

**DEUTSCHE GEODÄTISCHE KOMMISSION**  
bei der Bayerischen Akademie der Wissenschaften

---

Reihe C

Dissertationen

Heft Nr. 562

**Heiner Hild**

**Automatische Georeferenzierung  
von Fernerkundungsdaten**

**München 2003**

---

**Verlag der Bayerischen Akademie der Wissenschaften  
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München**

ISSN 0065-5325

ISBN 3 7696 5001 8

Automatische Georeferenzierung  
von Fernerkundungsdaten

Bei der Fakultät für Bauingenieur- und Vermessungswesen  
der Universität Stuttgart  
zur Erlangung der Würde eines  
Doktor-Ingenieurs (Dr.-Ing.)  
eingereichte Dissertation

vorgelegt von

Dipl.-Phys. Heiner Hild  
aus Ludwigsburg

München 2003

---

Verlag der Bayerischen Akademie der Wissenschaften  
in Kommission bei der C. H. Beck'schen Verlagsbuchhandlung München

Adresse der Deutschen Geodätischen Kommission:

**Deutsche Geodätische Kommission**

Marstallplatz 8 • D – 80 539 München

Telefon (089) 23 031 113 • Telefax (089) 23 031 – 283/– 100

E-mail hornik@dgfi.badw.de • <http://www.dgfi.badw.de/dgfi/DGK/dgk.html>

Hauptberichter: Prof. Dr.-Ing.habil. Dieter Fritsch

Mitberichter: Prof. Dr.-Ing. habil. Dr. tech. h.c.mult.Dr.-Ing. Erik Grafarend  
Prof. Dr.-Ing. habil. Dr.-Ing.h.c. Hans-Peter Bähr

Tag der Einreichung: 20.11.2001

Tag der mündlichen Prüfung: 28.02.2002

---

© 2003 Deutsche Geodätische Kommission, 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

## Kurzfassung

Die Weiterverarbeitung bildbasierter Information innerhalb eines Geoinformationssystems in Photogrammetrie und Fernerkundung setzt georeferenzierte Daten voraus. Während die direkten Verfahren zur Georeferenzierung von sich aus einen hohen Automationsgrad besitzen, sind bei den indirekten Verfahren nach wie vor manuelle Messungen unerlässlich.

Diese Arbeit stellt ein vollständig implementiertes Verfahren zur automatischen indirekten Georeferenzierung von Fernerkundungsdaten vor. Das entwickelte Verfahren basiert auf Merkmalen von Polygonen, die in den zuzuordnenden Datensätzen extrahiert wurden. Während Bilddaten naturgemäß rasterförmig sind, können Referenzdaten für das vorgestellte Verfahren raster- oder vektorförmig sein. Eine Einschränkung in Form von Näherungswerten besteht dabei grundsätzlich nicht.

Die Erzeugung polygonaler Objekte aus Bilddaten geschieht mit Hilfe multipler Segmentierung. Zur Bewertung der Ergebnisse des gewählten Bereichswachstumsverfahrens mit adaptivem Schwellwert werden verschiedene Untersuchungen durchgeführt. Die Extraktion geeigneter Polygone aus GIS-Referenzdaten erfolgt durch Verschmelzung benachbarter Polygone gleicher Objektart. In dieser Arbeit werden dabei ausschliesslich ATKIS-Daten benutzt.

Aufbauend auf der Theorie affiner Invarianten algebraischer Formen werden Merkmale entwickelt, die sich aus geometrischen Momenten von Raster- und Vektorpolygonen berechnen lassen. Alternativ dazu werden aus den sogenannten Cross-Weighted-Moments-Invarianten Merkmale hergeleitet, die ebenfalls aus geometrischen Momenten von Raster- und Vektorpolygonen berechnet werden können. Untersuchungen zur Robustheit der Merkmale und zur Unterscheidbarkeit ähnlicher Objekte werden an synthetischen und realen Daten durchgeführt.

Die Zuordnung von Bild- und Referenzdaten geschieht mit Hilfe eines Baumsuchverfahrens mit unären und binären Zwangsbedingungen. Zur Formulierung der Zwangsbedingungen wird paarweise eine affine Näherungstransformation aus den geometrischen Momenten berechnet. Das Verfahren basiert auf der Theorie der Standardlagen. Ein abschließender Verifikationsschritt liefert eine globale affine Transformation zwischen Bild- und Referenzdatensystem. Die berechnete Näherungstransformation wird individuell mit Hilfe des Iterative-Closest-Point-Algorithmus' verbessert. Basierend auf den paarweise optimalen Transformationen werden Passpunktpaare automatisch auf den Polygonkonturen erzeugt.

Aufbauend auf einer theoretischen Betrachtung werden anhand dreier synthetischer und realer Datensätze umfangreiche Untersuchungen zur erreichbaren Paßpunktgenauigkeit durchgeführt. Durch gezielte Anbringung von Störungen wird die Leistungsfähigkeit und Robustheit des entwickelten Verfahrens bei unterschiedlichen Eingangsdaten dargestellt.

## Abstract

In photogrammetry and remote sensing, the processing of image based information within a GIS is based on georeferenced data. The methods of direct georeferencing naturally have a high level of automation where indirect methods still need manual interaction.

Within this thesis, a completely implemented system for automatic indirect georeferencing is presented. The developed procedure is based on features of polygons being extracted in the data sets to be matched. Where image data are in raster form by nature, reference data for the outlined system can be in raster or in vector form. There do not exist any limitations in form of approximate values.

The extraction of polygonal objects from image data is carried out by multiple segmentation. In order to evaluate the results of the chosen region growing algorithm, several examinations are carried out. The extraction of suitable polygons from GIS reference data is done by merging of neighbored polygons with the same object class. Within the framework of this thesis, german ATKIS data are used exclusively.

Based on the theory of affine invariants of algebraic forms, features are developed which can be computed from geometric moments of raster and vector polygons. Alternatively, polygon features are derived from cross weighted moments invariants which can be obtained from geometric moments of raster and vector polygons in a similar way. Studies on the robustness of the features and their ability to discriminate similar objects are carried out by the use of synthetic and real data.

The matching of image and reference data is obtained by a backtracking-depth-first tree search algorithm with unary and binary constraints. In order to formulate the constraints, an affine transformation is computed from the geometric moments for each pair. The procedure is based on the theory of standard positions. A closing verification step leads to a global affine transformation between image and reference system. The approximate transformation is individually refined with the Iterative-Closest-Point algorithm. Based on the optimised affine transformation for each polygon pair, tie points are generated automatically on the borderline of the polygons.

Based on a theoretical consideration, extensive studies on the achievable tie point accuracy were carried out with three synthetic and real data sets. The performance and robustness of the developed system under various input data is demonstrated by controlled application of disturbances.

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>11</b>
1.1	Georeferenzierung von Fernerkundungsdaten	11
1.1.1	Strenge geometrische Transformation	11
1.1.2	Direkte Georeferenzierung	13
1.1.3	Vereinfachte geometrische Prozessierung	13
1.2	Bereitstellung von Paßpunkten	14
1.2.1	Originalmessungen	14
1.2.2	Paßpunkte aus georeferenzierten Daten	14
1.3	Zur Automatisierung der Bildregistrierung	15
1.3.1	Registrierung von Bilddaten gleichen Typs	17
1.3.2	Registrierung von Bilddaten unterschiedlichen Typs	18
1.4	Arbeiten zur automatischen Bild-Modell-Zuordnung	19
1.4.1	Automatische Georeferenzierung von Luftbildern	19
1.4.2	Automatische Georeferenzierung von Satellitenbildern	21
1.5	Zielsetzung und Übersicht über diese Arbeit	22
<b>2</b>	<b>Objektbildung</b>	<b>25</b>
2.1	Flächenobjekte aus Rasterdaten	25
2.1.1	Segmentierung von Rasterbildern	26
2.1.2	Ein Bereichswachstumsverfahren mit adaptivem Schwellwert	28
2.1.3	Segmentierung in verschiedenen Pyramidenstufen	30
2.2	Flächenobjekte aus Vektordaten	32
2.2.1	ATKIS	33
2.2.2	Verschmelzung zusammenhängender ATKIS-Objekten	34
<b>3</b>	<b>Merkmalsgewinnung</b>	<b>36</b>
3.1	Invarianten algebraischer Formen und Momentinvarianten	37
3.1.1	Algebraische Formen und Invarianten	37
3.1.2	Geometrische Momente und einfache Invarianten	39
3.1.3	Das Fundamentaltheorem der Momentinvarianten	41
3.1.4	Momentinvarianten unter allgemeiner Affintransformation	42
3.1.5	Robustheit gegen Rauschen	43
3.1.6	Statistisches Rauschen und Momentinvarianten	43
3.2	Kreuzmomente	46
3.3	Berechnung invarianter Merkmale	50
3.3.1	Geometrische Momente aus Objektkonturen	51
3.3.2	Zur Diskriminierbarkeit mit Kreuzmoment-Invarianten	54
3.4	Weitere Merkmale von Polygonobjekten	55
3.4.1	MBR-basierte Merkmale	55
3.4.2	Fourier-Deskriptoren	58

<b>4</b>	<b>Szene-Modell-Zuordnung</b>	<b>60</b>
4.1	Lösung des Korrespondenzproblems anhand des Interpretationsbaums . . . . .	60
4.2	Zwangsbedingungen zur Verkleinerung des Suchraums . . . . .	61
4.3	Die implementierten Zwangsbedingungen . . . . .	62
4.3.1	Mindestgröße (unär) . . . . .	62
4.3.2	Beziehung der Invariantenvektoren (unär) . . . . .	63
4.3.3	Schwerpunkte und MBRs (binär) . . . . .	63
4.3.4	Objektüberlappung (binär) . . . . .	63
4.4	Das gesamte implementierte Suchverfahren . . . . .	64
4.5	Bestimmung der affinen Transformation zwischen homologen Objekten . . . . .	66
4.5.1	Standardmethode zur Bestimmung einer Standardlage . . . . .	67
4.5.2	Polynommethode zur Bestimmung einer Standardlage . . . . .	67
4.5.3	Iterationsmethode zur Bestimmung einer Standardlage . . . . .	68
4.5.4	Ein Experiment zum praktischen Nutzen der Standardlage nach der Iterationsmethode . . . . .	70
4.6	Illustration des Zuordnungsverfahrens . . . . .	72
4.7	Verifikation der Hypothesen . . . . .	74
4.8	Optimierung individueller Transformationsparameter . . . . .	75
4.8.1	Der Iterative Closest Point (ICP)-Algorithmus . . . . .	76
4.8.2	Modifizierter ICP-Algorithmus . . . . .	78
4.8.3	Einfluß von x-y-Translation . . . . .	79
4.8.4	Einfluß von Rotation . . . . .	80
4.8.5	Einfluß des Maßstabs . . . . .	80
4.9	Bestimmung homologer Punktpaare . . . . .	82
<b>5</b>	<b>Untersuchungen mit dem Gesamtsystem</b>	<b>83</b>
5.1	Annäherung der Zentralperspektive durch eine lineare Transformation . . . . .	83
5.2	Das entwickelte System . . . . .	86
5.3	Untersuchung des Systems bei realer Topographie . . . . .	88
5.3.1	Der Testdatensatz . . . . .	89
5.3.2	Durchführung und Ergebnisse der Untersuchung . . . . .	90
5.3.3	Diskussion . . . . .	94
5.4	Georeferenzierung von SPOT-PAN- mit ATKIS-Daten . . . . .	98
5.4.1	Der Testdatensatz . . . . .	98
5.4.2	Durchführung und Ergebnisse der Untersuchung . . . . .	99
5.4.3	Diskussion . . . . .	103
5.5	Georeferenzierung von SPOT XS-Daten über ein SPOT-PAN-Orthophoto . . . . .	106
5.5.1	Der Testdatensatz . . . . .	106
5.5.2	Durchführung und Ergebnisse der Untersuchung . . . . .	107
5.5.3	Diskussion . . . . .	110
<b>6</b>	<b>Zusammenfassung und Ausblick</b>	<b>112</b>
<b>A</b>	<b>Geometrische Momente aus Objektkonturen</b>	<b>120</b>
<b>B</b>	<b>Zu absoluten Invarianten aus Kreuzmomenten</b>	<b>122</b>
<b>C</b>	<b>Einzelergbnisse der Korrelationsuntersuchungen</b>	<b>123</b>
<b>D</b>	<b>Rotation und Spiegelung ungerader Momente</b>	<b>127</b>
<b>E</b>	<b>Experimente zur Standardlage nach der Iterationsmethode</b>	<b>128</b>

## Abbildungsverzeichnis

1.1	Geometrische Prozessierung von Fernerkundungsdaten. . . . .	11
1.2	Veranschaulichung der zentralperspektiven Bildaufnahme. . . . .	12
1.3	Vereinfachte geometrische Prozessierung von Fernerkundungsdaten. . . . .	14
1.4	Grobes Flußdiagramm für das in dieser Arbeit entwickelte System. . . . .	23
2.1	Ausschnitt aus einer panchromatischen SPOT-Szene zur Segmentierung. . . . .	30
2.2	Segmentierung mit adaptivem Schwellwert bei variabler Lage des Startpunktes. . . . .	31
2.3	Segmentierung mit variablem adaptivem Schwellwert $h$ . . . . .	32
2.4	Ergebnisse der Segmentierung auf den Pyramidenstufen 0 bis 3. . . . .	33
2.5	Konzeption von ATKIS. . . . .	34
2.6	ATKS-Daten der Innenstadt von Karlsruhe. . . . .	35
3.1	Verschiedene Daten, für die Momente berechnet werden können. . . . .	41
3.2	Die Länder der BRD als Beispiel für Vektorpolygone. . . . .	44
3.3	Statistisches Rauschen am Beispiel der Grenze des Bundeslandes Berlin. . . . .	44
3.4	Verhalten der Momentinvarianten $\Psi_1$ bis $\Psi_4$ unter statistischem Rauschen. . . . .	45
3.5	Die Grenze des Bundeslandes Berlin in variablen Pixelgrößen. . . . .	45
3.6	Verhalten der Momentinvarianten bei einer Vektor-Raster-Konversion. . . . .	45
3.7	Affine Transformation eines Dreiecks zur Veranschaulichung der relativen Invarianten. . . . .	47
3.8	Verhalten der Momentinvarianten $\Phi(2, 4)$ bis $\Phi(8, 6)$ unter statistischem Rauschen. . . . .	50
3.9	Verhalten der Kreuzmoment-Invarianten unter Diskretisierung. . . . .	50
3.10	Zur Berechnung geometrischer Momente aus der Objektkontur. . . . .	52
3.11	Repräsentation von Objekten im 2D-Merkmalsraum. . . . .	55
3.12	Untersuchung der Diskriminierbarkeit von Vektorpolygonen. . . . .	55
3.13	Autokorrelationsmatrix und Paare mit minimaler quasi-euklidischer Distanz . . . . .	56
3.14	Egebnisse der Untersuchungen zur Diskriminierbarkeit mit CWM-Invarianten. . . . .	56
3.15	MBR und Quadranten für ein flächenhaftes Objekt. . . . .	57
3.16	Aufhebung der Invarianz MBR-basierter Merkmale unter Scherung. . . . .	58
3.17	Verhalten MBR-basierter Merkmale unter Scherung der Koordinatenachsen. . . . .	59
4.1	Suchbaum für alle möglichen 2er-Paarungen. . . . .	61
4.2	Erweiterter Suchbaum. . . . .	61
4.3	Veranschaulichung der binären MBR-Zwangsbedingung. . . . .	63
4.4	Reduktion der Suchmenge durch Vorsortierung von Objekten. . . . .	65
4.5	Rechenaufwand für die vorgezogene Prüfung. . . . .	65
4.6	Affine Kopien und Standardlage des Umrisses von Baden-Württemberg. . . . .	70
4.7	Erstes Experiment zur praktischen Verwendbarkeit von Standardlagen. . . . .	71
4.8	Zweites Experiment zur praktischen Verwendbarkeit von Standardlagen. . . . .	72
4.9	Anzahl der als akzeptabel eingestuften Rücktransformationen. . . . .	72
4.10	Veranschaulichung des gesamten Zuordnungsprozesses. . . . .	73
4.11	Zuordnungsergebnisse für den Beispieldatensatz. . . . .	73
4.12	Ergebnisse der Baumsuche mit Zwangsbedingungen. . . . .	74
4.13	Illustration des zweistufigen Verifikationsprozesses. . . . .	75
4.14	Flußdiagramm des Iterative-Closest-Point-Algorithmus'. . . . .	76

4.15	Verschiedene Zustände bei Anwendung des ICP-Algorithmus' . . . . .	77
4.16	Fragment des Zielpolygons. . . . .	78
4.17	Konvergenzverhalten des modifizierten ICP-Algorithmus' . . . . .	79
4.18	Polygone zur Untersuchung des modifizierten ICP-Algorithmus' . . . . .	80
4.19	Einfluß einer Translation des Quellobjekts. . . . .	80
4.20	Einfluß einer Rotation des Quellobjekts. . . . .	81
4.21	Einfluß einer Maßstabsänderung des Quellobjekts. . . . .	81
4.22	Vergleich der Ergebnisse aus Standard- und modifiziertem ICP-Algorithmus. . . . .	82
4.23	Veranschaulichung der Paßpunkterzeugung im Geländekoordinatensystem. . . . .	82
5.1	Zweidimensionale perspektive Abbildung. . . . .	84
5.2	Abschätzung für den Bildfehler. . . . .	84
5.3	Produkt aus Bildfehler und Bodenpixelgröße. . . . .	85
5.4	Flußdiagramm des gesamten im Rahmen dieser Arbeit entwickelten Systems. . . . .	87
5.5	ATKIS-Daten des Testgebiets Vaihingen/Enz. . . . .	89
5.6	Verschmolzene Polygone des Testdatensatzes Vaihingen/Enz. . . . .	89
5.7	Digitales Höhenmodell des Testdatensatzes Vaihingen/Enz. . . . .	90
5.8	3D-Ansicht der Luftbildsimulation des Untersuchungsgebiets Vaihingen/Enz. . . . .	91
5.9	Anzahl der Bildpolygone in Abhängigkeit von der Flughöhe und Kammerkonstante. . . . .	91
5.10	Flughöhe $h = 750\text{m}$ , Kammerkonstante $c = 300\text{mm}$ . . . . .	92
5.11	Flughöhe $h = 3000\text{m}$ , Kammerkonstante $c = 150\text{mm}$ . . . . .	92
5.12	Anzahl an Objektpaarhypothesen nach merkmalsbasierter Zuordnung. . . . .	93
5.13	Automatische Erzeugung und Bewertung von Paßpunkten. . . . .	93
5.14	Erzielte Paßpunktgenauigkeiten im Bildraum. . . . .	94
5.15	Abhängigkeit der Paßpunktgenauigkeiten von $\sigma_{NP}$ . . . . .	94
5.16	Abhängigkeit der Paßpunktgenauigkeiten von der Anzahl erzeugter Punkte. . . . .	95
5.17	Beispiele für automatisch erzeugte Paßpunkte in virtuellen Luftbildaufnahmen. . . . .	95
5.18	Beispiele für eine schlechte und eine gute globale affine Sensor-Welt-Transformation. . . . .	96
5.19	Der untersuchte Ausschnitt aus einer SPOT2 PAN-Szene. . . . .	98
5.20	ATKIS-Waldpolygone des Untersuchungsgebietes Elchingen/Großkuchen. . . . .	99
5.21	Ergebnisse der durchgeführten Segmentierungen der SPOT-Szene Elchingen/Großkuchen. . . . .	100
5.22	Statistische Daten zu den untersuchten Durchläufen der merkmalsbasierten Zuordnung. . . . .	100
5.23	Globale Transformation der SPOT-Szene in das ATKIS-System. . . . .	101
5.24	Anzahl erzeugter Paßpunkte in Abhängigkeit von der Nächsten-Nachbar-Standardabweichung. . . . .	102
5.25	Paßpunktgenauigkeit in Abhängigkeit von der Anzahl erzeugter Paßpunkte. . . . .	102
5.26	Beispiel für automatisch erzeugte Paßpunkte. . . . .	103
5.27	Ergebnis der globalen Transformation für künstlich gescherte ATKIS-Polygone. . . . .	103
5.28	Ergebnis für künstlich gescherte, rotierte und gespiegelte ATKIS-Polygone. . . . .	104
5.29	Automatisch erzeugte Paßpunkte für die SPOT-ATKIS-Registrierung. . . . .	105
5.30	SPOT XS Szene des Untersuchungsgebietes Elchingen/Großkuchen. . . . .	106
5.31	SPOT PAN-Orthophoto. . . . .	107
5.32	Alle segmentierten Regionen für die SPOT PAN- und die SPOT XS-Szene. . . . .	108
5.33	Verifizierte Polygonpaare für die SPOT XS-PAN-Registrierung . . . . .	109
5.34	Beispiele für die unterschiedliche Qualität der ICP-verfeinerten Zuordnung. . . . .	109
5.35	Ergebnis der automatischen Paßpunkterzeugung und Georeferenzierung der SPOT XS-Szene. . . . .	110

C.1	Korrelationen für Generalisierung der Stufe 1. . . . .	123
C.2	Korrelationen für Generalisierung der Stufe 2. . . . .	123
C.3	Korrelationen für Generalisierung der Stufe 3. . . . .	124
C.4	Korrelationen für Generalisierung der Stufe 4. . . . .	124
C.5	Korrelationen für Generalisierung der Stufe 5. . . . .	124
C.6	Korrelationen für Generalisierung der Stufe 6. . . . .	125
C.7	Korrelationen für Generalisierung der Stufe 7. . . . .	125
C.8	Korrelationen für Generalisierung der Stufe 8. . . . .	125
C.9	Korrelationen für Generalisierung der Stufe 9. . . . .	126
C.10	Korrelationen für Generalisierung der Stufe 10. . . . .	126
D.1	Die 64 Möglichkeiten für zwei ungerade zentrale Momente $m_{gu}$ und $m_{ug}$ . . . . .	127
E.1	Erstes Experiment: affine Kopie ohne statistisches Rauschen. . . . .	128
E.2	Zweites Experiment: affine Kopie mit statistischem Rauschen $\sigma_{a.u.} = 1$ . . . . .	128
E.3	Drittes Experiment: affine Kopie mit statistischem Rauschen $\sigma_{a.u.} = 2$ . . . . .	129
E.4	Viertes Experiment: affine Kopie mit statistischem Rauschen $\sigma_{a.u.} = 3$ . . . . .	129
E.5	Fünftes Experiment: affine Kopie mit statistischem Rauschen $\sigma_{a.u.} = 4$ . . . . .	129
E.6	Sechstes Experiment: affine Kopie mit statistischem Rauschen $\sigma_{a.u.} = 5$ . . . . .	130

## Tabellenverzeichnis

1.1	Zuordnungsfälle und mögliche Verfahrensgruppen. . . . .	17
3.1	Die vier einfachsten Invarianten binärer algebraischer Formen . . . . .	39
3.2	Physikalische Interpretation geometrischer Momente. . . . .	40
3.3	Invarianzen MBR-basierter Merkmale. . . . .	58
4.1	Die acht möglichen Fälle für Vorzeichen und Betrag zweier Momente. . . . .	69
4.2	Verhalten von Momenten ungerader Ordnung unter Rotationen um Vielfache von $90^\circ$ . . . . .	69
4.3	Anzubringende Rotation und Spiegelung an der $Y$ -Achse. . . . .	69
5.1	Einflußverschiedener Parameter auf den relativen Bildfehler. . . . .	85
5.2	Übersicht über alle Prozesse des entwickelten Verfahrens. . . . .	88
5.3	Anzahl automatisch erzeugter Paßpunkte für $c = 150\text{mm}$ . . . . .	97
5.4	Anzahl automatisch erzeugter Paßpunkte für $c = 300\text{mm}$ . . . . .	97
5.5	Anzahl erzeugter Segmente in Abhängigkeit des adaptiven Schwellwerts $h$ . . . . .	108
5.6	Reduzierte Anzahl erzeugter Segmente in Abhängigkeit des adaptiven Schwellwerts $AT$ . . . . .	108
5.7	Ergebnisse zur Georeferenzierung von SPOT XS- mit SPOT PAN-Daten. . . . .	110
E.1	Akzeptable und nicht akzeptable Rücktransformationen. . . . .	130

## Kapitel 1

# Einleitung

Mit der stetig wachsenden Verbreitung von Geoinformationssystemen (GIS) in den unterschiedlichsten Gesellschaftsbereichen wächst auch der Bedarf an raumbezogenen Informationen verschiedenster Art. Neben der terrestrischen Erfassung spielt dabei die Gewinnung aktueller Information aus Satelliten- oder Luftbilddaten eine wesentliche Rolle. Die heute existierenden Fernerkundungssatelliten stellen hoch auflösende Informationsquellen in einem weit über das sichtbare Spektrum hinaus reichenden Frequenzbereich dar, die in Zukunft noch höherwertigere Informationen liefern werden. Unter Fernerkundung versteht man nach [Lillesand & Kiefer 1987] die Gesamtheit der Verfahren, die Informationen über Objekte, Gebiete und Phänomene durch die berührungslose Aufnahme mit beliebigen Sensoren gewinnen und analysieren. Zur Nutzung des Potenzials von Fernerkundungsdaten innerhalb eines GIS müssen die aufgenommenen Szenen in einen Raumbezug gebracht werden. Jedem Bildpunkt wird dabei eine genaue Position in einem geodätischen Bezugssystem zugeordnet. Je nach Vorverarbeitungsgrad wird diese Aufgabe von den Vertreibern selbst oder vom Nutzer übernommen.

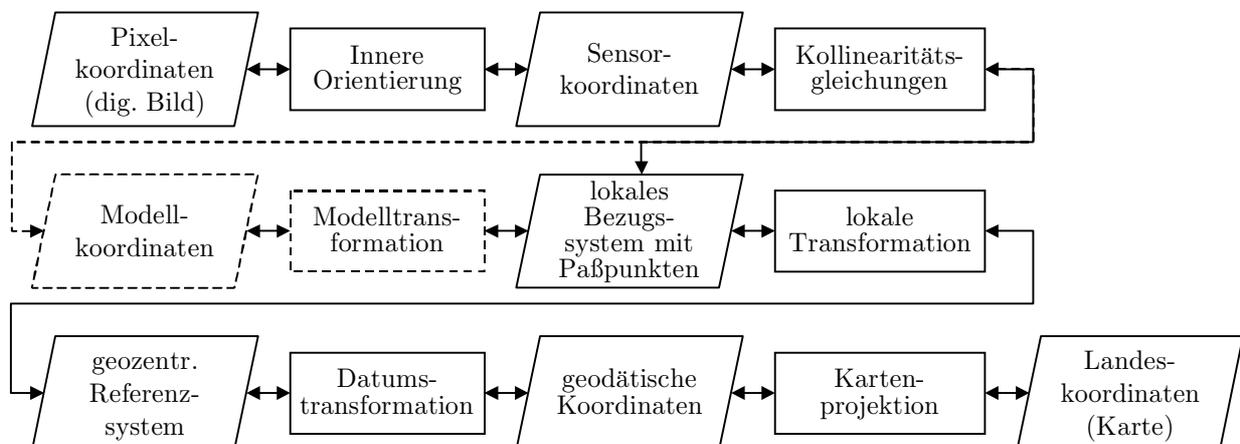
Ein breiter Konsens besteht darin, die gesamte Prozesskette von der Aufnahme bis zur Bereitstellung des gewünschten Endprodukts möglichst weit gehend zu automatisieren. Gerade auch in jüngster Zeit hat sich die Notwendigkeit einer schnellen und genauen Verfügbarkeit hochaktueller Fernerkundungsdaten zur Reaktion auf Katastrophen, Unruhen oder auch terroristische Aktivitäten gezeigt. Neben den zeitlichen und qualitativen Aspekten ist aber mittlerweile auch die Menge an zur Verarbeitung anstehenden Daten so hoch, daß am Übergang von manuellen auf automatische oder zumindest automatisierte Verfahren kein ernsthafter Weg vorbei führt. Der breite Einzug digitaler Systeme in Photogrammetrie und Fernerkundung [Heipke 1995] sowie die zunehmende digitale Aufzeichnung von Fernerkundungsdaten bilden dabei die wesentliche Grundlage für die Erschließung und Etablierung automatischer Verfahren.

## 1.1 Georeferenzierung von Fernerkundungsdaten

Die Herstellung des Raumbezugs von Fernerkundungsdaten jeglicher Art bezeichnet man als Georeferenzierung. Gegenstand der Georeferenzierung ist die Transformation von Bildkoordinaten in ein Landeskoordinatensystem. Die Georeferenzierung stellt somit eine geometrische Transformation der Bilddaten dar. Die Speicherung georeferenzierter Bilder in einem geeigneten Datenformat, etwa GeoTIFF [Ritter 1997], ermöglicht die Integration und Weiterverarbeitung bildbasierter Information in ein GIS.

### 1.1.1 Strenge geometrische Transformation

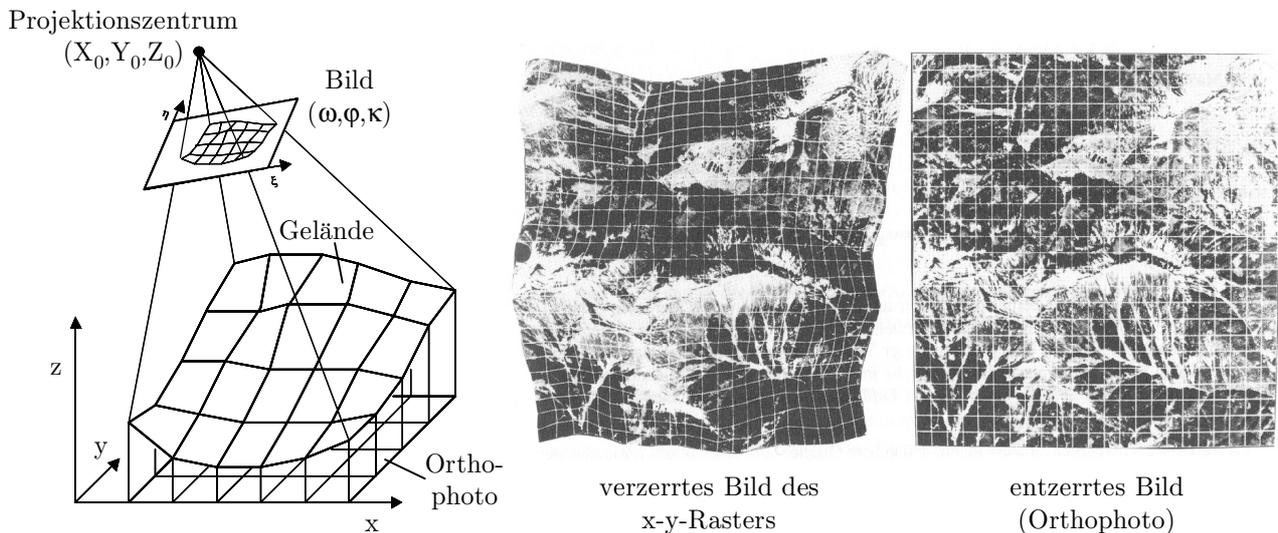
Die einzelnen Schritte, die zur strengen geometrischen Prozessierung digital vorliegender Fernerkundungsdaten durchgeführt werden, sind in Abbildung 1.1 dargestellt. Die Prozessierung beginnt mit der Herstellung der



**Abbildung 1.1:** Geometrische Prozessierung von Fernerkundungsdaten (angelehnt an die Darstellung von [Fritsch & Stallmann 2000]).

**inneren Orientierung** [Kraus 1997]. Dabei wird die reale Sensorgeometrie durch Labor- oder Selbstkalibrierung bestimmt und durch ein mathematisches Modell (Sensor- oder Kameramodell) beschrieben. Basierend auf diesem Modell kann das Bildkoordinatensystem, bestehend aus Zeilen und Spalten, in ein metrisches Sensorkoordinatensystem abgebildet werden.

Die **Kollinearitätsgleichungen** stellen die grundlegenden Abbildungsbeziehungen der Photogrammetrie dar. Sie beruhen auf dem Prinzip der geometrischen Optik, das besagt, daß das Projektionszentrum der Kamera, ein Objektpunkt und sein zugehöriger Bildpunkt auf einer Geraden liegen. Damit wird der Zusammenhang zwischen 3D-Objektkoordinaten und 2D-Bildkoordinaten hergestellt. Zur mathematischen Formulierung dieses Zusammenhangs ist die Position und die Orientierung der Kamera in einem frei wählbaren kartesischen 3D-Objektsystem notwendig. Umgekehrt können die Kollinearitätsgleichungen bei bekannten Objekt- und Bildkoordinaten von drei oder mehr Punkten dazu benutzt werden, die Position des Projektionszentrums und die Orientierung der Kamera in einem beliebigen Objektkoordinatensystem zu bestimmen, kurz: die **äußere Orientierung** der Kamera herzustellen. Punkte, für die sowohl 3D-Objektkoordinaten wie auch 2D-Sensorkoordinaten bekannt sind, nennt man **Paßpunkte**. Generell sollte ein sinnvoll gewähltes Paßpunktfeld aus Stabilitätsgründen räumlich gut über die gesamte Szene verteilt sein. Zur strengen Bildorientierung verwendet man Paßpunkte in einem kartesischen Koordinatensystem (z.B. dem **lokalen Bezugssystem** oder auch einem geozentrischen System wie WGS84). Werden Paßpunktkoordinaten in einem anderen System (z.B. einem Landeskoordinatensystem oder einem geodätischen Koordinatensystem) bestimmt, so müssen sie zunächst in ein kartesisches System transformiert werden (siehe Abbildung 1.1).



**Abbildung 1.2:** Zur Verzerrung eines 2D-Bilds durch die zentralperspektivische Aufnahme eines 3D-Objekts (entnommen aus [Kraus 1997]).

Die **Komplanaritätsbedingung** fordert für zwei Bilder, daß ein Objektpunkt, seine beiden Bildpunkte und die beiden Projektionszentren in einer Ebene liegen müssen. Anders ausgedrückt bedeutet dies, daß sich die zwei Projektionsstrahlen eines homologen Punktpaars in einem 3D-Punkt schneiden müssen. Diese Tatsache wird zur Orientierung von zwei oder mehr Bildern ausgenutzt. Durch manuelle oder automatische Bildmessung einer Anzahl von Punkten, die in mindestens zwei Bildern zu erkennen sind (homologe Punktpaare) kann die simultane Orientierung von zwei oder mehr Bildern mit Hilfe einer **Bündelblockausgleichung** hergestellt werden. Im Rahmen der Bündelblockausgleichung wird für den gesamten Bildverband eine Kleinste-Quadrate-Ausgleichung vorgenommen, bei der unter Einhaltung der Komplanaritätsbedingung für jedes homologe Punktpaar die Verbesserung der jeweiligen Bildkoordinaten minimiert wird. Nutzt man lediglich die Tatsache, daß sich die Projektionsstrahlen eines homologen Punktpaars in einem 3D-Punkt schneiden, ohne die Koordinaten dieses Punktes zu kennen (beliebiges Modellsystem), so spricht man von einer freien Netzausgleichung. Diese Freiheit der Wahl des Objektkoordinatensystems kann dazu benutzt werden, die numerische Stabilität der Bündelblockausgleichung zu verbessern. Liegen die 3D-Koordinaten von Schnittpunkten homologer Bildstrahlen explizit vor (Paßpunkte!), so spricht man im Luft- und Satellitenbildfall von einer **Aerotriangulation**. Die photogrammetrische Orientierung eines einzelnen Bildes ist ohne Paßpunkte prinzipiell nicht möglich.

Ist die Orientierung einer Kamera gegeben, so kann mit Hilfe eines digitalen Höhenmodells (DHM) durch geometrische Transformation des Bildes in ein beliebiges Zielkoordinatensystem (z.B. WGS84 oder Gauß-Krüger-Projektion) ein **Orthophoto** (georeferenziertes, entzerrtes Bildprodukt) hergestellt werden (Abbildung 1.2). Das gesamte beschriebene Verfahren bezeichnet man als indirekte Georeferenzierung. Es wird in vereinfachter Form (siehe Abschnitt 1.1.3) in dieser Arbeit angewandt.

### 1.1.2 Direkte Georeferenzierung

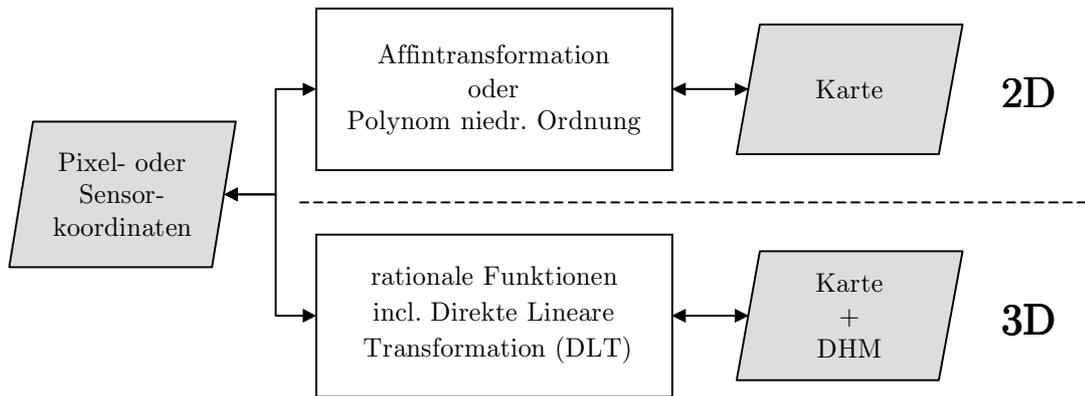
Einen alternativen Ansatz zur indirekten Georeferenzierung stellt die sich in jüngerer Zeit immer stärker verbreitende direkte Georeferenzierung dar [Cramer 2000]. Bei diesem Verfahren wird die Sensorposition mit Hilfe eines globalen, satellitengestützten Positionierungssystems wie GPS, GLONASS oder in Zukunft auch GALILEO erfasst. Die Orientierung wird für Satelliten meist mit Hilfe von Sternsensoren bestimmt (star tracker), die Winkelgenauigkeiten von etwa 2"-3" besitzen. Für Luftbildanwendungen, teilweise auch für Satelliten, werden mit Hilfe eines inertialen Meßsystems die Linear- und Drehbeschleunigungen bezüglich der drei Raumachsen gemessen. Durch zweifache Integration dieser Beobachtungen und geeignete Verrechnung mit den Positionsdaten ist es möglich, die äußere Orientierung des Sensors für jeden Zeitpunkt der Datenaufnahme herzustellen. Durch die Integration von GPS- und INS-Daten werden die Vorteile beider Verfahren – Langzeitstabilität bei GPS sowie Genauigkeit und hohe Datenrate bei INS – verbunden. Die erreichbaren geometrischen Genauigkeiten für photogrammetrische Punktbestimmungen am Boden liegen für Luftbildflüge mit einem Bildmaßstab von 1:5000 bei etwa 5-10cm in der Lage und 10-15cm in der Höhe [Heipke, Jacobsen & Wegmann 2001] und sind damit genau genug für viele Vermessungs- und Kartierungsaufgaben. Nachteilig für derartige Verfahren sind die hohen Kosten für die IMU (Inertial Measurement Unit), die man zur Erreichung solcher Genauigkeiten benötigt. Die direkte Georeferenzierung kann im Prinzip paßpunktfrei durchgeführt werden. Da in diesem Fall keine Redundanz besteht, werden zu Kontrollzwecken jedoch meist Paßpunkte herangezogen.

### 1.1.3 Vereinfachte geometrische Prozessierung

In der Praxis, speziell bei der Verarbeitung von Satellitendaten, wird vielfach mit vereinfachten Verfahren zur Georeferenzierung gearbeitet, da die Durchführung der Transformationsschritte für eine strenge Georeferenzierung sehr aufwendig sein kann bzw. oftmals nicht alle Daten zur Verfügung stehen. Bei geringem Einfluß der Topographie (kleine Öffnungswinkel, große Aufnahmeentfernung, schwache Perspektive) reicht oftmals eine 2D-Transformation aus, die das Bild- oder Sensorkoordinatensystem direkt mit einem 2D-Zielsystem (z.B. Landeskoordinaten) verknüpft. Zum Einsatz kommen hier die affine Transformation oder eine Polynomtransformation niedriger Ordnung. Bei stärkeren Verzerrungen (großer Öffnungswinkel, kleine Aufnahmeentfernung, starke Perspektive) ist eine dreidimensionale Behandlung erforderlich. Der Zusammenhang zwischen 3D-Objektkoordinaten und 2D-Bildkoordinaten läßt sich alternativ zu den Kollinearitätsgleichungen auch durch rationale Funktionen darstellen [Dowman & Dolloff 2000]. Aus einer hinreichenden Menge an Paßpunkten (abhängig von der Ordnung der benutzten rationalen Funktion) werden dabei die Koeffizienten der Zähler- und Nennerpolynome so geschätzt, daß die Bildpunktverbesserungen minimal werden. Die erreichbaren Restfehler an den Bildpaßpunkten liegen im Bereich von etwa 0.1 Pixel. Die derart bestimmte rationale Funktion ersetzt somit quasi das Sensormodell beziehungsweise "versteckt" es hinter einer universellen Schnittstelle. Im Gegensatz zu den Kollinearitätsgleichungen kommt den rationalen Funktionen wie auch der affinen oder polynomiellen Transformation keine physikalische Bedeutung zu. Ihr Nutzen liegt in der geschlossen durchführbaren und somit wesentlich vereinfachten Transformation zwischen Bild- und Zielsystem. Alle vereinfachten Transformationen sind optimiert für die vorliegenden Paßpunkte und erlauben eine näherungsweise Bildentzerrung durch Interpolation zwischen den Paßpunkten. Bei Benutzung rationaler Funktionen ist jedoch - im Gegensatz zu den 2D-Näherungen - auch die differenzielle Bildentzerrung mit Hilfe eines DHM möglich.

Die vereinfachten Verfahren laufen in der Regel folgendermaßen ab:

1. Bestimmung homologer Punktpaare im Satellitenbild und im Zielkoordinatensystem.
2. Schätzung einer 2D- oder 3D-Transformation, die das Bildkoordinatensystem und das Zielkoordinatensystem direkt miteinander verknüpft. Dabei wird der gesamte Abbildungsprozess in eine einzige Transformation gepackt.
3. Anwendung der geschätzten Transformation auf das Fernerkundungsbild und Resampling des Bildinhaltes in das Zielsystem mit einem gewünschten Verfahren wie z.B. dem Nächster-Nachbar-Verfahren oder der bilinearen Interpolation.



**Abbildung 1.3:** Vereinfachte geometrische Prozessierung von Fernerkundungsdaten, die auch in dieser Arbeit verwendet wird.

Die vereinfachten Verfahren weisen aufgrund ihrer globalen Natur Restklaffungen auf, die verringert werden können, wenn die Szene in kleinere Bereiche unterteilt wird, für die jeweils eine eigene Transformation des gleichen Typs geschätzt werden kann, vorausgesetzt, es sind in jedem Bereich genügend Paßpunkte vorhanden. Durch die Anwendung einer global oder bereichsweise uniformen Bildtransformation wird jedoch keine exakte Entzerrung erreicht. In dieser Arbeit wird mit dem vereinfachten Verfahren und einer affinen Näherungstranformation zwischen Bild- und projizierten Koordinaten gearbeitet.

Unabhängig von den beschriebenen Schritten zur geometrischen Prozessierung können auf jeder Verarbeitungsstufe vielfältige Verfahren zur Interpretation des Bildinhaltes und zur Extraktion höherwertiger Information durchgeführt werden, auf die hier jedoch nicht näher eingegangen werden soll.

## 1.2 Bereitstellung von Paßpunkten

Sowohl das strenge wie auch die vereinfachten Verfahren der indirekten Georeferenzierung benötigen Paßpunkte, deren Koordinaten im Bild und in einem Referenzsystem bekannt sein müssen. Zur Gewinnung dieser Koordinaten im Referenzsystem wird eine Unterscheidung in primäre Messungen und Koordinatenbestimmungen in bereits georeferenzierten Daten vorgenommen.

### 1.2.1 Originalmessungen

Traditionellerweise werden die zur Georeferenzierung von Fernerkundungsdaten benötigten Paßpunkte manuell im Bild gemessen, unabhängig davon, ob die Georeferenzierung in Einzelschritten oder über globale oder lokale Ersatztransformationen durchgeführt wird. Mit Hilfe einer Beschreibung der Paßpunkte ist es einem Operateur möglich, die entsprechenden Punkte im Bild mit einer Genauigkeit von etwa 1/3 Pixel zu messen. Die Genauigkeit automatischer Punktmessungen liegt im Bereich von 1/10 Pixel für korrelationsbasierte und 1/3 Pixel für merkmalsbasierte Verfahren.

Durch terrestrische Vermessung von markanten oder künstlich signalisierten Punkten können Paßpunkte mit Genauigkeiten im cm-Bereich bereitgestellt werden. Diese Methode stellt die genauesten Paßpunkte zur Verfügung, ist jedoch infolge der notwendigen Feldmessungen nicht immer möglich oder zu teuer. Die 3D-Koordinaten werden entweder im lokalen Bezugssystem oder in einem geozentrischen Referenzsystem (z.B. WGS84 bei GPS-Messungen) gewonnen.

Im Zuge der Aerotriangulation ist auch eine photogrammetrische Bestimmung von Punkten möglich, die bei weiteren Befliegungen als Paßpunkte genutzt werden können. Laut Kraus [1997] kommen dafür entweder vor der Bildaufnahme signalisierte Punkte oder natürliche oder künstliche, in den Meßbildern ausgewählte Punkte in Betracht.

### 1.2.2 Paßpunkte aus georeferenzierten Daten

Bei der Benutzung bereits georeferenzierter Daten werden Paßpunkte in einem projizierten System gewonnen. In der Regel liegen dabei Lage- und Höhendaten getrennt vor. Durch den Abriff von Lagekoordinaten aus

analogen Karten oder gescannten Rasterkarten können Lagepaßpunkte erhalten werden, die eine Georeferenzierung ohne Korrektur der Geländeform ermöglichen. Man spricht in diesem Fall von einer ebenen Entzerrung. Die Genauigkeit der so gemessenen Punkte hängt vom Kartenmaßstab ab. Für graphisch gemessene Punkte rechnet man mit einer Meßunsicherheit von 0.2mm. Damit besitzt beispielsweise ein manuell in einer topographischen Karte im Maßstab 1:25000 gemessener Paßpunkt eine Lageunsicherheit von 5m. Für die Gewinnung der Paßpunktkoordinaten in einem Referenzsystem kommen verschiedene Möglichkeiten in Betracht.

Existiert ein digitaler Kartenbestand, z.B. Daten des deutschen Amtlichen Topographisch-Kartographischen Informationssystems (ATKIS), so können je nach Natur der Daten (2D oder 3D) und Verfügbarkeit zusätzlicher Höheninformation Lage- oder Vollpaßpunkte in einem GIS gewonnen werden.

Bei Vorliegen eines DHM werden die Höhenwerte für die Lagepaßpunkte durch Interpolation bestimmt. Alternativ dazu können die Paßpunkthöhen auch durch Interpolation zwischen benachbarten Höhenlinien ermittelt werden. Damit stehen Vollpaßpunkte zur Verfügung, die, entsprechend dichte Geländehöheninformation vorausgesetzt, auch eine Korrektur des Einflusses der Geländeform ermöglichen.

Alternativ zum Abgriff aus analogen oder digitalen Kartendaten können Paßpunkte auch aus Orthophotos erhalten werden. Durch einmalige manuelle Messung im Orthophoto werden subpixelgenau prominente Paßpunkte in einem Kartensystem gemessen. Die Paßpunkthöhe wird auch hier aus einem DHM oder aus Höhenlinien bestimmt. Damit lassen sich sogenannte ground-chips (Ausschnitte fester Breite und Höhe aus dem Orthophoto, in deren Zentrum die Paßpunkte liegen) gewinnen, mit deren Hilfe die Georeferenzierung durchgeführt werden kann. Ground-chips werden im wesentlichen zur Georeferenzierung häufig wiederholter Aufnahmen eines Sensors, etwa in der Meteorologie, benutzt.

Die Gewinnung von Paßpunktkoordinaten geschieht nach wie vor weitgehend durch interaktive Verfahren, obwohl es vielfältige Verfahrensvorschläge zur Automation dieser Aufgabe gibt. Die Gründe hierfür sind im wesentlichen in der Spezialisierung auf bestimmte lokale oder thematische Verhältnisse oder der geringeren Genauigkeit oder in der fehlenden Serienreife dieser Verfahren zu sehen.

## 1.3 Zur Automatisierung der Bildregistrierung

Die wesentliche Leistung, die ein automatisches System zur Georeferenzierung von Fernerkundungsdaten erbringen muß, ist die Transformation und ggf. Entzerrung von Sensordaten in ein Referenzsystem. Die Hauptschwierigkeit besteht dabei darin, geometrische Primitive oder Objekte, die sich entsprechen, aufzufinden, also die kognitive Leistung des menschlichen Gehirns in geeigneter Weise nachzubilden. Objekte können in diesem Zusammenhang entweder einzelne geometrische Primitive oder auch Gruppierungen derselben sein. Ein im engeren Sinne objektorientierter Ansatz arbeitet mit geometrischen Objekten, die aus Primitiven gebildet werden, die anhand bestimmter Regeln zu Objekten gruppiert werden. Das rechnerbasierte Erkennen und Zuordnen geometrischer Objekte ist Gegenstand der Computer Vision (vgl. [Haralick & Shapiro 1992], [Haralick & Shapiro 1993] oder [Ballard & Brown 1982]). Damit automatische Orientierungsverfahren praxisrelevant sind, sollten sie nach Heipke [1997] idealerweise folgende Anforderungen erfüllen: Autonomie (Unabhängigkeit von Benutzereingriffen), höhere Schnelligkeit und Genauigkeit als manuelle Verfahren, Flexibilität bezüglich des Sensortyps und Art der Kontrollinformation, Robustheit und Zuverlässigkeit.

Generell besteht eine Zuordnungsaufgabe darin, korrespondierende Elemente zweier oder mehrerer Datensätze zu finden [WEU 1995]. Man bezeichnet dies auch als Lösung des Korrespondenzproblems. Dowman [1998] unterteilt die Aufgabe, Fernerkundungsdaten zu registrieren, in zwei Teile: Zum einen die Registrierung von Bildern gleichen Typs (relative Orientierung, Registrierung von zwei Satellitenbildern des selben Sensors, DHM-Generierung) und zum anderen die Registrierung von Daten unterschiedlichen Typs. Die Registrierung von Daten unterschiedlichen Typs läßt sich weiter unterteilen in die Registrierung von Daten unterschiedlicher Sensoren (z.B. [Fonseca & Manjunath 1996] oder [Schowengerdt 1997], Kapitel 8) und die Registrierung von Bildern mit Raster- oder Vektorkarten (absolute Orientierung). Die Klassifizierung der möglichen Zuordnungsverfahren findet in der in Photogrammetrie und Computer Vision oftmals in intensitätsbasierte (low level vision  $\Rightarrow$  area based matching), merkmalsbasierte (mid level vision  $\Rightarrow$  feature based matching) und relationale Zuordnungsverfahren statt (high level vision  $\Rightarrow$  relational matching) [Haralick & Shapiro 1992]. Alternative Klassifizierungen sind möglich, z.B. bezüglich des Suchverfahrens (Baumsuche, generalisierte Hough-Transformation, Kleinste-Quadrate, etc.) oder bezüglich des Bewertungsmaßes (Kostenfunktion, Distanzmaß, Ähnlichkeitsmaß, etc.). Während grauwertbasierte (oder ikonische) Zuordnungsverfahren nur zwischen zwei oder mehreren Bildern in Betracht kommen, sind merkmalsbasierte und relationale Verfahren sowohl für Bild-Bild- als auch für Bild-Modell- oder Modell-Modell-Zuordnungen geeignet. Wie von Heipke [1997] dargestellt, sollten erfolgversprechende Zuordnungsverfahren folgende Punkte berücksichtigen:

**Hierarchie.** Durch hierarchische Verfahren wird eine Grob-zu-Fein-Strategie verfolgt, mit der die Zahl der Mehrdeutigkeiten verringert und der Konvergenzradius vergrößert werden kann. Die Objekte, die zugeordnet werden, müssen dabei nicht notwendigerweise in jeder Ebene die selben sein.

**Redundanz.** Zur effizienten Beseitigung von Fehlzuordnungen, die bei jedem automatischen System unvermeidlich sind, sollte eine möglichst hohe Redundanz vorliegen.

**Selbstdiagnose.** Zur Vermeidung einer Akkumulation von Fehlern sollte möglichst in jedem Schritt eine Möglichkeit zur Selbstkontrolle eingebaut sein, z.B. die Prüfung von Konsistenzkriterien oder Schwellwerten.

Die digitale Bildzuordnung rein aufgrund der Grauwerte oder lokaler Grauwertverteilungen (z.B. Texturmaße) basiert in der Regel auf der Maximierung eines Ähnlichkeitsmaßes, das für zwei gleich große Fenster aus den zuzuordnenden Bildern berechnet wird. Gebräuchliche Ähnlichkeitsmaße sind der Korrelationskoeffizient [Rosenfeld & Kak 1976], oder die normalisierte Kreuzkorrelation [Pratt 1991]. Der selbe Autor berichtet in in einer früheren Arbeit [Pratt 1974] über die Registrierung simulierter Bilddaten mit Hilfe von Korrelationstechniken. Historisch wurden grauwertbasierte Zuordnungsverfahren in der Photogrammetrie zuerst zur DHM-Erzeugung mittels homologer Punkte aus Stereopaaren benutzt, z.B. [Kreiling 1976]. Eine Erweiterung der lokalen Maximierung von Korrelationsmaßen stellt die Bildzuordnung nach der Methode der kleinsten Quadrate dar [Förstner 1982], [Ackermann 1983]. Dabei wird durch Schätzung einer affinen und einer linearen radiometrischen Transformation zwischen zwei gleich großen Bildfenstern subpixelgenau die Position in der Suchmaske ermittelt, bei der die Quadratsumme der Grauwertdifferenzen der Pixel gleicher Lage minimiert wird. Gegenüber der Korrelation zeichnet sich dieses Verfahren durch höhere Genauigkeit, Invarianz unter radiometrischer Transformation und bessere Möglichkeiten zur Genauigkeitsanalyse aus. Allerdings ist der Konvergenzradius des Verfahrens beschränkt, weshalb meist eine Grob-zu-Fein-Strategie in Bildpyramiden verwendet wird [Ackermann & Hahn 1991]. Besteht die Georeferenzierungsaufgabe darin, ein Rasterbild eines Fernerkundungssensors einem Orthophoto oder Teilen eines Orthophotos zuzuordnen, so kommen prinzipiell alle grauwertbasierten Zuordnungsverfahren zur Identifikation homologer Punktepaare in Betracht. Über den Raumbezug der Bildelemente des Orthophotos können bei erfolgreicher Zuordnung die Landeskoordinaten der erzeugten Paßpunkte erhalten und mit Hilfe der o.g. Verfahren zur Georeferenzierung benutzt werden.

Liegen gescannte Karten als Referenzdaten in Rasterform vor, so können grauwertbasierte Zuordnungsverfahren nicht sinnvoll eingesetzt werden, da durch die kartographischen Regeln und die durchgeführten Generalisierungen die Grauwertverteilungen in den zuzuordnenden Datensätzen in der Regel nicht korreliert sind. In diesem Fall und auch im Falle der Raster-Vektor-Zuordnung ist der Übergang auf eine symbolische Bildbeschreibung zu wählen. Mit einer symbolischen Bildbeschreibung werden durch geeignete Operatoren Punkte, Linien oder Flächen nach bestimmten Kriterien aus den Rasterdaten extrahiert und gegebenenfalls zu höherwertigen Objekten gruppiert. Übersichten zur Merkmals- und Objektextraktion aus Bilddaten werden z.B. von Mayer [1998], Sester [1993] oder Schenk [1993] gegeben. Arbeiten zur Extraktion bedeutungsvoller Objekte aus gescannten Rasterkarten finden sich z.B. in Illert [1990], Stengele [1995] oder Nebiker & Carosio [1995].

Speziell für flächenhafte Objekte wurden seit den 60er Jahren vielfach invariante Merkmale zur Mustererkennung und Objektzuordnung benutzt [Hu 1962], [Flusser & Suk 1993]. Die Mustererkennung mit Hilfe invarianter Merkmale hat sich seither vielfach als mächtiges Instrument zur Beherrschung schwer lösbarer Zuordnungsaufgaben erwiesen. In dieser Arbeit wird, getrieben durch die allgemeine Fragestellung, insbesondere die Verwendbarkeit affin invarianter Merkmale für die Georeferenzierung untersucht.

Mit der intelligenten Punktauswahl mit Hilfe von Interest-Operatoren (z.B. Förstner-Operator [Förstner & Gülch 1987] oder Moravec-Operator [Moravec 1979]) hielten merkmalsbasierte Bildzuordnungsverfahren Einzug in die Stereo-Zuordnung [Hannah 1989] und DHM-Generierung [Hahn & Förstner 1988], später auch in die relative Orientierung [Hellwich, Heipke, Tang, Ebner & Mayr 1994], [Haala, Hahn & Schmidt 1994] und Aerotriangulation [Tsingas 1992]. Zuordnungsverfahren zur automatischen relativen Orientierung und Aerotriangulation sind bereits seit längerer Zeit in Standardsoftware integriert. [Dörstel, Tang & Madani 2001], [Sigle & Heuchel 2001]. Wie in Hahn & Förstner [1988] beschrieben, wird oftmals eine intelligente Punktauswahl mit Interest-Operatoren getroffen und die Feinzuordnung intensitätsbasiert mit einem Kleinst-Quadrate-Verfahren gelöst. Neben den bisher erwähnten punktbasierten Verfahren werden in der digitalen Photogrammetrie auch zunehmend Zuordnungsverfahren eingesetzt, die auf Linien oder Polygonen basieren.

Merkmalsbasierte Verfahren lassen sich nach Barnard & Thompson [1980] in drei Schritte zerlegen:

1. Merkmalsextraktion in beiden Partnern.
2. Aufstellung einer Liste möglicher Zuordnungen aufgrund nach bestimmten Kriterien passender Merkmale.
3. Ermittlung einer eindeutigen und konsistenten Lösung durch Ausdünnen (Verifikation) der Liste.

	<b>intensitätsbasiert</b>	<b>merkmalsbasiert</b>	<b>relational</b>
Bild-Bild	✓	✓	✓
Bild-Modell	-	✓	✓
Modell-Modell	-	✓	✓

**Tabelle 1.1:** Zuordnungsfälle und mögliche Verfahrensgruppen.

Eine Übersicht über merkmalsbasierte Zuordnungsverfahren wird z.B. von Förstner [1993], von Strunz [1993] oder für Nahbereichsanwendungen von Luhmann [2000] gegeben. Bei relationalen Verfahren werden neben den Objektmerkmalen Objektrelationen (z.B. "a ist links von b", "a liegt neben b", "a schneidet b", etc.) vom Zuordnungsverfahren berücksichtigt. Zunächst wird dabei eine strukturelle Beschreibung beider Datensätze aufgebaut [Wang 1994], ein anschließendes Zuordnungsverfahren sucht mit Hilfe der Objektmerkmale und Objektrelationen anhand einer bestimmten Bewertungsfunktion, etwa der gegenseitigen Information [Vosselmann 1992], in einer Baumsuche die bestmöglichen Zuordnungen heraus. Relationale Zuordnungsverfahren eignen sich auch für die Zuordnung von Vektordaten unterschiedlicher Natur, z.B. Straßendaten aus ATKIS und GDF [Walter 1997].

Grauwertbasierte Verfahren sind direkt mit der Bildgeometrie verknüpft und benötigen daher gute Näherungswerte um zu konvergieren. Bei merkmalsbasierten und insbesondere bei relationalen Verfahren findet die Zuordnung in einem Merkmalsraum und/oder über topologische Beziehungen statt und ist daher von der reinen Geometrie losgelöst. Merkmalsbasierte oder relationale Verfahren können prinzipiell auch ohne geeignete Näherungswerte durchgeführt werden, wenngleich sinnvolle Startwerte die Zahl möglicher Zuordnungen und damit die Komplexität der Suche von Beginn her wesentlich verkleinern können. In Tabelle 1.1 sind die möglichen Zuordnungsfälle und die in Frage kommenden Verfahrensgruppen dargestellt.

### 1.3.1 Registrierung von Bilddaten gleichen Typs

Das einfachste Beispiel für die automatische Registrierung gleicher Bilddaten ist die relative Orientierung eines Luftbildpaares, bei der mindestens fünf gut im Überlappungsbereich verteilte homologe Punktpaare bestimmt werden müssen [Dowman 1998]. Stereopaare sind sich in der Regel sehr ähnlich, d.h. die radiometrischen Verhältnisse und Bildmaßstäbe sind vergleichbar, es existiert nahezu keine Verdrehung und der Überlappungsbereich ist aufgrund sorgfältiger Flugplanung hoch (etwa 60%). Die Orientierungsbestimmung innerhalb automatischer Verfahren ist die selbe wie bei analytischen, der Unterschied in der digitalen Photogrammetrie liegt in der automatischen Extraktion und Zuordnung von homologen Primitiven und Objekten. Eine Übersicht über verschiedene Orientierungsverfahren gibt Heipke [1997].

Aufgrund des großen Öffnungswinkels und der stärkeren Perspektive ist der Geländeeinfluß bei Luftbildzuordnungen nicht vernachlässigbar. Die entsprechenden Verfahren arbeiten daher mit rigorosen Sensormodellen in drei Dimensionen. Bei der Zuordnung von Satellitendaten ergeben sich selbst bei Daten gleicher Sensoren und unkritischem Gelände andere Komplikationen, z.B. größere Verdrehungen (ergeben Maßstabsänderungen) oder größere Zeitdifferenzen zwischen den Aufnahmen (radiometrische Unterschiede, Veränderungen der Szene). Durch die hohe Redundanz automatischer Verfahren können diese Nachteile jedoch für Bilder gleicher Sensoren meist kompensiert werden. Aus diesem Grund können Stereopaare eines Satellitensensors meist mit den selben Methoden wie Luftbildpaare bearbeitet werden. Je nach Genauigkeitsanforderungen und Bewegtheit des Geländes liefert auch eine 2D-Transformation zur Registrierung von Satellitendaten hinreichend gute Ergebnisse.

Neben grauwert- und punktbasierten Verfahren existieren auch Ansätze zur Satellitenbildzuordnung mit Linien oder Flächen. Abbasi-Dezfouli & Freeman [1994] geben ein Verfahren zur Zuordnung von homogenen Grauwertpolygonen in einem SPOT-Stereopaar an ("patch matching"). Polygone werden in diesem Verfahren durch Fläche und Form charakterisiert. Die Formbeschreibung stützt sich auf ein minimales berandendes Rechteck (MBR) und auf den sogenannten "chain code" [Freeman 1961] der Kontur extrahierter Bildpolygone. Die Polygonzuordnung erfolgt ausgehend von einer Anfangstransformation aus drei zugeordneten Polygonen über die Minimierung einer Kostenfunktion. Die benutzte chain code frequency wird abgeleitet aus dem chain code, der aus den Konturpixeln eines Rasterpolygons berechnet wird. Gegenüber dem chain code ist die chain code frequency invariant gegenüber Rotation. Auf den Raster-Polygonkonturen werden nach der Methode von Newton, Gurney, Sloggett & Dowman [1994] ("dynamic programming") über ein Kantenzuordnungsverfahren homologe Punktpaare extrahiert, mit denen die Orientierungsparameter für die SPOT-Szenen geschätzt werden. Wang [1994] präsentiert ein Strukturzuordnungsverfahren (relationale Zuordnung), mit dem unter anderem auch ein MOMS02/D2-Paar registriert wird.

Für die Georeferenzierung von Fernerkundungsdaten sind reine Bild-Bild-Zuordnungsverfahren von Interesse, wenn der Referenzdatensatz ein bereits georeferenziertes Bild, z.B. ein Orthophoto, ist. Neben den bereits

erwähnten Ansätzen wird eine Möglichkeit, ein bereits existierendes Orthophoto zur Georeferenzierung neuer Luftbilder zu benutzen, von Höhle [1998] vorgestellt. Durch einen Interest-Operator werden im Orthophoto prominente Punkte ausgewählt und deren Geländehöhen aus dem DHM interpoliert. Zusammen mit der Umgebung der Punkte in einem gewissen Fenster des Orthobildes werden damit sog. "Ground Chips" gebildet. Durch lokale Maximierung des Korrelationskoeffizienten der Ground-Chip-Fenster in einem Suchbereich im neuen Luftbild (Näherung ist gegeben) wird die homologe Bildpunktposition ermittelt. Damit liegen Paßpunkte vor, mit denen die absolute Orientierung des neuen Bildes und - mit Hilfe des DHMs - ein neues Orthophoto berechnet werden kann.

Ein Verfahren zur Georeferenzierung von SAR-Daten wird von Strunz, Knöpfle & Roth [1994] gegeben. In diesem Verfahren wurden Ground-Chips aus georeferenzierten SAR-Szenen manuell erzeugt und mit weiteren Attributen in einer GCP-Datenbank abgelegt (GCP = Ground Control Point). Durch Korrelation können in neuen Szenen die entsprechenden Punkte ermittelt werden. In einem zweiten Verfahren wurden hier aus einem DHM simulierte SAR Ground-Chips erzeugt und zur Georeferenzierung benutzt.

Alternativ zur relativen oder absoluten Orientierung mittels Zuordnung von geometrischen Primitiven oder Objekten wurden von Rosenholm & Torlegaard [1988] und Ebner & Ohlhof [1994] Verfahren vorgeschlagen, die ohne Paßpunktmessungen im Bildraum auskommen. Für diese objektraumbasierten Verfahren reicht die Existenz eines DHM aus, sie sind daher insbesondere geeignet für Gebiete, in denen keine Merkmale oder Objekte extrahiert werden können.

Noch einen Schritt weiter geht das zeitgleich von Wrobel [1987] und Ebner, Fritsch, Gillessen & Heipke [1987] vorgestellte Verfahren, mit dem die Integration von Bildzuordnung, Punktbestimmung, Oberflächenrekonstruktion und Orthophotoprojektion in einem Schritt erreicht wird. In diesem Ansatz werden in einer nichtlinearen Ausgleichung nach der Methode der kleinsten Quadrate simultan die Parameter der absoluten Orientierung und der geometrischen und radiometrischen Parameter der Geländeoberfläche (z.B. Geländehöhe und Reflektanz) bestimmt. Ein ähnliches Verfahren, das bereits mit Hilfe eines Linux-Clusters parallelisiert wurde, wird von Holm [2001] präsentiert.

### 1.3.2 Registrierung von Bilddaten unterschiedlichen Typs

Die Betrachtung der Registrierung von Bilddaten unterschiedlicher Sensoren führt einen Schritt weiter in Richtung eines generellen Systems zur automatischen Georeferenzierung von Fernerkundungsdaten. Viele der in diesem Zusammenhang entwickelten Methoden lassen sich auch auf die Bild-Modell-Zuordnung übertragen, daher wird der gesamten Thematik hier ein etwas breiterer Raum gegeben.

Die Integration und Fusion unterschiedlicher Sensordaten führt zu Synergien, da die in den einzelnen Szenen enthaltene Information oft komplementär ist. Ein Beispiel stellt die Integration von panchromatischen SPOT-Daten mit multispektralen Landsat-Daten dar, bei der die hohe Auflösung der SPOT-Daten (10m Bodenpixel) mit der multispektralen Information der Landsat-Daten (bis zu 7 hyperspektrale Kanäle) kombiniert wird. Auf die Vielfalt der Verfahren zur Datenfusion soll hier allerdings nicht weiter eingegangen werden, eine Übersicht findet sich bei Pohl & Genderen [1998].

Unterschiedliche Fernerkundungssensoren unterscheiden sich in ihren spektralen, geometrischen und radiometrischen Eigenschaften, bilden in unterschiedlichen Maßstäben ab und bewegen sich auf unterschiedlichen Bahnen. Zusätzlich ändern sich bei multitemporalen Aufnahmen die radiometrischen Verhältnisse aufgrund variierender Vegetation etc. Aus diesen Gründen scheiden grauwertbasierte Verfahren zur Multisensor-Registrierung nahezu aus (mit Ausnahme von Lee, Carender, Knowlton, Bell & Bryan [1994], s.u.). Zuordnungsverfahren für Aufnahmen der selben Szene von unterschiedlichen Sensoren müssen all diesen Umständen Rechnung tragen und dürfen daher nicht von den Eigenschaften eines speziellen Sensors abhängen. Eine gute Übersicht über Multisensor-Registrierungsverfahren bieten Fonseca & Manjunath [1996]. Hier werden auch Verfahren besprochen, die die Zuordnung in einem transformierten Raum (z.B. Fourier-Transformation) durchführen.

Nach Dowman [1998] kommen Punkte für die merkmalsbasierte Zuordnung aus folgenden Gründen nicht in Betracht. Interest-Operatoren erkennen in Bildern optischer Sensoren aufgrund unterschiedlicher Auflösung unterschiedliche Punkte. In SAR-(Synthetic Aperture Radar)-Szenen führen Speckles zu zusätzlichen Verirrungen der Interest-Operatoren. Nach Aussage des selben Autors tendieren Linien, die auf Kantenextraktion beruhen, dazu, nicht eindeutig erkennbar zu sein, wenngleich für bestimmte Szenen Kreuzungen zwischen Kanten durchaus zu verlässlichen Merkmalen führen können. Die vielversprechendsten Objekte für ein möglichst universelles System zur Bild-Bild- und letztlich auch zur Bild-Modell-Zuordnung stellen Polygone dar, da deren Konturen auch bei leichten Störungen in der Regel gut unterscheidbar sind.

Die Extraktion von Polygonen aus Bildern geschieht in der Regel mit Hilfe von Segmentierungsverfahren (siehe auch Kapitel 2). Durch die Segmentierung werden in gewisser Weise homogene Pixelbereiche in Bildern zu

Regionen gruppiert. Aus der großen Fülle existierender Verfahren sei hier exemplarisch die Arbeit von Ruskoné & Dowman [1997] herausgegriffen, deren Segmentierungsverfahren in zwei Stufen abläuft. In der ersten Stufe werden Pixel nach einem Homogenitätskriterium zu Regionen verschmolzen, in einer zweiten Stufe werden die Regionen analysiert und ähnliche Regionen abermals verschmolzen. Kleine Regionen werden eliminiert.

Im EU-Projekt PAIRS (Prototype Automated Image Registration System) [Dowman & Ruskoné 1997] wurde neben der Zuordnung von Interest-Punkten ebenfalls das bereits in Abschnitt Polygonzuordnungsverfahren von Abbasi-Dezfouli & Freeman [1994] eingesetzt. Polygone werden in diesem Verfahren durch Fläche und Form charakterisiert. Die Formbeschreibung stützt sich auf ein minimales berandendes Rechteck (MBR) und auf die bereits beschriebenen chain codes der Konturen extrahierter Bildpolygone. Die Polygonzuordnung erfolgt ausgehend von einer Anfangstransformation aus drei zugeordneten Polygonen über die Minimierung einer Kostenfunktion. Das PAIRS-System unterscheidet zwischen ebenem Gelände (2D-Transformation, siehe Abschnitt 1.1.3) und bewegtem Gelände (3D-Transformation, siehe Abschnitt 1.1.1). Im 2D-Fall wird die Registrierung durch eine Polynomtransformation, die durch homologe Punkte geschätzt wurde, erreicht. Im 3D-Fall wird für jeden Sensortyp ein parametrisches Sensormodell benützt und mit Hilfe eines DHM eine volle 3D-Registrierung durchgeführt.

Mit einer modifizierten Kostenfunktion und ohne Verwendung der chain codes stellen Dowman & Dare [1999] ein Verfahren zur automatischen Registrierung von SPOT- und SAR-Daten dar, das ein robusteres Verhalten gegenüber Unterschieden in der Objektkontur zeigt. In einem zweiten Schritt werden die groben Polygonzuordnungen mit der in Newton et al. [1994] angegebenen Methode der dynamischen Programmierung mit Hilfe von Grauwertkanten verbessert. Im Gegensatz zu Newton et al. [1994], die Polygonkanten verwenden, benutzen Dowman & Dare [1999] jedoch Grauwertkanten, die direkt aus dem Bild abgeleitet wurden. Nach Durchführung der verfeinernden Kantenzuordnung wird mit diesem Verfahren für zwei ganze Szenen ein mittlerer quadratischer Fehler von 11 Pixeln bei Anwendung einer affinen Transformation erreicht. In einer weiterführenden Arbeit wird durch die Verwendung multipler Segmentierungen mit unterschiedlichen Parametern eine höhere Anzahl an Verknüpfungspunkten generiert und damit ein robusteres Verfahren erhalten [Dare & Dowman 2000].

Li, Manjunath & Mitra [1995] konnten durch Zuordnung von Regionenrändern auf der Basis von chain-code-Korrelationen und geometrischen Momenten mit einem aktiven Konturmodell ("snake") und einer Konsistenzprüfung im Parameterraum SPOT-, Landsat TM- und Seasat SAR-Daten registrieren.

Bartl & Schneider [1995] gelingt es mit Hilfe eines sogenannten "angle-chain"-Algorithmus', multispektrale MKF6-Daten der russischen MIR-Station mit Landsat TM-Daten bei einem Punktfehler von 1.5 Pixeln unter affiner Transformation zu registrieren. Mit Hilfe einer Segmentierung werden zunächst Gewässer- und Waldflächen aus beiden Szenen extrahiert. Für jeden Regionenschwerpunkt wird dann die Winkelsequenz, unter der die anderen Regionenschwerpunkte erscheinen, bestimmt. Das Zuordnungsverfahren findet mit einer Hypothesen-Test-Strategie aufgrund dieser Relationen homologe Punktpaare in den beiden Szenen. Dieses Verfahren zeichnet sich durch große Robustheit gegenüber fehlenden Objekten aus. Durch die Bildung der Winkelsequenzen werden die komplexen geometrischen Relationen der Objekte in eine lineare Form gebracht, die sich leichter zuordnen läßt.

## 1.4 Arbeiten zur automatischen Bild-Modell-Zuordnung

Die Zuordnung von Bild- und Modelldaten stellt ein schwierigeres Problem als die reine Bild-Bild-Zuordnung dar, da die Daten in unterschiedlicher Form vorliegen. In diesem Fall ist, entgegen den bisherigen Fällen, ein 2D-3D-Zuordnungsproblem zu lösen. Im Rahmen der weiteren Betrachtungen findet eine Beschränkung der Vielfalt möglicher Modelldaten auf Geodaten (Rasterkarten, Vektorkarten, DHM) statt. Geodaten liegen oftmals nicht in reiner 3D-Form sondern in  $2\frac{1}{2}$ D-Form oder getrennt in Lage- und Höheninformation vor, wodurch die Zuordnungsaufgabe unter Umständen erleichtert werden kann.

### 1.4.1 Automatische Georeferenzierung von Luftbildern

Drewniok & Rohr [1995] stellen einen Ansatz zur Registrierung großmaßstäbiger Luftbilder städtischer Bereiche mit Katasterkarten vor. Die Zuordnung erfolgt über Kanaldeckel, die in den Katasterkarten in 3D vorliegen, als Kontrollobjekte. In einer Lernphase werden die Parameter eines Bildmodells für als symmetrisch angenommene Kanaldeckel geschätzt. Mit den so gelernten Bildausschnitten werden per Kleinste-Quadrate-Korrelation Kanaldeckel im Bild identifiziert. Unter Annahme eines einheitlichen Bildmaßstabes wird - basierend auf den relativen Abständen der extrahierten Kanaldeckel - für jeden Datensatz eine strukturelle Beschreibung erzeugt

und über ein relationales Verfahren zugeordnet. Über einen räumlichen Rückwärtsschnitt wird das Luftbild orientiert. Ein kritischer Punkt in diesem Verfahren ist die Frage, wie gut die Modellannahmen für die Kanaldeckel tatsächlich stimmen. Zudem ist es in dieser Form nur in urbanen Szenen anwendbar.

**Pedersen [1996]** registriert Zeiss RMK-A15/23-Luftbilder im Maßstab 1:5000 mit digitalen Katasterkarten. Als Kontrollobjekte werden Polygone benutzt, die homogene Grauwertflächen (Straßenzüge, Gehsteige oder Parkplätze) repräsentieren. Die Objekte werden in dieser Arbeit manuell aus der Karte digitalisiert und über die genähert vorliegende äußere Orientierung mit Hilfe einer Vektor-Raster-Konversion in das Luftbild projiziert. Über mehrere Pyramidenstufen werden die so erzeugten templates mit Hilfe des Pakets MATCH-T der INPHO GmbH subpixelgenau durch Kleinste-Quadrate-Korrelation im Bild lokalisiert. Aus den so erzeugten Paßpunkten wird in jeder Pyramidenstufe die Orientierung des Luftbildes verbessert und in die nächsttiefere Stufe übernommen. In der höchsten Auflösungsstufe erreicht das Verfahren einen mittleren quadratischen Lagefehler von etwa 15cm. Das Verfahren benötigt in dieser Version manuelle Messungen, die jedoch künftig durch Benutzung einer GIS-Datenbasis entfallen könnten.

Das System **AMOR (Automatic Model-Based Orientation)** [**Schickler 1992**] stellt einen Ansatz zur absoluten Orientierung von Luftbildern im Maßstab 1:12000 dar, der auf der Zuordnung von 3D-Drahtmodellen von Gebäuden zu entsprechenden Grauwertkanten im Bild beruht. Unter Verwendung guter Näherungswerte für das Projektionszentrum ( $\pm 50\text{m}$ ) und für die Orientierungen ( $\pm 1^\circ$ ) werden die 3D-Drahtmodelle ins Bild projiziert. Innerhalb eines Bereichs um das projizierte Modell werden gerade Grauwertkanten im Bild extrahiert. Mit einem wahrscheinlichkeitsbasierten Gruppierungsverfahren ("probabilistic clustering") werden sich entsprechende Liniensegmente zugeordnet. Über ein robustes Kleinste-Quadrate-Ausgleichsverfahren werden die Orientierungsparameter direkt mit Hilfe zugeordneter Liniensegmente geschätzt. Das Verfahren wurde an über 50 Bildern erfolgreich getestet. Die Verfügbarkeit der Kontrollinformation in Form von 3D-Gebäudekanten ist nicht immer gegeben, was die globale Nutzbarkeit dieses Verfahrens beschränkt.

Von **Haala & Vosselmann [1992]** wurde ein relationales 2D-Zuordnungsverfahren zur automatischen äußeren Orientierung eines Echtfarben-Luftbildes (Maßstab 1:8000) und eines CIR-Bildes (Maßstab 1:6000) mit Hilfe einer gescannten Karte vorgestellt. Die Kontrollinformation besteht in dieser Arbeit aus manuell digitalisierten Straßen, Flüssen und Feldgrenzen. Über eine Klassifikation im IHS-Farbraum und anschließende Binarisierung bzw. Kantenextraktion, -skelettierung und Linienverfolgung wurden Flächen- bzw. Linienobjekte aus den Bildern erzeugt. Die Zuordnung erfolgte relational über die strukturelle Beschreibung der Objekte in den Datensätzen. Dabei wurde ein Baumsuchverfahren benutzt. Durch Abschätzung des potenziellen Nutzens eines bestimmten Suchpfades und Anwendung verschiedener Heuristiken konnte eine Reduktion des Suchaufwandes erreicht werden. Durch Berechnung einer Näherungstransformation aus den ersten drei zugeordneten Objekten konnte durch ein Ausschlußverfahren der Suchraum weiter verkleinert werden. Das Verfahren kommt prinzipiell ohne Näherungswerte aus. Es wurde nicht in erster Linie auf Genauigkeit, sondern auf möglichst breite Anwendbarkeit ausgelegt. Kritisch für dieses Verfahren ist die Segmentierung und die Extraktion von Straßenzügen. Durch die relationale Beschreibung und das robuste Baumsuchverfahren zeigt es jedoch einen möglichen Weg für globaleres Georeferenzierungsverfahren auf.

**Shahin & Novak [1994]** verfolgen ebenfalls einen straßenbasierten Ansatz zur automatischen äußeren Orientierung kleinmaßstäbiger Luftbilder. Hier werden Straßendaten, die mit einem mobilen GPS-Meßsystem aufgezeichnet wurden, als Kontrollinformation benutzt. Aus den Straßendaten werden mit Hilfe von kritischen Punkten 3D-B-Splines berechnet. B-Splines besitzen die Besonderheit, daß ihre Kontrollpunkte invariant unter affiner und projektiver Transformation sind. Aus dem Bild werden Straßen mit Hilfe des Duda-Straßenoperators extrahiert und durch Skelettierung und Linienverfolgung weiter verfeinert. Aus den extrahierten Straßen werden im Bild ebenfalls kubische B-Splines berechnet. Das verwendete relationale Zuordnungsverfahren basiert auf den Koeffizienten der B-Splines und bedient sich ebenfalls einer Baumsuche (backtracking with forward checking). In der Arbeit wurde neben einem Luftbild auch eine SPOT-Szene georeferenziert, womit eine breitere Anwendbarkeit des Verfahrens demonstriert werden konnte.

**Growe & Tönnies [1997]** registrieren Luftbilder und SAR-Daten mit ATKIS-Daten. Als Kontrollobjekte dienen Straßenkreuzungen. Eine genäherte äußere Orientierung ist durch GPS/INS-Beobachtungen bereits a priori gegeben. In einem blockweise vorgehenden Verfahren werden mit Hilfe eines semantischen Netzwerks (wissensbasiertes System AIDA) zunächst aus der GIS-Datenbasis Kreuzungen extrahiert. Die Kreuzungsbereiche werden über die grobe äußere Orientierung in das Bild transformiert und dienen zur Unterstützung eines kantenbasierten Straßenextraktionsalgorithmus'. Durch Verifikation der Kreuzungspunkte mit AIDA werden Paßpunkte erzeugt, mit denen eine verbesserte äußere Orientierung der Bilddaten berechnet wird. Das System ist nicht auf Kreuzungen beschränkt, da die verwendete Wissensbasis auf beliebige Objekte erweitert werden kann.

Einen ähnlichen Ansatz verfolgt **Bückner [1998]**. In dieser Arbeit wird das Straßennetzwerk jedoch ohne Wissensbasis, sondern mittels einer Expansion von zuverlässigen Straßensegmenten erzeugt. Zuverlässige

Straßensegmente werden durch eine kantenbasierte Segmentierung mit Zwangsbedingungen an Kantenlänge, Gradientenrichtung und Parallelität erzeugt. Das implizierte radiometrische Straßenkörpermodell ist flexibel genug, um in Luftbildern (VIS und IR), SAR-Szenen und in Rasterkarten zu funktionieren. Die Expansion der Straßensegmente geschieht mit einem modifizierten A\*-Algorithmus, mit dem parallele Kanten modellkonsistent durch Minimierung einer vorausgeschätzten Kostenfunktion verfolgt werden. Kreuzungen, die aus Modelldaten (Rasterkarte oder Vektor-GIS) gewonnen wurden, werden über eine genähert vorliegende absolute Orientierung in das Bild transformiert. Durch Relaxation werden Zuordnung im Bildraum gewonnen. Mit Hilfe der so erzeugten Paßpunkte kann damit die äußere Bildorientierung kartengestützt verbessert werden.

### 1.4.2 Automatische Georeferenzierung von Satellitenbildern

**Morris, Stevens & Muller [1988]** präsentieren mehrere automatisierte Verfahren zur Georeferenzierung von SPOT-Daten. Als Referenzdatensatz dienen digitale 2D-Karten mit trigonometrischen 3D-Punkten und Höhenlinien. Die 3D-Punkte liegen sowohl an Bergspitzen als auch entlang größerer Straßen vor. An Objekten, die im SPOT-Bild gut erkennbar sind (Uferlinien, Kreuzungen größerer Straßen) werden mit Hilfe eines kommerziellen DHM-Pakets und Spline-Interpolationen aus den vorhandenen Höhendaten Punkthöhen für die punkt- und linienförmigen Kontrollobjekte interpoliert. Die Höhengenaugkeit der erzeugten Kontrollpunkte wird durch photogrammetrische Auswertung einer Stereo-Unterfliegung mit etwa  $\pm 3m$  angegeben. Der Schwerpunkt dieser Arbeit liegt auf der automatischen Erzeugung von 3D-Kontrollinformation, explizite Verfahren zur automatischen Merkmalsextraktion im Bild und zur Zuordnung werden nicht genannt.

Das System **REGGIE [Lee et al. 1994]** erlaubt die vollautomatische Registrierung verschiedener Satellitendaten untereinander und mit Karten in Rasterform. Dieser Ansatz arbeitet mit geometrischen Modellen, deren Parameter für jeden Datentyp separat in einer iterativen Prozedur bestimmt werden. Die Selektion von Punktmerkmalen geschieht über eine normalisierte 2D-Grauwertkorrelation. Zur Reduktion des verzerrenden Einflusses der verschiedenen Datengeometrien findet die grauwertbasierte Zuordnung in einer gemeinsamen Perspektive statt, die aus den Einzeldaten mit Hilfe der a priori bekannten Näherungswerte der geometrischen Modelle berechnet wird. Durch die iterative Schätzung der Modellparameter werden für die Bild-Bild-Zuordnung (z.B. SPOT-Landsat TM) Genauigkeiten von etwa 1 Pixel angegeben. Für den Bild-Karte-Fall konnte eine grobe Registrierung erreicht werden, eine Genauigkeitsangabe fehlt hier jedoch.

Ein polygonbasierter Ansatz wird von **Holm, Parmes, Andersson & Vuorela [1995]** vorgestellt. In deren Verfahren werden SPOT-, Landsat TM- und NOAA AVHRR-Daten einem landesweiten Referenzsystem vollautomatisch zugeordnet. Dazu werden zunächst Polygonmerkmale von leicht segmentierbaren Wasserflächen berechnet. Neben Durchmesser, Fläche und Kompaktheit der Regionen werden auch Momente und Invarianten extrahiert. In einer vorläufigen Zuordnung werden Polygone aus den verschiedenen Sensordaten Polygonen eines Referenzdatensatzes zugeordnet. Als Referenzobjekte werden Seen, die aus einer Klassifikation einer bereits georeferenzierten Landsat TM-Szene stammen, benutzt. In einem zweiten Schritt werden Fehlzugeordnungen durch Konsistenzprüfungen eliminiert. Das Verfahren wurde zusätzlich zur Georeferenzierung von Landsat TM-Daten mit See-Polygonen aus gescannten Rasterkarten erfolgreich eingesetzt und ist damit auch für die Bild-Karte-Zuordnung geeignet. Durch die hohe Anzahl an gut segmentierbaren Seen und durch das ebene finnische Gelände konnte für eine NOAA AVHRR Szene ein Punktfehler von 0.46 Pixeln und für eine Landsat TM-Szene ein Punktfehler von 0.43 Pixeln erreicht werden. Dieses Verfahren weist trotz der Besonderheiten der finnischen Topographie einen Erfolg versprechenden Weg zu einem generelleren automatischen System zur Georeferenzierung von Fernerkundungsdaten.

**Flusser & Suk [1993]** präsentieren die Idee, absolute Affininvarianten aus geometrischen Momenten zur Registrierung von Satellitenbildern und Kartendaten zu benutzen. Ausgehend von der Annahme einer näherungsweise affinen Transformation zwischen Bild und Karte werden aus der Theorie der algebraischen Invarianten in Fortsetzung der Arbeiten von Hu [1962] und Dudani, Breeding & McGhee [1977] absolute Affininvarianten abgeleitet. Deren Potenzial zur merkmalsbasierten Zuordnung wird an mehreren Beispielen, unter anderem auch der Bild-Karte-Zuordnung eindrucksvoll demonstriert. Durch die Verwendung absoluter Invarianten sind keine Näherungswerte erforderlich, zwischen den Datensätzen können beliebige affine Transformationen vorliegen. Selbst bei nicht-affinen Transformationen, wie sie bei der perspektiven Bildaufnahme von 3D-Objekten auftreten, erweisen sich die benutzten Invarianten als robust genug für eine erfolgreiche Objektzuordnung.

Das EU-Projekt **ARCHANGEL (Automatic Registration and CHANGE Location) [Dowman & Ruskoné 1997], [Dowman 1998]**, stellt eine Erweiterung des PAIRS-Systems (s.o.) dar. Durch verbesserte Algorithmen zur Extraktion von Polygonen aus Satellitendaten verschiedener Sensoren und Vektorkarten wurde ein robustes System zur automatischen Registrierung von Satellitenbild und Karte realisiert. Satellitenbildpolygone werden darin mit einem zweistufigen Segmentierungsverfahren [Ruskoné & Dowman 1997] extrahiert.

Das Segmentierungsverfahren wurde an verschiedenen optischen und mikrowellenbasierten Sensoren durchgeführt. Die Polygonzuordnung stützt sich auf das Verfahren von Abbasi-Dezfouli & Freeman [1994], das von Newton et al. [1994] zur Zuordnung von Waldpolygonen aus Satellitenbild und Rasterkarte vorgestellt wurde. Die Polygonextraktion aus GIS-Daten geschieht durch eine Selektion nach der Objektart und die anschließende Verschmelzung benachbarter Objekte [Hild & Fritsch 1998]. Auf diese Weise wird sichergestellt, daß die GIS-Objekte auch in den Bildern auffindbar sind. Durch die Verwendung von Polygonen ist dieses Verfahren sehr robust, allerdings werden Näherungswerte für die Pixelgröße und die Orientierung benötigt. Zudem sind die verwendeten Merkmale lediglich translations- und rotationsinvariant.

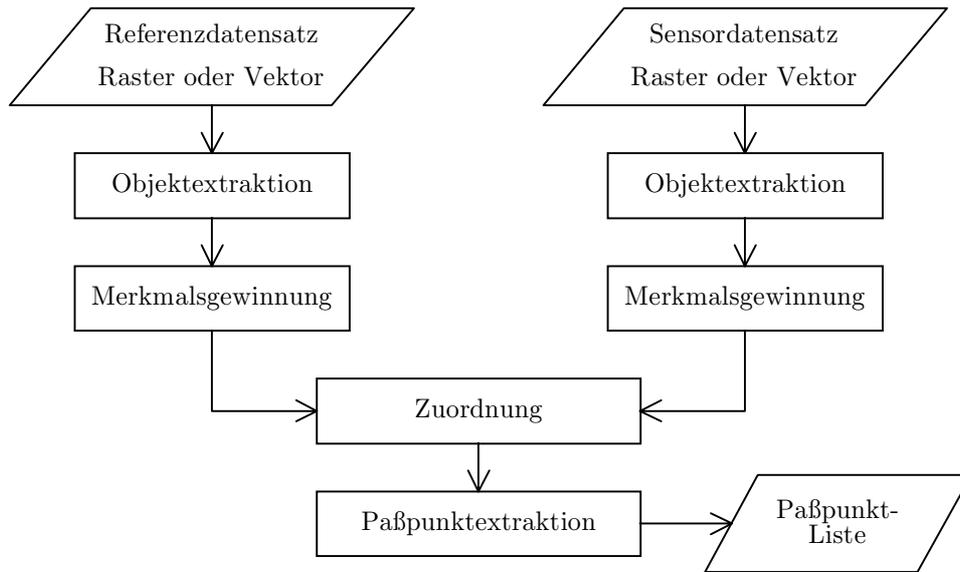
Eine Parallelentwicklung zu ARCHANGEL stellt das System **ARMIES** [Vohra & Dowman 2000] dar. Das ARMIES-System ist weitgehend mit ARCHANGEL identisch. Den wesentlichen Unterschied stellt die Eliminierung grober Punktfehler nach dem Zuordnungsschritt dar. Damit wird eine höhere Registrierungs Genauigkeit erreicht, ohne ein rigoroses Verfahren zu bemühen, das ein DHM voraussetzt. Die Eliminierung von Punktpaaren mit groben Fehlern wird iterativ mit Hilfe eines statistischen Modells durchgeführt. Dabei werden sukzessiv die Punktpaare mit den größten Residuen entnommen bis sich die Standardabweichung von einem zum nächsten Schritt nicht mehr wesentlich ändert. Für die verbleibenden Punkte wird im Falle der Registrierung von SPOT PAN-Daten mit einer Rasterkarte ein Fehler von 0.64 Pixel angegeben. Als Ausblick auf eine weitere Verbesserung der Registrierungs Genauigkeit wird die Benutzung eines für ARCHANGEL entwickelten Verfahrens zur näherungswertfreien Extraktion von Straßenkreuzungen in hochaufgelösten Satellitenszenen [Klang 1998] vorgeschlagen.

Von **Sester, Hild & Fritsch [1998]** wurde ebenfalls ein Ansatz, Polygone zur automatischen Georeferenzierung von Satellitendaten mit Hilfe von GIS-Daten zu benutzen vorgeschlagen. In dieser Arbeit werden die GIS-Waldpolygone aus ATKIS-Daten extrahiert, ähnlich dem Vorgehen in ARCHANGEL. Aus den Satellitendaten (SPOT XS) werden Waldpolygone mit Hilfe einer unüberwachten multispektralen Klassifikation aus den drei Multispektralkanälen (grün, rot, NIR) gewonnen. Für die Zuordnung werden invariante Polygonmerkmale extrahiert (Elongation der am besten passenden Ellipse, Anzahl der Löcher im Polygon, Polygonfüllungsgrad im MBR, Länge-Breite-Verhältnis des MBR und Kompaktheit des Polygons). Über ein Baumsuchverfahren mit unären und binären Zwangsbedingungen zwischen den Objekten (constrained depth-first backtracking tree search, [Grimson, Lozano-Pérez, White & Noble 1993]) werden konsistente Objektpaare ermittelt. Unter Anwendung einer Heuristik, mit der Paarungen großer Objekte zuerst untersucht werden wird der Suchbaum bis zu einer gewissen Tiefe nach einer konsistenten Gruppe von Objekten durchsucht. Aus den Objektschwerpunkten wird fortlaufend die aktuell beste Affintransformation geschätzt. In einem anschließenden Verifikationsschritt werden alle möglichen Objektpaare ermittelt. Ausgehend von der groben Transformation aus der Baumsuche wird die Polygontransformation für alle verifizierten Polygonpaare individuell mit Hilfe des Iterative-Closest-Point-Algorithmus' von Besl & McKay [1992] verbessert. Die erreichbaren Genauigkeiten werden zu 1–3 Pixel angegeben. Das System ist sehr robust und kann auch größere affine Störungen noch richtig handhaben.

Die Ausführungen in diesem und dem vorangehenden Abschnitt erheben keinen Anspruch auf Vollständigkeit, es wurde jedoch versucht, einen Querschnitt über die Möglichkeiten und den Stand der Verfahren zu geben, die für diese Arbeit von Bedeutung sind. Wie z.B. in der Arbeit von Holm et al. [1995] angeklungen ist, erlauben viele automatische Zuordnungsverfahren für Bilder unterschiedlicher Sensoren bereits eine - zumindest grobe - Georeferenzierung von Fernerkundungsdaten mit Kartendaten unterschiedlicher Natur.

## 1.5 Zielsetzung und Übersicht über diese Arbeit

In den vorigen Abschnitten wurde die Notwendigkeit zur Bereitstellung eines automatischen Verfahrens für die Georeferenzierung von Fernerkundungsdaten motiviert. Das Ziel dieser Arbeit war die Entwicklung, Implementierung und Bewertung eines automatisch arbeitenden Systems zur Zuordnung von Original-Bilddaten zu bereits georeferenzierten Daten. Das Hauptaugenmerk liegt dabei auf der Robustheit des Verfahrens gegenüber geometrischen und radiometrischen Schwankungen, sowie in der potenziellen Anwendbarkeit auf verschiedene Fälle (Raster-Vektor-Zuordnung, Raster-Raster-Zuordnung und Vektor-Vektor-Zuordnung). Daten unterschiedlicher Sensoren sollten mit dem zu entwickelnden System georeferenziert und ggf. auch koregistriert werden können. Während der gesamten Arbeit stand die Formulierung und Entwicklung eines durchgängigen Konzeptes im Vordergrund. Auf die Optimierung einzelner Systemteile kann daher nur in für die Gesamtaufgabe wichtigen Teilen eingegangen werden. Die Grundidee des entwickelten Systems besteht darin, in den Sensor- und Referenzdaten polygonale Objekte zu extrahieren und diese merkmalsbasiert zuzuordnen. Auf den Randlinien zugeordneter Polygonen können automatisch beliebig viele Paßpunkte erzeugt werden. Diese Idee geht zurück auf das EU-Projekt ARCHANGEL, an dem der Autor in den Jahren 1996-1999 mitgearbeitet hat. In Abbildung 1.4 ist ein grobes Flußdiagramm des entwickelten Systems zu sehen, das auch die Struktur der vorliegenden Arbeit reflektiert.



**Abbildung 1.4:** Grobes Flußdiagramm für das in dieser Arbeit entwickelte System zur automatischen Georeferenzierung von Fernerkundungsdaten.

Die Bildung geeigneter Objekte aus Raster- und Vektordaten ist Gegenstand des **Kapitels 2 (Objektbildung)**. Für die Gewinnung polygonaler Objekte aus Rasterdaten wird zur Segmentierung ein Bereichswachstumsverfahren mit adaptivem Schwellwert benutzt. Zu Beginn dieses Kapitels wird daher zunächst eine Übersicht über die gängigen Segmentierungsverfahren gegeben, bevor detaillierter auf das benutzte Verfahren eingegangen wird. Zur Bewertung des gewählten Segmentierungsverfahrens wurden Untersuchungen bei unterschiedlichen Startpunkten, unterschiedlichen Schwellwerten und auf verschiedenen Auflösungsstufen eines Bildes durchgeführt, anhand derer die prinzipielle Eignung des gewählten Verfahrens für die Georeferenzierungsaufgabe belegt wird. In einem kurzen Abschnitt wird die Bildung geeigneter Flächenobjekte aus Vektordaten exemplarisch für ATKIS-Daten erläutert.

Laut Vorgabe soll das entwickelte System auf Merkmalen polygonaler Objekte basieren. In **Kapitel 3 (Merkmalsgewinnung)** wird dazu eine Einführung in die Theorie invarianter Formen gegeben und der Weg zu absoluten Affinvarianten aus zentralen Momenten mit Hilfe des Fundamentaltheorems für Momentinvarianten beschrieben. Eine Methode zur Berechnung der Momente direkt aus der Objektkontur wird präsentiert. Zur Untersuchung der Stabilität der absoluten Affininvarianten unter statistischem Rauschen und dem Einfluß verschiedener Rasterweiten wurden verschiedene Experimente durchgeführt, die hier beschrieben werden. Alternativ zu den klassischen Affininvarianten werden sog. Cross-Weighted-Moment-Invarianten (CWM-Invarianten) abgeleitet, die ebenfalls über zentrale geometrische Momente direkt aus der Objektkontur berechnet werden können. Auch für die CWM-Invarianten wurden Untersuchungen zur Stabilität unter statistischem Rauschen und dem Einfluß verschiedener Rasterweiten durchgeführt, die hier beschrieben werden. Ein Vergleich zu den klassischen Affininvarianten wird gegeben. Anhand eines synthetischen Datensatzes wird das Potenzial von CWM-Invarianten zur Diskriminierbarkeit von Objekten untersucht. Abschließend wird ein Ausblick auf weitere mögliche Merkmale gegeben.

**Kapitel 4 (Szene-Modell-Zuordnung)** befasst sich mit der gesamten Problematik der merkmalsbasierten Zuordnung. Der Zuordnungsprozess wird auf eine Baumsuche mit unären und binären Zwangsbedingungen zurückgeführt. Eine Beschreibung der verwendeten Zwangsbedingungen und deren Realisierung wird gegeben. Da während der Baumsuche bereits Transformationen zwischen Objekten eines hypothetischen Paares benötigt werden, wird ein Formalismus zur Gewinnung dieser Transformationen aus zentralen Momenten vorgestellt, der sich auf die sogenannte Standardlage einer affinen Äquivalenzklasse stützt. Verfahren zur Gewinnung der Standardlage aus zentralen Momenten werden ebenfalls vorgestellt. Basierend auf den Ergebnissen der Baumsuche wurde ein Verfahren zur Verifikation von Objektpaarhypothesen entwickelt, das hier erläutert wird. Die verifizierten Objektpaare werden anhand eines für diese Arbeit modifizierten Algorithmus' aus der Computer Vision verbessert. Die angebrachte Modifikation und Untersuchungen zum Einfluß von Verschiebungen, Rotationen und Maßstabsänderungen werden beschrieben. Abschließend in diesem Kapitel wird ein selbst entwickeltes Verfahren zur Gewinnung von Paßpunkten auf den Konturen verifizierter und verfeinerter Objektpaare vorgestellt.

In **Kapitel 5 (Untersuchungen mit dem Gesamtsystem)** wird das Gesamtsystem und Untersuchungen mit verschiedenen Datensätzen vorgestellt. Dazu wird zunächst in einer eindimensionalen Näherung eine analytische Beziehung des Fehlers, der sich durch die Annahme einer ebenen Affintransformation anstelle einer perspektiven Abbildung zwischen Sensor- und Referenzsystem ergibt, abgeleitet. Dies dient der Abgrenzung der in diesem Verfahren getroffenen Annahmen. Anhand simulierter Luftbildaufnahmen aus ATKIS-Daten, die ideal segmentierte Bilder darstellen, wird in einer Vektor-Vektor-Zuordnung die Qualität der automatisch erzeugten Paßpunkte bei variabler Flughöhe und Kammerkonstante untersucht. Die Verwendung simulierter Luftaufnahmen erlaubt dabei eine Trennung in Fehler durch die Objektbildung (die somit wegfallen) und solche, die durch den Geländeformeinfluß entstehen. In einer zweiten Untersuchung wird die automatische Georeferenzierung einer realen SPOT PAN-Szene mit ATKIS-Daten beschrieben, untersucht und diskutiert. Dieser Fall stellt eine Vektor-Raster-Zuordnung dar. In einer dritten Untersuchung wird mit einer Raster-Raster-Zuordnung die automatische Georeferenzierung einer SPOT XS-Szene mit einem SPOT PAN-Orthophoto beschrieben, untersucht und diskutiert. Mit diesen drei Untersuchungen wird die Funktion des Systems in den drei wesentlichen Fällen anhand von realen Daten mit praxisrelevanter Ausdehnung demonstriert und bewertet.

Abschließend wird in **Kapitel 6 (Zusammenfassung und Ausblick)** die gesamte Arbeit und die erzielten Ergebnisse zusammengefasst und ein Ausblick auf weitere Verbesserungsmöglichkeiten gegeben.

## Kapitel 2

# Objektbildung

”Das Ganze ist die Summe seiner Teile.”

*Euclid*

In diesem Kapitel wird die Bildung flächenhafter Objekte aus Raster- und Vektordaten behandelt. Die Objekte müssen dabei so beschaffen sein, daß die zu extrahierenden Merkmale vergleichbar sind, d.h., daß eine merkmalsbasierte Zuordnung, wie sie für das entwickelte Verfahren eingesetzt wird (siehe Kapitel 4), möglich ist. Eine Bedingung für die Auswahl des Verfahrens und geeigneter Objektklassen ergibt sich aus der Forderung, daß die zu extrahierenden Objekte in beiden zu registrierenden Datensätzen vorhanden sind. Besteht die Registrierungsaufgabe z.B. zwischen einem Satellitenbild und einer GIS-Datenbasis im Vektorformat, so können im Satellitenbild nur Objekte gesucht werden, die auch als flächenhafte Objekte im GIS-Datensatz vorhanden bzw. extrahierbar sind. Beispiele hierfür sind Wald-, Wasser oder auch Vegetationsflächen, da diese bei Bodenpixelgrößen von mehr als etwa 1m eine relativ homogene Grauwertverteilung aufweisen und damit prinzipiell gut isolierbar sind. Bei gleicher Bodenpixelgröße erscheinen z.B. besiedelte Flächen relativ inhomogen, sie sind daher für ein grauwertbasiertes Extraktionsverfahren weniger gut geeignet. Prinzipiell hängt die Auswahl der Objektarten und des Extraktionsverfahrens insofern von der Bodenpixelgröße ab, als die betrachteten Objektklassen in bestimmter Hinsicht (Grauwert, Textur, etc.) homogen sein müssen. Das selbe gilt für eine Raster-Raster-Zuordnung, z.B. die Registrierung von SPOT- und LANDSAT-Daten. Bei reinen Vektor-Vektor-Zuordnungen existieren naturgemäß keine Pixel, der Homogenitätsaspekt erübrigt sich von daher in diesem Fall. Ein weiteres Kriterium, das von den zu extrahierenden Flächen neben der Homogenität erfüllt sein muß, ist die Seltenheit. Geeignete Objekte für ein objektbasiertes Zuordnungsverfahren sind demnach solche, die sich hinsichtlich ihrer Form oder ihres mittleren Grauwertes anhand der benutzten Merkmale deutlich von anderen trennen lassen.

## 2.1 Flächenobjekte aus Rasterdaten

Die Zerlegung eines ein- oder mehrkanaligen Rasterbildes in bedeutungsvolle Bereiche bezeichnet man nach Gonzalez & Wintz [1977] als Segmentierung. Dabei läßt sich die Frage, wann ein Bereich ”bedeutungsvoll” ist, nicht global, sondern nur im Zusammenhang mit einem gegebenen Problem beantworten. Beispiele für bedeutungsvolle Bildbereiche sind etwa: Flächen einheitlicher Vegetation oder Gebäude (Luft- oder Satellitenbildauswertung), Flächen eines 3D-Objekts (Nahbereichs-Bildverarbeitung) oder auch Flugzeuge (militärischer Bereich). Nach Bässmann & Behrlich [1991] wird unterschieden in Verfahren, die zusätzlich eine Interpretation der erhaltenen Segmente liefern und solchen, die lediglich in gewisser Weise homogene Bereiche geometrisch voneinander trennen. Bereiche eines Bildes, die gemäß gewisser Homogenitätskriterien oder Ähnlichkeitsmaße räumlich zusammenhängen, werden als Regionen bezeichnet. Homogenität kann dabei in 2D-Rasterdaten (ein- oder mehrkanalig) z.B. bezüglich der Farbe, des Grauwerts oder der Textur gefordert werden. Die Segmentierung, d.h. die Zerlegung eines Bildes in Regionen, besitzt nach Haberäcker [1987] und Haralick & Shapiro [1992] folgende wesentliche Eigenschaften:

1. Jedes Pixel gehört zu genau einer Region. Mehrfach- (Überlappungen) oder Null-Zugehörigkeiten sind nicht erlaubt (Gleichungen 2.1 und 2.2).
2. Jedes Pixel einer Region liefert ein logisches ”Wahr” im Sinne des angewandten Ähnlichkeitsmaßes (Gleichung 2.3).
3. Regionen sind maximal, d.h. jedes an die Region angrenzende Pixel würde ein logisches ”Falsch” im Sinne des angewandten Ähnlichkeitsmaßes liefern.
4. Benachbarte Regionen unterscheiden sich signifikant bezüglich des verwendeten Homogenitätskriteriums für eine Region (Gleichung 2.4).

Neben diesen eher mathematisch formulierten Eigenschaften sollten die erzeugten Regionen in ihrem Generalisierungsgrad nach Möglichkeit dem verwendeten Kartenmaßstab, z.B. 1:25000 bei SPOT PAN-Daten, entsprechen. Zur Klärung der Frage, wann zwei Pixel benachbart sind, verwendet man entweder die 4er- oder die 8er-Nachbarschaft (siehe z.B. [Demant, Streicher & Waskewitz 1998]). Im Falle der 4er Nachbarschaft gelten nur Pixel als Nachbarn, die in Zeilen- oder Spaltenrichtung an das interessierende Pixel angrenzen, nicht die vier

diagonalen Nachbarn. Im Falle der 8er-Nachbarschaft gelten die acht nächstliegenden Pixel als Nachbarn eines Pixels. Laut Haralick & Shapiro [1992] existiert keine geschlossene Theorie der Segmentierung. Die einzelnen Verfahren sind vielmehr ad hoc, d.h. angepaßt an die Eigenschaften der zu separierenden Objekte, formuliert. Das wesentliche Ziel einer Segmentierung besteht darin, interessierende Objekt zu separieren, von denen in weiteren Verarbeitungsschritten die gewünschten Merkmale berechnet werden können. Die zu erhaltenden Objekte können dabei in unterschiedlicher Weise repräsentiert sein: binär (Flächen, Punkte, Kanten), als Kontur (Flächen), oder als Skelett (Linien) [Rosenfeld & Kak 1976]. In einer etwas abstrakteren Formulierung läßt sich das allgemeine Segmentierungsproblem nach Pinz [1994] folgendermaßen beschreiben:

**Definition 1** Ein Bild  $B$  sei die Menge aller Pixel  $p(x, y)$  und eine Segmentierung  $S$  bestehe in einer Aufteilung von  $B$  in  $N_R$  verschiedene Regionen unter einem Homogenitätskriterium  $H$ . Für die Gesamtheit aller Regionen  $S = \{R_i \mid i = 1 \dots N_R\}$  müssen dann unter Annahme einer 4er-Nachbarschaft die folgenden Bedingungen erfüllt sein:

$$\bigcup_{i=1}^{N_R} R_i = B, \quad \text{wobei} \quad \forall i : R_i \subseteq B \quad (2.1)$$

$$\forall i \neq j : R_i \cap R_j = \emptyset \quad (2.2)$$

$$H(R_i) = WAHR \quad (2.3)$$

$$H(R_i \cup R_j) = FALSCH, \text{ wenn } R_i \text{ Nachbar von } R_j \text{ ist.} \quad (2.4)$$

### 2.1.1 Segmentierung von Rasterbildern

In der digitalen Bildverarbeitung und im Bereich des Maschinensehens wurde eine kaum mehr überschaubare Vielzahl von unterschiedlichsten Segmentierungsverfahren entwickelt. Es liegt außerhalb des Anspruchs dieser Arbeit, dieser Vielfalt ein weiteres Verfahren hinzuzufügen. Trotz der ungeheuren Anzahl lassen sich die Bildsegmentierungsverfahren jedoch in drei wesentliche Gruppen unterteilen: Punktorientierte, kantenorientierte und regionenorientierte Verfahren [Jähne 1991].

#### Punktorientierte Segmentierungsverfahren

Bei punktorientierten Segmentierungsverfahren wird das Ergebnis pixelweise anhand des Grauwerts oder der Grauwerte einer festen Umgebung bestimmt. Im einfachsten Fall wird aus einem Grauwertbild lediglich ein Binärbild durch Anwendung eines festen Schwellwertes  $c$  erzeugt. Dabei wird die Grauwertfunktion  $g(x, y)$  folgendermaßen in ein Binärbild transformiert:

$$g'(x, y) = \begin{cases} 1 & \text{falls } g(x, y) > c \\ 0 & \text{sonst} \end{cases} \quad (2.5)$$

Der Schwellwert  $c$  kann dabei entweder mit Hilfe von Vorwissen (z.B. bei bekannter Objektgröße oder bei konstanten Beleuchtungsverhältnissen) oder mit Hilfe einer Separation von Moden im globalen oder lokalen Histogramm bestimmt werden. Derartige Verfahren eignen sich zur Trennung zwischen Objekt und Hintergrund, wenn die jeweiligen Grauwertbereiche wesentlich verschieden sind, d.h. wenn Objekte und Hintergrund im Histogramm in getrennten Moden erscheinen. Gegebenenfalls kann mit individuellen Schwellwerten pro Objekt gearbeitet werden. Für den Fall, daß die Grauwerte der interessierenden Objekte in einem mittleren Bereich liegen, verwendet man eine sogenannte Bereichsbinarisierung mit zwei Schwellwerten  $c_1$  und  $c_2$  pro Objektklasse. Die Abbildung der Grauwertfunktion  $g(x, y)$  nimmt dann folgende Gestalt an:

$$g'(x, y) = \begin{cases} 1 & \text{falls } c_1 < g(x, y) < c_2 \\ 0 & \text{sonst} \end{cases} \quad (2.6)$$

Soll lediglich ein gewisser Hintergrund ausmaskiert werden, so kann nach Rosenfeld & Kak [1976] ein sog. *Semithresholding*-Verfahren benutzt werden, bei dem nur Pixel unterhalb eines Schwellwerts zu 0 gesetzt werden. Pixel mit höheren Grauwerten bleiben unverändert. Im Falle nichthomogener Grauwertverteilungen in Objekt und Hintergrund kann ein dynamisches Verfahren eingesetzt werden, bei dem die Schwellwerte ortsabhängig bestimmt werden [Haberäcker 1987]. Schwellwertverfahren können auch auf multidimensionalen Daten durchgeführt werden; die entsprechenden Homogenitätskriterien müssen dann für jede Region kanalweise erfüllt werden. Neben Segmentierungsverfahren, die auf dem Grauwert eines Pixels beruhen werden auch solche verwendet, die anhand bestimmter Grauwertverteilungen in der Umgebung eines Pixels eine Binarisierung vornehmen, z.B. auf Texturmerkmalen basierende Segmentierungen [Haralick & Shapiro 1992]. In der Regel wird dabei zunächst ein lokaler Fenster-Operator mit dem Grauwertbild gefaltet. Auf dem Faltungsergebnis (z.B. Texturmaß) wird anschließend ein punktorientiertes Segmentierungsverfahren durchgeführt.

### Kantenorientierte Segmentierungsverfahren

Eine andere Klasse von Segmentierungsverfahren basiert auf Grauwertkanten, die in der Regel durch geeignete Filteroperatoren aus einem Grauwertbild gewonnen werden. Eine Übersicht zu gängigen Kanten-Operatoren ist z.B. in Pratt [1991] zu finden. Kantenorientierte Verfahren versuchen, aus den Grauwert- oder Texturkanten im Bild geschlossene Kurven zu erzeugen, die dann als Objektrand interpretiert werden. Voraussetzung für einen Erfolg dieser Verfahren ist natürlich das Auftreten eines Grauwert- oder Textursprungs am Objektrand. Bässmann & Behrlich [1991] gliedern die kantenorientierte Segmentierung in folgende Schritte:

1. Konturpunktdetektion: Gradientenbildung mittels eines Kantenoperators. Bildbereiche mit hohen Gradienten werden dabei als potenzielle Objektkanten gesehen.
2. Konturverdünnung: Reduktion der potenziellen Objektkonturen auf Pixel mit lokal maximalem Gradienten.
3. Konturpunktverkettung: Einzelne Konturpunkte werden zu möglichst langen zusammenhängenden Linien verbunden.
4. Konturapproximation: Ausgleichung der ermittelten Liniensegmente nach gewissen Regeln (z.B. Parallelität oder Rechtwinkligkeit oder "weicher" Krümmungsverlauf).

Oftmals liegt jedoch nach Durchlaufen dieser Schritte noch keine geschlossene Kontur vor, weshalb sich eine manuelle Nachbearbeitung anschließen muß. Ein alternatives Verfahren stellt die sog. Konturverfolgung dar, wie sie in Parker [1997] beschrieben wird, die per Definition zu geschlossenen Linienzügen führt. Zur effizienteren Handhabung einer Objektkontur wird oft eine spezielle Repräsentation gewählt. Im sog. Chain-Code [Freeman 1961] wird - beginnend von einem Startpixel - die Richtung zum nächsten Konturpixel abgespeichert. Je nach verwendeter Nachbarschaft (4er- oder 8er-) gibt es vier (N-O-S-W) bzw. acht (N-NO-O-SO-S-SW-W-NW) mögliche (Himmels-)Richtungen. Eine andere Möglichkeit der Konturkodierung stellt der sog. RULI-Code dar. Dabei wird unter Annahme einer 4er-Nachbarschaft die Art kodiert, mit der eine virtuelle, reellwertige Kontur das aktuelle Pixel schneidet. Diese Kodierung hängt damit von beiden Nachbarn auf der skelettierten Kontur ab. Der Name rührt von den vier möglichen Fällen her: R (Kontur "biegt" im Pixel nach rechts ab), U (Kontur kehrt um = "U-turn"), L (Kontur "biegt" im Pixel nach links ab), I (Kontur durchquert Pixel geradlinig = "Intersection"). Durch Interpolation der Grauwerte längs des maximalen Richtungsgradienten läßt sich an der Stelle des Wendepunktes der interpolierten Kurve auch eine subpixelgenaue Kantenlokalisierung erreichen ("Zero-crossings" = Nullstellen der zweiten Ableitung). Eine weitere Art, subpixelgenau geschlossene Objektkonturen zu erhalten, basiert auf sog. "Edgels" (edge elements). Ein Edgel ist die reellwertige Repräsentation einer Kante durch einen Punkt, eine Richtung und die Gradientenstärke. Durch geeignete Gruppierung von Edgels läßt sich eine geschlossene Kontur in Vektorform erzeugen [Steinwendner & Schneider 1998]. Die Erzeugung einer geschlossenen Objektkontur kann auch durch die Benutzung von 2D-Modellen unterstützt werden. Im Falle parametrischer Modelle wird hier die Hough-Transformation verwendet [Hough 1962], bei Verwendung geometrischer Modelle kommen sog. Active Contour Models ("Snakes") zum Einsatz [Kass, Witkin & Terzopoulos 1987], mit denen initiale Konturen durch Minimierung der Energie an Grauwertkanten, Punkte oder Linien "herangezogen" werden. Für große, komplexe Szenen, wie sie in Fernerkundungsaufnahmen in der Regel vorliegen, sind kantenorientierte Verfahren zur Segmentierung von Flächenobjekten aufgrund ihres erhöhten Rechenaufwands gegenüber punkt- und regionenorientierten Verfahren benachteiligt.

### Regionenorientierte Segmentierungsverfahren

Während die kantenbasierten Verfahren besser für langgestreckte Objekte geeignet sind, wird bei kompakteren Objekten eher ein regionenorientiertes Verfahren eingesetzt [Pinz 1994]. In der Literatur werden verschiedene regionenorientierte Segmentierungsverfahren vorgestellt: Bereichswachstumsverfahren ("Region Growing") [Haralick & Shapiro 1992], [Jähne 1991], Breichs-Gruppierung ("Region-Clustering") [Gonzalez & Wintz 1977], Split-and-Merge-Verfahren [Rosenfeld & Kak 1976], [Haralick & Shapiro 1992] oder Scale-Space-Blob-Verfahren [Pinz 1994], um nur die wichtigsten zu nennen. Generell liefern die regionenorientierten Verfahren, gleich wie die punktorientierten, flächenhafte Objekte, die gewissen Homogenitätskriterien genügen, jedoch zu Ungenauigkeiten der Kontur neigen. Im Gegensatz dazu liefern kantenorientierte Verfahren genauere Konturen, deren Inneres jedoch nicht notwendigerweise eine homogene Grauwertverteilung oder Textur besitzt. Für diese Arbeit wurde ein Bereichswachstumsverfahren verwendet, das in den folgenden Abschnitten näher beschrieben und untersucht wird. Daher soll hier auf die anderen Verfahren nicht näher eingegangen werden.

Die Bildsegmentierung mittels eines Bereichswachstumsverfahrens läuft nach Pinz [1994] im wesentlichen in folgenden Schritten ab:

1. Erzeugung von Saatzellen im Bild, d.h. Bestimmung von Pixeln, an denen das Wachstumsverfahren beginnt. Diese Bestimmung kann statistisch oder modelldatengestützt erfolgen, es kann auch eine einfache Regel bestehen, z.B.: "Nächstes Saatzpixel = Das am weitesten links unten liegende Pixel, das noch zu keiner Region gehört".
2. entweder (iterativ, bis keine Änderung mehr eintritt):  
Paralleles Wachstum aller Saatregionen durch Betrachtung aller 4er-Nachbarn, die noch zu keiner Region gehören, und Hinzunahme zur jeweiligen Region, falls das Homogenitätskriterium damit erfüllt bleibt. Im Falle einer Berührung von Regionen: Verschmelzung, falls die vereinte Region das Homogenitätskriterium erfüllt.  
oder:  
Ausgehend vom am weitesten links unten liegenden Pixel wird die Region solange vergrößert, bis kein Nachbar mehr im Sinne des Homogenitätskriteriums hinzugefügt werden kann. Dann wird eine neue Region mit dem jetzt am weitesten links unten liegenden und nicht zu einer bereits vorhandenen Region gehörigen Pixel begründet und vergrößert, so weit es geht usw. bis das ganze Bild segmentiert ist.
3. Pixel, die zu keiner Region gehören, werden entweder als nicht zusammenhängender Hintergrund markiert oder es werden jeweils zusammenhängende Gruppen von Pixeln als neue Regionen gespeichert oder es werden automatisch neue Saatzpunkte in diese Bereiche gelegt und das Bereichswachstumsverfahren wird fortgesetzt.

### 2.1.2 Ein Bereichswachstumsverfahren mit adaptivem Schwellwert

Das für diese Arbeit verwendete Segmentierungsverfahren ist ein Bereichswachstumsverfahren mit adaptiver Schwellwertanpassung. Das Verfahren beruht auf den Arbeiten von Levine & Shaheen [1981] und ist in dem Software-Paket KBVision, in dem das Gesamtsystem implementiert wurde, in Form des Basis-Tasks GROWREG enthalten. Der Steuerparameter des Verfahrens ist die maximale Differenz zum mittleren Grauwert der Region. Der vorgegebene Schwellwert kann dabei der Grauwertstatistik einer Region angepasst werden (Adaption), um z.B. eine unerwünschte Vergrößerung inhomogener Regionen zu verhindern. Etwas mathematischer formuliert drückt sich der Sachverhalt folgendermaßen aus:

**Definition 2** Ein  $n$ -kanaliges Bild sei repräsentiert durch die  $n$  Bildebenen  $P_q(c, r)$ ,  $q = 1, \dots, n$ . Der mittlere Grauwert einer Region des Bildes ist kanalweise definiert durch:

$$F_{R_i}(q) = \frac{1}{N_i} \sum_{(c,r) \in R_i} [P_q(c, r)], \quad (2.7)$$

wobei  $R_i$  eine Region des Bildes bezeichnet und  $N_i$  die Anzahl der Pixel in  $R_i$  ist.

Die Standardabweichung der Grauwerte innerhalb einer Region  $R_i$  ist kanalweise gegeben durch

$$\hat{\sigma}_{R_i}(q) = \begin{cases} \frac{1}{\sqrt{N_i-1}} \left[ \sum_{(c,r) \in R_i} [P_q(c, r) - F_{R_i}(q)]^2 \right]^{\frac{1}{2}} & \text{für } N_i > 1 \\ 0 & \text{sonst} \end{cases}. \quad (2.8)$$

Das Bereichswachstum mit adaptivem Schwellwert läuft in folgenden Schritten ab:

1. Zu Beginn gehören alle Pixel keiner Region an.
2. In der linken unteren Ecke beginnend wird das Bild ausgelesen, bis ein Pixel, das zu keiner Region gehört, gefunden wird (dies ist beim ersten Durchlaufen natürlich das linke untere Pixel). Dieses Pixel begründet eine neue Region, der eine fortlaufende Nummer zugeordnet wird. Das aktuelle Pixel befinde sich an der Stelle  $(c^*, r^*)$ .
3. Falls das aktuelle Pixel in der 4er-Nachbarschaft einer oder mehrerer Regionen liegt, wird es hypothetisch jeder dieser Regionen zugeschlagen. Der neue mittlere Grauwert jeder dieser Regionen berechnet sich dann im  $q$ -ten Kanal zu:

$$F'_{R_i}(q) = \frac{1}{N_i + 1} \sum_{(c,r) \in R_i} [P_q(c^*, r^*) + N_i F_{R_i}(q)] \quad (2.9)$$

und die aktualisierten Standardabweichungen erhält man aus:

$$\hat{\sigma}'_{R_i}(q) = \frac{1}{\sqrt{N_i}} \left[ [P_q(c^*, r^*) - F'_{R_i}(q)]^2 + \sum_{(c,r) \in R_i} [P_q(c, r) - F_{R_i}(q)]^2 \right]^{\frac{1}{2}}. \quad (2.10)$$

Falls dies nicht der Fall ist, folgt direkt Schritt 7.

4. Der adaptierte Schwellwert für alle hypothetischen Regionen berechnet sich aus dem globalen Schwellwert  $h$  zu:

$$h'_{R_i}(q) = \left( 1 - \min(0.8, \frac{\sigma'_{R_i}(q)}{F'_{R_i}(q)}) \right) h. \quad (2.11)$$

5. Für jede hypothetische Region wird die Grauwertdifferenz zwischen aktuellem Pixel  $(c^*, r^*)$  und neuem Regionenmittelwert  $F'_{R_i}(q)$  in jedem Kanal berechnet:

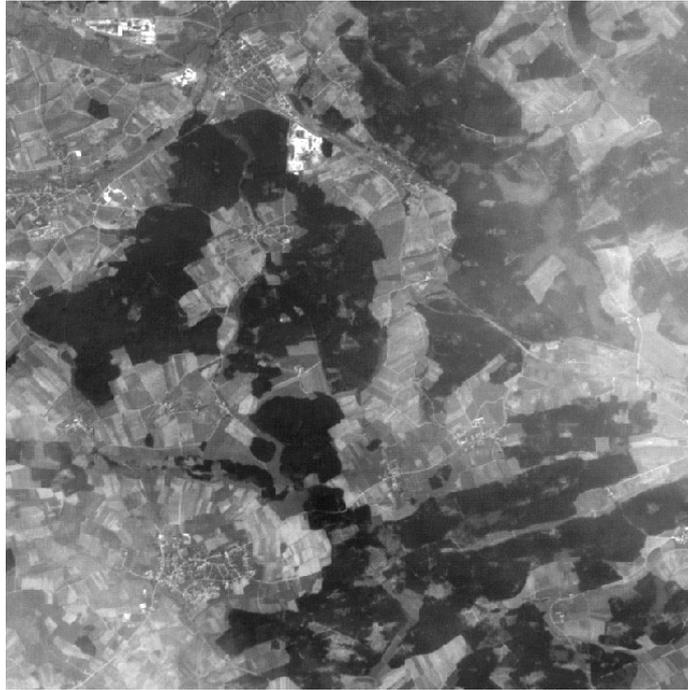
$$\Delta F_{R_i}(q) = |P_q(c^*, r^*) - F'_{R_i}(q)| \quad (2.12)$$

6. Falls  $\Delta F_{R_i}(q) \leq h'_{R_i}(q) \forall q = 1, \dots, n$  erfüllt ist, wird das aktuelle Pixel dieser Region zugeschlagen. Falls dieses Kriterium für mehrere Regionen erfüllt ist, fällt das Pixel an die Region mit der kleinsten Distanz  $\sum_q [\Delta F_{R_i}(q)]^2$ . Alle anderen Regionen werden wieder in ihren vorigen Zustand versetzt.
7. Falls keine der angrenzenden Regionen das Kriterium in Schritt 6 erfüllt oder falls es (mit Ausnahme des linken unteren Pixels) keine benachbarte Region gibt, so wird mit Pixel  $(c^*, r^*)$  eine neue Region mit der nächsten fortlaufenden Nummer erzeugt. Es schließt sich Schritt 3 an. Gibt es kein weiteres regionenfreies Pixel mehr, so ist das Verfahren beendet.

Die Ergebnisse dieses Verfahrens sind von der Wahl des Schwellwertes  $h$  abhängig. Kleinere Schwellwerte führen zu kleineren und mehr Regionen, also eher zu einer Übersegmentierung, größere Schwellwerte  $h$  führen zu größeren und weniger Regionen, also eher zu einer Untersegmentierung. Wie aus Gleichung 2.11 zu ersehen ist, kann sich der adaptierte Schwellwert  $h'$  nur innerhalb eines Bereichs  $[0.2h, \dots, h]$  bewegen. Der Adaptionsmechanismus aus Gleichung 2.11 bewirkt somit, daß der Schwellwert für Regionen mit relativ hoher Standardabweichung im Extremfall bis auf 20% verkleinert wird. Damit wird ein übermäßiges Wachstum inhomogener Regionen eingeschränkt. Bei sehr homogenen Regionen findet dagegen kaum eine Adaption statt, es wird annähernd der voreingestellte Schwellwert  $h$  verwendet.

Mit der Sequenz in Abbildung 2.2 wird die Abhängigkeit des Verfahrens vom Startpunkt (alle vier Bildecken) illustriert. Wie in dieser Abbildung gut zu erkennen ist, liefern die ersten drei Segmentierungen relativ ähnliche Ergebnisse für die Waldflächen, wogegen die Segmentierung, die rechts unten beginnt, ein wesentlich anderes Ergebnis liefert. Die Ursache hierfür ist darin zu finden, daß in den ersten drei Fällen das Startpixel in einem helleren Bereich liegt, während es im vierten Fall innerhalb eines dunkleren Bereichs liegt. Aufgrund ihrer klaren Abgrenzung im Bild sind die Waldflächen für diese Szene geeignete Zielobjekte für eine Segmentierung. Mit dieser Sequenz soll auch verdeutlicht werden, daß es selbst bei für den Menschen klar erkennbaren Objekten sehr schwierig ist, ein einigermaßen robustes Segmentierungsverfahren zu formulieren. Gewiss könnte man das benutzte Verfahren weiter raffinieren, etwa durch iterative Segmentierung mit Saatpunkten in den Schwerpunkten der Regionen des vorigen Durchlaufs. Dies würde jedoch den Rahmen dieser Arbeit bei weitem sprengen und würde mit hoher Wahrscheinlichkeit immer noch eine sehr proprietäre Lösung hervorbringen. Ein Ziel dieser Arbeit ist jedoch auch die Robustheit gegenüber variierenden Eingabedaten, von daher ist die hier angewandte Strategie eine andere. Um einen automatischen Ablauf des gesamten Verfahrens bei gleichzeitig weitgehender Unabhängigkeit von den radiometrischen Verhältnissen im Bild zu ermöglichen, wird das besprochene Segmentierungsverfahren hintereinander mit mehreren verschiedenen Schwellwerten  $h$  durchlaufen. Die dabei zu Grunde liegende Annahme besteht darin, daß der Schwellwertbereich breit genug ist, um von einer Über- bis zu einer Untersegmentierung bezüglich der gesuchten Objekte alle Fälle abzudecken. Des Weiteren wird angenommen, daß innerhalb dieses Parameterbereichs eine hinreichende Menge an interessierenden Objekten mit akzeptabler Genauigkeit segmentiert wird. Daraus ergibt sich eine Bedingung für das weitere Vorgehen: Die zu entwickelnde Strategie muß so ausgelegt sein, daß eine in weiten Grenzen variable Anzahl von Bildobjekten mit z.T. sehr unterschiedlicher Qualität zum Ziel führt. Dieser Aspekt wird in den Kapiteln 3 und 4 weiter vertieft.

Die in Abbildung 2.3 dargestellte Bildsequenz zeigt die Abhängigkeit der Segmentierung mit adaptivem Schwellwert von der Wahl des Schwellwertes  $h$  anhand eines Ausschnittes einer SPOT PAN-Szene des Testgebietes Elchingen-Großkuchen (1.4.1993). Wie in Abbildung 2.3 zu erkennen ist, ergeben sich für den verwendeten



**Abbildung 2.1:** Ausschnitt aus einer panchromatischen SPOT-Szene des Testgebietes Elchingen/Großkuchen vom 01.04.1993 als Basis für die Beispiele zur Segmentierung. Es wurde eine Histogrammanpassung vorgenommen.

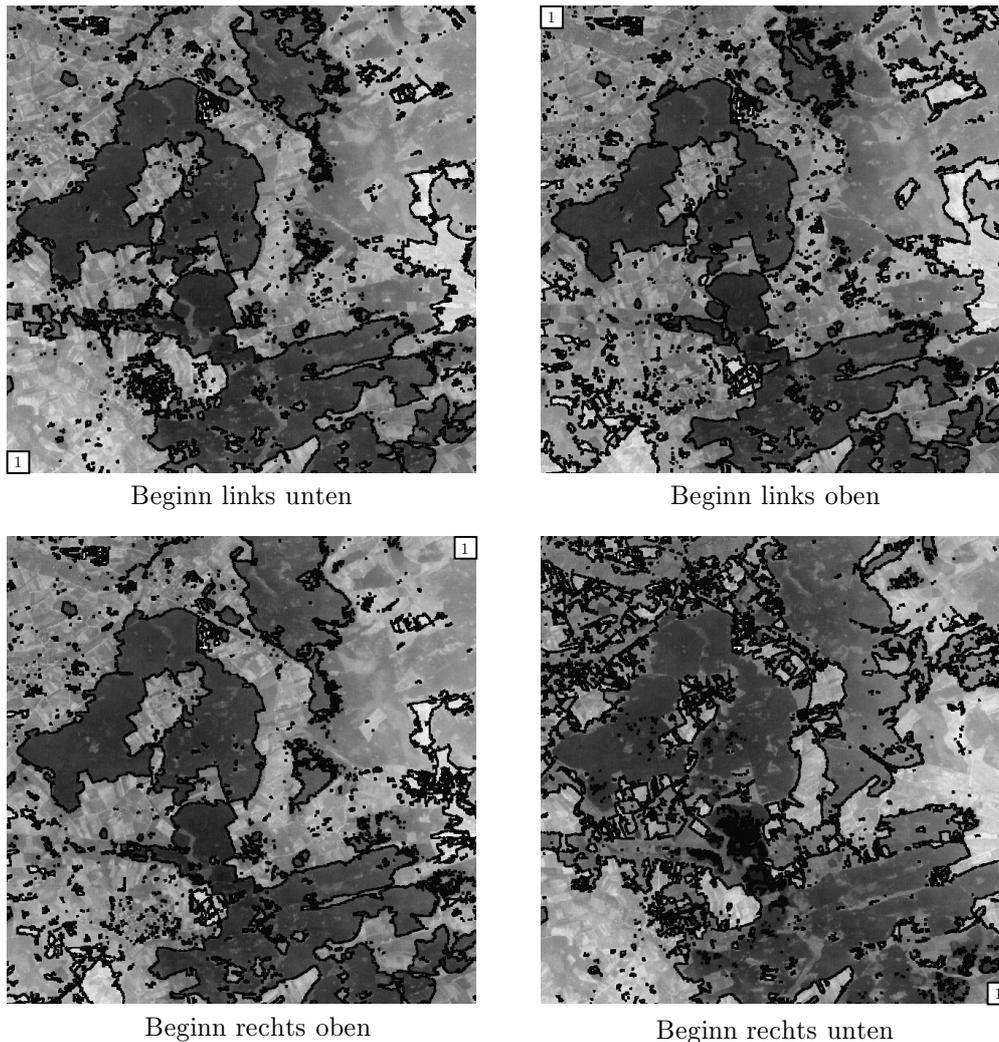
Schwellwertbereich von  $h \in [6, 12]$  verschiedene Resultate, die - der Annahme entsprechend - von einer Über- ( $h = 6$ ) bis zu einer Untersegmentierung ( $h = 12$ ) reichen. Aus der visuellen Beurteilung lässt sich zudem erkennen, daß eine größere Anzahl kleinerer und größerer flächiger Objekte in mindestens einem Segmentierungsdurchlauf in akzeptabler Weise segmentiert werden kann. Auf eine quantitative Analyse der Segmentierungsqualität einzelner Objekte sei hier verzichtet und für eine Rechtfertigung dieser Strategie auf die Beispiele in Kapitel 5 verwiesen.

### 2.1.3 Segmentierung in verschiedenen Pyramidenstufen

Bei der Georeferenzierung ganzer Satellitenszenen mit typischerweise mehreren Millionen Pixeln (z.B. SPOT PAN: etwa 36 Mio. Pixel) ist der zeitliche Aufwand für die Segmentierung sehr erheblich. Hier bietet sich die Verwendung von Bildpyramiden an (siehe dazu etwa [Rosenfeld & Kak 1976] oder [Gonzalez & Wintz 1977]). Bildpyramiden sind in der digitalen Photogrammetrie seit längerer Zeit intensiv im Gebrauch, eine Übersicht über deren Erzeugung und Nutzung für automatische Punktzuordnungsverfahren und zur Repräsentation von digitalen Geländemodellen ist in Ackermann & Hahn [1991] zu finden. Demnach stellt eine Bildpyramide eine Serie von Bildern mit jeweils gleichem Reduktionsfaktor dar. Das Bild nullter Stufe ist das Originalbild, verwendet man etwa einen Kantenreduktionsfaktor von 2, so besitzt das Bild der jeweils nächsten Pyramidenstufe nur noch die halbe Kantenlänge und ein Viertel des Flächeninhaltes. Das bedeutet, daß beim Übergang von einer Stufe zur nächsten aus vier Pixeln ein Pixel interpoliert wird. Bezüglich des Reduktionsfaktors  $k$  und der Interpolation existieren verschiedene Ansätze, auf die hier jedoch nicht näher eingegangen werden soll.

Der Nutzen von Bildpyramiden in der digitalen Photogrammetrie liegt einerseits in der schnelleren Verarbeitungszeit in Bildern höherer Stufen (die Pixelanzahl verringert sich mit jeder Stufe auf  $\frac{1}{k^2}$ ) und andererseits - wie etwa bei der intensitätsbasierten Punktzuordnung - in einer räumlichen Vergrößerung des Konvergenzradius' der verwendeten Verfahren. Mit Bildpyramiden läßt sich ein sogenanntes "coarse-to-fine"-Vorgehen verwirklichen. Dabei werden - ausgehend von einer niedrigen Auflösungsstufe - die Ergebnisse immer weiter verfeinert, um in der höchsten Auflösungsstufe zu Genauigkeiten zu kommen, die direkt nicht oder nur mit erheblich höherem Zeitaufwand zu erhalten wären.

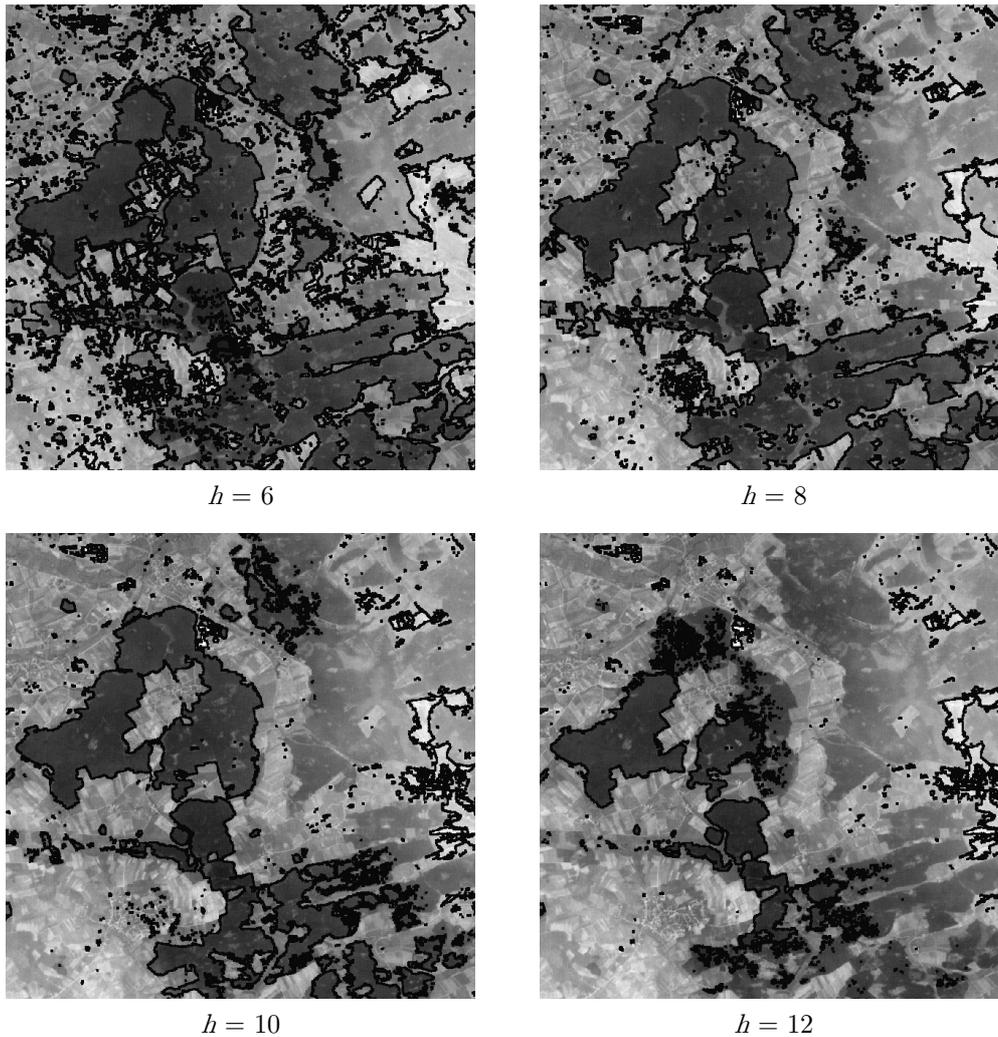
Das in dieser Arbeit entwickelte System soll die Georeferenzierungsaufgabe automatisch zumindest näherungsweise lösen. Je nach erforderter Genauigkeit der Georeferenzierung beziehungsweise Registrierung kann es dabei ausreichen, mit verringerter Bildauflösung zu arbeiten und so sehr viel Rechenzeit zu sparen. Aus diesem Grund



**Abbildung 2.2:** Abhängigkeit des Ergebnisses der Segmentierung eines Ausschnittes einer SPOT PAN Szene mit adaptivem Schwellwert  $h = 8$  von der Lage des Startpunkts.

wird im folgenden nun kurz das qualitative Verhalten der verwendeten Segmentierung in den einzelnen Stufen einer Bildpyramide beleuchtet. Auch hier wird auf eine eingehende quantitative Untersuchung verzichtet und auf die Beispiele in Kapitel 5 verwiesen. In Abbildung 2.4 sind die Ergebnisse der beschriebenen Segmentierung der Pyramidenstufen 0-3 des bereits mit Abbildung 2.1 eingeführten Bildbeispiels mit adaptivem Schwellwert  $h = 8$  zu sehen. Wie darin deutlich zu erkennen ist, sind viele der interessierenden flächenhaften Objekte in allen oder zumindest den ersten drei verwendeten Pyramidenstufen in akzeptabler Weise segmentiert. Damit liegt es nahe, die Segmentierung nicht auf der Originalauflösung, sondern auf einer höheren Pyramidenstufe durchzuführen. Allerdings ist durch die mit jeder Pyramidenstufe einhergehende Vergrößerung der Bodenpixelgröße und damit einer Generalisierung der Objektkontur selbst bei bestmöglicher Segmentierung mit einer Erhöhung der Fehler der Georeferenzierung zu rechnen. Es muß daher ein Kompromiss zwischen gewünschter Genauigkeit und eingesetzter Rechenzeit gefunden werden. Bei kommerziell verfügbaren Satellitenszenen wird sich die benutzte Pyramidenstufe im wesentlichen zwischen 0 und 3 bewegen, da sonst einzelne Objekte zu klein werden und da in der dritten Pyramidenstufe (Flächenverringerng auf  $\frac{1}{64}$  bei  $k = 2$  !) die benötigte Rechenzeit in akzeptablen Grenzen bleibt.

Das Resultat dieses Verarbeitungsschrittes (Extraktion von flächenhaften Objekten aus Rasterdaten) sind Vektorpolygone, die die äußere Umrißlinie der segmentierten Regionen darstellen. Löcher in den Regionen werden hier nicht betrachtet. In einer weiteren Ausbaustufe des Systems könnten Löcher jedoch letztlich zu einer Verdichtung der gewonnenen Paßpunkte und damit einer Verbesserung des gesamten Prozesses beitragen. Die gewonnenen Vektorpolygone folgen den Pixelkanten und besitzen daher viele Eckpunkte. Zur besseren Handhabung der Polygone bietet sich insbesondere bei großen Szenen eine Generalisierung, etwa mit Hilfe des Douglas-Peucker-



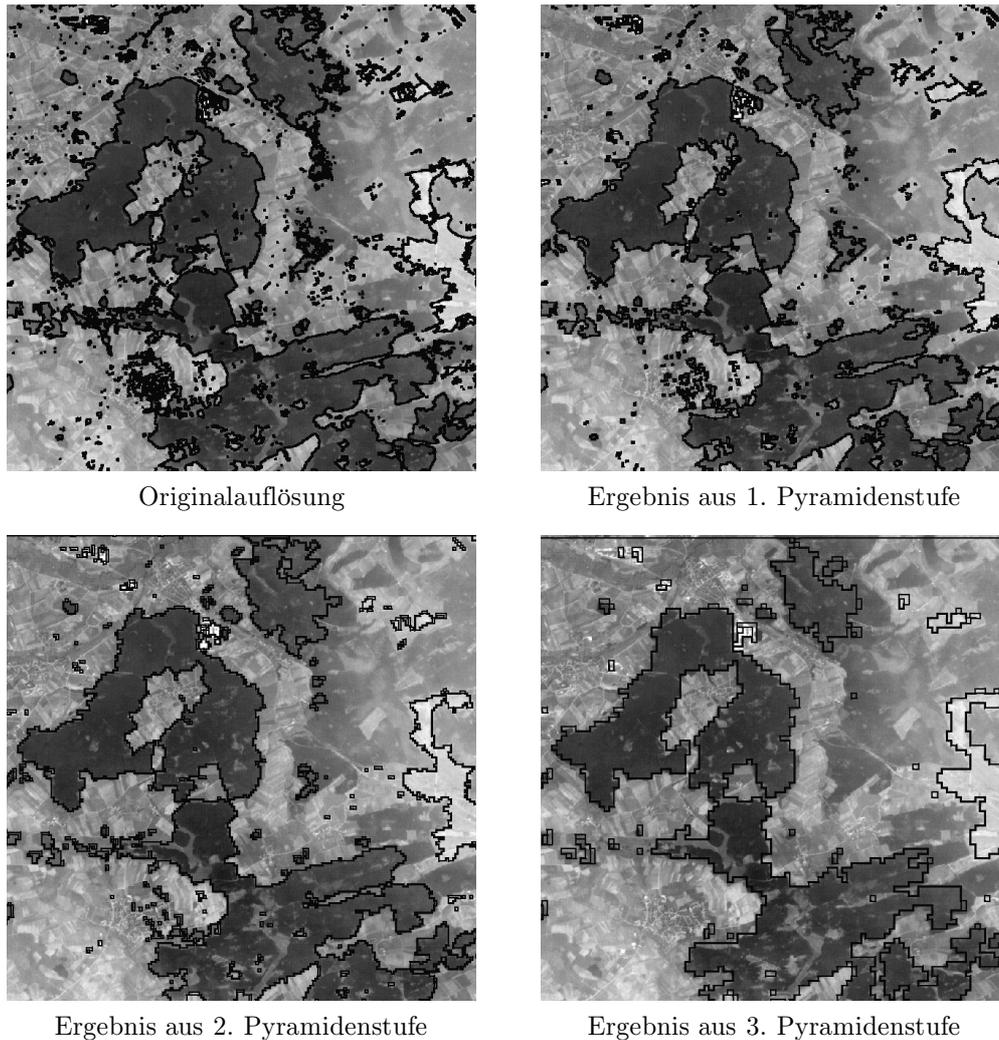
**Abbildung 2.3:** Segmentierung eines Ausschnittes einer SPOT PAN Szene mit variablem adaptivem Schwellwert  $h$ .

Algorithmus' [Douglas & Peucker 1973], an. Infolge der Generalisierung ist dann jedoch ebenfalls mit einer Zunahme der Punktfehler zu rechnen und es muß im Einzelfall geprüft werden, ob eine verringerte Genauigkeit des Ergebnisses noch ausreichend für die gewünschte Anwendung ist.

Eine Kombinationsstrategie für die automatische Georeferenzierung mit günstiger Rechenzeit und genaueren Ergebnissen wäre - ähnlich wie in einer früheren Arbeit vorgestellt [Hild 2000]- folgendermaßen denkbar: Zunächst wird bei der Segmentierung aus einer höheren Pyramidenstufe eine grobe Georeferenzierung erreicht. Mit der erhaltenen Transformation werden Referenzobjekte (Vektor- oder Rasterpolygone) in das Originalbild transformiert und zur Unterstützung der Segmentierung auf der Originalauflösung verwendet. Dabei könnte für jedes Objekt mit einem individuellen Schwellwert (aus den multiplen Durchläufen bei verringerter Auflösung) gearbeitet werden und somit auf höchster Auflösung nurmehr ein Segmentierungsdurchlauf nötig sein. Im Rahmen dieser Arbeit wird diese Strategie jedoch nicht weiter verfolgt, da hier die prinzipielle Machbarkeit des vorgeschlagenen Weges im Vordergrund stehen soll.

## 2.2 Flächenobjekte aus Vektordaten

Die Erzeugung von flächenhaften Objekten aus Vektordaten, die für das entwickelte Verfahren nutzbar sind, muß so beschaffen sein, daß die Objekte bezüglich Form und Objektart auch durch das im vorigen Abschnitt vorgestellte Segmentierungsverfahren isoliert werden können. Das bedeutet, es müssen Objektarten gewählt werden, die eine relativ homogene Grauwertverteilung im Ziel-Auflösungsbereich besitzen. Zudem müssen benachbarte Objekte gleicher Objektart, die aufgrund bestimmter Erfassungsregeln getrennt vorliegen, zunächst



**Abbildung 2.4:** Ergebnisse der Segmentierung mit adaptivem Schwellwert  $h = 8$  auf den Pyramidenstufen 0 bis 3 eines Ausschnitts der SPOT PAN-Szene Elchingen/Großkuchen. Im Hintergrund ist jeweils das vollaufgelöste Bild zu sehen.

verschmolzen werden, da sie sonst in aller Regel im Bild nicht isoliert werden können. In dieser Arbeit wird vektorseitig mit ATKIS-Daten gearbeitet [AdV 1988], daher wird im folgenden Abschnitt zunächst ein kurzer Überblick über ATKIS und das zu Grunde liegende Digitale Landschaftsmodell 1:25000 (DLM 25) gegeben.

### 2.2.1 ATKIS

Zur Vermeidung divergierender Tendenzen in der digitalen Erfassung topographischer Informationen wurde von der Arbeitsgemeinschaft der Vermessungsverwaltungen der Bundesrepublik Deutschland (AdV) im Jahr 1986 der Aufbau des Amtlichen Topographisch-Kartographischen Informationssystems (ATKIS) beschlossen. Nach Festlegung ist ATKIS ein weitestgehend redundanzfreies Basis-Informationssystem für raumbezogene Daten, in das gegebenenfalls Fachdaten integriert werden können. Durch die integrierende Funktion des ATKIS wird der Datenerfassungs- und -fortführungsaufwand, der nach Bill & Fritsch [1991] der kostenintensivste Aspekt eines GIS ist, wesentlich verringert. Die ATKIS-Daten liegen in einem projizierten Gauß-Krüger-Koordinatensystem, basierend auf dem internationalen Referenzellipsoid, vor. Zu allgemeinen transversalen Mercator-Projektionen sei auf die Arbeit von Grafarend [1995] verwiesen.

Um raumbezogene Daten digital vorhalten zu können, muß zunächst ein abstraktes Modell der Wirklichkeit angefertigt werden. Dieses Modell (Digitales Landschaftsmodell = DLM) wird zum einen aus der Landschaft selbst (Vermessung), zum anderen aber auch durch bereits existierende originäre Daten (Karten, Bilder etc.) gebildet. Als Grundlage für die Kartenerzeugung dient das Digitale Kartographische Modell (DKM), das aus

dem DLM durch Generalisierung unter Berücksichtigung kartographischer Richtlinien gewonnen wird. Da die Generalisierung mit wachsendem Maßstabsunterschied immer komplexer wird, wurde die Erfassung mehrerer DLM beschlossen (DLM 25, DLM 200 und DLM 1000). Die Modellbildung erfolgt in den verschiedenen DLM derart, daß der Generalisierungsgrad in etwa dem zugeordneten Kartenmaßstab entspricht, z.B. entspricht im DLM 25 der Generalisierungsgrad in etwa dem einer analogen Karte 1:25000 etc. Wie in Abbildung 2.5 zum ATKIS-Konzept dargestellt ist, geschieht die Primärmodellbildung in ATKIS mittels eines sogenannten Objektartenkatalogs (OK). Der OK gliedert die zu erfassenden Objekte in sieben Objektbereiche: Festpunkte, Siedlung, Verkehr, Vegetation, Gewässer, Relief und Gebiete, die wiederum in neunzehn Objektgruppen zerfallen. Die Objektgruppen zergliedern sich weiter in insgesamt 168 Objektarten. Nach [AdV 1988] ist die Objektbildung folgendermaßen definiert:

Objektart ist die zusammenfassende Bezeichnung für eine Klasse von gleichartigen topographischen Objekten. Ein Objekt selbst ist ein konkretes, geometrisch abgrenzbares, mit einem Namen benennbares Element einer solchen Klasse, z.B. Bundeshaus, (...) .

Den dergestalt hierarchisch organisierten Objekten können 111 Attributtypen mit insgesamt 664 möglichen Attributwerten zugewiesen werden. Die eigentlichen Träger der geometrischen Information sind die Objektteile, zu deren Bildung in [AdV 1988] folgende Definition getroffen wird:

Objektteil ist ein konkreter, geometrisch bergrenzter, durch einheitliche Attribute und Relationen bestimmter Teil eines Objekts.

In Abbildung 2.5 ist schematisch das Konzept von ATKIS dargestellt. Eine detaillierte Beschreibung findet sich in [AdV 1988] oder auch bei Walter [1997].

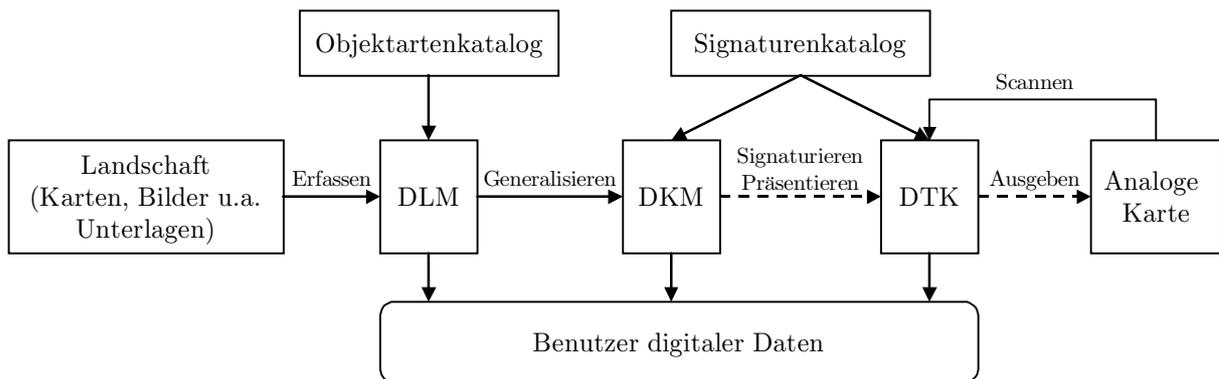


Abbildung 2.5: Konzeption von ATKIS.

ATKIS-Daten können blattschnittfrei nach benutzerspezifischen Gesichtspunkten von den Landesvermessungsämtern bezogen werden. Die Datenauszüge werden dabei mit Hilfe der Einheitlichen Datenbankschnittstelle (EDBS) ausgetauscht. Die EDBS ist eine plattformunabhängige und herstellerneutrale Schnittstelle, die den Austausch von hierarchisch aufgebauten, objektweise gespeicherten Daten ermöglicht. Neben der einfachen Abfrage ermöglicht die EDBS auch die Datenfortführung. Das deutsche ATKIS ist ein Beispiel für einen flächendeckenden GIS-Datenbestand in einem für Mitteleuropa typischen Maßstab. Ähnliche Bestrebungen gibt es mit leichten Unterschieden auch in anderen europäischen Ländern, so daß die beispielhafte Verwendung von ATKIS-Daten keine wesentliche Einschränkung für die Benutzbarkeit des entwickelten Verfahrens darstellt

## 2.2.2 Verschmelzung zusammenhängender ATKIS-Objekten

Die Modellbildung des DLM 25 (Abschnitt 2.2.1) für flächenhafte Objekte erfolgt nicht zwingend so, daß dabei Objekte entstehen, die durch eine Segmentierung aus Fernerkundungsdaten zu erhalten sind. Ein Grund hierfür sind Erfassungsregeln, die Trennungen aufgrund im Bild unsichtbarer Linien vornehmen (z.B. kleine Waldwege, Gemarkungsgrenzen etc.). Zur Vermeidung dieser unerwünschten Objekttrennungen werden mit dem hier entwickelten Verfahren in einem Vorverarbeitungsschritt räumlich zusammenhängende Objekte gleicher Objektart miteinander verschmolzen. Da diese Operation mittlerweile zum Standardumfang kommerzieller GIS-Produkte

gehört (z.B. Arc/View von ESRI Inc.), soll hier nur kurz auf das Resultat dieses Schrittes eingegangen werden. Im Beispiel in Abbildung 2.6 ist links eine ATKIS-Visualisierung des Stadtzentrums von Karlsruhe dargestellt (Flächen und Linien). Im rechten Teilbild ist das Ergebnis der Verschmelzung benachbarter Objekte gleicher Objektart dargestellt (nur Flächen).



Flächen und Linien (original)

Verschmelzung von Flächen  
gleicher Objektart

**Abbildung 2.6:** ATKIS-Daten der Innenstadt von Karlsruhe im Original (Flächen und Linienelemente, links) und nach Verschmelzung benachbarter Flächen gleicher Objektart.

Wie an der Waldfläche, die den ganzen unteren rechten Teil dominiert, sehr schön zu sehen ist, ist erst durch die Verschmelzung eine Vergleichsmöglichkeit zu Segmentierungsergebnissen aus Fernerkundungsdaten prinzipiell gegeben. Die Einzelflächen im linken Teilbild sind mit einer Bildsegmentierung nicht zu trennen.

## Kapitel 3

# Merkmalsgewinnung

”Die Invariantentheorie wurde bereits mehrere Male totgesagt, doch wie der Phoenix hat sie sich wieder und wieder aus der Asche erhoben.”

*Jean Dieudonné, 1968*

Objekterkennung mit Hilfe der Computer Vision, gleich ob in 2D oder 3D, ist immer mit der Verwendung eines oder mehrerer Objektmodelle verbunden, die in einem Vision-System implementiert sind. Im Rahmen dieser Arbeit werden Objekte mit Hilfe von Merkmalen anhand ihrer Form beschrieben und zugeordnet. Dieses Vorgehen bezeichnet man gemeinhin als Konturanalyse (”Shape Analysis”). Nach Haralick & Shapiro [1993], Kapitel 18, läßt sich die 2D-Konturanalyse entsprechend des verwendeten Objektmodells grob in fünf Kategorien unterteilen.

### Repräsentation durch globale Merkmale

In dieser Repräsentation werden Objekte binär beschrieben, d.h. ein Pixel beziehungsweise ein Raumpunkt ist entweder Teil eines Objekts oder nicht. Aus der Behandlung von Binärbildern bekannte und vielfach verwendete Merkmale sind Durchmesser, Umfang, Fläche, diverse Rundheits- und Elongationsmaße, Momente oder Fourier-Deskriptoren. Ein Verfahren zur Mustererkennung, das auf geometrischen Momenten basiert, wird von Flusser & Suk [1993] angegeben. Mit Hilfe von Fourier-Deskriptoren wurde von Arbter, Snyder & Burkhardt [1990] ein Verfahren zur 3D-affinvarianten Objekterkennung vorgeschlagen. Von Reiss [1993] wird eine Möglichkeit beschrieben, wie teilweise verdeckte Objekte mit Hilfe von Momenten erkannt und zugeordnet werden können. Geo- und Kartenobjekte, wie sie hier zugeordnet werden, weisen jedoch in der Regel keine gegenseitigen Überlappungen auf, eine Karte beziehungsweise ein Orthophoto ist ein 2D-Abbild der realen Welt.

### Repräsentation durch lokale Merkmale

Lokale Merkmale können z.B. Ecken oder Löcher sein. Zu einer Objektbeschreibung durch lokale Merkmale gehören auch die Relationen zwischen den einzelnen Merkmalen. Das gesamte Objektmodell wird in der Regel in einem Graphen abgelegt. Objekterkennung in diesem Modell bedeutet dann die Zuordnung von Graphen oder Subgraphen.

### Repräsentation durch die Kontur

Die Repräsentation von 2D-Objekten durch ihre Kontur ist am weitesten verbreitet, da sie in Bezug auf Invarianzen die größten Vorteile bietet. Mögliche Formen sind: Folge von Konturpixeln, ”Chain-code”-Repräsentation [Freeman 1961], [Abbasi-Dezfouli & Freeman 1994] oder eine Repräsentation durch Liniensegmente. Die meisten globalen Merkmale lassen sich direkt aus der Konturbeschreibung bestimmen, wodurch ein wesentlich geringerer Speicherbedarf und eine wesentliche effizientere Berechnung entsteht. Davon wird auch in dieser Arbeit Gebrauch gemacht (s. Abschnitt 3.3.1).

### Skelettrepräsentation

Mit der Skelettrepräsentation wird ein Objekt auf seine Mittelachsen zurückgeführt. Die Mittelachsen zeichnen sich aus als Orte mit gleichem Abstand zu den nächsten Kanten. Für langgestreckte oder eher linienhafte Objekte kann es günstiger für ein Zuordnungsverfahren sein, auf die skelettierte Darstellung eines Objekts aufzusetzen.

### Repräsentation durch 2D-Teile

Durch die Dekomposition eines komplexen Objekts in primitivere 2D-Teile gelangt man zu einer strukturellen Beschreibung des Objekts in Form eines Graphen, ähnlich der Repräsentation durch lokale Merkmale. Ein wesentlicher Vorteil dieser Repräsentation liegt in der Unempfindlichkeit gegenüber geometrischer Transformation oder Rauschen.

In diesem Kapitel findet sich eine Darstellung der Theorie der invarianten Merkmale unter geometrischer Transformation. Entsprechend dem Anspruch der vorliegenden Arbeit, ein System zur Bild-Karte-Registrierung aufzubauen, liegt der Schwerpunkt dabei auf Invarianten unter affiner Transformation. Wie in diesem Kapitel dargestellt wird, lassen sich Invarianten unter affiner Transformation mit Hilfe von Momenten aus der Kontur von Polygonen berechnen. Flächenhafte GIS-Objekte werden oft als Polygone abgelegt, insofern liegt die Verwendung dieser Art von Invarianten nahe. Die Berechnung von Invarianten unter perspektiver Transformation ist prinzipiell schwieriger, da die perspektiven Transformationen keine abgeschlossene Gruppe bilden [Lenz 1990]. Bekannte perspektive Invarianten sind etwa das Strecken-Doppelverhältnis und das Flächen-Doppelverhältnis ([Grafarend & Shan 1997a], [Grafarend & Shan 1997b]), mit denen jedoch eine Objekterkennung für nicht exakt gleiche Objekte schwierig zu formulieren ist. Zudem lassen sich nicht ohne weiteres perspektive Invarianten aus

Momenten berechnen [Reiss 1993]. Es gibt jedoch Ansätze, wie 3D-Momente von Polyedern effizient aus der Oberfläche berechnet werden können [Bing-Cheng 1993] oder Verfahren, in denen unter der Beschränkung auf 3D-Ähnlichkeitstransformationen 3D-Momentinvarianten zur Objekterkennung und -zuordnung benutzt werden. Die Benutzung perspektiver Invarianten wird jedoch aus den genannten Gründen hier nicht weiter verfolgt.

Zunächst wird im folgenden Abschnitt eine Einführung in die Theorie der algebraischen Invarianten gegeben. Die algebraischen Invarianten werden dann mit Hilfe des Theorems für Momentinvarianten mit den zentralen Momenten der gegebenen Objekte verknüpft. Zur praktischen Durchführung der Algorithmen wird ein Verfahren vorgestellt, mit dem die zentralen Momente direkt aus der Objektkontur berechnet werden können. Momentinvarianten haben für die Mustererkennung eine Bedeutung, wenn sie es erlauben, Objekte hinreichend gut zu trennen und wenn sie robust gegen Störungen des Objekts sind. Störungen können allgemein geometrischer (Form des Objekts) oder radiometrischer Natur (Grauwertverteilung) sein. Für diese Arbeit werden lediglich geometrische Aspekte betrachtet, da Objekte in GIS-Datenbeständen in der Regel ohne Textur vorliegen. Die klassischen Invarianten aus geometrischen Momenten erweisen sich jedoch als relativ anfällig gegen Störungen, was einer Verwendung für Objekterkennungszwecke erschwerend im Wege steht. Zur Erhöhung der Robustheit der Merkmale gegen Rauschen jeder Art werden die sogenannte "Cross-weighted moments" (CWMs) eingeführt. Anhand von Beispielen werden die Auswirkungen verschiedener Arten des Rauschens untersucht sowie Betrachtungen zur Diskriminierbarkeit angestellt. Abschließend wird noch ein kurzer Überblick über alternative Merkmale gegeben.

## 3.1 Invarianten algebraischer Formen und Momentinvarianten

Invariante Merkmale von Bildern oder Objekten unter affiner oder projektiver Transformation haben ihren Ursprung in den Invarianten algebraischer Formen, die im folgenden Abschnitt 3.1.1 eingeführt werden. Die Wurzeln der Theorie der algebraischen Formen reichen in die Mitte des 19. Jahrhunderts zurück. Die Gewinnung von invarianten Merkmalen aus geometrischen Momenten ist mit den Invarianten algebraischer Formen über das Fundamentaltheorem der Momentinvarianten verknüpft, das in Abschnitt 3.1.3 in der revidierten Fassung nach Reiss [1991] behandelt wird. Die ursprüngliche, aber nicht ganz korrekte Form des Fundamentaltheorems für Momentinvarianten wurde von Hu [1962] gegeben, das Fundamentaltheorem algebraischer Formen wurde von Sylvester [1878] eingeführt.

Das Studium der algebraischen Invarianten wurde Ende des 19. Jahrhunderts von Cayley [1856], Sylvester [1878] und Salmon [1885], sowie von Clebsch [1872] vorangetrieben. Die explizite Angabe einer größeren Anzahl algebraischer Invarianten findet sich bei Salmon [1885], der berühmte Mathematiker David Hilbert leistete abschließende Arbeiten [Hilbert 1890], [Hilbert 1893]. Die Arbeiten der genannten Wissenschaftler zur Gewinnung algebraischer Invarianten bezeichnet man zusammenfassend und rückblickend als historische Methoden. Demgegenüber stehen moderne und kompakter formulierte Tensormethoden, wie sie z.B. von Gurevich [1964] vorgestellt werden. Die Theorie in ihrer Gesamtheit ist jedoch zu umfangreich, um hier auch nur in Ansätzen dargestellt werden zu können. Daher begnügt sich die Darstellung in dieser Arbeit im wesentlichen auf die Auflistung der für das Gesamtvorhaben relevanten Sachverhalte. In Reiss [1993] findet sich eine gute Übersicht über die genannten Arbeiten, in deren Anhang sowohl die historischen wie auch die Tensormethoden zusammengefasst werden.

### 3.1.1 Algebraische Formen und Invarianten

**Definition 3** *Ein homogenes Polynom  $p$ -ten Grades von  $n$  unabhängigen Variablen bezeichnet man als (algebraische) Form  $p$ -ter Ordnung. Liegen zwei Variable vor, so spricht man von einer binären Form, im Falle dreier Variablen von einer ternären Form usw.*

Eine binäre Form  $p$ -ter Ordnung in  $u$  und  $v$  ist gegeben durch:

$$f_p(u, v) = a_{p,0}u^p + \binom{p}{1}a_{p-1,1}u^{p-1}v + \binom{p}{2}a_{p-2,2}u^{p-2}v^2 + \dots + a_{0,p}v^p \quad (3.1)$$

Gibt es nun Funktionen der Koeffizienten  $a_{i,p-i}$ , die sich unter einer linearen Koordinatentransformation nicht oder nur um einen bekannten konstanten Faktor ändern?

Zur Beantwortung trifft man die folgende

**Definition 4** Eine Funktion  $I(a_{p,0}, a_{p-1,1}, \dots, a_{0,p})$  der Koeffizienten einer algebraischen Form  $f_p(u, v, \dots)$  bezeichnet man als Invariante, wenn sie der Beziehung  $I(a'_{p,0}, a'_{p-1,1}, \dots, a'_{0,p}) = \Delta^g I(a_{p,0}, a_{p-1,1}, \dots, a_{0,p})$  genügt. Dabei bezeichnen  $a'_{p,0}, a'_{p-1,1}, \dots, a'_{0,p}$  die Koeffizienten der betreffenden Form nach Durchführung einer linearen Transformation der Koordinaten  $u, v, \dots$ . Für die Determinante  $\Delta$  des homogenen Teils der linearen Transformation muß dabei einschränkend  $\Delta \neq 0$  gelten. Die Variable  $g$  bezeichnet man als Gewicht der Invarianten. Im Fall  $g = 0$  spricht man von einer absoluten Invarianten, ansonsten von einer relativen Invarianten.

Die Koeffizienten  $(a'_{p,0}, a'_{p-1,1}, \dots, a'_{0,p})$  ergeben sich, wenn man in Gleichung 3.1 die Koordinaten  $(u, v)$  mittels der linearen Transformation

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix} \quad \text{mit} \quad \Delta = \begin{vmatrix} \alpha & \gamma \\ \beta & \delta \end{vmatrix} \neq 0 \quad (3.2)$$

ersetzt. Schreibt man die umgekehrte Transformation in folgender Weise:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad (3.3)$$

so gilt die Invarianzbeziehung

$$ux + vy = u'x' + v'y'. \quad (3.4)$$

Diese Invarianzbeziehung spielt bei der Ableitung des Fundamentaltheorems (Abschnitt 3.1.3) eine grundlegende Rolle.

Die Zusammenhänge werden veranschaulicht mit folgendem

**Beispiel 5** Eine binäre algebraische Form zweiter Ordnung ist gegeben durch  $f_2(x, y) = ax^2 + bxy + cy^2$ . Eine lineare Koordinatentransformation ist gegeben in Form einer homogenen affinen Transformation  $x = \alpha x' + \beta y'$  und  $y = \gamma x' + \delta y'$ .

Setzt man diese Beziehungen in die gegebene Form ein, so erhält man aus einem Koeffizientenvergleich die folgenden Beziehungen zwischen den Koeffizienten  $(a, b, c)$  der untransformierten und  $(a', b', c')$  der transformierten gegebenen binären Form:

$$a' = a\alpha^2 + 2b\alpha\gamma + c\gamma^2; \quad b' = a\alpha\beta + b(\alpha\delta + \gamma\beta) + c\beta\gamma; \quad c' = a\beta^2 + 2b\beta\delta + c\delta^2$$

Berechnet man damit den Ausdruck  $Q' = a'c' - b'^2$ , so findet man  $Q' = \Delta^2(ac - b^2)$  mit  $\Delta = (\alpha\delta - \beta\gamma)$ , also genau die Determinante der homogenen affinen Transformation. Die Funktion  $Q(a, b, c)$ , die sich mit der gegebenen Transformation zu  $Q(a', b', c')$  transformiert, genügt der Beziehung  $Q' = \Delta^2 Q$  und ist damit eine relative Invariante der gegebenen binären Form mit Gewicht 2.

In dem eben gegebenen Beispiel liefert der Koeffizientenvergleich drei lineare Gleichungen, die die transformierten und untransformierten Koeffizienten in Beziehung zueinander setzen. Allgemein bestehen bei einer Form von  $n$  Variablen  $n + 1$  solcher linearer Gleichungen. Eine lineare, homogene Koordinatentransformation besitzt 4 unabhängige Parameter, so daß man für Formen von mehr als drei Variablen  $(p + 1) - 4 = p - 3$  abhängige Koeffizienten hat und damit auch  $(p - 3)$  absolute Invarianten finden kann [Salmon 1885].

Neben Invarianten für eine bestimmte Form lassen sich auch Invarianten für beliebige Systeme von Formen bestimmen, z.B. für lineare und quadratische binäre Formen gleichzeitig oder für lineare, quadratische und kubische binäre Formen gleichzeitig. Ohne weitere Herleitung werden im folgenden die vier einfachsten Invarianten binärer algebraischer Formen gegeben, da sie für die weiteren Untersuchungen herangezogen werden. Eine Übersicht über weitere, kompliziertere Invarianten findet sich in Reiss [1993] und insbesondere in Salmon [1885].

Die verwendeten binären Formen lauten wie folgt:

$$\begin{array}{ll} f_2(x, y) = ax^2 + bx + c & \text{quadratische binäre Form} \\ f_3(x, y) = \alpha x^3 + \beta x^2 + \gamma x + \delta & \text{kubische binäre Form} \end{array}$$

Die vier einfachsten Invarianten binärer Formen sind in Tabelle 3.1 aufgeführt. Bereits bekannt aus Beispiel 5 ist die Invariante  $I_1$  einer binären quadratischen Form.

binäre Form(en)	$g$	$k$	Invariante $I$
quadratisch	2	2	$I_1 = ac - b^2$
kubisch	6	4	$I_2 = (\alpha\delta - \beta\gamma)^2 - 4(\alpha\gamma - \beta^2)(\beta\delta - \gamma^2)$
kubisch+quadratisch	4	3	$I_3 = a(\beta\delta - \gamma^2) - b(\alpha\delta - \beta\gamma) + c(\alpha\gamma - \beta^2)b$
kubisch+quadratisch	6	5	$I_4 = \alpha^2c^3 - 6\alpha\beta bc^2 + 6\alpha\gamma c(2b^2 - ac) + 6\beta\delta a(2b^2 - ac) + \alpha\delta(6abc - 8b^3) + 9\beta^2ac^2 - 18\beta\gamma abc + 9\gamma^2a^2c - 6\gamma\delta a^2b + \delta^2a^3$

**Tabelle 3.1:** Die vier einfachsten Invarianten binärer algebraischer Formen mit Gewicht  $g$  und Ordnung  $k$ .

An dieser Stelle sei deutlich festgehalten:

*Invarianten binärer algebraischer Formen sind homogene Polynome der Koeffizienten der betreffenden Formen.*

Mit der Definition 4 ist es immer möglich, aus zwei relativen Invarianten durch Division entsprechender Potenzen relativer Invarianten die Terme  $\Delta^g$  zu eliminieren und damit zu absoluten Invarianten zu gelangen. Dabei geht man folgendermaßen vor: Es liegen zwei algebraische Invarianten  $I_1$  und  $I_2$  mit den Gewichten  $g_1$  und  $g_2$  vor. Nach Definition 4 existieren dann die relativen Invarianten  $I'_1 = \Delta^{g_1} I_1$  und  $I'_2 = \Delta^{g_2} I_2$ . Zur Bildung einer absoluten Invarianten  $\Psi$  dividiert man die  $g_2$ -te Potenz der ersten Invariante durch die  $g_1$ -te Potenz der zweiten. Man erhält dann:

$$\Psi' = \frac{(I'_1)^{g_2}}{(I'_2)^{g_1}} = \frac{(\Delta^{g_1} I_1)^{g_2}}{(\Delta^{g_2} I_2)^{g_1}} = \frac{I_1 \Delta^{g_1 g_2}}{I_2 \Delta^{g_2 g_1}} = \frac{I_1}{I_2} = \Psi \tag{3.5}$$

Zusammenfassend läßt sich folgende für absolute Invarianten binärer Formen wichtige Aussage treffen:

*Absolute Invarianten binärer Formen sind Quotienten homogener Funktionen der Koeffizienten der betreffenden Formen.*

Die Benutzung binärer Formen führt zu Invarianten im  $\mathbb{R}^2$ . Um zu Invarianten in drei Dimensionen zu gelangen müssen ternäre algebraische Formen betrachtet werden.

In Abschnitt 3.1.3 wird der Zusammenhang zwischen den Invarianten algebraischer Formen und den geometrischen Momenten gegeben, der durch das Fundamentaltheorem der Momentinvarianten hergestellt wird. Dazu werden zunächst im folgenden Abschnitt 3.1.2 die geometrischen Momente eingeführt und einfache Invarianten aufgezeigt.

### 3.1.2 Geometrische Momente und einfache Invarianten

Momente finden in vielen Disziplinen Anwendung, z.B. in der Physik (Schwerpunkt, Trägheitsmomente) oder in der Statistik (Mittelwert, Streuung). Von Hu [1962] wurden Momente schon früh in der digitalen Bildverarbeitung und Objekterkennung eingesetzt. Heute finden sich Bildverarbeitungslösungen, die auf Momenten beruhen, in vielen Bereichen, z.B. in der industriellen Bildverarbeitung.

**Definition 6** Eine Funktion  $f(x, y)$  sei stückweise stetig und lediglich innerhalb eines beschränkten Bereichs  $B$  des  $\mathbb{R}^2$  von null verschieden. Für zwei nicht negative natürliche Zahlen  $p$  und  $q$  definiert man das Riemann-Integral

$$M_{p,q} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy = \iint_B x^p y^q f(x, y) dx dy \tag{3.6}$$

als Moment  $(p + q)$ -ter Ordnung der Funktion  $f(x, y)$ .

Es seien  $(\bar{x}, \bar{y})$  die Schwerpunktskoordinaten des Bereichs  $B$ . Als zentrales Moment  $(p+q)$ -ter Ordnung bezeichnet man

$$m_{p,q} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy = \iint_B (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy . \tag{3.7}$$

Im Falle diskreter Koordinaten gilt analog für die Funktion  $g(c, r)$  :

$$M_{p,q} = \sum_{c=-\infty}^{+\infty} \sum_{r=-\infty}^{+\infty} c^p r^q g(c, r) = \sum_{(c,r) \in B} c^p r^q g(c, r) \tag{3.8}$$

bzw.

$$m_{p,q} = \sum_{c=-\infty}^{+\infty} \sum_{r=-\infty}^{+\infty} (c - \bar{c})^p (r - \bar{r})^q g(c, r) = \sum_{(c,r) \in B} (c - \bar{c})^p (r - \bar{r})^q g(c, r). \quad (3.9)$$

Dabei wird die Spalte des aktuellen Rasterelements durch  $c$  angegeben, die Zeile durch  $r$ .

Zwischen den zentralen und den regulären Momenten besteht folgender Zusammenhang:

$$m_{pq} = \sum_{k=0}^p \sum_{l=0}^q \binom{p}{k} \binom{q}{l} (-1)^{k+l} M_{p-k, q-l} M_{10}^k M_{01}^l M_{00}^{-(k+l)} \quad (3.10)$$

Liegen keine baryzentrischen Koordinaten vor, so kann dieser Zusammenhang zur effizienteren Berechnung der zentralen Momente benutzt werden. Es läßt sich beweisen [Hu 1962], daß für Funktionen mit den in Definition 6 beschriebenen Eigenschaften die Momente für alle zulässigen Werte von  $p$  und  $q$  existieren (Vollständigkeit) und daß durch die Gesamtheit der Momente die Funktion  $f(x, y)$  eindeutig beschrieben ist (Eindeutigkeit). Damit läßt sich aus allen Momenten auch  $f(x, y)$  darstellen. Diese Tatsache wird von Teh & Chin [1988] zur Objektrekonstruktion benutzt. Durch die in Definition 6 definierten Momente wird damit die beschränkte, kontinuierliche Funktion  $f(x, y)$  auf die infinite, diskrete Momentenmatrix  $M$  mit den Elementen  $M_{pq}$  bzw.  $m_{pq}$  im schwerpunktreduzierten Fall, abgebildet. Der Funktion  $f(x, y)$  beziehungsweise  $g(c, r)$  in Definition 6 kann in der Bildverarbeitung, Objekterkennung und -rekonstruktion unterschiedliche Bedeutung zukommen.

Handelt es sich um ein Grauwertbild, wie in Abbildung 3.1, linkes Teilbild, so repräsentiert  $g(c, r)$  dessen diskrete Grauwertverteilung. Die berechneten Momente beziehen sich dann auf das gesamte Bild. Sie werden auch als *radiometrische Momente* bezeichnet. Derartige Momente dienen z.B. der Objekterkennung bei gleichbleibender Aufnahmesituation (z.B. Teile auf einem Fließband). Für Zuordnungen zwischen Bild- und GIS-Objekten, wie sie in dieser Arbeit vorliegen, sind radiometrische Momente aus zwei Gründen nicht geeignet. Zum einen liegen GIS-Objekte in der Regel noch ohne radiometrische Information vor (eine Ausnahme bilden hier die sich zunehmend verbreitenden virtuellen Stadtmodelle, bei denen GIS-Objekte zusammen mit ihrer Textur abgelegt sind). Zum anderen müssten, etwa im Falle einer Satellitenaufnahme und eines georeferenzierten Orthophotos, die Bildausschnitte identisch oder nahezu identisch sein. Dies stellt jedoch keinen praxisrelevanten Fall dar.

Handelt es sich um einzelne Objekte in Grauwertbildern, die durch eine vorangegangene Objektextraktion bereits in ihrer Form vorliegen, so wählt man objektweise  $g(c, r) = 1$  innerhalb des betreffenden Objekts und  $g(c, r) = 0$  außerhalb. In diesem Fall nennt man  $g(c, r)$  auch *Indikatorfunktion* des Objekts. Das jeweilige Objekt ist entsprechend Abbildung 3.1, rechtes Teilbild, als Bereich  $B$  im Bildraum repräsentiert. Die so berechneten Momente bezeichnet man auch als *geometrische Momente*, da sie infolge der binären Indikatorfunktion lediglich von der Form des Objekts abhängen.

Liegen Vektorobjekte in reellwertigen Koordinaten vor (Polygone, Abbildung 3.1, mittleres Teilbild), wählt man ebenfalls objektweise die Indikatorfunktion  $f(x, y) = 1$  innerhalb und  $f(x, y) = 0$  außerhalb des Objekts. In diesem Fall wird das Objekt durch den Bereich  $B$  im  $\mathbb{R}^2$  repräsentiert. Auch hier spricht man von geometrischen Momenten. Da die betrachteten Objekte beschränkt sind und da das Eindeutigkeitstheorem existiert, lassen sich Objekte aus Momenten rekonstruieren. Für die Rekonstruktion von Objekten ist es allerdings sinnvoll, ein orthogonales Funktionensystem zu wählen (s.u.). Geometrische Momente sind untereinander abhängig und stellen damit keine optimale Basis zur Objektrekonstruktion dar. Geometrische Momente niederer Ordnung bzw. Kombinationen aus ihnen lassen sich physikalisch interpretieren. Dazu versteht man die bereits beschriebene binäre Indikatorfunktion als Flächendichte. Das statistische Analogon ist die Wahrscheinlichkeitsdichte. Mit der Flächendichte 1 ergeben sich folgende Interpretationen:

Moment(e)	Physikalische Interpretation
$M_{0,0}$	Gesamtfläche des Objektes
$M_{1,0}$ und $M_{0,1}$	statische Momente bzgl. y- und x-Achse
$M_{1,0}/M_{0,0}$	x-Koordinate des Flächenschwerpunkts
$M_{0,1}/M_{0,0}$	y-Koordinate des Flächenschwerpunkts
$M_{2,0}$ und $M_{0,2}$	Trägheitsmomente bzgl. y-Achse und x-Achse
$M_{2,0} + M_{0,2}$	polares Trägheitsmoment des Objektes bzgl. Koordinatenursprung

**Tabelle 3.2:** Physikalische Interpretation geometrischer Momente.

Aus den Momenten zweiter Ordnung läßt sich eine objektfeste Trägheitsellipse bilden, die zur genäherten Lagebeschreibung eines Objektes benutzt werden kann. Ein Verfahren dazu wird von Süße [1999] gegeben.

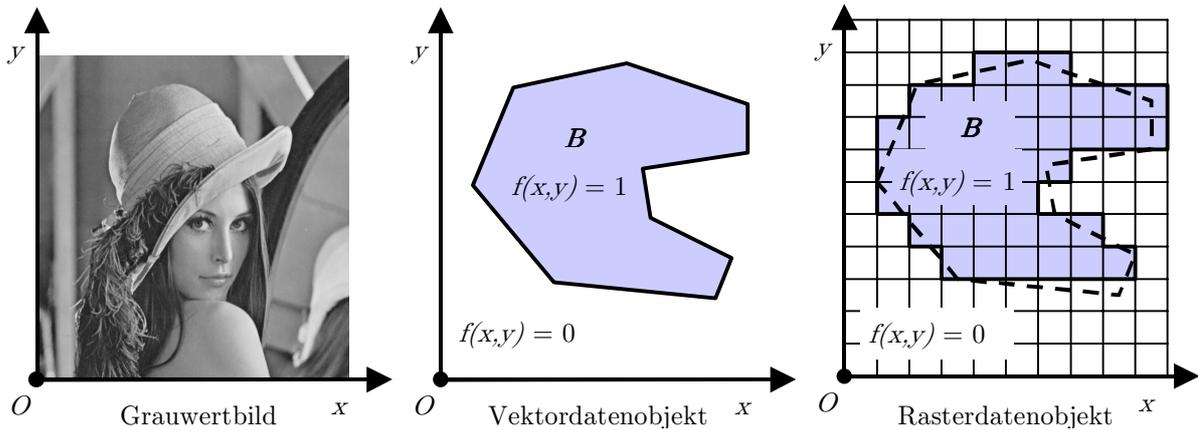


Abbildung 3.1: Verschiedene Daten, für die Momente berechnet werden können.

Im Hinblick auf das Ziel der vorliegenden Arbeit läßt sich folgende wichtige Aussage treffen:

*Geometrische Momente von flächenhaften Objekten lassen sich sowohl in Raster- wie auch in Vektorkoordinaten berechnen. Die Berechnung erfolgt über die objektweise Definition einer binären Indikatorfunktion. Damit beschreiben die geometrischen Momente die Form eines Objekts. Geometrische Momente erfüllen somit die Grundvoraussetzungen, um für ein merkmalsbasiertes Zuordnungsverfahren zwischen Raster- und Vektordatenobjekten in Frage zu kommen.*

Die zentralen Momente sind Invarianten unter Translation in  $x$  und  $y$ , da sich der Objektschwerpunkt unter Translation gleich wie die einzelnen Objektpunkte transformiert. Es sei eine beliebige Translation gegeben durch  $x' = x + a$  und  $y' = y + b$ . Der Objektschwerpunkt  $(\bar{x}, \bar{y})$  transformiert sich in gleicher Weise zu:  $\bar{x}' = \bar{x} + a$  und  $\bar{y}' = \bar{y} + b$ . Somit gilt für die zentralen Momente:

$$\begin{aligned}
 m'_{pq} &= \iint_B (x' - \bar{x}')^p (y' - \bar{y}')^q f'(x', y') dx' dy' = \iint_B (x + a - (\bar{x} + a))^p (y + b - (\bar{y} + b))^q f(x, y) dx dy \\
 &= \iint_B (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy = m_{pq} \quad (3.11)
 \end{aligned}$$

Neben den in Definition 6 eingeführten Momenten gibt es noch eine Reihe weiterer. Von Abu-Mostafa & Psaltis [1984] wurden komplexe Momente angegeben, mit denen auf einfache Art relative und absolute Invarianten unter Rotation erzeugt werden können. Eine weitere Klasse bilden die Rotationsmomente, wie sie von Teh & Chin [1988] angegeben werden. Allen bisher genannten Momentensystemen ist gemeinsam, daß sie jeweils nicht orthogonal sind, d.h. es bestehen Abhängigkeiten zwischen den einzelnen Momenten. Legendre-Polynome, die auch aus der Physik zur Bildung von Elektron-Wellenfunktionen bekannt sind [Courant & Hilbert 1953], stellen einen orthogonalen Funktionensatz auf dem Intervall  $[-1, 1]$  dar und erlauben die Formulierung von orthogonalen Legendre-Momenten. Legendre-Momente bestimmter Ordnung lassen sich durch geometrische Momente gleicher und kleinerer Ordnung ausdrücken. Ebenso bilden die komplexen Zernike-Momente, die aus den auf der Einheitskreis  $x^2 + y^2 \leq 1$  orthogonalen, komplexwertigen Zernike-Polynomen aufgebaut werden, einen Satz orthogonaler Momente. Auch sie lassen sich aus geometrischen Momenten aufbauen. Eine Darstellung eines kompletten Satzes findet sich bei Wallin & Kübler [1995]. Von Khotanzad & Hong [1990] wurde ein Verfahren zur invarianten Mustererkennung basierend auf Zernike-Momenten vorgeschlagen. Aus ähnlichen, jedoch reellwertigen Polynomen lassen sich die Pseudo-Zernike-Momente berechnen. Teh & Chin [1988] führten umfangreiche Untersuchungen zu Informationsgehalt, Rekonstruierbarkeit und Rauschverhalten der verschiedenen Sätze von Invarianten durch. Da in dieser Arbeit keine Objektrekonstruktion durchgeführt wird, ist die Forderung nach Orthogonalität nicht wesentlich. Im folgenden werden daher ausschliesslich geometrische Momente betrachtet, da deren Berechnung am einfachsten ist.

### 3.1.3 Das Fundamentaltheorem der Momentinvarianten

Von Hu [1962] wurde erstmals der Zusammenhang zwischen den im vorigen Abschnitt vorgestellten Invarianten algebraischer Formen und den geometrischen Momenten von Objekten (bzw. Momenten von Grauwertbildern)

hergeleitet. In der Ableitung ist jedoch ein Fehler enthalten, der mit der von Reiss [1991] gegebenen Herleitung korrigiert wird. Nach dessen Herleitung lautet das Fundamentaltheorem der Momentinvarianten folgendermaßen:

**Theorem 7** *Es sei  $|\Delta|$  der Betrag der Determinante einer linearen Transformation. Besitzt die binäre algebraische Form der Ordnung  $p$  eine algebraische Invariante  $I(a'_{p,0}, a'_{p-1,1}, \dots, a'_{0,p})$  mit Gewicht  $g$  und Ordnung  $k$ , also z.B.  $I(a'_{p,0}, a'_{p-1,1}, \dots, a'_{0,p}) = \Delta^g I(a_{p,0}, a_{p-1,1}, \dots, a_{0,p})$ , dann besitzen die zentralen Momente  $\{m_{p,0}, m_{p-1,1}, \dots, m_{0,p}\}$  der Ordnung  $p$  dieselbe Invariante, jedoch mit einem zusätzlichen Faktor  $|\Delta|^k$  versehen:*

$$I(m'_{p,0}, m'_{p-1,1}, \dots, m'_{0,p}) = \Delta^g |\Delta|^k I(m_{p,0}, m_{p-1,1}, \dots, m_{0,p}). \quad (3.12)$$

Dieser Satz gilt auch für algebraische Invarianten, die Koeffizienten zweier oder mehrerer algebraischer Invarianten unterschiedlicher Ordnung enthalten.<sup>1, 2</sup>

Für die detaillierte Herleitung des Fundamentaltheorems sei auf die Originalliteratur [Hu 1962], [Reiss 1991] verwiesen.

### 3.1.4 Momentinvarianten unter allgemeiner Affintransformation

Zur Ableitung absoluter Momentinvarianten aus zentralen geometrischen Momenten  $m_{pq}$  betrachtet man zunächst das Fundamentaltheorem aus Gleichung 3.12 für das Moment nullter Ordnung, wobei zu beachten ist, daß die Invariante für die nullte Ordnung gerade das zentrale geometrische Moment nullter Ordnung selbst ist. Vereinfacht man  $m_{0,0} = m$ , so gilt die Beziehung:

$$m' = |J| m \quad \text{bzw.} \quad m' = |\Delta| m. \quad (3.13)$$

Wie bei der Bildung von absoluten algebraischen Invarianten (vgl. Gleichung 3.5) bildet man zur Gewinnung absoluter Momentinvarianten den Quotient aus zwei relativen Momentinvarianten unter Berücksichtigung von Gleichung 3.12.

Abhängig davon, ob das Gewicht  $g$  einer relativen Momentinvarianten gerade oder ungerade ist, ergeben sich zwei Fälle.

#### Fall 1 $g$ ist gerade

In diesem Fall gilt  $\Delta^g = |\Delta|^g$ . Die erste relative Invariante sei  $I$  mit Gewicht  $g$  und Ordnung  $k$ . Als zweite relative Invariante benutzt man  $m^{g+k}$ . Potenziert man Gleichung 3.13 mit  $(g+k)$ , so zeigt sich, daß  $m^{g+k}$  eine relative Invariante der Ordnung  $g+k$  mit Gewicht  $g+k$  ist:

$$m'^{g+k} = |\Delta|^{g+k} m^{g+k}. \quad (3.14)$$

Die dem Vorgehen von Gleichung 3.5 äquivalente Quotientenbildung führt zu folgender absoluter Invariante:

$$\Psi' = \frac{I'}{m'^{g+k}} = \frac{|\Delta|^{g+k} I}{|\Delta|^{g+k} m^{g+k}} = \frac{I}{m^{g+k}} = \Psi. \quad (3.15)$$

#### Fall 2 $g$ ist ungerade

In diesem Fall gilt  $\Delta^g = \Delta |\Delta|^{g-1}$  damit für  $\Psi$ :

$$\Psi' = \frac{I'}{m'^{g+k}} = \frac{\Delta |\Delta|^{g-1} I}{|\Delta|^{g+k} m^{g+k}} = \frac{\Delta}{|\Delta|} \frac{I}{m^{g+k}} = \text{sign}(\Delta) \frac{I}{m^{g+k}}. \quad (3.16)$$

Damit ist  $\Psi$  eine sogenannte Spiegelinvariante, die das Vorzeichen abhängig von  $\Delta$  wechselt.

Aus den vier einfachsten, in Tabelle 3.1 angegebenen Invarianten algebraischer Formen lassen sich in der eben dargestellten Weise mit Hilfe des Fundamentaltheorems für Momentinvarianten die entsprechenden vier absoluten Momentinvarianten ableiten. Da die Gewichte aller vier in Tabelle 3.1 angegebenen Invarianten gerade sind,

<sup>1</sup>In der Originalversion von [Hu 1962] ist das Fundamentaltheorem in einschränkender Weise mit  $k \equiv 1$  formuliert.

<sup>2</sup>Der Term  $|\Delta|^k$  lautet eigentlich  $|J|^k$ , wobei  $J$  die Jacobi-Determinante der zu Grunde liegenden Koordinatentransformation ist (siehe Gl.3.3). Da diese Transformation jedoch linear ist, gilt  $J = \Delta$  (siehe Gleichung 3.2).

tritt hier nur Fall 1 in Kraft. Führt man das Fundamentaltheorem aus, so ergeben sich die folgenden absoluten Momentinvarianten:

$$\text{aus } I_1 \text{ mit } k = 2 \text{ und } g = 2 : \quad \Psi_1 = \frac{I_1}{m^4}, \quad (3.17a)$$

$$\text{aus } I_2 \text{ mit } k = 4 \text{ und } g = 6 : \quad \Psi_2 = \frac{I_2}{m^{10}}, \quad (3.17b)$$

$$\text{aus } I_3 \text{ mit } k = 3 \text{ und } g = 4 : \quad \Psi_3 = \frac{I_3}{m^7}, \quad (3.17c)$$

$$\text{aus } I_4 \text{ mit } k = 5 \text{ und } g = 6 : \quad \Psi_4 = \frac{I_4}{m^{11}}, \quad (3.17d)$$

wobei wiederum abkürzend  $m = m_{0,0}$  gesetzt ist.

Für absolute Momentinvarianten unter allgemeiner affiner Transformation läßt sich insgesamt folgende wichtige Aussage treffen:

*Absolute Momentinvarianten unter allgemeiner affiner Transformation sind Quotienten eines homogenen Polynoms zentraler Momente und einer entsprechenden Potenz des zentralen Moments nullter Ordnung. Das homogene Polynom hat dieselbe Gestalt wie die Invariante der entsprechenden algebraischen Form(en), anstelle der Koeffizienten der Form(en) stehen zentrale Momente. Die Potenz des zentralen Moments nullter Ordnung ist gleich der Summe von Ordnung und Gewicht der zu Grunde liegenden algebraischen Invarianten.*

**Beispiel 8** Als Beispiel diene die algebraische Invariante  $I_2 = (\alpha\delta - \beta\gamma)^2 - 4(\alpha\gamma - \beta^2)(\beta\delta - \gamma^2)$  der kubischen binären Form  $f_3(x, y) = \alpha x^3 + \beta x^2 + \gamma x + \delta$ . Die Invariante  $I_2$  ist von 4. Ordnung und besitzt das Gewicht 6. Damit lautet die absolute Momentinvariante  $\Psi_2$ :

$$\Psi_2 = \frac{(m_{3,0}m_{0,3} - m_{2,1}m_{1,2})^2 - 4(m_{3,0}m_{1,2} - m_{2,1}^2)(m_{2,1}m_{0,3} - m_{1,2}^2)}{m_{0,0}^{10}}. \quad (3.18)$$

### 3.1.5 Robustheit gegen Rauschen

Die bisher angestellten Betrachtungen galten für den Fall exakter Koordinaten und sind daher zunächst auch nur für diesen Idealfall gültig. In der Praxis treten jedoch immer statistische oder systematische Fehler in den Koordinaten auf, deren Einfluß auf die geometrischen Momente und die daraus gebildeten Invarianten (Gleichungen 3.17a bis 3.17d) nun beleuchtet werden soll. Dazu werden zwei Fälle von Koordinatenfehlern untersucht. Den ersten Fall stellen statistische Fehler dar, wie sie für Vektordaten z.B. durch die Erfassungsungenauigkeit oder durch Datenkonversion mit verringerter Genauigkeit vorliegen können. Fehler systematischer Natur kommen im zweiten behandelten Fall durch das sogenannte Diskretisierungsrauschen ins Spiel. Da in dieser Arbeit geometrische Objekte mittels invarianter Merkmale von Vektor- und Rasterdaten zugeordnet werden, ist dieser Aspekt hier von Bedeutung.

In der Literatur wurde bereits ausführlich der Einfluß von statistischem Rauschen auf geometrische Momente und Momentinvarianten von Grauwertbildern untersucht. Abu-Mostafa & Psaltis [1984] leiten in Betrachtungen zur Robustheit unter weißem Rauschen ab, daß sich das Signal-Rausch-Verhältnis zentraler Momente  $m_{pq}$  proportional zu  $1/\sqrt{p+q}$  verhält. Von Teh & Chin [1988] werden Signal-Rausch-Verhältnisse für geometrische, komplexe, Legendre-, Zernike- und Pseudo-Zernike-Momente angegeben. Übereinstimmend läßt sich für alle Arten von Momenten sagen, daß die Anfälligkeit für statistisches Rauschen mit der Ordnung der Momente zunimmt. Von daher ist eine obere Grenze für die Verwendbarkeit von geometrischen Momenten zur Objekterkennung und -rekonstruktion gegeben. Auf der anderen Seite tragen gerade die Momente höherer Ordnung die Information über feinere Details, die zur Diskriminierung ähnlicher Objekte oder zur exakten Rekonstruktion nötig ist. Je nach Stärke des auftretenden Rauschens gibt es folglich eine optimale höchste Ordnung. Die hier gewählten Ordnungen sind jedoch vorrangig aus rechentechnischen Gründen festgelegt, es bietet sich eine weiterführende Untersuchung zur Optimierung an.

Im Gegensatz zu den erwähnten Arbeiten, die von einem verrauschten Integranden (z.B. Grauwertfunktion eines Bildes) ausgehen, liegt bei geometrischen Momenten die Indikatorfunktion streng fest (s.o.). Vielmehr ist im vorliegenden Fall der Integrationsbereich, also die Randkurve eines Objekts, mit Rauschen behaftet. Eine Abschätzung des Verhaltens der Invarianten ließe sich mittels des Fehlerfortpflanzungsgesetzes ableiten.

Als realistisches Untersuchungsbeispiel werden hier die Ländergrenzen der BRD angeführt, die in Abbildung 3.2 dargestellt sind. Die Polygone in Abbildung 3.2 liegen in geographischen Koordinaten vor.

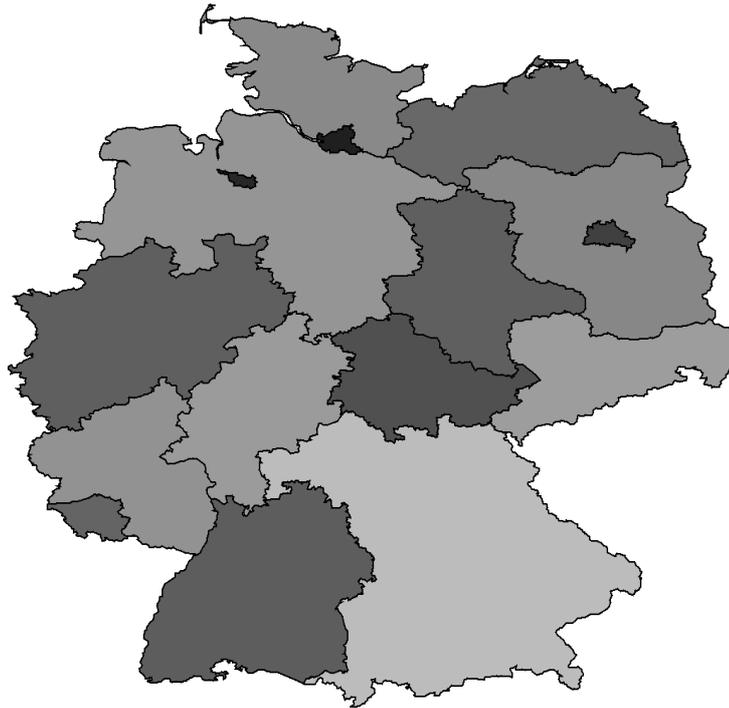


Abbildung 3.2: Die Länder der BRD als Beispiel für Vektorpolygone.

### 3.1.6 Statistisches Rauschen und Momentinvarianten

Zur Untersuchung des Einflusses von weißem Rauschen auf die Momentinvarianten  $\Psi_1$  bis  $\Psi_4$  (siehe Gleichungen 3.17a bis 3.17d) wurde an den Polygonecken statistisches Rauschen mit Standardabweichungen zwischen  $\sigma_{\min} = 0^\circ$  und  $\sigma_{\max} = 1/100^\circ$  (etwa 1000m) angebracht. Dieser Wert liegt mit Sicherheit höher als die Unsicherheit in Fernerkundungsaufnahmen. Die Auswirkungen des zugefügten Rauschens sind in Abbildung 3.3 am Beispiel der Grenze des Landes Berlin aufgezeigt.

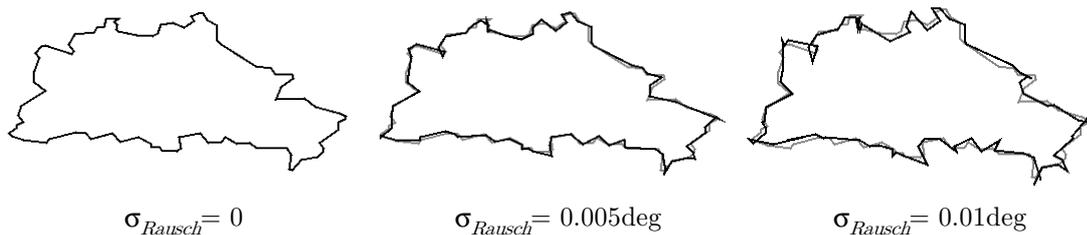
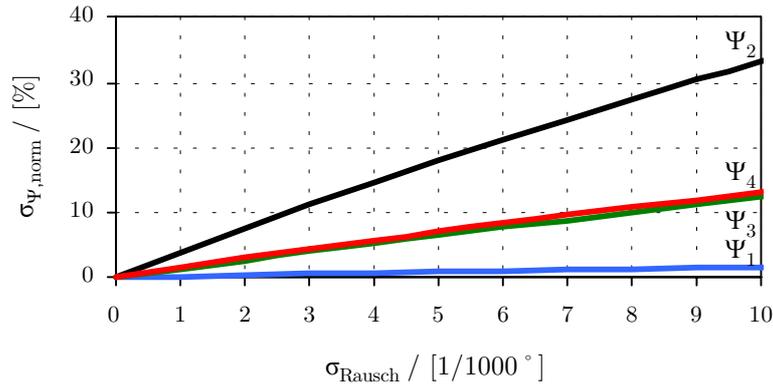


Abbildung 3.3: Statistisches Rauschen am Beispiel der Grenze des Bundeslandes Berlin. Das ursprüngliche Polygon ist hell, das verrauschte dunkel gezeichnet.

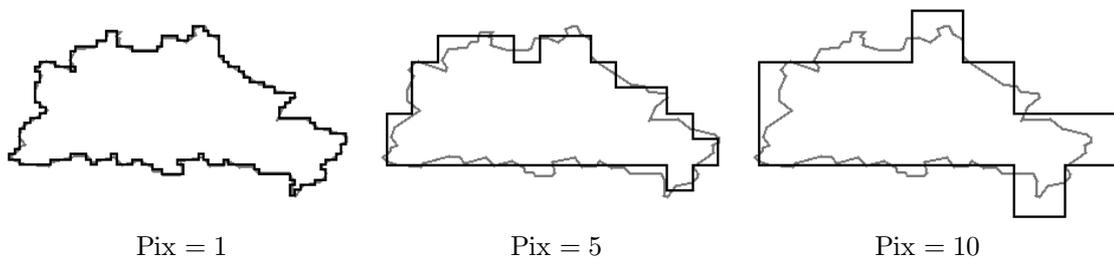
Von den verrauschten Länderkonturen wurden für alle Bundesländer die Momentinvarianten  $\Psi_1$  bis  $\Psi_4$  berechnet. Die Werte wurden auf den jeweiligen unverrauschten Wert normiert. Die relative Änderung der normierten Momentinvarianten gegenüber dem unverrauschten Wert ist in Abbildung 3.4 dargestellt. Wie aus Abbildung 3.4 ersichtlich ist, ändert sich die Standardabweichung der Momentinvarianten  $\Psi_1$  bis  $\Psi_4$  im Untersuchungsbereich nahezu proportional zur Standardabweichung des angebrachten Rauschens. Es zeigt sich ebenfalls, daß  $\Psi_1$  am wenigsten empfindlich reagiert ( $\sigma_{\Psi_1, \max} = 1.82\%$ ), für  $\Psi_2$  zeigt sich die größte relative Änderung ( $\sigma_{\Psi_2, \max} = 33.25\%$ ).



**Abbildung 3.4:** Verhalten der Momentinvarianten  $\Psi_1$  bis  $\Psi_4$  unter statistischem Rauschen. Die Standardabweichung ist in der Abszisse abgetragen. Auf der Ordinate ist die Standardabweichung der Invarianten, gemittelt über alle Objekte aus Abbildung 3.2, abgetragen.

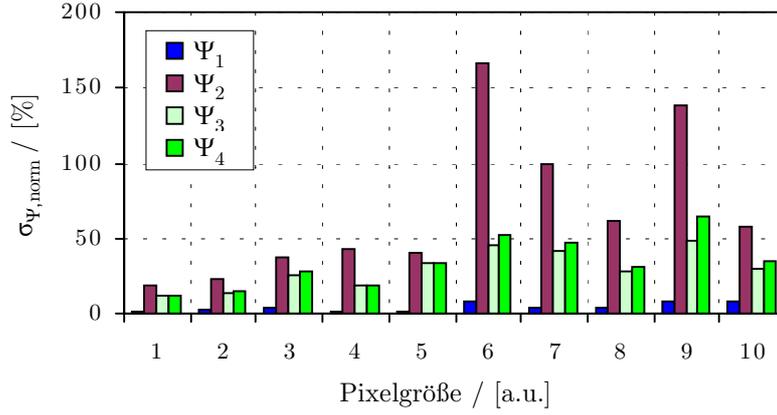
### Diskretisierungsrauschen und Momentinvarianten

Eine zweite Untersuchung am selben Beispieldatensatz (Bundesländer der BRD) beleuchtet den Einfluß der Pixelgröße bei der Diskretisierung der Polygone. Die Vektor-Raster-Konversion stellt ebenfalls eine Verrauschung der Koordinaten der Randpolygone dar. Im Unterschied zu statistischem Rauschen besteht jedoch eine Systematik, da jede Polygonecke nur auf einem Rasterpunkt des Basisrasters "einrasten" kann. Dieser Fall ist für die vorliegende Arbeit von besonderer Bedeutung, da die segmentierten Fernerkundungsdatenobjekte in Raster- und die Referenzobjekte in Vektorform vorliegen. Für eine sichere Zuordnung anhand invarianter Merkmale muß der Einfluß der Diskretisierung so gering wie möglich bleiben. Für diese Untersuchung wurde die ursprüngliche Pixelgröße bis auf einen zehnfachen Wert vergrößert. Ebenfalls am Beispiel der Landesgrenze Berlins ist in Abbildung 3.5 die Diskretisierung für ausgewählte Pixelgrößen dargestellt.



**Abbildung 3.5:** Die Grenze des Bundeslandes Berlin in Pixelgrößen von 1-10 (beliebige Einheiten).

Die Standardabweichung aller berechneten Momentinvarianten ist in Abbildung 3.6 zu sehen. Auch hier zeigt sich die geringste Empfindlichkeit für  $\Psi_1$  und die stärkste für  $\Psi_2$ . Im Gegensatz zum statistischen Rauschen liegt hier jedoch in Folge der bereits erwähnten Systematik keine monotone Zunahme vor. Vielmehr können offensichtlich besonders ungünstige Konstellationen wie im Fall der Pixelgrößen 6,7 und 9 zu besonders hoher Empfindlichkeit führen (im Beispiel bis zu  $\sigma_{\Psi_{\max}} = 166.58\%$ ). Insgesamt sind die Einflüsse von statistischem und Diskretisierungsrauschen als relativ hoch zu bewerten. Besonders im Fall der Diskretisierung ist der Einfluß mit relativen Standardabweichungen von teilweise über 150% (siehe Abbildung 3.6) unter Umständen zu groß, um mit den Momentinvarianten  $\Psi_1$  bis  $\Psi_4$  eine merkmalsbasierte Objekterkennung zuverlässig durchführen zu können. Insbesondere bei Verwendung von Bildpyramiden, bei denen die Pixelgröße in der Regel jeweils verdoppelt wird, ist die dadurch entstehende Unsicherheit in den Merkmalswerten signifikant. Diese Erkenntnis wurde für verrauschte Grauwertbilder bereits von Abu-Mostafa & Psaltis [1984] erlangt. Sie schlagen zur Umgehung der entstehenden Nachteile ein Normalisierungsverfahren vor, was hier jedoch nicht weiter verfolgt wird. Aufgrund der relativ hohen Empfindlichkeit der beschriebenen Invarianten gegenüber Schwankungen in der Objektkontur wird im folgenden Abschnitt ein Weg beschrieben, alternative Momentinvarianten, also solche, die nicht algebraischen Invarianten entstammen, zu finden.



**Abbildung 3.6:** Abhängigkeit der Momentinvarianten  $\Psi_1$  bis  $\Psi_4$  von der Pixelgröße bei einer Vektor-Raster-Konversion. Auf der Ordinate ist die Standardabweichung der Invarianten, normiert auf den unverrauschten Wert, abgetragen.

## 3.2 Kreuzmomente

**Definition 9** Es seien im  $\mathbb{R}^2$  zwei Bereiche gegeben mit  $(x, y) \in B_1$  und  $(x', y') \in B_2$ . Eine Funktion  $w(x, y, x', y')$ , die auf dem kartesischen Produkt  $B_1 \times B_2$  definiert ist, heißt Kreuzkorrelationsfunktion, wenn sie die Gebiete  $B_1$  und  $B_2$  miteinander in Beziehung setzt. Im Falle  $B_1 = B_2$  heißt sie Autokorrelationsfunktion. Als Kreuzmoment bezeichnet man den Ausdruck<sup>3</sup>

$$\mu_{k,l,p,q} = \iint_{B_1} \iint_{B_2} (x - \bar{x})^k (y - \bar{y})^l (x' - \bar{x}')^p (y' - \bar{y}')^q w(x, y, x', y') * f_1(x, y) * f_2(x', y') dx dy dx' dy' \quad (3.19)$$

wobei  $f_1(x, y)$  und  $f_2(x', y')$  als Dichtefunktionen über den Bereichen  $B_1$  und  $B_2$  zu verstehen sind und  $(\bar{x}, \bar{y})$  bzw.  $(\bar{x}', \bar{y}')$  die Schwerpunkte der Bereiche darstellen.

Im Falle geometrischer Objekte, wie sie in dieser Arbeit behandelt werden, setzt man die Dichten  $f_1(x, y) = f_2(x', y') = 1$ . Für den Autokorrelationsfall  $B_1 = B_2 = B$  schreibt sich dann das Auto-Kreuzmoment zu

$$\mu_{k,l,p,q} = \iint_B \iint_B (x - \bar{x})^k (y - \bar{y})^l (x' - \bar{x}')^p (y' - \bar{y}')^q w(x, y, x', y') dx dy dx' dy' . \quad (3.20)$$

Eine anschauliche physikalische Bedeutung des Kreuzmoments erhält man z.B. für den Fall, daß  $w(x, y, x', y')$  gleich dem Kehrwert des Abstandsquadrates der Punkte  $(x, y)$  und  $(x', y')$  ist. Dann nämlich stellt  $\mu_{k,l,p,q}$  das gravitative Wechselwirkungspotenzial zwischen den Bereichen  $B_1$  und  $B_2$  dar.

Im Falle der nullten Ordnung ( $k = l = p = q = 0$ ) lautet das Auto-Kreuzmoment für geometrische Objekte

$$\mu_{0,0,0,0} = \iint_B \iint_B w(x, y, x', y') dx dy dx' dy' . \quad (3.21)$$

Ein Punktpaar  $\mathbf{x}(x, y)$  und  $\mathbf{x}_a(x_a, y_a)$ , sei durch eine affine Transformation

$$\mathbf{x}_a^T = \mathbf{A}\mathbf{x} + \mathbf{b} \quad (3.22)$$

miteinander verknüpft.  $\mathbf{A}$  sei dabei eine Matrix mit vollem Rang. Für Mengen von Punktpaaren  $\{\mathbf{x}_i\}$  und  $\{\mathbf{x}_{a,i}\}$  bzw. Bereichen  $B$  und  $B_a$  gilt dann für die jeweiligen Schwerpunkte  $\bar{\mathbf{x}}$  und  $\bar{\mathbf{x}}_a$  die Beziehung (ohne Beweis)

$$\bar{\mathbf{x}}_a = \mathbf{A}\bar{\mathbf{x}} + \mathbf{b} . \quad (3.23)$$

<sup>3</sup>Die Ableitung der Invarianten aus Auto-Kreuzmomenten ist an die Darstellung in Yang & Cohen [1999] angelehnt. Dieser Beitrag ist jedoch insofern nicht ganz korrekt, als in der angegebenen Beziehung für die absoluten Invarianten aus Auto-Kreuzmomenten (Gleichung 3.10 in der o.g. Veröffentlichung) im Radikanden des Zählers und des Nenners jeweils ein Summand von 2 fehlt. Der Fehler wird mit entsprechendem Hinweis im Text mit der hier gegebenen Ableitung korrigiert.

Aus Gleichung 3.22 und Gleichung 3.23 lässt sich folgende relative affine Invariante ableiten:

$$[(x_a - \bar{x}_a)(y'_a - \bar{y}_a) - (x'_a - \bar{x}_a)(y_a - \bar{y}_a)] = \det \mathbf{A} [(x - \bar{x})(y' - \bar{y}) - (x' - \bar{x})(y - \bar{y})] . \quad (3.24)$$

Wenn die drei Punkte  $\mathbf{x}_i(x_i, y_i)$ ,  $\mathbf{x}_j(x_j, y_j)$  und  $\bar{\mathbf{x}}(\bar{x}, \bar{y})$  nicht kollinear sind und damit ein Dreieck bilden, so besagt Gleichung 3.24, daß für die Flächen  $S$  des originalen und  $S_a$  des affin transformierten Dreiecks die folgende relative Invarianzbeziehung gilt:

$$S_a = \det \mathbf{A} * S , \quad (3.25)$$

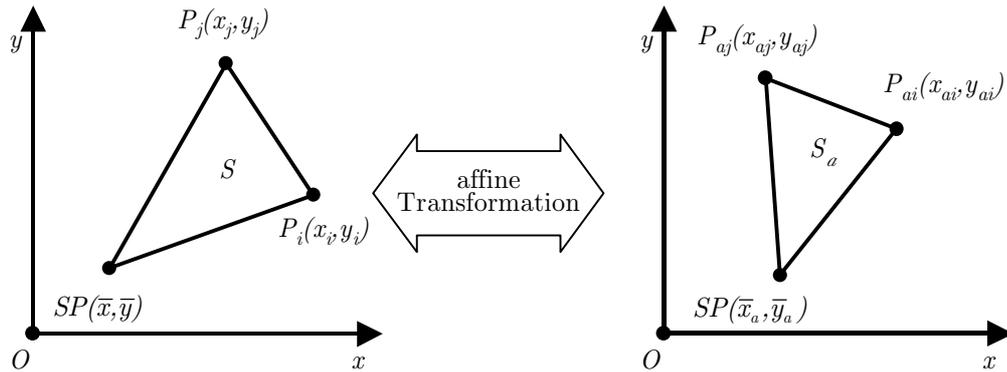
da sich die Dreiecksflächen auch durch

$$\begin{aligned} S &= \begin{vmatrix} \bar{x} & \bar{y} & 1 \\ x & y & 1 \\ x' & y' & 1 \end{vmatrix} = \bar{x}(y - y') + \bar{y}(x' - x) + xy' - x'y \\ &= (x - \bar{x})(y' - \bar{y}) - (x' - \bar{x})(y - \bar{y}) \end{aligned}$$

beziehungsweise

$$\begin{aligned} S_a &= \begin{vmatrix} \bar{x}_a & \bar{y}_a & 1 \\ x_a & y_a & 1 \\ x'_a & y'_a & 1 \end{vmatrix} = \bar{x}_a(y_a - y'_a) + \bar{y}_a(x'_a - x_a) + x_a y'_a - x'_a y_a \\ &= (x_a - \bar{x}_a)(y'_a - \bar{y}_a) - (x'_a - \bar{x}_a)(y_a - \bar{y}_a) \end{aligned}$$

ausdrücken lassen. Dieser Sachverhalt wird mit der folgenden Abbildung 3.7 (entnommen aus [Yang & Cohen 1999]) veranschaulicht.



**Abbildung 3.7:** Affine Transformation eines Dreiecks aus zwei aufeinanderfolgenden Konturpunkten und dem Objektschwerpunkt zur Veranschaulichung der relativen Invarianten aus Gl.3.24.

Die Invarianzbeziehung aus Gleichung 3.24 legt folgende beiden Definitionen der Kreuzkorrelationsfunktion nahe:

$$w_{1,s}(x, y, x', y') = [(x - \bar{x})(y' - \bar{y}) - (x' - \bar{x})(y - \bar{y})]^s \quad (3.26a)$$

$$w_{2,s}(x, y, x', y') = |[x - \bar{x})(y' - \bar{y}) - (x' - \bar{x})(y - \bar{y})]|^s . \quad (3.26b)$$

Den Exponenten  $s$  bezeichnet man als *Gewichtsfaktor*. Für diese Definitionen gelten infolge ihrer Linearität in  $x$  und  $y$  folgende Invarianzbeziehungen:

$$w_{1a,s}(x_a, y_a, x'_a, y'_a) = (\det \mathbf{A})^s w_{1,s}(x, y, x', y') \quad (3.26c)$$

$$w_{2a,s}(x_a, y_a, x'_a, y'_a) = |\det \mathbf{A}|^s w_{2,s}(x, y, x', y') . \quad (3.26d)$$

Die Beziehungen der Gleichungen 3.26c und 3.26d erfordern zur Berechnung die Kenntnis von Punktkorrespondenzen, die nur schwierig oder unter hohem Aufwand zu erhalten sind. Man kann nun versuchen, eine Invarianzbeziehung für die Kreuzmomente anstelle der Kreuzkorrelationsfunktionen zu gewinnen. Dazu setzt man die Kreuzkorrelationsfunktionen der Gleichungen 3.26a und 3.26b in die Definition der Auto-Kreuzmomente (Gleichung 3.20) ein. Für die nullte Ordnung steht dann:

$$\mu_{i,(0,0,0,0),s} = \iint_B \iint_B w_{i,s}(x, y, x', y') dx dy dx' dy' \quad \text{für } i \in \{1, 2\} \quad (3.27)$$

beziehungsweise nach affiner Transformation:

$$\mu_{ia,(0,0,0,0),s} = \iint_B \iint_B w_{ia,s}(x_a, y_a, x'_a, y'_a) dx_a dy_a dx'_a dy'_a \quad \text{für } i \in \{1, 2\} . \quad (3.28)$$

Für die weitere Ableitung soll ausschließlich die Alternative  $w_{1,s}(x_a, y_a, x'_a, y'_a)$  verfolgt werden, da für diese Alternative und für positive ganzzahlige Exponenten  $s$  die Auto-Kreuzmomente in Abhängigkeit der zentralen geometrischen Momente angegeben werden können. Darauf wird später in diesem Abschnitt näher eingegangen.

Die Frage lautet nun: In welcher Beziehung stehen die Kreuzmomente nullter Ordnung vor und nach Affintransformation zueinander ?

Nach Gleichung 3.22 in koordinatenweiser Darstellung stehen die transformierten Koordinaten zu den untransformierten in folgender funktionaler Beziehung:

$$x_a = x_a(x, y) = a_{11}x + a_{12}y + b_1 \quad (3.29a)$$

$$y_a = y_a(x, y) = a_{21}x + a_{22}y + b_2 . \quad (3.29b)$$

Zur Darstellung eines Integrals in transformierten Koordinaten muß man deren funktionale Beziehung in Form der Jacobi-Determinante einbringen.

Für den Ausdruck der affin transformierten Koordinaten in untransformierten lautet die Jacobi-Determinante:

$$\frac{\partial(x_a, y_a)}{\partial(x, y)} = \begin{vmatrix} \frac{\partial x_a}{\partial x} & \frac{\partial y_a}{\partial x} \\ \frac{\partial x_a}{\partial y} & \frac{\partial y_a}{\partial y} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{vmatrix} = \det \mathbf{A} . \quad (3.30)$$

Diese Beziehung gilt äquivalent auch für  $(x'_a, y'_a, x, y)$ .

Anschaulich ausgedrückt gibt die Jacobi-Determinante die Änderung der Größe einer Koordinaten-Einheitszelle (im  $\mathbb{R}^2$  eine Fläche) bei Durchführung einer gegebenen Transformation an. Im vorliegenden Fall der affinen Transformation im  $\mathbb{R}^2$  ist das Verhältnis von transformierter zu untransformierter Fläche eben gerade gleich der Determinante der Matrix des homogenen Teils der Affintransformation.

Formal muß man also beim Übergang von affin transformierten zu untransformierten Koordinaten für das Integral in Gleichung 3.28 folgende Ersetzungen vornehmen:

$$dx_a dy_a \longrightarrow \frac{\partial(x_a, y_a)}{\partial(x, y)} dx dy = (\det \mathbf{A}) dx dy \quad \text{bzw.} \quad dx'_a dy'_a \longrightarrow \frac{\partial(x'_a, y'_a)}{\partial(x', y')} dx' dy' = (\det \mathbf{A}) dx' dy' .$$

Damit und mit den Gleichungen 3.26c und 3.26d lassen sich nun die affin transformierten Kreuzmomente nullter Ordnung in Abhängigkeit der untransformierten Koordinaten angeben und damit direkt mit den untransformierten Kreuzmomenten nullter Ordnung in Beziehung setzen. Die affin transformierten Kreuzmomente nullter Ordnung nehmen dabei folgende Gestalt an:

$$\mu_{1a,(0,0,0,0),s} = \iint_B \iint_B w_{1a,s}(x_a, y_a, x'_a, y'_a) dx_a dy_a dx'_a dy'_a \quad (3.31)$$

$$= \iint_B \iint_B (\det \mathbf{A})^s w_{1,s}(x, y, x', y') (\det \mathbf{A})^2 dx dy dx' dy' \quad (3.32)$$

$$= (\det \mathbf{A})^{s+2} \iint_B \iint_B w_{1,s}(x, y, x', y') dx dy dx' dy' \quad (3.33)$$

und damit

$$\mu_{1a,(0,0,0,0),s} = (\det \mathbf{A})^{s+2} \mu_{1,(0,0,0,0),s} . \quad (3.34)$$

Damit ist der Zusammenhang zwischen transformierten und untransformierten Auto-Kreuzmomenten hergestellt. Die Kenntnis konkreter Punktkorrespondenzen ist dazu nicht mehr notwendig<sup>4</sup>.

Die abgeleitete Gleichung 3.34 stellt eine relative Invarianzbeziehung dar. Um daraus absolute affine Invarianten abzuleiten, muß die Abhängigkeit von  $(\det \mathbf{A})$  eliminiert werden. Dazu geht man vor wie in Abschnitt 3.1.4

<sup>4</sup>Zur Anmerkung sei gesagt, daß analog dazu die Beziehung  $\mu_{2a,(0,0,0,0),s} = |\det \mathbf{A}|^{s+2} \mu_{2,(0,0,0,0),s}$  gilt.

und betrachtet die Auto-Kreuzmomente für zwei unterschiedliche Exponenten  $s_0$  und  $s_1$ . Man benutzt also die relativen Invarianzbeziehungen

$$\mu_{1a,(0,0,0,0),s_0} = (\det \mathbf{A})^{s_0+2} \mu_{1,(0,0,0,0),s_0} \quad (3.35)$$

$$\mu_{1a,(0,0,0,0),s_1} = (\det \mathbf{A})^{s_1+2} \mu_{1,(0,0,0,0),s_1} . \quad (3.36)$$

Die Eliminierung von  $(\det \mathbf{A})$  aus den Gleichungen 3.35 und 3.36 liefert die folgende Schar absoluter affiner Invarianten.<sup>5</sup>

$$\phi(s_0, s_1) = \frac{s_1+2\sqrt{\mu_{1a,(0,0,0,0),s_1}}}{s_0+2\sqrt{\mu_{1a,(0,0,0,0),s_0}}} = \frac{s_1+2\sqrt{\mu_{1,(0,0,0,0),s_1}}}{s_0+2\sqrt{\mu_{1,(0,0,0,0),s_0}}} . \quad (3.37)$$

Bis hierher wurde an die Exponenten  $s_0$  und  $s_1$  keine einschränkende Bedingung gestellt. Wie man Gleichung 3.26a entnehmen kann, zeigt die Kreuzkorrelationsfunktion  $w_{1,s}(x, y, x', y')$  folgendes Symmetrieverhalten:

$$w_{1,s}(x', y', x, y) = \begin{cases} -w_{1,s}(x, y, x', y') \text{ (antisymmetrisch)} & \text{für } s \text{ ist natürliche, ungerade Zahl} \\ w_{1,s}(x, y, x', y') \text{ (symmetrisch)} & \text{für } s \text{ ist natürliche, gerade Zahl} . \end{cases} \quad (3.38)$$

Ohne Beweis wird das folgende Theorem gegeben:

**Theorem 10 (Symmetrietheorem)** *Wenn die Kreuzkorrelationsfunktion antisymmetrisch ist, also die Beziehung:  $w_{1,s}(x', y', x, y) = -w_{1,s}(x, y, x', y')$  gilt, dann ist auch das zugehörige Auto-Kreuzmoment antisymmetrisch, d.h.  $\mu_{1,(p,q,k,l),s} = -\mu_{1,(k,l,p,q),s}$ . Insbesondere gilt für die nullte Ordnung:  $\mu_{1,(0,0,0,0),s} = 0$ .*

Wie bereits weiter oben angedeutet, lassen sich die Auto-Kreuzmomente nullter Ordnung für positive ganze Zahlen durch die zentralen geometrischen Momente  $m_{p,q}$  ausdrücken:

$$\mu_{1,(0,0,0,0),s} = \sum_{g=0}^s (-1)^{s-g} \binom{s}{g} m_{g,s-g} m_{s-g,g} . \quad (3.39)$$

Zusammenfassend läßt sich also folgende für das Ziel der vorliegenden Arbeit wichtige Aussage treffen:

*Für positive, gerade Zahlen  $s_0, s_1$  lassen sich mit Hilfe von Auto-Kreuzmomenten nullter Ordnung absolute Invarianten unter affiner Transformation aus zentralen geometrischen Momenten  $m_{p,q}$  ableiten. Die Invarianten nehmen dabei folgende Gestalt an:*

$$\Phi(s_0, s_1) = \frac{s_1+2\sqrt{\sum_{g_1=0}^{s_1} (-1)^{s_1-g_1} \binom{s_1}{g_1} m_{g_1,s_1-g_1} m_{s_1-g_1,g_1}}}{s_0+2\sqrt{\sum_{g_0=0}^{s_0} (-1)^{s_0-g_0} \binom{s_0}{g_0} m_{g_0,s_0-g_0} m_{s_0-g_0,g_0}}} . \quad (3.40)$$

Wählt man  $s_0 = s_1$ , so ergibt sich der Trivialfall  $\Phi(s_0, s_1) = 1$ , der jedoch für die angestrebte Objekterkennungsaufgabe keine Bedeutung besitzt.

Die in Gleichung 3.20 definierten Auto-Kreuzmomente lassen sich naturgemäß nicht nur für ganze, sondern für beliebige reelle Zahlen  $s$  berechnen. Im allgemeinen Fall geht aber der Zusammenhang zu den zentralen geometrischen Momenten verloren. Aus diesem Grund werden hier nur positive, gerade  $s$  behandelt. Eine Diskussion zu reellen  $s$  findet sich bei Yang & Cohen [1999].

**Beispiel 11** *Als Beispiel sei hier die absolute Invariante für  $s_0 = 2$  und  $s_1 = 4$  angegeben, die den einfachsten nichttrivialen Fall darstellt. Sie nimmt folgende Form an:*

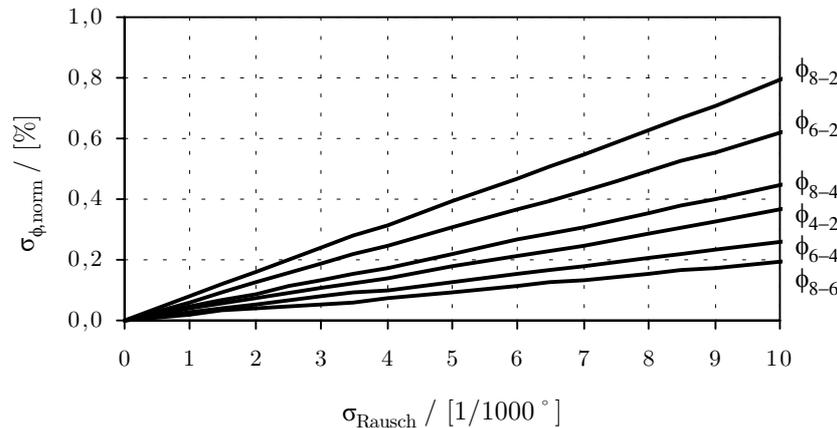
$$\Phi(2, 4) = \frac{\sqrt[6]{m_{4,0}m_{0,4} + 3m_{2,2}^2 - 4m_{3,1}m_{1,3}}}{\sqrt[4]{m_{2,0}m_{0,2} - m_{1,1}^2}} . \quad (3.41)$$

In Abschnitt 3.1.5 wurde für Momentinvarianten aus algebraischen Momenten der Einfluß von Rauschen (statistisch und aufgrund der Diskretisierung) untersucht. Eine entsprechende Untersuchung wurde auch für die soeben eingeführten Momentinvarianten aus Kreuzmomenten durchgeführt. Aus Gründen der direkten Vergleichbarkeit wurden ebenfalls die Bundesländer der BRD herangezogen (siehe Abbildung 3.2). Auch hier wird statistisches und Diskretisierungsrauschen unterschieden.

<sup>5</sup>Das Ergebnis in Yang & Cohen [1999] lautet abweichend davon:  $\phi(s_0, s_1) = \frac{s_1\sqrt{\mu_{1,(0,0,0,0),s_1}}}{s_0\sqrt{\mu_{1,(0,0,0,0),s_0}}}$ , d.h. die Radikanden sind jeweils um den Wert 2 kleiner. Die Ursache hierfür liegt in der Ableitung. Yang & Cohen [1999] leiten die absoluten Invarianten aus Kreuzmomenten in diskreten Koordinaten (Raster) ab. Dabei wird nach Ansicht des Autors fälschlicherweise die Rolle der Jacobi-Determinante vernachlässigt. Dies schlägt sich darin nieder, daß das Ergebnis von Yang & Cohen [1999] für alle Transformationen, deren Determinante 1 ist, richtig ist, für alle anderen jedoch nicht. Diese Aussage läßt sich leicht durch Einsetzen beweisen.

### Statistisches Rauschen und Kreuzmoment-Invarianten

Der Einfluß von weißem Rauschen auf die Kreuzmoment-Invarianten  $\Phi(2,4)$  bis  $\Phi(8,6)$  (s. Gleichung 3.40) wurde durch Anbringen von statistischem Rauschen an den Polygonecken mit Standardabweichungen zwischen  $\sigma_{\min} = 0$  und  $\sigma_{\max} = 1/100^\circ$  (etwa 1000m) untersucht. Die Auswirkungen des angebrachten Rauschens sind in Abbildung 3.3 am Beispiel der Grenze des Landes Berlin aufgezeigt. Hier wurden von den verrauschten Länderkonturen für alle Bundesländer die Momentinvarianten  $\Phi(2,4)$  bis  $\Phi(8,6)$  berechnet. Die Werte wurden normiert auf den jeweiligen unverrauschten Wert. Die relative Änderung der normierten Kreuzmoment-Invarianten gegenüber dem unverrauschten Wert zeigt Abbildung 3.8.



**Abbildung 3.8:** Verhalten der Momentinvarianten  $\Phi(2,4)$  bis  $\Phi(8,6)$  unter statistischem Rauschen. Die Standardabweichung ist in der Abszisse abgetragen, während auf der Ordinate die Standardabweichung der Invarianten, gemittelt über alle Objekte aus Abbildung 3.2, abgetragen ist.

Im Gegensatz zu den Momentinvarianten  $\Psi_1$  bis  $\Psi_4$ , bei denen im gleichen Untersuchungsintervall (0.001-0.01 $^\circ$ ) relative Fehler von bis über 30% auftraten, liegen hier die relativen Fehler über alle Objekte des Testdatensatzes gemittelt für alle Invarianten unter 1%. Für weißes Rauschen zeigt sich also eine wesentlich verringerte Empfindlichkeit der Kreuzmoment-Invarianten.

### Diskretisierungsrauschen und Kreuzmoment-Invarianten

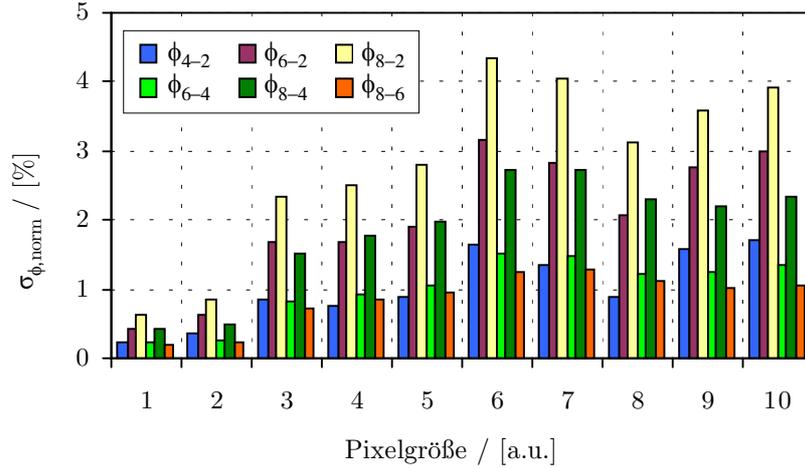
Für den Einfluß der Pixelgröße auf die Kreuzmoment-Invarianten  $\Phi(2,4)$  bis  $\Phi(8,6)$  wurde analog zu den Betrachtungen in Abschnitt 3.1.6 eine Aufrasterung der Länderpolygone in Pixelgrößen von 1-10 Einheiten untersucht. Die Ergebnisse zeigt Abbildung 3.9. Auch hier ist mit einem maximalen relativen Fehler von weniger als 5% im selben Untersuchungsbereich gegenüber mehr als 150% für Momentinvarianten (Abbildung 3.6) eine deutliche Abnahme der Empfindlichkeit gegen Diskretisierungsrauschen zu beobachten.

Insgesamt zeigen die Untersuchungen zum Einfluß von statistischem Rauschen und Diskretisierungsrauschen, daß Kreuzmoment-Invarianten infolge höherer Robustheit besser zur Objekterkennung geeignet sind als die Momentinvarianten, die auf algebraischen Invarianten beruhen. Für den weiteren Verlauf der Arbeit werden aufgrund dieses Ergebnisses lediglich die Kreuzmoment-Invarianten betrachtet.

## 3.3 Berechnung invarianter Merkmale

In den vorigen Abschnitten wurde die Theorie der Invarianten aus geometrischen Momenten vorgestellt und an Beispielen erläutert. Explizit durchgeführt wurde die Theorie für Invarianten binärer algebraischer Formen, die mit Hilfe des Fundamentaltheorems zu absoluten Invarianten unter affiner Transformation führten. Des weiteren wurden Kreuzmomente eingeführt und Invarianten aus Kreuzmomenten abgeleitet, die eine wesentlich höhere Robustheit gegenüber Rauschen in den Objektkoordinaten besitzen.

Mit den folgenden Abschnitten werden die rechentechnischen Aspekte beleuchtet und die Algorithmen und Methoden zur Berechnung geometrischer Momente aufgeführt, die für diese Arbeit implementiert und getestet wurden.



**Abbildung 3.9:** Verhalten der Kreuzmoment-Invarianten  $\Phi(2, 4)$  bis  $\Phi(8, 6)$  unter Diskretisierung mit unterschiedlicher Pixelgröße. Auf der Abszisse sind die angewandten Pixelgrößen abgetragen, die Ordinate zeigt die resultierenden relativen Fehler aller Invarianten.

### 3.3.1 Geometrische Momente aus Objektkonturen

Zur Berechnung der regulären wie auch der zentralen geometrischen Momente eines Objektes muß im kontinuierlichen Fall eine Integration und im diskreten Fall eine Summation über das gesamte Integrationsgebiet (die Objektfläche)  $G$  erfolgen. Eine derartige Berechnung ist sehr aufwendig, es ist daher erstrebenswert, eine elegantere und rechentechnisch effizientere Formulierung zur Berechnung geometrischer Momente aufzusuchen. Eine Möglichkeit zur Lösung der Gleichung 3.6 besteht in der Umwandlung des Gebietsintegrals in ein Kurvenintegral mit Hilfe des Integralsatzes von Green. Die folgenden Ausführungen gelten gleichermaßen für reguläre (nicht auf den Objektschwerpunkt bezogene) und für zentrale geometrische Momente. Das Symbol  $M$  steht von daher bis auf weiteres für beide Varianten.

Nach Bronstein [1987] lautet der Green'sche Integralsatz:

$$\iint_G \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dx dy = \oint_C (u dy + v dx) , \quad (3.42)$$

wobei  $C(G)$  der Rand des Integrationsgebietes  $G$  ist. Die Anwendung des Green'schen Integralsatzes erlaubt somit die Berechnung geometrischer Momente aus der Kontur eines Objekts. Zu beachten ist dabei allerdings, daß bei mehrfach zusammenhängenden Gebieten auch die Innenkonturen mitberücksichtigt werden müssen.

Um die Rechnung durchführen zu können, müssen Funktionen  $u(x, y)$  und  $v(x, y)$  derart gefunden werden, daß sie der Bedingung

$$\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = x^p y^q f(x, y) \quad (3.43)$$

genügen. Für eine beliebige Funktion  $f(x, y)$  ist dies extrem schwierig, der rechentechnische Aufwand würde durch den zu erwartenden Nutzen nicht gerechtfertigt werden. Im Falle einer binären Indikatorfunktion, in dem innerhalb des Integrationsgebiets  $f(x, y) = 1$  und außerhalb  $f(x, y) = 0$  ist, sind dagegen leicht Integranden anzugeben, die der Bedingung 3.43 genügen. Zwei einfache Möglichkeiten für Binärbilder oder geometrische Objekte sind z.B. gegeben durch

$$u(x, y) = \frac{x^{p+1}}{p+1} y^q \quad \text{und} \quad v(x, y) = 0 \quad (3.44)$$

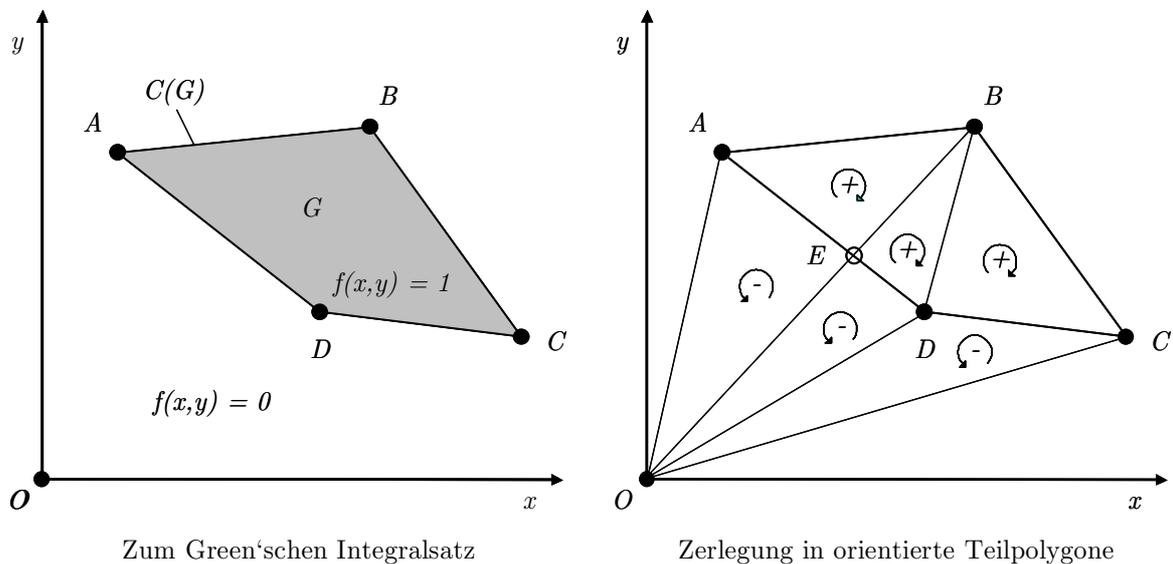
oder

$$u(x, y) = 0 \quad \text{und} \quad v(x, y) = -x^p \frac{y^{q+1}}{q+1} . \quad (3.45)$$

Ohne Beschränkung der Allgemeinheit wird im folgenden nur die Lösung 3.44 weiterverfolgt. Für diesen Ansatz ist das folgende Ringintegral zu lösen:

$$M_{pq} = \oint_{C(G)} \frac{x^{p+1}}{p+1} y^q dy. \quad (3.46)$$

In Abbildung 3.10, links, sind die Verhältnisse zur Anwendung des Green'schen Integralsatzes veranschaulicht. Da im Rahmen dieser Arbeit Momente geometrischer Objekte (Vektorpolygone) berechnet werden, ist die Indikatorfunktion  $f(x, y)$  binär: Innerhalb des Objekts nimmt sie den Wert 1 an, außerhalb den Wert 0. Von Singer [1993] wurde eine Methode vorgeschlagen, mit der man das Ringintegral aus Gleichung 3.46 für polygonale Objekte als Summe der Integrale längs einzelner Polygonseiten ausdrücken kann. Dazu stellt man folgende Überlegung an. Wie in Abbildung 3.10, rechts, dargestellt, zerlegt man das orientierte Polygon (hier: im Uhrzeigersinn zählt positiv) derart in Dreiecke, daß jede Polygonseite mit einem nicht auf  $C(G)$  liegenden Fixpunkt (hier: Der Koordinatenursprung  $O$ ) ein Dreieck bildet. Für den Fall, daß eine Dreiecksseite eine Polygonseite schneidet, wie in Abbildung 3.10, rechts, beim Dreieck  $OAB$ , wird die geschnittene Polygonseite mit einem Hilfspunkt in zwei lineare Teilstrecken unterteilt (Punkt  $E$  in Abbildung 3.10, rechts). Als Integrale bilden



**Abbildung 3.10:** Zur Berechnung geometrischer Momente aus der Objektkontur. Linkes Teilbild: Polygonobjekt mit Randkurve  $C(G)$  und Indikatorfunktion  $f(x, y)$  zur Veranschaulichung der Verwendung des Green'schen Integralsatzes. Rechtes Teilbild: Zerlegung des Polygonobjekts in Dreiecke mit Umlaufsinn zur Berechnung des Kurvenintegrals aus den einzelnen Liniensegmenten.

die Teilmomente der einzelnen Dreiecke lineare Ausdrücke. Damit läßt sich das Moment  $M_{ABCD}$  des Polygons als Summe der Teilmomente der einzelnen Dreiecke ausdrücken. Der Umlaufsinn ist dabei im Vorzeichen der Summanden zu berücksichtigen. Die Zerlegung sieht folgendermaßen aus:

$$\begin{aligned} M_{ABCD} &= M_{ABO} - M_{AEO} + M_{BCO} - M_{DCO} - M_{EDO} \\ &= M_{ABO} + M_{BCO} - M_{DCO} - M_{ADO} \\ &= M_{ABO} + M_{BCO} + M_{CDO} + M_{DAO}. \end{aligned} \quad (3.47)$$

Man wählt die Orientierung des Polygons derart, daß das Moment nullter Ordnung (die Polygonfläche) ein positives Vorzeichen besitzt.<sup>6</sup>

Mit dem Ergebnis von Gleichung 3.47 läßt sich folgende für diese Arbeit wichtige Aussage treffen:

*Geometrische Momente (auch zentrale) beliebiger Ordnung eines im Gauss'schen Sinne positiv orientierten Polygons lassen sich als Summe der Teilmomente jeder Polygonseite ausdrücken. Das Teilmoment jeder Polygonseite  $\overline{P_k P_{k+1}}$  wird infolge des Green'schen Integralsatzes aus dem Kurvenintegral längs der geschlossenen*

<sup>6</sup>Für das geometrische Moment nullter Ordnung ergibt sich gerade die wohlbekanntete Gauss'sche Formel zur Berechnung einer Polygonfläche.

Kurve  $\overline{P_k P_{k+1} O}$  berechnet. Mathematisch ausgedrückt lautet die Aussage für ein Polygon mit  $K$  Knoten und  $K$  Kanten:

$$M_{p,q} = \sum_{k=0}^{K-1} M_{p,q,k} = \sum_{k=0}^{K-1} M_{P_k P_{k+1} O, p, q} \quad \text{wobei} \quad P_K = P_0. \quad (3.48)$$

Mit diesem Ergebnis ist es ausreichend, den Beitrag einer einzelnen Polygonseite zum Gesamtmoment funktional zu kennen. Zur Berechnung des Gesamtmoments wird dann sukzessive das gesamte Polygon durchlaufen. Der Rand  $C$  des Integrationsgebiets für das Teilmoment einer Polygonseite besteht aus den Teilkurven  $C_1 = \overline{OP_n}$ ,  $C_2 = \overline{P_n P_{n+1}}$ , und  $C_3 = \overline{P_{n+1} O}$ , so daß das Integral aus Gleichung 3.46 folgende Gestalt annimmt:

$$\oint_C \frac{x^{p+1}}{p+1} y^q dy = \sum_{i=1}^3 \int_{C_i} \frac{x_i^{p+1}}{p+1} y_i^q dy_i. \quad (3.49a)$$

Zur Auflösung der einzelnen Summanden werden die jeweils geraden Teilkurven  $C_i$  linear in  $s$  parametrisiert. Gewählt wird folgender Ansatz:

$$x_i(s) = a_i s + b_i \quad (3.49b)$$

$$y_i(s) = c_i s + d_i. \quad (3.49c)$$

Die Koeffizienten sehen im einzelnen folgendermaßen aus:

$$a_1 = \frac{X_1}{S_1} \quad b_1 = 0 \quad c_1 = \frac{Y_1}{S_1} \quad d_1 = 0$$

$$a_2 = \frac{X_2 - X_1}{S_2} \quad b_2 = X_1 \quad c_2 = \frac{Y_2 - Y_1}{S_2} \quad d_2 = Y_1$$

$$a_3 = -\frac{X_2}{S_3} \quad b_3 = X_2 \quad c_3 = -\frac{Y_2}{S_3} \quad d_3 = Y_2$$

wobei  $S_1$  die Länge der Teilkurve  $C_1$ ,  $S_2$  die Länge der Teilkurve  $C_2$  und  $S_3$  die Länge der Teilkurve  $C_3$  ist.

Beim Übergang von den Variablen  $x$  und  $y$  auf die Parametrisierung  $x(s)$  und  $y(s)$  mit der unabhängigen Variablen  $s$  gemäß Gleichung 3.49b und Gleichung 3.49c muß in den Einzelintegralen von  $dy_i$  auf  $ds$  übergegangen werden gemäß  $dy = ds * \frac{dy_i}{ds} = \frac{d(c_i s)}{ds} ds = c_i ds$ . Die Integration über  $s$  wird dementsprechend von 0 bis  $S_i$  vorgenommen.

Mit der in Gleichung 3.49b und Gleichung 3.49c eingeführten Substitution schreibt sich Gleichung 3.46 zu:

$$M_{pq} = \sum_{i=1}^3 \int_0^{S_i} \frac{(a_i s + b_i)^{p+1}}{p+1} (c_i s + d_i)^q c_i ds. \quad (3.50)$$

Zur Berechnung des Integrals wird der Integrand mit Hilfe des Binomischen Satzes umgeformt. Nach Bronstein [1987] lautet der Binomische Satz:

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \quad \text{mit } a, b \in \mathfrak{R} \text{ und } 0^0 \stackrel{\text{def}}{=} 0. \quad (3.51)$$

Für die Integranden in Gleichung 3.50 ergibt sich daraus:

$$(a_i s + b_i)^{p+1} = \sum_{n=0}^{p+1} \binom{p+1}{n} (a_i s)^{p+1-n} b_i^n \quad \text{und} \quad (3.52a)$$

$$(c_i s + d_i)^q = \sum_{l=0}^q \binom{q}{l} (c_i s)^{q-l} d_i^l \quad (3.52b)$$

Setzt man die Integranden in Gleichung 3.50 ein, so erhält man:

$$\begin{aligned} M_{pq} &= \sum_{i=1}^3 \int_0^{S_i} \frac{1}{p+1} \sum_{n=0}^{p+1} \sum_{l=0}^q \binom{p+1}{n} \binom{q}{l} (a_i s)^{p+1-n} b_i^n (c_i s)^{q-l} d_i^l c_i ds \\ &= \sum_{i=1}^3 \frac{1}{p+1} \sum_{n=0}^{p+1} \sum_{l=0}^q \binom{p+1}{n} \binom{q}{l} \int_0^{S_i} a_i^{p+1-n} b_i^n c_i^{q+1-l} d_i^l s^{p+q-n-l+1} ds. \end{aligned} \quad (3.53)$$

Zu bemerken ist, daß Gleichung 3.53 symmetrisch in  $p + 1$  und  $q$  ist.<sup>7</sup>

Die Berechnung des gesamten Ausdrucks in Gleichung 3.53 wird für jeden Summanden in  $i$  separat vorgenommen. Dabei werden jeweils die oben angegebenen Koeffizienten  $a_i, b_i, c_i$ , und  $d_i$  eingesetzt. Die einzelnen Schritte sind im Anhang A ausgeführt. Für das zu berechnende Integral aus Gleichung 3.46 ergibt sich insgesamt für ein Polygon mit  $K$  Kanten und  $K$  Knoten:

$$M_{pq} = \frac{1}{p+1} \sum_{k=0}^{K-1} \left\{ \frac{1}{p+q+2} X_k^{p+1} Y_k^{q+1} + \sum_{n=0}^{p+1} \sum_{l=0}^q \frac{1}{p+q-n-l+2} \binom{p+1}{n} \binom{q}{l} F_{pqnl}(X_k, X_{k+1}, Y_k, Y_{k+1}) \right\}. \quad (3.54)$$

Für  $k = 1$  schreibt sich

$$F_{pqnl}(X_1, X_2, Y_1, Y_2) = X_1^n Y_1^l (X_2 - X_1)^{p+1-n} (Y_2 - Y_1)^{q+1-l} + (-1)^{p+q-n-l} X_2^{p+1} Y_2^{q+1}.$$

Auch hier ist zu beachten:  $P_K(X_K, Y_K) = P_0(X_0, Y_0)$ , d.h. die Kurve ist geschlossen.

Diese Beziehung stellt in allgemeiner Weise das Bildungsgesetz für beliebige geometrische Momente aus der Kontur eines polygonalen Objekts dar. Aufgrund ihrer Struktur ist diese Beziehung gut für eine numerische Berechnung geeignet und wurde für die vorliegende Arbeit auch in exakt der in Gleichung 3.54 angegebenen Weise implementiert.

### 3.3.2 Zur Diskriminierbarkeit mit Kreuzmoment-Invarianten

Wie in Abschnitt 3.2 nachgewiesen wurde, sind Kreuzmoment-Invarianten sehr robust gegenüber statistischem und systematischem Rauschen in der Objektkontur, womit sie für Objekterkennungszwecke bei unbekannter Objektorientierung prinzipiell geeignet sind. Eine zuverlässige Objekterkennungsstrategie erfordert jedoch zusätzlich zur Robustheit eine hinreichend gute Diskriminierbarkeit durch die verwendeten Merkmale (hier: Kreuzmoment-Invarianten). Verwendet man  $n$  unabhängige Merkmale, so wird - wie bei den bekannten Verfahren der multispektralen Klassifikation von Rasterdaten (z.B. [Schowengerdt 1997], [Lillesand & Kiefer 1987]) - jedes Objekt durch einen Punkt in einem  $n$ -dimensionalen Merkmalsraum repräsentiert. Analog zu den Verfahren der Fernerkundung bezeichnet man Objekte also in diesem Kontext als ähnlich, wenn ihre euklidische Distanz im  $n$ -dimensionalen Merkmalsraum kleiner als ein bestimmter Wert  $\epsilon$  ist. Zwei Objekte  $O_1$  und  $O_2$  seien beschrieben durch ihre  $n$ -dimensionalen Merkmalsvektoren:  $O_1 : \Phi_1 = (\Phi_{11}, \dots, \Phi_{n1})^T$  und  $O_2 : \Phi_2 = (\Phi_{12}, \dots, \Phi_{n2})$  wobei  $\Phi_{lm}$  die  $l$ -te Kreuzmoment-Invariante des  $m$ -ten Objekts bedeutet.

Der relative Unterschied zweier reeller Zahlen  $r$  und  $s$  sei folgendermaßen definiert:

$$U(r, s) = \begin{cases} 0 & r = s \\ \left| \frac{r}{s} - 1 \right| & \text{für } |r| < |s| \\ \left| \frac{s}{r} - 1 \right| & |s| \leq |r| \end{cases} \quad (3.55)$$

Zwei Objekte gelten in dieser Arbeit als ähnlich, wenn der nach Gleichung 3.55 definierte relative Unterschied für die  $i$ -te Dimension einen bestimmten Wert  $\epsilon_i$  nicht überschreitet<sup>8</sup>, d.h. der folgende Ausdruck ein logisches "WAHR" liefert:

$$(U(\Phi_{11}, \Phi_{12}) < \epsilon_1) \cap (U(\Phi_{21}, \Phi_{22}) < \epsilon_2) \cap \dots \cap (U(\Phi_{n1}, \Phi_{n2}) < \epsilon_n) \quad (3.56)$$

Als Abstand zwischen zwei Objekten im Merkmalsraum läßt sich eine quasi-euklidische Distanz definieren zu<sup>9</sup>:

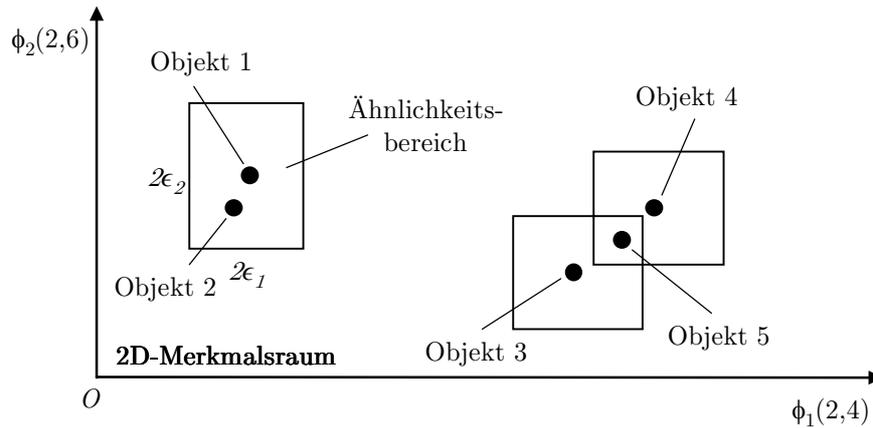
$$D(\Phi_1, \Phi_2) = \sqrt{\sum_{i=1}^n (U(\Phi_{i1}, \Phi_{i2}))^2}. \quad (3.57)$$

Die in Abschnitt 3.2 festgestellte Robustheit bedeutet somit, daß verrauschte Kopien des selben Objekts im Merkmalsraum eine kleine quasi-euklidische Distanz haben, also ähnlich im Sinne von Gleichung 3.56 sind. In Abbildung 3.11 sind die Verhältnisse in einem zweidimensionalen Merkmalsraum bei Verwendung der beiden einfachsten Invarianten  $\Phi(2, 4), \Phi(2, 6)$  dargestellt. Aufgrund der Symmetrie des relativen Unterschieds ist es

<sup>7</sup>Hätte man bei der Anwendung des Gauss'schen Integralsatzes nicht  $u(x, y) = \frac{x^{p+1}}{p+1} y^q$  und  $v(x, y) = 0$ , sondern  $u(x, y) = 0$  und  $v(x, y) = -x^p \frac{y^{q+1}}{q+1}$  gewählt, so bestünde dieselbe Symmetrie in  $p$  und  $q + 1$ .

<sup>8</sup>Bei der multispektralen Klassifikation von Rasterdaten, beziehungsweise ganz allgemein bei der Clusteranalyse, werden auch andere Distanzmaße verwendet, z.B. die euklidische Distanz (Minimum-Distance-Klassifikation mit Schwellwert), ein fester Wert (Box- oder Parallelepiped-Klassifikation) oder statistische Größen wie Mittelwert und Kovarianzmatrix. Es sei jedoch zu bedenken gegeben, daß hier zum einen der Definitionsbereich jeder Dimension nicht a priori bekannt ist (bei Multispektralaufnahmen z.B. 256 Grauwerte/Kanal). Zudem skalieren die Definitionsbereiche in der Regel unterschiedlich. Zum anderen stellt hier jedes Objekt ein Cluster für sich dar, so daß statistische Größen keine Rolle spielen können.

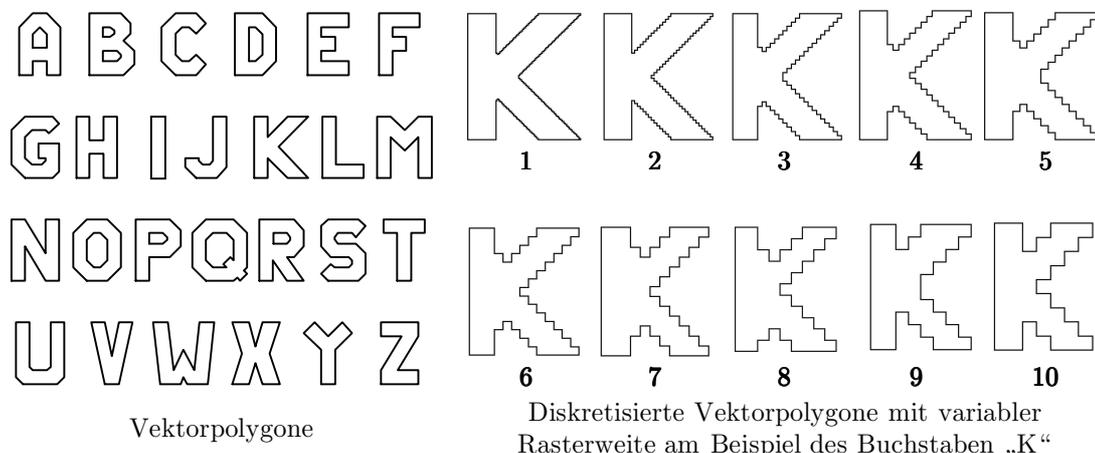
<sup>9</sup>Bei Verwendung von Kreuzmomentinvarianten steht der Index  $i$  für eine bestimmte, erlaubte Kombination  $(s_0, s_1)$ , z.B.  $\Phi_i = \Phi(2, 4)$ .



**Abbildung 3.11:** Repräsentation von Objekten im 2D-Merkmalraum. Objekte werden als verschieden betrachtet, wenn ihre Ähnlichkeitsbereiche sich nicht schneiden.

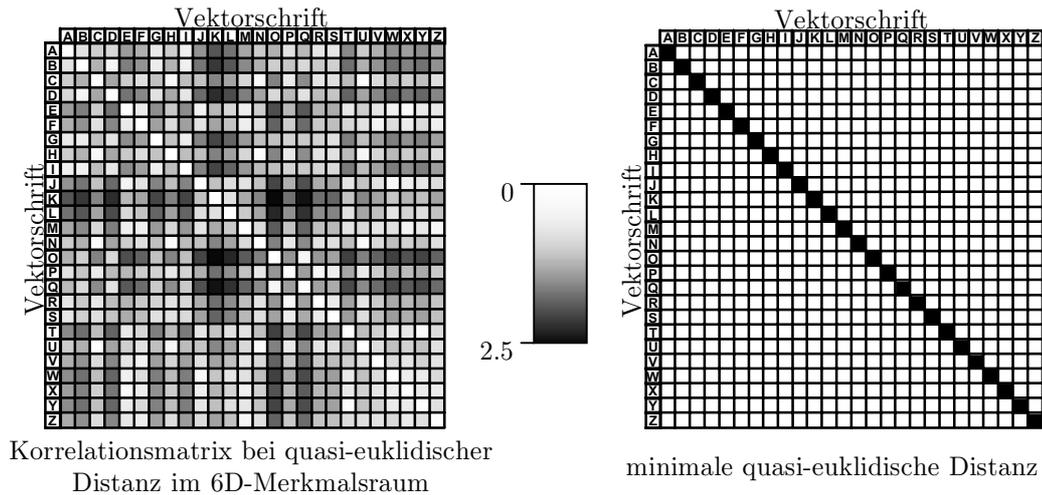
in Abbildung 3.11 gleichgültig, an welchem der beiden Objekte der Ähnlichkeitsbereich eingezeichnet wird. Im dargestellten Fall sind die Objekte  $O_1$  und  $O_2$ ,  $O_3$  und  $O_5$  sowie  $O_4$  und  $O_5$  paarweise ähnlich, alle anderen möglichen 2-er Paarungen jedoch nicht. Für diese Arbeit wird die beschriebene Ähnlichkeit bei der Suche nach konsistenten Korrespondenzen als sogenannte unäre Zwangsbedingung benutzt (Abschnitt 4.1).

Anhand eines Testdatensatzes, der aus den Buchstaben des Alphabets, dargestellt durch Vektorpolygone, besteht, wurde ein Experiment zur Objektdiskriminierbarkeit mit Kreuzmoment-Invarianten durchgeführt. Die Originalpolygone und die verschiedenen verwendeten Generalisierungen sind in Abbildung 3.12 dargestellt. Zur Modellierung realistischer Verhältnisse wurde der Polygondatensatz mit unterschiedlicher Rasterweite diskretisiert. Untersucht wurde die nach Gleichung 3.57 definierte quasi-euklidische Distanz zwischen den generalisierten und den Originalbuchstaben.

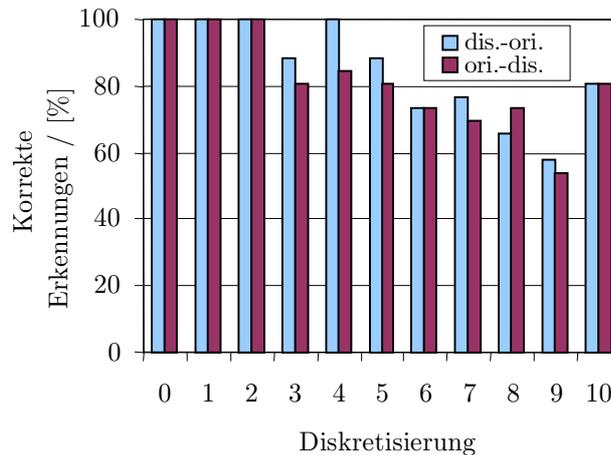


**Abbildung 3.12:** Beispieldatensatz zur Untersuchung der Diskriminierbarkeit von Vektorpolygonen unterschiedlichen Generalisierungsgrades mittels Kreuzmoment-Invarianten.

Für den Autokorrelationsfall der Originalpolygone ergeben sich die in Abbildung 3.13 dargestellten Ergebnisse. Im linken Teil ist die Autokorrelationsmatrix zu sehen, der rechte Teil zeigt in schwarz die Paare mit der geringsten quasi-euklidischen Distanz. Wie in diesem Autokorrelationsfall zu erwarten ist, bilden sie die Hauptdiagonale der Korrelationsmatrix. In Abbildung 3.14 ist eine Übersicht über die Ergebnisse der Korrelationsuntersuchungen mit variabler Generalisierung gegeben. Wie daraus ersichtlich ist, werden ab Generalisierungen der Stufe 3 nicht mehr 100% der Buchstabenpaare richtig erkannt. Die Anzahl korrekter Ergebnisse ist jedoch mit schlechtestenfalls etwa 55% immer noch hinreichend gut, so daß die Verwendbarkeit von CWM-Invarianten zur Objekterkennung gegeben ist. Das implementierte Verfahren ist ohnehin so ausgelegt, daß nur ein Teil der tatsächlich möglichen Zuordnungen zur Konvergenz ausreicht. Die Ergebnisse dieser Untersuchungen sind im Detail in Anhang C zu finden.



**Abbildung 3.13:** Autokorrelationsmatrix (links) und Paare mit minimaler quasi-euklidischer Distanz im 6D-Merkmalraum. Verwendet wurden die sechs einfachsten Kreuzmoment-Invarianten.



**Abbildung 3.14:** Ergebnisse der Untersuchungen zur Diskriminierbarkeit mit CWM-Invarianten am Beispiel des Vektoralphabets in Originalform und in Diskretisierungen der Stufen 1-10.

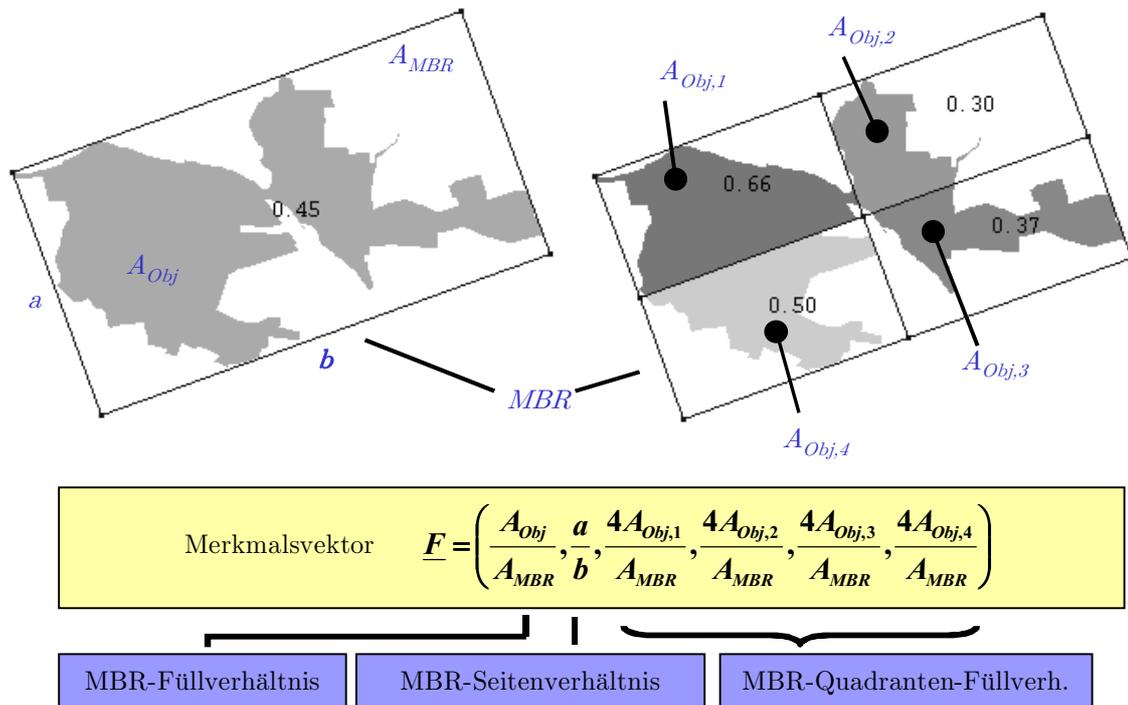
## 3.4 Weitere Merkmale von Polygonobjekten

### 3.4.1 MBR-basierte Merkmale

Im Rahmen dieser Arbeit wurde zusätzlich zu affinen Invarianten aus Kreuzmomenten und aus geometrischen Momenten eine Gruppe alternativer Invarianten untersucht, die im folgenden näher dargestellt werden.

Für jedes flächenhafte Objekt mit endlicher Ausdehnung läßt sich ein Rechteck bestimmen, das das Objekt ganz enthält und eine minimale Fläche besitzt. Um eine Unabhängigkeit von der Wahl der Koordinatenachsen und damit Invarianz bezüglich Rotation zu ermöglichen, wird dabei keine Einschränkung (z.B. Achsparallelität) vorgenommen. Ein solches Rechteck bezeichnet man als minimal berandendes Rechteck (im weiteren Verlauf als MBR bezeichnet).

Ein flächenhaftes Objekt und dessen MBR ist in Abbildung 3.15, links, abgebildet. Das rechte Teilbild zeigt das selbe Objekt mit dem MBR und seinen vier gleich großen Quadranten. Die Verwendung von MBRs besitzt einen wesentlichen Vorteil für die hybride Zuordnung von Raster- und Vektordaten: Ein MBR läßt sich in gleicher Weise für flächenhafte Vektor- und für Rasterdatenobjekte finden. Für eine allgemeingültige Formulierung ist jedoch dem Autor keine geschlossene Lösung bekannt. Vielmehr werden MBRs durch ein iteratives Verfahren



**Abbildung 3.15:** Minimal berandendes Rechteck (MBR) und Quadranten für ein flächenhaftes Objekt. Der sechsdimensionale Merkmalsvektor setzt sich aus den einzelnen MBR-basierten Merkmalen zusammen.

aufgesucht. Dazu wird zunächst ein Winkelinkrement bestimmt, beispielsweise  $5^\circ$ . Ausgehend von einem Startwinkel, in der Regel  $0^\circ$ , wird für jede Winkellage, die einem ganzen Vielfachen des Winkelinkrements entspricht, das MBR durch Berechnung der vier Seiten aufgebaut und die für diese Winkellage minimale Fläche berechnet. Das MBR, das bei Betrachtung aller erlaubten Winkellagen die minimale Fläche besitzt, ist das MBR des Objekts. Zu bemerken sind zwei Punkte: Zum einen steigt der Rechenaufwand für die MBR-Berechnung linear mit der Zahl der Polygonecken und reziprok mit der Größe des Winkelinkrements. Zum anderen kann in sehr seltenen Fällen abhängig von der verwendeten Rechengenauigkeit kein eindeutiges MBR existieren. Für diesen Fall muß eine Regel implementiert sein, welches der gleichwertigen MBRs zu wählen ist, beispielsweise das erste gefundene.

Aufgrund ihrer Berechenbarkeit für Vektor- und Rasterdatenobjekte lassen sich Merkmale, die man auf Basis des Objekts und seines MBRs, beziehungsweise Teilen seines MBRs, berechnet, zur hybriden Objekterkennung benutzen. Ein einfach zu berechnendes Merkmal ist das Verhältnis von Objektfläche zu MBR-Fläche, in Abbildung 3.15 (links) durch die Zahl im MBR angegeben. Der Wert dieses Merkmals bewegt sich immer zwischen 0 und 1, ist also von Natur aus bereits normiert. Dieses MBR-Füllverhältnis ist aufgrund der Natur des MBR unabhängig von anisotroper Skalierung, von Translation in der  $xy$ -Ebene und von Rotation. Eine mathematisch fundierte Ableitung dieser Invarianzen ist aufgrund der iterativen Natur der MBR-Bestimmung nicht möglich, die evidenten Aussagen lassen sich jedoch durch explizite Berechnung leicht verifizieren.

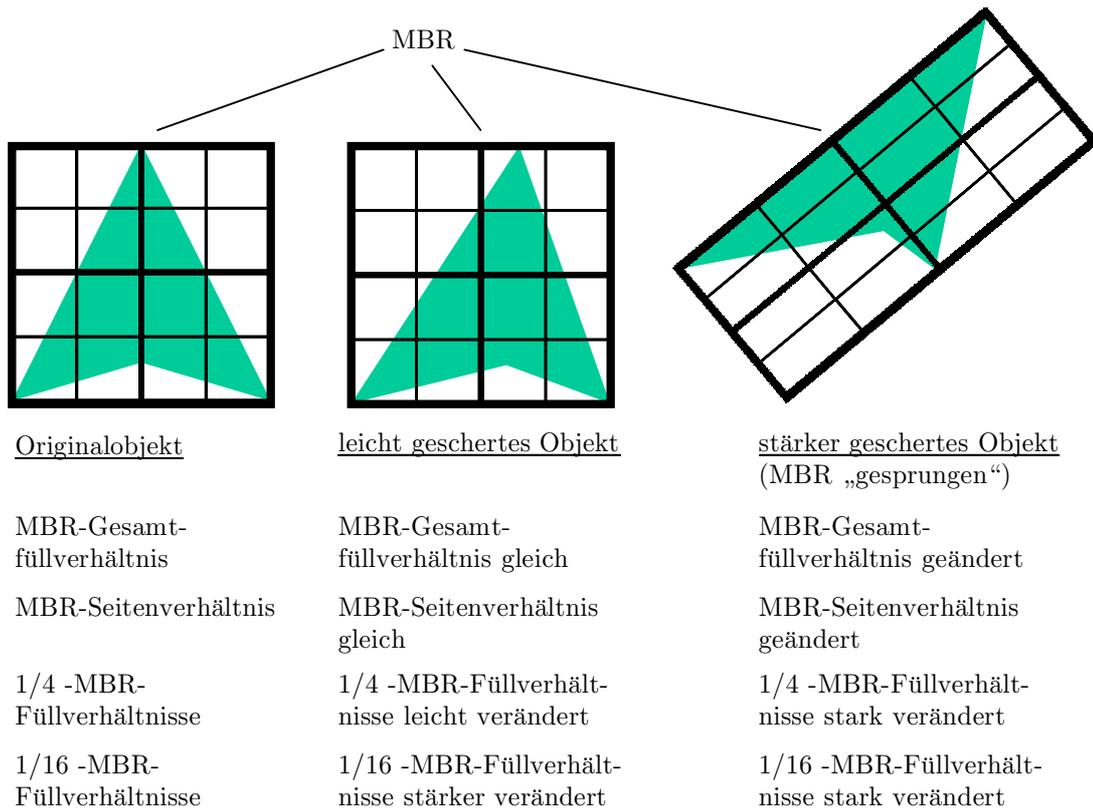
Nun sind diese Invarianzen einerseits zwar sehr nützlich, andererseits können jedoch Objekte mit sehr unterschiedlicher Form exakt das selbe MBR-Füllverhältnis besitzen. Zudem ist die Lagebestimmung eines Objekts aus dem MBR nur modulo  $180^\circ$ , bei quadratischem MBR sogar modulo  $90^\circ$  eindeutig. Zur besseren Diskriminierung läßt sich aus dem MBR selbst ein weiteres Merkmal definieren: Das MBR-Seitenverhältnis von kürzerer zu längerer Seite des MBR. Der Wert dieses Merkmals ist von Natur aus auf das Intervall  $[0, 1]$  normiert und ist unabhängig vom MBR-Füllverhältnis. Auch hier läßt sich durch explizites Berechnen oder aus den selben Evidenzbetrachtungen wie oben zeigen, daß die selben Invarianzen wie für das MBR-Füllungsverhältnis bestehen. Zur weiteren Verbesserung der Unterscheidbarkeit zwischen einzelnen Objekten lassen sich basierend auf den Quadranten des MBR vier zusätzliche Merkmale definieren. Dazu wird das MBR achsensymmetrisch (oder diagonalensymmetrisch) in vier gleich große Quadranten zerlegt. Für jeden der Quadranten läßt sich dann das Verhältnis von Objektfläche im Quadranten und der Fläche des Quadranten ( $\frac{1}{4}$  der MBR-Fläche) bilden. Da die Lage und Größe der Quadranten durch das MBR und die Bildungsvorschrift (seitenparallel oder diagonal, flächengleich) festliegt, weisen auch die Füllverhältnisse der MBR-Quadranten alle bereits beschriebenen Invarianzen auf. Zur Berechnung der Füllverhältnisse der MBR-Quadranten muß allerdings das Objekt mit

Merkmal	Invarianz unter					
	x-Transl.	y-Transl.	x-Streck.	y-Streck.	Rot.	Scher.
MBR-Füllverhältnis	✓	✓	✓	✓	✓	–
MBR-Seitenverhältnis	✓	✓	✓	✓	✓	–
MBR-Quadr.-Füllverhältnis.	✓	✓	✓	✓	✓	–

**Tabelle 3.3:** Invarianzen MBR-basierter Merkmale.

den Quadranten verschnitten werden, was einen erhöhten Rechenaufwand bedeutet. In Tabelle 3.3 sind die Merkmale aus MBRs und ihre Invarianzen noch einmal zusammengefasst.

Durch fortgesetzte Vierteilung jedes Quadranten, im Extremfall bis hinab zur Pixelgröße, ließen sich in der selben Weise noch etliche weitere Teil-Füllverhältnisse berechnen und somit Mehrdeutigkeiten in der Objektbeschreibung durch einen höherdimensionalen Merkmalsvektor immer weiter reduzieren. Da aber jedes merkmalsbasierte Zuordnungsverfahren eine gewisse Toleranz für jedes Merkmal besitzt, ist es sinnlos, die Merkmalsgenauigkeit wesentlich höher als die angewandte Toleranz zu legen. Im Fall der Teil-Füllverhältnisse gibt es für die beliebige affine Transformation, die zusätzlich eine Scherung beinhaltet, noch einen weiteren Grund, die Berechnung von Teilfüllverhältnissen nicht zu weit zu treiben. Beinhaltet die Transformation eine Scherung, so bleibt zwar bis zu einem gewissen Grad das Gesamtfüllverhältnis invariant, die Teil-Füllverhältnisse können sich aber empfindlich ändern und taugen daher nicht mehr unbedingt als invariante Merkmale. Diese Empfindlichkeit ist umso höher, je weiter die Teilung fortgeschritten ist. Aus diesem Grund werden hier nur die ersten Quadranten betrachtet. In Abbildung 3.16 wird dieser Sachverhalt anhand eines Beispiels illustriert.

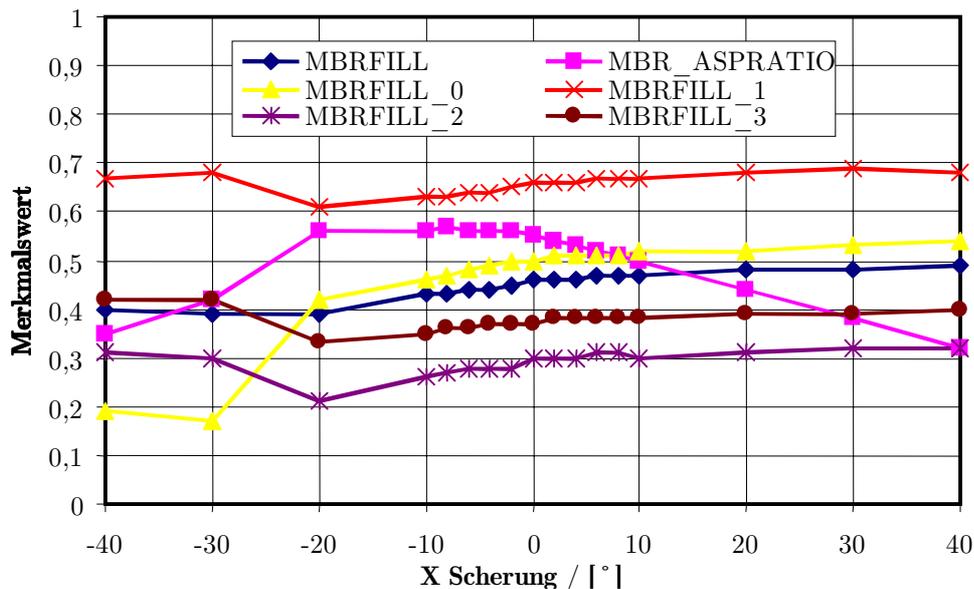


**Abbildung 3.16:** Teilweise bis vollständige Aufhebung der Invarianz MBR-basierter Merkmale unter Scherung.

Bei größeren Scherungen kann sich auch das Gesamtfüllverhältnis ändern, da sich die Lage des MBR relativ zum Objekt dann möglicherweise ändert. Ab diesem Punkt sind sowohl alle Füllverhältnisse wie auch das MBR-Seitenverhältnis nicht mehr als invariante Merkmale zu gebrauchen.

Zur Untermauerung der eben angestellten Betrachtungen wurde ein Satz von GIS-Objekten variablen Scherungen unterzogen und jeweils die sechs genannten MBR-basierten Merkmale berechnet. Die jeweilige Standardabweichung ist in Abbildung 3.17 in Abhängigkeit vom Scherwinkel abgetragen, wobei ein Scherwinkel von  $0^\circ$

senkrecht aufeinander stehende Koordinatenachsen bedeutet.



**Abbildung 3.17:** Verhalten MBR-basierter Merkmale unter Scherung der Koordinatenachsen. Die Abszisse gibt den Scherungswinkel an, die Ordinate die Standardabweichung der Merkmalswerte.

### 3.4.2 Fourier-Deskriptoren

Eine andere Art, zweidimensionale Objekte anhand ihrer Randkurve invariant zu beschreiben, stellen die sogenannten Fourier-Deskriptoren dar. Der Berechnung von Fourier-Deskriptoren liegt die Idee zugrunde, daß die geschlossene Randkurve eines 2D-Objekts durch eine periodische Funktion eines kontinuierlichen Parameters beziehungsweise durch die Fourier-Koeffizienten dieser periodischen Funktion dargestellt werden kann. Die so erhaltenen Fourier-Koeffizienten bezeichnet man als Fourier-Deskriptoren. Um zu einer invarianten Objektbeschreibung zu gelangen, muß zunächst die Randkurven-Parametrisierung auf die entsprechende Transformationsklasse normalisiert werden. Die Normalisierung kann im Orts- oder Frequenzraum erfolgen. Bei Durchführung der Normalisierung im Ortsraum werden Ableitungen der Randkurve benötigt, was bei Polygonen an den Ecken zu Schwierigkeiten führt. Von daher ist die Normalisierung im Frequenzraum zu bevorzugen. Bei exakter Normalisierung sind die Fourier-Deskriptoren unter der gegebenen Transformation invarianten. Für die Klasse der Ähnlichkeitstransformationen (Translation, Rotation, isotroper Maßstab) wurden in der Vergangenheit zahlreiche Untersuchungen zur Objekterkennung mit Fourier-Deskriptoren durchgeführt [Dudani et al. 1977], [Reeves 1988]. Arbter et al. [1990] leiten affin-invariante Fourier-Deskriptoren ab und benutzen sie als (näherungsweise invariante) Merkmale zur Silhouetten-Beschreibung von im 3D-Raum rotierten Flugzeugen. Diese Betrachtung ist ähnlich gelagert wie die Annahme in der vorliegenden Arbeit, die zentralperspektive Abbildung durch eine affine Transformation anzunähern. Infolge der nur näherungsweisen Invarianz der Fourier-Deskriptoren sind die Merkmalsvektoren verschiedener Ansichten des selben Objekts zwar im Merkmalsraum nicht exakt identisch, die Untersuchungen in Arbter et al. [1990] zeigen jedoch bei synthetischen Daten für affin-invariante Fourier-Deskriptoren eine deutlich klarere Clusterung als für Fourier-Deskriptoren der Ähnlichkeitstransformation. In Reiss [1993], Abschnitt 4.5, werden Ergebnisse aus einem Vergleich zwischen der Buchstabenerkennung mit Fourier-Deskriptoren und mit Momentinvarianten präsentiert. Dabei zeigen die Momentinvarianten das bessere Diskriminierungsverhalten. Dies ist ein Grund, warum für diese Arbeit Fourier-Deskriptoren nicht untersucht wurden. Ein weiterer Grund ist durch die hybride Datensituation gegeben. Fourier-Deskriptoren eines Polygons, das aus Bildpixeln aufgebaut ist, unterscheiden sich aufgrund der höheren Frequenzen auf der Randkurve von Natur aus von denjenigen eines "echten" Vektorpolygons und bringen daher a priori eine weitere Ungenauigkeit mit ins Spiel, die den ohnehin bestehenden Nachteil noch vergrößert.

## Kapitel 4

### Szene-Modell-Zuordnung

Die automatische Zuordnung von Sensor- zu Modelldaten, z.B. Fernerkundungsdaten zu digitalen Karten, läßt sich grob in zwei Bereiche unterteilen. Zum einen gilt es, das Korrespondenzproblem zu lösen, d.h. homologe Objekt- oder Punktpaare zu finden. Objekte können dabei allgemein linien- oder flächenhafter Natur oder 3D-Objekte sein. Zum anderen muß im Rahmen der gegebenen Freiheitsgrade eine Transformation berechnet werden, die die Sensordatenobjekte konsistent in das Modellkoordinatensystem abbildet. Letzteres Problem wird in der Fachliteratur auch als "pose estimation" bezeichnet, was bedeutet, daß sowohl die *Lage* wie auch die *Orientierung* der Sensordatenobjekte bestimmt werden muß. Etwas formaler läßt sich das Problem folgendermaßen ausdrücken: Es sei  $\{m_i \mid 1 \leq i \leq M\}$  ein Satz von Modellobjekten und  $\{s_j \mid 1 \leq j \leq S\}$  ein Satz von Objekten im Sensorsystem. Die Transformation  $\mathcal{T} : \mathcal{S} \rightarrow \mathcal{M}$  bilde das Sensor- in das Modellkoordinatensystem ab. Unter einer Korrespondenz  $I$  versteht man eine Teilmenge der  $M \cdot S$  möglichen Paarungen. Dabei sind alle Paare einer Korrespondenz dahingehend konsistent, daß sie durch eine gemeinsame Transformation  $\mathcal{T}_I$  näherungsweise aufeinander abgebildet werden. Die Qualität dieser Abbildung  $\mathcal{T}_I$  wird durch ein Qualitätsmaß  $\rho(\mathcal{T}_I s_j, m_i)$  definiert. Für den Fall flächenhafter Objekte wäre ein geeignetes Qualitätsmaß etwa der prozentuale Überlapp des transformierten Sensorobjekts  $\mathcal{T}_I s_j$  und des Modellobjekts  $m_i$  oder ein Formfaktor, der in irgendeiner Weise die Randpolygone des Paares vergleicht. Die Korrespondenz  $I$  beinhaltet demnach alle Paare  $(m_i, s_j)$ , für die das Qualitätsmaß  $\rho$  einen gewissen Schwellwert  $\epsilon$  unterschreitet:

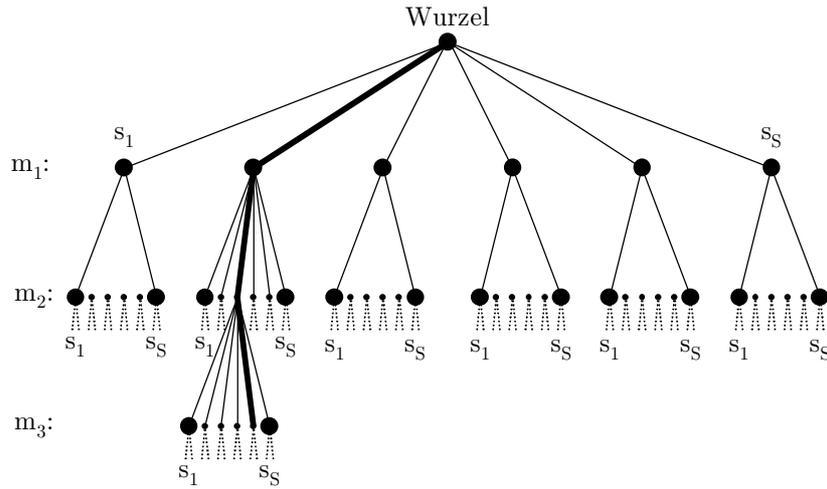
$$I = \{(m_i, s_j) \mid \rho(\mathcal{T}_I s_j, m_i) \leq \epsilon\}. \quad (4.1)$$

In dieser formalen Beschreibung besteht dabei das Ziel des gesamten Zuordnungsprozesses darin, die Korrespondenz mit der maximalen Zahl an Paaren zu finden. Aufgrund der zu erwartenden hohen Anzahl sowohl von Sensor- wie auch von Modellobjekten - und damit einhergehend eines hohen potenziellen Suchaufwands - wird hier wie in Anlehnung an Sester et al. [1998] ein hybrides Verfahren vorgeschlagen, welches aus einer Korrespondenz mit einer geringen Zahl von Paaren eine Näherungstransformation ermittelt, anhand derer dann alle anderen möglichen Paare verifiziert oder verworfen werden. Bevor das implementierte Verfahren näher beschrieben wird, widmet sich der folgende Abschnitt zunächst einmal der Lösung des Korrespondenzproblems.

#### 4.1 Lösung des Korrespondenzproblems anhand des Interpretationsbaums

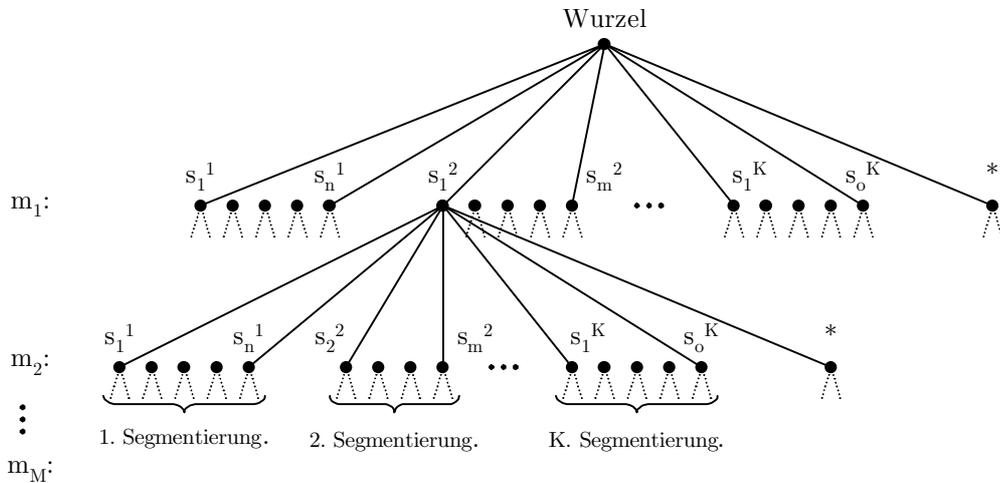
Das Ziel des Zuordnungsverfahrens besteht darin, richtige Korrespondenzen zu finden ohne dabei den gesamten Suchraum zu durchsuchen. Von Grimson et al. [1993] wird vorgeschlagen, dazu einen sogenannten Interpretationsbaum zu benutzen. In Abbildung 4.1 ist der Suchbaum für alle möglichen 2er-Paarungen zwischen  $S$  Sensor- und  $M$  Modellobjekten illustriert. Der im Suchbaum in Abbildung 4.1 verstärkt gezeichnete Pfad repräsentiert eine Interpretation, also eine bestimmte Korrespondenz. Im dargestellten Fall besteht diese Korrespondenz aus den drei Paaren  $(m_1, s_2)$ ,  $(m_2, s_3)$  und  $(m_3, s_5)$ . Im allgemeinen spricht man bei jeder Korrespondenz, die bis zur Tiefe  $n$  reicht von einer  $n$ -Interpretation, wobei die maximale Tiefe der Anzahl  $M$  der Modellobjekte entspricht:  $n_{\max} = M$ . Es ist zu beachten daß in jeder Ebene alle  $S$  Modellobjekte mögliche Partner für eine Zuordnung sind, es sind also zunächst Mehrfachzuordnungen möglich. Die Suche nach konsistenten  $n$ -Interpretationen erfolgt nun, von der Wurzel aus beginnend, anhand dieses Interpretationsbaums. Für jeden möglichen Ast des Baums wird während des Suchprozesses geprüft, ob für den  $n$ -ten Knoten Konsistenz mit den bereits in der aktuellen Interpretation enthaltenen  $n - 1$  Knoten besteht. Konsistenz besteht zum Beispiel dann, wenn sich alle Paare einer aktuellen Korrespondenz durch ein und dieselbe Transformation zumindest näherungsweise ineinander abbilden lassen. Darauf wird speziell im Hinblick auf das Ziel dieser Arbeit in Abschnitt 4.4 näher eingegangen. Die simple Abarbeitung aller möglicher Korrespondenzen stellt sehr schnell ein rechenzeitlich nicht mehr handhabbares Problem dar: Bei  $S$  Sensor- und  $M$  Modellobjekten ergeben sich für den in Abbildung 4.1 aufgebauten Suchbaum  $S^M$  Knoten<sup>1</sup>. Durch Einführung von Zwangsbedingungen läßt sich die Suche jedoch bedeutend verkürzen. Dies ist Gegenstand des folgenden Abschnitts. Wie in Kapitel 2 beschrieben wurde, werden im Rahmen dieser Arbeit Objekte in Sensordaten durch ein Segmentierungsverfahren gewonnen. Um möglichst unabhängig von radiometrischen Gegebenheiten zu sein, werden dabei mehrere Segmentierungsläufe mit unterschiedlichen Parametern durchgeführt. Konsistente Interpretationen müssen daher über die Objekte aller Segmentierungsläufe gesucht werden. Der Suchbaum in Abbildung 4.1 stellt lediglich die möglichen Paarungen

<sup>1</sup>Zur Behandlung des in der Praxis üblichen Falls, daß für ein Sensorobjekt kein passendes Modellobjekt existiert, wird in der Regel den Modellobjekten ein sog. Nullobjekt hinzugefügt. Dadurch erhöht sich die Zahl der Knoten im Suchbaum auf  $(S + 1)^M$ .



**Abbildung 4.1:** Suchbaum für alle möglichen 2er-Paarungen zwischen  $S$  Sensorobjekten  $s_j$  und  $M$  Modellobjekten  $m_i$ . Der dicker eingezeichnetete Pfad stellt eine mögliche Korrespondenz bestehend aus den Paaren  $(s_2, m_1), (s_3, m_2), (s_5, m_3), \dots$  dar.

für einen Segmentierungslauf dar. Führt man insgesamt  $K$  Segmentierungen durch, so bezeichne  $s_i^k$  das  $i$ -te Sensorobjekt der  $k$ -ten Segmentierung. Der entsprechend erweiterte zu durchsuchende Baum ist in Abbildung 4.2 dargestellt. Zusätzlich enthält der erweiterte Baum in jeder Ebene ein Nullelement (\*), das ein vorzeitiges Abbrechen der Suche verhindert, indem es einem Modellelement, für das kein passendes Sensordatenelement gefunden wurde, zugeordnet wird.



**Abbildung 4.2:** Erweiterter Suchbaum für  $M$  Modelldatenobjekte und alle Sensordatenobjekte aus  $K$  Segmentierungen. In jeder Ebene ist zusätzlich ein Nullelement (\*) eingeführt, um einen vorzeitigen Abbruch der Baumsuche zu verhindern.

## 4.2 Zwangsbedingungen zur Verkleinerung des Suchraums

Die Suche nach einer konsistenten  $n$ -Interpretation beginnt an der Wurzel des Suchbaums in Abbildung 4.2. In jeder Ebene des Suchbaums werden die Knoten auf Konsistenz mit der bereits vorliegenden  $(n-1)$ -Korrespondenz geprüft. In dieser Arbeit wird die sogenannte Tiefensuche verwendet ("depth-first-search"), d.h. es wird zuerst ein Ast in die Tiefe verfolgt. Alternativ dazu gibt es die Breitensuche, bei der zuerst alle Knoten einer Ebene bearbeitet werden. Durch Einführung von Zwangsbedingungen, die für eine konsistente Interpretation erfüllt sein müssen, lässt sich an jedem Knoten prüfen, ob der zugehörige Ast in die Korrespondenz aufgenommen

wird. Würde die Aufnahme zu einer Inkonsistenz führen, so wird der Ast verworfen, der gesamte Sub-Baum unterhalb dieses Knotens muß nicht mehr durchsucht werden. Die Konsistenzprüfung ist in der Regel relativ aufwendig, daher sucht man einfach abzuprüfende Zwangsbedingungen, mit denen nicht passende Knoten schnell identifiziert werden können. Es bezeichne  $i$  die Ebene des aktuellen Knotens im Suchbaum und damit das Sensorobjekt  $s_i$ ,  $p$  den Ast zwischen aktuellem Knoten und dem Knoten der vorigen Ebene  $i - 1$  und damit das Modellobjekt  $m_p$ ,  $j$  eine beliebige Ebene oberhalb von  $i$  und  $q$  den  $j$  zugeordneten Ast. Je nach Zahl der Knoten, die berücksichtigt werden, unterscheidet man folgende Arten von Zwangsbedingungen [Grimson et al. 1993]:

- **unäre Zwangsbedingung:**  $u(i, p)$ : Diese Bedingung liefert ein logisches "Wahr", wenn die Zuordnung des  $i$ -ten Sensorobjekts  $s_i$  zum  $p$ -ten Modellobjekt  $m_p$  konsistent ist.
- **binäre Zwangsbedingung:**  $b(i, j, p, q)$ : Diese Bedingung liefert ein logisches "Wahr", wenn die Zuordnung des  $i$ -ten Sensorobjekts  $s_i$  zum  $p$ -ten Modellobjekt  $m_p$  und gleichzeitig die Zuordnung des  $j$ -ten Sensorobjekts  $s_j$  zum  $q$ -ten Modellobjekt  $m_q$  konsistent ist.
- **$n$ -näre Zwangsbedingung:**  $n(i, j, \dots, p, q, \dots)$ : Diese Bedingung liefert ein logisches "Wahr", wenn gleichzeitig alle  $n$  paarweisen Zuordnungen von  $\{s_i, s_j, \dots\}$  zu  $\{m_p, m_q, \dots\}$  konsistent sind.

Im Rahmen dieser Arbeit wurde zur Baumsuche das sogenannte "Backtracking"-Verfahren implementiert, das in [Grimson et al. 1993] beschrieben ist. Dabei werden am aktuellen Knoten  $i$  die Zwangsbedingungen geprüft. Ergeben alle Zwangsbedingungen ein logisches "Wahr", so wird gemäß der "depth-first"-Strategie der am weitesten links liegende Ast zur nächsten Ebene  $i + 1$  des Suchbaums inspiziert. Wird dabei ein Knoten erreicht, an dem die Zwangsbedingungen nicht erfüllt sind (mindestens eine Zwangsbedingung, erzeugt ein logisches "Falsch"), so wird der gesamte darunter liegende Unterbaum aus dem Suchraum entfernt. Im "backtracking"-Schritt wird, ausgehend vom vorigen Knoten  $i$ , der rechts benachbarte Ast inspiziert. Existieren auf einer Ebene des Suchbaums keine weiteren Äste mehr, so wird ein weiterer "backtracking"-Schritt zur nächsthöheren Ebene unternommen. Wird ein Blatt, d.h. die tiefstmögliche Ebene  $S$ , erreicht und liefern alle Zwangsbedingungen Konsistenz, so ist eine mögliche  $n$ -Interpretation gefunden, die gespeichert wird. Mit einem "backtracking"-Schritt wird die Suche fortgesetzt bis der gesamte Baum durchsucht ist und alle möglichen Interpretationen gefunden sind. Dieses Vorgehen liefert alle möglichen  $S$ -Interpretationen. Zur Verkürzung des Suchverfahrens, insbesondere bei vielen Sensor- und Modellobjekten, ist es sinnvoll, bereits Interpretationen, die nicht bis zu einem Blatt reichen, zu akzeptieren, wenn sie für das spezielle Suchproblem hinreichend konsistent sind. Bei Verwendung von Null-Elementen wird üblicherweise die aktuell beste Interpretation gespeichert, da infolge von Null-Zuordnungen nicht sicher ist, daß eine  $S$ -Interpretation auch tatsächlich  $S$  Modellobjekte enthält. Mit anderen Worten: Die Zahl der Sensorobjekte in einer konsistenten Interpretation, die tatsächlich einem Modellobjekt zugeordnet werden können, ist nicht von vornherein bekannt, sie kann auch kleiner als  $S$  sein. In der Literatur werden verschiedene Verfahren zur Verbesserung des Suchverhaltens des Backtrackingverfahrens angegeben, die jedoch in dieser Arbeit keinen Niederschlag finden [Haralick & Elliott 1980].

## 4.3 Die implementierten Zwangsbedingungen

Im folgenden werden die für diese Arbeit implementierten unären und binären Zwangsbedingungen näher beschrieben. Um der Invarianzmaxime Rechnung zu tragen, werden ausschliesslich Zwangsbedingungen verwendet, die auch invariant formuliert werden können. Zur Minimierung der Rechenzeit wurden dabei möglichst einfache Zwangsbedingungen gesucht, die der Reihe nach von der einfachsten bis hin zur aufwendigsten geprüft werden. Diese Strategie führt dazu, daß die aufwendigeren Operationen während der Suche nur relativ selten durchzuführen sind, da in den meisten Fällen bereits die einfacheren Zwangsbedingungen die Inkonsistenz des aktuellen Knotens aufzeigen.

### 4.3.1 Mindestgröße (unär)

Sensordatenobjekte, die unterhalb einer gewisse Mindestgröße liegen, werden für die Suche nach konsistenten Interpretationen nicht berücksichtigt. Der Schwellwert wird jedoch nicht absolut, sondern relativ zur Bildgröße formuliert, um eine Maßstabsunabhängigkeit zu gewährleisten. Diese Zwangsbedingung ist die am einfachsten prüfbar, da die Objektfläche in Form des Moments  $m_{00}$  bereits vorliegt. Sie wird daher zuerst geprüft.

### 4.3.2 Beziehung der Invariantenvektoren (unär)

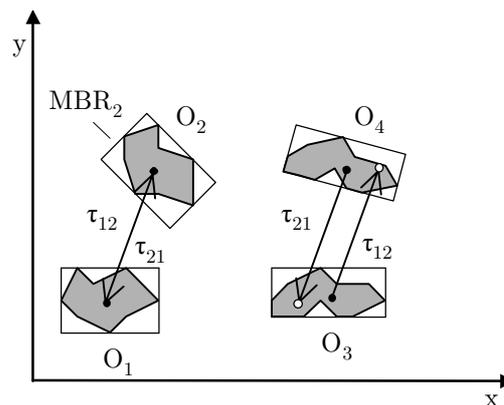
Wie in Kapitel 3 beschrieben wurde, lassen sich 2D-Polygone durch invariante Merkmale, die aus zentralen geometrischen Momenten oder Kreuzmomenten berechnet werden, klassifizieren. Mit den angestellten Untersuchungen wurde eine hinreichende Robustheit gegen statistisches Rauschen und Diskretisierungsrauschen sowie eine hinreichende Diskriminierbarkeit mittels affiner Invarianten festgestellt. Eine unäre Zwangsbedingung lässt sich damit aus der Forderung formulieren, dass sich homologe Objekte in ihren affinen Invarianten nur zu einem gewissen Grad unterscheiden dürfen. Dies setzt allerdings voraus, daß die Objekte in beiden Datensätzen vollständig vorhanden sind, da die hier benutzten invarianten Merkmale globaler Natur sind. Die Sensorobjekte  $s_i$  seien durch die sechs einfachsten affinen Invarianten aus Kreuzmomenten (siehe Gleichung 3.40) repräsentiert:  $s_i : \{\Phi_{s,i}(2,4), \Phi_{s,i}(2,6), \Phi_{s,i}(2,8), \Phi_{s,i}(4,6), \Phi_{s,i}(4,8), \Phi_{s,i}(6,8)\}$ . In gleicher Weise seien die Modellobjekte repräsentiert durch  $m_j : \{\Phi_{m,j}(2,4), \Phi_{m,j}(2,6), \Phi_{m,j}(2,8), \Phi_{m,j}(4,6), \Phi_{m,j}(4,8), \Phi_{m,j}(6,8)\}$ . Zur Bewertung der Ähnlichkeit zweier Invarianten wurde bereits mit Gleichung 3.56 ein Kriterium definiert, das sich auf den mit Gleichung 3.55 eingeführten relativen Unterschied zweier reeller Zahlen stützt. Benützt man, wie in dieser Arbeit, das Kriterium aus Gleichung 3.56, so ist das Verfahren analog zur Box-Klassifikation multispektraler Fernerkundungsdaten. Die linke Seite der Gleichung 3.57 läßt sich auch als quasi-euklidische Distanz im 6D-Merkmalraum verstehen. Benützt man dieses Kriterium, so ist das Vorgehen analog zu der in der Fernerkundung häufig angewandten Minimum-Distance-Klassifikation mit maximaler Varianz pro Cluster, bei der kugelförmige Cluster gebildet werden (z.B. [Lillesand & Kiefer 1987], [Sabins 1996] oder [Schowengerdt 1997]).

### 4.3.3 Schwerpunkte und MBRs (binär)

Eine binäre Zwangsbedingung wird durch die Forderung, daß sich mit der Hin- und Rücktransformation eines Paares die Schwerpunkte eines anderen Paares jeweils in das minimal berandende Rechteck (MBR, siehe Abschnitt 3.4.1) transformieren lassen, dargestellt. Wird die Forderung von beiden Paaren erfüllt, so ist diese binäre Zwangsbedingung befriedigt. Mit  $(O_1, O_2)$  und  $(O_3, O_4)$  seien zwei Objektpaare gegeben. Die Transformationen der Objekte eines Paares aufeinander seien durch  $\mathcal{T}_{ij} : O_i \rightarrow O_j$  gegeben. Der Operator  $CM(O_i)$  bilde den Schwerpunkt des Objekts  $O_i$ . Falls der Punkt  $CM$  innerhalb eines Polygons  $P$  liegt, so gelte:  $CM \in P$ . Das MBR des Objekts  $O_i$  sei  $MBR_i$ . Diese binäre Zwangsbedingung ist damit erfüllt, wenn der Ausdruck

$$(CM(\mathcal{T}_{34}(O_1)) \in MBR_2 \wedge CM(\mathcal{T}_{12}(O_3)) \in MBR_4) \quad (4.2)$$

ein logisches "Wahr" ergibt. Durch Vertauschung der Paare (1,2) und (3,4) läßt sich in der selben Weise eine weitere binäre Zwangsbedingung formulieren. Zur Veranschaulichung dieser binären Zwangsbedingung möge Abbildung 4.3 dienen.



**Abbildung 4.3:** Veranschaulichung der binären Zwangsbedingung, daß die Schwerpunkte eines Objektpaares durch die entsprechende Transformation des anderen Paares jeweils innerhalb der entsprechenden MBRs liegen.

### 4.3.4 Objektüberlappung (binär)

Eine weitere binäre Zwangsbedingung besteht darin, daß die Schnittmenge der beiden Objekte eines Paares nach Anwendung einer affinen Näherungstransformation auf eines der Objekte mindestens einen gewissen Prozentsatz

beider Einzelflächen betragen muß. Die anzubringende Transformation wird dabei vom jeweils anderen Paar über dessen Standardlagen (s. Abschnitt 4.5) bereitgestellt. Etwas formaler läßt sich diese Zwangsbedingung folgendermaßen formulieren. Es seien  $(O_1, O_2)$  und  $(O_3, O_4)$  zwei Objektpaare wobei jeweils das erste Objekt aus dem Sensor- und das zweite aus dem Modelldatensatz stamme. Wie in Abschnitt 4.3.3 transformiere  $T_{ij} : O_i \rightarrow O_j$  das Objekt  $O_i$  in das Objekt  $O_j$ . Der Operator  $\mathcal{A}(O_i)$  bilde die Fläche des Objekts  $O_i$ , das Schnittobjekt zweier Objekte  $O_i$  und  $O_j$  sei gegeben durch  $\mathcal{S}(O_i, O_j) = O_i \wedge O_j$ . Dann muß folgender Ausdruck zur Erfüllung dieser Zwangsbedingung ein logisches "WAHR" ergeben:

$$\min \left( \frac{\mathcal{A}(\mathcal{S}(T_{34}(O_1), O_2))}{\mathcal{A}(T_{34}(O_1))}, \frac{\mathcal{A}(\mathcal{S}(T_{34}(O_1), O_2))}{\mathcal{A}(O_2)} \right) > \delta \wedge \min \left( \frac{\mathcal{A}(\mathcal{S}(T_{12}(O_3), O_4))}{\mathcal{A}(T_{12}(O_3))}, \frac{\mathcal{A}(\mathcal{S}(T_{12}(O_3), O_4))}{\mathcal{A}(O_4)} \right) > \delta . \quad (4.3)$$

Die gleiche Bedingung könnte auch durch gleichzeitiges Vertauschen der Indices 1 und 2 sowie 3 und 4 aufgestellt werden. Für die festzulegende prozentuale Objektüberlappung  $\delta$  gilt dabei:  $\delta \in [0, 1]$ .

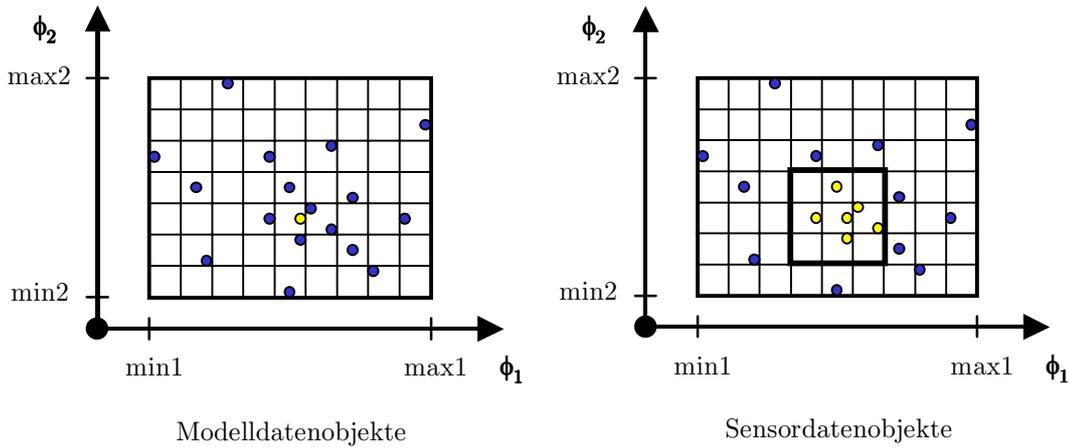
Die exakte Berechnung dieser Zwangsbedingung kann infolge einer großen Zahl an Vertices sehr aufwendig werden. Zur Vereinfachung wird daher die Verschneidung der Objekte in einem normierten Raster durchgeführt. Die Berechnung der Schnittfläche reduziert sich dann auf das Zählen gemeinsamer Pixel. Die Größe des Rasters wird so gewählt, daß der dabei entstehende Fehler kleiner als 1% und somit nicht erheblich ist.

## 4.4 Das gesamte implementierte Suchverfahren

Die in den vorigen Abschnitten beschriebene Baumsuche wird in der Regel für reale Datensätze sehr aufwendig, da infolge vieler Objekte in den Modell- und Sensordaten eine Unzahl möglicher Knoten entsteht. Da zunächst jedes Sensorobjekt in jeder Ebene des Suchbaumes mehrfach auftritt (s. Abbildung 4.2), läßt sich eine Reduktion des Rechenaufwandes erreichen, indem die unären Zwangsbedingungen vor Beginn der Baumsuche geprüft und nicht passende Objekte direkt eliminiert werden. Dadurch reduziert sich einerseits bereits vor der Baumsuche die Zahl der Objekte und damit der Knoten im Baum. Andererseits wird jede unäre Zwangsbedingung mit diesem Vorgehen genau einmal geprüft.

Während die Prüfung der relativen Mindestgröße keiner weiteren Erklärung bedarf, wird auf die Beziehung der Invariantenvektoren im folgenden noch etwas näher eingegangen. Die Zwangsbedingung für die absoluten Affinvarianten ist, wie bereits beschrieben, derart definiert, daß für jede Dimension (d.h. für jede Invariante) der nach Gleichung 3.55 definierte relative Unterschied kleiner als ein gegebener Schwellwert, z.B. 5%, ist. Da zunächst noch keine Objektpaarhypothesen vorliegen, müssen für diesen Schritt alle möglichen 2-er-Paarungen von  $M$  Modell- und  $S$  Sensordatenobjekten betrachtet werden. Auch hier ist durch überlegtes Handeln die Zahl der tatsächlich zu untersuchenden Paare deutlich reduzierbar. Dazu wird folgende Strategie vorgeschlagen: Vor der Prüfung der relativen Unterschiede aller Invarianten eines hypothetischen Paares werden die Objekte in jedem Datensatz für sich invariantenweise kategorisiert. Dazu werden Kategorien in Abhängigkeit des gegebenen Schwellwertes gebildet. Sind z.B. relative Unterschiede von maximal 10% pro Invariante für eine gültige Hypothese erlaubt, so wird der gesamte Wertebereich der betreffenden Invariante aller Objekte aus beiden Datensätzen in 10 gleich große Kategorien unterteilt. Ganz allgemein ist die Größe der einzelnen Kategorien umgekehrt proportional dem Schwellwert für den relativen Unterschied. In einem  $n$ -komplexen Schritt wird für jedes Objekt jede Invariante der entsprechenden Kategorie zugeordnet. Der Vorteil dieses Verfahrens besteht darin, daß zur Suche nach passenden Objekten nur noch Elemente derselben und der benachbarten Kategorie betrachtet werden müssen, da die Objekte aller anderen Kategorien per Definition nicht als gültige Hypothesen in Frage kommen. Aus den jeweils maximal drei in Frage kommenden Kategorien werden der Reihe nach alle möglichen Paare in jeder Invariante geprüft. Ist für eine Invariante eine Zwangsbedingung nicht erfüllt, so wird das Paar nicht weiter untersucht. Sind für die einzelnen Invarianten unterschiedliche Schwellwerte angegeben, so wird die Invariante mit dem kleinsten Schwellwert zuerst geprüft, da diese Zwangsbedingung am strengsten ist und damit so früh wie möglich ungültige Paare identifiziert und entfernt werden. Im Fall des gleichen Schwellwertes für alle Invarianten spielt die Reihenfolge der Prüfung keine Rolle für den Rechenaufwand. Die invariantenweise Einordnung von Sensor- und Modelldatenobjekten in Kategorien und die damit einhergehende Verkleinerung der Suchmenge werden mit Abbildung 4.4 illustriert.

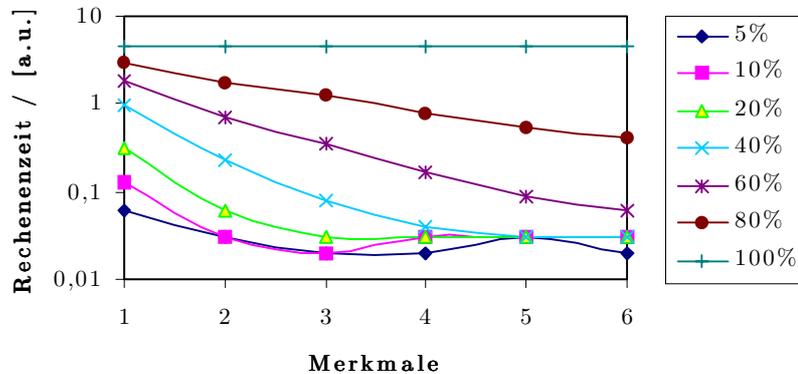
Es liege der Fall vor, daß zu dem hell gezeichneten Objekt des Modelldatensatzes (Abbildung 4.4, linkes Teilbild) alle passenden Sensordatenobjekte gefunden werden sollen. Man bestimmt zunächst für jede Invariante die Kategorie, in der das Objekt liegt (schraffierte Fläche im linken Teilbild) und betrachtet dann alle Objekte des Sensordatensatzes (bzw. aller Segmentierungen desselben), die innerhalb derselben oder einer benachbarten Kategorie liegen (dick eingezeichneter Bereich im rechten Teilbild von Abbildung 4.4). Alle anderen Objekte können die Zwangsbedingung infolge der angepaßten Wahl der Kategorien nicht erfüllen und brauchen daher



**Abbildung 4.4:** Reduktion der Suchmenge durch Einordnung von Objekten in Merkmalskategorien, die abhängig von den verwendeten Schwellwerten automatisch gebildet werden.

gar nicht erst geprüft zu werden. Für alle möglichen Paare muß nun noch explizit der relative Unterschied für jede Invariante berechnet und das Paar gegebenenfalls aussortiert werden. Dieses Verfahren liefert somit alle Sensor-Modell-Objektpaare, die die Zwangsbedingung, daß der relative Unterschied pro Invariante einen gewissen Wert nicht übersteigt, erfüllen.

Der Rechenaufwand des Verfahrens und die damit verbundenen Ersparnisse wurde anhand synthetischer Daten untersucht. Dazu wurden für jeden Meßpunkt zwei Datensätze mit je 400 Objekten erzeugt. Die Merkmalswerte wurden mit einem Pseudo-Zufallszahlengenerator erzeugt. Jedes der Datensatz-Paare wurde mit 1 bis 6 Merkmalen ausgestattet. Variiert wurde der maximale relative Unterschied in einem Bereich zwischen 5% und 100%. Die Ergebnisse des Tests sind in Abbildung 4.5 dargestellt (man beachte den logarithmischen Maßstab für die Rechenzeit !).



**Abbildung 4.5:** Rechenaufwand für die vorgezogene Prüfung der Zwangsbedingung für den relativen Unterschied pro Invariante.

Bei einem relativen Unterschied von 100% (1.0) besteht maximaler Rechenaufwand, da alle möglichen Paarungen (hier: 160 000) bewertet werden müssen. In Abbildung 4.5 ist gut zu erkennen, daß die Vorsortierung der Objekte in adaptive Kategorien eine erhebliche Zeitersparnis mit sich bringt. Im Fall von 5% Toleranz pro Merkmal ist der Rechenaufwand nur etwa 20% des Aufwands ohne Sortierung. Weiterhin ist zu beobachten, daß prinzipiell der Rechenaufwand umso kleiner wird, je mehr Merkmale betrachtet werden. Für relative Unterschiede kleiner als etwa 20% ist diese Tendenz jedoch nicht streng gegeben, dennoch wird der Nutzen der Vorsortierung umso größer, je kleiner der maximal erlaubte relative Unterschied ist und betsätigt somit den Sinn der Vorsortierung, die auch in der beschriebenen Weise implementiert wurde.

Die beschriebene vorgezogene Prüfung des relativen Unterschieds pro Invariante liefert gleichzeitig bereits eine Information für die nachfolgende Baumsuche: Im Suchbaum aus Abbildung 4.2 sind nicht mehr alle Knoten zu inspizieren, sondern lediglich diejenigen, die den Paaren entsprechen, die diese Zwangsbedingung (und zusätzlich

die Zwangsbedingung der minimalen relativen Objektfläche) erfüllen. Aus diesem Schritt lässt sich zudem ein Qualitätsmaß für jedes Paar bestimmen, indem eine normierte euklidische Distanz im Merkmalsraum, den die sechs benutzten absoluten Affininvarianten  $\phi_i$  (s.o.) aufgespannen, gemäß

$$Q = \sqrt{\sum_{i=1}^6 \left( \frac{\phi_{i,1} - \phi_{i,2}}{\phi_{i,\max} - \phi_{i,\min}} \right)^2} \quad (4.4)$$

berechnet wird. Die Werte  $\phi_{i,\max}$  und  $\phi_{i,\min}$  sind die Extremwerte der  $i$ . Invarianten aus beiden Datensätzen. Je kleiner dieses Qualitätsmaß ist, desto näher liegen die Objekte im Merkmalsraum zusammen. Dieses Qualitätsmaß findet in späteren Phasen des Prozesses weitere Verwendung (Abschnitt 4.7).

Nach der vorgezogenen Prüfung der unären Zwangsbedingungen erfolgt die Baumsuche auf den verbleibenden Objektpaaren mit den in Abschnitt 4.3 eingeführten binären Zwangsbedingungen. Die Suche liefert bei erfolgreichem Abschluß einen Satz konsistenter Objektpaare. Aus der so gewonnenen Menge an konsistenten Objektpaaren wird eine globale affine Näherungstransformation bestimmt, die Sensor- und Modelldaten ineinander transformiert. Zur Berechnung der gemittelten Transformation wird das in Gleichung 4.4 definierte Qualitätsmaß, das für jede Hypothese berechnet wurde, zur Gewichtung verwendet. Anschaulicherweise erhält die Transformation eines Objektpaares der konsistenten Interpretation ein um so höheres Gewicht, je kleiner der Wert  $Q$  ist (Gleichung 4.4), d.h. je näher die beiden Objekte des Paares im 6D-Merkmalsraum zusammen liegen. Dahinter steckt die Annahme, daß die absoluten Affininvarianten hinreichend gut zur Objektdiskriminierung geeignet sind, was durch die Untersuchungen aus Abschnitt 3.3.2 bereits belegt wurde. Das Gewicht  $W_k$  des  $k$ -ten Objektpaares wird folgendermaßen aus dem Qualitätsmaß  $Q$  bestimmt:

$$W_k = \begin{cases} \frac{1}{Q_k} & Q_k > 0 \\ 10^6 & Q_k = 0 \end{cases} \quad \text{für} \quad (4.5)$$

Es seien  $a_i^k \forall i \in \{1, 6\}$  die Parameter der Affintransformation des Objektpaares  $k$ . Die Parameter  $\bar{a}_i^k$  der gemittelten Transformation berechnen sich dann zu:

$$\bar{a}_i = \frac{1}{n} \sum_{k=1}^n a_i^k W_k \quad (4.6)$$

Während der Baumsuche liegt diese globale Transformation jedoch noch nicht vor, daher wird jeweils für das erste Paar einer möglichen Interpretation während des Suchprozesses die paarweise Affintransformation als globale Näherungstransformation interpretiert. Zur Bestimmung der paarweisen Affintransformation sei auf Abschnitt 4.5 verwiesen. Erfüllt ein weiteres Paar während des Suchprozesses alle Zwangsbedingungen, so wird es in die aktuelle Interpretation mit aufgenommen und die globale Affintransformation wird in eben beschriebener Weise aus allen Paaren der aktuellen Interpretation gemittelt und dabei aktualisiert.

## 4.5 Bestimmung der affinen Transformation zwischen homologen Objekten

Trotz Kenntnis der aufgrund passender Merkmalsvektoren möglichen Zuordnungen ist zunächst die Transformation vom Sensor- ins Modellkoordinatensystem unbestimmt. In diesem Abschnitt werden Verfahren vorgestellt, die es erlauben, aus den bereits für jedes Objekt berechneten geometrischen Momenten eine sogenannte Standardlage zu berechnen. Eine Standardlage wird durch Normierung bestimmter zentraler Momente willkürlich festgelegt. Über die damit eindeutige Standardlage jedes Objekts läßt sich die gesuchte Transformation für jedes Objektpaar bestimmen. Das Ziel dieses Vorgehens besteht dabei darin, die unvermeidbaren Fehlzuordnungen möglichst sicher und vollständig durch Konsistenzuntersuchungen zu eliminieren [Süße 1999].

Eine ebene affine Transformation ist eindeutig festgelegt durch Angabe dreier jeweils nicht kollinearere Punktpaare. Bei mehr als drei Punktkorrespondenzen ist das entsprechende Ausgleichsproblem zu lösen, z.B. [Albertz & Kreiling 1989]. Im vorliegenden Fall sind jedoch außer den Objektschwerpunkten ohne weiteres Wissen keine korrespondierenden Punkte bestimmbar, so daß eine alternative Lösung zu suchen ist. Betrachtet man das Verhalten der geometrischen Momente unter ebener affiner Transformation, so zeigt sich, daß die sechs Transformationsparameter prinzipiell auch aus diesen Momenten gewonnen werden können, da die Momente von den Transformationsparametern  $a_{ik}$  abhängen.

Zerlegt man die allgemeine ebene affine Transformation in einen homogenen Anteil  $\mathbf{A}$  und eine Translation  $\mathbf{b}$  :

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{b} = \begin{pmatrix} a_{00}x + a_{01}y + b_0 \\ a_{10}x + a_{11}y + b_1 \end{pmatrix}, \quad (4.7)$$

so transformieren sich die geometrischen Momente  $M_{01}$  und  $M_{10}$  mit den Gleichungen 4.7 und 3.6 unter reiner Translation  $\mathbf{b}$  zu

$$M'_{10} = M_{10} - b_0 M_{00} \quad (4.8)$$

$$M'_{01} = M_{01} - b_1 M_{01}. \quad (4.9)$$

Fordert man für die neue Lage  $M'_{10} \equiv 0$  und  $M'_{01} \equiv 0$ , so ergeben sich für die Translationsparameter  $b_0$  und  $b_1$  die Beziehungen

$$b_0 = \frac{M_{10}}{M_{00}} = x_S \quad \text{und} \quad b_1 = \frac{M_{01}}{M_{00}} = y_S. \quad (4.10)$$

Dies sind jedoch genau die Koordinaten des Objektschwerpunkts [Süße 1999] und damit gilt folgende Aussage: *Schwerpunktbezogene (zentrale) geometrische Momente sind translationsinvariant. Die zentralen Momente  $m_{01}$  und  $m_{10}$  sind immer gleich null.*

Führt man zunächst die Schwerpunktstransformation in der eben dargestellten Weise durch, so verbleibt zur Auflösung von Gleichung 4.7 die Bestimmung der vier Parameter der homogenen ebenen affinen Transformation  $\mathbf{A}$ . Vor der Behandlung möglicher Wege zur Lösung dieses Problems wird zunächst sichergestellt, daß durch eine homogene Affintransformation die bereits durchgeführte Schwerpunktnormierung ( $m_{10} = 0$  und  $m_{01} = 0$ ) nicht wieder zerstört wird. Unter homogener Affintransformation gilt für  $m'_{10}$  und  $m'_{01}$

$$m'_{10} = \det \mathbf{A} \iint_B (a_{00}x + a_{01}y) dx dy = \det \mathbf{A} * (a_{00}m_{10} + a_{01}m_{01}) = 0 \quad (4.11)$$

$$m'_{01} = \det \mathbf{A} \iint_B (a_{10}x + a_{11}y) dx dy = \det \mathbf{A} * (a_{10}m_{10} + a_{11}m_{01}) = 0, \quad (4.12)$$

womit die Schwerpunktnormierung gemäß Gleichung 4.10 in keinem Fall zerstört wird. Generell sollten zur Bestimmung der verbleibenden Parameter aus numerischen Gründen zentrale Momente mit möglichst niedriger Ordnung benutzt werden. Dies wären also  $m_{00}, m_{20}, m_{11}, m_{02}$ . Sie sind jedoch aufgrund der Invarianzbeziehung 3.17a voneinander abhängig, daher muß in jedem Fall auch auf Momente mindestens dritter Ordnung zurückgegriffen werden.

#### 4.5.1 Standardmethode zur Bestimmung einer Standardlage

Die Grundidee zur Bestimmung einer Standardlage aus zentralen Momenten wird in Udagawa, Toriwaki & Sugino [1964] vorgezeichnet. Mit der Standardmethode wird die homogene Affintransformation zerlegt in eine X-Scherung  $\mathbf{X}$ , eine anisotrope Skalierung  $\mathbf{S}$  und eine Rotation  $\mathbf{R}$  :

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} = \mathbf{R} \cdot \mathbf{S} \cdot \mathbf{X} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \beta & 0 \\ 0 & \gamma \end{pmatrix} \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \quad (4.13)$$

Die Parameter  $\alpha, \beta, \gamma$  und  $\varphi$  werden sukzessive durch die Normierungen  $m_{11}^{[X]} \equiv 0$  (liefert  $\alpha$ ),  $m_{20}^{[SX]} = m_{02}^{[SX]} \equiv 1$  (liefert  $\beta$  und  $\gamma$ ) und  $m_{30}^{[RSX]} + m_{12}^{[RSX]} \equiv 0$  (liefert  $\varphi$ ) bestimmt, wobei nach jeder Einzeltransformation zunächst alle Momente neu berechnet werden müssen. Hochgestellte Indices kennzeichnen im folgenden die bereits durchgeführte(n) Transformation(en), z.B. bedeutet  $m_{20}^{[SX]}$  : Das Moment  $m_{20}$  nach X-Scherung und Skalierung S. Die Standardmethode soll hier jedoch nicht weiter verfolgt werden, da sie für Objekte mit annähernd vierzähliger Symmetrie, bei denen nach X-Scherung und Skalierung  $m_{30}^{[SX]} + m_{12}^{[SX]} \approx 0$  ist, zu numerischen Instabilitäten bezüglich der Rotation in die Standardlage führt. Für eine detaillierte Betrachtung sei auf Süße [1999] verwiesen.

#### 4.5.2 Polynommethode zur Bestimmung einer Standardlage

Mit der Polynommethode, die ebenfalls in [Süße 1999] beschrieben ist, wird die homogene Affintransformation in folgender Weise separiert: X-Scherung  $\mathbf{X}$ , Y-Scherung  $\mathbf{Y}$  und anisotrope Skalierung  $\mathbf{S}$  :

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} = \mathbf{S} \cdot \mathbf{Y} \cdot \mathbf{X} = \begin{pmatrix} \gamma & 0 \\ 0 & \delta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \beta & 1 \end{pmatrix} \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \quad (4.14)$$

Mit der Normierung  $m_{30}^{[X]} = m_{30} + 3\alpha m_{21} + 3\alpha^2 m_{12} + \alpha^3 m_{03} \equiv 0$  erhält man ein Polynom dritten Grades in  $\alpha$ . Zunächst werden alle drei Lösungen weiterverfolgt. Zur Bestimmung von  $\beta$  fordert man  $m_{11}^{[XY]} \equiv 0$ . Die Skalenfaktoren  $\gamma$  und  $\delta$  werden wiederum durch die Normierungen  $m_{20}^{[SYX]} = m_{02}^{[SYX]} \equiv 0$  bestimmt. Aus den drei möglichen Standardlagen wird nun eine ausgewählt, etwa durch die Zusatzforderung  $m_{40}^{[SYX]} + m_{04}^{[SYX]} \rightarrow \min$ . Die Bestimmung der Standardlage nach dieser Methode ist unmöglich, wenn alle Momente dritter Ordnung gleich null sind. Für  $m_{30} \approx 0$  und  $m_{21} \approx 0$  und  $m_{12} \approx 0$  und  $m_{03} \approx 0$  ergeben sich auch hier Instabilitäten der Standardlage. Diese Einschränkung ist jedoch weit schwächer als die der Standardmethode.

### 4.5.3 Iterationsmethode zur Bestimmung einer Standardlage

Um der Ursache für die Instabilitäten der Polynommethode auszuweichen, muß die Bestimmung der Scherungsparameter  $\alpha$  und  $\beta$  durch eindeutig auflösbare Gleichungen erfolgen. Unter Scherung transformieren sich zentrale geometrische Momente zu:

$$\text{X-Scherung} : m_{pq}^{[X]} = \sum_{k=0}^p \binom{p}{k} \alpha^{p-k} m_{k,p+q-k} \quad (4.15)$$

$$\text{Y-Scherung} : m_{pq}^{[Y]} = \sum_{l=0}^q \binom{q}{l} \beta^{q-l} m_{l,p+q-l} . \quad (4.16)$$

Zur eindeutigen Bestimmung von  $\alpha$  und  $\beta$  taugen daher im allgemeinen alle  $m_{1q}$  beziehungsweise  $m_{p1}$ . Zur Auflösung wird dann jeweils durch  $m_{0,p+q}$  bzw.  $m_{p+q,0}$  dividiert. Für ungerade Ordnungen  $p+q$  können diese Nenner verschwinden, was wiederum Singularitäten bzw. Instabilitäten hervorrufen würde. Daher sind zur Bestimmung von  $\alpha$  und  $\beta$  Momente gerader Ordnung  $p+q$  besser geeignet. Hier verschwinden die Momente  $m_{0,p+q}$  bzw.  $m_{p+q,0}$  lediglich bei "leeren" Objekten, was in der Praxis jedoch keinerlei Einschränkung bedeutet. Zur konkreten Rechnung werden bei dieser Methode die Momente vierter Ordnung benutzt.

Zunächst wird der Parameter  $\alpha$  der X-Scherung aus Gleichung 4.15 mit der Normierung  $m_{13}^{[X]} \equiv 0$  berechnet:

$$\alpha = -\frac{m_{13}}{m_{04}} . \quad (4.17)$$

Damit berechnen sich die anderen Momente vierten Grades aus Gleichung 4.16 zu:

$$m_{04}^{[X]} = m_{04} \quad (4.18)$$

$$m_{22}^{[X]} = m_{22} + \alpha m_{13} \quad (4.19)$$

$$m_{31}^{[X]} = m_{31} + 3\alpha m_{22} + 2\alpha^2 m_{13} \quad (4.20)$$

$$m_{40}^{[X]} = m_{40} + 4\alpha m_{31} + 6\alpha^2 m_{22} + 3\alpha^3 m_{13} . \quad (4.21)$$

Den Parameter  $\beta$  der Y-Scherung berechnet man aus Gleichung 4.16 mit der Normierung  $m_{31}^{[YX]} \equiv 0$  zu

$$\beta = -\frac{m_{31}^{[X]}}{m_{40}^{[X]}} . \quad (4.22)$$

Die anderen Momente vierter Ordnung nehmen durch X- und Y-Scherung folgende Gestalt an:

$$m_{40}^{[YX]} = m_{40}^{[X]} \quad (4.23)$$

$$m_{22}^{[YX]} = m_{22}^{[X]} + \beta m_{31}^{[X]} \quad (4.24)$$

$$m_{13}^{[YX]} = m_{13}^{[X]} + 3\beta m_{22}^{[X]} + 2\beta^2 m_{31}^{[X]} \quad (4.25)$$

$$m_{04}^{[YX]} = m_{04}^{[X]} + 6\beta^2 m_{22}^{[X]} + 3\beta^3 m_{31}^{[X]} . \quad (4.26)$$

Dadurch wird jedoch unter Umständen die Normierung  $m_{13}^{[X]} \equiv 0$  wieder zerstört. Das bedeutet, es ist nicht möglich,  $\alpha$  und  $\beta$  mit jeweils einer X- und einer Y-Scherung so zu bestimmen, daß gleichzeitig beide Normierungen  $m_{13}^{[X]} \equiv 0$  und  $m_{31}^{[YX]} \equiv 0$  erfüllt sind. Wiederholt man jedoch die X- und Y-Scherungen fortwährend, so läßt sich nachweisen [Süße 1999], daß sowohl  $m_{13}^{[XYX\dots YX]}$  und  $m_{31}^{[YX\dots YX]}$  als auch  $\alpha$  und  $\beta$  eine monotone Nullfolge bilden. Bricht man den Prozeß bei gleichzeitigem Unterschreiten eines kleinen Schwellwertes (z.B.  $10^{-6}$ ) von  $m_{13}^{[XYX\dots YX]}$  und  $m_{31}^{[YX\dots YX]}$  ab, so verschwinden in diesem Endzustand auch die beiden Scherungsparameter nahezu. Die gesamte Zerlegung der affinen Abbildung schreibt sich damit für diese Methode zu:

$$\mathbf{A} = \mathbf{S} \cdot (\mathbf{Y} \cdot \mathbf{X})^n = \mathbf{S} \cdot \begin{pmatrix} 1 & 0 \\ \beta_n & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & \alpha_n \\ 0 & 1 \end{pmatrix} \cdot \dots \cdot \begin{pmatrix} 1 & 0 \\ \beta_1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & \alpha_1 \\ 0 & 1 \end{pmatrix} . \quad (4.27)$$

Nach jeder Scherung müssen selbstverständlich alle Momente gemäß Gleichungen 4.15 und 4.16 nachgeführt werden.

Zur Bestimmung der Skalierungsparameter  $\gamma$  und  $\delta$  werden die Normierungen  $m_{20} \equiv 1$  bzw.  $m_{02} \equiv 1$  vorgegeben. Diese Parameter berechnen sich dann mit den Gleichungen 4.15 und 4.16 zu:

$$\gamma = \sqrt[3]{\frac{m_{02}}{m_{20}^3}} \quad \text{bzw.} \quad \delta = \sqrt[3]{\frac{m_{20}}{m_{02}^3}}. \quad (4.28)$$

Die homogene Standardlage nach der Iterationsmethode ist zusammenfassend durch folgende Normierungen definiert:

$$\boxed{\text{Standardlage: } m_{13} \equiv 0 \quad \text{und} \quad m_{31} \equiv 0 \quad \text{und} \quad m_{20} \equiv 1 \quad \text{und} \quad m_{02} \equiv 1}. \quad (4.29)$$

Die so definierte Standardlage ist noch mehrdeutig modulo  $90^\circ$  und bezüglich Spiegelungen an der  $X$ - oder  $Y$ -Achse. Zur Beseitigung dieser Mehrdeutigkeiten werden die Vorzeichen von zwei Momenten ungerader Ordnung,  $m_{gu}$  und  $m_{ug}$  (z.B.  $m_{21}$  und  $m_{12}$ ) sowie deren betragsmäßige Verhältnisse betrachtet. Hierbei steht  $g$  für "gerade Zahl" und  $u$  für "ungerade Zahl". Bei ebenen Rotationen um ganze Vielfache von  $90^\circ$  und gegebenenfalls zusätzlicher Spiegelung an der  $X$ - oder  $Y$ -Achse existieren acht möglichen Lagen, die sich allein durch Betrachtung von  $m_{13}$ ,  $m_{31}$ ,  $m_{20}$  und  $m_{02}$  nicht unterscheiden lassen. Die Durchführung einer  $X$ - und einer  $Y$ -Achsen Spiegelung zugleich stellt keinen neuen Fall dar, da dieselbe Abbildung auch durch eine  $180^\circ$ -Drehung erreicht wird. Demgegenüber stehen für Vorzeichen und relative betragsmäßige Größe acht mögliche Fälle für  $m_{gu}$  und  $m_{ug}$ , die in Tabelle 5.1 aufgeführt sind. Ein doppeltes Vorzeichen kennzeichnet dabei die betragsmäßig größere Zahl.

Fall	1	2	3	4	5	6	7	8
$\mathbf{m}_{gu}$	++	+	++	+	--	-	--	-
$\mathbf{m}_{ug}$	+	++	-	--	+	++	-	--

**Tabelle 4.1:** Die acht möglichen Fälle, die sich für die Vorzeichen und die relative betragsmäßige Größe zweier Momente  $m_{gu}$  und  $m_{ug}$  ergeben.

Ausgehend von diesen acht Fällen besteht nach Durchführung der in Gleichung 4.29 dargestellten Normierungen demnach die Aufgabe, eine der acht möglichen Lagen durch entsprechende Wahl der  $m_{gu}$  und  $m_{ug}$  als Standardlage auszuzeichnen und die anzubringende Rotation und Spiegelung zur Erreichung dieses Zustands zu ermitteln. Bei Rotationen mit Vielfachen von  $90^\circ$  verhalten sich die Vorzeichen von Momenten ungerader Ordnung wie in Tabelle 4.2 dargestellt.

Rotation	$\mathbf{m}_{gu}^R$	$\mathbf{m}_{ug}^R$
$90^\circ$	$m_{ug}$	$-m_{gu}$
$180^\circ$	$-m_{gu}$	$-m_{ug}$
$270^\circ$	$-m_{ug}$	$m_{gu}$

**Tabelle 4.2:** Verhalten von Momenten ungerader Ordnung unter Rotationen um Vielfache von  $90^\circ$ .

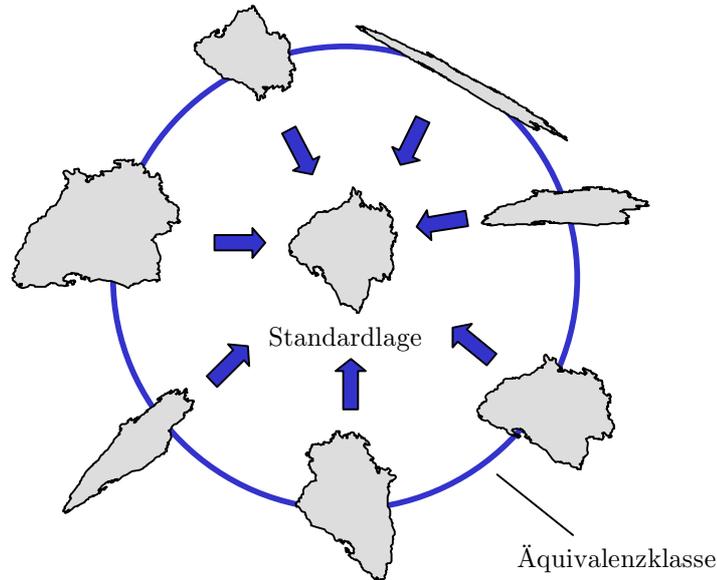
Zu jeder dieser möglichen Drehlagen existieren zwei spiegelsymmetrische Lagen. Im folgenden wird ohne Beschränkung der Allgemeinheit lediglich eine Spiegelung an der  $Y$ -Achse in Betracht gezogen. Dabei ändern ganz allgemein nur Momente  $m_{un}$  ihr Vorzeichen, wobei  $n$  eine beliebige und  $u$  eine beliebige ungerade positive ganze Zahl ist. Fordert man für die Standardlage z.B.  $m_{gu} \geq 0$  und  $m_{ug} \geq 0$  und  $m_{gu} \geq m_{ug}$ , so ergeben sich entsprechend den genannten Transformationsregeln aus den acht möglichen Fällen (Tabelle 5.1) die in Tabelle 4.3 angegebenen Transformationen in die oben definierte Standardlage.

$\mathbf{m}_{gu}^R$	$\mathbf{m}_{ug}^R$	Rotation	Spiegelung	$\mathbf{m}_{gu}^R$	$\mathbf{m}_{ug}^R$	Rotation	Spiegelung
++	+	$0^\circ$	nein	+	++	$90^\circ$	ja
++	-	$0^\circ$	ja	+	--	$270^\circ$	nein
--	+	$180^\circ$	ja	-	++	$90^\circ$	nein
--	-	$180^\circ$	nein	-	--	$270^\circ$	ja

**Tabelle 4.3:** Anzubringende Rotation und Spiegelung an der  $Y$ -Achse in Abhängigkeit von  $m_{gu}$  und  $m_{ug}$ .

Die in Tabelle 4.3 dargestellten Abbildungsvorschriften wurden aus der Analyse aller 64 möglichen Fälle (acht Konfigurationen aus Tabelle 5.1 mit jeweils vier Dreh- und zwei Spiegellagen kondensiert. Das Ergebnis der gesamten Analyse wird in Anhang D gegeben.

Abschließend läßt sich feststellen, daß die Iterationsmethode zwar etwas rechenaufwendiger ist, dafür jedoch zu sehr stabilen Standardlagen führt. Während bei der Standard- und der Polynommethode die Momente dritter Ordnung zu Instabilitäten bzw. Singularitäten führen können, wird hier zur Normierung auf deren Benutzung verzichtet. In Abbildung 4.6 ist die Standardlage für ein Geo-Objekt in verschiedenen affinen Kopien (einschließlich Spiegelungen) gegeben.



**Abbildung 4.6:** Der Umriss von Baden-Württemberg in sieben affinen Kopien (Repräsentanten der Äquivalenzklasse), sowie deren Standardlage.

Der Nutzen der Standardlage für die vorliegende Arbeit besteht nun darin, daß über die jeweiligen Standardlagen die Transformation von zwei korrespondierenden Objekten aufeinander ohne Kenntnis weiterer Punktkorrespondenzen berechnet werden kann. Maßgeblich für diese Transformation sind also lediglich die zentralen Momente der beiden Objekte. Es seien  $O_1$  und  $O_2$  zwei Objekte, die anhand ihrer Invarianten oder sonstiger Merkmale korrespondieren. Die Standardlagentransformation der Objekte sei  $S_1(O_1)$  bzw.  $S_2(O_2)$ . Die affine Transformation von  $O_1$  nach  $O_2$  ist dann gegeben durch:

$$T_{1 \rightarrow 2} = S_2^{-1}(O_2) * S_1(O_1) . \quad (4.30)$$

Entsprechend ist die umgekehrte Transformation gegeben durch:

$$T_{2 \rightarrow 1} = S_1^{-1}(O_1) * S_2(O_2) . \quad (4.31)$$

Sind die beiden Objekte lediglich affine Kopien voneinander, so ist dieses Verfahren exakt, die Objekte lassen sich über ihre Standardlage in der angegebenen Weise fehlerfrei ineinander transformieren. Von praktischem Interesse ist jedoch - auch im Hinblick auf das Ziel dieser Arbeit - der Fall, bei dem die Objekte lediglich ähnlich, nicht exakt gleich sind. Zur Untersuchung der Verwendbarkeit der Standardlagen wurde ein Experiment durchgeführt, das im folgenden Abschnitt beschrieben wird.

#### 4.5.4 Ein Experiment zum praktischen Nutzen der Standardlage nach der Iterationsmethode

In real vorkommenden Daten, insbesondere wenn die Sensorobjekte durch ein Segmentierungsverfahren gewonnen werden, sind homologe Objektpaare nie geometrisch genau gleich geartet. Daher ist es notwendig, zu untersuchen, ob sich die Standardlagentransformation robust gegen eine gewisse geometrische Varianz der Objekte verhält. Ist dies der Fall, so kann die Transformation von einem homologen Partner auf den anderen

mit Hilfe der Gleichungen 4.30 und 4.31 zumindest näherungsweise aus den jeweiligen Standardlagen berechnet werden. Somit lässt sich die Transformation ohne aufwendig zu bestimmende homologe Punktpaare berechnen.

Zur Untersuchung des Verhaltens der Standardlage bei geometrisch leicht verschiedenen Objekten wurde an einem Satz von Vektorpolygonen statistisches Rauschen der  $x$ - und  $y$ -Koordinaten mit variabler Standardabweichung angebracht. Mit der Iterationsmethode und den im vorigen Abschnitt angegebenen Normierungen wurde jeweils die Standardlage und über die Beziehungen 4.30 und 4.31 die affine Transformation von einem zum anderen Objekt berechnet. Aus zwei Gründen wurden die Buchstaben des Alphabets als Untersuchungsdatensatz ausgewählt. Zum einen stellen sie aufgrund verschiedener auftretender Symmetrien Extremfälle dar, die schwieriger als die bei Geo-Objekten auftretenden Verhältnisse zu handhaben sind. Zum anderen sollten die Buchstaben vertraute Objekte sein, Unterschiede oder Fehler sind damit leichter zu erkennen. Die verrauschten Daten wurden einer affinen Transformation einschließlich einer Spiegelung unterzogen, um einen möglichst allgemeinen Fall zu untersuchen. In Abbildung 4.7 ist links das verwendete Alphabet (grau) und das Ergebnis der individuellen Transformationen (schwarz) abgebildet, die rechte Seite zeigt den affin transformierten, jedoch unverrauschten Datensatz. Abbildung 4.8 zeigt die selbe Untersuchung mit verrauschten Daten.

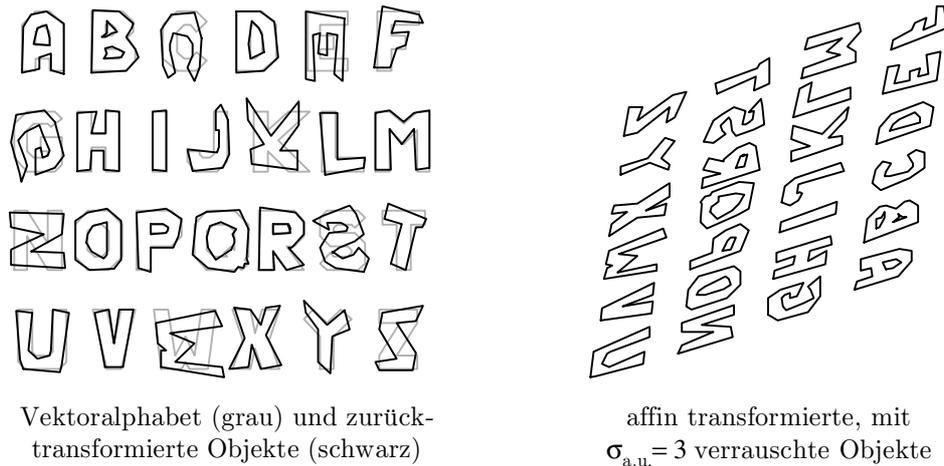


**Abbildung 4.7:** Experiment zur praktischen Verwendbarkeit von Standardlagen: Das Alphabet (links, grau) wird affin transformiert (rechts) und zurücktransformiert (links, schwarz).

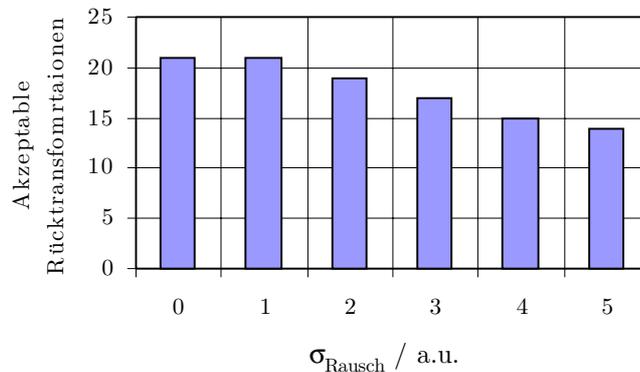
Bereits im unverrauschten Fall zeigt sich, daß einige Buchstaben nicht korrekt zurücktransformiert werden, was bedeutet, daß in diesen Fällen die Standardlage nicht robust zu berechnen ist. All diesen Fällen ist gemeinsam, daß die Momente  $m_{21}$  und  $m_{12} \approx 0$  sind (hier die Buchstaben H, N, O, S, und Z; das X ist zufälligerweise korrekt transformiert). Dies führt zu numerischen Unstabilitäten bei der Behandlung der nach der Normierung noch vorhandenen Mehrdeutigkeiten. Die Standardlage für derartige Objekte läßt sich demnach nicht eindeutig bestimmen, was jedoch für praxisrelevante Geo-Objekte keine allzu große Einschränkung bedeutet.

Unter dem Einfluß von statistischem Rauschen, der in Abbildung 4.8 exemplarisch dargestellt ist, lassen sich auch Objekte, die nicht die o.g. Eigenschaften besitzen, nicht mehr in jedem Fall korrekt zurücktransformieren (hier die Buchstaben C, E, G, J, K, N, S und W), wogegen z.B. der Buchstabe H in diesem Fall im Rahmen der angebrachten Fehler korrekt zurücktransformiert wird.

Insgesamt wurden sechs solcher Experimente mit unterschiedlichen Standardabweichungen durchgeführt, deren zusammengefasste Ergebnisse in Abbildung 4.9 veranschaulicht sind. Im Anhang E sind die Ergebnisse aller sechs Experimente im einzelnen abgebildet. Zusammenfassend ist zu diesen Untersuchungen festzustellen, daß mittels der Standardlagen, die mit der Iterationsmethode gewonnen wurden, auch stärker geometrisch unterschiedliche Objekte noch in annehmbarem Maße ineinander transformiert werden können. Wie später noch genauer beschrieben wird, ist die Gesamtstrategie zur automatischen Paßpunkterzeugung ohnehin so angelegt, daß nicht alle möglichen Objekte zwingend richtig erkannt und transformiert werden müssen. Vielmehr wird sich die Untermenge an richtigen Objektpaaren durchsetzen, die die stärkste Korrespondenz besitzt (vgl. dazu Abschnitt 4.1).



**Abbildung 4.8:** Experiment zur praktischen Verwendbarkeit von Standardlagen: Das Alphabet (links, grau) wird affin transformiert und verrauscht (rechts) und zurücktransformiert (links, schwarz).



**Abbildung 4.9:** Anzahl der als akzeptabel eingestuften Rücktransformationen bei variablem statistischem Rauschen in  $x$  und  $y$ .

## 4.6 Illustration des Zuordnungsverfahrens

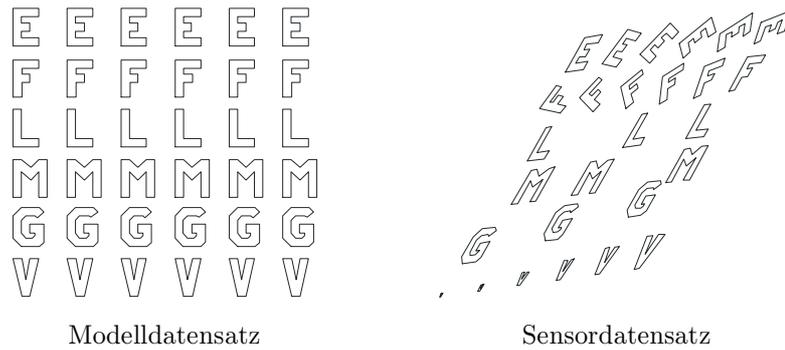
Zur Veranschaulichung der Schritte, die während des gesamten Zuordnungsverfahrens ablaufen, diene das in Abbildung 4.10 dargestellte Beispieldatensatzpaar. Als Modelldatensatz wurden dabei sechs Buchstaben jeweils sechs mal kopiert und in einem regelmäßigen Raster angeordnet (Abbildung 4.10, links). Der Sensordatensatz wurde aus dem Modelldatensatz folgendermaßen abgeleitet: Zunächst wurden einzelne Buchstaben verdreht ("E" und "F"), verschoben ("L", "M" und "G") oder skaliert ("V"). Anschließend wurden die  $x$ - und  $y$ -Koordinaten statistisch verrauscht. Abschließend wurde der ganze Datensatz noch folgender Affinttransformation unterzogen (Abbildung 4.10, rechts):

$$x' = 0.7x + 0.5y \quad (4.32)$$

$$y' = 0.1x + 0.9y. \quad (4.33)$$

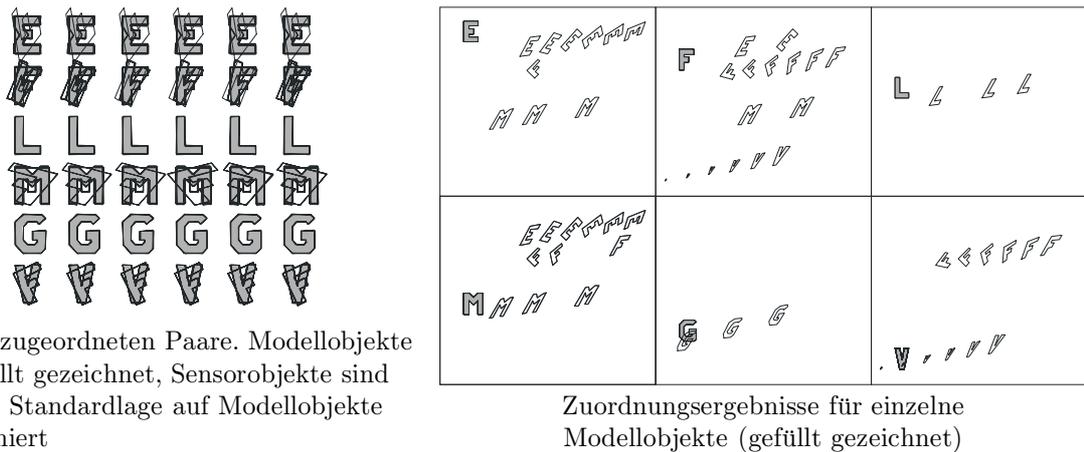
Damit stellt der Beispiel-Sensordatensatz einen allgemeinen Fall bei affiner Transformation dar.

Von beiden Datensätzen wurden zunächst die zentralen geometrischen Momente bis einschließlich achter Ordnung berechnet. Aus diesen Momenten wurden jeweils für alle Objekte die ersten sechs absoluten Affinvarianten  $\Phi(2, 4)$ ,  $\Phi(2, 6)$ ,  $\Phi(2, 8)$ ,  $\Phi(4, 6)$ ,  $\Phi(4, 8)$  und  $\Phi(6, 8)$  berechnet. Da für diese simulierten Daten keine reale Sensoraufnahme existiert, wurde die unäre Zwangsbedingung einer minimalen relativen Objektgröße nicht angewandt. Mit dem in Abschnitt 4.4 beschriebenen Verfahren wurden insgesamt 326 hypothetische Modell-Sensor-Paare ermittelt, deren relativer Unterschied in den einzelnen Invarianten jeweils kleiner als 5% ist. Alle



**Abbildung 4.10:** Beispieldatensätze (Modell- und Sensordaten) zur Veranschaulichung des gesamten Zuordnungsprozesses.

326 Paare sind in Abbildung 4.11, links abgebildet. Dabei sind die Sensorobjekte über ihre Standardlage individuell auf das jeweilige Modellobjekt transformiert (vgl. Gleichungen 4.30 und 4.31). Der rechte Teil von Abbildung 4.11 zeigt die untransformierten Zuordnungsergebnisse für die erste Spalte im Modelldatensatz. Neben allen jeweils gleichen Buchstaben existiert für die Buchstaben "E", "F", "M" und "V" noch eine Reihe von Fehlzuordnungen, die es mit der nachfolgenden Baumsuche nach konsistenten Objektpaaren zu eliminieren galt. Für die Baumsuche, wie sie in Abschnitt 4.1 beschrieben ist, wurde eine maximale Suchtiefe von 3 vorgege-



Alle 326 zugeordneten Paare. Modellobjekte sind gefüllt gezeichnet, Sensorobjekte sind über ihre Standardlage auf Modellobjekte transformiert

Zuordnungsergebnisse für einzelne Modellobjekte (gefüllt gezeichnet)

**Abbildung 4.11:** Zuordnungsergebnisse für den Beispieldatensatz.

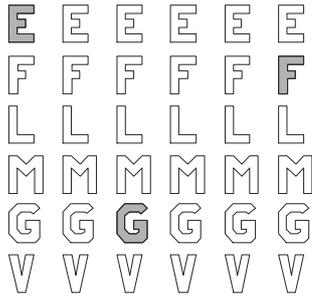
ben. Objektpaare, die sich bei Anwendung der aktuellen Transformation um mehr als 70% überlappen, wurden in die jeweilige Interpretation aufgenommen. Die drei konsistenten Objektpaare, die aus der Baumsuche mit Zwangsbedingungen hervorgingen, sind in Abbildung 4.12 dargestellt.

Im linken Teil der Abbildung 4.12 sind die konsistenten Objekte des Modelldatensatzes zu sehen, der mittlere Teil zeigt die entsprechenden Objekte des Sensordatensatzes. Im rechten Teil sind die Sensordaten mit der aus den konsistenten Paaren gemittelten Affintransformation auf die Modelldaten transformiert. Die gemittelte Affintransformation nimmt folgende Gestalt an:

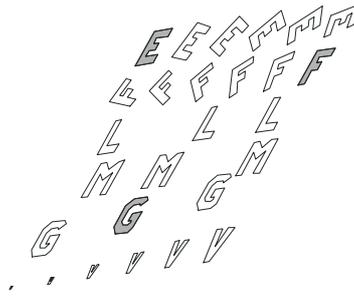
$$x' = 0,698x + 0,502y - 1,067 \tag{4.34}$$

$$y' = 0,104x + 0,905y - 0,511 \tag{4.35}$$

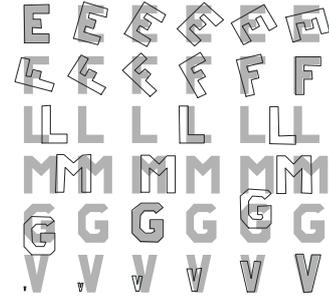
Die Verschiebungen in  $x$  und  $y$  sind vergleichsweise klein gegen die Objektausdehnungen. Das ermittelte Ergebnis liegt daher sehr nahe an der ursprünglich angewandten Transformation (Gleichungen 4.32 und 4.33). Die Objekte der konsistenten Paare ("E-E", "F-F" und "G-G") kommen gut übereinander zu liegen. Somit ist das Ziel dieses Schrittes, eine globale affine Transformation von Sensor- in Modelldaten zu finden, erreicht. Man sieht zusätzlich noch die Übereinstimmung eines vierten Paares ("V-V", rechts unten), das jedoch aufgrund



Die 3 gefundenen  
konsistenten Objekte  
im Modelldatensatz



Die 3 gefundenen  
konsistenten Objekte  
im Sensordatensatz



Transformation von Sensor- auf  
Modelldaten (grau). Gemittelte  
Transformation aus 3 konsistenten Paaren

**Abbildung 4.12:** Ergebnisse der Baumsuche mit Zwangsbedingungen, angewandt auf den Beispiel-Datensatz. Die maximale Suchtiefe von 3 liefert drei konsistente Objektpaare.

der maximalen Suchtiefe von 3 nicht in der konsistenten Interpretation enthalten ist. Ein abschließender Verifikationsschritt dient dazu, Paare, die konsistent, aber nicht in der konsistenten Interpretation enthalten sind, zu ermitteln. Im rechten Bild in Abbildung 4.12 sind alle Sensorobjekt-Hypothesen mit der zuvor ermittelten globalen Transformation auf ihre Modellobjekt-Partner transformiert.

## 4.7 Verifikation der Hypothesen

Das in Abschnitt 4.1 beschriebene Baumsuchverfahren ist so angelegt, daß die Suche nach einer konsistenten Interpretation bei einer gewissen Tiefe  $n$  abbricht und aus allen Objektpaaren dieser konsistenten  $n$ -Interpretation eine gemittelte globale Affintransformation bestimmt wird. Das Ziel des Verifikationsschrittes besteht nun darin, aus allen Hypothesen diejenigen Objektpaare zu finden, die mit dieser globalen Transformation in Einklang stehen<sup>2</sup>. Dazu wird ein zweistufiges Verfahren vorgeschlagen. In einer ersten Stufe wird die individuelle Transformation eines Paares mit der ermittelten globalen Transformation verglichen. Um Unwägbarkeiten im Bereich kleiner Zahlenwerte zu vermeiden, empfiehlt es sich hierbei, von den Matrixelementen der homogenen Affintransformation (ohne Translation) auf Dreh- und Scherwinkel und Maßstabsfaktoren überzugehen. Die gesamte Affintransformation sei durch Gleichung 4.7 gegeben. Die Maßstabsfaktoren einer derartigen Transformation  $k_x$  in  $x$ -Richtung und  $k_y$  in  $y$ -Richtung berechnen sich zu

$$k_x = \sqrt{a_{00}^2 + a_{10}^2} \quad (4.36)$$

$$k_y = \sqrt{a_{01}^2 + a_{11}^2}, \quad (4.37)$$

die starre Rotation  $\kappa$  des Koordinatensystems ergibt sich zu

$$\kappa = \begin{cases} \arctan\left(\frac{a_{10}}{a_{00}}\right) & a_{00} \neq 0 \wedge a_{10} \geq 0 \\ \pi + \arctan\left(\frac{a_{10}}{a_{00}}\right) & \text{für } a_{00} \neq 0 \wedge a_{10} < 0 \\ \frac{\pi}{2} & a_{00} = 0 \wedge a_{10} \geq 0 \\ -\frac{\pi}{2} & a_{00} = 0 \wedge a_{10} < 0 \end{cases} \quad (4.38)$$

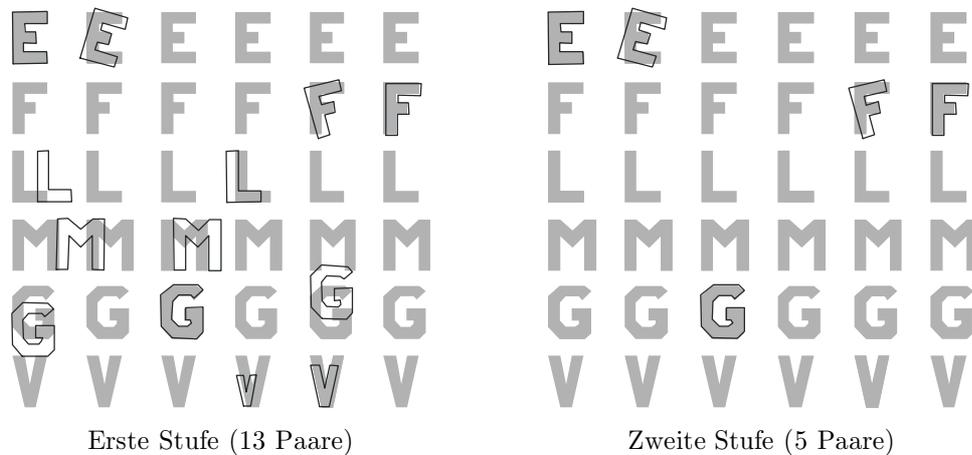
und die Scherung  $\alpha$  ergibt sich aus

$$\alpha = \begin{cases} -(\arctan\left(\frac{a_{11}}{a_{01}}\right) + \kappa) & a_{01} \neq 0 \wedge a_{11} \geq 0 \\ -(\pi + \arctan\left(\frac{a_{11}}{a_{01}}\right) + \kappa) & \text{für } a_{01} \neq 0 \wedge a_{11} < 0 \\ \kappa - \frac{\pi}{2} & a_{01} = 0 \wedge a_{11} \geq 0 \\ \frac{\pi}{2} - \kappa & a_{01} = 0 \wedge a_{11} < 0 \end{cases}. \quad (4.39)$$

<sup>2</sup>Würde die vorangegangene Baumsuche nicht in der Tiefe  $n$  abbrechen, so wäre der Verifikationsschritt nicht mehr nötig, da bereits alle möglichen konsistenten Paare identifiziert wären. Für praktische Berechnungen ist dies jedoch nicht durchführbar, da sich der Suchraum mit der zu durchsuchenden Tiefe exponentiell vergrößert.

In diesem Schritt wird für jede Objektpaarhypothese geprüft, ob  $k_x$ ,  $k_y$ ,  $\kappa$  und  $\alpha$  der paarweisen Affintransformation im Rahmen bestimmter Grenzen mit den entsprechenden Parametern der globalen Affintransformation übereinstimmen. Alle Paare, für die dies der Fall ist, werden in die zweite Stufe weitergeleitet, die restlichen werden als nicht verifizierte Paare aus der Liste entfernt. In einer zweiten Stufe werden die Objekte jedes Paares mit der globalen Affintransformation aufeinander transformiert. Damit ein Paar als endgültig verifiziert gilt, muß der Überlappungsbereich der transformierten Objekte einen gewissen Mindestanteil (z.B 70%) beider Einzelobjektflächen betragen. Sinnvollerweise ist dabei die Objektüberlappung ähnlich zu wählen wie bei der Baumsuche (siehe Abschnitt 4.3.4).

Zur Illustration des Verifikationsschrittes wird das Beispiel aus Abschnitt 4.6 nochmals aufgegriffen. Das rechte Bild in Abbildung 4.12 zeigt die Transformation aller hypothetischen Sensorobjekte auf ihre Modellobjekt-Partner mittels der zuvor bestimmten globalen Transformation. Bei Anwendung der ersten Stufe des Verifikationsschrittes ergibt sich das in Abbildung 4.13 links dargestellte Bild. Folgende Schwellwerte wurden dabei angewandt:  $\Delta\alpha < 45^\circ$ ,  $\Delta\kappa < 45^\circ$ ,  $\Delta(k_x) < 30\%$ ,  $\Delta(k_y) < 30\%$ .



**Abbildung 4.13:** Illustration des zweistufigen Verifikationsprozesses anhand des o.g. Beispiels.

Die erhaltenen 13 Paare stimmen im Rahmen der angewandten Schwellwerte in ihren Transformationen überein. In der zweiten Stufe wird wie in Abschnitt 4.3.4 der Anteil der Schnittfläche an den Einzelflächen berechnet. Dabei fallen weitere Paare aus der verifizierten Liste heraus, es verbleiben die fünf in Abbildung 4.13 rechts abgebildeten Paare. Wie oben sind auch hier die Sensorobjekte auf die Modellobjekte transformiert. Der Verifikationsschritt ermöglicht es also, aus allen Hypothesen und einer zuvor berechneten globalen Transformation diejenigen Paare zu eliminieren, die "in etwa" übereinstimmen. Genau diese Eigenschaft ist notwendig bei der Behandlung topographisch verzerrter oder leicht veränderter Daten, wie sie in der Praxis auftreten. Wie im Beispiel gezeigt, können in Extremfällen auch visuell "richtige" Objekte nicht verifiziert werden (hier das "V" in der rechten unteren Ecke), wenn die zu Grunde liegende paarweise Transformation nicht konsistent ist. Dieser Effekt tritt allerdings bei realen weniger symmetrischen GIS-Objekten kaum auf und ist somit eher als Besonderheit des gewählten synthetischen Datensatzes zu werten.

## 4.8 Optimierung individueller Transformationsparameter

Das bisher beschriebene Verfahren liefert eine globale Affintransformation, mittels der der Sensordatensatz auf den Modelldatensatz abgebildet wird. Da es letztlich Ziel des Gesamtverfahrens ist, möglichst genaue Passpunkte zu erzeugen, ist die so erreichbare Genauigkeit (mehrere Pixel) nicht ausreichend. Zur Verbesserung der Transformation vom Sensor- in das Modellsystem wird daher vorgeschlagen, die objektweise Transformation aus den Standardlagen individuell zu verbessern. In diesem Schritt muss demzufolge - ausgehend von der bereits vorliegenden individuellen Transformation als Näherung - eine objektweise optimale affine oder projektive Transformation bestimmt werden. Da es sich bei den Sensordaten ursprünglich um 3D-Polygone handelt, sollte zumindest die ebene projektive Transformation nicht ausser Acht gelassen werden. Der Algorithmus sollte in der Lage sein, Objekte, die lediglich teilweise vorhanden sind, richtig zu handhaben. Beispiele hierfür sind abgeschnittene Objekte am Bild- oder Kartenrand oder Objekte, deren Umriss sich aufgrund unterschiedlicher

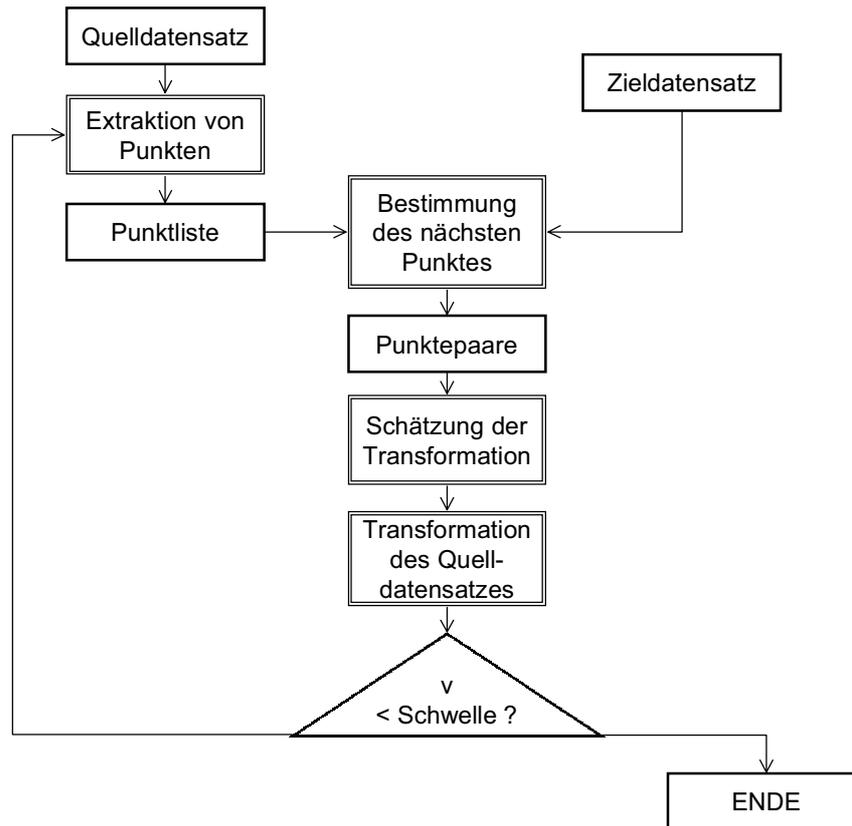


Abbildung 4.14: Flußdiagramm des Iterative-Closest-Point-Algorithmus<sup>3</sup>.

Aktualität von Sensor- und Modelldaten unterscheidet, etwa ein zwischen der Kartenerfassung und Sensordatenaufzeichnung gerodetes Waldstück. Das Problem stellt somit ein Registrierungsproblem von zumindest teilweise überlappenden 2D-Polylinien bei bekannter Näherungstransformation dar.

In der Literatur wird zur Lösung dieser Aufgabe von Besl & McKay [1992] der "Iterative Closest Point"-Algorithmus (ICP) vorgeschlagen. Der ICP-Algorithmus und zweckmäßige Erweiterungen werden im folgenden vorgestellt und untersucht.

#### 4.8.1 Der Iterative Closest Point (ICP)-Algorithmus

Das Grundprinzip des ICP-Algorithmus<sup>3</sup> besteht darin, für eine Punktmenge eines Quellobjekts (im hier behandelten Fall ein Sensordatenobjekt) jeweils den Punkt des Zielobjekts mit der kleinsten euklidischen Distanz zu finden. Aus allen Punktepaaren wird die gewünschte Transformation, z.B. affin oder projektiv, geschätzt. Das Quellobjekt wird dann mit der geschätzten Transformation transformiert. Im nächsten Iterationsschritt wird die Suche nach räumlich am nächsten liegenden Zielobjektpunkten erneut durchgeführt. Anschliessend wird eine verbesserte Transformation geschätzt und der nächste Iterationsschritt durchlaufen. Das Verfahren bricht ab, wenn entweder die vorgegebene Maximalzahl an Iterationen erreicht ist oder wenn die Verbesserung in den Transformationsparametern einen vorgegebenen Wert unterschritten hat. Die Bestimmung der Punktmenge des Quellobjektes ist abhängig von der Natur des Quellobjektes. Bei einem 2D- oder 3D-Punkthaufen wird sie von den Punkten selbst gebildet. Bei 2D- oder 3D-Linien bzw. 2D- oder 3D-Primitiven (Dreiecksmaschen o.Ä.) muß zuvor eine sinnvolle Auswahl getroffen werden. Dies kann bei einer Linie z.B. durch Extraktion von äquidistanten Punkten auf der Linie geschehen. Das Flussdiagramm des ICP-Algorithmus<sup>3</sup> wird illustriert durch Abbildung 4.14.

Besl & McKay [1992] unterscheiden zwischen direkter, parametrischer oder impliziter Darstellung der o.g. Quell- und Zielobjekte. Für diese Arbeit ist hiervon jedoch lediglich die direkte Darstellung von 2D-Polygonen relevant<sup>3</sup>.

<sup>3</sup>Der ICP-Algorithmus ist nicht auf 2 oder 3 Dimensionen beschränkt. Bei korrekter Verwendung der euklidischen Metrik ist er auf beliebig viele Dimensionen erweiterbar.

Es sei  $Q = \{l_{q,i}\}$  die Menge der  $N_q$  Liniensegmente des Quellobjektes für  $i = 1, \dots, N_l$  und  $Z = \{l_{z,j}\}$  die Menge der  $N_z$  Liniensegmente des Zielobjektes für  $j = 1, \dots, N_z$ . Eine Auswahl von Punkten auf  $Q$  sei gegeben durch  $P = \{\mathbf{p}_k\}$ . Seien  $\mathbf{r}_1$  der Anfangs- und  $\mathbf{r}_2$  der Endpunkt des Liniensegments  $l_{z,i}$ . Dann ist der euklidische Abstand eines Punktes  $\mathbf{p}_k$  des Quellobjektes zum Liniensegment  $l_{z,j}$  des Zielobjektes gegeben durch die Beziehung

$$d(\mathbf{p}_k, l_{z,i}) = \min_{u+v=1} \|u\mathbf{r}_1 + v\mathbf{r}_2\| \quad (4.40)$$

wobei  $u \in (0, 1)$  und  $v \in (0, 1)$ .

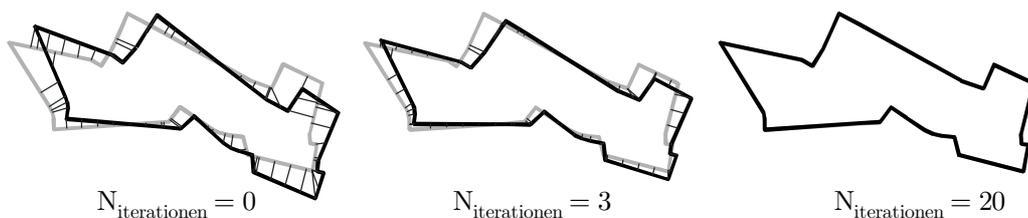
Der Abstand von  $\vec{p}_k$  zu  $Z$  ist dann gegeben durch

$$d(\mathbf{p}_k, Z) = \min_{i \in \{1, \dots, N_z\}} d(\mathbf{p}_k, l_{z,i}). \quad (4.41)$$

Für den räumlich nächsten Punkt  $\mathbf{y}_i$  aus  $\{Z\}$  zu  $\mathbf{p}_k$  ist die Beziehung  $d(\mathbf{p}_k, \mathbf{y}_i) = d(\mathbf{p}_k, Z)$  erfüllt.

Bei Anwendung einer Kleinsten-Quadrate-Minimierung zur Schätzung der Transformation aus den erhaltenen Punktpaaren konvergiert der ICP-Algorithmus monoton gegen ein Minimum. Der Beweis dazu wird von Besl & McKay [1992] geführt und daher hier nicht weiter erläutert.

Eine weitere interessante Frage ist, ob das Minimum, gegen das der Algorithmus konvergiert, globaler oder lediglich lokaler Natur ist. Die Hauptschwierigkeit dabei liegt nach Besl & McKay [1992] darin, Klassen von Näherungstransformationen für alle möglichen Objekte derart zu finden, dass der Algorithmus sicher zum richtigen, nämlich dem globalen Minimum hin konvergiert. Theoretisch besteht dazu die Möglichkeit, den ganzen Transformationsraum (im Falle affiner Transformation 6-dimensional) hinreichend dicht durchzuprobieren und somit Klassen, resp. Hyperflächen, in diesem Transformationsraum aufzufinden, die gegen das globale Minimum konvergieren. Der erforderliche Speicheraufwand sprengt jedoch aufgrund der höheren Dimensionalität schnell einen sinnvollen Rahmen. So beansprucht beispielsweise ein Hypervoxel-Raster mit 50 Punkten pro Dimension bei 6 Dimensionen einen Speicherplatz von  $50^6 = 1.5625 * 10^{10}$  bytes! Eine einfachere Bewertung einer Anfangstransformation hinsichtlich der Konvergenz in ein globales Minimum ist z.B. gegeben durch die prozentuale Überlappung der Objekte oder - bei hinreichender Unterscheidbarkeit der Hauptträgheitsmomente des Objektes - durch Vergleich der Hauptträgheitsachsen. Da in dem in dieser Arbeit entwickelten System die Näherungswerte für die Transformation in bereits beschriebener Weise beschafft werden, wird diese Diskussion nicht weiter vertieft. Untersuchungen zum Konvergenzverhalten des modifizierten ICP-Algorithmus' finden sich in Abschnitt 4.8.2. Mit Abbildung 4.15 wird der Ablauf des ICP-Algorithmus anhand dreier Zustände veranschaulicht. Der Ausgangszustand, in dem das schwarze Quellpolygon gegen das graue Zielpolygon verschoben und verdreht ist, ist im linken Teil dargestellt.



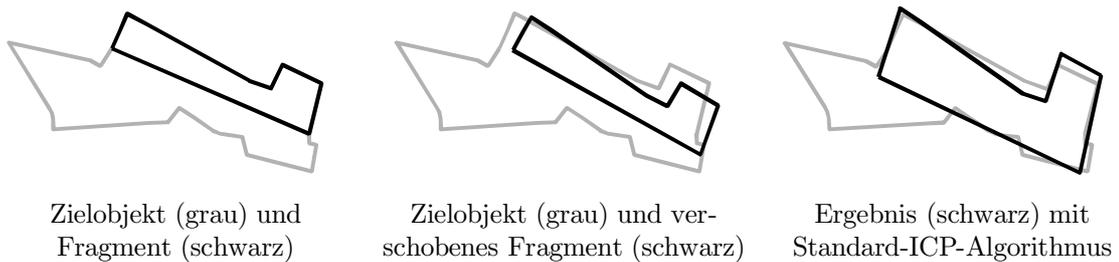
**Abbildung 4.15:** Verschiedene Zustände bei Anwendung des ICP-Algorithmus' auf ein zum Zielpolygon (dunkel) kongruentes Quellpolygon (hell).

Für äquidistante Punkte auf dem Rand des Quellpolygons wird jeweils der nächste Punkt auf dem Zielpolygon gesucht. Homologe Punktpaare sind durch schwarze Verbindungslinien dargestellt. Im mittleren Teilbild ist der Zustand nach 3 Iterationen dargestellt. Wie zu sehen ist, liegt das schwarze Quellpolygon bereits näher am Zielpolygon. Nach 20 Iterationen, dargestellt im rechten Teilbild von Abbildung 4.15, ist kein Unterschied mehr zwischen Quell- und Zielpolygon mehr zu erkennen. Das Ziel der Registrierung eines Polygons auf ein kongruentes Polygon wird damit vom Standard-ICP-Algorithmus bei genügend guter Näherungstransformation erreicht. Im Falle des mit dieser Arbeit entwickelten Systems sind jedoch die aufeinander zu registrierenden Polygone in der Regel weder völlig kongruent noch gleicher Herkunft. Vielmehr stammt eine Sorte Polygone aus einem GIS-Datensatz und besitzt daher von Natur aus reellwertige Koordinaten. Die andere Sorte Polygone entstammt einem Rasterdatensatz und besitzt daher zunächst diskrete Koordinaten mit der Rasterweite des zugrunde liegenden Bildes. Zudem kann selbst bei idealer Objektextraktion aus Rasterdaten ein Unterschied in der Objektberandung infolge unterschiedlicher Erfassungs- bzw. Aufnahmezeitpunkte oder radiometrischer

Verhältnisse bestehen. Diese Faktoren bewirken, dass der ICP-Algorithmus, wie er im vergangenen beschrieben wurde, nicht optimal für reale Verhältnisse geeignet ist. Eine Modifikation des ICP-Algorithmus' und Untersuchungen dazu werden im folgenden Abschnitt gegeben.

### 4.8.2 Modifizierter ICP-Algorithmus

Das entwickelte System liefert - wie bereits beschrieben - eine globale Transformation zwischen den beiden Eingabedatensätzen. Zur Erhöhung der Zahl homologer Punkte ist es daher wünschenswert, nicht nur vollständige Randpolygone sondern auch lediglich teilweise vorhandene bestmöglich zu integrieren. Wie Abbildung 4.16 zeigt, ist der Standard-ICP-Algorithmus dazu nicht oder nur bedingt in der Lage. Grund hierfür ist die Tatsache, dass die Schätzung der Transformation im ICP mittels einer Kleinste-Quadrate-Ausgleichung bewerkstelligt wird, die empfindlich gegenüber groben Fehlern und Ausreißern ist.



**Abbildung 4.16:** Fragment (schwarz) des Zielpolygons (grau).

Ausgehend von einer verbesserten globalen Affintransformation als Näherung bietet es sich an, im ICP eine robuste Ausgleichung zur Beseitigung des Einflusses der groben Fehler vorzunehmen. In der Literatur finden sich dazu eine Menge unterschiedlicher und ausgiebig erprobter Verfahren. Keines dieser Verfahren kann jedoch eine Fehlerrate von mehr als 50% verkraften. Da es in Randbereichen jedoch durchaus vorkommen kann, dass mehr als 50% eines Objektes in einem Eingabedatensatz nicht vorhanden sind, wird diese Möglichkeit hier nicht weiter verfolgt. Vielmehr wird das Grundprinzip der robusten Parameterschätzung, die Verringerung des Gewichtes von Beobachtungen mit großer Verbesserung, in der Weise auf die Spitze getrieben, daß homologe Punktpaare mit einem Abstand, der größer als ein sinnvoll zu wählender Maximalabstand ist, im ICP nicht mehr an der Schätzung der nächsten Transformation beteiligt werden. Die vorgeschlagene Modifikation des Standard-ICP-Algorithmus' besteht darin, den Maximalabstand aus der Standardabweichung aller homologen Punktpaare zu ermitteln. Es sei  $N$  die Anzahl der homologen Punktpaare. Die Menge der homologen Punktpaare sei gegeben durch  $H = \{(\mathbf{q}_i, \mathbf{z}_i) \mid i = (1, \dots, N)\}$ . Der Schwellwert  $d_{\max}$  für den Maximalabstand schreibt sich dann zu

$$d_{\max} = f * \sigma_H . \quad (4.42)$$

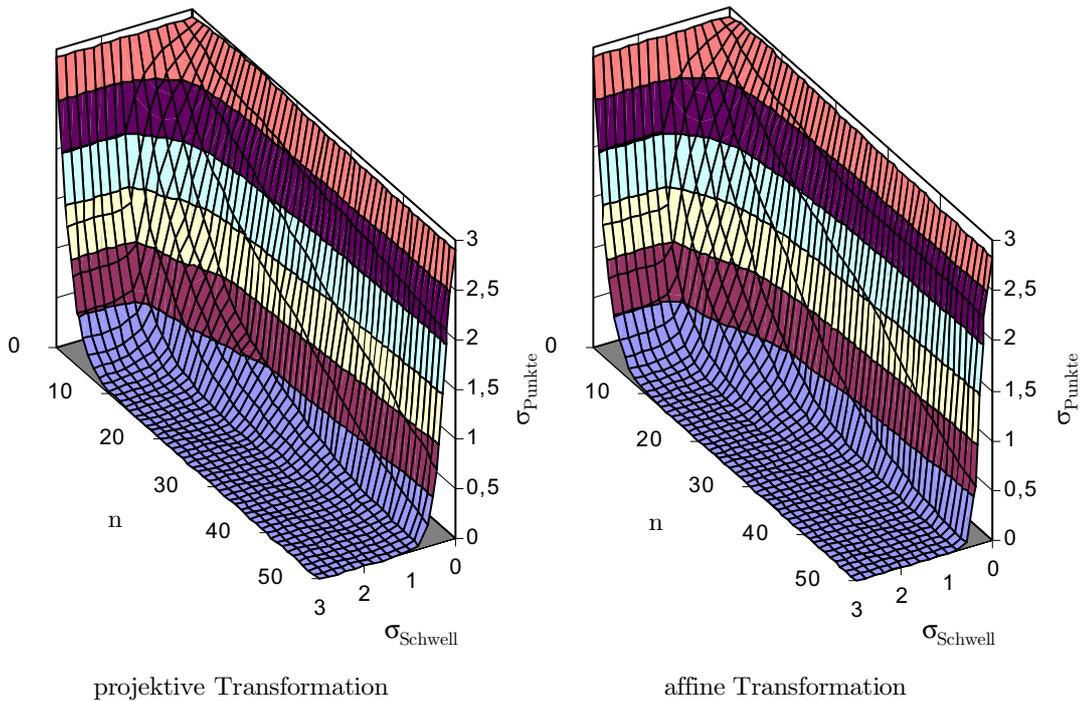
Die Standardabweichung  $\sigma_H$ , also der mittlere Punktfehler der homologen Punktpaare ist dabei

$$\sigma_H = \frac{\sum_{i=1}^N \|\mathbf{q}_i - \mathbf{z}_i\|}{\sqrt{N}} . \quad (4.43)$$

Mit der Modifikation des ICP-Algorithmus werden somit die ermittelten homologen Punktpaare nicht direkt der Schätzung der verbesserten Transformation zugeführt (siehe Abbildung 4.14). Vielmehr wird zunächst aus allen Paaren die Standardabweichung nach Gleichung 4.43 berechnet. Punktpaare, deren Abstand nach Gleichung 4.42 größer als  $d_{\max}$  ist, werden nicht in die Schätzung eingebracht. Zur Bestimmung des Faktors  $f$  aus Gleichung 4.42 wurde eine Meßreihe mit einem komplexen GIS-Objekt, das 10% gegen das Zielpolygon verschoben wurde, durchgeführt. Die Ergebnisse dieser Meßreihe sind in Abbildung 4.17 dargestellt.

Diese Untersuchung wurde folgendermaßen angestellt: Für jedes  $f \in \{0.2, 3.0\}$  wurde der modifizierte ICP-Algorithmus 50mal iterativ abgearbeitet. Das in jedem Iterationsschritt auftretende  $\sigma_H$  wurde berechnet und ist in z-Richtung in Abbildung 4.17 aufgetragen. Aus den Graphen lassen sich folgende Ergebnisse ablesen:

- Für  $f \rightarrow 0$  konvergiert der Algorithmus immer schlechter. Grund hierfür ist darin zu finden, daß immer weniger Punkte zur Schätzung beitragen und ein "Driften" des Objekts bzw. die Konvergenz in ein nicht globales Minimum eintreten kann.



**Abbildung 4.17:** Konvergenzverhalten des modifizierten ICP-Algorithmus'. Das  $\sigma_{Schwell}$ -fache der Standardabweichung aller Punkte gibt die Distanzgrenze, bis zu der Punkte noch zur Schätzung herangezogen werden.  $\sigma_{Punkte}$  gibt die Standardabweichung nach der  $n$ -ten Transformation an.

- Für  $f \rightarrow 3.0$  zeigt sich wie erwartet ein stabiles Konvergenzverhalten, da der Übergang  $f \rightarrow \infty$  den unmodifizierten Fall darstellt.
- Im Bereich um  $f = 2.0$  ist sowohl für affine wie auch für 2D-projektive Transformation die Grenze zwischen gut und schlecht konvergierendem Verhalten zu finden.

Im weiteren Verlauf der Arbeit wird, wenn nicht explizit anders erwähnt, mit  $f = 2.0$  gearbeitet, da mit diesem Wert einerseits ausreichend viele Ausreisser abgeschnitten werden und andererseits noch genügend Punkte zur Stabilisierung der Konvergenz vorhanden sind. Wie weit diese Lösung tatsächlich tragfähig ist lässt sich nur durch intensives Testen ermitteln. Zur Festlegung des günstigsten  $f$ -Wertes soll hier jedoch die angestellte Untersuchung ausreichen.

Zur Klärung des Verhaltens des in beschriebener Weise modifizierten ICP-Algorithmus' wurden eine Reihe Untersuchungen durchgeführt, die im folgenden näher erläutert werden. Alle Untersuchungen haben gemeinsam, daß sie mit einem komplexen und einem einfachen Polygon durchgeführt wurden. Des weiteren wurde sowohl eine affine wie auch eine 2D-projektive Transformation angewandt. Das Ergebnis der Transformation mit dem geringeren resultierenden Punktfehler ist jeweils aufgetragen. Es wurden keine Polygonfragmente betrachtet. Die verwendeten Polygone sind in Abbildung 4.18 dargestellt.

### 4.8.3 Einfluß von x-y-Translation

Mit dieser Untersuchung wird der Einfluß von Verschiebungen des Quellobjekts in x- und y-Richtung gegen das Zielobjekt beleuchtet. Als Untersuchungsobjekte dienen die beiden in Abbildung 4.18 dargestellten Polygone. Um eine Vergleichbarkeit der Ergebnisse zu ermöglichen, wurde die Verschiebung in x- und y-Richtung jeweils relativ zur entsprechenden Objektabmessung angebracht. Nach 50 Iterationen wurde für jede Verschiebung die Standardabweichung der Residuen in den Vertices berechnet. Für alle Untersuchungen gilt  $f = 2$ . Die Durchführung führte zu den in Abbildung 4.19 dargestellten Ergebnissen. Wie man an den Ergebnissen dieser Untersuchung sieht, zeigt der modifizierte ICP-Algorithmus bei Verschiebungen des Quell- gegen das (identische) Zielobjekt von etwa 10% der Objektausdehnung eine exzellente Konvergenz. Das Ergebnis ist nahezu isotrop, die Linien gleicher Standardabweichung folgen in etwa den Umrissen der Polygone. Für diese Arbeit bedeutet das Ergebnis: Wenn die affine Näherungstransformation aus der bereits besprochenen Baumsuche zwei Objekte

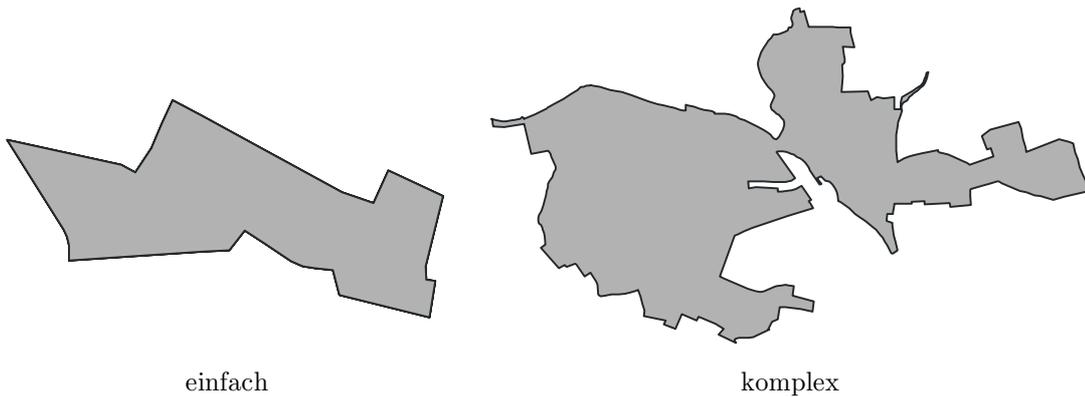


Abbildung 4.18: Polygone zur Untersuchung des modifizierten ICP-Algorithmus'.

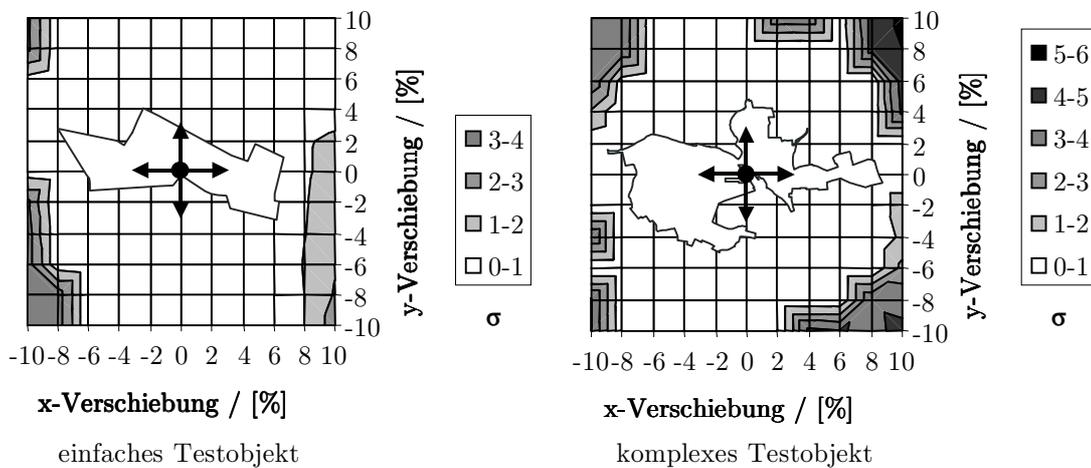


Abbildung 4.19: Ergebnisse der Untersuchungen zum Verhalten des modifizierten ICP-Algorithmus' unter Verschiebung der Objekte zueinander in  $x$ - und  $y$ -Richtung. Die Verschiebungen sind relativ zur Objektabmessung angegeben. Dargestellt ist die Standardabweichung nach 50 Iterationen in Abhängigkeit von der relativen Verschiebung.

näher als 10% ihrer Ausdehnung aneinander bringt, so ist bei Einsatz des modifizierten ICP-Algorithmus' mit Konvergenz zu rechnen, d.h., die individuelle Objekttransformation kann noch verbessert werden.

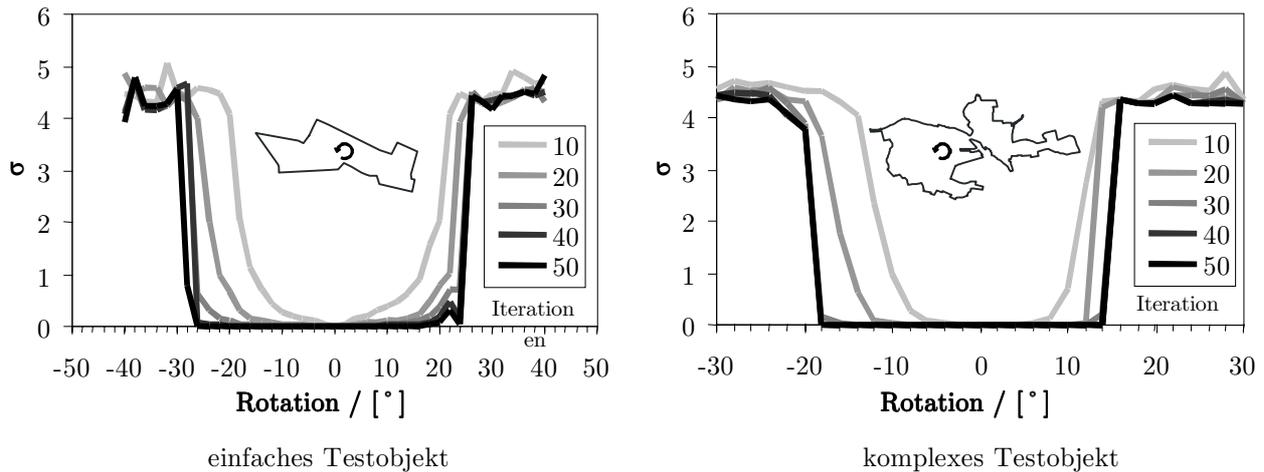
#### 4.8.4 Einfluß von Rotation

Ziel dieser Untersuchung war es, den Einfluß der Rotation zwischen Quell- und Zielobjekt auf den modifizierten ICP-Algorithmus zu untersuchen. Auch hier wurde für alle Rotationen mit  $f = 2$  gearbeitet. Die Rotationen wurden in beide Richtungen in Schritten von  $2^\circ$  angebracht. Die Ergebnisse für das einfache und das komplexe Testobjekt aus Abbildung 4.18 sind in Abbildung 4.20 dargestellt.

Aus den Ergebnissen dieser Untersuchung zeigt sich, daß der modifizierte ICP-Algorithmus für Rotationen zwischen Quell- und Zielobjekt von etwa  $\pm 15^\circ$  sehr gut konvergiert, d.h. die individuelle Transformation von Objekten, die durch eine Näherungstransformation um weniger als  $\pm 15^\circ$  gegeneinander verdreht sind, wird durch den modifizierten ICP-Algorithmus verbessert.

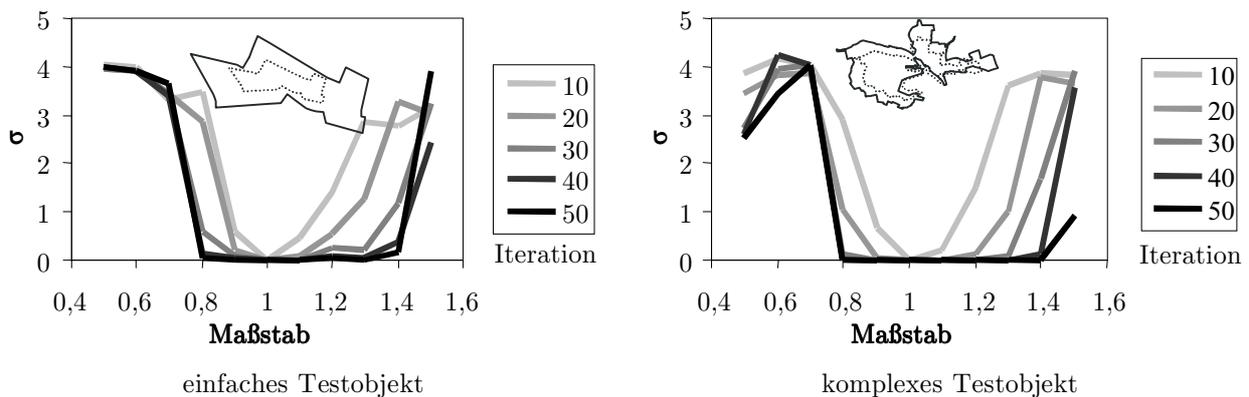
#### 4.8.5 Einfluß des Maßstabs

Der Einfluß einer Skalierung zwischen Quell- und Zielobjekt wird mit dieser Untersuchung anhand der Testobjekte aus Abbildung 4.18 untersucht. Hier gilt ebenfalls für alle Messungen  $f = 2$ . Ausgehend von einem



**Abbildung 4.20:** Einfluß einer Rotation des Quellobjekts gegen das Zielobjekt. Es sind jeweils die Verläufe der Standardabweichungen der Residuen in den Vertices für 10 - 50 Iterationen abgetragen.

Maßstab 1:1 wurden Vergrößerungen und Verkleinerungen jeweils in 10%-Schritten angebracht und die Standardabweichungen der Residuen in den Vertices berechnet. Die Ergebnisse dieser Untersuchung zeigt Abbildung 4.21.

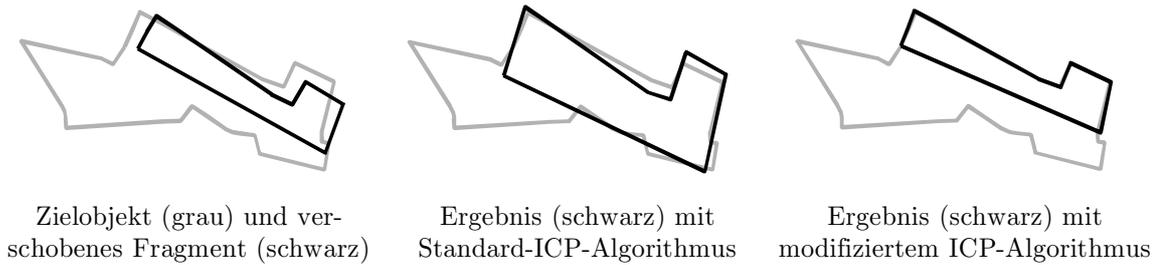


**Abbildung 4.21:** Einfluß einer Maßstabsänderung des Quellobjekts auf die Standardabweichung der Residuen in den Vertices. Dargestellt sind die Graphen für 10-50 ICP-Iterationen.

Für beide Testobjekte zeigt sich übereinstimmend eine exzellente Konvergenz des ICP-Algorithmus' für Skalenfaktoren im Bereich 0,8 – 1,4. Dieses Ergebnis bedeutet für die vorliegende Arbeit, daß der Maßstabsfaktor der Näherungstransformation pro Paar in diesem Bereich liegen muß, um sicher eine Verbesserung der individuellen Transformation bringen zu können.

Die Ergebnisse der angestellten Untersuchungen und das Beispiel aus Abschnitt 4.6 zeigen, daß die globale affine Näherungstransformation, die aus den konsistenten Paaren der Baumsuche gemittelt wird, innerhalb der Konvergenzbereiche des modifizierten ICP-Algorithmus' liegt und damit eine Optimierung der individuellen Affintransformation bewirkt.

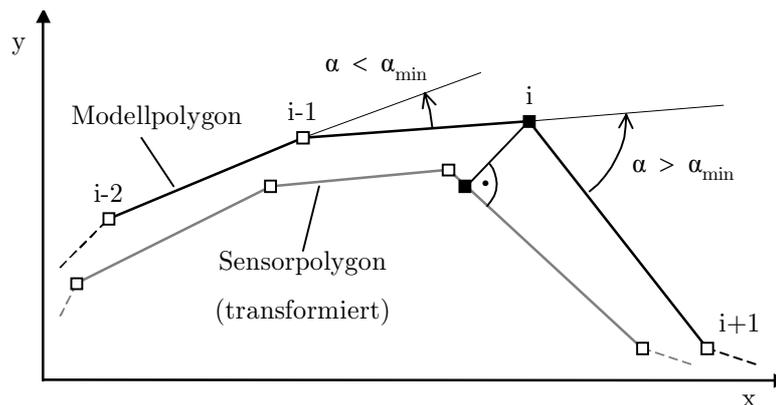
Abschließend zu diesem Abschnitt sei noch einmal das Beispiel für das Versagen des unmodifizierten ICP-Algorithmus' aus Abbildung 4.16 aufgegriffen. Wendet man den modifizierten ICP-Algorithmus an, so ergibt sich das in Abbildung 4.22 dargestellte Verhalten (50 Iterationen,  $f = 1$ ). Wie sehr gut zu erkennen ist, liefert die modifizierte Version nun die korrekte Registrierung für das Quellpolygon, das hier zu Untersuchungszwecken nur ein Fragment des Zielpolygons ist.



**Abbildung 4.22:** Vergleich der Ergebnisse aus Standard- (mittleres Bild) und modifiziertem ICP-Algorithmus (rechtes Bild). Als Testdatensatz wurde das Fragment (linkes Bild, schwarz) des Zielpolygons (grau) etwas verschoben und gedreht.

## 4.9 Bestimmung homologer Punktpaare

Abschließend wird in diesem Abschnitt die Gewinnung homologer Punktpaare (z.B. Paßpunkte im Fall von Sensor- und GIS-Daten) aus den verifizierten und bezüglich ihrer individuellen Transformation optimierten Objektpaaren behandelt. Infolge der polygonalen Natur der untersuchten Objekte stützt sich die angewandte Strategie dabei auf die Vertices der Modelldatenpolygone. Prinzipiell wird für jeden Vertex jedes Modelldatenpolygons der räumlich nächste Punkt auf dem Sensordatenpolygon gesucht. Da Polygonpunkte, die auf Geraden liegen, schlecht als Paßpunkte geeignet sind, werden zunächst nur diejenigen Punkte ausgewählt, an denen die Richtungsänderung der beiden Kanten einen gewissen Wert nicht unterschreitet (siehe  $\alpha_{\min}$  in Abbildung 4.23). Für jeden dieser Vertices wird dann wie bei der Durchführung des ICP-Algorithmus' der räumlich nächste Punkt auf dem Sensordatenpolygon bestimmt. Die letztlich ausgegebene Liste homologer Punktpaare kann noch weiter beschränkt werden. Dazu werden die homologen Paare nach zunehmendem Punktabstand im Modellkoordinatensystem sortiert. Die ausgegebene Punktliste erfolgt dann, beginnend mit dem Paar mit dem geringsten Punktabstand, entweder bis zu einer maximalen Punktzahl oder bis zu einer maximalen Standardabweichung der Residuen. Als Erweiterung bietet sich hier an, die Auswahl homologer Punktpaare zusätzlich bezüglich guter räumlicher Verteilung zu treffen, da diese ebenfalls ein wichtiges Kriterium für eine Paßpunktmenge darstellt. Im implementierten System wurde dies jedoch nicht realisiert, da die Auswahl an Objektpaaren mit den größten Paaren beginnt und daher - infolge nicht vorhandener Objektüberschneidung - von vornherein eine weitgehend gleichmäßige räumliche Verteilung der Objekte wahrscheinlich ist.



**Abbildung 4.23:** Veranschaulichung der Paßpunkterzeugung im Geländekoordinatensystem. Das Sensorpolygon wurde mit der ICP-optimierten Transformation in das Geländesystem übertragen.

## Kapitel 5

# Untersuchungen mit dem Gesamtsystem

## 5.1 Annäherung der Zentralperspektive durch eine lineare Transformation

Als Grundlage für die in den folgenden Abschnitten dargestellten Untersuchungen wird zunächst untersucht, inwieweit die Zentralprojektion des Raumes durch eine lineare Transformation angenähert werden kann. Für diese Arbeit ist insbesondere von Interesse, wann eine lineare Transformation zwischen Bildkoordinaten und der  $XY$ -Ebene in Landeskoordinaten hinreichend genau ist. In diesem Fall läßt sich ein Satelliten- oder Luftbild prinzipiell linear entzerren und georeferenzieren. Die Frage, ob und wann eine solche Näherung zulässig ist, ist von grundlegender Bedeutung für diese Arbeit, da von ihrer Beantwortung die zu wählende Strategie abhängt. Beschreibt eine lineare Transformation zwischen Bildebene und einer Ebene im Objektraum das vorliegende Problem hinreichend genau, dann reicht es für die merkmalsbasierte Objektzuordnung aus, 2D-Merkmale zu verwenden, die invariant unter linearer Transformation sind. Diese sind einfacher zu extrahieren. Im negativen Fall müssen die zu extrahierenden Merkmale für eine zuverlässige Objektzuordnung invariant unter der gegebenen nichtlinearen Transformation sein. Zur Beantwortung dieser zentralen Frage werden nun die Abbildungsverhältnisse quantitativ erfasst.

In der Photogrammetrie ist der Zusammenhang zwischen Bild- und Objektkoordinaten im 3D-Raum durch die Kollinearitätsgleichungen gegeben [Kraus 1997]. Für eine Zentralprojektion seien die drei Parameter der inneren Orientierung in Form der Kammerkonstante  $c$  und des Bildhauptpunkts  $H(\xi_0, \eta_0)$  gegeben. Die äußere Orientierung ist durch Angabe einer reinen Rotationsmatrix  $\mathbf{R}$  im Raum (drei unabhängige Parameter, z.B. die Drehwinkel  $\omega, \phi, \kappa$ , und  $\det \mathbf{R} = 1$ ) und die drei Raumkoordinaten des Projektionszentrums  $O(X_0, Y_0, Z_0)$  gegeben. Die räumliche Drehmatrix  $\underline{R}$  beschreibt die räumliche Stellung des Bildes zum Objektkoordinatensystem ( $XYZ$ ). Bei Benutzung der drei Drehwinkel  $\omega$  (primäre Drehung um die  $x$ -Achse),  $\phi$  (sekundäre Drehung um die  $y$ -Achse) und  $\kappa$  (tertiäre Drehung um die  $z$ -Achse) nimmt die räumliche Drehmatrix  $\underline{R}$  folgende Gestalt an:

$$\mathbf{R}_{\omega\phi\kappa} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} = \begin{pmatrix} \cos \phi \cos \kappa & -\cos \phi \sin \kappa & \sin \phi \\ \cos \omega \sin \kappa + \sin \omega \sin \phi \cos \kappa & \cos \omega \cos \kappa - \sin \omega \sin \phi \sin \kappa & -\sin \omega \cos \phi \\ \sin \omega \sin \kappa - \cos \omega \sin \phi \cos \kappa & \sin \omega \cos \kappa + \cos \omega \sin \phi \sin \kappa & \cos \omega \cos \phi \end{pmatrix}. \quad (5.1)$$

Jedem Objektpunkt  $P(X, Y, Z)$  ist durch die Kollinearitätsgleichungen ein eindeutiger zentralperspektiver Bildpunkt  $P'(\xi, \eta)$  zugeordnet. Für die Koordinaten im Bildraum gilt:

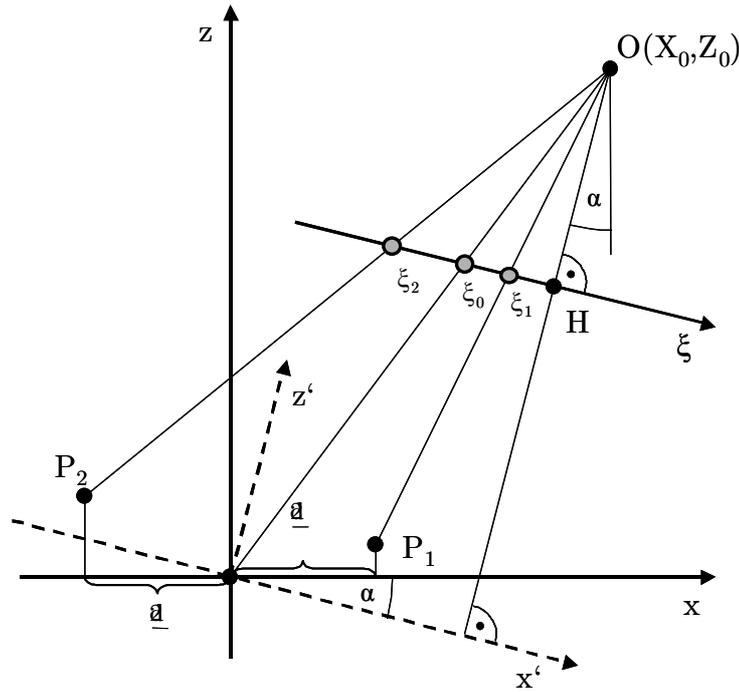
$$\xi = \xi_0 - c \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} \quad (5.2)$$

$$\eta = \eta_0 - c \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)}$$

Zur Klärung der Frage, inwieweit eine lineare Abbildung die zentralperspektive Abbildung hinreichend gut beschreibt ist es nützlich, den durch die Näherung entstehenden Fehler abzuschätzen. Für diese Abschätzung wird zugunsten der Übersichtlichkeit eine zweidimensionale Betrachtung der zentralprojektiven Abbildung gemäß Abbildung 5.1 durchgeführt.

In dieser vereinfachten Betrachtung sind jedoch die wesentlichen Tatsachen enthalten. Wie in Abbildung 5.1 zu sehen ist, reduziert sich die räumliche Drehmatrix auf eine Drehung um die (nicht eingezeichnete)  $y$ -Achse, hier repräsentiert durch den Drehwinkel  $\alpha$ . Es werden zwei Punkte  $P_1$  und  $P_2$  betrachtet, die mit ihren  $x$ -Koordinaten  $X_1 = \frac{d}{2}$  bzw.  $X_2 = -\frac{d}{2}$  symmetrisch zum Koordinatenursprung liegen. Die jeweiligen Höhen  $Z_1$  und  $Z_2$  seien beliebig, jedoch kleiner als die Aufnahmehöhe  $Z_0$ . Zur weiteren Vereinfachung habe der Bildhauptpunkt die Bildkoordinate  $\xi_0 = 0$ . Die linear transformierten Bilder der Punkte  $P_1$  und  $P_2$  haben die Bildkoordinaten  $\xi_{1,linear}$  bzw.  $\xi_{2,linear}$ . Das Bild des Ursprungs des  $XZ$ -Systems besitzt die Bildkoordinate  $\xi_{O,linear}$ . Unter einer beliebigen linearen Transformation zwischen Objekt- und Bildkoordinatensystem wird das Streckenverhältnis auf der  $x$ -Achse unverändert in den Bildraum übertragen. Es gilt also im linearen Abbildungsfall allgemein die Beziehung:

$$\frac{X_1}{X_2} = \frac{\xi_{1,linear} - \xi_{O,linear}}{\xi_{2,linear} - \xi_{O,linear}}. \quad (5.3)$$



**Abbildung 5.1:** Zweidimensionale perspektive Abbildung zweier auf der  $x$ -Achse zum Ursprung äquidistanter Punkte.

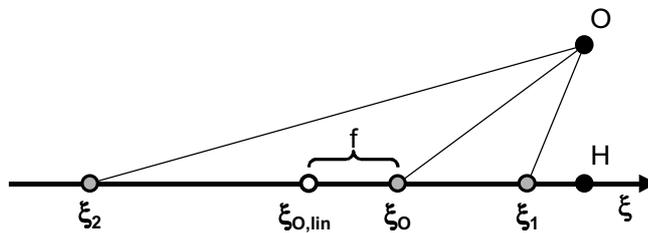
Speziell im gewählten Fall ist bei linearer Transformation zwischen Bildachse und  $x$ -Achse  $\frac{X_1}{X_2} = -1$  und damit

$$\xi_{1,linear} - \xi_{O,linear} = \xi_{O,linear} - \xi_{2,linear} \quad \text{bzw.} \quad \xi_{1,linear} + \xi_{2,linear} - 2\xi_{O,linear} = 0. \quad (5.4)$$

Im Falle der zentralperspektiven Abbildung besitzen die Bilder der Punkte  $P_1, P_2$  und  $O$  die Bildkoordinaten  $\xi_1, \xi_2$  und  $\xi_O$ . Als Maß für den Bildfehler, den man mit der Annahme einer linearen Transformation zwischen  $x$ -Achse und  $\xi$ -Bildachse erhält, läßt sich die folgende Größe definieren:

$$f = \frac{1}{2}(\xi_1 + \xi_2) - \xi_O. \quad (5.5)$$

Die Abbildung 5.2 illustriert das in Gleichung 5.5 definierte Fehlermaß  $f$ .



**Abbildung 5.2:** Abschätzung für den Bildfehler  $f$  zwischen 3D-zentralperspektiver und 2D-linearer Transformation.

Im vereinfachten Fall reduzieren sich die Kollinearitätsgleichungen 5.2 auf eine Gleichung für die Bildachse, die hier  $\xi$  sein soll:

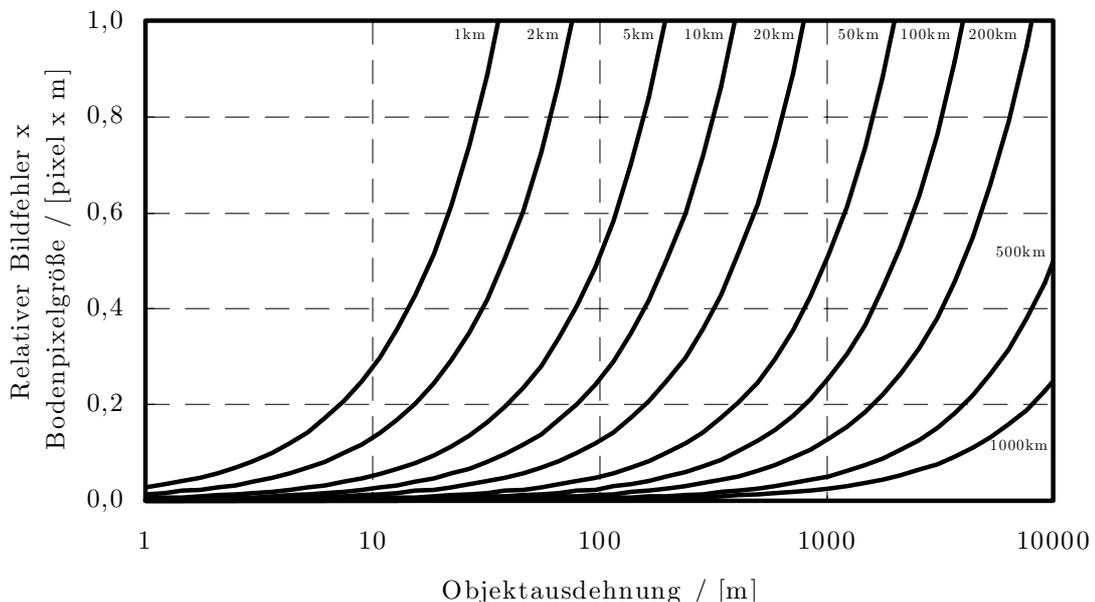
$$\xi = -c \frac{(X - X_0) \cos \alpha - (Z_1 - Z_0) \sin \alpha}{(X - X_0) \sin \alpha + (Z - Z_0) \cos \alpha}. \quad (5.6)$$

Setzt man die Punkte  $P_1(\frac{d}{2}, Z_1), P_2(-\frac{d}{2}, Z_2)$  und  $O(0,0)$  über die reduzierte Kollinearitätsgleichung 5.6 in die Fehlerbeziehung aus Gleichung 5.5 ein, so erhält man den Fehler  $f(X_0, Z_0, d, Z_1, Z_2, \alpha)$  in Abhängigkeit von den

Koordinaten des Projektionszentrums  $O(X_0, Z_0)$ , dem  $x$ -Punktabstand  $d$ , den Punkthöhen  $Z_1$  und  $Z_2$  sowie dem Sensorneigungswinkel  $\alpha$ :

$$f = c \frac{X_0 \cos \alpha - Z_0 \sin \alpha}{X_0 \sin \alpha + Z_0 \cos \alpha} - \frac{c}{2} \left( \frac{\left(\frac{d}{2} - X_0\right) \cos \alpha - (Z_2 - Z_0) \sin \alpha}{\left(\frac{d}{2} - X_0\right) \sin \alpha + (Z_2 - Z_0) \cos \alpha} + \frac{\left(\frac{d}{2} + X_0\right) \cos \alpha - (Z_2 - Z_0) \sin \alpha}{\left(\frac{d}{2} + X_0\right) \sin \alpha + (Z_2 - Z_0) \cos \alpha} \right). \quad (5.7)$$

Das Verhalten des Produktes aus Bildfehler und Bodenpixelgröße bei Annahme einer linearen Transformation zwischen  $x$ -Achse in Landeskoordinaten und  $\xi$ -Achse im Bild ist in Abbildung 5.3 dargestellt. In dem gewählten Beispiel ist auf der Abszisse der  $x$ -Punktabstand  $d$ , der als Objektausdehnung interpretiert wird, abgetragen. Als Parameter für die Kurvenschar wurde die Aufnahmehöhe  $Z_0$  gewählt. Festgehalten wurden für diesen Graphen die Punkthöhen  $Z_1 = 0\text{m}$  und  $Z_2 = 100\text{m}$  sowie die Kameraneigung  $\alpha = 0^\circ$  und die  $x$ -Koordinate des Projektionszentrums ( $X_0 = 0$ ).



**Abbildung 5.3:** Produkt aus Bildfehler in Pixeln und Bodenpixelgröße in m in Abhängigkeit von der Objektgröße. Scharparameter ist die Aufnahmehöhe.

Die Graphen in Abbildung 5.3 sind folgendermaßen zu lesen: Angenommen, es liegt eine Sensor vor, der eine Bodenpixelgröße von 0.25m bei einer Flughöhe von 2000m besitzt, wie es z.B. im panchromatischen Kanal der DPA-Kamera (DPA steht für Digital Photogrammetric Assembly) der Fall ist. Für ein Objekt, das auf 100m Ausdehnung einen Höhenunterschied von 100m aufweist (z.B. ein Waldgebiet an einem Steilhang) beträgt der Bildfehler bei Normalfall-Bedingungen:  $\frac{0.27[\text{pix}][\text{m}]}{0.25[\text{m}]} \text{ (Ablesewert aus der Kurve für } Z_0=2\text{km)} = 1.08\text{pixel}$ . Für das selbe Objekt unter den selben Bedingungen, jedoch bei einer Aufnahmehöhe von 832km und einer Bodenpixelgröße von 10m (panchromatischer SPOT-Kanal), liegt der Bildfehler bei etwa  $10^{-5}\text{pixel}$ . Eine Veränderung der bisher festgehaltenen Parameter wirkt sich folgendermaßen aus:

Aktion	Auswirkung
Vergößerung der Sensorneigung $\alpha$	Streckung der Kurvenschar in $y$ -Richtung, größerer Fehler
Vergößerung von $X_0$	Auseinanderziehen der Kurven in $y$ -Richtung, größerer Fehler
Vergößerung des Höhenunterschiedes $\Delta Z$	Streckung der Kurvenschar in $y$ -Richtung, größerer Fehler

**Tabelle 5.1:** Einflußverschiedener Parameter auf den relativen Bildfehler bei verschiedenen Flughöhen.

Mit dieser einfachen Betrachtung läßt sich zusammenfassend feststellen, daß zumindest für Satellitenaufnahmen bei gängigen Aufnahmekonfigurationen und Bodenpixelgrößen über 1m für viele Geo-Objekte (Ausdehnungen bis 10km, Höhenunterschied wenige 100m) eine lineare Transformation vom Bildkoordinatensystem in die  $XY$ -Ebene eines Landeskoordinatensystems mit Restfehlern von weniger als einem pixel durchführbar ist. Diese

Aussage bildet die Grundlage für eine Verwendung von Merkmalen, die invariant unter linearen Transformationen, insbesondere der Affintransformation, sind. Inwieweit die Ergebnisse, z.B. automatisch gemessene Paßpunkte, mit manuellen Messungen konkurrieren können ist Gegenstand der weiteren Untersuchungen.

## 5.2 Das entwickelte System

Aufbauend auf den in den vorigen Kapiteln angestellten Betrachtungen wurde im Rahmen dieser Arbeit ein Gesamtsystem zur automatischen Registrierung zweier Raster- oder Vektor-Datensätze anhand polygonaler Objekte implementiert. Je nach Art der Eingabedaten können mit diesem System unterschiedliche Registrierungsaufgaben gelöst werden. Es werden folgende Fälle unterschieden:

- Raster-Raster-Registrierung:  
z.B. Näherungstransformation für Stereo-Bildpaare, multisensorale Registrierung von Fernerkundungsdaten gleicher oder unterschiedlicher Auflösung (Data Fusion), multitemporale Registrierung von Fernerkundungsdaten, Georeferenzierung von Fernerkundungsdaten mit Rasterkarten.
- Raster-Vektor-Registrierung:  
z.B. Georeferenzierung von Fernerkundungsdaten mit Vektor-GIS-Daten (hier ATKIS).
- Vektor-Vektor-Registrierung:  
z.B. Räumliche Integration verschiedener Vektor-GIS-Daten, Simulationen.

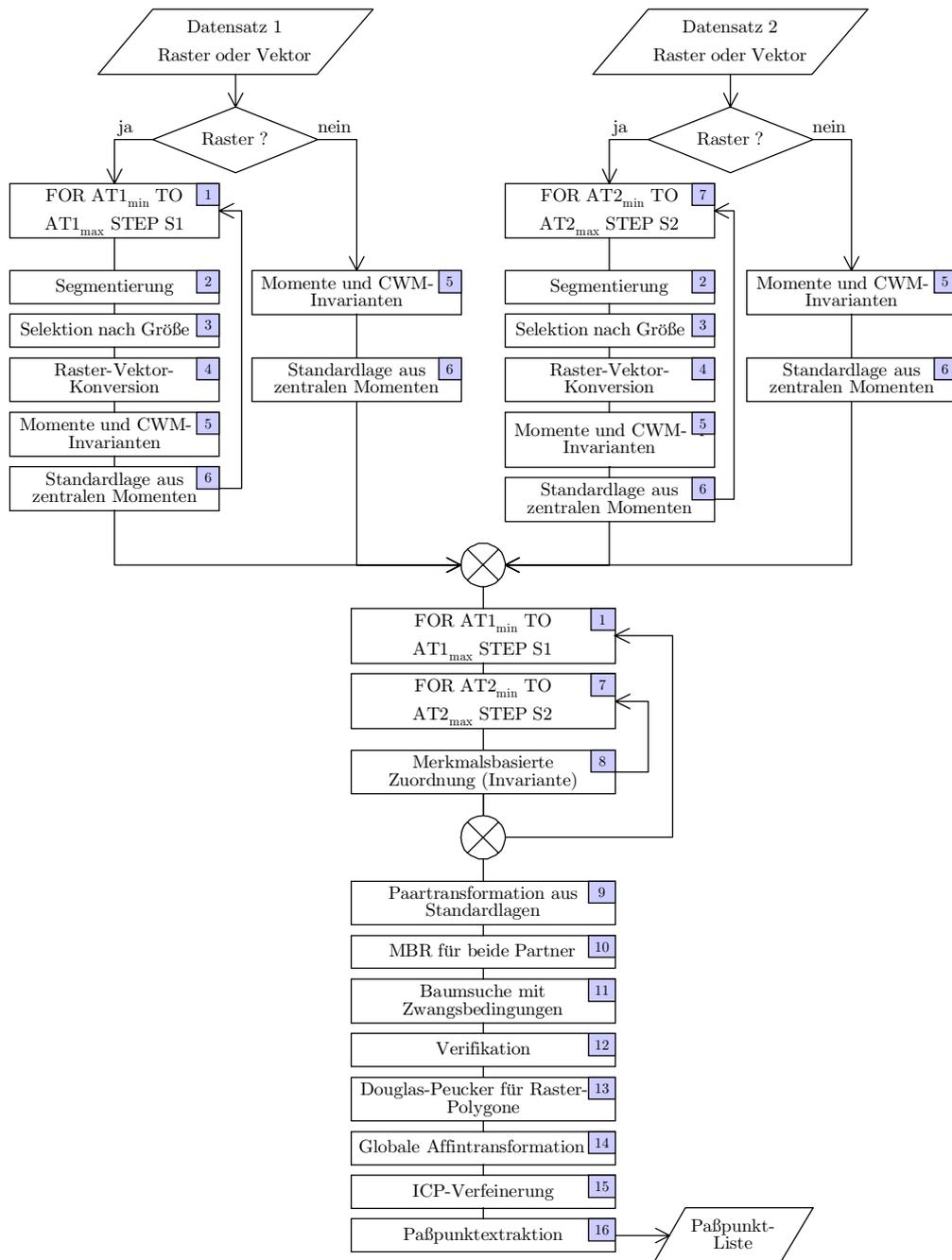
In Abbildung 5.4 ist ein Flußdiagramm des Gesamtsystems dargestellt, die einzelnen Prozesse und ihre für den Ablauf relevanten Parameter werden mit Tabelle 5.2 beschrieben. Zur besseren Übersicht ist jeweils das Kapitel oder der Abschnitt, in dem der betreffende Prozess detailliert beschrieben bzw. untersucht wurde, mit angegeben.

Die Eingabedaten können jeweils in Raster- oder in Vektorform vorliegen. Rasterdaten werden - je nach gewählten Parametern - zur Bildung von Regionen mit homogener Grauwertverteilung mehrfach segmentiert (Prozess 2). Dazu wird das in Abschnitt 2.1.2 beschriebene Bereichswachstumsverfahren mit adaptivem Schwellwert eingesetzt. Zur Reduktion der Objektanzahl werden daran anschließend Objekte, die kleiner als eine bestimmte Mindestgröße sind, eliminiert (Prozess 3). Die verbleibenden Regionen werden dann in Vektor-Polygone gewandelt (Prozess 4). Löcher bleiben dabei unberücksichtigt. Die Vektorpolygone folgen zunächst den äußeren Pixelkanten.

Für alle so aufbereiteten Segmentierungsdurchläufe und für die Polygone aus originären Vektordaten werden zuerst alle geometrischen und dann alle zentralen Momente bis zur achten Ordnung direkt aus der Kontur berechnet (vgl. Abschnitt 3.3.1). Aus den zentralen Momenten werden für alle Objekte die sechs einfachsten Kreuzmoment-Invarianten berechnet (Prozess 5, vgl. Abschnitt 3.2). Aus den zentralen Momenten bis einschliesslich vierter Ordnung wird für jedes Objekt die Transformation in eine Standardlage mit Hilfe der Iterationsmethode (Abschnitt 4.5.3) berechnet (Prozess 6).

Zur Bildung von Objektpaarhypothesen wird dann in allen möglichen Kombinationen von Segmentierungsergebnissen und Vektordatensätzen ein merkmalsbasiertes Zuordnungsverfahren eingesetzt (Prozess 8). Dabei werden Objektpaare erzeugt, deren relativer Unterschied für jede Affininvariante einen gewissen Wert nicht überschreitet (vgl. Abschnitt 4.4). Damit wird bereits vorab eine unäre Zwangsbedingung für die in Prozess 9 durchgeführte Baumsuche mit Zwangsbedingungen (vgl. Abschnitt 4.1 und 4.2) abgeprüft. Dadurch kann der Suchaufwand für die später folgende Baumsuche erheblich reduziert werden. Für alle Objektpaarhypothesen wird die affine Hin- und Rücktransformation von einem Objekt auf das andere über deren Standardlage berechnet (Prozess 9). Vor der Baumsuche wird in Prozess 10 noch für jedes Objekt das minimal berandende Rechteck (MBR) in iterativer Weise berechnet.

Im Rahmen der Baumsuche (Prozess 11) werden bis zu einer gewissen Suchtiefe Objektpaare ermittelt, die im Sinne der implementierten Zwangsbedingungen konsistent sind. Aus diesen Paaren wird aus der paarweise individuellen affinen Hin- und Rücktransformation eine globale affine Hin- und Rücktransformation zwischen Datensatz 1 und Datensatz 2 gemittelt. Die Mittelung erfolgt dabei gewichtet mit dem Kehrwert der Objektdistanz im 6D-Merkmalraum, der von den Kreuzmoment-Affininvarianten aufgespannt wird. Anhand dieser globalen Transformationen und Schwellwerte für Rotation, Scherung, Maßstab und Translation werden aus allen Hypothesen durch Prüfung der Überlappung transformierter Objekte die passenden Paare verifiziert und nicht passende eliminiert (Prozess 12). Damit liegen nur noch Objektpaare vor, die bezüglich individueller Rotation,



**Abbildung 5.4:** Flußdiagramm des gesamten im Rahmen dieser Arbeit entwickelten Systems zur vollautomatischen Registrierung von Raster- und Vektordaten.

Scherung, Translation und Skalierung nicht weiter als in spezifizierter Weise von den globalen affinen Transformationen abweichen und deren Flächenüberlappung nach Anwendung der passenden globalen Transformation auf ein Objekt mindestens einen gewissen Prozentsatz der kleineren Objektfläche beträgt.

Im Fall von Rasterdaten wird in Prozess 13 eine Glättung der Objektkontur mit Hilfe des Douglas-Peucker-Algorithmus' [Douglas & Peucker 1973] durchgeführt. Dies reduziert zum einen die Anzahl an Knoten auf der Kontur, zum anderen stimmt nach der Generalisierung auch die Normalenrichtung eher mit der Realität überein (vorher: Raster-Polygone folgen den Pixelkanten!). Beides wirkt sich verbessernd auf die ICP-Verfeinerung in Prozess 15 aus. Im Anschluss an die Generalisierung werden die Objekte eines der beiden Datensätze mit Hilfe der globalen Affintransformation grob auf ihre Partner transformiert.

Für alle verifizierten Objektpaare wird in Prozess 15 die individuelle Transformation ausgehend von der globalen Transformation mit Hilfe eines modifizierten ICP-Algorithmus' verbessert. Dabei wird nach einer gewissen

Prozess	Arbeitsschritt	Prozessrelevante Parameter
1	Schleife für Datensatz 1 (falls Raster)	Adaptiver Start-Schwellwert ( $AT_{\min}$ ), Adaptiver End-Schwellwert ( $AT_{\max}$ ), Schrittweite
2	Segmentierung	Adaptiver Schwellwert $AT$ für maximale Grauwertvarianz pro Region
3	Selektion nach Größe	keine
4	Raster-Vektor-Konversion	keine
5	Momente und CWM-Invarianten	keine
6	Standardlage aus zentralen Momenten	keine
7	Schleife für Datensatz 2 (falls Raster)	Adaptiver Start-Schwellwert ( $AT_{\min}$ ), Adaptiver End-Schwellwert ( $AT_{\max}$ ), Schrittweite
8	Merkmalsbasierte Zuordnung zur Bestimmung von Objektpaarhypothesen, basierend auf CWM- Invarianten	Maximaler relativer Unterschied (Gleichung 3.55) für jedes verwendete Merkmal
9	Paarweise affine Hin- und Rücktransformation aus Standardlagen beider Objekte	keine
10	Berechnung der minimalen berandenden Rechtecke (MBR) für beide Partner	Winkelinkrement in Grad (Abschnitt 3.4.1)
11	Baumsuche mit Zwangsbedingungen zur Ermittlung einer konsistenten Interpretation aus allen Objektpaarhypothesen und einer gemittelten, globalen Affintransformation zwischen Datensatz 1 und Datensatz 2. Die globale Transformation wird aus den Paartransformationen (Prozess 9) der konsistenten Interpretation gewichtet gemittelt	Minimale Überlappung für konsistente Paare, Suchtiefe (Kapitel 4)
12	Verifikation aller Objektpaarhypothesen mit Hilfe der globalen Transformation aus Prozess 11	Maximal zulässige Werte für Rotation und Scherung, $x$ - und $y$ -Maßstab, $x$ - und $y$ -Verschiebung, minimaler prozentualer-Überlapp
13	Generalisierung der Raster-Polygone mit Hilfe des Douglas-Peucker-Algorithmus <sup>7</sup>	Generalisierungsmaß
14	Transformation der generalisierten Raster-Polygone mit globaler Transformation aus Prozess 11	keine
15	Individuelle Verfeinerung der Transformation für jedes verifizierte Objektpaar. Ergebnis ist eine 2D-projektive Transformation	Sampling-Abstand auf Polygonrand, Anzahl der Iterationen, $\sigma$ -Faktor $f$ (Abschnitt 4.8.2)
16	Extraktion von homologen Punkten in Datensatz 1 und Datensatz 2 = Paßpunkte.	Entweder maximale Paßpunktzahl oder maximale Standardabweichung der Paßpunktresiduen, minimale Richtungsänderung am Vertex (Abschnitt 4.9)

**Tabelle 5.2:** Übersicht über alle Prozesse des entwickelten Verfahrens und deren prozessrelevante Parameter.

Zahl von Iterationen aus allen verbleibenden Punkten auf der Objektkontur und ihren nächsten Nachbarn auf der Partnerobjektkontur eine 2D-projektive Transformation mit acht Parametern geschätzt. Nach erfolgreich verlaufener ICP-Optimierung werden mit Prozess 16 homologe Punktpaare aus den beiden Datensätzen nach dem in Abschnitt 4.9 beschriebenen Verfahren extrahiert.

### 5.3 Untersuchung des Systems bei realer Topographie

In diesem Abschnitt wird das Verhalten des Gesamtsystems unter realen topographischen Verhältnissen untersucht. Zur Separation der Fehler, die durch eine fehlerhafte Segmentierung hervorgerufen werden, von Nichtaffinitäten infolge nichtebenen Geländes (perspektive Bildverzerrungen) wurde ein spezieller 3D-Testdatensatz

erzeugt. Von diesem Testdatensatz wurden systematisch virtuelle Luftaufnahmen in Vektorform berechnet, die mit dem entwickelten System bestmöglich mit den vorhandenen 2D-ATKIS-Daten registriert werden. Die gesamte Untersuchung stellt also eine Vektor-Vektor-Registrierung simulierter Bilddaten mit realen ATKIS-Daten dar.

### 5.3.1 Der Testdatensatz

Für die Simulation zentralprojektiver Aufnahmen wurden polygonale ATKIS-Objekte zunächst objektartenweise verschmolzen und dann über das DHM (10m-Raster) des Untersuchungsgebiets gelegt. Der vollaufgelöste Datensatz besteht aus 1733 Polygonen, durch die Verschmelzung benachbarter Polygone gleicher Objektart reduzierte sich die Zahl auf 433 Polygone.

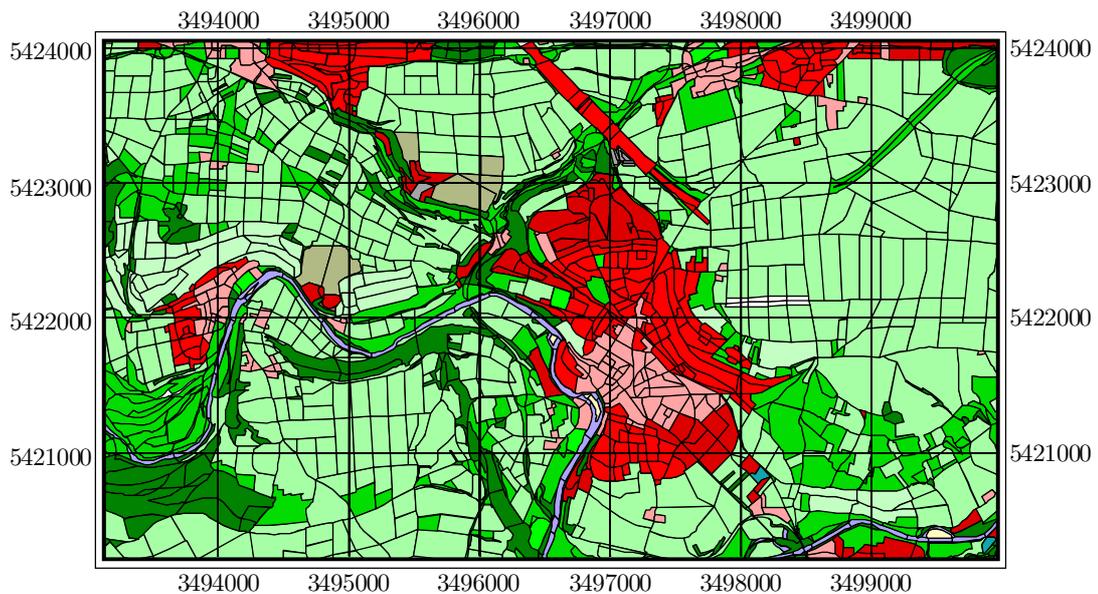


Abbildung 5.5: ATKIS-Daten des Testgebiets Vaihingen/Enz.

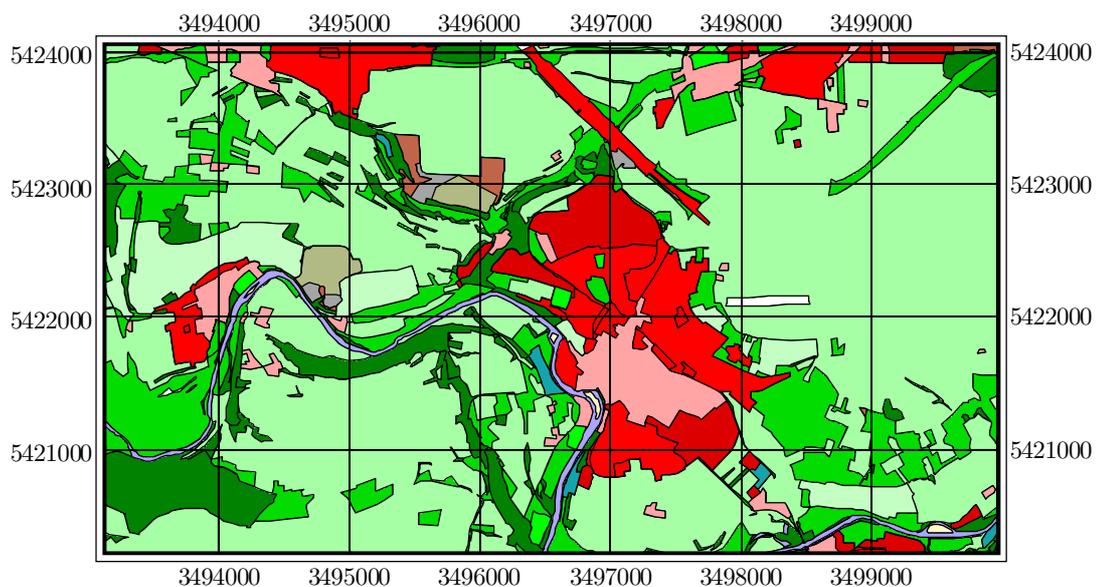
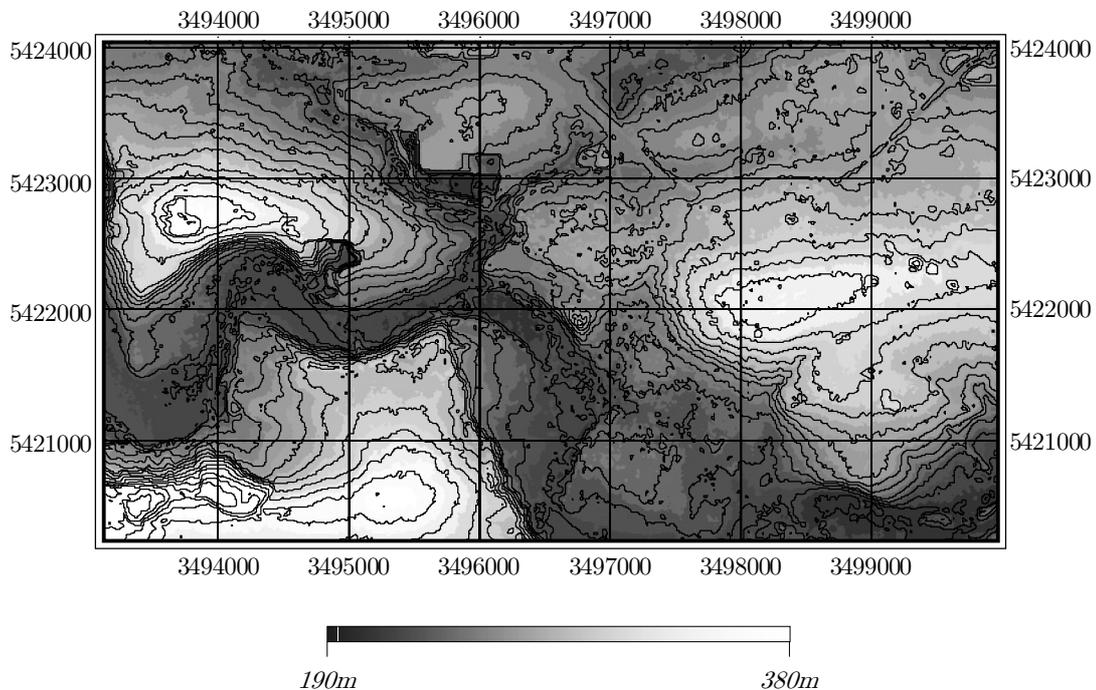


Abbildung 5.6: Objektartenweise verschmolzene polygonale ATKIS-Objekte des Testdatensatzes Vaihingen/Enz.

Dieser Datensatz wird im folgenden als Modelldatensatz bezeichnet. Das Testgebiet liegt bei Vaihingen/Enz in Baden-Württemberg. In Abbildung 5.5 ist der ATKIS-Ausgangsdatensatz in voller Objektartenaufösung abgebildet. Die Koordinaten sind in Gauß-Krüger-Projektion angegeben. Der Testdatensatz hat eine Ausdehnung von 6380m x 3840m und besitzt damit praxisrelevante Größe. Das Resultat der objektklassenweisen Verschmelzung von ATKIS-Polygonen ist in Abbildung 5.6 zu sehen. Das DHM des Geländeausschnittes für den ATKIS-Datensatz ist mit 10m-Höhenlinien in Abbildung 5.7 dargestellt.

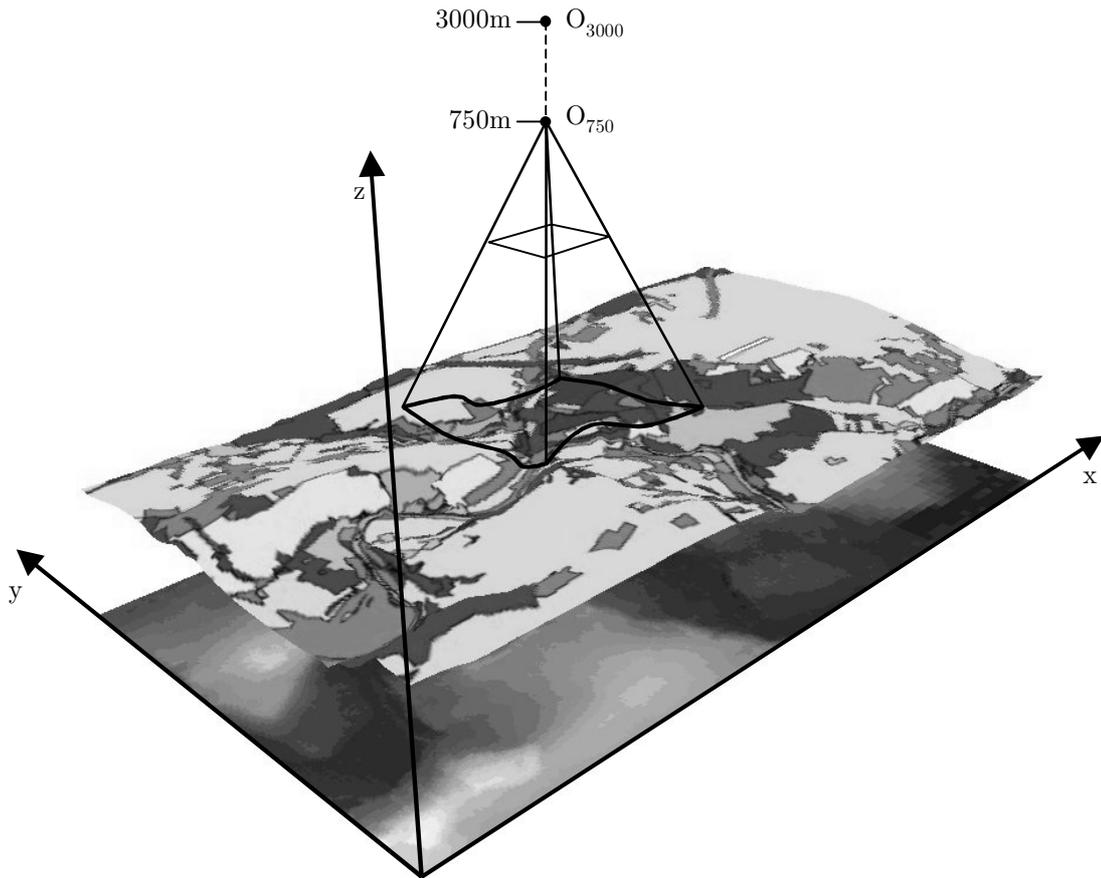


**Abbildung 5.7:** Digitales Höhenmodell des Testdatensatzes Vaihingen/Enz.

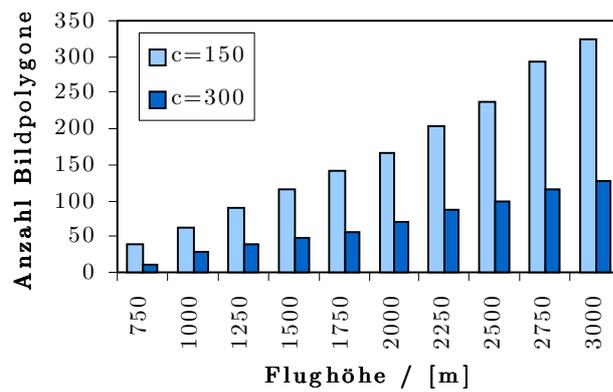
### 5.3.2 Durchführung und Ergebnisse der Untersuchung

Zur Untersuchung des entwickelten Verfahrens wurde von den in oben beschriebener Weise erzeugten 3D-ATKIS-Daten des Untersuchungsgebiets virtuelle Luftbildaufnahmen bei Flughöhen zwischen 750m und 3000m in 250m-Schritten und Kammerkonstanten von  $c = 150\text{mm}$  und  $c = 300\text{mm}$  berechnet. Die gewählten Werte sind realistisch für photogrammetrische Bildflüge, z.B. [Albertz 1991], S.38. Dieses Vorgehen liefert quasi ideal segmentierte Bilder mit realer Abbildungsgeometrie, der reine Geländeeinfluß kann damit unbeeinflusst von Segmentierungsfehlern studiert werden. Für alle berechneten Aufnahmen wurde der photogrammetrische Normalfall angenommen [Kraus 1997], das Projektionszentrum liegt in jedem Fall exakt über der Mitte des ATKIS-Datensatzes. Ferner liegt der Bildhauptpunkt exakt in der Bildmitte und es wurde eine verzeichnungsfreie Bildaufnahme angenommen. Die folgenden Untersuchungen sind trotz aller genannten Idealisierungen praxisrelevant, da die Topographie des Geländes letztlich willkürlich gewählt ist und sich Änderungen in der Topographie weit stärker in Bildverzerrungen niederschlagen würden als realistischere Werte für die innere und äußere Orientierung der virtuellen Kamera. Die Aufnahmegeometrie für die unterschiedlichen virtuellen Luftbildaufnahmen wird mit Abbildung 5.8 anhand der objektartenweise verschmolzenen ATKIS-Polygone veranschaulicht.

Für diese Untersuchung wurde für jede Flughöhe und jede Kammerkonstante zunächst der ganze ATKIS-Datensatz mit den o.g. Orientierungsparametern in die virtuelle Bildebene projiziert. In einem zweiten Schritt wurde dann jeweils der virtuelle Bildbereich von  $230\text{mm} \times 230\text{mm}$  ausgeschnitten. Die Anzahl der Bildpolygone - vollständig oder angeschnitten - für alle Flughöhen und Kammerkonstanten ist in Abbildung 5.9 dargestellt. Zur Veranschaulichung sind die beiden Extremfälle  $h = 750\text{m}$ ,  $c = 300\text{mm}$  und  $h = 3000\text{m}$ ,  $c = 150\text{mm}$  in Abbildung 5.10 und Abbildung 5.11 dargestellt. Bei einer mittleren Geländehöhe von 200m entsprechen diese Extreme einem Bildmaßstabsbereich von 1:1833 bis 1:18667 und beinhalten damit den Bildmaßstab für die DLM25-Kartierung und -Fortführung (1:10000, Quelle: [LVBW 2001]).

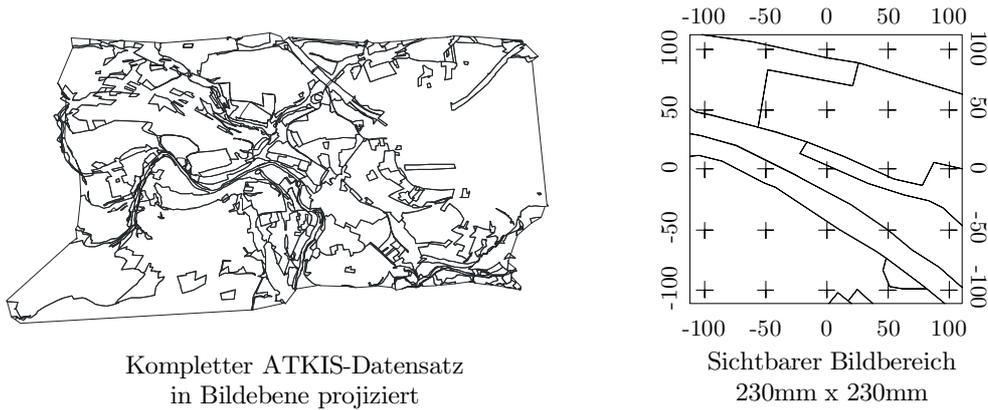


**Abbildung 5.8:** 3D-Ansicht der Geometrie der simulierten Luftbildaufnahmen des Untersuchungsgebiets Vaihingen/Enz. In der  $x$ - $y$ -Ebene ist das benutzte DHM (10m-Raster) abgebildet. Die Aufnahmehöhe variiert zwischen 750m und 3000m. Die Verhältnisse in  $z$ -Richtung sind nicht maßstäblich wiedergegeben.

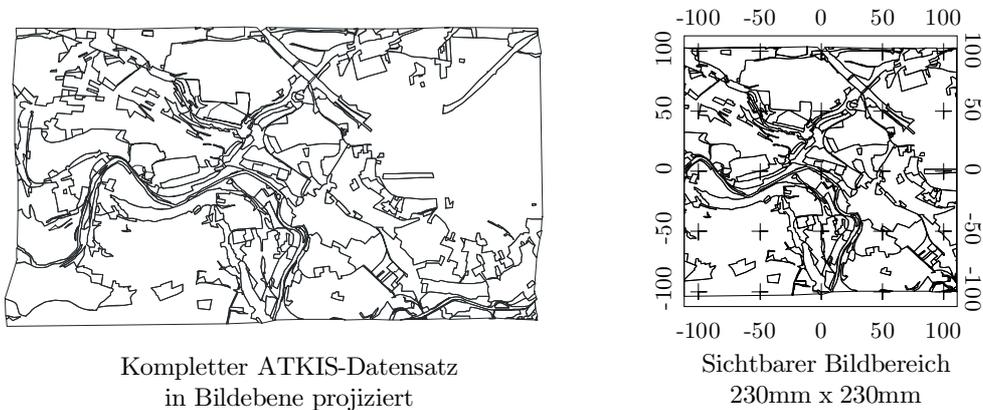


**Abbildung 5.9:** Anzahl der Bildpolygone in Abhängigkeit von der Flughöhe und Kammerkonstante für alle berechneten virtuellen Aufnahmen.

Gemäß der in Abschnitt 5.2 vorgestellten Strategie wurden zunächst für den Modelldatensatz und für alle virtuellen Ansichten für jedes Objekt die zentralen geometrischen Momente bis zur achten Ordnung und die daraus resultierenden sechs einfachsten Kreuzmoment-Affinvarianten berechnet. Zudem wurde, ausgehend von den zentralen Momenten bis vierter Ordnung, für alle Objekte die Standardlage mit der Iterationsmethode berechnet. Zur Erzeugung von Objektpaarhypothesen wurde für jede virtuelle Ansicht das in Abschnitt 4.4 beschriebene Zuordnungsverfahren zur Prüfung der unären Zwangsbedingungen für die Affinvarianten durchgeführt. Der maximal zulässige relative Unterschied wurde dabei für jedes Merkmal zu 3% gesetzt. Anschließend wurde für



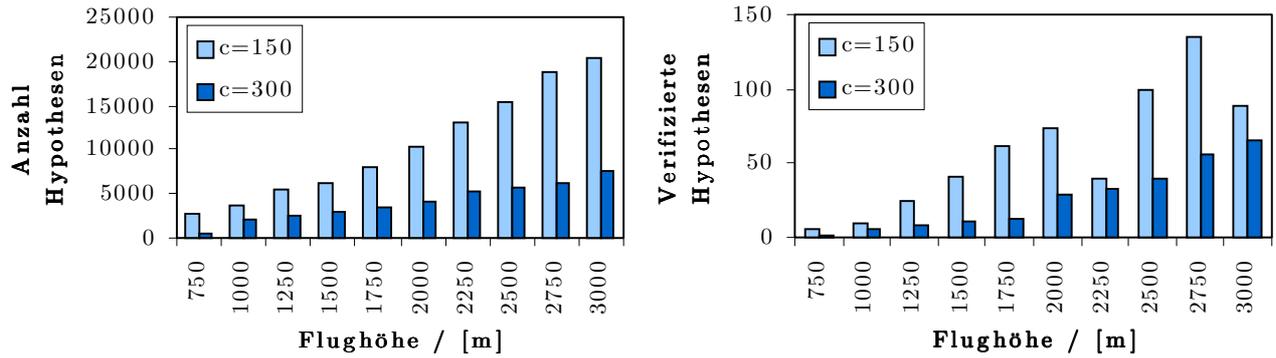
**Abbildung 5.10:** Projektion der verschmolzenen ATKIS-Polygone des Untersuchungsgebiets in die Bildebene. Flughöhe  $h = 750\text{m}$ , Kammerkonstante  $c = 300\text{mm}$ . Diese Aufnahme stellt den größten vorkommenden Bildmaßstab (1:1833) dar.



**Abbildung 5.11:** Projektion der verschmolzenen ATKIS-Polygone des Untersuchungsgebiets in die Bildebene. Flughöhe  $h = 3000\text{m}$ , Kammerkonstante  $c = 150\text{mm}$ . Diese Aufnahme stellt den kleinsten vorkommenden Bildmaßstab (1:18667) dar.

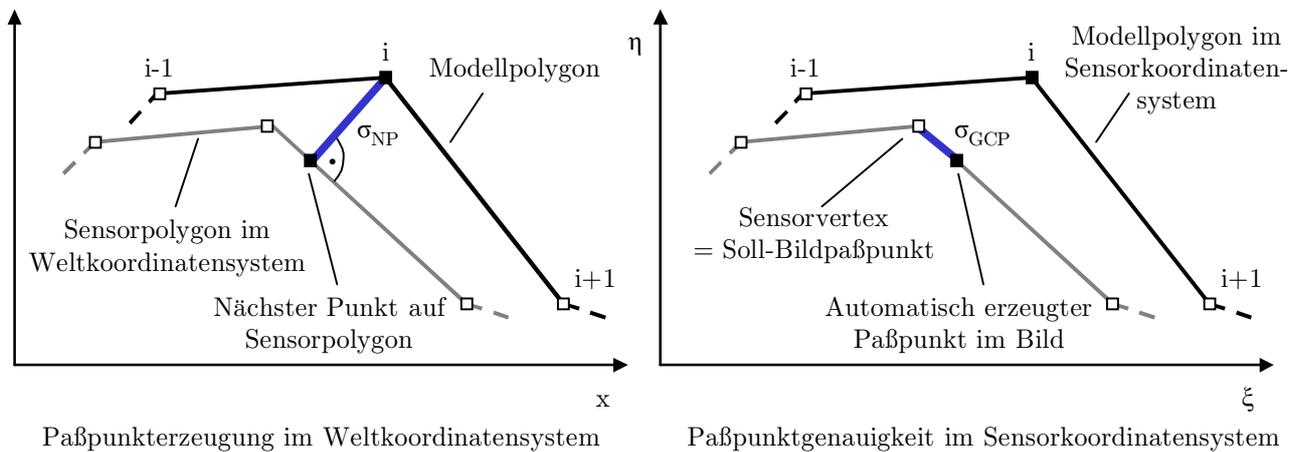
jedes hypothetische Paar aus den Standardlagen der Partner die individuelle Hin- und Rücktransformation sowie die minimal berandenden Rechtecke (MBR) berechnet. Ausgehend von den so erzeugten Objektpaarhypothesen wurde für jede Paarung aus Modelldatensatzobjekt und Objekt in virtueller Aufnahme das beschriebene Baumsuchverfahren mit Zwangsbedingungen zur Ermittlung konsistenter Paare und einer globalen affinen Näherungstransformation durchgeführt. Die maximale Suchtiefe wurde für alle Durchläufe auf 6 beschränkt, der Überlapp zur Aufnahme eines Paares in die Liste konsistenter Paare wurde für alle Läufe zu 70% gesetzt. Zur Verifikation der Hypothesen wurden die virtuellen Sensorobjekte jeweils individuell in das Weltkoordinatensystem transformiert. Ausgehend von der jeweiligen globalen Transformation wurden alle Hypothesen hinsichtlich ihrer Sensor-Welt-Transformation bewertet. Hypothesen mit weniger als  $\pm 30^\circ$  Unterschied in der Rotations- und Scherungskomponente, weniger als  $\pm 30\%$  Unterschied in den  $x$ - und  $y$ -Maßstäben, weniger als  $\pm 30\%$  Unterschied in der  $x$ - und  $y$ -Translation und einem Überlappungsbereich von mindestens 70% beider Objektflächen wurden als relevant verifiziert, alle anderen wurden verworfen. In Abbildung 5.12 ist links die Anzahl der Hypothesen nach der merkmalsbasierten Zuordnung abgebildet, die rechte Grafik zeigt die Anzahl verifizierter Hypothesen für jede virtuelle Aufnahme.

Mit Hilfe des modifizierten ICP-Algorithmus' (Abschnitt 4.8.2) wurden die verifizierten Sensorpolygone ausgehend von der globalen Sensor-Welt-Transformation im Weltkoordinatensystem bestmöglich auf die Modelldatenpolygone transformiert. Einheitlich für alle Aufnahmen wurde dabei mit einem Punkt-Samplingabstand von 10m, einer Iterationsrate von 100 und einer maximalen Punktdistanz von zwei Standardabweichungen ( $f = 2$ , siehe Abschnitt 4.8.2) gearbeitet.



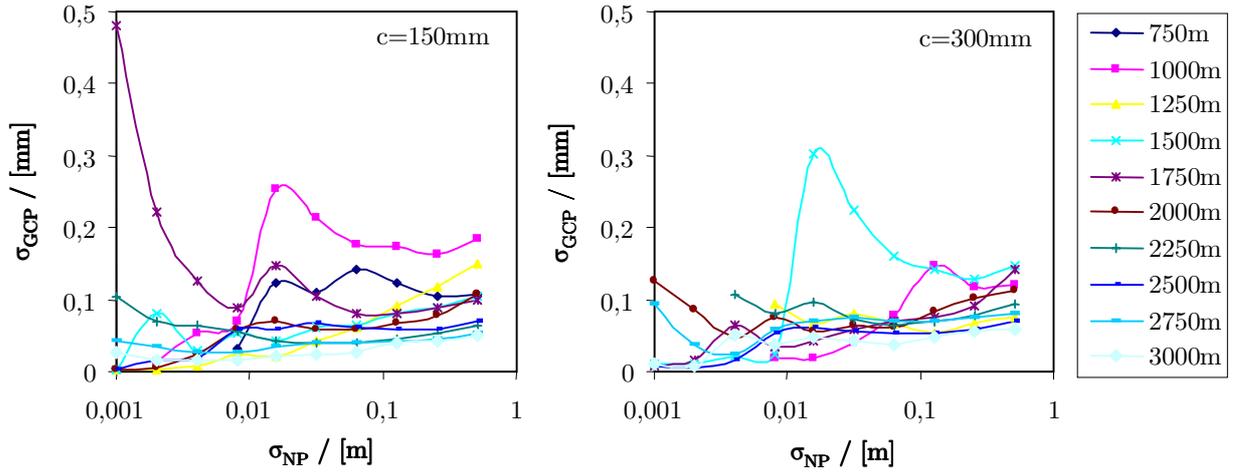
**Abbildung 5.12:** Anzahl an Objektpaarhypothesen nach merkmalsbasierter Zuordnung und nach Verifikation für alle virtuellen Aufnahmen des Untersuchungsgebiets Vaihingen/Enz.

Zur Untersuchung der automatischen Erzeugung von Paßpunkten wurde für jede virtuelle Ansicht das in Abschnitt 4.9 beschriebene Verfahren mit minimalen Richtungsänderungen von  $30^\circ$  pro Modelldatenvertex angewandt. Für eine analytische Betrachtung wurden dabei für jede Aufnahme Soll-Standardabweichungen im Bereich zwischen 1.0mm und 1.0m im Weltkoordinatensystem angewandt. Die Soll-Standardabweichung  $\sigma_{NP}$  bezieht sich auf die Residuen zwischen Modelldatenvertices und nächsten Punkten auf den transformierten Sensordatenobjekten (vgl. Abbildung 5.13). Im Gegensatz zu Messungen in realen Bildern läßt sich hier durch die Verwendung synthetischer Aufnahmen die Genauigkeit der erzeugten Paßpunkte infolge strukturell identischer Objekte exakt berechnen. Dazu werden die den Modelldatenvertices entsprechenden Sensordatenvertices mitgeführt. Die Standardabweichung  $\sigma_{GCP}$  aus den Residuen der berechneten Paßpunkte und der entsprechenden Sensorvertices kann dann im Bildraum bestimmt werden und ist direkt vergleichbar mit der manuellen Bildmeßgenauigkeit. Die Verhältnisse zur automatischen Erzeugung von Paßpunkten und zur analytischen Bewertung sind in Abbildung 5.13 dargestellt. In Abbildung 5.14 ist die Standardabweichung der Paßpunktresiduen im

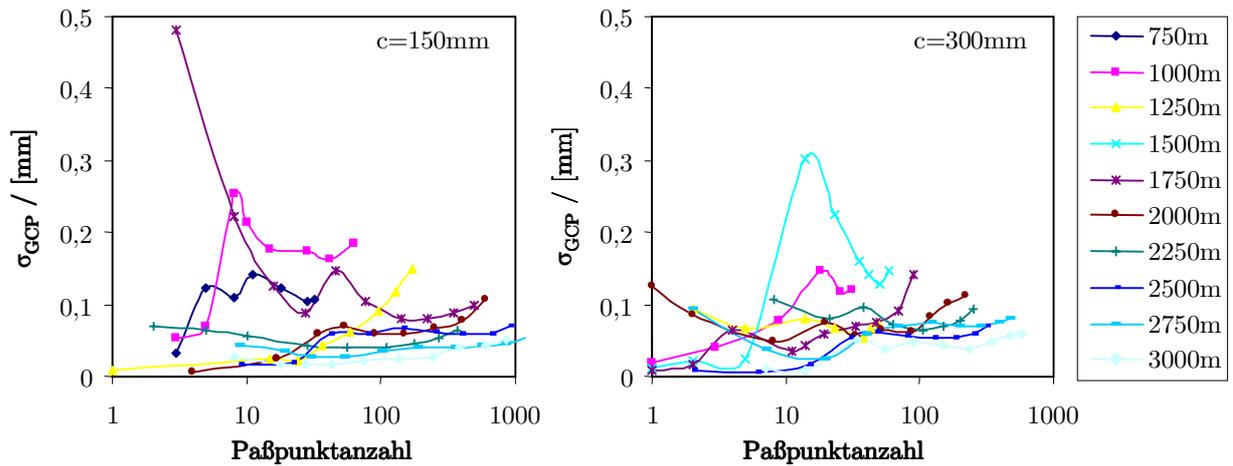


**Abbildung 5.13:** Veranschaulichung der automatischen Erzeugung und Bewertung von Paßpunkten.

Bildraum für alle Flughöhen in Abhängigkeit von der Soll-Standardabweichung  $\sigma_{NP}$  aufgetragen. Die Vorgabe der Standardabweichung der Nächste-Punkt-Abstände akzeptierter homologer Punktpaare beeinflusst auch die Anzahl an Paßpunkten, die für die jeweilige Aufnahme erzeugt wird. Je größer die vorgegebene Standardabweichung  $\sigma_{NP}$  ist, umso mehr Paßpunkte werden tendenziell auch erzeugt. Von Interesse ist dabei auch die Frage, welche Genauigkeiten sich im Bildraum in Abhängigkeit von der Anzahl erzeugter Paßpunkte erreichen lassen. Die Graphen dieser Abhängigkeit sind in Abbildung 5.15 dargestellt. Die Abhängigkeit der Anzahl an Paßpunkten von  $\sigma_{NP}$  ist implizit im Definitionsbereich von Abbildung 5.15 enthalten. Zur Beurteilung der Leistungsfähigkeit des entwickelten Verfahrens zur vollautomatischen Georeferenzierung und Paßpunkterzeugung sollten neben den Bildgenauigkeiten auch die Paßpunktgenauigkeiten im Weltkoordinatensystem bekannt



**Abbildung 5.14:** Erzielte Paßpunktgenauigkeiten im Bildraum für alle untersuchten Flughöhen und Kammerkonstanten bei Variation der Standardabweichung der Nächste-Punkt-Residuen akzeptierter homologer Punkt-paare zwischen 1.0mm und 0.512m in Weltkoordinaten.



**Abbildung 5.15:** Abhängigkeit der erreichten Paßpunktgenauigkeiten im Bildraum für alle untersuchten Flughöhen, Kammerkonstanten und Variationen von  $\sigma_{NP}$  für das Untersuchungsgebiet Vaihingen/Enz.

sein. Für die durchgeführten Untersuchungen wurde im Untersuchungsgebiet eine mittlere Geländehöhe von  $\bar{h} = 200\text{m}$  angenommen. Im verwendeten exakten photogrammetrischen Normalfall läßt sich die Paßpunktgenauigkeit  $\sigma_{GCP,world}$  im Gelände aus der Paßpunktgenauigkeit  $\sigma_{GCP}$  folgendermaßen berechnen:

$$\sigma_{GCP,world}[\text{m}] = \sigma_{GCP}[\text{mm}] * \frac{h - \bar{h} [\text{m}]}{c [\text{mm}]}, \quad (5.8)$$

wobei  $h$  die  $z$ -Koordinate des Projektionszentrums und  $c$  die Kammerkonstante der Kamera ist. Die erreichten lateralen Geländegenauigkeiten sind in Abbildung 5.16 wiedergegeben. Abschließend zu den Ergebnissen dieser Untersuchung sind in Abbildung 5.17 zwei erzeugte Paßpunktmenge im Bildkoordinatensystem dargestellt. Verifizierte Polygone, auf deren Rand Paßpunkte mit diesem Verfahren ausschließlich erzeugt werden können, sind durch die graue Einfärbung hervorgehoben. In gleicher Weise wurden Paßpunktmenge für alle betrachteten Konfigurationen erzeugt. Zur räumlichen Verteilung der Paßpunkte sei auf die anschließende Diskussion verwiesen.

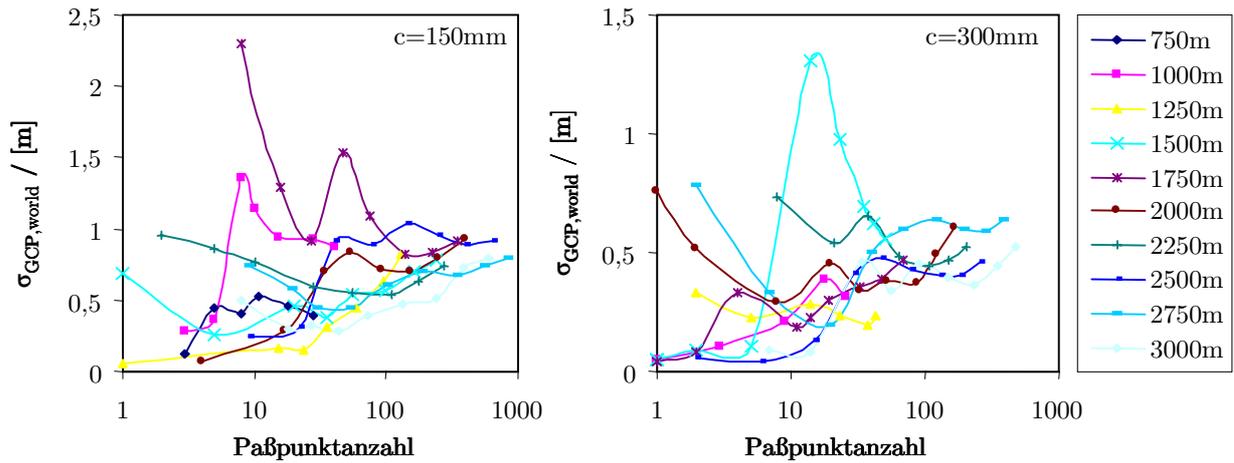


Abbildung 5.16: Abhängigkeit der erreichten Paßpunktgenauigkeiten im Gelände von der Anzahl erzeugter Punkte für alle untersuchten Flughöhen und Kammerkonstanten für das Untersuchungsgebiet Vaihingen/Enz.

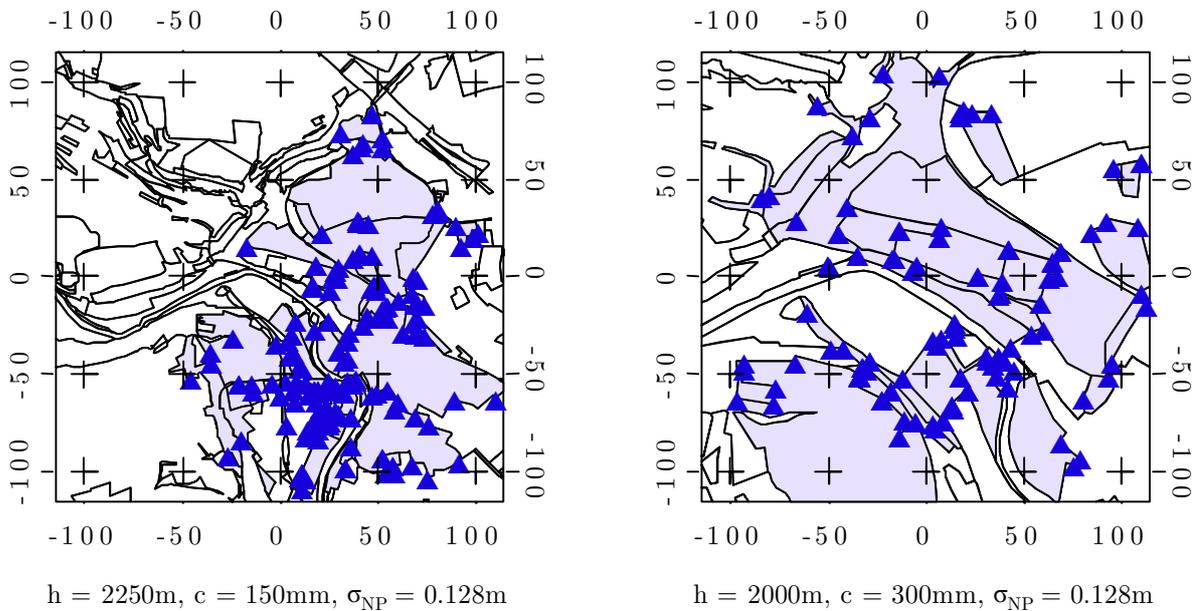


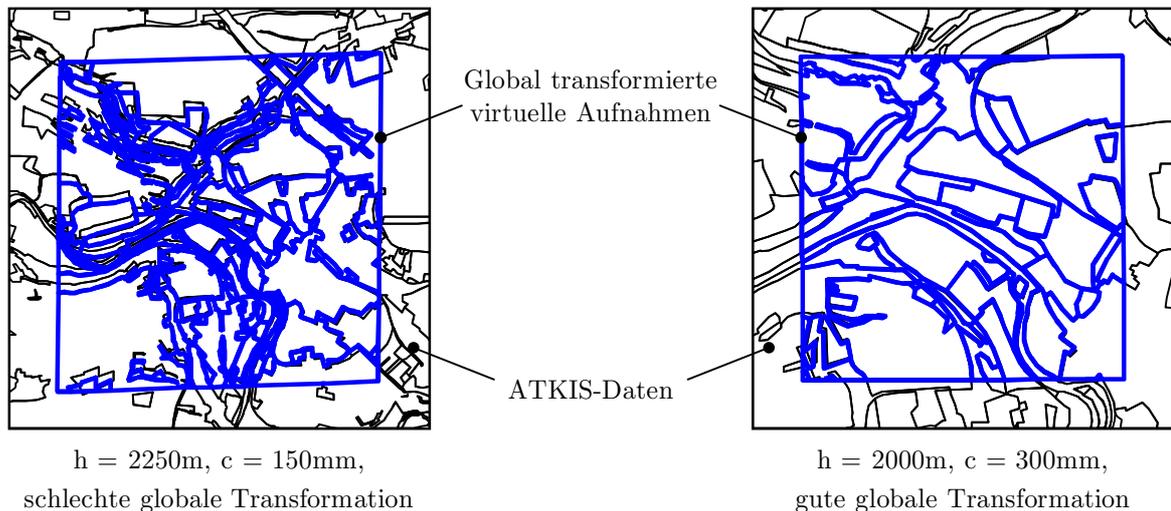
Abbildung 5.17: Zwei Beispiele für vollautomatisch erzeugte Paßpunkt mengen in virtuellen Luftbildaufnahmen des Untersuchungsgebiets Vaihingen/Enz.

### 5.3.3 Diskussion

Die durchgeführten Untersuchungen an synthetisch erzeugten Luftbildaufnahmen von 3D-ATKIS-Polygonen homogener Objektart zeigen die Leistungsfähigkeit der mit dieser Arbeit entwickelten Strategie zur vollautomatischen Georeferenzierung und Paßpunktextraktion bei real existierenden topographischen Verhältnissen im Luftbild-Normalfall. Im extremen Fall einer Flughöhe von 750m und einer Kammerkonstante von 300mm (siehe Abbildung 5.10) versagte das Verfahren, da insgesamt nur eine Objektpaarhypothese mit unzureichender Näherungstransformation erzeugt wurde. Der Grund dieses Versagens läßt sich im Zusammenspiel aus dem relativ großen Bildmaßstab (1:1833) und dem Generalisierungsgrad des den ATKIS-Daten zu Grunde liegenden DLM25 finden. Wie in Abbildung 5.10 zu erkennen ist, ist in dem betreffenden Bildausschnitt nur ein Polygon vollständig enthalten, alle anderen sind angeschnitten. Das einzige vollständig enthaltene Polygon hat jedoch die Zwangsbedingung für die Affinvarianten nicht befriedigt und wurde daher nicht zur Findung einer Näherungstransformation herangezogen. Vielmehr lieferte ein angeschnittenes Polygon hinreichende Konsistenz mit

den Zwangsbedingungen, führte jedoch zu einer unzureichenden Näherungstransformation. In allen anderen Fällen konnte das Verfahren mindestens fünf konsistente Polygone aus der Baumsuche extrahieren und eine hinreichend gute affine Näherungstransformation vom Bild- ins Weltkoordinatensystem ermitteln. Im Fall von nur fünf konsistenten Paaren ( $h = 1000\text{m}$ ,  $c = 300\text{mm}$ ) kam die Eigenschaft der implementierten Baumsuche zum Tragen, bei Nicht-Erreichen der maximalen Suchtiefe die bis dahin beste konsistente Interpretation zu verwenden.

Weiter ist zu beobachten, daß - wie infolge des sinkenden Bildmaßstabs erwartet - mit zunehmender Flughöhe mehr Hypothesen erzeugt und auch verifiziert wurden (Abbildung 5.12). Die Zahl der Hypothesen wächst stetig mit der Flughöhe, bei der Zahl der verifizierten Hypothesen treten zwei Abweichungen vom stetigen Wachstum auf: Bei Flughöhen von 2250m bzw 3000m, jeweils bei  $c = 150\text{mm}$ , werden weniger Hypothesen verifiziert als erwartet. Dieses Abweichen läßt sich damit erklären, daß hier die affine Näherungstransformation schlechter als in den restlichen Fällen ist. Damit werden mehr Hypothesen nicht verifiziert. Der Grund für die schlechtere Näherungstransformation liegt in der angewandten Baumsuche, die bei einer Tiefe von 6 konsistenten Paaren abgebrochen wurde. Dadurch kann die ermittelte globale Affintransformation unter Umständen nicht optimal sein, da die die Transformation bestimmenden Objekte zwar alle Zwangsbedingungen erfüllen, aber z.B. sehr klein sind oder nahe zusammen liegen. Damit ergeben sich an den Rändern des Bildausschnittes größere Fehler, die sich in einer geringeren Zahl verifizierter Polygone niederschlagen. Würde der Suchbaum ganz durchsucht, so wäre die ermittelte Transformation auch optimal, der Rechenaufwand jedoch deutlich höher. Zur Veranschaulichung der unterschiedlichen Verhältnisse sind in Abbildung 5.18 in zwei Fällen die Bildausschnitte in das ATKIS-System transformiert dargestellt.



**Abbildung 5.18:** Beispiele für eine schlechte und eine gute globale affine Sensor-Welt-Transformation aus der Baumsuche mit Zwangsbedingungen.

Für die Aufnahme mit  $h = 2000\text{m}$  und  $c = 300\text{mm}$  ist die Näherungstransformation gut, es werden erwartet viele Hypothesen verifiziert (siehe Abbildung 5.12). Die erzeugten Paßpunkte sind gut über den Bildausschnitt verteilt (siehe Abbildung 5.17). Für die Aufnahme mit  $h = 2250\text{m}$  und  $c = 150\text{mm}$  ist die Näherungstransformation wesentlich schlechter, es werden entsprechend weniger Hypothesen verifiziert, zudem sind die erzeugten Paßpunkte schlechter im Bild verteilt, was an den höheren Residuen an bestimmten verifizierten Polygonen liegt. Es läßt sich feststellen, daß die erzeugten Paßpunkte im wesentlichen in den Bereichen liegen, in denen bereits die globale Näherungstransformation eine hinreichend gute Korrespondenz zwischen ATKIS- und transformierten Sensorobjekten liefern. Wie aus Abbildung 5.17 ersichtlich ist, stimmt die globale Näherungstransformation nur im rechten unteren Bildbereich relativ gut. Entsprechend liefert die automatische Paßpunkterzeugung trotz individueller Verfeinerung der Transformationen Paßpunkte überwiegend aus diesem Bereich, obwohl die verifizierten Polygone ebenfalls gut über den Bildbereich verteilt sind.

Die Variationen der Soll-Standardabweichung  $\sigma_{NP}$  im Weltkoordinatensystem zwischen 1.0mm und 0.512m überträgt sich nicht über den ganzen Untersuchungsbereich linear auf die Standardabweichung  $\sigma_{GCP}$  im Bild. Vielmehr ist bei  $\sigma_{NP}$  kleiner als etwa 5cm kein Zusammenhang mit  $\sigma_{GCP}$  zu erkennen (Abbildung 5.14). Oberhalb dieses  $\sigma_{NP}$  läßt sich dagegen - wie erwartet - bis auf wenige Ausnahmen eine tendenzielle Zunahme von  $\sigma_{GCP}$  bei zunehmendem  $\sigma_{NP}$  beobachten. Die größeren Schwankungen bei kleinen  $\sigma_{NP}$  sind durch die

$\sigma_{NP}$ [mm] \ h[m]	750	1000	1250	1500	1750	2000	2250	2500	2750	3000
1	-	-	-	-	3	-	-	-	9	8
2	-	-	-	1	8	4	2	9	19	18
4	-	3	1	5	16	17	5	22	30	27
8	3	5	15	20	27	34	10	41	53	43
16	5	8	24	36	46	53	28	78	101	73
32	8	10	36	55	76	91	56	146	191	136
64	11	15	60	99	139	152	108	250	353	240
128	18	28	96	162	222	250	174	399	567	382
256	28	41	130	237	349	401	270	650	846	606
512	32	63	171	334	492	592	368	898	1179	861

**Tabelle 5.3:** Anzahl automatisch erzeugter Paßpunkte für  $c = 150\text{mm}$ , alle Flughöhen und alle angewandten Nächster-Punkt-Standardabweichungen.

$\sigma_{NP}$ [mm] \ h[m]	750	1000	1250	1500	1750	2000	2250	2500	2750	3000
1	-	-	-	1	1	1	-	2	2	7
2	-	-	-	1	2	2	-	6	7	14
4	-	-	-	2	4	8	8	15	20	34
8	-	-	2	5	11	20	21	30	40	56
16	-	1	5	14	14	33	38	48	70	91
32	-	3	14	23	19	52	64	80	123	151
64	-	9	23	35	33	86	106	129	199	234
128	-	18	38	42	47	121	151	185	287	347
256	-	26	43	51	70	164	203	254	394	478
512	-	31	49	60	90	221	255	315	482	574

**Tabelle 5.4:** Anzahl automatisch erzeugter Paßpunkte für  $c = 300\text{mm}$ , alle Flughöhen und alle angewandten Nächster-Punkt-Standardabweichungen.

damit einhergehende geringe Zahl an Paßpunkten zu erklären. In einigen Fällen (Flughöhen von 750m und 1000m) sind die untersuchten Werte  $\sigma_{NP}$  sogar so klein, daß überhaupt keine Paßpunkte erzeugt wurden. In Tabelle 5.3 ist die Anzahl der automatisch erzeugten Paßpunkte für  $c = 150\text{mm}$  und in Tabelle 5.4 für  $c = 300\text{mm}$  angegeben. Der besseren Übersichtlichkeit halber sind alle Felder mit 30 oder mehr erzeugten Paßpunkten grau hinterlegt.

Über die in den Tabelle 5.3 und 5.4 angegebenen Werte gelangt man aus der Abhängigkeit von  $\sigma_{GCP}$  von  $\sigma_{NP}$  zu der in Abbildung 5.15 dargestellten Abhängigkeit von  $\sigma_{GCP}$  von der Anzahl erzeugter Paßpunkte. Infolge der näherungsweise proportionalen Zunahme der Anzahl erzeugter Paßpunkte mit dem angewandten  $\sigma_{NP}$  zeigt sich ein ähnliches Verhalten wie im  $\sigma_{GCP}(\sigma_{NP})$ -Diagramm. Für eine kleine Anzahl von Paßpunkten zeigt sich kein genereller Zusammenhang zwischen der Standardabweichung  $\sigma_{GCP}$  der Paßpunktresiduen im Bildraum und ihrer Anzahl. Für mehr als etwa 30 Paßpunkte ergibt sich dagegen - wie auch hier erwartet - eine leichte Zunahme von  $\sigma_{GCP}$  mit der Anzahl. Für den Bereich zwischen 30 und 50 Paßpunkten bewegt sich der Punktfehler im Bildraum bei beiden untersuchten Kammerkonstanten überwiegend im Bereich um 0.1mm.

Abschließend zeigt Abbildung 5.16 die für praktische Zwecke wichtige Abhängigkeit der Standardabweichung der Paßpunktresiduen im Geländekoordinatensystem. Für das gesamte Untersuchungsgebiet wurde eine mittlere

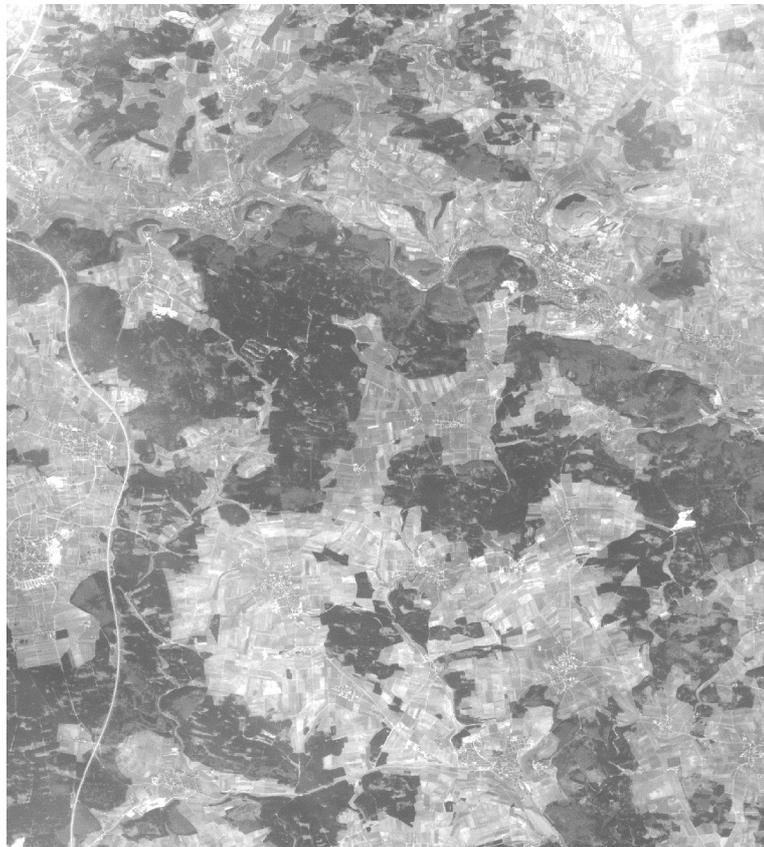
Geländehöhe von  $\bar{h} = 200\text{m}$  angenommen. Die Standardabweichungen  $\sigma_{GCP,world}$  der Paßpunktresiduen im Geländesystem (Abbildung 5.16) wurden durch die Beziehung 5.8 aus den entsprechenden Standardabweichungen  $\sigma_{GCP}$  im Bildraum berechnet. Es zeigt sich tendenziell dasselbe Verhalten wie für  $\sigma_{GCP}$ . Für die Kammerkonstante  $c = 150\text{mm}$  liegen die erreichbaren Geländegenauigkeiten bei 30 erzeugten Paßpunkten im Bereich um 50cm, für  $c = 300\text{mm}$  ergibt sich ein etwas besserer Wert von etwa 30cm. Es sei aber noch einmal betont, daß die in diesem Abschnitt angestellten Untersuchungen ideale Bildsegmentierungen und damit eine theoretische Obergrenze für die mit dem entwickelten Verfahren erreichbaren Genauigkeiten darstellen. Bei segmentierten Bilddaten kann der Fehler, abhängig von der Qualität der Segmentierung, größer werden. Um diesem Effekt jedoch weitgehend vorzubeugen wird in der Gesamtstrategie die bereits beschriebene Mehrfachsegmentierung mit unterschiedlichen Schwellwerten eingesetzt. Zudem ist zu bemerken, daß sich alle Geländegenauigkeiten aus Punkten auf ATKIS-Polygonen ermittelt wurden, die ihrerseits Lageungenauigkeiten im Bereich von 2-3m aufweisen.

## 5.4 Georeferenzierung von SPOT-PAN- mit ATKIS-Daten

Mit dieser Untersuchung wird das Verhalten des entwickelten Gesamtsystems Im Falle realer Fernerkundungsdaten untersucht. Hier wurden SPOT2-Daten des Testgebiets Elchingen/Großkuchen in Ost-Baden-Württemberg mit ATKIS-Daten registriert. Die Untersuchung stellt damit ein Georeferenzierungsproblem dar und ist dem Bereich Raster-Vektor-Registrierung zuzuordnen.

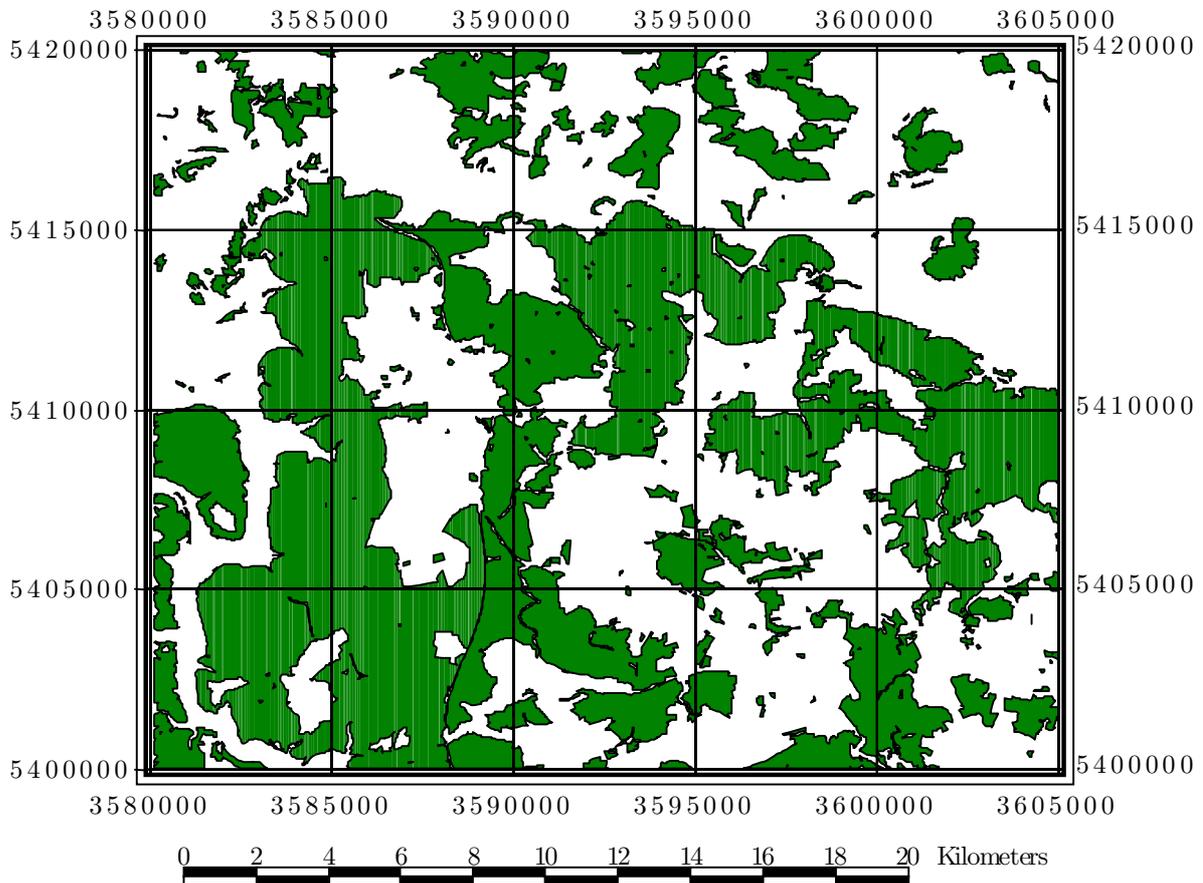
### 5.4.1 Der Testdatensatz

Für diese Untersuchung wurde ein Ausschnitt aus einer panchromatischen SPOT2-Szene (level 1a) vom 01.04.1993 herangezogen. Der Ausschnitt besitzt eine Größe von 1796pixel x 1962pixel und erfasst damit bei etwa 10m Bodenpixelgröße eine Fläche von etwa 18km x 19.5km. In Abbildung 5.19 ist der verwendete SPOT2-Ausschnitt dargestellt. Zur besseren Erkennbarkeit wurde eine Histogrammoptimierung vorgenommen. Wie bei der vorigen



**Abbildung 5.19:** Der untersuchte Ausschnitt aus einer SPOT2-PAN-Szene des Untersuchungsgebiets Elchingen/Großkuchen vom 01.04.1993.

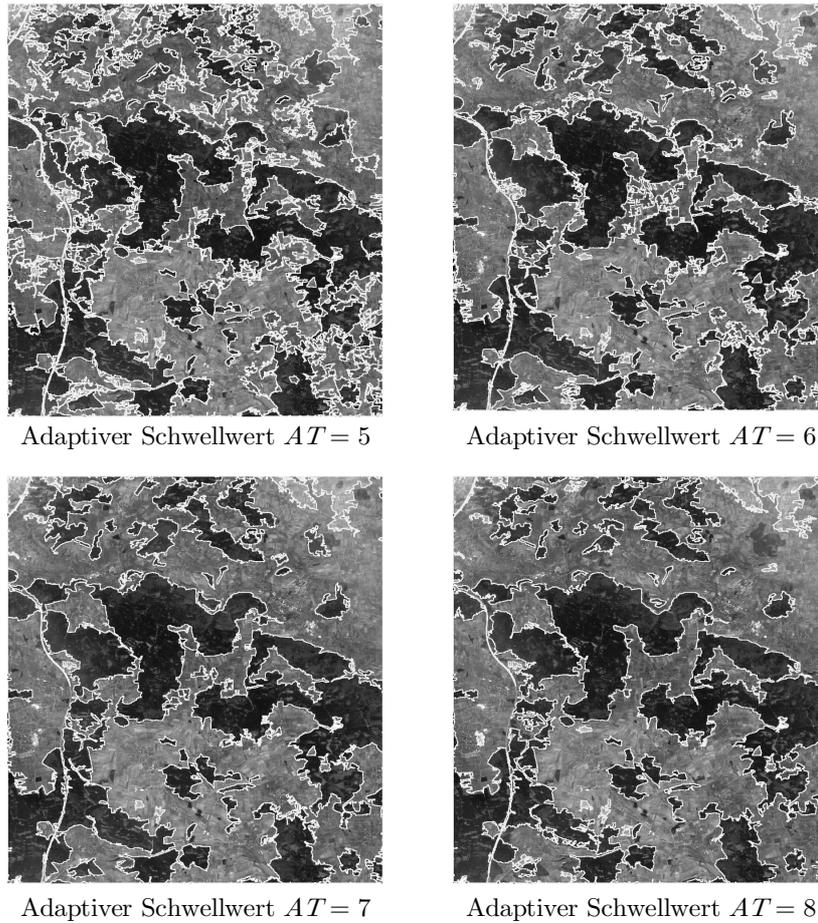
Untersuchung wurden die GIS-Daten aus dem ATKIS-Datensatz für Baden-Württemberg extrahiert. Dazu wurden zunächst in einem Bereich, der grob mit dem in Abbildung 5.19 abgebildeten Ausschnitt übereinstimmt, alle polygonalen ATKIS-Objekte der Objektart Wald (code 4107) selektiert. In einem anschließenden Schritt wurden benachbarte Waldpolygone zu zusammengefassten Objekten verschmolzen. Das Ergebnis der Selektion und Verschmelzung wurde als GIS-Eingabedatensatz verwendet und ist in Abbildung 5.20 abgebildet. Die Ausdehnung des GIS-Datensatzes von 25km x 20km ist als praxisrelevant zu betrachten.



**Abbildung 5.20:** ATKIS-Waldpolygone (Objektart 4107) des Untersuchungsgebiets Elchingen/Großkuchen. Benachbarte Objekte sind miteinander verschmolzen.

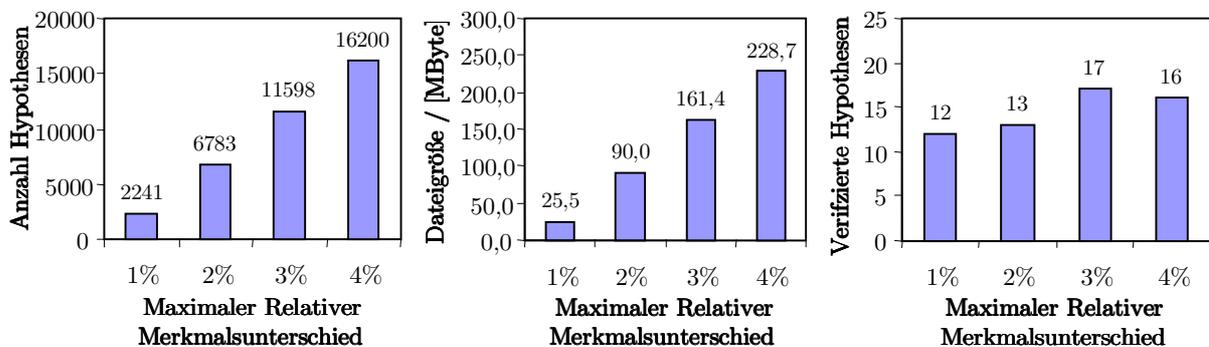
### 5.4.2 Durchführung und Ergebnisse der Untersuchung

Wie in der Systemübersicht in Abbildung 5.4 angegeben, wurde zunächst der SPOT-Ausschnitt mit unterschiedlichen adaptiven Schwellwerten mit dem in Abschnitt 2.1.2 beschriebenen Verfahren segmentiert. Es wurden adaptive Schwellwerte  $AT \in [5, 8]$  verwendet, d.h. es wurden insgesamt vier Segmentierungsergebnisse erhalten. Zur Reduktion der Zahl betrachteter Regionen wurden anschließend für jedes Segmentierungsergebnis alle Regionen, die kleiner als 1000pixel (etwa 0.03% der Bildgröße) sind, eliminiert. Für die verbleibenden Regionen wurden Vektor-Umrißpolygone, die zunächst exakt den Pixelkanten folgen, extrahiert. Anschließend wurden für alle Polygone aus allen Segmentierungen die zentralen geometrischen Momente bis achter Ordnung, die sechs einfachsten Kreuzmomentvarianten und die Transformation in die in Abschnitt 4.5.3 definierte Standardlage berechnet. Für die verschmolzenen ATKIS-Waldpolygone wurden ebenfalls die zentralen geometrischen Momente bis zur achten Ordnung, die sechs einfachsten Kreuzmomentvarianten und die Transformation in die Standardlage berechnet. Die Ergebnisse der Segmentierungsläufe sind in Abbildung 5.21 dargestellt. Regionen, die kleiner als 1000pixel sind, wurden bereits eliminiert, gezeigt sind die Umrißpolygone für die jeweils verbleibenden Regionen, also die Objekte, die zur merkmalsbasierten Zuordnung und zur Baumsuche in Betracht kamen. Für die Segmentierungen ergaben sich nach Eliminierung der kleinen Regionen insgesamt 421 Objekte (167 aus  $AT = 5$ , 104 aus  $AT = 6$ , 74 aus  $AT = 7$  und 76 aus  $AT = 8$ ). Da der implementierten Strategie zufolge alle Objekte aus allen Segmentierungen potenzielle Kandidaten für die merkmalsbasierte Zuordnung sind, ergibt



**Abbildung 5.21:** Ergebnisse der durchgeführten Segmentierungen der SPOT-Szene Elchingen/Großkuchen. Die verwendeten adaptiven Schwellwerte liegen im Bereich zwischen 5 und 8.

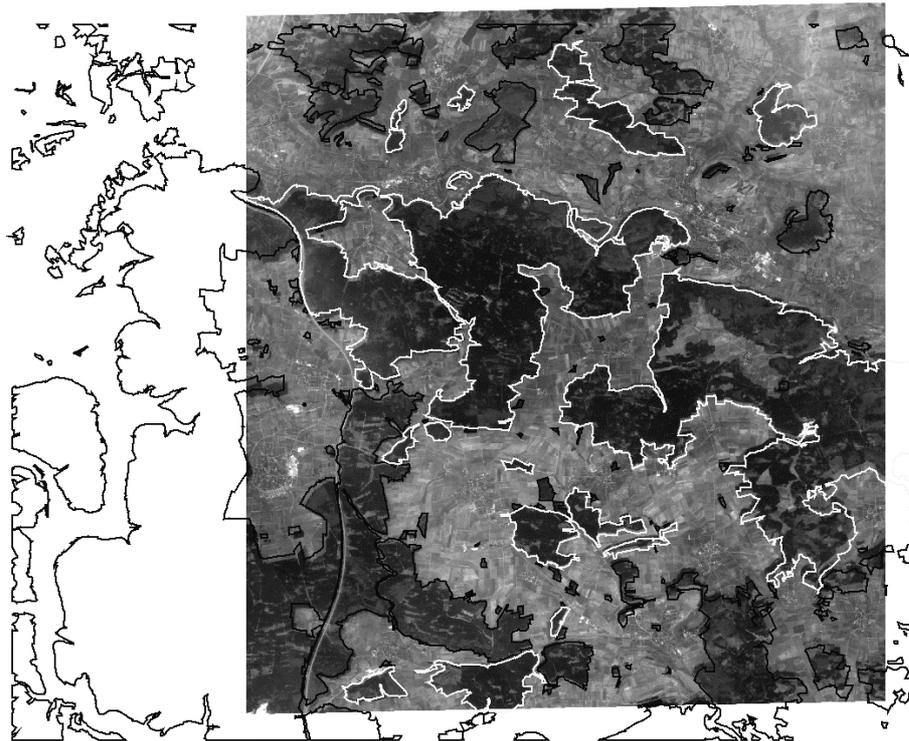
sich zusammen mit den 216 Objekten des GIS-Datensatzes die nicht unerhebliche Anzahl von  $216 \times 421 = 90936$  potenziellen Paaren. Die tatsächliche Anzahl an Hypothesen wird durch die Wahl des maximalen relativen Unterschieds je Merkmal bei der Durchführung der merkmalsbasierten Zuordnung gesteuert. Zur Kontrolle über den weiteren Verlauf der Untersuchung wurde daher das Zuordnungsverfahren aus Abschnitt 4.4 mit relativen Unterschieden (gleich für alle sechs Invarianten) im Bereich zwischen 1% und 4% durchgeführt. Die Ergebnisse der unterschiedlichen Durchläufe sind in Statistik 5.22 dargestellt.



**Abbildung 5.22:** Statistische Daten zu den untersuchten Durchläufen der merkmalsbasierten Zuordnung für alle Objekte des Testgebiets Elchingen/Großkuchen. Der relative Unterschied wurde für alle sechs Invarianten gleichermaßen im Bereich zwischen 1% und 4% gewählt.

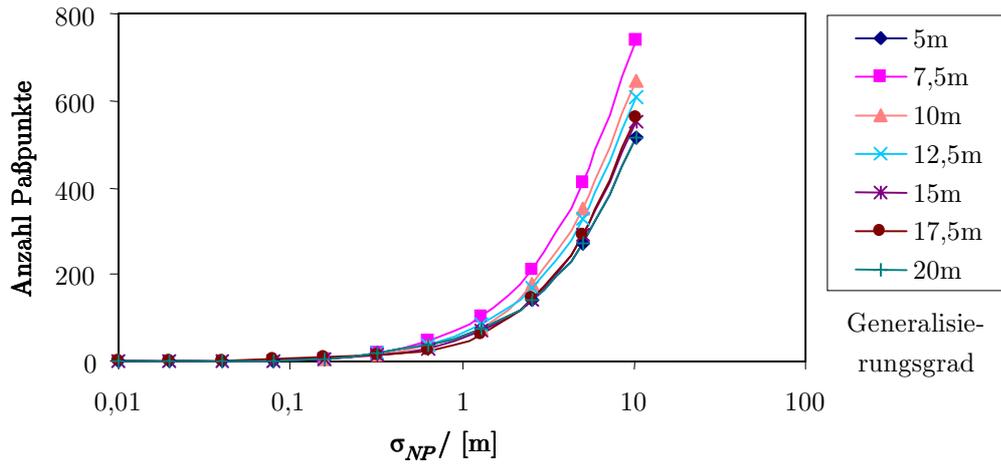
Die noch ausstehenden Prozessierungsschritte konnten für alle vier Ergebnisse der merkmalsbasierten Zuordnung erfolgreich durchgeführt werden. Speziell für den Satz mit 4% maximalem relativem Unterschied pro Invariante war jedoch die zu handhabende Dateigröße mit 228.7 MByte zu groß, um über die ganze Prozesskette im Arbeitsspeicher der verwendeten SGI O2-Workstation gehalten werden zu können, wodurch sich die Rechenzeit extrem verlängerte (insgesamt etwa 1 Stunde für  $RD = 4\%$ ). Da sich die Ergebnisse aus allen Läufen jedoch in allen Punkten nicht wesentlich unterscheiden, wird im folgenden nur mit  $RD = 3\%$  weitergearbeitet, da hier die größte Zahl an verifizierten Hypothesen (17) erzielt werden konnte. Vor der Baumsuche wurden für alle Hypothesen jeweils die MBRs und die affinen Hin- und Rücktransformationen aus den Standardlagen berechnet. Das Baumsuchverfahren wurde wie bei der Untersuchung in Abschnitt 5.3 mit einer maximalen Suchtiefe von 6 und einer minimalen Objektüberlappung von 80% durchgeführt.

Die Anwendung der so ermittelten globalen Affintransformation auf die SPOT-Szene ist in Abbildung 5.23 dargestellt.

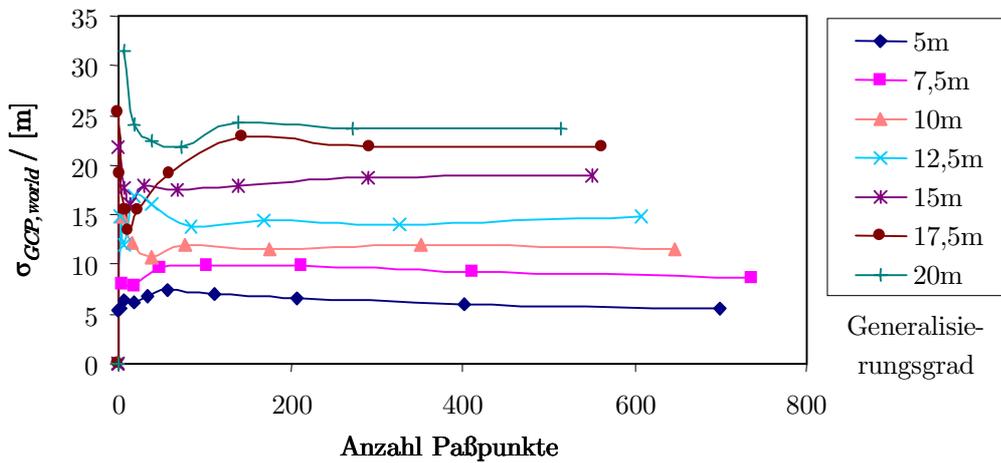


**Abbildung 5.23:** Transformation des SPOT-Testdatensatzes Elchingen/Großkuchen auf die ATKIS-Polygone mit Hilfe der globalen Affintransformation aus dem Baumsuchverfahren. Die verifizierten Polygone sind hell eingezeichnet.

Für den Verifikationsschritt wurden in dieser Untersuchung folgende Werte verwendet: Maximale Abweichung in Rotation und Scherung:  $\pm 45^\circ$ , maximaler Unterschied im  $x$ - und  $y$ -Maßstab: 20%, maximaler Unterschied in der  $x$ - und  $y$ -Verschiebung: 20% und minimale Überlappung: 40%. Die insgesamt 17 so verifizierten Polygone (GIS-Partner) sind in Abbildung 5.23 hell eingezeichnet. Vor der Durchführung der ICP-Verfeinerung wurden die Raster-Polygone zur besseren Konvergenz und zum besseren Zeitverhalten des ICP-Algorithmus' mit Hilfe des Douglas-Peucker-Algorithmus' generalisiert [Douglas & Peucker 1973]. Um zu untersuchen, wie sich die Generalisierung auf die Paßpunktgenauigkeit auswirkt, wurden hierbei Generalisierungen im Bereich von 5m bis 20m in 2.5m-Schritten an den ins Weltkoordinatensystem transformierten Raster-Polygonen angebracht. Abschließend wurden die Raster-Polygone aller verifizierten Hypothesen in das Weltkoordinatensystem transformiert. Mit Hilfe des modifizierten ICP-Algorithmus' wurde die individuell optimale 2D-projektive Transformation ermittelt. Zur Extraktion von Paßpunkten aus den GIS-Polygonen und den generalisierten Raster-Polygonen wurde das bereits beschriebene Verfahren mit Nächster-Nachbar-Standardabweichungen von  $\sigma_{NP} = 1\text{cm}$  bis  $\sigma_{NP} = 10.24\text{m}$ , durchgeführt. Die angebrachten  $\sigma_{NP}$  wurden dabei jeweils verdoppelt. Die Anzahl der erhaltenen Paßpunkte in Abhängigkeit vom angewandten  $\sigma_{NP}$  ist in Abbildung 5.24 für alle untersuchten Generalisierungen dargestellt. Die daraus resultierenden Paßpunktgenauigkeiten in Abhängigkeit von der Anzahl erzeugter Paßpunkte sind in Abbildung 5.25 zu sehen.



**Abbildung 5.24:** Anzahl erzeugter Paßpunkte in Abhängigkeit von der Nächsten-Nachbar-Standardabweichung in Weltkoordinaten für alle untersuchten Generalisierungsgrade.



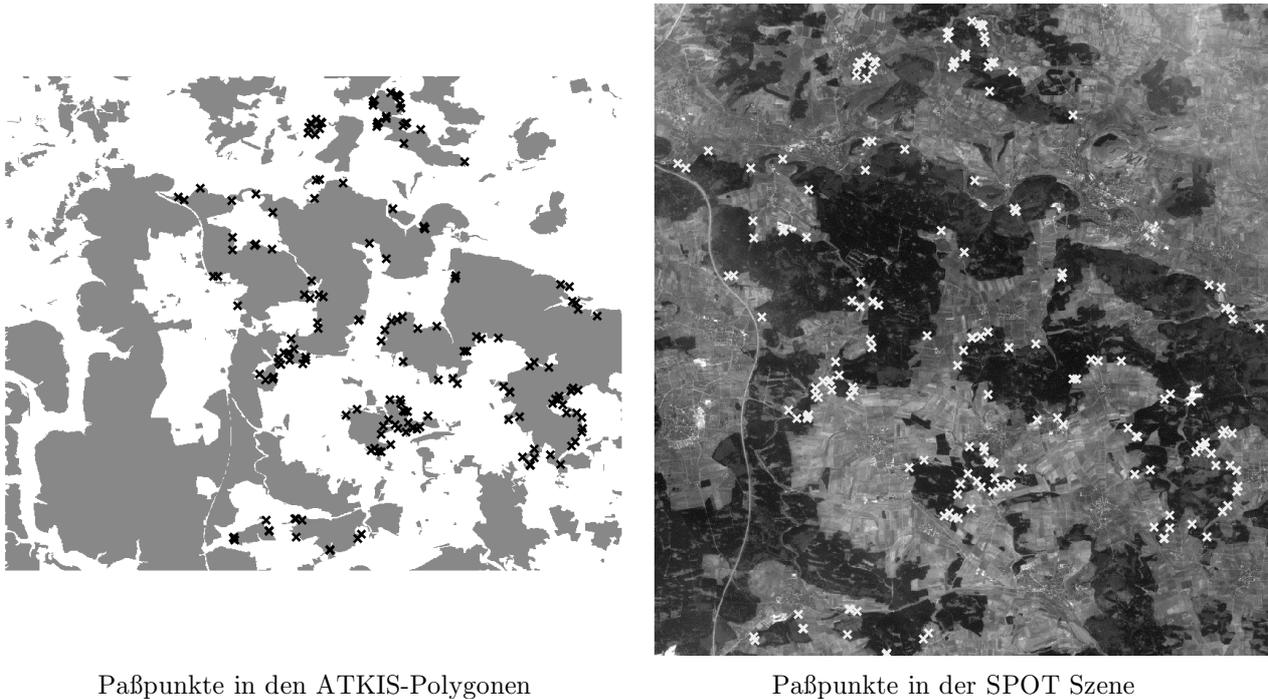
**Abbildung 5.25:** Paßpunktgenauigkeit in Abhängigkeit von der Anzahl erzeugter Paßpunkte für alle untersuchten Generalisierungsgrade.

In den Abbildungen 5.24 und 5.25 markiert jeder Meßpunkt einen ganzen Satz von Paßpunkten. Aus den Abbildungen zeigt sich, daß eine sinnvolle Wahl der untersuchten Parameter z.B. bei  $\sigma_{NP} = 2.56\text{m}$  und einem Generalisierungsgrad von  $10.0\text{m}$  liegt. In diesem Fall ergeben sich über den ganzen Datensatz verteilt 177 Paßpunkte mit einer lateralen Geländegenauigkeit von etwa  $11\text{m}$ . Dies ist ein für praktische Georeferenzierungsaufgaben realistischerweise erreichbarer Wert. In Abbildung 5.26 sind die erhaltenen Paßpunkte für diese Parameterwahl exemplarisch aufgezeigt. Abschließend zu den Untersuchungen im Testgebiet Elchingen/Großkuchen wurde in zwei Beispielen die Robustheit des Gesamtverfahrens qualitativ betrachtet. Dazu wurde der ATKIS-Datensatz künstlich zwei unterschiedlichen affinen Transformationen unterworfen und das gesamte Verfahren wie bereits beschrieben durchlaufen. Es ist dabei zu betonen, daß keiner der prozessrelevanten Parameter verändert wurde, alle Parameter wurden bei den bereits angegebenen Werten belassen. Im ersten Fall wurde lediglich eine  $x$ -Scherung an den ATKIS-Polygonen angebracht. Die zu Grunde liegende Affintransformation hatte folgende Gestalt:

$$x' = x + 0.5y \tag{5.9}$$

$$y' = y \tag{5.10}$$

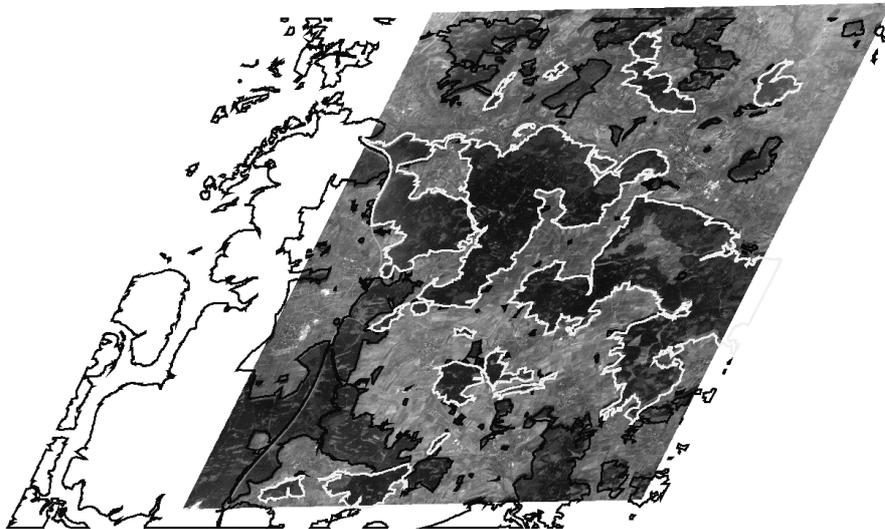
In Abbildung 5.27 sind die transformierten ATKIS-Polygone mit schwarzem Rand eingezeichnet. In einem zweiten Experiment wurden die ATKIS-Polygone einer Rotation, Scherung und einer Spiegelung unterworfen.



Paßpunkte in den ATKIS-Polygonen

Paßpunkte in der SPOT Szene

**Abbildung 5.26:** Beispiel für automatisch erzeugte Paßpunkte im Untersuchungsgebiet Elchingen/Großkuchen. Die Standardabweichung der nächsten Nachbarn wurde hier zu  $2.56m$  gesetzt, der Generalisierungsgrad lag bei  $10m$ .



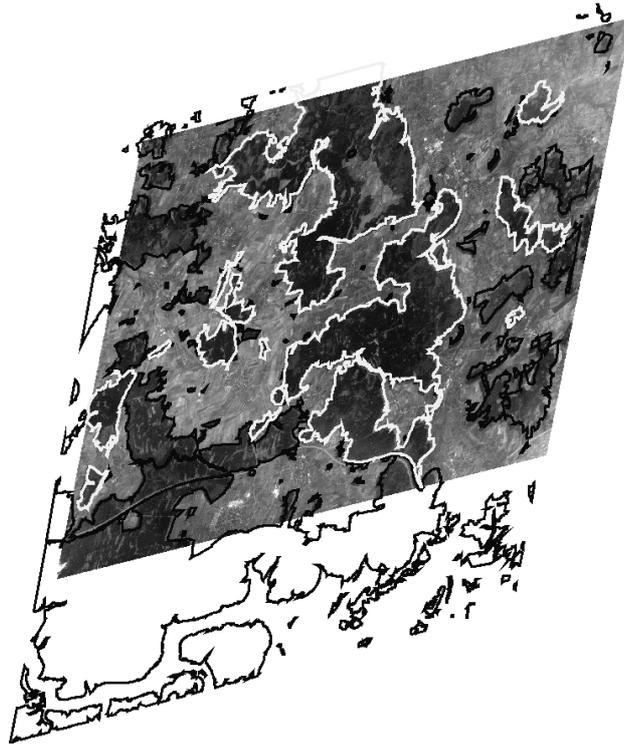
**Abbildung 5.27:** Ergebnis der globalen Affintransformation bei künstlich in  $x$ -Richtung gescherten ATKIS-Polygonen. Verifizierte Polygone sind hell gezeichnet.

Die angewandte Affintransformation hatte die Gestalt:

$$x' = 0.170x + 0.985y \quad (5.11)$$

$$y' = 0.964x + 0.267y . \quad (5.12)$$

Das mit der berechneten globalen Affintransformation transformierte SPOT-Bild und die ATKIS-Polygone sind in Abbildung 5.28 dargestellt. Auch hier sind verifizierte Polygone hell eingezeichnet. Beide Beispiele unterstreichen die Robustheit des Gesamtsystems selbst in extremen Fällen, die in der Praxis so nicht vorkommen.



**Abbildung 5.28:** Ergebnis für künstlich gescherte, rotierte und gespiegelte ATKIS-Polygone.

### 5.4.3 Diskussion

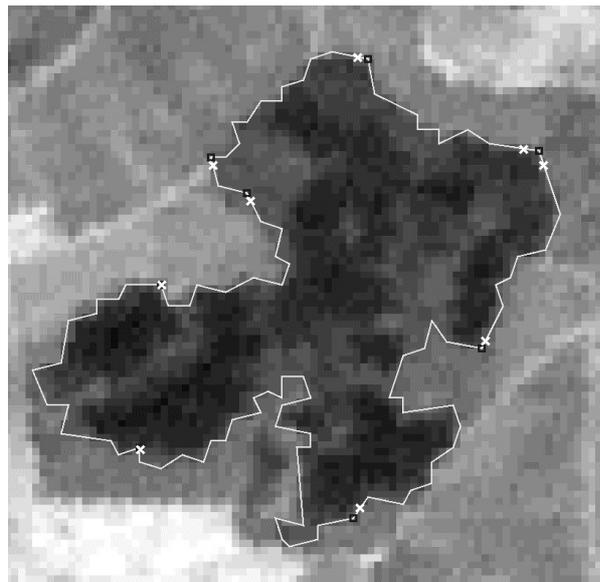
Mit dieser Untersuchung wurde das Verhalten des Gesamtsystems anhand einer Georeferenzierungsaufgabe für eine panchromatische SPOT2-Szene und ATKIS-Daten der Objektart Wald (code 4197) in einem Untersuchungsgebiet bei Elchingen/Großkuchen (Ost-Baden-Württemberg) betrachtet. Der Datensatz ist mit einer Ausdehnung von etwa  $18\text{km} \times 19.5\text{km}$  durchaus von praxisrelevanter Größe. Zur Abdeckung des Segmentierungsbereiches von "übersegmentiert" bis "untersegmentiert" wurden adaptive Schwellwerte  $AT = 5, 6, 7$  und  $8$  verwendet. Abbildung 5.21 zeigt für  $AT = 5$  eine deutliche Übersegmentierung, besonders im rechten unteren Bildbereich. Mit Vergrößerung des adaptiven Schwellwertes verringert sich die Zahl der Regionen im Bild, bei  $AT = 8$  ist für die durchzuführende Aufgabe bereits eine Untersegmentierung zu betrachten, die sich deutlich im rechten oberen Bildbereich bemerkbar macht. Hier sind einige Waldpolygone bereits mit der großen Hintergrundregion verschmolzen. Mit dieser Wahl von adaptiven Schwellwerten wird damit die dem Gesamtsystem zu Grunde liegende Annahme, innerhalb mehrerer Segmentierungsdurchläufe jedes gewünschte Objekt mindestens einmal "richtig" zu segmentieren, bestmöglich befriedigt.

Mit der Variation der maximalen relativen Unterschiede pro Kreuzmoment-Invariante zwischen 1% und 4% wurde gezeigt, wie sich die Zahl der erhaltenen Hypothesen und die Dateigröße verändert (siehe Abbildung 5.22). Es zeigt sich hierbei, daß auch für eine vernünftige Wahl dieses Parameters (4%) bereits eine praktikable Grenze in Form von zu handhabenden Dateien, die größer als 250MByte sind, erreicht werden kann. Für alle untersuchten Werte wurde die maximale Suchtiefe von 6 erreicht, die ermittelten globalen Transformationen waren in allen Fällen sehr ähnlich. Ein maximaler relativer Merkmalsunterschied von 3% stellte mit dem implementierten System den vernünftigsten Kompromiß aus Rechenzeit und Zuverlässigkeit in Form der Anzahl verifizierter Hypothesen dar. Es konnten damit 17 von insgesamt 90936 dem Baumsuchverfahren zugeführten Objektpaarhypothesen verifiziert werden. Dies zeigt deutlich die Leistungsfähigkeit der implementierten Baumsuche mit unären und binären Zwangsbedingungen. Wie in Abbildung 5.23 gut zu sehen ist, liefert die damit ermittelte globale Affintransformation bereits eine akzeptable Näherungstransformation der SPOT2-Szene in das ATKIS-Koordinatensystem, womit eine grobe Georeferenzierung bereits bewerkstelligt ist. Zudem sind die verifizierten Hypothesen (helle Polygone in Abbildung 5.23) gleichmäßig über die Szene verteilt und liefern damit die notwendige Bedingung für eine Gleichverteilung der zu extrahierenden Paßpunkte in der Szene. Hier zeigt sich der Nutzen der Sortierung der Hypothesen nach der Größe. Die Suche beginnt bei den größten Paaren und da sich sowohl die GIS-Objekte als auch die Objekte einer Segmentierung nicht überlappen, wird damit eine bestmögliche räumliche Verteilung erzwungen.

Die Generalisierung der verifizierten Raster-Polygone aus den einzelnen Segmentierungen mit Hilfe des Douglas-Peucker-Verfahrens wurde sinnvollerweise vor der individuellen ICP-Optimierung durchgeführt. Der Nutzen der Generalisierung besteht zum einen in der Verringerung der Zahl an Vertices und damit eine Verkürzung der Rechenzeit. Zum anderen liegen dadurch die Normalenrichtungen zu den Raster-Polygonen, längs derer ja während des ICP-Laufs die nächsten Punkte auf den GIS-Polygonen gesucht werden, realistischer als bei den ungeneralisierten Polygonen, wo es insgesamt nur vier verschiedene Normalenrichtungen gibt. Damit liegen die nächsten Punkte in der Regel jeweils näher an den Startpunkten, was eine besser Konvergenz des ICP-Verfahrens bewirkt.

Zur Untersuchung des Einflusses der im Weltkoordinatensystem durchgeführten Generalisierung auf die Anzahl und Genauigkeit extrahierbarer Paßpunkte wurden Generalisierungsgrade zwischen 5m und 20m in 2.5m-Schritten verwendet. Für alle Generalisierungsgrade wurden zur Paßpunkterzeugung Standardabweichungen für nächste Nachbarpunkte im Bereich zwischen  $\sigma_{NP} = 1\text{cm}$  bis  $\sigma_{NP} = 10.24\text{m}$  verwendet. Wie in Abbildung 5.24 zu sehen ist, verhält sich die Anzahl der erzeugten Paßpunkte prinzipiell ähnlich für alle Generalisierungen. Unterhalb von  $\sigma_{NP} = 10\text{cm}$  ist keine nennenswerte Zahl an Paßpunkten extrahierbar. Für  $\sigma_{NP} = 2.56\text{m}$  liegt die Zahl der Paßpunkte im Bereich zwischen 100 und 200, dieser Wert stellt somit einen guten Kompromiss zwischen kleinem  $\sigma_{NP}$  und hoher Paßpunktanzahl dar. Aus der Abhängigkeit der erreichbaren Genauigkeit in Weltkoordinaten von der Anzahl erzeugter Paßpunkte (Abbildung 5.25) erkennt man, daß sich für mehr als etwa 50 Paßpunkte für alle untersuchten Generalisierungen nahezu stabile Verhältnisse ergeben. Für eine Generalisierung von 10m liegt beispielsweise die Paßpunktgenauigkeit bei etwa 11m. Kleinere Generalisierungen als 10m liefern zwar noch bessere Genauigkeiten, jedoch sind hier die generalisierten Polygone noch sehr stark den ursprünglichen Raster-Polygonen ähnlich, von daher sind diese Genauigkeiten etwas mit Vorsicht zu genießen.

Die Genauigkeit der Paßpunkte wurde im Bildraum ermittelt. Dazu wurden die nächsten Nachbarn der in Frage kommenden GIS-Polygon-Vertices mit der jeweiligen individuellen Transformation in den Bildraum transformiert. Diese Punkte werden als Bildmessung betrachtet. Die Residuen, die sich durch  $\sigma_{GCP}$  ausdrücken, wurden bezüglich des jeweils nächsten Vertex des entsprechenden generalisierten Raster-Polygons gebildet. Damit ist eine automatische Berechnung der Genauigkeit möglich. Es besteht jedoch im Gegensatz zu den simulierten Aufnahmen die Gefahr, daß der jeweils betrachtete Raster-Polygon-Vertex nicht "der richtige" ist und damit das Residuum eventuell einen falschen Wert annimmt. Diese Gefahr ist wesentlich höher, solange die generalisierten Polygone größtenteils den ursprünglichen Raster-Polygonen entsprechen, also für Generalisierungsgrade kleiner als 10m ( $\approx 1\text{pixel}$  im Bild). Von daher scheint die Wahl eines Generalisierungsgrades von 10m der günstigste Fall, da hier bei hoher Paßpunktgenauigkeit die Wahrscheinlichkeit eines falschen Residuums wesentlich geringer als bei Generalisierungen von 5m oder 7.5m ist. In Abbildung 5.29 sind die eben beschriebenen Verhältnisse im Bildraum in einer Detailansicht eines verifizierten Objekts zu sehen.



**Abbildung 5.29:** Automatisch erzeugte Paßpunkte im Bildraum (weiße Kreuze) und zugehörige Vertices des generalisierten Raster-Polygons (dunkle offene Quadrate). Die Generalisierung betrug 10m, die Standardabweichung der Nächsten-Nachbar-Residuen im ATKIS-System wurde mit 2.56m vorgegeben. In diesem Fall ergaben sich 177 Paßpunkte in der gesamten Szene.

Hier sind die "Bildmessungen" durch weiße Kreuze dargestellt, die Vertices des generalisierten Polygons, bezüglich derer die Residuen gemessen werden, sind durch dunkle, offene Quadrate gekennzeichnet. In dieser Abbildung ist auch gut zu erkennen, daß sich das generalisierte Polygon weitgehend vom ursprünglichen, den Pixelkanten exakt folgenden Verlauf gelöst hat und somit die verbleibenden Vertices relativ zuverlässig für die Residuenmessung sind. Um letztlich zu den in Abbildung 5.25 angegebenen Paßpunktgenauigkeiten zu gelangen, wurde die Standardabweichung der Residuen im Bildraum (gemessen in pixel) lediglich mit dem Faktor 10 mutipliziert, da die Bodenpixelgröße der panchromatischen SPOT-Aufnahme 10m beträgt.

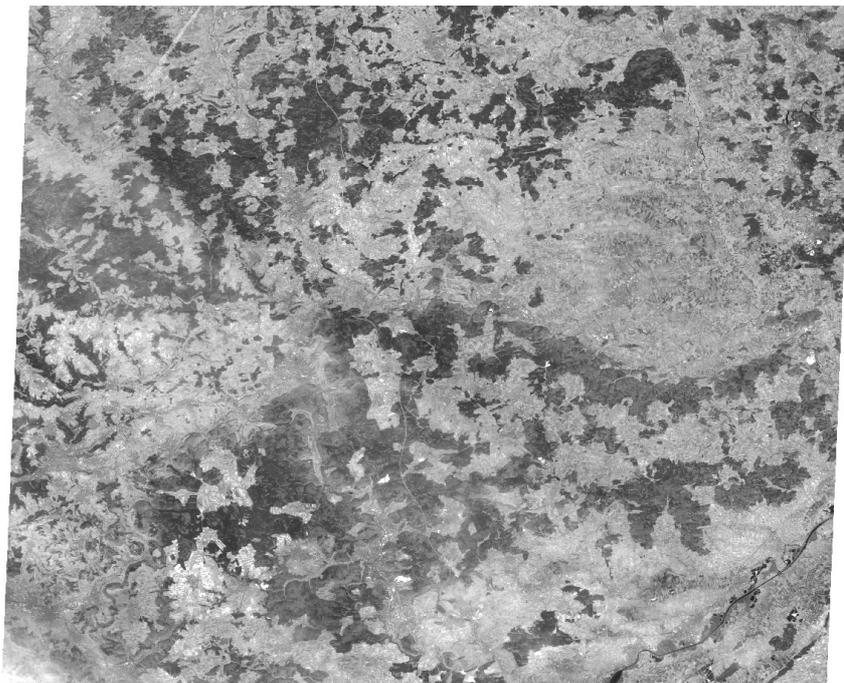
Mit zwei künstlich an die ATKIS-Polygone angebrachten Affintransformationen (Abbildungen 5.27 und 5.28 konnte die Robustheit des gesamten Verfahrens nachgewiesen werden. In beiden Fällen wurden bei exakt gleicher Parameterwahl genau die selben 17 Polygone wie im ungestörten Originalfall verifiziert. Damit läßt sich auch anschaulich der Nutzen der verwendeten affininvarianten Merkmale und der Standardlagentransformation erkennen: Jegliche affine Störung eines der beiden Datensätze beeinflusst das Ergebnis nicht wesentlich.

## 5.5 Georeferenzierung von SPOT XS-Daten über ein SPOT-PAN-Orthophoto

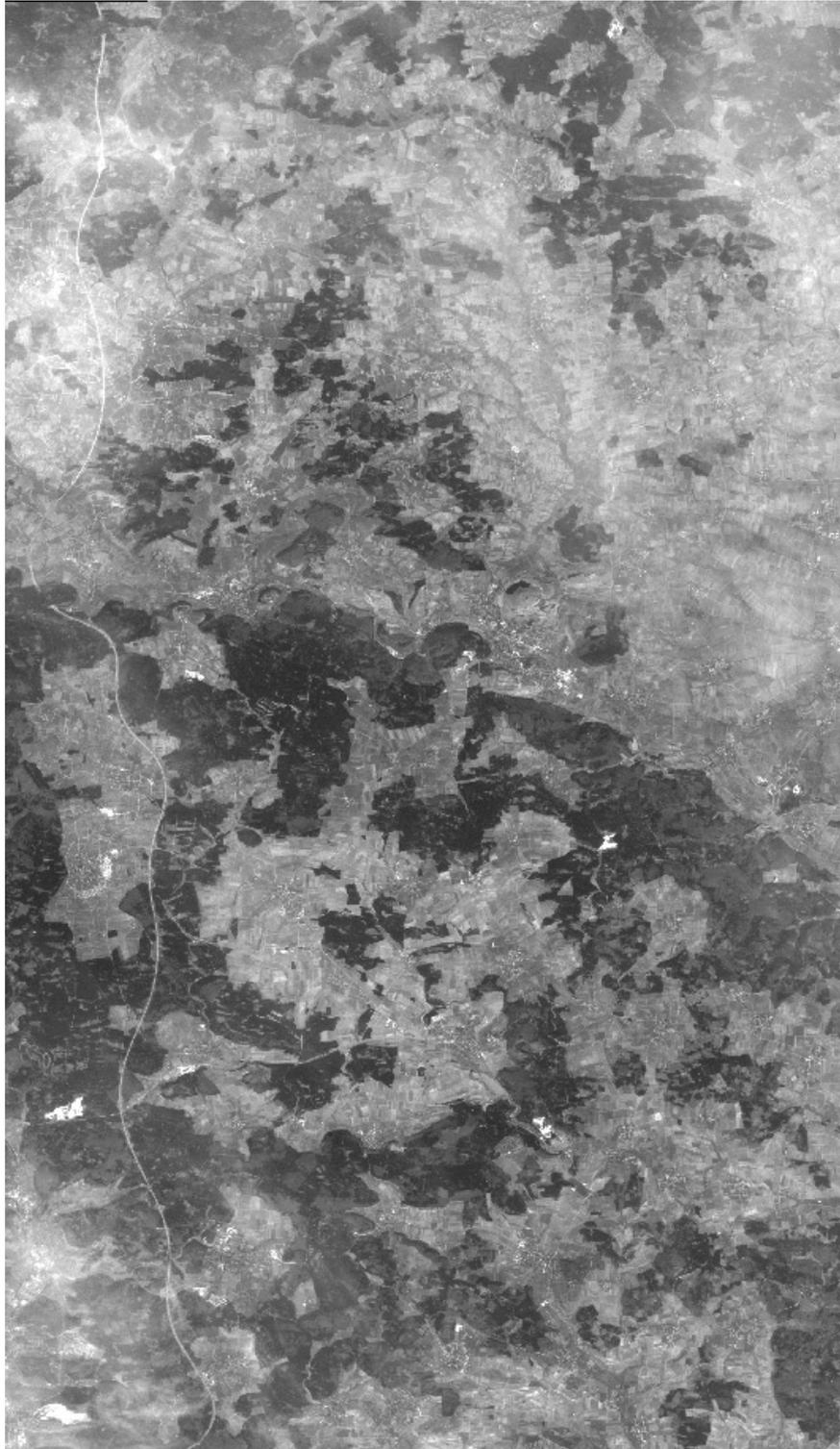
In Abschnitt 5.3 wurde die Fähigkeit des Gesamtsystems, Vektor-Vektor-Zuordnungen zu ermöglichen, anhand simulierter Luftbildaufnahmen aufgezeigt. In den Untersuchungen mit realen SPOT-Daten (Abschnitt 5.4) wurde dagegen eine automatische Raster-Vektor-Zuordnung durchgeführt. Der dritte mögliche Fall, die Zuordnung von Raster- zu Rasterdaten, wird in diesem Abschnitt behandelt.

### 5.5.1 Der Testdatensatz

Zur Untersuchung wurde eine multispektrale SPOT XS-Szene (level 1b) des Untersuchungsgebietes Elchingen/Großkuchen vom 8.5.1998 mit einem SPOT PAN-Orthophoto (1.4.1994) einer Teilszene georeferenziert. Üblicherweise sind die Verhältnisse umgekehrt, d.h. die zu georeferenzierende Szene ist kleiner als das zur Verfügung stehende Orthophoto. Hier wurde dennoch die gesamte SPOT XS-Szene verwendet, um die Leistungsfähigkeit des Verfahrens zu demonstrieren. Abbildung 5.30 zeigt das arithmetische Mittel der drei multispektralen SPOT XS Kanäle, mit dem in dieser Untersuchung gearbeitet wurde, in Abbildung 5.31 ist das verwendete Orthophoto abgebildet.



**Abbildung 5.30:** SPOT XS-Szene (arithmetisches Mittel aus den drei Kanälen) des Untersuchungsgebiets Elchingen/Großkuchen vom 8.5.1998.



**Abbildung 5.31:** Orthophoto aus einer SPOT PAN-Szene des Testgebietes Elchingen/Großkuchen vom 1.4.1994.

Die verwendeten Szenen sind in ihrer vollen Auflösung zu groß für das implementierte System (PAN 2426pixel  $\times$  4193pixel, XS 3709pixel  $\times$  2989pixel). Zur Reduktion der Größe der zu handhabenden Bilder wurde von beiden Szenen die erste Pyramidenstufe berechnet (Reduktionsfaktor 2). Dies nimmt auf das Verfahren nur geringen Einfluß, da die verwendeten Merkmale (vgl. Kapitel 3) vollkommen und die Segmentierungsergebnisse (vgl. Abschnitt 2.1.3) weitgehend maßstabsinvariant sind.

### 5.5.2 Durchführung und Ergebnisse der Untersuchung

Zur Erzeugung radiometrisch homogener Objekte wurde das Orthophoto mit adaptiven Schwellwerten  $AT = 6, 7, 8$  segmentiert. Im Falle der Multispektralszene wurde das Mittelwertbild mit adaptiven Schwellwerten  $AT = 6, 8, 10$  segmentiert. Eine Statistik für die Zahl der erhaltenen Regionen ist in Tabelle 5.5 gegeben.

AT	6	7	8	gesamt
PAN-Regionen	46394	38896	28536	113826
AT	6	8	10	gesamt
XS-Regionen	117816	68966	43582	230364

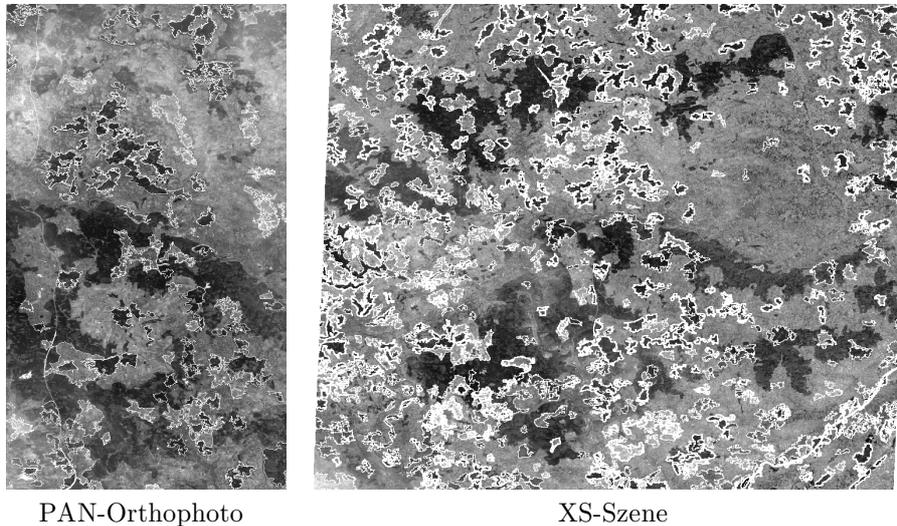
**Tabelle 5.5:** Anzahl erzeugter Segmente in Abhängigkeit des adaptiven Schwellwerts  $h$ .

Um die Rechenzeit für diese Untersuchung überhaupt in Grenzen halten zu können, wurde die Größe der verwendeten Regionen in der ersten Pyramidenstufe für das Orthophoto auf den Bereich [1000pixel, 20000pixel] und für die XS Szene auf den Bereich [250pixel, 5000pixel] beschränkt. In der Originalauflösung sind die Bereichsgrenzen jeweils vierfach so groß. Mit dieser Begrenzung ergibt sich eine vernünftig handhabbare Anzahl an Regionen, wie in Tabelle 5.6 dargestellt ist.

AT	6	7	8	gesamt
PAN-Regionen	80	61	58	199
AT	6	8	10	gesamt
XS-Regionen	496	353	274	1123

**Tabelle 5.6:** Reduzierte Anzahl erzeugter Segmente in Abhängigkeit des adaptiven Schwellwerts  $AT$ .

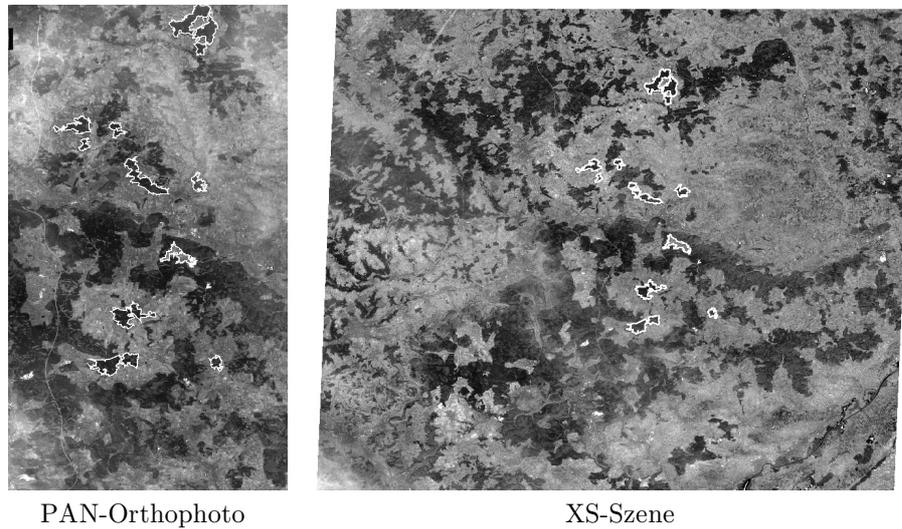
In Abbildung 5.32 sind alle Regionen für beide Szenen abgebildet.



**Abbildung 5.32:** Alle segmentierten Regionen für die SPOT PAN- und die SPOT XS-Szene.

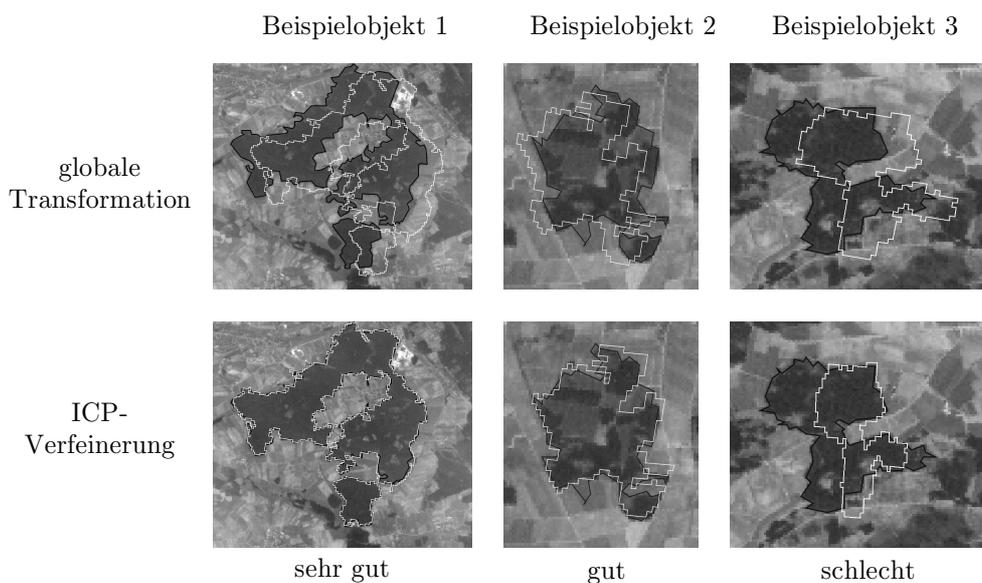
Wie im Flußdiagramm des Gesamtsystems in Abbildung 5.4 dargestellt, wurde zunächst für alle Regionen der Rand in ein Vektorpolygon konvertiert. Zur Reduktion der Zahl der Polygonknoten und für einen stabileren Lauf der später folgenden ICP-Verfeinerung wurden die so erzeugten Polygone aus den Orthophoto-Segmentierungen mit Hilfe des Douglas-Peucker-Algorithmus' [Douglas & Peucker 1973] generalisiert. Der Generalisierungsfaktor wurde zu 1.2 gewählt. Für alle Vektorpolygone der PAN- und XS-Segmentierungen wurden anschließend die zentralen Momente, die acht einfachsten CWM-Invarianten und die Standardlagen berechnet. Für die merkmalsbasierte Zuordnung der beiden Polygonmengen wurde mit einem maximalen relativen Unterschied von 2% gearbeitet. Aus den  $199 \times 1123 = 233477$  kombinatorischen Möglichkeiten wurden so 20254 hypothetische Polygon-Paare ermittelt. Für jedes Polygon dieser Paare wurde das MBR berechnet. Die Durchführung der

Baumsuche nach einer konsistenten Interpretation erfolgte bis zu einer Tiefe von 5 mit einer minimalen Objektüberlappung von 70% pro Paar. Es konnte eine konsistente Interpretation mit fünf Paaren gefunden werden, auf deren Basis eine globale Affintransformation berechnet wurde. Mit Hilfe dieser Transformation wurde der Verifikationsschritt durchgeführt. Die dabei verwendeten Parameter waren: maximale Winkeldifferenz:  $45^\circ$ , maximaler Maßstabsunterschied: 50%, maximaler Translationsunterschied: 50%, minimale Überlappung: 40%. Diese Werte sind relativ weit gefasst. Damit wird der durch die beidseitige Segmentierung erhöhten Unsicherheit der Objektkontur Rechnung getragen. In diesem Verifikationsschritt konnten 12 Paare verifiziert werden, die für beide Datensätze in Abbildung 5.33 dargestellt sind.



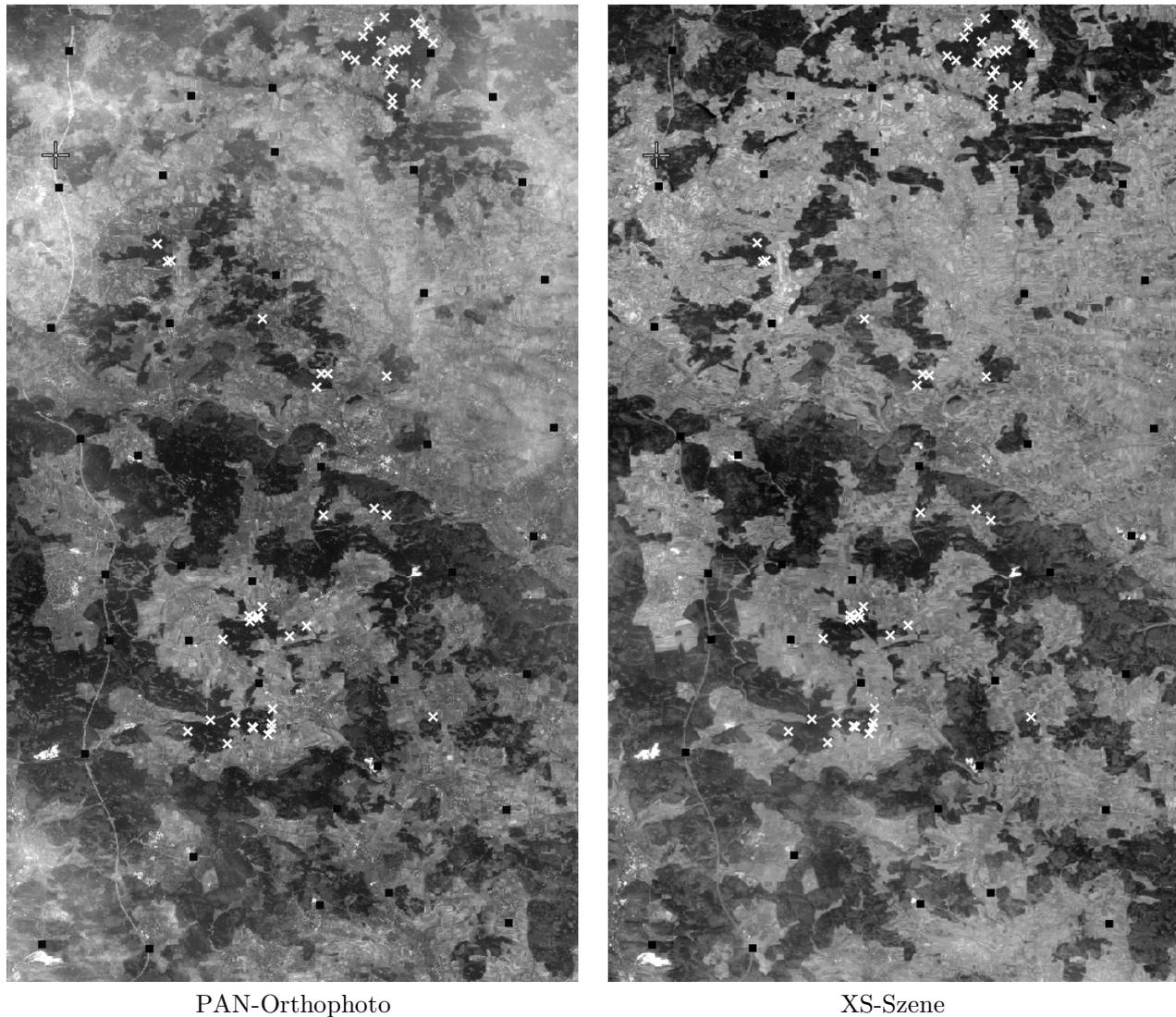
**Abbildung 5.33:** Verifizierte Polygonpaare für die Registrierung der SPOT XS-Szene Elchingen/Großkuchen mit einem SPOT PAN-Orthophoto.

Für die mittels der globalen Transformation in das Orthophoto transformierten XS-Polygone wurde die ICP-Verfeinerung durchgeführt. In dieser Untersuchung wurde mit einer Abtastweite von 10pix, einem Punktakzeptanzintervall von  $2\sigma$  und einer Iterationsanzahl von 100 gearbeitet (zur ICP-Verfeinerung vgl. Abschnitt 4.8.2). Infolge der beidseitigen Segmentierung und des Maßstabsunterschiedes der beiden Szenen ist die Qualität der verfeinerten Zuordnung nicht für alle Objekte gleich gut. Beispiele sind in Abbildung 5.34 angeführt.



**Abbildung 5.34:** Beispiele für die unterschiedliche Qualität der ICP-verfeinerten Zuordnung von SPOT XS-Polygonen zu SPOT PAN-Polygonen.

Auf den Konturen der verfeinert zugeordneten Polygone wurden im Orthobild 50 Paßpunkte generiert. Aus diesen 2D-Paßpunkten wurde eine affine Transformation geschätzt, mit der die gesamte XS-Szene letztendlich georeferenziert werden konnte. Zur Ermittlung der Genauigkeit der Zuordnung wurden 40 Kontrollpunkte manuell eingemessen. In Abbildung 5.35 ist das Orthophoto und die automatisch georeferenzierte SPOT XS-Szene abgebildet. Die automatisch erzeugten Paßpunkte sind durch weiße Kreuze, die manuell gemessenen Kontrollpunkte durch schwarze Quadrate dargestellt.



**Abbildung 5.35:** Ergebnis der automatischen Paßpunkterzeugung und Georeferenzierung der SPOT XS-Szene (Ausschnitt auf Orthophotogröße). 50 automatisch erzeugte Paßpunkte sind durch weiße Kreuze, die 40 manuell gemessenen Kontrollpunkte durch schwarze Quadrate dargestellt.

Die Ergebnisse der Genauigkeitsuntersuchung sind zusammengefasst in Tabelle 5.7 dargestellt.

	Alle Kontrollpunkte	Nur "innere" Kontrollpunkte
Minimaler Fehler	5.9m	5.9m
Mittlerer Fehler	61.9m	10.0m
Maximaler Fehler	366.4m	39.4m

**Tabelle 5.7:** Ergebnisse der Genauigkeitsuntersuchung für automatisch erzeugte Paßpunkte zur Georeferenzierung von SPOT XS- mit SPOT PAN-Daten.

### 5.5.3 Diskussion

Die Zuordnung von Raster- zu Rasterdaten, die von zwei verschiedenen Sensoren unterschiedlicher Auflösung zu vier Jahre auseinander liegenden Zeitpunkten aufgenommen wurden, stellt in dieser Untersuchungsreihe die schwierigste Aufgabe für das entwickelte System dar. Zusätzlich sind die hier verwendeten Szenen größer als in den vorhergehenden Untersuchungen. Durch die beidseitig notwendige Segmentierung zur Bildung radiometrisch homogener Objekte entsteht eine größere Unsicherheit hinsichtlich der Objektkontur als im Vektor-Raster-Fall. Dort werden zumindest die Vektorobjekte sicher und genau erzeugt. Durch die Natur der Segmentierung und die infolge fehlender Attributierung wegfallende Selektionsmöglichkeit stehen wesentlich mehr Objekte zur Verarbeitung an als in den anderen Fällen. Daher wurde eine Beschränkung auf Objekte eines gewissen Größenbereichs vorgenommen. Zudem wirkt sich die bewusst unterschiedlich gewählte räumliche Ausdehnung der Szenen erschwerend aus, da - wie in allen anderen Untersuchungen auch - ohne Näherungswerte gearbeitet wurde. Die Tatsache, daß trotz der hohen Zahl von 20254 erzeugten Hypothesen eine konsistente Interpretation gefunden werden konnte, spricht für die Robustheit des entwickelten Verfahrens. Die wesentlichen Gründe hierfür sind in der Verwendung affin invarianter Merkmale und der Baumsuche mit sinnvollen Zwangsbedingungen zu finden. Die zwölf verifizierten Objekte in Abbildung 5.33 sind nicht optimal über das Orthophoto verteilt, was sicherlich auch an den teilweise vorhandenen Schleiern in der PAN-Szene liegt. In diesen Bereichen liefert die Segmentierung keine zuverlässigen Objektkonturen und daher lassen sich hier auch keine Objekte verifizieren bzw. werden erst gar keine Hypothesen vom Zuordnungsverfahren gefunden.

Die aus den verifizierten Polygonen erzeugte Näherungstransformation liefert innerhalb der konvexen Hülle dieser Polygone bereits eine gute Näherung für die Georeferenzierung (vgl. die Beispiele in Abbildung 5.34, obere Reihe). Außerhalb dieser Bereiche sind die Fehler infolge der Extrapolation der Transformation naturgemäß größer. Die individuelle Verfeinerung der objektweisen Transformation mit Hilfe des modifizierten ICP-Algorithmus' führt jedoch nicht in allen Fällen zu einem optimalen Ergebnis (vgl. die Beispiele in Abbildung 5.34, untere Reihe). Dadurch stehen für die abschließende Paßpunkterzeugung weniger Linienstücke zur Verfügung. Wie man jedoch in Abbildung 5.35 erkennen kann, sind die erzeugten Paßpunkte hinreichend gut innerhalb der konvexen Hülle aller verifizierten Polygone (also des theoretisch möglichen Bereichs) verteilt. Die Qualität der aus den Paßpunkten geschätzten Affintransformation wurde mit Hilfe von 40 gleichmäßig über den Orthophotobereich verteilten Kontrollpunkten überprüft. Der besseren Vergleichbarkeit mit anderen Verfahren halber wurde die Genauigkeit in Weltkoordinaten ermittelt. Die Güte des Ergebnisses wird beeinflusst von der Tatsache, daß die gesamte Zuordnung auf der ersten Pyramidenstufe durchgeführt wurde. Damit verdoppeln sich die Fehler, die durch die affine Näherung und das Verfahren an sich eingebracht werden. Auf der tatsächlich verarbeiteten ersten Pyramidenstufe bedeutet der mittlere Fehler von 39.4m innerhalb der konvexen Hülle der verifizierten Polygone einen Bildfehler von knapp 2pixel im Orthophoto (10m Bodenpixelgröße) und knapp 1pix im Multispektralbild (20m Bodenpixelgröße). Diese Genauigkeit stellt eine hinreichend gute Näherung für ein automatisches Punktzuordnungsverfahren dar, mit dem die Zahl und Qualität der erzeugten Paßpunkte wesentlich erhöht werden könnte. Eine aussichtsreiche Strategie zur Gewinnung sehr vieler hochgenauer und räumlich gut verteilter Paßpunkte bestünde daher darin, in einem ersten Schritt das hier entwickelte Verfahren durchzuführen und auf der erhaltenen Näherung aufbauend in einem zweiten Schritt ein intensitätsbasiertes Punktzuordnungsverfahren durchzuführen.

## Kapitel 6

### Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurde ein System zur automatischen Georeferenzierung von Fernerkundungsdaten entwickelt, implementiert und getestet. Das Verfahren erlaubt die Zuordnung von Raster- zu Rasterdaten, Raster- zu Vektordaten und Vektor- zu Vektordaten und stellt somit einen flexiblen Ansatz für eine Vielzahl von indirekten Georeferenzierungsaufgaben dar. Es wurde ein merkmalsbasiertes Zuordnungsverfahren gewählt, das Polygone als Elementarobjekte benutzt. Extraktionsverfahren für punkt- oder linienförmige Objekte, die für eine breite Palette von Daten (Luftbilder, Radardaten, optische Daten verschiedenster Spektralbereiche und Pixelgrößen) zuverlässig funktionieren, sind kaum zu finden. Daher wurde für diese Arbeit auf die Benutzung von Punkten und Linien als Elementarobjekte verzichtet. Aufgrund der universellen Ausrichtung des Systems kommt die Verwendung einer strengen geometrischen Transformation hier nicht in Frage, da dafür jeder Sensor eine spezifische Behandlung verlangen würde.

Die Zuordnung der Datenquellen erfolgt in zwei Schritten. Im ersten Schritt wird eine globale Affintransformation ermittelt, mit der die Eingabedatensätze anhand zugeordneter Polygone bestmöglich ineinander transformiert werden. Im zweiten Schritt wird die globale Transformation für jedes Polygonpaar durch ein modifiziertes Konturzuordnungsverfahren, den Iterative-Closest-Point-Algorithmus, optimiert. Auf den Konturen der derart verfeinerten Polygonpaare werden automatisch homologe Punkte erzeugt, die für eine abschließende Bildentzerrung als 2D-Paßpunkte zur Verfügung stehen. Durch die Verwendung von Merkmalen, die absolut invariant unter einer beliebigen affinen Transformation sind, kommt das entwickelte Verfahren jedoch gänzlich ohne Näherungswerte aus. Zudem bestehen nahezu keine Einschränkungen bezüglich der Bodenpixelgröße oder des Kartenmaßstabs. Die wesentliche Voraussetzung für eine Konvergenz des Verfahrens stellt die Existenz gleicher oder ähnlicher Polygone in beiden Eingabedatensätzen dar.

In einer Einführung wird zunächst die photogrammetrisch strenge geometrische Transformation zwischen Fernerkundungsdaten und einem Zielkoordinatensystem dargelegt. Als Grundlage für die mit diesem System bestimmte Näherung werden die gebräuchlichen Verfahren zur vereinfachten Georeferenzierung erläutert. Ein Überblick über Arbeiten, die mit dem Ziel dieser Arbeit in Verbindung stehen, rundet die Einführung ab.

Die Objekterzeugung aus Rasterdaten erfolgt mit Hilfe einer Segmentierung. Aufbauend auf einer Übersicht über die gängigen Segmentierungsverfahren wird die benutzte Segmentierung, ein Bereichswachstumverfahren mit adaptivem Schwellwert, detailliert erläutert. Die prinzipielle Tauglichkeit des Verfahrens für die Einbindung in das entwickelte System wird durch Untersuchungen mit verschiedenen Schwellwerten und in verschiedenen Pyramidenstufen belegt. Durch automatische multiple Segmentierung wird die Anzahl benutzerspezifischer Parameter auf ein Mindestmaß reduziert. Mit dieser Strategie wird angenommen, daß im Rahmen der durchgeführten Segmentierungen jedes potenziell für eine Zuordnung in Frage kommende Polygon mindestens einmal "richtig" segmentiert wird. Das Resultat der multiplen Segmentierung ist eine Menge von Rasterpolygonen mit homogener Grauwertverteilung.

Für attributierte Vektorkarten erfolgt die Objekterzeugung einfach durch Verschmelzung benachbarter Polygone gleicher Objektart. Dadurch wird bestmöglich sichergestellt, daß Objekte, die im Bild als homogene Grauwertbereiche erscheinen, auch im Vektordatensatz vorhanden sind. Rasterkarten werden in dieser Arbeit aufgrund des zusätzlich notwendigen Aufwands für die Objektextraktion nicht als Referenzdaten verwendet. Prinzipiell kommen sie jedoch genauso wie Vektorkarten als Referenzdaten in Betracht.

Da die globale Beziehung zwischen den Datenquellen in dieser Arbeit durch eine affine Transformation hergestellt wird, stellt die Gewinnung affin invarianter Merkmale aus Polygonen einen Schwerpunkt der Arbeit dar. Basierend auf affinen Invarianten algebraischer Formen wird zunächst dargelegt, wie absolute Invarianten unter affiner Transformation mit Hilfe eines Fundamentalsatzes aus geometrischen Momenten eines Polygons berechnet werden können. Unter Benutzung des Green'schen Integralsatzes wird eine Möglichkeit dargestellt, geometrische Momente beliebiger Ordnung direkt aus der Objektkontur zu berechnen. Durch Untersuchungen zur Empfindlichkeit dieser Momentinvarianten unter statistischem Rauschen und dem Einfluß der Diskretisierung läßt sich die Robustheit dieser Merkmale belegen. Alternativ werden Affinvarianten aus Kreuzmomenten (CWM-Invarianten) abgeleitet und ebenfalls auf ihre Empfindlichkeit gegenüber statistischem Rauschen und Diskretisierung der Objektkontur untersucht. Im Vergleich erweisen sich die CWM-Invarianten um etwa einen Faktor 50 robuster als Momentinvarianten und werden daher als affin invariante Polygonmerkmale für das entwickelte Verfahren ausgewählt. Neben der Invarianz ist die Möglichkeit zur Unterscheidbarkeit von Objekten mit unterschiedlicher Kontur ein wesentliches Kriterium für sinnvolle Merkmale. Durch eine Untersuchung an synthetisch erzeugten Daten mit z.T. mehrfachen Symmetrien wird die Eignung der sechs einfachsten CWM-Invarianten als objektbeschreibende Merkmale bestätigt.

Einen weiteren Schwerpunkt stellt die Wahl eines geeigneten Zuordnungsverfahrens für die erzeugten Polygone auf der Basis ihrer invarianten Merkmale dar. Für diese Arbeit wurde dazu ein sogenanntes "backtracking-depth-first"-Baumsuchverfahren benutzt, das detailliert erläutert und motiviert wird. Zur Einschränkung des exponentiell wachsenden Suchraums werden unäre und binäre Zwangsbedingungen benutzt, die ebenfalls detailliert beschrieben werden. Zur weiteren Reduktion des Suchaufwands nach einer Gruppe konsistenter Polygonpaare wird die Baumsuche nach Erreichen einer bestimmten Tiefe abgebrochen. Zusätzlich wird eine Heuristik benutzt, durch die Paare großer Objekte zuerst bearbeitet werden. Für die Prüfung der Zwangsbedingungen wird bereits während der Baumsuche eine affine Transformation zwischen den Polygonen einer Hypothese benötigt. Da zu diesem Zeitpunkt außer den Schwerpunkten keine homologen Punkte bekannt sind, wird ein alternatives Verfahren benötigt. Aus den bereits berechneten Momenten läßt sich für jedes Polygon durch geschickte Normierung eindeutig eine sogenannte Standardlage definieren, über die die Transformation zwischen zwei Polygonen ohne Punktkorrespondenz hergestellt werden kann. Mögliche Verfahren zur Standardlagentransformation werden beschrieben. Untersuchungen des Verhaltens der so erzeugten Polygontransformation unter affiner Transformation und statistischem Rauschen in der Objektkontur zeigen eine hinreichende Robustheit dieser Methode. Anhand eines Beispiels wird das gesamte globale Zuordnungsverfahren einschließlich der Verifikation illustriert. Aus der Menge verifizierter Polygonpaare ergibt sich durch gewichtete Mittelung eine globale Affintransformation.

Zur Optimierung der globalen Transformation wurde der aus der Computer Vision bekannte Iterative-Closest-Point-(ICP)-Algorithmus dahingehend modifiziert, daß nur ein Teil der Konturpunkte zur Transformations-schätzung beiträgt. Der Algorithmus und die angebrachte Modifikation werden beschrieben. Durch verschiedene Untersuchungen wird gezeigt, daß dieser Algorithmus bei affinen Näherungen, die bezüglich einer Translation weniger als  $\pm 10\%$  der Objektausdehnung, bezüglich einer Rotation weniger als  $\pm 15^\circ$  und bezüglich einer Skalierung weniger als 20% abweichen, zu einer Verbesserung der Transformation benutzt werden kann.

Basierend auf einer Abschätzung des Fehlers, den die in dieser Arbeit benutzte affine Transformation gegenüber einer perspektiven Abbildung bewirkt, wird das Flußdiagramm des Gesamtsystems entwickelt. Die einzelnen Prozessierungsschritte werden detailliert erläutert. Anhand der drei möglichen Kombinationen von Raster- und Vektordaten wurde das Verhalten des Gesamtsystems untersucht.

Zur Untersuchung des reinen Einflusses der Topographie wurden virtuelle Luftbilder eines 3D-Datensatzes (ATKIS-Daten + DHM) mit Kammerkonstanten von 150mm und 300mm für Flughöhen von 750m bis 3000m berechnet. Die virtuellen Bilder stellen ideal segmentierte Bilder dar. Für eine Kammerkonstante von 150mm konnte der Fehler, der rein durch topographische Verzerrungen verursacht wird, in vollautomatisch erzeugten Paßpunkten zu etwa 80cm ermittelt werden. Im Falle der Kammerkonstanten von 300mm liegt die Standardabweichung bei etwa 50cm. Im Vergleich zur direkten Georeferenzierung sind diese Werte etwa um einen Faktor 10 schlechter, allerdings sind für dieses Verfahren keine Zusatzinformationen nötig.

Die Georeferenzierung einer SPOT PAN-Szene mit Hilfe von 2D-ATKIS-Daten fällt in den Bereich Raster-Vektor-Zuordnung. Hier konnten mit diesem Verfahren für eine  $18.5\text{km} \times 19\text{km}$  große Szene 177 gut verteilte Paßpunkte mit einer Genauigkeit von 11m vollautomatisch erzeugt werden. Durch verschiedene künstliche affine Transformationen der SPOT PAN-Szene konnte zudem mit nahezu identischen Ergebnissen die Robustheit des Verfahrens nachgewiesen werden.

In einer dritten Untersuchung wurde eine vollständige multispektrale SPOT XS-Szene über ein SPOT PAN-Orthophoto georeferenziert. Zur Reduktion des Rechenaufwandes wurden die Bilder jeweils um einen Faktor 2 verkleinert. Hier konnten 50 Paßpunkte mit einer Genauigkeit von 39.4m, also etwa einem Pixel in der ersten Pyramidenstufe der SPOT XS-Szene erzeugt werden.

Ein Schwerpunkt während der vorgestellten Arbeit lag in der Formulierung einer durchgängigen Lösung und nicht in der vollständigen Optimierung der Einzelschritte. Daher sind in mehreren Bereichen Verbesserungsmöglichkeiten für das vorgeschlagene Verfahren gegeben:

**Objektextraktion.** Das benutzte Segmentierungsverfahren stellt mit Sicherheit nicht die obere Grenze des Möglichen dar. Durch ein raffinierteres Verfahren, eventuell mit Rückkoppelung aus dem Referenzsystem ließe sich eine bessere Bildsegmentierung und damit auch genauere Paßpunkte erreichen.

**Performance.** Das gesamte Verfahren ist in ein kommerzielles System (KBVision) eingebettet und wurde in C/C++ implementiert. Durch die Architektur dieses Systems und durch die große Zahl an Funktionsaufrufen entstehen höhere Rechenzeiten als nötig. Im Bereich der Baumsuche ließe sich zudem durch diverse algorithmische Verbesserungen weitere Rechenzeit einsparen.

**Genauigkeit.** Durch die Hinzunahme linien- und punktförmiger Objekte, die mit Hilfe der Näherungstransformation aus diesem Verfahren extrahiert werden könnten, ließe sich die Qualität der erzeugten Paßpunkte ebenfalls erhöhen. Die erreichbaren Genauigkeiten bleiben dabei naturgemäß hinter denen einer strengen Transformation oder - zumindest im Luftbildfall - einer direkten Georeferenzierung zurück.

Die Stärken des entwickelten Verfahrens liegen in folgenden Punkten:

**Robustheit.** Polygone stellen eine in vielen Fernerkundungsdaten zuverlässig extrahierbare Objektklasse dar. Durch die ausschließliche Verwendung invarianter Merkmale ist die Konvergenz des Verfahrens stabil, selbst bei Fällen, die in der Praxis nicht auftreten (z.B. Spiegelungen).

**Flexibilität.** Die verwendeten Merkmale basieren auf der Polygonkontur. Radiometrische Unterschiede innerhalb der Polygone spielen dabei nur eine untergeordnete Rolle durch die Segmentierung, solange die relevanten Objekte die verwendeten Homogenitätskriterien erfüllen.

**Einfachheit.** Durch die Verwendung einer einfachen Transformation, durch multiple Segmentierung und durch eine Grob-zu-Fein-Strategie im Zuordnungsverfahren ist die Zahl benutzerdefinierbarer Parameter sehr beschränkt.

Mit den angestellten Untersuchungen konnte anhand eingehender Untersuchungen und einer Vielzahl an synthetischen und realen Beispielen nachgewiesen werden, daß der entwickelte Lösungsweg zum Ziel führt und damit das implementierte System für eine automatische Georeferenzierung von Fernerkundungsdaten unterschiedlicher Natur gut geeignet ist.

## Literaturverzeichnis

- Abbasi-Dezfouli, M. & Freeman, T. [1994], Patch matching in stereo-images based on shape, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXX, Part 3/1, München, pp. 1–8.
- Abu-Mostafa, Y. & Psaltis, D. [1984], ‘Recognitive aspects of moment invariants’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 698–706.
- Ackermann, F. [1983], High precision digital image correlation, *in* ‘Photogrammetric Week ’83’, Stuttgart, pp. 231–243.
- Ackermann, F. & Hahn, M. [1991], Image pyramids for digital photogrammetry, *in* ‘Digital Photogrammetric Systems’, Wichmann Verlag, Karlsruhe, pp. 43–57.
- AdV [1988], *Amtliches Topographisch-Kartographisches Informationssystem (ATKIS)*, Arbeitsgemeinschaft der Länder der Vermessungsverwaltungen der Bundesrepublik Deutschland, Bonn.
- Albertz, J. [1991], *Grundlagen der Interpretation von Luft- und Satellitenbildern: eine Einführung in die Fernerkundung*, Wiss. Buchgesellschaft, Darmstadt.
- Albertz, J. & Kreiling, W. [1989], *Photogrammetrisches Taschenbuch*, Wichmann Verlag, Karlsruhe.
- Arbter, K., Snyder, W. E. & Burkhardt, H. H. G. [1990], ‘Application of affine-invariant fourier descriptors to recognition of 3-D objects’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**, 640–647.
- Ballard, D. H. & Brown, C. M. [1982], *Computer Vision*, Prentice-Hall Inc., Englewood Cliffs, NJ.
- Barnard, S. T. & Thompson, W. B. [1980], ‘Disparity analysis of images’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2**, 333–340.
- Bartl, R. & Schneider, W. [1995], Information fusion based on geometrical arrangement, *Schriftenreihe der OCG*, Band 81, R. Oldenburg, Wien-München, pp. 50–58.
- Bückner, J. [1998], Model based road extraction for the registration and interpretation of remote sensing data, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, Part 4, Stuttgart, pp. 85–90.
- Besl, P. J. & McKay, N. D. [1992], ‘A method for registration of 3-D shapes’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**, 239–256.
- Bill, R. & Fritsch, D. [1991], *Grundlagen der Geo-Informationssysteme, Band I: Hardware, Software und Daten*, Wichmann, Karlsruhe.
- Bing-Cheng, L. [1993], ‘The moment calculation of polyhedra’, *Pattern Recognition* **26**, 1229–1233.
- Bronstein, I. N. [1987], *Taschenbuch der Mathematik*, Harri Deutsch Verlag, Thun.
- Bässmann, H. & Behrlich, P. W. [1991], *Bildverarbeitung Ad Oculos*, Springer-Verlag, Berlin.
- Cayley, A. [1856], ‘A second memoir upon quantics’, *Philosophical Transactions of the Royal Society of London* pp. 110–126.
- Clebsch, A. [1872], *Theorie der binären algebraischen Formen*, B.G. Teubner, Leipzig.
- Courant, R. & Hilbert, R. [1953], *Methods of Mathematical Physics*, Interscience, New York.
- Cramer, M. [2000], *Genauigkeitsuntersuchungen zur GPS/INS-Integration in der Aerophotogrammetrie*, Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 537, München.
- Dare, P. & Dowman, I. [2000], A new approach to automatic feature based registration of SAR and SPOT images, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII, Part B2, Amsterdam.
- Demant, C., Streicher, A. B. & Waskewitz, P. [1998], *Industrielle Bildverarbeitung*, Springer, Berlin.
- Douglas, D. H. & Peucker, T. K. [1973], ‘Algorithms for the reduction of the number of points to represent a digitized line or its caricature’, *Canadian Cartographer* **10**, 112–122.
- Dowman, I. [1998], ‘Automatic image registration and absolute orientation: Solutions and problems’, *Photogrammetric Record* **16**, 5–18.
- Dowman, I. & Dare, P. [1999], Automated procedures for multisensor registration and orthorectification of satellite images, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, Part 7-4-3W6, Valladolid, pp. 37–44.
- Dowman, I. & Dolloff, J. [2000], An evaluation of rational functions for photogrammetric restitution, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII, Part B3, Amsterdam, pp. 254–266.
- Dowman, I. & Ruskoné, R. [1997], Extraction of polygonal features from satellite images for automatic registration: The ARCHANGEL project, *in* ‘Automatic Extraction of Man-Made Objects from Aerial and Space Images (II) Monte Verità’, Birkhäuser Verlag, Bern, pp. 343–354.
- Drewniok, C. & Rohr, K. [1995], Automatic exterior orientation of aerial images in urban environments, *in* ‘Mustererkennung 1995, 17. DAGM Symposium’, Bielefeld, pp. 594–601.
- Dörstel, C., Tang, L. & Madani, M. [2001], Automatic aerial triangulation software of Z/I imaging, *in* ‘Photogrammetric Week ’01’, Wichmann Verlag, Heidelberg, pp. 177–181.

- Dudani, S. A., Breeding, K. J. & McGhee, R. B. [1977], 'Aircraft identification by moment invariants', *IEEE Transactions on Computers* **C**, 39–46.
- Ebner, H., Fritsch, D., Gillessen, W. & Heipke, C. [1987], 'Integration von Bildzuordnung und Objektrekonstruktion in der digitalen Photogrammetrie', *Bildmessung und Luftbildwesen* **55**(5), 194–203.
- Ebner, H. & Ohlhof, T. [1994], Utilisation of ground control points for image orientation without tie point identification in image space, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXX, Part 3/1, München, pp. 206–211.
- Flusser, J. & Suk, T. [1993], 'Pattern recognition by affine moment invariants', *Pattern Recognition* **26**, 167–174.
- Fonseca, L. M. G. & Manjunath, B. S. [1996], 'Registration techniques for multisensor remotely sensed imagery', *Photogrammetric Engineering & Remote Sensing* **62**, 1049–1056.
- Freeman, H. [1961], 'On the encoding of arbitrary geometric configurations', *IEEE Transactions on Electronic Computers* **10**, 260–268.
- Fritsch, D. & Stallmann, D. [2000], Rigorous photogrammetric processing of high resolution satellite imagery, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII, Part B1, Amsterdam, pp. 313–321.
- Förstner, W. [1982], On the geometric precision of digital correlation, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXIV, Part 3, Helsinki, pp. 176–189.
- Förstner, W. [1993], 'Feature extraction for digital photogrammetry', *Photogrammetric Record* **14**, 585–611.
- Förstner, W. & Gülch, E. [1987], A fast operator for detection and location of distinct points, corners and centers of circular features, in 'ISPRS Intercommision Workshop on Fast Processing of Photogrammetric Data', Interlaken, pp. 281–305.
- Gonzalez, R. C. & Wintz, P. [1977], *Digital Image Processing*, Addison-Wesley Publishing Company, Berlin.
- Grafarend, E. W. [1995], 'The optimal universal transverse mercator projection', *manuscripta geodaetica* **20**, 421–468.
- Grafarend, E. W. & Shan, J. [1997a], 'Estimable quantities in projective networks, I', *Zeitschrift für Vermessungswesen* **122**, 218–226.
- Grafarend, E. W. & Shan, J. [1997b], 'Estimable quantities in projective networks, II', *Zeitschrift für Vermessungswesen* **122**, 323–333.
- Grimson, W. E. L., Lozano-Pérez, T., White, S. J. & Noble, N. [1993], Recognizing 3d objects using constrained tree search, in A. K. Jain & P. J. Flynn, eds, 'Three-Dimensional Object Recognition Systems', Elsevier, pp. 259–282.
- Grove, S. & Tönnies, R. [1997], A knowledge based approach to automatic image registration, in 'IEEE International Conference on Image Processing', Santra Barbara, CA.
- Gurevich, G. B. [1964], *Foundation of the theory of algebraic invariants*, P. Noordhoff, Groningen, Niederlande.
- Haala, B., Hahn, M. & Schmidt, D. [1994], Quality and performance analysis of automatic relative orientation, in 'SPIE, Vol.1944', pp. 140–150.
- Haala, N. & Vosselmann, G. [1992], Recognition of road and river patterns by relational matching, *International Archives of Photogrammetry and Remote Sensing*, Vol. IXXX, Part B3, Washington, pp. 969–975.
- Haberäcker, P. [1987], *Digitale Bildverarbeitung*, Carl Hanser Verlag, München.
- Hahn, M. & Förstner, W. [1988], The applicability of a feature based and least squares matching algorithm for DEM-acquisition, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXVII, Part B9, Kyoto, pp. III – 137–150.
- Hannah, M. J. [1989], 'A system for digital stereo image matching', *Photogrammetric Engineering & Remote Sensing* **55**, 1765–1770.
- Haralick, R. M. & Elliott, G. L. [1980], 'Increasing tree search efficiency for constraint satisfaction problems', *Artificial Intelligence* **14**, 263–313.
- Haralick, R. & Shapiro, L. [1992], *Computer and Robot Vision, Vol.1*, Addison-Wesley, Reading, Mass.
- Haralick, R. & Shapiro, L. [1993], *Computer and Robot Vision, Vol.2*, Addison-Wesley, Reading, Mass.
- Heipke, C. [1995], *Digitale photogrammetrische Arbeitsstationen*, Habilitationsschrift, Deutsche Geodätische Kommission, Reihe C, Nr. 450, München.
- Heipke, C. [1997], 'Automation of interior, relative and absolute orientation', *ISPRS Journal of Photogrammetry & Remote Sensing* **52**, 1–19.
- Heipke, C., Jacobsen, K. & Wegmann, H. [2001], The OEEPE test on integrated sensor orientation - Results of phase I, in 'Photogrammetric Week '01', Wichmann Verlag, Heidelberg, pp. 195–204.
- Hellwich, O., Heipke, C., Tang, L., Ebner, H. & Mayr, W. [1994], Experiences with automatic relative orientation, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXX, Part 3/1, München, pp. 370–378.
- Höhle, J. [1998], Automatic orientation of aerial images by means of existing orthoimages and height data, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, Part 2, Cambridge, pp. 121–126.

- Hilbert, D. [1890], 'Über die Theorie der algebraischen Formen', *Math. Ann.* pp. 473–534.
- Hilbert, D. [1893], 'Über die vollen Invariantensysteme', *Math. Ann.* **42**, 313–373.
- Hild, H. [2000], A strategy for automatic image-to-map-registration, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B2, Amsterdam, pp. 287–294.
- Hild, H. & Fritsch, D. [1998], Integration of vector data and satellite imagery for geocoding, International Archives of Photogrammetry and Remote Sensing, Vol. XXXII, Part 4, Stuttgart, pp. 246–251.
- Holm, M. [2001], An integrated approach for orthophoto production, in 'Photogrammetric Week '01', Wichmann Verlag, Heidelberg, pp. 249–253.
- Holm, M., Parmes, E., Andersson, K. & Vuorela, A. [1995], A nationwide automatic satellite image registration system, in 'SPIE Aerosense '95 - Conference on Integrating Photogrammetric Techniques with Scene Analysis & Machine Vision II', SPIE Vol. 2486-18, Orlando, Florida, USA, pp. 156–167.
- Hough, P. V. C. [1962], 'A method and means for recognizing complex patterns', *U.S. Patent 3,069,654* .
- Hu, M. K. [1962], 'Visual pattern recognition by moment invariant', *IRE Transactions on Information Theory* **8**, 179–187.
- Illert, A. [1990], *Automatische Erfassung von Kartenschrift, Symbolen und Grundrißobjekten aus der Deutschen Grundkarte 1:5000*, Dissertation, Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 166.
- Jähne, B. [1991], *Digitale Bildverarbeitung*, Springer-Verlag, Berlin.
- Kass, M., Witkin, A. & Terzopoulos, D. [1987], Snakes: Active contour models, in '1st Int. Conf. on Computer Vision', IEEE Comp. Soc. Press, pp. 259–268.
- Khotanzad, A. & Hong, Y. H. [1990], 'Invariant image recognition by zernike moments', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**, 489–498.
- Klang, D. [1998], Automatic detection of changes in road databases using satellite imagery, International Archives of Photogrammetry and Remote Sensing, Vol. XXXII, Part 4, Stuttgart, pp. 293–298.
- Kraus, K. [1997], *Photogrammetrie, Band 1*, Dümmler, Bonn.
- Kreiling, W. [1976], *Automatische Erstellung von Höhenmodellen und Orthophotos durch digitale Korrelation*, Dissertation, Universität Karlsruhe.
- Lee, A. J., Carender, N. H., Knowlton, D. J., Bell, D. M. & Bryan, J. K. [1994], Fast autonomous registration of landsat, spot and digital map imagery, in 'SPIE Conference', pp. 68–79.
- Lenz, R. [1990], 'Group invariant pattern recognition', *Pattern Recognition* **23**, 199–217.
- Levine, M. D. & Shaheen, S. I. [1981], 'A modular computer vision system for picture segmentation and interpretation', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **3**, 540–556.
- Li, H., Manjunath, B. S. & Mitra, S. K. [1995], 'An approach to multisensor image registration', *IEEE Transactions on Image Processing* **4**, 320–334.
- Lillesand, T. M. & Kiefer, R. W. [1987], *Remote Sensing and Image Interpretation*, John Wiley and Sons, New York.
- Luhmann, T. [2000], *Nahbereichsphotogrammetrie*, Wichmann, Heidelberg.
- LVBW [2001], *Website des Landesvermessungsamtes Baden-Württemberg*, <http://www.lv-bw.de>.
- Mayer, H. [1998], *Automatische Objektextraktion aus digitalen Luftbildern*, Habilitationsschrift, Deutsche Geodätische Kommission, Reihe C, Nr. 494, München.
- Moravec, H. [1979], Visual mapping by a robot rover, in 'Proc. of 6th Joint Conference on Artificial Intelligence, Vol. 2', Tokyo.
- Morris, A. C., Stevens, A. & Muller, J. P. [1988], 'Ground control determination for registration of satellite imagery using digital map data', *Photogrammetric Record* **72**, 809–822.
- Nebiker, S. & Carosio, A. [1995], Automatic extraction and structuring of objects from scanned topographical maps - an alternative to the extraction from aerial and space images ?, in 'Automatic Extraction of Man-Made Objects from Aerial and Space Images', Birkhäuser Verlag, Basel, pp. 287–296.
- Newton, W., Gurney, D., Sloggett, D. & Dowman, I. [1994], An approach to automated identification of forests and forest change in remotely sensed images, International Archives of Photogrammetry and Remote Sensing, Vol. XXX, Part 3/2, München, pp. 607–614.
- Parker, J. R. [1997], *Algorithms for Image Processing and Computer Vision*, Wiley & Sons, Inc., New York.
- Pedersen, B. M. [1996], Automated measurement of ground control objects in large aerial photographs, International Archives of Photogrammetry and Remote Sensing, Vol. XXXI, Part B3, Wien, pp. 633–637.
- Pinz, A. [1994], *Bildverstehen*, Springer-Verlag, Wien.
- Pohl, C. & Genderen, J. L. [1998], 'Multisensor image fusion in remote sensing', *International Journal of Remote Sensing* **19**, 823–854.

- Pratt, P. [1991], *Digital Image Processing*, Wiley, New York.
- Pratt, W. C. [1974], 'Correlation techniques of image registration', *IEEE Transactions on Aerospace and Electronic Systems* **10**, 353–358.
- Reeves, A. P. [1988], 'Three dimensional shape analysis using moments and fourier descriptors', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **10**, 937–943.
- Reiss, T. H. [1991], 'The revised fundamental theorem of moment invariants', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**, 830–833.
- Reiss, T. H. [1993], *Recognizing Planar Objects Using Invariant Image Features*, Springer, Berlin, Heidelberg, New York.
- Ritter, N. [1997], 'The GeoTIFF data interchange standard for raster geographic images', *International Journal of Remote Sensing* **18**, 1637–1647.
- Rosenfeld, A. & Kak, A. [1976], *Digital Image Processing*, Academic Press, New York.
- Rosenholm, D. & Torlegaard, K. [1988], 'Three dimensional absolute orientation of stereo models using digital elevation models', *Photogrammetric Engineering & Remote Sensing* **54**, 1385–1398.
- Ruskoné, R. & Dowman, I. [1997], Segmentation design for an automatic multisource registration, in 'Integrating photogrammetric techniques with scene analysis and machine vision III.', SPIE 3072, pp. 307–317.
- Sabins, F. F. [1996], *Remote Sensing - Principles and Interpretation*, W.H. Freeman and Co., New York.
- Salmon, G. [1885], *Lessons introductory to the Modern Higher Algebra*, Hodges, Figgis & Co., Dublin.
- Schenk, T. [1993], Image understanding and digital photogrammetry, in 'Photogrammetric Week '93', Wichmann Verlag, Heidelberg, pp. 197–207.
- Schickler, W. [1992], Feature matching for outer orientation of single images using 3-D wireframe controlpoints, International Archives of Photogrammetry and Remote Sensing, Vol. XXIX, Part B3, Washington, pp. 591–598.
- Schowengerdt, R. A. [1997], *Remote Sensing - Models and methods for image processing*, Academic Press, San Diego.
- Sester, M. [1993], Modelling strategies in photogrammetry, in 'Photogrammetric Week '93', Wichmann Verlag, Heidelberg, pp. 21–32.
- Sester, M., Hild, H. & Fritsch, D. [1998], Definition of ground-control features for image registration using GIS-data, International of Photogrammetry and Remote Sensing, Vol. XXXII, Part 2, Columbus, Ohio, pp. 537–543.
- Shahin, F. & Novak, K. [1994], The exterior orientation of digital images by road matching, International Archives of Photogrammetry and Remote Sensing, Vol. XXX, Part 2, Ottawa, pp. 174–181.
- Sigle, M. & Heuchel, T. [2001], MATCH-AT: Recent developments and performance, in 'Photogrammetric Week '01', Wichmann Verlag, Heidelberg, pp. 189–194.
- Singer, M. H. [1993], 'A general approach to moment calculation for polygons and line segments', *Pattern Recognition* **26**, 1019–1028.
- Süße, H. [1999], 'Tutorium: Bildanalyse mittels Invarianten', *Mustererkennung 1999, 21. DAGM Symposium* pp. 1–93.
- Steinwendner, J. & Schneider, W. [1998], Algorithmic improvements in spatial subpixel analysis of remote sensing data, International Archives of Photogrammetry and Remote Sensing, Vol. XXXII, Part 3/1, Columbus, Ohio, pp. 265–271.
- Stengele, R. E. [1995], *Kartographische Mustererkennung - Rasterorientierte Verfahren zur Erfassung von Geo-Information*, Dissertation, Institut für Geodäsie und Photogrammetrie, ETH Zürich, Mitteilungen Nr. 54.
- Strunz, G. [1993], *Bildorientierung und Objektrekonstruktion mit Punkten, Linien und Flächen*, Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 408, München.
- Strunz, G., Knöpfle, W. & Roth, A. [1994], Automation of tie pointing procedure for the geocoding of satellite images, International Archives of Photogrammetry and Remote Sensing, Vol. XXX, Part 3/2, München, pp. 793–800.
- Sylvester, J. J. [1878], 'Proof of the hitherto undemonstrated fundamental theorem of invariants', *Philosophical Magazine* **5**, 178–188.
- Teh, C. H. & Chin, R. T. [1988], 'On image analysis by the method of moments', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **10**, 496–513.
- Tsingas, V. [1992], *Automatisierung der Punktübertragung in der Aerotriangulation durch mehrfache digitale Bildzuordnung*, Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 392, München.
- Udagawa, K., Toriwaki, J. & Sugino, K. [1964], 'Normalization and recognition of two-dimensional patterns with linear distortions by moments', *Electronic Communications Japan* **47**, 34–46.
- Vohra, V. & Dowman, I. [2000], Automatic registration of images to maps - the ARCHANGEL and ARMIES system, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B2, Amsterdam, pp. 590–597.
- Vosselmann, G. [1992], *Relational Matching*, Dissertation, Springer Verlag, Berlin.
- Wallin, A. & Kübler, O. [1995], 'Complete sets of complex zernike moment invariants and the role of pseudoinvariants', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**.

- Walter, V. [1997], *Zuordnung von raumbezogenen Daten - am Beispiel der Datenmodelle ATKIS und GDF*, Dissertation, Deutsche Geodätische Kommission, Reihe C, Nr. 480, München.
- Wang, Y. [1994], *Strukturzuordnung zur automatischen Oberflächenrekonstruktion*, Dissertation, Wissenschaftliche Arbeiten der Fachrichtung Vermessungstechnik der Universität Hannover, Nr. 207, Hannover.
- WEU [1995], Analysis of automated registration techniques - a study of registration of images to images and images to maps, ref. 36/199, Technical report, Madrid.
- Wrobel, B. [1987], Facets stereo vision (FAST vision) - a new approach to computer stereo vision and to digital photogrammetry, in 'Fast processing of photogrammetric data, Proceedings', Interlaken, pp. 230–258.
- Yang, Z. & Cohen, F. S. [1999], 'Cross-weighted moments and affine invariants for image registration and matching', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**, 804–814.

## Anhang A

## Geometrische Momente aus Objektkonturen

Zur Berechnung von geometrischen Momenten aus der Objektkontur ist der folgende Ausdruck zu lösen:

$$\begin{aligned}
M_{pq} &= \sum_{i=1}^3 \int_0^{S_i} \frac{1}{p+1} \sum_{n=0}^{p+1} \sum_{l=0}^q \binom{p+1}{n} \binom{q}{l} (a_i s)^{p+1-n} b_i^n (c_i s)^{q-l} d_i^l c_i ds \\
&= \sum_{i=1}^3 \frac{1}{p+1} \sum_{n=0}^{p+1} \sum_{l=0}^q \binom{p+1}{n} \binom{q}{l} \int_0^{S_i} a_i^{p+1-n} b_i^n c_i^{q+1-l} d_i^l s^{p+q-n-l+1} ds. \tag{A.1}
\end{aligned}$$

Die Koeffizienten  $a_i, b_i, c_i$  und  $d_i$  nehmen dabei folgende Gestalt an:

$$\begin{aligned}
a_1 &= \frac{X_1}{S_1} & b_1 &= 0 & c_1 &= \frac{Y_1}{S_1} & d_1 &= 0 \\
a_2 &= \frac{X_2 - X_1}{S_2} & b_2 &= X_1 & c_2 &= \frac{Y_2 - Y_1}{S_2} & d_2 &= Y_1 \\
a_3 &= -\frac{X_2}{S_3} & b_3 &= X_2 & c_3 &= -\frac{Y_2}{S_3} & d_3 &= Y_2. \tag{A.2}
\end{aligned}$$

Die Berechnung des gesamten Ausdrucks in Gleichung A.1 wird für jeden Summanden in  $i$  separat vorgenommen. Dabei werden jeweils die oben angegebenen Koeffizienten  $a_i, b_i, c_i$ , und  $d_i$  eingesetzt.

Auflösung für  $i = 1$ 

$$\begin{aligned}
&\int_0^{S_1} a_1^{p+1-n} b_1^n c_1^{q+1-l} d_1^l s^{p+q-n-l+1} ds \\
&= \begin{cases} 0 & \text{für } n \neq 0 \cup l \neq 0 \\ \int_0^{S_1} a_1^{p+1} c_1^{q+1} s^{p+q+1} ds & \text{für } n = 0 \cap l = 0 \end{cases} \\
&= \frac{1}{p+q+2} a_1^{p+1} c_1^{q+1} S_1^{p+q+2} ds \\
&= \frac{1}{p+q+2} \left( \frac{X_1}{S_1} \right)^{p+1} \left( \frac{Y_1}{S_1} \right)^{q+1} S_1^{p+q+2} \\
&= \frac{1}{p+q+2} X_1^{p+1} Y_1^{q+1} \tag{A.3}
\end{aligned}$$

Auflösung für  $i = 2$ 

$$\begin{aligned}
&\int_0^{S_2} a_2^{p+1-n} b_2^n c_2^{q+1-l} d_2^l s^{p+q-n-l+1} ds \\
&= \frac{1}{p+q-n-l+2} a_2^{p+1-n} b_2^n c_2^{q+1-l} d_2^l S_2^{p+q-n-l+2} ds \\
&= \frac{1}{p+q-n-l+2} \left( \frac{X_2 - X_1}{S_2} \right)^{p+1-n} X_1^n \left( \frac{Y_2 - Y_1}{S_2} \right)^{q+1-l} Y_1^l S_2^{p+q-n-l+2} \\
&= \frac{1}{p+q-n-l+2} X_1^n Y_1^l (X_2 - X_1)^{p+1-n} (Y_2 - Y_1)^{q+1-l} \tag{A.4}
\end{aligned}$$

Auflösung für  $i = 3$

$$\begin{aligned}
& \int_0^{S_3} a_3^{p+1-n} b_3^n c_3^{q+1-l} d_3^l s^{p+q-n-l+1} ds \\
&= \frac{1}{p+q-n-l+2} a_3^{p+1-n} b_3^n c_3^{q+1-l} d_3^l S_3^{p+q-n-l+2} \\
&= \frac{1}{p+q-n-l+2} (-1)^{p+1-n} \left(\frac{X_2}{S_3}\right)^{p+1-n} X_2^n (-1)^{q+1-l} \left(\frac{Y_2}{S_3}\right)^{q+1-l} Y_2^l S_3^{p+q-n-l+2} \\
&= \frac{1}{p+q-n-l+2} (-1)^{p+q-n-l} X_2^{p+1} Y_2^{q+1} \tag{A.5}
\end{aligned}$$

Damit ergibt sich durch Einsetzen der Gleichungen A.3, A.4 und A.5 in Gleichung A.1 insgesamt für das zu berechnende Integral aus Gleichung A.1:

$$\begin{aligned}
Mpq &= \frac{1}{p+1} \left\{ \frac{1}{p+q+2} X_1^{p+1} Y_1^{q+1} + \sum_{n=0}^{p+1} \sum_{l=0}^q \frac{1}{p+q-n-l+2} \binom{p+1}{n} \binom{q}{l} F_{pqnl}(X_1, X_2, Y_1, Y_2) \right\} \\
&\quad \text{mit} \\
F_{pqnl}(X_1, X_2, Y_1, Y_2) &= X_1^n Y_1^l (X_2 - X_1)^{p+1-n} (Y_2 - Y_1)^{q+1-l} + (-1)^{p+q-n-l} X_2^{p+1} Y_2^{q+1} . \tag{A.6}
\end{aligned}$$

## Anhang B

### Zu absoluten Invarianten aus Kreuzmomenten

In Kapitel 3 wird eine Schar absoluter Invarianten unter affiner Transformation aus Auto-Kreuzmomenten nullter Ordnung abgeleitet. Das Ergebnis lautet (Gleichung 3.37):

$$\Phi(s_0, s_1) = \frac{{}^{s_1+2}\sqrt{\mu_{1,(0,0,0,0),s_1}}}{{}^{s_0+2}\sqrt{\mu_{1,(0,0,0,0),s_0}}} . \quad (\text{B.1})$$

Abweichend davon kommen Yang & Cohen [1999] zu folgendem Ergebnis

$$\tilde{\Phi}(s_0, s_1) = \frac{{}^s\sqrt{\mu_{1,(0,0,0,0),s_1}}}{{}^s\sqrt{\mu_{1,(0,0,0,0),s_0}}} , \quad (\text{B.2})$$

d.h. die Radikanden sind jeweils um den Wert 2 kleiner. Die Ursache hierfür liegt in der Herleitung. Yang & Cohen [1999] leiten die absoluten Invarianten aus Kreuzmomenten in diskreten Koordinaten (Raster) ab. Dabei wird nach Ansicht des Autors fälschlicherweise die Rolle der Jacobi-Determinante vernachlässigt. Das schlägt sich darin nieder, daß der Ausdruck von Yang & Cohen [1999] nur absolut invariant unter affinen Transformationen mit Determinante 1 ist. Für alle anderen affinen Transformationen ist dies nicht unbedingt der Fall.

Die Einschränkung der Gültigkeit für beliebige affine Transformationen läßt sich leicht nachweisen für ganzzahlige  $s_0, s_1$  und isotrope Skalierung mit dem Maßstabsfaktor  $\alpha$ . Für ganzzahlige  $s_0, s_1$  lassen sich die Auto-Kreuzmomente aus zentralen geometrischen Momenten berechnen (Gleichung 3.39) und die absolute Invariantenschar lautet dann (Gleichung 3.40):

$$\Phi(s_0, s_1) = \frac{{}^{s_1+2}\sqrt{\sum_{g_1=0}^{s_1} (-1)^{s_1-g_1} \binom{s_1}{g_1} m_{g_1, s_1-g_1} m_{s_1-g_1, g_1}}}{{}^{s_0+2}\sqrt{\sum_{g_0=0}^{s_0} (-1)^{s_0-g_0} \binom{s_0}{g_0} m_{g_0, s_0-g_0} m_{s_0-g_0, g_0}}} \quad (\text{B.3})$$

bzw. nach Yang & Cohen [1999] :

$$\tilde{\Phi}(s_0, s_1) = \frac{{}^{s_1}\sqrt{\sum_{g_1=0}^{s_1} (-1)^{s_1-g_1} \binom{s_1}{g_1} m_{g_1, s_1-g_1} m_{s_1-g_1, g_1}}}{{}^{s_0}\sqrt{\sum_{g_0=0}^{s_0} (-1)^{s_0-g_0} \binom{s_0}{g_0} m_{g_0, s_0-g_0} m_{s_0-g_0, g_0}}} . \quad (\text{B.4})$$

Nach Hu [1962] gilt bei isotroper Skalierung mit dem Maßstabsfaktor  $\alpha$  für die zentralen geometrischen Momente:

$$m_{p,q}^\alpha = \alpha^{p+q+2} m_{p,q} . \quad (\text{B.5})$$

Setzt man diese Beziehung in Gleichung B.4 ein und separiert die Terme in  $\alpha$ , so erhält man für den Ausdruck nach Yang & Cohen [1999]:

$$\tilde{\Phi}^\alpha(s_0, s_1) = \frac{{}^{s_1}\sqrt{\alpha^{2(s_1+2)}}}{{}^{s_0}\sqrt{\alpha^{2(s_0+2)}}} \tilde{\Phi}(s_0, s_1) . \quad (\text{B.6})$$

Damit  $\tilde{\Phi}(s_0, s_1)$  eine absolute Invariante ist, muß die Beziehung

$$\frac{{}^{s_1}\sqrt{\alpha^{2(s_1+2)}}}{{}^{s_0}\sqrt{\alpha^{2(s_0+2)}}} = \alpha^{\frac{2(s_1+2)}{s_1} - \frac{2(s_0+2)}{s_0}} = \alpha^{4\left(\frac{1}{s_1} - \frac{1}{s_0}\right)} \equiv 1 \quad (\text{B.7})$$

erfüllt sein. Dies ist jedoch ganz offensichtlich für beliebige  $s_0, s_1$  mit  $s_1 \neq s_0$  nur dann möglich, wenn  $\alpha \equiv 1$  ist. Diese Einschränkung besteht jedoch nicht für eine generelle affine Transformation, q.e.d.

Im Falle des in Kapitel 3 abgeleiteten Ergebnisses aus Gleichung B.3 erhält man bei gleichem Vorgehen:

$$\Phi^\alpha(s_0, s_1) = \frac{{}^{s_1+2}\sqrt{\alpha^{2(s_1+2)}}}{{}^{s_0+2}\sqrt{\alpha^{2(s_0+2)}}} \Phi(s_0, s_1) \quad (\text{B.8})$$

und damit für die Forderung nach absoluter Invarianz von  $\Phi(s_0, s_1)$  :

$$\frac{{}^{s_1+2}\sqrt{\alpha^{2(s_1+2)}}}{{}^{s_0+2}\sqrt{\alpha^{2(s_0+2)}}} = \alpha^{\frac{2(s_1+2)}{s_1+2} - \frac{2(s_0+2)}{s_0+2}} = \alpha^0 \equiv 1 , \quad (\text{B.9})$$

was für beliebige  $\alpha$  erfüllt ist (der Fall  $\alpha = 0$  stellt als Singularität keine invertierbare Transformation dar).

## Anhang C

## Einzelerggebnisse der Korrelationsuntersuchungen

Im folgenden sind die Einzelergbnisse der Untersuchungen zur Diskriminierbarkeit mit CWM-Invarianten aus Abschnitt 3.3.2 dargestellt. Eine Zusammenfassung dieser Ergebnisse findet sich in Abbildung 3.14 im Hauptteil. In der jeweils linken Abbildung sind grauwertcodiert die einzelnen Korrelationen abgetragen (symmetrische Matrix). In der jeweils rechten Abbildung sind die Paarungen mit minimaler Merkmalsraum-Distanz durch schwarze Quadrate gekennzeichnet. Abweichungen von der Hauptdiagonalen bedeuten, daß die höchste Korrelation zwischen unterschiedlichen Buchstaben besteht. Bei Betrachtung der Spalten von links nach rechts liefert das erste schwarze Matrixelement das Objekt des Originaldatensatzes, das am besten mit dem der Zeile entsprechenden Element des diskretisierten Satzes korreliert (diskretisiert-original). Bei Betrachtung der Spalten von oben nach unten liefert das erste schwarze Matrixelement das Element des diskretisierten Datensatzes, das am besten mit dem der Spalte entsprechenden Element des Originaldatensatzes korreliert (original-diskretisiert).

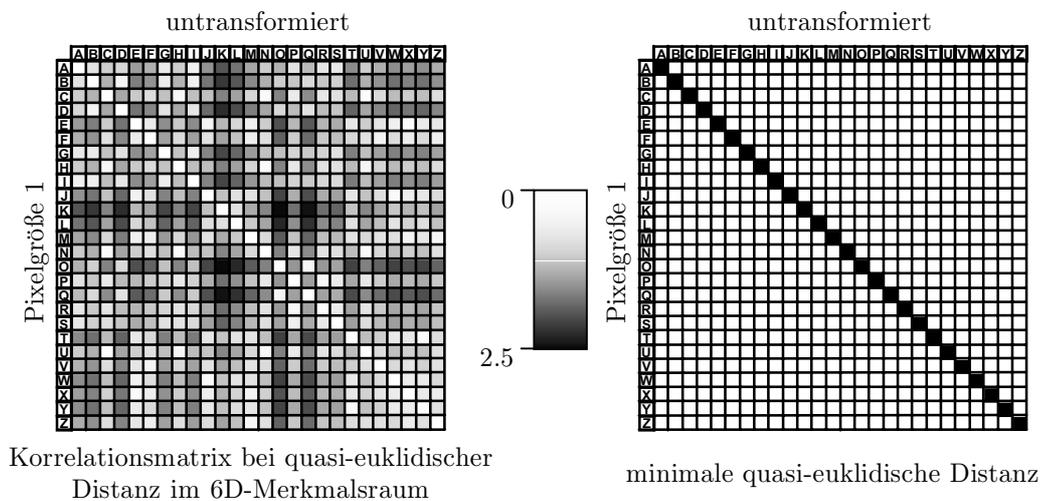


Abbildung C.1: Korrelationen für Generalisierung der Stufe 1.

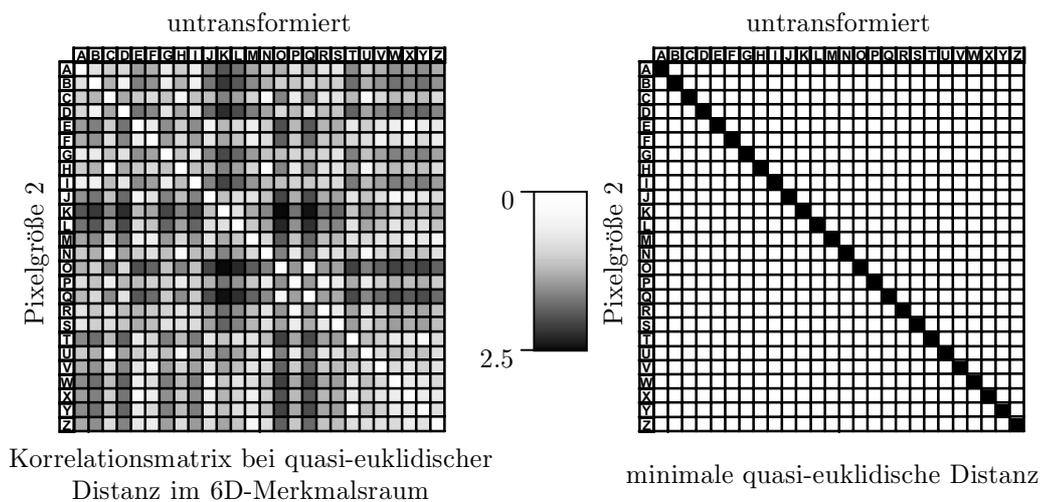


Abbildung C.2: Korrelationen für Generalisierung der Stufe 2.

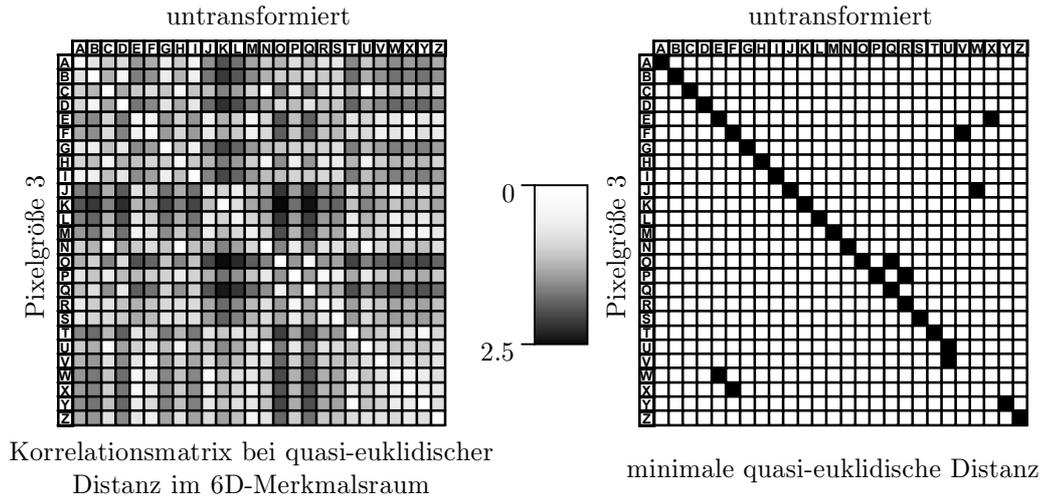


Abbildung C.3: Korrelationen für Generalisierung der Stufe 3.

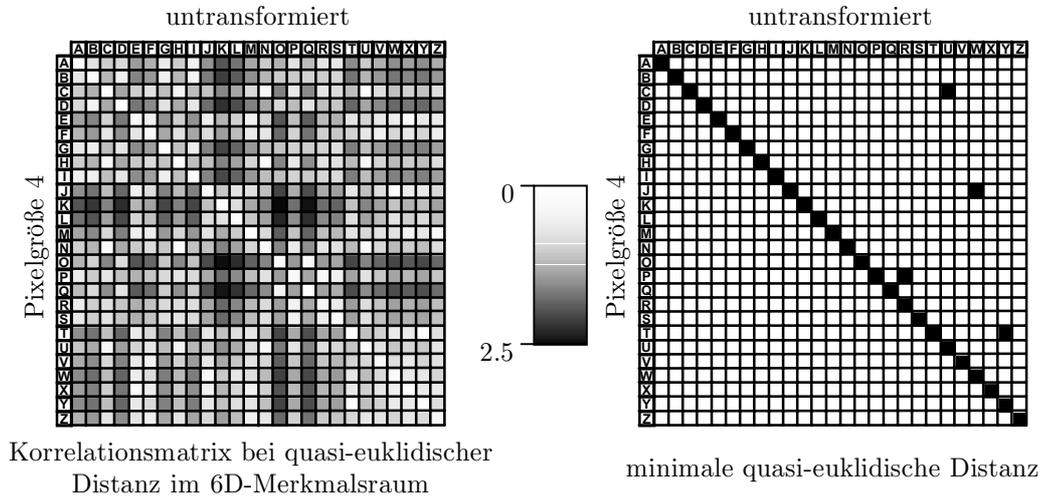


Abbildung C.4: Korrelationen für Generalisierung der Stufe 4.

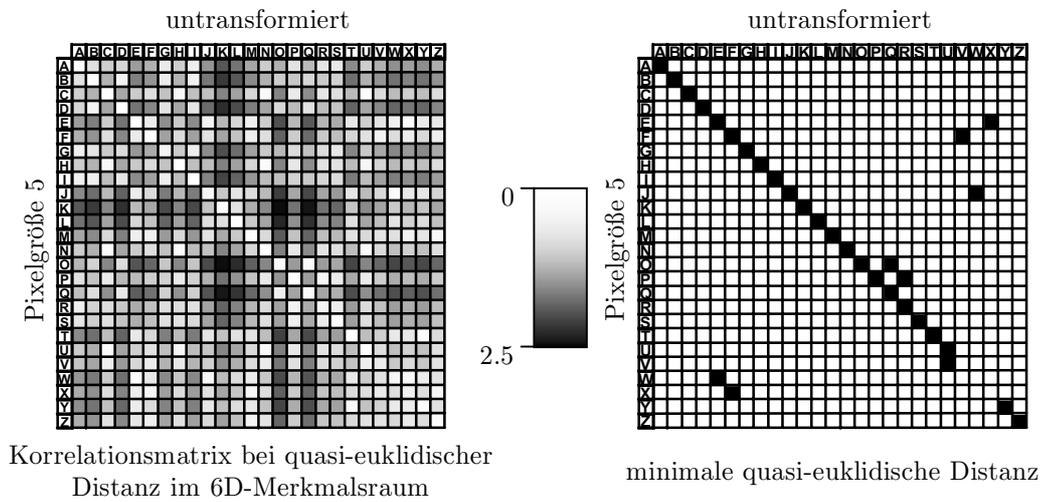


Abbildung C.5: Korrelationen für Generalisierung der Stufe 5.

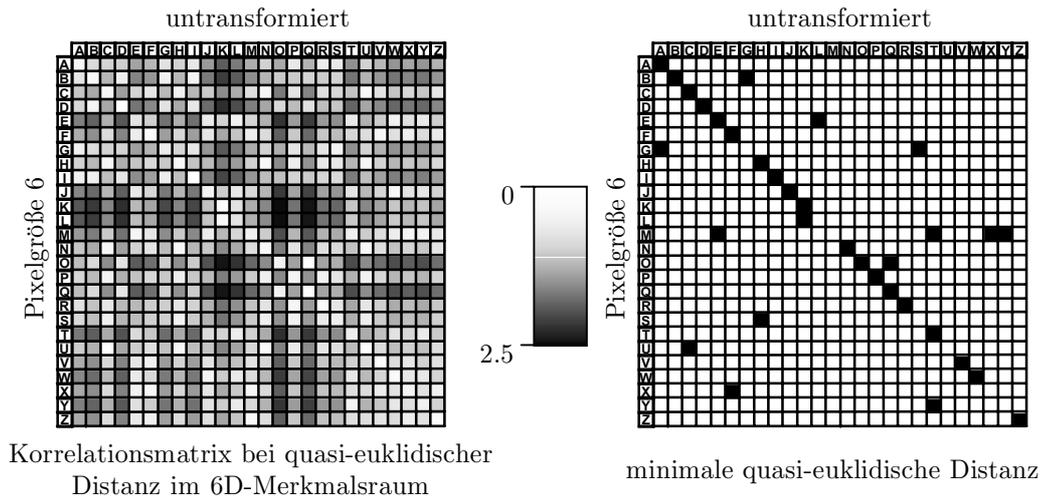


Abbildung C.6: Korrelationen für Generalisierung der Stufe 6.

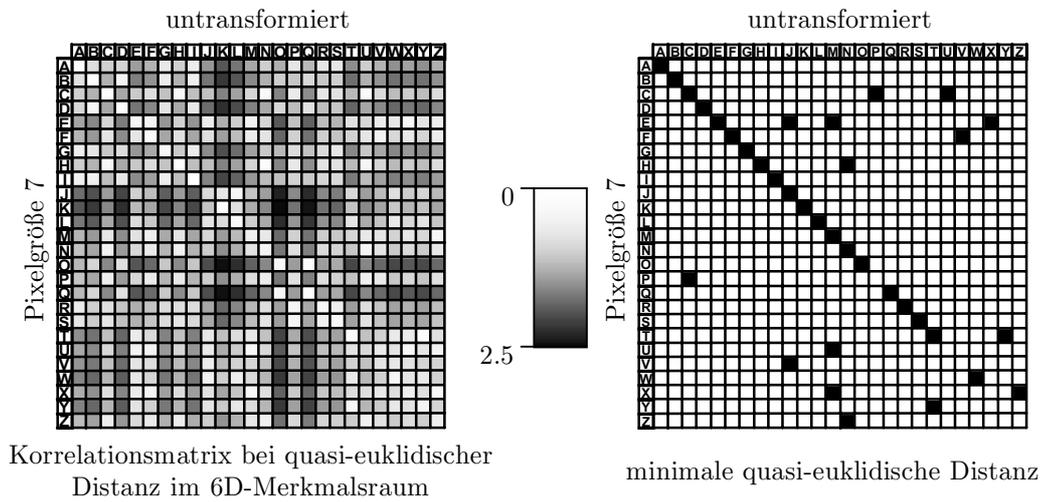


Abbildung C.7: Korrelationen für Generalisierung der Stufe 7.

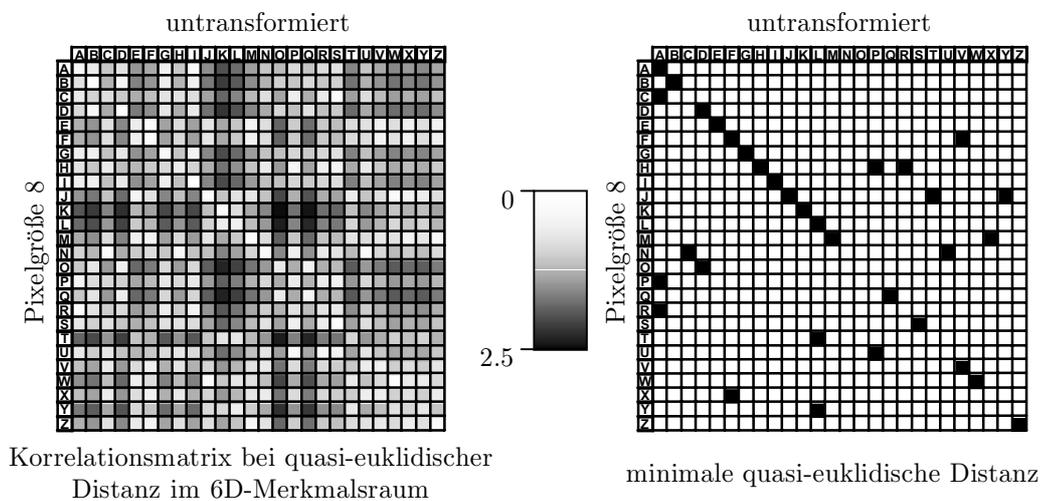


Abbildung C.8: Korrelationen für Generalisierung der Stufe 8.

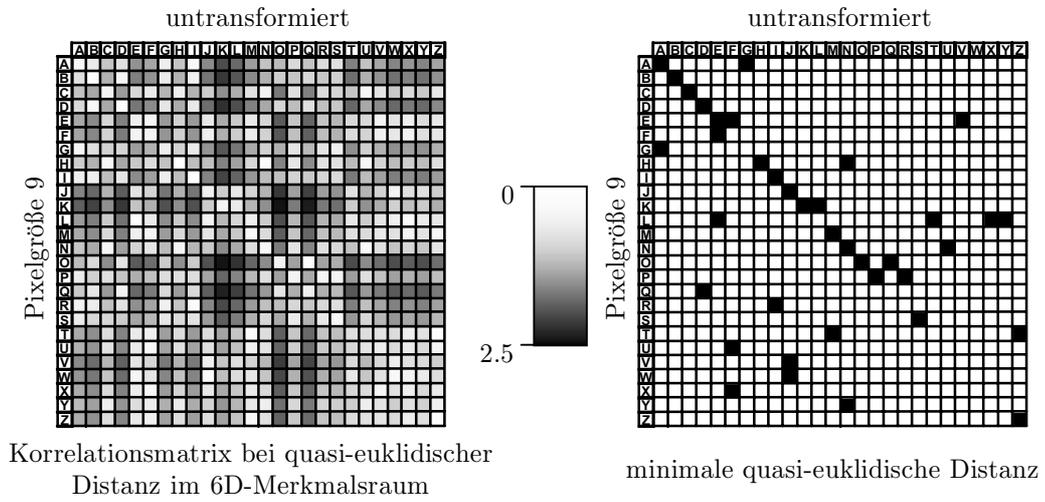


Abbildung C.9: Korrelationen für Generalisierung der Stufe 9.

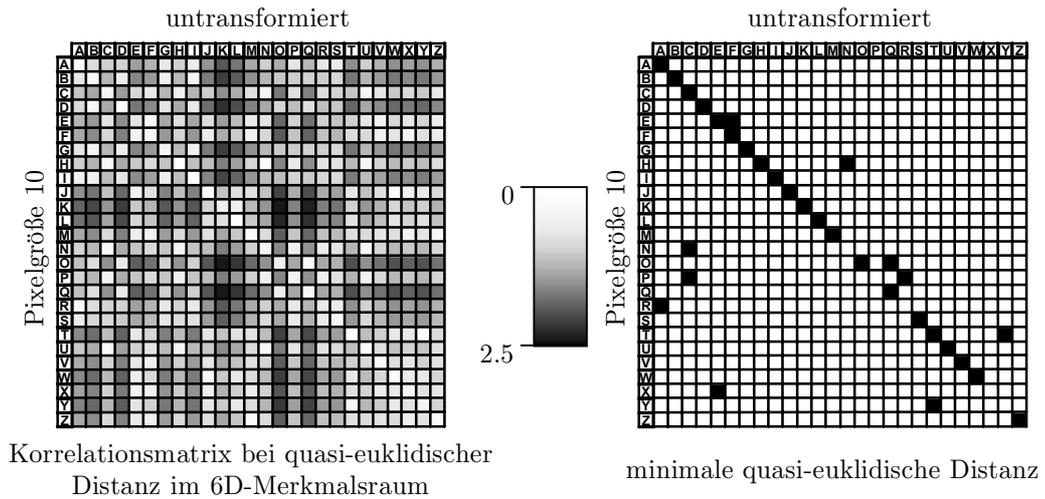


Abbildung C.10: Korrelationen für Generalisierung der Stufe 10.

Anhang D

Rotation und Spiegelung ungerader Momente

Die folgende Tabelle stellt alle 64 möglichen Fälle dar, zwei zentrale Momente ungerader Ordnung ( $m_{gu}$  und  $m_{ug}$ ) durch Rotationen um ganze Vielfache von  $90^\circ$  und Spiegelungen bezüglich der  $y$ -Achse zu transformieren. Hier stehen die Indices "g" für "positiv definite gerade ganze Zahl" und "u" für "positive ungerade ganze Zahl". Das doppelte Vorzeichen kennzeichnet dabei das Moment mit dem größeren Betrag. Dies Ergebnisse werden in der vorliegenden Arbeit dazu benutzt, die zur Erreichung einer Standardlage notwendige Transformation geometrischer Objekte mittels der im Text beschriebenen Iterationsmethode zu ermitteln. Die schraffierten Felder repräsentieren beispielhaft die Standardlage mit  $m_{gu} \geq 0, m_{ug} \geq 0$  und  $m_{gu} \geq m_{ug}$ , die so auch implementiert wurde.

$m_{gu}$	$m_{ug}$	Drehung	$m'_{gu}$	$m'_{ug}$	Y-Spieg.	$m''_{gu}$	$m''_{ug}$	$m_{gu}$	$m_{ug}$	Drehung	$m'_{gu}$	$m'_{ug}$	Y-Spieg.	$m''_{gu}$	$m''_{ug}$						
++	+	0	++	+	N	++	+	--	+	0	--	+	N	--	+						
					J	++	-						J	--	-						
		90	+	--	N	+	--			--	+	90	+	++	N	+	++				
					J	+	++								J	+	--				
		180	--	-	N	--	-					--	+	180	++	-	N	++	-		
					J	--	+										J	++	+		
		270	-	++	N	-	++							--	+	270	-	--	N	-	--
					J	-	--												J	-	++
+	++	0	+	++	N	+	++	-	++							0	-	++	N	-	++
					J	+	--												J	-	--
		90	++	-	N	++	-			-	++					90	++	+	N	++	+
					J	++	+												J	++	-
		180	-	--	N	-	--					-	++			180	+	--	N	+	--
					J	-	++												J	+	++
		270	--	+	N	--	+							-	++	270	--	-	N	--	-
					J	--	-												J	--	+
++	-	0	++	-	N	++	-	--	-							0	--	-	N	--	-
					J	++	+												J	--	+
		90	-	--	N	-	--			--	-					90	-	++	N	-	++
					J	-	++												J	-	--
		180	--	+	N	--	+					--	-			180	++	+	N	++	+
					J	--	-												J	++	-
		270	+	++	N	+	++							--	-	270	+	--	N	+	--
					J	+	--												J	+	++
+	--	0	+	--	N	+	--	-	--							0	-	--	N	-	--
					J	+	++												J	-	++
		90	--	-	N	--	-			-	--					90	--	+	N	--	+
					J	--	+												J	--	-
		180	-	++	N	-	++					-	--			180	+	++	N	+	++
					J	-	--												J	+	--
		270	++	+	N	++	+							-	--	270	++	-	N	++	-
					J	++	-												J	++	+

Abbildung D.1: Mögliche Symmetrieoperationen für ungerade zentrale Momente  $m_{gu}$  und  $m_{ug}$ .

## Anhang E

## Experimente zur Standardlage nach der Iterationsmethode

Hier sind alle Ergebnisse der in Abschnitt 4.5.4 beschriebenen Experimente zur Verwendbarkeit der Transformation in eine Standardlage mittels der Iterationsmethode (siehe Abschnitt 4.5.3) abgebildet.



Abbildung E.1: Erstes Experiment: affine Kopie ohne statistisches Rauschen.

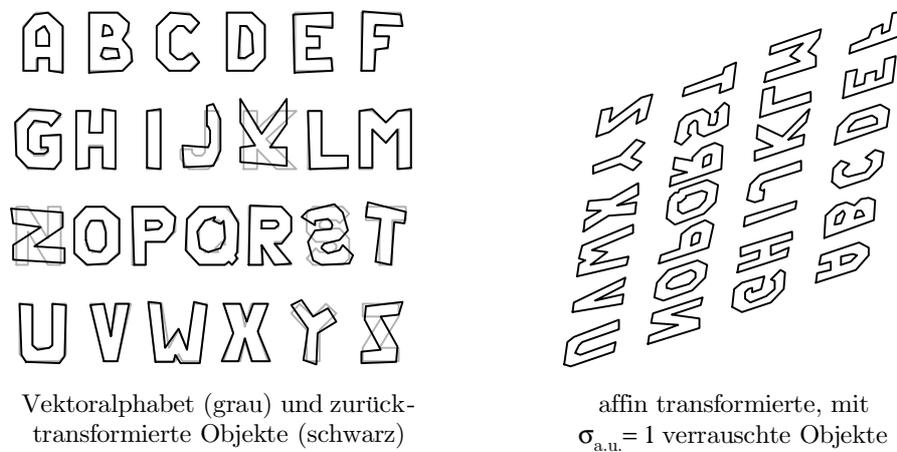
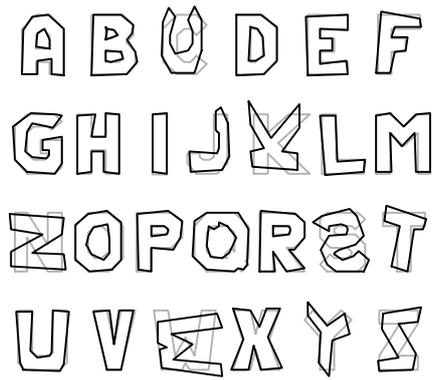


Abbildung E.2: Zweites Experiment: affine Kopie mit statistischem Rauschen.  $\sigma_{a.u.} = 1$ .

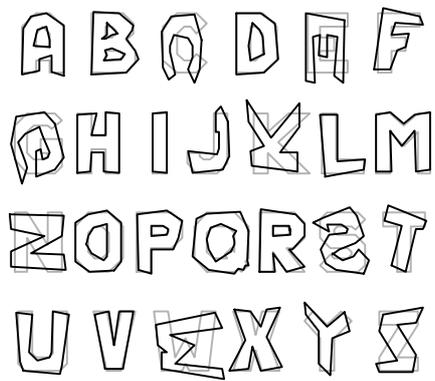


Vektoralphabet (grau) und zurücktransformierte Objekte (schwarz)



affin transformierte, mit  $\sigma_{a.u.} = 2$  verrauschte Objekte

**Abbildung E.3:** Drittes Experiment: affine Kopie mit statistischem Rauschen  $\sigma_{a.u.} = 2$ .



Vektoralphabet (grau) und zurücktransformierte Objekte (schwarz)



affin transformierte, mit  $\sigma_{a.u.} = 3$  verrauschte Objekte

**Abbildung E.4:** Viertes Experiment: affine Kopie mit statistischem Rauschen  $\sigma_{a.u.} = 3$ .

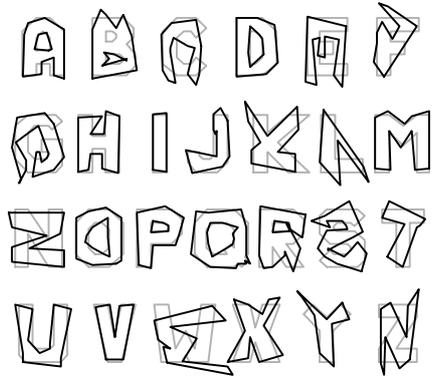


Vektoralphabet (grau) und zurücktransformierte Objekte (schwarz)



affin transformierte, mit  $\sigma_{a.u.} = 4$  verrauschte Objekte

**Abbildung E.5:** Fünftes Experiment: affine Kopie mit statistischem Rauschen  $\sigma_{a.u.} = 4$ .



Vektoralphabet (grau) und zurücktransformierte Objekte (schwarz)



affin transformierte, mit  $\sigma_{a.u.} = 5$  verrauschte Objekte

Abbildung E.6: Sechstes Experiment: affine Kopie mit statistischem Rauschen  $\sigma_{a.u.} = 5$ .

$\sigma$ in a.u.	A	B	C	D	E	F	G	H	I	J	K	L	M
0	j	j	j	j	j	j	j	n	j	j	j	j	j
1	j	j	j	j	j	j	j	j	j	n	n	j	j
2	j	j	n	j	j	j	j	j	j	n	n	j	j
3	j	j	n	j	n	j	n	j	j	n	n	j	j
4	j	j	n	j	n	n	n	j	j	n	n	n	j
5	j	n	n	j	n	n	n	j	j	n	n	n	j
$\sigma$ in a.u.	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	n	n	j	j	j	n	j	j	j	j	j	j	n
1	n	j	j	j	j	n	j	j	j	j	j	j	n
2	n	j	j	j	j	n	j	j	j	n	j	j	n
3	n	j	j	j	j	n	j	j	j	n	j	j	n
4	n	j	j	j	j	n	j	j	j	n	j	j	n
5	n	j	j	j	j	n	j	j	j	n	j	j	n

Tabelle E.1: Akzeptable und nicht akzeptable Rücktransformationen unterschiedlich stark verrauschter Buchstaben.

## Dank

Mein Dank gilt Herrn Prof. Dr.-Ing. habil. Dieter Fritsch, für die Iniziiierung der Fragestellung dieser Arbeit, die Übernahme des Hautberichts und für das angenehme und offene Arbeitsklima am ifp.

Ebenso bedanke ich mich bei meinen beiden Mitberichtern, Herrn Prof. Dr.-Ing. habil. Dr. tech. h.c. mult. Dr.-Ing. Erik Grafarend vom Geodätischen Institut der Universität Stuttgart und Herrn Prof. Dr.-Ing. habil. Hans-Peter Bähr vom Institut für Photogrammetrie und Fernerkundung der Universität Karlsruhe, für die freundliche Übernahme des Korreferats.

Meinen Zimmerkollegen Norbert Haala, Michael Kiefner, Christian Stätter und Babak Ameri danke ich herzlich für das freundschaftliche Verhältnis und die vielen angeregten, fruchtbaren und manchmal auch einfach nur entspannenden Gespräche.

Bei allen anderen Kollegen, die mit mir am ifp beschäftigt sind oder waren bedanke ich mich herzlich für die angenehme und entspannte Atmosphäre und den kollegialen Umgang. Die Zeit am ifp wird mir in sehr positiver Erinnerung bleiben.

Mein größter Dank gilt jedoch meiner lieben Frau Katja, die mich, wo immer es ging, bei meiner Arbeit unterstützt hat. Durch ihre Rücksicht und den Verzicht auf die eine oder andere Stunde gemeinsam mit unseren Söhnen Max und Lukas verbrachter Freizeit hat sie wesentlichen Anteil am Gelingen dieser Arbeit.

Ein letzter Dank gilt meinen Eltern für die großzügige Unterstützung während meines Studiums.

## Lebenslauf

Heiner Hild

- |                 |   |
|-----------------|---|
| 20. Juli 1967   | geboren in Marbach / Neckar   |
| 1973 - 1977     | Grundschule in Poppenweiler   |
| 1977 - 1986     | Goethe-Gymnasium in Ludwigsburg   |
| 13. Juni 1986   | Abitur  |
| 1986 - 1987     | Grundwehrdienst   |
| 1988 - 1989     | Studium Maschinenbau/Produktionstechnik FHT Esslingen/Neckar                      |
| 1989 - 1996     | Studium der Physik (Diplom) an der Universität Stuttgart                          |
| 6. Februar 1996 | Diplom in Physik  |
| 1996            | Wiss. Mitarbeiter am Uniklinikum Freiburg i.Br., Abteilung Röntgendiagnostik      |
| 1996 - 2002     | Wiss. Mitarbeiter am Institut für Photogrammetrie (ifp) der Universität Stuttgart |
| seit 2002       | Mitarbeiter der Robert Bosch GmbH   |