft, ob die Endpunkte der beiden Kandidaten-Linien im Puffer-Bereich des jeweils anderen liegen und zur Verifikation einer möglichen Zuordnung dann noch die Richtungswinkel der beiden Kandidaten verglichen. Dieses Verfahren wird zuerst auf Linienzüge angewendet, die von Punktpaaren ausgehen, welche in einem Vorbereitungsschritt bereits erfolgreich einander zugeordnet werden konnten. In einem zweiten Schritt werden auch solche Linien untersucht, die nach dem ersten Durchgang noch ohne Gegenstück geblieben waren. Darauf aufbauend wird in GABAY und DOYTSHER (1995) ein automatisches Verfahren zur geometrischen Verbesserung von Linien in Karten mit niedrigerer Genauigkeit vorgestellt. Die Genauigkeitssteigerung wird durch die Zuordnung der Linien zu Gegenständen in einer Karte mit höherer Genauigkeit erreicht.

BRUNS und EGENHOFER (1996) beschreiben ein Verfahren zur Bestimmung der Ähnlichkeit von räumlichen Objekten. Sie verwenden dazu topologische, richtungs- und distanzbezogene Relationen zwischen Objekten innerhalb verschiedener räumlicher Szenen. Sie definieren als Maß für die Ähnlichkeit von Objekten die Anzahl der Transformations-Schritte, die benötigt werden, um die eine Szene in die jeweils andere zu überführen. Je weniger solcher Änderungsschritte nötig sind, um Szenen ineinander zu überführen, desto ähnlicher sind sie sich. Interessant an diesem Ansatz ist die ausschließliche Verwendung von Relationen und der konzeptionellen Nachbarschaft zwischen ihnen. So wird beispielsweise der Abstand zwischen Objekten nicht etwa durch eine Distanzfunktion definiert. Es wird vielmehr eine distanzbezogene Relation mit den Stufen *null* < *sehr nahe* < *nahe* < *weit* eingeführt. Transformations-schritte sind nur zwischen angrenzenden Stufen möglich, sodass für eine Überführung zweier Objekte die *weit* auseinanderliegen, drei Transformationsschritte nötig sind.

COBB et al. (1998) beschreiben ein hierarchisches, regelbasiertes Verfahren zur Verschmelzung von Vektordaten. Die Beurteilung der Zuordnungskriterien erfolgt hierbei auf Grundlage des *reasoning under uncertainty* (z.B. Fuzzy Logik oder Dempster-Shafer Theorie, auch Evidenztheorie genannt), also unter der Annahme, dass explizite oder starre Attributzuweisungen nicht möglich sind und vielmehr ein Element lediglich mit einer bestimmten Wahrscheinlichkeit einer unscharfen (engl. fuzzy) Gruppe zugeordnet werden kann. Die Zuordnung selbst erfolgt mehrstufig. In einer Datenvorverarbeitung werden Objekte der Realwelt, welche durch mehrere Kartenobjekte repräsentiert werden, wieder zu einem logischen Objekt zusammengeführt, um eventuelle Modellierungsunterschiede in den verschiedenen Karten auszugleichen. Danach werden mögliche Zuordnungskandidaten für ein bestimmtes Objekt der Quellkarte aus der Zielkarte über eine Entfernungsfunktion bestimmt. Dann werden in einem ersten Schritt Objekte des gleichen Objekttyps gesucht sowie weitere Attributwerte mit Hilfe semantischer Zuordnungsfunktionen ausgewertet. In einem zweiten Schritt werden die geometrischen Eigenschaften der Objekte betrachtet. Dazu wird nach einer Zuordnung der Knoten, die Form der zu betrachtenden Objekte und ihre relative Position zueinander mit Hilfe der Frechët-Distanz untersucht. Zuletzt werden dann die topologischen Eigenschaften der beiden Objekte verglichen.

YUAN und TAO (1999) beschreiben einen komponentenbasierten Ansatz zur Verschmelzung von geographischen Daten (Geodaten). Sie gehen dabei davon aus, dass die Auswahl und Reihenfolge der

verwendeten Zuordnungsansätze abhängig ist von der konkreten Anwendung, d.h. von den konkret zu vergleichenden Datensätzen und der jeweiligen Fragestellung. Deshalb schlagen sie eine Reihe von Zuordnungsverfahren vor, die der Nutzer dann anwendungsspezifisch zusammenstellen und auf seine zu untersuchenden Datensätze anwenden kann. Die vorgeschlagenen Verfahren gliedern sich in *geometrische*, *topologische* und *semantische* Verfahren. Geometrische Zuordnungen können dabei mittels Distanzfunktionen (z.B. Hausdorff- oder Frechét-Distanz), Richtungswinkel der Linien, geometrischen Beziehungen zwischen Objekten (z.B. innerhalb eines Polygons) oder der Form der Objekte (z.B. Länge, Krümmung, Umfang, Fläche, etc.). Topologische Zuordnungen werden mithilfe von topologischen Beziehungen zwischen den Objekten (Konnektivität) oder Adjazenzbeziehungen vorgeschlagen. Allerdings wird darauf verwiesen, dass topologische Zuordnungsverfahren in der Regel nicht als alleinige Verfahren angewandt werden, sondern eher zur Suchraumbeschränkung oder Kontrolle der Ergebnisse aus der geometrischen Zuordnung. Topologische Verfahren können dazu verwendet werden, Zuordnungen auf das gesamte Netz auszudehnen, jedoch nur wenn die beiden Netze eine hohe topologische Ähnlichkeit aufweisen. Semantische Zuordnung erfolgt durch Attributvergleich, wenn die Semantiken der beiden zu vergleichenden Datensätze bekannt sind.

DOYTSHER und FILIN (2000) beschreiben ein relationales Verfahren zur Detektion von entsprechenden (Linien-)Objekten in Karten. Der Ansatz basiert zunächst auf einer Identifikation von zusammengehörigen Punktepaaren in den beiden Karten. Dies geschieht relativ simpel durch Bestimmung der relativen Nähe der Objekte mittels Suchradius. Die Autoren empfehlen die Suche nach Entsprechungen von beiden Karten aus durchzuführen, um auch alle  $n:m$  Zuordnungen zu finden (bi-direktionale Suche). Wird lediglich nur uni-direktional gesucht (z.B. von der Quellkarte ausgehend nach Übereinstimmungen in der Zielkarte) können nur  $1:n$  Zuordnungen zuverlässig aufgedeckt werden. Die dadurch gefundenen Kandidaten-Knoten werden dann anhand der Anzahl der von ihnen ausgehenden Kanten und deren Richtungswinkel mit den Charakteristiken aus der Quellkarte verglichen. Da diese Knoteneigenschaften aus verschiedenen Gründen als fehlerbehaftet angesehen werden müssen, also sich in den zu vergleichenden Karten durchaus unterscheiden können, schlagen die Autoren robuste Methoden zur Bestimmung von robusten Vergleichsmaßen vor. Weil die rein geometrische Suche nach der richtigen Zuordnung dennoch manchmal falsche Ergebnisse liefern kann, wird das Verfahren durch die Einbeziehung topologischer Eigenschaften ergänzt. Konkret wird vorgeschlagen, von bereits identifizierten Zuordnungspaaren ausgehend, in beiden Netzen nach den jeweils nächsten Knotenpunkten zu suchen. Dieser Schritt soll iterativ so lange wiederholt werden, bis keine weiteren Kandidaten mehr gefunden werden. Nach erfolgreicher Zuordnung der Knoten erfolgt nun die Berechnung von Routen zwischen den Knoten im Zielnetz. Der Vergleich von möglichen Routen mit der Referenz aus dem Quellnetz erfolgt dabei unter Berücksichtigung der Fläche zwischen den beiden Routen und deren Gesamtlänge.

DUEKER und BUTLER (2000) gehen einen anderen Weg und schlagen die Einführung eines neuen Datenmodells vor, um den Bedürfnissen von Verkehrsdaten mit vielen unterschiedlichen Daten- und Service-Providern und hohen Update-Zyklen gerecht zu werden. Dieses bezeichnen sie als Geographisches Informations System - Transportation (GIS-T). In ihrem Modell sind die Verkehrsinformationen (also z.B. Geschwindigkeitsbegrenzungen) entlang der Straßensegmente aus der digitalen Karte definiert, können jedoch über Kreuzungen hinaus gehen. Dadurch können die selben Verkehrsinformationen mit weniger aber längeren Objekten modelliert werden, was den Update-Prozess vereinfacht. Zum Austausch der Verkehrsinformationen zwischen den verschiedenen Systemen schlagen die Autoren die Verwendung von einheitlichen Identifikatoren (IDs) vor.

KRAUT (2003) hat ein relationales, knotenbasiertes Zuordnungsverfahren entwickelt und implementiert, welches zunächst mit einer Startzuordnung beginnt. Dabei werden solche Knoten in beiden Netzen gesucht, die aufgrund ihrer Eigenschaften mit sehr hoher Wahrscheinlichkeit korrespondierenden. Dies ist insbesondere bei Kreuzungen der Fall, deren geometrischer Abstand in beiden Karten ein bestimmtes Maß nicht übersteigt, die drei oder mehr ausgehende Kanten besitzen und die Winkeldifferenzen dieser Kanten auch innerhalb festgesetzter Grenzen liegen. Von dieser Startzuordnung ausgehend wird zunächst versucht, die noch nicht zugeordneten Knoten über eine Linienverfolgung in beiden Netzen

unter Einbeziehung der topologischen Nachbarschaftsbeziehungen sowie der semantischen Eigenschaften (Straßennamen und -Klassifizierung (L, B, A, etc.) sowie erlaubte Fahrtrichtung) zu identifizieren. Schlägt die Zuordnung fehl, wird alternativ ein merkmalsbasierter Zuordnungsansatz verfolgt, der ähnlich wie die Startzuordnung auf geometrischen Kriterien basiert. Zum Schluss werden alle gefundenen Zuordnungspaare anhand mehrerer Eigenschaften bewertet und so die wahrscheinlichste richtige Zuordnung identifiziert.

DUNKARS (2003) und DUNKARS (2004) beschreibt ein automatisches Zuordnungsverfahren von Elementen zweier Datensätze mit unterschiedlicher Auflösung zum Aufbau von Multi-Repräsentations-Datenbanken. Die verwendeten Ähnlichkeitsmaße basieren auf semantischen, geometrischen, topologischen und inter-Objekt-Beziehungen und können mittels einer eigens definierten n-dimensionalen, gewichteten Euklidischen Distanzfunktion zu einem globalen Ähnlichkeitsmaß zusammengeführt werden. Dazu werden die Ähnlichkeitsmaße in quantitative und qualitative Masse unterschieden, letztere müssen zur weiteren Berechnung noch in solche mit nominalen und ordinalen Skalen unterschieden werden. Das entwickelte Verfahren ist grundsätzlich für verschiedene Arten von Daten verwendbar. Es wurde mit verschiedenen topografischen Datensätzen getestet, darunter auch Straßendaten. Für die Zuordnung von Straßendaten wurden als geometrische Merkmale die Positionen der Anfangs- und Endpunkte der Linien, die Länge der Linien und Richtungswinkel verwendet. Als topologische Eigenschaften wurden die Anzahl der im Netz angrenzenden Straßenelemente sowie einige deren Eigenschaften (z.B. Straßenklasse) genutzt. Die semantische Ähnlichkeitsbetrachtung fußte auf dem Vergleich von Attributwerten (z.B. Straßenummer).

XIONG (2000) stellt einen dreistufigen relationalen Ansatz zur Zuordnung von Netzwerken aus unterschiedlichen Quellen vor, die Zuordnung erfolgt auf Knoten-, Segment- und Kanten-Ebene. Ein Merkmal dieses Verfahrens ist, dass die Bottom-up und Top-down Zuordnungstechniken kombiniert werden. Zunächst mit Bottom-up gestartet, d.h. es werden zuerst die Knoten zugeordnet, dann die Segmente und zuletzt dann die Kanten. Danach folgt der Top-Down Ansatz, bei dem zuerst die (noch nicht zugeordneten) Kanten bearbeitet werden, gefolgt von Segmenten und Knoten. Der Zweite Teil baut dabei auf den erfolgreichen Zuordnungen aus dem Bottom-up Durchgang auf. Die Knoten-Zuordnung erfolgt dabei mittels zweier Kriterien, nämlich der Distanz zwischen den beiden Kandidaten sowie der Ähnlichkeit der Winkelbilder, die sich aus den abgehenden Kanten ergeben. Die Zuordnung auf Segment-Ebene basiert auf drei Kriterien: Winkeldifferenz zwischen den Segmenten einer Kante, Abstand zwischen den Segmenten und Segmentlänge. Dabei wird die Länge als Gewichtungsfaktor für die beiden anderen Kriterien verwendet. Die Kantenzuordnung verwendet die gleichen Kriterien wie die der Segmente, allerdings auf die gesamte Kante angewendet: durchschnittliche Winkeldifferenz, durchschnittliche Distanz und Gesamtlänge. XIONG und SPERLING (2004) greifen diesen Ansatz auf und erweitern und verbessern ihn mit einem clusterbasierten Zuordnungsalgorithmus. Dieser beinhaltet eine Selektion von Knoten-Zuordnungen (engl. *seed nodes*) mit sehr hoher Zuordnungs-Wahrscheinlichkeit. Dazu werden Cluster um Knoten gebildet, die drei oder mehr ausgehende Kanten besitzen, da diese mit hoher Wahrscheinlichkeit im anderen Netz vorhanden sind und die entsprechenden Korrespondenzen auch erfolgreich identifiziert werden können. Von diesen Knotenpaaren ausgehend werden entlang beider Netze weitere Knoten und ihre Zuordnungen gesucht. Nach erfolgreicher Knotenzuordnung folgt dann die Segment- und Kantenzuordnung wie oben beschrieben.

SAMAL et al. (2004) beschreiben zwei verschiedene Zuordnungsansätze, einen kontextunabhängigen und einen kontextbasierten Ansatz. Unter geographischem Kontext verstehen die Autoren dabei die Beziehungen eines Objektes (z.B. einer Kante) zu seinen benachbarten Objekten. Der zunächst vorgestellte kontextunabhängige Ansatz vergleicht die Ähnlichkeit zweier Zuordnungskandidaten anhand ihrer Eigenschaften (*Attribute*). Dies sind konkret skalare Attributwerte, Zeichenketten (Vergleich anhand der Levenshtein-Distanz), aber auch anhand geometrischer Abstände in den beiden Karten mittels euklidischer Distanz und der Objektform nach GOODCHILD und HUNTER (1997). Alle diese Eigenschaften werden für jedes Objektpaar durch einen gewichteten Mittelwert zusammengeführt. Die Autoren schlagen darüberhinaus die Zusammenführung aller dieser gewichteten Ähnlichkeitsmittel für alle

Objektpaare zweier Datensätze in einer sogenannten *Ähnlichkeitsmatrix* vor. Der zweite, kontextbasierte, Zuordnungsansatz, basiert auf den Entfernungen und Richtungen der umgebenden Objekte. Der geographische Kontext wird als invariant und dadurch als besonders für den Kartenvergleich geeignet angesehen. Die Umsetzung erfolgt durch Berechnung sogenannter Ähnlichkeitsgraphen (engl. *proximity graphs*) aus den Richtungswinkeln und Entfernungen zu den umliegenden Objekten. Die Ähnlichkeit der Ähnlichkeitsgraphen in den verschiedenen Karten wird durch die Identifikation sog. *landmarks*, also Objekten, die mit hoher Wahrscheinlichkeit in allen Datensätzen vorhanden sind. Eine automatische Erkennung solcher Landmarken wird durch spezielle Auswertung der Ähnlichkeitsmatrizen vorgeschlagen. KIM et al. (2010) wandeln die Idee des geografischen Kontext ab und berechnen ihn mit Hilfe von Voronoi Diagrammen.

BEERI et al. (2004) konzentrieren sich auf die Zuordnung von Knoten und beschreiben drei Verfahren die ausschließlich auf Grundlage der Knoten-Positionen beruhen. Zusätzliche topologische oder semantische Informationen sehen sie allenfalls als Ergänzungsmöglichkeiten ihres vorgeschlagenen Verfahrens. Der erstgenannte Ansatz nutzt einseitige Nachbarschaftsbeziehungen (engl. *one-sided nearest neighbor*) für die Zuordnung, wohingegen der zweite Ansatz die Nachbarschaftsbeziehungen in beide Richtungen auswertet (engl. *mutually nearest neighbor*). Dieses wiederum wird nochmals weiterentwickelt und um eine probabilistische Zuordnungsfunktion ergänzt wird *Normalized-Weights-Join* genannt. Darin ist die Zuordnungswahrscheinlichkeit der beiden zu betrachtenden Punkte aus zwei verschiedenen Datensätzen umgekehrt proportional zur Distanz zwischen den beiden Punkten. Demnach steigt die Wahrscheinlichkeit  $P_{a_i}$ , dass ein Punkt  $a_i$  einem Punkt  $b_j$  zugeordnet wird, je kleiner die Distanz zwischen  $a_i$  und  $b_j$  ist. Diese Wahrscheinlichkeiten finden nun Eingang in eine Zuordnungsmatrix  $M$  und werden mit einem iterativen Algorithmus normalisiert, an dessen Ende die Wahrscheinlichkeiten für eine Zuordnung von  $a_i$  und  $b_j$  sowie für die Fälle, dass  $a_i$  und/oder  $b_j$  keine Entsprechung im jeweils anderen Datensatz enthalten, berechnet wurden. Das letztgenannte Verfahren mit normalisierten Gewichten erzielte bei der Anwendung sowohl auf Realdaten als auch mit simulierten Daten die besten Ergebnisse und wird von den Autoren daher empfohlen. BEERI et al. (2005) überführen die in BEERI et al. (2004) beschriebenen iterativen Verfahren in einen simultanen Ansatz, bei dem die Datensätze gleichzeitig verarbeitet werden. Diesen Ansatz bezeichnen die Autoren als ganzheitlich (engl. *holistic*). Beide Arten der Zuordnung werden in SAFRA et al. (2010) nochmals weiterentwickelt.

In SAFRA et al. (2006) wird die in BEERI et al. (2004) vorgestellte *Mutual Nearest Neighbor* Zuordnungsmethode für Punkte aufgegriffen und um ein Verfahren zur Zuordnung von Linien ergänzt. Dabei steht die Konzentration auf geometrische Eigenschaften, nämlich die Position der Linienendpunkte im Vordergrund. Nach einer Datenvorbereitung zur Identifikation aller topologischen Knoten im Straßennetzwerk, werden die Knoten der beiden Straßennetze einander zugeordnet. Dies erfolgt entweder anhand der klassischen einseitigen oder der gegenseitigen (engl. *mutual*) Nachbarschaftsbedingung.

In MUSTIÈRE (2006) wird ein automatisches Zuordnungsverfahren vorgestellt und anhand zweier Netzwerke mit unterschiedlichen Maßstäben aus der französischen IGN Datenbank ausführlich getestet. Dieses Verfahren kann grundsätzlich auch zur Anwendung auf Netze mit ähnlichen Maßstäben verallgemeinert werden. Zunächst werden in einer Vorverarbeitung die verwendeten Netze in eine einheitliche Struktur gebracht, wobei bei spezifischen Netzdaten (z.B. Energieversorgung), welche nicht zwangsläufig aus Kanten und Knoten bestehen, die entsprechenden Elemente durch Kanten und Knoten ersetzt. Dieser Schritt ist für die Anwendung auf Straßenkarten nicht relevant, wohl aber beispielsweise Ansätze, wonach komplexere Knotenstrukturen an Kreuzungen im Vorfeld bereinigt und zu einem komplexen Knoten zusammengefasst werden. Des Weiteren wird auch vorgeschlagen, Informationen, die nur implizit in der Datenbank vorhanden sind (z.B. Kreisverkehre) zu detektieren und explizit abzuspeichern (Im Gegensatz zu einigen gängigen Kartenformaten, in denen z.B. Kreisverkehre explizit ausgewiesen werden, war dies in den hier verwendeten Datensätzen offenbar nicht der Fall). Für die eigentliche Zuordnung findet zunächst eine grobe Zuordnung der Knoten statt. Die Knoten werden mit einem Distanz-Kriterium zugeordnet, wobei dieses in Abhängigkeit von Attributinformationen, z.B. der Valenz der Knoten (entspr. der Komplexität einer Kreuzung), gestaffelt werden kann. Danach findet eine

Grobzuordnung der Kanten mittels einer geometrischen Distanzfunktion (Hausdorff-Distanz) statt. Bei der Zuordnung von Karten unterschiedlichen Maßstabs wird die gerichtete Hausdorff-Distanz vom Objekt aus dem Detail-reicheren Datensatz zum Objekt des Detail-ärmeren Datensatzes verwendet. Im allgemeinen Fall mit Karten ähnlichen Maßstabs kann die normale Hausdorff-Distanz verwendet werden. Semantische Informationen (z.B. Straßenklasse) werden unterstützend verwendet. Im dritten Schritt werden dann zunächst die Knoten final zugeordnet. Dabei wird geprüft, ob die Zahl der eingehenden und abgehenden Kanten übereinstimmt. Für jede dieser Kanten wird dann noch die Fahrtrichtung (Einbahnstraße oder beidseitig befahrbar) verglichen. Die Identifikation der Kanten Kandidaten erfolgt mittels kürzester-Wege-Berechnung zwischen den zugeordneten Knoten im Zielnetz unter Berücksichtigung semantischer Eigenschaften (z.B. Straßenklasse). Als Zuordnungskriterium wird die räumliche Nähe zwischen den beiden zu vergleichenden Polylinien herangezogen. Nah im Sinne der Autoren bedeutet hier, dass die Fläche zwischen den zu vergleichenden Kanten in beiden Karten minimal sein soll. Das Verfahren erkennt automatisch vollständige, unvollständige oder unmögliche Zuordnungen.

MUSTIÈRE und DEVOGELE (2008) haben das in MUSTIÈRE (2006) vorgestellte Verfahren nochmals verfeinert und als sogenannten *NetMatcher* implementiert und ausführlich getestet. Dabei legen sie auf zwei Prinzipien besonderen Wert: zum einen die Komplementarität von Knoten und Kanten, wonach für die Zuordnung sowohl die Knoten als auch die Kanten herangezogen werden sollen und wichtig sind, um zwei Karten-Netze einander zuzuordnen. Zum zweiten ist es der Grundsatz des Ausdünnens und Fokussierens: Der Suchraum für potentielle Zuordnungskandidaten soll in mehreren Schritten immer weiter verkleinert werden und unter den verbleibenden Kandidaten soll dann der am wahrscheinlichsten passende gefunden werden. Explizit gehen die Autoren auf die Schwierigkeit der Bestimmung der verwendeten Distanzschwellwerte ein, denn diese Schwellwerte haben entscheidenden Einfluss auf die Effektivität der Zuordnungsverfahren. Im vorliegenden Ansatz sollen die Schwellwerte für die Grobzuordnung deshalb immer zu pessimistisch angesetzt werden, um eine (genügend) große Gruppe von Kandidaten zu finden, aus denen dann im nächsten Schritt die korrekte Zuordnung ausgewählt wird. Die Autoren, wie auch MASCRET et al. (2006) empfehlen, das Verfahren um die Nutzung der Fréchet-Distanz zu erweitern, da sie häufig bessere Ergebnisse erzielt als die Hausdorff-Distanz. Des Weiteren sehen sie weiteres Verbesserungspotenzial in der Zusammführung der einzelnen Ähnlichkeitsmaße mittels Fuzzy Logik oder Evidenztheorie (Dempster-Shafer-Theorie).

In OLTEANU (2007a) und OLTEANU (2007b) wird ein Verfahren zur Zuordnung von geographischen Daten (Geodaten) präsentiert, bei dem die Auswahl der Zuordnungspaare aus den Gruppen der Kandidaten mit Hilfe von Mitteln der Evidenztheorie getroffen wird. Die Zuordnung selbst erfolgt durch Vergleich von Geometrie und Semantik. Als geometrische Ähnlichkeitsmaße kommen die Euklidische Distanz zu Bestimmung der räumlichen Nähe der Kandidaten zum Einsatz. Bei den semantischen Kriterien wird weiter zwischen den Objektnamen und den sonstigen Attributen und ihren Werten unterschieden. Die Ähnlichkeit der Objektnamen wird anhand der Levenshtein-Distanz bestimmt. Das Verfahren wurde mit topografischen Daten aus Frankreich getestet und lieferte gute Ergebnisse. RAIMOND und MUSTIÈRE (2008) haben das vorgenannte Verfahren weiterentwickelt und zur Anwendung auf Straßennetze adaptiert. Zunächst wurde der Satz von Zuordnungskriterien auf fünf Kriterien erweitert. Die geometrische Zuordnung erfolgt nun mittels Hausdorff-Distanz. Zusätzlich wird nun auch die Orientierung der Liniensegmente mit in die Zuordnung einbezogen. Zur Bestimmung der semantischen Ähnlichkeit wird nun die semantische Distanz nach WU und PALMER (1994) verwendet, die auf Basis einer Ontologie aus Informationen über Straßenklasse und -nummer berechnet wird. Neben dem Vergleich der Straßennamen beziehen die Autoren als letztes Kriterium noch den nachbarschaftlichen Kontext eines Zuordnungskandidaten durch Vergleich dessen angrenzender Kanten mit ein.

SIRIBA et al. (2011) stellen ein Zuordnungsverfahren vor, bei dem zunächst Kandidaten für eine Knotenzuordnung mittels Fangkreis aus dem Zieldatensatz ausgewählt werden. Durch die Auswertung weiterer geometrischer und semantischer Kriterien werden unwahrscheinliche Kandidaten ausgefiltert und die Menge der durch diese geometrische Abstandsfunktion gefundenen Kandidaten verkleinert. Beim schlussendlichen Pfad-Matching werden zunächst auf Grundlage der zuvor gewonnenen Knoten-

Zuordnungen, kürzeste Wege zwischen den einzelnen Knoten berechnet. Korrespondierende Linien-segmente werden nun durch Vergleich von geometrischen, semantischen und topologischen Kriterien identifiziert. Ein Schwerpunkt der Arbeit ist die Einpassung des Routenverlaufs in der Zielkarte nach der Zuordnung. Dies wird durch die nicht auszuschließenden geometrischen Differenzen zwischen den Datensätzen und somit möglichen Längenunterschieden zwischen den Routenverläufen in den beiden Karten motiviert. Die Längseinpassung der Route in der Zielkarte soll dabei anhand von Punkten mit großer Richtungsänderung erfolgen. Nach Annahme der Autoren besitzen solche Punkte eine große Bedeutung zur Beschreibung einer Route und stellen deshalb gute Stützpunkte für die Einpassung dar.

DALYOT et al. (2012, 2013) und DAHINDEN et al. (2013) beschreiben ein geometriebasiertes Verfahren zur Zuordnung von Vektordaten. Kernstück ist eine kleinste-Quadrate-Ausgleichung zur Minimierung der Lotabstände aller Punkte aus dem Quelldatensatz auf die Linien des Zieldatensatzes. Zur Suchraumbeschränkung wird ein Puffer um den Routenverlauf im Quelldatensatz angewendet, innerhalb dessen mögliche Zuordnungskandidaten zugelassen werden. Das Verfahren wurde beispielhaft implementiert und mit ALKIS- und ATKIS-Daten unterschiedlicher Maßstäbe getestet.

HACKELÖER et al. (2015) beschreiben ein iteratives Zuordnungsverfahren, welches mehrere geometrische und topologische Zuordnungsschritte auf verschiedenen hierarchischen Struktur-Ebenen des Straßennetzwerkes anwendet. Dadurch werden nach und nach mögliche Zuordnungslösungen gefunden, bewertet und verbessert. Das Verfahren kombiniert einen mehrstufigen Ansatz mit dem Buffer-Growing Verfahren. Es setzt sich aus vier Stufen zusammen: Datenvorverarbeitung, Knoten-Matching, Elementares (Kanten-) Matching und Kombiniertes Kanten-Matching. Zunächst werden die Phasen in der genannten Reihenfolge angewendet, um Übereinstimmungen zwischen elementaren Strukturen zu identifizieren (bottom-up). Danach folgt eine Prozessierung in umgekehrter Reihenfolge (top-down), um Zuordnungen zwischen zusammengesetzten Objekten auf ihre einzelnen Bestandteile aufgetrennt werden. Die Datenvorverarbeitung zielt darauf ab, die zwei zu vergleichenden Kartendatensätze zu harmonisieren und so mögliche systematische geometrische Differenzen sowie unterschiedliche Shape-Point Auflösungen zu bereinigen und die Datenmodelle anzugleichen, so dass direkte geometrische und topologische Vergleiche möglich sind. Die Punktzuordnung erfolgt dann mittels des *Exact Angular Index (EAI)* Verfahrens aus HACKELÖER et al. (2013). Ausgehend von den gefundenen Punktzuordnungen werden im nächsten Schritt, der elementaren Zuordnung, Korrespondenzen zwischen einfachen Straßenobjekten, die nur aus einer Kante bestehen, gesucht. Es werden Kantenzuordnungen gefunden, wenn deren Anfangs- und auch Endpunkte in der Punktzuordnung ein Zuordnungspaar darstellen. Der folgende kombinierte Zuordnungsschritt sucht nun in den Ergebnissen der einfachen Zuordnung nach zugeordneten Kanten, die die gleichen Anfangs- oder Endpunkte haben und bildet daraus zusammengesetzte Kantenobjekte. Diese zusammengesetzten Kantenobjekte werden nun wieder zwischen den beiden Datensätzen auf übereinstimmende Anfangs- und Endknoten geprüft und einander zugeordnet. Um dadurch möglicherweise entstehende Mehrdeutigkeiten zu erkennen bzw. zu vermeiden, d.h. Kanten, die in mehr als einem zusammengesetzten Kantenobjekt enthalten sind, wird jedem Zuordnungspaar ein Ähnlichkeitsmaß zugewiesen. In diesem Fall wird dazu das Längenverhältnis der beiden Polylinien verwendet. Um die Mehrdeutigkeiten zu bereinigen und eine eindeutige (bijektive) Zuordnung zu erhalten, wird die Zuordnung mit vertauschten Rollen (Quell-/Zielkarte) wiederholt.

In der anschließenden top-down Phase werden die gefundenen Zuordnungen zwischen den zusammengesetzten Kanten-Objekten genutzt, um Knoten aus einem Netz auf ihre entsprechende Position im anderen Netz zu projizieren. Daraus können dann elementare Zuordnungen auf Kanten- und Knoten-Ebene abgeleitet werden.

## 2.4.6 Zusammenfassung

In diesem Abschnitt konnte ein Überblick über die in der Geoinformatik und auch der Kartographie entwickelten und eingesetzten Zuordnungsverfahren gegeben werden. Diese werden häufig zur Zusam-

menführung ganzer Karten bzw. Datensätze verwendet, die teilweise in unterschiedlichen, aber auch in ähnlichen Maßstäben bzw. Auflösungen vorliegen. Dabei ist festzustellen, dass viele der Zuordnungsverfahren die Ähnlichkeit als Maß für die Übereinstimmung zweier zu vergleichender Objekte verwenden. Diese Ähnlichkeit wird anhand der geometrischen, topologischen oder semantischen Objekteigenschaften, oder einer Kombination aus ihnen, bestimmt. Die Suche nach korrespondierenden Objekten zweier Datensätze kann in einer Richtung erfolgen (uni-direktional). Oftmals wird aber zur Steigerung der Genauigkeit und Zuverlässigkeit der Zuordnung in beide Richtungen gesucht (bi-direktional).

Im Gegensatz zum Location Referencing liegen in den genannten Anwendungsszenarien die zuzuordnenden Datensätze zum gleichen Zeitpunkt in einem Auswertesystem (z.B. Geoinformationssystem, GIS) vor und sämtliche Eigenschaften der Kartenobjekte können direkt miteinander verglichen werden. Beim Location Referencing hingegen liegen die beiden Kartennetzwerke nur in räumlich getrennten Systemen vor. Es werden lediglich die in das gemeinsame Austausch-Datenschema übertragenen Objekteigenschaften der Ortsreferenz übermittelt und im Zielsystem für die Zuordnung verwendet. Dadurch sind beim Location Referencing weder bi-direktionale Zuordnungsverfahren anzuwenden, noch ist dabei eine iterative Ausdehnung des zugeordneten Gebietes von bereits zugeordneten Objekten auf deren unmittelbare Nachbarn (z.B. seed-point Ansatz oder ICP) möglich.

Insbesondere die verschiedenen aufgezeigten Ähnlichkeitsmaße und deren beispielhafte Implementierung und Verwendung als Zuordnungskriterien dienen jedoch als fruchtbare Anregungen für die weitere Verbesserung von Location Referencing Verfahren. Dazu sollen im folgenden Teil die Zuordnungsprinzipien auf Schema- bzw. Objekt-Ebene sowie eine Auswahl möglicher Ähnlichkeitsmaße zur Objektzuordnung vorgestellt und diskutiert werden.

## 2.5 Zuordnung auf Schema-Ebene

Nach DEVOGELE et al. (1998), SHEEREN et al. (2004), MUSTIÈRE (2006) ist es wichtig, vor der Zuordnung auf Daten-Ebene, die Korrespondenzen der beiden Datenbank Schemas zu identifizieren. Die Beziehungen zwischen den beiden Schemas werden durch sogenannte Inter-Schema-Korrespondenzen dargestellt. BISHR (1998) beschreibt die schematische Heterogenität, d.h. die möglichen Modellierungsunterschiede der realen Welt in den sie abbildenden Datenbanken (z.B. Karten), als zwischen verschiedenen Disziplinen abweichende Klassifikation und hierarchische Strukturierung von Kategorien der realen Welt. Da unterschiedliche Datenbank-Schemata unabhängig voneinander entwickelt werden, haben sie häufig eine unterschiedliche Struktur und verwenden verschiedene Terminologien. Diese Effekte können auftreten, wenn die Datenbank-Schemata aus unterschiedlichen Fachdisziplinen stammen, aber eben auch wenn sie innerhalb der selben Disziplin jedoch aber von unterschiedlichen Personen entwickelt wurden (RAHM und BERNSTEIN 2001). Dabei können Mehrdeutigkeiten bei der Zuordnung sowohl von Entitäten als auch Attributen auftreten, Attribute können fehlen oder nur implizit vorhanden sein, d.h. sie können aus anderen Attributen abgeleitet werden. Es wird auch die Möglichkeit beschrieben, dass Objekte einer Datenbank lediglich einem Attribut in einer anderen Datenbank entsprechen. Als mögliche Lösung des Problems wird die Einführung eines gemeinsamen Datenschemas für die verschiedenen, heterogenen Datenbankschemata vorgeschlagen. Dieser Schritt wird als *Schema Integration* bezeichnet. Sie erfolgt durch die Identifikation und Charakterisierung der Inter-Schema-Korrespondenzen. Sind diese gefunden, können sich entsprechende Elemente unter einem gemeinsamen, integrierten Schema vereinigt werden (RAHM und BERNSTEIN 2001). EDWARDS und SIMPSON (2002) merken dazu an, dass Übersetzungen oder Umrechnungen zwischen sich unterscheidenden Schemata so ungenau sein können, dass ein sinnvoller Vergleich der Objekte zwischen den beiden Karten sehr schwierig oder womöglich nicht mehr möglich ist.

CONRAD (2002) beschreibt die Technik der zusicherungs-basierten Schema-Integration, wobei unter Zusicherungen die zuvor angesprochenen Inter-Schema-Korrespondenzen verstanden werden. Es werden folgende Arten von Korrespondenzen unterschieden:

- Element-Korrespondenzen,
- Element-und-Attribut-Korrespondenzen,
- Pfad-Korrespondenzen.

*Element-Korrespondenzen* treten dabei zwischen Objekttypen aus zwei verschiedenen Schemata auf, *Element-und-Attribut-Korrespondenzen* erweitern diese um Korrespondenzen für die zugehörigen Attribute, *Pfad-Korrespondenzen* beschreiben schließlich Beziehungen zwischen Schema-Elementen als Alternative zur direkten Zuordnung von Objekttypen eines Schemas zu Objekttypen eines anderen Schemas, basierend auf dem Konzept der Referenzattribute. Für alle diese Korrespondenz-Arten ist die Definition von vier extensionalen Beziehungen möglich, die in Abbildung 2.9 dargestellt sind, siehe dazu u.a. auch DEVOGELE et al. (1998), SHEEREN et al. (2004), MUSTIÈRE (2006).

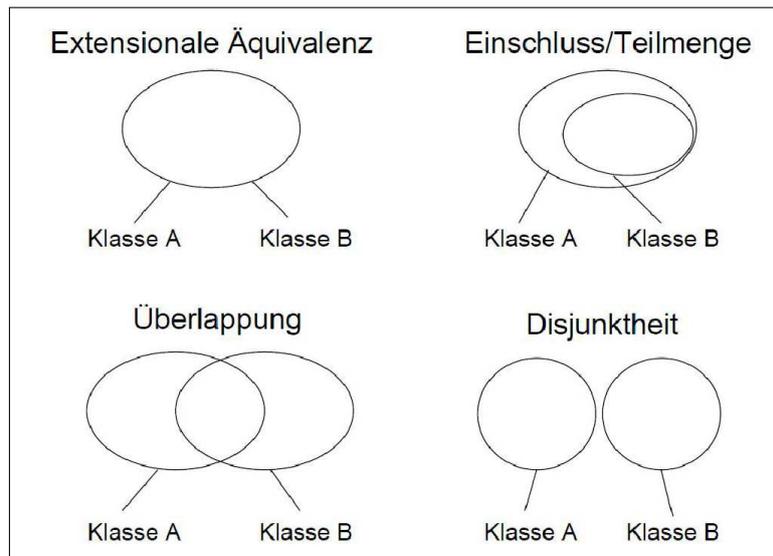


Abbildung 2.9: Die vier möglichen Extensionalen Beziehungen (aus CONRAD (2002))

## 2.6 Zuordnung auf Objekt-Ebene

Nach erfolgreicher Schema-Zuordnung kann laut DUNKARS (2003) und OLTEANU (2007b) eine Zuordnung auf Objekt oder auch Daten-Ebene erfolgen. Dabei lassen sich Zuordnungen durch Vergleich und Bewertung der *geometrischen*, *topologischen* und *semantischen* Eigenschaften der Objekte sowie Kombinationen daraus, unterscheiden (DEVOGELE et al. 1996, YUAN und TAO 1999, DOYTSHER und FILIN 2000, CECCONI 2003, VOLZ 2006b, ZHANG 2009, OLTEANU-RAIMOND et al. 2015). Geeignete Methoden zur Berechnung der Ähnlichkeit zweier zu vergleichender Objekte werden im Folgenden vorgestellt und ihre Eignung für eine Verwendung im Location Referencing diskutiert.

### 2.6.1 Geometrische Zuordnungsmaße

#### 2.6.1.1 Entfernung

Zur geometrischen Zuordnung von Objekten werden häufig sogenannte *Metriken* verwendet, also Abstandsfunktionen, die zwei Elementen des Raums einen nicht negativen reellen Wert zuordnen. Seien  $A$ ,  $B$  und  $C$  Kurven und  $d(A, B)$  eine Abstandsfunktion zwischen  $A$  und  $B$ , dann sind folgende metrischen Eigenschaften definiert (VELTKAMP 2001):

- Nicht-Negativität:  $d(A, B) \geq 0$ , für alle  $A$  und  $B$ ;
- Identität:  $d(A, A) = 0$ ;
- Gleichheit:  $d(A, B) = 0$  wenn  $A = B$ ;
- Symmetrie:  $d(A, B) = d(B, A)$ ;
- Dreiecks Ungleichheit:  $d(A, B) \leq d(A, C) + d(B, C)$ .

Abstandsfunktionen, welche die Identitäts-, Gleichheits- und Dreieck-Ungleichheitseigenschaft erfüllen, heißen *Metrik*. Abstandsfunktionen, welche nur die Eigenschaften der Identität und Dreieck-Ungleichheit erfüllen, heißen *Semimetrik* (VELTKAMP 2001).

Für lineare Objekte bietet der Vergleich der Entfernung eine erste und intuitive Möglichkeit zur Verifizierung der Zuordnung von homologen Objekten in zwei verschiedenen Datensätzen. Je näher die beiden Objekte einander sind, desto höher ist die Wahrscheinlichkeit, dass sie das gleiche Realweltobjekt repräsentieren (YUAN und TAO 1999). Drei Distanz-Funktionen werden häufig für unterschiedliche Anwendungsfälle verwendet: Die *Euklidische*-, *Hausdorff*- und *Frechét-Distanz* (ZHANG 2009).

### **$L_p$ -Distanz bzw. Minkowski-Distanz**

Viele formbasierte Ähnlichkeitsmaße basieren auf der sog.  $L_p$  Distanz zwischen zwei Punkten, sie wird auch *Minkowski Distanz* genannt. Für zwei Punkte  $x, y$  in  $\mathfrak{R}^k$  wird die  $L_p$  Distanz als  $L_p = (\sum_{i=0}^k |x_i - y_i|^p)^{1/p}$  definiert. Für  $p = 2$  wird die Distanz zur Euklidischen Distanz  $L_2$ . Für  $p = 1$  erhält man die *Manhattan, city block oder Taxicab Distanz*  $L_1$ . Die  $L_p$  Distanzen sind Metriken für alle  $p \geq 1$ , da für  $p < 1$  die Dreiecks Ungleichung nicht mehr gilt (VELTKAMP 2001).

### **Euklidische Distanz**

Die *Euklidische Distanz* wird verwendet, um Abstände zwischen zwei Punkten oder von Punkten zu einer Linie zu berechnen (YUAN und TAO 1999). Die Euklidische Distanz von einem Punkt  $p_i = (x_i, y_i)$  zu einem Punkt  $p_j = (x_j, y_j)$  berechnet sich wie folgt:

$$d_e(p_i, p_j) = [(x_j - x_i)^2 + (y_j - y_i)^2]^{1/2}. \quad (2.1)$$

Der Abstand eines Punktes  $P_0$  von einem Liniensegment  $l(p_i, p_j)$ , wird durch die Euklidische Distanz  $d_e$  zu seinem Lotfußpunkt auf dem Liniensegment bestimmt. Die Euklidische Distanz erfüllt folgende Eigenschaften der Nicht-Negativität, Gleichheit, Symmetrie und die Dreiecks Ungleichheit (MIN et al. 2007). Die Euklidische Distanz ist somit eine Metrik.

### **Hausdorff-Distanz**

Die *Hausdorff-Distanz* wird vor allem dazu verwendet, den Abstand von räumlichen Objekten zueinander zu bestimmen. Sie wurde 1918 vom deutschen Mathematiker Felix Hausdorff eingeführt und später nach ihm benannt (HAUSDORFF 1918). Da diese Objekte häufig aus mehreren Punkten bestehen und somit eine räumliche Ausdehnung besitzen, ist die Bestimmung des Abstandes zwischen ihnen nicht direkt möglich.

Seien  $A$  und  $B$  zwei Punktwolken, dann ist die Hausdorff-Distanz  $d_H$  zwischen  $A$  und  $B$  allgemein definiert als (RUCKLIDGE 1996):

$$d_H = \max(d_h(A, B), d_h(B, A)), \quad (2.2)$$

wobei  $d_h(A, B)$  und  $d_h(B, A)$  als gerichtete Hausdorff-Distanzen von  $A$  nach  $B$  bzw. von  $B$  nach  $A$  bezeichnet werden. Anschaulich beschreibt  $d_h(A, B)$  die größtmögliche Entfernung eines Punktes auf der äußeren Grenze von  $A$  zum nächstgelegenen Punkt in  $B$  und umgekehrt  $d_h(B, A)$  die größtmögliche Entfernung eines Punktes auf der äußeren Grenze von  $B$  zum nächstgelegenen Punkt in  $A$ . Die beiden gerichteten Hausdorff-Distanzen müssen nicht gleich sein (HANGOUËT 1995, VELTKAMP 2001) und

erfüllen daher auch nicht die Symmetrie-Eigenschaft (MIN et al. 2007). Daher wird die Hausdorff-Distanz  $D_H$  aus dem Maximum von  $d_1$  und  $d_2$  berechnet, siehe Gleichung 2.2. Die Hausdorff-Distanz ist sehr empfindlich gegenüber Ausreißern (VELTKAMP 2001). Lediglich ein weit außerhalb liegender Punkt kann die Hausdorff-Distanz eines Objektes bestehend aus vielen Punkten signifikant verfälschen.

Für geschlossene räumliche Objekte in GIS können die gerichteten Hausdorff-Distanzen wie folgt angegeben werden:

$$d_h(A,B) = \max_{p_a \in A} \left\{ \min_{p_b \in B} \{d(p_a, p_b)\} \right\}, \quad (2.3)$$

$$d_h(B,A) = \max_{p_b \in B} \left\{ \min_{p_a \in A} \{d(p_a, p_b)\} \right\}. \quad (2.4)$$

Die Hausdorff-Distanz erfüllt die selben Eigenschaften wie die Euklidische Distanz (MIN et al. 2007) und ist damit auch eine Metrik. Sie hat jedoch den Vorteil, dass sie den Abstand zwischen beliebigen räumlichen Objekten bestimmen kann (MIN et al. 2007). Aufgrund der Identitäts-Eigenschaft, sind Objekte genau dann geometrisch identisch, wenn ihre Hausdorff Distanz verschwindet (HANGOUËT 1995).

Abbildung 2.10 zeigt ein Beispiel für die Berechnung der gerichteten Hausdorff-Distanz  $d_h(A,B)$  von einer Punktwolke A zu einer Punktwolke B. Zunächst werden alle Distanzen von Punkt  $a_1$  zu allen Punkten in B berechnet und die kürzeste ( $d_{11}$ ) ausgewählt. Dann werden alle Distanzen von Punkt  $a_2$  zu allen Punkten in B berechnet und wiederum die kürzeste ausgewählt ( $d_{23}$ ). Von diesen beiden kürzesten Distanzen wird nun die längste ausgewählt, sie ist die gerichtete Hausdorff-Distanz von A nach B ( $d_{11}$ ). Man kann also sagen, ein beliebiger Punkt aus A hat maximal die Entfernung  $d_h(A,B) = d_{11}$  zu einem Punkt in B. Zur Berechnung dieser diskreten Form der Hausdorff-Distanz für linienhafte Datensätze

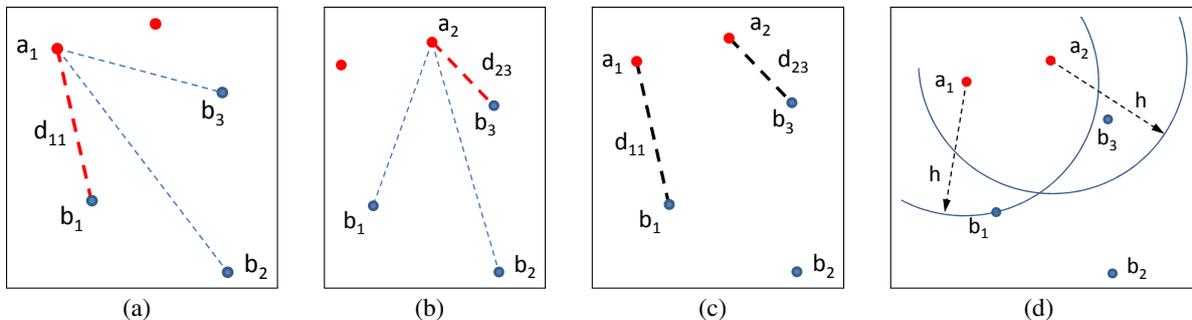


Abbildung 2.10: Beispiel Hausdorff-Distanz (nach GRÉGOIRE und BOUILLOT (1998))

gibt es zwei Möglichkeiten. Zum einen können nur die existierenden, formbeschreibenden Punkte als Repräsentation der Linien verwendet und zur Berechnung der Hausdorff-Distanz herangezogen werden, zum anderen können die Linien auch diskretisiert, also gleichabständige Zwischenpunkte eingerechnet, und die Hausdorff-Distanz mit Hilfe dieser Punkte berechnet werden.

### Fréchet-Distanz

Die *Fréchet-Distanz* ist eine Metrik, welche die maximale Distanz zwischen zwei orientierten oder gerichteten Kurven bzw. Linien beschreibt (DEVOGELE 2002). Sie geht auf die Arbeit des französischen Mathematikers Maurice René Fréchet zurück und wurde nach ihm benannt (FRÉCHET 1906, ALT und GODAU 1995). Sie wird sehr häufig wie folgt intuitiv illustriert: Ein Mann führt seinen Hund an der Leine spazieren. Der Mann geht entlang einer Kurve, der Hund entlang einer zweiten. Beide dürfen ihre Geschwindigkeit unabhängig voneinander verändern oder sogar stehen bleiben, aber nicht zurückgehen.

Dann ist die Fréchet-Distanz der beiden Kurven die minimale Länge der Leine, damit der Mann und sein Hund gerade noch entlang der beiden Kurven gehen können (ALT und GODAU 1995).

Die Fréchet-Distanz erfüllt die Eigenschaften der Symmetrie und Dreiecks-Ungleichheit, und kann somit als Metrik angesehen werden (ALT und GODAU 1995). Eine Polylinie kann durch einen geordneten Satz von Punkten beschrieben werden. Die Fréchet-Distanz berücksichtigt neben der Lage vor allem auch die Reihenfolge der Punkte der Kurven. Digitale Straßenkarten enthalten ausschließlich gerichtete Polylinien zur Repräsentation von linienhaften Straßenobjekten. Die Richtung ergibt sich dabei aus der gespeicherten Punktreihenfolge der Anfangs-, End- und gegebenenfalls auch Zwischenpunkte der Polylinien. Daher ist die Fréchet-Distanz geeigneter zum Vergleich zweier linienhafter Ortsreferenz-Objekte im Rahmen der vorliegenden Arbeit als die Hausdorff-Distanz, welche die Reihenfolge der Punkte einer Linie nicht berücksichtigt (EITER und MANNILA 1994, VELTKAMP und HAGEDOORN 2001, MASCRET et al. 2006).

Für den Vergleich von Polylinien mit diskreten Stützpunkten wurde von EITER und MANNILA (1994) die diskrete Fréchet-Distanz  $d_{dF}$  definiert. Sie bildet die obere Grenze für  $d_F$ , wobei der Unterschied zwischen diskreter und kontinuierlicher Form durch die längste Kante der Polylinie bestimmt wird.  $d_{dF}$  wird zwischen homologen Punkten der beiden Polylinien berechnet, also zwischen Punkten, die das selbe Realwelt-Objekt darstellen. Weisen die zu vergleichenden Polylinien allerdings zu große Unterschiede in ihrer geometrischen Modellierung auf, können auch mittels Resampling gleichabständige Zwischenpunkte eingerechnet, und diese dann zur Ermittlung der Fréchet-Distanz herangezogen werden (MASCRET et al. 2006). Im intuitiven Beispiel können sich sowohl der Mann als auch der Hund somit lediglich auf den Stützpunkten entlang der Polylinien aufhalten, aber nicht auf den Linien zwischen den Punkten. Es werden nur die Kombinationen ausgewertet, bei denen sich Mann und Hund entweder auf dem gegenwärtigen oder dem jeweils nächsten Punkt befinden (ARONOV et al. 2006). Nach EITER und MANNILA (1994) wird die diskrete Fréchet-Distanz wie folgt definiert: Seien  $P: [0, n] \rightarrow V$  und  $Q: [0, m] \rightarrow V$  zwei Polylinien und sei  $\sigma(P)$  die Sequenz  $(P(0), P(1), \dots, P(n))$  der Endpunkte der Liniensegmente von P, dann sind  $\sigma(P) = (u_1, \dots, u_p)$  und  $\sigma(Q) = (v_1, \dots, v_q)$  die Punktsequenzen der Polylinien  $P$  und  $Q$ . Die Verbindung (engl. coupling)  $\mathbf{L}$  zwischen  $P$  und  $Q$  ist dann eine Sequenz

$$(u_{a_1}, v_{b_1}), (u_{a_2}, v_{b_2}), \dots, (u_{a_m}, v_{b_m}) \quad (2.5)$$

eindeutiger Punktpaare von  $\sigma(P) \times \sigma(Q)$ . Die Länge  $\|\mathbf{L}\|$  der Verbindung  $\mathbf{L}$  ist die Länge des längsten Segmentes in  $\mathbf{L}$ , es ist,

$$\|\mathbf{L}\| = \max_{i=1, \dots, m} d(u_{a_i}, v_{b_i}). \quad (2.6)$$

Die diskrete Fréchet-Distanz  $d_{dF}(P, Q)$  der beiden gegebenen Polylinien  $P$  und  $Q$  wird dann wie folgt definiert:

$$d_{dF}(P, Q) = \min\{\|\mathbf{L}\|\}. \quad (2.7)$$

### Nächster Nachbar

Bei der Zuordnung von einander entsprechenden Knoten zweier digitaler Straßenkarten werden Distanzfunktionen häufig zur ersten Identifikation möglicher Zuordnungskandidaten verwendet (ROSEN und SAALFELD 1985). Mit dieser Methode der nächsten Nachbarn (engl. *Nearest Neighbour*) werden die dem jeweiligen Knoten in der Ausgangskarte geometrisch nächstgelegenen Knoten in der Zielkarte identifiziert. Der korrekte Zuordnungskandidat muss nicht notwendigerweise derjenige mit dem geringsten Abstandswert sein. Aufgrund der Kartenunterschiede können sich in nicht wenigen Fällen auch weiter entfernt liegende Knoten nach Überprüfung weiterer Zuordnungskriterien als wahrscheinlichste Zuordnung herausstellen.

### 2.6.1.2 Orientierung bzw. Richtungswinkel

Die Orientierung einer Linie oder eines Liniensegments kann in zweierlei Hinsicht als Zuordnungskriterium herangezogen werden: Erstens kann der Richtungswinkel als eigenständiges Kriterium zum Vergleich zweier Linien verwendet werden (YUAN und TAO 1999, RAIMOND und MUSTIÈRE 2008). Zum anderen kann das Richtungswinkel-Kriterium bei der Punktidentifikation auf mit einem Kreuzungspunkt verbundenen Linien angewandt werden.

ZHANG et al. (2005) betrachten als Orientierung einer linienhaften Ortsreferenz (Polylinie) den Richtungswinkel einer gedachten geraden Linie vom Anfangs- zum Endpunkt der Ortsreferenz. Berechnet man nun diesen Orientierungswinkel für die beiden zu vergleichenden linienhaften Ortsreferenzen, so ist ihre Differenz  $\Delta\theta = \theta_1 - \theta_2$  ein Maß für die Übereinstimmung der Ortsreferenzen bezüglich ihrer grundsätzlichen Orientierung.

### 2.6.1.3 Lage

In manchen Fällen kann die Zuordnung auch anhand der relativen Lage der zu vergleichenden Objekte erfolgen. Wenn die Karten keine allzu großen geometrischen Abweichungen aufweisen, können einander entsprechende Punkte aufgrund ihrer relativen geometrischen Nähe paarweise einander zugeordnet werden (ROSEN und SAALFELD 1985). Mit entsprechenden räumlichen Abfragen kann beispielsweise bestimmt werden, ob ein Punkt innerhalb oder außerhalb eines Polygons liegt (YUAN und TAO 1999) oder ob ein Polygon ganz oder teilweise mit einem anderen überlappt (VELTKAMP 2001). Für zwei Polygone  $A$  und  $B$  ist der Überlappungsbereich definiert als  $A \cap B$ . Als weiteres Zuordnungsmaß kann auch die *symmetrische Differenz*  $((A - B) \cup (B - A))$  berechnet werden, welche im Gegensatz zur Überlappung eine Metrik ist (VELTKAMP und HAGEDOORN 2001). Die relative Lage zweier Objekte zueinander wird hauptsächlich für Punkt-zu-Polygon und Polygon-zu-Polygon Zuordnungen verwendet (YUAN und TAO 1999).

### 2.6.1.4 Form

Betrachtet man nicht nur einzelne Liniensegmente, sondern ganze Polylinien mit ein oder optimalerweise mehreren Richtungsänderungen, so kann deren Form auch als Zuordnungsmaß verwendet werden. Ein weitverbreitetes Maß für die Form einer Kurve oder Polylinie ist die Richtungswinkelkurve oder englisch *Turning Function Distance*  $\Theta_A(s)$  aus der Kategorie der Signatur-Funktionen (engl. *signature functions*) (VELTKAMP und HAGEDOORN 2001, VELTKAMP 2001, ZHANG und LU 2004, VOLOTÃO et al. 2010). Sie beschreibt den Richtungswinkel der Kanten einer Polylinie als Funktion der Länge  $s$ . VELTKAMP (2001) beschreibt den Richtungswinkel mathematisch entgegen dem Uhrzeigersinn, sodass der Funktionswert bei einer Linksdrehung ansteigt und bei einer Rechtsdrehung abnimmt (siehe auch Abbildung 2.11).

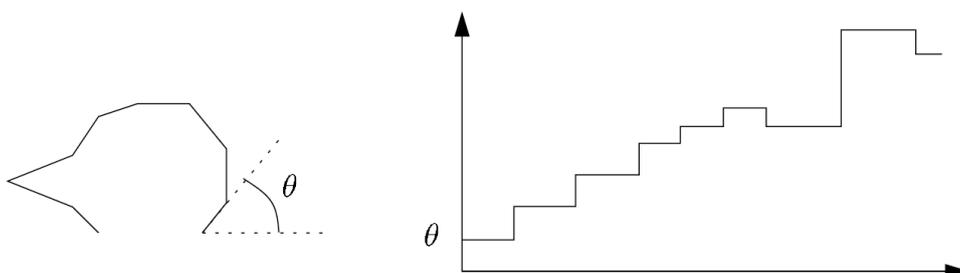


Abbildung 2.11: Richtungswinkelkurve (aus VELTKAMP (2001))

ARKIN et al. (1991) verwenden die Richtungswinkelfunktion für die Zuordnung von Polygonen. Seien  $A$  und  $B$  zwei Polygone mit ihren zugeordneten Richtungswinkelfunktionen  $\Theta_A(s)$  und  $\Theta_B(s)$ . Dann kann die Ähnlichkeit zwischen  $A$  und  $B$  mithilfe der Distanz zwischen  $\Theta_A(s)$  und  $\Theta_B(s)$  entsprechend der  $L_p$  Metrik ausgedrückt werden:

$$L_p(A, B) = \|\Theta_A - \Theta_B\|_p = \left( \int |\Theta_A - \Theta_B|^p ds \right)^{1/p}. \quad (2.8)$$

Hierbei ist  $\|\bullet\|$  die  $L_p$  Distanz, oder auch *Minkowski Distanz* genannt. Je größer der Funktionswert von  $L_p$ , desto kleiner wird die Ähnlichkeit zwischen den beiden zu vergleichenden Polygonen. In der Praxis wird der Parameter  $p$  oft mit dem Wert '1' verwendet, sodass  $L_p(A, B)$  die von den Kurven  $\Theta_A(s)$  und  $\Theta_B(s)$  umschlossene Fläche darstellt (ZHANG 2009), siehe auch Abbildung 2.12.

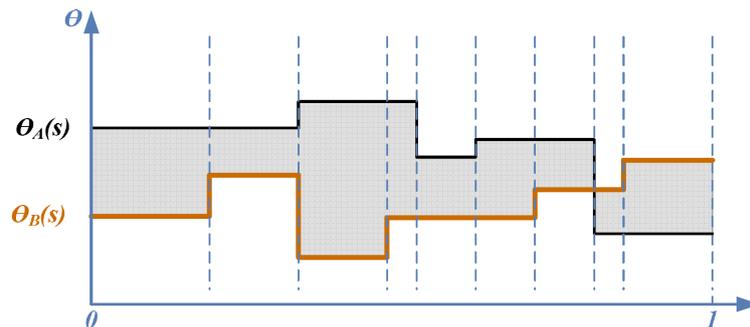


Abbildung 2.12: Funktionswerte von  $\Theta_A(s)$  und  $\Theta_B(s)$  (aus ZHANG (2009) und ARKIN et al. (1991))

Eine Abwandlung der Richtungswinkelfunktion ist die kumulierte Richtungswinkelfunktion (ZHANG und LU 2004). Sie ist ebenfalls eine Signatur-Funktion und ist definiert als die Summe der Richtungswinkel der Kanten einer Polylinie.

### 2.6.1.5 Ausdehnung

Bei linienhaften Objekten wird häufig deren Länge und auch die Euklidische Distanz zwischen dem Anfangs- und Endpunkt sowie bei Polygonen der Umfang oder die Fläche für die Zuordnung verwendet (ZHANG 2009).

Ein flächenbasiertes Zuordnungsmaß wird in ZHANG et al. (2005) beschrieben. Die Autoren setzen die von den beiden zu vergleichenden Polylinien  $P$  und  $Q$  beschriebene Fläche  $S_{P,Q}$  in Bezug zur Summe der Gesamtlänge der beiden Polylinien  $L_P$  und  $L_Q$ :

$$\bar{S} = \left| \frac{S_{P,Q}}{L_P + L_Q} \right|. \quad (2.9)$$

Je kleiner das resultierende Ähnlichkeitsmaß  $\bar{S}$ , desto ähnlicher sind sich die beiden zu vergleichenden Polylinien.

## 2.6.2 Topologische Zuordnungsmaße

Topologische Zuordnungsverfahren verwenden die Topologie, also die „nichtmetrischen räumlichen Verbindungen zwischen Objekten“ (BILL und ZEHNER 2001) zur Erkennung von übereinstimmenden Objekten in verschiedenen Karten. Die Topologie selbst ist ein Zweig der Mathematik, „der sich mit Eigenschaften geometrischer Gebilde befasst, die bei umkehrbar eindeutigen stetigen Abbildungen invariant bleiben. Die metrischen Verhältnisse spielen dabei keine Rolle; es kommt lediglich auf die

gegenseitige Lage und Anordnung der Gebilde im Raum an“ (BILL und ZEHNER 2001). Die Topologischen Beziehungen oder Relationen bleiben also erhalten, wenn Objekte beispielsweise verschoben, rotiert oder der Maßstab verändert wird (EGENHOFER 1991). In der Geoinformatik versteht man unter Topologie heute generell die räumlichen Beziehungen zwischen angrenzenden oder benachbarten Objekten (THEOBALD 2001). In einem zweidimensionalen Datensatz, beispielsweise einer digitalen Straßenkarte, sind dies insbesondere die Adjazenz, die Inzidenz sowie die Konnektivität zwischen einzelnen Kanten (ROSEN und SAALFELD 1985) und auch das Aneinandergrenzen von Maschen sowie daraus zusammengesetzte Beziehungen. Bei der Zuordnung zwischen linienhaften (Straßen-) Netzwerken spielen demnach vor allem die Beziehungen der Kanten und mit ihnen die sie definierenden Start- und Endknoten eine große Rolle (ZHANG 2009). Nach EGENHOFER und MARK (1995) ist die Topologie die höchstklassige Information zur Beschreibung räumlicher Zusammenhänge. Die metrischen Informationen wie beispielsweise Distanz oder Form werden demnach lediglich als Verfeinerung der topologischen Zusammenhänge genutzt, ganz nach dem Motto:

„*Topology Matters, Metric Refines*“.

WALTER (1997) schlägt zur Bestimmung der Ähnlichkeit die Verwendung der Relationen zwischen den Elementen (z.B. bei Daten mit Raumbezug: topologische Relationen) zusätzlich zur Ähnlichkeit der Attribute (merkmalsbasierte Zuordnung) vor. Die Zuordnung mittels topologischer Beziehungen zwischen den Objekten wird daher auch als *Relationale Zuordnung* bezeichnet (VOSSELMANN 1992, DOYTSHER und FILIN 2000).

In STIGMAR (2005) wird die Signifikanz der topologischen Beziehungen zwischen den einzelnen Objekten hervorgehoben. Konkret wird in einem ersten Schritt versucht, Objekte mittels Suchradius und Richtungswinkelbetrachtung einander zuzuordnen. Objekte, die durch dieses Matching nicht zugeordnet werden konnten, können jedoch durch ihre mögliche topologische Verbindung zu bereits zugeordneten Elementen in einem zweiten Schritt dennoch zugeordnet werden.

### 2.6.2.1 Adjazenz und Inzidenz

Nach BILL und ZEHNER (2001) ist die *Adjazenz* ein Grundbegriff aus der Graphentheorie und bezeichnet das Aneinandergrenzen oder Berühren gleichartiger Elemente. Zwei Knoten sind z.B. adjazent, wenn sie über eine Kante verbunden sind. Zwei Kanten sind adjazent, wenn sie in einem gemeinsamen Knoten enden. Die Adjazenz bezeichnet also die Beziehungen zwischen *gleichartigen* Elementen eines Graphen. Die *Inzidenz* bezeichnet nach BILL und ZEHNER (2001) das „Ineinander fallen“ oder das „Ineinander verschachtelt sein“ der Elemente eines Graphen. So sind die Anfangs- und Endknoten mit der sie verbindenden Kante inzident, ebenso wie alle von einem Knoten abgehenden Kanten mit dem Knoten inzident sind (KRUMKE und NOLTEMEIER 2012). Die Inzidenz bezeichnet also die Beziehungen zwischen *verschiedenartigen* Elementen eines Graphen.

Bei der topologischen Zuordnung werden die Struktur oder die topologischen Beziehungen zwischen den verschiedenen Objekten für die Zuordnung verwendet. Wenn zwei dieser Beziehungen übereinstimmen, kann diese Übereinstimmung genutzt werden, um homologe Objekte zu identifizieren, die durch die Beziehungen verbunden sind (DEVOGELE et al. 1996). Die topologischen Beziehungen zu benachbarten Objekten werden häufig auch als *geographischer Kontext* bezeichnet (SAMAL et al. 2004, ZHANG und MENG 2008, ZHANG 2009, KIM et al. 2010)

### 2.6.2.2 Valenz

Bei der Zuordnung von linienhaften Netzwerken wird üblicherweise die topologische Anordnung der Knoten und Kanten verwendet (ZHANG 2009). Die Knoten können dabei durch ihre *Valenz* charakterisiert werden, d.h. durch die Anzahl der mit ihnen verbundenen Kanten (WALTER und FRITSCH 1999, ZHANG et al. 2005, VOLZ 2006b, MUSTIÈRE und DEVOGELE 2008). Sie wird oft auch als Knoten-Index (ROSEN und SAALFELD 1985) bezeichnet. Der Knoten-Index ist eine diskrete Maßzahl, da er nur positiv ganzzahlige Werte annehmen kann. Ein Index von 1 bezeichnet dabei den Endpunkt einer Sackgasse, ein Index-Wert von 2 einen topologisch nicht relevanten Zwischenpunkt, oft auch als *Shape Point* bezeichnet. Erst ein Knoten-Index von 3 oder höher bezeichnet echte Kreuzungs-Knoten, an denen 3 oder mehr Kanten abgehen. Diese sind insbesondere für die Zuordnung von Knoten in verschiedenen Karten geeignet.

### 2.6.2.3 Topologische Bäume

In WARTENBERG (2008) wird die Verwendung von topologischen Bäumen bzw. Baumstrukturen als Muster zur Wiedererkennung von Ortsreferenzen in der Zielkarte vorgeschlagen. Eine Baumstruktur repräsentiert einen Netzausschnitt der Umgebung der eigentlich zu codierenden Ortsreferenz in der Quellkarte. Zur Einpassung in das Zielnetz wird der Baum nun mit Bäumen aus der Zielkarte verglichen. In der Graphentheorie ist dieses Problem unter Graph-Isomorphismus bzw. auch Subgraph-Isomorphismus bekannt. Zwei Graphen heißen Isomorph, wenn sie gleiche topologischen Eigenschaften (Adjazenz und Inzidenz) haben (BILL 2010). Geometrisch können die Graphen dabei durchaus Unterschiede aufweisen. Aufgrund von Kartenunterschieden werden die zu vergleichenden Bäume mit hoher Wahrscheinlichkeit nicht identisch sein, vielmehr wird ein gemeinsamer Sub-Baum der beiden zu vergleichenden Bäume gesucht. Dieses Zuordnungsproblem kann mit einem Algorithmus zum allgemeinen Sub-Baum Isomorphismus (engl. *common subtree isomorphism*) gelöst werden.

## 2.6.3 Gemischte geometrische und topologische Zuordnungsmaße

### 2.6.3.1 Spider Function

Neben der Anzahl der in einem Knoten zusammenlaufenden Kanten kann als weiteres Zuordnungskriterium auch der jeweilige Richtungswinkel, unter dem die Kanten den Knoten verlassen, herangezogen werden. Die Richtungen der Kanten bilden dabei ein für jeden Knoten typisches Strahlenmuster. Die Richtungswinkel werden gewöhnlich als kontinuierliche Zahlen aus den Koordinaten der Anfangs- und Endknoten der jeweiligen Kanten berechnet. Durch die Kartenunterschiede zwischen Ausgangs- und Zielkarte ergeben sich jedoch fast zwangsläufig auch Unterschiede in den Richtungswinkelmustern für den jeweiligen Knoten. Zur Vereinfachung des Vergleichs der Richtungswinkelmuster kann die sogenannte *Spider Function* verwendet werden (ROSEN und SAALFELD 1985, SAALFELD 1988). Dabei wird der Vollkreis in eine feste Anzahl von Sektoren eingeteilt, wie beispielhaft in Abbildung 2.13 dargestellt. Jeder Sektor erhält dabei einen Bezeichner (im angesprochenen Beispiel eine hexadezimale Zahl). Der jeweilige Richtungswinkel wird fortan nicht mehr durch einen kontinuierlichen Winkelbetrag sondern nur durch die Belegung des entsprechenden Sektors repräsentiert. Dadurch wird aus der kontinuierlichen Beschreibung der Richtungswinkel eine diskrete Belegung von Sektoren, die sehr einfach automatisiert verglichen werden kann.

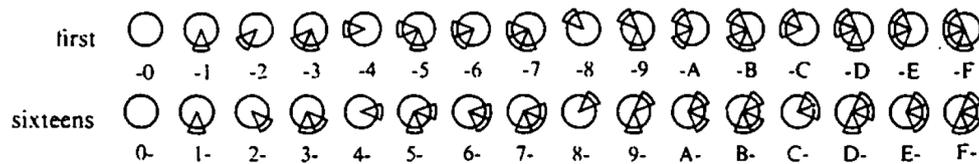


Abbildung 2.13: Hexadezimale und Sektoren Muster der *Spider Function* (aus SAALFELD (1988))

### 2.6.3.2 Exact Angular Index

HACKELÖER et al. (2013) wählen für die Zuordnung ein kombiniertes Verfahren, dass die topologische Valenz eines Knotens (also die Anzahl der verbundenen Kanten) als auch die geometrische Winkeldifferenz der ausgehenden Kantenpaare in beiden Karten berücksichtigt. Fehlende Kanten werden mit einem hohen Strafzuschlag bei der Berechnung der Summe der Winkeldifferenzen für einen Knoten mit einbezogen. Das Verfahren kann auch noch um die Einbeziehung einer geometrischen Abstandsfunktion (z.B. Euklidischer Distanz) erweitert werden.

### 2.6.4 Thematische bzw. semantische Zuordnungsmaße

Die Semantik ist die Lehre von der Sinnhaftigkeit oder Bedeutung von Wörtern oder Sätzen, deshalb wird sie auch Bedeutungslehre genannt (BILL und ZEHNER 2001). Im Bereich Geoinformatik versteht man unter der Semantik die thematischen Objekteigenschaften, also beschreibende Merkmale, Sachdaten oder Attribute (BILL 2010). Entgegen dieser aus praktischen Erwägungen getätigten Beschränkung auf den Vergleich von Attributwerten (numerische Werte oder Zeichenketten), umfasst die Semantik eigentlich auch den Vergleich ihrer *Bedeutung*, also z.B. ein Tal mit einem Gipfel oder eine Autobahn mit einer Nebenstraße (OLTEANU-RAIMOND et al. 2015). Die semantische Zuordnung ordnet Objekte einander entsprechend der Ähnlichkeit ihrer (semantischen) Attribute und deren unterschiedlichen Attributwerte zu (DEVOGELE 2002, COHEN 2000). Damit die Attribute von Objekten verschiedener Datensätze vergleichen bzw. einander zuordnen zu können, müssen sie im einfachsten Fall die gleichen Attributfelder mit der selben Definition besitzen. Andernfalls müssen die Datensätze gleiche oder entsprechende Datenfelder enthalten und die Semantik der beiden Datensätze muss bekannt sein, um eine Zuordnungsvorschrift oder auch Zuordnungsmatrix zwischen den sich entsprechenden Attributen der beiden Datensätze erstellen zu können (YUAN und TAO 1999). Aufgrund möglicher Unterschiede in den Semantiken der Datensätze müssen auftretende Unterschiede beim Vergleich der semantischen Attribute nicht notwendigerweise auch bedeuten, dass die verglichenen Objekte einander nicht zuzuordnen sind. Deshalb sollte die Semantik nicht als alleiniges Zuordnungsverfahren, sondern immer zusammen mit geometrischen und/oder topologischen Eigenschaften ausgewertet werden (WALTER und FRITSCH 1999, ZHANG et al. 2005, VOLZ 2006b, HACKELÖER et al. 2014). ZHANG (2009) hält semantische Eigenschaften für sehr wertvoll für die Zuordnung von Kartenobjekten.

Nach KASHYAP und SHETH (1996) lässt sich die semantische Ähnlichkeit zweier Objekte aus verschiedenen Datensätzen in vier Klassen ordnen:

- *Semantische Äquivalenz*: Zwei Attribute sind semantisch äquivalent, wenn sie das selbe Objekt der realen Welt repräsentieren, also eine exakte 1-1 Zuordnung der Attributwerte möglich ist.
- *Semantische Beziehung*: Zwei Attribute haben eine semantische Beziehung, wenn zumindest eine teilweise Zuordnung ihrer Attributwerte, eine Generalisierung oder eine Aggregation zwischen ihren Wertebereichen möglich ist.

- *Semantische Relevanz*: Zwei Attribute gelten als semantisch relevant, wenn eine Verbindung mithilfe einer Abstraktion in einem bestimmten Kontext zwischen ihnen hergestellt werden kann. Demnach ist der Begriff der semantischen Relevanz zweier Attribute kontextabhängig.
- *Semantische Ähnlichkeit*: Wann immer zwei Attribute unter keiner Abstraktion und in keinem Kontext einander zugeordnet werden können, sie aber dennoch in ihrem jeweiligen Kontext die selben Rollen innehaben und die Kontext-Definitionen zueinander kohärent sind, dann gelten sie als einander semantisch ähnlich.

Im Optimalfall haben also zwei zu vergleichende Datensätze semantisch äquivalente Attribute und diese sind vollständig mit Werten gefüllt. In diesem Fall sind die Entsprechungen zwischen den Attributen klar und die Zuordnung kann durch Vergleich der jeweiligen Attributwerte sehr effizient und einfach erfolgen, und die einander entsprechenden (homologen) Objekte aus den beiden Datensätzen einander zugeordnet werden. Fallen die Attribute in die zweite und dritte der oben genannten Kategorien können sie noch als eine Art Filter im Zuordnungsprozess verwendet werden, um zumindest diejenigen Objekte zu identifizieren, die nicht als potentielle Zuordnungs-Kandidaten angesehen werden sollten. Dazu müssen im Vorfeld die beiden Datensätze analysiert werden, um die semantisch ähnlichen Objekttypen zu identifizieren.

#### 2.6.4.1 Unterscheidung objektiver und subjektiver Attribute

Für eine Zuordnung anhand semantischer Attribute unterscheidet ZHANG (2009) zwischen *objektiven* semantischen Attributen und *subjektiven* semantischen Attributen.

##### Zuordnung mittels objektiven semantischen Attributen

Zur Kategorie der *objektiven* semantischen Attribute werden alle Objektattribute gezählt, die inhärente und objektiv messbare Eigenschaften eines räumlichen Objektes beschreiben. Demnach sind sie nicht abhängig von der jeweiligen Datenstruktur oder Datenspezifikation und eine 1-1 Zuordnung ist möglich. Im Bereich der Zuordnung von Elementen aus digitalen Straßenkarten gehören zu den objektiven semantischen Attributen beispielsweise:

- Straßename
- Straßenbreite
- Anzahl der Fahrspuren
- Fahrtrichtung

##### Zuordnung mittels subjektiven semantischen Attributen

*Subjektive* semantische Attribute repräsentieren entweder unscharfe Objekteigenschaften oder künstlich eingeführte Eigenschaften für bestimmte Anwendungen. Im Bereich der digitalen Straßenkarten sind beispielsweise der Straßentyp und die Straßenklasse als unscharfe Objekteigenschaften anzusehen. Da es in diesen Fällen eigentlich immer gewisse Zusammenhänge zwischen den Definitionen in den einzelnen Datensätzen gibt, sie aber meist nicht konsistent sind, müssen die semantischen Zuordnungen sehr viel komplizierter modelliert und analysiert werden, als das für die objektiven Attribute der Fall ist. Bei künstlich eingeführten Objekteigenschaften, wie zum Beispiel einer Objekt-ID, gibt es in den seltensten Fällen einen Zusammenhang zwischen den verschiedenen Datensätzen und somit ist eine Zuordnung auf Basis dieser Attribute äußerst schwierig und nur selten möglich.

#### 2.6.4.2 Vergleich von Zeichenketten

Beim Vergleich von Parametern bzw. Parameterwerten nimmt der Straßename eine Sonderrolle ein, da er vom Typ Zeichenkette (engl. String) ist und nicht so einfach wie die anderen numerischen Werte verglichen werden kann. Die Schwierigkeit des Vergleichs ist dabei nicht allein auf den Datentyp zurückzuführen, es gibt in jeder Programmierumgebung effiziente und komfortable Möglichkeiten, Strings

miteinander zu vergleichen. Es ist vielmehr die mögliche Unschärfe in der Schreibweise der Straßennamen, die einen Vergleich schwer macht. Vor allem, wenn man den Vergleich robust gegenüber möglichen Unterschieden in der Schreibweise oder kleinen Schreibfehlern machen will. Dann soll der Vergleich bzw. die Zuordnung trotz dieser kleinen Unterschiede zwischen den beiden Namen möglich sein. In der Praxis gibt es sehr viele verschiedene Verfahren und Ansätze zum Vergleich von Zeichenketten, COHEN et al. (2003) geben eine sehr gute Übersicht. Aus dem Bereich der Kartenzuordnungen gibt ZHANG (2009) ein zweistufiges Verfahren an: zunächst werden die Straßennamen normalisiert und in einem zweiten Schritt die Ähnlichkeit anhand der *Levenshtein Distanz* (LEVENSHTEIN 1966) berechnet. Die beiden Schritte werden im Folgenden näher betrachtet:

#### *Schritt 1 - Normalisierung der Straßennamen*

In einem ersten Schritt werden zum einen Länder/Sprach-spezifische Buchstaben durch solche aus dem englischen Alphabet ersetzt. Beispielsweise werden die deutschen Umlaute *ä, ö, ü* ersetzt durch *ae, oe, ue* oder das *ß* durch *ss*. Dann werden allgemein gebräuchliche Abkürzungen wieder durch die jeweilige vollständige Schreibweise ersetzt, also z.B. *'str'* durch *'strasse'* sowie alle Leerzeichen oder Bindestriche gelöscht. In einem letzten Vorbereitungsschritt werden alle Buchstaben in Großbuchstaben konvertiert (ZHANG 2009).

#### *Schritt 2 - Vergleich von Straßennamen mithilfe der Levenshtein Distanz*

Die Levenshtein Distanz, auch *Edit-Distanz* genannt, ist ein Maß für die Ähnlichkeit zweier Zeichenketten (Strings). Sie ist definiert als die minimale Anzahl von Operationen, die dazu benötigt werden, eine Zeichenkette in die andere zu überführen. Die möglichen Operationen umfassen das Löschen, das Hinzufügen oder das Ersetzen von Zeichen. Daraus folgt, je größer die Levenshtein Distanz, desto weniger ähnlich sind sich die beiden zu vergleichenden Zeichenketten (LEVENSHTEIN 1966, NAVARRO 2001, SAMAL et al. 2004, MUSTIÈRE und DEVOGELE 2008, ZHANG 2009, OLTEANU-RAIMOND et al. 2015).

Eine andere Möglichkeit des Vergleichs von Zeichenketten ist der sogenannte *Soundex-Algorithmus* (HALL und DOWLING 1980). Dieser verwendet eine phonetische Repräsentation der Zeichenketten und ist somit unempfindlicher gegen falsche Schreibweisen.

## 2.7 Bewertungsfunktionen

Nach OLTEANU-RAIMOND et al. (2015) gibt es drei Ansätze zur Auswahl des am besten geeigneten Zuordnungskandidaten:

Beim **sequentiellen Ansatz** wird ein Zuordnungskriterium nach dem anderen anwenden angewendet, um die Auswahl nach und nach einzugrenzen und schlussendlich den besten Kandidaten ausgewählt zu haben. Dieser Ansatz wird beispielsweise von DEVOGELE et al. (1998) implementiert. SAALFELD (1988) und COBB et al. (1998) verfeinern den Ansatz und beginnen mit dem strengsten Auswahlkriterium und lassen dann immer schwächere Kriterien folgen, bis das Zuordnungspaar gefunden ist.

Für den **kombinatorischen Ansatz** wird jeder Kandidat mit Hilfe einer Summenfunktion bewertet. Diese gewichtet die einzelnen Bewertungskriterien und führt sie in einer einzelnen Größe zusammen. Der Kandidat mit der besten Bewertung wird für die Zuordnung ausgewählt, z.B. in DUNKARS (2003), SAMAL et al. (2004), VOLZ (2006b).

Schließlich gibt es noch einen **gemischten Ansatz**, der die beiden vorgenannten Ansätze verbindet. So werden schlecht passende oder einzelnen Bedingungen widersprechende Kandidaten in einer Vorauswahl aussortiert und die verbleibenden dann einer Bewertung und Auswahl zugeführt (ZHANG 2009).

### 2.7.1 Kosten- vs. Leistungsfunktion

Grundsätzlich bestehen zwei Möglichkeiten, Bewertungsfunktionen zu gestalten. Eine *Kostenfunktion* summiert die Kosten, welche unter verschiedenen Gesichtspunkten entstehen. Gesucht wird dabei nach demjenigen Kandidaten, der die *geringsten* und im Optimalfall gar keine Kosten verursacht. Kostenfunktionen werden zum Beispiel bei der Suche nach kürzesten Wegen in Straßennetzwerken eingesetzt (vgl. BASELAU (2010)). Als Kosten wird hier den Kanten des Netzwerks ihre jeweilige Länge zugeordnet. Beim Vergleich mehrerer Routenalternativen wird diejenige als kürzester Weg ausgewählt, deren Kosten (Länge) minimal sind.

Eine andere Möglichkeit der Modellierung von Bewertungsfunktionen bietet die *Leistungsfunktion*. Hier wird im Gegensatz zur Kostenfunktion die Leistung eines Objekts im Hinblick auf verschiedene mögliche Kriterien gewürdigt. Gesucht wird also dasjenige Objekt, dessen Leistungsfunktionswert *maximal* ist.

Nach WALTER (1997) ist zur Lösung von Zuordnungsproblemen die Leistungsfunktion einer Kostenfunktion vorzuziehen, da bei letzterer im Falle einer nicht möglichen Zuordnung eine falsche Entscheidung getroffen werden kann. Bei der Bewertung mittels einer Kostenfunktion entstehen nämlich für die nicht mögliche Zuordnung keine Kosten, sie erhält somit das kleinste Gewicht. Daher wird sie bei der Suche nach dem Kandidaten mit den geringsten Kosten allen anderen, tatsächlich erfolgten Zuordnungen vorgezogen. Bei Verwendung einer Leistungsfunktion erbringt eine nicht mögliche Zuordnung keine Leistung und unterliegt somit allen anderen, tatsächlich möglichen Zuordnungen. Diese Aussage des Autors ist sicherlich für sich genommen korrekt, jedoch erscheint die Behandlung von nicht möglichen Zuordnungen mittels einer Kostenfunktion beherrschbar. So könnten beispielsweise für den Fall, dass kein Zuordnungspartner im Zielnetz gefunden werden kann, vorher festzulegende Strafkosten fällig werden.

### 2.7.2 Gesamt-Ähnlichkeitsmaß

Leistungsfunktionen werden häufig im Bereich der Kartenzuordnungen oder dem Location Referencing eingesetzt. WALTER (1997) verwendet zur Bewertung der einzelnen Zuordnungspaare die *gegenseitige Information* der beiden zu vergleichenden Elemente. Die gegenseitige Information ist die Differenz der Selbstinformation und der bedingten Information. Die Leistungsfunktion wird nach dem Verfahren von VOSSELMANN (1992) berechnet und kommt ohne datenabhängige Werte wie z.B. Schwellwerte oder Gewichtsfaktoren aus. DUNKARS (2003) berechnet die *gewichtete Euklidische Distanz* zur Zusammenführung aller betrachteter Zuordnungskriterien zu einem Zuordnungs- bzw. Ähnlichkeitsmaß. SAMAL et al. (2004) verwendet ein gewichtetes Mittel aus den Ähnlichkeitswerten aller Objekteigenschaften. ZHANG et al. (2005), ZHANG und MENG (2007) und ZHANG (2009) fassen die einzelnen berechneten Ähnlichkeitsmaße basierend auf Geometrie und Topologie zu einem Gesamtmaß der Ähnlichkeit eines Zuordnungspaars zusammen. Die OpenLR Musterimplementierung TOMTOM (2014) verwendet einen gemischten Ansatz nach der Klassifikation in Abschnitt 2.7. Zunächst wird mittels Vorauswahl eine Suchraumbeschränkung durchgeführt (siehe Kapitel 3.4.2.2). Für die verbleibenden Kandidaten wird dann eine Leistungsfunktion zur Identifikation des am besten passenden Zuordnungskandidaten berechnet, siehe Kapitel 3.4.1 bzw. 3.4.2.2.

VOLZ (2006a) und VOLZ (2006b) beschreiben die Zusammenführung verschiedener Ähnlichkeitsmaße  $IM_i$  für Kartenobjekte zu einem Gesamtmaß  $GM$  für die Ähnlichkeit einer Zuordnung bzw. der Ähnlichkeit zweier Objekt-Repräsentationen mit Hilfe einer gewichteten Summe. Jedem Einzelmaß  $IM_i$  wird dabei ein Gewicht  $\omega_i$  zugeordnet. Der individuelle Beitrag jedes einzelnen Ähnlichkeitsmaßes zum Gesamtmaß kann so anwendungsabhängig reguliert werden:

$$GM = \sum_{i=0}^k IM_i \cdot \omega_i. \quad (2.10)$$

## 2.8 Zusammenfassung

Nach einer Einführung in die Abbildung der realen Welt mit Hilfe von Karten wurde gezeigt, dass durch die dafür notwendige Erfassung und Modellierung Unterschiede zwischen Karten entstehenden, die das gleiche Gebiet abbilden. Danach wurden Verfahren vorgestellt, mit Hilfe derer korrespondierende Elemente, die die selben Objekte der realen Welt darstellen, in den Karten einander zugeordnet werden können. Daran anschließend folgt eine Zusammenstellung von geometrischen, topologischen und semantischen Objekteigenschaften, die für eine Bewertung der Ähnlichkeit einer Gruppe von vorausgewählten Kandidaten mit dem gesuchten Objekt herangezogen werden können. Für die weitere Arbeit erscheint die Verwendung von topologischen und geometrischen Zuordnungskriterien sinnvoll, da für diese Eigenschaften die Entstehung von Mehrdeutigkeiten bei den Übergängen zwischen den einzelnen Datenbank-Schemata nicht zu erwarten sind. Auch die Verwendung semantischer Eigenschaften für die Zuordnung, wie z.B. dem Straßennamen ist theoretisch interessant. Durch die angesprochene Vielzahl von Schema-Übergängen entstehen aber große Unsicherheiten, die letztlich negative Auswirkungen auf das Zuordnungsergebnis erwarten lassen.

Um unter allen bewerteten Zuordnungspaaren eine Zuordnungsentscheidung treffen zu können, müssen die einzelnen Ähnlichkeitsmaße abschließend zu einem Gesamtmaß zusammengeführt werden. Dieses Maß kann, wie gesehen, als Kosten- oder Leistungsfunktion berechnet werden. Auf seiner Grundlage erfolgt letztlich die Entscheidung für das Zuordnungspaar mit der höchsten Wahrscheinlichkeit der Korrespondenz. Für die vorliegende Arbeit erscheint die Verwendung einer Leistungsfunktion für die Zusammenführung der einzeln ermittelten Ähnlichkeitsparameter und das Treffen der Zuordnungsentscheidung zweckmäßig. Aus den einzelnen Ähnlichkeitsmaßen können Leistungsbeiträge ermittelt werden, sodass eine höhere Leistung auch eine bessere Zuordnung bedingt.

## 3 Location Referencing

### 3.1 Statisches und dynamisches Location Referencing

Als Ortsreferenzierung (engl. Location Referencing) werden im Bereich der Verkehrstelematik ganz allgemein Verfahren zur Referenzierung von Objekten bezeichnet. Diese Funktionalität wird in der Praxis hauptsächlich von kartenbasierten Telematiksystemen verwendet, welche Informationen mit anderen Systemen über Kommunikationseinrichtungen austauschen. Dabei werden Informationen ausgetauscht, die ihren Ortsbezug relativ zum Straßennetz bzw. den Elementen in einer Straßenkarte haben. Um diese Objekte nun in anderen kartenbasierten Systemen verorten zu können, muss vor einer Übertragung die Georeferenzierung verallgemeinert werden. Die Georeferenzierung bleibt dabei kartenbasiert, wird aber unabhängig von der spezifischen Ausgangskarte definiert. Diese verallgemeinerte Georeferenz wird *Location Reference*, oder auch *Ortsreferenz* genannt.

Abhängig von der Art der Referenzierung unterscheiden WEVERS und HENDRIKS (2006), WEVERS (2012) zwischen dem *statischen* und dem *dynamischen* Location Referencing. Einen guten Überblick über verfügbare statische und dynamische Ortsreferenzierungssysteme bietet ein von der Bundesanstalt für Straßenwesen (BASt) in Auftrag gegebenes Gutachten (KOCHS et al. 2011). Anfänglich und teilweise bis in die heutige Zeit wurden zur Referenzierung der zu übermittelnden Informationen die einzelnen Straßenabschnitte in einem landesweit einheitlichen und zentralen Verzeichnis referenziert. Dadurch erhielt jeder spezifizierte Abschnitt des in diesem Verzeichnis abgebildeten Straßennetzwerkes einen anwendungs- und kartenübergreifend einheitlichen Identifikator. Dieses Verfahren wurde als *statisches Location Referencing* bezeichnet. Für die Referenzierung von Verkehrsmeldungen wird bis heute der digitale Verkehrskanal (engl. Traffic Message Channel, TMC) im Radio-Daten-System (RDS) verwendet (DIN EN ISO 14819-3 2014). Eine TMC-Verkehrsmeldung enthält Angaben zum Ort und zur Art des Ereignisses in codierter Form. Für Deutschland führt die Bundesanstalt für Straßenwesen (BASt) je eine Tabelle mit den Codes der zu referenzierenden Straßenabschnitten (*Location Code List, LCL*) und den zu übermittelnden Ereignissen (*Event Code List, ECL*). Die in diesen Listen genannten Referenzierungs-Codes sind die Grundlage für die Übermittlung der Verkehrsinformationen. Sie müssen sowohl den Sendern (i.d.R. Verkehrsleitzentralen und Rundfunkanstalten) aber vor allem auch den Empfängern (z.B. Radio- und Fahrerassistenzsysteme in den Fahrzeugen) bekannt sein. Die Listen werden zwar in regelmäßigen Abständen von der BASt überarbeitet und neu herausgegeben, eine Aktualisierung in den Endgeräten ist jedoch meist nicht möglich, da diese Systeme noch über keine aktive Datenverbindung verfügen. In Japan wurde mit dem *Vehicle Information and Communication System, VICS* ein ähnliches System entwickelt und implementiert (YAMADA 1996).

Im Gegensatz dazu verwendet das *dynamische Location Referencing* keine festen Location-Codes zum Referenzieren der zu übertragenden Informationen. Es wird vielmehr dynamisch, d.h. zum Zeitpunkt der Generierung der Nachricht, eine Ortsreferenz für das entsprechende Objekt generiert. Daher spricht man auch von ad-hoc, on-the-fly oder kartenbasiertem Location Referencing. WEVERS (2012) gibt eine gute Übersicht über die verschiedenen Verfahren zum dynamischen Location Referencing. Seit der Entwicklung der ersten dynamischen Verfahren werden die Ortsreferenzen im sendenden System zum jeweils benötigten Zeitpunkt auf der aktuell vorhandenen Kartengrundlage erzeugt. Dieses System wird auch als *Ausgangs- oder Quellkarte* bezeichnet. Nach der Übertragung werden die Ortsreferenzen im empfangenden System auf der zum aktuellen Zeitpunkt vorhandenen Kartengrundlage decodiert.

Dies wird im Folgenden auch als *Zielkarte* bezeichnet. Nach erfolgreicher Decodierung werden die Referenzen nicht mehr benötigt und können gelöscht werden.

Ein dynamisches Location Referencing System umfasst also im Allgemeinen die folgenden drei Prozessschritte:

1. Generierung (Codierung) der Ortsreferenz im Quellnetz,
2. Übermittlung der Location Codes in das Zielsystem,
3. Decodieren des Location Codes im Zielnetz.

Im Rahmen der vorliegenden Arbeit wird lediglich das dynamische, kartenbasierte Location Referencing betrachtet.

In der Praxis werden sehr häufig linienhafte Ortsreferenzen verwendet, wie etwa der Gültigkeitsbereich einer Geschwindigkeitsbeschränkung. Es existieren aber auch punktförmige (zum Beispiel die Position eines Stopp-Schildes) oder auch flächenförmige Ortsreferenzen. In den originären Kartendatenbanken der verschiedenen Hersteller sind die zu übertragenden Informationen entweder als eigenständige Objekte bzw. Features oder aber als Eigenschaften bzw. Attribute von Straßenobjekten abgespeichert. Dies hängt von der Modellierung der entsprechenden Kartendatenbank sowie der jeweiligen projektspezifischen Objektart ab. Beispielsweise werden Geschwindigkeitsbegrenzungen in Navigationskarten üblicherweise als Attribute der jeweiligen Straßen-Objekte abgespeichert. Soll ein Bereich einer Geschwindigkeitsbegrenzung dann übertragen werden, so muss ein eigenständiges Objekt mit eigener Georeferenzierung für diese Geschwindigkeitsbegrenzung generiert werden.

Neben der geometrischen Position sind dabei stets auch die (topologischen) Beziehungen zum Straßennetz zu berücksichtigen. Aus diesem Grund genügt es in diesem Anwendungsbereich nicht, wie beispielsweise in der Geodäsie üblich, die Ortsreferenzierung nur durch Koordinaten zu beschreiben. Die nachfolgend beschriebenen Verfahren beschreiben daher unterschiedliche Versuche der Referenzierung mit Hilfe digitaler Straßenkarten.

## 3.2 Anwendungen

Nach WEVERS und HENDRIKS (2006) ist Location Referencing eine wichtige Methodik für Anwendungen in der Verkehrstelematik und speziell für alle ortsbezogenen Dienste (*engl. Location-based Services*). Ein typische Anwendung von Location Referencing Verfahren ist die Übermittlung von Verkehrsinformationen (WEVERS 2012). In der Vergangenheit waren dies oft sogenannte *Broadcast*-Anwendungen, bei denen Verkehrsinformationen zentral von Verkehrslagezentren bereitgestellt und über Rundfunk an die Fahrzeuge und deren Systeme ausgeliefert wurden. Dabei handelte es sich um uni-direktionale Systeme, es gab keinen Rückkanal von den Fahrzeugen zum Lagezentrum bzw. anderen Fahrzeugen. Mit Verbreitung der Mobilfunktechnologie und fortschreitender Verfügbarkeit der Netze und Bandbreiten für die Datenübermittlung entstehen immer mehr Systeme, die eine aktive Kommunikation der Fahrzeuge bzw. fahrzeugbasierten Systeme mit ihrer Umgebung nutzen. Die dabei übermittelten Informationen benötigen häufig einen systemübergreifenden aber dennoch kartenbasierten Ortsbezug. Dieser kann durch das dynamische Location Referencing zur Verfügung gestellt werden. Im Folgenden soll die Verwendung von Location Referencing Verfahren anhand einiger Anwendungen aus verschiedensten Bereichen motiviert werden.

### 3.2.1 Digitale Straßenkarten in ITS Anwendungen

In den zurückliegenden Jahren untersuchten eine Vielzahl von Europäischen Forschungsprojekten die verschiedensten Aspekte des Einsatzes digitaler Straßenkarten im Bereich der Intelligenten Verkehrs-

systeme (engl. Intelligent Transportation Systems, ITS). Das *PReVENT / MAPS&ADAS* Projekt erarbeitete beispielsweise Definitionen von sicherheitsrelevanten Straßenattributen für Fahrerassistenzsysteme. Das *SpeedAlert* Projekt wiederum verfolgte die Harmonisierung von fahrzeuggebundenen Geschwindigkeits-Warn Anwendungen. In diesem Zusammenhang wurde die Verfügbarkeit von Informationen über Geschwindigkeitsbeschränkungen in digitalen Straßenkarten als eine der wichtigen Voraussetzungen für kartenbasierte Geschwindigkeitswarner identifiziert.

Das von der EU-Kommission geförderte Forschungsprojekt *ROSATTE (ROad Safety ATtribute exchange infrastructure in Europe)* hatte zum Ziel, eine effiziente und qualitätsgesicherte Datenaustauschinfrastruktur für sicherheitsrelevante Verkehrsinformationen zu entwickeln (ROSATTE 2008). Im Projektkonsortium waren neben den beiden großen kommerziellen Kartenherstellern und verschiedenen Straßenbehörden und Betreiberfirmen unterschiedlicher Europäischer Nationen auch Institutionen zur wissenschaftlichen Begleitung vertreten. Der hauptsächlich betrachtete Anwendungsfall war die Kommunikation von Änderungen in sicherheitsrelevanten Straßenattributen (z.B. Geschwindigkeitsbegrenzung, Baustellen, etc.) von den Straßenbetreibern zu den Kartenherstellern mittels einer einheitlichen Datenschnittstelle (T'SIOBBEL et al. 2008). Die Kartenhersteller sollten die erhaltenen Informationen dann in ihre jeweiligen Datenbestände einpflegen und in geeigneter Weise an ihre Kunden bzw. Nutzer digitaler Straßenkarten weiterreichen. Dies war jedoch nicht mehr Teil des Projektes. Abbildung 3.1 zeigt die Prozesskette inklusive der angedachten Weiterreichung der Informationen durch die Kartenhersteller.

Die entwickelte Schnittstelle (WIKSTRÖM et al. 2009) definierte sogenannte *Road Safety Features* zum Austausch von Änderungsinformationen (Updates) für digitale Straßenkarten. Eine solche Update-Nachricht umfasste einen Teil (Abschnitt) eines Straßennetzwerkes, für den diese Aktualisierung durchgeführt werden sollte. Die Referenzierung dieser Objekte war bewusst nicht auf eine bestimmte Kartengrundlage ausgelegt, sondern explizit auf die Verwendung von dynamischen Location Referencing Verfahren ausgelegt. Bei den beteiligten Straßenbetreibern kamen AGORA-C und OpenLR als Referenzierungsverfahren zum Einsatz. Dazu wurden teilweise bereits existierende Encoder weiterentwickelt, teilweise neue Encoder implementiert oder auf Online Services der Kartenhersteller zurückgegriffen. Die Kartenhersteller als Empfänger konnten Nachrichten mit AGORA-C empfangen, OpenLR wurde leider nur von einem Kartenhersteller unterstützt.

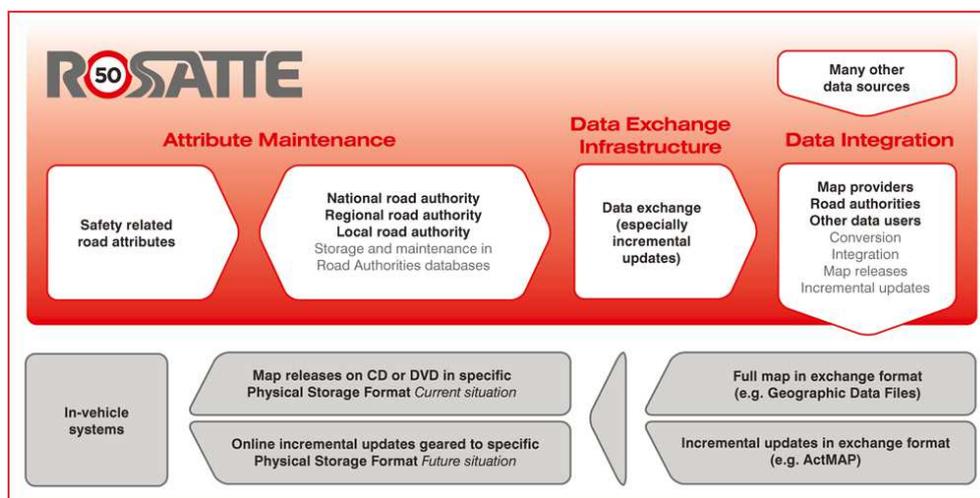


Abbildung 3.1: ROSATTE Datenprozessierung (roter Teil) (ROSATTE 2008)

Nach erfolgreichem Abschluss von ROSATTE fanden sich Anbieter (überwiegend Straßenbehörden) und Nutzer (kommerzielle Kartenhersteller u.a.) von ITS-Verkehrsdaten in der *Transport Network ITS Spatial Data Deployment Platform (TN-ITS)* zusammen, um weiter an der Umsetzung der entwickelten Datenaustausch-Struktur zu arbeiten und diese in möglichst vielen Ländern auch zu implementieren

(TN-ITS 2015). TN-ITS ist in Arbeitsgruppen gegliedert, von denen die beiden wichtigsten die Themen *Location Referencing* und *Standardisierung* bearbeiten. Ziel ist es, Location Referencing für ihre praktische Verwendung in diesem Kontext weiterzuentwickeln und die in ROSATTE identifizierten Schwächen zu eliminieren. Die zweite Arbeitsgruppe arbeitet an der Einbettung der in ROSATTE spezifizierten Austausch-Struktur als Erweiterung des Transport Networks (TN) Themes für ITS-Daten in *INSPIRE (Infrastructure for Spatial Information in the European Community)*. Parallel soll die ROSATTE Spezifikation inklusive der INSPIRE-Erweiterung als *CEN Technical Specification* weiterentwickelt werden.

### 3.2.2 Kartenupdate bei Service-Providern

Eine weitere Anwendungsmöglichkeit von Location Referencing besteht beim Update der Basis-Karten bei Service Providern. Diese unterhalten als Basis der von ihnen angebotenen, kartenbasierten Dienste eigene Kartendatenbanken, die von kommerziell erhältlichen digitalen Straßenkarten abgeleitet werden. Darauf werden dann die für die jeweiligen Dienst (z.B. Verkehrsinformationsdienst, Verknüpfung von Straßennavigation mit Angeboten des Öffentlichen Personen Nahverkehrs (ÖPNV), Radfahrer-Navigation mit speziellen Angeboten für Radfahrer, etc.) benötigten Sachdaten referenziert. In gewissen Intervallen ist jedoch eine Aktualisierung der Basiskarte notwendig. In seltenen Fällen wird der Anbieter der Basiskarte gewechselt. Dabei sind alle auf die seitherige Basiskarte referenzierten Sachdaten auf die neue Version der Basiskarte abzubilden. Für Gebiete mit veränderter Modellierung der Straßen-segmente ergeben sich dann Zuordnungsprobleme, die mit Hilfe von Location Referencing Verfahren bewältigt werden können.

DUEKER und BUTLER (2000) beschreiben den Austausch von Verkehrsdaten nicht als einmaliges Problem. Vielmehr ist es ein Mittel zur Verteilung von Änderungen in den Verkehrsdaten. Datenbanken bei verschiedenen Daten- und/oder Service-Providern auf einem einheitlichen und aktuellen Stand zu halten, impliziert die Möglichkeit, Daten, die eine dieser Stellen akquiriert und in ihren Datenbestand eingepflegt hat, auch an alle anderen beteiligten Stellen weiterzugeben. Nachdem in der Vergangenheit mehr der Aufbau von Datenbanksystemen und damit auch die Ersterfassung der benötigten Daten im Vordergrund stand, rückt immer mehr das Update-Management in den Vordergrund.

### 3.2.3 Elektronischer Horizont

Beide großen kommerziellen Kartenhersteller bieten mittlerweile in Kooperation mit großen Automobilzulieferern einen sogenannten *elektronischen Horizont*, eine kartenbasierte Streckenvorausschau und erweitert damit den für den Fahrer sichtbaren Bereich, vgl. LUDWIG (2014). Bosch entwickelt zusammen mit TomTom den *Connected Horizon* (BOSCH SOFTTEC 2015), Continental zusammen mit Nokia/HERE ein vergleichbares System namens *eHorizon* (CONTINENTAL 2015). Der Austausch und die Nutzung von Fahrzeug- und Sensordaten wurden schon früher unter dem Begriff *Local Dynamic Map* im Rahmen verschiedener Europäischer Forschungsprojekte erforscht, z.B. *SAFESPOT* (SAFESPOT 2015). Daimler bietet in Fahrzeugen seiner Marke Mercedes-Benz seit 2013 bereits einen elektronischen Horizont (DAIMLER 2015b). Abbildung 3.2 zeigt ein typisches Anwendungsszenario: die Verortung und Übermittlung einer Gefahrenstelle.

Herzstück dieser Systeme ist die Integration von topographischen und digitalen Straßenkarten mit vom Fahrzeug erfassten, ortsgebundenen Fahrzeug- und Umgebungsinformationen (Geschwindigkeit, Bremsseingriffe, detektierte Verkehrszeichen). Diese Informationen werden über eine Mobilfunkverbindung an einen zentralen Server gesendet. Diese zentralen Dienste sammeln Daten aller Fahrzeuge, integrieren diese in die Kartendatenbank und generieren inkrementelle Updates für andere Nutzer/Fahrzeuge.

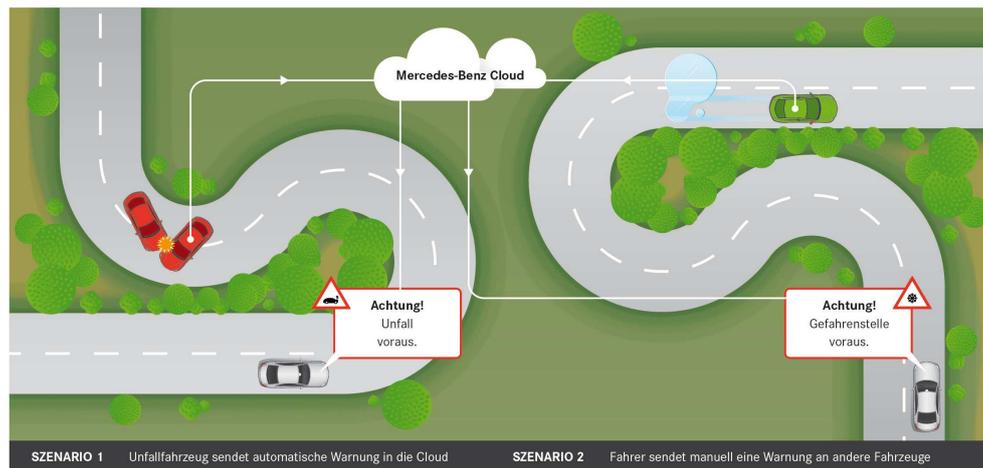


Abbildung 3.2: Car-to-X Kommunikation von Daimler (DAIMLER 2015b)

Dieser Austausch von lokal erfassten Daten mit denen anderer Fahrzeuge ermöglicht eine höhere Aktualität und gesteigerter Detailreichtum der digitalen Karten sowie die Kartierung neuer Straßen aus von Kundenfahrzeugen erfassten/befahrenen Trajektorien.

Durch die Interpretation dieser dynamischen Kartendaten agiert der elektronische Horizont als sekundärer Sensor im Fahrzeug, der zusammen mit den primären Fahrzeugsensoren (z.B. GPS, ESP, Kamera, Radar, Lidar) die Fahrerassistenzsysteme mit zusätzlichen Informationen über die Fahrzeugumgebung versorgt.

Übertragene Informationen:

- Geschwindigkeitsbeschränkung (Änderung)
- Dynamische Verkehrszeichen
- Aktuelle (verkehrsbedingte) Geschwindigkeit
- Staubeginn
- Gefahrenstelle / Unfall
- Wetterbedingungen
- Straßenzustand (Schnee, Glatteis, Hitzeschäden, ...)
- Änderung Straßengefälle

Dies ermöglicht die prädiktive Steuerung verschiedenster Fahrzeugsysteme, beispielsweise Kurvenwarnung, Verkehrszeichenerkennung oder Lichtsteuerung. Der Abstandsregeltempomat (*Adaptive Cruise Control, ACC*) mit aktiver Prädiktion trägt durch automatische Anpassung der Fahrzeuggeschwindigkeit abhängig vom 3D-Profil der vorausliegenden Route zur Verbrauchsreduktion durch Eingriff in die Motor- und Getriebe-Steuerung (Antriebsstrang) bei, vgl. z.B. MARKSCHLÄGER et al. (2014). Dies hilft insbesondere bei reichweitenbeschränkten Hybrid- oder batteriebetriebenen Fahrzeugen durch optimales Ausnutzen der Rekuperationsphasen die Reichweite zu erhöhen. Darüber hinaus hilft das rechtzeitige Abbremsen vor Gefahrenstellen oder Stauenden, die durch Kuppen, Kurven oder andere Hindernisse verdeckt und somit für den Fahrer nicht sichtbar sind, Auffahrunfälle zu vermeiden.

Der elektronische Horizont und die dadurch erstellten dynamischen Karten sind die Grundlage für das automatisierte (autonome) Fahren. Außerdem ermöglichen die dadurch gewonnenen Daten aus den Fahrzeugen die weitere Verbesserung und Erweiterung von Intelligenten Verkehrssystemen (*Intelligent Transportation Systems (ITS)*).

Die derzeitige Beschreibung der Systeme deutet darauf hin, dass zunächst lediglich proprietäre zentrale Kartenserver betrieben werden sollen. Diese würden dann nur Informationen der eigenen Systeme

sammeln und verarbeiten. Dadurch können die übermittelten Objekte relativ einfach auf die jeweiligen Kartenobjekte (z.B. Knoten oder Kanten) mit deren Identifikatoren referenziert werden. Die Verwendung von Location Referencing Verfahren, wie sie in dieser Arbeit betrachtet werden, wäre somit nicht notwendig. Für einen künftigen Informationsaustausch über die Herstellergrenzen hinweg ist die Verwendung eines kartenbasierten aber kartenunabhängigen Location Referencing Verfahrens jedoch unabdingbar.

### 3.2.4 Automatisiertes Fahren

Durch die aktuellen Entwicklungen im Bereich des automatisierten oder autonomen Fahrens und die entsprechende Berichterstattung rücken hochpräzise und aktuelle digitale Straßenkarten als notwendige Voraussetzung zur Referenzierung und Übertragung von Verkehrsinformationen immer mehr in den Fokus einer breiteren Öffentlichkeit (TÖNNESMANN 2015).

Nokia gab im August 2015 bekannt, seinen Kartendienst *HERE* (vormals Navteq) im Zuge einer internen Umstrukturierung und Konzentration auf das Geschäftsfeld Netzwerkausrüstung an ein Konsortium führender Automobilhersteller bis zum ersten Quartal 2016 zu verkaufen. Das Konsortium besteht aus der Audi AG, der BMW Group und der Daimler AG (NOKIA 2015). Die neuen Eigner möchten mit den nun im gemeinsamen Besitz befindlichen hochpräzisen Karten von HERE verbunden mit Echtzeit-Fahrzeugdaten ihren jeweiligen Kunden differenzierte und markenspezifische Produkte für eine neue Generation von Fahrerassistenzsystemen bis hin zum vollautomatisierten Fahren anbieten (DAIMLER 2015a). Neben den meisten führenden Automobilherstellern weltweit zählt HERE auch zahlreiche Software- und Internetfirmen wie Microsoft, Samsung und SAP zu seinen Kunden und möchte seine Kartendaten und Services auch weiterhin für deren Dienste anbieten.

HERE entwickelt derzeit einen *location cloud service*, der die von Fahrzeugen, Geräten und Infrastruktur-Systemen generierten Daten zusammenführen und für echtzeitfähige, prädiktive und personalisierte Location Services nutzbar machen möchte (HERE 2015).

Im Juli 2015 gaben Bosch und TomTom bekannt, eine Kooperation im Bereich Kartendaten für das hoch-automatisierte Fahren eingegangen zu sein (BOSCH 2015). Im Rahmen dieser Kooperation leitet Bosch die Anforderungen an Genauigkeit und Inhalte der digitalen Karten aus der Systementwicklung ab. TomTom erstellt verbesserte, digitale Straßenkarten mit Dezimeter- statt wie bisher Meter-Genauigkeit. Erste Autobahn-Abschnitte in der Region Stuttgart sind bereits erfasst, die Autobahnen innerhalb Deutschlands sollen bis Ende 2015 folgen, Karten für Europa und Nordamerika danach.

Neben der Genauigkeit soll auch die Aktualität der Kartendaten verbessert werden, denn hochautomatisiertes Fahren ohne aktuelle Karten-Informationen ist nicht möglich. Daher erfolgt die Erfassung zusätzlich zur weiterhin aktiven TomTom Kartierungsflotte durch speziell mit entsprechenden Sensoren ausgerüsteten, weiteren Fahrzeugflotten. Dadurch können neben Veränderungen der Straßen oder ihrer Eigenschaften (z.B. neu gebaute Strecken oder neu aufgestellte Verkehrszeichen) auch temporäre Verkehrsinformationen, wie z.B. Staus erfasst werden. Diese Informationen werden in Echtzeit an einen zentralen Server bei TomTom übertragen, validiert und in die digitale Kartendatenbank eingetragen. Diese werden dann den hochautomatisiert fahrenden Fahrzeugen aktualisiert zur Verfügung gestellt. Somit ist es möglich, Verkehrsinformationen, die über die Reichweite der Umfeldsensorik hinausgeht, zu verbreiten.

Die verbesserten TomTom Karten bestehen nunmehr aus drei Schichten: über die *Basis-Navigationsschicht* erfolgt die Berechnung der Fahrtrouten. Mit Hilfe der *Lokalisierungsschicht* kann ein automatisiertes Fahrzeug die Genauigkeit der Bestimmung seiner Fahrzeugposition verbessern. Dazu werden beispielsweise über die Umfeldsensoren des Fahrzeugs erfasste Objekte mit entsprechenden Informationen der Lokalisierungsschicht erfasst und seine Position relativ zu diesen

Objekten bestimmt. Die darüber liegende *Planungsschicht* beinhaltet verschiedenste Straßeneigenschaften, wie z.B. Fahrbahnmarkierungen, Verkehrsschilder oder Geschwindigkeitsbeschränkungen oder auch Informationen zur Fahrbahngeometrie, wie beispielsweise Kurvenradien oder Gefälle. Diese Informationen können von einem automatisiert fahrenden Fahrzeug verwendet werden, um zu entscheiden, wann und wo z.B. eine Fahrspur gewechselt werden soll.

## 3.3 Dynamische Location Referencing Verfahren

### 3.3.1 ILOC

Anfang der 1990er Jahre begann mit dem SOCRATES Projekt (CATLING et al. 1994) die Entwicklung von Systemen zur Übermittlung von Verkehrsnachrichten und zur Erfassung von Verkehrsdaten. Die übertragenen bzw. erfassten Informationen sollten dabei auf Grundlage von RDS-TMC Codes bzw. Subcodes auf den jeweiligen Straßenabschnitt bezogen werden. Dies gelang jedoch nicht in zufriedenstellendem Maße (WEVERS 2012). Um die begonnene Arbeit fortzusetzen, wurde 1995 das *ERTICO Committee on Location Referencing* gegründet. Ziel war es zunächst, die in SOCRATES aufgetretenen Probleme mit dem statischen Location Referencing weiter zu analysieren und nach einer geeigneten Lösung zu suchen. Im Ergebnis würde die Verwendung von statischen Location Codes jedoch die Zuweisung von dauerhaften eindeutigen Bezeichnern für alle Straßenabschnitte innerhalb Europas oder gar weltweit notwendig machen. Solche Tabellen würden sehr groß werden und immense Kosten für ihre Erstellung und Laufendhaltung hervorrufen. Schlussendlich wurde also mit der Entwicklung eines dynamischen Location Referencing Verfahrens begonnen. Das entwickelte Verfahren basiert auf der Verwendung von Kreuzungspunkten (engl. Intersection Location) und wurde unter dem Namen ILOC-Verfahren in DUCKECK et al. (1998) veröffentlicht.

Eine Intersection Location (ILOC) wird durch die geographischen Koordinaten des Kreuzungsmittelpunktes beschrieben. Sie wird ergänzt durch drei beschreibende Attribute der verbundenen Straßenelemente, um eventuelle Mehrdeutigkeiten bei der koordinatenbasierten Suche nach der entsprechenden Kreuzung in der Zielkarte aufzulösen. Da Kreuzungssituationen in verschiedenen digitalen Straßenkarten häufig unterschiedlich modelliert sind, werden bei der ILOC Berechnung die Kreuzungspunkte auf einen Mittelpunkt reduziert. Dies ist in der Regel der Schwerpunkt der Schnittpunkte der sich kreuzenden Straßenachsen. Zur Reduzierung des Datenvolumens wird der Kreuzungspunkt nur mit Informationen über drei der verbundenen Straßenelemente versehen. Dies können z.B. Straßen-Nummer oder Straßennamen sein. Sollten mehr Informationen in der Karte zur Verfügung stehen, muss eine Auswahl getroffen werden. Ein erweitertes ILOC-Format kann darüber hinaus noch Angaben zum Location Typ (z.B. einfache/komplexe Kreuzung oder Kreisverkehr) und der verwendeten Sprache für die Straßennamen versehen werden, um die Zuverlässigkeit des Verfahrens zu erhöhen.

Es bleibt zu erwähnen, dass in DUCKECK et al. (1998) lediglich eine Methodik zur Erstellung von Intersection Locations beschrieben wird. Es wird kein Verfahren zum Decodieren der Ortsreferenz in der Zielkarte beschrieben.

Innerhalb des EVIDENCE Projektes wurde das ILOC Verfahren bei mehreren Projektpartnern, darunter auch kommerzielle Kartenhersteller, implementiert und ausgiebig getestet. Die durch die Tests gewonnenen Erfahrungen im Umgang mit ILOCs führten zu einer weiteren Konkretisierung der Definition und zur klaren Empfehlung zur Einführung von Location Codes, welche in DORENBECK und HERES (1999) veröffentlicht wurden.

BOFINGER (2001) implementiert ein vereinfachtes ILOC Verfahren und nutzt dabei die Geometrie (Punktkoordinaten), Semantik (Straßennummern und Straßennamen) und Topologie (Anzahl der abgehenden Kanten eines Knotens) der ILOC-Punkte für die Zuordnung.

Während verschiedener Diskussionen nach Abschluss des EVIDENCE Projektes bildete sich die sogenannte *Ad-hoc Group on Location Referencing*, um das ILOC Verfahren noch weiter zu verfeinern. Insbesondere die bereits erwähnten Location Codes sowie die explizite Codierung von Straßensegmenten und die damit verbundene Steigerung des Informationsgehalts der Ortsreferenz sollen zu weiteren Verbesserung des Location Referencing führen. Neben dem in EVIDENCE angedachten *Intersection Type* sollen nun auch die *Functional Road Class (FC)* und *Form of Way (FW)* entsprechend des GDF Standards (ISO 14825 2005) zur Auflösung von Mehrdeutigkeiten in der Zielkarte verwendet werden. Durch die Einführung von expliziten Straßensegmenten, welche durch die ILOC Codes der sie umschließenden Kreuzungspunkte gebildet werden, wurde es nun möglich, auch die Richtung und Länge des Segments mit zu codieren (WEVERS 2000). Das weiterentwickelte Verfahren wird auch als *Extended ILOC* bezeichnet (WEVERS et al. 2001). Abbildung 3.3 zeigt ein Beispiel für die Repräsentation von komplexen Kreuzungen durch ILOC Locations. Dabei werden die Kreuzungen 1, 2 und 3 als komplex, Kreuzung 4 als einfach angesehen. Des Weiteren wird auch ein Teil eines Straßenstücks mit separat digitalisierten Richtungsfahrbahnen zwischen zwei (notwendigerweise) komplexen Kreuzungen als komplexes Objekt repräsentiert (im Beispiel der Bereich zwischen 1 und 2).

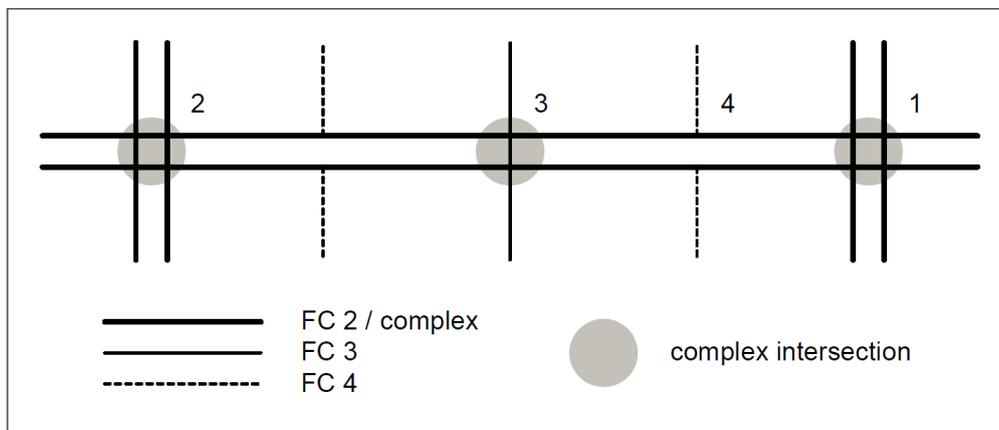


Abbildung 3.3: Extended ILOC Location (WEVERS 2000)

### 3.3.2 SPOT

Nach der Entwicklung des ILOC Verfahrens wurde von Tele Atlas, einst auch ILOC Projektpartner, ein weiterentwickeltes Verfahren namens *Spatial Object Tag (SPOT)* zum Patent angemeldet (EP 1 078 346 2002). Mit SPOT sollen mögliche Klassifizierungsunterschiede bei den Kartenattributen durch im Vorfeld definierte Übersetzungstabellen verhindert werden. Es werden dabei eindeutige Zuordnungen von Attributen der Ausgangskarte zu denen der Zielkarte definiert. Dadurch könnte ein entscheidender Schwachpunkt des ILOC-Verfahrens beseitigt werden. Sind nämlich Objekteigenschaften (z.B. die Straßenklasse) in den beteiligten Karten unterschiedlich modelliert, dann muss spätestens beim Decodieren eine Zuordnung zwischen den beiden Klassifizierungen hergestellt werden. Dies setzt jedoch die Kenntnis der jeweiligen Klassendefinitionen voraus, was in der Praxis nur sehr selten der Fall sein dürfte.

SPOT verallgemeinert den ILOC Ansatz in der Hinsicht, dass alle möglichen Attribute aus der Ausgangskarte mit der Ortsreferenz codiert und zur Auflösung von Mehrdeutigkeiten beim Decodieren auf der Zielkarte verwendet werden können. Die verwendeten Objekteigenschaften müssen vor der Verwendung von SPOT in einem Objektartenkatalog beschrieben werden, der dann in der Quell- und Zielkarte die Grundlage für das Encoding und Decoding bildet. Ein *spatial object tag* besteht aus einer räumlichen Referenz (Punktkoordinaten), einem Verweis auf die Objektart aus dem zuvor vereinbarten Objektartenkatalog und schließlich einem beschreibenden Element entsprechend der Objektart (z.B.

der Straßennamen für ein Straßen-Objekt). Auch Daimler war auf dem Gebiet des Location Referencing aktiv und hat mit ROSA (Reconstruction of Objects on a Second mAp) eine Erweiterung von SPOT in DEMIR (2002) veröffentlicht, bei der mehr Informationen über die zu referenzierenden Objekte übertragen werden. Damit soll eine wichtige Voraussetzung von SPOT, ein im Voraus verabredeter Objektkatalog, obsolet und das Verfahren somit insgesamt flexibler einsetzbar sein. Das ROSA Location Referencing Verfahren wurde im Rahmen des ActMAP-Projekts implementiert und zur Referenzierung von Kartenupdates für Navigationsanwendungen eingesetzt (ALEKSIĆ 2004).

### 3.3.3 GoodLane

Unabhängig von ILOC beschäftigte sich auch Bosch mit dem Problem des Location Referencing. Im Gegensatz zum ILOC Verfahren wählte Bosch einen allgemeineren, geometriebasierten Ansatz namens *Geographical Objects: Dynamical Linking and Navigable Embedding (GoodLane)* (HAHLWEG et al. 2000). Es werden nur Informationen zur Objektgeometrie verwendet. Das Verfahren verzichtet vollständig auf die Verwendung semantischer oder topologischer Objekteigenschaften für die Zuordnung. Dadurch soll das Verfahren auch mit Karten verwendet werden können, die größere Modellierungsunterschiede gerade in Bezug auf die Objektattribute aufweisen. Bei GoodLane wird die zu codierende Objektgeometrie bei Bedarf um zusätzliche, umliegende Geometrielemente erweitert, um eine möglichst zuverlässige Decodierung im Zielnetz zu ermöglichen. Die codierte Geometrie soll soweit eindeutig bzw. einzigartig in einem gewissen lokalen Umfeld sein, dass das ihr entsprechende Stück mittels Muster-/ Formerkennung bzw. Map Matching Verfahren zuverlässig in der Zielkarte identifiziert werden kann. Je mehr Geometrieinformationen mit der Ortsreferenz codiert werden, desto robuster wird das Verfahren auch gegenüber Abweichungen zwischen Ausgangs- und Zielkarte. Abbildung 3.4 zeigt ein einfaches Beispiel.

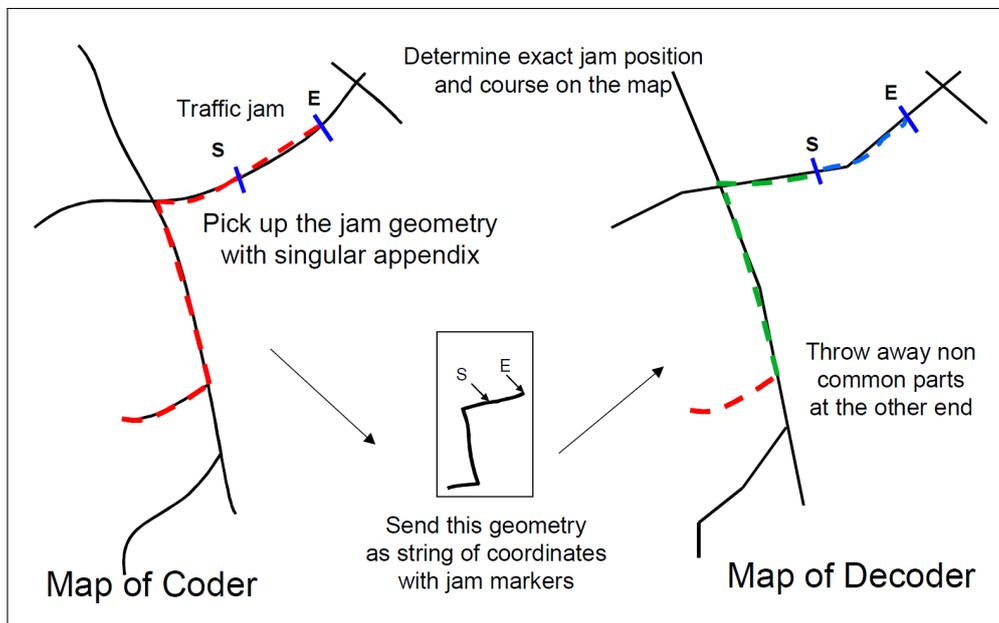


Abbildung 3.4: Codierung eines Stau-Objektes mit GoodLane (HAHLWEG et al. 2000)

Das Prinzip des GoodLane Verfahrens wurde von der Robert Bosch GmbH beim Europäischen Patentamt zum Patent angemeldet (EP 1 224 645 2003). Des Weiteren wurden auch die Verfahren zur Erweiterung der Geometrie (EP 1 336 079 2004) und Korrelation von Punktketten zur Decodierung (EP 1 122 518 2008) patentrechtlich geschützt. Eine allgemeine Verwendung bzw. Weiterentwicklung dieses Verfahrens wird somit erschwert (WEVERS 2012).

### 3.3.4 PivotPoint

In etwa zur selben Zeit wie auch GoodLane, arbeitete Siemens VDO ebenfalls einem eigenen Location Referencing Verfahren. Interessanterweise argumentieren die Entwickler bei der Auswahl der verwendeten Verfahren genau entgegengesetzt dem GoodLane Verfahren. Um mögliche Rekonstruktionsfehler oder Mehrdeutigkeiten beim Decodieren der Ortsreferenzen zu vermeiden, sollen bevorzugt relative Zuordnungsmethoden verwendet werden. Dazu werden beispielsweise die Kriterien zur Bewertung der Übereinstimmung als Prozentwert angegeben.

Das Verfahren selbst stützt sich hauptsächlich auf die Netzwerk-Topologie in den Karten. Kreuzungssituationen sind häufig von einer gewissen Komplexität und werden unterschiedlich in den verschiedenen Karten modelliert. Diese Kreuzungspunkte sind somit nicht optimal zur Referenzierung von Informationen auf dem Straßennetz. Referenzen auf Straßenverbindungen zwischen den Kreuzungen sind hingegen meist eindeutig, vor allem wenn es sich um Straßen höherer Klassifikation handelt. Dementsprechend sollen beim PivotPoint Verfahren sogenannte Angelpunkte (engl. pivot points) abseits von Kreuzungen ausgewählt und codiert werden. Mit den geographischen Koordinaten des Angelpunktes sollen auch der Richtungswinkel der Straßenachse, der Straßentyp entsprechend dem GDF-Standard (ISO 14825 2005) sowie ein Parallelspuranzeiger codiert werden. Abbildung 3.5 zeigt beispielhaft die Codierung eines Straßensegments mit außerhalb liegenden Pivot Points (PP).

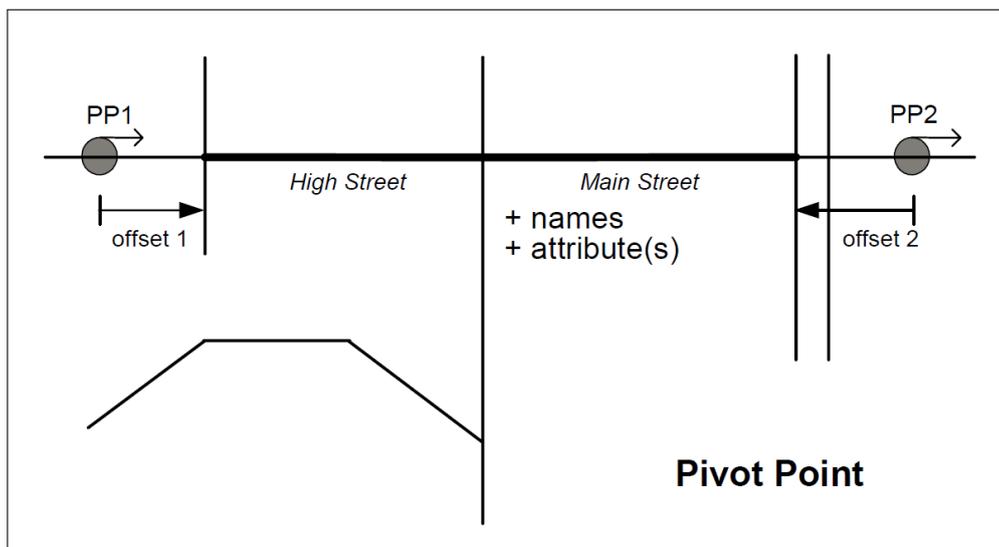


Abbildung 3.5: Codierung einer Ortsreferenz mit PivotPoints (WEVERS et al. 2002)

Um nun eine Ortsreferenz entlang einer Straße zu codieren, werden ein oder mehrere (sequenzielle) Angelpunkte zusammen mit ihren Abständen untereinander sowie den gerichteten Abständen zum Anfangs- und Endpunkt der Ortsreferenz gespeichert. Zusätzlich werden auch noch eventuelle Punkte an denen die Straßenklasse wechselt durch ihren Abstand zum nächsten Angelpunkt referenziert.

Um Mehrdeutigkeiten durch evtl. parallel verlaufende Routen zu vermeiden, werden die berechneten Entfernungen entsprechend der jeweiligen Straßenklasse gewichtet. Die Angelpunkte müssen dann so gewählt werden, dass die zu codierende Verbindung die niedrigste gewichtete Entfernung zwischen den Angelpunkten im Netz hat. Darüber hinaus müssen die Angelpunkte so gewählt werden, dass jede mögliche Alternativroute mindestens 25 % länger ist, als die zu codierende Ortsreferenz (WEVERS et al. 2001).

Das PivotPoint Verfahren wurde von Siemens ebenfalls zum Patent angemeldet (EP 1 225 552 2004).

### 3.3.5 AGORA

#### AGORA Projekt

Zur Fortsetzung der Arbeiten innerhalb des EVIDENCE Projektes und zur weiteren Steigerung der Robustheit und Zuverlässigkeit der Location Referencing Verfahren förderte die Europäische Union das Projekt *Implementation of global location referencing approach* oder kurz *AGORA* (WEVERS 2012). *AGORA* kommt aus dem Griechischen und bedeutet so viel wie *Marktplatz* (WEVERS und HENDRIKS 2006). Die Anspielung auf einen Platz, auf dem üblicherweise Gegenstände oder Objekte (aus)getauscht werden, zeigt schon die antizipierte Verwendung von *AGORA* als Grundlage für den Austausch von Verkehrsinformationen. Zunächst sollte es auf dem Extended ILOC Verfahren aus EVIDENCE aufbauen. Allerdings wurden auch die zur selben Zeit entwickelten Verfahren GoodLane von Bosch sowie PivotPoint von Siemens VDO eingereicht und mussten daher in *AGORA* berücksichtigt werden (WEVERS et al. 2001). Das zu entwickelnde Verfahren sollte also alle drei Ansätze in einem vereinen, was nicht ganz einfach war. Schließlich sind der geometriebasierte Ansatz von Bosch und der topologiebasierte Ansatz von Siemens VDO völlig konträr.

Das Ergebnis war dann auch mehr eine Kombination aus den drei zuvor genannten Verfahren als eine wirkliche Neuentwicklung. Eine *AGORA* Location Reference besteht aus einer PivotPoint Referenz, einer oder mehreren ILOC Referenzen und einer optionalen erweiterten Geometrie Referenz, die auf GoodLane basiert. Die *AGORA* Referenz wird durch Punkte verschiedenen Typs definiert. Neben Punkten vom Typ ILOC, PivotPoint und GoodLane sind auch Zwischenpunkte ohne besondere thematische Bedeutung vorgesehen. Um die Effizienz zu steigern, kann ein Punkt auch mehr als eine Punktart besitzen, es ist also zum Beispiel möglich, einen GoodLane Point in einen Intersection Point zu legen. Mit einem Satz von 28 Codier-Regeln wird beschrieben, wie aus einem gegebenen Routenverlauf, eine *AGORA* Ortsreferenz gebildet werden soll. Dies beinhaltet zunächst die Auswahl von Punkten der entsprechenden Punkttypen sowie die Zuweisung von Attributen zu den jeweiligen Punkten (WEVERS et al. 2003). Abbildung 3.6 zeigt die Codierung eines Straßensegments mittels *AGORA*.

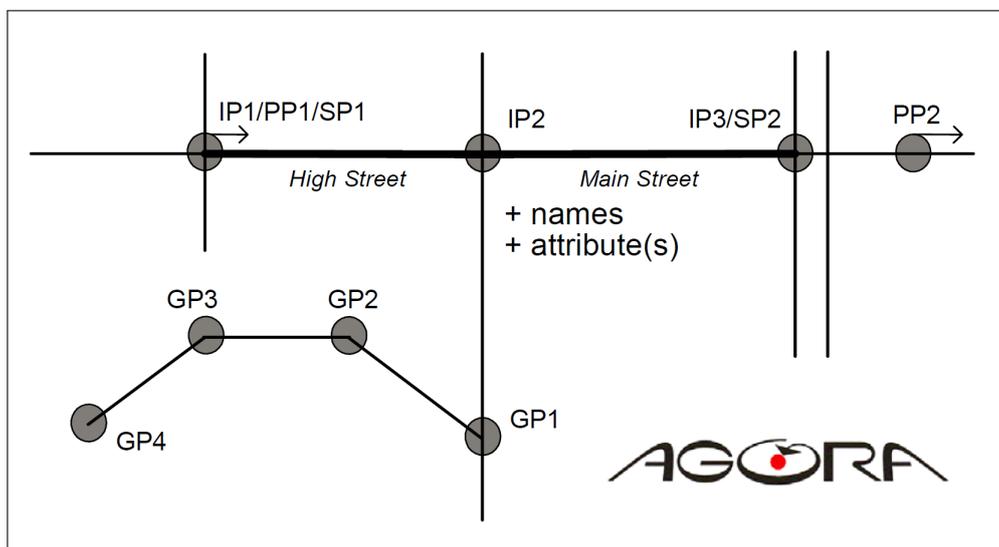


Abbildung 3.6: AGORA als Kombination von ILOC, PivotPoint und GoodLane (WEVERS et al. 2002)

#### AGORA-C

Aufgrund der oben genannten Schwachstellen von *AGORA* setzten zwei der Projektpartner nach dem offiziellen Abschluss des *AGORA*-Projektes die Arbeit an einem Location Referencing Verfahren fort. *AGORA* war als Kombination der drei ursprünglichen Location Referencing Ansätze zwar grundsätzlich geeignet, jedoch waren die zu übertragenden Datenpakete viel zu groß für kommerzielle Anwendungen. Zum anderen waren die aufgestellten Codier-Regeln vielen Anwendern zu kompliziert und

dadurch einer weiteren Verbreitung von AGORA eher abträglich. Deshalb wurde vor allem die weitere Integration der drei Ursprungs-Methoden forciert, um den Anteil an redundant übertragenen Informationen weiter zu minimieren (WEVERS und HENDRIKS 2006).

Entstanden ist daraus schließlich das sogenannte AGORA-C Verfahren, wobei das *C* für *Compact* steht. In der ursprünglichen Spezifikation (WEVERS und HENDRIKS 2003) wurde auf den GoodLane Ansatz verzichtet. Durch die Beschränkung auf Extended ILOC und PivotPoint sowie einige Vereinfachungen im physischen Datenformat sollte der Standard weiter vereinfacht und die benötigte Bandbreite für die Datenübertragung stark reduziert werden, bei ungefähr gleichen Ansprüchen an die Performance. Nachdem NAVTEQ und Siemens VDO zunächst alleine an AGORA-C gearbeitet hatten, konnte das Mobile.Info Projektkonsortiums im Jahr 2005 schließlich Gespräche mit Bosch vermitteln, welche die Integration von GoodLane in AGORA-C zum Ziel hatten (WEVERS und HENDRIKS 2006). Nach erfolgreichem Abschluss der Gespräche wurde dann die finale Spezifikation von AGORA-C im April 2005 veröffentlicht (HENDRIKS et al. 2005).

AGORA-C besteht aus zwei Anwendungsprofilen: Das *Core Profile* für Problem bzw. Status Nachrichten, wie z.B. Verkehrslageinformationen. Zum anderen das *Extended Profile*, welches insbesondere zur Codierung von Ziel-Ortsreferenzen für Routing-Anwendungen verwendet werden kann.

Der Übergang von AGORA nach AGORA-C stellte eine komplette Überarbeitung des Verfahrens inklusive sämtlicher seiner Begrifflichkeiten dar. So entstand das *Core Profile* aus den ILOC und PivotPoint Ansätzen. Drei verschiedene Punktarten wurden hierfür definiert: Location-, Intersection- und Routing Points. Eine Ortsreferenz besteht immer aus einer geordneten Liste solcher Punkte. Punkte können auch mehrere der genannten Punkttypen repräsentieren. Jedem dieser Punkte, aber auch der Ortsreferenz selbst, können ein oder mehrere Attribute zugeordnet werden.

Ein Location Point bezeichnet den Start-, Zwischen- oder Endpunkt eines Realwelt-Objektes. Ein Intersection Point beschreibt eine Kreuzung, wenn sich darin die sogenannte *Road Section Signature* ändert. Diese besteht aus den Attributen Functional Road Class, Form-of-Way und dem sogenannten *Road descriptor*. Dieser enthält die offizielle Straßenummer oder, falls nicht vergeben, ihren Namen. Routing Points können verwendet werden, um eine Ortsreferenz durch Routenberechnung zu rekonstruieren. Der erste und letzte Core-Point einer Location sind per Definition immer auch Routing Points. Routing Points sollen außerhalb von komplizierten Kreuzungsbereichen liegen. Im Gegensatz zu AGORA wurde dies in AGORA-C zu einer "soll"-Regelung herabgestuft (HENDRIKS et al. 2005).

Das *Extended Profile* kommt zur Anwendung, wenn Ortsreferenzen als Ziele für Routing-Verfahren codiert werden sollen. Es wurde auf Grundlage des GoodLane Ansatzes von Bosch entwickelt und in AGORA-C eingefügt. Die Idee ist, die zu codierende Ortsreferenz durch zusätzlich zu codierende Punkte über einen kürzesten Weg mit dem übergeordneten Straßennetz zu verbinden. Dieses übergeordnete Netz wird in der Regel mit einer höheren Priorität in den Karten erfasst und aktualisiert. Es sollte also in möglichst vielen Kartendatenbanken vollständig und aktuell enthalten sein. Über die zusätzlich codierten *Extension Points* kann dann vom übergeordneten Netz ausgehend zum Zielpunkt navigiert werden.

Wie schon bei AGORA wird auch AGORA-C durch einen Satz von Codier-Regeln definiert. Äußerst hilfreich für die praktische Anwendung und Implementierung ist jedoch die zusätzlich enthaltene Prozessbeschreibung für das Codieren von Core- und Extended Locations.

### ISO 17572-3

Nach der Fertigstellung der AGORA-C Spezifikation wurde schließlich auch eine ISO-Standardisierung des Verfahrens angestoßen. Daraus entstand dann Teil 3 *Dynamic Location References* der ISO 17572 *Intelligent Transport Systems - Location Referencing for Geographic Databases* (ISO 17572-3 2008). Jedoch gelangten während des Standardisierungs-Verfahrens nicht nur die Ansätze aus Europa in den Standard. Es wurden auch einige Ergänzungen, hauptsächlich aus Japan, in den Standard eingearbeitet. Nichtsdestotrotz ist die Ähnlichkeit mit AGORA-C immer noch sehr groß.

### 3.3.6 MEI-LIN

WARTENBERG (2008) beschreibt ein topologiebasiertes Verfahren zur Referenzierung von linienhaften Ortsreferenzen namens *MEI-LIN* (*MEthod for Identifying Locations In road Networks*). Die Grundidee dazu wurde in WARTENBERG (2006a) und später auch in WARTENBERG (2006b) veröffentlicht. Grundlage ist die Verwendung topologischer Bäume als Muster, welche in der Zielkarte wiedergefunden werden sollen. Dazu müssen aus der zu referenzierenden Ortsreferenz oder aus deren näherer Umgebung ein oder mehrere Punkte ausgewählt werden, um die innerhalb einer bestimmten Entfernung ein Baum aufgebaut wird. Dieser enthält die kürzesten Wege aller Knoten innerhalb des Suchbereichs zum Ausgangsknoten. Diese Baumstruktur kann leicht und effizient codiert und übertragen werden. Diese Stammunkte müssen so gewählt sein, dass die gesamte Ortsreferenz später durch die Baumstruktur abgedeckt ist. In der Zielkarte werden dann mittels Suchradius Punkte in der Umgebung des Ausgangspunktes gesucht und für alle gefundenen Punkte ebenfalls Baumstrukturen mit kürzesten Wegen des umliegenden Netzes berechnet. Die Identifikation des am besten passenden Punktes erfolgt durch Vergleich der Baumstrukturen.

### 3.3.7 TPEG2-ULR

Unter dem Akronym TPEG (Transport Protocol Experts Group) wird eine Protokoll-Spezifikation zum Austausch von Echtzeit Verkehrsinformationen (RTTI, Real Time Traffic Information) verstanden. Mit Hilfe von TPEG sollen Reisenden umfassende und aktuelle Verkehrs- und Reiseinformationen zur Verfügung gestellt werden. TPEG besteht aus einer Reihe von ISO-Standards, die bereits eine Vielzahl von Anwendungen abdecken, aber auch durch neue zukünftige Anwendungen erweitert werden kann. Diese Standards werden von der Traveller Information Services Association (TISA), einer gemeinnützigen Organisation, ständig weiter entwickelt.

Zur Übertragung der Verkehrsinformationen müssen diese mit einer Georeferenzierung versehen werden. Bislang geschah dies mit Hilfe von statischen Location Referencing Methoden wie RDS-TMC (DIN EN ISO 14819-3 2014). Im Zuge der Einführung des digitalen Rundfunks DAB+ soll jedoch TPEG als neue Art der Übertragung von Verkehrsinformationen mit erweiterten Anwendungsmöglichkeiten eingeführt werden. Zur (dynamischen) Georeferenzierung der TPEG-Nachrichten wurde ein eigenes dynamisches Location Referencing Verfahren unter dem Namen TPEG-Loc entwickelt, welches auf dem ILOC Verfahren basiert (DIN ISO/TS 18234-6 2006). TPEG-Loc erzielte jedoch in der Praxis nur ungenügende Ergebnisse und wurde im Rahmen der Entwicklung der zweiten Generation von TPEG (TPEG2) durch ein erweitertes Location Referencing Konzept abgelöst. Demnach können nun neben dem herkömmlichen RDS TMC auch verschiedene dynamische Location Referencing Verfahren verwendet werden. Dies ist zum einen AGORA-C, welches aber wegen der fehlenden Möglichkeit der kostenlosen Nutzung für den Einsatz in DAB+ nicht infrage kommt. Zusätzlich wurde als Nachfolger von TPEG-Loc ein neues Verfahren namens TPEG2-ULR (Universal Location Referencing) entwickelt (SCHRAMM et al. 2012).

Ein zunächst verfolgter Ansatz zur Darstellung eines Locationverlaufs durch kürzeste Wege zwischen einer Liste von sogenannten *Segment Identification Points (SIP)* (ERNST et al. 2013b) wurde wegen Bedenken bzgl. der möglichen Verletzung geistigen Eigentums Dritter später wieder verworfen und durch einen komplett neuen Ansatz abgelöst (ERNST et al. 2013a). Dieser neue Ansatz basiert auf der Übertragung linearer Ortsreferenzen (engl. Locations) mittels formalisierter Richtungsangaben (engl. *formalized direction instructions*). Die Beschreibung einer Ortsreferenz erfolgt durch absolute Richtungswinkelangaben an jeder signifikanten Richtungsänderung. Zusätzlich werden die Koordinaten des Anfangs- und Endpunktes und optional weitere Punktattribute der Ortsreferenz übertragen. Der Decoder kann also ausgehend vom Startpunkt den Richtungsangaben folgen, um den Routenverlauf zum

Endpunkt in der Zielkarte zu identifizieren. Im Zielnetz werden mögliche Zuordnungskandidaten identifiziert und mittels einer Kostenfunktion bewertet. Darin fließen die gewichteten Winkel- und Strecken-Abweichungen für alle Liniensegmente der Ortsreferenz ein. Der Kandidat mit den geringsten Kosten ist der mit der höchsten Zuordnungswahrscheinlichkeit. Abbildung 3.7 zeigt beispielhaft eine langgestreckte Ortsreferenz und ihre *Direction Instructions* an den Punkten mit signifikanten Richtungsänderungen. Um das Verfahren robust zu machen gegen mögliche lokale topologische oder geometrische Fragmentierung (z.B. durch Kreisverkehre oder allgemein unterschiedlich modellierte Kreuzungssituationen in Quell- und Zielkarte), bleiben Richtungsänderungen von kurzen Segmenten im Encoding und Decoding unberücksichtigt (sog. Ultimatum-Konzept). Das Zuordnungsverfahren wurde mit verschiedenen Karten getestet und lieferte eine Zuordnungsrate (Anteil korrekt zugeordneter Ortsreferenzen) von 80-90 % (ERNST et al. 2014).

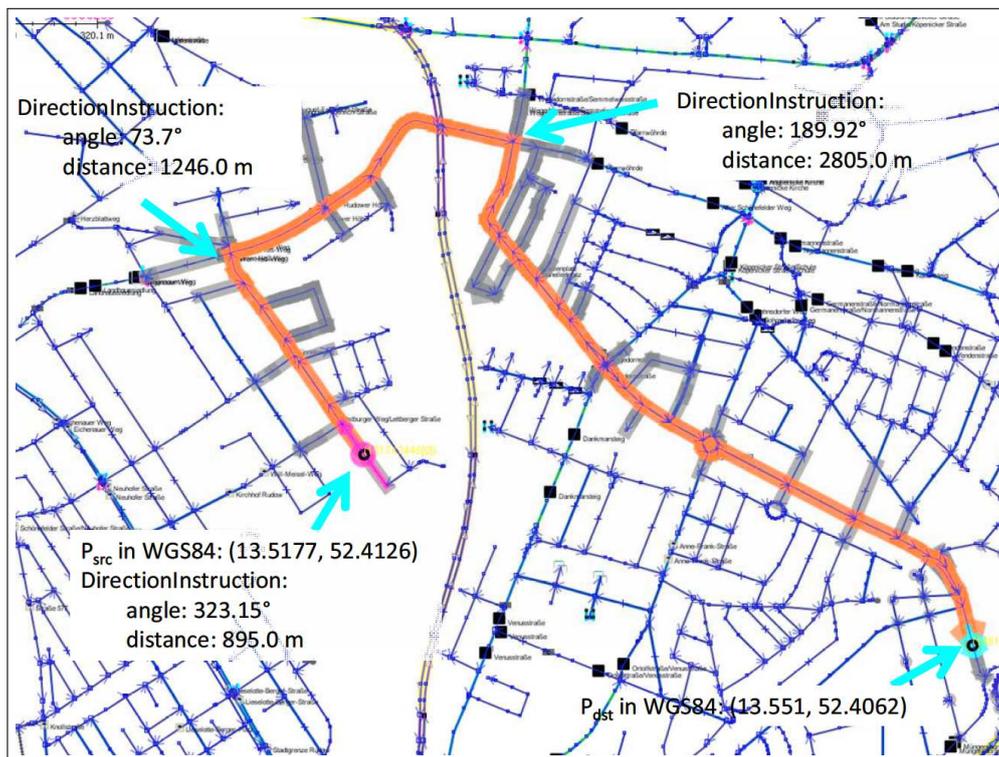


Abbildung 3.7: Repräsentation einer Ortsreferenz durch Direction Instructions (ERNST et al. 2013a)

## 3.4 OpenLR

### 3.4.1 Standard

OpenLR ist ein von TomTom International B.V. entwickeltes Location Referencing Verfahren. Im Gegensatz zu AGORA wurde OpenLR als offener Standard entwickelt. TomTom pflegt und entwickelt ihn weiter, jedermann ist jedoch grundsätzlich eingeladen, an der Weiterentwicklung mitzuarbeiten und seine Entwicklungen einzureichen. Um dies zu ermöglichen, veröffentlicht TomTom den OpenLR Standard in Form eines White Papers, aktuell in der Version 1.5, Revision 2 (TOMTOM 2012). Dieses White Paper ist durch die *Creative Commons Public License - Namensnennung-Keine Bearbeitungen* (CCPL 2007) geschützt. Darin wird die kostenfreie Verbreitung des OpenLR White Papers gewährt, sofern der Urheber genannt wird. Änderungen dürfen aber nur von TomTom selbst vorgenommen werden.

OpenLR selbst wird als Referenzimplementierung angeboten. Die Veröffentlichung unter der Apache License 2.0 (APACHE 2004) ermöglicht jedermann die kostenfreie und uneingeschränkte Verwendung, Verbreitung sowie Modifikation der OpenLR Software.

OpenLR wurde ursprünglich zur Übertragung von Verkehrsinformationen von einem Verkehrslagezentrum zu fahrzeugbasierten Systemen oder sonstigen mobilen Endgeräten entwickelt. Voraussetzung zur Verwendung von OpenLR ist generell das Vorhandensein von navigationsfähigen digitalen Straßenkarten auf der Sender- und Empfängerseite. Inhaltlich müssen die Karten des Weiteren folgende Anforderungen erfüllen: Wie auch bei den vorgenannten Verfahren müssen alle Punkte durch ihre geographischen Koordinaten bezogen auf WGS84 festgelegt sein. Jedes Liniensegment muss eine Angabe zu seiner wahren Länge enthalten, sowie die Stützpunkte der Linien deren wahre Geometrie widerspiegeln. Die Karten müssen darüber hinaus Angaben zur Straßenklasse (Functional Road Class, FRC) und zur technischen Beschaffenheit der Straße (Form of way, FOW) beinhalten. Da gerade bei diesen beiden Attributen unterschiedliche Klassifizierungen in den verschiedenen Karten auf Sender- und Empfängerseite möglich sind, gibt OpenLR eine Klassifizierung vor, die an den GDF-Standard (ISO 14825 2011) angelehnt ist. Für den Encoder und Decoder müssen dann jeweils Schnittstellen (Übersetzungstabellen) gefunden bzw. definiert werden, mit denen die Attribute in das OpenLR-Format übersetzt werden können. Dies ist nicht immer eindeutig möglich. Etwaige Mehrdeutigkeiten können manchmal, jedoch nicht immer, unter Zuhilfenahme anderer Attributinformationen aufgelöst werden.

#### 3.4.1.1 Location String

Eine Ortsreferenz wird im Quellnetz mit Hilfe eines sogenannten *Location Strings* definiert. Dieser enthält alle relevanten Informationen für den OpenLR Encoder, um daraus eine OpenLR Ortsreferenz zu generieren, siehe folgendes Beispiel:

```
LIN;Loc_0057;0,0,7948,53453,53454
```

Das Format des Location Strings selbst ist nicht im OpenLR Whitepaper spezifiziert, wird jedoch in der hier angegebenen Form von der OpenLR Musterimplementierung verwendet. Der Location String besteht aus folgenden Teilen:

- Typ (LIN für linienhafte Ortsreferenz),
- Identifikator,
- Startoffset,
- Endoffset,
- Liste mit Kanten (spezifiziert durch ihre jeweilige ID) aus der Quellkarte.

In unserem Beispiel handelt es sich also um eine linienhafte Ortsreferenz mit der Bezeichnung Loc\_0057, sie hat weder Start- noch Endoffset und besteht in der Quellkarte aus drei Kanten mit den angegebenen Kanten-IDs.

#### 3.4.1.2 Location Reference Points (LRP)

Grundsätzlich können mit OpenLR punkt-, linien- und flächenhafte Ortsreferenzen codiert werden. Da für Anwendungen im Rahmen der vorliegenden Arbeit vor allem linienhafte Objekte relevant sind, soll speziell deren Codierung in OpenLR im Folgenden näher betrachtet werden. Das Prinzip von OpenLR basiert auf der Verkettung von kürzesten Wegen (engl. *shortest paths*) in einem Straßennetz. Die gesamte zu codierende Ortsreferenz soll mit diesen verketteten kürzesten Wegen abgedeckt sein. Jeder der kürzesten Wege wird durch Informationen über die Liniensegmente am Start und am Ende definiert. Diese Informationen werden in Form von sogenannten *Location Reference Points (LRP)* gespeichert.

Jeder LRP bezeichnet das erste Liniensegment eines kürzesten Weges, mit Ausnahme des letzten LRPs, der das letzte Liniensegment des letzten kürzesten Weges der Ortsreferenz bezeichnet. Der Anfangs- und Endpunkt einer Ortsreferenz werden durch Location Reference Points beschrieben. Weitere dazwischenliegende LRPs müssen dort eingefügt werden, wo die Ortsreferenz nicht durch den kürzesten Weg zwischen zwei LRPs beschrieben wird. LRPs müssen auf einem validen Netzknoten mit mehr als zwei abgehenden Kanten liegen. Trifft dies für die der Ortsreferenz am nächsten gelegenen Knoten nicht zu, so ist die Ortsreferenz so lange um zusätzliche Kanten zu erweitern, bis ein entsprechender echter Knoten erreicht wird. Neben den Punktkoordinaten des LRP werden noch weitere Eigenschaften der Bezugskante, auch *building blocks* genannt, für diesen LRP gespeichert, siehe Tabelle 3.1.

Tabelle 3.1: Location Reference Point building blocks

LRP Element	Beschreibung
Lon/Lat	Punktkoordinaten bezogen auf WGS84 [°]
Bear	Richtungswinkel des entspr. Straßensegments [°]
FRC	Functional Road Class
FOW	Form of Way
IFRC	Niedrigste Straßenklasse bis zum nächsten LRP
DNP	Distanz bis zum nächsten LRP [m]

Die Ortsreferenz kann durch optionale Offsets am Start- und Endpunkt ergänzt werden, wenn die Ortsreferenz nicht direkt in einem Knotenpunkt beginnt bzw. endet. Wurde die Ortsreferenz bis zu einem echten Netzknoten erweitert, so vergrößern sich auch die entsprechenden Offsets, sodass der tatsächliche Beginn bzw. das Ende der Location unverändert codiert wird. Die niedrigste Straßenklasse (FRC) bis zum nächsten LRP kann helfen, das Routing im Netz zu beschleunigen bzw. robuster zu machen. Dadurch können niedrigere, in der Ortsreferenz nicht vorkommende Straßenklassen von vornherein vom Routing ausgeschlossen werden. Jede OpenLR Ortsreferenz besteht somit aus folgenden Elementen:

1. Location ID
2. Offsets
3. Liste mit Location Reference Points

### 3.4.1.3 Location Reference

Die für alle Location Reference Points zusammen getragenen Informationen bilden die sogenannte *Location Reference*. In obigem Beispiel wurden zwei Location Reference Points identifiziert, es sind der erste und letzte Knoten der Ortsreferenz. Für sie wurden die zusätzlichen Eigenschaften aus der Quellkarte abgefragt und eine Location Reference erstellt, siehe Tabelle 3.2.

Tabelle 3.2: Location Reference im Encoder

Knoten	LRP	Lon	Lat	FRC	FOW	Bear	IFRC	DNP
7948	1	9.20381	48.77819	3	SINGLE_CW	197	3	674
53454	2	9.19839	48.77319	3	SINGLE_CW	28	-	-

Diese Location Reference Informationen werden dann in das binäre Übertragungsformat (Base64) umgewandelt. Daraus entsteht dann der folgende Location Reference Code:

```
CwaLgSKvzRtxC/3i/gwbAg==
```

Dieser wird dann mittels Datenverbindung in das Zielsystem übertragen und in der dortigen Karte decodiert.

### 3.4.2 Implementierung

Das OpenLR Whitepaper enthält neben den OpenLR Format-Definitionen auch eine Empfehlung für den Aufbau eines En- und Decoders. Diese Empfehlungen beschreiben auch den Ablauf der von TomTom kostenlos erhältlichen Muster-Implementierung (TOMTOM 2014). Sie sind jedoch nicht Teil des OpenLR Standards. Die wesentlichen Prozessschritte dieser Muster-Implementierung, bestehend aus dem Encodieren einer Ortsreferenz im Quellnetz, der Übertragung des erzeugten Location Reference Codes in das Zielsystem und das dortige Decodieren, sind in Abbildung 3.8 vereinfacht dargestellt. Dabei wird im Encoder zunächst geprüft, ob die Ausgangs-Ortsreferenz ggfs. bis zum nächsten echten Netzknoten erweitert muss. Dann folgt eine Prüfung, ob die Ortsreferenz vollständig durch eine kürzeste-Wege-Verbindung im Netz abgedeckt ist. Ist das nicht der Fall, so müssen weitere Zwischenpunkte eingefügt werden. Dann folgt die eigentliche Codierung und Übertragung der Ortsreferenz ins Zielnetz. Dort wird die empfangene Ortsreferenz decodiert und für jeden LRP geeignete Kandidaten im Zielnetz gesucht, bewertet und der jeweils Beste ausgewählt. Aus diesen kann dann schließlich die decodierte Ortsreferenz im Zielnetz zusammengesetzt werden, an der ggfs. noch die vorhandenen geometrischen Offsets am Beginn bzw. Ende angebracht werden müssen. Diese Prozessschritte des Codierens bzw. Decodierens werden in den folgenden Abschnitten detailliert anhand einem Beispiel erläutert. Die dabei dargestellten numerischen Werte sind ebenso den *OpenLR-Encoder-Properties* bzw. den *OpenLR-Decoder-Properties* aus der OpenLR Musterimplementierung (TOMTOM 2014) entnommen.

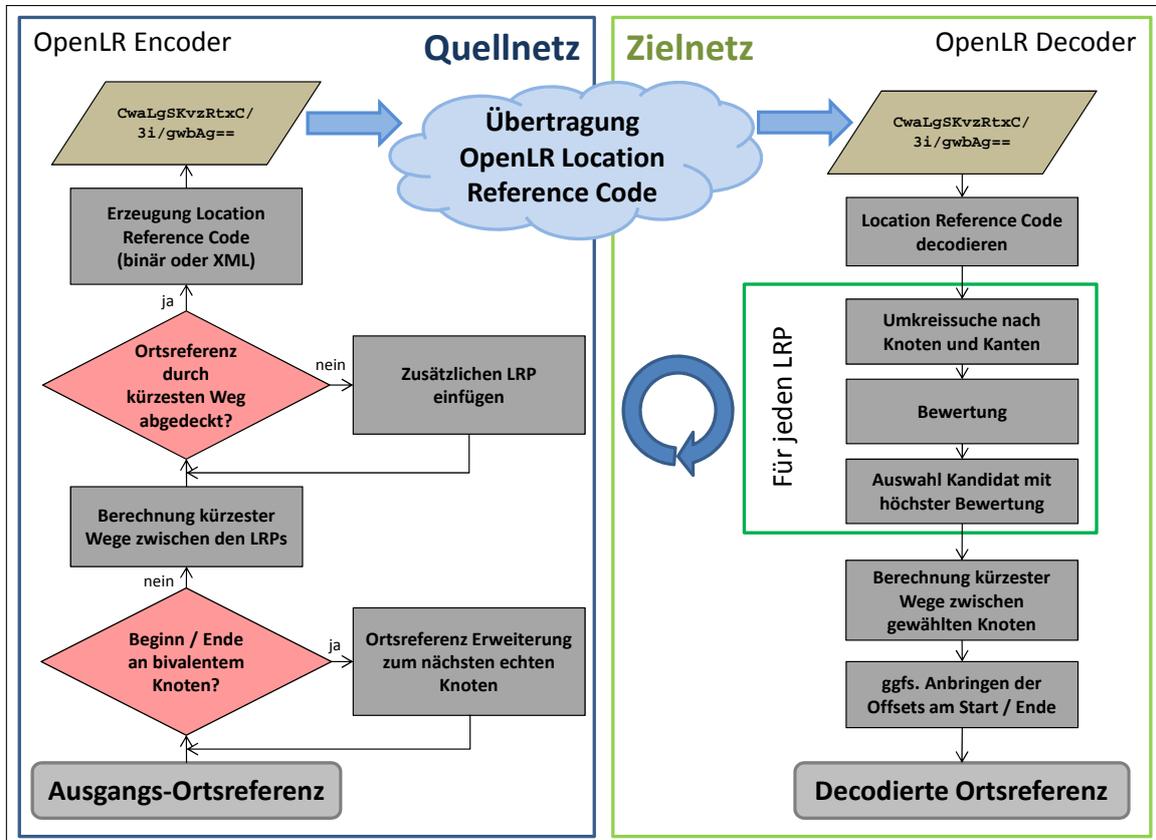


Abbildung 3.8: Ablaufdiagramm OpenLR Encoding/Decoding (vereinfacht)

### 3.4.2.1 Encoder

Abbildung 3.9 zeigt ein Beispiel für die Codierung einer OpenLR Line Location aus dem OpenLR Whitepaper (TOMTOM 2012). Das erste Bild 3.9(a) zeigt die zu codierende Line Location, sie beginnt an Knoten 3 und endet in Knoten 15. Da der Startknoten 1 nur eine eingehende und eine ausgehende Kante besitzt, also bivalent ist, sucht der Encoder entsprechend dem OpenLR Standard nach dem nächsten echten Netzknoten. Diesen findet er in Knoten 1 und wählt diesen als Startknoten für die zu codierende Ortsreferenz. Der geometrische Abstand zum eigentlichen Beginn der Ortsreferenz bei Knoten 3 wird als Offset übertragen und bleibt erhalten. Knoten 1 und 15 stehen somit als Location Reference Point fest. Als nächstes berechnet der Encoder den kürzesten Weg zwischen dem Startknoten 1 mit Kante nach 3 und dem Endknoten 15 (Abbildung 3.9(b)). Bei der darauffolgenden Prüfung wird festgestellt, dass der gewünschte Locationverlauf und der berechnete kürzeste Weg nicht übereinstimmen. Daher ist es notwendig, bei Knoten 10 einen zusätzlichen Location Reference Point (LRP) in Richtung Knoten 11 einzufügen. Berechnet man nun die beiden kürzesten Wege zwischen Knoten 1 und 10 sowie zwischen 10 und 15, so wird die gewünschte Ortsreferenz vollständig durch diese verketteten kürzesten Wege abgedeckt. Die Ortsreferenz kann also durch drei LRPs (1, 10 und 15) in OpenLR referenziert werden (Abbildung 3.9(c)).

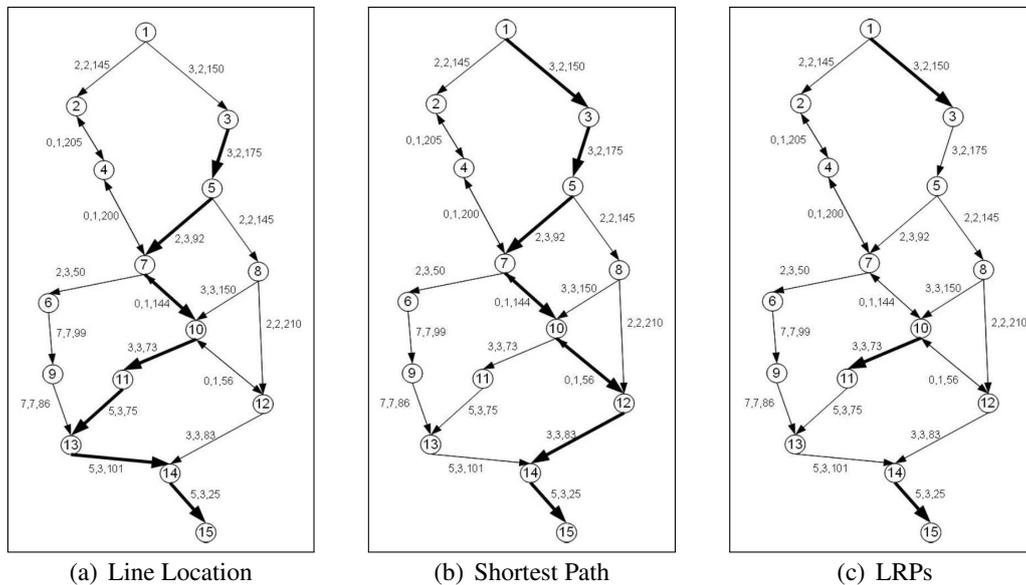


Abbildung 3.9: Beispiel OpenLR Location Encoding (TOMTOM 2012)

### 3.4.2.2 Identifikation der Stützpunkte (LRPs) im Zielnetz

Die Rekonstruktion der gesuchten Ortsreferenz im Zielnetz erfolgt durch die Suche nach der optimalen Zuordnung der zu decodierenden Ortsreferenz zu Objekten (Kanten) des Zielnetzes. In OpenLR geschieht dies bislang mittels eines punktbasierten Verfahrens: Die Ortsreferenz im Quell- oder Ausgangsnetz wird durch sog. *Location Reference Points (LRP)* repräsentiert. Diese werden dann zusammen mit Geometrie- und Attributinformationen der von ihnen ausgehenden Bezugskante in OpenLR codiert und übertragen. Als Location Reference Points werden in der Regel Kreuzungspunkte mit drei oder mehreren abgehenden Kanten (oder topologisch gesprochen: echte Knoten im Netzwerk) verwendet. Unechte (oder auch: bivalente) Knoten, mit nur zwei abgehenden Kanten, besitzen meist nicht genügend einzigartige Eigenschaften, um sie von umliegenden anderen bivalenten Knoten zu unterscheiden und mit genügend hoher Wahrscheinlichkeit eindeutig im Zielnetz identifizieren zu können.

### Suche nach Zuordnungskandidaten

Im Zielnetz muss diejenige Kante identifiziert werden, die am wahrscheinlichsten der gesuchten entspricht. Dazu werden zunächst in einem Suchradius (standardmäßig:  $MaxNodeDist = 100\text{ m}$ ) um die Koordinaten des übertragenen Location Reference Points sämtliche Knoten in der Zielkarte und die mit ihnen verbundenen Kanten bestimmt. Diese werden einer Vorauswahl unterzogen (siehe nächster Abschnitt). Für das gewählte Beispiel zeigt Abbildung 3.10(a) die im Zielnetz identifizierten Knoten-Kandidaten.

In einem zweiten Suchlauf werden nun Kanten, die innerhalb eines Suchradius um die Koordinaten des gesuchten Location Reference Point liegen, identifiziert. So können Kanten gefunden werden, deren Anfangs- oder Endpunkt nicht im zuvor abgesuchten Suchraum liegen, deren Verlauf aber dennoch eine gewisse räumliche Nähe zum Suchgebiet aufweist. Für diese direkt gefundenen Kanten, wird nicht ihr Start-/Endknoten für das Decoding betrachtet, es wird der Location Reference Point auf die Kante projiziert (gelotet).

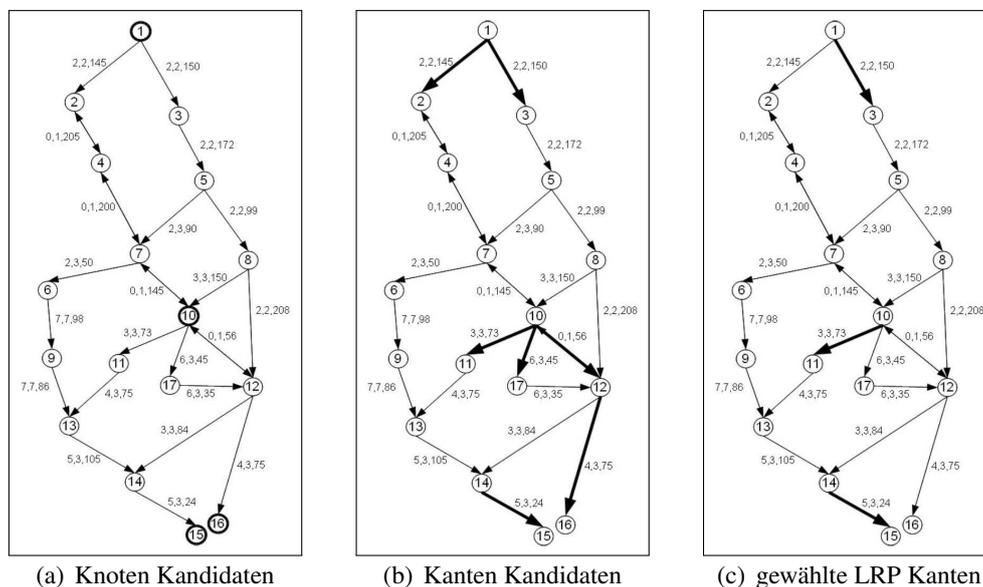


Abbildung 3.10: Beispiel OpenLR Location Decoding - 1 (TOMTOM 2012)

### Vorauswahl (Suchraumreduktion)

Kanten, die entgegengesetzt der gesuchten Richtung verlaufen oder eine zu niedrige Straßenklasse haben, werden in einer Vorauswahl aus der Menge der gefundenen Kanten verworfen. Die verbleibenden Kanten werden dann mit den mittels OpenLR aus der Ausgangskarte übertragenen Geometrie- und Attributinformationen verglichen und bewertet. Dies erfolgt über die Rating bzw. Bewertungsfunktion, die auch zur finalen Entscheidungsfindung herangezogen wird (siehe nächster Abschnitt). Alle Kanten, die nicht mindestens eine Bewertung von  $MinimumAcceptedRating = 800$  erlangt haben, werden verworfen. Bei Kanten, die nicht mit ihrem Start-/Endknoten, sondern über einen Lotpunkt gefunden wurden, wird eben dieser Lotpunkt auch für die Distanz-Bewertung verwendet. Da dieser Lotpunkt dem Location Reference Point sehr nahe liegt, erhalten die 'direkt' gefundenen Kanten in der Regel eine sehr gute Distanz-Bewertung. Eigentlich sind die Location Reference Points aber als echte Netzknoten definiert, ein Decoding auf einen eingerechneten Zwischenpunkt ist also lediglich 'zweite Wahl'. Um dies auch im Ergebnis zu erreichen, wird die Bewertung aller 'direkt' gefundenen Kanten noch mit

einem Maßstab  $LinesDirectlyFactor = 0,95$  reduziert. Dies soll sicherstellen, dass die über Knoten gefundenen Kanten bevorzugt in die finale Bewertung kommen. Abbildung 3.10(b) zeigt die im Beispiel ausgewählten Kanten Kandidaten.

### Entscheidungsfindung / Rating

Für die Decodierung jedes einzelnen Location Reference Points wird die am besten übereinstimmende Kante im Zielnetz gesucht. Für diese Auswahl werden alle für den jeweiligen LRP gefundenen Kanten Kandidaten entsprechend dem Grad ihrer Übereinstimmung mit dem gesuchten Location Reference Point bewertet. Dabei wird jeder LRP einer Ortsreferenz einzeln und nur für sich betrachtet. Die Bewertung stützt sich auf den Vergleich der mittels OpenLR von der Ausgangslocation übertragenen Informationen mit denen des jeweiligen Kandidaten. Für jede Teilwertung wird ein Qualitätsmaß berechnet, das Werte zwischen 0 und 100 annehmen kann.

**Entfernung:** Zunächst gilt die Euklidische Distanz der Start-/Endknoten (bzw. des Lotpunktes) zwischen gesuchtem LRP aus dem Quellnetz und gefundenen Kandidaten im Zielnetz  $d_{LRP,Cand}$  als wichtiges geometrisches Bewertungsmaß (vgl. Kapitel 2.6.1.1). Die Differenz zur maximal akzeptierten Entfernung  $d_{max}$  ( $= MaxNodeDist$ , s.o.) bildet die Bewertung  $R_{Node}$ :

$$R_{Node} = d_{max} - d_{LRP,Cand} \quad (3.1)$$

**Richtungswinkel:** Daneben wird auch der Richtungswinkel der Kanten als geometrisches Ähnlichkeitsmaß verglichen (vgl. Kapitel 2.6.1.2). Für jeden Location Reference Point (LRP) wird auch der Richtungswinkel  $t_{LRP}$  der entlang des Locationverlaufs abgehenden (beim letzten LRP: eingehenden) Kante übertragen. Dieser Richtungswinkel kann mit denjenigen der Kanten-Kandidaten  $t_{Cand}$  des jeweiligen LRPs verglichen werden. Um durch Modellierungsunterschiede möglicherweise auftretende Richtungswinkeldifferenzen kurzer Kanten im Kreuzungsbereich auszuschließen, wird zur Richtungswinkelbestimmung immer ein Punkt mit einer anwendungsspezifisch zu definierenden Entfernung  $d_{bear}$  in den Locationverlauf eingerechnet. Der Vergleich erfolgt durch Differenzbildung:

$$\Delta t = |t_{LRP} - t_{Cand}| \quad (3.2)$$

Die Richtungswinkeldifferenzen werden entsprechend den in Tabelle 3.3(a) beschriebenen Klassen in Rating-Werte überführt. Die Einteilung der Bewertungsskala für die Richtungswinkeldifferenzen  $\Delta t$  von  $6^\circ$  entsprechen ungefähr der halben OpenLR Winkelauflösung von  $11,25^\circ$  bei binärer Übertragung (vgl. Abschnitt 3.4.1). Da die Richtungswinkeldifferenz als Betrag angegeben wird (vgl. Abschnitt 3.4.2.2) liegt ein Kandidat mit einer Richtungswinkeldifferenz von  $\Delta t = \pm 6^\circ$  im ersten Sektor um den gesuchten Richtungswinkel des entsprechenden Location Reference Points (LRP) aus der Quellkarte.

**Functional Road Class (FRC):** Als semantisches Ähnlichkeitsmaß (siehe Kapitel 2.6.4) wird zunächst die im Quellnetz encodierte *Functional Road Class* mit den entsprechenden Attributwerten der Kandidaten im Zielnetz verglichen. Dieser Vergleich erfolgt durch Bildung der Differenz  $\Delta_{FRC}$  zwischen der numerisch modellierten Straßenklasse des übertragenen LRPs  $FRC_{LRP}$  und dem Kandidaten in der Zielkarte  $FRC_{Cand}$ :

$$\Delta_{FRC} = |FRC_{LRP} - FRC_{Cand}|, \quad FRC \in \{1, \dots, 7\} \quad (3.3)$$

Tabelle 3.3(b) zeigt die Definition der Bewertungsklassen.

Tabelle 3.3: Rating - OpenLR Musterimplementierung

(a) Bearing		(b) FRC	
$\Delta t$	$R_t$	$\Delta FRC$	$R_{FRC}$
$< 6^\circ$	100	0	100
$< 12^\circ$	50	1	75
$< 18^\circ$	25	2	50
$\geq 18^\circ$	0	$> 2$	0

**Form of Way (FOW):** Das Attribut *Form of Way* wird als weiteres semantisches Ähnlichkeitsmaß ausgewertet. Da die Attributwerte bei Form of Way eine dezidierte Bedeutung haben, die nicht wie bei FRC numerisch modelliert werden kann (z.B. *multiple Carriageway* oder *Roundabout*), erfolgt der Vergleich zwischen den FOW-Attributwerten der Ausgangs-Location aus der Quellkarte  $FOW_{LRP}$  und den Routenkandidaten in der Zielkarte  $FOW_{Cand}$  mit Hilfe einer Bewertungstabelle (engl. *Rating Table*). In dieser werden alle möglichen Attributwert-Kombinationen abgebildet und ihnen jeweils ein Qualitätsmaß zugeordnet. Entsprechend dem Grad der Übereinstimmung der beiden zu vergleichenden FOW-Attributwerte erfolgt die Zuweisung des Bewertungsmaßes  $R_{FOW}$  von 100 für eine sehr gute Übereinstimmung, 50 für gute und 25 für eine schlechte Übereinstimmung der zu vergleichenden FOW-Attributwerte, siehe Tabelle 3.4.

Tabelle 3.4: Bewertungstabelle „Form of Way“ (TOMTOM 2014)

	Undef.	MW	MCW	SCW	RA	TS	SR	OTHER
UNDEFINED	50	50	50	50	50	50	50	50
MOTORWAY	50	100	50	25	25	25	25	25
MULTIPLE_CARRIAGEWAY	50	50	100	50	50	25	25	25
SINGLE_CARRIAGEWAY	50	25	50	100	50	50	25	25
ROUNDABOUT	50	25	50	50	100	50	25	25
TRAFFIC_SQUARE	50	25	25	50	50	100	25	25
SLIPROAD	50	25	25	25	25	25	100	25
OTHER	50	25	25	25	25	25	25	100

### Berechnung der Leistungsfunktion

Die vorgenannten einzelnen Bewertungsmaße werden dann für jeden LRP Decoding-Kandidaten mit Hilfe einer Leistungsfunktion (siehe Kapitel 2.7.2) zu einer OpenLR Gesamtbewertung  $R_{OpenLR}$  zusammengeführt. Neben dem Anteil für Knoten  $R_{Node}$  werden die Einzelbewertungen für den Richtungswinkel  $R_t$ , die Straßenklasse  $R_{FRC}$  und für Form of Way  $R_{FOW}$  zu einer Kanten-Bewertung zusammengefasst und jeweils mit einem Faktor  $f_n$  und  $f_e$  gewichtet:

$$R_{OpenLR} = R_{Node} \cdot f_n + (R_t + R_{FRC} + R_{FOW}) \cdot f_e. \quad (3.4)$$

Die OpenLR Musterimplementierung verwendet eine gleiche Gewichtung von Knoten- und Kanten-Bewertung, die Gewichtungsfaktoren werden mit  $f_n = 3$  und  $f_e = 3$  gewählt. Es werden nur Kandidaten ab einem Leistungsfunktionswert von  $MinimumAcceptedRating = 800$  für die finale Auswahl herangezogen. Die Kante mit dem höchsten Wert der Leistungsfunktion wird als die am wahrscheinlichsten dem LRP in der Quellkarte entsprechende Kante ausgewählt und für die im Weiteren näher beschriebene

Routenberechnung zwischen den einzelnen LRPs verwendet. In Abbildung 3.10(c) sind die ausgewählten Kanten mit der höchsten Bewertung hervorgehoben.

### 3.4.2.3 Rekonstruktion des Location-Verlaufs und Anbringen der Offsets

Zwischen den im vorigen Kapitel genannten, im Zielnetz decodierten Location Reference Points wird (bei linienhaften Ortsreferenzen) zur Rekonstruktion des Location-Verlaufs mit Hilfe eines Routing Algorithmus die kürzeste Verbindung im Netz berechnet. Abbildung 3.11(a) zeigt den für das Beispiel im Zielnetz mittels kürzester Wege zwischen den gewählten LRPs rekonstruierten Locationverlauf.

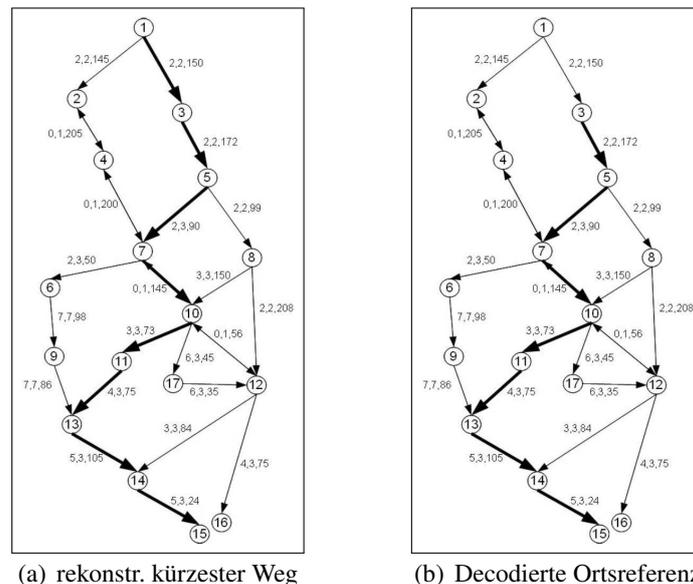


Abbildung 3.11: Beispiel OpenLR Location Decoding - 2 (TOMTOM 2012)

Wurde der gesamte (linienhafte) Location Verlauf im Zielnetz erfolgreich rekonstruiert, müssen zuletzt noch die aus dem Quellnetz mittels OpenLR übermittelten Offsets am Beginn und Ende des Route angebracht werden. Bei Übertragung der OpenLR Location Strings im Binärformat (Standardfall) werden die Offsets nicht in ihrem ursprünglichen metrischen Format übertragen. Laut der Beschreibung für das Binärformat im OpenLR Whitepaper TOMTOM (2012) werden die Offsets in diesem Fall relativ zur Länge des Pfades zwischen den beiden relevanten Location Reference Points übertragen.

Dementsprechend sind die metrischen Offsets beim Decodieren aus den Pfadlängen der decodierten Ortsreferenz zu berechnen. Sind die erhaltenen Offsets länger als die erste bzw. letzte Kante, so ist die Ortsreferenz um diese Kanten zu kürzen. Dieser Vorgang wird so oft wiederholt, bis die Offsets kürzer als die Länge der entsprechenden Kanten sind. Dann liegen die Offsets auf der ersten bzw. der letzten Kante der Ortsreferenz. Dieses Kürzen der Ortsreferenz ist insbesondere dann notwendig, wenn die Ortsreferenz 'erweitert' werden musste, da sie nicht auf einem validen Netzknoten begann bzw. endete (siehe Abschnitt 3.4.3.2). Die berechneten Offsets werden dann der Kantenfolge der decodierten Ortsreferenz beigelegt. Im Beispiel war ein Startoffset aus der Erweiterung der Ortsreferenz zum nächsten echten Knoten entstanden, der nun in Abbildung 3.11(b) seine Berücksichtigung fand. Somit beginnt die decodierte Ortsreferenz nun bei Knoten 3 im Zielnetz.

### 3.4.3 Analyse

In diesem Abschnitt soll die im vorigen Kapitel beschriebenen OpenLR Spezifikation und ihre Umsetzung in der OpenLR Musterimplementierung kritisch analysiert und Schwachstellen identifiziert werden. Diese bieten dann Möglichkeiten zur weiteren Verbesserung des Verfahrens. Diese Verbesserung ist eines der Ziele der vorliegenden Arbeit. Dazu wurden umfangreiche Experimente mit der OpenLR Muster-Implementierung innerhalb der LR-Testumgebung durchgeführt. Dabei konnten einige 'Schwachstellen' identifiziert werden. Diese werden im Folgenden erläutert.

#### 3.4.3.1 Attribut-Mapping bei differierenden Definitionen

##### Schema-Matching (mit LR Zwischenschicht)

Location Referencing / OpenLR ermöglicht den Austausch von Informationen zwischen Kartensystemen unterschiedlicher Hersteller und Versionen (TOMTOM 2012, WEVERS et al. 2001, WEVERS und HENDRIKS 2006, WEVERS 2012). Bei einer direkten Übertragung von Informationen zwischen zwei unterschiedlichen Karten müssten Zuordnungsvorschriften für die verwendeten Attribute aufgestellt werden. Damit können die Informationen aus dem Schema der Ausgangs- bzw. Quellkarte in das Datenschema der Zielkarte überführt werden und sind mit den Objekten und deren Attribute in der Zielkarte vergleichbar. Eine Übertragung mittels Location Referencing darf nicht an eine bestimmte Sender-/Empfänger Kombination gebunden sein. Somit ist zum Zeitpunkt des Codierens in der Zielkarte das verwendete Kartensystem auf der Empfängerseite nicht bekannt. Deshalb definieren Location Referencing Verfahren jeweils ein eigenes, proprietäres Attributschema. Dieses dient sozusagen als Zwischenschicht für die Übertragung. Für alle Sender und Empfänger müssen dann Zuordnungsvorschriften vom jeweils eigenen Datenschema in das des verwendeten Location Referencing Verfahrens spezifiziert werden. Damit können die für die Übertragung und vor allem die Rekonstruktion einer Ortsreferenz im Zielnetz benötigten Objekt-/Ortsreferenz-Eigenschaften (wie z.B. Functional Road Class) in das standardisierte LR-Attribut-Schema überführt, übertragen und vor dem Decodieren in das Schema des Zielsystems überführt werden.

##### Verwendung der OpenLR Attribute

Bei den im Rahmen dieser Arbeit mit TomTom MultiNet und OpenStreetMap durchgeführten Tests sowie bei LR-Anwendungen im Rahmen des ROSATTE-Projektes (SCHÜTZLE und FRANK 2010) wurden Probleme vor allem mit den in OpenLR spezifizierten Attributen *Functional Road Class (FRC)* und *Form of Way (FOW)* (TOMTOM 2012) offenkundig. Da die Attribut-Modellierung von OpenLR derjenigen der TomTom MultiNet sehr ähnelt, treten die Schwierigkeiten vor allem bei der Attributzuordnung zwischen OpenLR und Karten anderer Anbieter (wie beispielsweise HERE oder OpenStreetMap) auf.

Die OpenLR Attribute *Functional Road Class (FRC)* und *Form of Way (FOW)* sind kartenherstellerspezifische Attribute. Sie sind in dieser Form nicht mit den von den jeweils zuständigen Straßenbetreibern / Behörden vergebenen Straßenklassifizierern identisch. Attribute wie FRC oder FOW wurden von den Kartenherstellern beim Aufbau ihrer Kartendatenbanken eingeführt, um eine länderübergreifende und innerhalb der jeweiligen Kartendatenbank einheitliche Beschreibung der Straßenklasse und des Ausbaustandes zu ermöglichen. Dies ist für alle Anwendungen notwendig oder zumindest hilfreich, die lediglich mit nur einer digitalen Straßenkarte arbeiten (wie z.B. Routing-Algorithmen in Navigationssystemen). Diese können auf das jeweils der verwendeten Karte zugrunde liegende Datenmodell angepasst und dann für das gesamte in der Datenbank abgebildete Gebiet unabhängig von Ländergrenzen

verwendet werden. Man kann sagen, dass solche Hersteller-spezifischen Attribute eine hohe Verfügbarkeit haben (außer bei nicht-kommerziellen Karten wie beispielsweise FOW in OpenStreetMap, siehe Kapitel 4.2.2.2). Die Attribute sind jedoch (teilweise stark) unterschiedlich in den verschiedenen Datenbank Schemata modelliert. So wird beispielsweise die FRC in OpenLR mit acht Klassen modelliert. Es ist jedoch bekannt, dass HERE (früher: Navteq) für seine Karten nur fünf Straßenklassen verwendet (CHEN et al. 2008). Für jede dieser Datenquellen müsste also eine gesonderte Zuordnungsvorschrift zu der von OpenLR spezifizierten *Functional Road Class* definiert werden.

Die Modellierung der Zuordnungsvorschriften ist also wie zuvor beschrieben, ein komplexes Thema. Unschärfen, Mehrdeutigkeiten oder gar Fehler in den Zuordnungen führen zu einer verfälschten Klassifizierung der gesuchten Straßenklasse im Zielnetz und dadurch zu einer Verfälschung der Ortsreferenz-Zuordnung.

### Verwendung staatlich definierter Attribute

Weisen die zur Zuordnung herangezogenen Attribute große Modellierungsunterschiede auf, so kann sich dies, ähnlich wie auch im vorigen Abschnitt beschrieben, je nach Konstellation aus Sender- und Empfänger-Karte äußerst nachteilig auf die Zuordnungsqualität auswirken (siehe auch Kapitel 4.2). Für das Location Referencing nahezu ideal wären hingegen Attribute, die in den verschiedenen digitalen Straßenkarten möglichst gleich modelliert sind. Das ist vor allem für solche Attribute der Fall, die die offiziellen, vom jeweiligen Straßenbetreiber vergebenen, Eigenschaften abbilden. Das können beispielsweise der *Straßenname* und die *Straßenklassifizierung* (in Deutschland also z.B. 'A' oder 'B' für Autobahn oder Bundesstraße) mit der jeweiligen Nummer sein. Schon AGORA und AGORA-C verwendeten Straßenummer und Straßenname im sogenannten *road descriptor* (WEVERS et al. 2003, HENDRIKS et al. 2005). Allerdings wird je nach Verfügbarkeit nur die Straßenummer oder der Straßenname übertragen.

In OpenStreetMap werden diese Informationen unter dem Tag: 'ref' geführt, bei TomTom MultiNet unter den Attributen 'NAME' bzw. 'SHIELDNUM'. Diese sind direkt mit dem Realwelt-Objekt verbunden und sollten somit keine Modellierungsunterschiede zwischen den einzelnen Straßenkarten ermöglichen. Jedoch ist die flächendeckende und kartenübergreifende Verfügbarkeit nicht garantiert. Sie ist abhängig davon, ob der jeweilige Hersteller die Attribute der Realwelt-Objekte auch tatsächlich erfasst und in seinen Kartendatenbanken abbildet. Attributinformationen, die von kommerziellen Kartenherstellern geführt und in ihren Produkten angeboten werden, sind meist auch flächendeckend erfasst. Ganz anders ist die Situation hingegen bei OpenStreetMap. Hier liegt der Fokus der Erfassung zunächst klar auf der Geometrie und wichtigen Tags wie 'highway' (Straßenklasse). Andere Tags wie das hier angesprochene 'ref' sind zwar im Datenmodell definiert, nach den Erfahrungen des Autors jedoch nur sehr nachrangig erfasst. Möglicherweise führen jedoch die andauernden Arbeiten zur Vervollständigung der OpenStreetMap zu einer zukünftigen Verbesserung der Situation und einer miteinhergehenden Verwendbarkeit dieser Informationen auch für LR-Anwendungen.

#### 3.4.3.2 Locationerweiterung bei bivalenten Knoten

Wird eine zu encodierende Ortsreferenz nur von bivalenten Knoten begrenzt, dann erweitert der Encoder die Ortsreferenz bis zum nächsten validen Netzknoten mit mindestens 3 abgehenden Kanten und setzt dort einen Location Reference Point für diese Ortsreferenz (vgl. Abschnitt 3.4.1.2). Dabei wird der Location-Offset an dieser Stelle entsprechend um die Länge der hinzugefügten Kanten im Quellnetz vergrößert. Dadurch können zum Teil sehr lange Offsets entstehen, vor allem außerhalb geschlossener Ortschaften oder auf Autobahnen. Durch das Abtragen der langen Offsets und den variierenden Kantenlängen entstehen jedoch Ungenauigkeiten an den jeweils betroffenen Start- und Endpunkten der Ortsreferenz im Zielnetz. Darüber hinaus entfällt der für die ursprüngliche Definition der Ortsreferenz

relevante Bezug zu diesem bivalenten Knoten, soweit dieser im Zielnetz vorhanden wäre. Dadurch entstehen Punkte, die nach Abtragen des Offsets in die dortige Kante eingerechnet werden, d.h. der ursprünglich definierte Knoten wird i.d.R. nicht aufgefunden. Dies führt zu Versätzen und/oder Überständen dem Decodieren im Zielnetz.

Bivalente Knoten am Location-Beginn/Ende entstehen hauptsächlich aus 2 Gründen:

**(1) Shape Points:** Formbeschreibende Punkte werden zur geometrischen Beschreibung einer Kante eingefügt, meist jedoch ohne das explizite Aufbrechen einer Kante. Des Weiteren können sie auch durch die Datenerfassung durch mehrere, unabhängig arbeitende Einheiten beispielsweise an den jeweiligen Bearbeitungsgrenzen entstehen. Häufig sind solche Zwischenpunkte Reste früherer Kanten, die später geändert oder ergänzt wurden. Bei Karten kommerzieller Hersteller ist dies durch das professionelle Qualitätsmanagement und die Arbeit von automatischen Algorithmen zur Fehlerdetektion und Fehlerkorrektur bzw. den Einsatz von Editoren eher selten der Fall. Bei Karten, die durch freie Communities erstellt bzw. gepflegt werden (wie z.B. OSM) ist das Auftreten solcher Punkte eher möglich bzw. auch im Rahmen dieser Arbeit zu Tage getreten.

**(2) Wechsel von Attributwerten:** Häufig werden bivalente Knoten eingefügt, um den Wechsel von Attributwerten in der Karte zu modellieren. Ändert sich an einer Stelle beispielsweise die Geschwindigkeitsbegrenzung oder verläuft die Straße in einem Tunnel oder einer Brücke, muss dafür der jeweilige Attributwert geändert werden. Aufgrund des Datenmodells kann aber für jedes Attribut eines Objektes (Kante) immer nur ein Wert angegeben werden. Daher müssen Objekte zur Modellierung von Attributwertänderungen aufgeteilt werden.

Das Beispiel in Abbildung 3.12 zeigt die Ausgangs-Location im Quellnetz [blau], welche an einem bivalenten Knoten [roter Punkt] beginnt. Dieser markiert das Ende eines Straßentunnels. In der Zielkarte [schwarz] ist dieses Tunnelende ebenfalls als bivalenter Knoten vorhanden, jedoch mit einem Abstand von ca. 12 m [roter Pfeil]. Der Location Reference Point zu diesem Locationbeginn liegt in ca. 2,7 km Entfernung am anderen Ende des Tunnels. Geometrisch gesehen wurde der Location Beginn mit guter Genauigkeit ins Zielnetz decodiert. Jedoch bietet OpenLR in seiner bisherigen Form keine Möglichkeit, eine semantische Bedeutung der codierten Punkte, wie in diesem Beispiel, zu übertragen. Daher kann im Zielnetz der durchaus vorhandene semantisch entsprechende Knoten nicht gefunden werden, sondern lediglich ein eingerechneter Zwischenpunkt in der Nähe.

Grundsätzlich erscheint es sinnvoll und richtig, mit OpenLR zu übertragende Ortsreferenzen bis an echte Netzknoten mit drei oder mehr abgehenden Kanten zu erweitern. Dies erhöht die Wahrscheinlichkeit, diesen Knoten im Zielnetz aufzufinden. Den Bezug zum eigentlich gewählten Start-/Endpunkt der Ortsreferenz (wie von OpenLR im Whitepaper definiert und so auch implementiert) aber lediglich durch einen metrischen Abstand (Offset) herzustellen, ist für ein präzises Decodieren im Zielnetz nicht ausreichend. Es entfällt daher die für das Decodieren relevante Information, ob die so übertragene Ortsreferenz an einem bivalenten Knoten begonnen oder geendet hat, oder ob der gewählte Offset einen Locationbeginn abseits eines Netzknotens beschreibt (Normalfall) und der Decoder den entsprechenden Punkt in die entsprechende Kante einzurechnen hat. Das Anfügen eines entsprechenden Flags zur Festlegung, ob der durch den Offset beschriebene Location-Beginn bzw. das Location-Ende einen Knoten oder einen (einzurechnenden) Punkt auf einer Kante beschreibt, könnte dieses Problem lösen.

### 3.4.3.3 Suchradius zur Knoten- und Kantendetektion

Bei den im Rahmen der vorliegenden Arbeit durchgeführten Untersuchungen kam es häufiger vor, dass die Kanten der korrekten Zuordnungslösungen gar nicht in der Menge der zu untersuchenden Kandidaten enthalten waren. Dies war hauptsächlich in den Datensätzen *Inter-Urban* und *Motorway* (vgl.

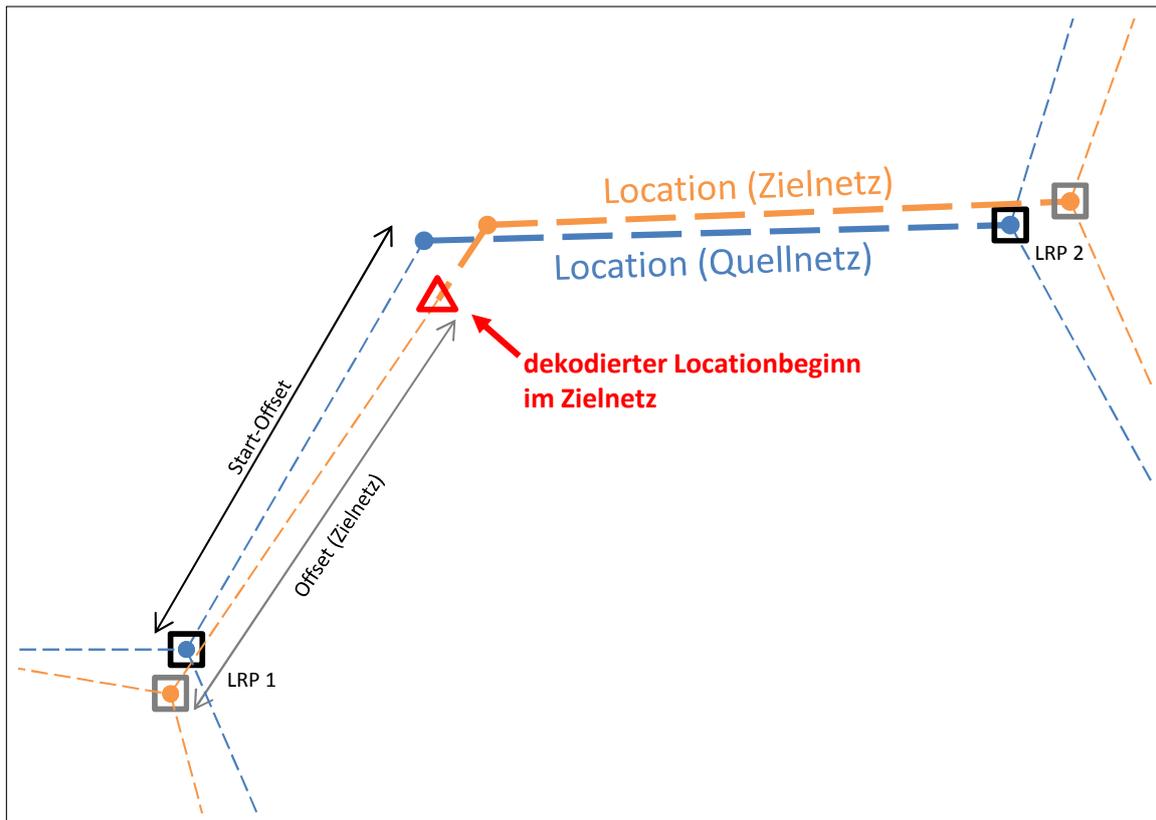


Abbildung 3.12: Locationbeginn an bivalentem Knoten

Kapitel 7.2.2) der Fall. Bei näherer Betrachtung stellte sich der in der OpenLR Musterimplementierung statisch definierte Suchradius *Maximum Node Distance* als Ursache für das Problem heraus. Dieser Wert ist in den Standard-Einstellungen mit 100 m gewählt. Aus den für die vorliegende Arbeit durchgeführten praktischen Untersuchungen erscheint dies ein pragmatisch gewählter Kompromiss zu sein. Jedoch werden die Kandidatenmengen innerorts eher zu groß, weil hier wegen dichter Bebauung die Straßen und Kreuzungen eine größere räumliche Dichte aufweisen. Eine solche Vergrößerung der Kandidatenmenge eröffnet die Möglichkeit einer falschen Zuordnung, da unter den vielen Kandidaten möglicherweise auch solche sind, deren vom LR-Verfahren untersuchte Eigenschaften besser zu denen der gesuchten Kante passen, als die eigentlich korrekt Zuzuordnende. Außerorts hingegen besitzen die Verkehrsanlagen typischerweise eine größere räumliche Ausdehnung, als dies in bebauten Gebieten der Fall ist. Dies hängt auch mit den höheren Geschwindigkeiten zusammen, die auf außerörtlichen Straßen gefahren werden können. Dadurch sind auch größere geometrische Kartenunterschiede zu beobachten, denn die Lage der sie repräsentierenden Knoten ist beispielsweise für solche langgestreckten Kreuzungs- und Verzweigungsanlagen vor Ort nicht immer eindeutig, was zu Unterschieden in der Erfassung durch die Kartenhersteller führt. Dies führt im Endeffekt dazu, dass die Kandidatenmengen, die ein fester Suchradius von 100 m liefert, häufig zu klein und in manchen Fällen die korrekten Lösungen überhaupt nicht unter den Kandidaten sind. In solchen Fällen wird dann oft die Zielkante mit projiziertem Locationpunkt durch die geometrische Suche gefunden, denn dieser Punkt (entspricht der kürzesten Entfernung des LRP aus der Quellkarte zur Zielkante) liegt meist innerhalb des 100 m Suchradius. Dadurch wird in solchen Fällen genau dieser projizierte Punkt als Decoding-Lösung angegeben.

Dieses Problem könnte mit einer dynamischen Berechnung des Suchradius  $r$ , beispielsweise auf Grundlage der Straßenklasse oder der zulässigen Höchstgeschwindigkeit, gelöst oder wenigstens vermindert werden. Ziel wäre eine Verkleinerung des Suchradius in bebauten Gebieten auf der einen Seite, sowie eine Vergrößerung des Suchradius außerorts, etwa auf Landstraßen oder Autobahnen.

#### 3.4.3.4 Knoten Valenz der LRPs

Location Reference Points werden entsprechend dem OpenLR Whitepaper TOMTOM (2012) nur auf echten Knoten mit Valenz  $> 2$  gebildet, da diese 'echten' Netzknoten im Gegensatz zu zweiwertigen Knoten tatsächlich auch echte Verzweigungen im Netz repräsentieren und darum mit einer höheren Wahrscheinlichkeit auch im Zielnetz vorhanden und für den LR-Decoder auffindbar sind (vgl. Kapitel 3.4.1.2). Beginnt oder endet eine Ortsreferenz nicht an einem mindestens zweiwertigen Knoten, so wird diese bis zum nächsten echten Netzknoten erweitert, wie in Abschnitt 3.4.3.2 beschrieben. Diese Eigenschaft der LRPs sollte deshalb auch beim Decoding ausgenutzt werden. Die OpenLR Musterimplementierung macht davon allerdings keinen Gebrauch (TOMTOM 2014). Alle im Zielnetz innerhalb des Suchradius gefundenen Knoten werden gleich behandelt und lediglich anhand ihrer Distanz zur LRP-Position und der mittels OpenLR übertragenen Eigenschaften bewertet.

Dadurch treten in der Praxis häufig Fälle auf, bei denen nicht die vom Encoder als LRP gewählten drei- oder vierwertigen Knoten vom Decoder im Zielnetz gefunden und als homologe Punkte ausgewählt werden. Stattdessen werden oft zweiwertige Knoten ausgewählt, die möglicherweise näher an der gesuchten LRP-Position liegen, oder deren andere Eigenschaften (z.B. FRC, FOW, Richtungswinkel) besser passen. Die Tatsache, dass sie aber aufgrund ihrer topologischen Eigenschaften im Netz nicht als LRP-Entsprechung im Zielnetz in Frage kommen, findet bislang im Decoder der Musterimplementierung keine Berücksichtigung.

Eine Möglichkeit zur Abhilfe besteht beispielsweise darin, bei der Bewertung und Vorauswahl (Suchraumreduktion) der mittels Suchradius vom Decoder im Zielnetz gefundenen Knoten, zunächst die topologischen Eigenschaften dieser Knoten festzustellen. Bei der anschließenden Bewertung der Zuordnungskandidaten jedes LRPs im Zielnetz könnten dann diejenigen Knoten bevorzugt werden, die mehr als zwei abgehende Kanten besitzen. Alternativ könnten solche Kandidaten mit Wertigkeit  $\leq 2$  von der weiteren Betrachtung ausgeschlossen werden.

#### 3.4.3.5 Location Form

Die stützpunktbasierte Rekonstruktion der Ortsreferenzen im Zielnetz geht direkt auf die OpenLR Spezifikation zurück, da im Whitepaper (TOMTOM 2012) die Beschreibung der Ortsreferenzen durch *Location Reference Points (LRP)* und kürzeste Wege Verbindungen im Netz zwischen ihnen explizit vorgegeben ist. Dies bildet ein wesentliches Grundprinzip von OpenLR. In den vorliegenden Untersuchungen hat sich jedoch gezeigt, dass durch die Reduktion des Locationverlaufs auf die durch die kürzeste Wege Bedingung verbundenen Location Reference Points (LRPs), der Informationsgehalt und somit auch die Wiedererkennbarkeit der Ortsreferenzen im Zielnetz wesentlich reduziert wird. Praktisch gesehen werden entsprechend dem OpenLR Standard die Koordinaten der LRPs, der Richtungswinkel der im Locationverlauf ausgehenden Kante, deren Functional Road Class und Form of Way sowie die Distanz und die niedrigste im Quellnetz auftretende Functional Road Class entlang des Locationverlaufs zum nächsten LRP, übertragen. Die zu codierende Ortsreferenz wird so lediglich durch Stützpunkte repräsentiert, die mit weiteren Informationen über Eigenschaften ihrer ausgehenden (beim letzten: eingehenden) Kanten versehen werden.

Bei der Suche bzw. der Rekonstruktion dieser Stützpunkte im Zielnetz wird die Entfernung möglicher Kandidaten zum gesuchten Stützpunkt berücksichtigt, aber natürlich auch die Übereinstimmung mit den gesuchten Attributinformationen. In der Praxis können dadurch Fälle auftreten, bei denen weit vom gesuchten Stützpunkt entfernte Knoten die beste Bewertung erzielen, weil beispielsweise ihre semantischen Eigenschaften aufgrund der in Abschnitt 3.4.3.1 diskutierten möglichen Unsicherheiten beim Schema-Matching oder schlicht eines Datenbank- oder Erfassungsfehlers besser mit den in der OpenLR Ortsreferenz übermittelten Werten übereinstimmen, als diejenigen der eigentlich korrekten Zuordnung. Für diese Fälle erscheint es sinnvoll, der Location Reference weitere Informationen über

die Form des Locationverlaufs hinzuzufügen, um damit den Auswahlprozess im Zielnetz zu optimieren. So kann es schon hilfreich sein, zu wissen, ob die Ortsreferenz eine langgestreckte und ohne signifikante Richtungsänderungen geprägte Form besitzt, oder ob solche (z.B. rechtwinklige) Richtungsänderungen vorliegen.

### 3.4.3.6 Einfluss der Datenerfassung der Karte

#### Routingfähiges Netz

Voraussetzung für das Auffinden einer Route zwischen zwei Punkten in einem Straßennetz ist, dass zwischen diesen Punkte eine durchgehende Verbindung im Netz besteht. Kanten dürfen dabei nur in echten Knoten aufeinandertreffen und jeder Kreuzungspunkt muss durch einen solchen Knoten repräsentiert sein. Überschneidungen bzw. Unterquerungen sind nur bei Brücken oder Tunneln zulässig. Schon kleinste Unterbrechungen oder fehlende Verbindungen führen zu einer Nichtauffindbarkeit einer Routenverbindung in diesem Bereich. Oft sind Kanten auch augenscheinlich korrekt verbunden, sie laufen scheinbar visuell in einem Knoten zusammen. Liegen an dieser Stelle jedoch zwei Knoten mit exakt den selben Koordinaten übereinander und dienen jeweils als Start- oder Endpunkt für eine Kante, so kann topologisch keine Verbindung gefunden werden und die Routensuche scheitert an dieser Stelle. Da außer der Gesamtlänge der Ortsreferenz, der Teillängen der jeweiligen kürzesten Wege, sowie den Richtungswinkeln der vom jeweiligen LRP ausgehenden Kanten, keine Informationen über die Form der Ortsreferenz übertragen werden, kann eine solche falsche oder fehlende topologische Verbindung im Straßennetz der Zielkarte in manchen Fällen zu fehlerhaften oder nicht möglichen Decodierungen führen.

#### Fehlerhafte Erfassung des Fahrtrichtungsattributs

Neben der Konnektivität benachbarter Kanten ist für eine erfolgreiche Routensuche auch die Fahrtrichtungsangabe (engl. *Direction of Traffic Flow*) entscheidend. OpenLR kennt nur gerichtete Ortsreferenzen, d.h. sie haben einen Start- und einen Endknoten. Daraus ergibt sich eine implizite Richtungsangabe, die Kanten sind nur in diese Richtung befahrbar. Beim Decodieren der Ortsreferenz ist es entscheidend, einen Routenkandidaten zu finden, welcher in der durch die Ortsreferenz vorgegebenen Richtung auch befahrbar ist. Informationen zur richtungsweisen Befahrbarkeit sind in den verschiedenen verfügbaren digitalen Straßenkarten unterschiedlich modelliert und bezeichnet. Es wird jedoch bei den im Rahmen dieser Arbeit verwendeten Karten angegeben, in welcher Richtung ein Straßensegment (eine Kante) befahren werden kann, ausgehend von der Richtung der Digitalisierung. Diese ist implizit durch die Reihenfolge der formbeschreibenden Punkte gespeichert. Mögliche Angaben sind befahrbar in oder gegen die digitalisierte Richtung oder (der Normalfall) in beide Richtungen. Diese Informationen werden Routing-Algorithmen aus, um keine Routen entgegen der vorgegebenen Fahrtrichtung (z.B. bei Einbahnstraßen oder “Durchfahrt verboten“ Anweisungen) zu generieren. Für das Location Referencing stellt die Fahrtrichtungsangabe jedoch eine weitere Fehlerquelle dar. Vor allem in durch freie Projekte erstellten Karten (z.B. OpenStreetMap) wird meist vorrangig die Straßengeometrie erfasst und Wert auf deren Korrektheit gelegt. Angaben zur richtungsweisen Befahrbarkeit eines Abschnittes werden oft nur mit nachrangiger Priorität erfasst und kartiert. Im Rahmen der vorliegenden Arbeit sind diesbezüglich vermehrt Fälle offenkundig geworden, bei denen die Fahrtrichtungsangabe fehlte. Solche Straßen werden dann vom Encoder bzw. Decoder entsprechend dem OpenStreetMap Datenmodell (OPENSTREETMAP 2015c) als in beide Richtungen befahrbar angenommen. Sind die Straßen jedoch in Wirklichkeit nur in eine Richtung befahrbar oder ist die Durchfahrt sogar gesperrt, können Kanten fälschlicherweise als Kandidaten in Betracht kommen, die normalerweise aufgrund beispielsweise der Durchfahrtsbeschränkung nicht in Betracht kämen.

### 3.4.3.7 Zusammenfassung

Die Identifikation der LRPs im Zielnetz ist im OpenLR Whitepaper (TOMTOM 2012) nicht direkt spezifiziert. Sie basiert jedoch auf den im Standard definierten und übertragenen Location Informationen. Die Gewichtung und Bewertung der Übereinstimmung der einzelnen Parameter für jeden Kandidaten mit der übertragenen Ortsreferenz erfolgt proprietär in jeder Encoder / Decoder Implementierung.

Das Attribut-Mapping ist nicht unmittelbar Teil der OpenLR-Übertragung und ihrer Rekonstruktion im Zielnetz. Es ist jedoch ein wesentlicher Bestandteil der Informationskette vom Quellnetz über OpenLR ins Zielnetz. Werden die Attributinformationen beim jeweiligen Übergang zwischen den einzelnen Attribut-Schemata nicht exakt dem Zielschema zugeordnet, so gehen wertvolle Informationen nicht nur für eine einzelne Ortsreferenz Übertragung sondern bei allen Übertragungen von bzw. in das jeweilige Kartensystem, verloren. Sind die im OpenLR Standard vorgesehenen Attribute jedoch nicht im jeweiligen Kartensystem vorhanden und können diese auch nicht aus anderen Attributen bzw. aus Kombinationen von ihnen abgeleitet werden, führt dies zu einer wesentlichen Beeinträchtigung der Rekonstruierbarkeit der Ortsreferenz Übertragungen. Um etwaige Mehrdeutigkeiten bei der Attribut-Zuordnung in das LR-Übertragungsschema auf Senderseite (Encoder) der Empfängerseite (Decoder) verfügbar zu machen, könnte die Angabe und Übermittlung zusätzlicher qualitätsbeschreibender Metadaten sinnvoll sein. Zusammen mit dem Wissen über die Güte der Schema-Zuordnung im Empfängersystem könnte der Decoder solche Attributinformationen, die auf ihrem Weg vom Sender zum Empfänger mit Mehrdeutigkeiten behaftet wurden, in der Bewertung möglicher Ortsreferenz Kandidaten mit einem geringeren Gewicht versehen oder gar ganz aus der Bewertung nehmen, um eine falsche Zuordnung zu vermeiden.

Beginnen oder enden Ortsreferenzen an bivalenten (zweiwertigen) Knoten, so werden sie entsprechend dem OpenLR-Standard bis zum nächsten echten Netzknoten mit mehr als zwei verbundenen Kanten erweitert. Der eigentliche Locationbeginn wird dann durch den metrischen Abstand (Offset) zu diesem Knoten definiert. Dadurch ist für das empfangende System (Decoder) nicht ersichtlich, dass eine solche Ortsreferenz ursprünglich an einem zweiwertigen Knoten beginnen sollte. Der Decoder kann lediglich den übermittelten, geometrischen Abstand vom erweiterten und im Zielnetz identifizierten Startknoten abtragen. Aufgrund der zu beobachtenden geometrischen Kartenunterschiede, weichen die Distanzen zwischen den jeweiligen Knoten im Quell- und Zielnetz häufig voneinander ab. Dies führt dazu, dass der ursprünglich intendierte (bivalente) Startknoten oft nicht erreicht und stattdessen ein Punkt in der Umgebung dieses Knotens in eine nebenliegende Kante eingerechnet wird. Abhilfe könnte die zusätzliche Übertragung von Informationen zum genauen Beginn bzw. Ende der Ortsreferenz schaffen. Diese zusätzlichen Referenzierungs-Eigenschaften könnten beispielsweise beinhalten, dass die Ortsreferenz an einem zweiwertigem Knoten beginnt oder noch etwas detaillierter, dass dieser einen Attributwechsel der zulässigen Höchstgeschwindigkeit von 100 km/h auf 70 km/h darstellt.

Dieser Bezug einer Ortsreferenz auf Stützpunkte (LRP), die durch kürzeste Wege im Netz verbunden werden, ist einer der grundlegenden Ansätze der OpenLR-Referenzierung. Die Stützpunkte werden vom Encoder nur auf echten Netzknoten mit einer Wertigkeit  $> 2$  gewählt. Auf der Decoder-Seite gilt es dann natürlich auch, diese wertvolle topologische Eigenschaft der übermittelten LRPs bei der Suche nach ihrer Entsprechung im Zielnetz einzusetzen. Dadurch können falsche Zuordnungen vermieden werden, die ansonsten aufgrund besser passender geometrischer oder attributiver Ähnlichkeitsparameter ausgewählt worden wären.

Die Verwendung eines statischen Suchradius zum Auffinden potenzieller Knoten- und Kantenkandidaten im Zielnetz führt in der Praxis dazu, dass in dicht besiedelten Gebieten mit kurzen Abständen zwischen Kreuzungen tendenziell zu viele und an überörtlichen Verkehrswegen mit großen Abständen zwischen den einzelnen Kartenobjekten tendenziell zu wenige Kandidaten gefunden werden. Eine Dynamisierung des Suchradius in Abhängigkeit beispielsweise der Straßenklasse jedes Location Refe-

rence Points in der Quellkarte könnte eine einfache aber effektive Maßnahme sein, um diesem Problem zu begegnen.

Durch die Reduzierung der Ortsreferenz Beschreibung auf die Stützpunkte (LRPs) mit ihrer Position und weiterer Informationen können beim Decodieren im Zielnetz teilweise weit von der Suchposition entfernte Knoten ausgewählt und durch das Routing dahin die Location-Form stark verfälscht werden. Durch das Hinzufügen von formbeschreibenden Informationen zur Location Reference könnten solche Fälle im Zielsystem ausgeschlossen und die Rekonstruktion des beabsichtigten Locationverlaufs deutlich robuster gestaltet werden.

## 4 Automatische Zuordnung von Kartenelementen zwischen TomTom und OpenStreetMap

### 4.1 Motivation

Ein Ziel der vorliegenden Arbeit ist die Verbesserung von Location Referencing Verfahren hinsichtlich der in Kapitel 5 und 6 abgeleiteten Qualitätsmerkmale *Vollständigkeit*, *Genauigkeit* und *Korrektheit*. Aus den in Kapitel 3 vorgestellten existierenden Location Referencing Verfahren eignet sich vor allem OpenLR (siehe Kapitel 3.4.1) als Grundlage für weitere Entwicklungen, da es in der Praxis erfolgreich angewendet wurde und wird, wie beispielsweise innerhalb des ROSATTE-Projektes (SCHÜTZLE und FRANK 2010). Die Spezifikationen von OpenLR sind öffentlich zugänglich (TOMTOM 2012) und sogar als Musterimplementierung verfügbar (TOMTOM 2014). Das ebenfalls in ROSATTE zum Einsatz gekommene AGORA-C Location Referencing Verfahren (siehe Kapitel 3.3.5) eignet sich grundsätzlich ebenfalls. Allerdings wird dieses Verfahren aufgrund existierender Patentansprüche in der Praxis kaum angewendet, eine vollständige Implementierung ist ebenso nicht verfügbar. Alle weiteren in Kapitel 3 genannten Verfahren sind entweder durch die Entwicklung von OpenLR oder AGORA(-C) überholt oder sie werden wie im Falle von TPEG2-ULR (siehe Kapitel 3.3.7) noch nicht in der Praxis eingesetzt. Aus diesen Überlegungen heraus wurde OpenLR (Kapitel 3.4.1) als Grundlage für die im Rahmen der vorliegenden Arbeit vorgenommenen Analysen und Entwicklungen bestimmt.

Im Sinne der in Kapitel 2.4 vorgestellten Zuordnungsverfahren kann OpenLR grundsätzlich als *mehrstufiges* Verfahren betrachtet werden, auch wenn Zuordnungen mittels Location Referencing unter anderen Rahmenbedingungen als vollständige Kartenzuordnungen (beispielsweise in der Geoinformatik) stattfinden. So liegen beim Location Referencing die beiden zuzuordnenden Karten nicht in einem gemeinsamen Auswertesystem vor, wie dies bei Zuordnungen in der Geoinformatik der Fall ist. Dort werden beide Karten häufig zeitgleich in ein Geographisches Informations-System (GIS) geladen, um weiterreichende Auswertungen und Abfragen für die Zuordnungen vornehmen zu können. Des Weiteren wird beim Location Referencing lediglich ein kleiner Ausschnitt aus dem Netz (im Extremfall lediglich ein Punkt bzw. Knoten) übermittelt und im Zielnetz einem korrespondierenden Objekt zugeordnet. Bei den Kartenzuordnungen in der Geoinformatik werden hingegen häufig ganze Kartennetze einander zugeordnet. Dabei können vor allem auch die topologischen Beziehungen zwischen den einzelnen Kartenelementen für die Zuordnung herangezogen werden, beispielsweise beim *Seed Node*-Ansatz (vgl. KRAUT (2003), XIONG und SPERLING (2004), VOLZ (2006a)). Beim Location Referencing stehen lediglich die einmalig übermittelten Informationen der Ortsreferenz für die Zuordnung im Zielnetz zur Verfügung.

Da das in diesem Kapitel entwickelte Location Referencing Verfahren vor der eigentlichen Bewertung der Zuordnungs-Kandidaten auch eine vorgeschaltete Suchraumbeschränkung einsetzt, ist es grundsätzlich der Kategorie der *mehrstufigen* Zuordnungsverfahren zuzurechnen (vgl. Kapitel 2.4.3). Im Gegensatz zu anderen kombinierten Verfahren (wie etwa aus VOLZ (2006a)) stehen für den vorliegenden Ansatz jedoch weit weniger Objekteigenschaften für die Zuordnung zur Verfügung. Im typischen Anwendungsfall des Location Referencing kann nur eine Auswahl von Eigenschaften der zu übertragenden Objekte im Quellsystem ausgewählt werden, in die Ortsreferenz codiert und ins Zielsystem übertragen

werden. Des Weiteren ist eine Zuordnung in beide Richtungen (bi-direktional) beim Location Referencing ebenso wenig möglich, denn die Übermittlung der Ortsreferenz erfolgt nur in eine Richtung. Eine iterative Vorgehensweise ist in den typischen Anwendungsfällen des Location Referencing nicht vorstellbar.

Die von VOLZ (2006a) verwendete, auf dem Seed Point-Ansatz basierende Suche nach Zuordnungen, von denen ausgehend dann weitere Netzelemente untersucht und zugeordnet werden können, erscheint für eine Verwendung beim Location Referencing aus ähnlichen Gründen nicht praktikabel. Beim Location Referencing wird nur ein kleiner Netzausschnitt, nämlich gerade die Ortsreferenz (ggfs. erweitert bis zum nächsten echten Netzknoten), übertragen. Die Zuordnung muss dann genau für dieses Netzsegment erfolgen, unabhängig von der Güte der Kartenübereinstimmung an dieser Stelle, und kann nicht auf andere, vermeintlich besser passende Netzsegmente ausgedehnt werden.

Die von OpenLR übertragenen Informationen und insbesondere die Aufteilung des Location-Verlaufs in voneinander unabhängige kürzeste-Wege (engl. *shortest-path*) führen dazu, dass beim Decodieren lediglich die Location Reference Points im Zielnetz anhand der zusammen mit ihren Koordinaten übertragenen Eigenschaften unabhängig voneinander identifiziert werden. Diese decodierten LRPs (Knoten mit abgehender Kante) werden dann durch ein shortest-path Routing miteinander verbunden. Die so gebildete Route ergibt die decodierte Ortsreferenz (siehe Kapitel 3.4.2.3).

Durch die in Kapitel 3.4.3 durchgeführte Analyse wurde ersichtlich, dass einige Möglichkeiten der Erweiterung und Qualitätssteigerung von OpenLR existieren. Dazu sollen in diesem Kapitel einzelne Vorschläge gemacht werden, um die Vollständigkeit, Genauigkeit und Korrektheit des Decoding Ergebnisses zu verbessern. Die bislang von OpenLR verwendeten Zuordnungsmaße werden teilweise angepasst und weitere Zuordnungsmaße aus Kapitel 2.6 hinzugefügt. OpenLR verwendet mit der Euklidischen Distanz bereits ein einfaches geometrisches Zuordnungsmaß. Daher bietet es sich an, eine Verwendung weiterer, in Kapitel 2.6.1.1 vorgestellten Metriken für die Zuordnung zu untersuchen. Darüber hinaus sind die in Kapitel 2.6.1.4 vorgestellten, *formbasierten* Zuordnungsansätze für eine Verwendung im neu zu entwickelnden Zuordnungsverfahren interessant. Topologische Beziehungen werden in OpenLR nur implizit für die Zuordnung herangezogen. Die zur Beschreibung der Ortsreferenzen maßgebenden Location Reference Points müssen vom Encoder auf echte Netzknoten gesetzt werden (vgl. Kapitel 3.4.1.2). Diese topologische Eigenschaft der Knotenvalenz (siehe Kapitel 4.4.2) muss auf Seiten des Decoders aber auch konsequent genutzt werden. Als semantische Zuordnungsmaße kommen die numerischen Attributwerte für Functional Road Class und Form of Way zum Einsatz. Daher werden vorwiegend die Möglichkeiten der Verbesserung der geometrisch/topologischen Zuordnungsmaße untersucht. Des Weiteren verspricht auch die in der OpenLR Analyse angedachte Bewertung vollständiger Routen-Kandidaten neue Möglichkeiten der Ähnlichkeitsbetrachtung zwischen der Ausgangslocation im Quellnetz und den möglichen Routen-Kandidaten im Zielnetz.

## 4.2 Schema-Matching

Beim Aufsetzen eines Location Referencing Systems bzw. der Anbindung eines Anwendungssystems (mit zugehöriger Kartendatenbank) muss zunächst das Datenbank-Schema der verwendeten digitalen Straßenkarte dem LR-Datenschema zugeordnet werden. Im Rahmen der vorliegenden Arbeit wird ausschließlich das Location Referencing ohne Anbindung an eine Applikation und Übertragung zusätzlicher Nutzdaten (z.B. Verkehrsinformationen) getestet. Dazu werden als Kartendaten TomTom MultiNet und OpenStreetMap verwendet. Die Grundlage für die Anbindung dieser beiden Kartendatenbanken wurde in ALHESSI (2013) gelegt und im Zuge der Untersuchungen zur vorliegenden Arbeit weiterentwickelt.

Bei der Anbindung der Kartendatenbanken müssen nicht sämtliche verfügbaren Attribute in das Schema der LR-Testumgebung zugeordnet werden, sondern nur diejenigen, die zur Erstellung einer OpenLR

Location Reference benötigt werden (z.B. FRC, FOW). Des Weiteren werden noch andere Attributinformationen zur technischen Anbindung der Kartensysteme benötigt. Dies sind beispielsweise Angaben über die richtungswise Befahrbarkeit (*oneway*). Mit dieser Information lässt sich für jede Kante ermitteln, ob sie in beide Richtungen bzw. nur in oder entgegen der Digitalisier-Richtung befahren werden kann. Dementsprechend kann der Decoder Kanten, die der Location-Richtung entgegenlaufen, von vornherein als nicht geeignet verwerfen, vgl. Kapitel 3.4.3.6.

Die vollständige Schema-Zuordnung beider verwendeter Kartensysteme mit dem OpenLR-Schema wird im Folgenden detailliert dargestellt.

### 4.2.1 TomTom MultiNet

Das OpenLR-Datenschema kann als Untermenge des TomTom MultiNet-Datenschemas bezeichnet werden. Dies ist insofern nicht verwunderlich, als OpenLR zwar ein OpenSource Projekt ist, jedoch maßgeblich von TomTom entwickelt wurde. Die im OpenLR Whitepaper (TOMTOM 2012) definierten Attribute entsprechen denen von MultiNet, beinhalten aber nicht immer alle Attributwerte. Die exakte Ausgestaltung der Schema-Zuordnung von TomTom MultiNet zum für die vorliegende Arbeit verwendeten Datenschema ist in den Tabelle 4.1 dargestellt.

Tabelle 4.1: Schema-Zuordnung TomTom zu OpenLR

(a) Übersicht		(b) FRC	
LR-Attribut	TT MultiNet	OpenLR	TT MultiNet (TOMTOM 2011)
FRC	FRC	FRC-0	0: Motorway, Freeway, or Other Major Road
FOW	FOW		
Road Length	METERS	FRC-1	1: Major Road less important than Motorway
Road Identifier	ID		
Road name	NAME	FRC-2	2: Other Major Road
Oneway	ONEWAY	FRC-3	3: Secondary Road
Max Speed	KPH	FRC-4	4: Local Connecting Road
		FRC-5	5: Local Road (high importance)
		FRC-6	6: Local Road
		FRC-7	7: Local Road (minor importance)

Für das Attribut *Form of Way* enthält die MultiNet-Spezifikation (TOMTOM 2011) beispielsweise 22 unterschiedliche Werte, OpenLR spezifiziert jedoch nur sieben. Diese sieben OpenLR FOW Attributwerte finden in TomTom MultiNet jeweils eine exakte Entsprechung, siehe Tabelle 4.2. Die restlichen, von OpenLR nicht abgedeckten FOW-Attributwerte in TomTom MultiNet beinhalten spezielle Straßenformen, etwa in Parkhäusern oder Fußwege, die für eine Nutzung von OpenLR nicht relevant sind. Daher erfolgt die Anbindung von TomTom MultiNet an die OpenLR-basierte Implementierung faktisch mittels eines 1:1 Mappings.

### 4.2.2 OpenStreetMap

Bei der Anbindung von OpenStreetMap Karten können die Attribute Straßename, Kanten-ID und zulässige Höchstgeschwindigkeit direkt aus den OSM-Daten übernommen werden. Die Kantenlänge wird in OSM nicht explizit gespeichert, kann jedoch aus der Geometrie berechnet werden. Tabelle 4.3(a)

Tabelle 4.2: FOW-Zuordnung (TomTom zu OpenLR)

OpenLR FOW	TT MultiNet FOW (TOMTOM 2011)
MOTORWAY	1: Part of Motorway
SLIPROAD	10: Part of a Slip Road
MULTIPLE_CARRIAGEWAY	2: Part of Multi Carriageway (not Motorway)
SINGLE_CARRIAGEWAY	3: Part of a Single Carriageway
ROUNDABOUT	4: Part of a Roundabout
UNDEFINED	<sofern nicht im TT-MN Datenschema definiert>
OTHER	<alle anderen definierten Klassen>

zeigt die Zuordnung der OpenStreetMap Attribute zum verwendeten Location Referencing Datenschema.

Tabelle 4.3: Schema-Zuordnung OpenStreetMap zu OpenLR

(a) Übersicht		(b) FRC	
OpenLR Attribut	OSM tag	OpenLR FRC	OSM <i>highway</i>
FRC	highway	0 - Main Road	motorway, motorway_link
FOW	highway + junction	1 - First Class	trunk, trunk_link
Road Length	<berechnet>	2 - Second Class	primary, primary_link
Road Identifier	id	3 - Third Class	secondary, secondary_link
Road name	name	4 - Fourth Class	tertiary, tertiary_link
Oneway	oneway, highway, junction	5 - Fifth Class	residential, unclassified
Max speed	maxspeed	6 - Sixth Class	living_street
		7 - Other	-

Das Attribut *oneway* ist im OpenStreetMap Datenschema mit gleichem Namen enthalten, jedoch werden bestimmte Straßentypen automatisch als *oneway* angenommen. Die Angabe von *oneway* ist in diesen Fällen nicht obligatorisch. So entspricht bei *oneway=yes* die Fahrtrichtung der Digitalisierungsrichtung. Bei *oneway=-1* ist die Straße in entgegengesetzter Digitalisierungsrichtung befahrbar. Zur eindeutigen Klärung müssen dann neben *oneway* noch weitere Attribute herangezogen werden, um herauszufinden, ob eine Straße nur in eine oder beide Richtungen befahrbar ist. So wurden entsprechend den OpenStreetMap Richtlinien für die *tag*-Verwendung für alle Kreisverkehre mit *junction=roundabout* sowie für Autobahnen mit *highway=motorway* oder *highway=motorway\_link* angenommen, dass es sich hierbei um Einbahnstraßen in Digitalisierungsrichtung handelt. Dabei muss natürlich sichergestellt sein, dass diese Straßen tatsächlich auch in Fahrtrichtung digitalisiert sind.

Die beiden OpenLR-Attribute *Functional Road Class* und *Form of Way* sind abweichend oder nicht in OpenStreetMap definiert und müssen aus anderen Informationen abgeleitet werden. Dies wird in den folgenden beiden Abschnitten detailliert erläutert.

#### 4.2.2.1 Functional Road Class (FRC)

Die Functional Road Class ist einer der *building blocks* (vgl. Kapitel 3.4.1.2) für jede OpenLR Ortsreferenz bzw. für jeden OpenLR Location Reference Point (LRP). In OpenStreetMap wird die Straßenklasse

über das *highway* Tag modelliert, wobei der Begriff etwas weiter gefasst ist und ebenso Fußwege und auch nicht öffentlich zugängliche Zufahrten u.ä. beinhaltet.

Die Definition von OpenStreetMap beinhaltet sieben Straßenklassen und eine Klasse *unclassified* (OPENSTREETMAP 2015b), dies entspricht der Anzahl von acht Straßenklassen in OpenLR. Die Klassendefinitionen in OpenLR und OpenStreetMap unterscheiden sich jedoch. So sind im OpenLR Whitepaper (TOMTOM 2012) die Straßenklassen durchnummeriert, Klasse '0' wird als *Main Road* und Klasse '7' als *Other* bezeichnet. Es sind also lediglich die oberste und unterste Klasse eindeutig definiert, die restlichen Zuordnungsbeziehungen müssen durch empirische Versuche bestimmt werden. Insbesondere für die Straßen niedrigerer Bedeutung war dies nicht ganz einfach. Die Zuordnung der OSM-Klasse *residential* zu OpenLR-Klasse 5 und OSM-Klasse *living\_street* zu OpenLR-Klasse 6 hat sich dabei in der Praxis als am besten passend herausgestellt. Die entwickelte Klassenzuordnung für OpenStreetMap ist in Tabelle 4.3(b) dargestellt.

#### 4.2.2.2 Form of Way (FOW)

Noch schwieriger wird die Umsetzung einer Zuordnungsvorschrift beim OpenLR-Attribut *Form of Way*. Es beschreibt laut Definition 'den physischen Straßentyp' der jeweiligen Kante. Man könnte auch sagen: die bauliche Ausprägung. Leider ist diese Information in dieser Form nicht in OpenStreetMap spezifiziert und somit auch nicht verfügbar. OpenLR ist jedoch auf diese Information angewiesen, da das Attribut FOW Teil des OpenLR Bewertungsschemas ist, mit dem mögliche Zuordnungskandidaten in der Zielkarte bewertet und der am besten passende ausgewählt werden. Deshalb ist es erforderlich, eine Angabe zu Form of Way aus anderen, in OpenStreetMap verfügbaren Attributen, abzuleiten. Im Rahmen dieser Arbeit wurden verschiedene Möglichkeiten der Ableitung von FOW-Informationen aus OSM-Attributen getestet.

Es besteht das Problem, dass in den OpenStreetMap Daten nicht genügend Informationen vorhanden sind, um eine fehlerfreie Zuordnung abzuleiten. Deshalb ist es nicht möglich, eine mehrdeutigkeitsfreie FOW-Zuordnung von OSM nach OpenLR zu entwickeln. Dies betrifft vor allem die Unterscheidung zwischen den FOW Attributwerten *Single Carriageway* und *Multiple Carriageway*, also ob eine Straße zwischen den Richtungsfahrbahnen eine bauliche Trennung aufweist, oder nicht. So gibt es beispielsweise mit dem OSM-Tag *lanes* eine Angabe der für diese Kante vorhandenen Fahrstreifen. Es ist jedoch durchaus möglich, dass Straßen mit zwei Fahrstreifen je Fahrtrichtung keine bauliche Trennung zwischen den Richtungsfahrbahnen aufweisen. Diese wäre dann als FOW=Single Carriageway zu kennzeichnen. Genauso kann es aber auch sein, dass baulich getrennte Richtungsfahrbahnen jeweils nur einen Fahrstreifen je Fahrtrichtung besitzen, dennoch aber als FOW = Multiple Carriageway zu klassifizieren sind. Dadurch entstehen in bestimmten Fällen Mehrdeutigkeiten. Da aber in jedem Fall eine Entscheidung für eine FOW-Klasse getroffen werden muss, kommt es in konkreten Fällen zum Teil zu falschen Zuordnungen.

Um also überhaupt eine Entscheidung über das Vorhandensein einer baulichen Trennung treffen zu können, wird für FOW = *Multiple Carriageway* folgende Annahme getroffen: In OSM als *highway=trunk* klassifizierte Straßen besitzen eine übergeordnete Bedeutung und sind somit meistens mehrspurig. Diese Annahme muss natürlich nicht immer zutreffen. Genauso ist aber festzustellen, dass auch Straßen beispielsweise der Klassen *primary* oder *secondary* mehrspurig sein können, aber in der Regel doch einspurig sind (also FOW=Single Carriageway).

Eine Möglichkeit der weiteren Unterscheidung besteht in der zusätzlich Auswertung des OSM Attributs *oneway*. Bei einer höherklassigen und zusätzlich als Einbahnstraße gekennzeichneten Straße handelt es sich mit großer Wahrscheinlichkeit auch um eine mehrspurige Straße. Jedoch haben die empirischen Untersuchungen gezeigt, dass *oneway* bei OSM nicht in jeder Region flächendeckend oder nur mit nachrangiger Priorität erfasst wird. Eine Verwendung bei der Zuordnung des FOW erscheint

damit fraglich und wird daher nicht weiter verfolgt. Tabelle 4.4 zeigt eine Zuordnungsmatrix, die sich empirisch als am besten herausgestellt hat.

Tabelle 4.4: Zuordnung der Form of Way (OSM-OpenLR)

OpenLR FOW	OSM 'highway'	OSM 'junction'
MOTORWAY	motorway	-
SLIPROAD	*_link	-
MULTIPLE_CARRIAGEWAY	trunk	-
SINGLE_CARRIAGEWAY	primary, secondary, tertiary, residential, living street, unclassified	-
ROUNDABOUT	-	roundabout
UNDEFINED	road	-
OTHER	<andere>	-

### 4.2.3 Dynamisches Rating mit Schema-Matching Informationen

Durch die zuvor genannten Unsicherheiten bei der Schema-Zuordnung zwischen Sender und LR-Zwischenschicht sowie von der LR-Zwischenschicht in das Zielschema kann es auch zu Fehlern beim Vergleich der mittels Location Referencing übertragenen Attributinformationen mit denen der Kanten-Kandidaten im Zielsystem führen. Die dadurch entstehenden falschen Entscheidungen bei der Decodierung von Ortsreferenzen ließen sich in vielen Fällen umgehen, wenn dem Empfänger einer LR-Übertragung bekannt wäre, ob Mehrdeutigkeiten in der Schema-Zuordnung bestehen und wenn ja, bei welchen übermittelten Attributinformationen solche Unsicherheiten auftreten können.

Solche etwaigen Schwierigkeiten in der Schema-Zuordnung sind für den Decoder lediglich im Zielsystem verfügbar. Informationen aus dem Quellsystem werden in bisherigen LR-Verfahren nicht übertragen. Diese könnten durch die Angabe zusätzlicher Metadaten für die Location Referencing-Übertragungen dem Zielsystem bekannt gemacht werden. Mit Hilfe dieser Metadaten ließe sich die Güte der Schema-Zuordnungen auf Seiten des LR Senders dokumentieren und an den Empfängern übermitteln, wie auch beispielsweise in GAN und SHI (2002) dargestellt. Der Decoder kann dann anhand dieser Informationen beim Attribut-Rating der einzelnen Zuordnungskandidaten die entsprechend als sicher markierten Attribute höher und diejenigen mit einer bekannten Unsicherheit (sei es auf Sender- oder Empfängerseite) bei der Schema-Zuordnung niedriger gewichten.

Konkret betrifft das für OpenLR-Übertragungen die Attribute *Functional Road Class (FRC)* und *Form of Way (FOW)*. Die Güte der Zuordnung dieser Attribute in bzw. aus dem OpenLR-Datenschema kann in drei Klassen eingeteilt werden:

1. eindeutig,
2. mehrdeutig,
3. nicht möglich.

In der praktischen Umsetzung wird im Decoder jedem dieser Attribute ein Gewichtungsfaktor  $\omega_{src}$  für die Güte der Attributzuordnung im Ausgangssystem und  $\omega_{tar}$  im Zielsystem zugewiesen. Die Informationen über die Schema-Zuordnung im Quell- und Zielsystem werden dann über eine Multiplikation zu einem gemeinsamen Gewichtungsfaktor  $f_{dyn}$  zusammengeführt, siehe Gleichung 4.1. Dieser Faktor

kann dann wiederum an das jeweilige Rating für die LR-Zuordnungsentscheidung (siehe Abschnitt 4.4) angebracht werden.

$$f_{dyn} = \omega_{src} \cdot \omega_{tar} \quad (4.1)$$

mit

$$\omega_{src}, \omega_{tar} = \begin{cases} 1 & : \text{eindeutig} \\ 0,5 & : \text{mehrdeutig} \\ 0 & : \text{nicht möglich} \end{cases}$$

Mit der Wahl dieser numerischen Gewichtungsfaktoren ist gewährleistet, dass das entsprechende Attribut aus dem Rating niedriger gewichtet wird, sobald auf Sender- oder Empfängerseite eine Attributzuordnung nicht oder nur mehrdeutig möglich ist. Ist diese hingegen auf beiden Seiten eindeutig möglich, wird der Faktor  $f_{dyn} = 1$  und das Rating für dieses Attribut bleibt unverändert. Bei mehrdeutigen Zuordnungen wird das Rating mit  $\omega = 0,5$  jeweils halbiert bzw. bei beiderseitiger Unsicherheit sogar mit  $f_{dyn} = 0,25$  weiter reduziert.

### Implementierung und Ergebnisse

Bei den durchgeführten empirischen Untersuchungen wurde ein solches dynamisches Rating für die beiden Attribute *Functional Road Class (FRC)* und *Form of Way (FOW)* implementiert. Da wie in Kapitel 4.2.1 dargestellt, das Datenbank-Schema von TomTom MultiNet eindeutig den im OpenLR-Schema verwendeten Attributen zugeordnet werden kann, erhalten alle TomTom-Zuordnungen die Klassifikation *eindeutig*. Für die OpenStreetMap-Karten hingegen besteht, wie in Kapitel 4.2.2 gezeigt, nicht immer eine eindeutige Zuordnungsmöglichkeit der Attribute zum OpenLR-Datenschema. Daher wurde die Zuordnungsgüte für jeden einzelnen möglichen Attributwert für FRC und FOW wie in Tabelle 4.5 dargestellt festgesetzt, und daraus die dynamischen Gewichtungsfaktoren  $\omega$  wie oben beschrieben abgeleitet und in die Bewertung eingeführt.

Tabelle 4.5: Gewichtungsfaktoren für dynamisches Rating (OSM)

(a) FRC		(b) FOW	
FRC	Zuordnung	FOW	Zuordnung
0	eindeutig	0_UNDEFINED	mehrdeutig
1	mehrdeutig	1_MOTORWAY	eindeutig
2	mehrdeutig	2_MULTIP. CARR.WAY	mehrdeutig
3	mehrdeutig	3_SINGLE CARR.WAY	mehrdeutig
4	mehrdeutig	4_ROUNDABOUT	eindeutig
5	mehrdeutig	5_TRAFFICSQUARE	nicht möglich
6	mehrdeutig	6_SLIPROAD	eindeutig
7	nicht möglich	7_OTHER	mehrdeutig

Bei der praktischen Anwendung dieser dynamischen Gewichtung der Attribute mit den verwendeten Testdatensätzen konnte an einzelnen Ortsreferenzen eine Verbesserung der LR-Übertragung festgestellt werden. Bei dem in Abbildung 7.8 (Seite 137) gezeigten Beispiel erfolgt die LR-Übertragung von einem OpenStreetMap Quellnetz nach TomTom MultiNet als Zielnetz. Entsprechend der zuvor beschriebenen Implementierung der dynamischen Gewichtungsfaktoren, werden alle Schema-Zuordnungen zwischen OpenLR und TomTom MultiNet als *eindeutig* klassifiziert. Bei der Schema-Zuordnung zwischen OpenStreetMap und OpenLR ist dies jedoch nicht so pauschal möglich. Die Attributwerte der

Location Reference Points im Quellnetz (OSM) werden entsprechend Tabelle 4.5 sowohl für die Functional Road Class (2), als auch für Form of Way (3) als *mehrdeutig* klassifiziert. Dadurch ergibt sich jeweils der dynamische Gewichtungsfaktor  $f_{dyn,Bsp}$  entsprechend folgender Gleichung:

$$\begin{aligned} f_{dyn,Bsp} &= \omega_{src} \cdot \omega_{tar} \\ &= 0,5 \cdot 1 \\ &= 0,5. \end{aligned}$$

Tabelle 4.6: Beispiel 'dynamisches Rating'

Kandidat	SRC	Wert		Rating		dynRating	
		A	B	A	B	A	B
Distanz	-	11 m	60 m	89	40	89	40
FRC	2	1	2	75	100	38	50
FOW	3 (SCW)	2 (MCW)	3 (SCW)	50	100	25	50
$\Delta t$	-	-2,22 °	0,81 °	100	100	100	100
Summe (gew.)	-	-	-	942	<b>1020</b>	<b>755</b>	720

Die Bewertungsergebnisse des (klassischen) statischen OpenLR-Ratings sind für das vorgenannte Beispiel dem mittels der dynamischen Gewichtsanteile veränderten OpenLR-Rating in Tabelle 4.6 gegenübergestellt. Wie Abbildung 7.8 auf Seite 137 zeigt, wählt der Standard-OpenLR Decoder dabei einen vom ursprünglichen Startknoten im Quellnetz [roter Kreis] weit entfernt liegenden Kandidaten [blaues Quadrat], der in der Ergebnis Tabelle als Kandidat B bezeichnet ist. Die korrekte Zuordnungslösung (Kandidat A) entspricht jedoch dem in der Abbildung durch ein grünes Quadrat markierten Knoten. Betrachtet man die von OpenLR zur Zuordnung verwendeten Eigenschaften der beiden Kandidaten, so wird klar, dass Kandidat B deutlich weiter vom gesuchten Punkt entfernt liegt, seine Attributwerte für FRC und FOW aber besser mit den gesuchten Werten aus der Ortsreferenz in der Quellkarte übereinstimmen. Daher erhält Kandidat B für beide Attribute die volle Bewertung 100 und gewinnt schließlich mit einem Ergebnis von 1020 zu 942 die Bewertung und kann damit den Entfernungs-Nachteil mehr als wettmachen. Berücksichtigt man die Unsicherheit der Attribute in der Schema-Zuordnung mit den zuvor genannten Gewichtungsfaktoren mittels des dynamischen Bewertungsverfahrens, so reduziert sich der Einfluss der Attribute FRC und FOW auf die Gesamtbewertung, beide erhalten bei Kandidat B im dynamischen Rating nur noch den Wert 50. Das dynamische Bewertungsverfahren führt dann auch zum korrekten Ergebnis, der korrekte Kandidat A kann diese Bewertung mit einem Ergebnis von 755 zu 720 gewinnen.

In den Gesamtauswertungen der Datensätze sind Verbesserungen der Ergebnisse durch Hinzunahme der dynamischen Gewichtungsfaktoren sowohl bei OpenLR als auch beim Form-Matcher nicht zu beobachten. Die Ergebnisse liegen mit und ohne dynamischen Anteilen auf gleichem Niveau (fast ohne Veränderungen!). Eine mögliche Ursache kann in der im Rahmen dieser Arbeit in mehreren Iterationsschritten immer weiter verbesserten Schema-Zuordnung für die verwendeten OpenStreetMap-Karten liegen. Des Weiteren sind Unsicherheiten durch eine mehrdeutige Schema-Zuordnung nicht in allen Fällen die Ursache für eine falsche oder nicht mögliche LR-Zuordnung. Daher hat eine Verbesserung durch die dynamische Gewichtung der Attributbewertungen, abgesehen von Einzelfällen, nur einen begrenzten Einfluss auf das Gesamtergebnis.

## 4.3 Ermittlung der Decoding-Kandidaten

Ein erster Schritt beim Decodieren einer Ortsreferenz, ist die Abfrage aller Kanten, die eine räumliche Nähe zur Ausgangs-Location aufweisen. Das in Abschnitt 3.4.2.2 beschriebene Verfahren zur Identifikation von Knoten und Kanten innerhalb eines Suchradius um die Position des jeweiligen LRP wurde grundsätzlich auch im hier vorgestellten Verfahren beibehalten. In den nächsten Abschnitten werden die vorgenommenen Erweiterungen und Verbesserungen im Vergleich zu OpenLR beschrieben.

### 4.3.1 Dynamischer Suchradius

Wie schon in Kapitel 3.4.3.3 diskutiert, erscheint eine Dynamisierung des Suchradius, der für die Suche nach den nächstgelegenen Knoten und Kanten verwendet wird, sehr sinnvoll. Daher wurde für den Form-Matcher der Suchradius im Decoder für jeden Location Reference Point dynamisch bestimmt. Der Suchradius lässt sich etwa aus der jeweiligen zulässigen Höchstgeschwindigkeit oder der Functional Road Class ableiten.

#### Suchradius in Abhängigkeit der Höchstgeschwindigkeit

Der dynamische Suchradius  $r$  zur Knoten- und Kanten-Detektion im Zielnetz könnte mit Hilfe eines Faktors  $f$  direkt aus der jeweils gültigen Höchstgeschwindigkeit  $v_{max}$  abgeleitet werden:

$$r = f \cdot |v_{max}| [m]. \quad (4.2)$$

Wählt man beispielsweise einen Faktor  $f = 2$ , so würde der Suchradius innerorts in Wohnstraßen mit  $v_{max} = 30 \text{ km/h} \rightarrow r = 60 \text{ m}$  kleiner ausfallen als bisher, auf Hauptstraßen innerorts bliebe er mit  $v_{max} = 50 \text{ km/h} \rightarrow r = 100 \text{ m}$  gleich wie bisher und außerorts würde sich der Suchradius deutlich vergrößern, wie die Zusammenstellung in Tabelle 4.7 zeigt.

Tabelle 4.7: Vorschlag dynamische Suchradien

$v_{max} [km/h]$	$\leq 30$	$\leq 50$	$\leq 70$	$\leq 100$	$> 100$
$r [m]$	60	100	140	200	250

Dadurch würden die in der Schwachstellenanalyse in Kapitel 3.4.3.3 diskutierten Randbedingungen für die Dynamisierung des Suchradius erreicht. Die erzielten Größenordnungen der Suchradien haben sich empirisch als zielführend erwiesen. Die maßgebliche Höchstgeschwindigkeit oder auch direkt der Suchradius müsste dann zusammen mit den anderen in OpenLR spezifizierten Informationen für jeden Location Reference Point aus der Quellkarte abgeleitet und in der Ortsreferenz übertragen werden, was faktisch zu einer Erweiterung des Datenformats der OpenLR Übertragungen führen würde. Alternativ könnte der Suchradius auch ohne Erweiterung des Übertragungsformats aus den Geschwindigkeits-Informationen der Zielkarte abgeleitet werden. Dieser Ansatz ist jedoch wegen der Abhängigkeit des Suchradius vom jeweiligen Kandidaten und damit eines nicht einheitlichen Suchradius nicht praktikabel. Da diese Attributinformation in der Regel für Kanten gespeichert wird, ist die Höchstgeschwindigkeit beim Encoding der dem LRP zugehörigen Kante zu entnehmen. Dies würde die allgemeine Verfügbarkeit der jeweils zulässigen Höchstgeschwindigkeit für die auf Encoder-Seite verwendeten Karten voraussetzen. Bei kommerziellen Karten beispielsweise von TomTom oder HERE kann davon

ausgegangen werden. Diese Karten und deren Attribute werden entsprechend der Kundenanforderungen aufwendig gepflegt und es sind bereits Systeme am Markt verfügbar, die die zulässige Höchstgeschwindigkeit aus Kartendaten verwenden. Allerdings sieht die Situation bei Community-basierten Karten (wie z.B. OpenStreetMap) anders aus. Bei diesen ist zwar die Erfassung der zulässigen Höchstgeschwindigkeit durch die explizite Definition eines entsprechenden Attributs vorgesehen. Die tatsächlich lückenlose Erfassung kann aber niemand garantieren. In solchen Fällen wäre alternativ die Verwendung eines Standard-Suchradius bei fehlender Verfügbarkeit der Höchstgeschwindigkeit für einzelne Kanten denkbar. Dies dann jedoch mit dem jetzigen, statischen Zustand vergleichbar.

### Suchradius in Abhängigkeit der Straßenklasse

Eine andere Möglichkeit wäre die Definition des Suchradius in Abhängigkeit von der Straßenklasse. Daraus würden sich möglicherweise Unsicherheiten ergeben, denn die Straßenklassen lassen im Gegensatz zur zulässigen Höchstgeschwindigkeit eine weniger granulare Unterscheidung zu. Des Weiteren variieren der Ausbauzustand der Straßen sowie die tatsächliche physische Ausprägung innerhalb der Klassen. Diese Faktoren haben sicherlich einen Einfluss auf die Erfassungsgenauigkeit und somit auch auf die möglichen geometrischen Unterschiede zwischen den Karten, die ja eine Dynamisierung des Suchradius notwendig erscheinen ließen.

Der Vorteil bei der Verwendung der Straßenklasse ist jedoch die Verfügbarkeit ihrer Information. Bei OpenLR wird für jeden Location Reference Point obligatorisch auch seine Functional Road Class übertragen. Die übertragene FRC kann je nach verwendetem Kartentyp und Schema-Matching möglicherweise Mehrdeutigkeiten aufweisen (vgl. Abschnitt 3.4.3.1 bzw. Kapitel 4.2.2.1). Dennoch erscheinen die verschiedenen Schemata ähnlich genug, um daraus einen geeigneten Suchradius ableiten zu können. Um wiederum einen für jeden LRP einer Ortsreferenz einheitlichen Suchradius zu erhalten, wird die Verwendung der mittels OpenLR für jeden LRP übertragenen Functional Road Class vorgeschlagen. Eine Erweiterung des OpenLR Datenformats wie zuvor für die zulässige Höchstgeschwindigkeit angedacht, wäre hier nicht notwendig.

Der dynamische Suchradius  $r$  wird im Rahmen der vorliegenden Arbeit entsprechend Tabelle 4.8 relativ zur Functional Road Class gewählt. Damit werden für die Suchradien ähnliche Größenordnungen wie bei der zuerst vorgestellten Ableitung von der zulässigen Höchstgeschwindigkeit erzielt. Dieser Ansatz für eine Dynamisierung des Suchradius wurde für die empirischen Untersuchungen im Form-Matcher implementiert und erzielte zufriedenstellende Ergebnisse (siehe Kapitel 7).

Tabelle 4.8: Vorschlag dynamische Suchradien auf Basis von FRC

OpenLR FRC	OpenStreetMap ( <i>highway=*</i> )	$r[m]$
0 - Main Road	motorway, motorway_link	200
1 - First Class	trunk, trunk_link	150
2 - Second Class	primary, primary_link	100
3 - Third Class	secondary, secondary_link	75
4 - Fourth Class	tertiary, tertiary_link	75
5 - Fifth Class	residential, unclassified	50
6 - Sixth Class	living_street	50
7 - Other	-	100

### 4.3.2 Bewertung und Vorauswahl der gefundenen Kanten

Wie bei der OpenLR Beispiel-Implementierung, werden von den über Knoten gefundenen Kanten diejenigen verworfen, die nicht von diesem Knoten ausgehen, beim letzten LRP einer Ortsreferenz werden nur die eingehenden Kanten behalten. Ebenso werden Kanten verworfen, deren Attribut 'Functional Road Class' (FRC) die niedrigste FRC um mehr als eine zu definierende Anzahl von Klassen übersteigt.

Ein drittes Kriterium zur Vorauswahl der im Zielgebiet gefundenen Kanten ist ein Rating. Die OpenLR-Musterimplementierung (TOMTOM 2014) wendet hier das OpenLR Standard-Rating (Punktabstand, Richtungswinkel, FRC, FOW) an. Bleibt die Bewertung der zu untersuchenden Kante unter einem in den Decoder-Einstellungen festzulegenden Wert, dann wird der Kandidat verworfen.

In den empirischen Untersuchungen traten jedoch häufiger Fälle auf, in denen Kanten durch dieses Rating in der Vorauswahl aussortiert wurden, obwohl sie eigentlich die korrekte Lösung darstellten. Die Gründe lagen oft in einer zu großen Differenz der Straßenklassen (FRC). Dies wiederum war die Folge einer nicht ganz mehrdeutigkeitsfreien Schema-Zuordnung für dieses Attribut (vergleiche Kapitel 4.2.2.1). Daher wurde das Rating in der Vorauswahl durch ein weniger restriktives ersetzt.

In diesem neuen Rating  $R_{pre}$  für die Vorauswahl geeigneter Kanten wird lediglich der Abstand des Startknotens zu dem gesuchten LRP ( $R_{Node}$ ) sowie der Richtungswinkelunterschied ( $R_t$ ) bewertet (vgl. Abschnitt 4.4):

$$R_{pre} = R_{Node} + R_t. \quad (4.3)$$

Alle Kanten mit einem Bewertungsergebnis von weniger als 60 % werden verworfen. Dieser Schwellwert wurde empirisch bestimmt und hat sich im Rahmen der vorliegenden Arbeit als zweckmäßig erwiesen. Ziel des neuen Vorgehens ist, die Kandidaten-Menge an diesem Punkt der Prozessierung noch möglichst groß zu halten, damit beim finalen Gesamt-Rating der richtige Routenverlauf mit sehr hoher Wahrscheinlichkeit noch in der Menge der Kandidaten enthalten ist.

### 4.3.3 Decoding ohne direkte Knoten-Zuordnung

Die Referenzierung einer (linienhaften) Ortsreferenz in OpenLR basiert auf den nächstgelegenen validen Netzknoten, diese werden als Location Reference Point ausgewählt (siehe Abschnitt 3.4.1.2). Diese Vorgabe der OpenLR Spezifikation soll sicherstellen, dass die LRPs genügend geometrisch/topologische Eigenschaften (Alleinstellungsmerkmale) besitzen, um sie mit hoher Wahrscheinlichkeit eindeutig im Zielnetz auffinden zu können.

In manchen Fällen ist es jedoch möglich, dass im Suchgebiet keine entsprechenden validen Netzknoten vorhanden sind. Gründe dafür können Kartenunterschiede, mangelnde Vollständigkeit oder Kartenfehler sein. Um in einem solchen Fall trotzdem eine Decoding-Lösung anbieten zu können, bildet die OpenLR Musterimplementierung (TOMTOM 2014) das Lot vom LRP auf die Kante mit dem besten Rating und verwendet diesen eingerechneten Zwischenpunkt als decodierten Location Reference Point.

Die Entscheidung, ob für eine Kante ein Kandidat mit Lotpunkt in die Menge der Zuordnungskandidaten eingefügt werden soll, wird anhand des Vorauswahl-Ratings getroffen. Dabei ist zu beachten, dass die Kante mit eingefügtem Lotpunkt ein deutlich besseres Distanz-Rating besitzt, als die gleiche Kante bei Betrachtung ihres Anfangs- bzw. Endknotens. Daher wird das Kanten-Rating mit einem Faktor *LinesDirectlyFactor* reduziert. Dieser Faktor lag bei der OpenLR Musterimplementierung bei  $LinesDirectlyFactor_{OpenLR} = 95\%$ . Da das Bewertungsverfahren für die Vorauswahl der Kandidaten im Vergleich zu OpenLR geändert wurde, muss auch der Faktor für direkt gefundenen Kanten

für den Form-Matcher angepasst werden. Aufgrund empirischer Untersuchungen wird der Faktor mit  $LinesDirectlyFactor_{Form-Matcher} = 90\%$  festgelegt.

Abbildung 4.1(a) zeigt das Beispiel einer Ortsreferenz, deren im Zielnetz decodierter Endpunkt [rotes Dreieck] einen geometrischen Abstand von 12 m zum Endpunkt der Ausgangslocation in der Quellkarte [fett blau] besitzt. Der Sollpunkt im Zielnetz [grünes Quadrat] liegt 21 m vom Ausgangspunkt entfernt. Der zugehörige Location Reference Point im Quellnetz liegt ca. 3 km süd-östlich [roter Kreis], siehe Abbildung 4.1(b). Der entsprechende Kreuzungspunkt im Zielnetz ist ca. 105 m entfernt in süd-östlicher Richtung [rote Pfeile]. Aufgrund dieser großen Entfernung wurde der LRP vom Form-Matcher auf die Kante im Zielnetz [schwarz] gelotet [blauer Kreis] und der End-Offset von diesem Lotpunkt zum Endpunkt der Ortsreferenz abgetragen. Dadurch konnte eine Übertragung der relativ großen Kartenunterschiede im Gebiet des Location Reference Points auf das Ortsreferenz-Ende verhindert werden.

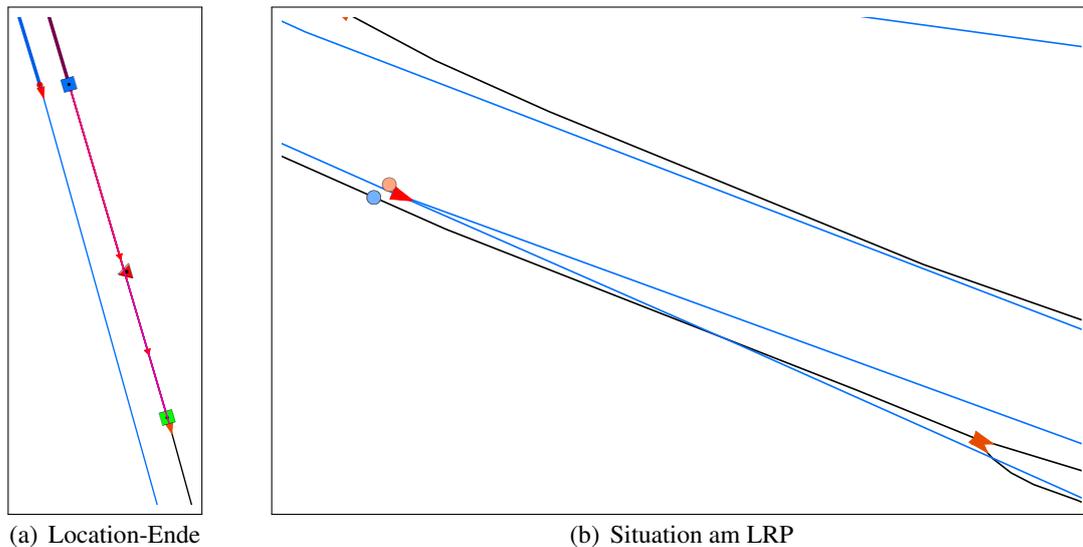


Abbildung 4.1: Beispiel Projektion auf Kante

#### 4.3.4 Zusammensetzen aller möglichen Routenverläufe

Um später die Form der Ortsreferenz mit denen der Kandidaten für die Zuordnung vergleichen zu können, müssen diese Kandidaten schon vor der Bewertung als die Ortsreferenz vollständig abdeckende Routenverläufe vorliegen. Dazu werden die möglichen Verbindungen zwischen allen Kandidaten jeweils zweier aufeinander folgender LRPs im Zielnetz gesucht. Dadurch kann, im Gegensatz zu OpenLR, jede Ortsreferenz vom ersten bis zum letzten LRP als Ganzes betrachtet und alle im Zielnetz möglichen Zuordnungskandidaten für den gesamten Locationverlauf gebildet und bewertet werden. Bei OpenLR hingegen erfolgt die Bewertung für jeden Location-Abschnitt zwischen zwei aufeinanderfolgenden LRPs separat. Die Bestimmung der kürzesten Wege zwischen den einzelnen LRP-Kandidaten erfolgt mittels der in der OpenLR Musterimplementierung (TOMTOM 2014) enthaltenen *shortest-path* Routensuche. Diese verwendet eine Implementierung des *A\*-Algorithmus* zur Bestimmung des kürzesten Weges zwischen zwei Knoten im Netz, siehe auch BASELAU (2010). Aufgrund der in Kapitel 2.2.3 beschriebenen Problematik der Kartenunterschiede können die im Zielnetz gefundenen Routenkandidaten sämtliche der in Kapitel 2.3.2 genannten Kardinalitäten aufweisen. Im einfachsten Falle ist eine 1:1 Zuordnung möglich, bei längeren Ortsreferenzen ist jedoch die m:n Zuordnung eher die Regel, bei der m Kanten aus dem Quellnetz n Kanten des Zielnetzes zugeordnet werden.

Besteht eine Ortsreferenz aus mehr als zwei LRPs, dann müssen aus den im Zielnetz gefundenen Routen-Segmenten zwischen den jeweils aufeinander folgenden LRPs alle möglichen Kombinationen erkannt und die Segmente dementsprechend zu größeren Ketten zusammengesetzt werden. Dadurch entstehen alle im Zielnetz möglichen Routen, die den Locationverlauf vollständig abbilden. Dieser Schritt ist in der praktischen Umsetzung nicht trivial, da die Routen Segmente ja nicht mehr wie die Ausgangs-Location im Quellnetz nur an einem LRP (Knoten mit angehängter Kante) enden bzw. beginnen. Es werden unter Umständen für jeden LRP mehrere Kandidaten im Zielnetz gefunden. Dementsprechend gibt es auch sehr viele Kombinationsmöglichkeiten aus diesen einzelnen Segmenten. Es ist also für jedes zwischen zwei LRP-Kandidaten liegende kürzeste-Wege-Segment zu prüfen, ob für den nächsten Abschnitt der Ortsreferenz Routen-Kandidaten existieren, die am Endknoten dieses kürzesten-Wege-Segments beginnen. Abbildung 4.2 illustriert ein fiktives Beispiel mit drei Location Reference Points. Zu jedem LRP sind die jeweiligen Kanten-Kandidaten mit ihrem Identifikator (z.B. 4711) dargestellt. Die rot-gestrichelten Linien stellen sämtliche möglichen Kombinationen dieser LRP-Kandidaten dar, die dicken blauen Linien symbolisieren die im Zielnetz gefundene kürzeste Wege zwischen den LRP Kandidaten (Kante mit Startknoten). Die möglichen Routenkombinationen für die beiden Routensegmente der Ortsreferenz sowie die Kombinationsmöglichkeiten für die gesamte Ortsreferenz sind in Tabelle 4.9 durch die Kanten-Bezeichner im Zielnetz dargestellt.

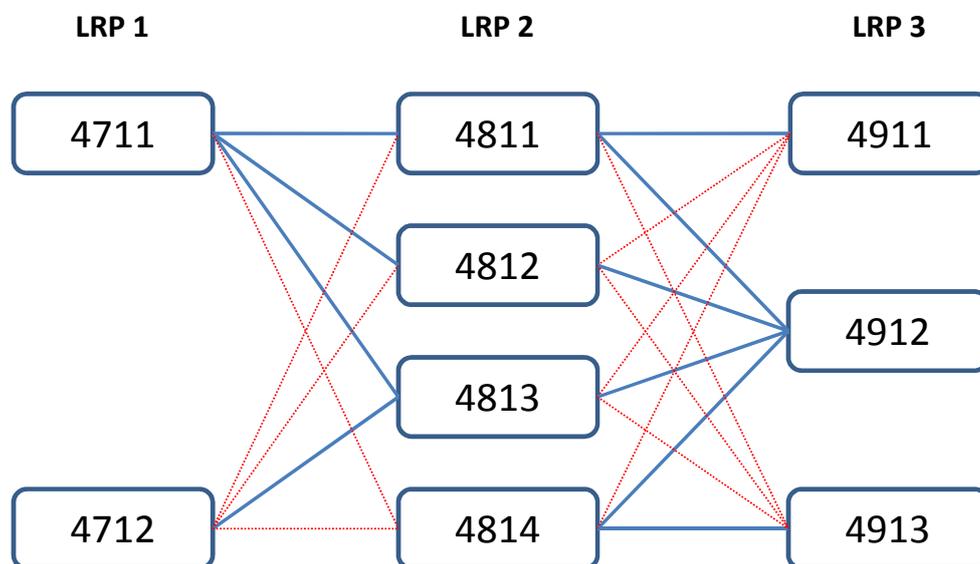


Abbildung 4.2: Zusammensetzen der im Zielnetz gefundenen Kanten Kandidaten

Tabelle 4.9: Mögliche Routenkombinationen

LRP 1-2	LRP 2-3	gesamt
4711 - 4811	4811 - 4911	4711 - 4811 - 4911
4711 - 4812	4811 - 4912	4711 - 4811 - 4912
4711 - 4813	4812 - 4912	4711 - 4812 - 4912
4712 - 4813	4813 - 4912	4711 - 4813 - 4912
	4814 - 4912	4712 - 4813 - 4912
	4814 - 4913	

Durch das Bilden aller möglichen Routenkombinationen zwischen den Kandidaten aller Location Reference Points entstehen sehr oft auch Routen Duplikate, also mehrere Routen, die jeweils den selben Verlauf haben. Solche Duplikate können bei Ortsreferenzen mit drei oder mehr LRPs entstehen. Beinhalten

Routenkandidaten an LRPs im Verlauf der Ortsreferenz (nicht am Start und Ende) die selben Kantenfolgen am Ende der eingehenden und am Anfang der ausgehenden Routenkandidaten, so entstehen beim Zusammensetzen Routenverläufe mit identischen Kantenfolgen, also Duplikate. Diese gilt es in einem nächsten Verarbeitungsschritt zu eliminieren, sodass in der verbleibenden Menge von Routen Kandidaten jede Route nur noch einmal vorkommt. Zum Vergleich der einzelnen Routen werden die jeweilige Folge von Kanten-IDs verglichen. Ein weiterer Vergleichsoperator prüft, ob die Start-/Endkanten der beiden Routenverläufe Unterschiede bzgl. ihrem Betrachtungspunkt (Knoten oder Projektionspunkt, vgl. Kapitel 3.4.2.2) aufweisen. Dies kann über eine Verkettung von logischen Ausdrücken erreicht werden, wie in Abbildung 4.3 graphisch dargestellt. Zunächst werden dabei die Start- und Endkanten der beiden Routen Kandidaten jeweils mit einem logischen Exklusiv-Oder (XOR) verglichen. Dieser Ausdruck wird nur dann 'wahr', wenn ein Unterschied zwischen den beiden Vergleichswerten besteht, d.h. einer der beiden Location-Punkte ein eingerechneter Zwischenpunkt (Projektionspunkt) ist und der andere direkt auf dem Endknoten der Kandidatenkante sitzt. Haben die beiden zu vergleichenden Punkte die gleichen Eigenschaften (also Projektionspunkt oder Knoten), dann wird der XOR-Ausdruck 'falsch'. Die Ergebnis-Ausdrücke werden dann abschließend mit einem einfachen 'Oder' verknüpft, sodass das Gesamtergebnis immer dann 'wahr' wird, wenn eine der beiden Gruppen (Start- oder Endkante) einen Unterschied aufweist. In diesem Fall werden beide Routenverläufe in der Menge der Kandidaten belassen. Sind sie jedoch nicht unterschiedlich, wird eines der beiden Duplikate verworfen.

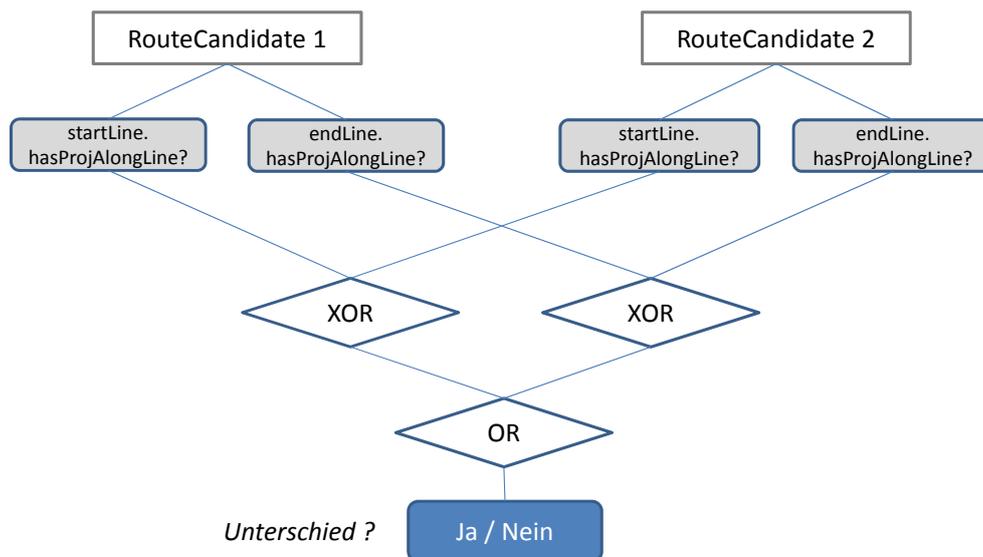


Abbildung 4.3: Identifikation von Alternativrouten unter Nutzung von Projektionspunkten

In einem weiteren Schritt wird überprüft, ob die Routen-Kandidaten dieselben Kanten mit entgegengesetzter Fahrtrichtung direkt nacheinander enthalten. Durch die Verkettung der einzelnen Teilrouten im vorigen Schritt kann es vorkommen, dass die resultierenden Routen eine solche 'Wende' enthalten. Es ist nahezu ausgeschlossen, dass diese in der ursprünglichen Ortsreferenz so beabsichtigt waren. Da solche Artefakte zu Problemen im weiteren Verlauf der Bewertungen führen, werden sie an diesem Punkt der Prozessierung eliminiert. Durch Prüfung der Identifikatoren der jeweils aufeinanderfolgenden Kanten können solche Situationen erkannt und der entsprechende Routen-Kandidat von der weiteren Verarbeitung ausgeschlossen werden. Die so gebildete Menge von Kandidaten wird im nächsten dann im nächsten Schritt bezüglich ihrer Übereinstimmung mit der Ausgangslocation bewertet.

## 4.4 Bewertung der LRP-Kandidaten

Zunächst werden die zuvor im Zielnetz gebildeten Ortsreferenz Kandidaten an ihren Stützpunkten (LRPs) auf die Übereinstimmung mit den mittels OpenLR übertragenen Eigenschaften der Location Reference Points im Quellnetz untersucht und bewertet. Dafür kann das in Kapitel 3.4.2.2 beschriebene Bewertungsverfahren der OpenLR Musterimplementierung (TOMTOM 2014) als Grundlage herangezogen werden. In ihm werden die Distanz zum gesuchten LRP, sowie Unterschiede im Richtungswinkel, Functional Road Class (FRC) und Form of Way (FOW) bewertet. Im Rahmen der vorliegenden Arbeit wurde das Bewertungsverfahren weiterentwickelt und an die eigenen Anforderungen angepasst. Diese Anpassungen werden im Folgenden beschrieben. Das Punktdistanz-Rating, das FRC-Rating sowie das FOW-Rating wurden unverändert belassen (siehe Kapitel 3.4.2.2).

### 4.4.1 Richtungswinkel

Mit der OpenLR Ortsreferenz wird für jeden Location Reference Point (LRP) auch der Richtungswinkel (vgl. Kapitel 2.6.1.2) der zugehörigen ausgehenden Kante (bzw. beim letzten LRP einer Ortsreferenz der eingehenden Kante) übertragen. Zur Bewertung von Kandidaten-Kanten im Zielnetz kann nun die Richtungswinkeldifferenz  $\Delta t$  aus dem aus der Quellkarte übertragenen Richtungswinkel  $t_{src}$  und dem des jeweiligen Kandidaten  $t_{cand}$  nach Gleichung 4.4 berechnet werden:

$$\Delta t = t_{src} - t_{cand}. \quad (4.4)$$

Die in der OpenLR Musterimplementierung (TOMTOM 2014) verwendete Klassifizierung der Richtungswinkeldifferenzen (vgl. Tabelle 3.3(a)) wurde nicht beibehalten, da zwischen den verwendeten Klassen große Bewertungsunterschiede auftreten. Fällt eine berechnete Richtungswinkeldifferenz äußerst knapp in eine dieser Klassen, so kann sich das Rating dadurch drastisch verändern (z.B.  $\Delta t = 6,001^\circ$  führt zu einer Abwertung um die Hälfte). Stattdessen wird zur Beurteilung die Differenz der Richtungswinkel zur maximal akzeptierten Richtungswinkelabweichung  $\Delta t_{max}$  gebildet. Damit kann entsprechend Gleichung 4.5 eine kontinuierliche Bewertung der Richtungswinkeldifferenz  $R_t$  ohne abrupte Sprünge der Bewertung an den Klassengrenzen erzielt werden. Ist die Richtungswinkeldifferenz  $\Delta t$  größer als der maximal akzeptierte Wert, so liefert das Rating den Wert 0:

$$R_t = \max(0, \Delta t_{max} - |\Delta t|). \quad (4.5)$$

### 4.4.2 Valenz der Knoten

Wie in Kapitel 3.4.3.4 diskutiert, soll die Vorgabe aus dem OpenLR Whitepaper, wonach im Encoder als LRPs nur echte, mindestens dreiwertige Knoten ausgewählt werden sollen, auch bei der Auswahl der LRP Decoding Kandidaten ausgenutzt werden. Da die tatsächliche Valenz (Anzahl der abgehenden Kanten eines Knotens, siehe Kapitel 2.6.2.2) eines LRP im Quellnetz nicht mittels OpenLR übertragen wird, bleibt diese im decodierenden Zielsystem unbekannt. Dennoch ist eine Prüfung auf Valenz  $> 2$  und entsprechende Aufwertung aller echter Knoten Kandidaten im Zielnetz möglich. Dazu wird eine neue Bewertung der Knotenvalenz  $R_{val}$  für jeden LRP Kandidaten eingeführt. Eine zu implementierende Valenzfunktion  $val()$  ermittelt die Anzahl der abgehenden Kanten des zu betrachtenden Knotens  $N_i$ . Ist dessen Wertigkeit  $val(N_i) > 2$ , so wird der Wert der Bewertungsfunktion maximal, sodass diese Knoten bei der Gesamtbewertung bevorzugt werden. Für ein- oder zweiwertige Knoten gibt es keinen Bewertungsaufschlag:

$$R_{val} = \begin{cases} 100 & : val(N_i) > 2 \\ 0 & : val(N_i) \leq 2 \end{cases}. \quad (4.6)$$

Abbildung 4.4 zeigt ein Beispiel auf einer Autobahn. Die Ausgangs-Location [fett-blau] geht über eine Anschlussstelle hinaus. Der OpenLR Encoder hat hierzu einen zusätzlichen LRP (in diesem Fall LRP 2) eingefügt (roter Kreis). Zufälligerweise ist jedoch im Zielsystem [schwarz] ein unechter Knoten deutlich näher am LRP als der entsprechende Knoten der Autobahn-Abfahrt, der vom Decoder der OpenLR Musterimplementierung bevorzugt wird [blauer Kreis]. Von diesem Knoten aus wird dann ein Routing zum letzten LRP durchgeführt. Dieses führt die Route dann fälschlicherweise über die südlich abzweigende Abfahrt und an der nächsten Auffahrt wieder auf die Autobahn zum Ziel. Der Form-Matcher decodiert den LRP hingegen korrekt [gelber Kreis], die Ortsreferenz [rot] nimmt dadurch den korrekten Verlauf entlang der Autobahn.

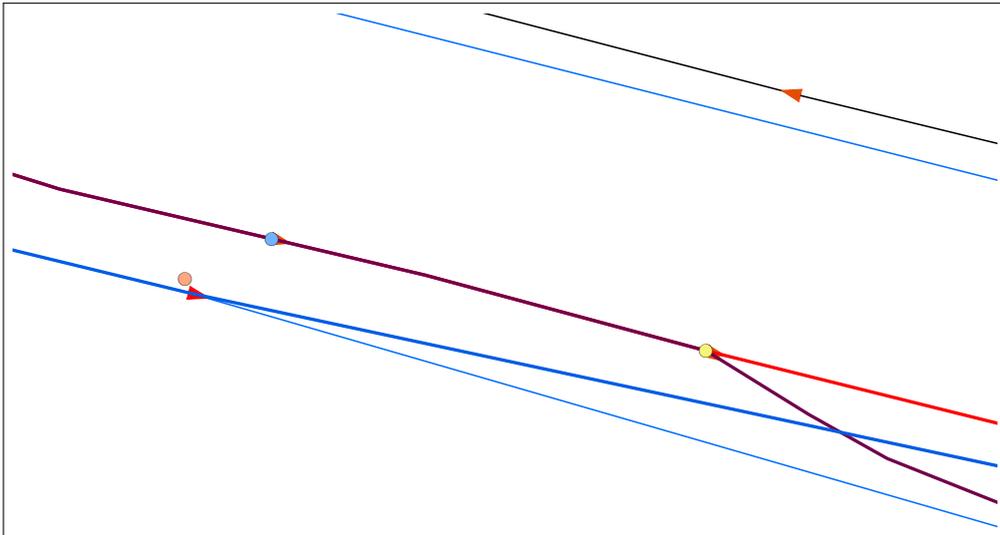


Abbildung 4.4: Beispiel für Relevanz der Knoten-Valenz beim Decoding

## 4.5 Bewertung der Gesamtlocation – Form-Matching

Die Analyse des OpenLR-Verfahrens in Kapitel 3.4.3.5 zeigt Verbesserungsmöglichkeiten durch Einbeziehung von formbasierten Zuordnungsmaßen auf. Mittels der in Abschnitt 4.3.4 beschriebenen Zusammenführung vollständiger Routen-Kandidaten im Zielnetz besteht die Möglichkeit, die Eigenschaften dieser Kandidaten für die Gesamt-Location bei der Ähnlichkeitsbetrachtung zu berücksichtigen. Neben den in den vorangegangenen Abschnitten besprochenen Eigenschaften an den Stützpunkten ist dies vor allem die *Form* einer Ortsreferenz. Wie generell in der vorliegenden Arbeit, gilt dies natürlich nur für linienhafte Ortsreferenzen. Mit formbasierten Zuordnungsparametern ist es möglich, den geometrischen Verlauf einer Ortsreferenz vollständig zu betrachten. Dabei ist es für das menschliche Auge intuitiv, die Ähnlichkeit geometrischer Objekte anhand ihrer Form zu beurteilen. Kleinere Abweichungen im Detail werden dabei toleriert. Wichtig sind vor allem Phänomene mit großen Änderungen der Form, wie z.B. Kreuzungen von Linien, Ecken, große Radien usw. Durch das Betrachten des gesamten Location-Verlaufs und seiner Form kann das Wiederauffinden der Ortsreferenz im Zielnetz deutlich erleichtert werden.

Die Form einer Ortsreferenz, also eines Linienvverlaufs in einem Straßennetzwerk, kann durch sehr viele unterschiedliche Parameter und Eigenschaften beschrieben werden. Bei der Auswahl von Parametern, die einer näheren Untersuchung zugeführt werden sollen, sind zwei Gesichtspunkte / Randbedingungen zu berücksichtigen. Dies ist zum einen die Frage nach der praktischen Umsetzbarkeit sowohl was die softwaretechnische Implementierung auf Encoder- und Decoder-Seite, als auch die Erweiterung

der zu übertragenden Informationen (also dem OpenLR-Datenformat) anbelangt. Unter diesem Gesichtspunkt erscheinen solche Eigenschaften für eine Verwendung zur Formbeschreibung vorteilhaft, bei deren softwaretechnischer Umsetzung wenig Interpretationsspielraum verbleibt. Dadurch können Abweichungen oder Unterschiede zwischen den Software-Implementierungen verschiedener möglicher LR-Anwendungen von vornherein vermieden bzw. reduziert werden. Zum anderen sollen die zu übertragenden Daten möglichst wenig Bandbreite beanspruchen, da die LR-Daten bei vielen Anwendungsszenarien via Mobilfunk übertragen werden. Bei der Entwicklung der etablierten LR-Standards AGORA(-C) bzw. OpenLR war dies noch eine nicht zu vernachlässigende Einschränkung (WEVERS et al. 2003, HENDRIKS et al. 2005, TOMTOM 2012). Dieser Aspekt verliert jedoch aufgrund der zunehmend steigenden Datenübertragungs-Geschwindigkeiten im Mobilfunk sowie stetig fallenden Preisen pro Daten-Einheit, zunehmend an Bedeutung. Eine Erweiterung der zu übertragenden Daten für die hier vorgestellte formbasierte Zuordnung ist deshalb vertretbar.

Grundlage für eine formbasierte Zuordnung sind die in Kapitel 2.6.1.4 vorgestellten Ansätze. Die darin beschriebene Richtungswinkelfunktion zur formbasierten Zuordnung von Polylinien ist für eine Verwendung beim Location Referencing grundsätzlich zweckmäßig. Für die praktische Umsetzung ist allerdings zu beachten, dass die jeweils zu vergleichenden Polylinien nicht gleichzeitig für einen direkten Vergleich zur Verfügung stehen. Vielmehr müssen die für einen formbasierten Vergleich mit den Kandidaten im Zielnetz benötigten Eigenschaften der zu encodierenden Ortsreferenz im Quellnetz ermittelt, codiert und zusammen mit den anderen Referenzierungsparametern in das Zielsystem übertragen werden.

Ausgehend von diesen Überlegungen werden im Folgenden zunächst solche formbeschreibenden Eigenschaften beschrieben, die für mögliche LR-Anwendungen und deren verwendete Kartensysteme sicher und mehrdeutigkeitsfrei umzusetzen und zu implementieren sind und nur wenig Bandbreite für ihre Übertragung beanspruchen.

#### 4.5.1 Anzahl signifikanter Richtungsänderungen

Nach COHEN et al. (1992) besitzen Punkte bzw. Knoten, bei denen die ein- oder ausgehenden Kanten eine große Richtungsänderung aufweisen, grundsätzlich eine anatomische Bedeutung und eignen sich daher besonders als Stützpunkte (engl. *landmarks*) für den Zuordnungsprozess. Diese Eigenschaft nutzen u.a. auch SIRIBA et al. (2011) und verwenden diese Stützpunkte mit großer Richtungsänderung zum longitudinalen Alignment (gegenseitige Einpassung) zweier Polylinien.

Für die Zwecke des Location Referencing wird hier die Detektion von Punkten mit signifikanter Richtungsänderung als *qualitatives* Maß zur Formbeschreibung einer Ortsreferenz vorgeschlagen. Augenscheinlich tragen solche Punkte in besonderer Weise zur Formbeschreibung einer Ortsreferenz (eines Linienzuges) bei.

Die Anzahl der signifikanten Richtungsänderungen  $nDC$  einer Ortsreferenz wird aus allen ihren formbeschreibenden Punkten  $P_i$ ,  $i = 1 \dots n$  berechnet. Diese unterscheidet sich möglicherweise von der Anzahl der echten Knoten eines Linienverlaufs. Dabei werden zunächst für alle Punkte außer dem ersten und letzten die Differenzen der Richtungswinkel der jeweils abgehenden Linie  $\Delta t_i = t_i - t_{i-1}$ ,  $i = \{2, 3, \dots, n-1\}$  berechnet. Diese müssen dann noch so normalisiert werden, dass  $\bar{\Delta t}_i \in (-180, 180]$ , siehe Gleichung 4.7. Damit ist gewährleistet, dass die beiden Drehrichtungen getrennt erfasst und gezählt werden können. Alle nach der Normalisierung positiven Richtungsänderungen von  $\Delta t \leq 180^\circ$  signalisieren also eine Rechtsdrehung im Uhrzeigersinn, alle negativen eine Linksdrehung gegen den Uhrzeigersinn.

$$\bar{\Delta t} = \begin{cases} \Delta t - 360 & : \Delta t > 180[^\circ] \\ \Delta t + 360 & : \Delta t \leq -180[^\circ] \\ \Delta t & : -180[^\circ] < \Delta t \leq 180[^\circ] \end{cases} \quad (4.7)$$

Es wird dann die Anzahl aller Richtungswinkeländerungen  $nDC$  gezählt, die größer als ein festzulegender Grenzwert  $t_{min}$  sind, siehe Gleichung 4.8. Im Rahmen dieser Arbeit wurde  $t_{min} = 50^\circ$  angenommen.

$$nDC = \sum_{i=2}^{n-1} (f(\bar{\Delta t}_i)) \quad \text{mit} \quad f(\bar{\Delta t}_i) = \begin{cases} 1 & : |\bar{\Delta t}| > t_{min} \\ 0 & : |\bar{\Delta t}| \leq t_{min} \end{cases} \quad (4.8)$$

Da OpenLR bislang keine formbeschreibenden Eigenschaften der Ortsreferenzen für die Zuordnung verwendet, wird der Wert  $nDC_{src}$  im Encoding-Prozess für die ursprüngliche Ortsreferenz im Quellnetz nach Gleichung 4.8 bestimmt und als Ergänzung der Location Reference ins Zielsystem übertragen. Dort bestimmt der Decoder  $nDC_{cand}$  für alle Routen-Kandidaten und vergleicht diese mit dem Soll-Wert aus dem Quellsystem.

### 4.5.2 Kumulierte Richtungsänderungen

Für eine etwas detailliertere Formbeschreibung bei gleichzeitig geringer Übertragungs-Bandbreite sei hier die kumulierte Richtungsänderung  $\delta t$  einer Ortsreferenz als *quantitatives* Maß entsprechend Gleichung 4.9 vorgeschlagen. Dadurch kann die im vorigen Abschnitt erläuterte Anzahl der signifikanten Richtungsänderungen weiter quantifiziert werden. Sie basiert auf der in Kapitel 2.6.1.4 vorgestellten kumulierten Richtungswinkelfunktion. Dabei werden die Richtungswinkeländerungen  $\bar{\Delta t}_i$  für alle Bruchpunkte einer Ortsreferenz aufsummiert.

$$\Delta t_{sum} = \sum_{i=1}^{n-1} \bar{\Delta t}_i, \quad \bar{\Delta t}_i \in (-180, 180] \quad (4.9)$$

Als Alternative lassen sich die Anteile der positiven und negativen Richtungsänderungen  $\Delta t_{sum}^+$  und  $\Delta t_{sum}^-$  getrennt aufsummieren. Dies hat den Vorteil, dass sich positive und negative Richtungsänderungen nicht gegenseitig ausmitteln können. Dadurch kann beispielsweise eine langgestreckte Ortsreferenz ohne größere Richtungswechsel von einer anderen unterschieden werden, deren Richtungswechsel im als auch gegen den Uhrzeigersinn einen ähnlichen Betrag aufweisen.

Ähnlich wie im vorigen Abschnitt werden auch für die Nutzung der kumulierten Richtungsänderungen zur Zuordnung die Parameter  $\Delta t_{sum}^+$  und  $\Delta t_{sum}^-$  im Encoder des Quellsystems für jede Ortsreferenz entsprechend Gleichung 4.9 bestimmt und zusammen mit der Location Reference ins Zielsystem übertragen. Dort erfolgt der Vergleich mit den entsprechend für die Routen-Kandidaten bestimmten Größen.

## 4.6 Berechnung der Leistungsfunktion

Die in den vorangegangenen Abschnitten vorgestellten Zuordnungs-/Ähnlichkeitsmaße werden für jeden Kandidaten individuell berechnet. Für die Gesamtbeurteilung des jeweiligen Routen-Kandidaten werden sämtliche seiner Zuordnungsmaße mittels einer Leistungsfunktion (vgl. Kapitel 2.7) zusammengeführt. Die dazu in diesem Abschnitt dargestellten Gewichtungsfaktoren und Schwellwerte wurden durch empirische Untersuchungen festgelegt und haben sich als zweckmäßig erwiesen. Als bester Kandidat wird derjenige mit dem höchsten / maximalen Wert der Leistungsfunktion ausgewählt. Dieser ist dann die decodierte Ortsreferenz  $\hat{\mathbf{L}}_i^T$  im Zielnetz.

### 4.6.1 Bewertungsanteile je LRP

Zunächst wird für alle Decoding Kandidaten (vgl. Kapitel 4.3) die Bewertung aus der OpenLR Musterimplementierung, Version 1.4.0 (TOMTOM 2014) berechnet. Beim klassischen OpenLR werden Kandidaten für jede Location Reference Point (LRP) separat mittels Suchradius gesucht, einzeln bewertet und der beste Kandidat dann ausgewählt. Beim hier vorgestellten Ansatz werden jedoch Routenverläufe, welche die zu decodierende Ortsreferenz vollständig abdecken, als Kandidaten betrachtet. Daher müssen die Rating-Anteile  $R_{LRP}$ , die Location Reference Point-Informationen verwenden, zunächst für jeden LRP gesondert berechnet und dann für die Gesamt-Bewertung zusammengeführt werden (siehe Abschnitt 4.6.3). Im Gegensatz zur OpenLR Musterimplementierung ist es in der hier vorgestellten Implementierung möglich, individuelle Gewichte  $w_i$  für alle Bewertungsanteile zu vergeben.

$$R_{LRP} = w_{dist} \cdot R_{dist} + w_t \cdot R_t + w_{val} \cdot R_{val} + w_{FRC} \cdot R_{FRC} + w_{FOW} \cdot R_{FOW} \quad (4.10)$$

Tabelle 4.10 zeigt eine Übersicht der empirisch bestimmten und verwendeten Gewichte der Bewertungsanteile für die einzelnen Location Reference Points.

Tabelle 4.10: Verwendete Gewichte für das Knoten-Rating

Parameter	$w_{dist}$	$w_{val}$	$w_t$	$w_{FRC}$	$w_{FOW}$
Wert	4	1	1	1	1

### 4.6.2 Bewertung der Form für die gesamte Ortsreferenz

Der Form-Matching Anteil des Ratings besteht aus den Teilen für die Anzahl der Differenz signifikanten Richtungswinkeländerungen und der Differenz der kumulierten Richtungswinkeländerungen.

#### Differenz signifikanter Richtungswinkeländerungen

Die Funktionswert für die Bewertung der Differenz der signifikanten Richtungswinkeländerungen  $R_{nDC}$  soll kleiner werden, je größer die Differenz der signifikanten Richtungswinkeländerungen aus dem Quell- und Zielnetz wird. Weisen im Optimalfall beide Ortsreferenzen die selbe Anzahl an Richtungsänderungen auf, so liefert die Bewertungsfunktion den maximalen Wert  $R_{nDC} = 100$ . Übersteigt die Differenz einen Schwellwert  $nDC_{max}$ , so ist der Funktionswert  $R_{nDC} = 0$  und leistet keinen Beitrag zum Gesamtrating mehr. Für die vorliegende Arbeit hat sich ein Wert  $nDC_{max} = 3$  als zweckmäßig erwiesen. Die Bewertungsfunktion lässt sich wie folgt definieren:

$$R_{nDC} = \frac{\max(0, nDC_{max} - |nDC_{src} - nDC_{cand}|)}{nDC_{max}} \cdot 100. \quad (4.11)$$

#### Differenz kumulierter Richtungswinkeländerungen

Analog können auch für die kumulierten positiven und negativen Richtungsänderungen  $\Delta t_{sum}^+$  und  $\Delta t_{sum}^-$  die Bewertungsfunktionen  $R_{DC}^+$  und  $R_{DC}^-$  definiert werden. Die Differenz der kumulierten Richtungswinkeländerungen wird allerdings nochmals mit der Anzahl der Kanten der Quell-Ortsreferenz  $nL_{src}$  normiert, um eine nachteilige Bewertung durch mit steigender Kantenzahl ansteigendem Funktionswert auszugleichen. Zur Berechnung einer Leistungsfunktion wird der Schwellwert  $DC_{max}$  für die maximal akzeptierte Differenz der kumulierten Richtungswinkeländerungen eingeführt. Im Rahmen der vorliegenden Arbeit wurde  $DC_{max} = 20^\circ$  empirisch bestimmt und für die weiteren Auswertungen verwendet. Da die Funktion für positive und negative Richtungswinkeländerungen getrennt berechnet wird, aber

beide Teile zusammen zu einem maximalen Funktionswert von 100 aufsummieren sollen, ist jede Teilfunktion zunächst mit dem Multiplikator 50 versehen.

$$R_{DC}^+ = \frac{\max\left(0, DC_{max} - \frac{|\Delta_{sum,src}^+ - \Delta_{sum,cand}^+|}{nL_{src}}\right)}{DC_{max}} \cdot 50 \quad (4.12)$$

$$R_{DC}^- = \frac{\max\left(0, DC_{max} - \frac{|\Delta_{sum,src}^- - \Delta_{sum,cand}^-|}{nL_{src}}\right)}{DC_{max}} \cdot 50 \quad (4.13)$$

### Gesamtmaß für die Ortsreferenz Form

Die zuvor vorgestellten formbeschreibenden Bewertungsanteile können nun wie folgt zum Bewertungsmaß für die Übereinstimmung der Ortsreferenz Form zusammengefasst werden. Der individuelle Einfluss jedes einzelnen Bewertungsanteils kann über die Gewichte  $w_{nDC}$ ,  $w_{DC+}$  und  $w_{DC-}$  gesteuert werden:

$$R_{form} = w_{nDC} \cdot R_{nDC} + w_{DC+} \cdot R_{DC}^+ + w_{DC-} \cdot R_{DC}^- \quad (4.14)$$

Die verwendeten Gewichte für die Bewertung der Ortsreferenz Form sind in Tabelle 4.11 zusammengestellt, sie haben sich bei empirischen Untersuchungen als geeignet erwiesen.

Tabelle 4.11: Verwendete Gewichte für das Form-Rating

Parameter	$w_{nDC}$	$w_{DC+}$	$w_{DC-}$
Wert	1	0.2	0.2

### 4.6.3 Zusammenführung zur Gesamt-Leistungsfunktion

Aus dem Mittelwert der Ratings der einzelnen Location Reference Points  $R_{LRP}$  (Gleichung 4.10) sowie dem Form-Rating  $R_{form}$  der Gesamtlocation (Gleichung 4.14) wird die Gesamt-Leistungsfunktion für jeden Ortsreferenz-Kandidaten berechnet:

$$R = \frac{\sum(R_{LRP})}{nLRP} + R_{form} \quad (4.15)$$

## 4.7 Zuordnung mittels Frechét-Distanz

In Kapitel 2.6.1.1 wurden verschiedene Metriken als geometrische Zuordnungsmaße vorgestellt. Die euklidische Distanz wird bereits in vielen Location Referencing Verfahren verwendet, so auch bei OpenLR zur Berechnung der Punktabstände und des darauf basierenden Knoten-Ratings. Die in vielen Veröffentlichungen genannte *Hausdorff*-Distanz kann zur Ähnlichkeitsbetrachtung von Graphen verwendet werden. Allerdings werden nur ungerichtete Graphen betrachtet, sodass die Hausdorff-Metrik eher zum Vergleich von Punktwolken geeignet ist. Eine Ähnlichkeitsbetrachtung von Polylinien für das hier diskutierte Location Referencing ist nicht möglich, da die Punktfolgenfolge für die Polylinien ein entscheidendes Definitionsmerkmal ist. Die Frechét-Distanz hingegen eignet sich sehr gut zur Ähnlichkeitsbestimmung von gerichteten Graphen, also auch von diskreten Punktfolgen, wie sie hier bei Polylinien zum Einsatz kommen. Für eine Verwendung der diskreten Frechét-Distanz für eine Zuordnung in Location Referencing Anwendungen müssen allerdings einige Voraussetzungen erfüllt sein.

Zunächst muss die Geometrie der zu übertragenden Ortsreferenz aus der Quellkarte für die Berechnung der Frechét-Distanz im Zielsystem vollständig bekannt sein. Das bedeutet, die Koordinaten aller formbeschreibenden Punkte müssen mittels LR in das Zielsystem übermittelt werden. Dies ist bislang bei LR-Anwendungen nicht der Fall, da die Geometrie der zu übertragenden Ortsreferenzen durch Stützpunkte diskretisiert und repräsentiert wird. Nichtsdestotrotz bietet die Frechét-Distanz die Möglichkeit, die geometrische Nähe zweier gerichteter Polylinien (in dem hier diskutierten Fall: der linienhaften Ortsreferenzen), zu bestimmen.

### 4.7.1 Diskrete Frechét-Distanz

Die in Kapitel 2.6.1.1 beschriebene *Frechét-Distanz* eignet sich sehr gut zur Berechnung des Abstandes zwischen zwei gerichteten Graphen. Die im Rahmen der vorliegenden Arbeit verwendeten linienhaften Ortsreferenzen sind Polylinien und bestehen aus einer Folge diskreter Einzelpunkte. Dementsprechend kann die von EITER und MANNILA (1994) vorgeschlagene *diskrete* Frechét-Distanz zwischen zwei Ortsreferenzen als Maß für die Ähnlichkeit der Form betrachtet werden.

Die diskrete Frechét-Distanz kann in der Form von MASCRET et al. (2006) mittels zweier Matrizen berechnet werden, welche alle Punkte  $P_{i,k}, k = \{1 \dots m\}$  der ursprünglichen Ortsreferenz  $L_i^S$  im Quellnetz und die Punkte  $P_{j,l}, l = \{1 \dots n\}$  des Ortsreferenz Kandidaten  $L_{i,j}^T$  im Zielnetz abdecken. Die erste Matrix  $\mathbf{M}_d$  beinhaltet die Euklidische Distanzen (vgl. Kapitel 2.6.1.1) zwischen den Punkten der beiden zu vergleichenden Ortsreferenzen:

$$m_d(k, l) = d_e(P_{i,k}, P_{j,l}). \quad (4.16)$$

Die zweite Matrix  $\mathbf{M}_{dF}$  dient nun zur Berechnung der diskreten Frechét-Distanz, sie hat die Dimension  $m \times n$ . Ihre Elemente  $m_{dF}(k, l)$  werden sequentiell nach Gleichung 4.17 berechnet, beginnend mit Element  $m_{dF}(0, 0)$ :

$$m_{dF}(k, l) = \max(m_d(k, l), \min(m_{dF}(k-1, l), m_{dF}(k, l-1), m_{dF}(k-1, l-1))). \quad (4.17)$$

In dem in Abbildung 4.5 dargestellten Beispiel besteht die Ausgangs-Location in der Quellkarte [blau] aus zwei Punkten, der Ortsreferenz-Kandidat in der Zielkarte [schwarz] aus sechs Punkten. Es werden zunächst die Euklidischen Distanzen zwischen allen Punkten berechnet und in die Matrix  $M_d$  eingetragen, siehe Tabelle 4.12. Dabei werden die Punkte in der durch einen Pfeil dargestellten Richtung der Ortsreferenz durchlaufen.

Tabelle 4.12: Matrix mit Euklidischen Distanzen [m]

	$p_{2,1}$	$p_{2,2}$	$p_{2,3}$	$p_{2,4}$	$p_{2,5}$	$p_{2,6}$
$p_{1,1}$	6.06	8.57	21.88	35.86	75.61	95.62
$p_{1,2}$	89.46	81.94	67.91	53.97	15.32	10.86

Die Frechét-Distanz wird nun in einer neuen Matrix  $M_{dF}$  (siehe Tabelle 4.13) iterativ aus den Elementen der Tabelle mit den Euklidischen Distanzen berechnet.

Tabelle 4.13: Matrix zur iterativen Berechnung der Frechét-Distanz [m]

	$p_{2,1}$	$p_{2,2}$	$p_{2,3}$	$p_{2,4}$	$p_{2,5}$	$p_{2,6}$
$p_{1,1}$	6.06	8.57	21.88	35.86	75.61	95.62
$p_{1,2}$	89.46	81.94	67.91	53.97	35.86	35.86

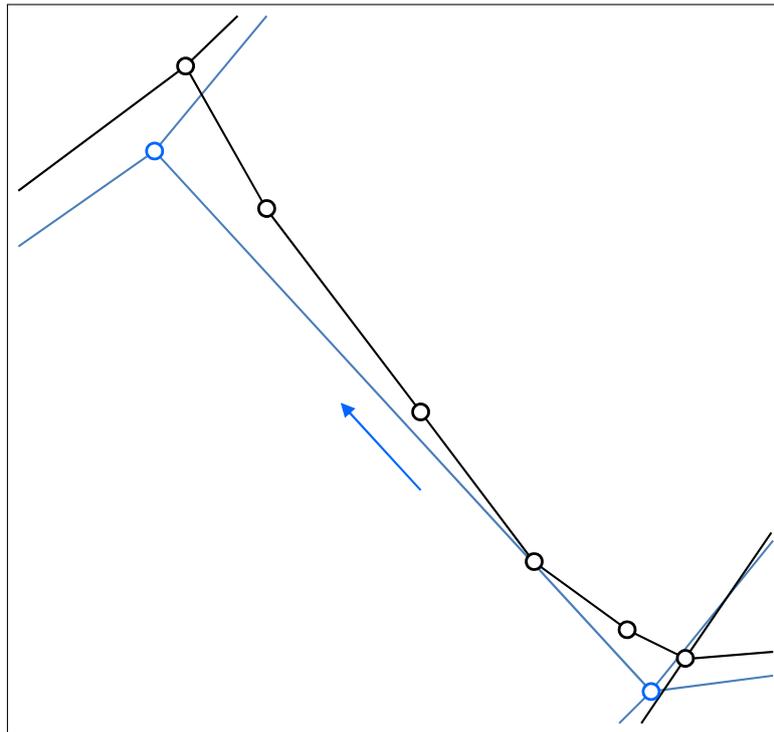


Abbildung 4.5: Beispiel Frechet-Distanz

Die diskrete Frechét-Distanz  $d_{dF}$  wird dem letzten Element der Berechnungsmatrix  $m_{dF}(m, n)$  entnommen und hat in diesem Beispiel den Wert 35.86 m.

In diesem konkreten Fall handelt es sich um zwei Ortsreferenzen mit einem signifikanten Unterschied der Zahl der Zwischenpunkte (Shape Points). Die Ausgangs-Location ist nur durch ihren Anfangs- und den Endpunkt definiert. Der Decodier-Kandidat in der Zielkarte hat aus Anfangs- und Endpunkt noch vier formbeschreibende Zwischenpunkte.

Eigentlich ist die diskrete Frechét-Distanz ohne Resampling der beiden Polylinien nur zum Vergleich von Linien mit ungefähr gleichabständigen Stützpunkten geeignet (EITER und MANNILA 1994). Im vorliegenden Fall soll jedoch nicht das *absolute* Minimum der Frechét-Distanz zwischen zwei Kurven berechnet werden. Es ist vielmehr das Ziel, den am besten passenden Decodier-Kandidaten durch das *relative* Minimum seiner diskreten Frechét-Distanz zu der Ausgangs-Location, zu finden.

#### 4.7.2 Bewertungsmaß

Wie gesehen, ist die diskrete Frechét-Distanz ein geometrisches Abstandsmaß zwischen zwei gerichteten Polylinien. Je größer ihr Funktionswert, desto weiter sind die Polylinien voneinander entfernt bzw. desto unähnlicher ist ihre Form. Je kleiner hingegen die diskrete Frechét-Distanz, desto ähnlicher sind sich die beiden Ortsreferenzen.

Die OpenLR Musterimplementierung enthält zur Bewertung von LRP Zuordnungskandidaten bereits einen Ansatz zur Ableitung eines Bewertungsmaßes aus berechneten Euklidischen Distanzen (Kapitel 3.4.2.2). Er basiert auf einem Vergleich mit einem Grenzwert, der maximal akzeptierten Entfernung zwischen den zu vergleichenden Punkten. Überschreitet die berechnete Distanz diesen Maximalwert, so wird das Bewertungsmaß zu Null. Durch die Bildung der Differenz der Entfernung zum akzeptierten Grenzwert, kann das Entfernungsmaß für die zur Zusammenführung aller Ähnlichkeitsparameter verwendete Leistungsfunktion verwendet werden. In dieser wird diejenige Bewertung, die den höchsten

Funktionswert aufweist, als bestmögliche Zuordnung ausgewählt. Da in diesem Fall einzelne Punktkandidaten bewertet werden, ist die Verwendung eines Grenzwertes sinnvoll. Punkte, die weiter entfernt liegen, als diese Grenze, sollen eine niedrige Bewertung erhalten.

Die Wahl eines solchen Grenzwertes zur Bestimmung eines Bewertungsmaßes der Zuordnungskandidaten aus der diskreten Frechét-Distanz ist nicht einfach. Wie bei ihrer Definition im vorigen Abschnitt gesehen, ist der Funktionswert der Frechét-Distanz nämlich nicht immer intuitiv 'greifbar'. Es können mitunter auch sehr große Funktionswerte entstehen, die einen zu definierenden Grenzwert übersteigen, aber trotzdem auf eine relative Ähnlichkeit der verglichenen Ortsreferenzen hindeuten. Es ist sinnvoller, die Bewertung mit Hilfe einer Minimum-Funktion durchzuführen. Damit soll derjenige Kandidat den höchsten Wert für die Leistungsfunktion erhalten, der die *kleinste* diskrete Frechét-Distanz aller Kandidaten  $dF_{min}$  für diese Ortsreferenz aufweist. Es handelt sich also um ein *relatives* Ähnlichkeitsmaß. Das Bewertungsmaß  $R_{dF}$  wird dann aus der Differenz der diskreten Frechét-Distanz des zu bewertenden Ortsreferenz Kandidaten  $dF_{cand}$  zu  $dF_{min}$  bestimmt. Die Bewertungsfunktion kann maximal einen Wert von  $R_{dF} = 100$  annehmen. Im schlechtesten Fall wird die Bewertung zu 0 und trägt nicht zum Rating des Ortsreferenz Kandidaten bei. Durch den eingeführten Faktor  $f_{dF}$  kann implizit definiert werden, ab welcher Differenz zum Minimum die Bewertung zu 0 wird. Für  $f_{dF} = 1$  wird  $R_{dF} = 0$  für alle  $\Delta dF \geq 100m$ .

$$R_{dF} = \max(0, 100 - f_{dF} \cdot (dF_{cand} - dF_{min})) \quad (4.18)$$

In Anlehnung an das Gesamtbewertungsmaß aus Kapitel 4.6.3 kann eine Gesamtbewertung auf Grundlage des Mittelwerts der Einzelbewertungen der Location Reference Points berechnet werden. Dieser wird als  $R_{LRP^*}$  bezeichnet, da die Punktdistanzen nicht berücksichtigt werden (siehe Gleichung 4.19). Dazu werden die Bewertung der Ortsreferenz Form  $R_{form}$  sowie die Bewertung auf Grundlage der diskreten Frechét-Distanz  $R_{dF}$  addiert. Zur individuellen Gewichtung der jeweiligen Teilbewertungen werden die Faktoren  $w_N$  für die LRPs,  $w_{form}$  für die Ortsreferenz Form sowie  $w_{dF}$  für die diskrete Frechét-Distanz eingeführt:

$$R_{LRP^*} = w_{val} \cdot R_{val} + w_t \cdot R_t + w_{FRC} \cdot R_{FRC} + w_{FOW} \cdot R_{FOW}, \quad (4.19)$$

$$R = w_N \cdot \frac{\sum(R_{LRP^*})}{nLRP} + w_{form} \cdot R_{form} + w_{dF} \cdot R_{dF}. \quad (4.20)$$

## 4.8 Zusammenfassung

Die in diesem Kapitel vorgestellten Location Referencing Verfahren greifen den bestehenden Ansatz von OpenLR auf und verwenden zum Teil die darin definierten Zuordnungsmerkmale. Dieser bestehende Bewertungsansatz wird um formbeschreibende, geometrische Merkmale erweitert und topologische Kriterien werden beim Decodieren im Zielnetz entsprechend ihrer Definition angewendet. Die Form linienhafter Ortsreferenzen wird durch die Anzahl der als signifikant gewerteten Richtungsänderungen im Quellnetz sowie durch die Kumulation der positiven und negativen Richtungsänderungen beschrieben. Dadurch wird eine auch im Hinblick auf die Übertragungskapazitäten in realen Anwendungen optimierte Möglichkeit der Formbeschreibung gefunden. Das Bewertungsverfahren zur Identifikation der Zuordnungen wurde entsprechend ergänzt. Für diese neuen Ähnlichkeitsmerkmale ist es allerdings erforderlich, das bestehende OpenLR-Datenformat um zusätzliche formbeschreibende Eigenschaften zu erweitern, die vom Encoder im Quellnetz erfasst und dann mit dem Location Reference Code in das Zielsystem übertragen werden. Daneben werden die Bewertungsfunktionen für die bestehenden geometrischen und topologischen Zuordnungsmerkmale optimiert. Bestehende Anforderungen an die Knotenvalenz bei der Auswahl der Location Reference Points zur Erstellung von Ortsreferenzen im Quellnetz, die dem OpenLR Whitepaper zu entnehmen sind, werden in dem hier vorgestellten Verfahren beim Decodieren im Zielnetz strikter als in der OpenLR Musterimplementierung angewendet.

Dabei werden LRP-Kandidaten, die keine echten Netzknoten sind (Wertigkeit  $\leq 2$ ) abgewertet und dadurch echte Netzknoten bei der Zuordnungsentscheidung bevorzugt.

Die Einführung formbeschreibender Zuordnungsmerkmale bedingt im Gegensatz zur bislang von OpenLR verwendeten Betrachtung der einzelnen Stützpunkte unabhängig von einander, die Bewertung vollständiger Routenverläufe als Zuordnungskandidaten. Erst die Beurteilung der gesamten Ortsreferenz erlaubt eine sichere Zuordnungsentscheidung. Bei der in der OpenLR Musterimplementierung angewandten individuellen Betrachtung der einzelnen Stützpunkte (LRPs) kann es vorkommen, dass für diese sehr gut passende Zuordnungskandidaten im Zielnetz gefunden werden, die jedoch in der Gesamtschau der Ortsreferenz keine große Übereinstimmung mit der Ausgangslocation aus dem Quellnetz aufweisen. Der gesamte Bewertungs- und Auswahlprozess der Kandidaten wird deshalb im neuen Verfahren so gestaltet, dass alle zwischen den gesuchten Stützpunkten (LRP) im Zielnetz möglichen Routenkombinationen zunächst gesucht, dann bewertet und schließlich die am besten passende Gesamtroute als Decoding-Lösung ausgewählt wird. Dadurch ist es mit dem Form-Matcher auch möglich, unrealistische Ergebnisse der Routenberechnung zu erkennen (beispielsweise wenn diese eine Kante nachfolgend in beide Richtungen durchläuft) und solche Routen von der weiteren Betrachtung auszuschließen.

Eine ausschließlich formbasierte Zuordnung bietet jedoch nicht immer genügend Informationen, um eine eindeutige und auch korrekte Zuordnung vorzunehmen. So weisen oft mehrere in der Kandidaten Menge enthaltene Ortsreferenzen eine ähnliche Form auf. Erst die Betrachtung ihrer weiteren Eigenschaften hilft, eine fundierte und schließlich korrekte Zuordnungsentscheidung zu treffen. Deshalb ist die Semantik neben der Form ein wichtiger Faktor für die Ähnlichkeitsbetrachtung und die Zuordnungsentscheidung. So können dank des neuen Ansatzes solche Stützpunkt-Kandidaten, die eine vermeintlich große Ähnlichkeit mit den übermittelten Eigenschaften der LRPs aufweisen, aus der Menge der Zuordnungskandidaten ausgeschlossen werden, wenn sie große Abweichungen in den formbeschreibenden Ähnlichkeitsparametern aufweisen.

In einem weiteren Teil wird ein geometrischer Zuordnungsansatz basierend auf der diskreten Frechét-Distanz entwickelt. Die Frechét-Metrik kann zur Ähnlichkeitsbestimmung gerichteter Graphen herangezogen werden. Ein daraus abgeleitetes Ähnlichkeitsmaß wird im Rahmen der vorliegenden Arbeit als weiterer Baustein zur Robustifizierung der Location Referencing Zuordnung entwickelt und implementiert.

Die Implementierung des hier vorgestellten Location Referencing Verfahrens sowie die sich daran anschließenden empirischen Untersuchungen und deren Ergebnisse werden ausführlich in Kapitel 7 dargestellt.

# 5 Qualität von Geodaten

## 5.1 Einführung

DIN EN ISO 9000 (2005) bezeichnet Qualität als

**„Grad, in dem ein Satz inhärenter Merkmale Anforderungen erfüllt“.**

Etwas ausführlicher und intuitiver zugänglich definiert BIROLINI (1991) in einer älteren Definition die Qualität als „Grad, in welchem die Betrachtungseinheit den durch den Verwendungszweck gestellten funktionellen, operationellen und physikalischen Eigenschaften und Anforderungen genügt“. Dies kommt unserem alltäglichen Gebrauch des Begriffes *Qualität* sehr nahe, wonach wir mit *Qualität* meist den Erfüllungsgrad unserer Erwartungen an ein Produkt meinen. Demnach ist ein Produkt für uns von hoher Qualität, wenn es die von uns (meist unbewusst) aufgestellten Anforderungen in hohem Maße erfüllt. Produkte in diesem Sinne können natürlich auch nicht-materielle Güter, wie zum Beispiel Dienstleistungen oder Informationen bzw. Daten sein.

Um dieses rein subjektive Qualitätsempfinden auf mehrere Produkte bzw. von einer Gruppe von Kunden oder Nutzern zu verallgemeinern bzw. zu formalisieren, werden Qualitätskonzepte aufgestellt und verwendet. Qualitätskonzepte bestehen im Allgemeinen aus einem *Qualitätsmodell* zur Beschreibung der Qualität mittels spezifischer Kriterien und einem *Analyseverfahren* zur Ermittlung der spezifischen Qualität eines Produktes (WILTSCHKO 2004, WILTSCHKO und KAUFMANN 2005). Im Sinne der Überwachung und Verbesserung der Qualität eines Produktes können Prozesse der Qualitätssicherung dienen, die beispielsweise in der ISO 9000 Normenfamilie beschrieben sind (DIN EN ISO 9000 2005, DIN EN ISO 9001 2008). Häufig bilden diese Qualitätssicherungsverfahren einen wiederkehrenden Prozess der Qualitätsermittlung mit Hilfe des zuvor genannten Analyseverfahrens auf Grundlage des für die jeweilige Anwendung aufgestellten Qualitätsmodells. Das so ermittelte Qualitätsniveau wird mit den Qualitätsanforderungen verglichen, die beispielsweise aus Kundenbefragungen oder anderen Vorgaben abgeleitet sein können. Durch diesen Vergleich wird dann sichtbar, ob die aktuelle Produktqualität den Vorgaben entspricht, oder ob und ggfs. wie nachgesteuert werden muss, um die Qualitätsziele zu erreichen. Dieser Gesamtprozess muss in vorher festgelegten, regelmäßigen Zyklen wiederholt werden. Man spricht daher auch von einem *kontinuierlichen Verbesserungsprozess*, japanisch *Kaizen* genannt. Die Anfänge des Kaizen wurden in den 1950er Jahren von dem Amerikaner W.E. Deming in Japan gelegt (siehe u.a. DEMING (1982)). Eine Anwendung der kontinuierlichen Verbesserung in einem Qualitätsmanagement für Verkehrssysteme ist in SCHÜTZLE und SCHEIDER (2009) beschrieben. Allgemein werden die zur Beschreibung, Erfassung, Sicherung und Verbesserung der Qualität eines Produktes unternommenen Anstrengungen unter dem Begriff des *Qualitätsmanagements* zusammengefasst. Für einen umfassenden Überblick über Qualitätsmanagementsysteme sei an dieser Stelle auf ZOLLONDZ (2011) verwiesen. In der Arbeit von LAUFER (2011) findet sich eine interessante Anwendung in der Verkehrstechnik.

Ein Qualitätsmodell besteht aus einem Satz festgelegter, inhärenter Qualitätsmerkmale, mit deren Hilfe sich die Qualität der zu betrachtenden Daten bzw. Informationen vollumfänglich beschreiben lässt. Die (abstrakten) Qualitätsmerkmale müssen für die jeweilige Fragestellung bzw. Anwendung durch individuell abgeleitete Qualitätsparameter konkretisiert werden. Diese Qualitätsparameter sind mess- oder

bestimmbare Größen, anhand derer sich die Qualitätseigenschaften eines Produkts bzw. einer Information direkt bestimmen lassen. Es ist darauf zu achten, dass das Modell die Qualitätseigenschaften vollständig abbildet, d.h. sämtliche an das entsprechende Produkt gerichteten Qualitätsanforderungen müssen sich mit den Merkmalen und Parametern des Modells beschreiben lassen. Ebenso müssen diese Merkmale und Parameter unabhängig voneinander definiert sein und die Parameter dürfen sich auch nicht gegenseitig beeinflussen.

Eine solche zweistufige Qualitätsmodellierung durch Qualitätsmerkmale und ihre beschreibenden Parameter wurde beispielsweise von WILTSCHKO (2004) und WILTSCHKO und KAUFMANN (2005) vorgeschlagen. Andere Modellierungsansätze (z.B. ISO 19157 (DIN EN ISO 19157 2014)) verwenden noch eine dritte Modellierungsschicht, und unterteilen die Qualitätsmerkmale in Teilmerkmale. Es können immer mehrere Parameter für ein Qualitätsmerkmal herangezogen werden. Qualitäts(an)forderungen, also Definitionen eines Mindestmaßes an die Qualität eines Produkts oder einer Information, lassen sich mithilfe der Qualitätsparameter formulieren. Abbildung 5.1 zeigt den zugrunde gelegten Zusammenhang zwischen der Qualität einer Information, den beschreibenden Qualitätsmerkmalen und deren Parametern sowie den daraus formulierten Qualitätsforderungen.

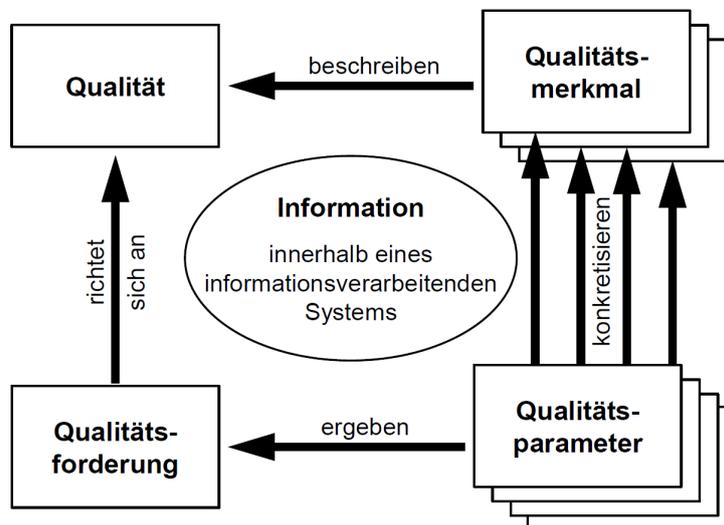


Abbildung 5.1: Qualitätsmodell für Informationen (aus WILTSCHKO (2004))

Nach SCHWEITZER und SCHWIEGER (2011) ist ein Qualitätsmodell sowohl auf Produkte als auch auf Prozesse anzuwenden. Sie verwenden dazu ein zweistufiges Qualitätsmodell nach WILTSCHKO (2004) und unterscheiden die Qualitätsparameter unter Bezug auf ISO/TS 19138 (2006) bzw. den Nachfolgestandard DIN EN ISO 19157 (2014) in vier Typen: *Messwert*, *Indikator*, *Zähler* und *Rate*, siehe Tabelle 5.1.

BEEZ et al. (2013) und auch ZHANG und SCHWIEGER (2011) beschreiben die Zusammenführung der einzelnen Qualitätsparameter zu einem globalen Qualitätsparameter, der auch als *Total Quality Index* bezeichnet wird. Dieser wird als gewichteter Mittelwert aller Einzelparameter gebildet. Qualitätsparameter, die für das Produkt von besonderer Wichtigkeit sind, können darüber hinaus zur Abwertung des globalen Qualitätsparameters bzw. zur Nicht-Erfüllung der globalen Qualitätsanforderungen führen. Dies kann beispielsweise durch die Einführung sehr hoher Gewichte für diese Parameter bei der Berechnung des globalen Mittelwerts erzielt werden. Alternativ lassen sich diese wichtigen Parameter auch separat abfragen und bei Nicht-Einhaltung der vorgegebenen Anforderungen wird das Gesamtqualitätsmaß minimal, was dem schlechtest möglichen Wert entspricht.

Tabelle 5.1: Typunterscheidung für Qualitätsparameter (nach DIN EN ISO 19157 (2014))

Parameter Typ	Definition	Werttyp
Messwert	Direkt mess-/beobachtbare Größe bzw. Performanz	Realwert
Indikator	Binäre ja/nein Entscheidung	Boole'sch
Zähler	Anzahl Elemente aus ja/nein Entscheidungen	Ganzzahl
Rate	Anzahl ja/nein Entscheidungen im Verhältnis zur Gesamtzahl der Elemente	Prozent

## 5.2 Qualitätsmodelle für Geodaten

### 5.2.1 Datenqualität in der Geoinformatik

Unter dem Begriff 'Qualität von geographischen Daten (Geodaten)' wird üblicherweise die Qualitätsbeschreibung eines Datensatzes oder einer ganzen Datenquelle in Bezug auf ihre Übereinstimmung mit der *konzeptionellen Realität* (WILTSCHKO 2004) oder auch Gegenstandsbereich (engl. *Universe of Discourse*, siehe DIN EN ISO 19157 (2014)) verstanden. Abweichungen in geometrischer, topologischer und semantischer Sicht können durch falsche oder fehlende Erfassung bzw. durch fehlerhafte Verarbeitung oder Modellierung entstehen.

Die konzeptionelle Realität stellt einen idealen Datensatz als korrektes Abbild der realen Welt entsprechend der jeweiligen Spezifikation für den Datensatz dar (DIN EN ISO 19157 2014). Sie kann als gewünschten Grad der Abstraktion und Generalisierung definiert werden (VEREGIN 1999). Die Modellierung der realen Welt ist die Voraussetzung der rechnergestützten Darstellung und Verarbeitung von Phänomenen der realen Welt in Form von Daten (vgl. WILTSCHKO (2004), BILL (2010)). Dabei erfolgt in der konzeptionellen Modellierung die Festlegung, welche Objekte der realen Welt für eine Erfassung relevant sind und wie diese dargestellt bzw. abstrahiert werden. Ergebnis ist eine fachspezifische Abstraktion der realen Welt unter Berücksichtigung der speziellen Anforderungen der jeweiligen Anwendung.

Ähnlich wie WILTSCHKO (2004), wertete auch VAN OORT (2005) verschiedene existierende Normen und Standards bzgl. Datenqualität aus und hat daraus einen Satz von acht Qualitätselementen gebildet, die in den meisten der ausgewerteten Standards explizit genannt sind:

- Lineage (Herkunft),
- Lagegenauigkeit,
- Attributgenauigkeit,
- Logische Konsistenz,
- Vollständigkeit,
- Semantische Genauigkeit,
- Verwendung, Zweck, Bedingungen,
- Zeitliche Qualität.

Eine der ersten Normen zur Qualität von raumbezogenen Daten war der durch das amerikanische Institut für Standardisierung und Technologie (NIST) veröffentlichte *Federal Information Processing Standard FIPS 173* (MORRISON 1995, VEREGIN 1999). Das in FIPS 173 vorgestellte Qualitätsmodell definiert sämtliche bei VAN OORT (2005) genannten Qualitätsmerkmale, bis auf die *Verwendung*.

Nach VEREGIN (1999) sind die zur Erstellung einer räumlichen Datenbank (Karte) auf die einzelnen Objekte angewendeten Auswahlkriterien, Definitionen und Zuordnungsvorschriften sehr wichtig zur

späteren Beurteilung ihrer Qualität. Beispielsweise können zur Beurteilung der Vollständigkeit nur solche Objekte herangezogen werden, die entsprechend der Datenbankspezifikation auch erfasst wurden. Die Qualität von Datenbanken mit Raumbezug wird durch die vier Merkmale *Genauigkeit*, *Auflösung*, *Konsistenz* und *Vollständigkeit* beschrieben. Diese werden dann noch weiter entsprechend ihrer räumlichen, thematischen, topologischen oder zeitlichen Ausprägung in Teilmerkmale untergliedert.

JOOS (2000) beschreibt ein Qualitätsmodell für „objektstrukturierte Geodaten“. Ziel ist die „Festlegung von Kriterien und Methoden zur vollständigen Beschreibung der Übereinstimmung zwischen Daten und ihrem Pendant in der realen Welt“. Es wird dabei ein Satz von vier Qualitätsmerkmalen, ähnlich dem von Veregin, definiert. Anstelle der *Auflösung* verwendet Joos allerdings das Qualitätsmerkmal der *Korrektheit*.

Die konkrete Bewertung von Daten anhand der definierten Qualitätsmerkmale geschieht durch Aufstellen von Qualitätsparametern. Diese können entweder für ein einzelnes Objekt, einen einzelnen Attributwert eines Objektes oder eine Sammlung von Objekten (Datensatz) definiert werden. Dabei sind die Qualitätsparameter für die Merkmale Vollständigkeit, Richtigkeit und Konsistenz vom Typ Boole'sch. Sie können entweder wahr oder unwahr sein. So kann beispielsweise zur Berechnung der Vollständigkeit ein Objekt in einem Datensatz entweder verfügbar oder nicht verfügbar sein. Für die Genauigkeit empfiehlt Joos die Einführung eines Grenzwertes für die geometrische Unsicherheit eines Objektes, anhand dessen eine Verletzung der Genauigkeitsforderung festgestellt werden kann.

Daneben gibt Joos auch globale Qualitätsparameter für die Bewertung ganzer Datensätze an. Diese werden jeweils als Rate aus der Anzahl der Elemente, die die jeweilige Qualitätsanforderung erfüllen im Verhältnis zur Gesamtzahl der Elemente im Datensatz angegeben. Somit erfüllen bei einem globalen Parameterwert von 100 % alle Objekten des Datensatzes den vorgegebenen Grenzwert. Dies stellt eine absolute obere Schranke dar. Jeder auftretende Fehler bzw. jedes Datensatzelement, welches die Anforderung nicht erfüllt, führt zu einer Absenkung des Wertes. Dies qualifiziert zum direkten Vergleich der Qualität von Datensätzen.

DIN EN ISO 19157 (2014) definiert ein Modell zur Beschreibung der Qualität geographischer (raumbezogener) Daten. Sie löst damit die beiden Normen ISO 19113 (Geoinformation - Qualitätsgrundsätze) und ISO 19114 (Verfahren zur Ermittlung der Datenqualität) ab und fasst deren Inhalt in teilweise überarbeiteter Form zusammen. Ziel ist es, die Qualität geographischer Datensätze in einheitlicher Form beschreibbar und die Datensätze anhand ihrer Qualität vergleichbar zu machen. Die Norm definiert ein Modell mit Qualitätsmerkmalen, eine Liste mit Qualitätsparametern sowie Verfahren zur Ermittlung und Dokumentation der Datenqualität.

Es werden fünf (abstrakte) Qualitätsmerkmale definiert, die jeweils durch Teilmerkmale konkretisiert werden, siehe Tabelle 5.2. Allgemein wird darin die *Korrektheit* als Übereinstimmung mit der konzeptionellen Realität (auch: Gegenstandsbereich) definiert. Die *Genauigkeit* wird entsprechend ihrer allgemeinen Definition als das Maß der Übereinstimmung eines Ermittlungs- oder Messergebnisses mit dem Erwartungswert definiert und umfasst nur zufällige Abweichungen.

### 5.2.2 Qualitätsmodell für Verkehrsinformationen

Auf der Arbeit von JOOS (2000) aufbauend hat WILTSCHKO (2004) sein *Qualitätsmodell für Informationen* entwickelt. Es wurde später im Rahmen des EU-geförderten Forschungsprojektes *EuroRoadS* für die Anwendung in einer Austauschplattform für Straßendaten weiterentwickelt (WILTSCHKO und KAUFMANN 2005). Dem Qualitätsmodell liegt ein Satz von sechs inhärenten Qualitätsmerkmalen zugrunde, die im Folgenden kurz eingeführt werden. Die nachfolgenden Definitionen sind WILTSCHKO (2004) entnommen.

Tabelle 5.2: Qualitätsmodell nach DIN EN ISO 19157 (2014)

Q-Merkmale	Q-Teilmerkmale
Vollständigkeit	Übervollständigkeit Untervollständigkeit
Logische Konsistenz	Konzeptuelle Konsistenz Domänen Konsistenz Format Konsistenz Topologische Konsistenz
Geometrische Genauigkeit	Absolute (externe) Genauigkeit Relative (interne) Genauigkeit Lagegenauigkeit von Rasterdaten
Thematische Genauigkeit	Korrektheit der Klassifikation Korrektheit qualitativer Attribute Genauigkeit quantitativer Attributen
Zeitliche Genauigkeit	Genauigkeit der Zeitmessung Zeitliche Konsistenz Zeitliche Gültigkeit / Validität

**Verfügbarkeit:** Die Verfügbarkeit gibt das Ausmaß des Vorhandenseins der Information zu einem definierten Zeitpunkt an einem bestimmten Ort an. Dieses Merkmal gibt an, ob die gewünschte Information bzw. das Datenobjekt zum gewünschten Zeitpunkt und am gewünschten Ort verfügbar ist. So kann beispielsweise eine Verkehrslageinformation zu einem bestimmten Zeitpunkt in der Verkehrszentrale vorliegen. Wenn sie aber aufgrund einer gestörten Kommunikationsschnittstelle nicht an die Klientensysteme (z.B. Navigationssysteme) ausgeliefert werden kann, so ist die Verfügbarkeit verletzt.

**Aktualität:** Die Aktualität gibt das Ausmaß der Übereinstimmung der Information mit der sich zeitlich ändernden konzeptionellen Realität an. Durch die Aktualität wird also die zeitliche Änderung der Realität und damit das Auseinanderdriften der Realität und der zu einem diskreten Zeitpunkt erstellten Information, beschrieben. Die konzeptuelle Realität beinhaltet dabei nur die im Datenmodell enthaltenen Elemente und ist somit eine Untermenge der Realität.

**Vollständigkeit:** Die Vollständigkeit gibt das Ausmaß des Vorhandenseins sämtlicher zur Beschreibung der konzeptionellen Realität erforderlichen Informationen an. Die Vollständigkeit beschreibt, ob alle Elemente der konzeptuellen Realität mit allen ihren im Datenmodell definierten Attributen und Relationen im zu untersuchenden Datenbestand enthalten sind.

**Konsistenz:** Die Konsistenz gibt das Ausmaß der Übereinstimmung der Information mit dem Informationsmodell an. Mit Hilfe der Konsistenz kann geprüft werden, ob die Information der im Informationsmodell spezifizierten Form entspricht. Diese umfasst die im Datenbestand enthaltenen Objekte, Attribute und Relationen sowie deren Verknüpfung und die Einheit der Attribute.

**Korrektheit:** Die Korrektheit gibt das Ausmaß der Übereinstimmung der Information mit der konzeptionellen Realität bei vorausgesetzter Aktualität an. Die Korrektheit einer Information ist verletzt, wenn ein grober oder systematischer Fehler festgestellt wurde. Diese Entscheidung ist abhängig von

der geforderten Genauigkeit, welche durch die nachfolgend beschriebenen Genauigkeitsmerkmale beschrieben werden kann.

**Metrische Genauigkeit:** Die metrische Genauigkeit gibt den Zusammenhang zwischen dem gemessenen und dem wahren bzw. plausibelsten Wert an. Anders als die bisher genannten Merkmale, kann die metrische Genauigkeit nur auf kontinuierlich darstellbare Informationen und Daten angewendet werden. Diese Daten müssen messbar oder aus gemessenen Daten über einen funktionalen Zusammenhang berechenbar sein.

**Semantische Genauigkeit:** Die semantische Genauigkeit gibt den Diskretisierungsgrad der Objekt- und Sachdaten an. Aus der Definition der semantischen Genauigkeit folgt direkt, dass sie nur auf diskret darstellbare Daten und Informationen Anwendung findet. Sie gibt an, wie genau die Objekte, Attribute und Relationen der konzeptionellen Realität entsprechend dem gewählten Abstraktionsgrad der Modellierung erfasst werden können.

Diese Qualitätsmerkmale werden in drei Merkmalsgruppen eingeteilt, wie aus Abbildung 5.2 ersichtlich.

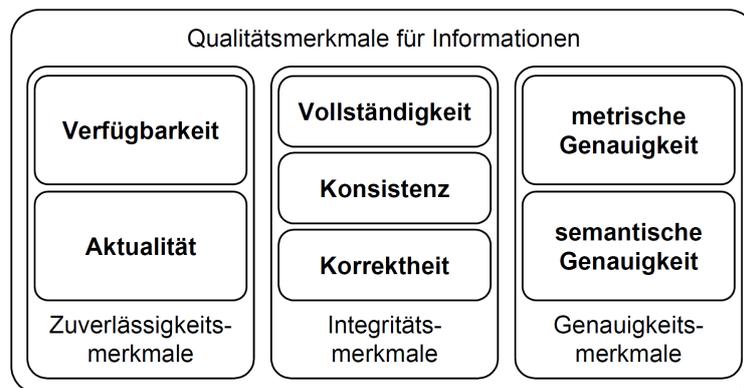


Abbildung 5.2: Qualitätsmerkmale für Informationen (aus WILTSCHKO (2004))

### 5.2.3 Zusammenfassung

Die in diesem Kapitel vorgestellten Qualitätsmodelle beschreiben die Qualität von geographischen Daten auf unterschiedliche Weise, entstammen sie doch verschiedener Gebiete der Geoinformatik und einiger ihrer Anwendungen im Verkehrsbereich.

Alle vorgestellten Modelle beschreiben die Qualität jedoch durch einen übergeordneten, unveränderlichen Satz von Qualitätsmerkmalen, die dann anwendungsspezifisch durch mess- oder zählbare Qualitätsparameter konkretisiert werden. Die Qualitätsmodelle aus der Geoinformatik sind sehr allgemein formuliert und bilden ein großes Spektrum an möglichen Anwendungen ab. Qualitätsmerkmale wie beispielsweise *Lineage* oder *Auflösung* werden in der Vektordatenverarbeitung eher seltener verwendet. Sie haben ihre Bedeutung hauptsächlich in kartographischen Anwendungen mit Raster- oder Bilddaten. Von großer Relevanz für die vorliegende Arbeit sind hingegen die Qualitätsmerkmale, die auch in dem vorgestellten Qualitätsmodell für Verkehrsinformationen Verwendung finden.

Aus dem so gewonnenen Portfolio an Modellierungsmöglichkeiten der Datenqualität können nun für das im Rahmen dieser Arbeit zu entwickelnde Qualitätsmodell für Kartenzuordnungen geeignete Ansätze ausgewählt und im weiteren Verlauf dann adaptiert und konkretisiert werden.

## 5.3 Qualitätsmodelle für Kartenzuordnungen

### 5.3.1 Einführung

Zunächst lässt sich feststellen, dass das im Rahmen der vorliegenden Arbeit qualitativ zu beschreibende Location Referencing Verfahren deutliche Unterschiede zu den Anwendungsszenarien der zuvor beschriebenen Qualitätsmodelle der allgemeinen Geoinformatik sowie der Verkehrsmodelle aufweist. Im Gegensatz zu den Qualitätsmodellen für geographische Daten (-banken) soll beim Location Referencing gerade nicht die Übereinstimmung einer Karte mit der (konzeptuellen) Realität bestimmt werden. Die Qualitätsmodelle der Geoinformatik beschreiben stets die Qualität eines Datensatzes bzw. einer Karte als Qualität der Erfassung bzw. der Weiterverarbeitung der erfassten Informationen bis zum fertigen Produkt. Bei der Ermittlung der Qualität von Location Referencing Übertragungen ist aber vielmehr die Ermittlung der Güte der Übertragung einer Information von einer Karte in eine andere Karte das Ziel. Es soll der Einfluss des Verarbeitungsschrittes auf das Produkt qualitativ bewertet werden und nicht das entstandene Endprodukt (in diesem Fall die Zielkarte nach erfolgter Decodierung). Die Realität kann also nicht als Vergleichsmaßstab für die Qualitätsbewertung herangezogen werden. Im Falle der LR-Bewertung muss deshalb das im Zielnetz decodierte LR-Ergebnis mit der Ausgangslocation im Quellnetz verglichen werden. Dazu ist die der Ausgangslocation entsprechende Ortsreferenz im Zielnetz mit übergeordneter Qualität zu definieren und vorhandene Kartenunterschiede zu berücksichtigen.

Daraus ergibt sich direkt, dass möglicherweise nicht alle der zuvor beschriebenen Qualitätsmerkmale bzw. Qualitätsmaße für eine Beschreibung der Qualität einer Location Referencing Übertragung tatsächlich Anwendung finden können.

Während solche vollständigen referenzierten Übertragungen beispielsweise im EU-geförderten Forschungsprojekt ROSATTE qualitativ beschrieben wurden (vgl. Abschnitt 5.3.4), konzentriert sich die vorliegende Arbeit mit dem Location Referencing nur auf einen Teilbereich dieser Übertragung, nämlich der Kartenzuordnung. Diese besteht lediglich aus der Generierung der Referenzierung in der Quellkarte und deren Umkehr beim Decodieren in der Zielkarte.

### 5.3.2 Relevante Qualitätsmerkmale

Die Qualität dieses Referenzierungs- und Decodierungs-Prozesses lässt sich am besten anhand der decodierten Ortsreferenzen ('Locations') im Zielnetz beurteilen. Diese umfasst also alle drei Teile des Location Referencing: Codierung, Übermittlung und Decodierung des Location Reference Codes im Zielnetz, siehe auch Kapitel 3.3. Eine separate qualitative Beschreibung der einzelnen Teile ist prinzipiell möglich. So ließe sich die Qualität bzw. der mögliche Qualitätsverlust während des Erzeugens der Ortsreferenz im Quellnetz beschreiben. Ebenso kann auch der Einfluss der physikalischen Übertragung des Location Codes in das Zielsystem bestimmt werden. Da diese Einflüsse der individuellen Umsetzung/Implementierung des Location Referencing Verfahrens zugerechnet werden müssen, erscheint ihre qualitative Beschreibung im Rahmen der vorliegenden Arbeit nicht sinnvoll. In dieser Arbeit soll ausschließlich der Einfluss des Location Referencing Verfahrens auf das Ergebnis qualitativ bewertet werden.

Aus diesen Gründen finden die Qualitätsmerkmale *Konsistenz*, *Verfügbarkeit* und *Aktualität* keine Anwendung. Die *Konsistenz* beschreibt die Übereinstimmung von Objekten mit den Festlegungen in ihrem Datenmodell. Übertragen auf das Location Referencing könnten Beeinträchtigungen der Konsistenz durch fehlerhafte Implementierung auf Encoder-Seite oder durch fehlerhafte Übertragungen entstehen. Da diese Teile jedoch hier nicht betrachtet werden sollen, wird auch die Konsistenz nicht weiter untersucht.

Die *Verfügbarkeit* eines Location Codes könnte durch Fehler in der Datenübermittlung beeinträchtigt werden. Die Aktualität der in einer mithilfe Location Referencing übermittelten Nutzdaten könnte dahingehend verletzt sein, dass die codierten Nutzdaten nicht mehr dem aktuellen Stand der Realität zum Zeitpunkt der Codierung entsprechen. Beide Szenarien sind jedoch für die vorliegende Arbeit nicht relevant und werden daher auch nicht weiter vertieft.

Für die qualitative Bewertung eines Location Referencing Verfahrens erscheinen jedoch die Merkmale *Vollständigkeit*, *Genauigkeit* und *Korrektheit* relevant. Diese Entscheidung soll im Folgenden diskutiert werden.

Die *Vollständigkeit* kann lokal für ein einzelnes Objekt (z.B. eine Ortsreferenz) oder auch global für einen ganzen Datensatz untersucht werden. Da eine Ortsreferenz mehrere Kanten abdecken kann, erscheint es sinnvoll zu prüfen, ob alle erforderlichen Kanten auch im decodierten Ergebnis enthalten sind. Gerade durch eine falsche Entscheidung bei der Zuordnung der Start- und Endknoten können dadurch zusätzliche Kanten in der decodierten Ortsreferenz enthalten sein (entspricht einer *Übervollständigkeit*). Andererseits können aber auch Kanten im decodierten Ergebnis fehlen (*Untervollständigkeit*). Für eine globale Vollständigkeitsbetrachtung könnte beispielsweise die Anzahl der im Zielnetz decodierten Ortsreferenzen ins Verhältnis zu ihrer Gesamtzahl gesetzt werden.

Da eine Ortsreferenz geometrisch durch einen Start- und einen Endpunkt definiert wird, kann deren Lage auch zur Qualitätsanalyse der geometrischen *Genauigkeit* herangezogen werden. Denkbar wäre hier, die (euklidischen) Distanzen zum jeweiligen Sollpunkt im Zielnetz zu berechnen. Ebenso kann die Gesamtlänge einer decodierten Ortsreferenz mit der Länge der ursprünglich codierten Ortsreferenz verglichen werden.

Für die *Korrektheit* wurden in den zuvor genannten Qualitätsmodellen unterschiedliche Definitionen gefunden. Von JOOS (2000) wurde sie als *Richtigkeit* bezeichnet, allerdings bezogen auf eine Zuordnung eines Objektes zu einer bestimmten Klasse. WILTSCHKO (2004) beschreibt die Korrektheit in Bezug auf das Auftreten von groben oder systematischen Fehlern. Diese können durch die Überprüfung von Genauigkeitsbedingungen festgestellt werden. Somit können Objekte (hier: Ortsreferenzen) als geometrisch korrekt definiert werden, wenn ihre geometrische Genauigkeit unterhalb eines zu definierenden Schwellwertes bleibt. Erweitert man den Begriff der Korrektheit, so könnte ebenso eine topologische Korrektheit angegeben werden, die auf die (topologische) Vollständigkeit einer Ortsreferenz prüft, also ob alle der erwarteten Kanten auch im decodierten Ergebnis gefunden werden können. In dem in Abbildung 5.3 dargestellten Beispiel ist dieser akzeptierte Genauigkeitsbereich als grau schattierter Bereich um die Straßenachsen des konzeptuellen Modells (also die Abbildung der Realität) dargestellt. Die in der digitalen Straßenkarte enthaltenen Straßenachsen sind in rot eingezeichnet. Fehlt nun eine Straße in der Karte, wie in Teil (I), so ist die Karte nicht vollständig. Übersteigt die geometrische Abweichung einer digitalisierten Straßenachse den akzeptierten Wert, wie in Teilbild (II) gezeigt, so ist diese Straße inkorrekt erfasst. Der Fall (III) zeigt nun eine Straße, deren geometrische Genauigkeit innerhalb des akzeptierten Bereichs liegt, sie wird daher als korrekt bezeichnet.

### 5.3.3 Übersicht existierende Ansätze

CHEN et al. (2008) und CHEN (2011) definieren ein Qualitätsmodell zur Evaluierung der Qualität von Verfahren zur Integration digitaler Straßenkarten für Navigationssysteme. Ausgehend von internationalen Standards (vgl. die inzwischen aktualisierte und umbenannte ISO 19157 (DIN EN ISO 19157 2014)) sowie den Qualitätsbeschreibungen in der Geoinformatik (JOOS 2000) wird die Qualitätsbeschreibung in eine *globale* für den gesamten Datensatz sowie in eine *lokale* für jedes Zuordnungspaar unterteilt.

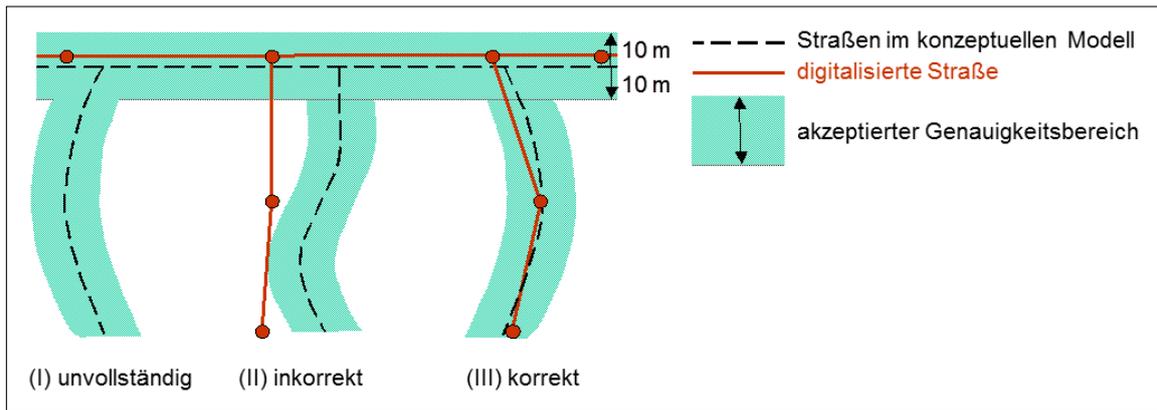


Abbildung 5.3: Korrektheit und Genauigkeit (nach WILTSCHEK und KAUFMANN (2004))

Für die globale Qualitätsauswertung werden die drei Ähnlichkeitsmaße *Geometrische Ähnlichkeit*, *Vollständigkeit* und *Topologische Ähnlichkeit* ausgewählt. Die Bestimmung der geometrischen und topologischen Ähnlichkeit basiert auf dem Vergleich der Adjazenzmatrizen der zu vergleichenden Datensätze. Damit diese Matrizen gleich dimensioniert sind, müssen in einer Vorverarbeitung komplexe, aus mehreren Teilen bestehende Objekte zu sog. Komplexknoten und Komplexkanten zusammengefasst werden. Für die geometrische Ähnlichkeit wird für jede Zelle der Matrizen die Entfernung der Knoten in den jeweiligen Karten bestimmt und daraus der Maximal- und Mittelwert für den gesamten Datensatz. Die topologische Ähnlichkeit wird durch Korrelation der Exzentrizitäten von Knoten bestimmt. Die Exzentrizität eines Knotens ist die maximale Distanz von kürzesten Wegen zu allen anderen Knoten im Netz. Die Vollständigkeit wird durch Vergleich der Gesamtlängen der Kanten in den jeweiligen Datensätzen berechnet.

Zur lokalen Bewertung der einzelnen Zuordnungspaare wird die Ähnlichkeit der Form durch Erkennung und Vergleich, der Geometrie durch die Hausdorff-Distanz, der Topologie durch Vergleich der Erreichbarkeit (Vektor mit kürzesten Wegen von allen anderen Knoten aus) sowie die Ähnlichkeit der Attribute durch Vergleich der Attributwerte ermittelt.

ZHANG (2009) (Kapitel 2.4.4) verwendet für die Qualitätsbeschreibung der Zuordnungsergebnisse die für die Zuordnung berechneten Ähnlichkeitsmaße. Er nutzt dazu zum einen ein Maß für die geometrische und topologische Ähnlichkeit der Anfangs- und Endpunkte einer Kante sowie ein Maß für die semantischen Unterschiede zwischen den beiden Kanten im Hinblick auf Name, Breite, Anzahl der Fahrspuren, durchschnittliche Fahrtzeit, Fahrtrichtung und andere. Dieses Qualitätsmaß wird dann noch in Abhängigkeit von der Anzahl der Zuordnungs-Kandidaten reduziert, da nach Ansicht des Autors die Qualität der Zuordnung mit zunehmender Zahl von wahrscheinlichen Zuordnungskandidaten leidet.

WALTER (1997) (Kapitel 2.4.1) beschreibt die Qualität von Kartenzuordnungen als 'Sicherheit' einer einzelnen oder der gesamten Zuordnung zweier Datensätze. Dazu betrachtet er die Attribute und Relationen der Elemente der Zuordnungen. Abhängig von der jeweiligen Anwendung kann dann definiert werden, ab wann eine Zuordnung nicht mehr als sicher einzustufen ist. Dies können z.B. Abstandsüberschreitungen als geometrische Bedingung oder der Vergleich von topologischen Eigenschaften sein. Die Qualitätsmaße werden aus den während des Zuordnungsprozesses berechneten Ähnlichkeitsmaßen abgeleitet. Zur Bewertung der Qualität des einzelnen Zuordnungspaares wird die bedingte Information genutzt, sie beschreibt den Informationsverlust pro Übertragung (Zuordnung) und ist somit ein Maß für die Ähnlichkeit der Zuordnungspartner. Ebenfalls kommt hier die gegenseitige Information zur Anwendung und gibt an, mit welcher Wahrscheinlichkeit es sich bei der Zuordnung um eine Fehlzugewandlung handelt. Als Qualitätsmaß für die Gesamtzuordnung zwischen zwei Datensätzen wird die durchschnittliche bedingte Information aus allen Zuordnungen dieses Datensatzes angegeben.

Alle verwendeten Qualitätsmaße stammen also direkt aus der Ähnlichkeitsberechnung des Zuordnungsprozesses und können ohne weitere Erhebungen zur Bestimmung der Güte einer einzelnen Zuordnung aber auch der Qualität des Gesamtdatensatzes verwendet werden.

### 5.3.4 Qualitätsmodell im Projekt ROSATTE

Im EU-geförderten Forschungsprojekt ROSATTE (vgl. Kapitel 3.2.1) wurde ein Qualitätsmodell zur Beschreibung und Beurteilung der Qualität der mittels einer Austauschinfrastruktur für Verkehrsdaten übermittelten Informationen entwickelt (SCHÜTZLE 2009, SCHÜTZLE und FRANK 2010, SCHÜTZLE et al. 2011). Grundlage dafür war die Arbeit von WILTSCHKO (2004) und WILTSCHKO und KAUFMANN (2005), siehe Kapitel 5.2.2, da deren Qualitätsmodell im Rahmen des Vorgängerprojektes *Euro-RoadS* für einen ähnlichen Anwendungsfall abgeleitet worden war und somit für eine Verwendung in ROSATTE prädestiniert erschien. Das Qualitätsmodell war eingebettet in ein umfassendes Qualitätsmanagement innerhalb des Projektes (SCHÜTZLE und SCHEIDER 2009). Die mit ihm abzudeckende Prozesskette umfasste neben der Referenzierung und Decodierung mittels (mehrerer) Location Referencing Verfahren auch die Beurteilung möglicher Qualitätsveränderungen der zu übertragenden Nutzdaten. Tabelle 5.3 zeigt die verwendeten Qualitätsmerkmale, Teilmerkmale und die zu ihrer Konkretisierung definierten globalen Qualitätsparameter.

Tabelle 5.3: Qualitätsmodell in ROSATTE (SCHÜTZLE und FRANK 2010)

Q-Merkmal	Q-Teilmerkmal	Q-Parameter
Verfügbarkeit	-	Service Verfügbarkeitsrate
Aktualität	-	Datenprozessierungszeit
Vollständigkeit	eines empfangenen Objekts eines empfangenen Datensatzes eines decodierten Objektes eines decodierten Datensatzes	Vollst. Rate empf. Obj. Vollst. Rate empf. DS Vollst. Rate dek. Obj. Vollst. Rate dek. DS
Korrektheit	Attributive Korrektheit Topologische Korrektheit	Attr. Korr. Rate Topol. Korr. Rate
Konsistenz	Attributkonsistenz eines empfangenen Objekts	Attr. Kons. Rate
Genauigkeit	Geometrische Genauigkeit	Mittlere geom. Abweichung des Start-/Endpunktes

Zur Qualitätsevaluierung standen neben den Quell- und Zielkarten auch die jeweils auf diese referenzierten Ausgangs- und decodierte Ortsreferenz zur Verfügung. Alle diese Daten konnten daher für eine bestmögliche Analyse mittels eines geographischen Informationssystems (GIS) überlagert und direkt verglichen werden (SCHÜTZLE und FRANK 2010). Für das Location Referencing waren dabei vor allem die Qualitätsparameter *Topologische Korrektheit* und *Geometrische Genauigkeit* relevant (SCHÜTZLE et al. 2012). Diese wurden im Rahmen der Evaluierung der Projektergebnisse im Besonderen betrachtet. Die Topologische Korrektheit der decodierten Ortsreferenz war dabei durch drei Bedingungen definiert:

1. Sie muss demjenigen Routenverlauf in der Zielkarte folgen, der mit dem Routenverlauf der Ausgangslocation in der Quellkarte korrespondiert.
2. Sie muss den selben Richtungsverlauf haben.

3. Ihre Start- und Endpunkte müssen jeweils auf der mit der Ausgangslocation in der Quellkarte korrespondierenden Kante in der Zielkarte zu liegen kommen.

Sind alle diese Bedingungen erfüllt, so wird die in der Zielkarte decodierte Ortsreferenz als topologisch korrekt betrachtet.

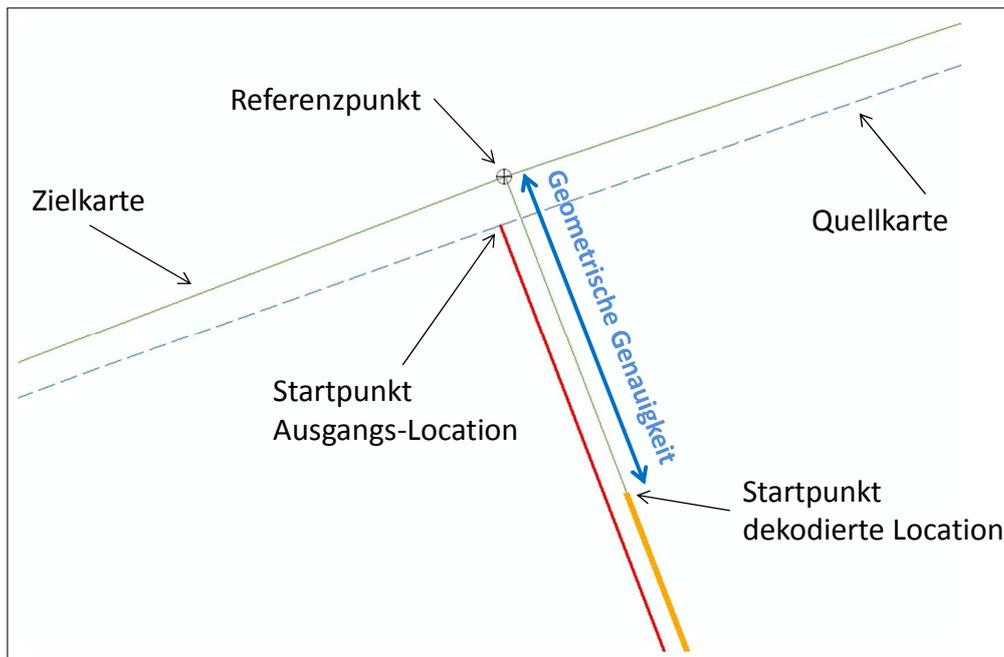


Abbildung 5.4: Geometrische Genauigkeit an Kreuzungen (SCHÜTZLE und FRANK (2010))

Zur Evaluierung der geometrischen Genauigkeit werden im Zielnetz sogenannte *Referenzpunkte* eingeführt. Diese stellen die korrekt decodierte Lösung im Zielnetz dar (=Soll). Für Start- oder Endpunkte einer Ortsreferenz, die in der Umgebung einer Kreuzung liegen, werden die Referenzpunkte direkt am oder mit der gleichen Entfernung (Offset) zum Kreuzungspunkt im Zielnetz ermittelt. Abbildung 5.4 zeigt hierfür ein Beispiel. Liegen die Punkte aber weit entfernt und ohne thematischen Bezug zur nächsten Kreuzung, so wird der Referenzpunkt durch Projektion (Ablotung) des Ortsreferenz Punktes auf die entsprechende Kante im Zielnetz erzeugt, siehe auch Abbildung 5.5.

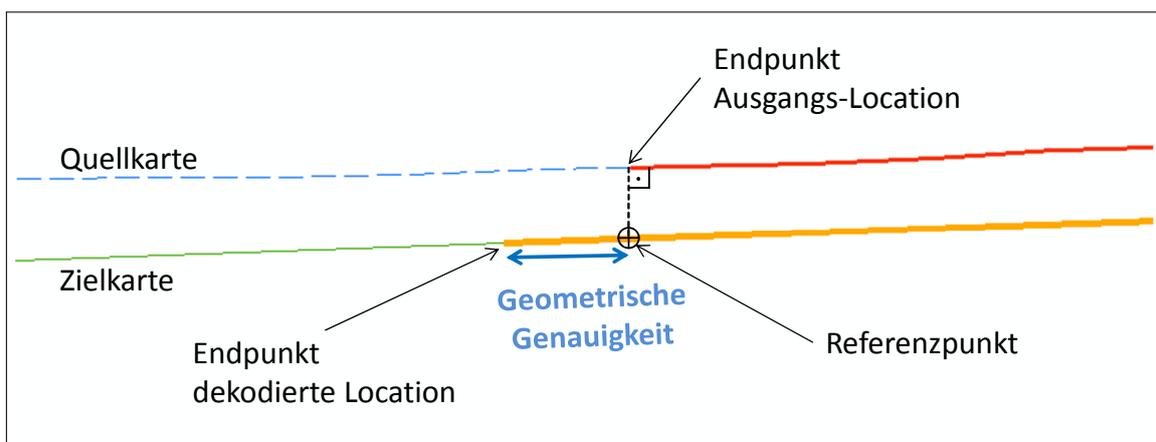


Abbildung 5.5: Geometrische Genauigkeit abseits von Kreuzungen (SCHÜTZLE und FRANK (2010))

Zur Bewertung der Ergebnisse auf Datensatz-Ebene wurden *Güte-Kriterien* für vier gestaffelte Fahrerassistenz (ADAS) Anwendungsszenarien definiert. Beginnend mit einem Basis-Niveau folgten dann Szenarien für informierende, warnende oder schließlich Anwendungen, die selbstständig kontrollierend

eingreifen könnten. Diese sind in Tabelle 5.4 dargestellt. Dadurch ist es möglich, die Tauglichkeit eines Datensatzes für ein bestimmtes Anwendungs-Szenario auf Grundlage der auf Objektebene durchgeführten Evaluierung zu bestimmen.

Tabelle 5.4: ROSATTE Qualitäts-Anforderungen entsprechend vier Anwendungsszenarien (SCHÜTZLE und FRANK 2010)

Q-Merkmal	Basic	Information	Warning	Control
Verfügbarkeit	> 80%	> 90%	> 95%	> 99%
Aktualität	3 Monate	1 Monat	1 Tag	1 Stunde
Vollständigkeit	> 80%	> 90%	> 95%	> 99%
Korrektheit	> 80%	> 90%	> 95%	> 99%
Geom. Genauigkeit	50m	20m	10m	5m

### 5.3.5 Klassifikation von Kartenzuordnungen

Zur Beurteilung der Ergebnisse einer Zuordnung wird jedes einzelne Zuordnungspaar mit Hilfe eines Klassifikators einer Ergebnisklasse zugeteilt. Für die Zuordnung selbst kann unterschieden werden, ob ein korrespondierendes Kartenelement identifiziert werden konnte, oder nicht. Betrachtet man die Situation in der Zielkarte, so ist zu unterscheiden, ob das gesuchte Element überhaupt existiert oder nicht. Daraus ergeben sich vier Kombinationsmöglichkeiten der Zuordnung: *richtig positiv*: Ortsreferenz im Zielnetz existiert und wurde zugeordnet; *falsch negativ*: korrespondierende Ortsreferenz im Zielnetz existiert nicht, es wurde aber dennoch eine Zuordnung vorgenommen; *falsch positiv*: Ortsreferenz im Zielnetz existiert, konnte aber nicht gefunden und zugeordnet werden; *richtig negativ*: korrespondierende Ortsreferenz im Zielnetz existiert nicht und es wurde richtigerweise auch keine Zuordnung vorgenommen.

Diese Klassifikation der Zuordnungsentscheidungen ist vergleichbar mit den bei Testentscheidungen in der Statistik auftretenden Möglichkeiten. Ein statistischer Test ist ein Verfahren, das für jede Stichprobe eine Entscheidung trifft, ob das Stichprobenergebnis eine zuvor aufgestellte Hypothese stützt oder nicht (SACHS und HEDDERICH 2006). Bezogen auf die Kartenzuordnung könnte eine Nullhypothese  $H_0$  als 'Ortsreferenz wurde korrekt zugeordnet' aufgestellt werden. Da eine Nullhypothese nie direkt bestätigt werden kann, wird als Gegenhypothese die sogenannte Alternativhypothese  $H_A$  aufgestellt. Diese gilt es dann zu widerlegen. Für eine Kartenzuordnung könnte  $H_A$  dann als 'Ortsreferenz wurde nicht korrekt zugeordnet' formuliert werden. Existiert eine Zuordnungslösung im Zielnetz und diese wurde auch korrekt gewählt, so kann  $H_A$  verworfen werden, die Zuordnung war korrekt. Wird jedoch für den selben Fall keine oder eine falsche Zuordnung vorgenommen, so ist ein Fehler 1. Art aufgetreten:  $H_0$  wurde unberechtigt abgelehnt. Existiert hingegen keine Zuordnungslösung im Zielnetz, es wurde aber dennoch eine Zuordnung vorgenommen, so ist ein Fehler 2. Art aufgetreten:  $H_0$  wurde unberechtigt beibehalten. Existiert im Zielnetz keine Zuordnungslösung und es wurde auch keine Zuordnung vorgenommen, so ist dies eine korrekte Entscheidung und die Nullhypothese kann angenommen werden, siehe auch Tabelle 5.5.

Ein solcher Klassifizierungsansatz wurde auch im Bereich des *Information Retrieval* verwendet und dann verschiedentlich zur Bewertung von Kartenzuordnungen herangezogen (SAMAL et al. 2004, BERRI et al. 2004, 2005, SAFRA et al. 2006, MUSTIÈRE und DEVOGELE 2008, SAFRA et al. 2010, HACKELÖER et al. 2015). Beim *Information Retrieval* geht es um die Klassifizierung von Suchmaschinen Ergebnissen. Dabei ist oft die Frage zu beantworten, ob zu einer bestimmten Suchanfrage ein bestimmtes Ergebnis gefunden wurde und ob überhaupt eine Internetseite mit dem gesuchten Inhalt existiert.

Übertragen auf die Zuordnung von Kartenelementen bzw. das Location Referencing wird jedoch eine Schwäche dieses Ansatzes offensichtlich: es wird bisher nur in zwei Klassen *zugeordnet* oder *nicht zugeordnet* unterschieden. Dadurch kann der Fall einer erfolgreichen aber fehlerhaften Zuordnung (z.B. aufgrund geometrischer oder topologischer Abweichungen am Anfangs- oder Endpunkt) nicht erfasst werden. Abhilfe schafft der Ansatz von ZHANG (2009), der unter Berufung auf COBB et al. (1998) und SAALFELD (1988) eine detaillierte Unterscheidung zwischen korrekt und falsch zugeordnet einführt und somit zwischen fünf auftretenden Kategorien unterscheidet:

**Korrekte Zuordnung** ( $r_p^+$ ) - eine Ortsreferenz aus der Quellkarte wurde mit ausreichender Qualität ihrem korrespondierenden Element im Zieldatensatz zugeordnet;

**Inkorrekte Zuordnung** ( $r_p^-$ ) - eine Ortsreferenz aus der Quellkarte wurde einem Element der Zielkarte zugeordnet. Die Zuordnung erfüllt jedoch nicht die an sie gestellten Qualitätsanforderungen. Es liegt also eine inkorrekte Zuordnung vor;

**Falsche Zuordnung** ( $f_p$ ) - eine Ortsreferenz aus der Quellkarte wird fälschlicherweise einem Objekt im Zieldatensatz zugeordnet, obwohl ein ihm korrespondierendes Element in der Zielkarte überhaupt nicht existiert;

**Korrekte Nicht-Zuordnung** ( $r_n$ ) - für eine Ortsreferenz in der Quellkarte wird richtigerweise erkannt, dass kein korrespondierendes Element im Zieldatensatz existiert und somit auch keine Zuordnung vorgenommen;

**Falsche Nicht-Zuordnung** ( $f_n$ ) - für eine Location in der Quellkarte wird fälschlicherweise erkannt, dass kein korrespondierendes Element in der Zielkarte existiert, obwohl es tatsächlich doch existiert.

Neu ist dabei die Aufteilung des *richtig positiven* ( $r_p$ ) Anteils in zwei Teile für eine richtige (genaue) Zuordnung ( $r_p^+$ ) und für den Fall, dass zwar eine Zuordnung erfolgte, diese jedoch inkorrekt ist oder ein falsches Element ausgewählt wurde ( $r_p^-$ ).

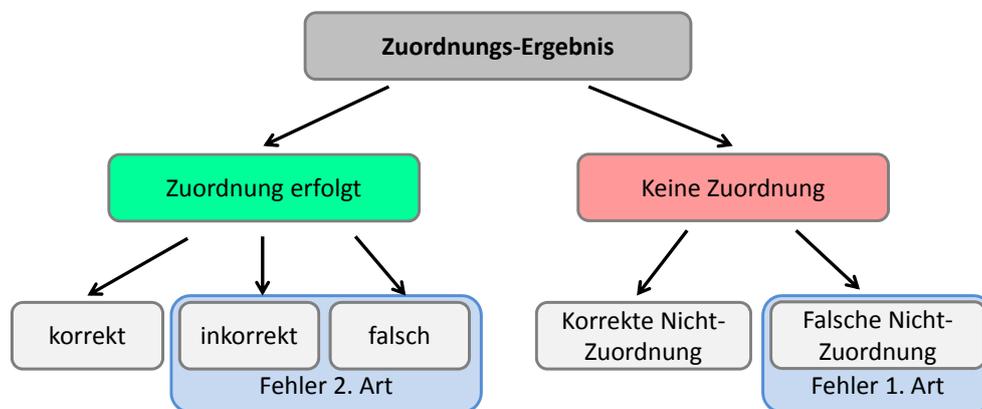


Abbildung 5.6: Klassifizierung von Zuordnungsergebnissen (nach ZHANG (2009))

Zur Auswertung werden die Auftretungshäufigkeiten der oben genannten Fälle für einen Satz von Zuordnungen bestimmt. Eine gebräuchliche Form der Ergebnisdarstellung ist die sogenannte Wahrheitsmatrix, die auch Konfusionsmatrix genannt wird (HACKELÖER et al. 2015). Eine Konfusionsmatrix ist die einfache Form einer Kreuztabelle und enthält die Häufigkeiten von Kombinationen bestimmter Merkmalsausprägungen. In den Zeilen stehen dabei die tatsächlichen Zustände (in der Zielkarte existiert entweder eine entsprechende Ortsreferenz oder nicht). Die Spalten der Matrix enthalten die Urteile des Klassifikators (eine Ortsreferenz aus der Quellkarte wurde einer Ortsreferenz in der Zielkarte zugeordnet oder nicht). Dies ist allerdings eine quadratische  $2 \times 2$  Matrix. Um die erweiterte Klassifikation dennoch in der Konfusionsmatrix darstellen zu können, wird die *falsche Zuordnung* der *falsch negativen Zuordnung* zugeordnet, siehe Tabelle 5.5.

Tabelle 5.5: Wahrheitsmatrix / Konfusionsmatrix

Ortsreferenz im Zielnetz	existiert	existiert nicht
wurde zugeordnet	richtig positiv ( $r_p^+$ )	falsch positiv ( $f_p$ ) [Fehler 2. Art]
wurde nicht / falsch zugeordnet	falsch negativ ( $f_n$ ) + falsch zugeordnet ( $r_p^-$ ) [Fehler 1. Art]	richtig negativ ( $r_n$ )

### 5.3.6 Beurteilung von Kartenzuordnungen

Aus den in der Wahrheitsmatrix eingetragenen Häufigkeiten des Auftretens der einzelnen Klassen können nun verschiedene Kenngrößen zur Beurteilung des Klassifikators durch Berechnung von relativen Häufigkeiten abgeleitet werden (POWERS 2007).

#### 5.3.6.1 Precision und Recall

Neben dem eigentlichen Klassifizierungsschema stammt auch der sog. *Precision / Recall*-Ansatz zur Beurteilung des Klassifikators aus dem Information Retrieval (MUSTIÈRE und DEVOGELE 2008). Bei der Bewertung der empirischen Ergebnisse von Kartenzuordnungen findet dieser Ansatz auch sehr häufig Anwendung (SAMAL et al. 2004, BEERI et al. 2004, 2005, OLTEANU 2007b, SAFRA et al. 2006, 2010, OLTEANU-RAIMOND et al. 2015, HACKELÖER et al. 2015). Der von ZHANG (2009) und ZHANG et al. (2011) für die Unterscheidung nach korrekter und inkorrektter Zuordnung adaptierte Precision/Recall Ansatz in leicht veränderter Definition wird im Folgenden dargestellt und kurz erläutert.

Der Begriff *Recall*, auch als Richtig-Positiv-Rate bezeichnet, beschreibt den Anteil der zugeordneten Ortsreferenzen im Verhältnis zur Gesamtheit aller existierender Zuordnungspaare und damit die Vollständigkeit der Zuordnung (OLTEANU-RAIMOND et al. 2015). Im Sinne des im Rahmen dieser Arbeit zur Anwendung kommenden Qualitätsmodells kann dieser Parameter auch als *Vollständigkeitsrate* bezeichnet werden:

$$\text{Recall} = P(\text{zugeordnet} | \text{existiert}) = \frac{r_p^+ + r_p^-}{r_p^+ + r_p^- + f_n} \times 100 \%. \quad (5.1)$$

Unter *Precision*, auch Wirksamkeit oder positiver Vorhersagewert genannt, versteht man in diesem Fall den Anteil der korrekt zugeordneten Ortsreferenzen im Verhältnis zur Gesamtheit aller zugeordneter Ortsreferenzen an (OLTEANU-RAIMOND et al. 2015). Dieser Qualitätsparameter gibt somit innerhalb des Qualitätsmodells die *Korrektzuordnungsrate* an:

$$\text{Precision} = P(\text{korrekt zugeordnet} | \text{zugeordnet}) = \frac{r_p^+}{r_p^+ + r_p^- + f_p} \times 100 \%. \quad (5.2)$$

Der Begriff *Precision* wird dabei oft unterschiedlich oder unpräzise definiert. In MUSTIÈRE und DEVOGELE (2008) wird die Precision ähnlich der Korrekturklassifikationsrate (siehe Abschnitt 5.3.6.2) definiert.

Seltener wird auch die *Falsch-Negativ-Rate* angegeben, sie beschreibt den Anteil der nicht zugeordneten Ortsreferenzen im Verhältnis zur Gesamtheit aller existierender Zuordnungspaare und stellt im Qualitätsmodell eine *Fehlerrate* bzw. Inkorrektheitsrate dar:

$$\text{FNR} = P(\text{nicht zugeordnet} | \text{existiert}) = \frac{f_n}{r_p^+ + r_p^- + f_n}. \quad (5.3)$$

Übertragen auf das Problemfeld des Location Referencing scheint es so, dass die Verwendung von *Precision* und *Recall* nicht ausreicht, um den Erfolg des Zuordnungsverfahrens bzw. des Location Referencing Decoders zu bewerten. Es fällt auf, dass insbesondere die *Precision* nur die korrekten positiven Zuordnungen als Erfolg wertet. Wird jedoch bei einem nicht existierenden Gegenstück in der Zielkarte auch keine Zuordnung durch den Decoder vorgenommen (richtig negative Zuordnung), so wird dieses im Prinzip positive Ergebnis nicht berücksichtigt.

### 5.3.6.2 Hit Rate

Zur Bewertung von Location Referencing Methoden besser geeignet ist die *Trefferquote* (engl. *Hit Rate*), auch Korrektorklassifikationsrate oder nach RAND (1971) *Rand Accuracy* bzw. *Rand Index* genannt. Sie beschreibt den Anteil aller korrekt klassifizierten (also der korrekt zugeordneten und der korrekt nicht zugeordneten) Ortsreferenzen im Verhältnis zur Gesamtheit aller zuzuordnenden Ortsreferenzen aus dem Quelldatensatz. Im verwendeten Qualitätsmodell kann dieser Wert als *Korrekttheitsrate* bezeichnet werden:

$$\text{Hit Rate} = P(\text{korrekt klassifiziert} | \text{klassifiziert}) = \frac{r_p^+ + r_n}{r_p^+ + r_p^- + f_p + r_n + f_n}. \quad (5.4)$$

Dieser Wert kann auch als Erfolgsquote des Decoders angesehen werden, da er in beiden berücksichtigten Fällen (korrekte Zuordnung  $r_p^+$  und korrekte Nicht-Zuordnung  $r_n$ ) der Decoder die richtige Entscheidung getroffen hat. Im Falle einer richtig erkannten Nicht-Zuordnung ist die Ursache in fehlenden Straßenobjekten (Kanten) in einem der Datensätze, letztlich also in Kartenfehlern bzw. -Unterschieden, und nicht im Versagen des Location Referencing Verfahrens zu suchen.

Nur in seltenen Fällen werden zur Berechnung der *Hit Rate* nur die richtig positiven Zuordnungen herangezogen (z.B. in DEMIR (2002)). Damit werden richtig negative Zuordnungen als Decoderfehler gewertet, was nach Auffassung des Autors nicht korrekt ist.

Die Angabe einer Trefferquote ist im Bereich des Location Referencing weit verbreitet. Allerdings wird sie in manchen Fällen nicht klar definiert. Ebenso verwenden nicht alle Veröffentlichungen eine einheitliche Definition der Trefferquote. Dennoch ist eine Angabe zu Vergleichszwecken und auch Gegenüberstellung zu den anderen in dieser Arbeit eingeführten Qualitätsmaßen hilfreich.

Im Rahmen der Entwicklung von AGORA und später AGORA-C wurde auch die *Hit Rate* zur Bewertung der Zuordnungsergebnisse verwendet (WEVERS und HENDRIKS 2006, WARTENBERG 2008). Es wird jedoch nicht genau spezifiziert, wann eine Ortsreferenz als 'korrekt' decodiert angesehen wird. HIESTERMANN (2008) lockert diese Definition etwas und lässt auch Ortsreferenzen als korrekt gelten, wenn sie nur einen Teil der ursprünglichen Ortsreferenz aus der Quellkarte abdecken. Dieser Bewertungsansatz ist sicher für manche Anwendungen praktikabel. Es sind aber ebenso Anwendungen denkbar, für die eine nur teilweise Decodierung einer Ortsreferenz in der Zielkarte als nicht ausreichend und damit als Misserfolg bewertet werden müssten.

### 5.3.6.3 Matching Rate

ERNST et al. (2014) definieren eine *matching rate* bezogen auf die Anzahl der überhaupt möglichen Zuordnungen. Sie wählen dazu ihre Testdatensätze so, dass alle zu testenden Ortsreferenzen auch in der Zielkarte existieren. Damit können sie den Erfolg ihres Zuordnungsverfahrens zwar relativ einfach bestimmen (denn alle Zuordnungsfehler sind dem Verfahren und nicht möglicher Kartenunterschiede zuzuordnen). Diese matching rate ist jedoch nicht mit einer unter realen Anwendungsbedingungen gewonnenen Trefferquote vergleichbar.

### 5.3.7 Zusammenfassung

In diesem Kapitel wurde zunächst ein Überblick allgemeiner Qualitätsgrundlagen sowie der in der Geoinformatik gebräuchlichen Qualitätsmodellierungen gegeben. Auf deren Grundlage wurden dann Qualitätsbeschreibungen und Qualitätsmaße existierender Zuordnungsverfahren vorgestellt und diskutiert. Bei diesen werden häufig neben einer globalen Vollständigkeitsbetrachtung auch geometrische, topologische und semantische Ähnlichkeitsbetrachtungen zur lokalen Bewertung der Zuordnungsgüte herangezogen. Offenbar wird aber nur selten ein Vergleich der Zuordnungsergebnisse mit dem tatsächlich zu erwartenden Ergebnis durchgeführt, da dies in der Regel die aufwändige Definition einer Soll-Lösung bedingt. Fast alle Verfahren nutzen die Informationen aus der Zuordnung selbst und vergleichen sie mit definierten Grenzwerten. Eine Überschreitung des Grenzwertes führt dann dazu, dass eine Zuordnung als nicht korrekt betrachtet wird.

Das im Rahmen des EU-geförderten Forschungsprojektes ROSATTE entwickelte Qualitätsmodell ist für die Anwendung der Qualitätsanalyse im Rahmen der vorliegenden Arbeit geeignet. Wie im vorigen Abschnitt gesehen, kann auf einen Teil der im ROSATTE-Qualitätsmodell enthaltenen Qualitätsmerkmale verzichtet werden, da hier durch die Konzentration auf das Location Referencing nur ein begrenzter Teil der Datenprozessierungskette untersucht wird. Aus dem gewählten Qualitätsmodell sind daher insbesondere die Qualitätsmerkmale *Vollständigkeit*, *Genauigkeit* und *Korrektheit* von Bedeutung. Dabei sind die in ROSATTE vorgeschlagenen und für die praktische Auswertung detailliert abgeleiteten Definitionen der lokalen Qualitätsparameter *geometrische Genauigkeit* und *topologische Korrektheit* auch für die vorliegende Arbeit anwendbar. Sie bilden die Basis für die globale Auswertung auf Datensatz-Ebene.

Bei Zuordnungen von Elementen aus verschiedenen Karten ist beim aktuellen Erfassungs- und Qualitätsstand der verfügbaren Kartensysteme davon auszugehen, dass ein entsprechendes Gegenstück in der empfangenden Karte auch existiert. Um aber auch dem unwahrscheinlichen Fall eines fehlenden Gegenstückes Rechnung zu tragen, erscheint die vorgestellte Klassifizierung der Zuordnungsergebnisse anhand einer Konfusionsmatrix als sinnvoll. Durch die vorgestellte Erweiterung dieser Klassifikation können als Ergebnis einer Qualitätsanalyse die Korrektheit der Zuordnungen sowie das Vorhandensein einer Lösung im Zielnetz berücksichtigt werden. Die vorgestellten, aus den Werten der Konfusionsmatrix abgeleiteten globalen Qualitätsparameter *Recall* und *Precision* erscheinen für die Anwendung im Rahmen der vorliegenden Arbeit als nicht unbedingt hilfreich. Die Angabe des *Recall*-Wertes zeigt lediglich die zugeordneten Objekte im Verhältnis zur Gesamtzahl der Objekte im Datensatz. Dazu ist keine tiefere qualitative Untersuchung des Zuordnungsergebnisses vorgesehen. Gerade diese ist jedoch eines der Auswertungs-Ziele der vorliegenden Arbeit. Entsprechend seiner Definition gibt der *Precision*-Wert nur den Anteil der korrekt zugeordneten Objekte im Verhältnis zu allen zugeordneten Objekten an. Die korrekt nicht-zugeordneten Objekte werden nicht berücksichtigt und somit eine wichtige Größe bei der Bewertung von Kartenzuordnungen nicht berücksichtigt. Auch wenn dieser Fall selten auftritt, sollte er dennoch bei der Definition der Qualitätsparameter berücksichtigt werden.

Für die Berechnung eines globalen Gesamtqualitätsmaßes kommen zwei Möglichkeiten in Betracht. Zum einen ist die auch unter der Bezeichnung *Hit Rate* bekannte Korrekturklassifikationsrate im Bereich des Location Referencing eine eingeführte Größe zur Beschreibung der globalen Performanz eines Location Referencing Systems. Sie berücksichtigt sowohl die korrekt zugeordneten als auch die korrekt nicht-zugeordneten Objekte und stellt somit ein Maß für die Korrektheit des Zuordnungsverfahrens insgesamt dar. Die für die Bestimmung der Hit Rate notwendige Erkennung von korrekten Zuordnungen ließe sich aus den Qualitätsparametern der im vorigen Abschnitt genannten Qualitätsmerkmale ableiten. Zusätzlich könnte in Anlehnung an die Berechnung von Leistungsfunktionen zur Ähnlichkeitsbestimmung (vgl. Kapitel 2.7.2) sowie die Qualitätsbeurteilung mittels Ähnlichkeitsmaßen (vgl. Abschnitt 5.3.3) ein globales Gesamtqualitätsmaß als gewichtete Summe der Einzelmaße angegeben werden.

Im Kapitel 6 werden die hier vorgestellten und diskutierten Grundlagen der Qualitätsmodellierung für die durchgeführten Location Referencing Übertragungen adaptiert und spezifische Qualitätsparameter entsprechend dem Qualitätsmodell für Verkehrsinformationen abgeleitet.



## 6 Evaluierungsgrundlagen zur Bewertung von Location Referencing Übertragungen

Das Qualitätsmodell bildet die Grundlage für die Überprüfung und Bewertung sowohl bestehender Location Referencing Verfahren als auch des im Rahmen dieser Arbeit neu entwickelten Verfahrens. Mit Hilfe des im vorangegangenen Kapitel 5 vorgestellten Qualitätsmodells für Verkehrsinformationen sollen Qualitätsparameter definiert werden, anhand derer die mittels LR übertragenen Ortsreferenzen nach dem Decodieren in der Zielkarte bewertet werden können. Je nach Anwendungsfall werden die einzelnen Qualitätsparameter gegeneinander gewichtet und bilden die unterschiedlichen Qualitätseigenschaften im Gesamtqualitätsmaß einer Ortsreferenz ab.

Nach der Definition der Vergleichsgrundlage für die Qualitätsbewertung (Referenzdaten) und einem Überblick über den sich bietenden Lösungsraum für Kartenzuordnungen kann das vorgenannte Qualitätsmodell zur Verwendung im Rahmen dieser Arbeit für die Bewertung von Location Referencing Übertragungen konkretisiert werden. Dazu werden alle Qualitätsmerkmale dahingehend überprüft, ob und wenn ja, welche Qualitätsparameter für ihre Konkretisierung definiert werden können.

Dazu werden zunächst lokale Qualitätsparameter auf *Objektebene* definiert (Kapitel 6.2). Bei der Bewertung jedes einzelnen Objektes (jeder einzelnen Ortsreferenz) geht es darum, herauszufinden, ob und inwieweit das aus der LR-Übertragung auf der Zielkarte decodierte Objekt (Ortsreferenz) mit dem zu erwartenden Objekt/Ortsreferenz (=Soll) übereinstimmt. Diese Einzelergebnisse werden dann global auf *Datensatzebene* zu einem Gesamtergebnis zusammengeführt (Kapitel 6.3).

### 6.1 Definition von SOLL-Ortsreferenzen

#### 6.1.1 Grundlagen

Wie bereits in Kapitel 5.3.1 angesprochen, kann zur qualitativen Bewertung von Kartenzuordnungen nicht die (konzeptuelle) Realität als Vergleich herangezogen werden, da die Qualität der Kartenzuordnung relativ zur Ausgangslocation im Quellnetz zu beschreiben ist. Die Qualität einer Kartenzuordnung ist vielmehr in Bezug auf die der Ausgangslocation aus dem Quellnetz im Zielnetz tatsächlich entsprechenden Ortsreferenz, also einer Soll- oder Referenzzuordnung, zu bewerten.

Die Ausgangslocation wird durch eine Kantenfolge im Quellnetz zunächst topologisch definiert. Die Start- und Endpunkte der Ortsreferenz müssen nicht unbedingt mit den Start- und Endknoten dieser Kantenfolge identisch sein. Über optional anzugebende Offsets können die Start- und Endpunkte der Ortsreferenz geometrisch von den Start- und am Endknoten der Kantenfolge abweichen. Somit ist eine OpenLR Ortsreferenz durch topologische und geometrische Bezüge zum jeweiligen Straßennetz definiert.

Zur Qualitätsbewertung müssen nun die korrespondierenden Ortsreferenzen im Zielnetz gefunden und in einem Referenzdatensatz zusammengestellt werden. Wichtig ist dabei vor allem die Identifikation derjenigen Knoten im Zielnetz, die den Start- und Endknoten sowie den weiteren Location Reference Points der Ortsreferenz *topologisch* entsprechen. Von diesen ausgehend können dann etwaige Offsets

geometrisch ins Zielnetz übertragen werden. Aufgrund der Kartenunterschiede ist die Identifikation der entsprechenden Start-/Endpunkte der Ortsreferenz im Zielnetz nicht immer eindeutig möglich.

### 6.1.2 Lösungsraum von Kartenzuordnungen

Aufgrund der in Kapitel 2.2.3 beschriebenen Kartenunterschiede gibt es bei Kartenzuordnungen nicht immer eine eindeutige Möglichkeit der Zuordnung. Abbildung 2.3 zeigt beispielhaft einige solcher Fälle mit nicht eindeutig möglicher Zuordnung. Neben den in Kapitel 2.3.2 genannten Kardinalitäten der Zuordnungen, bei denen eine Zuordnung stets möglich und auch eindeutig ist, lassen sich aber grundsätzlich noch drei weitere Möglichkeiten der Kartenzuordnung im Zielnetz unterscheiden:

1. eindeutige Lösung,
2. mehrdeutige Lösung,
3. keine Lösung.

Im Normalfall gibt es zu einer Ortsreferenz aus dem Quellnetz genau eine ihr entsprechende Kante bzw. genau eine Folge von ihr entsprechenden Kanten im Zielnetz. Haben neben dem Routen-Verlauf auch der Start- und Endpunkt der Ortsreferenz eine eindeutige Entsprechung im Zielnetz, so bedeutet dies, dass eine eindeutige Zuordnungslösung existiert, wie beispielhaft in Abbildung 6.1(a) dargestellt.

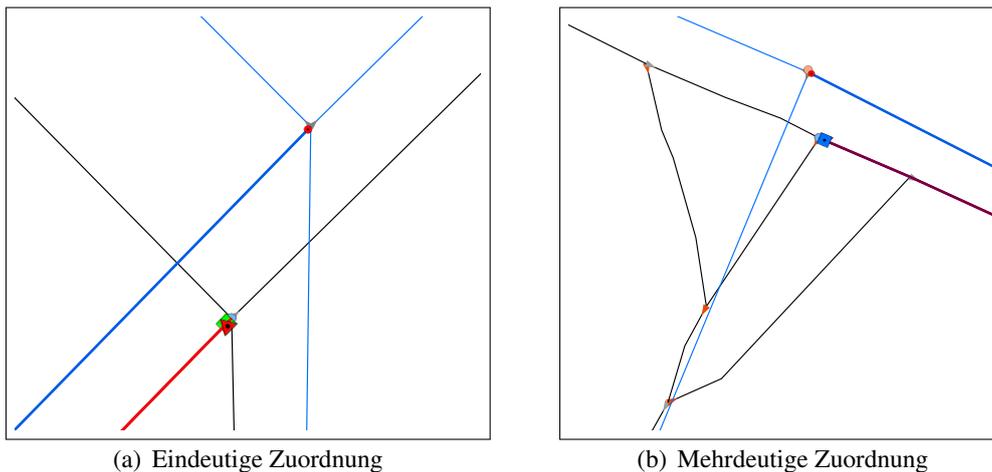


Abbildung 6.1: Lösungsraum für Zuordnungen

In manchen Fällen kann jedoch keine eindeutige Lösung gefunden werden. Meist existieren für die Start- und Endpunkte der Ortsreferenzen mehrere Lösungen im Zielnetz, die der originären Ortsreferenz aus dem Quellnetz entsprechen. Abbildung 6.1(b) zeigt ein Beispiel. Der in der Quellkarte [blau] als Startpunkt angehaltene Kreuzungspunkt (Knoten) wird in der Zielkarte [schwarz] durch drei Knoten repräsentiert. Ohne weitere Informationen zur genaueren Differenzierung müssen alle drei Knoten als korrekte Repräsentanten des Location Startpunktes im Zielnetz angesehen werden.

Fälle ohne existierende Lösung im Zielnetz sind fast immer durch Kartenunterschiede zwischen Quell- und Zielkarte oder Kartenfehler bedingt. Im Beispiel von Abbildung 6.2 fehlt im Zielnetz [blau] der Abzweig in Richtung Süd-Ost. Somit kann keine vollständige Entsprechung der Ortsreferenz im Zielnetz identifiziert werden. Da in diesem konkreten Beispiel jedoch ein Teilabschnitt der Ortsreferenz im Zielnetz vorhanden ist, kann dieser allenfalls als teilweise Lösung betrachtet werden.

Bei den im Rahmen der vorliegenden Arbeit durchgeführten Untersuchungen wurden sämtliche Ortsreferenzen in den verwendeten Datensätzen (vgl. Kapitel 7) auf ihren Lösungsraum untersucht und

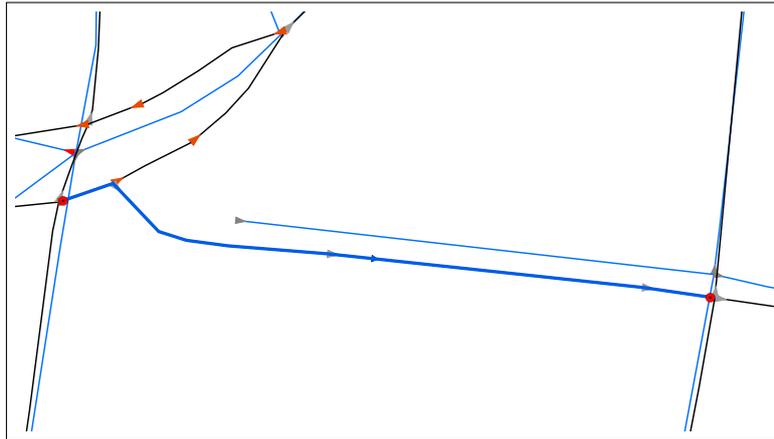


Abbildung 6.2: Beispiel für eine nicht mögliche Zuordnung

sichergestellt, dass bei mehrdeutigen Lösungen alle möglichen Lösungen als 'korrekt' gewertet werden. Ebenso werden bei nicht möglichen Zuordnungen entsprechende Entscheidungen des Decoders als 'korrekte' Entscheidung gewertet. Wird dennoch eine Zuordnung vorgenommen, so führt dies zur Wertung 'inkorrekt' (vgl. Kapitel 5.3.5).

### 6.1.3 Realisierung

Ähnlich wie bei der Qualitätsmodellierung für das Forschungsprojekt ROSATTE (Abschnitt 5.3.4), wird bei der Definition der Soll-Ortsreferenzen im Zielnetz auch für die vorliegende Arbeit zwischen zwei Arten der Definition von Ortsreferenzen unterschieden: (1) Locationpunkt mit Bezug zu einer Kreuzung, d.h. der Ortsreferenz Start- oder Endpunkt liegt direkt auf oder steht in direktem Bezug zu einem Netzknoten. (2) ohne Bezug zu einer Kreuzung, d.h. der Ortsreferenz Start- oder Endpunkt liegt auf einem bivalenten Knoten abseits einer Kreuzung oder ist auf freier Strecke mit Hilfe eines Offsets auf einer Kante eingerechnet. Als Erweiterung dazu sei noch der Fall eines Locationbeginns an bivalenten Knoten diskutiert, die einen Attributwechsel in der Karte beschreiben.

Für diese Szenarien gelten die im Folgenden beschriebenen Festlegungen und Definitionen.

#### 6.1.3.1 Locationpunkt mit Bezug zu einer Kreuzung

Die Lage der Ortsreferenzen sollte zur Kompensation der geometrischen Kartenunterschiede nicht absolut entsprechend den Koordinaten ihrer Anfangs-/Endpunkte betrachtet werden, sondern relativ zu den umliegenden Kartenelementen im Kartenkontext.

Liegt der Start- bzw. Endpunkt einer Ortsreferenz in unmittelbarer Nähe zu einer Kreuzung (echter Knoten mit Valenz  $> 2$ ) oder steht der Start-/Endpunkt in einem offensichtlichen Zusammenhang mit dem Kreuzungspunkt, so definiert dieser den Kartenkontext und wird als Referenzpunkt der Ortsreferenz betrachtet. In diesen Fällen ist die Distanz zum Kreuzungspunkt (Referenzpunkt) maßgeblich für die Bestimmung des Beginns bzw. Endes der Soll-Ortsreferenz.

Hierbei ist jedoch zu beachten, dass für die Location Referencing Übertragung nicht der tatsächliche Ortsreferenz Beginn/Ende maßgeblich ist, sondern der vom Encoder definierte Location Reference Point. Dieser ist innerorts häufig mit dem tatsächlichen Ortsreferenz Beginn identisch, kann aber gerade auf langgestreckten Straßen außerorts oder auf Autobahnen weit außerhalb der Ortsreferenz liegen und dadurch geometrische Differenzen verursachen.

Beim Decoding wird immer die 'entsprechende' Kante im Zielnetz gesucht. Diese 'Entsprechung' stützt sich vor allem auf die topologische Funktion der jeweiligen Kante im Netz. Dies kann beispielsweise ihre relative Lage vor oder hinter einer Kreuzung sein. Manchmal sind Kreuzungen in verschiedenen Karten unterschiedlich modelliert wie z.B. in Abbildung 6.1(b). Um solche Fälle aufzulösen, ist es notwendig zu wissen, auf welche ein- oder abgehende Kante (Straße) sich die Ortsreferenz genau bezieht. Nur dann kann auch in der Zielkarte auf diesen topologischen Ort referenziert werden.

### 6.1.3.2 Locationpunkt abseits von Kreuzungen aber mit Bezug zu Attributwechsel

Da digitale Straßenkarten, wie andere vektorbasierte Daten auch, häufig nur einen Attributwert für jedes Objektattribut speichern können, müssen zur Darstellung von Wechseln in diesen Attributwerten die entsprechende Kanten unterbrochen und durch zwei an einem bivalenten Knoten verbundene Kanten ersetzt werden. So kann beispielsweise ein Wechsel in der zulässigen Höchstgeschwindigkeit für eine Straße modelliert werden. Anstatt einer vorher durchgehenden Kante mit nur einem Wert für die Höchstgeschwindigkeit können die beiden sie ersetzenden Kanten jeweils einen anderen Wert aufnehmen.

In der Praxis kommt es häufig vor, dass Ortsreferenzen auf genau solche Attributwechsel in einer Karte referenziert sind. Um für diese die genaue Lage der Soll-Ortsreferenz im Zielnetz zu ermitteln, ist die Identifikation des entsprechenden Attributwechsels im Zielnetz erforderlich. Ist dies nicht möglich, da beispielsweise der angesprochene Geschwindigkeitswechsel in der Zielkarte nicht erfasst ist, so ist die Ortsreferenz anhand ihrer topologischen und geometrischen Eigenschaften ins Zielnetz zu verorten. Wurde der Attributwechsel im Zielnetz identifiziert, so beginnt bzw. endet die Soll-Ortsreferenz mit dem zuvor in der Quellkarte gemessenen Abstand zu diesem. Oftmals ist dieser Offset gleich Null.

### 6.1.3.3 Locationpunkt ohne Bezug zu einer Kreuzung

Für Ortsreferenz Start-/Endpunkte, die weit entfernt und ohne direkten Bezug zur nächsten Kreuzung entweder auf einem zweiwertigen Knoten oder auf einer Kante liegen, sind die zuvor genannten Regeln für die Definition der Soll-Ortsreferenzen im Zielnetz nicht optimal. Dies betrifft vor allem Ortsreferenzen auf langgestreckten Straßen, z.B. auf Autobahnen oder Landstraßen. In diesen Situationen weichen die in den Karten erfassten Kreuzungssituationen mitunter stark voneinander ab. Durch die hohen Fahrgeschwindigkeiten auf diesen Straßen haben bei der Erfassung mit Messfahrzeugen schon kurze zeitliche Auslöseverzögerungen einen großen Einfluss auf die geometrische Lage des erfassten Punktes. Zum anderen sind die baulichen Anlagen wie Auf- und Abfahrten oft sehr lang gestreckt. Weil Abzweige in den Karten jedoch durch einen einzelnen Knoten definiert werden müssen, kommt es zu Unterschieden in der Erfassung zwischen den verschiedenen Kartenanbietern.

Würde man wie zuvor den metrischen Abstand zum nächsten Kreuzungspunkt im Quellnetz messen und diesen im Zielnetz dann vom entsprechenden Kreuzungsknoten wieder abtragen, könnte aufgrund der großen Distanzen und der bestehenden Kartenunterschiede Soll-Punkte entstehen, die nicht der Ausgangslocation entsprechen. Es erscheint also in diesen Fällen besser, der geometrischen Punktlage eine höhere Bedeutung zuzumessen, als ihren topologischen Eigenschaften im Netz. Deshalb sollen im Rahmen der vorliegenden Arbeit, in solchen Situationen die Locationpunkte aus dem Quellnetz auf die entsprechende Kante im Zielnetz gelotet werden.

## 6.2 Qualitätsparameter auf Objektebene

Aus den in Kapitel 5.2 vorgestellten Qualitätsmodellen wurden wie in Kapitel 5.3.2 diskutiert, die Qualitätsmerkmale *Genauigkeit*, *Vollständigkeit* und *Korrektheit* für die Ermittlung der Qualität der Karten-

zuordnungen in dieser Arbeit ausgewählt. Für diese sollen im Folgenden konkrete Qualitätsparameter abgeleitet und definiert werden.

Dabei ist vor allem auf die engen Zusammenhang zwischen Genauigkeit und Korrektheit hinzuweisen. Nach ZHANG und SCHWIEGER (2011) und LAUFER (2011) ist bei der Konkretisierung der Korrektheit zwischen *messbaren* und *nicht-messbaren* Korrektheitsparametern zu unterscheiden. Nicht messbare Korrektheitsparameter können beispielsweise auf einer Experteneinschätzung beruhen, ob geforderte Anforderungen eingehalten werden oder nicht, etwa auf Basis einer visuellen Kontrolle. Aus Gründen der Objektivität soll jedoch im Rahmen der vorliegenden Arbeit auf die Verwendung von nicht messbaren Korrektheitsparametern verzichtet werden. Um zu einer Korrektheitsaussage vom Typ Boole'sch (ja/nein) für messbare Korrektheitsparameter zu kommen, können messbare Qualitätsparameter des Merkmals *Genauigkeit* mit einem Schwellwert auf Einhaltung der im Vorfeld festgelegten Anforderungen geprüft werden. Dadurch ist jedoch die Unabhängigkeit der entsprechenden Korrektheits- und Genauigkeitsparameter nicht mehr gegeben.

Wie in Kapitel 6.1 beschrieben, erfolgt die Qualitätsbewertung durch Vergleich mit zuvor und in übergeordneter Qualität bestimmten Referenzlösungen (Soll). Daher beziehen sich alle im Folgenden definierten Qualitätsparameter auf diese Soll-Ortsreferenzen im Zielnetz. Entsprechend der in Tabelle 5.1 vorgestellten Typen von Qualitätsparametern sind alle im Folgenden beschriebenen Genauigkeitsparameter vom Typ *Messwert* und werden dementsprechend durch einen Realwert repräsentiert. Sämtliche Parameter des Qualitätsmerkmals Korrektheit sind vom Typ *Indikator*, sie repräsentieren eine ja/nein Entscheidung und werden somit mit dem Datentyp Boole'sch dargestellt. Die Vollständigkeitsparameter sind vom Typ *Rate* und werden deshalb durch ganze Zahlen dargestellt.

## 6.2.1 Vollständigkeit

### 6.2.1.1 Anzahl der fehlenden Kanten (Untervollständigkeit)

Eine Ortsreferenz wird durch einen Kanten-Vektor  $\mathbf{L}_j = \{\mathbf{E}_{j,1}, \dots, \mathbf{E}_{j,n}\}$  beschrieben. Vergleicht man die beiden Vektoren dahingehend, welche Einträge (Kanten) nur in der SOLL-Ortsreferenz  $\mathbf{L}_j^T$ , nicht aber in der auf der Zielkarte decodierten Ortsreferenz  $\hat{\mathbf{L}}_j^T$  enthalten sind, so erhält man die Anzahl der in der decodierten Ortsreferenz fehlenden Kanten  $nE_j^-$ :

$$nE_j^- = |\mathbf{L}_j^T - \hat{\mathbf{L}}_j^T|. \quad (6.1)$$

### 6.2.1.2 Anzahl der überzähligen Kanten (Übervollständigkeit)

Umgekehrt erhält man die Anzahl der überzähligen Kanten  $nE_j^+$ , die zwar in der decodierten Ortsreferenz enthalten sind, jedoch nicht in der SOLL-Ortsreferenz:

$$nE_j^+ = |\hat{\mathbf{L}}_j^T - \mathbf{L}_j^T|. \quad (6.2)$$

## 6.2.2 Geometrische Genauigkeit

*Hinweis: Sämtliche Punktkoordinaten werden üblicherweise im Bereich der digitalen Straßenkarten, des Location Referencing und im Speziellen bei OpenLR in geographischer Länge und Breite bezogen auf das WGS-84 angegeben. Die Streckenberechnung der OpenLR-Implementierung erfolgt mittels WGS-84 Orthodrome und bestimmt somit die kürzeste Entfernung zwischen beiden Punkten auf dem WGS-84 Ellipsoid. Um die Übersichtlichkeit zu wahren wird im Folgenden jedoch die Streckenberechnung vereinfacht für den Fall ebener kartesischer Koordinaten angegeben.*

Die geometrische Genauigkeit in Bezug zu einer SOLL-Lösung entspricht der *absoluten (externen) Positions-Genauigkeit* aus ISO 19157 (DIN EN ISO 19157 2014), siehe auch Kapitel 5.2.1.

### 6.2.2.1 Distanz der Start- und Endpunkte

Die Distanz  $d_{j,s}$  zwischen dem Startpunkt einer Soll-Ortsreferenz  $\mathbf{L}_j$  im Zielnetz  $M^T$  und dem decodierten Startpunkt der entsprechenden Ortsreferenz  $\hat{\mathbf{L}}_j$  ist ein Maß für die geometrische Genauigkeit der decodierten Ortsreferenz. Sie wird als Euklidische Distanz aus den decodierten Punktkoordinaten  $\hat{x}^T, \hat{y}^T$  und den Soll-Punktkoordinaten  $x^T, y^T$  im Zielnetz wie folgt berechnet:

$$d_{j,s} = \sqrt{\left(x_{j,s}^T - \hat{x}_{j,s}^T\right)^2 + \left(y_{j,s}^T - \hat{y}_{j,s}^T\right)^2}. \quad (6.3)$$

Für die geometrische Genauigkeit der Endpunktzurordnung der Ortsreferenz  $\mathbf{L}^j$  kann entsprechend die Euklidische Distanz zwischen den beiden Endpunkten im Quell- und Zielnetz wie folgt definiert werden:

$$d_{j,e} = \sqrt{\left(x_{j,e}^T - \hat{x}_{j,e}^T\right)^2 + \left(y_{j,e}^T - \hat{y}_{j,e}^T\right)^2}. \quad (6.4)$$

### 6.2.2.2 Längendifferenz

Die Länge  $l_j$  einer Ortsreferenz  $\mathbf{L}_j$  ist definiert als die Summe der Längen der einzelnen Kanten  $\mathbf{E}_{j,k}$ :

$$l_j = \sum l(\mathbf{E}_{j,k}). \quad (6.5)$$

Die Längendifferenz der Soll-Ortsreferenz im Zielnetz  $\mathbf{L}_j^T$  und der entsprechenden im Zielnetz decodierten Location Reference im Zielnetz  $\hat{\mathbf{L}}_j^T$  ist ebenso ein Qualitätsparameter für die geometrische Genauigkeit:

$$\Delta l_j = |l_j^T - \hat{l}_j^T|. \quad (6.6)$$

### 6.2.3 Geometrische Korrektheit

Zur Bestimmung der geometrischen Korrektheit der Start- bzw. Endpunktzurordnung  $CG_j^S$  bzw.  $CG_j^E$  werden die Qualitätsmaße *Start-/Endpunktdistanz*  $d_{j,s}$  bzw.  $d_{j,e}$  mit ihrem Toleranzwert  $t_d$  verglichen. Bleiben die berechneten Qualitätsmaße innerhalb der Toleranz, wird die Start- bzw. Endpunktzurordnung als korrekt betrachtet. Das Überschreiten der Genauigkeits-Grenzwerte bedeutet die Verletzung der Geometrischen Korrektheit:

$$CG_j^S = \begin{cases} 1 & : d_{j,s} \leq t_d \\ 0 & : d_{j,s} > t_d \end{cases}, \quad (6.7)$$

$$CG_j^E = \begin{cases} 1 & : d_{j,e} \leq t_d \\ 0 & : d_{j,e} > t_d \end{cases}. \quad (6.8)$$

Der Toleranzwert (Schwellwert)  $t_d$  muss dynamisch festgelegt werden und die jeweiligen spezifischen Anforderungen an die geometrische Genauigkeit der Ortsreferenzen widerspiegeln. Bei den praktischen Untersuchungen hat sich gezeigt, dass Straßensegmente in Bereichen mit niedriger Höchstgeschwindigkeit (z.B. innerorts in Wohngebieten) kürzer und die Straßenmodellierung insgesamt kleinteiliger ist,

als dies in Bereichen mit höherer zulässiger Geschwindigkeit (z.B. außerorts oder auf Autobahnen) der Fall ist. Dies hat auch Auswirkungen auf die Genauigkeitsanforderungen der Ortsreferenzen. In Bereichen mit geringerer Geschwindigkeit sollten sie genauer übertragen werden, als in Bereichen in denen mit hoher Geschwindigkeit gefahren wird.

Im Rahmen dieser Arbeit wird daher der Wert  $t_d$  in Abhängigkeit der zulässigen Höchstgeschwindigkeit  $v_{max}$  der jeweiligen Kante festgelegt. Dabei sind die Höchstgeschwindigkeiten der Kanten der Ausgangs-Location im Quellnetz maßgebend. Wird eine Reaktionszeit des Fahrers von 1 Sekunde zugrunde gelegt, ergeben sich abhängig von der jeweiligen Fahrzeuggeschwindigkeit unterschiedlich lange Reaktionswege. Daraus lassen sich die in Tabelle 6.1 dargestellten und auf volle 5 Meter gerundete Schwellwerte für die geometrische Genauigkeit  $t_d$  für verschiedene, in Deutschland üblicherweise verwendete Geschwindigkeitsbereiche ableiten.

Tabelle 6.1: Geschwindigkeitsabhängige Toleranzwerte für geometrische Korrektheit

$v_{max}[km/h]$	$\leq 30$	$\leq 50$	$\leq 70$	$\leq 100$	$\leq 120$	$> 120$
$t_d[m]$	10	15	20	30	35	40

## 6.2.4 Topologische Korrektheit

### 6.2.4.1 Korrektheit der Start-/Endkanten

Für manche Anwendungen ist es neben der geometrischen Beschreibung der Zuordnungsqualität vor allem wichtig zu wissen, ob die Ortsreferenz nach der Location Referencing Übertragung auf denjenigen Kanten beginnt bzw. endet, die den Kanten der Ausgangs-Location im Quellnetz entsprechen. Diese 'Entsprechung' stützt sich vor allem auf die topologische Funktion der jeweiligen Kante im Netz (vgl. Kapitel 6.1).

Stimmen die Start- und Endkanten der decodierten Ortsreferenz  $\hat{\mathbf{L}}_j^T = \{\mathbf{E}_{j,1}^T, \dots, \mathbf{E}_{j,n}^T\}$  mit den Start- und Endkanten der SOLL-Ortsreferenz  $\mathbf{L}_j^T = \{\mathbf{E}_{j,1}^T, \dots, \mathbf{E}_{j,m}^T\}$  überein, so beginnt bzw. endet die decodierte Ortsreferenz auf der jeweils korrekten Kante. Die Korrektheit der Startkante  $CE_j^S$  bzw. der Endkante  $CE_j^E$  ist vom Typ Boole'sch und kann lediglich die Werte *wahr* oder *unwahr* annehmen, die Zuordnung der Start- bzw. Endkante kann somit entweder korrekt oder inkorrekt sein. Sie berechnet sich wie folgt:

$$CE_j^S = \begin{cases} 1 & : \mathbf{E}_{j,1}^T = \hat{\mathbf{E}}_{j,1}^T \\ 0 & : \mathbf{E}_{j,1}^T \neq \hat{\mathbf{E}}_{j,1}^T \end{cases}, \quad (6.9)$$

$$CE_j^E = \begin{cases} 1 & : \mathbf{E}_{j,n}^T = \hat{\mathbf{E}}_{j,m}^T \\ 0 & : \mathbf{E}_{j,n}^T \neq \hat{\mathbf{E}}_{j,m}^T \end{cases}. \quad (6.10)$$

### 6.2.4.2 Korrektheit des Ortsreferenz-Verlaufs

Die Kanten-Folge der decodierten Ortsreferenz muss mit der Kanten-Folge der Referenz-Location (SOLL) übereinstimmen. In diesem Fall verläuft die decodierte Ortsreferenz exakt entlang dem SOLL. Wenn die beiden vorausgegangenen Vollständigkeits-Prüfungen auf zusätzliche ( $nE_j^+$ ) oder fehlende ( $nE_j^-$ ) Kanten jeweils Null ergeben, so hat die decodierte Ortsreferenz  $\hat{\mathbf{L}}_j^T$  den selben Verlauf wie die

Referenz/Soll-Ortsreferenz im Zielnetz  $\mathbf{L}_j^T$ . Die Korrektheit des Location Verlaufs  $CR_j$  ist vom Typ Boole'sch und kann lediglich die Werte *wahr* oder *unwahr* annehmen. Sie berechnet sich wie folgt:

$$CR_j = \begin{cases} 1 & : nE_j^+ = 0 \wedge nE_j^- = 0 \\ 0 & : nE_j^+ > 0 \vee nE_j^- > 0. \end{cases} \quad (6.11)$$

## 6.2.5 Gesamtbetrachtung der Zuordnung

### 6.2.5.1 Gesamtqualitätsmaß

Die zuvor beschriebenen lokalen Qualitätsparameter können ähnlich wie in BEETZ et al. (2013) und ZHANG und SCHWIEGER (2011) bzw. SCHWIEGER et al. (2013) zu einem globalen Gesamtqualitätsmaß  $QM_j$  für eine Zuordnung zusammengefasst werden, wie in Gleichung 6.12 beschrieben. Allerdings wird im Rahmen der vorliegenden Arbeit abweichend nicht ein gewichteter Mittelwert, sondern die Bildung einer gewichteten Summe umgesetzt. Dabei wird das Qualitätsmerkmal *geometrische Genauigkeit* durch die Qualitätsparameter *Start-/Endpunktdistanz*  $d_{j,s}$  und  $d_{j,e}$  sowie der *Längendifferenz*  $\Delta l_j$  beschrieben. Da die im Rahmen dieser Arbeit verwendeten Qualitätsparameter des Qualitätsmerkmals *Korrektheit* alle durch Auswertung zuvor bestimmter Qualitätsparameter bestimmt werden, genügt an dieser Stelle die Einbeziehung der Qualitätsparameter der Merkmale *Vollständigkeit* und *Genauigkeit* in das Gesamtqualitätsmaß. Für eine detailliertere Betrachtung des Zusammenhangs zwischen Korrektheit und Genauigkeit sei an dieser Stelle nochmals auf Abschnitt 5.3.2 verwiesen. Die Korrektheitsparameter werden in einem nächsten Schritt zu einer Gesamtbetrachtung der Zuordnungskorrektheit zusammengeführt.

Implizit beschreiben die Parameter *Übervollständigkeit*  $nE_j^+$  und *Untervollständigkeit*  $nE_j^-$  neben dem Merkmal *Vollständigkeit* auch die *Korrektheit der Start-/Endkanten*  $CE_j^S$  bzw.  $CE_j^E$  sowie die *Korrektheit des Locationverlaufs*  $CL_j$ . Beginnt oder endet eine decodierte Ortsreferenz nämlich nicht auf der vorgesehenen Kante, so hat dies auch Einfluss auf die Über- oder Untervollständigkeit. Die Korrektheit des Locationverlaufs ist direkt abhängig und wird direkt aus der Über- bzw. Untervollständigkeit abgeleitet.

$$QM_j = w_{ds} \cdot \frac{d_{j,s}}{t_d} + w_{de} \cdot \frac{d_{j,e}}{t_d} + w_{\Delta l} \cdot \frac{\Delta l_j}{t_l} + w_{nE^+} \cdot \frac{nE_j^+}{t_{nE^+}} + w_{nE^-} \cdot \frac{nE_j^-}{t_{nE^-}} \quad (6.12)$$

Die einzelnen Maße müssen mit ihrem jeweiligen Toleranzwert / Schwellwert (engl. threshold)  $t_i$  normalisiert werden, um sie in Bezug auf ihre Einheit und Größenordnung vergleichbar zu machen. Zusätzlich können Gewichte  $w_i$  angebracht werden, um das Gesamtqualitätsmaß für verschiedene Anwendungen anpassen zu können. Dies ist vor allem hilfreich, da einige Anwendungen einen größeren Fokus auf die geometrische Qualität der decodierten Ortsreferenzen legen, für andere Anwendungen hingegen ist mehr die topologische Qualität wichtig. Bei der resultierenden Funktion handelt es sich um eine Kostenfunktion, d.h. je größer der Funktionswert, desto schlechter die Qualitätsbewertung der Zuordnung.

### 6.2.5.2 Korrektheit der Zuordnung

Die Korrektheit der Zuordnung  $CL_j$  leitet sich aus den einzelnen Teilbetrachtungen der Korrektheit ab. Dazu werden die *geometrische Korrektheit der Start-/Endpunkte*  $CG_j^S$  und  $CG_j^E$ , die *topologische Korrektheit der Start-/Endkanten*  $CE_j^S$  und  $CE_j^E$  sowie die Korrektheit des Location Verlaufs  $CR_j$  als Produkt ihrer Funktionswerte ausgewertet. Ist nur einer der Faktoren Null (also inkorrekt), so wird die Gesamtkorrektheit der Zuordnung auch als inkorrekt angenommen:

$$CL_j = CG_j^S \cdot CG_j^E \cdot CE_j^S \cdot CE_j^E \cdot CR_j, \quad CL_j \in \{0, 1\}. \quad (6.13)$$

### 6.2.5.3 Klassifizierung der Zuordnung

Die Beurteilung der Zuordnung kann mit Hilfe des Klassifizierungsschemas aus Kapitel 5.3.5 erfolgen. Aus den fünf möglichen Fällen (Klassen) kann die Unterscheidung als korrekte bzw. nicht korrekte Entscheidung des Location Referencing Decoders getroffen werden, siehe Tabelle 6.2. Die Entscheidung beruht darauf, ob der Decoder eine Lösung  $\hat{\mathbf{L}}_j^T$  im Zielnetz gefunden hat und ob eine korrekte Lösung  $\mathbf{L}_j^T$  überhaupt existiert. Ferner wird basierend auf den zuvor eingeführten Qualitätsparametern untersucht, ob die decodierte Lösung die Anforderungen an eine korrekte Zuordnung  $CL_j$  aus Kapitel 6.2.5.2 erfüllt.

Tabelle 6.2: Klassifizierung der Zuordnung

LR-Entscheidung	Zuordnung	Bedingung
<i>korrekt</i>	korrekte Zuordnung ( $r_p^+$ )	$\hat{\mathbf{L}}_j^T \neq \emptyset, \mathbf{L}_j^T \neq \emptyset, CL_j = 1$
	korr. nicht-Zuordnung ( $r_n$ )	$\hat{\mathbf{L}}_j^T = \emptyset, \mathbf{L}_j^T = \emptyset$
<i>nicht korrekt</i>	Falsche Zuordnung ( $r_p^-$ )	$\hat{\mathbf{L}}_j^T \neq \emptyset, \mathbf{L}_j^T \neq \emptyset, CL_j = 0$
	Falsch positive Zuordnung ( $f_p$ )	$\hat{\mathbf{L}}_j^T \neq \emptyset, \mathbf{L}_j^T = \emptyset$
	Falsch negative Zuordnung ( $f_n$ )	$\hat{\mathbf{L}}_j^T = \emptyset, \mathbf{L}_j^T \neq \emptyset$

## 6.3 Qualitätsparameter auf Datensatz-Ebene

Die im vorigen Kapitel 6.2 vorgestellten, lokalen Qualitätsparameter beziehen sich jeweils auf ein konkretes Zuordnungsobjekt (Ortsreferenz) aus einem Datensatz. Zur Beurteilung der Qualität des eingesetzten Zuordnungsverfahrens können diese Einzelbewertungen für eine Gruppe von Objekten oder den ganzen Datensatz zusammengefasst werden. Dadurch erhöht sich die Aussagefähigkeit. In diesem Abschnitt werden die im Rahmen dieser Arbeit abgeleiteten und bei den praktischen LR-Tests auch verwendeten Qualitätsparameter auf Datensatz-Ebene vorgestellt. Diese sind entweder auf die Anzahl der Ortsreferenzen im Ausgangsdatsatz  $N^S$  oder auf die Anzahl der tatsächlich im Zielnetz decodierten Ortsreferenzen  $N_{dec}^T$  normiert. Entsprechend der in Tabelle 5.1 dargestellten Typen von Qualitätsparametern, sind alle in diesem Kapitel vorgestellten globalen Genauigkeitsparameter vom Typ *Messwert*, wobei es sich konkret um aus Messwerten abgeleitete Mittelwerte handelt. Der globale Vollständigkeitsparameter sowie die globalen Korrektheitsparameter sind vom Typ *Rate*, da sie jeweils von ja/nein Entscheidungen repräsentierenden Indikatoren abgeleitet sind.

### 6.3.1 Vollständigkeit

Der Anzahl  $N_{dec}^T$  der im Zielnetz decodierten Ortsreferenzen  $\hat{\mathbf{L}}_j^T$  umfasst nach den Gleichungen 6.14 und 6.15 alle Ortsreferenzen, für die der Decoder *eine* Zuordnungslösung im Zielnetz gefunden hat. Dies beinhaltet noch keine Aussage über die Güte dieser Zuordnung, dazu müssen die Qualitätsparameter der Merkmale *Genauigkeit* und *Korrektheit* ausgewertet werden. Die Vollständigkeit der decodierten Zuordnungen  $Com_{dec}$  lässt sich in eine absolute, also im Verhältnis zur Anzahl der Ortsreferenzen im Ausgangsdatsatz  $N^S$  und eine relative, d.h. zur Gesamtzahl der im Zielnetz existierenden Ortsrefe-

renzen  $N_{exist}^T$ , unterscheiden. Dazu wird die Anzahl der decodierten Ortsreferenzen  $N_{dec}^T$  relativ zu  $N^S$  (Gleichung 6.16) bzw. zu  $N_{exist}^T$  (Gleichung 6.17) angeben:

$$LD_j = \begin{cases} 1 & : \hat{\mathbf{L}}_j^T \neq \emptyset \\ 0 & : \hat{\mathbf{L}}_j^T = \emptyset \end{cases}, \quad (6.14)$$

$$N_{dec}^T = \sum LD_j, \quad (6.15)$$

$$Com_{dec}^{abs} = \frac{N_{dec}^T}{N^S}, \quad (6.16)$$

$$Com_{dec}^{rel} = \frac{N_{dec}^T}{N_{exist}^T}. \quad (6.17)$$

### 6.3.2 Geometrische Genauigkeit

#### 6.3.2.1 Mittlere Abweichung der Start-/Endpunkte

Die Mittelwerte  $\bar{d}_s$  und  $\bar{d}_e$  der Euklidischen Distanzen  $d_{j,s}$  bzw.  $d_{j,e}$  zwischen den Anfangs- und Endpunkten der decodierten und der Soll-Ortsreferenz bilden einen Qualitätsparameter für die geometrische Genauigkeit der Start-/Endpunktzuordnung:

$$\bar{d}_s = \frac{\sum d_{j,s}}{N_{dec}^T}, \quad (6.18)$$

$$\bar{d}_e = \frac{\sum d_{j,e}}{N_{dec}^T}. \quad (6.19)$$

#### 6.3.2.2 Mittlere Abweichung der Ortsreferenz Längen

Der Mittelwert  $\bar{\Delta}l$  aus den Längendifferenzen  $\Delta l_j$  der decodierten Ortsreferenzen mit den Soll-Werten bildet einen weiteren Qualitätsparameter für die geometrische Genauigkeit des eingesetzten LR-Verfahrens:

$$\bar{\Delta}l = \frac{\sum \Delta l_j}{N_{dec}^T}. \quad (6.20)$$

### 6.3.3 Geometrische Korrektheit

Zur Beurteilung der geometrischen Zuordnungsqualität ist die Angabe hilfreich, welcher Anteil des Gesamtdatensatzes eine festgelegte Genauigkeitsanforderung nicht überschreitet. Dies kann beispielsweise für die zuvor ermittelten geometrischen Genauigkeitsparameter der Location Anfangs- und Endpunkte  $d_{j,s}$  bzw.  $d_{j,e}$  ermittelt werden. Zur Untersuchung beispielsweise systematischer Effekte im Decoder kann es hilfreich sein, die Auswertung getrennt für die Abstände am Beginn bzw. Ende der Ortsreferenzen durchzuführen. Dabei werden die Anzahl der geometrisch korrekten Startpunktzuordnung  $nCG_s$  und Endpunktzuordnung  $nCG_e$  bestimmt. Dazu dürfen die jeweiligen Abstandsmaße höchstens bis zur Akzeptanzschwelle  $t_d$  ansteigen. Für die Anzahl der geometrisch korrekten Zuordnungen  $nCG$  müssen die vorgenannten Teilanforderungen an beiden Enden der Ortsreferenz erfüllt sein. Zur Ermittlung der geometrischen Korrektheit  $CG$  wird dieser Anteil der korrekten Zuordnungen  $nCG$  ins Verhältnis zur Anzahl der im Ausgangsdatensatz enthaltenen Ortsreferenzen  $N^S$  gesetzt.

$$CG_j = \begin{cases} 1 & : (d_{j,s} \leq t_d) \wedge (d_{j,e} \leq t_d) \\ 0 & : (d_{j,s} > t_d) \vee (d_{j,e} > t_d) \end{cases} \quad (6.21)$$

$$CG = \frac{\sum CG_j}{N^S} \quad (6.22)$$

### 6.3.4 Topologische Korrektheit

#### 6.3.4.1 Anzahl/Anteil der decodierten Ortsreferenzen mit korrekter Start- und Endkante

Alle Ortsreferenzen, deren decodierte Startkante  $\hat{\mathbf{E}}_{j,1}^T$  der Soll-Startkante  $\mathbf{E}_{j,1}^T$  und deren decodierte Endkante  $\hat{\mathbf{E}}_{j,n}^T$  der Soll-Endkante  $\mathbf{E}_{j,m}^T$  entspricht, werden als *topologisch korrekt* bezeichnet. Für die Ermittlung des Qualitätsparameters *Anteil decodierter Ortsreferenzen mit korrekter Start- und Endkante* wird die Anzahl dieser topologisch korrekten Ortsreferenzen *topologischkorrekt* mit der Anzahl der im Ausgangsdatsatz enthaltenen Ortsreferenzen  $N^S$  normiert:

$$CE_j = \begin{cases} 1 & : \left( \hat{\mathbf{E}}_{j,1}^T = \mathbf{E}_{j,1}^T \right) \wedge \left( \hat{\mathbf{E}}_{j,n}^T = \mathbf{E}_{j,m}^T \right) \\ 0 & : \left( \hat{\mathbf{E}}_{j,1}^T \neq \mathbf{E}_{j,1}^T \right) \vee \left( \hat{\mathbf{E}}_{j,n}^T \neq \mathbf{E}_{j,m}^T \right) \end{cases}, \quad (6.23)$$

$$CE = \frac{\sum CE_j}{N^S}. \quad (6.24)$$

#### 6.3.4.2 Anzahl/Anteil der decodierten Ortsreferenzen mit korrektem Verlauf

Der für jede Zuordnung bestimmte lokale Qualitätsparameter *topologische Korrektheit des Location-Verlaufs*  $CR_j$  kann für die Gesamtbetrachtung des Datensatzes als Summe  $nCL$  aus allen Einzelzuordnungen bzw. als globaler Qualitätsparameter  $CR$  für den ganzen Datensatz relativ zur Gesamtzahl der Ausgangs-Locations im Quellnetz  $N^S$  angegeben werden:

$$nCR = \sum (CR_j), \quad (6.25)$$

$$CR = \frac{nCR}{N^S}. \quad (6.26)$$

### 6.3.5 Gesamtbetrachtung des Datensatzes

#### 6.3.5.1 Gesamtqualitätsmaß

Aus den Qualitätsmaßen der einzelnen Zuordnungen  $QM_j$  (vgl. Gleichung 6.12) kann für den gesamten Datensatz ein Gesamtqualitätsmaß  $QM$  als Mittelwert über der Anzahl der decodierten Zuordnungen  $N_{dec}^T$  formuliert werden:

$$QM = \frac{\sum QM_j}{N_{dec}^T}. \quad (6.27)$$

#### 6.3.5.2 Gesamtkorrektheit

Dieser Parameter entspricht der Klasse *richtig positiv* aus dem Klassifizierungsschema in Kapitel 5.3.5. Zunächst wird die Anzahl  $nC_{dec}$  aller Ortsreferenzen ermittelt, die die Anforderungen an eine korrekte Zuordnung  $CL_j$  aus Kapitel 6.2.5.2 erfüllen. Das Ergebnis in Relation zur Anzahl der Ausgangslocations  $N^S$  gesetzt, ergibt den Qualitätsparameter *Anteil der korrekt decodierten Ortsreferenzen*:

$$nC_{dec} = \sum (LC_j), \quad (6.28)$$

$$C_{dec} = \frac{nC_{dec}}{N^S}. \quad (6.29)$$

Dieser Qualitätsparameter umfasst die Anzahl  $nTN$  der Ortsreferenzen, für die keine Zuordnung im Zielnetz gefunden wurde, weil eine entsprechende Ortsreferenz im Zielnetz nicht existiert. Er entspricht

der Klasse *richtig negativ* aus dem in Kapitel 5.3.5 vorgestellten Klassifizierungsschema. Den Qualitätsparameter *Anteil der korrekt nicht zugeordneten Ortsreferenzen*  $TN$  erhält man durch Normierung auf die Anzahl aller Ausgangslocations im Quellnetz  $N^S$ :

$$nTN = \begin{cases} 1 & : \left( \hat{\mathbf{L}}_j^T = \emptyset \right) \wedge \left( \mathbf{L}_j^T = \emptyset \right) \\ 0 & : \left( \hat{\mathbf{L}}_j^T \neq \emptyset \right) \wedge \left( \mathbf{L}_j^T = \emptyset \right) , \\ 0 & : \left( \hat{\mathbf{L}}_j^T = \emptyset \right) \wedge \left( \mathbf{L}_j^T \neq \emptyset \right) \end{cases}, \quad (6.30)$$

$$TN = \frac{nTN}{N^S}. \quad (6.31)$$

Die auch als Trefferquote (engl. *Hit Rate*) bekannte Korrektklassifikationsrate  $CCR$  berechnet sich aus der Anzahl der erfolgreich zugeordneten Ortsreferenzen, siehe Kapitel 5.3.6.2. Dies beinhaltet neben den korrekt decodierten Ortsreferenzen  $nC_{dec}$  (Abschnitt 6.3.5.2) auch den Anteil der korrekt nicht zugeordneten Ortsreferenzen  $nTN$  (Abschnitt 6.3.5.2). Die Summe dieser korrekten Decoder-Entscheidungen wird dann mit der Anzahl der Ausgangslocations im Quellnetz  $N^S$  normiert:

$$CCR = \frac{nC_{dec} + nTN}{N^S}. \quad (6.32)$$

## 7 Testumgebung, Test Szenarien und Empirische Untersuchungen

### 7.1 Testumgebung zur Analyse von Location Referencing Verfahren

#### 7.1.1 Batch-Prozessierung

Die Grundidee der im Rahmen dieser Arbeit entwickelten und implementierten LR-Testumgebung ist die stapelweise Prozessierung von Ortsreferenzen mit anschließender automatisierter Auswertung und Erstellung eines Berichts (engl. Reporting). Aus Gründen der Verfügbarkeit wurde die OpenLR Musterimplementierung in Version 1.4.0 (TOMTOM 2014) als Basis für die Entwicklung der hier vorgestellten Testumgebung verwendet. Dies sind insbesondere der OpenLR Encoder, der Decoder sowie Klassen und Funktionen zum Umgang mit Karten und Ortsreferenzen etc. Die Prozesse zum Codieren und Decodieren von Ortsreferenzen wurden später verändert und um eigene Zuordnungsansätze erweitert (siehe Kapitel 4).

Zur Definition von Ortsreferenzen, die mittels Location Referencing codiert und übertragen werden, werden diese als Location Strings in einer Textdatei (Location String File) abgelegt. Das Format der Location Strings wird in Klassen des Pakets *openlr.location.utils* spezifiziert und in Kapitel 3.4.1.1 näher erläutert.

#### 7.1.2 Anbindung von Kartendatenbanken

Zur Verwendung der OpenLR Implementierung mit den im Shapefile-Format vorliegenden Kartendaten von TomTom und OpenStreetMap (vgl. Kapitel 7.2.1) musste jeweils eine Software-Schnittstelle in JAVA implementiert werden (ALHESSI 2013). Diese liest die Shapefiles ein und speichert sämtliche Objekte (engl. *Features*) und macht die darin gespeicherten Knoten und Kanten innerhalb der Testumgebung verfügbar. Dabei werden die Objekt-Attribute entsprechend der zuvor festgelegten Schema-Zuordnung (vgl. Kapitel 4.2) in das interne Schema importiert. Beim Einlesen der Geometrie ist vor allem die Unterscheidung der Fahrtrichtungsinformationen wichtig. Innerhalb der OpenLR Implementierung haben alle Kanten nur eine Richtung. Ist eine Kante aus der einzulesenden Karte in beiden Richtungen befahrbar (Normalfall), so sind für diese zwei interne Kanten für die jeweiligen Richtungen anzulegen. Handelt es sich um eine Einbahnstraße, so ist entsprechend nur eine Kante in eben diese Richtung anzulegen. Die Richtungen der Kanten wird dabei implizit durch die Reihenfolge der formbeschreibenden Punkte spezifiziert.

#### 7.1.3 Graphische Benutzeroberfläche

Die LR-Testumgebung verfügt zur Nutzerinteraktion über eine graphische Benutzeroberfläche (engl. *Graphical User Interface (GUI)*). Nach dem Starten der Testumgebung aus der JAVA-Entwicklungsumgebung erscheint ein Dialog zur Auswahl Karten, siehe Abbildung 7.1. Nach Auswahl des jeweiligen Typs (TomTom bzw. OpenStreetMap) ist der Pfad zum entsprechenden Shapefile zu

wählen. Dieser Dialog mit allen seinen Funktionen entstammt der OpenLR Musterimplementierung, Version 1.4.0 (TOMTOM 2014) und wurde für die Zwecke der LR-Testumgebung adaptiert. Nach Bestätigung der Eingaben schließt der Dialog und es werden die gewählten Karten mit den im vorigen Abschnitt beschriebenen Modulen geladen.

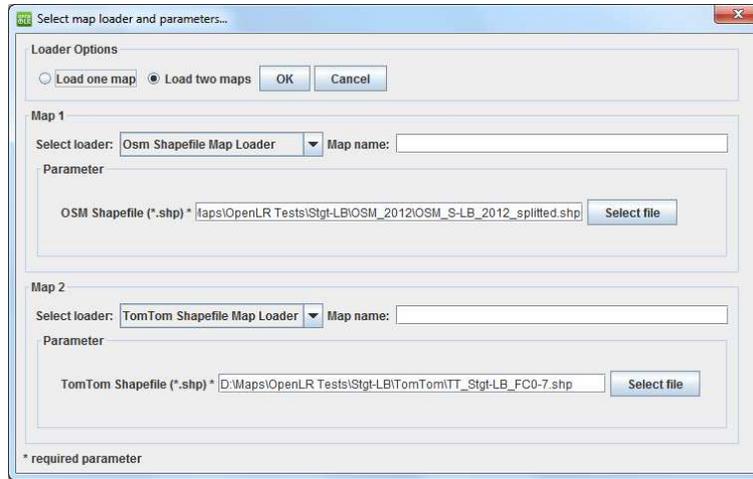


Abbildung 7.1: LR Testumgebung - MapLoader (Quelle: TomTom / OpenLR)

Sind die Karten erfolgreich geladen, erscheint das Hauptfenster der LR-Testumgebung, siehe Abbildung 7.2. Zuerst werden die im vorigen Schritt geladenen Quell- und Zielkarten genannt, damit der Nutzer darunter den passenden Location String File mit der Definition der Test-Ortsreferenzen auswählen kann. Der Button *Run tests* startet den Testdurchlauf, als Statusinformationen werden die Gesamtzahl an Ortsreferenzen sowie die Zahl der momentan prozessierten numerisch und als Fortschrittsbalken angezeigt. Neben einem *Logger*-Feld für die Protokollierung des Testverlaufs können verschiedene Optionen für die Tests über die entsprechenden Checkboxes gewählt werden. Damit ist es möglich, eine Location String Datei mit Referenz-Locations (=SOLL) im Zielnetz zu laden sowie den Test um eine Re-Encodierung und LR-Rückübertragung ins Quellnetz zu erweitern. Ebenso können verschiedene Ergänzungsmodul für das Form-Matching ausgewählt und schlussendlich Toleranzwerte für die Überprüfung der Ergebnisse definiert werden.

#### 7.1.4 Reporting

Zur Beurteilung der LR Ergebnisse sowie um die internen Abläufe (z.B. das Rating und die Entscheidung zwischen den Decoding-Kandidaten) für jede prozessierte Ortsreferenz nachvollziehbar zu machen, wurde die bestehende OpenLR Implementierung auch um eine umfangreiche Protokoll- und Reporting Funktionalität erweitert.

Die wichtigen Bausteine eines LR Prozesses sind das Auffinden möglicher Decoding-Kandidaten im Zielnetz, deren Bewertung anhand der mit der Ortsreferenz übertragenen Parameter, sowie schließlich die Auswahl des vermeintlich am besten passenden Kandidaten. Dieser ist dann die decodierte Ortsreferenz im Zielnetz. In realen LR-Anwendungen wären diese internen Abläufe für den Nutzer völlig uninteressant, ihn würde nur die schlussendlich decodierte Ortsreferenz mit ihrem Verlauf im Zielnetz interessieren. Zur Beurteilung eines LR-Verfahrens bzw. zum Vergleich mehrerer Verfahren oder verschiedener Varianten ist es erforderlich, diesen Bewertungs- und Entscheidungsprozess beim Decoding für jede einzelne Ortsreferenz und alle ihre Kandidaten zu dokumentieren und einer späteren Auswertung zuzuführen. Dazu wurde die bestehende OpenLR Implementierung um ein neu entwickeltes Reporting System erweitert. In den Rating Funktionen werden für jeden untersuchten Kandidaten dessen für das Rating relevante Eigenschaften, wie z.B. Richtungswinkel, Functional Road Class und

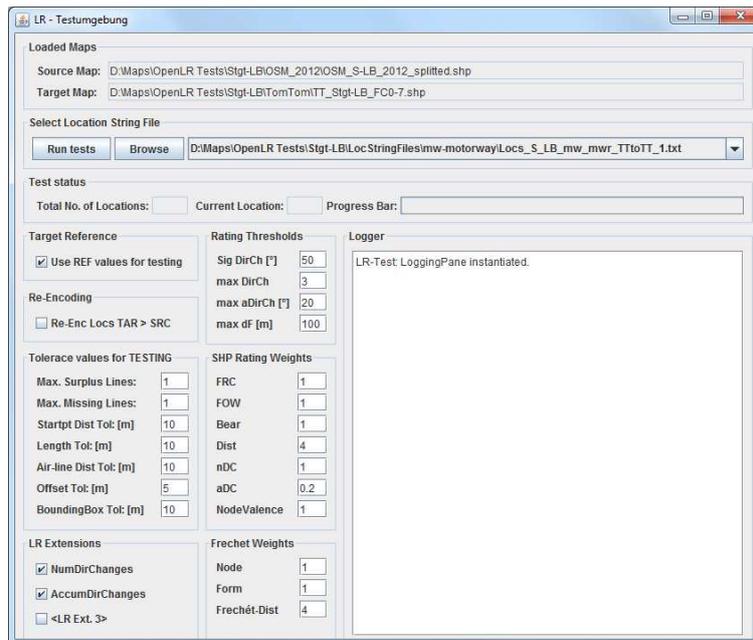


Abbildung 7.2: LR Testumgebung - Hauptmenü

Form of Way sowie die jeweils bestimmten Einzelbewertungen und das Gesamtergebnis jedes Kandidaten festgehalten. Im Decoder werden zusätzlich noch etwaige Fehlermeldungen oder Ausschlussgründe protokolliert.

Nach Durchlaufen aller Ortsreferenzen eines Location String Files wird dann ein Testbericht in Form einer Microsoft Excel Arbeitsmappe (\*.xls) automatisiert erstellt. Das Excel Format wurde gewählt, weil damit die numerischen Ergebnisse übersichtlich visualisiert werden können und es gleichzeitig eine Vielzahl von Auswerte- und (graphischen) Darstellungsfunktionen für die weitere manuelle Analyse bietet. Der Bericht dokumentiert zum einen die zuvor genannten und gespeicherten Zwischenergebnisse in tabellarischer Form, so dass die Entscheidungsfindung für jeden Kandidaten im Detail nachvollzogen werden kann. Des Weiteren erfolgt auch die vollständige numerische Auswertung des Testlaufs inklusive der Dokumentation der Ergebnisse. Dazu werden die decodierten Ortsreferenzen mit denen aus dem Quellnetz sowie eventuell vorhandener Referenzdaten im Zielnetz verglichen.

Die decodierten Ortsreferenzen im Zielnetz werden ebenfalls in ein Location String File geschrieben und liegen somit wie die originalen Ortsreferenzen aus dem Quellnetz im selben Format für weitere Auswertungen (z.B. Visualisierung, siehe nächster Abschnitt) vor.

### 7.1.5 Visualisierung der Ergebnisse

Zur Beurteilung der Zuordnungsqualität und somit auch der Güte des eingesetzten Location Referencing bzw. Zuordnungsverfahrens ist neben den zuvor genannten numerischen Auswerteverfahren eine grafische Visualisierung der originären Ortsreferenzen im Quellnetz sowie der entsprechenden im Zielnetz decodierten Ortsreferenzen hilfreich. Dadurch sind die Kartenunterschiede im Gebiet um die jeweilige Ortsreferenz, sowie die Decoding-Unterschiede der jeweiligen LR-Verfahren sofort ersichtlich. Als für die Auswertung überaus hilfreich hat sich die Visualisierung der *Location Reference Points* herausgestellt. Ohne das Betrachten der codierten und auch decodierten LRPs kann keine vollständige Bewertung einer LR-Übertragung erfolgen. Bei Ortsreferenzen, die nicht direkt am LRP beginnen oder enden, ist eine detaillierte Ursachenforschung für falsche bzw. inkorrekte LR-Übertragungen ohne Betrachten der LRPs nicht möglich.

Die Darstellung dieser Daten erfolgt in einem Geoinformationssystem, im Rahmen der vorliegenden Arbeit wurde *ESRI ArcGIS Desktop 10.0 (SP 5)* verwendet. Zur automatisierten Verarbeitung der Ausgabedaten der LR-Testumgebung und Generierung der Visualisierung wurden verschiedene *Python* Skripte entwickelt, welche in die *ArcToolbox* eingebunden und von dort aus aufgerufen werden können. Das sogenannte *LocationViewer*-Tool wertet die Dateien mit den *Location Strings* (siehe Kapitel 3.4.1.1) der Original-, Decodierten- und ggfs. Referenz-Locations aus. Darin sind die Ortsreferenzen als Kanten-Folge mit etwaigen Offsets am Anfang und Ende definiert. Diese Kanten werden anhand ihrer Identifikatoren der zugehörigen Karte (Shapefile) entnommen und die formbeschreibenden Punkte aller Kanten einer Ortsreferenz zu einer Polylinie zusammengefügt. Diese Polylinien werden dann noch um die etwaigen Offsets gekürzt und als neues Shapefile gespeichert. Sämtliche so gewonnenen Location Shapefiles sowie die Location Reference Points und die Quell- und Zielkarten werden nach Ende der Prozessierung in den aktuellen ArcGIS Arbeitsbereich geladen. Zur Steigerung der Übersichtlichkeit und Handhabbarkeit werden die jeweils gleichen Ortsreferenzen aus den verschiedenen Datensätzen (Original, Decodiert, Referenz) sowie die zugehörigen Location Reference Points zu einem Gruppenlayer zusammengefasst. Die Darstellungsart der einzelnen Layers wird jeweils von einem als Vorlage hinterlegten Layer-File übernommen.

Anhang B enthält eine detailliertere Beschreibung des *LocationViewer*-Tools.

## 7.2 Praktische Durchführung der Untersuchungen

### 7.2.1 Verwendete Kartendaten

Sämtliche im Rahmen der vorliegenden Arbeit durchgeführten Tests wurden mit digitalen Straßenkarten von TomTom und OpenStreetMap durchgeführt.

Da die Erfassung von Fußgängerwegen zum Testzeitpunkt noch nicht flächendeckend abgeschlossen ist, konzentrieren sich die Tests auf das Straßennetzwerk. Grundsätzlich sollten die gewonnenen Erkenntnisse aufgrund der geometrischen Ähnlichkeit von Straßen- und Fußgängernetzwerken im Wesentlichen übertragbar sein.

#### 7.2.1.1 TomTom

TomTom ist ein kommerzieller Hersteller von digitalen Straßenkarten und Navigationsgeräten für Consumer Anwendungen (engl. *Personal Navigation Device* PND) und als Zulieferer von Karten, Verkehrsinformationen und Software-Komponenten im Automotive Bereich. Der Geschäftsbereich zur Herstellung digitaler Straßenkarten wurde 2007 durch die Übernahme von Tele Atlas akquiriert (TOMTOM 2015).

Für die Tests wurde die TomTom *MultiNet* Kartendatenbank Deutschland, Version 2012.06 verwendet. Die Kartendaten liegen als *ESRI Shapefile (\*.shp)* (ESRI 1998) vor. Das für die Location Referencing Anwendung benötigte Straßennetzwerk ist im sog. *Network Theme (\_NW)* des TomTom MultiNet Kartendatensatzes enthalten. Die *Network* Shapefiles enthalten die Straßengeometrie als routingfähiges Netzwerk und müssen somit für ihre Verwendung mit OpenLR nicht mehr weiter prozessiert werden. Daneben enthalten die Shapefiles auch einen Satz von Basisattributen (z. B. Straßename, Straßenklasse, usw.), welcher für die Verwendung mit OpenLR ausreicht.

### 7.2.1.2 OpenStreetMap

Als zweiter Datensatz dient ein Auszug aus der frei verfügbaren OpenStreetMap (vgl. Kapitel 2.2.4.3). Um für die im Rahmen dieser Arbeit durchzuführenden Tests eine gleiche zeitliche Situation wie im zuvor beschriebenen TomTom-Datensatz darzustellen, wurde eine OSM-Karte vom 13.06.2012 verwendet. OpenStreetMap Kartendaten liegen im Original in einem XML-basierten Format vor (\*.osm). Zur Nutzung mit der LR-Testumgebung müssen die OSM Daten allerdings in das ESRI Shapefile-Format konvertiert werden. Außerdem enthalten die originären OSM Daten sehr viele Informationen, die für die Location Referencing Anwendung nicht gebraucht werden (z.B. Fußgängerwege, Bus- und Straßenbahnlinien sowie Eisenbahnverbindungen, Wasserwege).

Im Internet ist eine Vielzahl von bereits vorkonfektionierten OSM-Ausschnitten in unterschiedlichen Datenformaten, darunter auch das hier benötigte Shapefile Format, verfügbar. Allerdings ist zur Verwendung der Kartendaten mit OpenLR ein Mindestmaß an Attributinformationen erforderlich, welches sich aus der OpenLR Definition ergibt (TOMTOM 2012). Leider entsprach keiner der verfügbaren vorkonfektionierten OSM-Datensätze diesen Anforderungen, was eine manuelle Prozessierung der OSM-Daten erforderlich machte. Zur Nachvollziehbarkeit der einzelnen Prozessierungsschritte werden diese im Anhang A detailliert erläutert.

### 7.2.1.3 Kartenausschnitt

Für die Tests wurde aus den zuvor genannten TomTom und OpenStreetMap Karten ein Ausschnitt der Region Stuttgart-Ludwigsburg ausgeschnitten und in Form von Shapefiles für die weitere Verwendung gespeichert. Die genaue Abgrenzung des verwendeten Kartenausschnittes ist anhand der WGS-84 Koordinaten in Tabelle 7.1 dargestellt. Abbildung 7.3 zeigt die Straßen höherer Ordnung im verwendeten Testgebiet grafisch.

Tabelle 7.1: Lage Kartenausschnitt

Kartenrand	Nord	Süd	West	Ost
WGS 84-Koord. [°]	49.0280	48.7095	8.8756	9.7339

## 7.2.2 Verwendete Testdatensätze / Testscenarien

Für den zuvor genannten Kartenausschnitt wurden zur detaillierten empirischen Untersuchung des im Rahmen der vorliegenden Arbeit entwickelten Verfahrens und zum Vergleich mit der bestehenden OpenLR Implementierung vier Testscenarien definiert. Da die räumliche Situation in den Karten von Klasse und Nutzungsart der Straßen abhängt, wurden zunächst drei dieser Testscenarien für die Bereiche Autobahn (engl. Motorway), Landstraße (engl. Inter-Urban) und Innerstädtisch (engl. Urban) erstellt. Kreuzungssituationen stellen für das Location Referencing ebenfalls eine große Herausforderung dar. Deshalb wurde ein vierter Datensatz erstellt, um gezielt das Verhalten der Location Referencing Verfahren an Kreuzungen untersuchen zu können. Für jedes dieser vier Testscenarien wurde für die Durchführung der Untersuchungen ein Datensatz mit Testlocations erstellt. Tabelle 7.2 gibt dazu eine Übersicht über die Szenarien und die Größe der zugehörigen Datensätze.

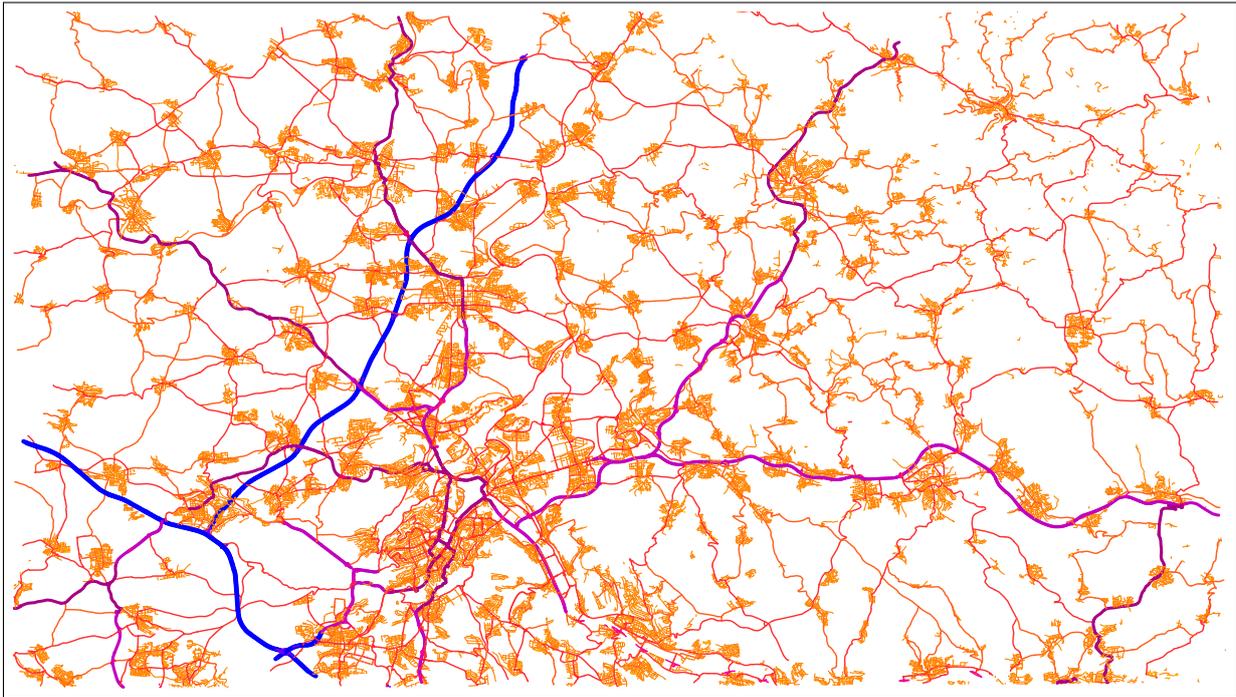


Abbildung 7.3: Übersicht Kartenausschnitt

Tabelle 7.2: Übersicht Testdatensätze im Gebiet Stuttgart-Ludwigsburg

Szenario	Anzahl Ortsreferenzen
Autobahn	64
Inter-Urban	23
Urban	120
Kreuzungen	82

### 7.3 Ergebnisse

Die im vorigen Abschnitt genannten Testdatensätze wurden mit Hilfe der in Abschnitt 7.1 beschriebenen Location Referencing Testumgebung verarbeitet. Für die einzelnen Ausgangs-Locations aus der Quellkarte wurden dabei Location References (Ortsreferenzen) durch den OpenLR Encoder erstellt. Diese wurden dann zum einen mittels OpenLR im Zielnetz decodiert. Zum anderen wurden Ortsreferenzen mit einem erweiterten Encoder für die Verwendung mit dem Form-Matcher codiert. Diese erweiterten Ortsreferenzen enthalten zusätzlich Angaben zur Anzahl signifikanter Richtungswechsel sowie die kumulierten Richtungsänderungen (vgl. Kapitel 4.5) und dann im Zielnetz mittels *Form-Matcher* decodiert. Dabei wurden echte Netzknoten bei der Suche nach den Location Reference Points im Zielnetz bevorzugt. Des Weiteren wurden die formbeschreibenden Eigenschaften für alle möglichen Routenverläufe der Kandidaten jeder Ortsreferenz im Zielnetz ermittelt und sind in die Bewertung der jeweiligen Kandidaten-Route eingeflossen.

Die dadurch erhaltenen Ortsreferenzen im Zielnetz wurden anhand der Qualitätsparameter des in Kapitel 6 abgeleiteten Qualitätsmodells evaluiert. Die folgenden Kapitel enthalten eine Zusammenstellung der numerischen Ergebnisse aller Testszenarien der Karten-Konstellation OSM->TomTom sowie eine detaillierte Beschreibung und Analyse der relevanten Fälle. Die Ergebnisse der entgegengesetzten Übertragungsrichtung (TomTom->OSM) sind in Kapitel 7.4.2 beschrieben. Die Darstellung sämtlicher abgebildeter Kartenausschnitte orientiert sich an der Legende in Abbildung 7.4.

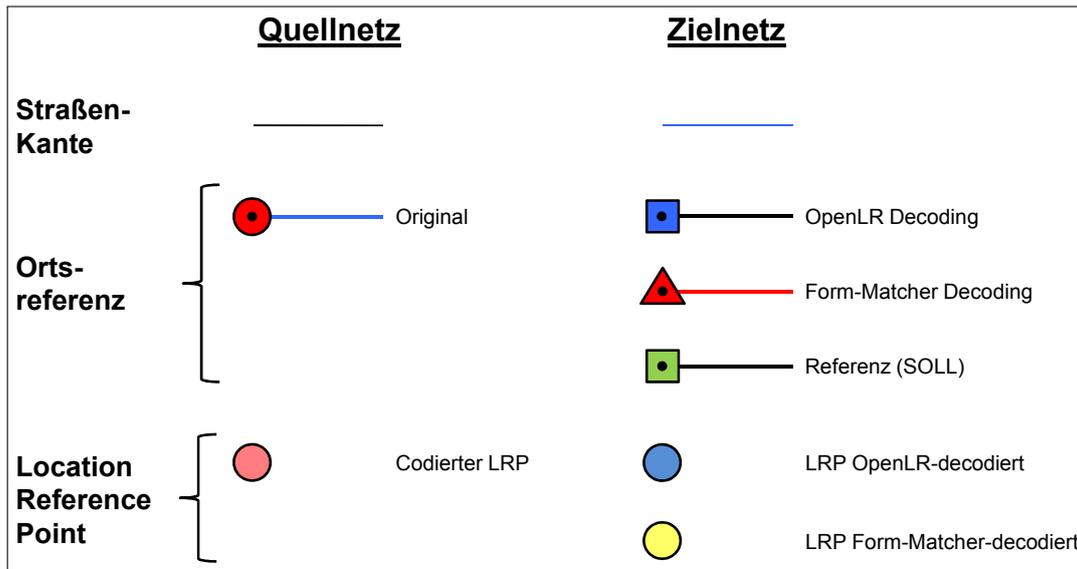


Abbildung 7.4: Legende für Kartendarstellungen

### 7.3.1 Szenario: Autobahn

Der Datensatz *Autobahn* besteht aus 64 Ortsreferenzen. Diese befinden sich sowohl auf freier Strecke, als auch im Bereich von Anschlussstellen und an Autobahnkreuzen.

Charakteristisch für diesen Datensatz sind langgestreckte Kanten, die häufig zur Modellierung von Attributwechseln (z.B. bei Änderungen der Höchstgeschwindigkeit oder an Brücken bzw. Tunnels) durch zweiwertige Knoten verbunden sind. Das erschwert eine direkte Knotenzuordnung in diesen Bereichen. Ortsreferenzen auf freier Strecke weisen fast immer sehr große Entfernungen zu den Location Reference Points auf, denn sie dürfen vom LR-Encoder nur auf echten Netzknoten mit Valenz  $> 2$  gesetzt werden. Diese sind auf Autobahnen praktisch nur an den Anschlussstellen vorhanden, und somit werden Ortsreferenzen in solchen Fällen bis an die jeweils nächsten Anschlussstellen referenziert.

Die Anschlussstellen bzw. Übergänge an Autobahnkreuzen sind aufgrund der hohen Fahrgeschwindigkeiten sehr häufig langgestreckt. Entsprechend schwer ist es, die Positionen der Verzweigungsknoten eindeutig zu bestimmen. Die Karten weisen in diesen Bereichen häufig sehr große geometrische Unterschiede auf. Dies erschwert die korrekte Zuordnung von Objekten zwischen den beiden Karten in solchen Bereichen ebenfalls.

Tabelle 7.3 zeigt die numerischen Evaluierungs-Ergebnisse der Zuordnungen mit OpenLR und dem im Rahmen dieser Arbeit entwickelten Form-Matcher. Dabei lässt der Form-Matcher allgemein eine Verbesserung der Zuordnungsrate ('Hit Rate') auf 64 % erkennen. Dennoch sind diese Ergebnisse zunächst nicht ganz zufriedenstellend und sollen im Folgenden diskutiert werden.

Eine detaillierte Betrachtung der Ergebnisse zeigt, dass die geometrische Korrektheit der Zuordnungen an den Anfangs- und Endpunkten auf hohem Niveau um 2 auf 52 Ortsreferenzen (81 %) gesteigert werden konnte. Die topologischen Bewertungen für die Identifikation der Start- und Endkante und des Locationverlaufs weisen jedoch schlechtere Erfolgsraten auf. Dies liegt vor allem an der Lage der decodierten Start-/Endpunkten der Ortsreferenzen. Wie bereits angesprochen, weist der Autobahn-Datensatz im Vergleich zu den anderen Szenarien große Modellierungsunterschiede bezüglich der einzelnen Kanten auf. Immerhin kann der Form-Matcher in diesem schwierigen Umfeld die topologische Korrektheit der Zuordnungen auf 67 % der Testfälle steigern. Die Gründe für die schwierigen Bedingungen für Location Referencing auf Autobahnen sollen im Folgenden diskutiert werden.

Tabelle 7.3: Auswertung Datensatz 'Autobahnen' [OSM-&gt;TT]

Parameter		OpenLR	Form-Matcher	Veränderung (OpenLR=100%)
Ortsreferenzen		64 (100 %)	64 (100 %)	-
decodiert	$Com_{dec}$	64 (100 %)	64 (100 %)	-
geom. korrekt	CG	50 (78 %)	52 (81 %)	+ 2 (+ 4 %)
topol. korr. (Start/Ende)	CE	40 (63 %)	43 (67 %)	+ 3 (+ 8 %)
topol. korr. (Verlauf)	CR	38 (59 %)	43 (67 %)	+ 5 (+ 13 %)
korrekt zugeordnet	$r_p^+$	35 (55 %)	41 (64 %)	+ 6 (+ 17 %)
korrekt nicht zugeordn.	$r_n$	0 (0 %)	0 (0 %)	-
korrekt klassifiziert	CCR	35 (55 %)	41 (64 %)	+ 6 (+ 17 %)
Mittel Dist. Start/Endpkt.	$\bar{d}$	26,3 m	26,6 m	
Mittel Längendifferenz	$\bar{\Delta}l$	-7,1 m	-10,6 m	

Die Anschlussstellen der Autobahnen sind durch relativ kurze Kanten gekennzeichnet, da hier neben den topologischen Verzweigungen häufig auch Attributwechsel mit zweiwertigen Zwischenknoten in den Karten dargestellt werden müssen. Dadurch kann es in manchen Fällen vorkommen, dass ein LRP, wie in Abbildung 7.5(a) gezeigt, an einer Anschlussstelle im Quellnetz [blaues Netz, roter Punkt] von OpenLR aufgrund der kürzeren Entfernung auf einen nahegelegenen zweiwertigen Knoten im Zielnetz decodiert wird [schwarzes Netz, blauer Punkt]. Der Form-Matcher decodiert den LRP hingegen aufgrund der Berücksichtigung der Knoten-Valenz korrekterweise auf den etwas weiter entfernten dreiwertigen Knoten im Zielnetz [schwarzes Netz, gelber Punkt]. Das zu diesem LRP zugehörige Location-Ende liegt mit einem großen Offset auf offener Strecke, siehe Abbildung 7.5(b). Durch die korrekte Zuordnung des LRPs durch den Form-Matcher wird auch der entsprechende Location-Endpunkt mit einer geometrischen Abweichung von zwei Metern zum Sollpunkt [grünes Quadrat] auf der korrekten Kante in das Zielnetz decodiert [rotes Dreieck]. Diese Lösung ist somit topologisch wie geometrisch korrekt. Der von OpenLR decodierte Endpunkt [blaues Quadrat] liegt aufgrund der falschen LRP-Zuordnung auf der nächsten anschließenden Kante und wird somit als topologisch nicht korrekt zugeordnet bewertet.

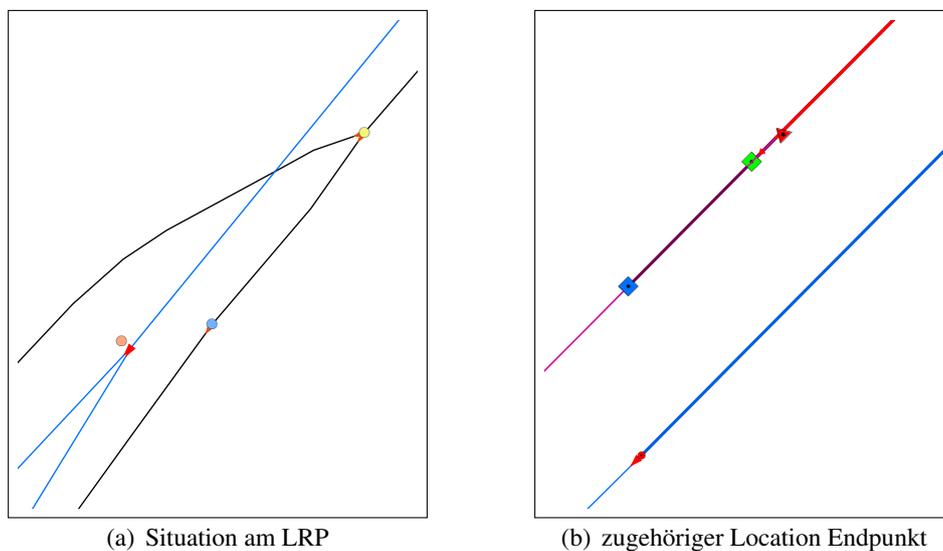


Abbildung 7.5: Beispiel korrektes LRP-Decoding dank Knoten-Valenz

Auf-, Ab- und Überfahrten sind, wie bereits erwähnt, in der Regel langgestreckt, sodass ihre digitalen Repräsentationen oft große geometrische Unterschiede zwischen den verschiedenen Karten aufweisen. Dadurch kommt es häufig vor, dass der gesuchte LRP im Zielnetz von OpenLR nicht innerhalb des Suchradius gefunden werden kann, und deshalb eine Projektion auf die nächstgelegene Kante erfolgt, um überhaupt eine Lösung zu finden. In diesen Fällen wird häufig eine an die eigentliche Ortsreferenz anschließende Kante im Zielnetz ausgewählt, so dass diese Lösungen als topologisch nicht korrekt bewertet werden müssen. Durch die Dynamisierung des Suchradius (Kap. 4.3.1) und die Berücksichtigung der Knotenvalenz (Kap. 4.4.2) bei der Suche nach den LRPs im Zielnetz konnte die topologische Korrektheit der Zuordnung deutlich gesteigert werden.

Auf freier Strecke werden Kanten oft zur Darstellung von Attributwechseln (z.B. Änderung der zulässigen Höchstgeschwindigkeit oder Beginn/Ende von Tunneln bzw. Brücken etc.) unterbrochen. Ähnlich wie an den Anschlussstellen, weisen auch diese zweiwertigen Verbindungsknoten teilweise große geometrische Unterschiede zwischen den Karten auf. Dies ist vermutlich der hohen Fahrgeschwindigkeit bei der Datenerfassung geschuldet, denn dadurch schlagen sich schon geringe Auslöseverzögerungen in großen geometrischen Unterschieden nieder.

Die Location Reference Points von Ortsreferenzen, die auf freier Strecke mit großem Abstand zur nächsten Anschlussstelle beginnen oder enden, müssen vom Encoder auf den nächstgelegenen Knoten mit Wertigkeit  $\geq 2$  gelegt werden. Diese liegen in der Regel an den nächsten Anschlussstellen. Dadurch entstehen teilweise sehr große geometrische Offsets von einigen Kilometern zu den eigentlichen Start-/Endpunkten der Ortsreferenzen, abhängig von ihrer genauen Definition. Selbst wenn diese LRPs korrekt in der Zielkarte identifiziert werden können, entstehen aufgrund der geometrischen Kartenunterschiede beim Abtragen der langen geometrischen Offsets bis zum tatsächlichen Location-Beginn bzw. -Ende geometrische Abweichungen in der Lage dieser Punkte. Bedenkt man nun die zuvor beschriebenen Unterschiede in den Kantenmodellierungen, so wird klar, dass es bei der Auswertung des vorliegenden Autobahn-Szenarios selbst bei kleinen geometrischen Abweichungen zu Einbußen bei der topologischen Korrektheit kommen kann. Damit können zumindest teilweise die im Vergleich zur geometrischen Korrektheit (81 %) schlechteren Raten der topologischen Korrektheit der Start-/Endpunkte (67 %) erklärt werden.

Es fällt auf, dass insbesondere die Ergebnisse der Korrektheit des topologischen Location Verlaufs um fünf auf 43 Ortsreferenzen (67 %) gesteigert werden konnte. Hier macht sich die Betrachtung des gesamten Locationverlaufs des Form-Matchers im Vergleich zur Abschnittweisen Bewertung durch OpenLR und die dadurch mögliche Beurteilung der Form der Ortsreferenz-Kandidaten positiv bemerkbar. Abbildung 7.6 zeigt einen solchen Fall beispielhaft. Die Ausgangslocation im Quellnetz [fett blau] folgt dem Verlauf einer Autobahn an einer Anschlussstelle. Während das mit OpenLR decodierte Ergebnis fälschlicherweise die Autobahn an dieser Anschlussstelle auf die Parallelfahrbahn verlässt [fett schwarz] bleibt die mit dem Form-Matcher decodierte Ortsreferenz korrekterweise auf der Autobahn [fett rot]. Der Grund hierfür ist wiederum in der Auswahl des LRP Kandidaten im Zielnetz zu suchen. Da an dieser Stelle die Ortsreferenz im Ausgangsnetz vom kürzesten Weg abweicht, hat der Encoder einen zusätzlichen LRP eingefügt [roter Punkt], um die Richtungsabweichung von der kürzesten-Wege-Verbindung zu beschreiben. Anders als der OpenLR Decoder mit seiner individuellen Betrachtung der LRPs, wählt der Form-Matcher richtigerweise den weiter entfernten aber topologisch korrekten LRP-Kandidaten [gelber Punkt], anstatt dem näher gelegenen, aber topologisch inkorrekten zweiwertigen Knoten [blauer Punkt]. Dadurch kann der Form-Matcher an dieser Stelle die decodierte Ortsreferenz im Zielnetz nicht entlang der kürzesten Verbindung, sondern korrekterweise entlang der intendierten Route führen. Zwei der vom Form-Matcher topologisch korrekt decodierten Fälle erfüllen jedoch die geometrischen Anforderungen nicht, daher gelten insgesamt 64 % der Ortsreferenzen als korrekt zugeordnet.

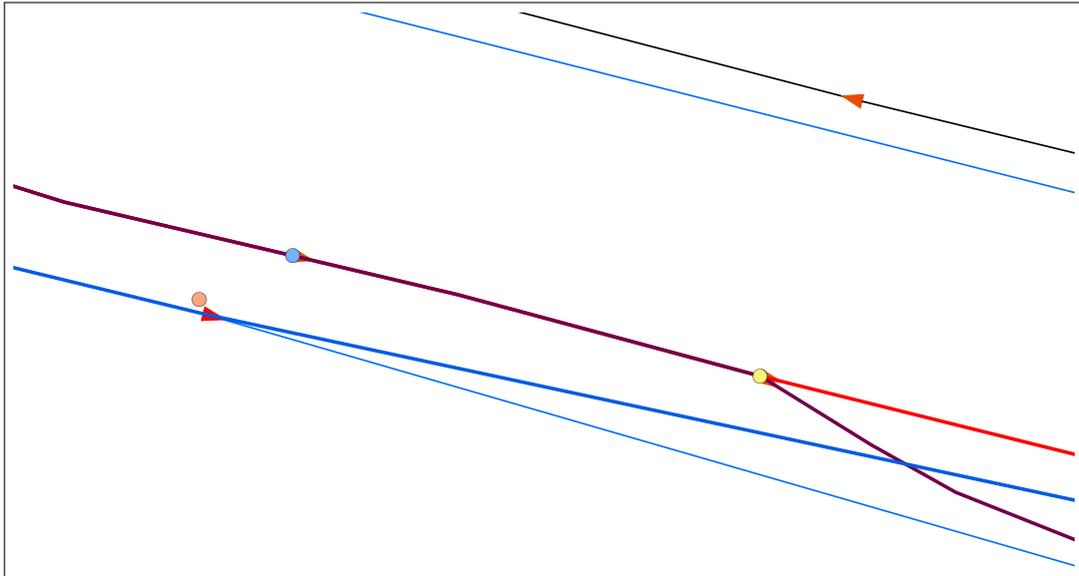


Abbildung 7.6: Beispiel korrekter Routenverlauf mit Form-Matcher

### 7.3.2 Szenario: Inter-Urban

Der Datensatz *Inter-Urban* besteht aus 23 Ortsreferenzen, die hauptsächlich auf Landstraßen zwischen bebauten Gebieten (Ortschaften) liegen. Die Ortsreferenzen sind meist langgestreckt, sie beginnen und enden im Bereich der Ortsausgänge oder auch auf freier Strecke. Anders als beim Datensatz *Autobahn* haben die Kreuzungen im überörtlichen Landstraßenbereich üblicherweise eine geringere geometrische Ausdehnung. Daher stimmen die Abbildungen der Kreuzungssituationen zwischen den verwendeten Karten in diesem Datensatz besser miteinander überein. Bei Ortsreferenzen mit großer Entfernung zur nächsten Kreuzung bleibt das Problem der teilweise stark abweichenden Kanten-Aufteilung und damit die Schwierigkeiten beim Zuordnen übereinstimmender Knotenpaare, bestehen.

Diese Veränderungen in den örtlichen Gegebenheiten bzw. den Kartenabbildungen spiegeln sich auch in den numerischen Ergebnissen der Validierung der Zuordnung für diesen Datensatz wider, die in Tabelle 7.4 dargestellt sind. Das Niveau der Zuordnungs-Korrektheit liegt im Vergleich zum *Autobahn*-Datensatz bei beiden Zuordnungsverfahren etwas höher. Auch im vorliegenden *Inter-Urban* Datensatz konnte das formbasierte Matching die Korrektheit der Zuordnungen im Vergleich zu OpenLR steigern.

Tabelle 7.4: Auswertung Datensatz 'Inter-Urban' [OSM-&gt;TT]

Parameter		OpenLR	Form-Matcher
Ortsreferenzen		23 (100 %)	23 (100 %)
decodiert	$Com_{dec}$	23 (100 %)	23 (100 %)
geom. korrekt	CG	20 (87 %)	23 (100 %)
topol. korr. (Start/Ende)	CE	17 (74 %)	18 (78 %)
topol. korr. (Verlauf)	CR	17 (74 %)	18 (78 %)
korrekt zugeordnet	$r_p^+$	16 (70 %)	18 (78 %)
korrekt nicht zugeordn.	$r_n$	0 (0 %)	0 (0 %)
korrekt klassifiziert	CCR	16 (70 %)	18 (78 %)
Mittel Dist. Start/Endpkt	$\bar{d}$	5,3 m	3,0 m
Mittel Längendifferenz	$\bar{\Delta}l$	2,1 m	-0,9 m

Die geometrische Korrektheit der Start- bzw. Endpunkte der Ortsreferenzen erreichte beim Form-Matcher 100 %, d.h. die geometrischen Abweichungen aller decodierten Ortsreferenzen lagen innerhalb der zulässigen Schwellwerte (diese werden in Abhängigkeit der jeweils zulässigen Höchstgeschwindigkeit bestimmt, vgl. Kapitel 6.2.3). Diese deutliche Verbesserung gegenüber dem *Autobahn*-Szenario kommt hauptsächlich durch die bessere Zuordnung der Location Reference Points zustande. Beim Blick auf die typischen LRP-Situationen im überörtlichen Bereich fällt auf, dass diese entgegen dem *Autobahn*-Szenario deutlich weniger Modellierungsunterschiede in Bezug auf Kantenunterbrechungen aufweisen. Die Berücksichtigung der Knoten-Valenz hilft dem Form-Matcher, die von OpenLR häufig ausgewählten, zweiwertigen Knoten zu vermeiden und stattdessen die korrekten Zuordnungen zu den LRPs zu identifizieren. Durch diese korrekte LRP-Zuordnung des Form-Matchers bleiben die geometrischen Abweichungen nach dem Abtragen der Offsets gering.

Auch bei der Auswertung der topologischen Korrektheit schneidet der Form-Matcher mit 78 % besser ab als OpenLR. Dass im Gegensatz zur geometrischen Auswertung dennoch keine 100 % topologischer Korrektheit erreicht wurden, liegt vor allem am Abtragen der geometrischen Location Offsets. Wie bereits erwähnt, konnten die LRPs vom Form-Matcher sehr häufig korrekt im Zielnetz identifiziert werden. Besitzt die Ortsreferenz jedoch einen großen Offset zum eigentlichen Location-Beginn, kann dieser, bedingt durch die geometrischen Kartenunterschiede, beim Abtragen des Offsets in die Zielkarte auf einer angrenzenden Kante zu liegen kommen. Dadurch weist dieses Decoding Ergebnis zwar wie oben erwähnt, geringe geometrische Abweichungen auf, muss aber nach den hier angewendeten Bewertungskriterien als topologisch inkorrekt klassifiziert werden. Abbildung 7.7 zeigt beispielhaft eine solche typische Situation. Der im Quellnetz [blau] an der Einfahrt zum Kreisverkehr positionierte LRP [roter Punkt] wurde vom Form-Matcher im Gegensatz zu OpenLR korrekt decodiert [gelber Punkt]. Durch das Abtragen des Offsets liegt das tatsächliche Location-Ende [rotes Dreieck] zwar näher an der Soll-Lage [grünes Quadrat] als die OpenLR-Lösung [blaues Quadrat]. Allerdings liegt die Lösung des Form-Matchers durch die geometrischen Kartenunterschiede auf der nächsten Kante und ist somit topologisch inkorrekt.

### 7.3.3 Szenario: Urban

Für das Szenario *Urban* wurden 120 Ortsreferenzen im innerstädtischen Bereich ausgewählt. Dieser umfasst sowohl kleinräumige Wohnstraßen mit beidseitiger Befahrbarkeit als auch Durchgangsstraßen mit überörtlicher Bedeutung und in den Karten teilweise getrennt digitalisierten Richtungsfahrbahnen. Die Ortsreferenzen sind überwiegend langgestreckt und simulieren so beispielsweise den Gültigkeitsbereich einer Geschwindigkeitsbegrenzung. Sie beginnen und enden häufig direkt nach/vor einer Kreuzung, teilweise mit geometrischen Abständen (Offsets) zu diesen.

Die Betrachtung der numerischen Evaluierungsergebnisse in Tabelle 7.5 zeigt, dass der Form-Matcher auch im innerörtlichen Bereich durchweg gute Ergebnisse erzielt. Es ist eine deutliche Steigerung der Zuordnungsqualität gegenüber OpenLR erkennbar. Fast alle (93 %) der zuzuordnenden Ortsreferenzen konnte der Form-Matcher geometrisch korrekt zuordnen. 82 % der Ortsreferenzen wurden topologisch korrekt zugeordnet. Das Qualitäts-Niveau der Zuordnungen liegt in etwa gleich hoch wie beim Szenario *Inter-Urban*. Die Ursache ist vermutlich darin zu suchen, dass viele der Ortsreferenzen direkt an nicht komplexen Kreuzungen beginnen bzw. enden, und diese wiederum relativ gut im Zielnetz identifizierbar sind.

Dennoch bleibt ein signifikanter Anteil der Ortsreferenzen, die mit beiden Verfahren nicht korrekt zugeordnet werden können. Durch das von beiden Verfahren verwendeten LRP-Prinzip ist dies beispielsweise häufig dort der Fall, wo die als LRP gewählten Kreuzungspunkte durch Kartenunterschiede einen großen geometrischen Abstand aufweisen. Dadurch können die korrekten Zuordnungslösungen nicht in der Zielkarte identifiziert werden, oder erhalten aufgrund des großen geometrischen Abstandes eine sehr schlechte Knoten-Bewertung und können die Auswahlprozedur nicht gewinnen. Ebenso kann

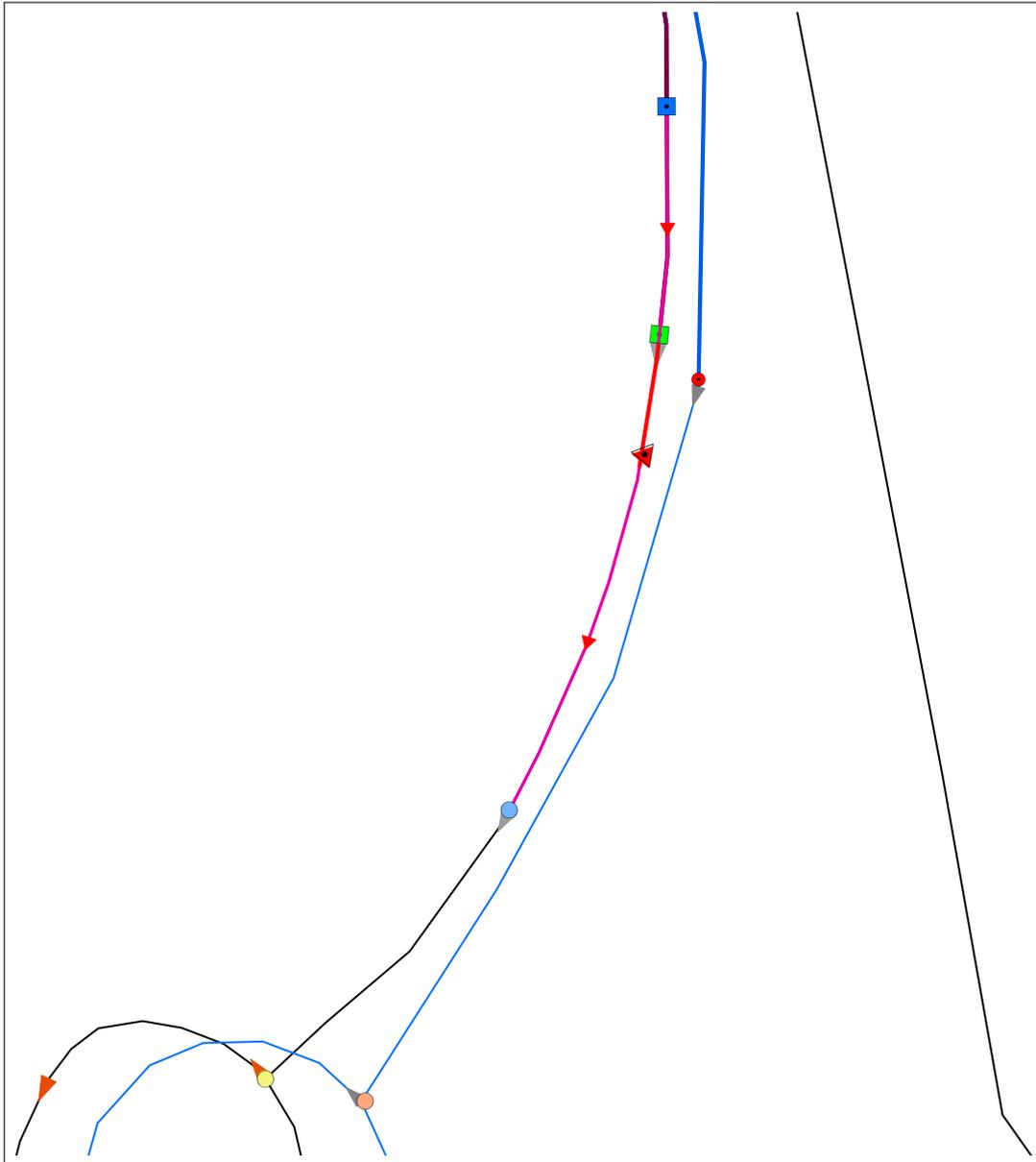


Abbildung 7.7: Beispiel topologisch inkorrekte Ortsreferenz trotz korrekter LRP-Zuordnung

bei Ortsreferenzen, die an komplexen Kreuzungen beginnen oder enden, eine erhöhte Fehlerhäufigkeit beobachtet werden. Komplexe Kreuzungen zeichnen sich oft durch die Anbindung an mehrspurig erfasste Straßen und zur Modellierung von Abbiegebeziehungen separat erfasste Querspannen aus, siehe auch Abbildung 7.8. Dadurch entstehen im Bereich der Kreuzung viele, relativ kurze Kanten. Dies führt zu einer großen Steigerung der Zahl der potenziellen Zuordnungskandidaten in diesem Bereich und erschwert die Auswahl der korrekten Kante durch den Decoder aber auch die Evaluierung der Ergebnisse. Durch die angesprochene Komplexität der Kreuzungen treten in diesen Bereichen mit hoher Wahrscheinlichkeit auch große Modellierungsunterschiede zwischen den Karten auf. Dadurch ist es in vielen Fällen schwer oder überhaupt nicht möglich, die korrekte Zuordnungslösung im Zielnetz zu identifizieren (vgl. Kapitel 6.1.2).

Abbildung 7.8 zeigt ein typisches Beispiel für eine komplexe Kreuzungssituation. Die Ausgangslocation im Quellnetz [blau] beginnt am Süd-westlichen Ausgang dieser komplexen Kreuzung [fett blau]. Da die Ortsreferenz an einem echten Netzknoten beginnt, wird dieser auch vom Encoder als LRP ausgewählt. Der OpenLR-Decoder hat diesen Start-LRP auf die andere Kreuzungsseite decodiert [blauer Kreis]. Dieser Knoten hat zwar aufgrund der weiteren Entfernung ein schlechteres Knoten-Rating, die

Tabelle 7.5: Auswertung Datensatz 'Urban' [OSM-&gt;TT]

Parameter		OpenLR		Form-Matcher	
Ortsreferenzen		120	(100 %)	120	(100 %)
decodiert	$Com_{dec}$	116	(97 %)	120	(100 %)
geom. korrekt	CG	103	(86 %)	111	(93 %)
topol. korr. (Start/Ende)	CE	85	(71 %)	98	(82 %)
topol. korr. (Verlauf)	CR	85	(71 %)	98	(82 %)
korrekt zugeordnet	$r_p^+$	85	(71 %)	95	(79 %)
korrekt nicht zugeordn.	$r_n$	0	(0 %)	0	(0 %)
korrekt klassifiziert	CCR	85	(71 %)	95	(79 %)
Mittel Dist. Start/Endpkt	$\bar{d}$	2,8 m		2,0 m	
Mittel Längendifferenz	$\bar{\Delta}l$	2,2 m		-0,2 m	

Straßenklasse und vor allem aber das Form of Way passen bei der zugehörigen Kante besser, als bei der eigentlich der Ausgangslocation entsprechenden. Gesucht wird eine Kante ohne bauliche Trennung zwischen den Richtungsfahrbahnen, also mit FOW = Single Carriageway. Da die Straße aber in beiden Karten getrennt erfasst ist, handelt es sich vermutlich um einen Kartenfehler in der Quellkarte bzw. eine nicht korrekte Schema-Zuordnung für dieses Attribut an dieser Stelle. Da auch die Straßenklasse der korrespondierenden Kante im Zielnetz von der gesuchten abweicht, kann sie die Gesamtbewertung nicht gewinnen. Der von OpenLR tatsächlich ausgewählte Location-Beginn ist die erste Kante in dieser Umgebung mit FOW = Single Carriageway. Aufgrund des ausgewogeneren Bewertungsverfahrens kann der Form-Matcher die vorhandenen Kartenunterschiede an dieser Stelle ausgleichen und den LRP am Locationbeginn korrekt decodieren [gelber Kreis].

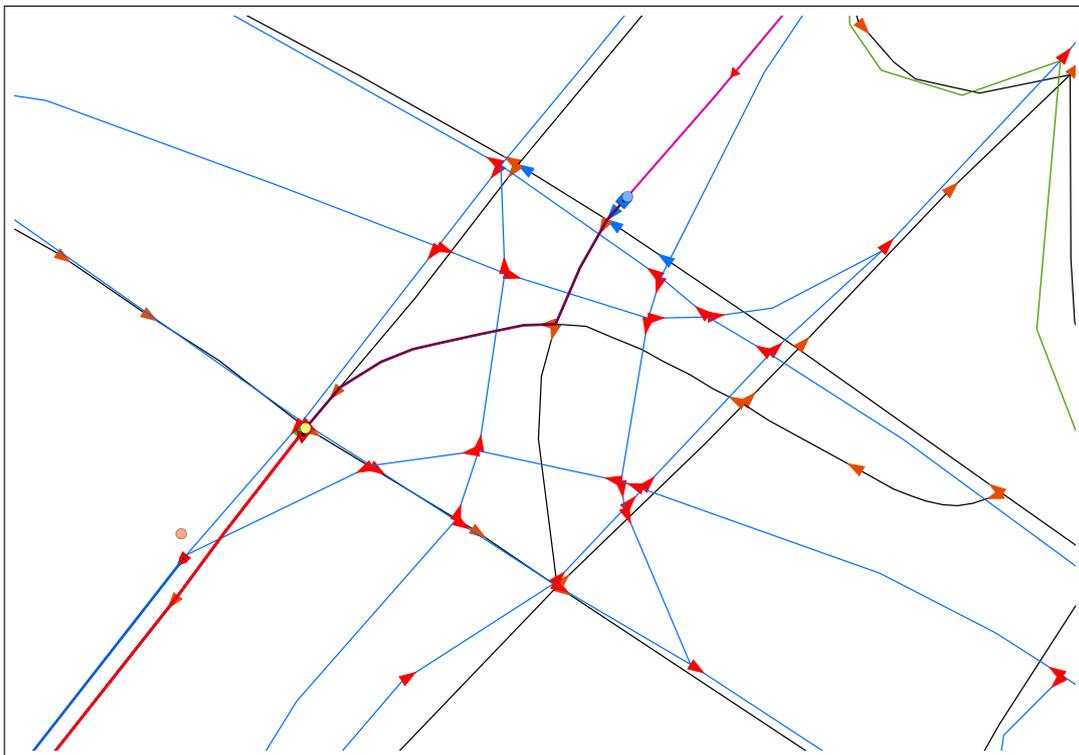


Abbildung 7.8: Beispiel korrekte Zuordnung an komplexer Kreuzung

Dass der Form-Matcher, verglichen mit der OpenLR Implementierung, auch in diesem Datensatz eine deutlich höhere Zahl von Ortsreferenzen korrekt im Zielnetz zuordnen konnte, kann auf die Verbesserungen bei der Identifikation echter Knoten sowie das Einbeziehen der Location Form in die Gesamtbewertung zurückgeführt werden. Für insgesamt 4 Ortsreferenzen konnte OpenLR keine Lösung im Zielnetz finden, da in diesen Fällen für mindestens einen Location Reference Point kein Zuordnungskandidat im Zielnetz gefunden werden konnte, der eine zufriedenstellende Bewertung erzielt hätte (vgl. Kapitel 3.4.2.2). Der Form-Matcher konnte für diese Fälle die korrekte Entsprechung im Zielnetz identifizieren und zuordnen, da hier der Fokus nicht auf einer strengen Vorauswahl für jeden einzelnen LRP, sondern mehr auf der Gesamtbetrachtung des möglichen Location-Verlaufs im Zielnetz liegt.

### 7.3.4 Szenario: Kreuzungen

Für das Szenario *Kreuzungen* wurden 82 Ortsreferenzen an und um Kreuzungssituationen und Kreisverkehren im Testgebiet erstellt. Die Ortsreferenzen sind nicht repräsentativ für das gesamte Straßenbild, sondern wurden speziell zum Testen bestimmter und erwartungsgemäß für eine Location Referencing Übertragung schwieriger Situationen ausgewählt.

Bei der Betrachtung der numerischen Evaluierungsergebnisse in Tabelle 7.6 ist die große Steigerung der Zuordnungskorrektheit des Form-Matchers gegenüber der OpenLR Implementierung auffällig. Die gute geometrische Korrektheit von OpenLR konnte vom Form-Matching Verfahren auf sehr gute 90 % gesteigert werden. Vor allem konnte aber die topologische Bewertung der Start-/Endkante im Vergleich zu OpenLR erheblich auf 80 % verbessert werden.

Tabelle 7.6: Auswertung Datensatz 'Kreuzungen' [OSM->TT]

Parameter		OpenLR	Form-Matcher
Ortsreferenzen		82 (100 %)	82 (100 %)
decodiert	$Com_{dec}$	81 (99 %)	82 (100 %)
geom. korrekt	CG	67 (82 %)	74 (90 %)
topol. korr. (Start/Ende)	CE	50 (61 %)	66 (80 %)
topol. korr. (Verlauf)	CR	50 (61 %)	66 (80 %)
korrekt zugeordnet	$r_p^+$	47 (57 %)	63 (77 %)
korrekt nicht zugeordn.	$r_n$	0 (0 %)	0 (0 %)
korrekt klassifiziert	CCR	47 (57 %)	63 (77 %)
Mittel Dist. Start/Endpkt	$\bar{d}$	6,1 m	2,9 m
Mittel Längendifferenz	$\bar{\Delta}l$	2,7 m	2,9 m

Die große Steigerung bei der topologischen Bewertung der Ortsreferenzen rührt daher, dass der Form-Matcher offenbar die ausgewogenere Bewertung und Auswahl der Kandidaten für seine Zuordnungsentscheidungen verwendet. So kann die Auswahl von Kanten in der Regel vermieden werden, die lediglich in einer Teilbewertung, dort aber in besonderem Maße, den Suchanforderungen entsprechen, jedoch nicht das korrekte Zuordnungsergebnis darstellen.

Die beim Form-Matcher vorgenommenen Veränderungen und Verbesserungen gegenüber OpenLR und dessen Musterimplementierung kommen offenbar besonders im Kreuzungsbereich zum tragen. In diesen Bereichen ist zu beobachten, dass sich auf engem Raum häufig viele Kanten befinden, die durch den Suchradius in die Menge der Zuordnungskandidaten aufgenommen werden. Durch die von OpenLR verwendete, individuelle Betrachtung jedes einzelnen LRPs anhand der OpenLR Location Reference

Eigenschaften kommt es in diesen engen Kreuzungssituationen häufig zu falschen Zuordnungsentscheidungen. Dies soll im Folgenden anhand einiger Beispiele erläutert werden.

Abbildung 7.9(a) zeigt eine Ortsreferenz, die in der Zielkarte [blau] an einem Kreuzungspunkt beginnt [roter Punkt] und Richtung Westen führt. Der Startpunkt ist auch Location Reference Point. Der OpenLR-Decoder wählt bei der Zuordnung den geometrisch näheren Knoten [blaues Quadrat], der jedoch topologisch nicht korrekt ist, da er keinen echten Kreuzungspunkt darstellt. Die beiden in der Zielkarte in Location-Richtung aufeinander folgenden Kanten stimmen in allen zur Bewertung herangezogenen Kriterien (FRC, FOW und Richtungswinkel) überein. Daher gibt für OpenLR lediglich die kürzere Entfernung zum gesuchten Start-LRP [roter Kreis] den Ausschlag für die (allerdings inkorrekte) Wahl des westlichen Startknotens. Auch in dieser Situation ist der Form-Matcher durch die zusätzliche Prüfung der Knoten-Valenz im Vorteil und ordnet den korrespondierenden Kreuzungspunkt korrekt zu [rotes Dreieck].

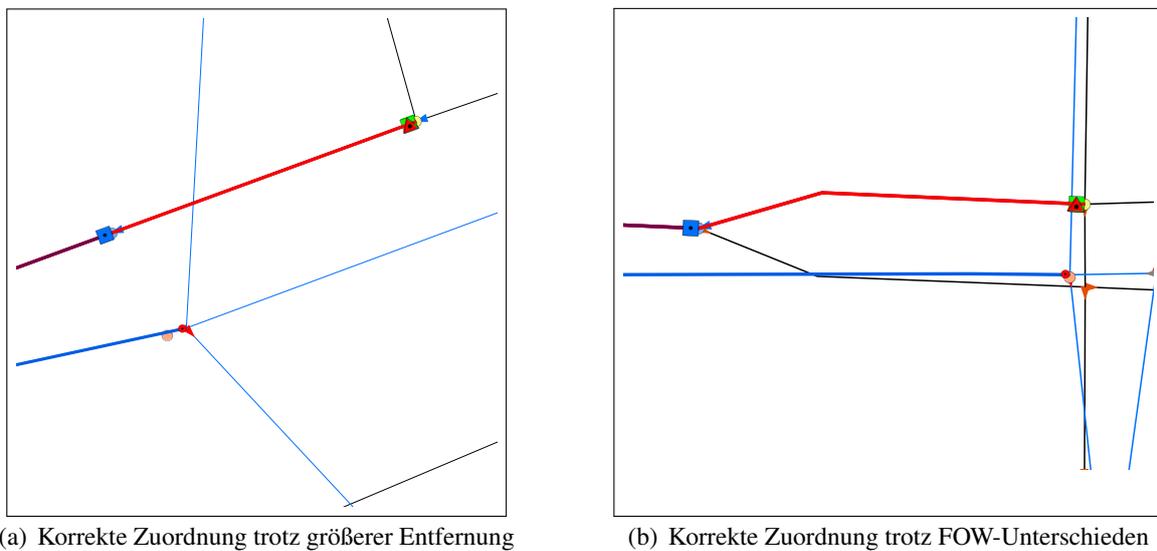


Abbildung 7.9: Beispiel korrekte Zuordnungen an Kreuzungen

Auch im folgenden Beispiel wird die Zuordnung durch unterschiedliche Modellierung der Form of Way in den beiden Karten erschwert. Abbildung 7.9(b) zeigt den Location Beginn an einer Kreuzung, deren in Ost-West-Richtung verlaufende Straße in der Quellkarte [blau] durch eine in beide Richtungen befahrbare Kante, in der Zielkarte [schwarz] jedoch für beide Fahrbahnen getrennt modelliert ist. Dadurch ergibt sich für das erste Teilstück der Ortsreferenz ein Unterschied für das Attribut FOW. Dies führt beim OpenLR Decoder zu der Auswahl der nächstgelegenen Kante, die als beidseitig befahrbar modelliert ist [blaues Quadrat]. Da im Form-Matcher die Bewertung des FOW durch die Hinzunahme weiterer Bewertungskriterien kein so hoher Stellenwert eingeräumt wird, kann hier der korrekte Startknoten im Zielnetz identifiziert werden [rotes Dreieck].

Das Beispiel in Abbildung 7.10(a) zeigt eine Situation an einem Kreisverkehr, an dessen nördlicher Ausfahrt die Ausgangslocation beginnt [fett blau]. Der Knoten ist auch gleichzeitig Location Reference Point. Der OpenLR-Decoder wählt dafür die gegenüberliegende, südliche Ausfahrt des Kreisverkehrs und führt die Ortsreferenz dann durch den Kreisverkehr Richtung Norden zum regulären Location-Verlauf. Ein Blick in die Bewertungen der einzelnen Kandidaten zeigt, dass die tatsächlich gewählte Kante eine größere Entfernung und eine größere Differenz in der Straßenklasse aufweist. Lediglich ein geringfügig besser passender Richtungswinkel kann diese Nachteile letztlich mehr als ausgleichen. Da die eigentlich korrekte Kante in der groben Bewertungsskala des Richtungswinkels (siehe Tabelle 3.3(a)) leider nur in die zweitbeste Kategorie mit einer Abweichung vom gesuchten Wert  $< 12^\circ$  fällt, erhält sie nur das halbe Rating. Dadurch kann sie in diesem Fall den Vergleich mit den umliegenden

Kanten nicht gewinnen. Der Form-Matcher ordnet den Startpunkt der Ortsreferenz hingegen korrekt dem entsprechenden Abzweig im Zielnetz [schwarz] zu [rotes Dreieck]. Vor allem durch die Berücksichtigung der formbeschreibenden Merkmale konnte in diesem Fall eine korrekte Lösung gefunden werden.

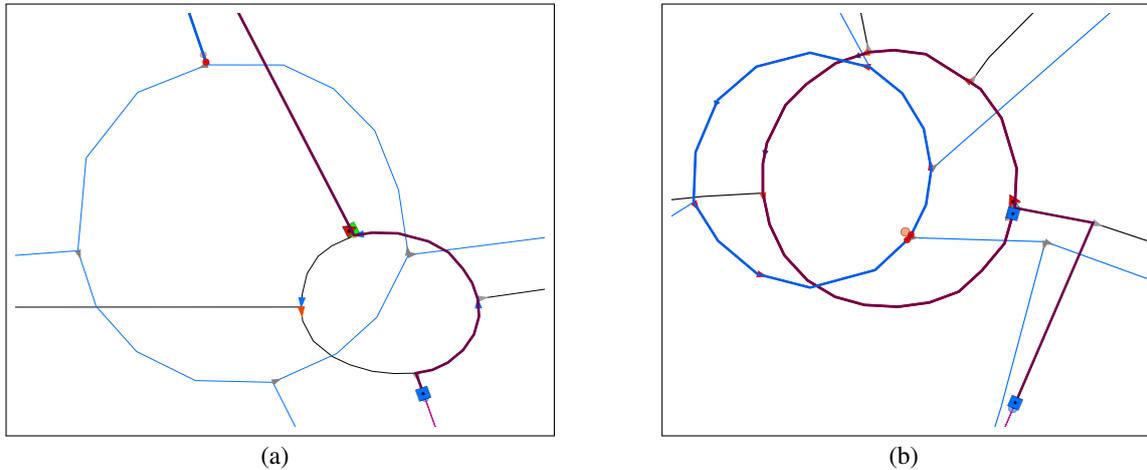


Abbildung 7.10: Beispiel korrekte formbasierte Zuordnungen an Kreisverkehren

Das Beispiel in Abbildung 7.10(b) zeigt eine Situation an einem anderen Kreisverkehr, in der die Ausgangs-Location am selben Knoten eines Kreisverkehrs beginnt und endet [roter Kreis] und diesen einmal umrundet [fett blaue Linie]. Der OpenLR-Decoder projiziert den Startknoten fälschlicherweise auf eine Kante außerhalb des Kreisverkehrs, deren Richtungswinkel dem gesuchten Wert sehr nahe kommt [blauer Kreis und blaues Quadrat, überlagert]. Dadurch enthält der decodierte Locationverlauf am Beginn zwei zusätzliche Kanten, bis die Route den Kreisverkehr erreicht und zum korrekten Endpunkt weiterführt. Dank seiner Form-Betrachtung ordnet der Form-Matcher Start- und Endknoten der Ortsreferenz korrekt zu [rotes Dreieck, teilweise überlagert] und wählt den korrekten Verlauf der Ortsreferenz im Zielnetz.

Das nächste Beispiel in Abbildung 7.11 zeigt einen durch OpenLR komplett falsch decodierten Routenverlauf [fett schwarze Linien, blaue Quadrate]. Die Ausgangs-Location beginnt an der westlichen Einfahrt des nördlichen Kreisverkehrs [fett blau, roter Kreis]. Die Ortsreferenz führt durch diesen hindurch und erstreckt sich weiter Richtung Osten bis zu einer Kreuzung [roter Punkt]. Beginn und Ende der Ortsreferenz liegen direkt auf dem jeweiligen Location Reference Point. Kurioserweise decodiert OpenLR den ersten LRP auf das westliche Ende eines Kreisverkehrs, der ca. 80 m südlich des eigentlich gemeinten Kreisverkehrs liegt [blauer Kreis]. Der Grund hierfür ist, dass dieser Kandidat sowohl bezüglich FRC als auch dem Richtungswinkel besser zu den gesuchten Eigenschaften passt und dadurch trotz der schlechteren Entfernungsbewertung das Rating gewinnt. Der Endpunkt wurde auf einen Kreuzungspunkt zwischen den beiden Kreisverkehren decodiert. Damit das Location-Ende auf diesem zu liegen kommen kann, wird die entsprechende, in Nord-Süd-Richtung verlaufende Kante im Zielnetz zweimal, also in jede Richtung einmal durchlaufen. Der Form-Matcher kann in diesem Beispiel mehrere seiner Verbesserungen ausspielen. Aufgrund des dynamischen Suchradius zur Ermittlung potentieller Decoding-Kandidaten (vgl. Kapitel 4.3.1) werden im Vergleich zu OpenLR Kandidaten in einem kleineren Umkreis von 50 m überhaupt erst für die weitere Untersuchung gesucht. Des Weiteren hilft die formbasierte Bewertung vermeintliche Lösungen mit besser passenden Attributwerten aufgrund ihrer Formunterschiede als Lösungsmöglichkeit auszuschließen. Im konkreten Beispiel wurde der korrekte Startknoten [rote Dreieck, gelber Kreis] und Location-Verlauf gewählt. Der Endknoten wurde jedoch aufgrund der geringeren Entfernung auf eine Kreuzung nördlich des korrekten Location-Endes im Zielnetz [grünes Quadrat] decodiert.

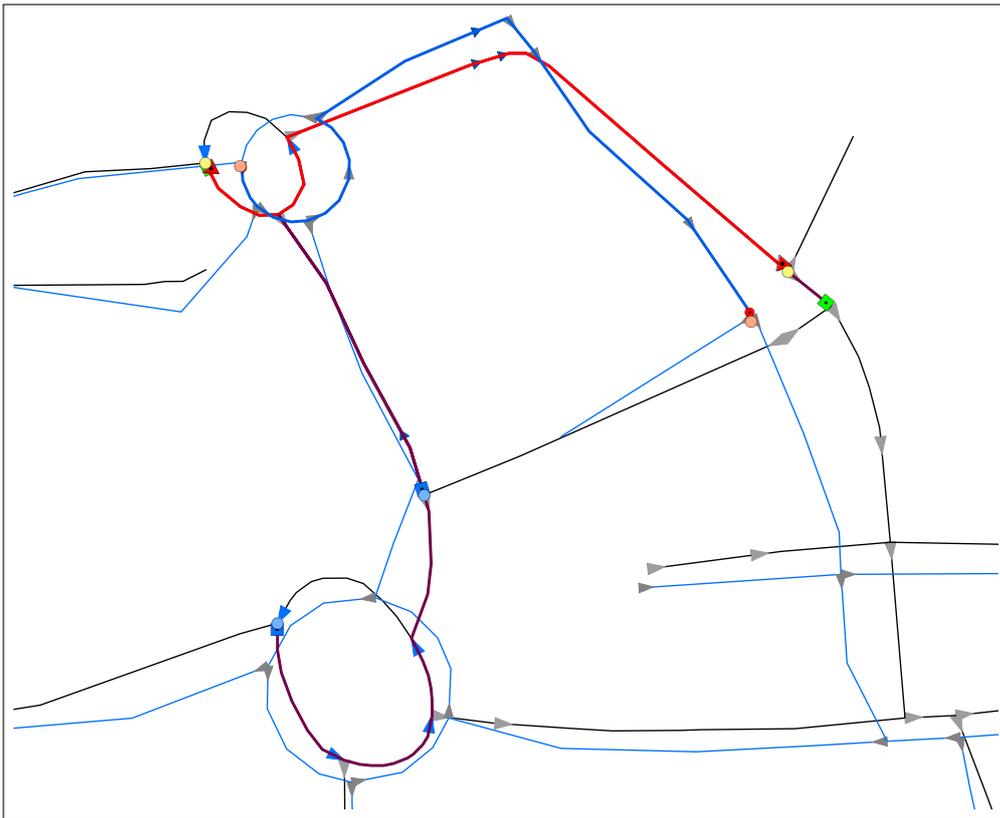


Abbildung 7.11: Beispiel korrekter Routenverlauf mit Form-Matching

Die große Steigerung der topologischen Korrektheit an Kreuzungen durch den Form-Matcher kann also auf die Zuordnungsschwächen von OpenLR in diesen Bereichen zurückgeführt werden, die der Form-Matcher im Wesentlichen durch die Hinzunahme der Knotenvalenz und der Formbetrachtung ausgleichen kann. Dadurch steigt seine Performanz auf 80 % topologischer Korrektheit im Vergleich zu OpenLR (61 %) deutlich.

Die gute topologische Zuordnungskorrektheit des Form-Matchers äußert sich auch in sehr guten Ergebnissen für die geometrische Korrektheit (90 %). Somit wurden bei lediglich 10 % der vom Form-Matcher zugeordneten Ortsreferenzen die geometrischen Anforderungen nicht eingehalten. Aufgrund der kleinräumigen Gliederung in den Kreuzungsbereichen hat die Auswahl nebenliegender und dadurch topologisch inkorrekturer Kanten lediglich eine geringe Auswirkung auf die geometrischen Abweichungen. Damit ist die im Vergleich zur topologischen Korrektheit (61 %) mit 82 % relativ gute geometrische Korrektheit von OpenLR zu erklären.

## 7.4 Diskussion der Ergebnisse

### 7.4.1 Performanz von OpenLR und Form-Matcher im Vergleich

Bei der Betrachtung der in Tabelle 7.7 zusammengestellten Gesamtergebnisse der Zuordnungen von OpenLR und Form-Matcher wird deutlich, dass der Form-Matcher die Performanz im Vergleich zu OpenLR in allen getesteten Szenarien steigern konnte. Über alle verwendeten Testdatensätze konnte der Form-Matcher die Anzahl der korrekt zugeordneten Ortsreferenzen um durchschnittlich 12 % erhöhen. In den Szenarien *Autobahn*, *Inter-Urban* und *Urban* lag die Verbesserung bei ungefähr 10 %. Auffällig ist die große Steigerung um 20 % im Datensatz *Kreuzungen*. Eine grafische Auswertung aller Detailergebnisse aller Szenarien findet sich in Abbildung 7.12. Darin ist beispielsweise zu erkennen,

dass die Korrekturklassifikationsrate (letzte Säule) des Form-Matchers (zweite Gruppe für jedes Szenario) gegenüber OpenLR (erste Gruppe jedes Szenarios) in allen Szenarien gesteigert werden konnte.

Tabelle 7.7: Gesamtergebnis (Hit Rate) [OSM->TT]

Datensatz	Anzahl	Korrekte Zuordnung (CCR)	
		OpenLR	Form-Matcher
Autobahn	64	35 (55 %)	41 (64 %)
Inter-Urban	23	16 (70 %)	18 (78 %)
Urban	120	85 (71 %)	95 (79 %)
Kreuzungen	82	47 (57 %)	63 (77 %)
GESAMT	289	183 (63 %)	217 (75 %)

Bei den empirischen Untersuchungen war insbesondere auffällig, dass die Attribute *Functional Road Class* und *Form of Way* bei OpenLR einen relativ großen Einfluss auf das Bewertungsergebnis der einzelnen Kandidaten haben und dadurch häufig inkorrekte Zuordnungsentscheidungen getroffen wurden. Der große Einfluss der beiden angesprochenen Attribute auf die OpenLR Zuordnungsentscheidung folgt direkt aus der Definition der Leistungsfunktion (vgl. Kapitel 3.4.2.2). Neben FRC und FOW wird darin nur die Entfernung und die Richtungswinkeldifferenz bewertet.

Durch die aufgezeigten Schwächen bzw. Fehlermöglichkeiten bei der Schema-Zuordnung (vor allem bei FOW) können Mehrdeutigkeiten entstehen, durch die Kandidaten-Kanten falsche Attributwerte erhalten können. Durch Fehler in der Kartenerstellung können diese Werte natürlich ebenfalls falsch erfasst oder schlicht nicht vorhanden sein (insbesondere bei OpenStreetMap). Solche Kanten stimmen dann in der Folge beim LR-Rating nicht mit den gesuchten Kriterien überein, und erhalten in der Bewertung ein entsprechend schlechtes Ergebnis.

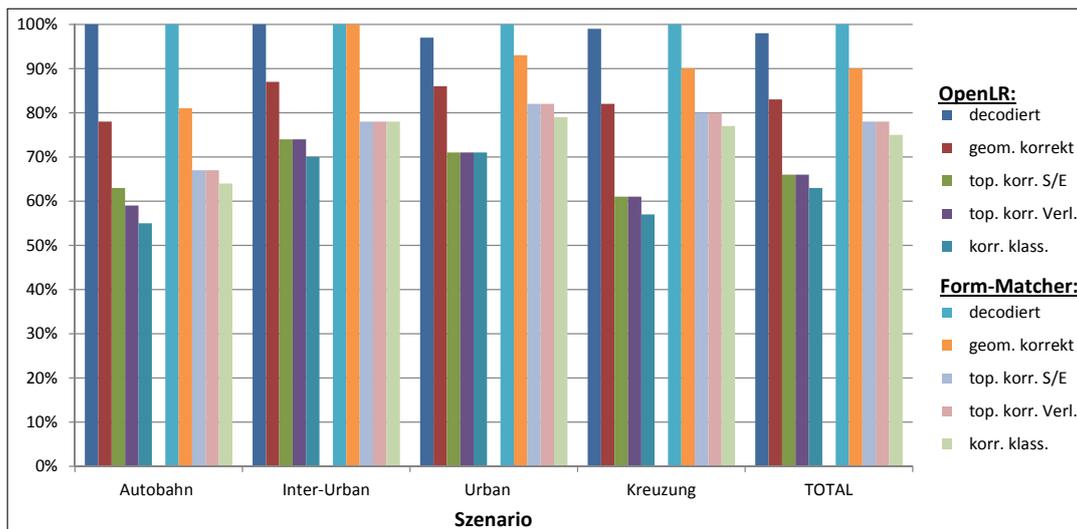


Abbildung 7.12: Gesamtergebnis OSM zu TomTom

Neben FRC und FOW basiert OpenLR auf dem Richtungswinkel und der Distanz der zu betrachtenden Kante. Vor allem das relativ grobe Bewertungsschema der Richtungswinkeldifferenzen konnte häufig als Fehlerquelle bei der Zuordnung identifiziert werden. Durch die Einteilung in Klassen führt die knappe

Überschreitung einer Klassengrenze zu einer deutlichen Abwertung einer Kante. Stimmen die Bewertungen der anderen Kriterien für zwei Kanten überein, so kann eine falsche Richtungswinkelbewertung die Auswahl der eigentlich korrekten Kante verhindern.

Durch die Betrachtung des gesamten Location-Routenverlaufs beim Form-Matcher wird verhindert, dass Routen-Kandidaten oder Teile von ihnen, deren attributive Eigenschaften zwar mit den gesuchten übereinstimmen, ihre Form aber stark von der erwarteten abweicht, die jeweilige Bewertung und somit die Wahl des Decoding-Kandidaten gewinnen können.

Auch durch die Dynamisierung des Suchradius für Decodier-Kandidaten werden in innerörtlichen Bereichen die Anzahl der Kandidaten reduziert und somit die versehentliche Auswahl weit entfernt liegender Ortsreferenzen verhindert. In überörtlichen Lagen und auf Autobahnen wird die Kandidatenmenge erhöht, was den in diesen Bereichen beobachteten geometrischen Kartenunterschieden Rechnung trägt. Dadurch kann sichergestellt werden, dass die korrekten Lösungen sehr häufig auch in der Menge der zu bewertenden Kandidaten enthalten sind. Das Erkennen von 'Wenden', also dem zweimaligen durchlaufen derselben Kante in Hin- und Rückrichtung (siehe Kapitel 4.3.4), und das Aussortieren solcher Routen-Kandidaten in der Vorauswahl leistet ebenfalls einen Beitrag zur Steigerung der Korrektheit der Decoding-Ergebnisse des Form-Matchers.

Tabelle 7.8: Anteile am Gesamtergebnis (Hit Rate) [OSM->TT]

Bewertungs-Anteil	Autobahn	Inter-Urban	Urban	Kreuzungen	GESAMT
Ortsreferenzen	64	23	120	82	289
Basis-Bewertung	30 (47 %)	14 (61 %)	80 (67 %)	54 (66 %)	178 (62 %)
Anzahl sig. Ri-Änderungen	30 (47 %)	16 (70 %)	85 (71 %)	53 (65 %)	184 (64 %)
Kum. Ri-Änderungen	29 (45 %)	14 (61 %)	89 (74 %)	52 (63 %)	184 (64 %)
Knotenvalenz	40 (63 %)	17 (74 %)	91 (76 %)	64 (78 %)	212 (73 %)
1 / 1 / 1	39 (61 %)	17 (74 %)	97 (81 %)	60 (73 %)	213 (74 %)
1 / 0 / 1	40 (63 %)	17 (74 %)	94 (78 %)	63 (77 %)	214 (74 %)
Form-Matcher (Standard)	41 (64 %)	18 (78 %)	95 (79 %)	63 (77 %)	217 (75 %)

Durch die Hinzunahme von weiteren, formbeschreibenden Faktoren für die Bewertung der Kandidaten ist der Form-Matcher weniger anfällig gegenüber Ausreißern in der Attribut-Bewertung. In Tabelle 7.8 sind die individuellen Beiträge der beim Form-Matcher neu eingeführten Teilbewertungen dargestellt. Zunächst wird eine Zuordnung mittels *Basis-Bewertung* angegeben. Dies entspricht der neu entwickelten Kandidaten-Vorauswahl sowie die Anwendung der Gesamtbetrachtung der Routen-Kandidaten, jedoch ausschließlich unter Verwendung der aus OpenLR bekannten Zuordnungsparameter. Die Zuordnungskorrektheit für alle Datensätze zusammen liegt mit 62 % auf dem Niveau von OpenLR (vgl. Tabelle 7.7), unterscheidet sich jedoch in den einzelnen Datensätzen. Durch die Hinzunahme der Anzahl signifikanter Richtungsänderungen bzw. der kumulierten Richtungsänderungen als weitere Zuordnungskriterien lässt sich die Anzahl korrekt zugeordneter Ortsreferenzen in einzelnen Fällen steigern. Es wird aber deutlich, dass die gesteigerte Zuordnungskorrektheit des neuen Verfahrens vor allem durch die Aufwertung der echten Netzknoten im Vergleich zu bivalenten Knoten erreicht wird. Dies ist ein entscheidender Vorteil. Dadurch kann die sehr oft beobachtete Auswahl von zweiwertigen Knoten im Zielnetz verhindert und eine deutliche Steigerung der Zuordnungskorrektheit von 62 % auf 73 % insgesamt erzielt werden. Durch die gleichgewichtete Kombination aller drei Zuordnungsparameter (in Tabelle 7.8 mit '1 / 1 / 1' bezeichnet) mit den OpenLR-Parametern lässt sich die Anzahl korrekt zugeordneter Ortsreferenzen abermals steigern. Werden jedoch nur die Anzahl signifikanter Richtungsänderungen und die Knotenvalenz für die Zuordnung berücksichtigt ('1 / 0 / 1'), so ist für den Datensatz *Kreuzungen* eine leichte Verbesserung festzustellen, wohingegen die Ergebnisse für alle anderen Datensätze auf gleichem Niveau bleiben oder sich leicht verschlechtern. Wählt man schließlich eine Kombination,

bei der die kumulierten Richtungsänderungen lediglich eine Gewichtung von 20 % erhalten, so wird ein noch besseres Ergebnis erzielt. Diese Kombination von Parametern und Gewichten wird deshalb standardmäßig beim Form-Matcher verwendet.

## 7.4.2 Abhängigkeit von der Übertragungsrichtung

Die empirischen Location Referencing Übertragungen lassen sich auch in umgekehrter Richtung von TomTom MultiNet als Quellkarte zur OpenStreetMap als Zielkarte durchführen. Die Ergebnisse von OpenLR sind in Tabelle 7.9 und die des Form-Matchers in Tabelle 7.10 dargestellt.

Tabelle 7.9: Ergebnisse von OpenLR [TT->OSM]

Parameter	Autobahn	Inter-Urban	Urban	Kreuzung	TOTAL
Ortsreferenzen	64 (100 %)	23 (100 %)	120 (100 %)	80 (100 %)	287 (100 %)
decodiert	58 (91 %)	22 (96 %)	113 (94 %)	76 (95 %)	269 (94 %)
geom. korrekt	48 (75 %)	21 (91 %)	101 (84 %)	59 (74 %)	229 (80 %)
top. korr. S/E	48 (75 %)	20 (87 %)	98 (82 %)	57 (71 %)	223 (78 %)
top. korr. Verl.	48 (75 %)	20 (87 %)	98 (82 %)	57 (71 %)	223 (78 %)
korr. klass.	40 (63 %)	20 (87 %)	92 (77 %)	53 (66 %)	205 (71 %)

Betrachtet man zunächst die OpenLR-Ergebnisse für diese Konstellation in Tabelle 7.9, so zeigt sich, dass die Anzahl der korrekt klassifizierten Ortsreferenzen in allen vier Testdatensätzen in der Größenordnung von ungefähr 10 % über den Ergebnissen der Referenzierung von OSM zu TomTom liegen. Die relativen Unterschiede zwischen den einzelnen Szenarien in der vorangegangenen Auswertung (vgl. Tabelle 7.7) finden sich auch in diesem Ergebnissatz wieder. So zeigt der *Autobahn*-Datensatz mit 63 % die niedrigste Korrektheit der Zuordnungen. Ähnlich wie in der vorangegangenen Auswertung liegt die Performanz des Datensatzes *Kreuzungen* mit 66 % korrekt zugeordneten Ortsreferenzen in einer vergleichbaren Größenordnung. Die Referenzierungen im *Inter-Urban* Bereich waren dagegen mit Abstand am besten (87 %). Die innerörtlichen Ergebnisse liegen mit 77 % korrekten Zuordnungen 10 % unter dem Resultat des *Inter-Urban* Datensatzes. Diese beiden Datensätze lieferten in der vorangegangenen Auswertung (OSM->TomTom) noch vergleichbare Ergebnisse.

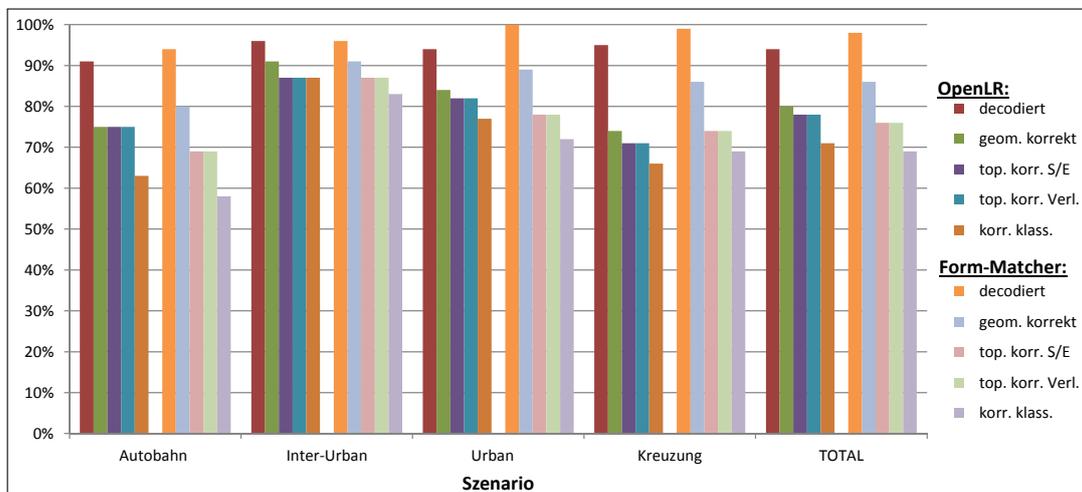


Abbildung 7.13: Gesamtergebnis TomTom zu OSM

OpenLR scheint also insgesamt besser zu funktionieren, wenn das Datenbank-Schema des Quellnetzes sehr gut zu dem der LR-Austauschschicht passt. Dies ist der einzig auffällige Unterschied zu der vorangegangenen Location Referencing Konstellation.

Tabelle 7.10: Ergebnisse des Form-Matcher [TT-&gt;OSM]

Parameter	Autobahn	Inter-Urban	Urban	Kreuzung	TOTAL
Ortsreferenzen decodiert	64 (100 %)	23 (100 %)	120 (100 %)	80 (100 %)	287 (100 %)
geom. korrekt	60 (94 %)	22 (96 %)	120 (100 %)	79 (99 %)	281 (98 %)
top. korr. S/E	51 (80 %)	21 (91 %)	107 (89 %)	69 (86 %)	248 (86 %)
top. korr. Verl.	44 (69 %)	20 (87 %)	94 (78 %)	59 (74 %)	217 (76 %)
korr. klass.	44 (69 %)	20 (87 %)	94 (78 %)	59 (74 %)	217 (76 %)
	37 (58 %)	19 (83 %)	86 (72 %)	55 (69 %)	197 (69 %)

Die Ergebnisse des *Form-Matchers* in Tabelle 7.10 liegen bei der Karten-Konstellation TomTom->OSM liegen in etwa auf gleichem Niveau, wie die der Konstellation OSM->TomTom (vgl. Tabelle 7.7). Das Szenario *Inter-Urban* zeigt mit 83 % korrekten Zuordnungen die besten Ergebnisse, die Szenarien *Urban* und *Kreuzungen* liegen beide bei ungefähr 70 % korrekten Zuordnungen. Diese beiden Datensätze erzielten bei der Location Referencing Übertragung von OSM zu TomTom noch um etwa 10 % höhere Korrektheitsraten. Die Ergebnisse für den Datensatz *Autobahn* liegen mit 58 % korrekten Zuordnungen wie schon bei der vorigen Auswertung am niedrigsten.

Abbildung 7.13 stellt die Ergebnisse beider Verfahren für diese Übertragungsrichtung nochmals grafisch gegenüber. Vergleicht man nun die Performanz des Form-Matchers mit derjenigen des OpenLR-Verfahrens, so fällt zunächst auf, dass der Form-Matcher mit einer Vollständigkeit von 98 % fast alle Ortsreferenzen im Zielnetz decodieren konnte, OpenLR erreicht hier nur eine Rate von 94 %. Auch bei der geometrischen Korrektheit kann der Form-Matcher beim Szenario *Inter-Urban* gleich hohe und bei den anderen Szenarien verglichen mit OpenLR klar verbesserte Ergebnisse liefern. Die topologische Korrektheit, sowohl der Anfangs- und Endpunkte, also auch des Location-Verlaufs liegen beim Szenario *Inter-Urban* gleichauf mit OpenLR, bei den *Kreuzungen* konnten sie leicht verbessert werden. Bei den Datensätzen *Autobahn* und *Urban* liefert der Form-Matcher jedoch leicht verschlechterte topologische Korrektheiten.

### 7.4.3 Untersuchung Codierung und Decodierung auf einer Karte

Neben den bislang beschriebenen, an reale Übertragungsszenarien angelehnte Untersuchungen, bietet es sich auch an, testweise Location Referencing Übertragungen auf der selben Karten zu decodieren, auf der sie auch codiert wurden. Dies wurde für alle im Rahmen der vorliegenden Arbeit verwendeten Datensätze durchgeführt. Die Ergebnisse sind mit den ermittelten Korrektklassifikationsraten (CCR) in Tabelle 7.11 zusammengestellt.

Beim Betrachten der Ergebnisse ist zu erkennen, dass die Korrektklassifikationsraten insgesamt auf einem sehr hohen Niveau liegen. Intuitiv wären aber Ergebnisse von 100 % zu erwarten gewesen. Auffällig ist, dass die Ergebnisse der Referenzierungen von/nach OSM eine leicht höhere Korrektheit aufweisen, als solche von/nach TomTom MultiNet. Innerhalb der Szenarien weist der Datensatz *Urban* die weitaus beste Korrektheit auf, aber selbst der am schlechtesten bewertete Datensatz *Autobahn* erreicht noch Korrektheitsraten von knapp 90 %. Zwischen OpenLR und Form-Matcher ist eine in etwa gleiche Leistungsfähigkeit zu beobachten.

Eine detailliertere Analyse der Ergebnisse hat gezeigt, dass beide Verfahren bis auf wenige Ausnahmen alle LRPs korrekt zuordnen und insofern die erwartete Zuordnungskorrektheit von 100 % erfüllt

Tabelle 7.11: Ergebnisse (CCR) 'gleiche Karten'

Datensatz	Anzahl	OpenLR		Form-Matcher	
		OSM->OSM	TT->TT	OSM->OSM	TT->TT
Autobahn	64	57 (89 %)	56 (88 %)	56 (88 %)	56 (88 %)
Inter-Urban	23	21 (91 %)	19 (83 %)	20 (87 %)	19 (83 %)
Urban	120	118 (98 %)	118 (98 %)	115 (96 %)	112 (93 %)
Kreuzungen	82	78 (95 %)	71 (88 %)	78 (95 %)	66 (81 %)
GESAMT	289	274 (95 %)	264 (91 %)	269 (93 %)	253 (88 %)

wäre. Die oben genannten Abweichungen traten in der Auswertung der tatsächlich decodierten und somit um evtl. vorhandene Offsets am Beginn / Ende korrigierte Ortsreferenzen bzw. bei zum nächsten echten Netzknoten erweiterten Ortsreferenzen auf. Diese Offsets konnten als Ursache für die in Tabelle 7.11 genannten Abweichungen identifiziert werden. Entsprechend dem OpenLR-Whitepaper TOMTOM (2012) werden Offsets im physikalischen OpenLR-Datenformat prozentual zur Routenlänge zwischen den ersten bzw. letzten beiden LRPs einer Ortsreferenz übertragen. Durch die physikalische Darstellung dieses Prozentwertes durch 1 Byte (= 8 Bit) können lediglich 256 Zustände übertragen werden. Daraus folgt eine Auflösung von 0,390625 %. Diese Auflösung der Offset-Übertragung führt dazu, dass es bei langen Ortsreferenzen und selbst bei korrekter LRP-Zuordnung zu geometrischen Versätzen und in deren Folge auch zu topologischen Abweichungen zwischen Codierter und Decodierter Ortsreferenz kommen kann. Dabei bleiben die angesprochenen geometrischen Abweichungen, bedingt durch die Offset-Auflösung, meist innerhalb der im Rahmen der vorliegenden Arbeit gewählten Akzeptanzgrenzen (vgl. Kapitel 6.2.3). Topologische Inkorrektheiten treten meist dann auf, wenn der durch den Offset eingerechnete Start- oder Endpunkt einer Ortsreferenz, bedingt durch die geometrische Abweichung, auf einer angrenzenden Kante zu liegen kommt. Damit lässt sich auch der zunächst aus den numerischen Ergebnissen gewonnene Eindruck bestätigen, dass die Abweichungen vermehrt bei den Datensätzen *Autobahn* und *Inter-Urban* auftreten, deren Ortsreferenzen deutlich länger sind, als in den innerstädtischen Datensätzen, die sich durch relativ kurze Ortsreferenzen auszeichnen.

#### 7.4.4 Signifikanz der OpenLR Zuordnungs-Attribute

Die OpenLR-Zuordnungsentscheidung wird wie in Kapitel 3.4.2.2 ausführlicher beschrieben, nach einer erfolgten Vorauswahl der Kandidaten schließlich aufgrund einer Bewertungsfunktion getroffen. Diese ist als Leistungsfunktion ausgelegt. Für jeden gesuchten Location Reference Point wird derjenige Kandidaten-Knoten mit abgehender Kante ausgewählt, der die höchste Bewertung (Leistungsfunktionswert) erreicht. Die Bewertung der Kandidaten erfolgt dabei anhand ihrer *Punktdistanz* bzw. der Differenz mit den mittels OpenLR übertragenen Eigenschaften *Richtungswinkel*, *Functional Road Class* und *Form of Way*. Die beiden letztgenannten Werte sind attributive Eigenschaften und die gesuchten Werte beziehen sich auf die jeweilige LRP-Kante aus dem Quellnetz. Durch die Schema-Zuordnung zwischen den einzelnen Kartensystemen (vgl. Kapitel 4.2) entstehen jedoch Unsicherheiten bei der Übersetzung der jeweiligen Schemata. Dies hat auch bei den im Rahmen dieser Arbeit durchgeführten Tests für Probleme bei Zuordnungen in Form von falschen oder gar unmöglichen Zuordnungen geführt.

Daher soll im Folgenden untersucht werden, wie groß der Einfluss der beiden Attribute FRC und FOW tatsächlich auf das OpenLR Zuordnungsergebnis ist, und ob eine Zuordnung möglicherweise auch ohne diese semantischen Informationen mit zufriedenstellendem Ergebnis möglich wäre. Dazu werden die Testdatensätze abwechselnd mit vollen Attributinformationen und dann ohne Berücksichtigung von FRC und FOW beim Kandidaten-Rating prozessiert. Da dadurch die Leistungsfunktion aus den verbleibenden zwei Anteilen *Distanz* und *Richtungswinkel* nicht mehr den gleichen Maximalwert einnehmen

kann, muss auch das im OpenLR Vorauswahlschritt abgeprüfte *MinimumAcceptedRating* entsprechend angepasst werden (vgl. Kapitel 3.4.2.2).

Tabelle 7.12: Korrekte Zuordnung (CCR) OpenLR mit/ohne Attributinformationen [OSM->TT]

Datensatz	OpenLR	ohne FRC/FOW
Autobahn	35 (55 %)	33 (52 %)
Inter-Urban	16 (70 %)	16 (70 %)
Urban	85 (71 %)	87 (73 %)
Kreuzungen	47 (57 %)	39 (48 %)
GESAMT	183 (63 %)	175 (61 %)

Die numerischen Ergebnisse sind in Tabelle 7.12 zusammengestellt. Es zeigt sich, dass beim Datensatz *Autobahn* eine leichte Verschlechterung der Zuordnungskorrektheit festzustellen ist, wenn FRC und FOW nicht für die Zuordnung herangezogen werden. Beim Datensatz *Inter-Urban* bleibt die Zahl der korrekten Zuordnungen gleich. Im durch den Datensatz *Urban* abgedeckten innerstädtischen Bereich erhöht sich die Zahl der korrekten Zuordnungen leicht. Die Auswertung des Datensatzes *Kreuzungen* zeigt einen deutlichen Rückgang der korrekt zugeordneten Ortsreferenzen. Insgesamt lässt sich also sagen, dass für die außerstädtischen Szenarien keine Veränderung der Zuordnung festgestellt werden kann. Im urbanen Umfeld hingegen zeigt sich, dass ein gewisser Anteil der Ortsreferenzen ohne Hinzunahme von FRC und FOW gar nicht mehr zugeordnet werden kann. Überraschenderweise zeigen die Ergebnisse aber auch, dass ein großer Anteil der Ortsreferenzen lediglich durch Betrachtung von Distanz und Richtungswinkel korrekt zugeordnet werden können. Eine vielleicht aus den im vorigen Kapitel analysierten Zuordnungsergebnissen erwartete Verbesserung der Zuordnungsleistung durch Weglassen der fehleranfälligen Attribute FRC und FOW konnte bei dieser Untersuchung aber nicht beobachtet werden.

#### 7.4.5 Qualitätsermittlung ohne SOLL-Werte

Im Rahmen der vorliegenden Arbeit werden die Zuordnungsergebnisse durch Vergleich mit den zu erwartenden Zuordnungsergebnissen im Zielnetz (=Soll) evaluiert. Dieses Vorgehen bietet eine hohe Sicherheit bei der Bewertung der Zuordnungsverfahren, stellt jedoch hohe Anforderungen an das Auswertesystem.

Ähnliche Evaluierungen sind auch für reale Location Referencing Anwendungen hilfreich, jedoch ist eine Bereitstellung von Soll-Lösungen im Zielnetz sehr aufwändig und kostenintensiv. Daher soll mit dieser Untersuchung herausgefunden werden, ob andere mögliche Zuordnungsverfahren eine Lösung mit übergeordneter Genauigkeit liefern können, die dann als Vergleichswert für die Location Referencing Ergebnisse dienen könnten.

Für eine solche Auswerte-Umgebung müssen jedoch, ähnlich wie in der im Rahmen dieser Arbeit verwendeten LR-Testumgebung, die Quell- und Zielkarten und die jeweiligen Ortsreferenzen verfügbar sein. Durch den direkten Zugriff auf die Karten und die auf deren Knoten und Kanten referenzierte Ortsreferenzen liegen sämtliche geometrischen, topologischen und semantischen Eigenschaften der von den Ortsreferenzen abgedeckten Bereiche im gleichen Auswertesystem vor. Damit stehen in einem solchen Auswertesystem wesentlich mehr Informationen zur Verfügung, als in einer realen Location Referencing Anwendung.

#### 7.4.5.1 Untersuchung Zuordnung mittels Frechét Distanz

Die diskrete Frechét-Distanz wurde in Kapitel 4.7.1 als geometrisches Zuordnungs- bzw. Ähnlichkeitsmaß vorgestellt. Im Rahmen der LR-Testumgebung wurde dann ein auf der Frechét-Distanz basierendes Zuordnungsmaß wie in Kapitel 4.7.2 vorgeschlagen, implementiert. Das Bewertungsmaß setzt sich aus von OpenLR übernommenen Parametern FRC, FOW und Richtungswinkel sowie den Erweiterungen des Form-Matchers zusammen. Anstelle der Punktdistanz aus OpenLR wurde jedoch eine Bewertung auf Grundlage der diskreten Frechét-Distanz eingeführt, siehe Gleichungen 4.18 und 4.20. Die Ergebnisse sind in Tabelle 7.13 den beiden anderen Zuordnungsverfahren gegenübergestellt.

Tabelle 7.13: Vergleich Zuordnung mittels Frechét-Distanz [OSM->TT]

Datensatz	Ortsreferenzen	Zuordnung korr. Loc.-Verlauf		
		OpenLR	Form-Matcher	Form+Frechét
Autobahn	64	38 (59 %)	43 (67 %)	37 (58 %)
Inter-Urban	23	17 (74 %)	18 (78 %)	14 (61 %)
Urban	120	85 (71 %)	98 (82 %)	99 (83 %)
Kreuzungen	82	50 (61 %)	66 (80 %)	64 (78 %)
GESAMT	289	190 (66 %)	225 (78 %)	214 (74 %)

Aus den numerischen Ergebnissen ist zu erkennen, dass die Zuordnungsraten mit Hilfe der Frechét-Distanz in den urbanen Datensätzen (einschl. Kreuzungen) in etwa gleich auf mit dem Form-Matcher liegen. Bei den Datensätzen *Autobahn* und *Inter-Urban* liegen die Resultate mit 58 % bzw. 61 % jedoch deutlich niedriger.

Eine detailliertere Auswertung zeigt, dass die Frechét-Distanz basierte Zuordnungslösung immer eine Kandidaten-Route im Zielnetz auswählt, die die korrekte Route zumindest enthält. In den Fällen nicht korrekter Zuordnung werden jedoch Routen gewählt, die am Beginn und/oder am Ende eine oder mehrere Kanten zusätzlich zur korrekten Lösung enthalten und deshalb in der Qualitätsbewertung als nicht korrekt klassifiziert wurden.

Zum zweiten ist auffällig, dass sehr häufig mehrere Routen-Kandidaten die gleiche Frechét-Distanz zur ursprünglichen Ortsreferenz aufweisen. Wie zuvor beschrieben, enthalten alle diese Kandidaten den korrekten Routenverlauf, jedoch meist mit zusätzlichen Kanten zu Beginn und/oder am Ende. Man kann also sagen, dass die Frechét-Distanz ein gutes Maß zur Identifikation des Routenverlaufs einer Ortsreferenz im Zielnetz ist. Sie liefert jedoch nicht genügend Unterscheidungsmerkmale, um im Vergleich zum Form-Matcher höhere Korrektheitsraten zu erzielen.

#### 7.4.5.2 Bewertung mittels Durchschnitts-Rating

Die Bewertungsfunktionen von Zuordnungsverfahren stellen ein Maß für die Ähnlichkeit zweier zu vergleichenden Ortsreferenzen dar. Je höher also der Funktionswert, desto ähnlicher sind sich die beiden Routenverläufe im Hinblick auf die untersuchten Parameter.

Für die reale Anwendung kann es sinnvoll sein, solche Angaben zur Sicherheit der Zuordnung auch an den "Nutzer"/ die Anwendung der OpenLR-Zuordnung als Qualitätsinformation zugänglich zu machen. Für die Bestimmung dieses Qualitätsmaßes sind keine über die mittels LR-Verfahren übertragenen Daten hinausgehenden Informationen notwendig. Insbesondere ist keine Kenntnis der Original-Location im Quellnetz sowie vor allem auch der Quellkarte notwendig. Das Verfahren eignet sich daher zur Abschätzung der Übertragungsqualität in Echtzeit für jede Ortsreferenz direkt in der decodierenden

Applikation/Anwendungssystem. Es kann aber auch ein Mittelwert der Bewertungen/Ratings für einen ganzen Datensatz gebildet werden.

Jedoch ist die Aussagekraft solcher Mittelwerte der Decoder-Ratings begrenzt. So sind die Werte immer nur innerhalb eines Location Referencing Verfahrens vergleichbar, da die jeweiligen Bewertungsfunktionen voneinander abweichen. Das beispielsweise im Rahmen dieser Arbeit berechnete OpenLR-Rating kann nicht mit demjenigen des Form-Matchings verglichen werden. Ein Mittelwert für eine Gruppe von Ortsreferenzen oder eines Datensatzes kann jedoch mit denen anderer Datensätze verglichen werden und es ließe sich daraus ein Maß für die relative Güte der Zuordnung in den verschiedenen Szenarien ableiten.

Im Folgenden soll daher, ausgehend von den vorliegenden Datensätzen, die Frage beantwortet werden, ob eine Auswertung der Bewertungsergebnisse aller in einem Datensatz decodierten Ortsreferenzen (also die 'Gewinner' des jeweiligen Vergleichs) ein Maß für die Performanz eines Zuordnungsverfahrens für einen bestimmten Datensatz liefern kann. Zu diesem Zwecke wurden die mittels Soll-Daten ermittelte Korrektheitsrate den jeweiligen mittleren Ratings, die ohne Soll-Daten ermittelt wurden, gegenübergestellt. Die Ergebnisse dieser Auswertung der im Rahmen der vorliegenden Arbeit durchgeführten Location Referencing Übertragungen sind in Tabelle 7.14 dargestellt.

Tabelle 7.14: Vergleich Mittelwert Ratings [OSM->TT]

Datensatz	Anzahl	OpenLR		Form-Matching	
		Korrektheit	Ø Rating	Korrektheit	Ø Rating
Autobahn	64	59%	1108	67%	827
Inter-Urban	23	74%	1117	78%	843
Urban	120	71%	1009	82%	827
Kreuzungen	82	61%	1027	80%	817

Aus den Ergebnissen dieser empirischen Untersuchungen ist abzulesen, dass für beide untersuchte LR-Verfahren (OpenLR und Form-Matcher) kein Zusammenhang zwischen dem mittels Qualitätsmodell und Soll-Lösungen ermittelten Anteil der korrekt zugeordneten Ortsreferenzen und dem durchschnittlichen Rating aus dem Zuordnungsverfahren für diesen Datensatz existiert. Die OpenLR Zuordnungen erzielten bei den Datensätzen *Autobahn* und *Inter-Urban* ein durchschnittlich um ca. 100 Bewertungspunkte höheres Rating. Dem steht jedoch kein erkennbarer Anstieg der Zuordnungskorrektheit gegenüber, beim Datensatz *Autobahn* ist die Korrektheit sogar geringer als bei den anderen Datensätzen. Der Form-Matcher erzielt lediglich für den Datensatz *Inter-Urban* eine leicht höhere durchschnittliche Bewertung. Diese schlägt sich ebenfalls nicht in einer höheren Zuordnungskorrektheit nieder. Dadurch kann geschlussfolgert werden, dass aus den für die Zuordnungen berechneten Bewertungsmaßen keine Rückschlüsse auf die Güte bzw. Korrektheit der Zuordnungen gezogen werden können.



## 8 Zusammenfassung und Ausblick

Im Rahmen der vorliegenden Arbeit wurde eine Zusammenstellung bestehender Zuordnungsansätze aus der Geoinformatik und der dabei zum Einsatz kommenden Zuordnungsmaße als Grundlage für die weitere Entwicklung aufbereitet. Dabei wurde ersichtlich, dass nahezu alle benannten Verfahren geometrische, topologische und semantische Objekteigenschaften oder eine aus ihnen für die Zuordnung von Objekten aus unterschiedlichen Kartensystemen verwenden.

Daran anschließend folgte eine Übersicht sämtlicher dynamischer Location Referencing Verfahren und deren historische Entwicklung. Darin wurde OpenLR aufgrund seines offenen Standards und der Verfügbarkeit einer Implementierung als für diese Arbeit relevantes Verfahren identifiziert. Sowohl der OpenLR-Standard, welcher hauptsächlich das zu Grunde gelegte Datenübertragungsformat beschreibt, und die Referenzimplementierung wurden ausführlich dargestellt und einer anschließenden Analyse unterzogen. Durch die Rückführung der Location Referencing Verfahren auf ein Zuordnungsproblem war die Zusammenführung der dabei identifizierten Verbesserungsmöglichkeiten mit den zuvor beschriebenen Zuordnungsansätzen und Zuordnungsmaßen möglich. Damit wurde erstmals eine Verbindung der Zuordnungsansätze aus der Geoinformatik und Kartographie zu den bestehenden Location Referencing Verfahren hergestellt. Aus dieser gemeinsamen Sicht entstand im Rahmen der vorliegenden Arbeit ein neues, verbessertes Location Referencing Verfahren, der *Form-Matcher*.

Dieses neue Verfahren erweitert das bestehende OpenLR-Datenformat um zwei formbeschreibende Eigenschaften der Ortsreferenz. Bei der Decodierung der Ortsreferenzen im Zielnetz zeichnet es sich vor allen Dingen durch die Betrachtung vollständiger Routenverläufe der Decoding-Kandidaten aus. Dadurch wird die formbasierte Zuordnung erst möglich. Zusätzlich werden in dem neu entwickelten Verfahren die topologischen Eigenschaften der Location Reference Points beim Decodieren genutzt. Dadurch konnte die Identifikation der korrekten LRP's im Zielnetz deutlich verbessert werden. Durch verschiedene Maßnahmen, wie etwa der Dynamisierung des Suchraumes, konnte die Zahl der zu bewertenden Kandidaten im Zielnetz in innerstädtischen Gebieten reduziert und in überörtlichen Lagen erhöht werden, sodass stets eine angemessene Menge an Kandidaten für die Bewertung zur Verfügung steht.

Ausgangspunkt für die Verbesserung eines Location Referencing Verfahrens ist die Möglichkeit, die Performanz bestehender Verfahren zu bestimmen. Dazu wurden die Grundlagen der Qualitätsmodellierung aus der Geoinformatik sowie deren verkehrstechnischer Anwendungen dargelegt. Ebenso wurde ein Überblick bestehender Evaluierungsansätze für Kartenzuordnungen gegeben. Darauf aufbauend wurde ein Qualitätsmodell für die Bewertung von Kartenzuordnungen für die folgenden empirischen Untersuchungen abgeleitet und detailliert beschrieben. Maßgeblich sind dabei die Qualitätsmerkmale *Vollständigkeit*, *Genauigkeit* und *Korrektheit*. Diese wurden für die konkrete Qualitätsbewertung durch messbare Qualitätsparameter konkretisiert. Auf Objektebene sind das die Anzahl der fehlenden bzw. überzähligen Kanten, die Distanz der Start-/Endpunkte und die Längendifferenz der Ortsreferenzen. Die geometrische Korrektheit der Ortsreferenz wird durch Vergleich der Start-/Endpunktdistanzen mit einem Schwellwert ermittelt, die topologische Korrektheit durch Vergleich der Start-/Endkanten sowie des Routenverlaufs mit dem Soll-Verlauf.

Zur Durchführung empirischer Untersuchungen wurde der *Form-Matcher* als Erweiterung der bestehenden OpenLR-Implementierung konzipiert. Dadurch war auch ein direkter Vergleich der Performanz beider Verfahren möglich. Datengrundlage bildeten die Ausschnitte aus zwei digitalen Straßennetzwerken. Dies war zum einen die kommerziell verfügbare digitale Straßenkarte *TomTom MultiNet* sowie die

frei verfügbare *OpenStreetMap*, die beide im Shapefile-Format vorlagen bzw. dahin transformiert wurden. Um eine detaillierte numerische Auswertung zu ermöglichen, wurde die Software-Umgebung um eine *LR-Testumgebung* ergänzt. Dadurch war es möglich, spezifische Einblicke in die Bewertungs- und Auswahlprozesse der Location Referencing-Implementierungen zu gewinnen und diese entsprechend zu dokumentieren. Die Testumgebung ermöglichte es außerdem, die von den Decodern getroffenen Zuordnungsentscheidungen exakt nachzuvollziehen und, wo nötig, auch Gründe für falsch getroffene Entscheidungen offenzulegen.

Ein ebenfalls im Rahmen dieser Arbeit entwickeltes Werkzeug für das Geoinformationssystem *ArcGIS* ermöglicht die Visualisierung sämtlicher Karten, Location Reference Points und Ortsreferenzen des Zuordnungsprozesses. Diese gemeinsame grafische Darstellung ermöglichte den direkten visuellen Vergleich der codierten und decodierten Ortsreferenzen im Quell- und Zielnetz sowie der zugehörigen Location Reference Points. Dies trug zu einem besseren Verständnis der numerisch erzielten Evaluierungsergebnisse bei und erleichterte die Ursachenforschung bei inkorrekt durchgeführten Location Referencing-Übertragungen.

Die empirischen Ergebnisse zeigen grundsätzlich eine Steigerung der Performanz des *Form-Matchers* im Vergleich zu OpenLR. Besonders auffällig war diese Steigerung bei Location Referencing Übertragungen von der OSM-Karte zu TomTom MultiNet. Hier erzielte der Form-Matcher eine Korrekturklassifikationsrate (Korrektheit der Zuordnung) von 75 % für alle Testszenarien im Vergleich zu 63 % von OpenLR. Die erzielten Steigerungen ließen sich auf die vorgenommenen Erweiterungen und Verbesserungen des Location Referencing-Verfahrens zurückführen. Bei entgegengesetzter Übertragung war lediglich nur eine leichte Steigerung zu beobachten.

Ebenso wurde untersucht, ob die Wahl einer alternativen Metrik zur Berechnung der geometrischen Abstände zwischen den zu vergleichenden Routen einen Einfluss auf das Zuordnungsergebnis besitzt. Dazu wurden zunächst Hausdorff- und Frechét-Metriken vorgestellt und dabei die diskrete Form der Frechét-Distanz als grundsätzlich für die Ähnlichkeitsbestimmung linienhafter Ortsreferenzen geeignet identifiziert, da diese im Gegensatz zur Hausdorff-Distanz den Richtungsverlauf der Ortsreferenz mit berücksichtigt. Bei den durchgeführten empirischen Untersuchungen konnte diesbezüglich allerdings keine Verbesserung der Ergebnisse festgestellt werden. Dies ist vermutlich im typischerweise Vorkommen mehrerer, einen ähnlichen Verlauf abdeckender Routen in der Menge der zu bewertenden Kandidaten, begründet. Immerhin war die korrekte Zuordnungslösung stets in der Menge der Kandidaten mit der minimalen Frechét-Distanz enthalten. Die Verwendung der Frechét-Metrik für das Location Referencing erscheint jedoch für zukünftige Arbeiten überaus interessant. Vielleicht ließe sich mit Hilfe einer spezifisch abgestimmten Nachprozessierung, die verbleibende Menge an Routenkandidaten mit gleicher Frechét-Distanz weiter bis zur korrekten Lösung eingrenzen.

## Ausblick

Durch die im Rahmen der vorliegenden Arbeit durchgeführte Analyse des OpenLR Location Referencing-Verfahrens sowie der Ergebnisse der empirischen Untersuchungen können im Folgenden Vorschläge zur weiteren Entwicklung und Verbesserung von Location Referencing-Verfahren gegeben werden.

OpenLR definiert die Start-/Endpunkte der zu übertragenden Ortsreferenzen lediglich durch geometrische und topologische Eigenschaften im Quellnetz. Solange Ortsreferenzen im Bereich von Kreuzungen bzw. echter Netzknoten beginnen oder enden, führt diese Herangehensweise zu guten Resultaten. Der hier vorgestellte Ansatz zur Bewertung der Knotenvalenz könnte jedoch weiter verbessert werden, indem ein wichtiger Sonderfall zusätzlich berücksichtigt wird. So sind einwertige Knoten (Valenz 1) beispielsweise am Ende einer Sackgasse auch valide Zuordnungslösungen und sollten als solche ebenso wie Knoten mit Valenz  $\geq 2$  gewertet werden. Beginnt oder endet eine Ortsreferenz aber abseits eines echten Knotens oder sogar an einem zweiwertigen (bivalenten) Knoten, so wird die Zuordnung bzw. die

Rekonstruktion dieses Start-/Endpunktes im Zielnetz mit den derzeitigen Verfahren in manchen Fällen nicht korrekt gelingen. Grund dafür ist die Reduktion des Location-Beginns auf einen geometrischen Abstand zum nächstgelegenen, echten Netzknoten. Die beobachteten geometrischen Unterschiede zwischen verschiedenen Kartensystemen können so zu einer ungenauen Rekonstruktion im Zielnetz führen. Eine wesentliche Verbesserung ließe sich in solchen Fällen erzielen, wenn bei der Codierung von Ortsreferenzen nicht nur deren geometrische/topologische Referenzierung im Netz herangezogen, sondern auch ihr *semantischer* Kontext mit einbezogen werden würde. So wäre es in benanntem Beispiel relevant zu wissen, welche semantische Information, etwa einen Attributwechsel, der besagte zweiwertige Knoten besitzt. Im Zielnetz wäre dann nicht so sehr der geometrische Abstand zum nächsten Vollknoten, sondern die Suche nach dem referenzierten Attributwechsel, beispielsweise einer Änderung der zulässigen Höchstgeschwindigkeit oder der Beginn eines Tunnels, relevant. Der entsprechende Knoten im Zielnetz könnte somit unabhängig von möglicherweise auftretenden geometrischen Abweichungen zwischen den Karten korrekt identifiziert werden. Durch Hinzunahme dieses semantischen Kontextes müsste angegeben werden, mit welcher Gewichtung die in der Ortsreferenz codierten geometrischen, topologischen und semantischen Eigenschaften bei der Decodierung im Zielnetz betrachtet werden müssen.

Ein ähnlicher, wenn auch nicht ganz so feinkörniger Effekt könnte erzielt werden, indem Information zur Ortsreferenz hinzugefügt werden, ob sich der angegebene geometrische Offset auf einen zweiwertigen Knoten bezieht, oder die Ortsreferenz tatsächlich auf freier Strecke im Quellnetz definiert wurde. Damit ließe sich bei der Decodierung im Zielnetz immerhin eine Präferenz für bzw. gegen mögliche zweiwertige Knotenkandidaten erzielen.

Wie in der Arbeit gezeigt, besitzt die Schema-Zuordnung zwischen den verschiedenen Datenbank-(Karten-)Schemata und der Location Referencing-Austausch-Schema ebenso einen systematischen Einfluss auf die Güte der Zuordnungen einer bestimmten Kombination von Quell- und Zielkarte. Dabei kann es zu falschen Zuordnungen kommen, wenn die Attributinformationen aus einer oder beider Karten einander nicht exakt zugeordnet werden können. Insbesondere beim Attribut *Form of Way* war dies häufiger zu beobachten, da dieses in OpenStreetMap nicht in dieser Form modelliert ist und daher aus anderen Attributen abgeleitet werden muss. Die Zuverlässigkeit der Schema-Zuordnung für bestimmte Attribute ist jedoch bei der Implementierung der Location Referencing-Anbindung an ein Kartensystem bekannt. Diese Information wird aber bislang nicht genutzt. Im Rahmen der vorliegenden Arbeit wurde eine Möglichkeit zur Modellierung dieser Unsicherheiten bei der Schema-Zuordnung und die Veröffentlichung dieser Informationen mit Hilfe beschreibender Metadaten vorgeschlagen. Dieses Verfahren wurde testweise implementiert und daraus eine dynamisches Gewicht für die entsprechenden Attribute bei der Ähnlichkeitsbetrachtung zweier Ortsreferenzen abgeleitet. Bei den empirischen Untersuchungen zeigte sich allerdings keine Verbesserung der Zuordnungsgüte, möglicherweise aufgrund der sehr gut übereinstimmenden Modellierung von OpenLR und TomTom MultiNet. Diese Betrachtung erfolgte jedoch lediglich als Randaspekt der vorliegenden Arbeit. Das entwickelte Konzept ist nach Ansicht des Autors jedoch grundsätzlich geeignet, die Unsicherheiten bei der Schema-Zuordnung zu modellieren. Für zukünftige Arbeiten wird eine eingehendere Untersuchung, vor allem auch mit mehreren, unterschiedlichen Kartensystemen, angeregt.

Bei den empirischen Untersuchungen wurde offenbar, dass die Qualität der Location Referencing Übertragungen nicht unabhängig von der Quell- und Zielkarte ist. Diese Erkenntnis regt weiterführende Untersuchungen an, insbesondere um nach den Gründen dieses Verhaltens zu suchen. Dabei bietet es sich an, die Qualität digitaler Straßenkarten weiter zu untersuchen und zu formalisieren. Für das Location Referencing ist sicherlich die immer geringer werdende Bedeutung von Bandbreitenbeschränkungen der Übertragungskanäle eine große Chance mehr Informationen übertragen zu können und somit die Qualität und Zuverlässigkeit der Zuordnungen weiter zu steigern. Im Rahmen dieser Arbeit hat sich gezeigt, dass die Location Referencing Übertragungen und dabei vor allem die Zuordnungsentscheidungen beim Decodieren im Zielnetz mit Unsicherheiten behaftet sind. Diese könnten in weiterführenden Arbeiten untersucht, und nach Möglichkeiten der Modellierung gesucht werden. Fuzzy-basierte oder

stochastische Ansätze wären hier sicher denkbar. Sicherlich bieten die im Rahmen der vorliegenden Arbeit gefundenen Ergebnisse, wie auch die analytischen Überlegungen eine gute Grundlage dafür.

# Literaturverzeichnis

- ALEKSIĆ, M. (2004): *Validation Report*. Deliverable 6.2, ActMAP-Consortium.
- ALHESSI, M. (2013): *Development and Implementation of the OpenLR Map Interface for Shapefiles*. Masterarbeit, Institut für Ingenieurgeodäsie, Universität Stuttgart, unveröffentlicht.
- ALT, H. und GODAU, M. (1995): *Computing the Fréchet Distance between two Polygonal Curves*. In: *International Journal of Computational Geometry & Applications*, Jahrgang 5, Heft 1-2: S. 75–91.
- ANDERS, F. (2007): *Mustererkennung in Straßennetzwerken – Verfahren zur Interpretation von Vektordaten*. Dissertation, Deutsche Geodätische Kommission (DGK), Reihe C, Nr. 607, München.
- APACHE (2004): *Apache License Version 2.0*. The Apache Software Foundation. URL <http://www.openlr.org/data/docs/OpenLR-Software-License2.pdf>.
- ARKIN, E., CHEW, L., HUTTENLOCHER, D., KEDEM, K. und MITCHELL, J. (1991): *An efficiently computable metric for comparing polygonal shapes*. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Jahrgang 13, Heft 3: S. 209–216.
- ARONOV, B., HAR-PELED, S., KNAUER, C., WANG, Y. und WENK, C. (2006): *Fréchet distance for curves, revisited*. In: *Proceedings of the 14th conference on Annual European Symposium – Volume 14, ESA'06*, S. 52–63. Springer-Verlag, London, UK.
- BASELAU, S. (2010): *Routensuche in zeitlich variablen Netzen, Schriftenreihe Studien zur Wirtschaftsinformatik*, Bd. 48. Verlag Dr. Kovač, Hamburg.
- BEERI, C., DOYTHER, Y., KANZA, Y., SAFRA, E. und SAGIV, Y. (2005): *Finding corresponding objects when integrating several geo-spatial datasets*. In: SHAHABI, C. und BOUCELMA, O. [Hrsg.]: *Proceedings of the 13th ACM International Workshop on Geographic Information Systems, ACM-GIS 2005*, 87–96. ACM, Bremen.
- BEERI, C., KANZA, Y., SAFRA, E. und SAGIV, Y. (2004): *Object Fusion in Geographic Information Systems*. In: NASCIMENTO, M. A., ÖZSU, M. T., KOSSMANN, D., MILLER, R. J., BLAKELEY, J. A. und SCHIEFER, K. B. [Hrsg.]: *Proceedings of the Thirtieth International Conference on Very Large Data Bases*, S. 816–827. Toronto, Canada.
- BEETZ, A., SCHÜTZLE, R. und METZNER, M. (2013): *Quality of Service*. Contracted study final report, eMaPS-Consortium. URL <http://tn-its.eu/docs/emaps/eMaPS-D2.43-USTUTT-Quality-of-Service-v20-130305.pdf>.
- BEHRENS, R., KLEINE-BESTEN, T., PÖCHMÜLLER, W. und ENGELSBURG, A. (2015): *Digitale Karten im Navigation Data Standard Format*. In: WINNER, H., HAKULI, S., LOTZ, F. und SINGER, C. [Hrsg.]: *Handbuch Fahrerassistenzsysteme*, ATZ/MTZ-Fachbuch, S. 513–523. Springer Fachmedien Wiesbaden.
- BESL, P. und MCKAY, N. D. (1992): *A method for registration of 3-D shapes*. In: *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, Jahrgang 14, Heft 2: S. 239–256.
- BILL, R. (2010): *Grundlagen der Geo-Informationssysteme*. 5. Aufl. Wichmann, Berlin.
- BILL, R. und ZEHNER, M. L. (2001): *Lexikon der Geoinformatik*. Wichmann, Heidelberg.
- BIROLINI, A. (1991): *Qualität und Zuverlässigkeit technischer Systeme: Theorie, Praxis, Management*. 3. Aufl. Springer, Berlin.
- BISHR, Y. A. (1998): *Overcoming the Semantic and Other Barriers to GIS Interoperability*. In: *International Journal of Geographical Information Science*, Jahrgang 12, Heft 4: S. 299–314.
- BLASBY, D., DAVIS, M., KIM, D. und RAMSEY, P. (2003): *GIS Conflation using Open Source Tools*. White paper, The JUMP Project.

- BOFINGER, J.-M. (2001): *Analyse und Implementierung eines Verfahrens zur Referenzierung geographischer Objekte*. Diplomarbeit, Institut für Photogrammetrie, Universität Stuttgart, unveröffentlicht.
- BOSCH (2015): *Bosch und TomTom arbeiten an Kartentechnologie für das automatisierte Fahren*. Presse-Information, Robert Bosch GmbH, Stuttgart, Deutschland.
- BOSCH SOFTTEC (2015): *360°logiQ – Connected Horizon Website*. Bosch SoftTec GmbH. URL [www.bosch-softtec.com/360logiq.html](http://www.bosch-softtec.com/360logiq.html), letzter Zugriff: 26.08.2015.
- BRUNS, H. T. und EGENHOFER, M. J. (1996): *Similarity of Spatial Scenes*. In: KRAAK, J.-M. und MOLENAAR, M. [Hrsg.]: *Proceedings of the 7th Symposium on Spatial Data Handling, Delft, The Netherlands*, S. 173–184. Taylor & Francis, London.
- BUTENUTH, M., v. GÖSSELN, G., TIEDGE, M., HEIPKE, C., LIPECK, U. und SESTER, M. (2007): *Integration of heterogeneous geospatial data in a federated database*. In: Theme Issue: Distributed Geoinformatics – From Sensors to Systems, Jahrgang 62, Heft 5: S. 328–346.
- CATLING, I., HARRIS, R. und ZIJDERHAND, F. (1994): *The SOCRATES projects: progress towards an integrated road transport environment*. In: *Vehicle Navigation and Information Systems Conference, 1994. Proceedings., 1994*, S. 651–655.
- CCPL (2007): *Creative Commons Attribution-No Derivative Works 3.0 Unported License*. Creative Commons. URL <http://creativecommons.org/licenses/by-nd/3.0/de/legalcode>.
- CECCONI, A. (2003): *Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping*. Dissertation, ETH Zürich. URL [http://www.geo.uzh.ch/fileadmin/files/content/abteilungen/gis/research/phd\\_theses/thesis\\_AllessandroCecconi\\_2003.pdf](http://www.geo.uzh.ch/fileadmin/files/content/abteilungen/gis/research/phd_theses/thesis_AllessandroCecconi_2003.pdf), letzter Zugriff: 30.05.2013.
- CHEN, H. (2011): *Entwicklung von Verfahren zur Beurteilung und Verbesserung der Qualität von Navigationsdaten*. Dissertation, Institut für Photogrammetrie, Universität Stuttgart. URL <http://elib.uni-stuttgart.de/opus/volltexte/2011/6899>, letzter Zugriff: 19.02.2013.
- CHEN, H., WALTER, V. und FRITSCH, D. (2008): *Quality Inspection and Quality Improvement by Map Fusion*. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVII, Part B2: S. 467–472.
- COBB, M., CHUNG, M., III, H. F., PETRY, F. E., SHAW, K. und MILLER, V. (1998): *A Rule-based Approach for the Conflation of Attributed Vector Data*. In: *GeoInformatica*, Jahrgang 2, Heft 1: S. 7–35.
- COHEN, I., AYACHE, N. und SULGER, P. (1992): *Tracking points on deformable objects using curvature information*. In: SANDINI, G. [Hrsg.]: *Computer Vision – ECCV’92, Lecture Notes in Computer Science*, Bd. 588, S. 458–466. Springer Berlin Heidelberg.
- COHEN, W. W. (2000): *Data integration using similarity joins and a word-based information representation language*. In: *ACM Transactions on Information Systems*, Jahrgang 18, Heft 3: S. 288–321.
- COHEN, W. W., RAVIKUMAR, P. D. und FIENBERG, S. E. (2003): *A Comparison of String Distance Metrics for Name-Matching Tasks*. In: KAMBHAMPATI, S. und KNOBLOCK, C. A. [Hrsg.]: *Proceedings of IJCAI-03 Workshop on Information Integration on the Web (IIWeb-03)*, S. 73–78. Acapulco, Mexico.
- CONRAD, S. (2002): *Schemaintegration Integrationskonflikte, Lösungsansätze, aktuelle Herausforderungen*. In: *Informatik, Forschung und Entwicklung*, Jahrgang 17, Heft 3: S. 101–111.
- CONTINENTAL (2015): *eHorizon Website*. Continental AG. URL [www.continental-automotive.com/www/automotive\\_de/en/themes/passenger\\_cars/interior/connectivity/pi\\_ehorizon\\_en.html](http://www.continental-automotive.com/www/automotive_de/en/themes/passenger_cars/interior/connectivity/pi_ehorizon_en.html), letzter Zugriff: 26.08.2015.
- DAHINDEN, T., BOLJEN, J., RUHE, N., DALYOT, S., SCHULZE, M. J. und SESTER, M. (2013): *Geodatenintegration – Semantischer und geometrischer Abgleich von Vektordaten*. In: *Zeitschrift für das Vermessungswesen*, Jahrgang 138, Heft 1: S. 85–94.
- DAIMLER (2015a): *AUDI AG, BMW Group und Daimler AG einigen sich mit Nokia Corporation über gemeinsamen Kauf des digitalen Kartengeschäfts HERE*. Presse-Information, Daimler AG, Stuttgart, Deutschland.

- DAIMLER (2015b): *Car-to-X Kommunikation – Den eigenen Horizont erweitern*. Presse-Information, Daimler AG, Stuttgart, Deutschland.
- DALYOT, S., DAHINDEN, T., SCHULZE, M. J., BOLJEN, J. und SESTER, M. (2012): *Geometrical Adjustment Towards the Alignment of Vector Databases*. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, I-4: S. 13–18.
- DALYOT, S., DAHINDEN, T., SCHULZE, M. J., BOLJEN, J. und SESTER, M. (2013): *Integrating network structures of different geometric representations*. In: Survey Review, Jahrgang 45, Heft 333: S. 428–440.
- DEMING, W. E. (1982): *Out of the crisis*. MIT Press, Cambridge, MA, USA.
- DEMIR, C. (2002): *A New Location Referencing Method for Unique Reconstruction of an Object on a Second Map*. In: *Proceedings of the 9th World Congress on Intelligent Transport Systems*. Chicago, USA.
- DERETSKY, Z. und RODNY, U. (1993): *Automatic conflation of digital maps*. In: *Vehicle Navigation and Information Systems Conference, 1993., Proceedings of the IEEE-IEE*, S. A27–A29.
- DEVOGELE, T. (2002): *A new merging process for data integration based on the discrete Fréchet distance*. In: *Proceedings of the 10th International Symposium on Spatial Data Handling (SDH)*. Ottawa, Canada.
- DEVOGELE, T., PARENT, C. und SPACCAPIETRA, S. (1998): *On spatial database integration*. In: *International Journal of Geographic Information Systems, Special Issue on System Integration*, Jahrgang 12, Heft 4: S. 335–352.
- DEVOGELE, T., TREVISAN, J. und RAYNAL, L. (1996): *Building a Multi-Scale Database with Scale-Transition Relationships*. In: *Proceedings of the 7th International Symposium on Spatial Data Handling (SDH)*. Delft, The Netherlands.
- DIN EN ISO 14819-3 (2014): *Intelligente Transportsysteme – Verkehrs- und Reiseinformationen über Verkehrsmeldungskodierung – Teil 3: Ortsreferenzierung für Radiodatensysteme für den digitalen Radiokanal für Verkehrsmeldungen (RDS-TMC) unter Nutzung von ALERT-C*. Deutsches Institut für Normung e.V., Beuth Verlag, Berlin.
- DIN EN ISO 19157 (2014): *Geoinformation – Datenqualität*. Deutsches Institut für Normung e.V., Beuth Verlag, Berlin.
- DIN EN ISO 9000 (2005): *Qualitätsmanagementsysteme – Grundlagen und Begriffe*. Deutsches Institut für Normung e.V., Beuth Verlag, Berlin.
- DIN EN ISO 9001 (2008): *Qualitätsmanagementsysteme – Anforderungen*. Deutsches Institut für Normung e.V., Beuth Verlag, Berlin.
- DIN ISO/TS 18234-6 (2006): *Reise- und Verkehrsinformation (TTI) – TTI über Datenströme der Transportprotokoll Expertengruppe (TPEG) – Teil 6: Ortskodierung für Anwendungen (TPEG-Loc), Vornorm*. Deutsches Institut für Normung e.V., Beuth Verlag, Berlin.
- DORENBECK, C. und HERES, L. (1999): *Final Location Referencing Rules Specifications*. Project Report, EVIDENCE Consortium, Brussels, Belgium.
- DOYTSHER, Y. und FILIN, S. (2000): *Detection of corresponding objects in a linear-based map conflation*. In: *Surveying and Land Information Systems*, Jahrgang 60, Heft 2: S. 117–128.
- DUCKECK, R., HIESTERMANN, V., MILTON, H., SENNA, M. L. und WEVERS, K. (1998): *Rules for Defining and Referencing an Intersection Location (ILOC): Detailed Location Referencing (DLR) for ITS based on ILOCs*. Final report, ERTICO Committee on Location Referencing, Brussels, Belgium.
- DUEKER, K. J. und BUTLER, J. A. (2000): *A geographic information system framework for transportation data sharing*. In: *Transportation Research Part C: Emerging Technologies*, Jahrgang 8, Heft 1–6: S. 13–36.
- DUNKARS, M. (2003): *Matching of Datasets*. In: VIRRANTAUS, K. und TVEITE, H. [Hrsg.]: *Proceedings of ScanGIS'2003 – The 9th Scandinavian Research Conference on Geographical Information Science*, S. 67–78. Department of Surveying, Helsinki University of Technology, Espoo, Finland.
- DUNKARS, M. (2004): *Multiple Representation Databases for Topographic Information*. Dissertation, Royal Institute of Technology, Stockholm, Schweden. URL <http://www.dissertations.se/dissertation/38256665ba>, letzter Zugriff: 13.05.2013.
- EDWARDS, D. und SIMPSON, J. (2002): *Integration and Access of Multi-source Vector Data*. In: *Proceedings of the Joint International Symposium on Geospatial Theory, Processing and Applications*. Ottawa, Canada.

- EGENHOFER, M. J. (1991): *Reasoning about binary topological relations*. In: GÜNTHER, O. und SCHEK, H.-J. [Hrsg.]: *Advances in Spatial Databases, Lecture Notes in Computer Science*, Bd. 525, S. 141–160. Springer, Berlin, Heidelberg.
- EGENHOFER, M. J. und MARK, D. M. (1995): *Naive Geography*. In: FRANK, A. U. und KUHN, W. [Hrsg.]: *Spatial Information Theory – A Theoretical Basis for GIS, Lecture Notes in Computer Science*, Bd. 988, S. 1–15. Springer, Berlin, Heidelberg.
- EITER, T. und MANNILA, H. (1994): *Computing Discrete Fréchet Distance*. Technical Report, Christian Doppler Labor für Expertensysteme, Technische Universität Wien, Österreich.
- EP 1 078 346 (2002): *Method for Generating and Interpreting a Location Reference Instance within a Map*. Patentblatt 2002/15. Europäisches Patentamt.
- EP 1 122 518 (2008): *Verfahren zur Decodierung von linienförmigen geographischen Objekten*. Patentblatt 2008/05. Europäisches Patentamt.
- EP 1 224 645 (2003): *Verfahren zur Codierung und Decodierung von Objekten in einem Verkehrsnetz*. Patentblatt 2003/45. Europäisches Patentamt.
- EP 1 225 552 (2004): *A method and system for referencing locations in transport telematics*. Patentblatt 2004/45. Europäisches Patentamt.
- EP 1 336 079 (2004): *Verfahren zur Erstellung von Ein Objekt geographisch eundeutig referenzierender Appendices*. Patentblatt 2004/49. Europäisches Patentamt.
- ERNST, T., KWELLA, B., SCHMIDT, M. und RADUSCH, I. (2014): *Cross-vendor testing TPEG2-ULR linear location referencing*. In: ERTICO [Hrsg.]: *Proceedings of the 10th Intelligent Transportation System (ITS) European congress and exhibition*. Helsinki, Finland.
- ERNST, T., KWELLA, B., SCHRAMM, A. und SCHMIDT, M. (2013a): *Direction Instructions for Communicating Linear Locations in TPEG2-ULR*. Progress Report, TPEG-LOC2 Study, Fraunhofer FOKUS, Berlin, Germany.
- ERNST, T., POHL, H.-W., KWELLA, B. und SCHMIDT, M. (2013b): *A Multipath Routing Algorithm for Encoding and Decoding Linear Locations in TPEG2-ULR*. Progress Report, TPEG-LOC2 Study, Fraunhofer FOKUS, Berlin, Germany.
- ESRI (1998): *ESRI Shapefile Technical Description*. White paper, Environmental Systems Research Institute, Inc. URL <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>, letzter Zugriff: 23.03.2015.
- FRÉCHET, M. M. (1906): *Sur quelques points du calcul fonctionnel*. In: Rendiconti del Circolo Matematico di Palermo, Jahrgang 22, Heft 1: S. 1–72.
- GABAY, Y. und DOYTSHER, Y. (1994): *Automatic adjustment of line maps*. In: *Proceedings of the GIS/LIS094 annual convention*, S. 333–341. Phoenix, USA.
- GABAY, Y. und DOYTSHER, Y. (1995): *Automatic feature correction in merging of line maps*. In: *Proceedings of the ACSM-ASPRS Annual Convention 2*, S. 404–410.
- GABAY, Y. und DOYTSHER, Y. (2000): *An approach to matching lines in partly similar engineering maps*. In: *Geomatica*, Jahrgang 54, Heft 3: S. 297–310.
- GAN, E. und SHI, W. (2002): *Error Metadata Management System*. In: SHI, W., FISHER, P. F. und GOODCHILD, M. F. [Hrsg.]: *Geographical information systems*, Kap. 17, S. 251–266. Taylor & Francis, London.
- GOODCHILD, M. F. (2007): *Citizens as sensors: the world of volunteered geography*. In: *GeoJournal*, 69, 4: S. 211–221.
- GOODCHILD, M. F. und HUNTER, G. J. (1997): *A simple positional accuracy measure for linear features*. In: *International Journal of Geographical Information Science*, Jahrgang 11, Heft 3: S. 299–306.
- GRÉGOIRE, N. und BOUILLOT, M. (1998): *Hausdorff Distance between convex polygons*. Web project. URL <http://cgm.cs.mcgill.ca/~godfried/teaching/cg-projects/98/normand/main.html>, letzter Zugriff: 03.08.2015.
- GÖSSELN, G. v. (2005): *A matching approach for the integration, change detection and adaptation of heterogeneous vector data sets*. In: *Proceedings of the XXII International Cartographic Conference*. A Coruña, Spanien.
- GÖSSELN, G. v. und SESTER, M. (2003): *Semantic and geometric integration of geoscientific data sets with ATKIS-applied to geo-objects from geology and soil science*. In: *Proceedings of ISPRS commission IV joint workshop*, S. 111–116. Stuttgart.

- HACKELÖER, A., KLASING, K., KRISP, J. M. und MENG, L. (2013): *Comparison of Point Matching Techniques for Road Network Matching*. In: ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-2/W1: S. 87–92.
- HACKELÖER, A., KLASING, K., KRISP, J. M. und MENG, L. (2014): *Georeferencing: a review of methods and applications*. In: Annals of GIS, Jahrgang 20, Heft 1: S. 61–69.
- HACKELÖER, A., KLASING, K., KRISP, J. M. und MENG, L. (2015): *Road Network Conflation: An Iterative Hierarchical Approach*. In: GARTNER, G. und HUANG, H. [Hrsg.]: *Progress in Location-Based Services 2014*, S. 137–151. Springer.
- HAHLWEG, C., HESSING, B., KERSKEN, U., BRANDES, H. und HESSLING, M. (2000): *Goodlane – An Approach to Location Referencing for Telematic Applications*. In: *Proceedings of the 7th World Congress on Intelligent Transport Systems*. Turin, Italy.
- HALL, P. A. V. und DOWLING, G. R. (1980): *Approximate String Matching*. In: ACM Comput. Surv., Jahrgang 12, Heft 4: S. 381–402.
- HANGOUËT, J.-F. (1995): *Computation of the Hausdorff Distance between Plane Vector Polylines*. In: *Proceedings of the Auto-Carto12, ACSM/ASPRS Annual Convention and Exposition Technical Papers*. Charlotte, NC, USA.
- HAUSDORFF, F. (1918): *Dimension und äußeres Maß*. In: Mathematische Annalen, Jahrgang 79, Heft 1–2: S. 157–179.
- HEINZLE, F. und ANDERS, K.-H. (2007): *Characterising Space via Pattern Recognition Techniques: Identifying Patterns in Road Networks*. In: MACKANESS, W. A., RUAS, A. und SARJAKOSKI, L. T. [Hrsg.]: *Generalisation of Geographic Information*, International Cartographic Association, S. 233–253. Elsevier Science B.V., Amsterdam, The Netherlands.
- HENDRIKS, T., WEVERS, K., PFEIFFER, H.-W. und HESSLING, M. (2005): *Specification of the AGORA-C on-the-fly location referencing method*. Version 1.0.
- HERE (2015): *Vehicle Sensor Data Cloud Ingestion Interface Specification*. Techn. Ber. 2.0.2, HERE, A Nokia Company, Espoo, Finland.
- HIESTERMANN, V. (2008): *Map-independent location matching certified by the AGORA-C standard*. In: Transportation Research Part C: Emerging Technologies, Jahrgang 16, Heft 3: S. 307–319. Emerging Commercial Technologies.
- HOLT, A. (1998): *The Matching and Ranking of Surface and Deep Similarities of Spatial Data*. In: *Proceedings of the 10th Annual Colloquium of the Spatial Information Research Centre, University of Otago, New Zealand*, S. 133–143.
- ISO 14825 (2005): *Intelligent Transport Systems – Geographic Data Files (GDF) – Overall Data Specification*. International Organization for Standardization.
- ISO 14825 (2011): *Intelligent Transport Systems – Geographic Data Files (GDF) – GDF5.0*. International Organization for Standardization.
- ISO 17572-3 (2008): *Intelligent Transport Systems – Location Referencing for Geographic Databases – Part 3: Dynamic Location Referencing (Dynamic Profile)*. International Organization for Standardization.
- ISO/TS 19138 (2006): *Geographic information – Data quality measures (Technical Specification)*. International Organization for Standardization.
- JOOS, G. (2000): *Zur Qualität von objektstrukturierter Geodaten*. Schriftenreihe des Studienganges Geodäsie und Geoinformation der Universität der Bundeswehr München, Heft 66. München.
- KAMPSHOFF, S. (2005): *Integration heterogener raumbezogener Objekte aus fragmentierten Geodatenbeständen*. Dissertation, RWTH Aachen, Aachen. URL <http://publications.rwth-aachen.de/record/51912>, letzter Zugriff: 04.06.2013.
- KASHYAP, V. und SHETH, A. (1996): *Semantic and schematic similarities between database objects: a context-based approach*. In: International Journal on Very Large Data Bases, Jahrgang 5, Heft 4: S. 276–304.
- KIM, J. O., YU, K., HEO, J. und LEE, W. H. (2010): *A new method for matching objects in two different geospatial datasets based on the geographic context*. In: Computers & Geosciences, Jahrgang 36, Heft 9: S. 1115–1122.

- KOCHS, A., WEIDNER, B. und SCHOLTES, W. (2011): *Analyse der Regelwerke und Ordnungssysteme/Ortsreferenzierungssysteme der Bereiche Straßeninformationssysteme und Verkehrssysteme*. Techn. Ber., im Auftrag der Bundesanstalt für Straßenwesen.
- KRAUT, M. (2003): *Zuordnung und Conflation heterogener Straßendaten*. Diplomarbeit, Institut für Photogrammetrie, Universität Stuttgart, unveröffentlicht.
- KRUMKE, S. O. und NOLTEMEIER, H. (2012): *Graphentheoretische Konzepte und Algorithmen*. 3. Aufl. Springer Vieweg, Wiesbaden.
- LAUFER, R. (2011): *Prozedurale Qualitätsmodellierung und -management für Daten: ingenieurgeodätische und verkehrstechnische Anwendungen*. Dissertation, Deutsche Geodätische Kommission (DGK), Reihe C, Nr. 662, München.
- LEMARIÉ, C. und RAYNAL, L. (1996): *Geographic Data Matching: First investigations for a generic tool*. In: *Proceedings of the GIS/LIS096 annual convention*, S. 405–420. Denver, Colorado, USA.
- LEVENSHTAIN, V. I. (1966): *Binary codes capable of correcting deletions, insertions, and reversals*. In: *Soviet Physics Doklady*, Jahrgang 10, Heft 8: S. 707–710. Übersetzt von *Doklady Akademii Nauk SSSR*, Vol. 163, Nr. 4, S. 845–848, 1965.
- LUDWIG, J. (2014): *Elektronischer Horizont – Vorausschauende Systeme und deren Anbindung an Navigationseinheiten*. In: SIEBENPFEIFFER, W. [Hrsg.]: *Vernetztes Automobil*, ATZ/MTZ-Fachbuch, S. 223–229. Springer Fachmedien Wiesbaden.
- LYNCH, M. P. und SAALFELD, A. J. (1985): *Conflation: automated map compilation – a video game approach*. In: *Proceedings of Autocarto 7*, S. 343–352. Washington, DC, USA.
- MANTEL, D. und LIPECK, U. (2004): *Matching Cartographic Objects in Spatial Databases*. In: *Proceedings of the XX th ISPRS Congress, Comm. IV*, S. 172–176. Istanbul, Türkei.
- MARKSCHLÄGER, P., WAHL, H.-G., WEBERBAUER, F. und LEDERER, M. (2014): *Assistenzsystem für mehr Kraftstoffeffizienz*. In: SIEBENPFEIFFER, W. [Hrsg.]: *Vernetztes Automobil*, ATZ/MTZ-Fachbuch, S. 146–153. Springer Fachmedien Wiesbaden.
- MASCRET, A., DEVOGELE, T., BERRE, I. L. und HÉNAFF, A. (2006): *Coastline Matching Process based on the discrete Fréchet Distance*. In: *Proceedings of the 12th International Symposium on Spatial Data Handling (SDH)*. Vienna, Austria.
- MIN, D., ZHILIN, L. und XIAOYONG, C. (2007): *Extended Hausdorff distance for spatial objects in GIS*. In: *International Journal of Geographical Information Science*, Jahrgang 21, Heft 4: S. 459–475.
- MÖLLER, C. (2012): *OSM2PO: OpenStreetMap-Data To Topology Converter and Routing Engine for Java*. Schnelleinstieg, Kurzdokumentation, Referenz für Version 4.x. URL <http://osm2po.de/dld/osm2po-4.8.8.zip>.
- MÜLLER, T. M. (2010): *Navigation Data Standard (NDS): Bald Industriestandard?* In: *Automobil-Elektronik*, , Dez. 2010: S. 30–31.
- MORRISON, J. L. (1995): *Spatial data quality*. In: GUPTILL, S. C. und MORRISON, J. L. [Hrsg.]: *Elements of spatial data quality*, Kap. 1, S. 1–12. Elsevier Science, Oxford.
- MUSTIÈRE, S. (2006): *Results of experiments on automated matching of networks*. In: *Proceedings of the ISPRS Workshop on Multiple Representation and Interoperability of Spatial Data*, S. 92–100. Hannover.
- MUSTIÈRE, S. und DEVOGELE, T. (2008): *Matching Networks with Different Levels of Detail*. In: *GeoInformatica*, Jahrgang 12, Heft 4: S. 435–453.
- NAVARRO, G. (2001): *A guided tour to approximate string matching*. In: *ACM Computing Surveys*, Jahrgang 33, Heft 1: S. 31–88.
- NDS (2015): *NDS Website*. Navigation Data Standard e.V. URL [www.nds-association.org](http://www.nds-association.org), letzter Zugriff: 07.10.2015.
- NOKIA (2015): *Nokia completes next stage of transformation with agreement to sell HERE to automotive industry consortium*. Börsenmitteilung, Nokia Corporation, Espoo, Finnland.
- OGR2OGR (2014): URL <http://www.gdal.org/ogr2ogr.html>, letzter Zugriff: 25.07.2014.
- OLTEANU, A.-M. (2007a): *A multi-criteria fusion approach for geographical data matching*. In: *Proceedings of the 5th International Symposium On Spatial Data Quality (ISSDQ)*. Springer Verlag, Enschede, Netherlands.

- OLTEANU, A.-M. (2007b): *Matching geographical data using the Theory of Evidence*. In: *Proceedings of the XXIII International Cartographic Conference*. Moscow, Russia.
- OLTEANU, A.-M., MUSTIÈRE, S. und RUAS, A. (2006): *Matching imperfect spatial data*. In: CAETANO, M. und PAINHO, M. [Hrsg.]: *7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. Lisbon, Portugal.
- OLTEANU-RAIMOND, A.-M., MUSTIÈRE, S. und RUAS, A. (2015): *Knowledge formalisation for vector data matching using Belief Theory*. In: *Journal of Spatial Information Science*, Nr. 10: S. 21–46.
- OPENSTREETMAP (2015a): *History of OpenStreetMap – OpenStreetMap Wiki*. URL [https://wiki.openstreetmap.org/wiki/History\\_of\\_OpenStreetMap](https://wiki.openstreetmap.org/wiki/History_of_OpenStreetMap), letzter Zugriff: 25.08.2015.
- OPENSTREETMAP (2015b): *Key:highway – OpenStreetMap Wiki*. URL <http://wiki.openstreetmap.org/wiki/Key:highway>, letzter Zugriff: 25.03.2015.
- OPENSTREETMAP (2015c): *Map Features – OpenStreetMap Wiki*. URL [https://wiki.openstreetmap.org/wiki/Map\\_Features](https://wiki.openstreetmap.org/wiki/Map_Features), letzter Zugriff: 24.08.2015.
- OPENSTREETMAP (2015d): *Statistics – OpenStreetMap Wiki*. URL <http://wiki.openstreetmap.org/wiki/Stats>, letzter Zugriff: 24.08.2015.
- OSGEO4W (2014): URL <http://trac.osgeo.org/osgeo4w/>, letzter Zugriff: 25.07.2014.
- OSMOSIS (2014): *OSMOSIS – OpenStreetMap Wiki*. URL <http://wiki.openstreetmap.org/wiki/Osmosis>, letzter Zugriff: 07.05.2014.
- POLLESCH, P. (2011): *Intermodale Netzreferenzierung im Kontext eines europaweiten Telematikansatzes*. In: *NWSIB-Tage 2011*. URL [http://www.nwsib.nrw.de/downloads/nwsib\\_tage\\_2011/Intermodale\\_Netzreferenzierung.pdf](http://www.nwsib.nrw.de/downloads/nwsib_tage_2011/Intermodale_Netzreferenzierung.pdf), letzter Zugriff: 02.12.2014.
- POWERS, D. M. W. (2007): *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation*. Techn. Ber. SIE-07-001, School of Informatics and Engineering, Flinders University, Adelaide, Australia.
- QGIS (2014): *QGIS – Ein freies Open-Source-Geographisches-Informationssystem*. URL [www.qgis.org](http://www.qgis.org), letzter Zugriff: 08.10.2014.
- RAHM, E. und BERNSTEIN, P. A. (2001): *A survey of approaches to automatic schema matching*. In: *The VLDB Journal*, Jahrgang 10, Heft 4: S. 334–350.
- RAIMOND, A.-M. O. und MUSTIÈRE, S. (2008): *Data Matching – A Matter of Belief*. In: *Headway in Spatial Data Handling*, 13th International Symposium on Spatial Data Handling (SDH). Springer Verlag, Montpellier, France.
- RAND, W. M. (1971): *Objective Criteria for the Evaluation of Clustering Methods*. In: *Journal of the American Statistical Association*, Jahrgang 66, Heft 336: S. 846–850.
- ROSATTE (2008): *ROSATTE Projektübersicht*. Faltblatt, ERTICO ITS Europe, Brüssel, Belgien.
- ROSEN, B. und SAALFELD, A. (1985): *Match Criteria For Automatic Alignment*. In: *Proceedings of 7th international symposium on computer-assisted cartography (Auto-Carto 7)*. Washington, DC, USA.
- RUCKLIDGE, W. [Hrsg.] (1996): *Efficient Visual Recognition Using the Hausdorff Distance*. Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg.
- RUIZ, J. J., ARIZA, F. J., UREÑA, M. A. und BLÁZQUEZ, E. B. (2011): *Digital map conflation: a review of the process and a proposal for classification*. In: *International Journal of Geographical Information Science*, Jahrgang 25, Heft 9: S. 1439–1466.
- SAALFELD, A. (1988): *Conflation – Automated map compilation*. In: *International Journal of Geographical Information Science*, Jahrgang 2, Heft 3: S. 217–228.
- SACHS, L. und HEDDERICH, J. (2006): *Angewandte Statistik*. 12. Aufl. Springer-Verlag, Berlin, Heidelberg.
- SAFESPOT (2015): *Safespot Project Website*. URL <http://www.safespot-eu.org>, letzter Zugriff: 26.08.2015.
- SAFRA, E., KANZA, Y., SAGIV, Y., BEERI, C. und DOYTSHER, Y. (2010): *Location-based algorithms for finding sets of corresponding objects over several geo-spatial data sets*. In: *International Journal of Geographical Information Science*, Jahrgang 24, Heft 1: S. 69–106.

- SAFRA, E., KANZA, Y., SAGIV, Y. und DOYTSHER, Y. (2006): *Efficient integration of road maps*. In: DE BY, R. A. und NITTEL, S. [Hrsg.]: *Proceedings of the 14th ACM International Symposium on Geographic Information Systems, ACM-GIS 2006*, S. 59–66. ACM, Arlington, Virginia, USA.
- SAMAL, A., SETH, S. und CUETO, K. (2004): *A feature-based approach to conflation of geospatial sources*. In: *International Journal of Geographical Information Science*, Jahrgang 18, Heft 5: S. 459–489.
- SCHRAMM, A., KWELLA, B., SCHMIDT, M., PIETH, N. und ERNST, T. (2012): *Universal Location Referencing: A New Approach for Dynamic Location Referencing in TPEG*. Progress Report, TPEG-LOC2 Study, Fraunhofer FIRS, Berlin, Germany.
- SCHÜTZLE, R. (2009): *Quality Management in ROSATTE*. In: *Proceedings of the 16th World Congress on Intelligent Transportation Systems*. ERTICO ITS Europe, Stockholm, Schweden.
- SCHÜTZLE, R. und FRANK, J. (2010): *Aggregated test report*. Deliverable 5.4, ROSATTE-Consortium.
- SCHÜTZLE, R., METZNER, M. und SCHWIEGER, V. (2012): *Evaluation of dynamic location referencing algorithms*. In: *Proceedings of the 6th International Symposium on Networks for Mobility*. University of Stuttgart – Center for Transportation Research, Stuttgart.
- SCHÜTZLE, R., PANDAZIS, J.-C., LANDWEHR, M., WEVERS, K., T'SIOBBEL, S., DELORME, C., MAHIOU, R., PETIT, F., SVENSK, P.-O., WIKSTRÖM, L., ISAKSSON, P., BOTERBERGH, B., BLAIVE, L., KEITH, H., BERGMAN, O. und HASPEL, U. (2011): *Test and validation plan*. Deliverable 5.1, ROSATTE-Consortium.
- SCHÜTZLE, R. und SCHEIDER, A. (2009): *Report on data quality management concept*. Deliverable 5.2, ROSATTE-Consortium.
- SCHWEITZER, J. und SCHWIEGER, V. (2011): *Modeling of quality for engineering geodesy processes in civil engineering*. In: *Journal of Applied Geodesy*, Jahrgang 5, Heft 1: S. 13–22.
- SCHWIEGER, V., ZHANG, L. und SCHWEITZER, J. (2013): *Ingenieurgeodätische Qualitätsmodellierung im Bauprozess*. In: DVW E.V. [Hrsg.]: *Qualitätssicherung geodätischer Mess- und Auswerteverfahren, Schriftenreihe des DVW*, Bd. 71. Beiträge zum 125. DVW Seminar, Hannover.
- SESTER, M. (1995): *Lernen struktureller Modelle für die Bildanalyse*. Dissertation, Deutsche Geodätische Kommission (DGK), Reihe C, Nr. 441, München.
- SESTER, M., ANDERS, K.-H. und WALTER, V. (1998): *Linking Objects of Different Spatial Data Sets by Integration and Aggregation*. In: *Geoinformatica*, Jahrgang 2, Heft 4: S. 335–358.
- SESTER, M., VON GOESSELN, G. und KIELER, B. (2007): *Identification and adjustment of corresponding objects in data sets of different origin*. In: *10th AGILE International Conference on Geographic Information Science 2007*.
- SHEEREN, D., MUSTIÈRE, S. und ZUCKER, J. (2004): *How to Integrate Heterogeneous Spatial Databases in a Consistent Way?* In: GOTTLOB, G., BENCZÚR, A. A. und DEMETROVICS, J. [Hrsg.]: *Advances in Databases and Information Systems, 8th East European Conference, ADBIS 2004, Lecture Notes in Computer Science*, Bd. 3255, S. 364–378. Springer, Budapest, Hungary.
- SHEEREN, D., MUSTIÈRE, S. und ZUCKER, J.-D. (2009): *A data-mining approach for assessing consistency between multiple representations in spatial databases*. In: *International Journal of Geographical Information Science*, Jahrgang 23, Heft 8: S. 961–992.
- SIRIBA, D., EGGERT, D. und SESTER, M. (2011): *Relative Curve Orientation in the Alignment of Inconsistent Linear Datasets*. In: *Proceedings of AGILE 2011*. Utrecht, The Netherlands.
- STIGMAR, H. (2005): *Matching Route Data and Topographic Data in a Real-Time Environment*. In: *Proceedings of the Scandinavian Research Conference on Geographical Information Sciences*. Stockholm, Sweden.
- THAKKAR, S., KNOBLOCK, C. A. und AMBITE, J. L. (2007): *Quality-driven geospatial data integration*. In: SAMET, H., SHAHABI, C. und SCHNEIDER, M. [Hrsg.]: *Proceedings of the 15th ACM International Symposium on Geographic Information Systems*. Seattle, Washington, USA.
- THEOBALD, D. M. (2001): *Understanding Topology and Shapefiles*. In: ArcUser. URL <http://www.esri.com/news/arcuser/0401/topo.html>, letzter Zugriff: 19.07.2013.
- TN-ITS (2015): *TN-ITS Project Website*. ERTICO Partnership Activities. URL [www.tn-its.eu](http://www.tn-its.eu), letzter Zugriff: 25.08.2015.

- TÖNNESMANN, J. (2015): *Wenn möglich, bitte nicht wenden!* ZEIT ONLINE. URL [www.zeit.de/2015/32/hererkartendienst-autoindustrie-bmw-audi-daimler-nokia](http://www.zeit.de/2015/32/hererkartendienst-autoindustrie-bmw-audi-daimler-nokia), letzter Zugriff: 16.08.2015.
- TOMTOM (2011): *MultiNet Shapefile Format Specification 4.6*. TomTom Global Content B.V. CONFIDENTIAL.
- TOMTOM (2012): *OpenLR White Paper*. Version 1.5 Revision 2. TomTom International B.V., Amsterdam, The Netherlands.
- TOMTOM (2014): *OpenLR Website – Downloads*. TomTom International B.V. URL <http://openlr.org/download.html>, letzter Zugriff: 24.02.2014.
- TOMTOM (2015): *TomTom Corporate Website – Unternehmensdaten*. TomTom International B.V. URL <http://corporate.tomtom.com/keyfacts.cfm>, letzter Zugriff: 20.03.2015.
- T' SIOBBEL, S., REINI, J. und SCHÜTZLE, R. (2008): *Integration of European Public Authorities Safety Attributes in Commercial Digital Maps Databases*. In: *Proceedings of the 15th World Congress on Intelligent Transportation Systems*. ITS America, New York, USA.
- VAN OORT, P. (2005): *Spatial data quality: from description to application*. Dissertation, Netherlands Geodetic Commission, Publications on Geodesy 60, Delft, Niederlande. URL [www.ncgeo.nl/phocadownload/60oort.pdf](http://www.ncgeo.nl/phocadownload/60oort.pdf), letzter Zugriff: 19.02.2013.
- VELTKAMP, R. C. (2001): *Shape Matching: Similarity Measures and Algorithms*. Technical Report, Department of Computing Science, Utrecht University, Utrecht, The Netherlands.
- VELTKAMP, R. C. und HAGEDOORN, M. (2001): *State of the art in shape matching*. In: LEW, M. S. [Hrsg.]: *Principles of visual information retrieval*, S. 87–119. Springer-Verlag, London, UK.
- VEREGIN, H. (1999): *Data Quality Parameters*. In: LONGLEY, P. A., GOODCHILD, M. F., MAGUIRE, D. J. und RHIND, D. W. [Hrsg.]: *Geographical Information Systems*, Bd. 1 – Principles and Technical Issues, 2. Aufl., Kap. 12, S. 177–189. John Wiley & Sons, New York, USA.
- VOLOTÃO, C. F. S., SANTOS, R. D. C., ERTHAL, G. J. und DUTRA, L. V. (2010): *Shape characterization with turning functions*. In: *Proceedings of the 17th international Conference on Systems, Signals and Image Processing IWSSIP*, S. 554–457. Rio de Janeiro, Brasilien.
- VOLZ, S. (2006a): *An iterative approach for matching multiple representations of street data*. In: *Proceedings of the joint ISPRS workshop on multiple representations and interoperability of spatial data*. Hannover.
- VOLZ, S. (2006b): *Modellierung und Nutzung von Relationen zwischen Mehrfachrepräsentationen in Geo-Informationssystemen*. Dissertation, Institut für Photogrammetrie, Universität Stuttgart. URL <http://elib.uni-stuttgart.de/opus/volltexte/2006/2785>, letzter Zugriff: 28.02.2013.
- VOSSELMANN, G. (1992): *Relational Matching, Lecture Notes in Computer Science*, Bd. 628. Springer Berlin Heidelberg.
- WALTER, V. (1997): *Zuordnung von raumbezogenen Daten – am Beispiel ATKIS und GDF*. Dissertation, Deutsche Geodätische Kommission (DGK), Reihe C, Nr. 480, München.
- WALTER, V. und FRITSCH, D. (1995): *Matching techniques for road network data in different data models*. In: SOLIMAN, J. und ROLLER, D. [Hrsg.]: *Proceedings of the 28th International Symposium on Automotive Technology and Automation*, S. 633–640. Croydon: Automotive Automation Limited, Stuttgart.
- WALTER, V. und FRITSCH, D. (1999): *Matching spatial data sets: a statistical approach*. In: *International Journal of Geographical Information Science*, Jahrgang 13, Heft 5: S. 445–473.
- WARTENBERG, M. (2006a): *Algorithms for location Referencing*. In: *Jahresbericht der Deutschen Mathematikervereinigung (DMV)*, Jahrgang 108, Heft 1: S. 33–44.
- WARTENBERG, M. (2006b): *On the choice of information to be transmitted using methods for location referencing*. In: *Proceedings of the 13th World Congress on Intelligent Transport Systems*. London, Great Britain.
- WARTENBERG, M. (2008): *Mathematical Methods for Location Referencing*. Shaker Verlag, Aachen.
- WEVERS, K. (2000): *Improved on-the-fly location referencing – a proposal for road section coding*. In: *Proceedings of the 7th World Congress on Intelligent Transport Systems*. Turin, Italy.

- WEVERS, K. (2012): *Map-based location referencing: status and prospects*. In: *Proceedings of the 19th World Congress on Intelligent Transport Systems*. Vienna, Austria.
- WEVERS, K., DUCKECK, R., HENDRIKS, T., OTTO, H.-U. und PFEIFFER, H.-W. (2001): *Specifications of algorithms for implementation in the trials*. Deliverable 2.1, AGORA-Consortium.
- WEVERS, K., DUCKECK, R., HENDRIKS, T., OTTO, H.-U. und PFEIFFER, H.-W. (2003): *Specification of the AGORA location referencing method*. Deliverable 2.2, AGORA-Consortium.
- WEVERS, K. und HENDRIKS, T. (2003): *Compact on-the-fly location referencing – the AGORA-C specification*. In: *Proceedings of the 10th World Congress on Intelligent Transport Systems*. Madrid, Spain.
- WEVERS, K. und HENDRIKS, T. (2006): *AGORA-C map-based location referencing*. In: *Transportation Research Record*, Nr. 1972: S. 115–122.
- WEVERS, K., HENDRIKS, T., PFEIFFER, H.-W., DUCKECK, R., HESSLING, M. und OTTO, H.-U. (2002): *Map-based on-the-fly Location Referencing*. In: *Proceedings of the 9th World Congress on Intelligent Transport Systems*. Chicago, USA.
- WIKSTRÖM, L., LANDWEHR, M., BOCK, H., NASR, A., WEVERS, K., SVENSK, P.-O., T'SIOBBEL, S., BOTERBERGH, B., BLAIVE, L., KEITH, H., ISAKSSON, P., BERGMAN, O., HASPEL, U. und SCHÜTZLE, R. (2009): *Specification of data exchange methods*. Deliverable 3.1, ROSATTE-Consortium.
- WILTSCHKO, T. (2004): *Sichere Information durch infrastrukturegestützte Fahrerassistenzsysteme zur Steigerung der Verkehrssicherheit an Straßennotenpunkten*. Nr. 570 in *Fortschritt-Berichte VDI*, Reihe 12. VDI Verlag, Düsseldorf.
- WILTSCHKO, T. und KAUFMANN, T. (2004): *Report on quality frame for information*. Deliverable 2.2, EuroRoadS-Consortium.
- WILTSCHKO, T. und KAUFMANN, T. (2005): *Modellierung und Bewertung der Qualität in Prozessen der Geodatenverarbeitung*. In: *Beiträge zum AGIT-Symposium 2005*. Salzburg, Österreich.
- WU, Z. und PALMER, M. (1994): *Verbs Semantics and Lexical Selection*. In: *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics, ACL '94*, 133–138. Association for Computational Linguistics, Stroudsburg, PA, USA.
- XIONG, D. (2000): *A three-stage computational approach to network matching*. In: *Transportation Research Part C: Emerging Technologies*, Jahrgang 8, Heft 1–6: S. 71–89.
- XIONG, D. und SPERLING, J. (2004): *Semiautomated matching for network database integration*. In: *ISPRS Journal of Photogrammetry and Remote Sensing, Special Issue on Advanced Techniques for Analysis of Geo-spatial Data*, Jahrgang 59, Heft 1–2: S. 35–46.
- YAMADA, S. (1996): *The strategy and deployment plan for VICS*. In: *Communications Magazine, IEEE*, Jahrgang 34, Heft 10: S. 94–97.
- YUAN, S. und TAO, C. (1999): *Development of Conflation Components*. In: *Geoinformatics and Socioinformatics – The Proceedings of Geoinformatics'99 Conference*. Ann Harbor, Michigan, USA.
- ZHANG, D. und LU, G. (2004): *Review of shape representation and description techniques*. In: *Pattern Recognition*, Jahrgang 37, Nr. 1: S. 1–19.
- ZHANG, J., ZHU, Y. und MENG, L. (2011): *Conflation of road network and geo-referenced image using sparse matching*. In: CRUZ, I. F., AGRAWAL, D., JENSEN, C. S., OFEK, E. und TANIN, E. [Hrsg.]: *19th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems, ACM-GIS 2011*, S. 281–288. ACM, Chicago, IL, USA.
- ZHANG, L. und SCHWIEGER, V. (2011): *Ein Echtzeit-Qualitätskontrollsystem für Wohnhausbauprozesse*. In: *Allgemeine Vermessungsnachrichten*, Jahrgang 118, Heft 11–12: S. 368–380.
- ZHANG, M. (2009): *Methods and Implementations of Road-Network-Matching*. Dissertation, Technische Universität München. URL <http://mediatum.ub.tum.de?id=820125>, letzter Zugriff: 02.05.2013.
- ZHANG, M. und MENG, L. (2007): *An iterative road-matching approach for the integration of postal data*. In: *Computers, Environment and Urban Systems*, Jahrgang 31, Heft 5: S. 597–615.
- ZHANG, M. und MENG, L. (2008): *Delimited Stroke Oriented Algorithm-Working Principle and Implementation for the Matching of Road Networks*. In: *Journal of Geographic Information Sciences*, Jahrgang 14, Heft 1: S. 44–53.
- ZHANG, M., MENG, L. und BOBRICH, J. (2010): *A road-network matching approach guided by 'structure'*. In: *Annals of GIS*, Jahrgang 16, Heft 3: S. 165–176.

- ZHANG, M., SHI, W. und MENG, L. (2005): *A generic matching algorithm for line networks of different resolutions*. In: *Proceedings of 8th ICA Workshop on Generalisation and Multiple Representation*. A Coruña, Spain.
- ZHANG, M., YAO, W. und MENG, L. (2012): *Enrichment of topographic road database for the purpose of routing and navigation*. In: *International Journal of Digital Earth*, Jahrgang 7, Heft 5: S. 411–431.
- ZOLLONDZ, H.-D. (2011): *Grundlagen Qualitätsmanagement*. 3. Aufl. Oldenbourg, München.



# Anhang A OpenStreetMap Datenprozessierung

## A.1 Datenausschnitt mit osmosis

Zunächst muss der räumlich wie auch thematisch für die weiteren Arbeiten benötigte Inhalt aus der originären OSM-Datei ausgeschnitten werden. Dazu bietet sich das Freeware-Tool *osmosis* (OSMOSIS 2014) an, mit dem man OSM-Daten prozessieren kann. Zur Verwendung in der LR-Testumgebung wurde für das jeweilige Testgebiet ein Ausschnitt über eine BoundingBox ausgeschnitten. Da sich die Tests im Rahmen dieser Arbeit lediglich auf das Straßennetzwerk konzentrieren, wurden nur Kanten mit dem OSM-Tag *highway* zugelassen und alle anders gekennzeichneten Kanten aus dem Datensatz verworfen. Zur weiteren Reduzierung der Dateigröße wurden nur solche Knoten belassen, die mit einer der akzeptierten Kanten verbunden sind. Durch die Beschränkung auf die für die jeweiligen Tests notwendigen Inhalte und die damit verbundene Reduzierung des Speicherbedarfs der jeweiligen OSM-Karte, kann die Prozessierung innerhalb der LR-Testumgebung wesentlich effizienter und schneller erfolgen.

## A.2 Routingfähiges Netzwerk erzeugen mit osm2po

OSM-Daten sind von sich aus nicht notwendigerweise auch routingfähig. Ein (Straßen-) Netzwerk wird als routingfähig bezeichnet, wenn alle aneinander grenzenden Kanten, die eine Verbindung im Netz darstellen, durch einen Knoten verbunden sind. In der Regel ist dies auch bei OSM der Fall. In Ausnahmefällen gibt es aber auch Kanten, die sich schneiden ohne durch einen Knoten verbunden zu sein. In diesem Fall kann ein Routingalgorithmus keine Verbindung zwischen den Kanten herstellen. Kreuzen sich Straßen etwa über eine Brücke oder einen Tunnel, so ist diese Überlagerung gewollt und die Kanten sollen nicht durch einen Knoten aufgetrennt werden. Es existieren jedoch Fälle, bei denen die sich überlagernden Kanten nicht als Brücken oder Tunnels durch entsprechende Attributwerte gekennzeichnet sind. An einer solchen 'unechten' Kreuzung kann ein Routingalgorithmus beispielsweise nicht auf die kreuzende Straße abbiegen, was eine korrekte Routenberechnung unmöglich macht. Diese Fälle gilt es im Folgenden im Datensatz zu identifizieren und zu bereinigen.

Dazu ist es notwendig, alle Kanten eines Datensatzes auf mögliche räumliche Überlagerung zu überprüfen. Ist dies der Fall, so werden die Kanten am Schnittpunkt aufgetrennt und an dieser Stelle ein neuer Knoten eingefügt, in dem sich dann alle Schnittkanten treffen. Da diese neu eingefügten Kanten auch einen Bezeichner benötigen, sind die von OSM vergebenen Objekt-Identifikatoren nicht mehr eindeutig und müssen durch neu einzuführende IDs ersetzt werden. Bei der Verarbeitung ist dafür Sorge zu tragen, dass Fälle, in denen sich Kanten beabsichtigter Weise schneiden (etwa bei Brücken oder Tunnels) nicht aufgetrennt werden. Diese Situationen sind durch das OSM Tag 'layer' sowie 'bridge' und 'tunnel' in den Daten gekennzeichnet.

Diese Auftrennung der Kanten kann sehr effizient mit dem Freeware-Tool *osm2po* (MÖLLER 2012) erfolgen. Dieses benötigt als Eingangsdaten den zuvor mit *osmosis* erzeugten OSM-Ausschnitt als \*.osm-Datei. Das Ergebnis wird als PostgreSQL/ PostGIS Datenbank gespeichert. Diese kann dann beispielsweise mit Hilfe des Freeware Geoinformationssystems *QGIS* (QGIS 2014) geöffnet und zur weiteren Verarbeitung in der LR-Testumgebung in ein ESRI Shapefile konvertiert werden, wie im nächsten Abschnitt beschrieben.

### A.3 Erzeugen von Shapefiles mit spezifizierten Attributinformationen mit ogr2ogr

Die mit *osm2po* erzeugten Dateien enthalten lediglich eine rudimentäre und nicht zu beeinflussende Auswahl der ursprünglich in der OSM-Datei enthaltenen Attribute, die für die Verwendung mit OpenLR nicht ausreichend ist. Deshalb ist es notwendig, mit Hilfe des Freeware-Tools *ogr2ogr* (OGR2OGR 2014) ein weiteres Shapefile-Extrakt des mit *osmosis* erzeugten OSM-Ausschnittes zu erzeugen. Dabei kann in einer Steuerdatei genau spezifiziert werden, welche Attributinformationen aus der OSM-Datei mit in das Shapefile übernommen werden sollen. OGR2OGR ist ein Konverter-Tool aus der GDAL/OGR Bibliothek zur Konvertierung von Raster- bzw. Vektordaten, das OSM-Daten lesen und in verschiedene andere Vektor-Formate konvertieren kann. Es steht für Microsoft Windows Systeme als Teil der OSGeo4W-Distribution zur Verfügung (OSGEO4W 2014).

### A.4 Zusammenführen von Geometrie und Attributinformationen

Die Geometrie dieses mit *ogr2ogr* erzeugten Shapefiles entspricht der originären OSM-Datei und ist wie zuvor beschrieben, nicht routingfähig. Daher müssen die Geometrie-Informationen aus *osm2po* mit den Attribut-Informationen aus *ogr2ogr* zusammengeführt werden. Den Geometrie-Objekten des Datensatzes mit dem routingfähigen Straßennetzwerk können über die neu vergebenen Objekt-Identifikatoren die Attributinformationen des zweiten Datensatzes mittels Attribut-Join, beispielsweise in ArcGIS, hinzugefügt werden. Das Ergebnis ist ein Shapefile, das sowohl eine routingfähige Geometrie des Straßennetzes als auch die für die weitere Verwendung der Daten für das Location Referencing notwendigen Attributinformationen verfügt.

## Anhang B LocationViewer

### B.1 Umsetzung in ArcGIS Toolbox

Wie bereits in Kapitel 7.1.5 beschrieben, werden sämtliche in der LR-Testumgebung prozessierten Ortsreferenzen zusammen mit ihrer Quell- und Zielkarte zu Auswerte- und Analysezielen in einem Geoinformationssystem visualisiert. Im Rahmen der vorliegenden Arbeit kam *ESRI ArcGIS Desktop 10.0 (SP 5)* zum Einsatz. Über die darin enthaltene *ArcToolbox* können u.a. *Python*-Skripte zur Automatisierung von Aufgaben erstellt, aufgerufen und ausgeführt werden. Für die LR-Testumgebung wurden solche Skripte zum automatisierten Einlesen von Daten-Dateien und zur Aufbereitung der Darstellung erstellt. Diese Skripte wurden als *LocationViewer*-Tool in die *ArcToolbox* eingebunden und können von dort aufgerufen und gestartet werden. Nach dem Aufruf erscheint die in Abbildung B.1 dargestellte Benutzeroberfläche, über welche die notwendigen Pfade zu den Eingangsdaten sowie die zur Ausführung notwendigen Optionen angegeben bzw. gesetzt werden können.

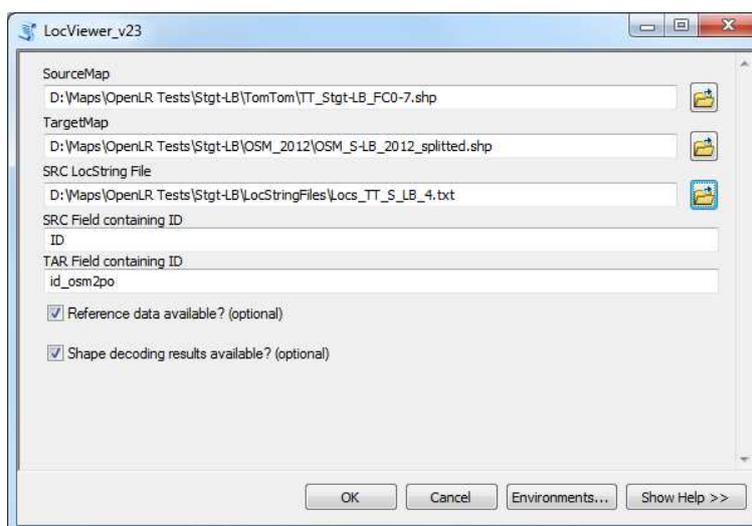


Abbildung B.1: Benutzeroberfläche Location Viewer

### B.2 Eingangsdaten

Die Visualisierung basiert auf den Ergebnis-Dateien der LR-Testumgebung: neben den Ortsreferenzen werden auch die zugehörigen Location Reference Points und für das gesamte Testgebiet die Quell- und Zielkarte dargestellt. Neben den Pfaden zu den Shape-Files der beiden Karten ist der Dateipfad zum Location String File der Ausgangslocations (engl. Source) anzugeben. In diesem Ordner müssen sich ebenfalls alle weiteren Datendateien befinden. Sie werden vom LocationViewer automatisch eingelesen. Da die Objekt-Identifikatoren in den verwendeten Karten unterschiedliche Attributbezeichnungen tragen, müssen diese ebenfalls in der LocationViewer Benutzeroberfläche angegeben werden. Anhand

dieser Bezeichner ist es dem Programm möglich, die in den Location String Files angegebenen Kanten den jeweiligen Ortsreferenzen zuzuordnen.

Konkret werden die folgenden Dateien eingelesen:

1. Location-String Dateien:
  - Ausgangs-Locations in Quellkarte
  - Soll-Location in der Zielkarte
  - Decodier-Ergebnis (OpenLR)
  - Decodier-Ergebnis (Form-Matching)
2. Location Reference Dateien:
  - Encodiert
  - Decodiert (OpenLR)
  - Decodiert (Form-Matching)
3. ShapeFiles für Quell- und Zielkarte

## B.3 Programmablauf

Neben den originären Python-Befehlen, z.B. für den Dateizugriff und den internen Ablauf der Skripte (z.B. Schleifen, Abfragen, etc.) ist vor allem der Zugriff auf die zusammen mit ArcGIS zur Verfügung gestellte *arcpy*-Bibliothek für die Automatisierung von GIS-Aufgaben sehr hilfreich. Diese Bibliothek enthält umfangreiche GIS-Funktionalitäten zur Bearbeitung und Analyse von Geodaten.

### Dateien laden

Zunächst werden die ShapeFiles der Quell- und Zielkarten sowie die Location-String- und Location Reference Point-Dateien geladen. Das Einlesen der Kartendaten erfolgt über spezielle Funktionen aus der *arcpy*-Bibliothek. Der Inhalt der Daten-Dateien muss jedoch mittels eines eigens entwickelten Parsers entsprechend ihrem definierten Datenformat aufgetrennt und die einzelnen Datenblöcke entsprechend verarbeitet und zwischengespeichert werden.

### Neue Feature Classes erstellen

Für die zu visualisierenden Ortsreferenzen und Location Reference Points müssen jeweils separat eine neue Feature Class vom Typ *Polylinie* bzw. *Punkt* erstellt werden, in welche die entsprechenden Geometrie-Elemente in einem späteren Verarbeitungsschritt gespeichert werden können. Das Anlegen der neuen Feature Classes erfolgt mittels arcpy-Funktion *CreateFeatureclass\_management()*.

### Prozessierung der Ortsreferenzen

Zum Verarbeiten der Ortsreferenzen werden die Daten-Dateien für die Quellkarte sowie die Dateien mit den mittels OpenLR und Form-Matcher in der Zielkarte decodierten Ortsreferenzen sequenziell durchlaufen. Dabei wird für jede Ortsreferenz ihre ID, ggfs. definierte Offsets sowie die Liste mit Kanten-IDs eingelesen. Diese Kanten werden in der jeweiligen Karte gesucht und die Koordinaten der sie beschreibenden Punkte ausgelesen. Aus diesen Punktkoordinaten wird dann eine neue Polylinie zur Beschreibung des Verlaufs der Ortsreferenz erstellt. Besteht eine Ortsreferenz aus mehr als einer Kante, dann ist beim Zusammensetzen der formbeschreibenden Punkte auf deren richtige Reihenfolge zu achten: Ist die Ortsreferenz entgegen der Digitalisier-Richtung einer Kante definiert, werden die formbeschreibenden

Punkte in umgekehrter Reihenfolge prozessiert. Enthält die Ortsreferenz Offsets am Beginn und/oder an ihrem Ende, so sind die Koordinaten der entsprechenden Punkte durch Einrechnen in die jeweilige Linie zu bestimmen. Für jede Ortsreferenz werden nun drei Features in der entsprechenden Feature Class angelegt: für den eigentlichen Verlauf der Ortsreferenz sowie für den Start- und End-Offset. Für die Location Reference Points werden die entsprechenden Datendateien in ähnlicher Weise durchlaufen und mit den darin abgelegten Punktkoordinaten und seinem Bezeichner für jeden LRP ein Point-Feature angelegt.

### Darstellung und Symbolisierung

Die so erzeugten Feature Classes sollen im aktiven ArcGIS-Arbeitsbereich (Workspace) entsprechend festgelegter Vorlagen dargestellt werden. Die Ortsreferenzen und Location Reference Points sollen darüber hinaus mit ihrem jeweiligen Bezeichner beschriftet werden.

Zunächst wird mit *arcpy.MakeFeatureLayer\_management()* für jede darzustellende Ortsreferenz ein temporäres Layer-File angelegt und die entsprechende Geometrie (Polylinie) aus der oben erzeugten Feature Class eingefügt. Mit *arcpy.mapping.Layer* wird aus diesem Layer-File ein Layer-Objekt erzeugt, dessen Darstellungsweise mit *arcpy.ApplySymbologyFromLayer\_management()* von einem als Vorlage dienenden Layer-File übernommen werden kann.

Aus Gründen der Übersichtlichkeit wird für jede Ortsreferenz und jeden Location Reference Point ein separater Layer angelegt. Alle zu einer Ortsreferenz zugehörigen Layer werden dann mittels Gruppen-Layer zusammengeführt.

## B.4 Ergebnis / Darstellung in ArcGIS

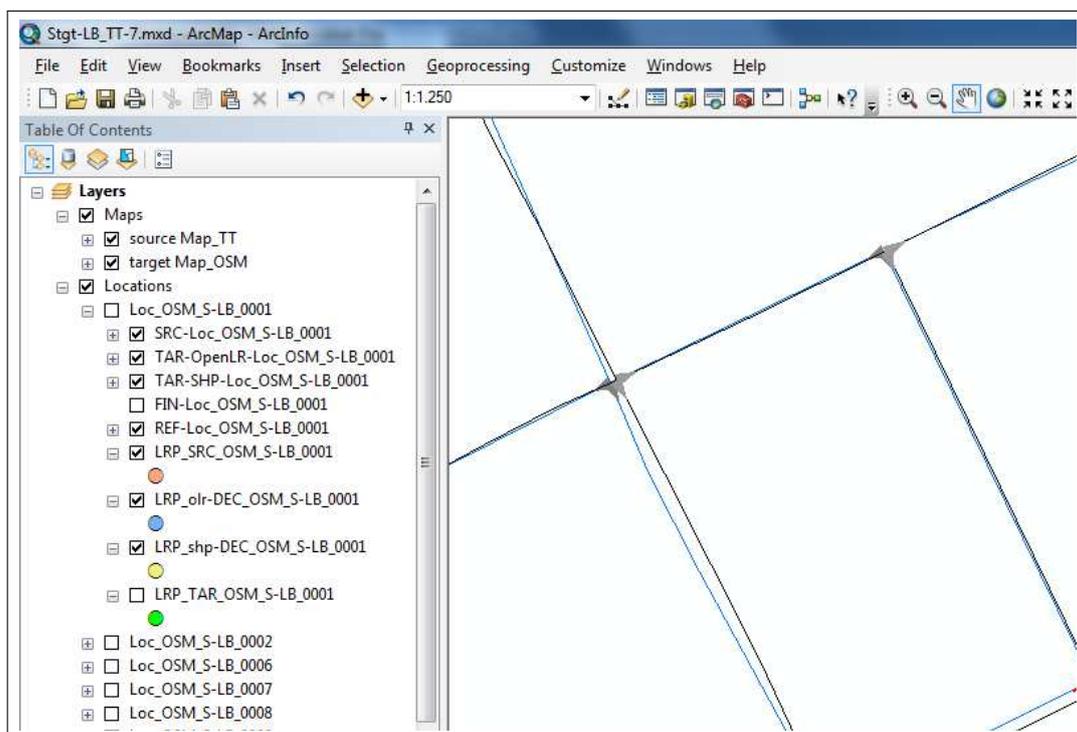


Abbildung B.2: Erzeugter ArcGIS Arbeitsbereich

Nach erfolgreichem Programmdurchlauf werden die so erzeugten Layer und Gruppen-Layer automatisch im aktiven Arbeitsbereich von ArcGIS geladen. Ein solcher Arbeitsbereich ist in Abbildung B.2 beispielhaft und auszugsweise dargestellt. Die so gewonnene Gliederung der Daten ermöglicht die getrennte Darstellung aller verarbeiteten und decodierten Ortsreferenzen und Location Reference Points. Sämtliche Layer können unabhängig voneinander aktiviert bzw. deaktiviert und das so erzeugte *Overlay* individuell für die jeweilige Analyse-Aufgabe zusammengestellt werden.

## Danksagung

Ich danke allen, die mich beim Erstellen dieser Arbeit unterstützt haben. Ganz besonders möchte ich Herrn Prof. Dr.-Ing. habil. Volker Schwieger für die Übernahme der Betreuung dieser Arbeit sowie das mir stets entgegengebrachte Vertrauen und die wissenschaftliche Förderung und Begleitung in den zurückliegenden Jahren danken.

Ebenso danke ich Frau Prof. Dr.-Ing. Liqiu Meng sowie Herrn Prof. Dr.-Ing. habil. Dieter Fritsch für die freundliche Übernahme der Koreferate.

Herrn Dr.-Ing. Martin Metzner danke ich besonders für seine Unterstützung während meiner gesamten Zeit am Institut und die vielen anregenden Gespräche und Diskussionen.

Nicht zuletzt danke ich meiner Familie, die oft auf mich verzichten musste und mich in unglaublichem Maße auf meinem beruflichen Weg unterstützt hat. Dieser Rückhalt hat entscheidend zum Gelingen dieser Arbeit beigetragen.



# Lebenslauf

## Zur Person

Name Rainer Schützle  
 Geburtsdatum 22.05.1979  
 Geburtsort Heilbronn

## Berufstätigkeit

seit 02/2016 Entwicklungsingenieur, Daimler AG, Sindelfingen  
 04/2014 – 01/2016 Doktorand am Institut für Ingenieurgeodäsie, Universität Stuttgart  
 04/2008 – 03/2014 Wissenschaftlicher Mitarbeiter am Institut für Ingenieurgeodäsie, Universität Stuttgart  
 11/1999 – 08/2000 Grundwehrdienst, Topographie-Batterie 850, Ulm

## Ausbildung

10/2006 – 03/2008 Universität Stuttgart, Studium *Geomatics Engineering*  
 Abschluss: Master of Science  
 10/2007 – 03/2008 Robert Bosch GmbH, Leonberg  
 Master Thesis: Eigenbewegungsschätzung eines Kraftfahrzeuges  
 03/2005 – 09/2006 DaimlerChrysler AG, Böblingen  
 Praktikum und Werkstudententätigkeit in der Forschungsabteilung  
 kartenbasierte Assistenzsysteme  
 10/2000 – 07/2004 Fachhochschule Stuttgart – Hochschule für Technik  
 Studium *Vermessung und Geoinformatik*,  
 Abschluss: Diplom-Ingenieur (FH)  
 03/2004 – 06/2004 Landesvermessungsamt Baden-Württemberg, Karlsruhe  
 Diplomarbeit: Entwicklung eines Testmanagements für SAPOS  
 09/2002 – 02/2003 DaimlerChrysler AG, Esslingen  
 Praxissemester in der Forschungsabteilung Assistenzsysteme  
 09/1998 – 07/1999 Wilhelm-Maybach-Schule, Heilbronn  
 Abschluss: Fachhochschulreife  
 09/1995 – 08/1998 Staatliches Vermessungsamt Heilbronn  
 Berufsausbildung zum Vermessungstechniker  
 09/1989 – 07/1995 Mörike-Realschule Heilbronn  
 Abschluss: Mittlere Reife