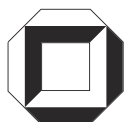


Astrid Laubenheimer

Automatische Registrierung adaptiver Modelle zur Typerkennung technischer Objekte



Dissertation, Universität Karlsruhe (TH), Fakultät für Elektrotechnik und
Informationstechnik, 2004

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe

www.uvka.de

© Universitätsverlag Karlsruhe 2004
Print on Demand
Druck: Digital Print Group Erlangen

ISBN 3-937300-14-7

Automatische Registrierung adaptiver Modelle zur Typerkennung technischer Objekte

Zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

von der Fakultät für Elektrotechnik und Informationstechnik

der Universität Karlsruhe (TH)

genehmigte

Dissertation

von

Dipl.-Math. Astrid Laubenheimer

aus Karlsruhe

Karlsruhe, im August 2004

Tag der mündlichen Prüfung:

18. Mai 2004

Hauptreferent:

Prof. Dr.-Ing. K. Kroschel

Korreferent:

Prof. Dr.-Ing. R. Dillmann

Kurzfassung

Anwendungen der Bildauswertung werden in zunehmendem Maße unter Verwendung dreidimensionaler Modelle realisiert und fusionieren auf diese Weise Methoden der Computergrafik und der Bildauswertung. In den letzten Jahren ist zudem ein vermehrter Einsatz adaptiver Modelle zu beobachten, d. h. solcher Modelle, die nicht nur ein einziges Objekt, sondern eine ganze Klasse von ähnlichen Objekten darstellen können. In der vorliegenden Arbeit wird die Flexibilität solcher adaptiver Modelle eingesetzt, um unvollständige Datenbasen und geometrisch ungenaue Systemeingaben zu kompensieren.

Hierfür wird ein neu entwickeltes Verfahren zur automatischen Konstruktion adaptiver Modelle vorgestellt. Das Besondere an diesem Verfahren ist, dass es auch für die Modellierung von technischen Objekten geeignet ist. Insbesondere werden scharfe Objektkanten und dünne Objektteile, die für technische Objekte charakteristisch sind, gut modelliert. Das Konstruktionsverfahren kombiniert Techniken der volumen- und oberflächenbasierten Methoden der *computational geometry*, um so zu einer für die Bildauswertung geeigneten Objektdarstellung zu gelangen. Ferner werden Module zur Anpassung dieser adaptiven Modelle an Grauwertbilder vorgestellt, die durch Synthese-Analyse-Iterationen die Brücke zur Bildanalyse schlagen. Diese Module sind auf Basis von extrahierten Bildkanten realisiert, wodurch eine Unabhängigkeit von der Oberflächenbeschaffenheit der modellierten Objekte erreicht wird.

Die zentrale Stärke dieser Komponenten liegt darin, dass sie aus Einzelbildern dreidimensionale Rekonstruktionen für unbekannte Objektvarianten liefern. Wie experimentell gezeigt wird, sind diese Rekonstruktionen geometrisch genauer als handelsübliche Modelle. Insbesondere können die auf dem Verfahren aufbauenden Systeme im laufenden Betrieb um neu auftretende Objektvarianten erweitert werden, und die manuelle Nachführung von dreidimensionalen Modellen für noch nicht modellierte Objektvarianten entfällt.

Die Leistungsfähigkeit der entwickelten Verfahren wird am Beispiel der Flugzeugtypisierung gezeigt. Darüber hinaus wird die Anwendbarkeit der Module im Bereich der Lageschätzung und der Objektrekonstruktion demonstriert.

Danksagung

Die Anfertigung einer Doktorarbeit ist nicht möglich ohne die wohlwollende Unterstützung vieler Personen im beruflichen und privaten Umfeld, sei es bewusst oder unbewusst. Diesen Personen ist diese Seite gewidmet.

Mein erster Dank gilt Herrn Prof. Dr. K. Kroschel, der sich am Beginn meiner Tätigkeit am Fraunhofer IITB als wissenschaftliche Angestellte und Doktorandin spontan bereit erklärt hat, meine Arbeit zu betreuen. Ebenso danke ich Herrn Prof. Dr. R. Dillmann für die Übernahme des Korreferats.

Herrn Prof. Dr. Norbert Link danke ich dafür, dass er mir durch sein Angebot, in seiner damaligen Abteilung zu promovieren, diese Arbeit überhaupt erst ermöglicht hat. Seine Gabe, mir zum richtigen Zeitpunkt die richtigen drei Sätze zu sagen, hat mir über mindestens eine Durststrecke hinweggeholfen. In ähnlicher Weise hat sich meine Mentorin, Frau Dr. Bärbel Hüsing, um diese Arbeit verdient gemacht. Meinen Vorgesetzten, Dr. Josef Pauli und Markus Müller, danke ich für die ungetrübte Freude, mich während meiner Arbeitszeit um meine Doktorarbeit kümmern zu können.

Meinen Kollegen danke ich für die Bereitschaft, meine vielen Fragen zu beantworten. Namentlich danken möchte ich Norbert Heinze, Dr. Thomas Müller und Dr. Nils Rehfeld für ihre Diskussionsbereitschaft sowie Heiko Widak und Martin Esswein für ihre technische Unterstützung. Dr. Hermann Fehrenbach verdient ein Dankeschön für die Erläuterung der Motivationskurve während der Doktorandenzeit: Die „Badewannenkurve“ war mir bis dato kein Begriff.

Mein ganz besonderer Dank gilt meinen Eltern Irmgard und Jürgen Laubenheimer. Ihnen bin ich inzwischen einiges schuldig, vor allem für ihre moralische Unterstützung, als ich nach der Geburt meines Sohnes Zweifel hatte, ob und wie eine Mutter promovieren kann. Abgesehen davon haben sich meine Eltern durch viele Stunden des Babysittings und Tausende von Autobahnkilometern um diese Arbeit verdient gemacht.

Ohne meinen Lebensgefährten, Holger Weindel, wäre diese Arbeit ebenfalls nicht zustande gekommen. Neben vielen anderen familiären Aufgaben hat er in den letzten Monaten die Kochlöffel geschwungen und so dafür gesorgt, dass unsere Familie mindestens einmal am Tag ein anständiges Essen auf den Tisch bekam. Ich danke ihm ganz besonders dafür, dass er nie in Frage gestellt hat, dass meine Arbeit, mit all ihren Auswirkungen auf unser Privatleben, mein Weg ist. An letzter und doch an erster Stelle der Danksagung steht mein Sohn Sören Weindel. Er kann noch nicht verstehen, warum man tagelang in Instituten oder Arbeitszimmern verschwinden muss, um eine Doktorarbeit zu schreiben. Dennoch, er hat in letzter Zeit mehr Geduld mit mir gehabt als ich mit ihm. Dafür gilt auch ihm ein großes Dankeschön.

Inhaltsverzeichnis

Symbolverzeichnis	v
Abbildungsverzeichnis	vii
1 Einleitung	1
1.1 Luftbildauswertung	1
1.2 Die Aufgabenstellung	3
1.3 Frühere Arbeiten	4
1.4 Alternative Ansätze	5
1.4.1 Modellbasierte Ansätze	5
1.4.2 Trainingsbasierte Ansätze	8
1.4.3 Deformierbare Modelle	9
1.4.4 Zusammenfassung	11
1.5 Der Lösungsansatz: Verfahrensübersicht	11
1.5.1 Das Typerkennungssystem	11
1.5.2 Die Modellkonstruktion	13
1.5.3 Die Modellanpassungsmodule	13
1.5.4 Die Merkmalsextraktion und Typerkennung	15
1.6 Kapitelübersicht	15
2 Grundlagen und Vorbereitungen	17
2.1 Darstellung dreidimensionaler Objekte	17
2.1.1 Oberflächenbasierte Darstellung	18
2.1.2 Volumenbasierte Darstellung	19
2.1.2.1 Radiale Summen	20
2.1.2.2 Signierte Distanztransformation	22
2.2 Polygonisierung und Netzoptimierung	27

2.2.1	Polygonisierung	27
2.2.2	Netzoptimierung nach Ohtake & Belyaev	27
2.3	Polygonreduktion nach Garland & Heckbert	31
3	Das adaptive Modell	33
3.1	Definition des adaptiven Modells	33
3.2	Konstruktion des adaptiven Modells	36
3.2.1	Konstruktion des Meta-Modells	41
3.2.2	Hinzufügen der Basismodelle	47
3.2.3	Polygonreduktion des adaptiven Modells	48
3.3	Erweiterte adaptive Modelle	50
3.4	Zusammenfassung und Ergebnisse	54
3.4.1	Flugzeugmodelle	55
3.4.2	Weitere Objektklassen	56
4	Modellanpassung	59
4.1	Die Abbildungskette	59
4.2	Robuste Modellanpassung	60
4.2.1	Merkmalsextraktion	60
4.2.2	Die Gütefunktion	62
4.2.3	Lokale Extremwerte und Mehrdeutigkeiten	65
4.2.4	Grob-zu-Fein Strategie	67
4.2.5	Die Optimierung der Gütefunktion	67
4.2.6	Laufzeitoptimierungen	70
4.3	Der Passpunktschätzer	72
4.4	Iterative Optimierung	74
4.5	Zusammenfassung	77
5	Anwendungen	81
5.1	Flugzeugerkennung	81
5.1.1	Die Kollateraldaten	81
5.1.2	Die Bilddaten	83
5.1.3	Das adaptive Modell und die Repräsentanten	84
5.1.4	Merkmalsextraktion, Datenbasis und Typerkennung	87
5.1.5	Die Modellanpassung	90

5.1.6	Ergebnisse der Modellanpassung	95
5.1.7	Die Typerkennung	98
5.1.8	Zusammenfassung der Flugzeuigerkennung	102
5.2	Lageschätzung	103
5.2.1	Die Versuchsdaten	104
5.2.2	Testbeschreibungen	105
5.2.3	Die Leistungsmessung	106
5.2.4	Die Ergebnisse	107
5.2.5	Zusammenfassung der Lageschätzung	110
6	Zusammenfassung und Ausblick	115
6.1	Zusammenfassung	115
6.2	Ausblick	116
A	Die Hauptkomponentenanalyse	119
B	Die Abbildungskette	121
C	Bilddatenübersicht	123
	Literaturverzeichnis	125

Symbolverzeichnis

Symbol	Bedeutung und Seitennummer
\mathcal{A}	Objekt, Menge
$\partial\mathcal{A}$	Rand von \mathcal{A}
\mathcal{A}^0	Inneres von \mathcal{A}
α	Winkelschwelle für die Dreiecksunterteilung 46
α_n	Parameter des adaptiven Modells 33
C	Kovarianzmatrix 34
c	Isowert einer impliziten Funktion 20
c_0	Radius der morphologischen Öffnung 44
c_x, c_y	Bildhauptpunkt 59
χ^2	Gütefunktion des robusten Schätzers 65
$\chi^2(\sigma_L, \hat{\theta})$	Vertrauenswert der Modellanpassung 67
$d(\mathbf{x}, \mathcal{A})$	Abstand von \mathbf{x} zu \mathcal{A} 22
d_{ij}	Orthogonaler Abstand Modell- und Bildkantensegment 63
d_M	Mahalanobis-Distanz 90
d_x, d_y	Bildpunktgröße 59
Δ_i	Kontraktionsfehler der Polygonreduktion 31
$\Delta_{i,n}$	Kontraktionsfehler für das Basismodell \mathcal{M}_n 49
Δ'_i	Kontraktionsfehler für das adaptive Modell 50
$\Delta\mathbf{g}_n$	Zentrierte geometrische Realisierung 34
δ_n	Parameter des orthogonalen adaptiven Modells 34
\mathbf{e}_n	Eigenvektor der Kovarianzmatrix C 34
f	Implizite Funktion 19
\mathbf{f}	Brennweite 59
∇f	Gradient der Funktion f
f_n	Implizite Funktion für das Basismodell \mathcal{M}_n 41
f_n^w	Implizite Funktion für das verzerrte Basismodell \mathcal{M}_n^w 43
θ	Parametervektor der Modellanpassung 60
$\tilde{\theta}$	Parametervektor des orthogonalen adaptiven Modells 34
\mathbf{g}	Geometrische Realisierung 18
$\tilde{\mathbf{g}}$	Geometrische Realisierung des Meta-Modells 37
$\bar{\mathbf{g}}$	Mittlere geometrische Realisierung 33

\mathbf{g}_n	Geometrische Realisierung eines Basismodells 33
\mathcal{H}	Menge der Bildkantensegmente 61
K	Anzahl Iterationen für die Korrespondenzberechnung 47
κ	Gierwinkel 59
\mathcal{L}_θ	Menge der Modellkantensegmente 61
l	Mindestlänge der Bildkantensegmente 61
λ_n	Eigenwert der Kovarianzmatrix C 34
LGS	Lineares Gleichungssystem
$\mathcal{M}, \mathcal{M} = (\mathcal{T}, \mathbf{g})$	Modell in Oberflächendarstellung 19
$\tilde{\mathcal{M}}$	Meta-Modell 46
\mathcal{M}_n	Basismodell 33
\mathcal{M}_n^w	Verzerrtes Basismodell 43
μ	Gradientenbetragsschwelle für die Bildkantensegmente 61
\mathcal{O}	Operatorkette 47
\mathbf{o}	Nullvektor
o_{ij}	Überlappung von Modell- und Bildkantensegment 63
ω	Rollwinkel 59
\mathcal{P} -Operator	Duale Netzoptimierung 28
$\mathcal{P}_{perspective}$	Projektionsmatrix perspektivisch 121
\mathcal{P}_{ortho}	Projektionsmatrix orthogonal 122
φ	Neigungswinkel 59
Q_i	Fehlerfunktion für die Polygonreduktion 31
$Q_{i,n}$	Fehlerfunktion für das Basismodell \mathcal{M}_n 49
q	Vertrauenswert der Klassifikation 90
\mathcal{R} -Operator	Regularisierungsoperator 28
ρ	Filterfunktion 63
$\bar{\rho}$	<i>Tukey's biweight</i> 63
s	Streckungslänge 50
\mathcal{S} -Operator	Netzunterteilung 28
σ	Skalierung der Filterfunktion 64
σ'	Normierte Skalierung der Filterfunktion 64
σ_n^2	Eigenwert der Kovarianzmatrix C 35
SDT	Signierte Distanztransformation 22
\mathcal{T}	Netztopologie 19
\mathcal{T}_0	Meta-Topologie 33
$(t_x, t_y, t_z)^T$	Translationsvektor 59
τ	Winkelschwelle für die Modellkantensegmente 62
w_n	Basismodellspezifische Verzerrung 43
\hat{w}_n	Inverse basismodellspezifische Verzerrung 43
\mathcal{Z} -Operator	Projektion auf die Isofläche 28
$\ \cdot\ $	Euklidische Norm

Abbildungsverzeichnis

1.1	Varianten der Boeing 737	3
1.2	Beispielbilder vom Frankfurter Flughafen	5
1.3	Anpassung eines handelsüblichen 3D-Modells	7
1.4	Anwendungsübersicht	12
2.1	Radiale Summen zur Objektdarstellung	20
2.2	Die signierte Distanztransformation	23
2.3	Minimaler Punktabstand zu Dreiecken	25
2.4	Einflussbereiche von Kanten und Dreiecken	26
2.5	Beispiel für das duale Netz	29
2.6	Netzoptimierung	30
3.1	Ein adaptives Modell	34
3.2	Hauptkomponentenanalyse für adaptive Modelle	36
3.3	Diskrete Merkmale in adaptiven Modellen	38
3.4	Ankerpunkte für Flugzeuge	42
3.5	Fehler der invertierten Verzerrung	44
3.6	Konstruktion des Meta-Modells	45
3.7	Dreiecksreduktion	46
3.8	Hinzufügen der Basismodelle	49
3.9	Streckung rigider Modelle	51
3.10	Ergebnisse der Modellkonstruktion	55
3.11	Ein adaptives Fahrzeugmodell	56
4.1	Orthogonaler Abstand und Überlappung	63
4.2	Graph der Filterfunktion	64
4.3	Ein Beispiel für die Filterfunktion	66

4.4	Lokale Extremwerte der Gütefunktion: Lage	66
4.5	Lokale Extremwerte der Gütefunktion: Schatten und Form	66
4.6	Die Gütefunktion	68
4.7	Gradienten der Kostenfunktion	70
4.8	Schablonenanpassung auf dem Gradientenbetragsbild	75
4.9	Iterative Optimierung	77
4.10	Modellrekonstruktion	78
4.11	Modellanpassung für ein Werkstück	78
4.12	Werkzeug zur Einstellung der Verfahrensparameter	80
5.1	Korrelation zwischen Rollwinkel und Form	82
5.2	Übersicht über die Luftaufnahmen	85
5.3	Perspektivische Verzerrungen	85
5.4	Diagramm der <i>ground truth</i>	86
5.5	Modellausprägung an Schattenrisse	88
5.6	Merkmale aus Segmenten	89
5.7	Initiale Parameterschätzung: Übersicht	91
5.8	Initiale Parameterschätzung: Beispiele	92
5.9	Modelldiskretisierung	94
5.10	Beispielsergebnisse für die Modellanpassung	97
5.11	Rekonstruktion unbekannter Flugzeuge	101
5.12	Falsch-Identifikationen	101
5.13	Die Hantel mit Modell	104
5.14	Hantelbilder für den Konvergenztest	105
5.15	Vertrauenswerte für einen Hantelversuch	107
5.16	Vergleich der Lageschätzungen: Standardabweichungen	111
5.17	Vergleich der Lageschätzungen: Winkel	112
5.18	Vergleich der Lageschätzungen: Position	113

Kapitel 1

Einleitung

Die maschinelle Bildauswertung hat sich während der letzten Jahre zu einer Basistechnologie mit den unterschiedlichsten Anwendungsgebieten entwickelt. So unterschiedlich wie die Anwendungsgebiete, so unterschiedlich sind auch die Anforderungen, die an das jeweilige System gestellt werden. In einigen Bereichen, wie z. B. der Qualitätskontrolle, der Fertigungssteuerung oder dem Bereich der autonomen Roboter, stehen die Echtzeitfähigkeit und eine vollständige Automatisierung der bildverarbeitenden Prozesse an erster Stelle. In anderen Gebieten, z. B. der Medizin oder im Überwachungsbereich, werden andere Schwerpunkte gesetzt. In diesem Zusammenhang denke man beispielsweise an eine Krebszellendetektion in Röntgenbildern, die auch im Zusammenhang mit bildverarbeitenden Systemen Gegenstand wissenschaftlicher Untersuchungen geworden ist [Klausmann '00]. Diesen Systemen wird man beim derzeitigen Stand der Technik keine Diagnosestellung zugestehen. Man wird aber, mit dem Ziel den Menschen in seinen Entscheidungen zu unterstützen, von ihnen die Detektion von Indikatoren einer Erkrankung erwarten. In [McInerney & Terzopoulos '96] wird die automatische Interpretation medizinischer Bilder im Rahmen einer Diskussion „Autonomie versus Kontrolle“ gar als Langzeit-Ziel bezeichnet. Kurz, die kognitiven Fähigkeiten des Menschen werden von Bildauswertungssystemen nicht erreicht und werden den Menschen bei sicherheitsrelevanten Entscheidungen in naher Zukunft nicht ersetzen.

1.1 Luftbildauswertung

Völlig analog zur Auswertung medizinischer Daten ist der Bereich der Luftbildauswertung einzuordnen. Auch hier besteht die Aufgabe darin, aus gegebenen Bilddaten relevante Bildbereiche zu extrahieren und die dort abgebildeten Objekte entsprechend vorgegebener Kriterien einzuordnen, d. h. zu klassifizieren. Diese Aufgabe wird bisher ausschließlich von so genannten Bildauswertern durchgeführt. Sie detektieren, erkennen, analysieren und melden relevante Objekte aus dem jeweiligen Bildmaterial. Dieses Bildmaterial unterscheidet sich gegenüber vielen anderen Anwendungen der Bildauswertung bereits durch seine Größe. Mitunter sind Bildauswerter mit Material

in Größenordnungen von 12.000×100.000 Bildpunkten¹ konfrontiert, die einen Bodenbereich von ca. 16 Quadratkilometern abdecken können. Berücksichtigt man dann noch die Sicherheitsrelevanz und den in der Regel strengen zeitlichen Rahmen der Aufgabe, so zeigt sich das Anforderungsprofil an den Auswerter in etwa vergleichbar mit der Arbeitssituation eines Fluglotsen. In beiden Fällen verbietet jedoch die Sicherheitsrelevanz bei dem derzeitigen Stand der Technik die vollständige Automatisierung. Nun stellt sich die Frage, inwieweit und in welchem Rahmen die Maschine den Menschen unterstützen kann. Wie kann also eine geeignete Aufgabenteilung zwischen Mensch und Maschine gefunden werden, so dass optimale Ergebnisse erreicht werden? Dieser Frage ist das Fraunhofer IITB bereits vor einigen Jahren nachgegangen und hat infolge dieser Untersuchungen die Entwicklung eines Experimentalsystems SAFIR anvisiert, eines maschinellen Assistenten für Auswerter zur Integration leistungsfähiger interaktiver und teilautomatischer Komponenten.² Der Einsatz einiger solcher Komponenten zur Auswerterunterstützung ist bereits in Reichweite. So existieren z. B. inzwischen robuste und zuverlässige *screeener*, die durch Signaturanalyse Bildbereiche detektieren, in denen mit hoher Wahrscheinlichkeit Objekte von Interesse abgebildet sind [Berger et al. '00].

Bleibt die Frage nach der Erkennung des jeweiligen Objekttyps. An dieser Stelle sei bemerkt, dass im Rahmen der Luftbildauswertung zwischen individuell auftretenden Objekten, also konkreten Gebäuden, Industrieanlagen, etc. und wiederholt auftretenden Objekten aus einer Serienfertigung unterschieden werden muss. Zu den wiederholt auftretenden Objekten sind die meisten mobilen Objekte zu zählen, von denen im Folgenden die Klasse der Flugzeuge fokussiert wird. Die übrigen Klassen der mobilen Objekte werden dabei jedoch nicht außer Acht gelassen. Eine Differenzierung zwischen militärischen und zivilen Flugzeugen ist zunächst von minderer Bedeutung, da diverse zivile Flugzeugtypen zu Aufklärungs- und Transportzwecken im militärischen Bereich eingesetzt werden, also gleichermaßen von Bedeutung sind.

Allgemein ist in Bezug auf die Typerkennung bzw. Identifikation technischer Objekte zunächst Folgendes festzustellen: Die menschliche Erkennungsfähigkeit konzentriert sich sehr stark auf diskrete, verbal benennbare Merkmale. Dies ist letztlich ein Zeichen für die hohe semantische Ebene, auf der der Mensch agiert und die das maschinelle Sehen nicht erreicht. Dem gegenüber werden kontinuierliche Merkmale, wie z. B. relative und vor allem absolute Größen vom Menschen vergleichsweise schlecht erfasst. Das folgende Beispiel belegt diese Aussage. Abbildung 1.1 zeigt links eine Aufnahme einer Boeing 737, die ein geübter Beobachter unter Verwendung diskreter Merkmale, wie z. B. der Anzahl der Triebwerke, des Kreuzleitwerks und groben Angaben über die Größenordnung erkennen kann. Die Frage nach der vorliegenden Variante lässt sich jedoch nur mit Kenntnis relativer und absoluter Größenordnungen ermitteln - immerhin stehen neun Varianten zur Auswahl (Abbildung 1.1 rechts). Spätestens dann, wenn die Aufnahme perspektivischen oder anderen optischen Verzerrungen unterwor-

¹Wollte man Bilder dieser Größe 1:1 auf Bildschirmen darstellen, so würden bei einer Bildschirmauflösung von 1280×1024 Bildpunkten ca. 10×100 Bildschirme benötigt.

²Projektbeschreibungen für SAFIR sind beispielsweise unter <http://vision.fraunhofer.de> und <http://www.iitb.fraunhofer.de> veröffentlicht.

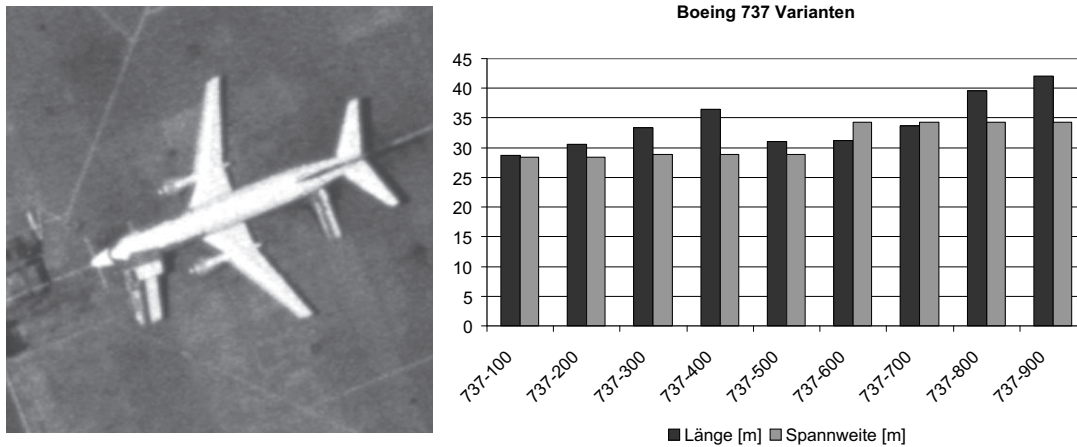


Abbildung 1.1: Die linke Aufnahme zeigt eine Boeing 737-300 auf dem Frankfurter Flughafen. Die Grafik rechts verdeutlicht die Vielzahl der Varianten, die von diesem Flugzeugtyp bekannt sind (Quelle technische Angaben: [Figgen et al. '00]).

fen ist, die bei der Vermessung berücksichtigt werden müssen, gerät der Mensch an seine Grenzen. Genau hier besteht Bedarf an maschineller Unterstützung durch Extraktion kontinuierlicher Merkmale durch eine Bildauswertungs-Komponente. Hierdurch ist das Anwendungsziel der hier vorliegenden Arbeit definiert.

1.2 Die Aufgabenstellung

Das Ziel dieser Arbeit ist die Entwicklung eines Verfahrens zur automatischen Extraktion kontinuierlicher Merkmale aus optischen Grauwertbildern im Hinblick auf eine Typerkennung technischer Objekte. Das Verfahren soll in Form eines Moduls zur Unterstützung von Bildauswertern am Beispiel von Verkehrsflugzeugen demonstriert werden. Ausgehend von einem Aufmerksamkeitsbereich, also einer *region of interest*, sollen dem Anwender - unter Angabe eines Vertrauenswertes - Vorschläge für den vorliegenden Flugzeugtyp gemacht werden. Insbesondere ist die Bereitstellung einer Datenbank mit trennungswirksamen Merkmalen notwendig. Eine Visualisierung der Ergebnisse und die Bereitstellung der extrahierten Größen ist in jedem Fall als vorteilhaft einzuordnen und daher gewünscht.

Durch diese Anwendung werden die Anforderungen an das Verfahren definiert, die sich in zwei Gruppen gliedern. Die erste Anforderungsgruppe bezieht sich auf die Objektklassen, auf die das Verfahren anwendbar sein muss:

- **Allgemeingültigkeit bezüglich der Objektklassen:** Das Verfahren soll zunächst für die Klasse der Verkehrsflugzeuge entwickelt werden. Die Übertragbarkeit auf andere Klassen mobiler Objekte muss jedoch gewährleistet sein und soll mit möglichst geringem Aufwand durchführbar sein.

- **Invarianz gegenüber den Materialeigenschaften:** Die Extraktion der Merkmale muss unabhängig von den Materialeigenschaften, also von der Objektfarbe, den Reflektionseigenschaften etc. sein, da diese auch innerhalb einer Klasse stark variieren können.

Die zweite Gruppe von Anforderungen folgt aus der Tatsache, dass das Verfahren auf Außenaufnahmen anwendbar sein muss:

- **Bildhintergrund:** Der Hintergrund, also der nicht zum Objekt gehörende Bildbereich, kann nicht als homogener Bildbereich angenommen werden. Vielmehr muss das Verfahren robust gegenüber systematischen Störungen sein, die durch andere Objekte, Gebäude oder Verkehrswege verursacht werden. Auch die durch den Untergrund entstehenden Texturen müssen toleriert werden, da dieser nicht als bekannt vorausgesetzt werden kann.
- **Beleuchtungsinvarianz:** Für Außenaufnahmen muss mit unbekanntem Beleuchtungssituationen gerechnet werden. Das Verfahren muss also Lichtschwankungen und lichtbedingte Artefakte wie Eigenbeschattung oder Schlagschatten tolerieren.

Die Tatsache, dass es sich bei dieser Anwendung um die Auswertung von Luftbildaufnahmen handelt, ist bei der Verfahrensentwicklung sekundär. Wie vielleicht angenommen werden könnte, handelt es sich bei Luftbildaufnahmen keinesfalls ausschließlich um Senkrechtsichten, wodurch die Verfahrensentwicklung auf eine einzige Blickrichtung eingeschränkt werden könnte. Vielmehr wird in Abbildung 1.2 (c) gezeigt, dass selbst bei senkrecht nach unten gerichteten Sensoren durch projektive Verzerrungen mehr oder weniger stark ausgeprägte Seitenansichten der Objekte erzeugt werden. Aus diesem Grund wird die Tatsache, dass es sich bei der Anwendung um Luft- und nicht um Bodenaufnahmen handelt, nicht gesondert berücksichtigt.

1.3 Frühere Arbeiten

Nun stellt sich die Frage, welche Ansätze geeignet sind, aus Luftbildern trennungswirksame Merkmale zu extrahieren, also solche Größen bereitzustellen, die die Unterscheidung der in Frage kommenden Flugzeugtypen ermöglicht. Im Vorfeld der hier vorliegenden Arbeit wurden am Fraunhofer IITB Untersuchungen zum Thema Flugzeugerkennung durchgeführt, die in dem Projektbericht [Berger et al. '00] publiziert sind. Das dort vorgestellte Verfahren basiert auf einer stochastischen, texturbasierten Segmentierung, also auf einem Verfahren, das die zum Flugzeug gehörenden Bildpunkte markiert. Aus diesen Segmenten werden unter Verwendung der Kollateraldaten, wie z. B. der Flughöhe und der vorliegenden Bildauflösung, die für die Identifikation erforderlichen Merkmale extrahiert. Abbildung 1.2 zeigt exemplarisch drei Bildausschnitte, die Gegenstand dieser Untersuchungen sind.

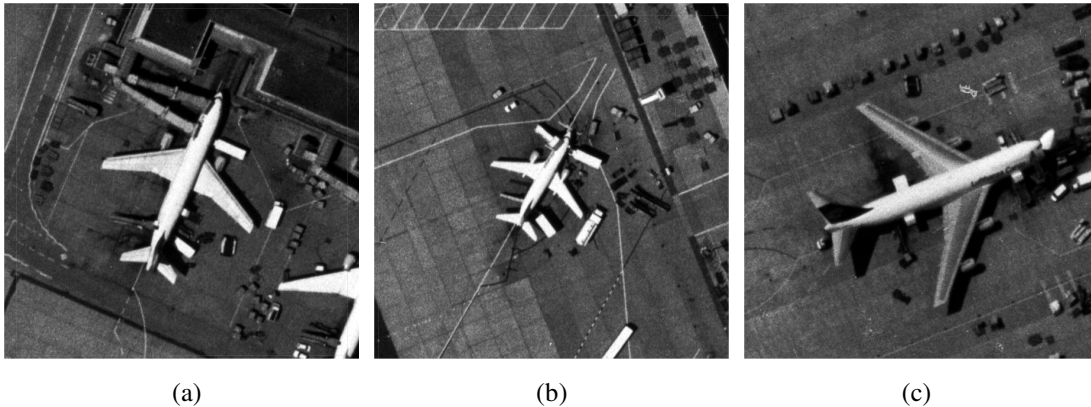


Abbildung 1.2: Beispielbildausschnitte vom Frankfurter Flughafen. Die Seitenansicht in Bild (c) wird durch perspektivische Verzerrung im Randbereich des Bildes verursacht.

Allerdings haben diese Untersuchungen gezeigt, dass die texturbasierte Segmentierung kaum ohne Verwendung von weiterem Szenenwissen zu lösen ist. Diese Aussage wird an den folgenden Fakten manifestiert:

- In der Nähe der Flugzeuge befinden sich häufig andere Objekte, wie z. B. Fahrzeuge und Gangways, die ähnliche Strukturen aufweisen wie das Flugzeug selbst.
- Die Bemalungsschemata der Flugzeuge führen dazu, dass die Flugzeuge nicht als homogene Strukturen im Bild erfasst werden können.
- Eigenbeschattung zergliedert den Flugzeugrumpf.
- Geringer Kontrast zwischen Flugzeug und Hintergrund erschwert die Segmentierung.

Auch bei einer gelungenen Segmentierung muss bei der Extraktion der Merkmale berücksichtigt werden, dass durch optische Verzerrungen keine Senkrechtansichten vorliegen (vgl. Abbildung 1.2 (c)). Deshalb muss ein segmentbasierter 2D-Ansatz in Frage gestellt werden.

1.4 Alternative Ansätze

1.4.1 Modellbasierte Ansätze

Eine für die Aufgabenstellung denkbare Alternative zur rein bildbasierten Segmentierung ist die Verwendung von modellbasiertem Szenenwissen. Die Anforderung nach Unabhängigkeit von der Blickrichtung des bildgebenden Sensors legt die Verwendung

dreidimensionaler Modelle nahe.³ Diese Überlegung verhält sich konform zu einem aktuellen Trend, nämlich Anwendungen im Bildverarbeitungsbereich als modellbasierte Sichtsysteme zu realisieren. Ein gutes Beispiel für die praktische Anwendbarkeit solcher 3D-modellbasierter Ansätze ist das am Fraunhofer IITB entwickelte Flugzeug-Andocksystem, das - mit dem Joseph-von-Fraunhofer-Preis 2001 ausgezeichnet - von Honeywell Airport Systems vertrieben wird und bereits an zahlreichen Flughäfen im Einsatz ist [Gengenbach et al. '98]. In diesem System ist eine so genannte Modellanpassung realisiert, d. h. ein Verfahren, das mit Hilfe eines gegebenen Modells und einem kalibrierten Bild die relative Orientierung zwischen Objekt und Kamera berechnet, also letztlich die Position des Flugzeugs auf dem Rollfeld bestimmt.

Andererseits stellt sich die Frage, wie Modelle zum Einsatz kommen können, wenn der Typ des beobachteten Objekts nicht bekannt ist. Eine nahe liegende Möglichkeit ist die Durchführung einer solchen Modellanpassung für alle in Frage kommenden Modelle und eine Entscheidung zugunsten desjenigen Modells bzw. Objekts, das am Besten zum Bild passt, für das also beispielsweise die Modellanpassung den geringsten Restfehler geliefert hat. Ein solches Vorgehen wird jedoch mit zunehmender Zahl an in Frage kommender Objekte ineffizient, da der Rechenaufwand direkt proportional zur Anzahl der zu testenden Objekte steht. Dieser Thematik haben sich in den letzten Jahren einige wissenschaftliche Untersuchungen gewidmet. Kölzow beispielsweise hat in [Kölzow '03] eine 3D-modellbasierte Klassifikation, Lokalisation und Verfolgung von Objekten am Beispiel von Flugzeugen auf dem Rollfeld vorgestellt und eigens zur Reduktion des Rechenaufwands ein Hypothesenmanagement entwickelt. In der modellbasierten Fahrzeugdetektion von [Schlosser et al. '03] wird hingegen vorgeschlagen, eine baumähnliche Hierarchie von Modellen zu konstruieren, mit einem 3D-Würfel an der Wurzel.

Modellakquisition

Eine wichtige Frage, die man sich bei jeder 3D-modellbasierten Anwendung, also unabhängig von der speziellen Methode stellen muss, ist die Frage nach der Bereitstellung der benötigten Modelle. Diese Frage gewinnt vor allem außerhalb kontrollierbarer Bereiche an Bedeutung: Während etwa in der Fertigungssteuerung oder der Qualitätskontrolle in der Regel vorausgesetzt werden kann, dass die Anzahl der relevanten Objekte klein ist und dass für diese Objekte schon von Seiten der Konstruktion ein entsprechendes CAD-Modell vorliegt, ist dies in anderen Anwendungsbereichen nicht unbedingt der Fall. So ist auch in der hier anvisierten Anwendung keinesfalls davon auszugehen, dass für alle beobachteten Objekte CAD-Modelle zur Verfügung gestellt werden können. Erschwerend kommt hinzu, dass die in der Luftbildauswertung interessanten Objekte Ausmaße besitzen, die eine Rekonstruktion mit handelsüblichen Modellerzeugern, wie z. B. von Streifenprojektoren, Volumenschnittverfahren und auch Lasersystemen, unmöglich machen.⁴

³Im Folgenden wird der Begriff „Modell“, sofern nicht ausdrücklich anderweitig erwähnt, immer mit einer dreidimensionalen Objektdarstellung assoziiert.

⁴Eine ausführliche Beschreibung solcher Modellerzeuger findet man in [Fautz '03].

[Kölzow '03] löst die Bereitstellung von Modellen, indem er für seine Versuche drei aus dem Internet frei verfügbare Modelle verwendet. In der Tat sind Modelle dieser Art für sehr viele technische Objekte verfügbar. Allerdings muss man sich bei diesen Modellen, die üblicherweise in der Bilderzeugung, also dem Bereich der Computergrafik zum Einsatz kommen, unbedingt die Frage nach der geometrischen Genauigkeit stellen. Dies gilt vor allem dann, wenn sich die zu unterscheidenden Objekte geometrisch sehr ähnlich sind. Abbildung 1.3 beispielsweise zeigt ein handelsübliches Modell, angepasst an einen digitalisierten Schattenriss und eine Luftaufnahme. Die Wiederholung der Nichtübereinstimmung von Modell und Bild im Bereich der Tragflächen und des Leitwerks auf beiden Bildern lässt nur einen Schluss zu, nämlich den einer Ungenauigkeit des Modells von mindestens drei Metern.⁵



Abbildung 1.3: Die linke Aufnahme zeigt das Modell einer Boeing 747-400, das mittlere Bild die Anpassung dieses Modells an einen digitalisierten Schattenriss (Quelle Schattenriss: [Green '92]) und das rechte Bild eine Modellanpassung an eine Szene des Frankfurter Flughafens.

In Verbindung mit dem im vorangegangenen Abschnitt gezeigten Diagramm der Varianten der Boeing 737 (Abbildung 1.1) ist klar, dass geometrische Fehler in dieser Größenordnung zumindest sehr nahe an der Grenze liegen, die für die Unterscheidung zweier Varianten notwendig ist. Darüber hinaus muss bei der Verwendung solcher Modelle sichergestellt werden, dass die Modellanpassung derart ungenaue Eingaben toleriert. Diese Toleranz ist jedoch immer gegenläufig zu der Genauigkeit, mit der ein entsprechendes Verfahren trennungswirksame Merkmale extrahieren kann.

Die Strategie, handelsübliche Modelle zu verwenden, wird weiterhin durch die Tatsache erschwert, dass Anbieter solcher Modelle in der Regel keine Angaben über die Variante machen, die durch das Modell repräsentiert wird. Zudem wird in der Regel nicht mehr als ein Modell pro Flugzeugtyp angeboten. Alles in allem muss festgehalten werden, dass der Modellakquisition in der Anwendung modellbasierter Verfahren erhebliche Bedeutung zukommt.

⁵Die maximale Abweichung auf dem Schattenriss beträgt 22 Bildpunkte, wobei die Position der Triebwerke vernachlässigt wurde. Bei einer Auflösung von circa 0,15 Metern pro Bildpunkt beträgt die Abweichung damit 3,3 Meter.

1.4.2 Trainingsbasierte Ansätze

Um einen Vergleich zu alternativen Methoden ziehen zu können, darf eine Gruppe von Objekterkennungsansätzen nicht unerwähnt bleiben: Die Gruppe der erscheinungsbasierten Methoden (in der englischsprachigen Literatur *appearance-based recognition* genannt) kodiert individuelle Objekte und Objektansichten in Form von Merkmalsvektoren, wobei die Basis für den Merkmalsraum durch eine statistische Analyse einer Trainingsbildmenge gewonnen wird.

Die klassische Vorgehensweise für erscheinungsbasierte Verfahren ist die direkte Anwendung einer Hauptkomponentenanalyse auf die Grau- oder Farbwerte der Trainingsbilder. Dieses Verfahren ist vielfach im Bereich der Gesichtserkennung anzutreffen und führte dort bereits zu einer eigenen Begriffsbildung für die Basisvektoren, die in diesem Zusammenhang als *eigenfaces* bezeichnet werden [Turk & Pentland '91]. Dieses Verfahren setzt jedoch eine Segmentierung des Eingabebildes voraus, wodurch die Anwendung auf einfach segmentierbare Objekte eingeschränkt ist [Campbell & Flynn '01]. Dem ist hinzuzufügen, dass die direkte Anwendung einer Hauptkomponentenanalyse auf den Bildbereich voraussetzt, dass sich die Bilder mehr oder weniger in Korrespondenz befinden. Insbesondere müssen identische Objektmerkmale (wie z. B. Augen oder Mundwinkel) auf nahezu denselben Bildpositionen abgebildet sein. In [Vasilescu & Terzopoulos '02] werden so genannte *tensor-faces* eingeführt, die nach Angabe der Autoren zu besseren Erkennungsraten führen.

Eine weitere Gruppe bilden diejenigen Verfahren, die nicht auf den Grau- oder Farbwerten der Bilder aufsetzen, sondern auf extrahierten Bildmerkmalen operieren, wie z. B. auf Bildkantensegmenten oder auf daraus aggregierten Bildmerkmalen.⁶ Diese Verfahren arbeiten nicht erscheinungsbasiert sondern formbasiert. In der Arbeit [Pope & Lowe '96] wird geschildert, wie solche Bildmerkmale - aus Beispielbildern unterschiedlicher Ansichten gewonnen und in Graphen eingeordnet - Objekte repräsentieren können. Die Erkennung erfolgt dann über die Extraktion der entsprechenden Bildmerkmale im Eingabebild und einer stochastischen Suche ähnlicher Bildmerkmale im Graphen. Der Suchaufwand wird bereits im Vorfeld reduziert, indem Ansichten mit ähnlichen Eigenschaften gruppiert und durch einen Repräsentanten ersetzt werden. Ein verwandter Ansatz ist das Verfahren von [Selinger & Nelson '00], in dem ebenfalls Bildmerkmale zur Objekterkennung eingesetzt werden.

Der entscheidende Punkt ist jedoch, dass sowohl bei den erscheinungsbasierten als auch bei den formbasierten Verfahren das Lernen im Vordergrund steht und immer eine Trainingsbildmenge vorausgesetzt wird, die - im Hinblick auf den statistischen Aspekt - nicht zu klein sein darf, um repräsentativ zu sein.

Die Anwendbarkeit eines solchen Verfahrens auf die Aufgabenstellung dieser Arbeit ist aus diesem Grund in Frage zu stellen, vor allem in Bezug auf die erscheinungsbasierten Methoden: Die in dieser Arbeit angestrebten Invarianzen gegenüber den Ma-

⁶Der Begriff des Merkmals taucht im Rahmen dieser Arbeit in zweifacher Hinsicht auf: Auf der einen Seite stehen die zur Erkennung benötigten Größen als Merkmale und auf der anderen Seite die aus Bildern extrahierten Merkmale wie Kantensegmente etc. Letztere werden als Bildmerkmale bezeichnet, wohingegen der Begriff „Merkmal“ mit den erkenntnisrelevanten Größen assoziiert wird.

terialeigenschaften, die Anwendung auf Außenaufnahmen mit den daraus abgeleiteten Systemanforderungen und die Unabhängigkeit von der Blickrichtung würden eine erhebliche Anzahl an Trainingsbildern voraussetzen. Ob eine solche Menge an Trainingsbildern bereitgestellt werden kann, ist fraglich.

1.4.3 Deformierbare Modelle

Wenn bisher die Rede von dreidimensionalen Modellen war, so war implizit vorausgesetzt, dass es sich bei diesen Modellen um rigide, also formstarre Modelle handelt. Den Gegenpol zu diesen rigiden Modellen stellen die so genannten dynamischen Modelle dar. Sie werden im Kontext der Bildauswertung auch adaptive oder generische Modelle genannt. Dynamische Modelle sind parametrisiert und besitzen, an die Parameter gekoppelt, eine gewisse Elastizität oder Flexibilität. Stellt man nun das Modell mit einem speziellen Parametersatz dar, so spricht man von einer Ausprägung des Modells, die wiederum durch den Parametersatz kodiert ist.

Bei der Anpassung eines solchen Modells an ein Bild muss unterschieden werden, ob das Modell a priori ausgeprägt war, d. h. ob die Parameter des Modells bereits bekannt waren, oder ob im Zuge der Modellanpassung die Position und die Modellparameter simultan bestimmt werden. Im letztgenannten Fall setzt sich die Modellanpassung aus einer Positionsbestimmung, die auch als Lageschätzung bezeichnet wird, und einer Modellausprägung, die auch Formschätzung genannt wird, zusammen.

Zweifelsohne stellt die Anpassung dynamischer Modelle an Bilder höhere Anforderungen an die einzusetzende Algorithmik als die Anpassung rigider Modelle. Dies ist schon alleine darin begründet, dass sich der Freiheitsgrad, also die Anzahl der unbekannt Parameter, erhöht. Zudem vervielfacht sich die Anzahl der Mehrdeutigkeiten, mit denen der Anpassungsprozess konfrontiert wird, denn durch die Dynamik des Modells bieten sich gegebenenfalls mehrere, nahezu gleich gute Möglichkeiten, wie das Modell und das Bild in Übereinstimmung zu bringen sind. Solche Mehrdeutigkeiten werden auch als *Fata Morganas* bezeichnet. Die Ursache hierfür sind in den Korrelationen zwischen den Lage- und den Modellparametern zu suchen. Dieser Sachverhalt wird im späteren Verlauf dieser Arbeit ausführlich diskutiert. Andererseits bieten dynamische Modelle gute Chancen, die Bereitstellung einer umfassenden Menge an geometrisch genauen Modellen zu umgehen, die bei der Verwendung rigider Modelle erforderlich ist.

Die bisher bekannten dynamischen Modelle und Anwendungen lassen sich in zwei Gruppen unterteilen. Eine Untergruppe bilden die so genannten Gelenkmodelle, also dynamische Verknüpfungen mehrerer Modelle, wobei diese wiederum dynamisch sein können. Gelenkmodelle werden insbesondere zur Modellierung von Personen eingesetzt. In [Wachter '97] beispielsweise kommt ein Gelenkmodell in Form einer Kollektion von Kegelstümpfen zur Personenverfolgung zum Einsatz. Auch in [Felzenszwalb & Huttenlocher '00, Theobalt et al. '03] wurden Gelenkmodelle zur Personenmodellierung verwendet.

Die zweite Untergruppe der dynamischen Modelle sind die deformierbaren Modelle. Ihr klassischer Anwendungsbereich ist die medizinische Bild- und Volumendatenanalyse, da es sich bei den dort modellierten „Objekten“ im Allgemeinen um Körperteile handelt, deren Form von Mensch zu Mensch variiert und die zudem temporalen Deformationen unterworfen sein können. In [McInerney & Terzopoulos '96] ist eine umfassende Übersicht über deformierbare Modelle in medizinischen Anwendungsbereichen nachzulesen. Während viele dieser Ansätze auf Volumendaten operieren, werden nur einige der dort aufgeführten Ansätze auf Bilddaten angewandt. Diese verwenden dann häufig zweidimensionale Ansätze [Cootes & Taylor '01] oder parametrisierte Oberflächen [Terzopoulos & Metaxas '91] zur Objektmodellierung, deren Parameter nach Anpassung der Modelle an die Bild- oder Volumendaten das vorliegende Objekt kodieren. Bildverarbeitende Verfahren dieser Art setzen in der Regel gut ausgeprägte Objektkonturen oder gar die ganze Silhouette als bekannt voraus, wie z. B. [Rosenhahn et al. '03], und sind somit nur schwer auf Außenaufnahmen zu übertragen.

Auch für die Extraktion von Lippenparametern [Kroschel & Heckmann '02] und im Zusammenhang von Gesichtserkennungs- und Gesichtsverfolgungstechniken [Zhang & Kambhamettu '00, DeCarlo & Metaxas '02] kommen deformierbare Modelle zum Einsatz. Im technischen Bereich konzentrieren sich deformierbare Modelle weitgehend auf parametrisierte geometrische Primitive [Kecici '01, Kostianin et al. '01]. Für die Verkehrsüberwachung wurden interaktiv definierte generische Fahrzeugmodelle verwendet. In [Koller et al. '93] kommt ein solches Modell zur Objektverfolgung zum Einsatz, ebenso in [Zhao & Nevatia '01] zur Fahrzeugdetektion und in [Sullivan et al. '95, Ferryman et al. '95] zur Kategorisierung von Fahrzeugen in drei Klassen. Die Modellierung allgemeinerer Objektklassen bieten aktive Konturen [Blake et al. '01], aktive Formen [Cootes et al. '00] und die aktiven erscheinungsbasierten Verfahren [Stegmann '01]. Eine Übersicht über weitere solcher Ansätze findet man in [Buxton '03].

Ein dreidimensionaler Ansatz, der letztlich die Motivation für den in dieser Arbeit gewählten Lösungsweg geliefert hat, ist das adaptive lineare 3D-Gesichtsmodell aus [Blanz & Vetter '99, Blanz '00]. Dieser Ansatz ist sowohl den erscheinungs- als auch den formbasierten Verfahren zuzuordnen. Blanz und Vetter haben zunächst mit Hilfe eines Laserscanners digitale Bilder von circa 200 Gesichtern akquiriert. Bei diesen Bildern handelt es sich um strahlenförmige Projektionen der jeweiligen Farb- und Forminformation auf eine Zylindermantelfläche. Diese Projektionen werden zu vierkanaligen Bildern aufgeklappt und dann mit Bildverarbeitungsmethoden bearbeitet. Mit Hilfe von optischem Fluss, einem *bootstrapping*-Mechanismus und einer Hauptkomponentenanalyse gelangen die Autoren zu einem linearen, deformierbaren Modell. Dieses Modell kann alle zuvor in Form von Scans eingebrachten Gesichter sowie deren Linarkombinationen darstellen. Durch einen stochastischen Modellanpassungsalgorithmus sind Blanz und Vetter in der Lage, Gesichter aus Einzelbildern dreidimensional zu rekonstruieren.

In der neueren Arbeit [Blanz & Vetter '03] zeigen sie, dass ihr Modell über die Synthese und Rekonstruktion hinaus auch im Bereich der Gesichtserkennung anwendbar ist. Das Bedeutsame am Verfahren von Blanz und Vetter ist die Tatsache, dass ihnen

die automatische Konstruktion eines erscheinungsbasierenden adaptiven Modells gelungen ist, das nur aus Linearkombinationen von Vektoren besteht.

1.4.4 Zusammenfassung

Fasst man die Überlegungen der vorangegangenen Abschnitte zusammen, so kommt man - der Aufgabenstellung folgend - zu dem folgenden Schluss: Segmentbasierte Ansätze erwiesen sich in der im Vorfeld dieser Arbeit angestellten Untersuchung [Berger et al. '00] als ungeeignet und werden zusätzlich durch die Tatsache erschwert, dass keine reinen Senkrechtansichten vorliegen. Trainingsbasierte Ansätze sind schon wegen der großen Anzahl an a priori bereitzustellenden Trainingsbildern auszuklamern. Ansätze mit rigiden Modellen wären denkbar, verlangen aber entweder nach einer vollständigen Datenbasis an geometrisch genauen Modellen, die nicht verfügbar ist, oder sie müssen so große Toleranzen gegenüber den Modellen besitzen, dass die Extraktion trennungswirksamer Merkmale in Frage gestellt werden muss. Ansätze mit dynamischen Modellen sind entweder an sehr spezielle Objektklassen gebunden oder werden aus einer Menge an „Erfahrungen“, seien es Bildern oder Scans, konstruiert.

Die Lösung ist die Verwendung der zur Verfügung stehenden rigiden, geometrisch ungenauen Modelle als Eingabe für die Konstruktion dynamischer Modelle, wobei die rigiden Modelle die Rolle synthetischer „Erfahrungen“ übernehmen. Auf diese Weise kann die Forderung nach einer vollständigen Datenbasis abgeschwächt und Ungenauigkeiten können durch die Elastizität kompensiert werden.

1.5 Der Lösungsansatz: Verfahrensübersicht

Die Lösung der Aufgabe setzt sich im Kern aus drei Komponenten zusammen. Zum einen wurde im Rahmen dieser Arbeit ein automatischer Mechanismus entwickelt, der aus einer gegebenen Menge an rigiden Modellen (Basismodelle) ein deformierbares Modell erzeugt. Jedes dieser deformierbaren Modelle repräsentiert eine Reihe von geometrisch ähnlichen Objekten. Der zweite Schritt ist die Bereitstellung von Modulen zur Anpassung solcher Modelle an Bilder und deren Kombination zu einem automatischen Modellanpassungssystem. Die dritte Komponente ist schließlich die Extraktion trennungswirksamer Merkmale aus den gewonnenen Ausprägungen. Zur besseren Einordnung dieser Module in den gesamten Verarbeitungsprozess erfolgt zunächst eine Übersicht über das zur Typerkennung entwickelte System. Im Anschluss werden die drei oben genannten Verfahrenskomponenten unter Angabe der eingesetzten Techniken skizziert. Teile des Verfahrens wurden in [Laubenheimer & Link '03] veröffentlicht.

1.5.1 Das Typerkennungssystem

Abbildung 1.4 gibt eine Übersicht über das im Rahmen dieser Arbeit implementierte System zur Erkennung von Verkehrsflugzeugen. Die Eingaben sind jeweils ein adapti-

ves Modell und ein Bild mit Kollateraldaten, aus denen eine dreidimensionale Rekonstruktion des abgebildeten Flugzeugs berechnet wird. Aus der Rekonstruktion bzw. aus einer daraus berechneten zweidimensionalen Projektion werden die gesuchten Merkmale extrahiert. Der Merkmalsvektor wird mit einer prototypischen Datenbasis verglichen, die ebenfalls durch Anpassung des Modells an Schattenrisse oder Bilddaten erzeugt wird. Unter Angabe von Vertrauenswerten werden dem Anwender dann die drei wahrscheinlichsten Objekte vorgeschlagen.

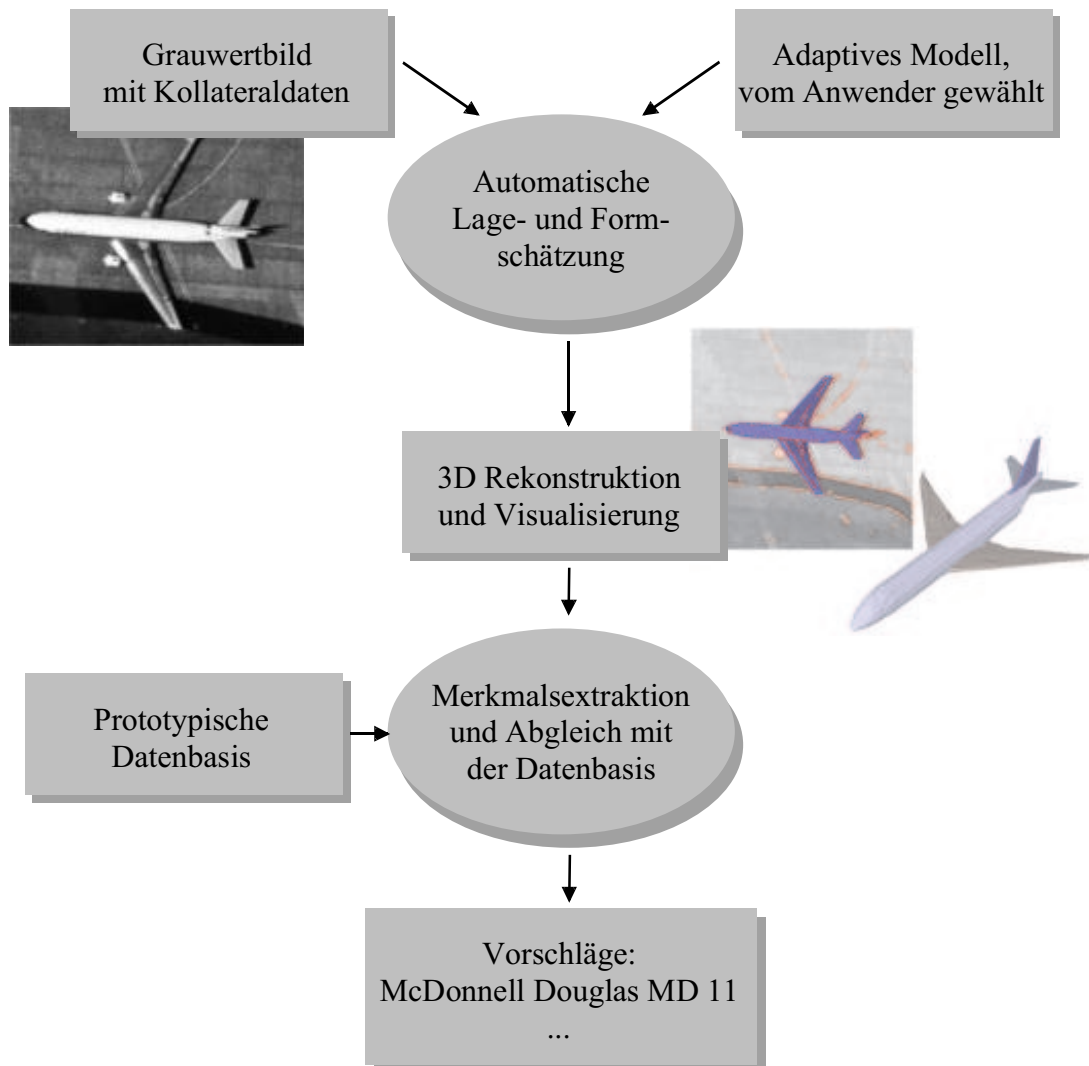


Abbildung 1.4: Eine Anwendung des in dieser Arbeit entwickelten Verfahrens im Bereich der Flugzeugerkennung. Nach Angabe des adaptiven Modells durch den Benutzer anhand diskreter Merkmale, wie z. B. dem Leitwerkstyp, wird eine automatische Merkmalsextraktion und Erkennung durchgeführt.

Das System verfügt über eine grafische Oberfläche, die dem Anwender gestattet, das Ergebnis der Modellanpassung - unter Angabe der Vertrauenswerte - zu kontrollieren und gegebenenfalls Maße aus der Rekonstruktion abzurufen.

1.5.2 Die Modellkonstruktion

Wie die in Abschnitt 1.4.3 aufgeführten Aspekte zum Thema Anpassung dynamischer Modelle im Vergleich zur Anpassung rigider Modelle bereits andeuten, ist die Modellanpassung die wesentliche Herausforderung des gewählten Lösungswegs. Ein entscheidender Punkt für das Gelingen dieser Anpassung ist die Elastizität, die das deformierbare Modell besitzt. Einerseits wird gewünscht, durch Deformationen Objekte ausprägen zu können, die nicht durch die Basismodelle vertreten sind. Dies setzt ein gewisses Mindestmaß an Dynamik voraus. Andererseits erhöht die Dynamik bei der Modellanpassung die Wahrscheinlichkeit, mit Artefakten wie Fata Morganas konfrontiert zu werden. In dieser Arbeit wird ein linearer Ansatz verwendet, von dem sich gezeigt hat, dass er für Flugzeugmodelle einen Mittelweg auf der Gratwanderung zwischen beiden Anforderungen darstellt.

Die Konstruktion dieser linearen Modelle ist eine Aufgabe, die vollständig im Bereich der *computational geometry* angesiedelt und aus zwei Schritten zusammengesetzt ist. Zunächst wird aus allen Basismodellen ein Dreiecksnetz konstruiert, die so genannte Meta-Topologie. Im zweiten Schritt wird für jedes Basismodell eine geometrische Darstellung in eben dieser Meta-Topologie berechnet. Das Ergebnis ist ein Dreiecksnetz, das jedes der Basismodelle darstellen kann. Analog zu [Blanz & Vetter '99, Shelton '00] wird das resultierende Modell in eine Linearkombination von Vektoren übertragen.

Die Meta-Topologie wird aus einer Erweiterung eines klassischen volumenbasierten Ansatzes der Metamorphosetechniken erzeugt. Die Konstruktion der geometrischen Darstellungen erfolgt durch Berechnung von Punktkorrespondenzen. Hierfür wird eine Netzoptimierung eingesetzt, welche den Ansatz aus [Ohtake & Belyaev '02] aufgreift und an die hier vorliegende Aufgabenstellung angepasst wurde. Zur Stabilisierung der Modellkonstruktion bedarf es einer Verzerrung der Basismodelle. Für diese Verzerrung werden interaktiv auf jedem Basismodell eine kleine Anzahl von so genannten Ankerpunkten definiert. Diese Vorbereitung für den Modellkonstruktionsprozess ist ohne Expertenwissen möglich und die einzige Interaktion, die der Anwender im Vorfeld zu leisten hat. Von dieser Interaktion und der Parametrierung des Verfahrens abgesehen, zeigt sich der Prozess dem Anwender gegenüber als vollautomatisch durchgeführter Prozess.

1.5.3 Die Modellanpassungsmodule

Die Module zur Modellanpassung sind Komponenten, die - auf eine zunächst nicht weiter präzierte Weise - die Parameter des Modells und der Lage so schätzen bzw. optimieren, dass Bild und Modell zu einer bestmöglichen Übereinstimmung kommen. Diese so genannte Registrierung kann auf Basis diverser Bildmerkmale durchgeführt werden, wie z. B. der iterativen Optimierung der quadratischen Differenz zwischen den Grauwerten im Bild und den Grauwerten einer synthetischen Ansicht des Modells. Ein solches Synthese-Analyse-Verfahren wird in [Blanz & Vetter '99] für Farbwerte eingesetzt, bedarf allerdings eines Beleuchtungsmodells, was die Anzahl der Freiheitsgrade

zusätzlich erhöht. Zudem würde die Forderung nach der Unabhängigkeit von den Bemalungsschemata bei einem solchen Ansatz die Bereitstellung eines Texturmodells erfordern, d. h. eines Modells, das neben Form auch Farbe bzw. Grauwerte enthält. Beide Fakten werden für Systeme verzichtbar, die auf Bildkanten aufsetzen. Aus diesem Grund wurde für die hier vorliegende Aufgabenstellung eine Modellanpassung entwickelt, die auf Gradientenbetragsbildern aufsetzt.

Ein weiterer, bei der Konstruktion eines Modellanpassers zu berücksichtigender Aspekt, ist die Laufzeiteffizienz. Während die Konstruktion des Modells in einem Vorverarbeitungsschritt durchgeführt wird, werden die Modellanpassungsmodule zur Systemlaufzeit eingesetzt. Bedingt durch diese Tatsache kommt der Rechengeschwindigkeit hier eine größere Bedeutung zu, als bei der Modellkonstruktion. Dieser Aspekt lässt bildpunktbasierte Optimierungsverfahren zunächst in den Hintergrund treten, wengleich sich herausgestellt hat, dass sie aus Gründen der Robustheit nicht verzichtbar sind. Bevorzugt werden aber schnellere Verfahren genutzt, die auf Bildkantensegmenten, also auf einer höheren semantischen Ebene, aufsetzen.

Die Modellanpassung ist durch die Optimierung einer Gütefunktion realisiert, die durch eine Quadratsumme gegeben ist. Die Gütefunktion wurde bereits in [Krüger '01] erfolgreich zur 2D-Registrierung von Karten an Bildmaterial eingesetzt. Sie ist robust, wobei das robuste Verhalten durch eine skalierbare Filterfunktion erreicht wird, die aus *Tukey's biweight* [Steward '99] abgeleitet ist. Entgegen der Arbeit von Krüger wird die Gütefunktion für die hier vorliegende Aufgabe nicht linearisiert, sondern durch eine (über die Skalierung der Filterfunktion) iterative Anwendung der Levenberg-Marquardt Optimierung realisiert. Die Anwendung dieses Schätzers auf ein Bild und ein Modell liefert zum einen den gesuchten Parametersatz und zum anderen einen Vertrauenswert, der durch die Gütefunktion mit der kleinsten Filtergröße, also der Zielgröße, definiert ist.

Um lokale Nebenmaxima zu überwinden, wird noch ein weiterer - nicht robuster - Schätzer eingesetzt, der durch Korrelation von Gradientenbildschablonen korrespondierende Bildpunkte bestimmt und die gesuchten Parameter nach einer Passpunktmethode optimiert. Bedingt durch die Tatsache, dass der Schätzer nicht robust ist, muss der Schätzer durch den Algorithmus überwacht werden. Insbesondere wird nach der Anwendung des Schätzers geprüft, ob sich die Anpassung tatsächlich verbessert hat.

An dieser Stelle ist zu betonen, dass die Modellkonstruktion und die Modellanpassung nur durch eine einzige Schnittstelle aneinander gekoppelt sind, nämlich durch die des linearen Modells. Dies bedeutet, dass die Modellanpassung für jedes andere lineare Modell eingesetzt werden kann, das in Form einer Linearkombination von Vektoren vorliegt.

Initialisierung der Modellanpassung

Während die Module zur Parameterschätzung anwendungsunabhängig gestaltet sind, muss der Modellanpasser, d. h. die spezifische Kombination dieser Module zu einem System, abhängig vom vorliegenden Bildmaterial eingestellt werden. Neben der Pa-

rametrierung der Anpassungmodule tritt auch die Frage nach den initialen Parametern auf, mit der die Modellanpassung gestartet wird. Wie alle Optimierungsverfahren nichtkonvexer Kosten- oder Gütefunktionen steht und fällt die Qualität des Ergebnisses mit der Qualität der Initialisierung, weil mit zunehmender Entfernung der initialen Parameter zum Funktionsoptimum mehr und mehr lokale Extremwerte überwunden werden müssen.

Die Initialisierung könnte beispielsweise durch eine äquidistante Abtastung des Parameterraums oder durch ein *random sampling* realisiert werden. Da die hier vorliegende Aufgabe jedoch einen hohen Freiheitsgrad besitzt, sind beide Techniken wegen der anfallenden Rechenzeit abzulehnen. In der vorliegenden Arbeit wird die Initialisierung durch Verwendung von Szenenwissen in Verbindung mit der äquidistanten Abtastung eines der zu schätzenden Lageparameter gelöst. Alternativ wäre die Verwendung eines fehlertoleranten Verfahrens denkbar. Diese Alternative wird im Rahmen der abschließenden Diskussion nochmals aufgegriffen.

1.5.4 Die Merkmalsextraktion und Typerkennung

Die dritte Komponente, die Merkmalsextraktion und Typerkennung, die den Auswerter in seiner Entscheidung unterstützen soll, ist ebenfalls abhängig von der Anwendung, und zwar von der jeweils vorliegenden Objektklasse. Für Verkehrsflugzeuge führen einige wenige Größen zu charakteristischen Merkmalsvektoren. Dazu gehört der Pfeilungswinkel, also der Winkel zwischen der Längsachse des Flugzeugs und der Tragfläche. In Verbindung mit der Länge, der Spannweite und dem Längenverhältnis des vorderen und hinteren Flugzeugteils, das durch den Tragflächenansatzpunkt gegeben ist, sind Verkehrsflugzeuge charakterisierbar. Diese Maße werden durch eine normalisierte Projektion der virtuellen Objektrekonstruktion auf eine Ebene - vergleichbar mit dem Schattenriss aus Abbildung 1.3 (b) - bestimmt.

1.6 Kapitelübersicht

In Kapitel 2 werden zunächst einige Begriffe definiert und Methoden der *computational geometry* erläutert, auf der die in Kapitel 3 beschriebene Konstruktion der adaptiven Modelle aufbaut. In Kapitel 4 werden die zur Modellanpassung entwickelten Module vorgestellt. Da das Verfahren zur Modellanpassung vielseitig anwendbar ist, erfolgt die Beschreibung der Module zunächst ohne Angabe der anwendungsspezifischen Parametrierung etc. Der Bezug zu den im Rahmen dieser Arbeit demonstrierten Anwendungen wird in Kapitel 5 hergestellt, in dem die Anwendung der Module zur Typerkennung und Lageschätzung beschrieben wird. In Kapitel 6 wird schließlich eine Zusammenfassung und ein Ausblick auf mögliche Erweiterungen und weitere Anwendungsfelder gegeben.

Kapitel 2

Grundlagen und Vorbereitungen

Die Konstruktion des adaptiven Modells setzt sich aus Komponenten der *computational geometry* zusammen. In diesem Kapitel werden einige Begriffe dieses Gebiets erläutert und Verfahren erklärt, die im nächsten Kapitel zur Definition und Konstruktion des adaptiven Modells verwendet oder für die Anwendung auf adaptive Modelle erweitert werden.

2.1 Darstellung dreidimensionaler Objekte

Techniken der *computational geometry* sind in der Regel an die Darstellung der vorliegenden Modelle gebunden. Je nachdem, ob die Objekte durch eine volumen- oder eine oberflächenbasierte Darstellung gegeben sind, sind unterschiedliche Verfahren anzuwenden. Häufig wird die Darstellungsform der Eingabemodelle gar als oberstes Kriterium für die Charakterisierung von Verfahren herangezogen (siehe z. B. [Lazarus & Verroust '98]).

Die in dieser Arbeit vorliegenden Basismodelle sind alle in einer Oberflächendarstellung gegeben. Andererseits wird von den Vorzügen der volumenbasierten Methoden Gebrauch gemacht. Für die Basismodelle werden also Volumendarstellungen und Methoden zur Umrechnung der Darstellungsformen benötigt. Im Folgenden wird eine Erläuterung zu den beiden Darstellungsformen und den später verwendeten Techniken zur Umrechnung gegeben. Um das Anwendungsziel einer weitgehenden Unabhängigkeit des Verfahrens von der jeweils vorliegenden Objektklasse zu erreichen, werden nur Darstellungen berücksichtigt, die zur Modellierung von Freiformobjekten¹ eingesetzt werden können.

Bei der Frage nach einer geeigneten Objektdarstellung für adaptive Modelle muss man sich vor Augen halten, dass die Modellkonstruktion im Vorfeld der Anwendung erfolgt und die Laufzeiteffizienz für den Konstruktionsschritt folglich sekundär ist. Insbesondere können innerhalb dieses Konstruktionsschritts auch vergleichsweise rechen-

¹Als Freiformobjekt werden hier solche Objekte bezeichnet, die sich nicht durch geometrische Primitive wie Quader oder Quadriken modellieren lassen. In [Campbell & Flynn '01] werden weitere Beschreibungen dieses nicht einheitlich verwendeten Begriffs angegeben.

intensive Darstellungsformen eingesetzt werden. Die Ausprägung des Modells und die Extraktion von Merkmalen zur Modellanpassung aus diesen Ausprägungen hingegen wird zur Systemlaufzeit durchgeführt. Da die Modellanpassung iterativ vorgenommen wird, ist mit einer erheblichen Anzahl von Ausprägungs- und Extraktionsvorgängen innerhalb der Anpassung des Modells an ein Bild zu rechnen. Aus der Anwendbarkeit der adaptiven Modelle in der Bildanalyse folgt also, dass die Modelle - nach der Konstruktion - in einer Darstellung vorliegen müssen, die die effiziente Modellausprägung und Merkmalsextraktion ermöglicht. Ein weiterer Punkt, der insbesondere im Zusammenhang mit der Modellierung technischer Objekte berücksichtigt werden muss, ist die mit der Darstellungsform erzielbare Genauigkeit.

2.1.1 Oberflächenbasierte Darstellung

Oberflächendarstellungen sind in der Regel in Form von segmentierten Flächen² gegeben. Sie sind also aus so genannten *patches* zusammengesetzt. Häufig sind die *patches* eben, die Darstellung ist also polyedrisch und dadurch in der Regel eine Approximation des darzustellenden Objekts. Handelt es sich bei den Polygonen ausschließlich um Dreiecke, so ist die Rede von einer Dreiecksdarstellung oder einer triangulierten Oberfläche. Im Folgenden werden nur noch triangulierte Oberflächen behandelt, da sich jede polyedrische Darstellung in eine Dreiecksdarstellung übertragen lässt.

Für die Definition des adaptiven Modells ist es wichtig, zwischen der so genannten Netztopologie und der geometrischen Realisierung einer triangulierten Oberfläche zu differenzieren.³ Die Netztopologie eines rigiden Modells besteht aus einem indizierten Graphen, der beschreibt, welche Ecken der triangulierten Fläche durch Kanten und Dreiecke verbunden sind. Ein solcher Graph beschreibt also die Konnektivität von Ecken, Kanten und Dreiecken. Die Graphen werden üblicherweise in einer Mengenschreibweise angegeben, die z. B. für ein einzelnes Dreieck wie folgt lautet:

Ecken: $\{v_1\}, \{v_2\}, \{v_3\}$
 Kanten: $\{v_1, v_2\}, \{v_2, v_3\}, \{v_3, v_1\}$
 Polygone: $\{v_1, v_2, v_3\}$

Die geometrische Realisierung hingegen ordnet jedem Knoten v_i des Graphen einen dreidimensionalen Vektor $(x_i, y_i, z_i)^T$ zu, identifiziert also jede Ecke mit einer geometrischen Position. Erst durch diese Zuordnung wird das obige Beispiel zu einem Dreieck im geometrischen Sinn. Analog zu [Blanz & Vetter '99] werden die geometrischen Realisierungen im Folgenden durch einen einzigen Vektor dargestellt, indem die Punktpositionen entsprechend der Indizierung der Ecken hintereinander geschrieben werden

$$\mathbf{g} = (x_0, y_0, z_0, x_1, y_1, z_1, \dots)^T. \quad (2.1)$$

²Eine Definition und Charakteristika von segmentierten Flächen findet man in [Aumann & Spitzmüller '92].

³In [Hoppe et al. '93] werden mathematische Definitionen für die Begriffe Netztopologie und geometrische Realisierung angegeben, die aus der Theorie der simplizialen Komplexe entwickelt wurden.

Die Notation für ein Modell \mathcal{M} mit Netztopologie \mathcal{T} und geometrischer Realisierung \mathbf{g} lautet dann $\mathcal{M} = (\mathcal{T}, \mathbf{g})$. Offensichtlich ist diese Form der Darstellung nicht eindeutig, da viele Topologien dieselbe Oberfläche modellieren können.

2.1.2 Volumenbasierte Darstellung

Dreidimensionale Objekte lassen sich auch durch ihr Volumen beschreiben. Solche volumenbasierten Darstellungen lassen sich in kontinuierliche und diskrete Darstellungen unterteilen. Eine gängige diskrete Volumenrepräsentation ist die so genannte Voxel- bzw. Rasterdarstellung. Sie entsteht durch eine Diskretisierung des 3D-Raums und zwar durch Definition eines regulären Gitters von Voxeln (Voxel = *volume pixel element*), die numerische Informationen über das darzustellende Objekt tragen. Ist der Wertebereich der Voxel auf eine endliche Menge eingeschränkt, spricht man von einem quantisierten Voxelmodell bzw. von einem binären Voxelmodell, wenn nur zwei Werte zugelassen sind. In beiden Fällen trägt das Modell in Verbindung mit einem Schwellwert Information über das Innere und Äußere des dargestellten Objekts. Allgemein lässt sich jedoch feststellen, dass gerasterte, diskrete Volumendarstellungen zur Darstellung technischer Objekte wegen ihres approximativen Charakters nur schlecht geeignet sind. Dies gilt insbesondere dann, wenn die Objekte dünne oder sich verjüngende Objektteile besitzen.

Zur Darstellungsoptimierung bzw. zur Reduktion des Speicheraufwands können Zellzerlegungen in Form von hierarchisch organisierten Oktalbäumen eingesetzt werden (siehe z. B. [Lerios et al. '95]). Zur Konstruktion von adaptiven Modellen sind diese jedoch ungeeignet, da die lokale Auflösung von der jeweiligen Ausprägung des adaptiven Modells abhängig ist. Das bedeutet, dass die Zellzerlegung entweder für jede Ausprägung separat berechnet werden muss oder aber das Zellraster insgesamt sehr häufig bis in die feinste Auflösungsstufe unterteilt sein muss. Im ersten Fall ist mit einer nicht akzeptablen Rechenzeit für die Modellausprägung zu rechnen und im zweiten Fall geht der Vorteil der Speichersparnis gegenüber den Voxeln verloren, bei gleichzeitig erhöhter Darstellungskomplexität.

Den diskreten Volumendarstellungen stehen die kontinuierlichen gegenüber, die durch implizite Funktionen realisiert werden. Im Rahmen dieser Arbeit sind vor allem diejenigen Volumendarstellungen von Interesse, die aus Oberflächendarstellungen konstruiert werden können. Zur Konstruktion solcher impliziter Darstellungen stehen mehrere Techniken zur Verfügung, von denen in den folgenden Abschnitten zwei beschrieben werden. Beide Techniken liefern zu einem gegebenen Objekt \mathcal{A} eine Funktion der Form $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

$$f(\mathbf{x}) \quad \begin{cases} > c & \mathbf{x} \text{ liegt im Inneren von } \mathcal{A} \\ = c & \mathbf{x} \text{ liegt auf der Oberfläche von } \mathcal{A} \\ < c & \text{sonst} \end{cases}, \quad (2.2)$$

wobei die reelle Konstante c den so genannten Isowert bezeichnet. Die entsprechende Isofläche (bzgl. c) ist dann durch die Menge

$$\{\mathbf{x} \mid f(\mathbf{x}) = c\}$$

gegeben.

2.1.2.1 Radiale Summen

Eine gebräuchliche Technik ist die Approximation von Oberflächen durch gewichtete Summen radialer Basisfunktionen, kurz radiale Summen genannt. Radiale Summen gehen auf [Duchon '76] zurück und sind den Methoden der multivariaten Interpolationstechniken zuzuordnen. Im Verlauf der Verfahrensentwicklung im Rahmen dieser Arbeit hat sich jedoch gezeigt, dass radiale Summen ungeeignet sind, die hier vorliegenden Flugzeugmodelle mit ausreichender Genauigkeit darzustellen. Diese Beobachtung ist darauf zurückzuführen, dass radiale Summen dazu neigen, an scharfen Kanten aufzuschwingen, wie Abbildung 2.1 zeigt. Andererseits sind radiale Summen für andere Objektklassen durchaus als Konstruktionsgrundlage für adaptive Modelle in Betracht zu ziehen, besitzen also im Zuge der Allgemeinheit des hier vorgestellten Ansatzes eine nicht zu vernachlässigende Relevanz. Darüber hinaus spielen radiale Summen in dieser Arbeit zur Konstruktion des adaptiven Modells eine von der Objektdarstellung unabhängige Rolle, werden daher in diesem Abschnitt kurz erklärt.

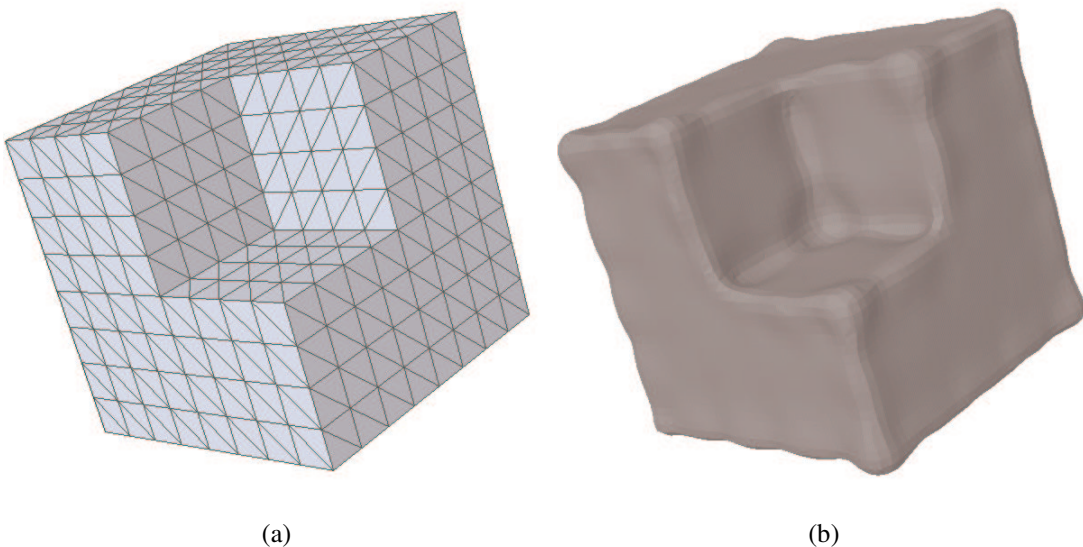


Abbildung 2.1: Radiale Summen zur Objektdarstellung. In (a) ist das originale Objekt abgebildet und in (b) eine polygonisierte Approximation durch radiale Summen. Als Stützstellen wurden hier die Ecken des Voronoi-Mosaiks verwendet.

Für eine Menge von K d -dimensionalen Punkten bzw. Stützstellen $\mathbf{c}_1, \dots, \mathbf{c}_K$ und genauso vielen skalaren Mess- bzw. Sollwerten h_1, \dots, h_K wird eine reellwertige Funk-

tion f gesucht, so dass $f(\mathbf{c}_k) = h_k$ für $k = 1, \dots, K$ gilt. Die gesuchte Funktion f wird wie folgt bestimmt: Für eine a priori festgelegte, univariate und stetige Funktion $\varphi : \mathbb{R}_+^0 \rightarrow \mathbb{R}$ wird zunächst die so genannte radiale Basisfunktion

$$\Phi(\mathbf{x}) := \varphi(\|\mathbf{x}\|)$$

erklärt, wobei $\|\cdot\|$ die euklidische Norm bezeichnet. Beispiele für radiale Basisfunktionen sind die so genannten *thin plate splines* $\varphi(d) = d^2 \ln d$ für radiale Summen des \mathbb{R}^2 und $\varphi(d) = d^3$ im \mathbb{R}^3 [Duchon '76].⁴ Die radiale Summe f ist dann für $\mathbf{x} = (x^1, \dots, x^d)^T$ durch

$$f(\mathbf{x}) := \sum_{k=1}^K d_k \Phi(\mathbf{x} - \mathbf{c}_k) + p(\mathbf{x})$$

definiert, wobei $p(\mathbf{x}) = \alpha_0 + \alpha_1 x^1 + \dots + \alpha_d x^d$ ein Polynom ersten Grades bezeichnet.⁵ Die Aufgabe der Interpolation reduziert sich nun auf die Bestimmung der Gewichte d_1, \dots, d_K und der Polynomkoeffizienten $\alpha_0, \dots, \alpha_d$, so dass die Stützstellen \mathbf{c}_k auf ihre Sollwerte h_k abgebildet werden. Da die Forderungen $f(\mathbf{c}_k) = h_k$ für *thin plate splines* ein unterbestimmtes Gleichungssystem erzeugen, wird zusätzlich gefordert, dass gilt

$$\sum_{k=1}^K d_k q(\mathbf{c}_k) = 0 \text{ für alle Polynome } q \text{ vom Grad eins.}$$

Diese Forderung hat für die resultierende Funktion ein gewisses Maß an Glattheit zur Folge (vgl. [Carr et al. '01]). Die Berechnung der Gewichte erfolgt durch Lösung des folgenden linearen Gleichungssystems (LGS)

$$\begin{pmatrix} \Phi(\mathbf{c}_1 - \mathbf{c}_1) & \dots & \Phi(\mathbf{c}_1 - \mathbf{c}_K) & 1 & c_1^1 & \dots & c_1^d \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Phi(\mathbf{c}_K - \mathbf{c}_1) & \dots & \Phi(\mathbf{c}_K - \mathbf{c}_K) & 1 & c_K^1 & \dots & c_K^d \\ 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ c_1^1 & \dots & c_K^1 & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ c_d^1 & \dots & c_K^d & 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ \vdots \\ d_K \\ \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} = \begin{pmatrix} h_1 \\ \vdots \\ h_K \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

wobei c_k^i die i -te Komponente von \mathbf{c}_k bezeichnet. Für *thin plate splines* ist die symmetrische Matrix, auch Interpolationsmatrix genannt, positiv definit und das Gleichungssystem folglich eindeutig lösbar.

Doch nun zurück zur volumenbasieren Objektdarstellung. Ist ein Objekt in Form einer triangulierten Oberfläche \mathcal{M} gegeben, so kann eine radiale Summe, deren Isofläche

⁴Weitere Basisfunktionen findet man in [Wendland '95, Buhmann '01].

⁵Ob dieses Polynom benötigt wird und welchen Grad es besitzen muss, ist abhängig von der gewählten Basisfunktion.

die Oberfläche des Objekts approximiert, wie folgt konstruiert werden: Aus der geometrischen Realisierung des Dreiecksnetzes wird eine geeignete Menge von Punkten $\mathbf{c}_1, \dots, \mathbf{c}_K$ ausgewählt und die entsprechenden Sollwerte h_1, \dots, h_K werden auf Null gesetzt. Anschließend wird jeder Punkt \mathbf{c}_k in Normalenrichtung um eine feste Konstante $\varepsilon > 0$ nach außen geschoben und so ein Punkt \mathbf{c}_{K+k} definiert. Für die so verschobenen Punkte werden die Sollwerte h_{K+1}, \dots, h_{2K} auf einen festen Wert $\delta < 0$ gesetzt. Konstruiert man aus diesen Stützstellen mit den selbstdefinierten Sollwerten durch Lösung des entsprechenden Gleichungssystems eine radiale Summe, so ist dies eine implizite Funktion, deren Isofläche mit Isowert $c = 0$ die Objektoberfläche approximiert.

In [Turk & O'Brian '99] wird diese Technik zur Definition eines Metamorphoseverfahrens eingesetzt und in [Carr et al. '01] zur Konstruktion von Oberflächen aus dreidimensionalen Punktwolken. Im Zusammenhang mit technischen Objekten stellt sich allerdings die Frage, wie gut die Isoflächen solcher radialen Summen die Objektoberflächen approximieren. An den Stützstellen (in [Turk & O'Brian '99] werden die Ecken des Dreiecksnetzes als Stützstellen verwendet) fällt die Isofläche mit dem Dreiecksnetz zusammen, über das Verhalten der Funktion in den Zwischenräumen kann jedoch nur schwer eine Aussage gemacht werden. Zwar gibt es für viele Basisfunktionen Aussagen über deren Approximationsgüte in Abhängigkeit von der Dichte der Stützstellen. Diese Aussagen sind jedoch immer an gewisse Glattheitsforderungen an die zu approximierende Oberfläche gebunden. Im Fall von polyedrischen Oberflächen sind diese Glattheitsforderungen nicht erfüllt. In Bezug auf eine adäquate Dichte bzw. Anzahl von Stützstellen ist zudem zu berücksichtigen, dass die Interpolationsmatrix in etwa so viele Zeilen und Reihen besitzt wie Stützstellen vorliegen. Dies führt, abhängig von der Implementierung des LGS-Lösers, zu einer oberen Schranke der möglichen Anzahl an Stützstellen. Ferner muss bedacht werden, dass die numerische Stabilität der Interpolationsmatrix abnimmt, wenn die Anzahl der Stützstellen erhöht wird. In [Narcowich et al. '99] wird dieser Sachverhalt wie folgt kommentiert: „*In short, one has to sacrifice good reproduction quality for good stability or vice versa.*“.

2.1.2.2 Signierte Distanztransformation

Eine Alternative zu radialen Summen ist die so genannte signierte Distanztransformation (SDT), die eine im Bereich der Gleitkommandarstellung genaue implizite Darstellung ist. Für die Definition der SDT wird zunächst der Abstand eines Punktes \mathbf{x} zu einer abgeschlossenen Menge \mathcal{A} benötigt, der durch

$$d(\mathbf{x}, \mathcal{A}) = \min_{\mathbf{y} \in \mathcal{A}} \|\mathbf{x} - \mathbf{y}\| \quad (2.3)$$

gegeben ist. Die SDT ist dann für jedes Objekt \mathcal{A} mit einer abgeschlossenen Oberfläche $\partial\mathcal{A}$ erklärt durch

$$f_{\mathcal{A}}(\mathbf{x}) = \begin{cases} d(\mathbf{x}, \partial\mathcal{A}) & \mathbf{x} \in \mathcal{A} \\ -d(\mathbf{x}, \partial\mathcal{A}) & \text{sonst} \end{cases} .$$

Im Gegensatz zu radialen Summen mit *thin plate splines* ist die Distanztransformation exakt, aber für polyedrische Flächen nur stückweise differenzierbar. Abbildung 2.2 zeigt die SDT für eine zweidimensionale Kontur mit den nicht differenzierbaren Knicken. Die Unstetigkeitsstellen treten genau an den Punkten auf, die zu zwei verschiedenen Kanten denselben Abstand haben. Die im späteren Verlauf benötigten Gradienten der SDT werden an solchen Stellen durch eine einseitig stetige Fortsetzung approximiert.

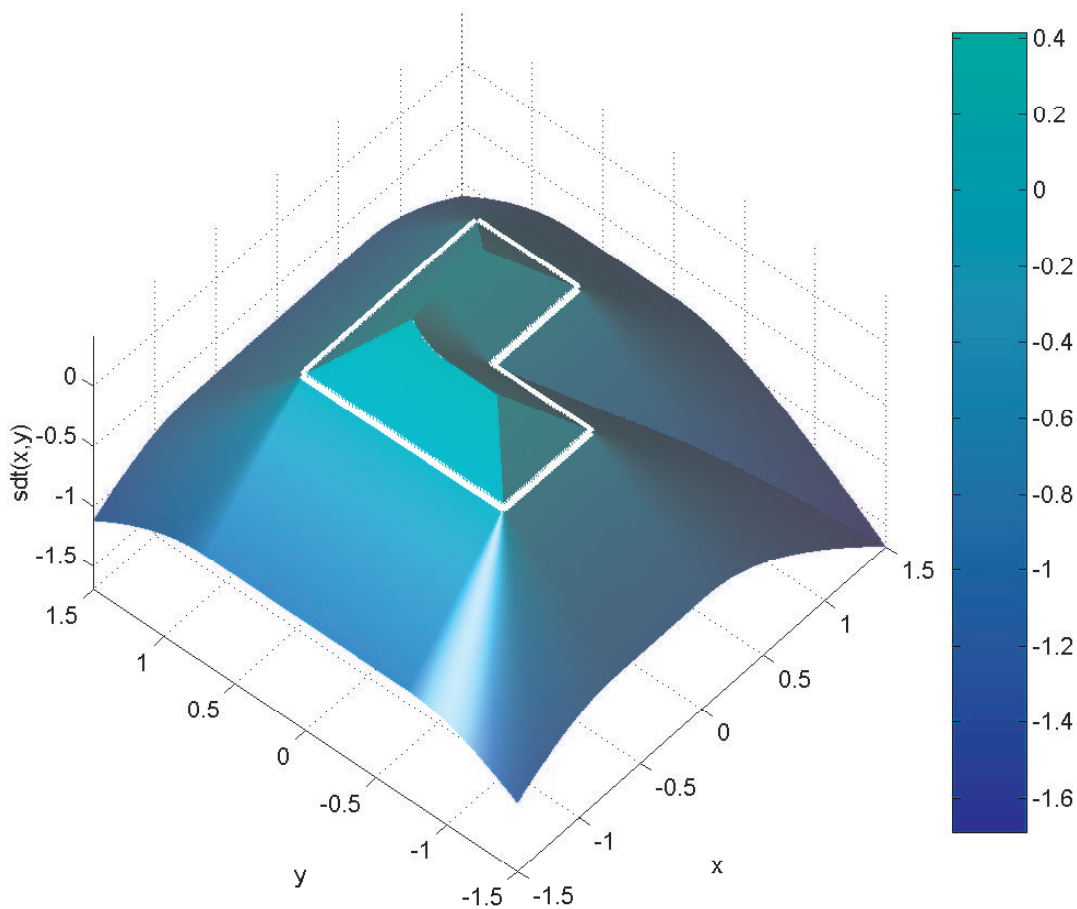


Abbildung 2.2: Ein Beispiel für die SDT einer zweidimensionalen Kontur (weiß).

Berechnung der signierten Distanz

Die SDT kann, wieder im Gegensatz zu radialen Summen, nicht in einer geschlossenen Form angegeben werden. Statt dessen muss sie den aufrufenden Instanzen in Form eines Algorithmus zur Verfügung gestellt werden, der die signierte Distanz für einen einzelnen Punkt berechnet.

Eine solche Funktion, die für einen gegebenen Punkt x und ein Modell \mathcal{M} die signierte Distanz berechnet, ist in Algorithmus 2.1 skizziert. Das Prinzip der Berechnung der

signedDistance ()

Eingabe: Modell \mathcal{M} in triangulierter Oberflächendarstellung
Punkt x

Ausgabe: signierter Abstand s von x zu \mathcal{M}

```

minDist = maximal möglicher Abstand
sign = 0

for (jedes Dreieck  $T$  von  $\mathcal{M}$ )
  if (triangleIsRelevant( $x$ ,  $T$ ))
    d = signedOrthogonalDistance( $x$ ,  $T$ )
    if (fabs(d) < minDist)
      minDist = fabs(d)
      sign = sgn(d)
end for

for (jede Kante  $K$  von  $\mathcal{M}$ )
  if (edgeIsRelevant( $x$ ,  $K$ ))
    d = signedOrthogonalDistance( $x$ ,  $K$ )
    if (fabs(d) < minDist)
      minDist = fabs(d)
      sign = sgn(d)
end for

for (jede Ecke  $E$  von  $\mathcal{M}$ )
  d = distance( $x$ ,  $E$ )
  if (d < minDist)
    minDist = d
    sign = 0
end for

if (sign == 0)
  // Minimaler Abstand ist durch eine Ecke gegeben
  sign = pointInPolyeder( $x$ ,  $\mathcal{M}$ )

s = sign * minDist

```

Algorithmus 2.1: Berechnung der signierten Distanz für einen gegebenen Punkt x zu einem Modell \mathcal{M} in triangulierter Oberflächendarstellung. Die Funktion **fabs ()** gibt den Absolutbetrag, **sgn ()** das Vorzeichen und **distance ()** den euklidischen Abstand zurück. Die übrigen Funktionen werden im Text erklärt.

signierten Distanz beruht darauf, zunächst den minimalen Abstand des Punkts zu allen Dreiecken des Modells zu berechnen. Das Vorzeichen wird anschließend durch einen Punkt-in-Polyeder-Test bestimmt. Für die Berechnung des Punktabstands zu einem Dreieck muss geprüft werden, ob der gesuchte minimale Abstand durch den Abstand des Punkts zum Inneren des Dreiecks, zu einer Kante oder zu einer Ecke gegeben ist (vgl. Abbildung 2.3). Da die Ecken und Kanten des Modells jeweils zu mehre-

ren Dreiecken gehören, würde die sequentielle Abarbeitung aller Dreiecke samt ihrer Ecken und Kanten mehrfache Zugriffe auf diese Ecken und Kanten zur Folge haben. Der Algorithmus 2.1 umgeht diese mehrfachen Zugriffe, indem die Abstandsberechnungen zuerst für das Innere aller Dreiecke des Modells, dann für alle Kanten und anschließend für alle Ecken durchgeführt werden.

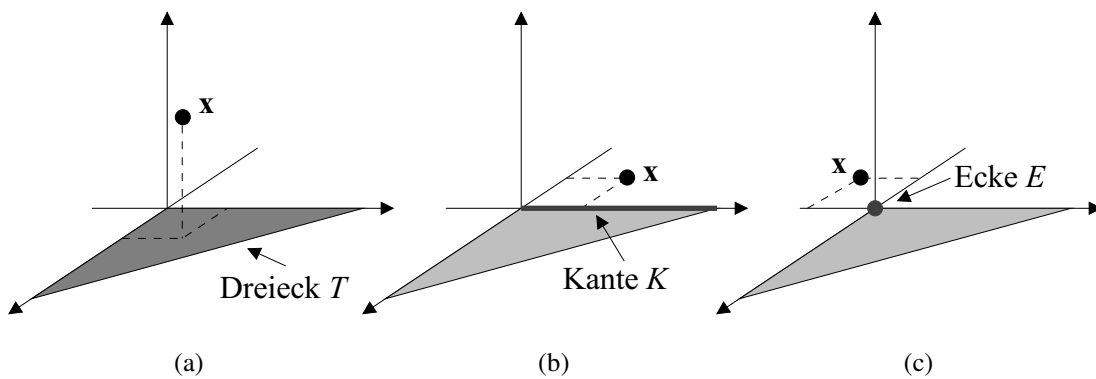


Abbildung 2.3: Der minimale Abstand eines Punktes x zu einem Dreieck ist, je nach Lage von x , durch den Abstand zum Inneren des Dreiecks (a), zu einer Kante (b) oder zu einer Ecke (c) gegeben.

Zudem wird von der Tatsache Gebrauch gemacht, dass für einen gegebenen Punkt nur bestimmte Dreiecke und Kanten den minimalen Abstand liefern können. Für ein Dreieck bzw. eine Kante ist diese Menge durch den Schnitt von drei bzw. vier Halbräumen definiert (vgl. Abbildung 2.4), die a priori berechnet werden können. Die Abfrage, ob ein gegebener Punkt in diesen so genannten Einflussbereichen liegt, wird im Algorithmus 2.1 mit den Funktionen `triangleIsRelevant()` bzw. `edgeIsRelevant()` bezeichnet. Fällt dieser Test positiv aus, so entspricht der gesuchte Abstand des Punktes zu dem jeweiligen Element dem orthogonalen Abstand zur linearen Hülle dieses Elements. Sonderfälle müssen nicht mehr berücksichtigt werden. Das Vorzeichen kann dann aus der Normalen des Dreiecks bzw. aus den Normalen der an die Kante angrenzenden Dreiecke berechnet werden, wodurch die überladene Funktion `signedOrthogonalDistance()` erklärt ist.

Eine Ecke kommt zur Berechnung des minimalen Abstands nur dann in Frage, wenn der Punkt in dem (in die Ecke verschobenen) Normalenkegel der Ecke liegt. Deshalb lassen sich auch für Ecken solche Einflussbereiche bestimmen. Da dieser Normalenkegel jedoch nicht notwendigerweise konvex ist, wird der Test, ob ein Punkt in diesem Kegel liegt, erheblich komplexer als im Fall von Dreiecken und Kanten. Aus diesem Grund wurde in der hier vorliegenden Implementierung auf die Berücksichtigung der Einflussbereiche von Ecken verzichtet. Für die Berechnung des Vorzeichens muss dann allerdings ein Punkt-in-Polyeder-Test durchgeführt werden.

Für diese Arbeit wurde für den Punkt-in-Polyeder-Test eine Halbgerade definiert, die im Punkt x startet und keine Kanten oder Ecken der Modelloberfläche trifft. Zur Konstruktion dieser Halbgeraden wird ein zufälliger Richtungsvektor generiert und anschließend ein Test durchgeführt, ob die Halbgerade Kanten oder Ecken trifft. Ist dies

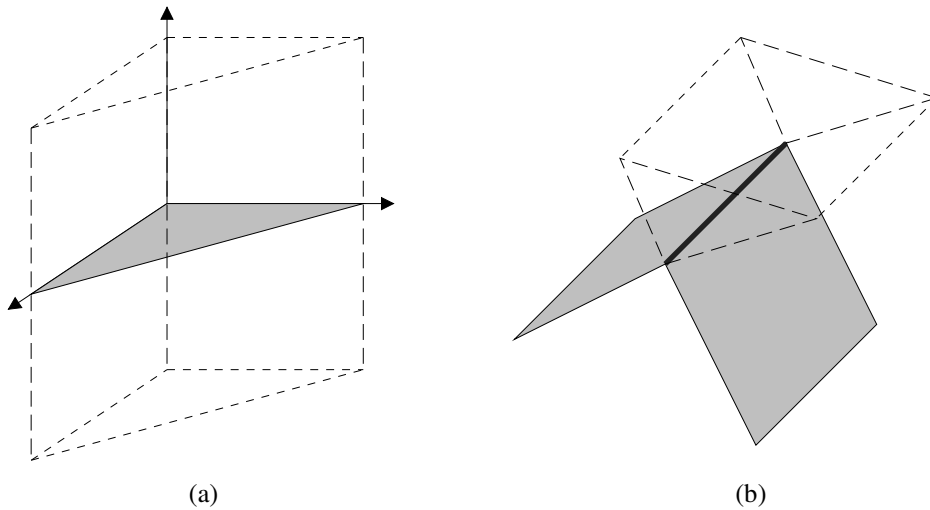


Abbildung 2.4: Die Einflussbereiche von Dreiecken lassen sich durch den Schnitt von drei Halbräumen angeben (a). Die Einflussbereiche von Kanten sind durch den Schnitt von vier Halbräumen gegeben (b).

der Fall, so wird ein neuer Richtungsvektor generiert usw. Das Verfahren konvergiert, da die Ecken und Kanten auf der Modelloberfläche eine Nullmenge bilden.⁶ Liegt eine solche Halbgerade vor, wird die Anzahl der Dreiecke gezählt, die von der Halbgeraden durchstoßen werden. Ist die Anzahl gerade, so liegt der Punkt außerhalb, ist sie ungerade, so liegt der Punkt innerhalb der Modelloberfläche. Diese Methode ist eine Abwandlung einer in [Aumann & Spitzmüller '92] aufgeführten Methode zur Lokalisierung von Punkten in zweidimensionalen Vielecken.

Die Berechnung der Distanztransformation kann in Bezug auf die Rechenzeit weiter optimiert werden, wenn auch die Einflussbereiche der Ecken berücksichtigt werden und dadurch der Punkt-in-Polyeder-Test entfallen kann. Eine zusätzliche Optimierung kann erreicht werden, wenn ein Zellraster angelegt wird, in dem jede Zelle eine Liste mit allen Einflussbereichen besitzt, die die Zelle treffen. Die Berechnung der signierten Distanz eines gegebenen Punktes kann dann auf diejenigen Dreiecke, Kanten und Ecken eingeschränkt werden, deren Einflussbereich diejenige Zelle trifft, in der auch der Punkt liegt. In [Mauch '03] ist ein Verfahren zur Berechnung der Distanztransformation beschrieben, das von einer solchen Zellrastertechnik Gebrauch macht.

⁶Eine Nullmenge ist eine Menge \mathcal{A} , für die gilt $\mathcal{A}^0 = \emptyset$, deren Inneres also leer ist.

2.2 Polygonisierung und Netzoptimierung

Im letzten Abschnitt wurde erklärt, wie zu einem in Oberflächendarstellung gegebenen Modell eine entsprechende implizite Funktion konstruiert werden kann. Auch die Umkehrung ist - in der Regel verlustbehaftet - durchführbar, d. h. zu einer gegebenen impliziten Funktion kann ein entsprechendes Dreiecksnetz konstruiert werden. Man spricht in diesem Zusammenhang von der Polygonisierung impliziter Flächen.

2.2.1 Polygonisierung

Zur Polygonisierung impliziter Flächen stehen diverse Verfahren zur Verfügung. Sie werden anhand der eingesetzten Technik zur Überdeckung der Isofläche charakterisiert. Die Überdeckung kann durch Konvergenz realisiert sein, also durch eine rekursive Zerlegung einer initialen Zelle in immer feinere Zellstrukturen, gegebenenfalls unter Berücksichtigung der lokalen Krümmung. Schneller jedoch sind Verfahren, die auf fortgesetzter Überdeckung basieren. Das bedeutet, dass nur eine einzige Zelle mit einer vorab definierten Größe instantiiert wird, die im Verlauf der Polygonisierung über die Oberfläche des Objekts geschoben wird. Die einfachste Form der Polygonisierung wird als *marching cube* bezeichnet, einem Synonym für die Polygonisierung durch Fortsetzung mit einer Würfelzelle.⁷ In dieser Arbeit wird das *marching cube*-Verfahren aus [Bloomenthal '94] verwendet, für das eine frei verfügbare C-Implementierung erhältlich ist.

Das Verfahren setzt allerdings voraus, dass die Isofläche jede Zelle höchstens einmal trifft. Dies hat für dünne Objektteile, wie z. B. die hinteren Kanten der Tragflächen zur Konsequenz, dass die Zelle sehr klein gewählt werden muss oder aber erhebliche Ungenauigkeiten bezüglich der Rekonstruktion in Kauf genommen werden müssen. In [Bloomenthal & Ferguson '95] wird diese Voraussetzung durch die Berücksichtigung zahlreicher Sonderfälle fallen gelassen. Dennoch werden auch bei dieser Verfahrensvariante nicht alle Besonderheiten technischer Objekte, wie z. B. Schleifen innerhalb der Zellgröße, berücksichtigt.

2.2.2 Netzoptimierung nach Ohtake & Belyaev

Eine Alternative zur Bearbeitung solcher Besonderheiten lässt sich auch durch eine Nachbearbeitung der Ausgabe des Polygonisierers lösen. In [Ohtake & Belyaev '02] wird eine solche Netzoptimierung vorgestellt, die insbesondere Objekte mit scharfkantigen Oberflächenmerkmalen gut rekonstruiert. Die Eingaben für das Verfahren von Ohtake und Belyaev sind eine implizite Funktion sowie ein Dreiecksnetz, das die Isofläche der Funktion bereits approximiert, also beispielsweise die Ausgabe ei-

⁷ In [Bloomenthal '97] werden Fortsetzungs-Verfahren weiter in stückweise lineare und Prediktor-Korrektor Methoden unterteilt. Andere Verfahren, wie z. B. die langsamen Partikelverfahren, die keiner der beiden Gruppen zugeordnet werden können, sind in [Karkanis & Stewart '01] in einer kurzen Übersicht dargestellt.

nes *marching cube*. Teile dieses Verfahrens werden zur Konstruktion des adaptiven Modells eingesetzt und aus diesem Grund näher erläutert.

Die Netzoptimierung von Ohtake und Belyaev basiert auf der Anwendung von vier Operatoren, die das Modell entsprechend der gegebenen Isofläche optimieren:

- **Krümmungsabhängige Regularisierung (\mathcal{R} -Operator):** In Abhängigkeit von der lokalen Krümmung der impliziten Funktion werden die Ecken regularisiert. Dabei werden die Ecken des Netzes auf der Isofläche in Richtung starker Krümmungen verschoben. Der Operator wird durch eine skalare Größe parametrisiert, die beeinflusst, wie stark die lokale Krümmung bei der Regularisierung berücksichtigt wird.
- **Duale Netzoptimierung (\mathcal{P} -Operator):** Die Ecken des Netzes werden durch die Berechnung von Schnittpunkten angrenzender tangentialer Ebenen optimiert. Der Operator wird durch einen Schwellwert kontrolliert, der weiter unten detailliert erläutert wird.
- **Projektion auf die Isofläche (\mathcal{Z} -Operator):** Die Ecken werden längs des Gradienten der impliziten Funktion auf die andere Seite der Isofläche verschoben und anschließend durch ein Bisektionsverfahren auf der Isofläche positioniert.
- **Krümmungsabhängige Netzunterteilung (\mathcal{S} -Operator):** Dreiecke, auf denen die implizite Funktion stark gekrümmt ist, werden in vier Subdreiecke unterteilt.⁸ Die Definition, welche Krümmung zu einer Dreiecksunterteilung führt, wird durch einen Schwellwert festgelegt.

Die ersten beiden Operatoren stellen die Kernelemente der Optimierung dar. Beide verwenden ein duales Netz, also ein Netz, dessen Ecken durch die Schwerpunkte der Dreiecke des primären (originalen) Netzes gegeben sind (siehe Abbildung 2.5). Jede Ecke des primären Netzes kann mit einem (nicht notwendigerweise ebenen) Vieleck des dualen Netzes identifiziert werden, das sich aus den Schwerpunkten der an die Ecke angrenzenden Dreiecke zusammensetzt. Das duale Netz wird noch durch Anwendung des \mathcal{Z} -Operators optimiert, wodurch die Ecken des dualen Netzes auf der Isofläche der impliziten Funktion positioniert werden. Die Dreiecksunterteilung durch den \mathcal{S} -Operator sorgt schließlich dafür, dass die Topologie des Netzes fein genug ist, um starke lokale Krümmungen modellieren zu können.

Die Autoren Ohtake und Belyaev schlagen die iterative Anwendung der Operatoren in der Form

$$(\mathcal{R}^m + \mathcal{P} + \mathcal{S})^n + (\mathcal{R}^m + \mathcal{P})$$

vor, also eine m -fache Anwendung des \mathcal{R} -Operators, gefolgt von der Anwendung des \mathcal{P} -Operators usw. Die Autoren haben experimentell ermittelt, dass $m = 3$ und $n = 2, 3, 4, 5$ gute Triangulierungen der Isofläche liefern.

⁸Eine alternative Technik zu dieser 1:4 Unterteilung, die auch von Ohtake und Belyaev vorgeschlagen wird, findet man in [Kobbelt '00].

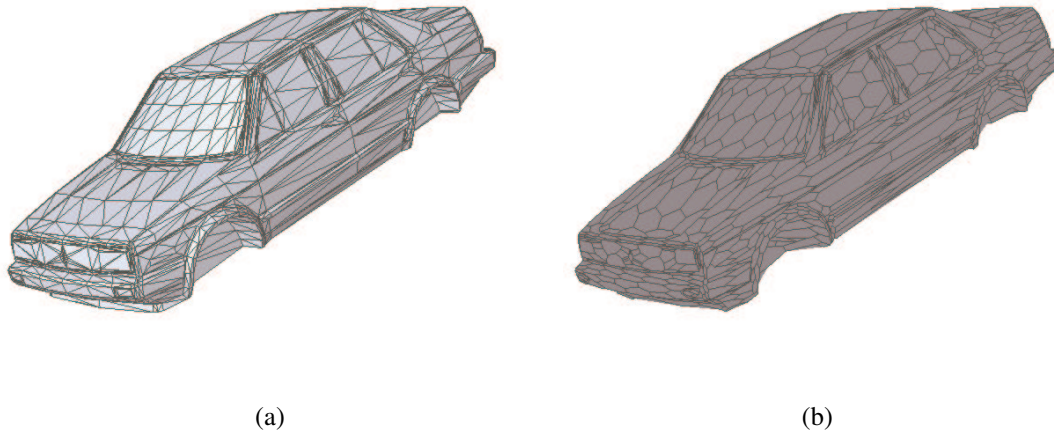


Abbildung 2.5: Das duale Netz für ein PKW-Modell. In (a) ist das Original, in (b) das duale Netz abgebildet.

Die wesentliche Errungenschaft der Arbeit von Ohtake und Belyaev ist die gute Rekonstruktionsqualität von scharfen Objektmerkmalen, die auf die krümmungsabhängige Regularisierung, aber mehr noch auf die duale Netzoptimierung zurückzuführen ist. Für die duale Netzoptimierung werden in den Ecken des optimierten dualen Netzes die Tangentialebenen bezüglich der Isofläche berechnet. Zur Optimierung eines Eckpunktes des primären Netzes werden dann die angrenzenden Tangentialebenen geschnitten und die neue Position wird durch den Schnittpunkt der Ebenen definiert. In Abbildung 2.6 ist ein zweidimensionales Beispiel für das Prinzip des \mathcal{P} -Operators dargestellt: Aus einem gegebenen Netz wird das optimierte duale Netz berechnet und an dessen Ecken werden Tangentialgeraden bestimmt, deren Schnittpunkte die neue Eckenpositionen definieren. An diesem Beispiel lässt sich auch erkennen, warum eine Optimierung durch alleinige Minimierung des Eckenabstands zur Objektkontur für scharfkantige Objekte unzureichende Ergebnisse liefert. Wenn die Ecken des Netzes bereits auf der Kontur oder in deren unmittelbarer Nähe liegen, so wird noch nicht erschlossenes Volumen durch die Minimierung des Eckenabstands nicht ausgefüllt. Die Energiefunktion solcher Verfahren muss in jedem Fall um weitere Terme erweitert werden, wenn die im Beispiel gezeigte, vom gegebenen Netz nicht erfasste Ecke modelliert werden soll.

Im Dreidimensionalen wird der Ebenenschnitt durch ein System von drei linearen Gleichungen beschrieben. Liegt die ursprüngliche Ecke auf einer nahezu ebenen Fläche oder an einer scharfen Kante, so kann die entsprechende Matrix schlecht konditioniert sein. Aus diesem Grund wird die Lösung des Gleichungssystems durch Anwendung einer Singulärwertzerlegung realisiert (siehe z. B. [Press et al. '92]) und Singulärwerte, die unterhalb einer gegebenen Schwelle liegen, werden unterdrückt, also auf Null gesetzt. In [Ohtake & Belyaev '02] wird gezeigt, dass dieser Schwellwert als

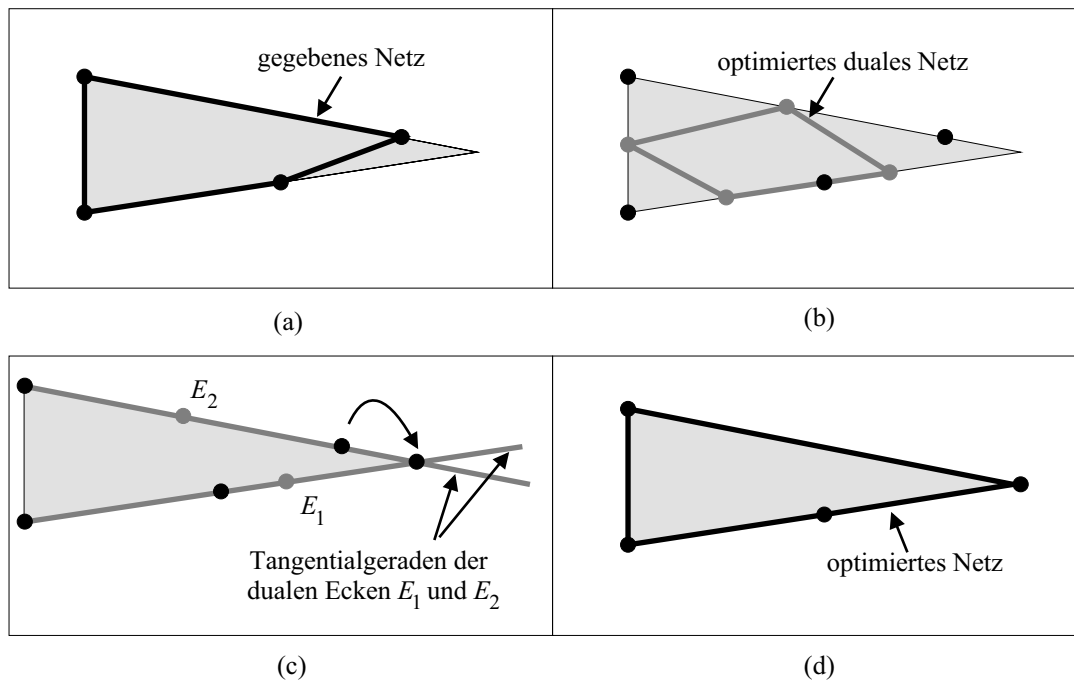


Abbildung 2.6: Netzoptimierung mit dem \mathcal{P} -Operator für ein zweidimensionales Dreieck. Zu einem gegebenes Netz (a) wird das optimierte duale Netz (b) berechnet. Für die dualen Ecken werden Tangentialgeraden berechnet (c) und deren Schnitte sind die neuen Punktpositionen für das optimierte Netz (d). Eine Minimierung des Punktabstands zur Kontur des Dreiecks würde das gegebene Netz nicht verändern.

kritisch einzuordnen ist. Wird der Schwellwert zu hoch angesetzt, so werden nicht alle Objektmerkmale rekonstruiert, das Netz läuft also nicht in die Ecken und Kanten, die durch die Isofläche beschrieben werden. Wird der Schwellwert jedoch zu niedrig angesetzt, so führt dies zu numerischen Instabilitäten.

Bezüglich dieser Netzoptimierung bleibt ein weiterer Punkt zu besprechen, und zwar die Anforderungen, die das Verfahren an die implizite Funktion stellt. Die Anwendung des \mathcal{Z} -Operators verlangt die Berechnung der Gradienten der impliziten Funktion. Also muss die Funktion differenzierbar sein.⁹ Für radiale Summen mit *thin plate splines* ist dies unproblematisch, da die partiellen Ableitungen überall existieren und analytisch angegeben werden können. Wird die Netzoptimierung jedoch auf eine SDT angewandt, so muss wegen der Approximation der Gradienten und der partiellen Linearität der SDT damit gerechnet werden, dass der \mathcal{Z} -Operator und damit auch die duale Netzoptimierung dünne Objektteile kollabieren lassen, also die obere und untere Begrenzungsfläche zusammenkleben. Versuche haben gezeigt, dass sich dieser Effekt durch eine ausreichend feine Unterteilung der Dreiecke an diesen Stellen vermeiden

⁹In der früheren Arbeit [Ohtake et al. '02] schlägt Ohtake die Verwendung so genannter \mathcal{R} -Funktionen vor, die in [Shapiro & Tsukanov '99] diskutiert werden und entsprechende differentielle Eigenschaften besitzen, aber für die hier vorliegenden Freiformobjekte nicht konstruierbar sind.

lässt. Letztlich bleibt zu erwähnen, dass die Methode von Ohtake und Belyaev eine große Anzahl an Funktionsaufrufen erfordert. Das hat zur Folge, dass die Anwendung der Methode auf rechenintensive Funktionen wie die SDT einen erheblichen Rechenaufwand mit sich bringt.

2.3 Polygonreduktion nach Garland & Heckbert

Eine weitere Technik, von der im nächsten Kapitel Gebrauch gemacht wird, ist die Polygonreduktion aus [Garland & Heckbert '97], für die unter dem Namen QSlim eine frei verfügbare C-Implementierung erhältlich ist.¹⁰ Das Verfahren von Garland und Heckbert kann für adaptive Modelle erweitert werden, was im nächsten Kapitel - nach Definition des adaptiven Modells - gezeigt wird. Zunächst wird kurz die Methode erklärt.

Das Verfahren ist den Kantenkontraktionsmethoden zuzuordnen. Das bedeutet, dass die Anzahl der Polygone bzw. Dreiecke eines Modells reduziert wird, indem Kanten des Modells zu Ecken kontrahiert werden. Dies muss unter weitestmöglichem Erhalt der geometrischen Merkmale geschehen. In erster Instanz sollen also solche Kanten kontrahiert werden, an denen die Oberfläche des Modells möglichst eben ist. In [Garland & Heckbert '97] wird zu diesem Zweck jede Modellkante K_i mit einer quadratischen Fehlerfunktion $Q_i : \mathbb{R}^3 \rightarrow \mathbb{R}$ identifiziert. Diese Fehlerfunktion geht aus den an die Kante angrenzenden Ebenen hervor und beschreibt den Modellierungsfehler, der entsteht, wenn die Kante K_i zu einer Ecke mit der Position \mathbf{x} kontrahiert wird. Für jede Kante K_i wird eine geometrische Position $\tilde{\mathbf{x}}_i$ bestimmt, an der die Fehlerfunktion ihr Minimum annimmt, für die also gilt

$$Q_i(\tilde{\mathbf{x}}_i) = \min_{\mathbf{x} \in \mathbb{R}^3} Q_i(\mathbf{x}). \quad (2.4)$$

Zur Polygonreduktion werden zu diesen geometrischen Positionen die minimalen Fehler

$$\Delta_i := Q_i(\tilde{\mathbf{x}}_i) \quad (2.5)$$

für alle Kanten K_i des Modells berechnet. Anschließend werden die Kanten entsprechend der minimalen Fehler Δ_i absteigend sortiert. In der Reihenfolge dieser Sortierung werden nun so lange Kanten kontrahiert (und die Fehlerfunktionen der angrenzenden Dreiecke entsprechend aktualisiert), bis die durch die Anwendung vorgegebene Anzahl an Dreiecken erreicht ist.

¹⁰Eine Übersicht über weitere Polygonreduktionsverfahren findet man in [Heckbert & Garland '97].

Kapitel 3

Das adaptive Modell

Dieses Kapitel beschreibt das entwickelte Verfahren zur automatischen Konstruktion adaptiver Modelle aus einer gegebenen Menge von Basismodellen. Zunächst wird die formale Definition angegeben und im Anschluss die Konstruktion motiviert und erläutert. In einem abschließenden Abschnitt werden Ergebnisse präsentiert.

3.1 Definition des adaptiven Modells

Nach den Vorbereitungen im vorangegangenen Kapitel kann nun das adaptive Modell definiert werden. Für eine Menge von Basismodellen $\mathcal{M}_0, \dots, \mathcal{M}_N$ ist das adaptive Modell durch

$$\left(\mathcal{T}_0, \sum_{n=0}^N \alpha_n \mathbf{g}_n\right) \quad \text{mit } \alpha_n \in \mathbb{R} \text{ und } \sum_{n=0}^N \alpha_n = 1 \quad (3.1)$$

gegeben, wobei \mathcal{T}_0 eine Netztopologie, die so genannte Meta-Topologie, bezeichnet und \mathbf{g}_n eine geometrische Realisierung, so dass das Tupel $(\mathcal{T}_0, \mathbf{g}_n)$ das Basismodell \mathcal{M}_n darstellt, $n = 0, \dots, N$. Abbildung 3.1 zeigt ein solches Modell, das aus drei Basismodellen konstruiert ist. Dieses generische Modell ist kontinuierlich und kann folglich unendlich viele Objekte darstellen, die durch wenige Parameter kodiert sind. Die Bedingung $\sum_{n=0}^N \alpha_n = 1$ stellt eine einheitliche Skalierung der Objekte sicher. Sie verhindert, dass zu einer Ausprägung auch deren Vielfaches enthalten sein kann [Blanz '00].

Andererseits ist diese Kodierung durch eventuelle lineare Abhängigkeiten der geometrischen Realisierung nicht eindeutig. Hier schafft die Anwendung einer Hauptkomponentenanalyse (siehe z. B. [Jolliffe '86]) Abhilfe, die das Modell orthogonalisiert. Für diese Orthogonalisierung wird zunächst die mittlere geometrische Realisierung

$$\bar{\mathbf{g}} = \frac{1}{N+1} \sum_{n_0}^N \mathbf{g}_n \quad (3.2)$$

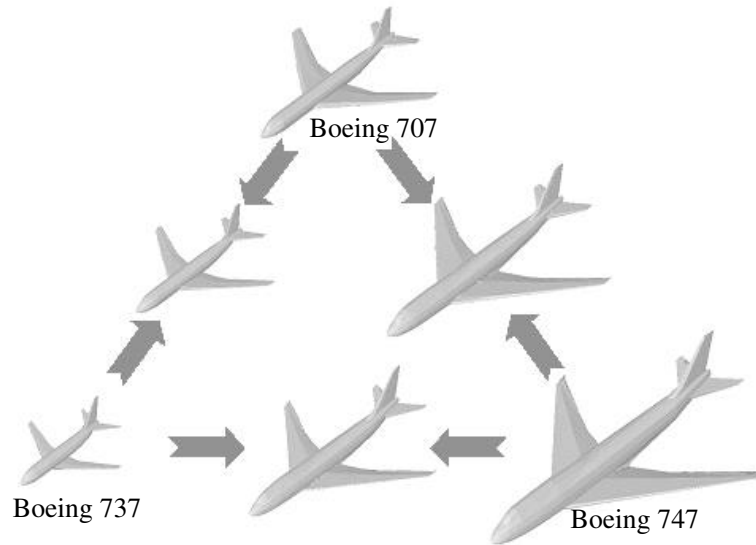


Abbildung 3.1: Die Realisierungen eines adaptiven Modells, das aus drei Basismodellen berechnet wurde. Die Ecken des Dreiecks zeigen die Repräsentationen der Basismodelle und die Modelle auf den Kanten entsprechende Zwischenformen.

berechnet und die Realisierungen werden durch $\Delta \mathbf{g}_n := \mathbf{g}_n - \bar{\mathbf{g}}$ für $n = 0, \dots, N$ zentriert. Die Diagonalisierung der Kovarianzmatrix

$$C = \frac{1}{N+1} (\Delta \mathbf{g}_0, \dots, \Delta \mathbf{g}_N) (\Delta \mathbf{g}_0, \dots, \Delta \mathbf{g}_N)^T \quad (3.3)$$

liefert $N + 1$ orthogonale Eigenvektoren $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_N$. Darüber hinaus liefert die Hauptkomponentenanalyse auch die Eigenwerte $\lambda_0, \lambda_1, \dots, \lambda_N$ der Kovarianzmatrix, die die Variation der geometrischen Realisierungen längs der Eigenvektoren beschreiben. Durch die Zentrierung sind die $\Delta \mathbf{g}_0, \dots, \Delta \mathbf{g}_N$ linear abhängig und es ist mindestens ein Eigenwert gleich Null und der entsprechende Eigenvektor demzufolge nicht relevant. Insgesamt sind also höchstens N Eigenvektoren bzw. Eigenwerte von Bedeutung. Die Indizierung der relevanten Eigenvektoren wird im Folgenden immer so vorausgesetzt, dass für die entsprechenden Eigenwerte $\lambda_1 \geq \dots \geq \lambda_N$ gilt. In dieser Sortierung besitzen die geometrischen Realisierungen längs \mathbf{e}_1 die größte und längs \mathbf{e}_N die kleinste Variation.

Mit Hilfe der Eigenvektoren wird die orthogonale Version des adaptiven Modells definiert:

$$\mathcal{M}_{\delta_1, \dots, \delta_N} = (\mathcal{I}_0, \bar{\mathbf{g}} + \sum_{n=1}^N \delta_n \mathbf{e}_n) \text{ mit } \delta_n \in \mathbb{R}. \quad (3.4)$$

Die orthogonale Version des adaptiven Modells ist folglich durch einen (höchstens) N -dimensionalen Vektor $\tilde{\theta} = (\delta_1, \dots, \delta_N)^T$ parametrisiert, besitzt also N Freiheitsgrade.

Unterstellt man den zentrierten geometrischen Realisierungen eine multivariate Normalverteilung mit der Kovarianzmatrix C , so gilt für die Dichte

$$\begin{aligned} p(\tilde{\theta}) &= \frac{1}{\sqrt{(2\pi)^N |C|}} \exp \left[-0.5 \sum_{n=1}^N \frac{\delta_n^2}{\lambda_n} \right] \\ &= \frac{1}{\sqrt{(2\pi)^N |C|}} \prod_{n=1}^N \exp \left[-0.5 \frac{\delta_n^2}{\lambda_n} \right], \end{aligned}$$

wobei $|C| = \prod_{n=1}^N \lambda_n$ die Determinante der Kovarianzmatrix bezeichnet. Unter Vernachlässigung der Skalierung ist die Dichte also ein Produkt eindimensionaler Normalverteilungen mit den Varianzen $\sigma_n^2 := \lambda_n$.

Bezüglich der Annahme einer Normalverteilung muss jedoch bedacht werden, dass diese einen stochastischen Prozess unterstellt, durch den die „Stichprobe“ der geometrischen Realisierungen entstanden sein soll. Da die Form technischer Objekte nur in Ausnahmefällen durch einen stochastischen Prozess modellierbar ist, kann man kaum davon ausgehen, dass diese Voraussetzung gegeben ist. Zudem ist die hier vorliegende Menge an Basismodellen zu klein, um Rückschlüsse auf die zugrundeliegende Verteilung eines stochastischen Prozesses zu ermöglichen. Die Interpretation der Eigenwerte der Kovarianzmatrix als Varianzen ist also im hier gegebenen Fall in Frage zu stellen.

Dennoch wird die Terminologie beibehalten und die Eigenwerte werden als Varianzen bzw. deren Wurzeln als Standardabweichungen bezeichnet. Immerhin beschreiben diese Werte die Variationen der Basismodelle längs der Eigenvektoren und liefern damit wichtige Informationen für den Modellanpassungsprozess. In Abbildung 3.2 wird beispielsweise die Variation eines Flugzeugmodells längs dreier Eigenvektoren gezeigt. Dieses Beispiel verdeutlicht zum einen, dass adaptive Flugzeugmodelle erhebliche Elastizität bezüglich der Objektgröße besitzen können. Zum anderen ist an diesem Beispiel erkennbar, dass die Modellparameter ab einem gewissen Index für die Modellanpassung keine Relevanz mehr besitzen, da sie die Form nur noch innerhalb des Bildpunktbereichs beeinflussen.

Bei der Durchführung der Hauptkomponentenanalyse ist zu berücksichtigen, dass die zentrierten geometrischen Realisierungen jeweils eine Dimension von

$$3 \times \text{Anzahl Ecken der Meta-Topologie}$$

besitzen. Berücksichtigt man nun, dass z. B. für die Modellierung von Flugzeugen adaptive Modelle mit mindestens 1.000 Ecken herangezogen werden, so bedeutet dies, dass Kovarianzmatrizen von der Dimension 3.000×3.000 und mehr zu diagonalisieren sind. Im Anhang A werden - neben dem Standardverfahren zur Hauptkomponentenanalyse - zwei weitere Verfahren skizziert, mit deren Hilfe sich die Dimension der auftretenden Matrizen erheblich reduzieren lässt.

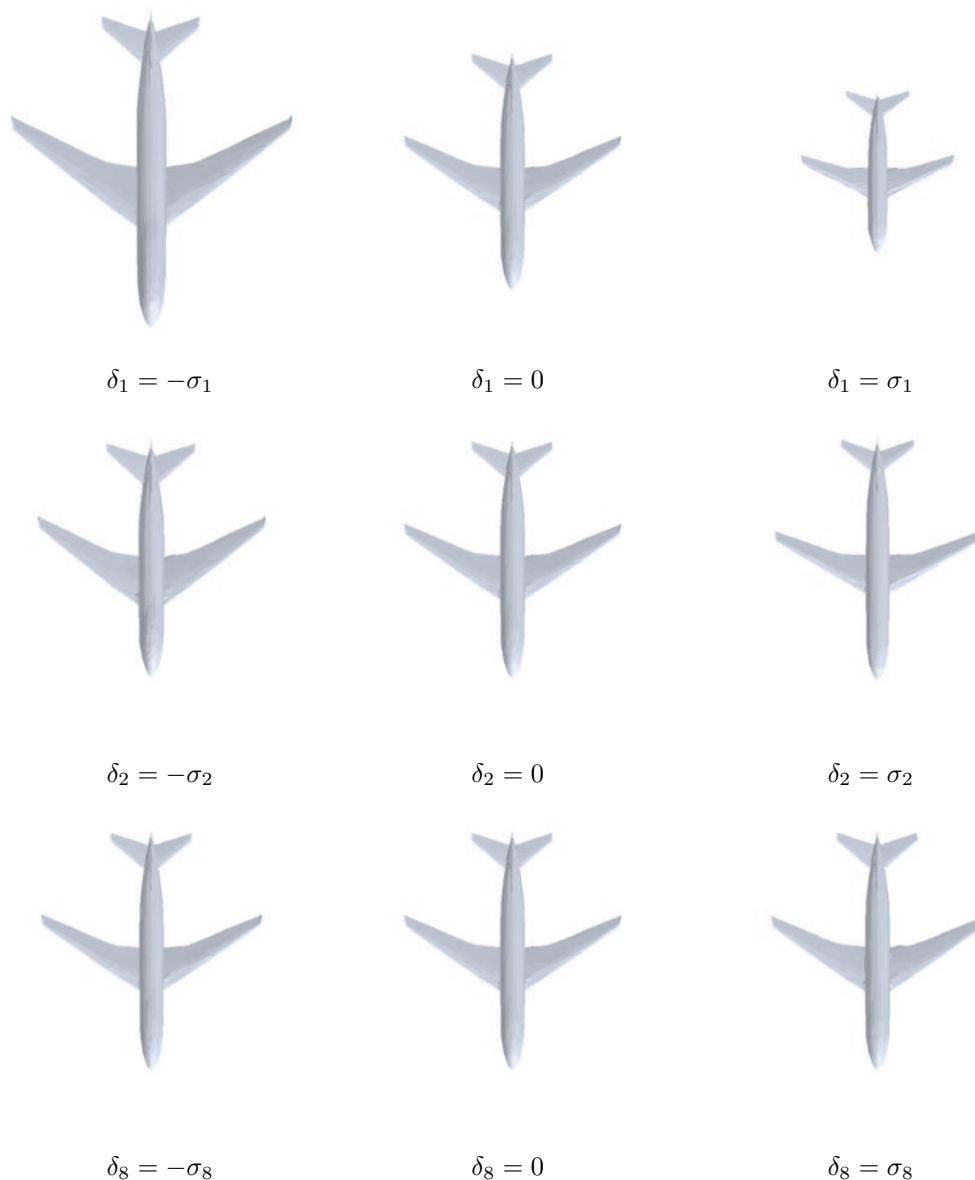


Abbildung 3.2: Variation längs dreier Eigenvektoren für ein orthogonales adaptives Flugzeugmodell (konstruiert aus 15 Basismodellen). Alle nicht angegebenen Parameter des adaptiven Modells sind auf Null gesetzt und die Größen σ_n bezeichnen die Standardabweichungen bezüglich des n -ten Eigenvektors. Die Abbildung zeigt, dass der erste Modellparameter hauptsächlich die Objektgröße verändert, während der zweite die Größe bereits kaum mehr beeinflusst. Der achte Parameter zeigt kaum mehr sichtbare Variationen.

3.2 Konstruktion des adaptiven Modells

Zur Konstruktion eines adaptiven Modells muss zum einen eine Meta-Topologie \mathcal{T}_0 berechnet werden und zum anderen $N + 1$ geometrische Realisierungen \mathbf{g}_n , so dass

das Tupel $(\mathcal{T}_0, \mathbf{g}_n)$ das Basismodell \mathcal{M}_n darstellt, $n = 0, \dots, N$. Diese Konstruktion wird in zwei Schritten vorgenommen:

1. Die Konstruktion eines Meta-Modells $(\mathcal{T}_0, \tilde{\mathbf{g}})$ in einer zunächst nicht näher präzisierten geometrischen Realisierung $\tilde{\mathbf{g}}$.
2. Hinzufügen der Basismodelle $\mathcal{M}_0, \dots, \mathcal{M}_N$ durch Berechnung der geometrischen Realisierungen $\mathbf{g}_0, \dots, \mathbf{g}_N$.

So weit die formale Definition. Es stellt sich nun die Frage nach der Bedeutung dieser Konstruktionsschritte. Eine Meta-Topologie bereitzustellen bedeutet, ein Netz zu konstruieren, mit dem alle Basismodelle dargestellt werden können. Die Forderung, dass jedes der Basismodelle durch ein und dasselbe Netz dargestellt werden soll, hat zur Folge, dass nur topologisch äquivalente Objekte¹ durch ein und dasselbe adaptive Modell modelliert werden können. Ein Beispiel für die Verletzung dieser Forderung sind die Objekte Kugel und Torus: Beide bestehen aus einer Zusammenhangskomponente, besitzen aber kein bzw. ein Loch und sind demnach nicht topologisch äquivalent. Für die hier vorliegende Aufgabe hat diese Forderung jedoch kaum Auswirkungen, da die Objektklassen der Flugzeuge, Fahrzeuge etc. in der Regel keine Löcher besitzen. Auch setzen sie sich - geometrisch betrachtet - nicht aus mehreren unverbundenen Komponenten zusammen.

Das Korrespondenzproblem

Die größere Schwierigkeit bei der Konstruktion des Meta-Modells ist die geometrische Realisierung desselben, da sie die Grundlage für den zweiten Konstruktionsschritt ist, nämlich die Aufnahme der Basismodelle in das adaptive Modell. Dieser Schritt kann nur dann gelingen, wenn für das Meta-Modell eine geometrische Realisierung vorliegt, die bereits nah genug an den Basismodellen ist. Nah genug deshalb, weil die zu lösende Aufgabe bedeutet, die Ecken des Meta-Modells auf die Oberfläche des jeweiligen Basismodells zu verschieben und dabei korrespondierende Oberflächenpunkte zu berücksichtigen.²

Die Berücksichtigung dieser Punktkorrespondenzen wird mit zunehmendem Abstand zwischen Meta-Modell und Basismodell immer schwieriger. Für Flugzeuge ist es z. B. eine erhebliche Herausforderung, die Punkte, die auf der oberen Hälfte der Tragfläche liegen, auch auf der oberen Hälfte der Tragfläche des Zielmodells unterzubringen und nicht mit der Unterseite zu verkleben. Auch muss dafür gesorgt werden, dass die geometrische Realisierung des Meta-Modells bereits alle markanten Teile der Objektklasse modelliert. Alles in allem ist festzustellen, dass der geometrischen Realisierung

¹Topologische Äquivalenz ist z. B. in [Jänich '94] definiert und bedeutet im Wesentlichen, dass nur Objekte mit derselben Anzahl zusammenhängender Komponenten und derselben Anzahl an Löchern dargestellt werden können.

²Das Korrespondenzproblem wird in [Shelton '00] wie folgt definiert: „given two instances of a class of objects, find a relation from parts of one object to 'corresponding' parts on the other object“.

bei der Konstruktion des Meta-Modells eine erheblich größere Bedeutung zukommt als der Meta-Topologie.

Doch zurück zum Korrespondenzproblem, das in [Shelton '00] als *ill-posed problem*³ bezeichnet wird. Die Punktkorrespondenzen sind insbesondere dann undefiniert, wenn Doppeldeutigkeiten existieren oder wenn Objektpunkte auf einer Instanz vorhanden sind, auf der anderen jedoch nicht. Letzteres hat erhebliche Auswirkungen auf die hier vorliegende Anwendung. Betrachtet man beispielsweise zwei Flugzeuge, eines mit zwei und das andere mit vier Triebwerken, so existieren für die zwei überzähligen Triebwerke keine korrespondierenden Objektpunkte. Wird die Punktkorrespondenz umgangen, indem zum Beispiel die Oberfläche der zwei überzähligen Triebwerke des Vierstrahlers mit der Tragfläche des Zweistrahlers identifiziert wird, so muss man sich die resultierenden Zwischenformen vor Augen halten: Die Interpolationen würden einen Zweistrahler mit zwei weiteren, partiell ausgeprägten Triebwerken liefern. Abbildung 3.3 zeigt ein vergleichbares Beispiel für die Interpolation zwischen zwei durch implizite Funktionen modellierten Flugzeugen mit unterschiedlicher Triebwerksanzahl und -anordnung. Die Zwischenform modelliert mehr oder weniger stark ausgeprägte Ansätze aller sechs Triebwerke.



Abbildung 3.3: Diskrete Merkmale in adaptiven Modellen. In (a) ist ein skaliertes Schattenriss einer Boeing 747 und in (c) der einer Boeing 727 abgebildet. Für die Berechnung der Zwischenform in (b) wurden die beiden Binärbilder aus (a) und (c) mit jeweils einer zweidimensionalen SDT identifiziert, diese wurden interpoliert und der Interpolant anschließend binarisiert. Diese Zwischenform modelliert alle sechs Triebwerke der beiden Flugzeuge in mehr oder weniger stark ausgeprägter Form.

Konsequenz für adaptive Modelle

Die Konsequenz für adaptive Modelle kann also nur sein, für Fälle dieser Art entweder zwei verschiedene adaptive Modelle zu konstruieren oder aber auf die Modellierung

³Unsachgemäßes (inkorrekt gestelltes) Problem [Eisenreich & Sube '94].

der Triebwerke zu verzichten. An dieser Stelle muss betont werden, dass ein solcher Verzicht auf die Modellierung diskreter Objektmerkmale nur dann von Bedeutung ist, wenn diese Merkmale für die Stabilisierung der Modellanpassung benötigt werden. Sollten diese Merkmale hingegen für die Typerkennung relevant sein, so kann die Extraktion dieser Merkmale im Anschluss an die Modellanpassung erfolgen. Dieses Vorgehen ist empfehlenswert, da nach der Modellanpassung zum einen eine Segmentierung des Eingabebildes vorliegt und zum anderen die Objektlage im Bildbereich bekannt ist. Durch dieses Wissen kann die Extraktion diskreter Merkmale erheblich erleichtert werden.

Auch für das Beispiel von T- und Kreuzleitwerken von Verkehrsflugzeugen muss eine Überlegung dieser Art angestellt werden. Da das Leitwerk im einen Fall zwischen Rumpf und Höhenruder liegt, im anderen Fall aber das Höhenruder zwischen Rumpf und Leitwerk, würden Punktkorrespondenzen zu Überlappungen führen und unsinnige Zwischenformen produzieren. Auch hier gilt es wieder, zwei verschiedene adaptive Modelle - eines für jede Objektklasse - zur Verfügung zu stellen. Völlig analog ist die Modellierung von Merkmalen einzuordnen, die sich nicht geometrisch abheben, sondern durch ihre Farbwerte (z. B. Fenster, Scheinwerfer etc.). Zusammenfassend ist festzuhalten, dass nur solche Objektmerkmale durch ein adaptives Modell dargestellt werden können, die für alle Basismodelle in derselben Anzahl und derselben geometrischen Anordnung vorhanden sind.

Techniken zur Korrespondenzberechnung

Wie wird nun die Korrespondenz berechnet? [Blanz & Vetter '99] lösen die Aufgabe durch die Projektion der Gesichter auf die Bildebene und berechnen die Korrespondenz dann im Zweidimensionalen. Dieses Verfahren kommt für die hier vorliegenden Objektklassen nicht in Frage, da die Oberfläche von Flugzeugen nicht sinnvoll auf eine Ebene zu projizieren ist. In [Shelton '00] wird eine Energieminimierungsmethode vorgestellt, die jedoch auf der Berechnung des minimalen Punktabstands zur Oberfläche des Zielmodells basiert und daher zur Modellierung dünner Objektteile - wie z. B. der Tragflächen - nicht geeignet ist. Auch scharfkantige Objektmerkmale werden durch solche Ansätze nicht gut modelliert, da der minimale Punktabstand die Ecken des Modells auf den nächstliegenden Oberflächenpunkt verschiebt und dadurch die Objektkanten nicht ausfüllt (vgl. Abbildung 2.6).

Verwandte Algorithmen findet man unter dem Begriff der Metamorphose in Veröffentlichungen aus dem Bereich der Computergrafik. Die Aufgabe einer Metamorphose besteht darin, kontinuierliche Übergänge von einem Start- zu einem Zielobjekt zu definieren. Diese Aufgabe entspricht Teilen der hier gestellten Aufgabe der Konstruktion eines adaptiven Modells. Wenngleich die meisten dieser Verfahren nur zwei Objekte berücksichtigen, so lassen sie sich in der Regel auf die Verwendung mehrerer Basismodelle übertragen. Eine Ausnahme bilden diejenigen Verfahren, die die Korrespondenz durch eine Vereinigung der Netze von Start- und Zielobjekt berechnen, da dies bei mehreren Modellen zu einer erheblichen Anzahl von Dreiecken führen würde.

Eine gute Übersicht über diverse Techniken zur Konstruktion von Metamorphosen bietet die Veröffentlichung [Lazarus & Verroust '98]. Dort werden die Verfahren in erster Instanz in volumenbasierte Techniken und oberflächenbasierte Techniken unterteilt. Rein volumenbasierte Techniken [He et al. '94, Pasko & Savchenko '95, Larios et al. '95, Turk & O'Brian '99, Morse et al. '01] zeichnen sich dadurch aus, dass sie keine topologische Äquivalenz der Basismodelle erfordern. Für die hier vorliegende Aufgabe kommen sie allerdings nicht in Frage, da diskrete Verfahren aus genannten Gründen auszuschließen sind und die kontinuierlichen Verfahren die Zwischenformen als implizite Funktionen zur Verfügung stellen. Diese impliziten Darstellungen sind im Hinblick auf die Modellanpassung jedoch nicht geeignet, da die Zwischenformen zu Laufzeit mehrfach polygonisiert oder durch einen *ray-tracer* dargestellt werden müssten, um die benötigten Modellmerkmale extrahieren zu können. Beides führt zu erheblichen Rechenzeiten der Modellanpassung und ist aus diesem Grund nicht empfehlenswert.

Die in der Berechnung der Zwischenformen effizienteren Techniken der oberflächenbasierten Verfahren werden weiter unterteilt, in projektions- und korrespondenzbasierte Techniken. Die projektionsbasierten Techniken setzen die sternförmige Projizierbarkeit der Objekte auf eine Kugel [Kent et al. '92], die sternförmige Projizierbarkeit bezüglich einer dreidimensionalen Achse [Lazarus & Verroust '97] oder die Projizierbarkeit auf die Einheitskugel [Kanai et al. '97, Kanai et al. '00, Gregory et al. '98] voraus. Zu den korrespondenzbasierten Verfahren ist die Methode [Parent '92] zu zählen, in der eine Fläche-zu-Fläche Korrespondenz definiert wird. In [Lee et al. '99] wird die Korrespondenz hierarchisch durch eine Dreiecksreduktion erklärt.

Der Lösungsansatz

Der in dieser Arbeit eingeschlagene Weg ist eine Kombination aus beiden Kategorien, der oberflächen- und volumenbasierten Darstellungen. Die volumenbasierte Darstellung wird unter anderem dazu eingesetzt, die oberflächenbasierte Korrespondenz zwischen dem Meta-Modell und dem Basismodell zu unterstützen, indem die Ecken des Meta-Modells langsam über die Isoflächen interpolierter impliziter Darstellungen in Richtung Basismodell geschoben werden. Auf diese Weise erfolgt die Berechnung der Punktkorrespondenzen in kleinen Teilschritten. Durch die Berücksichtigung der Gradienten der Interpolanten wird bei der Netzoptimierung zudem verhindert, dass dünne Objektteile zusammenkleben. Eine weitere Anwendung der volumenbasierten Techniken ist die Konstruktion des Meta-Modells selbst, die im nächsten Abschnitt beschrieben wird.

Abschließend sei betont, dass die Konstruktion des adaptiven Modells von der Tatsache Gebrauch macht, dass jedes Basismodell mit einer impliziten Funktion identifiziert werden kann. Um welche spezielle Form von impliziten Funktionen es sich dabei handelt, ist sekundär, solange die Netzoptimierung von [Ohtake et al. '02] für diese Funktionstypen anwendbar ist und die Basismodelle in der für die Anwendung erforderlichen Genauigkeit repräsentierbar sind.

3.2.1 Konstruktion des Meta-Modells

Die Konstruktion des Meta-Modells basiert auf einer Abwandlung der klassischen, volumenbasierten Metamorphose, die wie folgt durchgeführt werden könnte (siehe z. B. [Pasko & Savchenko '95]): Zunächst kann jedes Basismodell \mathcal{M}_n mit einer impliziten Funktion f_n identifiziert werden und aus diesen Funktionen die gemittelte Funktion

$$\frac{1}{N+1} \sum_{n=0}^N f_n$$

bereit gestellt werden. Die Anwendung des in Abschnitt 2.2.1 beschriebenen Polygonisierers auf diese Funktion kann schließlich das Meta-Modell liefern. Ein solches Vorgehen hätte jedoch einen erheblichen Verlust an Volumen zur Folge, da die Summation in etwa vergleichbar ist mit einer Durchschnittsbildung über die Basismodelle. Insbesondere besteht eine große Wahrscheinlichkeit, dass das Volumen markanter Objektteile von dieser Funktionsinterpolation nicht erfasst wird, wenn sich diese Objektteile räumlich nicht überdecken, wodurch wiederum die Berechnung der Punktkorrespondenzen erschwert wird.

Aus diesem Grund werden die Basismodelle durch eine vorab berechnete Verzerrung grob zur Deckung gebracht. Für Objekttypen, die ausreichend Volumen besitzen, wie z. B. Personenkraftwagen, wäre eine Skalierung der Objekte auf eine einheitliche Ausdehnung völlig ausreichend, um Volumenverlust zu vermeiden. Bei Flugzeugen jedoch muss dafür gesorgt werden, dass sich die Tragflächen, Leitwerke und Höhenruder der einzelnen Basismodelle bestmöglich überdecken. Dies kann weder durch eine Skalierung noch durch eine andere rigide Transformation erreicht werden. Aus diesem Grund wird eine elastische Verzerrung in der Art von [Cohen-Or et al. '98] durchgeführt, um die Objekte zur Deckung zu bringen.

Die Verzerrung nach Cohen-Or et al.

Für die Verzerrung nach [Cohen-Or et al. '98] werden für jedes der Basismodelle $\mathcal{M}_0, \dots, \mathcal{M}_N$ eine kleine Anzahl L von korrespondierenden Objektpunkten $\mathbf{x}_1^n, \dots, \mathbf{x}_L^n$, $n = 0, \dots, N$ benötigt (vgl. Abbildung 3.4 mit $L = 16$), die im Vorfeld der Konstruktion des adaptiven Modells interaktiv festgelegt werden müssen. Zur Definition dieser so genannten Ankerpunkte muss keine übertriebene Sorgfalt an den Tag gelegt werden, da die Verzerrung lediglich eine grobe Vorverarbeitungsstufe darstellt. Für die im Rahmen dieser Arbeit berechneten adaptiven Modelle konnten die Ankerpunkte mit Hilfe einer grafischen Oberfläche in weniger als fünf Minuten pro Basismodell festgelegt werden.

Die in [Cohen-Or et al. '98] beschriebene Verzerrung setzt sich aus einem rigiden Teil (Rotation und Translation) und einem elastischen Teil zusammen, wobei letzterer durch radiale Summen realisiert ist. Für den rigiden Teil wird im Rahmen dieser Arbeit auf die Rotation verzichtet, da die Basismodelle alle in derselben Orientierung vorliegen. Stattdessen werden - für jede Achse separat - drei Skalierungen durchgeführt.

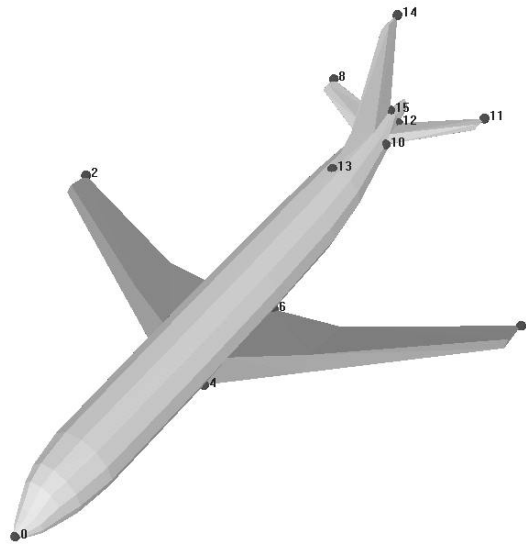


Abbildung 3.4: Für die Verzerrung der Flugzeuge werden auf jedem Basismodell 16 Ankerpunkte definiert.

Zur Konstruktion der Verzerrungen werden zunächst aus den modellspezifischen Ankerpunkten die Schwerpunkte

$$\bar{\mathbf{x}}_l := \frac{1}{N+1} \sum_{n=0}^N \mathbf{x}_l^n, l = 1, \dots, L$$

über alle Basismodelle berechnet. Für jedes Modell \mathcal{M}_n wird dann eine Abbildung $w_n : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ konstruiert, die die modellspezifischen Ankerpunkte auf diese Schwerpunkte abbildet, für die also gilt

$$w_n(\mathbf{x}_l^n) = \bar{\mathbf{x}}_l, l = 1, \dots, L.$$

Hierfür werden zunächst die Parameter

$$\vartheta := (\alpha_x^n, \alpha_y^n, \alpha_z^n, t_x^n, t_y^n, t_z^n)^T$$

der rigiden Abbildung

$$g_n(\mathbf{x}) = \begin{pmatrix} \alpha_x^n & 0 & 0 \\ 0 & \alpha_y^n & 0 \\ 0 & 0 & \alpha_z^n \end{pmatrix} \mathbf{x} + \begin{pmatrix} t_x^n \\ t_y^n \\ t_z^n \end{pmatrix}$$

bestimmt, die durch die Minimierungsaufgabe

$$\min_{\vartheta} \sum_{l=1}^L \|g_n(\mathbf{x}_l^n) - \bar{\mathbf{x}}_l\|^2$$

definiert sind.⁴ Diese Abbildung liefert in der Regel Restfehler, bildet also die modell-spezifischen Ankerpunkte nicht genau auf die Schwerpunkte ab. Zur Kompensation dieser Restfehler wird noch der elastische Teil der Verzerrung angewandt, um die Basismodelle zur Deckung zu bringen. Dieser elastische Teil setzt sich aus drei radialen Summen s_x , s_y und s_z zusammen, eine für jede Koordinatenachse. Für die x -Achse wird nach dem in Abschnitt 2.1.2.1 beschriebenen Verfahren für die Punkte

$$\Delta \mathbf{x}_l^n := \bar{\mathbf{x}}_l - g_n(\mathbf{x}_l^n), l = 1, \dots, L$$

eine radiale Summe s_x berechnet und zwar mit den Stützstellen $g(\mathbf{x}_1^n), \dots, g_n(\mathbf{x}_L^n)$ und den Sollwerten

$$h_l := x\text{-Komponente von } \Delta \mathbf{x}_l^n, l = 1, \dots, L.$$

Auf denselben Stützstellen werden zwei weitere radiale Summen s_y bzw. s_z berechnet, allerdings mit den y - bzw. z -Komponenten der Punkte $\Delta \mathbf{x}_l^n$. Die Verzerrung für das Modell \mathcal{M}_n ist dann durch Hintereinanderausführung der rigiden und elastischen Verzerrung gegeben, also durch die Funktion

$$w_n(\mathbf{x}) := g_n(\mathbf{x}) + \begin{pmatrix} s_x(g_n(\mathbf{x})) \\ s_y(g_n(\mathbf{x})) \\ s_z(g_n(\mathbf{x})) \end{pmatrix}. \quad (3.5)$$

Diese Verzerrung muss im weiteren Konstruktionsprozess rückgängig gemacht werden. Da die Verzerrung, bedingt durch die radialen Summen, nicht invertierbar ist, muss die Rücktransformation approximiert werden. Hierfür wird - auf dieselbe Art und Weise wie oben beschrieben - für jedes Basismodell eine Abbildung \hat{w}_n mit vertauschten Start- und Zielpunkten konstruiert. Die Fehler, die durch diese Approximation der Umkehrung gemacht werden, sind geringfügig (vgl. Abbildung 3.5) und werden im Laufe der Weiterverarbeitung kompensiert.

Das Meta-Modell

Die Konstruktion des Meta-Modells erfolgt entsprechend der grafischen Übersicht in Abbildung 3.6. Zunächst werden die basismodellspezifischen Verzerrungen auf die jeweiligen Basismodelle angewandt. Diese verzerrten Modelle werden mit \mathcal{M}_n^w bezeichnet und die entsprechenden impliziten Funktionen mit f_n^w . Mit diesen Funktionen ist durch Polygonisierung der Isofläche von

$$f = c_0 + \frac{1}{N+1} \sum_{n=0}^N f_n^w \quad (3.6)$$

das Meta-Modell definiert. Ein Beispiel für ein Meta-Modell ist in Abbildung 3.7 (a) abgebildet. Dieses Meta-Modell ist den verzerrten Basismodellen ausreichend ähnlich und bringt keinen Volumenverlust mit sich.

⁴Numerische Verfahren zur Lösung solcher Minimierungsverfahren findet man z.B. in [Press et al. '92].

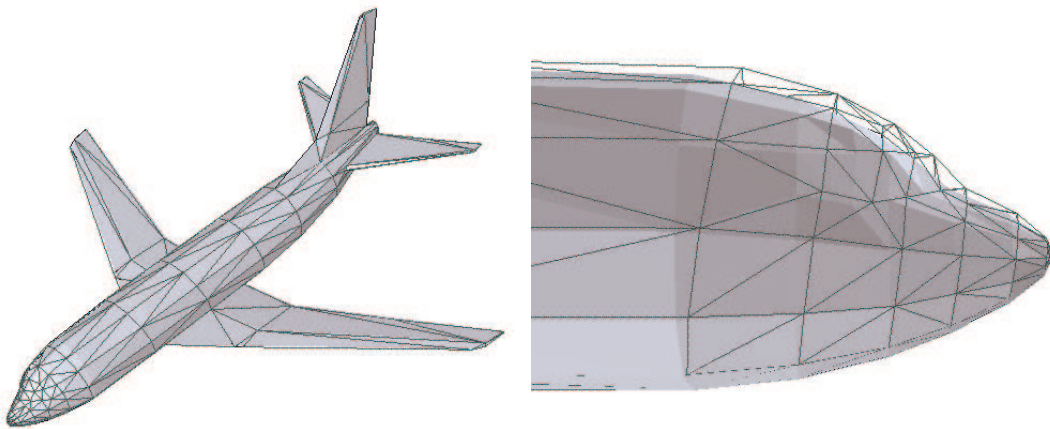


Abbildung 3.5: Fehler der invertierten Verzerrung. Das Basismodell ist in grauer Flächendarstellung dargestellt. Auf das Modell wurde zunächst die Verzerrung w angewandt und anschließend die approximierte inverse Verzerrung \hat{w} . Das Ergebnis ist in schwarzer Kantendarstellung eingezeichnet.

Die Addition der Konstanten c_0 gewährleistet, dass auch dünne Objektteile, die innerhalb des Zellrasters des Polygonisierers liegen, modelliert werden. Sie wird auf die Hälfte der Zellgröße des Polygonisierers gesetzt. Dadurch ist garantiert, dass das durch die implizite Funktion beschriebene Volumen vollständig vom Polygonisierer erfasst wird. Dieses Vorgehen entspricht einer morphologischen Operation, nämlich der Öffnung des Volumens um eine Einheitskugel mit Radius c_0 . Analog zur Verzerrung wird diese Operation in der letzten Iteration des Konstruktionsprozesses rückgängig gemacht und verursacht daher keine Fehler in der Darstellung der Basismodelle durch das adaptive Modell.

Die Zellgröße ist in Abhängigkeit von der vorliegenden Objektklasse zu wählen. Dies ist aber weniger kritisch, da die Ausgabe des Polygonisierers, wie im Folgenden beschrieben wird, noch weiter bearbeitet wird und schlecht gewählte Schwellen durch diese Weiterverarbeitung kompensiert werden. Für die im Rahmen dieser Arbeit konstruierten Flugzeugmodelle ist die Zellgröße auf 0,5 Meter festgesetzt, wodurch die resultierenden Netze etwa 15.000 Dreiecke besitzen.

Adaption an die lokale Krümmung

Die Berechnung der geometrischen Realisierungen wird unter Verwendung der Netzoptimierung aus [Ohtake & Belyaev '02] durchgeführt. Die von den Autoren vorgeschlagene Zusammenstellung der Operatoren beinhaltet die Verfeinerung der Netztopologie durch Dreiecksunterteilung, um starke lokale Krümmungen modellieren zu können. Grundsätzlich wäre es möglich, das Netz des adaptiven Modells während der

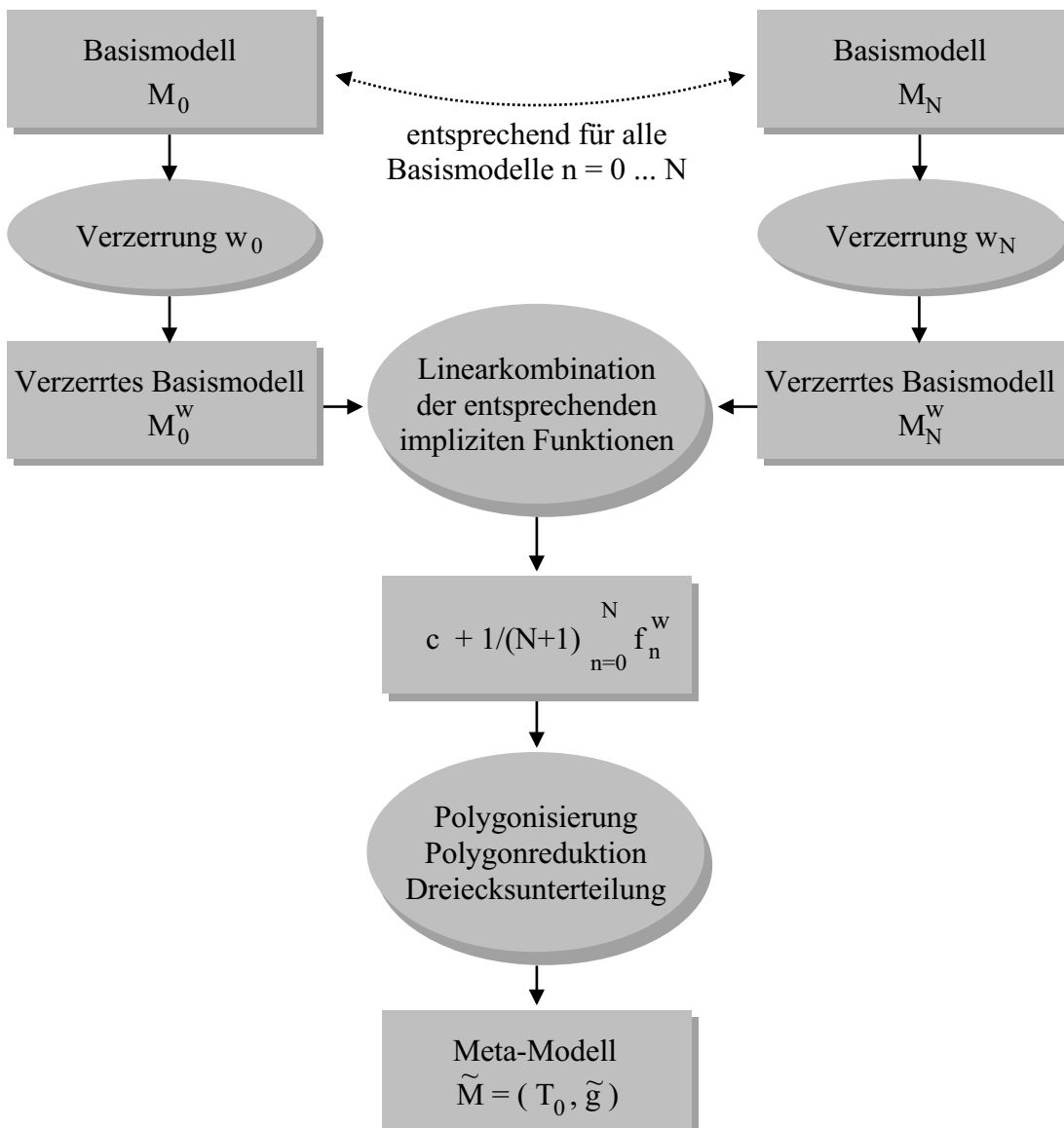


Abbildung 3.6: Die Grafik zeigt eine Übersicht über die Konstruktion des Meta-Modells. Aus den Basismodellen wird eine implizite Funktion abgeleitet, deren Isofläche polygonisiert wird. Die Ausgabe des Polygonisierers wird durch Polygonreduktion und Dreiecksunterteilung an die lokalen Krümmungen angepasst.

Berechnung der geometrischen Realisierungen noch zu verändern. Dies hätte jedoch zur Folge, dass neu eingefügte Ecken für alle bereits berechneten geometrischen Realisierungen berücksichtigt werden müssen. Einfacher ist es, die Netztopologie bereits im Vorfeld an die lokalen Krümmungen anzupassen. Möglich ist dies, da die Geometrie des Meta-Modells durch die oben beschriebene Konstruktion bereits über die entscheidenden geometrischen Merkmale verfügt. Insbesondere stimmen die Bereiche mit schwacher und starker Krümmung bereits mit den Basismodellen überein.

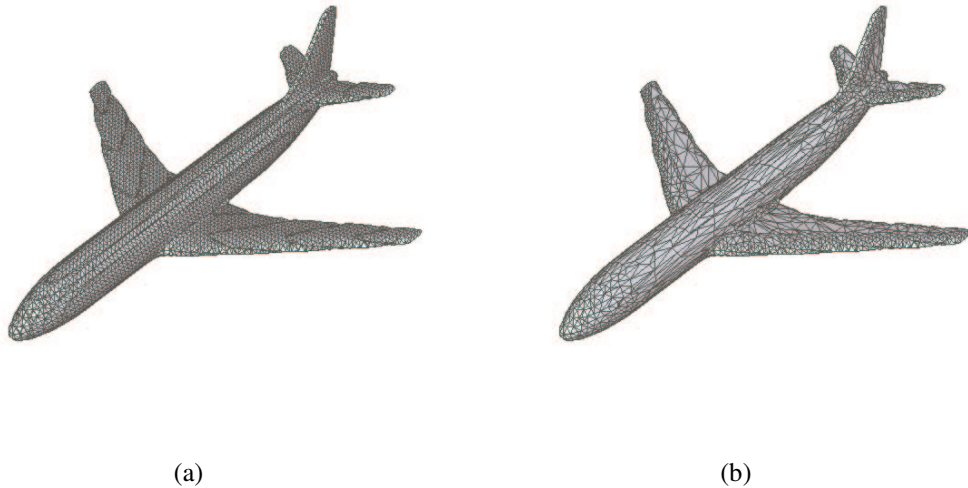


Abbildung 3.7: Abbildung (a) zeigt die Ausgabe des Polygonisierers für die Funktion f aus Gleichung (3.6) und zwar mit einer deutlich sichtbaren rasterförmigen Triangulierung. Abbildung (b) zeigt die Anwendung des Polygonreduktionsverfahrens auf das Modell aus (a), durch das die Anzahl der Dreiecke auf den wenig gekrümmten Flächen von circa 15.000 auf 5.000 reduziert wurde.

Betrachtet man die Ausgabe des Polygonisierers auf eine Funktion nach Gleichung 3.6 in Abbildung 3.7 (a), so stellt man fest, dass dieser ein Dreiecksnetz entsprechend der zugrunde liegenden Zelle liefert, also mehr oder weniger eine rasterförmige Triangulierung der Isofläche. Um die Triangulierung nun an die lokalen Krümmungen anzupassen, wird zunächst eine Polygonreduktion nach [Garland & Heckbert '97] durchgeführt. Dadurch werden die Dreiecke auf flachen Oberflächenstücken zusammengefasst und unnötiger Rechenaufwand wird vermieden. Anschließend werden diejenigen Dreiecke, an denen die implizite Funktion eine starke Krümmung aufweist, in der Art des S -Operators 1:4 unterteilt. Die Entscheidung, ob ein Dreieck unterteilt wird oder nicht, wird hier - im Gegensatz zu Ohtake und Belyaev - nicht in Abhängigkeit von der jeweiligen Dreiecksgröße getroffen, sondern durch die (approximierte) maximale Winkelabweichung der Dreiecksnormalen zu den Gradienten der impliziten Funktion. Formal lautet das Unterteilungskriterium

$$\max_{\mathbf{c} \in \mathcal{C}(T)} \mathbf{n}^T \frac{\nabla f(\mathbf{c})}{\|\nabla f(\mathbf{c})\|} > \cos(\alpha), \quad (3.7)$$

wobei \mathbf{n} die Normale des Dreiecks T bezeichnet und $\mathcal{C}(T)$ die Menge der Schwerpunkte der vier Subdreiecke. Das resultierende Modell definiert dann das Meta-Modell

$$\tilde{\mathcal{M}} = (\mathcal{T}_0, \tilde{\mathbf{g}}). \quad (3.8)$$

Das Unterteilungskriterium aus Gleichung (3.7) liefert im Gegensatz zu dem Unterteilungskriterium aus [Ohtake & Belyaev '02] eine anschauliche Schwellwertangabe in

Form der Winkeldifferenz α zwischen den Oberflächennormalen des Modells und den Gradienten Isofläche. Dies ist vorteilhaft, da im Rahmen dieser Arbeit durchgeführte Experimente gezeigt haben, dass die Dreiecksunterteilung in jedem Fall abhängig vom vorliegenden Objekttyp definiert werden sollte. Insbesondere für Objektklassen, die dünne Objektteile besitzen, darf die Schwelle nicht zu hoch angesetzt werden. Für Flugzeugmodelle werden erst mit einer Schwelle von $\alpha = \pi/8$ gute Ergebnisse erzielt. Für Fahrzeugmodelle hingegen hat sich bereits eine Schwelle von $\alpha = \pi/4$ als ausreichend gezeigt.

Durch die Adaption des Meta-Modells an die lokale Krümmung wird die Anzahl der Dreiecke modifiziert. Für die im Rahmen dieser Arbeit berechneten Flugzeugmodelle liefert der Polygonisierer etwa 15.000 Dreiecke, die zunächst durch die Polygonreduktion auf 5.000 reduziert und anschließend durch die Dreiecksunterteilung wieder auf circa 8.500 Dreiecke erhöht werden.

3.2.2 Hinzufügen der Basismodelle

Es wird nun erklärt, wie die geometrischen Realisierungen für die Basismodelle berechnet werden. Dabei wird von der Tatsache Gebrauch gemacht, dass die Netzoptimierung von [Ohtake & Belyaev '02] nicht voraussetzt, dass die Ecken des Eingabernetzes auf der Isofläche liegen. Aus den Operatoren von Ohtake und Belyaev wird zunächst durch Hintereinanderschaltung die folgende Operatorkette zusammengestellt

$$\mathcal{O} := \mathcal{R} + \mathcal{P} + \mathcal{Z}. \quad (3.9)$$

Im Gegensatz zu Ohtake und Belyaev wird hier auf die Anwendung des \mathcal{S} -Operators verzichtet. Dieser wird durch den \mathcal{Z} -Operator, also die Projektion der Tangentialschnittpunkte auf die Isofläche, ersetzt. Auf diese Weise liegen die resultierenden Ecken auf der Isofläche der Interpolanten.

Mit Hilfe dieser Operatorkette ist der Korrespondenz-Algorithmus 3.1 definiert, der die Ecken eines Modells iterativ auf die Isofläche der Zielfunktion schiebt, die durch das Zielmodell gegeben ist. Der Algorithmus wird durch die Parameter der Netzoptimierung nach Ohtake und Belyaev sowie die Anzahl K der Interpolationsschritte parametrisiert. In dieser Arbeit wurde für die Netzoptimierung auf die Vorschläge von Ohtake und Belyaev zurückgegriffen. Insbesondere wurde der Parameter für die krümmungsabhängige Regularisierung entsprechend deren Empfehlung auf 2 gesetzt. Bezüglich der Anzahl der Interpolationsschritte wurde die Zahl $K = 4$ experimentell als geeignete Wahl ermittelt und zwar sowohl für Fahrzeug- als auch für Flugzeugmodelle.

Mit Hilfe dieses Algorithmus lässt sich die Berechnung der geometrischen Realisierung g_n für ein Basismodell \mathcal{M}_n wie folgt ausführen (vgl. Abbildung 3.8):

1. Wende den Korrespondenzalgorithmus mit K Interpolationen auf das Meta-Modell $(\mathcal{T}_0, \tilde{g})$ und die Zielfunktion $c_0 + f_n^w$ an.
2. Invertiere die basismodellspezifische Verzerrung durch Anwendung der approximierten inversen Verzerrung \hat{w}_n auf das im ersten Schritt berechnete Modell.

3. Wende den Korrespondenzalgorithmus auf das invertierte Modell und die Zielfunktion f_n an.

Durch den letzten Schritt wird die morphologische Öffnung zurückgerechnet, die durch die Addition der Konstanten c_0 in Gleichung (3.6) realisiert wurde. Abgesehen davon wird in diesem Schritt der Fehler kompensiert, der durch die Approximation der inversen Verzerrung entsteht.

correspondence ()

Eingabe: Anzahl der Interpolationsschritte K

Startmodell $\mathcal{M}_s = (\mathcal{T}, \mathbf{g}_s)$

Zielfunktion f_d

Ausgabe: Modell \mathcal{M}_d : Approximation der Isofläche von f_d in Netztopologie \mathcal{T}

$\mathcal{M}^0 = \mathcal{M}_s$

f_s : implizite Darstellung von \mathcal{M}_s

for ($k = 1$ bis K)

// k -te Interpolation zwischen Start- und Zielfunktion

$f_k = (1 - \frac{k}{K})f_s + \frac{k}{K}f_d$

// Anwendung des Operators \mathcal{O}

$\mathbf{g}^k = \text{OperatorKette}(f_k, \mathcal{M}^{k-1})$

$\mathcal{M}^k = (\mathcal{T}, \mathbf{g}^k)$

end for

$\mathcal{M}_d = \mathcal{M}^K$

Algorithmus 3.1: Die Eckenverschiebung von Modell \mathcal{M}_s auf die Isofläche der Funktion f_d . Die Funktion **OperatorKette**(f, \mathcal{M}) liefert die geometrische Realisierung, die durch Anwendung der Operatorkette \mathcal{O} aus Gleichung (3.9) auf das Modell \mathcal{M} mit Zielfunktion f entsteht.

3.2.3 Polygonreduktion des adaptiven Modells

Für Anwendungen in der Bildanalyse ist es sinnvoll, die Anzahl der Dreiecke des adaptiven Modells zu reduzieren. Dies ist insbesondere dann der Fall, wenn die Anwendung auf der Extraktion von Kanten des Modells operiert, da eine zu hohe Anzahl von Dreiecken zu kurzen und damit unverlässlichen Kanten führt und zudem die Rechenzeit für die Modellanpassung unnötig erhöht. Für die Repräsentation von Flugzeugmodellen

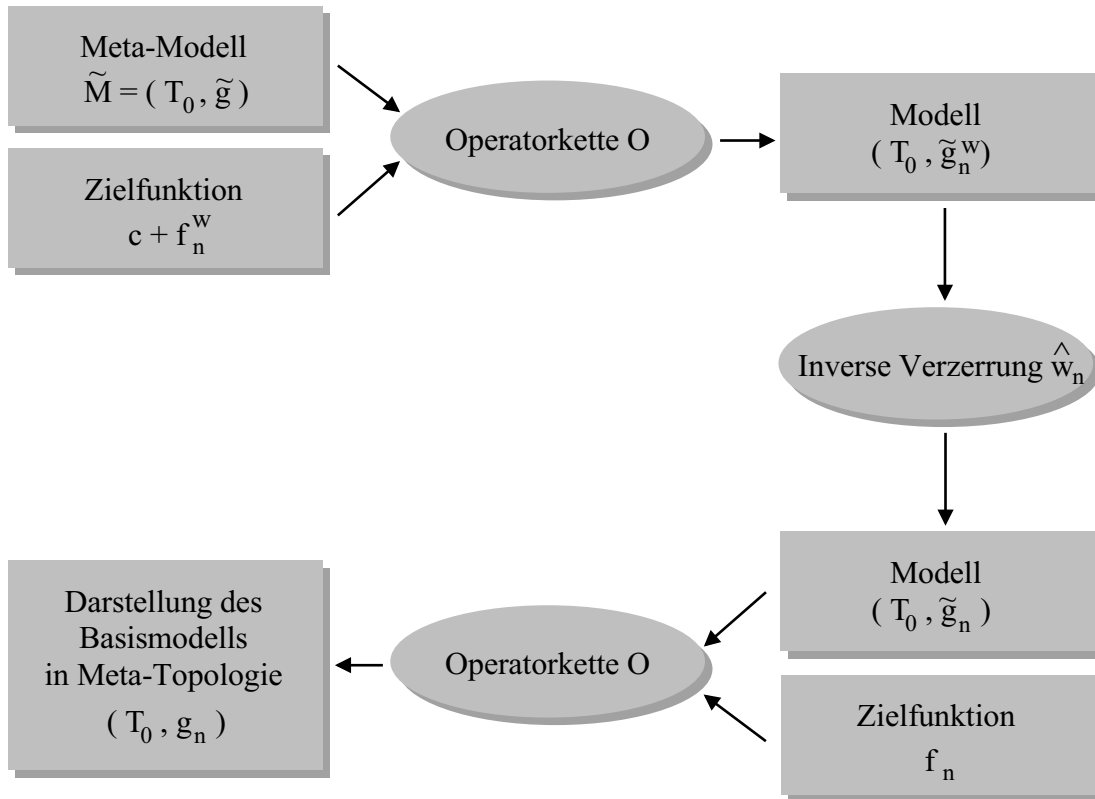


Abbildung 3.8: Die Grafik zeigt die Folge von Einzelschritten, die zur Berechnung einer geometrischen Realisierung durchgeführt werden.

durch ein adaptives Modell sind beispielsweise 2.000 Dreiecke ausreichend, um alle Basismodelle in geometrisch ausreichender Qualität darzustellen. Dies ist weniger als ein Viertel der Dreiecke des adaptiven Modells nach dessen Konstruktion. Aus diesem Grund wird in dieser Arbeit der in Abschnitt 2.3 beschriebene Polygonreduktionsalgorithmus aus [Heckbert & Garland '97] auf die folgende Weise für adaptive Modelle erweitert:

Anstatt jede Kante K_i mit einer einzelnen Fehlerfunktion Q_i zu identifizieren, die von der geometrischen Realisierung abhängig ist, werden für jede Kante $N + 1$ solcher Funktionen $Q_{i,n}$, $n = 0, \dots, N$ zur Verfügung gestellt, nämlich eine für jedes Basismodell des adaptiven Modells. Bezeichnen nun die Punkte $\tilde{\mathbf{x}}_{i,n}$, $n = 0, \dots, N$, die geometrischen Positionen, an denen die Fehlerfunktionen ihr Minimum annehmen, für die also gilt

$$Q_{i,n}(\tilde{\mathbf{x}}_{i,n}) = \min_{\mathbf{x} \in \mathbb{R}^3} Q_{i,n}(\mathbf{x}), n = 0, \dots, N,$$

so sind für die Kante K_i durch

$$\Delta_{i,n} := Q_{i,n}(\tilde{\mathbf{x}}_{i,n}), n = 0, \dots, N \quad (3.10)$$

$N + 1$ basismodellspezifische Kontraktionsfehler erklärt. Der Kontraktionsfehler Δ'_i einer Kante K_i des adaptiven Modells wird aus dem Maximum über alle Basismodelle definiert, also durch

$$\Delta'_i := \max_{n=0, \dots, N} \Delta_{i,n}. \quad (3.11)$$

Wie in [Heckbert & Garland '97] werden die Kanten dann entsprechend dieser Kontraktionsfehler absteigend sortiert und kontrahiert. Die Modifikation der Netztopologie für die Kantenkontraktion kann für adaptive Modelle unverändert nach Heckbert und Garland erfolgen, da das adaptive Modell nur eine Netztopologie besitzt. Für die geometrischen Realisierungen der Basismodelle werden jedoch die geometrischen Positionen der neu eingefügten Ecke jeweils auf die Punkte $\tilde{\mathbf{x}}_{i,0}, \dots, \tilde{\mathbf{x}}_{i,N}$ gesetzt.

3.3 Erweiterte adaptive Modelle

Die Elastizität des Modells ist von entscheidender Bedeutung, wenn der Modellanpassungsprozess gelingen soll. Die Elastizität wiederum ist entsprechend der Konstruktion der adaptiven Modelle alleine durch die Auswahl der Basismodelle bestimmt. Dies wird bereits aus der niedrigen Dimension des orthogonalen Modells ersichtlich, denn es besitzt weniger Parameter als Basismodelle zur Konstruktion verwendet wurden. Betrachtet man nun beispielsweise ein Modell, das aus Flugzeugmodellen mit ein und derselben Tragflächenform erzeugt wurde, so wird auch das adaptive Modell keine andere Tragflächenformen beschreiben können. Die im einleitenden Kapitel in Abbildung 1.1 dargestellten Varianten der Boeing 737 bieten ein anschauliches Beispiel für diesen Sachverhalt: Die vier Varianten 600 bis 900 besitzen alle dieselbe Tragfläche. Demzufolge kann aus diesen Varianten keine der Varianten 100 bis 500 dargestellt werden, die eine andere Spannweite besitzen. Eine Regel, welche Basismodelle zur Verfügung gestellt werden müssen, kann für Freiformobjekte nicht angegeben werden. Garantiert wäre die Darstellbarkeit aller Objekte nur dann, wenn man für jedes darzustellende Objekt auch ein Basismodell zur Verfügung stellen würde. Dies ist jedoch wegen der eingeschränkten Verfügbarkeit von Modellen nicht ohne weiteres möglich.

Nun kann man sich bei vielen technischen Objekten und insbesondere bei Flugzeugen zu Nutze machen, dass die Variationen zwischen einigen Objekten durch einfache, lineare Punkttransformationen definierbar sind. Für Flugzeuge z. B. sind die Varianten eines Flugzeugtyps häufig durch Verlängerungen bzw. Verkürzungen des Rumpfs und der Tragflächen gegeben. Dass solche linearen Punkttransformationen leicht zu definieren und zu implementieren sind, zeigt die folgende Definition (vgl. Abbildung 3.9):

Eine Streckung eines rigidem Modells $(\mathcal{T}, \mathbf{g})$ ist gegeben durch zwei Ecken E_1 und E_0 des Netzes und den Betrag $s > 0$, um den das Modell gestreckt werden soll. Gegebenenfalls muss noch das Objektteil (z. B. eine Tragfläche) angegeben werden, auf das die Transformation angewandt werden soll.

Seien nun \mathbf{x}_0 und \mathbf{x}_1 die geometrischen Positionen der Ecken E_0 und E_1 sowie $\mathbf{e} = (\mathbf{x}_1 - \mathbf{x}_0) / \|\mathbf{x}_1 - \mathbf{x}_0\|$ der entsprechende normierte Richtungsvektor. Für die

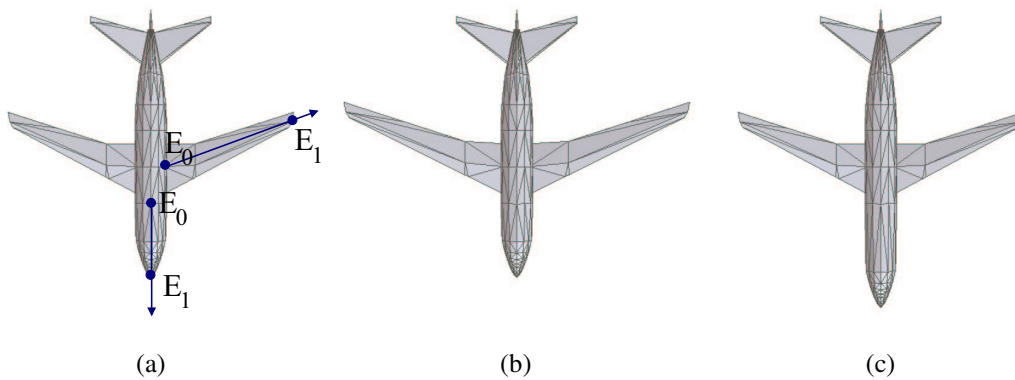


Abbildung 3.9: Streckung rigider Modelle. In (a) ist ein Modell mit vier Punkten abgebildet. Je zwei davon definieren die Streckungen, die in (b) und (c) dargestellt sind. (b) Zeigt eine Streckung der Tragflächen und (c) eine Verlängerung des vorderen Rumpfteils.

Ecken E_i des Objektteils, auf das die Transformation angewandt werden soll, seien die geometrischen Positionen mit \mathbf{x}_i bezeichnet. Zu diesen Positionen wiederum sei $\beta_i = \mathbf{e}^T(\mathbf{x}_i - \mathbf{x}_0)$. Die Verschiebungsvektoren sind dann durch

$$\delta \mathbf{x}_i = \begin{cases} \mathbf{o} & \beta_i < 0 \\ s\beta_i \mathbf{e} & 0 \leq \beta_i \leq 1 \\ s \mathbf{e} & 1 < \beta_i \end{cases}$$

definiert. Für alle übrigen Ecken wird der Verschiebungsvektor auf \mathbf{o} gesetzt. Analog zu den geometrischen Realisierungen werden die Verschiebungsvektoren $\delta \mathbf{x}_i = (\delta x_i, \delta y_i, \delta z_i)^T$ entsprechend der Indizierung der Ecken hintereinander geschrieben:

$$\Delta \mathbf{g} = (\delta x_1, \delta y_1, \delta z_1, \delta x_2, \delta y_2, \delta z_2, \dots)^T$$

Durch $\mathbf{g} + \Delta \mathbf{g}$ ist dann eine neue geometrische Realisierung entstanden, mit einem entsprechend gestreckten Objektteil. Solche Streckungen können z. B. für beide Tragflächen addiert werden.

Die Anwendung dieser Methode für alle Flugzeugtypen führt allerdings zu einer erheblichen Interaktion und der Aufwand zur Konstruktion des adaptiven Modells steigt mit jedem Basismodell erheblich. Sowohl die Interaktion als auch der Rechenaufwand können deutlich reduziert werden, wenn die erforderlichen Punkttransformationen direkt auf dem adaptiven Modell durchgeführt werden. Dazu wird der folgende Weg eingeschlagen: Zunächst wird das gegebene adaptive Modell $(\mathcal{T}_0, \sum_{n=0}^N \alpha_n \mathbf{g}_n)$ durch $\alpha_n = \frac{1}{N+1}$, $n = 0, \dots, N$ in seiner mittleren geometrischen Realisierung $\bar{\mathbf{g}}$ ausgeprägt. Dann werden nacheinander alle gewünschten Punkttransformationen berechnet, also beispielsweise je eine Streckung zur Verlängerung der Tragflächen, eine zur Verlängerung des vorderen und eine zur Verlängerung des hinteren Rumpfteils. Ist nun

eine Reihe von L Punkttransformationen $\Delta_1 := \Delta_1 \bar{\mathbf{g}}, \dots, \Delta_L := \Delta_L \bar{\mathbf{g}}$ gegeben, so werden zusätzlich zu den bereits vorhandenen $N + 1$ Realisierungen $2L + 1$ neue geometrische Realisierungen erzeugt:

$$\begin{aligned} \mathbf{g}_{N+1} &= \bar{\mathbf{g}} + \Delta_1 \\ \mathbf{g}_{N+2} &= \bar{\mathbf{g}} - \Delta_1 \\ &\vdots \\ \mathbf{g}_{N+2L-1} &= \bar{\mathbf{g}} + \Delta_L \\ \mathbf{g}_{N+2L} &= \bar{\mathbf{g}} - \Delta_L \\ \mathbf{g}_{N+2L+1} &= \bar{\mathbf{g}}. \end{aligned}$$

Das erweiterte adaptive Modell wird mit Hilfe dieser neu konstruierten geometrischen Realisierungen wie folgt definiert:

$$\left(\mathcal{T}_0, \sum_{n=0}^{N+2L+1} \alpha_n \mathbf{g}_n \right) \quad \text{mit } \alpha_n \in \mathbb{R} \text{ und } \sum_{n=0}^{N+2L+1} \alpha_n = 1.$$

Durch die Subtraktionen der Vektoren Δ_l für $l = 1, \dots, L$ werden auch die inversen Punkttransformationen zum Modell hinzugefügt. Mit jeder Streckung wird das adaptive Modell also auch um eine Stauchung erweitert. Diese inversen Punkttransformationen werden benötigt, um die mittlere geometrische Realisierung beizubehalten, denn mit

$$\sum_{n=N+1}^{N+2L+1} \mathbf{g}_n = (2L + 1) \bar{\mathbf{g}}$$

gilt für den Mittelwert über alle, also auch die neuen geometrischen Realisierungen:

$$\begin{aligned} \mathbf{g}' &= \frac{1}{(N + 1) + (2L + 1)} \left(\sum_{n=0}^{N+1} \mathbf{g}_n + \sum_{n=N+1}^{N+2L+1} \mathbf{g}_n \right) \\ &= \frac{N + 1}{(N + 1) + (2L + 1)} \bar{\mathbf{g}} + \frac{2L + 1}{(N + 1) + (2L + 1)} \bar{\mathbf{g}} \\ &= \bar{\mathbf{g}}. \end{aligned} \quad (3.12)$$

Die mittlere geometrische Realisierung wiederum wird zu dem adaptiven Modell hinzugefügt, um die Punkttransformationen für alle Basismodelle verfügbar zu machen. Andernfalls könnten die Punkttransformationen nur in der Form $\mathbf{g}_n + \alpha \Delta_l$ mit einem $\alpha \in \mathbb{R}$ auf die jeweiligen Basismodelle angewandt werden, was aber die Forderung $\sum_{n=0}^{N+2L+1} \alpha_n = 1$ an die Parameter des Modells verletzt. Steht jedoch die mittlere geometrische Realisierung für die Linearkombination zur Verfügung, so ist

$$\mathbf{g}_n + \alpha \Delta_l = \mathbf{g}_n + \alpha(\bar{\mathbf{g}} + \Delta_L) - \alpha \bar{\mathbf{g}}$$

und die Forderung ist wegen $1 + \alpha - \alpha = 1$ erfüllt. Diese Argumentation gilt analog, wenn mehr als eine Punkttransformation hinzugefügt wird.

Die neuen zentrierten geometrischen Realisierungen sind nach Gleichung (3.12) durch die Verschiebungsvektoren bzw. den Nullvektor gegeben

$$\begin{aligned}\Delta \bar{\mathbf{g}}_{N+1} &= \Delta_1 \\ \Delta \bar{\mathbf{g}}_{N+2} &= -\Delta_1 \\ &\vdots \\ \Delta \bar{\mathbf{g}}_{N+2L-1} &= \Delta_L \\ \Delta \bar{\mathbf{g}}_{N+2L} &= -\Delta_L \\ \Delta \bar{\mathbf{g}}_{N+2L+1} &= \mathbf{o}.\end{aligned}$$

Damit erfolgt die Berechnung des erweiterten orthogonales Modells auf den zentrierten Realisierungen

$$(\Delta \mathbf{g}_0, \dots, \Delta \mathbf{g}_N, \Delta_1, -\Delta_1, \dots, \Delta_L, -\Delta_L, \mathbf{o}).$$

Die Anwendung von Punkttransformationen auf das adaptive Modell anstelle der Anwendung auf die einzelnen Basismodelle hat drei erhebliche Vorteile. Erstens muss jede Punkttransformation nur ein einziges Mal für das Modell definiert werden und steht dennoch allen Basismodellen zur Variantenausprägung zur Verfügung. Zum Zweiten müssen die Parameter linearer Punkttransformationen, also z. B. die Streckungslängen s , nicht für jede einzelne Variante angegeben werden, denn sie sind zum Parameter des adaptiven Modells geworden. Zum Dritten müssen nicht alle Varianten in Form von Basismodellen zum adaptiven Modell hinzugefügt werden, wodurch die Rechenzeit zur Konstruktion des Modells erheblich verringert wird.

Bezüglich der Streckungslängen s ist noch zu bemerken, dass diese für Stauchungen zu Überlappungen der Dreiecke führen können. Dies ist auszuschließen, solange

$$s < \|\mathbf{x}_1 - \mathbf{x}_0\|$$

gilt. Da der Parameter s jedoch zu einem Modellparameter geworden ist, können diese Artefakte nur dann ausgeschlossen werden, wenn die Punkttransformationen keine übertriebenen Variationen beschreiben müssen. Schon aus diesem Grund kann die Erweiterung durch Streckungen bzw. Stauchungen keinesfalls die Auswahl geeigneter Basismodelle ersetzen.

In diesem Zusammenhang sei erwähnt, dass sich auch die charakteristische Form des Rumpfes einer Boeing 747 auf diese Weise nicht auf einem zylinderförmigen Rumpf zurückführen ließe. Auch die Variation zwischen den eckigen Karosserien älterer Fahrzeugmodelle und den abgerundeten Formen der heutigen Produkte lassen sich nicht durch einfach formulierbare und universelle Punkttransformationen definieren. Für die Modellierung von Flugzeugen bedeutet das beispielsweise, dass nach Möglichkeit für jeden Flugzeugtyp mindestens eine Variante als Basismodell zur Verfügung zu stellen ist, bei starken Längendifferenzen zwischen den Varianten entsprechend mehr.

3.4 Zusammenfassung und Ergebnisse

In den ersten Abschnitten dieses Kapitels wurde ein automatisches Verfahren zur Konstruktion adaptiver Modelle vorgestellt. Das automatisch konstruierte adaptive Modell kann durch geringfügige Interaktion des Benutzers in seiner Elastizität erweitert werden. Die entscheidende Neuerung bei der Konstruktion des adaptiven Modells ist, dass die Punktkorrespondenzen nicht direkt vom jeweiligen Start- zum Zielobjekt berechnet werden, sondern - gestützt durch Funktionsinterpolationen - schrittweise über interpolierte Isoflächen auf das Zielobjekt zubewegt werden. Auf diese Weise wird die Möglichkeit geschaffen, das adaptive Modell auch für dünne Objektteile zu berechnen.

Die einzige Interaktion, die der Benutzer im Vorfeld der Konstruktion zu treffen hat, besteht in der Definition der Ankerpunkte, die in Abhängigkeit der Objektklasse zu wählen sind. Darüber hinaus muss für die Anwendung des Verfahrens eine Reihe von Parametern angegeben werden, die im Folgenden aufgelistet sind:

- Die Zellgröße des Polygonisierers.
- Die Anzahl der Dreiecke, die der Polygonreduktionsalgorithmus liefern soll.
- Der Winkelabstand für die Dreiecksunterteilung (\mathcal{S} -Operator).
- Die Parameter für die Netzoptimierung nach Ohtake und Belyaev:
 1. Der Parameter für den Regularisierungsoperator \mathcal{R} .
 2. Die Schwelle zur numerischen Stabilisierung der dualen Netzoptimierung (\mathcal{P} -Operator).
- Die Anzahl der Funktionsinterpolationen K für die Korrespondenzberechnung.

Die ersten drei Parameter sind eng miteinander verbunden, da sie die Feinheit der Meta-Topologie bestimmen. Im Zweifelsfall sind die Parameter so zu wählen, dass das resultierende Netz sehr fein ist. Eine zu grobe Netztopologie gefährdet die Konstruktion der geometrischen Realisierungen, während eine zu feine Netztopologie lediglich einen erhöhten Rechenaufwand mit sich bringt. Die Anzahl der Dreiecke des adaptiven Modells kann nach Konstruktion des Modells auf eine für die Anwendung des Modells sinnvolle Zahl reduziert werden.

Die verbleibenden drei Parameter gehen in die Berechnung der geometrischen Realisierung ein. Dabei hat sich der Parameter für den Regularisierungsoperator als unkritisch gezeigt, während die anderen beiden Parameter entscheidend zum Gelingen der Berechnung beitragen. Für die Parametrierung der dualen Netzoptimierung kann nach [Ohtake & Belyaev '02] keine Regel angegeben werden. Die Schwelle muss experimentell bestimmt werden. Dasselbe gilt bezüglich der Iterationsschritte K zur Korrespondenzberechnung. Auch diese Größe muss experimentell ermittelt werden. Versuche mit Flugzeug- und Fahrzeugmodellen haben jedoch gezeigt, dass $K = 4$ Interpolationsschritte gute Ergebnisse liefern.

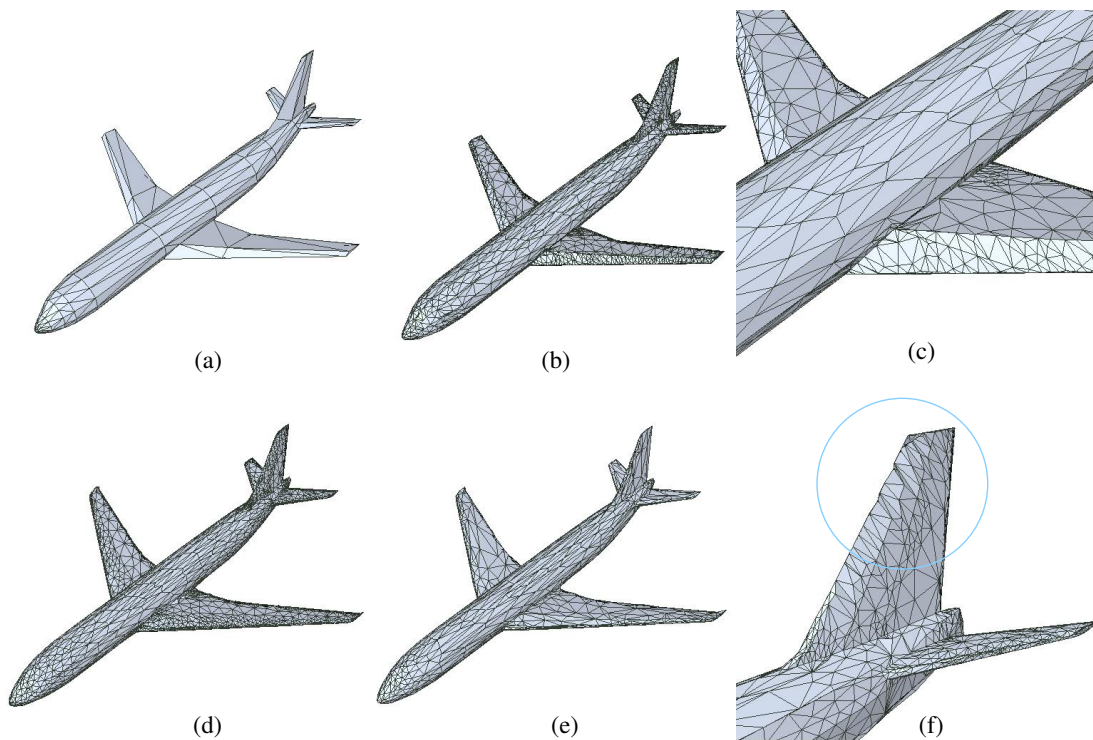


Abbildung 3.10: Ergebnisse der Modellkonstruktion aus sechs Basismodellen: (a) Ein Basismodell in originaler Netztopologie und (b) dasselbe Modell repräsentiert durch die Meta-Topologie. (c) Detail von (b). (d) Das orthogonale adaptive Modell in mittlerer geometrischer Realisierung $\tilde{\theta} = (0, \dots, 0)^T$, (e) dasselbe Modell reduziert auf 1.500 Dreiecke und (f) Fehler verursacht durch Unterabtastung.

3.4.1 Flugzeugmodelle

Mit dem Konstruktionsverfahren wurden mehrere Modelle von Verkehrsflugzeugen konstruiert. Dabei wurde entsprechend der zuvor angestellten Überlegungen auf die Modellierung von diskreten Merkmalen wie Triebwerke, Fahrwerk etc. verzichtet.

Wie aus Abbildung 3.10 ersichtlich ist, ist die Rekonstruktionsqualität der Basismodelle sehr gut. Der Abstand der Ecken des resultierenden Modells zum Basismodell beträgt weniger als 10^{-4} Meter, was der Genauigkeit entspricht, mit der das Bisektionsverfahren des \mathcal{Z} -Operators die Punkte auf der Isofläche positioniert. Vereinzelt treten Kanten auf, die das Innere des Basismodells durchdringen, wie in Abbildung 3.10 (f) dargestellt ist. Solche Fehler lassen sich vermeiden, indem man die Adaption der lokalen Krümmung bei der Konstruktion des Meta-Modells mit einer kleineren Winkelschwelle durchführt. Da diese Fehler für die Anwendung in der Bildanalyse jedoch ohne Bedeutung sind, kann für diesen Anwendungsbereich auf diese Erhöhung des Rechenaufwands verzichtet werden.

Abbildung 3.10 (d) zeigt ferner eine Realisierung nach Anwendung der Hauptkomponentenanalyse mit $\tilde{\theta} = (0, \dots, 0)^T$, also in mittlerer geometrischer Realisierung,

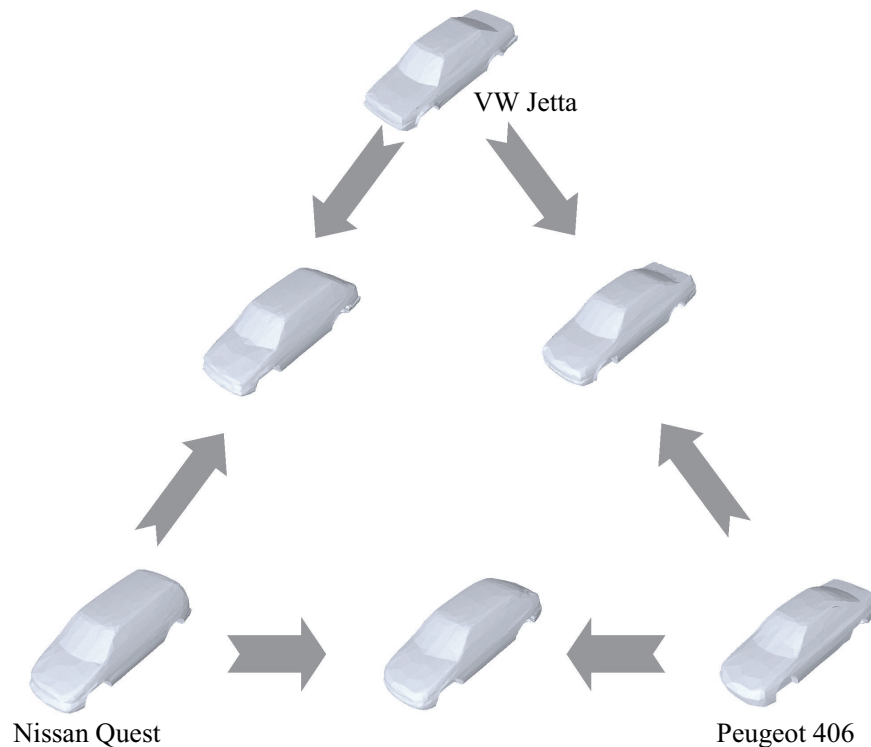


Abbildung 3.11: Die abgebildeten Modelle sind Ausprägungen eines Fahrzeugmodells, das aus drei im Internet frei verfügbaren Modellen erzeugt wurde. Die Ecken des Dreiecks zeigen wieder die Repräsentationen der Basismodelle und die Modelle auf den Kanten entsprechende Zwischenformen. Man beachte die Übergänge der Stufenhecks in die Kastenform des Nissans und den Übergang von der eckigen Karosserie des Jettas in die anderen beiden Modelle (Quelle der Modelle: http://www.maxnet.ru/aks/models_R.html).

und (e) die entsprechende Realisierung des auf 1.500 Dreiecke reduzierten adaptiven Modells.

Bedingt durch die Verwendung der SDT bringt die Konstruktion des adaptiven Modells erheblichen Rechenaufwand mit sich. Bei den hier genannten Auflösungen ist - auf einem Pentium IV mit 1,9 GHz - mit circa einer Stunde Rechenzeit für die Konstruktion des Meta-Modells und mit jeweils vier Stunden zur Berechnung einer geometrischen Realisierung zu rechnen.

3.4.2 Weitere Objektklassen

Die Übertragbarkeit des Verfahrens auf andere Objektklassen wird durch die Anwendung der Konstruktion auf drei Fahrzeugmodelle gezeigt (siehe Abbildung 3.11). Für die Konstruktion wird aus den folgenden Gründen auf die Modellierung der Räder verzichtet. Zum Ersten sind die Räder von handelsüblichen Modellen in der Regel nicht mit der Karosserie verbunden. Dies würde die Berücksichtigung mehrerer Zu-

sammenhangskomponenten notwendig machen, worauf die prototypische Implementierung des Verfahrens jedoch nicht ausgelegt ist. Zum Zweiten sollten zylinderförmige Objekte nicht in Form von Dreiecksnetzen modelliert werden, da diese entweder durch eine hohe Anzahl von Dreiecken dargestellt werden müssen (die Anzahl der für die Räder verwendeten Dreiecke beträgt nicht selten das Zweifache der zur Modellierung der Karosserie benötigten Dreiecke) oder aber schlecht approximiert werden und dann zur Merkmalsextraktion keinen Beitrag leisten. Drittens liefern die Räder kaum trennungswirksame Merkmale für die Typerkennung und müssen, wenn überhaupt, dann nur zur Stabilisierung der Modellanpassung modelliert werden. Für den Fall, dass die Modellierung der Räder für die Modellanpassung tatsächlich von Bedeutung ist, sollten diese in Form von parametrisierten Zylindern und nicht als Dreiecksnetze modelliert werden und auch in der Modellanpassung gesondert behandelt werden.

Für das Fahrzeugmodell ist die Zellgröße auf 0,05 Meter gesetzt und die resultierenden circa 30.000 Dreiecke der Meta-Topologie wurden auf 1.000 reduziert. Die Dreiecksunterteilung wird mit einer Winkelschwelle von $\alpha = \pi/4$ durchgeführt, wodurch das adaptive Modell 3.016 Dreiecke besitzt. Die übrigen Parameter sind unverändert aus der Konstruktion der Flugzeugmodelle übernommen.

Kapitel 4

Modellanpassung

In diesem Kapitel werden zwei für die Modellanpassung entwickelte Module beschrieben, der robuste und der nichtrobuste Schätzer. Da das Verfahren vielseitig anwendbar ist, erfolgt zunächst eine allgemein gültige Beschreibung, also ohne Angabe der anwendungsspezifischen Kombination der Module und deren Parametrierung.

4.1 Die Abbildungskette

Die Verwendung von dreidimensionalen Modellen in einem Bildauswertesystem verlangt immer die Definition einer Abbildung vom Objekt- ins Bildkoordinatensystem. Diese Abbildung startet mit einer Transformation vom Objekt- ins Kamerakoordinatensystem, die durch sechs Parameter (auch äußere Parameter genannt) gegeben ist:

- Die Rotation vom Objekt- ins Kamerakoordinatensystem wird durch die so genannten Eulerwinkel parametrisiert. Die Eulerwinkel werden auch als Roll-, Neigungs- und Gierwinkel (ω , φ und κ) bezeichnet. Die Angabe der Reihenfolge entspricht den Achsen x , y und z .
- Die Verschiebung ist durch einen Translationsvektor (t_x, t_y, t_z) gegeben.

Diese Transformation legt die relative Orientierung und die Verschiebung zwischen Objekt- und Kamerakoordinatensystem fest. Nach Anwendung dieser Transformation erfolgt die Projektion vom Kamera- ins Bildkoordinatensystem durch die Bereitstellung eines mathematischen Modells für den physikalischen Abbildungsvorgang. Typischerweise wird die Orthogonal- oder Zentralprojektion herangezogen, die unter Umständen um radiale Verzeichnungen, Scherungen oder Skalierungen erweitert werden, um Konstruktionsungenauigkeiten des abbildenden Systems erfassen zu können. Die Parameter dieser Projektionsmodelle werden in der Regel durch einen Kalibriervorgang bestimmt, etwa in der Art von [Tsai '87], in dem eine Kalibrierung für handelsübliche CCD-Kameras vorgestellt wird. Für die hier vorgestellte Modellanpassung werden das Lochkameramodell, also die Zentralprojektion, und die Orthogonalprojektion verwendet. Sie sind in Abhängigkeit von den Kameraparametern (auch innere Parameter genannt) definiert, die wie folgt lauten:

- Die Brennweite f (nur im Fall des Lochkameranagements),
- die Bildpunktbreite bzw. -höhe d_x bzw. d_y und
- der Bildhauptpunkt $(c_x, c_y)^T$.

Die Brennweite und die Bildpunktbreite bzw. -höhe werden in der Einheit des Kamerakoordinatensystems angegeben und der Bildhauptpunkt in Bildpunkten. Die entsprechenden Abbildungen werden in Anhang B in Form von homogenen Matrizen angegeben (Gleichung B.1 und B.2). Im Prinzip aber kann für das hier vorliegende Verfahren jedes Kameramodell verwendet werden, das zu einer differenzierbaren mathematischen Darstellung führt. Insbesondere können radiale Verzeichnungen, Scheunungen und Skalierungen modelliert werden.

Im Gegensatz zu den Kameraparametern, die im Folgenden immer als gegeben angenommen werden, müssen die äußeren Parameter im Rahmen der Modellanpassung geschätzt werden. Der Parametervektor für die Modellanpassung wird damit um die sechs Lageparameter erweitert und lautet folglich

$$\theta = (\delta_1, \dots, \delta_N, \omega, \varphi, \kappa, t_x, t_y, t_z)^T. \quad (4.1)$$

4.2 Robuste Modellanpassung

Zur Konstruktion eines Modellanpassungsverfahrens sind drei Schritte notwendig. Zum einen muss ein Merkmalstyp festgelegt werden, auf dem die Modellanpassung operiert. Zum anderen muss eine Kosten- oder Gütefunktion definiert werden, die die Übereinstimmung zwischen den extrahierten Bild- und Modellmerkmalen misst. Drittens muss ein Algorithmus zur Verfügung gestellt werden, der die Kosten- bzw. Gütefunktion minimiert bzw. maximiert.

4.2.1 Merkmalsextraktion

Zur Registrierung, also zur Anpassung eines Modells an ein Bild, stehen viele verschiedene Bildmerkmale zur Verfügung. Typische Merkmale sind Kanten [Theobalt et al. '03], bildpunkt-basierte Intensitäts- oder Farbwerte [Blanz & Vetter '99] oder homogene Flächen [Kecici '01]. Auch die Aggregation solcher Bildmerkmale zu Kantensegmenten, Ecken oder Kurven [Selinger & Nelson '00] lassen sich zur Modellanpassung verwenden. Ebenso möglich ist die Integration mehrerer solcher Merkmalstypen in ein Modellanpassungsverfahren, wie in [Kostiainen et al. '01] geschehen.

Homogene Flächen sind vor allem in der Bildfolgenauswertung von Interesse, wenn aus zeitlich zurückliegenden Aufnahmen Annahmen über die aktuelle Erscheinung des Objekts gemacht werden können. Bildpunkt-basierte Intensitäts- oder Farbwerte können dann eingesetzt werden, wenn diese Merkmale innerhalb der Objektklasse relativ homogen sind und / oder die Helligkeits- bzw. Farbwerte ebenfalls modelliert und

parametriert werden. Modellanpassungen, die von solchen Bildmerkmalen Gebrauch machen, minimieren in der Regel den quadratischen Abstand zwischen den Intensitäten oder Farbwerten des Bildes I und einem synthetisch erzeugten Abbild des Modells I_{syn} etwa in der Art

$$\min_{\theta} \sum_{(x,y)} \|I(x,y) - I_{\text{syn},\theta}(x,y)\|^2, \quad (4.2)$$

wobei die Summation über die Bildpunktpositionen geht. Dieses Vorgehen bedingt allerdings die Definition eines Beleuchtungsmodells zur Bereitstellung des synthetischen Bilds. Für adäquat gewählte Beleuchtungsmodelle können dessen Parameter ebenfalls in der Modellanpassung geschätzt werden, erhöhen damit aber den Freiheitsgrad der Schätzaufgabe.

Die in der Aufgabenstellung geforderte Unabhängigkeit von den Materialeigenschaften der Objektoberfläche lässt sich jedoch auch ohne Parametrierung oder Modellierung von Helligkeitswerten realisieren, wenn kantenbasierte Bildmerkmale eingesetzt werden. Sie benötigen darüber hinaus kein Beleuchtungsmodell. Bei der Verwendung von Kanteninformationen aus Bildmaterial bleibt zu entscheiden, ob diese bildpunktweise oder in Form von aggregierten Kantenelementen eingesetzt werden. Wird die Optimierung bildpunktweise vorgenommen, so entstehen Kostenfunktionen in der Art von Gleichung (4.2), wobei die Grauwertbilder durch Gradientenbetragsbilder oder binarisierte Gradientenbetragsbilder ersetzt werden. Die Verwendung von aggregierten Kantenelementen führt in der Regel zu schnelleren Systemen, da die Aggregation die Anzahl der Bildmerkmale reduziert. Aus diesem Grund werden in dieser Arbeit in erster Linie Kantensegmente eingesetzt.

Für die Extraktion von Bildkantensegmenten wird auf ein im Vorfeld der Arbeit entwickeltes Modul zurückgegriffen. Sie wird mit Standardverfahren der Bildverarbeitung durchgeführt, die beispielsweise in [Korn '99, Gengenbach '94, Jähne '97] nachzulesen sind. An dieser Stelle wird das in dieser Arbeit verwendete Verfahren kurz skizziert: Zunächst wird durch Anwendung von Kantenoperatoren aus [Hashimoto & Skalansky '87] ein Gradientenbetrags- und ein Gradientenrichtungsbild berechnet. Das Gradientenbetragsbild wird anhand einer Schwelle μ binarisiert. Dabei werden nur solche Bildpunkte berücksichtigt, die in einer 3×3 Umgebung in Gradientenrichtung den maximalen Gradientenbetrag besitzen. Auf diese Weise werden Doppelkanten vermieden. Die so definierten Kantenbildpunkte werden wieder unter Berücksichtigung des Gradientenrichtungsbildes zu Kantensegmenten aggregiert. Die Mindestlänge der Bildkantensegmente wird durch einen Schwellwert l angegeben. Das Ergebnis dieser Merkmalsextraktion, also die Menge der Bildkantensegmente, wird mit \mathcal{H} bezeichnet.

Im Gegensatz zu den Bildkantensegmenten sind die Modellkantensegmente von der Abbildungskette und der aktuellen geometrischen Realisierung des Modells abhängig und müssen in iterativen Anpassungsprozessen fortlaufend aktualisiert werden. Die Menge der Modellkantensegmente wird folglich in Abhängigkeit vom Parametervektor θ mit \mathcal{L}_{θ} bezeichnet.

Die Modellkantensegmente werden für einen gegebenen Parametersatz θ von einem *hidden line* Algorithmus in Bildkoordinaten geliefert:¹ Zunächst wird für alle Dreiecke anhand der Normalen ein Sichtbarkeitstest durchgeführt. Kanten mit mindestens einem angrenzenden sichtbaren Dreieck sind potentiell sichtbare Kanten. Diese Kanten können ganz oder partiell verdeckt sein, was in der letzten Stufe des *hidden line*-Algorithmus in der Bildebene getestet wird. Dabei werden partiell verdeckte Kanten entsprechend gekürzt. Die Menge der Modellkantensegmente enthält also die sichtbaren (Teil-)Kanten des Modells, wobei diese Menge noch weiter eingeschränkt werden kann, z. B. auf Kanten mit einem Winkel von mehr als einer gegebenen Schwelle τ . Diese Einschränkung ist insbesondere dann sinnvoll, wenn die modellierten Objekte runde Formen besitzen, die in der polygonalen Darstellung in Form von flach aneinander grenzenden Flächen approximiert werden. In diesem Fall kann nicht erwartet werden, dass das Pendant von Modellkanten mit flachem Winkel im Bild existiert. Diese Modellkanten besitzen somit für die Modellanpassung keine Relevanz. Je nach Objektklasse kann es sinnvoll sein, die Menge der Modellkantensegmente noch weiter auf anpassungsrelevante Modellkanten einzuschränken, z. B. auf die Silhouettenkanten.

4.2.2 Die Gütefunktion

Wie in [Blanz & Vetter '03] demonstriert wird, reagieren adaptive Modelle bei der Ausprägung sehr empfindlich auf Messrauschen, d. h. auf Störungen in der Extraktion der Bild- und Modellmerkmale. Aus diesem Grund müssen zur Schätzung der Parameter entweder die a priori Wahrscheinlichkeiten der Modellparameter herangezogen oder aber es muss ein robuster Schätzer eingesetzt werden. Robust bedeutet hier, dass nicht zum Objekt gehörende Bildkantensegmente nicht oder nur geringfügig berücksichtigt werden.² Bezüglich der a priori Wahrscheinlichkeiten muss auf die Überlegungen aus Abschnitt 3.1 verwiesen werden, in dem die Aussagekraft dieser Wahrscheinlichkeiten in Frage gestellt wurde. Aus diesem Grund ist die Verwendung eines robusten Schätzers vorzuziehen.

Im Gegensatz zu den bildpunktbasierten Methoden verlangt die Optimierung auf Basis aggregierter Merkmale die Lösung des Korrespondenzproblems. Korrespondenzproblem bedeutet hier, dass bestimmt werden muss, welches Bildmerkmal welchem Modellmerkmal entspricht. Die Korrespondenz ist erforderlich, um die Übereinstimmung zwischen Bild und projiziertem Modell messen zu können. Diese Korrespondenzen automatisch festzulegen führt leicht zu Instabilitäten des Systems, es sei denn, es wird ein robuster Schätzer eingesetzt. In [Krüger '01] wird ein Schätzer vorgestellt, der durch die Ausprägung einer Kollineation eine zweidimensionale Karte-zu-Bild Registrierung realisiert, das Korrespondenzproblem in die Optimierung integriert und zudem robust ist.

¹Diese Merkmalsextraktion wird ebenfalls durch ein im Vorfeld der Arbeit implementiertes Modul realisiert.

²Mehr zum Thema robuster Schätzer in der Bildverarbeitung findet man in [Steward '99].

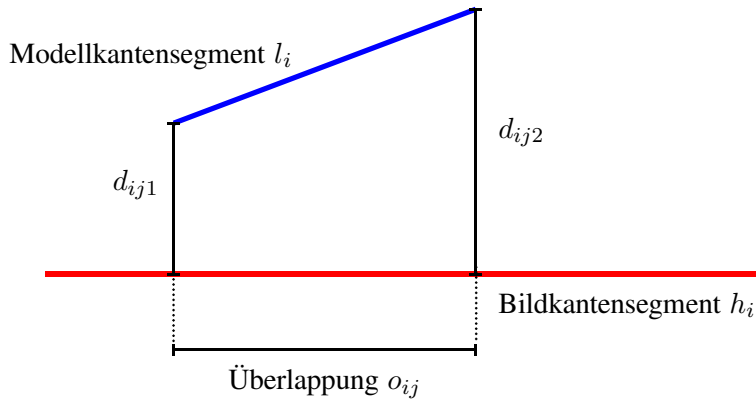


Abbildung 4.1: Die Grafik skizziert die Komponenten des orthogonalen Abstands und die Überlappung anhand eines Bild- und eines Modellkantensegments. Die orthogonale Projektion des Modellkantensegments auf das Bildkantensegment liefert die Überlappung. Der orthogonale Abstand ist durch den Abstand der Endpunkte des Modellkantensegments zu den entsprechenden Lotfußpunkten gegeben.

Für die hier entwickelte Modellanpassung wurde die Gütefunktion von Krüger aufgegriffen. Sie wird im Folgenden näher erläutert, da die Konstruktion der Gütefunktion die Grundlage der im weiteren Verlauf beschriebenen Optimierung ist.

Gegeben sei das Modellsegment $l_i = (\mathbf{q}_{i1}, \mathbf{q}_{i2})$ mit den Endpunkten \mathbf{q}_{i1} und \mathbf{q}_{i2} sowie ein Bildsegment $h_j = (\mathbf{p}_{j1}, \mathbf{p}_{j2})$ mit normierter Normalen \mathbf{n}_j und dem normalisierten Richtungsvektor $\mathbf{e}_j = (\mathbf{p}_{j2} - \mathbf{p}_{j1}) / \|\mathbf{p}_{j2} - \mathbf{p}_{j1}\|$. Der Abstand zwischen den beiden Segmenten wird mit Hilfe der Überlappung (vgl. Abbildung 4.1)

$$o_{ij} = | [0, \|\mathbf{p}_{j2} - \mathbf{p}_{j1}\|] \cap [\mathbf{e}_j^T (\mathbf{q}_{i1} - \mathbf{p}_{j1}), \mathbf{e}_j^T (\mathbf{q}_{i2} - \mathbf{p}_{j1})] | \quad (4.3)$$

und dem orthogonalen Abstand

$$d_{ij}^2 = d_{ij1}^2 + d_{ij2}^2 \quad (4.4)$$

$$= \|\mathbf{n}_j^T (\mathbf{p}_{j1} - \mathbf{q}_{i1})\|^2 + \|\mathbf{n}_j^T (\mathbf{p}_{j2} - \mathbf{q}_{i2})\|^2 \quad (4.5)$$

definiert. Unter Verwendung von *Tukey's biweight* $\bar{\rho}$ [Steward '99]

$$\bar{\rho}(d) = \begin{cases} \frac{1}{6}(1 - (1 - d^2))^3 & |d| \leq 1 \\ \frac{1}{6} & |d| > 1 \end{cases} \quad (4.6)$$

definiert [Krüger '01] die in Abbildung 4.2 dargestellte Funktion

$$\rho(d) = 1 - 6\bar{\rho}(d) \quad (4.7)$$

$$= \begin{cases} (1 - d^2)^3 & |d| \leq 1 \\ 0 & |d| > 1 \end{cases} \quad (4.8)$$

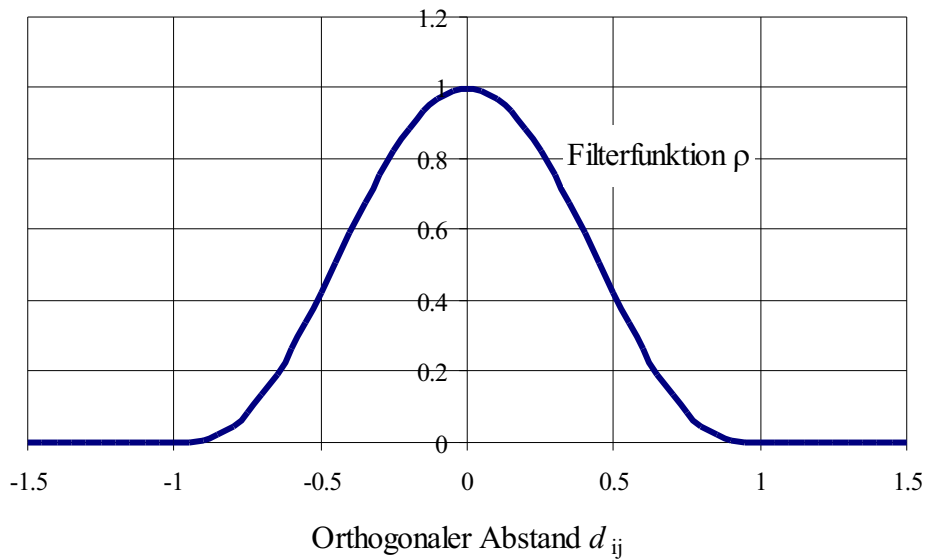


Abbildung 4.2: Die Filterfunktion ρ gewichtet die Paare von Modell- und Bildkantensegmenten entsprechend ihres skalierten orthogonalen Abstands.

Zu einem gegebenen Skalierungsparameter σ ist der Abstand \tilde{d}_{ij} des Modellkantensegments l_i zum Bildkantensegment h_j dann gegeben durch

$$\tilde{d}_{ij} = o_{ij} \rho(d_{ij}/\sigma).$$

Dieser Abstand ist offensichtlich nicht symmetrisch, liefert also keine Metrik auf der Menge aller Kantensegmente. Andererseits hängen nur die Größen \mathbf{q}_{i1} und \mathbf{q}_{i2} vom Parametervektor θ ab, was zu kompakten analytischen Ableitungen führt, die effizient ausgewertet werden können.

Die Skalierung von ρ wird gelegentlich nicht durch den Parameter σ sondern durch eine Größe σ' angegeben. Die Skalierung von ρ erfolgt dann durch

$$\sigma = \sqrt{2}\sigma'.$$

Diese Angabe führt zu einer anschaulichen Skalierung, denn für parallele Modell- und Bildkantensegmente gilt $d := d_{ij1} = d_{ij2}$ und es ist

$$d_{ij} = \sqrt{d_{ij1}^2 + d_{ij2}^2} = \sqrt{2}d.$$

Damit gilt $\rho(\frac{d_{ij}}{\sqrt{2}\sigma'}) = \rho(\frac{d}{\sigma'})$ und es ist

$$\rho(\frac{d_{ij}}{\sigma}) = 0 \Leftrightarrow d > \sigma'.$$

Die Größe σ' ist der kleinste Abstand d , für den parallele Kantenpaare keinen Beitrag zur Gütefunktion leisten.

Die Gütefunktion kann nun in Abhängigkeit vom Parametervektor θ als Summe dieser Abstände über alle Paare von Bild- und Modellkantensegmenten definiert werden:

$$\chi^2(\sigma, \theta) = \sum_{(l_i, h_j) \in \mathcal{L}_\theta \times \mathcal{H}} o_{ij}(\theta) \rho\left(\frac{d_{ij}(\theta)}{\sigma}\right). \quad (4.9)$$

Die Summation realisiert insbesondere die so genannten viele-zu-viele Korrespondenzen. Das bedeutet unter anderem, dass eine Bildkante mehreren Modellkanten entsprechen kann und umgekehrt. Liegen die entsprechenden Bildkanten auch noch parallel zu den Modellkanten, was bei kleinen Skalierungen und einer guten Übereinstimmung von Modell und Bild der Fall ist, so ist die Gütefunktion unabhängig von der Unterteilung der Modellkanten. Sie ist also weitgehend unabhängig von der Triangulierung des Modells und von unterbrochenen und unvollständig extrahierten Bildkantensegmenten.

Wie in [Krüger '01] erwähnt wird, löst die Maximierung dieser Funktion das Korrespondenzproblem und zwar simultan zur Parameterschätzung. Dabei wird keine scharfe Grenze zwischen korrekten und nicht korrekten Korrespondenzen gezogen, wodurch letztlich die Robustheit des Schätzers realisiert ist. In der Tat kann die Funktion $\rho(\frac{\cdot}{\sigma})$ als Tiefpassfilter für die orthogonalen Abstände d_{ij} interpretiert werden, der auf alle Bildkanten angewandt wird. Der *cut-off* und die Steilheit des Filters werden dabei durch den Skalierungsfaktor σ kontrolliert (vgl. Abbildung 4.3).

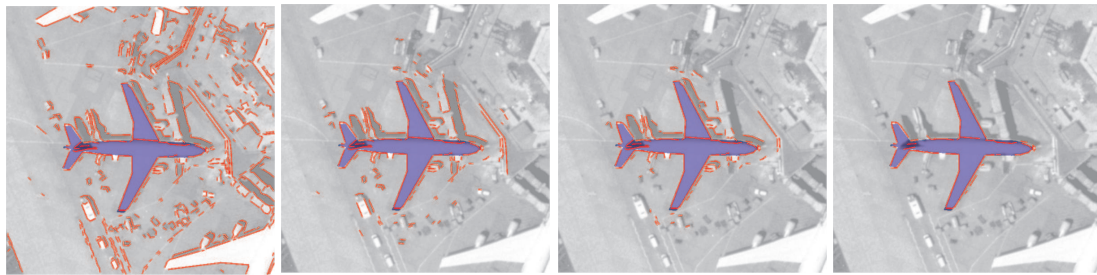
4.2.3 Lokale Extremwerte und Mehrdeutigkeiten

Die Gütefunktion ist, bedingt durch die eingesetzte Filterfunktion, nicht konvex, besitzt also lokale Nebenmaxima. Solche Extremwerte werden in Abbildung 4.4 für eine Drehung des Gierwinkels gezeigt. Die Deformierbarkeit des adaptiven Modells erhöht die Anzahl an Mehrdeutigkeiten für die Modellanpassung noch weiter. Ein Beispiel für solche Mehrdeutigkeiten ist in Abbildung 4.5 (d) dargestellt.

Trifft eine Funktionsoptimierung auf ein solches lokales Maximum, so bleibt die Optimierung an dieser Stelle stehen, da alle Richtungen, in die sich der Parametervektor bewegen könnte, mit einer Verschlechterung der Anpassungsqualität verbunden sind.³ In [Press et al. '92] wird dieser Sachverhalt wie folgt kommentiert: „*Finding a global extremum is, in general, a very difficult problem.*“ In der Praxis bedeutet dies, dass entweder die Parameter so initialisiert werden müssen, dass sie bereits in der Nähe des Optimums liegen, oder aber es bedarf einer weiteren Technik, um diese Extremwerte zu überwinden.

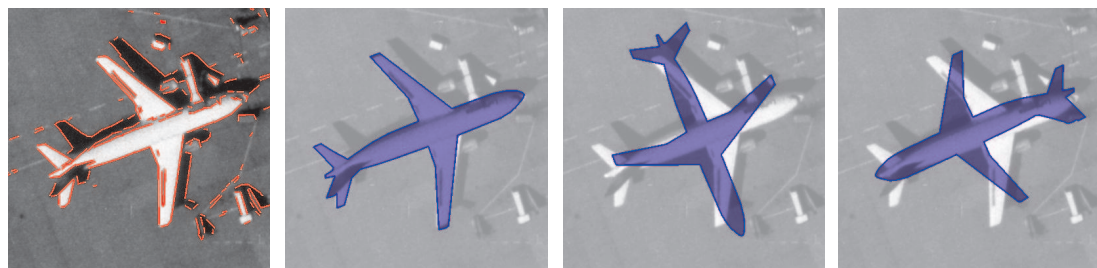
Sind die Mehrdeutigkeiten durch Schatten verursacht, so liegt die Überlegung nahe, den Schatten zu modellieren, was bei dreidimensionalen Modellen möglich ist. Durch die Verwendung von Kantensegmenten anstelle von Grauwerten könnte diese Modellierung durch ein sehr einfaches Modell für Schlagschatten erfolgen. Ein solches

³Eine Ausnahme bilden hier Verfahren, die eine äquidistante Abtastung des Parameterraums durchführen, die hier jedoch - wegen der hohen Anzahl der Parameter - nicht in Frage kommen.



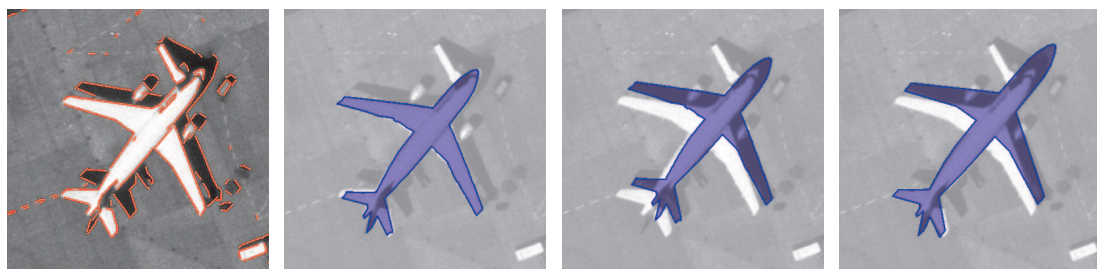
(a) (b) (c) (d)

Abbildung 4.3: Ein Beispiel für die Filterfunktion. In (a) ist ein Bild mit einem eingeblendeten Modell und allen für die Modellanpassung extrahierten Bildkantensegmenten abgebildet. In (b)-(d) sind nur noch diejenigen Bildkantensegmente gezeigt, die für die Modellkontur bei einer Filtergröße von $\sigma' = 100$ (b), 50 (c) bzw. 5 (d) einen Beitrag zur Gütefunktion liefern. Dabei reduziert sich die Anzahl der Bildkantensegmente von 791 auf 243, 163 bzw. 49. Dem stehen 125 Modellkantensegmente gegenüber.



(a) (b) (c) (d)

Abbildung 4.4: Lokale Extremwerte der Gütefunktion: Lage. Für das Bild (a) besitzt die Gütefunktion mit dem Modell und der Lage entsprechend Bild (b) einen Extremwert. Für dasselbe Modell besitzt die Gütefunktion auch für die Lagen aus Bild (c) und (d) lokale Extremwerte, wobei sich die Lagen nur durch den Gierwinkel von Bild (b) unterscheiden.



(a) (b) (c) (d)

Abbildung 4.5: Lokale Extremwerte der Gütefunktion: Schatten und Form. Für das Bild (a) besitzt die Gütefunktion mit dem Modell und der Lage entsprechend Bild (b) einen Extremwert. Die Bilder (c) und (d) zeigen das Modell in einer anderen Position, wobei sich in Bild (d) auch das Modell gegenüber Bild (b) verändert hat. Beide Lagen sind, bedingt durch die Schattenkanten, Extremstellen der Gütefunktion.

Modell ist durch die Ebene, auf die der Schatten fällt, den Abstand des Objekts zu dieser Ebene und die Richtung des Lichteinfalls gegeben. Da die Aufgabenstellung verlangt, keine Annahmen über die Beleuchtungssituation vorauszusetzen, müsste mindestens die Richtung des Lichteinfalls in die Liste der zu optimierenden Parameter aufgenommen werden.

Eine solche Methodik wurde hier implementiert, konnte aber aus folgendem Grund für die definierte Gütefunktion nicht weiter eingesetzt werden: Nach Konstruktion der Gütefunktion ist die Modellanpassung genau dann optimal, wenn jedes Modellkanten-segment - auch die Schattenkanten - mit möglichst vielen Bildkanten überdeckt ist. Da der Schatten im Bild in der Regel vom Objekt partiell überdeckt ist, bieten die Bildkanten-segmente des Objekts für die Schattenmodellkanten die bessere Position als die Bildkanten-segmente des Schattenwurfs. Kurz, die Optimierung setzt, so wie sie konstruiert ist, die Schattenkanten auf die Silhouette des Objekts. Für Anwendungen, die den Sonnenstand als bekannt voraussetzen können und so die Richtung des Lichteinfalls als Systemeingabe zur Verfügung haben, ist die Modellierung des Schlagschattens jedoch eine geeignete Möglichkeit, um die Optimierung der Gütefunktion zu stabilisieren.

4.2.4 Grob-zu-Fein Strategie

Wie Abbildung 4.6 zeigt, erhöht sich die Anzahl der lokalen Extremwerte drastisch mit abnehmendem Skalierungsparameter σ' . Andererseits ist die Optimierung mit großen Skalierungen sehr ungenau, weil die Gütefunktion dann noch durch viele „Ausreißer“ beeinflusst wird, sie also auch solche Bildkanten berücksichtigt, die weit von den Modellkanten entfernt sind.

Diese Beobachtung motiviert die folgende Grob-zu-Fein Strategie: Für eine absteigend sortierte Menge von Skalierungsparametern $\sigma_1 > \dots > \sigma_L$ wird jeweils eine Optimierung von $\chi^2(\sigma_l, \theta)$, $l = 1, \dots, L$ durchgeführt. Die Zusammenstellung der Skalierungsparameter ist der Parametrierung der Modellanpassung zuzuordnen und muss bild- und objektabhängig vorgenommen werden. Die Anwendung dieser Optimierungen auf ein Bild und ein Modell liefert zum einen den gesuchten Parametervektor $\hat{\theta}$ und zum anderen den Vertrauenswert, der durch die Gütefunktion mit der kleinsten Filterskalierung definiert ist:

$$\chi^2(\sigma_L, \hat{\theta}) = \max_{\theta} \chi^2(\sigma_L, \theta). \quad (4.10)$$

4.2.5 Die Optimierung der Gütefunktion

Bleibt die Frage der Optimierung der Gütefunktion für eine gegebene Skalierung σ . Krüger hat dieses Maximierungsproblem in [Krüger '01] auf ein iterativ gewichtetes Kleinste-Quadrate-Problem zurückgeführt. Für die Linearisierung hat er erstens angenommen, dass die Überlappung o_{ij} bezüglich der Parameter lokal konstant ist (also die partiellen Ableitungen der o_{ij} verschwinden) und zweitens, dass die Kollineation auf kleine projektive Störungen eingeschränkt ist. Projektive Annahmen dieser Art können

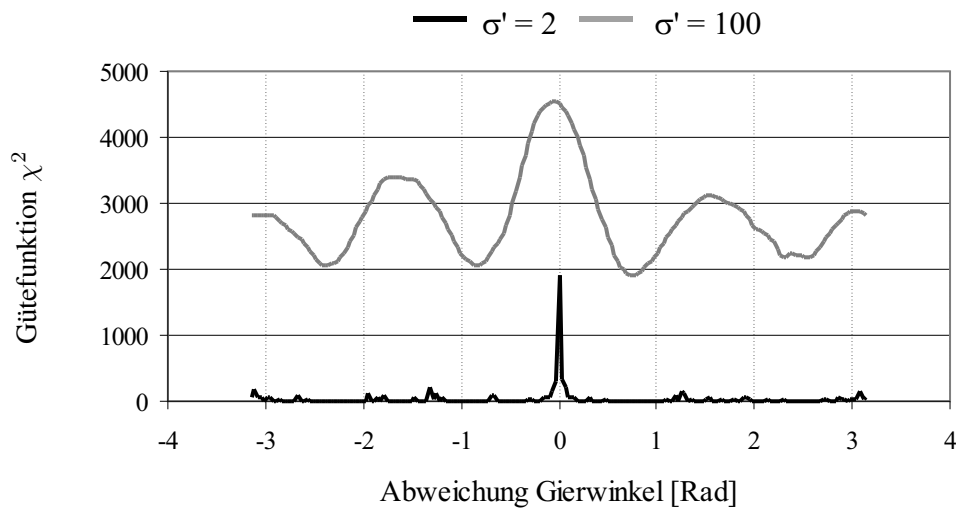


Abbildung 4.6: Die Gütefunktion für zwei Skalierungen mit einem Freiheitsgrad. Die Gütefunktion wurde entsprechend Abbildung 4.4 erzeugt, also durch Variation des Gierwinkels. Die übrigen Parameter waren konstant und wurden entsprechend der Abbildung 4.4 (a) gesetzt. Die Gütefunktion mit der Skalierung $\sigma' = 100$ zeigt vier lokale Extremwerte, von denen drei in Abbildung 4.4 dargestellt sind. Die Skalierung $\sigma' = 2$ zeigt erheblich mehr lokale Extremwerte. Aus Darstellungsgründen wurde die Gütefunktion für $\sigma' = 2$ mit Faktor fünf skaliert.

für die hier vorliegende Projektionskette, insbesondere für die Zentralprojektion, nicht gemacht werden. Aus diesem Grund wurde die Maximierungsaufgabe nichtlinear gelöst.

Zur Lösung solcher Aufgaben stehen viele Verfahren zur Verfügung. Sie unterscheiden sich im Wesentlichen dadurch, ob sie die Gradienten der zu optimierenden Funktion benötigen oder nicht. Die Verwendung der Gradienten liefert in der Regel eine schnellere Konvergenz. Unter der Annahme, dass die Überlappungen o_{ij} bezüglich der Parameter konstant sind (im Folgenden Konstantannahme genannt), sind die partiellen Ableitungen von Gleichung (4.9) für alle Parameter analytisch berechenbar. Daher ist es naheliegend, ein Optimierungsverfahren einzusetzen, das diese Gradienten auch verwendet. Typische Vertreter solcher Verfahren sind der direkte Gradientenabstieg oder das Gauß-Newton-Verfahren.

Für den Gradientenabstieg wird die Schrittweite a priori festgelegt, wohingegen das Gauß-Newton-Verfahren durch die Verwendung der (aus der ersten Ableitung approximierten) Hessematrix die Schrittweite krümmungsabhängig bestimmt. Die Konvergenz dieses Verfahrens ist quadratisch. Die Approximation der zweiten Ableitung aus der Steigung führt jedoch an Stellen mit schwacher Steigung (also insbesondere in der Nähe flach ausgeprägter Extremwerte) zu einem schlecht konditionierten Gleichungssystem, was wiederum die Konvergenz der Optimierung gefährdet. Abhilfe schafft hier das Levenberg-Marquardt-Verfahren, das einen Mittelweg zwischen dem Gradienten-

abstieg und dem Gauß-Newton-Verfahren realisiert. Durch einen Regularisierungsparameter, der sich im Verlauf der Optimierung selbst adaptiert, wird zwischen den beiden Verfahrenscharakteristika hin- und hergeschaltet. Eine ausführliche Erläuterung dieser Interpretation des Levenberg-Marquardt-Verfahrens findet man in [Wachter '97] und zwar im Rahmen einer Übersicht über Verfahren zur Minimierung von Quadratsummen. Eine numerische Beschreibung des Levenberg-Marquardt-Verfahrens ist in [Press et al. '92] nachzulesen.

Ein weiterer Vorteil des Levenberg-Marquardt-Verfahrens ist die Tatsache, dass nach jeder Iteration kontrolliert wird, ob der neue Parametersatz die Optimierung tatsächlich verbessert hat. Dies ist nämlich unter anderem wegen numerischer Instabilitäten nicht immer gegeben. Liegt keine Verbesserung vor, so wird der neue Parametersatz zurückgewiesen und statt dessen der Regularisierungsparameter adaptiert. Es werden also nur Iterationsschritte angenommen, die die Güte der Optimierung tatsächlich verbessern.

Da das Levenberg-Marquardt-Verfahren Quadratsummen voraussetzt, muss die in Gleichung 4.9 definierte Gütefunktion umgeschrieben werden in

$$\chi^2(\sigma, \theta) = \sum_{(l_i, h_j) \in \mathcal{L}_\theta \times \mathcal{H}} \left[\sqrt{o_{ij}(\theta) \rho \left(\frac{d_{ij}(\theta)}{\sigma} \right)} \right]^2. \quad (4.11)$$

Dies ist möglich, da die Summanden nicht negativ sind. Die Gütefunktion wird durch diese Umformung nicht verändert, die Wurzel muss jedoch bei der Bereitstellung der Funktionswerte und der Gradienten innerhalb des Optimierungsverfahrens berücksichtigt werden.

In der Praxis hat sich gezeigt, dass das Verfahren durch die Konstantannahme gelegentlich an Stellen konvergiert, die noch weit vom Optimum entfernt sind. Andererseits ist die Gütefunktion nur stückweise differenzierbar, wenn diese Annahme nicht gemacht wird. Diese Stellen der Nichtdifferenzierbarkeit lassen sich durch eine einseitig stetige Fortsetzung überbrücken, was allerdings zu unstetigen Gradienten führt (siehe Abbildung 4.7). Die unstetigen Gradienten haben zudem geringere Beträge, was kleinere Schrittweiten zur Folge hat. Das Verfahren konvergiert folglich langsamer. Die mit einem Pfeil markierte Stelle in Abbildung 4.7 entspricht einem (schwach ausgeprägten) Extremwert aus Abbildung 4.6. Dieser kann mit der Ableitung unter der Konstantannahme nicht gefunden werden, da der Gradient der Konstantannahme hier keinen Nulldurchgang besitzt. In dem hier gezeigten Fall ist dies von Vorteil. Gelegentlich jedoch resultiert die Nichtübereinstimmung von Gütefunktion und approximiertem Gradienten in einer unerwünschten Konvergenz.

Aus diesem Grund wurde für die Optimierung der folgende Weg eingeschlagen: Wann immer die Optimierung unter der Konstantannahme konvergiert, wird auf die unstetigen Gradienten, die diese Annahme nicht machen, umgeschaltet. Bringen diese Iterationen eine Verbesserung der Gütefunktion, so wird wieder zurückgeschaltet und die Optimierung unter der Konstantannahme fortgesetzt. Bringen die Iterationen keine Verbesserung, so wird die Konvergenz des Verfahrens angenommen und die Optimierung beendet.

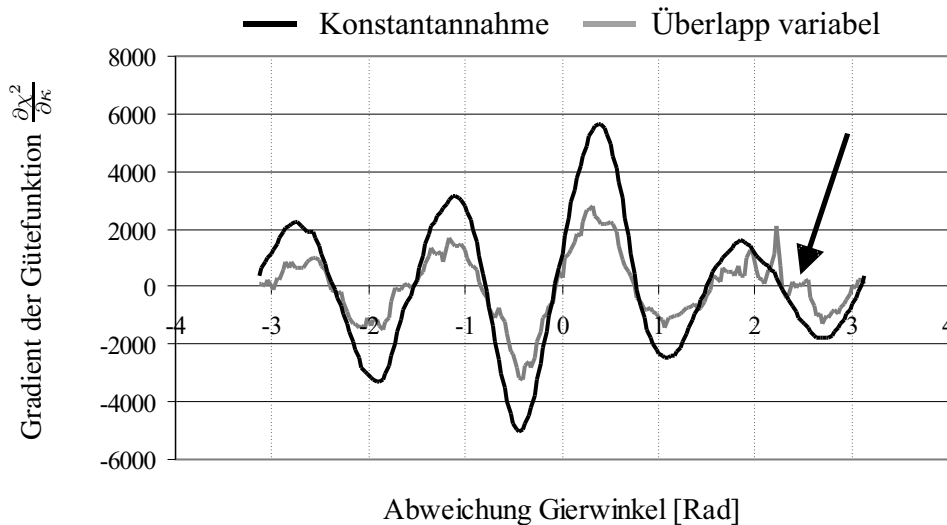


Abbildung 4.7: Die Ableitungen der Gütefunktion entsprechend dem in Abbildung 4.6 gezeigten Diagramm für $\sigma' = 100$. Während die Konstantannahme zu glatten Gradienten führt, zeigt die Verwendung der Gradienten der Überlappung eine ungünstigere Gradientenkurve. Der Pfeil markiert ein schwach ausgeprägtes lokales Maximum, für das der Gradient unter Konstantannahme keinen Nulldurchgang besitzt.

Der Klarheit halber sei nochmals erwähnt, dass diese unterschiedlichen Gradienten keine Auswirkung auf die Gütefunktion haben. Lediglich die Schrittweite und die Richtung, in die die Optimierung den aktuellen Parametersatz schiebt, verändert sich. In Algorithmus 4.1 ist eine Übersicht über die in den letzten Abschnitten definierte Optimierung dargestellt.

4.2.6 Laufzeitoptimierungen

Bevor im nächsten Abschnitt ein weiterer Schätzer erklärt wird, noch ein Wort zum Rechenaufwand, den die Optimierung der Gütefunktion mit sich bringt. Da die Summation der Gütefunktion alle Paare von Bild- und Modellkantensegmenten berücksichtigt, entsteht der Eindruck eines erheblichen Rechenaufwands für die Optimierung. In der Tat müssen in jeder Iteration die „Anzahl der Bildkantensegmente mal Anzahl der Modellkantensegmente“ ausgewertet werden. In dem in Abbildung 4.3 gezeigte Bild beispielsweise setzt sich die Gütefunktion aus 791×125 , also aus 98.875 Summanden zusammen.

Die Auswertung bezieht sich aber lediglich auf die Berechnung der Summanden. Die Gradienten, deren Berechnung erheblich aufwändiger ist, müssen nur für diejenigen Paare berechnet werden, für die der Summand einen positiven Wert geliefert hat. Dies sind höchstens diejenigen Paare, für die der orthogonale Abstand kleiner oder gleich

optimize ()

Eingabe: Bildkanten \mathcal{L}
 initialer Parametervektor θ_0
 adaptives Modell \mathcal{M}
 Skalierungen der Gütefunktion $\sigma_1 > \dots > \sigma_L$

Ausgabe: optimierter Parametervektor θ

```

 $\theta = \theta_0$ 
// für jede Skalierung eine Iteration
for ( $l = 1$  bis  $L$ )

    convergence = false
    repeat

        // Levenberg-Marquardt unter Konstantannahme
         $\theta = \mathbf{LMConstantOverlap}(\sigma_l, \mathcal{L}, \mathcal{M}, \theta)$ 

        // aktuelle Parameter merken
         $\theta' = \theta$ 

        // Levenberg-Marquardt mit variabler Überlappung
         $\theta = \mathbf{LMVariableOverlap}(\sigma_l, \mathcal{L}, \mathcal{M}, \theta)$ 

        if ( $\chi^2(\sigma_l, \theta) < \chi^2(\sigma_l, \theta')$ )
            // durch die variable Überlappung ist keine Verbesserung eingetreten
            convergence = true
             $\theta = \theta'$ 

    until convergence

end for

```

Algorithmus 4.1: Die Optimierung der Gütefunktion erfolgt durch eine Iteration über immer feiner werdende Skalierungen der Filterfunktion. Die Funktionen **LMConstantOverlap()** bzw. **LMVariableOverlap()** wenden die Levenberg-Marquardt Optimierung unter der Annahme der konstanten bzw. variablen Überlappung an. Sie sind durch die Skalierung der Gütefunktion, die Bildkanten, den aktuellen Parametervektor sowie das adaptive Modell parametrisiert. Die Modellkanten werden innerhalb der Levenberg-Marquardt Optimierungen aus dem adaptiven Modell unter Verwendung des jeweils aktuellen Parametervektors extrahiert.

dem Skalierungsfaktor ist, da die Filterfunktion für alle anderen Paare den Wert Null liefert. Dadurch wird die Anzahl der Paare, für die Gradienten berechnet werden müssen, erheblich reduziert. Für das Beispiel aus Abbildung 4.3 beispielsweise beträgt die Anzahl der Paare, für die Gradienten berechnet werden müssen, 30.357 ($\sigma'=100$), 20.375 ($\sigma'=50$) und 6.125 ($\sigma'=5$).

Auch die Anzahl der Paare, für die die Summanden berechnet werden müssen, lässt sich einschränken: Wird nach Beendigung einer Optimierung mit der Skalierung σ_l auf die nächstkleinere Filtergröße σ_{l+1} umgeschaltet, so können zu Beginn dieser Optimierung alle Bildkanten von der weiteren Berechnung ausgeschlossen werden, die für die Skalierung σ_l keinen Beitrag zur Gütefunktion geliefert haben. Dieses Vorgehen wird dadurch legitimiert, dass das Ziel der Grob-zu-Fein Strategie klar definiert ist, nämlich die zu erwartende Anzahl lokaler Extremwerte für kleine Filtergrößen zu umgehen. Zu diesem Zweck sind die Bildkanten überflüssig, die unter der vorangegangenen Filtergröße σ_{l+1} keinen Beitrag zur Gütefunktion mehr leisten.

4.3 Der Passpunktschätzer

Der im letzten Abschnitt vorgestellte robuste Schätzer ist in der Lage Extremwerte der Gütefunktion zu finden. Durch die Grob-zu-Fein Strategie ist gewährleistet, dass die Optimierung nicht auf den zahlreichen lokalen Extremwerten der kleinen Skalierungen konvergiert. Liegen die lokalen Extremwerte jedoch für alle Filtergrößen vor, so kann durch die bisher vorgestellten Techniken nicht verhindert werden, dass die Optimierung auf diesen lokalen Extremwerten konvergiert.

Eine Methode, mit der solche lokalen Extremwerte überwunden werden können, ist die zufällige Verschiebung des Parametervektors, gefolgt vom Neustart der Optimierung. Wird diese Methode mehrfach angewandt und liefern diese Optimierungen keine Verbesserung des Vertrauenswertes, so kann davon ausgegangen werden, dass das vorliegende Extremum ein globales Maximum der Gütefunktion ist. Wird hingegen eine Verbesserung des Vertrauenswertes erreicht, so kann die Prozedur mit dem entsprechenden Parametervektor wiederholt werden. Dieser Vorschlag wird in [Press et al. '92] in der Einleitung zum Kapitel Datenmodellierung gemacht. Andererseits ist die Dimension des hier vorliegenden Parametervektors recht hoch. Dadurch müssten viele Versuche gestartet werden, bis ein Extremum als global deklariert werden könnte.

Eine Alternative ist, an dieser Stelle andere Bildmerkmale ins Spiel zu bringen und einen zweiten Schätzer anzuwenden. Das Ergebnis dieses Schätzers kann dann mit der Gütefunktion des robusten Schätzers verifiziert werden. Wenn eine Verbesserung eingetreten ist, wird der neue Parametervektor angenommen und die Optimierung mit dem robusten Schätzer neu gestartet. Liefert der zweite Schätzer keine Verbesserung des Vertrauenswertes, so werden die neu berechneten Parameter zurückgewiesen und die Optimierung beendet.

Ein solcher zweiter Schätzer wurde für die hier vorliegende Anwendung implementiert und zwar als Minimierung der quadratischen Differenz von korrespondierenden Bildpunkten. Zur Definition einer solchen Passpunktmethode wird eine Menge von gemessenen Bildpositionen benötigt sowie eine Menge korrespondierender Soll-Bildpositionen. Zur Definition der Ist-Werte, also der gemessenen Bildpositionen, müssen auf dem adaptiven Modell einige wenige Oberflächenpunkte (z. B. fünf) definiert werden. Bei der Auswahl dieser Oberflächenpunkte ist - wie bei jeder Methode zur Berechnung von Punktkorrespondenzen - darauf zu achten, dass diese an markan-

ten Stellen angebracht werden, die sich gut und so eindeutig wie möglich zuordnen lassen. Eine schlechte Wahl wäre die Definition der Oberflächenpunkte inmitten einer geradlinigen Modellkante, da diese sicher nicht eindeutig zu einem Bildmerkmal zuzuordnen ist. In jedem Fall sind charakteristische Objektecken geeignet. Auch ist darauf zu achten, dass eine ausreichende Menge an Punkten zur Verfügung steht, da nur für die jeweils sichtbaren Oberflächenpunkte korrespondierende Bildpositionen gefunden werden können. Ausreichend bedeutet hier, dass zu jedem sichtbaren Bildpunkt zwei Gleichungen aufgestellt werden, also für je zwei Parameter ein sichtbarer Oberflächenpunkt zur Verfügung stehen muss. Andernfalls ist die Schätzung der Parameter unterbestimmt.

Theoretisch könnten die Oberflächenpunkte auch automatisch anhand starker Oberflächenkrümmungen des adaptiven Modells in mittlerer geometrischer Realisierung bestimmt werden. Der Aufwand der Implementierung steht allerdings in keinem Verhältnis zur einmaligen interaktiven Definition solcher Punkte, die mit dem adaptiven Modell gespeichert und später abgerufen werden können.

Kostenfunktion und Optimierung

Von einer gegebenen Menge von Oberflächenpunkten werden nun unter Anwendung des aktuellen Parametervektors die sichtbaren Punkte auf die Bildebene projiziert und liefern so die Bildpositionen der Ist-Werte $\mathbf{x}_1, \dots, \mathbf{x}_M$. Zu jedem der Ist-Werte \mathbf{x}_m wird dann abhängig von den gewählten Merkmalen eine Soll-Position \mathbf{y}_m bestimmt. Der neue Parametervektor wird dann durch eine Minimierung der Quadratsumme

$$\sum_{m=1}^M \|\mathbf{x}_m(\theta) - \mathbf{y}_m\|^2$$

aktualisiert. In dieser Arbeit ist die Minimierung der Summe durch eine Kleinste-Quadrate Optimierung realisiert, wobei lokal linearisiert und die Optimierung iterativ angewandt wird.

Diese Schätzmethode ist sicherlich nicht robust. Für eine reine Lagegeschätzung ist diese Tatsache von geringer Bedeutung, für die Formschätzung des adaptiven Modells jedoch nicht. Werden nur die Lageparameter geschätzt, so sorgt z. B. ein einzelner Ausreißer unter den Soll-Werten für eine (gegenüber korrekten Messungen) leicht veränderte Lage und die Ist- und Soll-Werte kommen nicht zur Deckung. Werden jedoch auch Formparameter geschätzt, so können durch die Elastizität des Modells alle Ist- und Soll-Werte, also insbesondere die Ausreißer zur Deckung kommen, wodurch völlig unpassende Ausprägungen entstehen können. Die in dieser Arbeit gezogene Konsequenz ist, den nicht robusten Schätzer auf maximal den ersten Parameter des Modells anzuwenden und die verbleibenden Modellparameter fest zu lassen. Auf diese Weise kommt der Schätzer seiner angedachten Aufgabe nach, nämlich lokale Extremwerte der Gütefunktion des robusten Schätzers zu überwinden, ohne das Modell in nicht zu den Bildsegmenten passende Ausprägungen zu versetzen. Für die bereits gezeigten Flugzeugmodelle bedeutet dies, dass dem nicht-robusten Schätzer nur die Objektgröße zur Optimierung zur Verfügung gestellt wird.

Berechnung der Punktkorrespondenzen

Bleibt die Frage nach den Merkmalen und der Methode, die zur Berechnung der Punktkorrespondenzen, also der Soll-Werte herangezogen werden. Grundsätzlich können für die Berechnung der Soll-Werte beliebige Bildmerkmale eingesetzt werden, so lange sie die Definition korrespondierender Bildpunkte erlauben. Im Rahmen dieser Arbeit wurde weiter auf die Gradientenbetragsbilder gesetzt, allerdings im Gegensatz zu dem robusten Schätzer auf Bildpunktbasis anstelle aggregierter Kantenelemente. Diese Entscheidung wurde motiviert durch die Tatsache, dass die hier relevanten Objekte neben vielen geraden Objektkanten auch Krümmungen besitzen (z. B. die Flugzeugnase), die durch Kantensegmente nur schlecht approximiert werden. Andere, ebenfalls anpassungsrelevante Objektkanten sind zu kurz, um verlässliche Bildkantensegmente zu liefern (z. B. die Begrenzung der Tragflächenspitze).

Die Berechnung wurde durch eine klassische Schablonenanpassung auf dem Gradientenbetragsbild realisiert (vgl. Abbildung 4.8). Für die Schablonen wird zunächst ein synthetisches Binärbild berechnet, indem die Modellkanten - unter Verwendung des aktuellen Parametervektors - in den Bildbereich projiziert werden. Dieses Binärbild wird mit einer Gaußglocke der Varianz $\sigma^2 = 1$ geglättet. Anschließend wird um jede Ist-Position herum eine Schablone ausgeschnitten. Durch Korrelation dieser Schablonen mit dem Gradientenbetragsbild des originalen Bildes werden letztlich die Soll-Positionen bestimmt.

Der Einsatz der Schablonenanpassung zur Berechnung der Punktkorrespondenzen erhöht die Rechenzeit erheblich. Ein Rechenzeitvergleich, der im Zusammenhang mit der Lageschätzung vorgestellt wird, beweist diese Aussage. Das Verfahren kann jedoch durch den Austausch dieses Submoduls, d. h. durch die Verwendung von aggregierten Merkmalen, optimiert werden. Gute Aussichten auf Verbesserung verspricht die Extraktion von Kurven in der Art von [Selinger & Nelson '00]. Aus geradlinigen Segmenten aggregierte Ecken, die in [Krüger '01] für die Initialisierung eingesetzt werden, sind wegen der anzustrebenden Extraktion von kurzen und krummlinigen Bildmerkmalen weniger empfehlenswert.

4.4 Iterative Optimierung

Für adaptive Modelle hat die simultane Optimierung aller Modell- und Lageparameter einen zu hohen Freiheitsgrad, um die Parameter auf einmal schätzen zu können. Da die Parameter orthogonaler adaptiver Modelle in einer - bezogen auf die Variation der Basismodelle - absteigenden Sortierung vorliegen, wird die Optimierung iterativ über die Anzahl der Modellparameter vorgenommen. In der ersten Iteration wird nur eine kleine Anzahl n_0 an Parametern geschätzt, also die Parameter $\delta_1, \dots, \delta_{n_0}$. Die übrigen Modellparameter werden entsprechend der Initialisierung fest gelassen. In der zweiten Iteration wird die Anzahl der freigelassenen Parameter um eine feste Zahl, z. B. zwei, erhöht und diese Prozedur so lange fortgesetzt, bis die maximale Anzahl an relevanten Parametern erreicht ist.

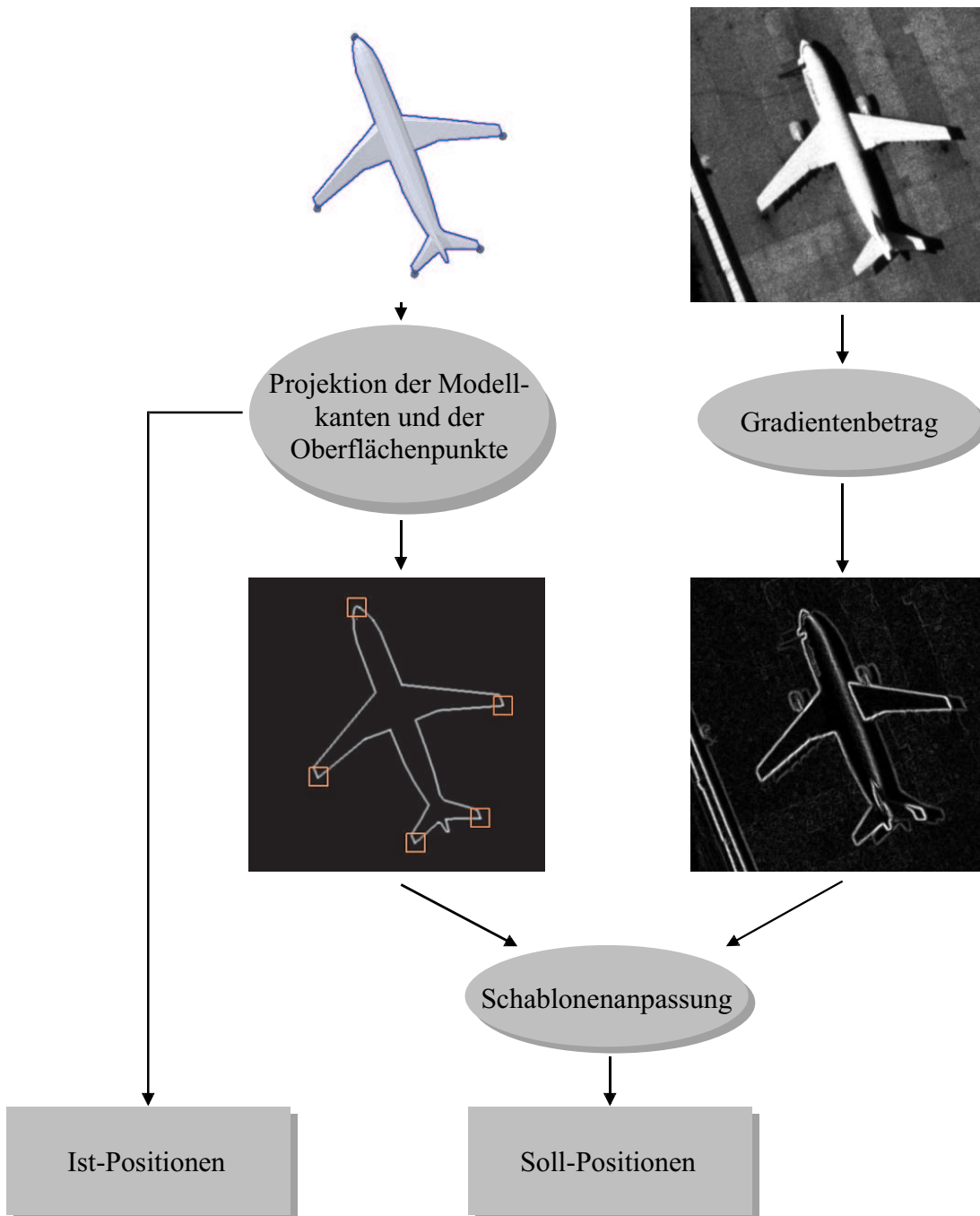


Abbildung 4.8: Die Berechnung der Passpunkte erfolgt durch Schablonenanpassung auf dem Gradientenbetragsbild. Die Schablonen werden aus einem synthetischen Bild erzeugt, das aus den Modellkantensegmenten berechnet wird.

In allen im Rahmen dieser Arbeit durchgeführten Optimierungen wurden in der ersten Iteration genau diejenigen δ_n geschätzt, für die $\sigma_n > 0,1 \sigma_1$ gilt, wobei σ_1 bzw. σ_n die zu δ_1 bzw. δ_n korrespondierende Standardabweichung bezeichnet. In der ersten Iteration werden also nur diejenigen n_0 Parameter geschätzt, deren zugehörige Standardab-

Parameterindex i	Standardabweichung σ_i
erste Iteration	
1	144,06
2	28,37
3	23,73
4	19,06
weitere Iterationen	
5	12,42
6	9,89
7	6,28
8	5,43
9	3,27
10	2,63
11	1,83
unberücksichtigt	
12	1,00
13	0,68
14	0,49

Tabelle 4.1: Die Tabelle zeigt die Standardabweichungen eines orthogonalen adaptiven Flugzeugmodells mit 14 Parametern. In der ersten Iteration werden vier Parameter geschätzt und die Anzahl der Freiheitsgrade wird in der Modellanpassung auf elf beschränkt.

weichung ein Zehntel der größten auftretenden Standardabweichung σ_1 überschreitet. Zudem wurde die Anzahl der zu schätzenden Parameter nach oben beschränkt. Parameter, deren entsprechende Standardabweichung σ_n die Größe $0,01 \sigma_1$ unterschreitet, deren Standardabweichung also kleiner ist als ein Hundertstel der größten auftretenden Standardabweichung, besitzen keine Relevanz für die Modellausprägung. Entsprechend ist der Index n_{\max} zu verstehen.

Tabelle 4.1 beispielsweise zeigt die Standardabweichungen eines in der Flugzeugerkennung verwendeten adaptiven Modells, das im Zusammenhang mit der Typerkennung in Abschnitt 5.1.3 beschrieben wird. Das Modell setzt sich aus 15 Basismodellen zusammen und besitzt demzufolge 14 Freiheitsgrade. Für dieses Modell werden in der ersten Iteration $n = 4$ Parameter geschätzt und für die maximale Anzahl gilt $n_{\max} = 11$. Die Optimierung umfasst insgesamt fünf Iterationen.

In Abbildung 4.9 ist eine Übersicht über die iterative Optimierung dargestellt, die in dieser Form in zweifacher Hinsicht eine Grob-zu-Fein Strategie ist. In einer äußeren Schleife wird das adaptive Modell zunehmend verfeinert und in der inneren Schleife wird die Skalierung der Filterfunktion sukzessive herabgesetzt.

In Abbildung 4.10 wird eine Anwendung der iterativen Optimierung für eine Boeing 747-400 vorgestellt. Da für das Bild keine Kalibrierdaten zur Verfügung standen, diese jedoch als gegeben vorausgesetzt wurden, wurden diese zuvor durch Anpassung

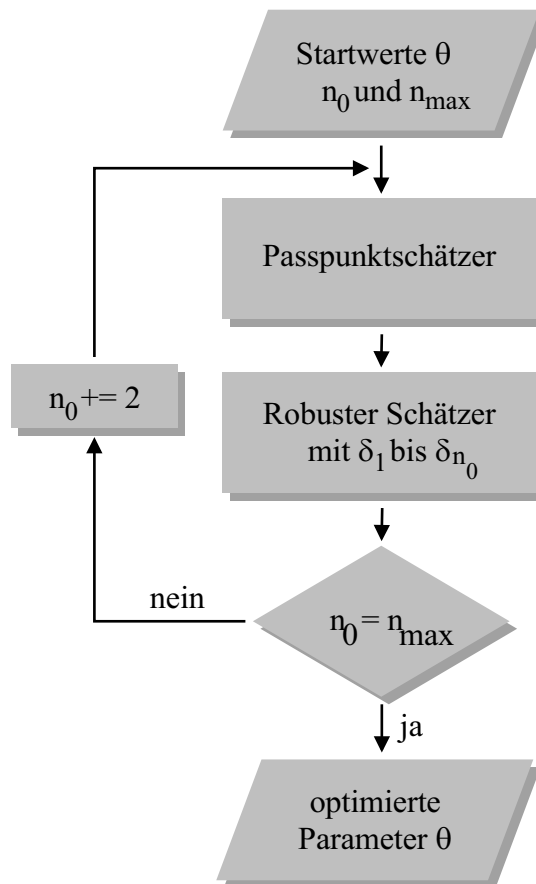


Abbildung 4.9: Iterative Optimierung. Das Flussdiagramm zeigt die äußere Schleife der Optimierung für die Form- und Lageschätzung. Der robuste Schätzer lässt nur die angegebenen Modellparameter frei. Die übrigen Modellparameter werden entsprechend der Initialisierung festgehalten.

einer ausgeprägten Boeing 747-400 bestimmt. Die initiale Lage wurde interaktiv - mit Hilfe einer grafischen Oberfläche - vorgegeben. Die initialen Modellparameter wurden auf einen Airbus A340-300 gesetzt. Das Beispiel ist als Beweis für die Anwendbarkeit der iterativen Optimierung auf Seitenansichten einzuordnen.

4.5 Zusammenfassung

In den letzten Abschnitten wurden zwei Module zur Modellanpassung vorgestellt, von denen eines zur Überwindung lokaler Extremwerte eingesetzt wird. Dieses Modul ist ein Passpunktschätzer, wobei die erforderlichen Punktkorrespondenzen durch Schablonenanpassung auf dem Gradientenbetragsbild bestimmt werden. Die Parametrierung erfolgt demnach über



Abbildung 4.10: Rekonstruktion einer Boeing 747-400. Die initiale Lage sowie die initiale Modellausprägung (Airbus A340-300) sind links oben abgebildet. Rechts oben ist das Ergebnis der Modellanpassung zu sehen, wobei dem Modell die Textur des Bildes aufgeprägt wurde. In der zweiten Reihe ist links das Ergebnis der Rekonstruktion in einer anderen Lage zu sehen und rechts sind die Komponenten der Hausdorff-Distanz gegenüber der in der Datenbank abgelegten Boeing 747-400 farblich codiert. Dabei sind Dreiecke mit geringem Fehler blau und Dreiecke mit großem Fehler rot dargestellt.



(a) Initialisierung

(b) Geschätzte Lage

Abbildung 4.11: Modellanpassung für rigide Modelle. Abbildung (b) zeigt das Ergebnis der Anpassung eines Modells an ein Bild durch den robusten Schätzer. Die Modellkanten sind blau eingezeichnet. Die Initialisierung war entsprechend Bild (a) gegeben. (Quelle des Modells: Benteler AG).

1. die Definition charakteristischer Oberflächenpunkte,
2. die Schablonengröße und
3. den Suchbereich.

Das eigentliche Optimierungsmodul ist jedoch die robuste Schätzung der Parameter auf Basis von Modell- und Bildkantensegmenten. Hierbei werden die Korrespondenzen der Merkmale simultan zur Parameterschätzung bestimmt. Dieser Schätzer kann über seine Skalierung in einer Grob-zu-Fein Strategie eingesetzt werden und wird durch die folgenden Parameter eingestellt:

1. Die Parameter zur Extraktion der Bildkantensegmente, also die Schwelle μ für den Gradientenbetrag und die Mindestkantenlänge l .
2. Die minimale Winkelschwelle τ für die Extraktion der Modellkanten.
3. Eine absteigende Folge von Skalierungen der Filterfunktion $\sigma_1 > \dots > \sigma_L$.

Zur Einstellung der Verfahrensparameter wurde ein Werkzeug mit grafischer Benutzungsoberfläche implementiert (siehe Abbildung 4.12). Die Oberfläche erlaubt unter anderem eine interaktive Initialisierung, die Eingabe der internen Kameraparameter und die Anwendung der implementierten Module auf die jeweilige Szene.

Die in diesem Kapitel vorgestellte Optimierung kann selbstverständlich auch auf rigide Modelle zur Lageschätzung angewandt werden, da diese eine Spezialform adaptiver Modelle darstellen. Ein Beispiel für eine solche Modellanpassung wird in Abbildung 4.11 gezeigt. Auf eine separate Implementierung für rigide Modelle kann verzichtet werden, da rigide Modelle einen Spezialfall adaptiver Modelle mit nur einer geometrischen Realisierung darstellen. In der Regel kann für rigide Modelle auf die Anwendung des Passpunktschätzers verzichtet werden.

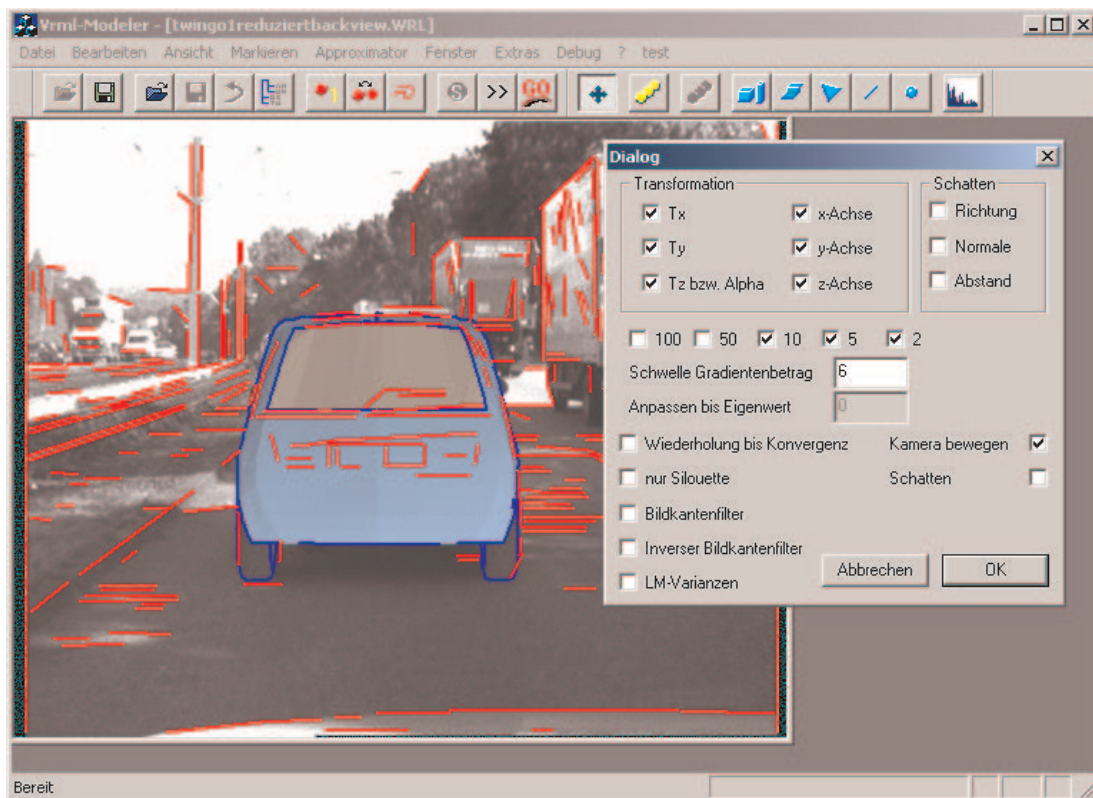


Abbildung 4.12: Ein *screenshot* des Werkzeugs zur Einstellung der Verfahrensparameter und zur Zusammenstellung der Optimierungsmodule (Bildquelle Verkehrsszene: Robert Bosch GmbH).

Kapitel 5

Anwendungen

Das in den letzten Kapiteln beschriebene Verfahren zur Konstruktion und Registrierung adaptiver Modelle ist in vielfältigen Anwendungen einsetzbar. Eine davon ist das in der Einleitung skizzierte Modul zur Unterstützung von Luftbilddauswertern, das im Folgenden beschrieben und diskutiert wird. Darüber hinaus kann das Verfahren zur Lageschätzung genutzt werden, was ebenfalls in diesem Kapitel gezeigt wird.

5.1 Flugzeuigerkennung

Die in dieser Arbeit entwickelte Flugzeuigerkennung arbeitet als vollständig automatisiertes System (vgl. Abbildung 1.4). Dem System wird ein Grauwertbild mit Kollateraldaten, ein adaptives Modell sowie eine Datenbasis zur Verfügung gestellt. Nach einer Vorverarbeitung des Bildes werden die Lage- und Modellparameter ausgeprägt. Diese Ausprägung erfolgt in zwei Stufen, nämlich einer Initialisierung und einer anschließend ausgeführten Optimierung mit Hilfe der im vorangegangenen Kapitel beschriebenen Modellanpassungsmodule. In einem letzten Schritt werden aus der 3D-Rekonstruktion Merkmale extrahiert, die durch eine prototypische Nächste-Nachbar Klassifikation mit einer Datenbasis verglichen werden. Diese Datenbasis wird aus dreidimensionalen Modellen erzeugt. Dem Anwender werden dann die besten Übereinstimmungen aus der Datenbasis präsentiert. Die grafische Oberfläche des Systems erlaubt dem Anwender zudem, die Modellanpassung visuell zu überprüfen. Im Folgenden wird der Aufbau dieses Systems beschrieben und anschließend die Versuchsergebnisse diskutiert. Zuvor werden jedoch die erforderlichen Kollateraldaten präzisiert.

5.1.1 Die Kollateraldaten

Betrachtet man das Abbild einer Kugel, das durch eine zentralperspektivische Projektion entstanden ist, so gibt das Bild nur dann Aufschluss über die Größe der Kugel, wenn der Abstand der Kugel zur Kamera bekannt ist. Umgekehrt kann man den Abstand nur dann bestimmen, wenn man die Größe der Kugel kennt. In anderen Worten, der Aufnahmeabstand und die Objektgröße sind korreliert. Dieselbe Argumentation gilt für die Korrelation zwischen den inneren Kameraparametern und der Objektgröße.

In Abbildung 3.2 wurde die Variation eines Flugzeugmodells längs des ersten Eigenvektors gezeigt, die verdeutlicht, dass die in dieser Arbeit hergeleiteten adaptiven Flugzeugmodelle (sofern sie nicht aus Basismodellen mit identischer Größe erzeugt wurden) erhebliche Elastizität in Bezug auf die Objektgröße besitzen. Zwar ist durch die Konstruktion des adaptiven Modells theoretisch eine einheitliche Skalierung gewährleistet. Die Ausprägung sehr ähnlicher Objekte in unterschiedlichen Größen ist dadurch jedoch nicht ausgeschlossen. In der Praxis hat das zur Folge, dass der Aufnahmeabstand und die Modellparameter auch für adaptive Flugzeugmodelle korreliert sein können.

Auch in Bezug auf Drehungen um die Längs- und Querachse der Flugzeuge lassen sich Korrelationen zu den Modellparametern feststellen: Betrachtet man ein Flugzeug von oben, so verringert eine Drehung des Flugzeugs um die Längsachse (z -Achse entsprechend Abbildung 5.1 (a)) die abgebildete Spannweite. Erfolgt die Rotation um die Querachse (x -Achse entsprechend Abbildung 5.1 (a)), so verkürzt sich die Darstellung des Rumpfes. Während die Rotation um die Längsachse in einer qualitativ hochwertigen Aufnahme durch das Leitwerk und die asymmetrische Abbildung der Tragflächen detektiert werden könnte, ist eine Drehung um die Querachse nur anhand der Stauchung oder Streckung der Tragflächen oder des Höhenruders erkennbar. Abbildung 5.1 (b) beispielsweise zeigt eine parallel projizierte Draufsicht auf ein Boeing 767-Modell. Für Bild (c) wurde das Modell um 20 Grad um dessen Querachse gedreht, wodurch der Rumpf um 2,4 Meter bzw. 5,5 Prozent kürzer abgebildet wird.

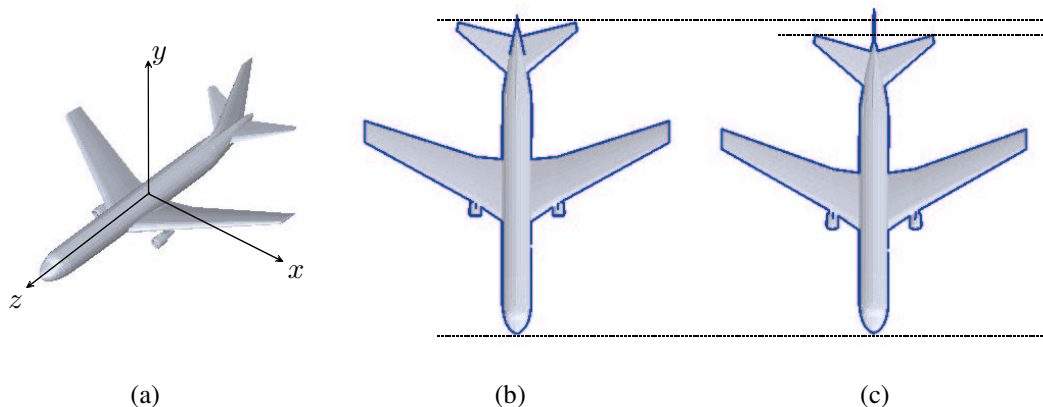


Abbildung 5.1: Korrelation zwischen den Formparametern und einer Drehung um die Querachse (x -Achse). In (b) ist eine parallel projizierte Draufsicht auf ein Boeing 767-Modell dargestellt. Für (c) wurde das Modell um 20 Grad um die x -Achse gedreht, wodurch der Rumpf um 2,4 Meter bzw. 5,5 Prozent kürzer abgebildet wird.

Da adaptive Modelle für Flugzeuge sowohl im Leitwerk als auch in den Tragflächen Elastizität besitzen müssen, hat das für die Erkennung des Flugzeugs die folgende Konsequenz: Entweder man kennt die Ebene, auf der das Flugzeug steht oder man kann die Spannweite und die Rumpflänge nicht messen. Mit der Anpassung rigider Modelle könnte diese Voraussetzung fallen gelassen werden, allerdings nur, wenn die ver-

fügbaren Modelle diese lagebestimmenden Merkmale in ausreichender Genauigkeit modellieren und die Modellanpassung diese Merkmale auch korrekt zuordnet. Dies ist jedoch wegen der unzureichenden Genauigkeit handelsüblicher Modelle nicht zu erwarten.

Die Voraussetzung, dass die relative Orientierung zwischen der Kamera und der Bodenebene bekannt ist, wird als *ground plane constraint* bezeichnet und wird in Anwendungen der Verkehrsüberwachung sowie im Bereich der intelligenten Fahrzeuge häufig genutzt. Exemplarisch sei die Fahrzeugklassifikation aus [Sullivan et al. '95] genannt, die ebenfalls von dieser Voraussetzung ausgeht.

In Anwendungen mit stationärer Kamera, wie z. B. im Flugzeug-Andocksystem von Honeywell Airport Systems, wird die Lage üblicherweise im Rahmen der Kamerakalibrierung ermittelt. In der hier vorliegenden Anwendung mit nichtstationären Kameras muss die Bodenebene entweder aus den Kollateraldaten hervorgehen oder durch Methoden der Georeferenzierung bestimmt werden. Für die direkte Bestimmung der Orientierung existieren inzwischen handelsübliche Flugzeug-getragene Sensoren mit integrierten GPS-Systemen. In [Cramer '03] wird die Qualität eines solchen Sensors anhand der Auswertung zahlreicher Testfeldbefliegungen untersucht: Sie unterscheidet sich nur um Faktor 1,5 bis 2 von der Aerotriangulation, einer klassischen Methode zur Georeferenzierung. Für den Fall, dass Karten für das abgebildete Gelände zur Verfügung stehen, kann alternativ die Karte-zu-Bild Registrierung aus [Krüger '01] herangezogen werden. Für Aufnahmen, die einige bekannte Objekte zeigen, kann diese Orientierung auch aus der Anpassung von einigen rigiden Modellen ermittelt werden. Zusammenfassend wird festgehalten, dass dem System die folgenden Angaben zur Verfügung gestellt werden müssen:

- Die internen Kameraparameter und
- die relative Orientierung der Bodenebene zur Kamera.

Unter dem *ground plane constraint* reduziert sich die Schätzung der Lageparameter auf zwei Translationen und eine Rotation im Bildbereich. Die übrigen drei Lageparameter gehen dann aus der relativen Lage zwischen Bodenebene und Kamera hervor.

5.1.2 Die Bilddaten

Die für die Versuche herangezogenen Bilder stammen von einem Vermessungsflug des Hessischen Landesvermessungsamtes (HVA) aus dem Jahr 1997 über dem Frankfurter Flughafen.¹ Die Bilder wurden nach Angaben des HVA mit einer Messkammer mit einer Brennweite von zwölf Zentimetern bei einer Bildfläche von ca. 23×23 Zentimetern² aufgenommen. Diese Luftbilder liegen in Form von Positiv-Kontaktabzügen vor,

¹Aktennummer 5670-LA 122-67/99 des Hessischen Landesvermessungsamtes.

²Auf die Größe eines handelsüblichen CCD-Sensors umgerechnet entspricht dies einer Brennweite von ca. sechs Millimetern.

von denen jeweils ein zehn Zentimeter breiter Streifen mit 2.500 dpi und einer linearen Kennlinie digitalisiert wurde. Da die Aufnahmen nicht kalibriert sind, müssen die erforderlichen internen Kameraparameter aus diesen kollateralen Angaben berechnet werden.

Die relative Orientierung zwischen der Kamera und der Bodenebene wird als Senkrechtsansicht vorausgesetzt und die Flughöhe durch Anpassung einiger rigider Modelle auf 1.930 Meter geschätzt.³ Dadurch ergibt sich eine mittlere Bodenauflösung von 16 Zentimetern pro Bildpunkt, was hochaufgelösten Luftaufnahmen entspricht [Schlosser et al. '03].

Weiterhin liegen vom Landesvermessungsamt angefertigte Digitalisierungen der kompletten Bildbereiche vor, jeweils in einer Auflösung von 2.000 dpi, wodurch Bildgrößen von circa 18.000×18.000 Bildpunkten erreicht werden. Über weitere Parameter der Digitalisierung bzw. der Quantisierung existieren keine Angaben.

Die Bilder wurden in kurzem zeitlichen Versatz aufgenommen, wodurch eine räumliche Überlappung entsteht (siehe Abbildung 5.2). Diese räumliche Überlappung bietet die Möglichkeit, identische Szenenausschnitte, also beispielsweise ein und dasselbe Flugzeug, in aufeinanderfolgenden Bildern zu vergleichen, um so einen Eindruck der perspektivischen Verzerrungen zu erhalten, die im Randbereich der Bilder auftreten. In Abbildung 5.3 ist eine solche Szene dargestellt, die klar zeigt, wie stark perspektivische Verzerrungen die Abbildung ein und derselben Szene verändern können: Die Drehung eines Flugzeugs im Randbereich des Bildes verlängert die Darstellung des Rumpfes um 18 Bildpunkte bzw. 5,4 Prozent.

Die zur Verfügung stehenden Aufnahmen zeigen hauptsächlich Verkehrsflugzeuge der Hersteller Airbus und Boeing. Nur vereinzelt sind andere Flugzeugtypen abgebildet. Aus den digitalen Bildern wurden mehr als 80 Ausschnitte von Verkehrsflugzeugen bereitgestellt, jeweils in einer Größe von 600×600 Bildpunkten. Mit wenigen Ausnahmen konnte für diese Ausschnitte eine *ground truth*⁴ bezüglich des abgebildeten Flugzeugs erstellt werden, d. h., der jeweilige Flugzeugtyp konnte bestimmt werden. Eine Übersicht über die *ground truth* der vorhandenen Bilder ist in Abbildung 5.4 dargestellt.

5.1.3 Das adaptive Modell und die Repräsentanten

Dem adaptiven Modell kommen in dieser Anwendung zwei Aufgaben zu. Zum einen wird das Modell zur Systemlaufzeit zur Typerkennung eingesetzt, zum anderen werden aus dem Modell die Referenzdaten für den Nächste-Nachbar Klassifikator erzeugt. In beiden Fällen, also sowohl zum Abgleich mit der Datenbasis als auch zur Erzeugung derselben, werden Merkmale aus einer entsprechenden Realisierung des Modells extrahiert. Diese Methode hat den erheblichen Vorteil, dass die Datenbasis kontinuierlich

³Diese rigiden Modelle wurden aus Ausprägungen des adaptiven Modells erstellt, die zuvor durch Anpassung des adaptiven Modells an Schattenrisse erzeugt wurden. Eine ausführliche Beschreibung dieser Ausprägungen folgt in Abschnitt 5.1.3.

⁴Unter der *ground truth* versteht man die wahren Referenzdaten.



Abbildung 5.2: Drei sich überlappende Bilder von einem Vermessungsflug über dem Frankfurter Flughafen.

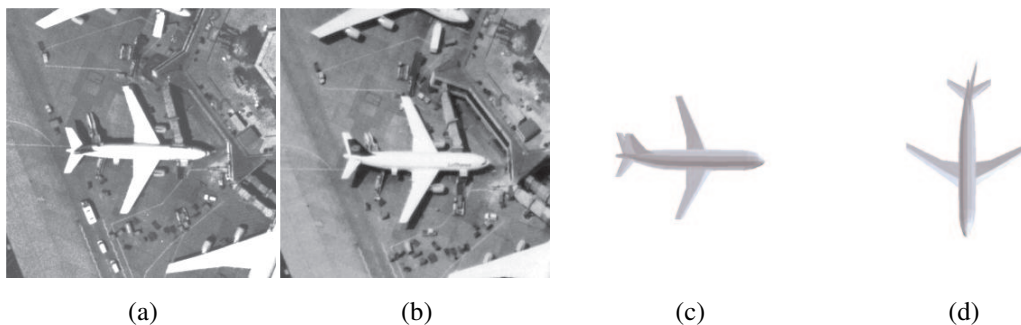


Abbildung 5.3: Perspektivische Verzerrung. Die Bildausschnitte (a) und (b) zeigen dieselbe Szene. Ausschnitt (a) ist der Bildmitte entnommen und Ausschnitt (b) dem Randbereich, wodurch die Schrägansicht entsteht. In (c) ist das Ergebnis einer Modellanpassung an Bild (b) zu sehen. In Bild (d) ist *dasselbe* Modell in *derselben* Lage wie in (c) abgebildet, lediglich der Gierwinkel wurde auf 0 Grad gesetzt. Die Abbildung des Flugzeugrumpfes wird durch diese Drehung um 5,4 Prozent verlängert.

und vor allem automatisch um weitere Repräsentanten erweitert werden kann. Die Interaktion zur Erweiterung der Datenbasis beschränkt sich damit auf die Angabe der Flugzeugbezeichnung. Auf diese Weise wurde z. B. der Airbus A340-200 in die Datenbasis aufgenommen (vgl. Anhang Tabelle C.1).

Für die Initialisierung des Systems, also für die Bereitstellung der initialen Datenbasis, bedeutet dies, dass für alle zu erkennenden Flugzeugvarianten ein Prototyp ausgeprägt werden muss. Bereits im einleitenden Kapitel wurde gezeigt, dass die handelsüblichen Modelle sehr ungenau sind, also nur bedingt den realen Gegebenheiten entsprechen.

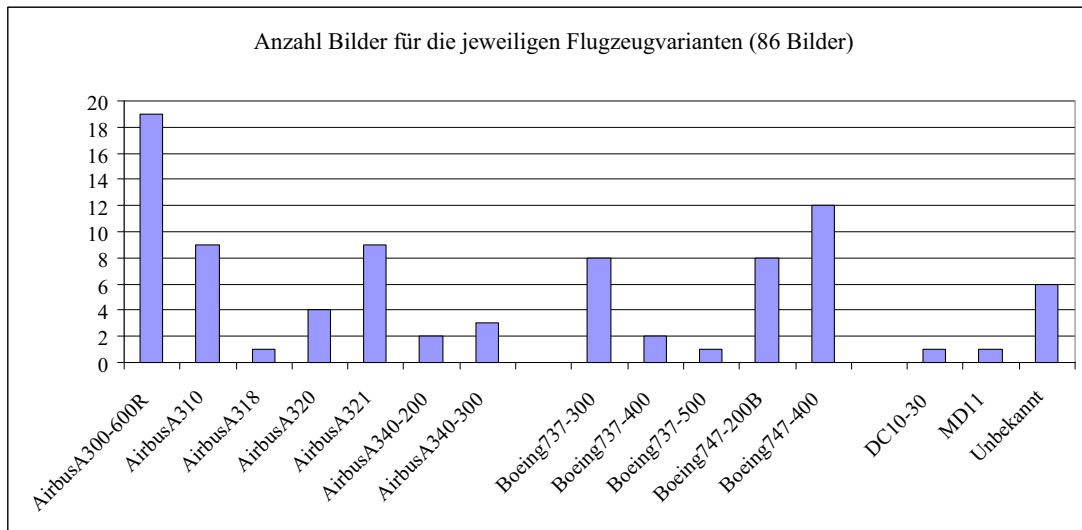


Abbildung 5.4: Die Anzahl der für diese Arbeit verfügbaren Bilder in Abhängigkeit von der abgebildeten Flugzeugvariante.

Infolgedessen können die Basismodelle nicht direkt als Repräsentanten für die Datenbasis herangezogen werden: Ihre Fehler würden sich auf die Datenbasis übertragen. Darüber hinaus wird von jeder Flugzeugvariante, die erkannt werden soll, ein Repräsentant benötigt.

Für das im Rahmen der Versuche eingesetzte adaptive Modell wurden sechs Basismodelle zusammengestellt, deren Auswahl entscheidend ist für die Elastizität des resultierenden Modells. Aus diesem Grund besteht die Basismodellmenge sowohl aus Flugzeugen von Boeing als auch von Airbus. Von beiden Herstellern geht jeweils ein kleiner, ein mittlerer und großer Flugzeugtyp in die Basismodellmenge ein:⁵

- Airbus A300
- Airbus A320-200
- Airbus A340
- Boeing 737
- Boeing 707
- Boeing 747-400

Da die vorhandenen Basismodelle erwartungsgemäß keine ausreichende Elastizität liefern, um verschiedene Varianten dieser Flugzeugtypen auszuprägern, wird das Modell durch die Streckungen des Rumpfes und der Tragflächen erweitert (vgl. Abschnitt 3.3).

⁵Die Flugzeugvarianten der Modelle sind, sofern sie hier nicht angegeben wurden, nicht bekannt.

Die erforderliche Korrektur der Basismodelle und die Ausprägung der fehlenden Varianten wird durch ein *bootstrapping* mit Hilfe des erweiterten adaptiven Modells durchgeführt. Dieses *bootstrapping* wird durchgeführt, indem für jeden benötigten Repräsentanten eine entsprechende Abbildung zur Verfügung gestellt wird und das (orthogonale) erweiterte adaptive Modell an dieses Bild angepasst wird. Die so erhaltenen, neuen geometrischen Realisierungen bilden mit der ursprünglich gegebenen Netztopologie ein neues adaptives Modell und die geometrischen Realisierungen entsprechen den erforderlichen Repräsentanten.

In diesem Zusammenhang muss man sich die Frage stellen, welches Bildmaterial für die Modellausprägungen zur Bereitstellung der Datenbasis herangezogen werden kann. Im Prinzip können sowohl digitale Schattenrisse als auch kalibrierte Realaufnahmen eingesetzt werden. Im Allgemeinen jedoch sind Schattenrisse den realen Aufnahmen vorzuziehen, da sie keine systematischen Störungen wie etwa Schatten oder vergleichbare Artefakte aufweisen. Darüber hinaus besitzen sie den Vorteil, dass eine Orthogonalprojektion vorausgesetzt werden kann und die Bildpunktauflösung aus den technischen Angaben über die Spannweite gewonnen werden kann. Andererseits darf an dieser Stelle nicht verschwiegen werden, dass sowohl die Prototypenausprägung auf realen Bildern als auch die auf Schattenrissen mit geometrischen Unsicherheiten behaftet ist, die nicht kontrollierbar sind. Bei den Schattenrissen ist dies zum einen durch die Ungenauigkeiten begründet, die der Schattenriss selbst mit sich bringt, und zum anderen können auch hier Digitalisierungsartefakte nicht ausgeschlossen werden.

Für das in den Versuchen eingesetzte Modell wurden Schattenrisse aus [Green '92, Green '80] und von der Internetseite *airliners.net*⁶ sowie eine Realaufnahme eines Airbus A318 verwendet, um 15 Repräsentanten der Hersteller Airbus und Boeing auszuprägen. Die entsprechenden Tyangaben sind im Anhang in Tabelle C.1 aufgeführt. Die Modellanpassung an die Schattenrisse wurde mit der iterativen Optimierung durchgeführt, die in Abschnitt 4.4 eingeführt wurde. Die Startwerte für die Modellausprägung wurden manuell auf einen ähnlichen Flugzeugtyp gesetzt.

Zwei Ergebnisse dieser Modellausprägungen sind in Abbildung 5.5 (a) und (b) zu sehen. Abbildung 5.5 (c) zeigt zum Vergleich nochmals die Anpassung eines handelsüblichen Modells. Die Verbesserung der Modellqualität durch das adaptive Modell ist offensichtlich, da die Anpassungen durch das adaptive Modell Fehler von weniger als vier Bildpunkten hat, während das handelsübliche Modell einen Fehler von 22 Bildpunkten liefert (siehe auch Abbildung 1.3). Eine quantitative Bewertung der Repräsentanten wird in Abschnitt 5.1.4 im Zusammenhang mit der Merkmalsextraktion vorgenommen.

5.1.4 Merkmalsextraktion, Datenbasis und Typerkennung

In dieser Arbeit werden drei Merkmalstypen verglichen. Der erste Merkmalstyp wird im Folgenden mit dem Begriff „3D-Merkmale“ bezeichnet, denn er wird durch eine normalisierte Projektion der Repräsentanten erzeugt, die nur durch die Verwendung

⁶<http://www.airliners.net/>

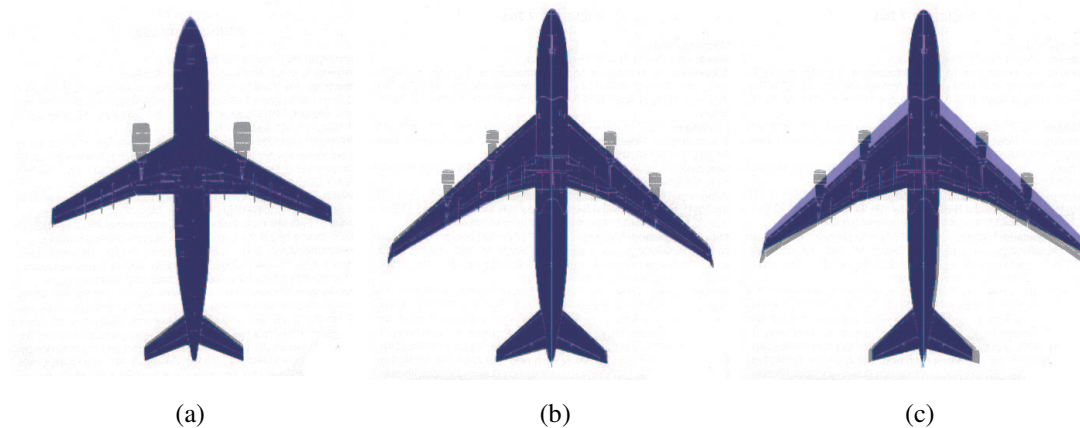


Abbildung 5.5: In (a) ist das Ergebnis der Modellausprägung für den Airbus A300 und in (b) für die Boeing 747-400 zu sehen. Abbildung (c) zeigt die Anpassung eines handelsüblichen Modells. Die Fehler bei den Anpassungen des adaptiven Modells betragen maximal vier, die des handelsüblichen Modells bis zu 22 Bildpunkte.

dreidimensionaler Modelle möglich ist. Zur Extraktion der Merkmale wird zunächst ein zeichnungsfreier, synthetischer Schattenriss in einer Senkrechtsansicht erzeugt. Aus diesem normalisierten Schattenriss werden dann mit Hilfe eines im Vorfeld dieser Arbeit implementierten Verfahrens (siehe [Berger et al. '00]) fünf trennungswirksame Größen extrahiert (vgl. Abbildung 5.6). Darunter sind vier metrische Größen sowie der Pfeilungswinkel α . Um den Nächsten-Nachbar Klassifikator anwenden zu können, müssen diese Merkmale skaliert werden. Andernfalls beeinflusst z. B. die Maßeinheit, in der die Merkmale vorliegen, das Klassifikationsergebnis. In der Regel wird diese Skalierung aus den Merkmalsvektoren selbst bestimmt. Dies setzt jedoch voraus, dass eine ausreichende Anzahl an Merkmalsvektoren, also letztlich eine große Anzahl an Trainingsbeispielen vorliegt. Da dies hier jedoch nicht der Fall ist, wird eine andere Skalierung vorgenommen, die den Pfeilungswinkel in eine metrische Größe umrechnet: Bezeichnet b die Tragflächenlänge, so wird der Pfeilungswinkel via $b \cdot \cos(\alpha)$ und $b \cdot \sin(\alpha)$ in den Merkmalsvektor integriert. Der Pfeilungswinkel geht also zweifach in den Merkmalsvektor ein und es liegen insgesamt sechs metrische Merkmale vor.

Die Fehler, die durch die Konstruktion der Repräsentanten und der anschließenden Merkmalsextraktion entstehen, lassen sich nur durch einen Vergleich der extrahierten Rumpflängen mit entsprechenden technischen Angaben ermitteln, da für die übrigen Merkmale keine Vergleichsdaten vorliegen. In Tabelle 5.1 sind die Ergebnisse eines solchen Längenvergleichs festgehalten. Auffällig ist hierbei, dass die Differenzen zwischen den Rumpflängen der 3D-Rekonstruktion und den aus dem synthetischen Segment extrahierten Längen gering ist, wohingegen die 3D-Rekonstruktionen teilweise deutliche Differenzen zu den technischen Angaben aufweisen. Für den Airbus A300 beispielsweise beträgt der Fehler in der Modellrekonstruktion 1,3 Meter. Diese Fehler lassen sich nur auf die Ungenauigkeiten des Schattenrisses zurückführen, da die Modellanpassung einen Fehler von maximal vier Bildpunkten hat (vgl. Abbil-

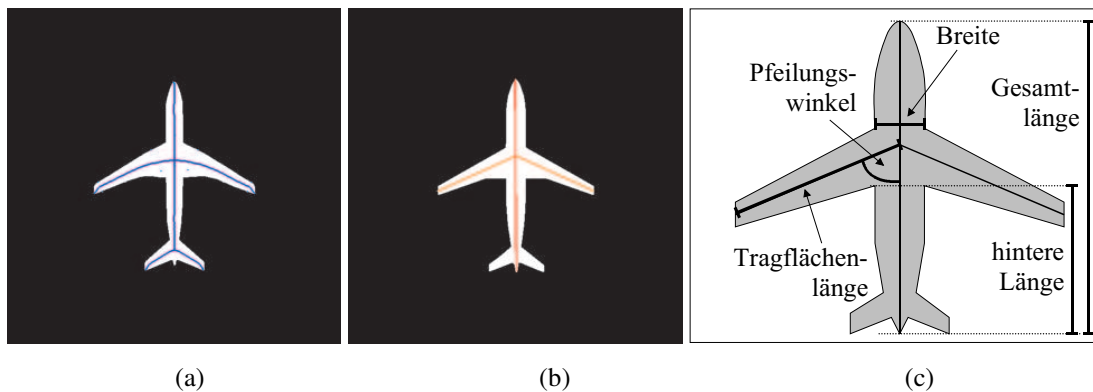


Abbildung 5.6: Merkmale aus Segmenten. Aus dem synthetisch erzeugten Segment wird ein Skelett berechnet (a). Aus diesem Skelett werden Geradensegmente extrahiert (b) und daraus wiederum die fünf benannten Merkmale extrahiert (c).

[m]	technische Angaben	Modell	Datenbank	Fehler
Airbus A300-600R	54,1	55,4	55,5	-1,4
Airbus A310	47,7	47,9	47,8	0,1
Airbus A318	31,5	32,0	31,8	0,3
Airbus A320	37,6	37,7	37,6	0,0
Airbus A321	44,5	44,6	44,6	0,1
Airbus A340-200	58,4	59,5	59,2	0,8
Airbus A340-300	63,7	63,1	62,6	-1,1
Boeing 737-300	33,4	33,5	33,1	-0,3
Boeing 737-400	36,4	36,6	36,2	-0,2
Boeing 737-500	31,0	30,7	30,4	-0,6
Boeing 737-600	31,2	31,4	31,6	0,4
Boeing 737-700	33,6	33,7	33,2	-0,4
Boeing 707-320B	46,6	47,5	47,7	1,1
Boeing 747-200B	70,7	71,0	70,6	-0,1
Boeing 747SP	56,3	56,7	56,2	-0,1
Boeing 747-400	70,7	70,1	70,5	-0,2

Tabelle 5.1: Die extrahierten Längen der Flugzeuge im Vergleich zu den technischen Daten. Die Spalte „Modell“ zeigt die Länge der dreidimensionalen Rekonstruktion und die Spalte „Datenbank“ die Längen, die aus der zweidimensionalen Merkmalsextraktion hervorgehen. Die Spalte „Fehler“ bezeichnet die signierte Differenz zwischen den technischen Angaben und dem Datenbankeintrag. (Quelle technische Angaben: [Figgen et al. '00]).

dung 5.5). Bei einer Bildauflösung von 0,12 Meter pro Bildpunkt bedeutet das, dass die Modellanpassung einen Fehler von maximal 0,48 Metern geliefert haben kann. Der ebenfalls vergleichsweise große Fehler für den Airbus A340-200 lässt sich dadurch erklären, dass die Modellrekonstruktion aus unkalibriertem Bildmaterial erfolgt ist.

Um einen Vergleich zu rein bildbasierten Ansätzen ziehen zu können, wird derselbe Extraktionsalgorithmus auf ein weiteres synthetisches Segment angewandt. Dieses Segment wird durch Projektion des Modells in der geschätzten Lage auf die Bildebene erzeugt. Es entspricht also dem Ergebnis, das durch einen rein bildbasierten Ansatz erreicht werden kann. Der Merkmalstyp wird entsprechend mit „2D-Merkmale“ bezeichnet. Einen dritten Merkmalstyp, die „Modell-Merkmale“, liefern letztlich die Parameter des orthogonalen Modells.

Während die metrischen Merkmale anhand des euklidischen Abstandes verglichen werden, wird für die Modellmerkmale die Mahalanobis-Distanz herangezogen. Das bedeutet, dass der Abstand zweier Modellmerkmalsvektoren θ und θ' gegeben ist durch

$$d_M(\theta, \theta') = \sum_{n=1}^N \frac{(\delta_n - \delta'_n)^2}{\sigma_n^2}, \quad (5.1)$$

wobei die σ_n^2 die Varianzen bezüglich der Hauptachsen bezeichnen (vgl. Abschnitt 3.1). Der Vertrauenswert bezüglich eines Merkmalstyps ist wie folgt erklärt: Bezeichnen d_0, d_1, \dots die aufsteigend sortierten Abstände des extrahierten Merkmalsvektors zu den in der Datenbasis abgelegten Merkmalsvektoren, so ist der Vertrauenswert der Typerkennung durch

$$q = 1 - \frac{d_0}{d_1} \quad (5.2)$$

definiert. Wird ein Merkmalsvektor extrahiert, der zu zwei Einträgen der Datenbasis denselben minimalen Abstand hat, so ist dieser Vertrauenswert Null. Stimmt der extrahierte Merkmalsvektor genau mit einem in der Datenbasis abgelegten Merkmalsvektor überein, so ist der Vertrauenswert Eins. In anderen Worten: Der Vertrauenswert ist klein, wenn die Erkennung unsicher ist, weil das beobachtete Objekt zu zwei der in der Datenbasis vorhandenen Objekten eine vergleichbare Ähnlichkeit aufweist. Dieser Fall tritt z. B. dann auf, wenn sich zwei in der Datenbank eingetragene Flugzeugvarianten nur in der Rumpflänge unterscheiden und die aus dem Bild gemessene Rumpflänge mittig zwischen den beiden Varianten liegt. Ein Vertrauenswert nahe Eins hingegen indiziert, dass die Erkennung als sicher einzustufen ist.

5.1.5 Die Modellanpassung

Das für die Anpassung adaptiver Modelle an Realaufnahmen entwickelte System besteht aus zwei Komponenten. Die erste dieser Komponenten ist die Initialisierung der Modell- und Lageparameter und die zweite die Optimierung der initialen Parameter. Letztere ist durch die Anwendung der in dieser Arbeit entwickelten iterativen Optimierung realisiert (vgl. Abschnitt 4.4). Die Initialisierung ist in Abbildung 5.7 skizziert und wird im Folgenden besprochen.

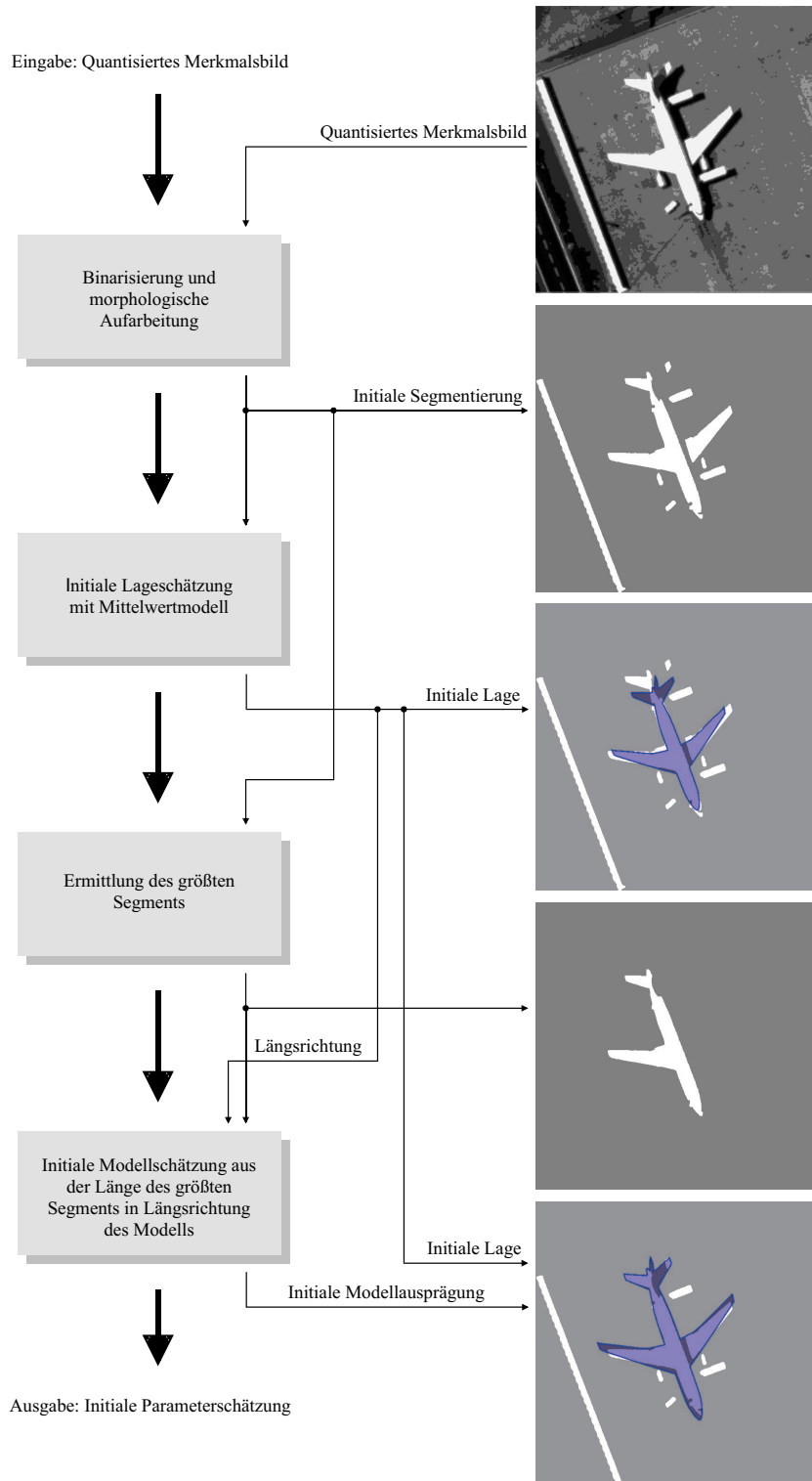
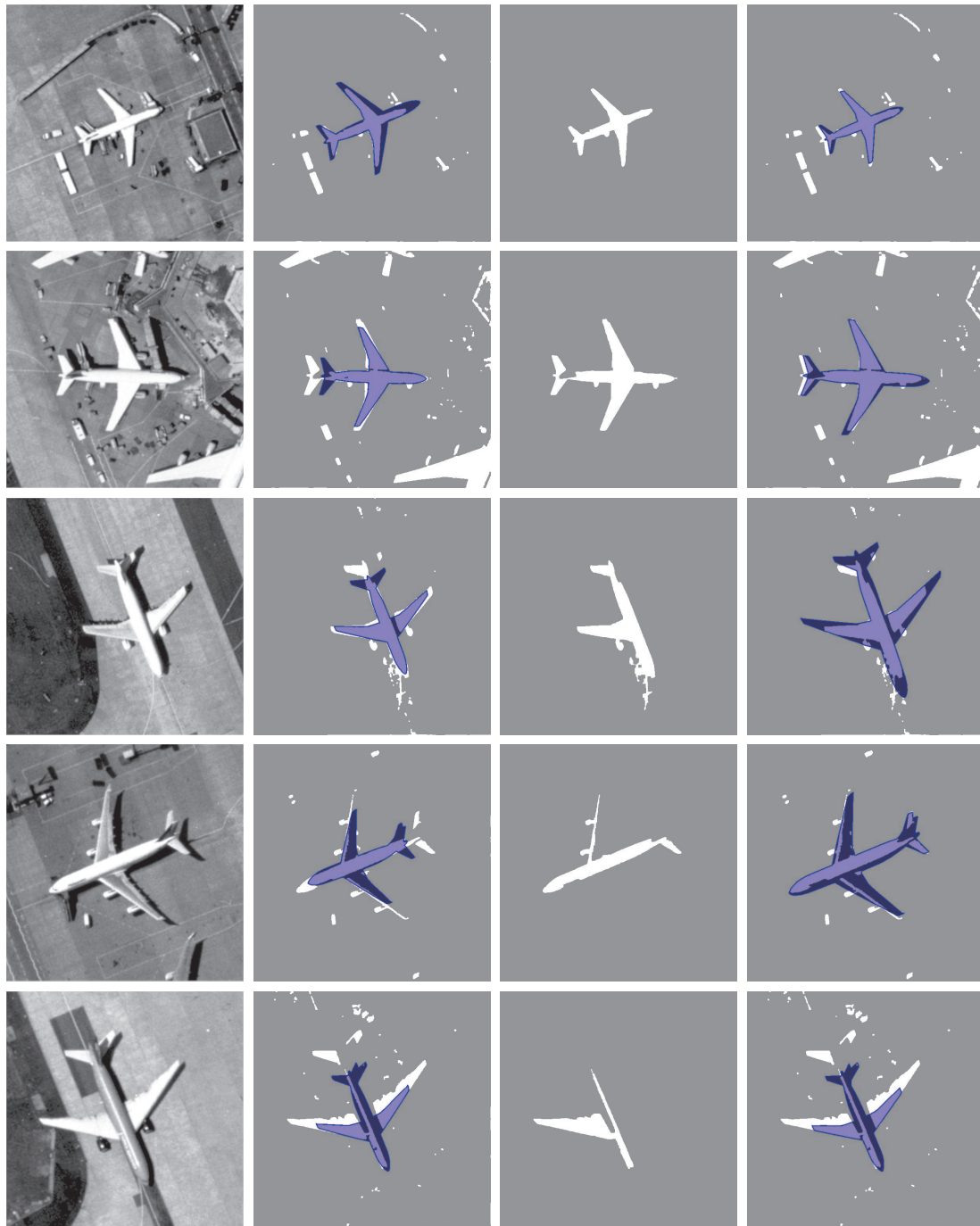


Abbildung 5.7: Initiale Parameterschätzung. Ausgehend von einem quantisiertem Merkmalsbild (hier aus dem Intensitätsbild berechnet), werden die initialen Lage- und Modellparameter geschätzt.



(a) Eingangsbild (b) Initiale Lage (c) Größtes Segment (d) Initiales Modell

Abbildung 5.8: Beispiele für die initiale Parameterschätzung. Die oberen zwei Reihen zeigen typische Bilder und Ergebnisse. In der dritten Reihe ist das größte Segment „ausgelaufen“ und das initiale Modell daher deutlich zu lang. In der vierten Reihe ist ein Beispiel für vergleichsweise dunkle Tragflächen zu sehen und in der fünften ein Beispiel für einen dunklen Rumpf. Letzteres resultiert als einziges der hier gezeigten Bilder in einer fehlerhaften Modellanpassung.

Lageparameter

Die Initialisierung der Lage ist unter Verwendung des robusten Schätzers realisiert und zwar durch Diskretisierung des Parameterraums in einer Dimension. Aus den Diskussionen über die Gütefunktion in Kapitel 4 geht hervor, dass die Kostenfunktion mit einem groß eingestellten Filter für Flugzeuge vier lokale Extremwerte besitzt, wenn sich das Flugzeug um die eigene Achse dreht (vgl. Abbildungen 4.4 und 4.6). Diese Extremwerte sind nahezu äquidistant, da sie durch die Kreuzform der Flugzeuge verursacht werden. Daraus lässt sich folgern, dass sich das Optimum mit sechs äquidistanten Abtastungen dieser Drehung finden lässt (sofern die translativen Anteile der Lage und die Parameter des Modells bekannt sind). Das daraus abgeleitete Prinzip der Initialisierung lautet nun wie folgt: Für die Initialisierung wird das adaptive Modell in seiner mittleren geometrischen Realisierung ausgeprägt und die robuste Parameterschätzung auf sechs Startwerte angewandt. Diese Startwerte werden so gesetzt, dass das Modell in der Mitte des Bildausschnitts positioniert ist und in einer Draufsicht abgebildet wird. Die sechs Startlagen unterscheiden sich nur in Bezug auf eine Drehung um den Sehstrahl und diese Drehung wird auf die sechs Winkel $k \cdot 60$ Grad, $k = 0, \dots, 5$ gesetzt. In den sechs Optimierungsläufen werden dann nur die drei Lageparameter, die unter dem *ground plane constraint* zu schätzen sind, frei gelassen und die Modellparameter bleiben fest. Das Ergebnis mit dem besten Vertrauenswert (der Modellanpassung) wird als Initialschätzung für die Lage verwendet.

Unterdrückung von Schattenkanten

Wie jedoch ebenfalls in Kapitel 4 am Beispiel von Abbildung 4.5 erklärt wurde, verursachen Schlagschatten ebenfalls Extremwerte in der Gütefunktion. Wird die Lage schätzung in der oben beschriebenen Weise durchgeführt, so kann nicht ausgeschlossen werden, dass die initiale Lage ganz oder teilweise auf den Schattenkanten liegt. Die Idee, durch Schattenwurf verursachte Kanten anhand der angrenzenden Grauwerte zu identifizieren und für die Lageschätzung auszublenden, wäre die logische Konsequenz. Nun muss man sich vor Augen halten, dass zwar die beschatteten Bereiche anhand der Intensität detektierbar sind, diese Bereiche jedoch sowohl an den Hintergrund als auch an das Objekt angrenzen. Das bedeutet, dass ein Teil der an den Schatten grenzenden Kanten ausgeblendet werden soll, der andere Teil, nämlich die Kanten, die an das Objekt grenzen, hingegen nicht. Daraus wiederum folgt, dass a priori Wissen über die Beschaffenheit von Hintergrund und Objekt eingesetzt werden muss. Andernfalls können die Kanten nicht unterschieden werden. Genau von solchem a priori Wissen wird zur Initialisierung Gebrauch gemacht, allerdings nicht in der oben beschriebenen Form, sondern durch Anwendung der Initialisierung auf ein Binärbild:

Zunächst wird vorausgesetzt, dass zu dem gegebenen Bild ein Merkmalsbild existiert, d. h. ein Bild, dessen Einträge einen großen Wert für diejenigen Bildpunkte haben, deren Textur charakteristisch ist für das Objekt. Aus welchen Bildmerkmalen dieses Bild letztlich erzeugt wurde, ist sekundär. Für das hier vorliegende Bildmaterial wird die Intensität eingesetzt, also das Bild selbst. Die in der Aufgabenstellung genannte Sys-

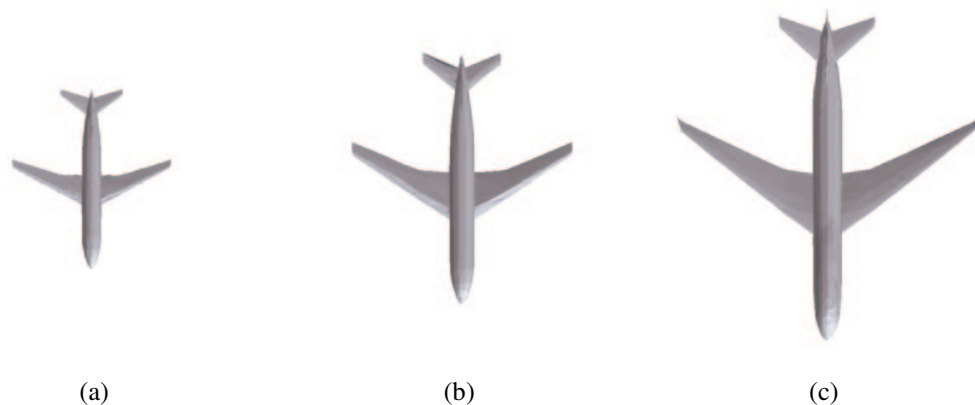


Abbildung 5.9: Modelldiskretisierung für die Parameterinitialisierung. In (a) ist die Interpolation zwischen einer Boeing 737 und einem Airbus A320 zu sehen, in (b) die zwischen einer Boeing 707 und einem Airbus A300 und in (c) die zwischen einer Boeing 747 und einem Airbus A340. Diese Ausprägungen werden zur Modellinitialisierung noch weiter interpoliert, so dass für die Initialisierung der Modellparameter insgesamt 13 Ausprägungen unterschiedlicher Länge zur Verfügung stehen.

temanforderung nach Unabhängigkeit von der Materialbeschaffenheit wird dadurch nur insofern verletzt, dass keine allgemeingültige Methode zur Auswahl geeigneter Bildmerkmale bereitgestellt wurde.

Anstatt nun die Initialisierung der Lageparameter auf dem gegebenen Bild vorzunehmen, wird das Merkmalsbild mit dem Verfahren aus [Comaniciu & Meer '97] in maximal acht Stufen quantisiert. Dieses quantisierte Bild wird binarisiert, also auf zwei Quantisierungsstufen reduziert, indem alle Bildpunkte, die den Maximalwert der Quantisierung besitzen, auf Eins gesetzt werden und alle übrigen auf Null.

Formparameter

Bleibt noch die Initialisierung der Modellparameter zu erklären, die im Anschluss an die initiale Lageschätzung erfolgt. Hierfür wird aus dem Binärbild nach einer morphologischen Öffnung das größte Segment bestimmt. In der Richtung, in der das Flugzeugmodell nach der initialen Lageschätzung liegt, wird nun die Länge des Segments bestimmt. Diese Länge wird unter Verwendung der Kollateraldaten in Meter umgerechnet und anhand dieser Länge eine Parametrierung des adaptiven Modells ausgewählt. Für diese initiale Parametrierung werden die Modellparameter im Vorfeld der Modellanpassung diskretisiert.

Für das im Rahmen der Flugzeugerkennung eingesetzte adaptive Modell beispielsweise wurde zunächst ein kleines, ein mittleres und ein großes virtuelles Flugzeugmodell ausgeprägt (vgl. Abbildung 5.9). Für diese Ausprägungen wurden jeweils zwei Flugzeugtypen gemittelt, wobei eines vom Hersteller Airbus und das andere von Boeing ausgewählt wurde. Auf diese Weise geht die Charakteristik beider Flugzeugherstel-

ler in jede der initialen Modellausprägungen ein. Das bedeutet z. B. für die Tragflächenform, dass diese eine Mittelung aus den für Boeing und Airbus charakteristischen Merkmalen ist. Auf diese Weise wird die nachfolgende Modellanpassung erleichtert, da die Anpassung der Tragflächen damit für alle vorkommenden Flugzeugtypen in etwa dieselbe Deformation erfordert. Um die Initialisierung weiter zu verbessern, wurden in einem zweiten Schritt noch fünf lineare Interpolationen zwischen der kleinen und der mittleren Ausprägung und fünf zwischen der mittleren und der großen Ausprägung erzeugt. Auf diese Weise stehen zur Modellinitialisierung insgesamt 13 Modellausprägungen mit einer mittleren Längendifferenz von 2,4 Metern zur Verfügung. Aus diesen Ausprägungen wird also anhand der aus dem Bild geschätzten Flugzeuglänge diejenige ausgewählt, für die die Längendifferenz am kleinsten ist.

In Abbildung 5.8 werden einige Beispielergebnisse der Initialisierungsmethode präsentiert. Diese Beispiele zeigen, dass die Lageschätzung sehr robust ist, auch wenn das Binärbild keine gute Segmentierung liefert. Tatsächlich wird die Orientierung des Flugzeugs für den vollständigen Bilddatensatz korrekt gefunden. Die Schätzung der Modellparameter aus der Segmentlänge hingegen führt mitunter zu Initialisierungen, die die Optimierung in einem lokalen Optimum konvergieren lässt, wie im folgenden Abschnitt gezeigt wird.

5.1.6 Ergebnisse der Modellanpassung

Abbildung 5.10 zeigt fünf Beispielergebnisse für die Modellanpassung. Die oberen drei Reihen zeigen typische Szenarien, wobei die ersten beiden Beispiele komplexe Szenarien abbilden, wobei die ersten beiden Beispiele komplexe Szenarien abbilden, was an den in Spalte (c) dargestellten Bildkantensegmenten deutlich wird. Das erste Beispiel zeigt zudem in der Nähe des Flugzeugs befindliche Fahrzeuge, die ähnliche Strukturen aufweisen wie das Flugzeug selbst. Dies wird bereits dadurch deutlich, dass die Segmentierung für die Initialisierung für diese Fahrzeuge ebenfalls große Segmente extrahiert. Solche Effekte sind in texturbasierten Segmentierungsansätzen nur schwer zu berücksichtigen, wohingegen diese bei dem hier gewählten Ansatz keinen Einfluss auf die Modellanpassung haben. Dies ist einerseits auf die Verwendung von Kanteninformationen und andererseits auf das in Form eines Modells eingesetzte Szenenwissen zurückzuführen.

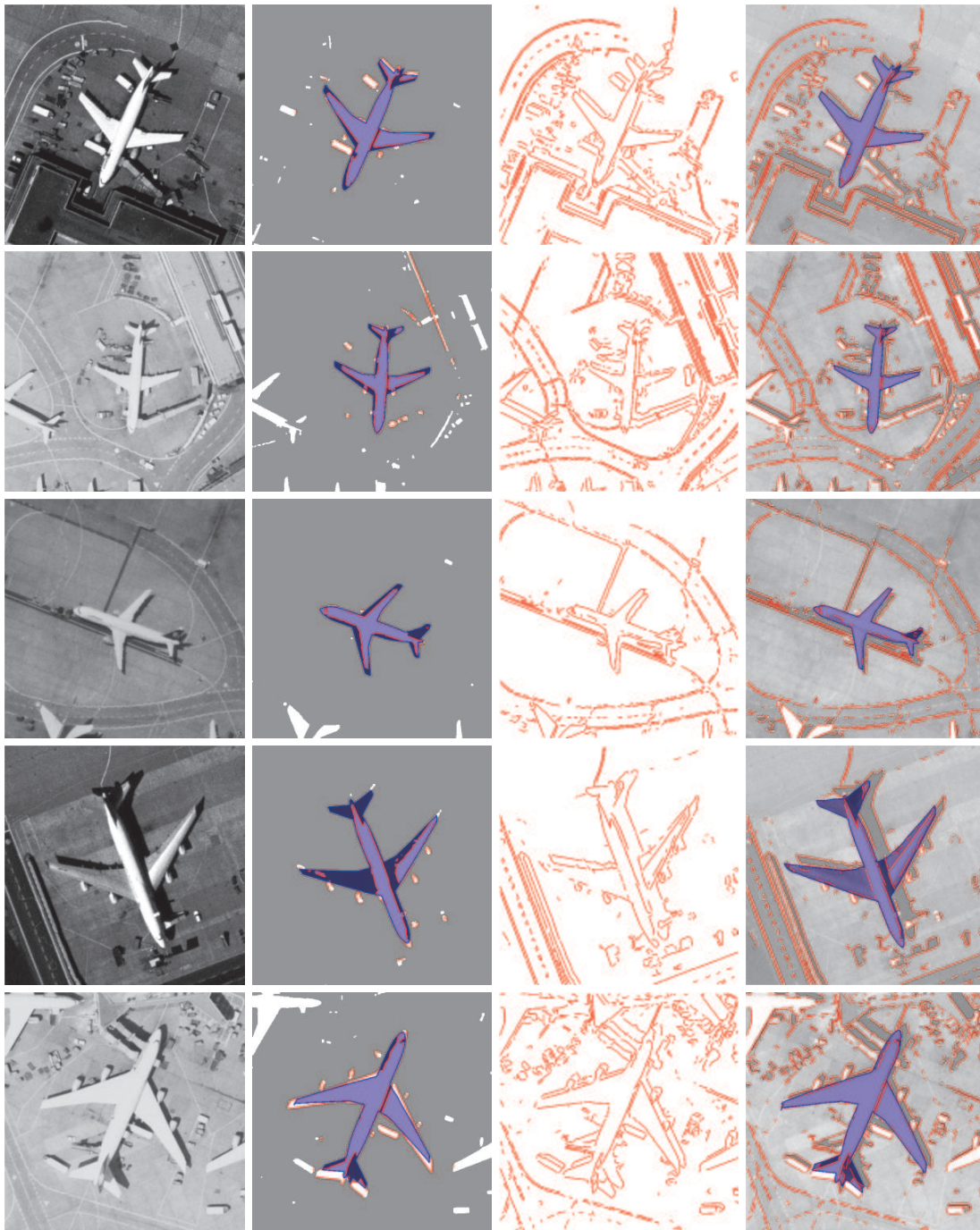
Die dritte Reihe ist ein gutes Beispiel für die Modellanpassung an eine Schrägansicht. Das Beispiel in der vierten Reihe hingegen zeigt deutlich, dass die Modellanpassung robust ist gegenüber unvollständigen Bildkantensegmenten. Unvollständig bedeutet hier, dass die Modellkantensegmente nicht vollständig durch Bildkantensegmente abgedeckt sein müssen. In diesem Beispiel ist diese partielle Extraktion der Objektkontur auf den geringen Kontrast zwischen Hintergrund und Objekt zurückzuführen: Für das links abgebildete Höhenruder existieren keine Bildkantensegmente und auch die Tragfläche ist nur lückenhaft durch die Bildkantensegmente gegenüber dem Hintergrund abgegrenzt. Die in diesem Beispiel stark ausgeprägte Eigenbeschattung auf dem Tragflächenansatz ist für den in dieser Arbeit vorgestellten Ansatz völlig bedeutungslos.

Das letzte Beispiel hingegen zeigt die größte Schwäche des Verfahrens: Die Initialisierung der Modellparameter liefert eine Modellausprägung mit zu kurzem Rumpf, die durch die Optimierung nicht kompensiert wird. Vielmehr konvergiert die Optimierung auf Teilen des Schlagschattens vom Höhenruder. Die Tragflächen hingegen wurden gut an die extrahierten Bildkanten angepasst. Diese Modellanpassung lässt sich *nicht* anhand des Vertrauenswertes der Modellanpassung von den gelungenen Rekonstruktionen unterscheiden. Diese Tatsache ist darin begründet, dass der Vertrauenswert als prozentuale Überdeckung der Modellkantensegmente durch Bildkantensegmente interpretiert werden kann. Diese Überdeckung ist in dem hier gezeigten Fall jedoch sehr gut, vor allem, wenn man sie mit der lückenhaften Überdeckung des Beispiels in der vierten Reihe vergleicht. Kurz, der Vertrauenswert der Modellanpassung ist ein bildabhängiges Maß und bietet keinen bildunabhängigen Aufschluss über das Gelingen der Modellanpassung.

Eine große Stärke des hier vorgestellten Verfahrens liegt also in der Toleranz der Modellanpassung gegenüber unvollständig extrahierten Bildkantensegmenten, und zwar unabhängig von den Ursachen. Als Ursachen kommen hier unter anderem schlecht eingestellte Merkmalsextraktionsverfahren, aber auch geringer Kontrast zwischen Objekt und Hintergrund sowie partielle Verdeckungen des Objekts in Frage. Andererseits zeigt die Komplexität einiger Szenen, dass auch Störkanten, seien sie systematischer Natur oder durch Texturen verursacht, von der Modellanpassung gut toleriert werden. Beide Toleranzen sind auf zwei Tatsachen zurückzuführen: Zum einen verlangt die Gütefunktion des robusten Schätzers keine vollständig extrahierten Bildkantensegmente. Zum anderen sorgt die Filterfunktion der Gütefunktion in Verbindung mit der durch die Basismodelle definierten Elastizität des adaptiven Modells dafür, dass Störkanten herausgefiltert und zum Objekt gehörende Bildkantensegmente als solche erkannt werden können.

Im Zusammenhang mit der Modellanpassung kann nun auch die Frage beantwortet werden, ob die Elastizität der adaptiven Modelle ausreichend ist, um den Basismodellen ähnliche Objekte darstellen zu können. Diese Frage lässt sich durch die Anpassung des Modells an unbekannte Flugzeugtypen beurteilen, also an solche Flugzeugtypen, die nicht in Form eines Basismodells in das adaptive Modell eingegangen sind. Abbildung 5.11 zeigt vier Beispiele solcher Flugzeugtypen. Die Qualität dieser Modellanpassungen ist sehr gut, insbesondere wenn man den Aspekt berücksichtigt, dass der Hersteller McDonnell Douglas (erste und letzte Zeile) nicht in Form eines Basismodells vertreten war. Offensichtlich lässt sich die Form dieser Flugzeuge aus den Formen der Airbus- und Boeing-Produkte linear interpolieren. Zudem ist in der untersten Zeile von Abbildung 5.11 ein Flugzeug mit T-Leitwerk abgebildet. Das Ergebnis der Modellanpassung zeigt, dass das adaptive Modell in dieser Draufsicht über die ihm unbekannte Leitwerksform hinweggeht, was in einer Seitenansicht nicht der Fall sein kann. Zudem wird an diesen Beispielen deutlich, dass eine schlechte Modellinitialisierung nicht in einer Fehlanpassung resultieren *muss*, sondern durchaus von der Modellanpassung kompensiert werden kann.

Zusammenfassend kann festgehalten werden, dass die Linearität des adaptiven Modells eine ausreichende Elastizität bietet um auch unbekannte Objekttypen modellie-



(a) Eingangsbild

(b) Initiale Lage

(c) Bildkanten

(d) Ergebnis

Abbildung 5.10: Beispielergebnisse für die Modellanpassung. Die oberen zwei Reihen zeigen typische Eingangsbilder und Ergebnisse, die dritte Reihe eine Seitenansicht. Die vierte Reihe ist ein Beispiel für die Toleranz des Verfahrens gegenüber unvollständig extrahierten Bildkantensegmenten. Die fünfte Reihe zeigt eine Modellanpassung, die nach schlechter Modellinitialisierung die Länge falsch rekonstruiert.

ren zu können. Andererseits ist die Elastizität kontrolliert. Insbesondere werden unvollständig extrahierte Bildkantensegmente von dem Modell ergänzt, ohne dabei die Charakteristik der Basismodelle zu verlieren. Dies ist ein entscheidender Vorteil des hier vorgestellten Verfahrens gegenüber anderen deformierbaren Modellen, wie z. B. den aktiven Konturen und texturbasierten Segmentierungsansätzen. Das Szenenwissen, das dem Modell in Form von Basismodellen zur Verfügung gestellt wird, wird genutzt, um partiell schwach ausgeprägte texturelle Unterschiede zwischen Objekt und Hintergrund zu tolerieren. Aus dieser Argumentation folgt direkt die Begründung für die Toleranz des Modellanpassungssystems gegenüber den Schwellwerten der Bildkantenextraktion.

Die mittlere Rechenzeit der im Rahmen dieser Arbeit durchgeführten Modellanpassungen liegt bei 118 Sekunden auf einem Pentium IV 1,9 GHz. Diese hohe Rechenzeit ist hauptsächlich auf zwei Faktoren zurückzuführen, nämlich auf die hohe Anzahl von Modellkantensegmenten und auf die Verwendung des Passpunktschätzers. Im Zusammenhang mit der im späteren Verlauf erklärten Lageschätzung wird diese Aussage detaillierter besprochen. In jedem Fall könnte die Rechenzeit drastisch reduziert werden, wenn die Schablonenanpassung in der Passpunktschätzmethode ausgetauscht würde.

5.1.7 Die Typerkennung

Grundsätzlich wird die Leistung einer automatischen Typerkennung an dem Prozentsatz der korrekt erkannten Objekte gemessen. Betrachtet man nun die Varianten des Airbus A300, dann stellt man fest, dass sich diese nur um einen halben Meter in der Länge unterscheiden. Auch bei größeren Längendifferenzen, wie z. B. den 2,4 Metern zwischen der Boeing 737-600 und der Boeing 737-700, ist es wünschenswert, bei der Leistungsmessung zwischen einer Falsch-Identifikation einerseits und einer fälschlicherweise als 600er Variante erkannten Boeing 737-700 zu differenzieren. Dieser Tatsache wurde Rechnung getragen, indem eine weitere Hierarchieebene eingeführt wurde, bezüglich derer die Leistung - zusätzlich zur Variantenerkennung - bewertet wurde. Diese Hierarchieebene gruppiert ähnliche Varianten desselben Flugzeugtyps. Eine entsprechende Gruppenübersicht ist im Anhang in Tabelle C.1 zusammengefasst.

Für den Versuch standen nun insgesamt 86 Bilder zur Verfügung, wobei auf acht Bildern Flugzeugtypen abgebildet sind, für die kein Eintrag in der Datenbank existiert, die dem System also unbekannt sind. Auf den verbleibenden 78 Bildern wurde eine Typerkennung durchgeführt, wobei die Modellanpassung für ein Bild aufgrund eines zu kleinen Startsegments abgebrochen wurde. Die Ergebnisse für die 77 Bilder sind in Tabelle 5.2 festgehalten und zwar für alle drei Merkmalstypen 3D, 2D und Modell. Die Angaben beziehen sich jeweils auf die korrekte Erkennung der Flugzeugvariante und der Gruppe.

Zunächst ist auffällig, dass die Erkennungsrate unter Verwendung der 2D-Merkmale extrem viel niedriger ist als die Rate, die mit den 3D-Merkmalen erzielt wurde. Daraus muss man folgern, dass die Typerkennung auf den hier vorliegenden Bildern mit einem rein zweidimensionalen Ansatz nicht möglich ist. Auch die Erkennungsrate an-

	3D-Merkmale		2D-Merkmale		Modell-Merkmale	
	Variante	Gruppe	Variante	Gruppe	Variante	Gruppe
Korrekt erkannt	64	70	2	48	21	27
Falsch identifiziert	13	7	75	29	56	50
Erk.-Rate (Prozent)	80,3	90,9	2,6	62,3	27,3	35,1

Tabelle 5.2: Die Tabelle zeigt die Ergebnisse der Typerkennung mit drei verschiedenen Merkmalstypen im Vergleich. Insgesamt wurden 77 Bilder ausgewertet. In den Zeilen „Korrekt erkannt“ und „Falsch identifiziert“ ist die entsprechende Anzahl an Bildern angegeben.

hand der Modellparameter ist extrem niedrig. Dieser Sachverhalt ist auf die bereits angesprochene Empfindlichkeit der Modellparameter gegenüber den zur Anpassung extrahierten Bildmerkmalen zurückzuführen.

Betrachtet man beispielsweise einen Flugzeugrumpf, der durch schräg einfallendes Licht seitlich abschattiert wird, so wird die Stelle des größten Grauwertanstiegs nicht auf der Silhouette liegen, sondern leicht versetzt auf dem Rumpf selbst. Auf diese Weise können Bildkantensegmente leicht um einige Bildpunkte verfälscht werden. Für die Modellanpassung bedeutet dies, dass ein Modell mit einem Rumpf ausgeprägt wird, der vielleicht zwei Bildpunkte schmaler ist als tatsächlich abgebildet. Bei der hier vorliegenden Bildauflösung würde das einen Größenunterschied von 0,32 Metern ausmachen. Diese Größenordnung darf aber für die Typerkennung keine Rolle spielen, da die Repräsentanten in der Datenbasis diese Ungenauigkeiten ebenfalls besitzen. Betrachtet man aber nun die Zusammensetzung von Basismodellen, die notwendig ist, um einen schmaleren Rumpf mit denselben Tragflächen zu erzeugen, so ist davon auszugehen, dass eine völlig andere Linearkombination von Basismodellen notwendig wird, um das Flugzeug entsprechend auszuprägen. In anderen Worten, die Modellparameter sind für Erkennungsaufgaben mit nur je einem Repräsentanten in der Datenbasis ungeeignet. Der Sachverhalt könnte sich durchaus von einer anderen Seite zeigen, wenn das adaptive Modell im Laufe des Systembetriebs an mehr und mehr Bilder angepasst wird und diese Ausprägungen in die Datenbasis aufgenommen werden. Auf diese Weise kann man eine große Datenbasis schaffen, die als Grundlage für eine stochastische Klassifikation auf Basis der Modell-Merkmale dienen kann.

Doch nun zu der besten Erkennungsrate von 80,3 Prozent für die Varianten und 90,9 Prozent für die Gruppe. Sie wurden für die Typerkennung auf Basis der 3D-Merkmale erzielt. Auf der Suche nach der Ursache für die falsch identifizierten 19,7 Prozent der Varianten wurde zunächst ein weiterer Test durchgeführt. Für diesen Test wurde die aus dem größten Segment berechnete Länge durch die Rumpflänge aus der Datenbasis ersetzt, d. h. es wurde getestet, inwieweit sich schlechte Modellinitialisierungen auf die Typerkennung auswirken. Das Ergebnis ist in Tabelle 5.3 festgehalten und zeigt eine deutliche Erhöhung der Erkennungsrate im Vergleich zu der automatischen Modellinitialisierung. Aus diesem Versuch kann man nur schließen, dass die derzeit implementierte Initialisierung der Modellparameter auf Basis der Rumpf-

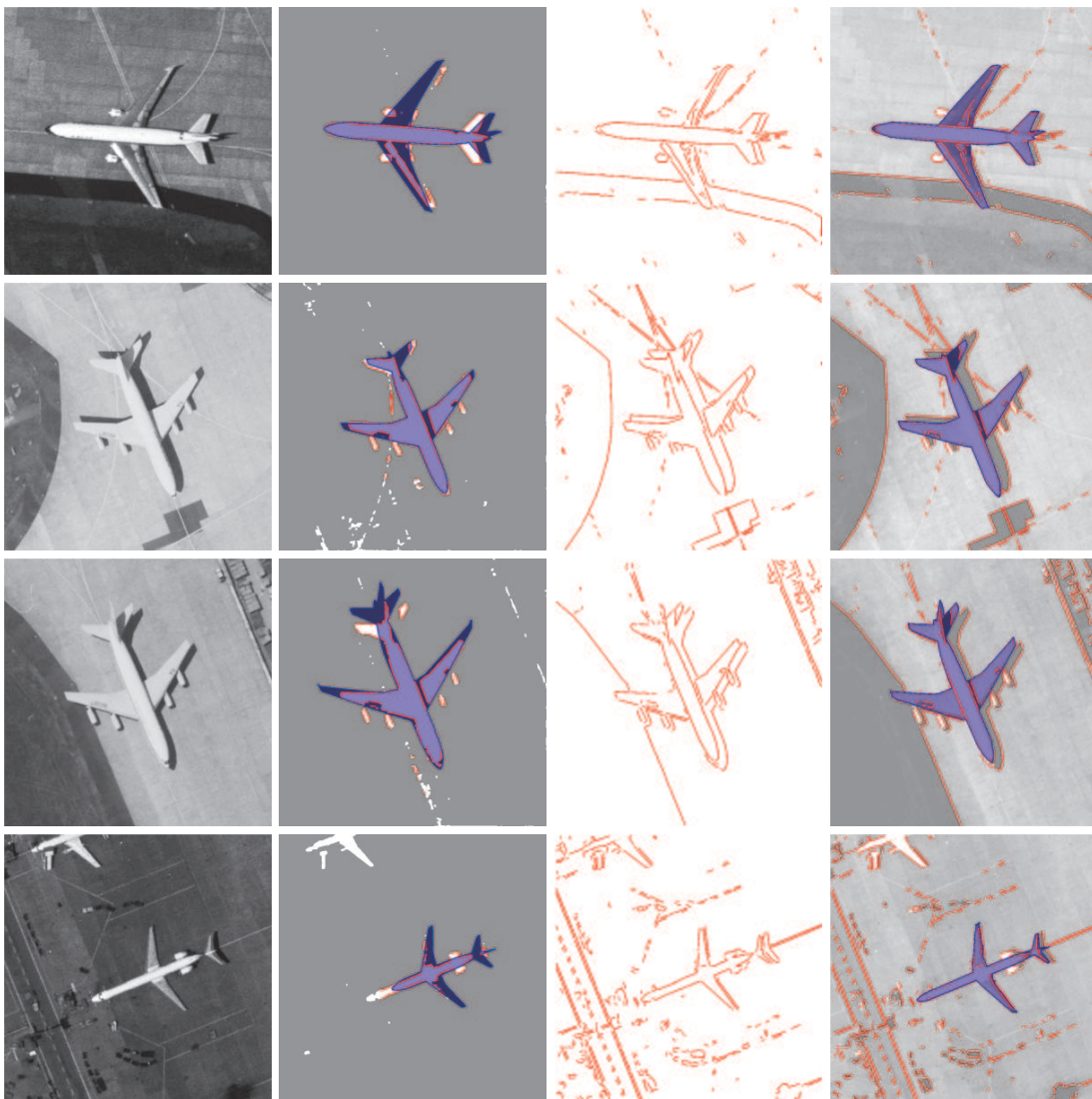
	3D-Merkmale	
	Variante	Gruppe
Korrekt erkannt	68	73
Falsch identifiziert	9	4
Erk.-Rate (Prozent)	88,3	94,8

Tabelle 5.3: Die Tabelle zeigt Ergebnisse der Erkennung entsprechend Tabelle 5.2, wobei hier die Länge zur Schätzung der initialen Modellparameter nicht aus dem größten Segment, sondern aus der *ground truth* entnommen wurden.

länge zu guten Ergebnissen führt, sofern diese Länge genau genug gemessen wurde. Andererseits ist nicht zu erwarten, dass auf der Grundlage von rein ansichtsbasierten Methoden eine allgemeingültige Verbesserung der Längenschätzung zu erzielen ist, da die falsch gemessenen Längen häufig auf Eigenbeschattung des Leitwerks oder dunkle Rumpfe zurückzuführen sind.

Eine Alternative zu rein ansichtsbasierten Methoden, die in diesem Zusammenhang unbedingt ins Auge gefasst werden muss, ist der Ansatz von [Pope & Lowe '96]. Für diesen Ansatz werden aus Trainingsbildern Bildmerkmale extrahiert, die in einen Graphen eingeordnet und dort unter Verwendung eines Ähnlichkeitsmaßes gruppiert werden. Liegt nun ein Bild vor, so werden dieselben Bildmerkmale extrahiert, der Graph nach ähnlichen Mustern durchsucht und schließlich ein Repräsentant ausgewählt. Zurzeit wird am Fraunhofer IITB eine Ergänzung dieses Verfahrens implementiert, so dass die Trainingsdaten nicht aus Realbildern, sondern aus synthetisch erzeugten Modellansichten extrahiert werden können [Plagemann '04]. Da das Verfahren von Pope tolerant ist gegenüber ungenauen Eingaben, können die synthetischen Ansichten aus einigen wenigen Ausprägungen des adaptiven Modells gewonnen und das erweiterte Verfahren somit zur Initialisierung für die Modellparameter herangezogen werden. Es ist zu erwarten, dass sich mit dieser Initialisierung Erkennungsraten erreichen lassen, die im Bereich des durchgeführten Versuchs liegen. Zudem kann der Ansatz auch für die Auswahl des adaptiven Modells herangezogen werden, wenn der Graph nicht nur aus synthetischen Ansichten eines einzigen adaptiven Modells erzeugt wird, sondern aus allen in Frage kommenden adaptiven Modellen zusammengestellt wird.

Die verbleibenden neun falsch identifizierten Varianten wurden einzeln auf die Ursache der Falsch-Identifikation untersucht. In drei Fällen war eine schlechte Startlage (nicht Länge!) die Ursache, die dazu geführt hat, dass die Tragflächen an Schattenkanten angepasst wurden. In zwei Fällen lag eine schlechte Modellanpassung vor. In den verbleibenden vier Falsch-Identifikationen wurde schließlich die Boeing 747-200 fälschlicherweise als Boeing 747-400 identifiziert. Die entsprechenden Modellanpassungsergebnisse sind in Abbildung 5.12 dargestellt. Wie aus der Abbildung hervorgeht, liegt der Fehler nicht in der Modellanpassung, sondern in der großen Ähnlichkeit der beiden Flugzeugtypen. Sie unterscheiden sich im Wesentlichen durch ihre Tragflächenform und auch durch ihren Pfeilungswinkel, was für Varianten eines Flug-



(a) Eingangsbild

(b) Initiale Lage

(c) Bildkanten

(d) Ergebnis

Abbildung 5.11: Diese Abbildung zeigt die Initialisierung und die Ergebnisse der Modellanpassung für unbekannte Flugzeugtypen, d. h. die abgebildeten Flugzeugtypen sind nicht in Form eines Basismodells in das adaptive Modell eingegangen.

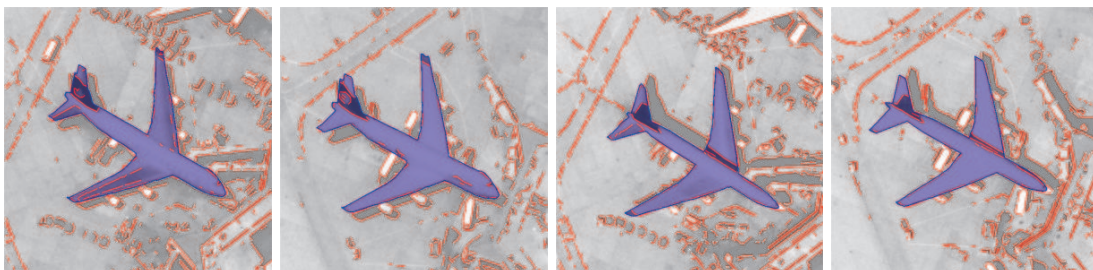


Abbildung 5.12: Falsch-Identifikationen. Alle vier Abbildungen zeigen den Flugzeugtyp Boeing 747-200 und alle vier wurden fälschlicherweise als Boeing 747-400 identifiziert.

zeugtyps sehr ungewöhnlich ist. Dies ist im Prinzip als Anlass zu nehmen, neben der Boeing 747-400 auch eine Boeing 747-200 als Basismodell in das (nicht erweiterte) adaptive Modell aufzunehmen. Dies konnte jedoch in dieser Arbeit nicht durchgeführt werden, da kein geeignetes Modell vorlag.

Über diese Versuche hinaus wurden die verbleibenden acht Bilder untersucht, die nicht durch einen Eintrag in der Datenbasis repräsentiert sind. Vier der Bilder konnten nicht ausgewertet werden, weil die dunklen Materialeigenschaften zu einer Initialisierung der Modellparameter führten, die eine fehlerhafte Modellanpassung nach sich zogen. An die übrigen vier wurde das adaptive Modell erfolgreich angepasst. Die Ergebnisse waren bereits in Abbildung 5.11 zu sehen.

5.1.8 Zusammenfassung der Flugzeugerkennung

In den letzten Abschnitten wurde ein System zur automatischen Typerkennung von Verkehrsflugzeugen auf Basis adaptiver Modelle präsentiert. Dieses System wurde an mehr als 80 Luftbildern getestet, die aus Senkrechtaufnahmen entstanden sind. Insgesamt wurden drei verschiedene Merkmalstypen eingesetzt, einer davon auf Basis von 2D-Informationen. Die Erkennung anhand der Modell-Merkmale hat sich, zumindest in Verbindung mit einer prototypischen Klassifikation, als nicht möglich erwiesen. Die Erkennung mit den aus der 3D-Rekonstruktion gewonnenen Merkmalen hingegen zeigt gute Ergebnisse, die noch entscheidend verbessert werden, wenn die Initialisierung der Modellparameter ebenfalls durch einen formbasierten Ansatz gewonnen wird. In diesem Zusammenhang wurde auf eine Erweiterung des formbasierten Ansatzes aus [Pope & Lowe '96] verwiesen, die auch für die Auswahl des adaptiven Modells in Betracht gezogen werden kann. Auf diese Weise ließe sich die Automatisierungslücke zwischen der Detektion und der Typerkennung schließen.

Eine entscheidende Erkenntnis der durchgeführten Versuche ist jedoch, dass eine Identifikation der Flugzeugvariante ohne Berücksichtigung der projektiven Verzerrungen, also durch rein segmentbasierte 2D-Verfahren, dem hier gezeigten Ansatz gegenüber deutlich unterlegen ist. Gegenüber dem Verfahren, das in der einleitend erwähnten Machbarkeitsanalyse durchgeführt wurde, ist also ein entscheidender Fortschritt erreicht worden.

Ein direkter Leistungsvergleich zu anderen Verfahren wird erschwert durch die Tatsache, dass - soweit bekannt - kein anderes Verfahren auf vergleichbaren Daten getestet worden ist, insbesondere nicht für technische Objekte mit einer Ähnlichkeit, die vergleichbar ist mit der der Flugzeugvarianten. Insofern ist ein quantitativer Vergleich der erreichten Erkennungsrate mit den Quoten anderer Verfahren nur unter großem Vorbehalt zu ziehen. Dennoch wird an dieser Stelle auf die in [Campbell & Flynn '01] angegebene tabellarische Übersicht über erscheinungsbasierte Erkennungsmethoden verwiesen, die für einige Verfahren auch Erkennungsraten beinhaltet. Die dort angegebenen Quoten liegen zwischen 81 und 100 Prozent, werden aber aus besagten Gründen nicht weiter interpretiert.

Insgesamt lässt sich feststellen, dass das entwickelte Modul ein ausgesprochen hilfreiches Werkzeug zur Erkennung von Flugzeugtypen und -varianten darstellt. Durch die visuelle Ausgabe, die Angabe des Vertrauenswertes der Typerkennung und letztlich durch die Option, die extrahierten Größen einzusehen, ist eine erhebliche Unterstützung für die Bildauswertung geschaffen worden.

5.2 Lageschätzung

Adaptive Modelle können im Kontext der Bildanalyse nicht nur zur Klassifikation herangezogen werden. Tatsächlich ist die Lageschätzung eine weitere Anwendung für adaptive Modelle. Im Zusammenhang mit der Lageschätzung sind adaptive Modelle genau dann von Interesse, wenn bestimmte Parameter der beobachteten Objekte nicht bekannt sind, z. B. weil die beobachteten Objekte deformierbar sind oder wenn den verwendeten Modellen erhebliche geometrische Ungenauigkeiten unterstellt werden müssen.

In diesem Abschnitt wird die Lageschätzung mit ungenauen, rigiden Modellen der Lageschätzung mit adaptiven Modellen gegenübergestellt. Es stellt sich heraus, dass adaptive Modelle schon bei geringfügigen Modellfehlern bessere Messgenauigkeiten liefern als rigide Modelle. Um eine Beeinflussung der Messergebnisse durch das zur Anpassung herangezogene Verfahren auszuschließen, wurde die Modellanpassung für die rigiden Modelle mit dem robusten Parameterschätzer realisiert.

Die Lageschätzung mit adaptiven Modellen unterliegt einer Reihe von Voraussetzungen. Die meisten dieser Voraussetzungen sind darin begründet, dass die Lageschätzung dem Bereich der Messaufgaben zuzuordnen ist. Im Gegensatz zur Typerkennung, die auf der Trennbarkeit von Merkmalen basiert und damit einen objektabhängigen aber vergleichsweise großen Toleranzbereich bezüglich der extrahierten Größen besitzen kann, muss die Leistung eines Lageschätzers an den effektiv erzielten Genauigkeiten gemessen werden.

Eine der wichtigsten Voraussetzungen ist daher, dass die Modell- und Lageparameter nicht oder nur in geringem Maße korreliert sind. Andernfalls muss gerade bezüglich dieser Parameter mit Fehlmessungen gerechnet werden. Im Fall von Flugzeugen bestehen z. B. erhebliche Zusammenhänge zwischen der Modellausprägung und der Objektskalierung, die wiederum mit dem Aufnahmeabstand korreliert ist. Darüber hinaus wurde auf das Zusammenspiel zwischen Modellausprägung und den Eulerwinkeln hingewiesen. An dieser Stelle sei jedoch ausdrücklich betont, dass sich diese Korrelationen durch die Verwendung von multiokularen Kamerasystemen und einer simultanen Anpassung über mehrere Bilder reduzieren, wenn nicht sogar vollständig unterdrücken lassen. Insofern kann diese Voraussetzung für die Anwendbarkeit adaptiver Modelle in der Praxis abgeschwächt werden.

Neben diesen Anforderungen an die Objektklasse müssen auch an das Bildmaterial Anforderungen gestellt werden, um ein gutes Testszenario zu erzielen. Insbesondere werden präzise Kenntnisse bezüglich der internen Kameraparameter benötigt, denn nur

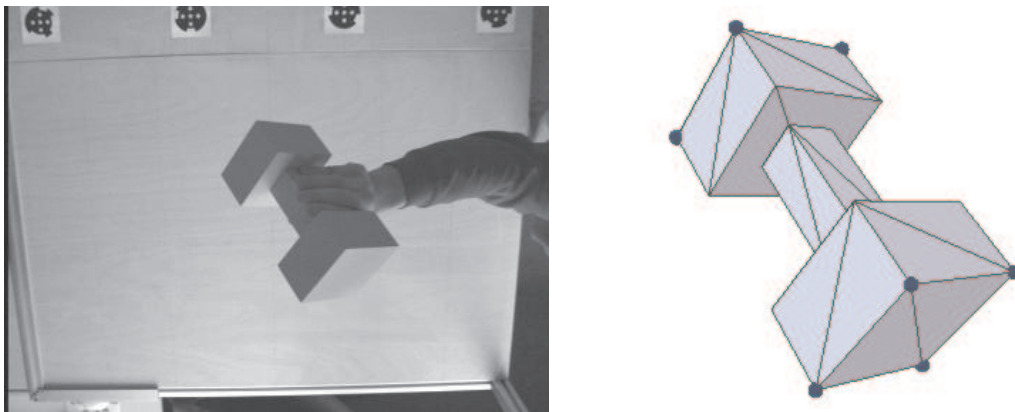


Abbildung 5.13: Eine Aufnahme der Hantel sowie das entsprechende Modell. Auf dem Modell sind die charakteristischen Oberflächenpunkte für den Passpunktschätzer markiert.

so können korrelationsbedingte Effekte, z. B. zwischen dem Aufnahmeabstand und der Brennweite oder der Bildpunktgröße, ausgeschlossen werden.

Alles in allem bilden weder die Klasse der kommerziellen Flugzeuge noch die im vorangegangenen Abschnitt untersuchten Bilder ein geeignetes Szenario für Untersuchungen bezüglich erzielbarer Messgenauigkeiten. Aus diesem Grund wurde für die folgenden Untersuchungen auf eine Innenbereichsszene zurückgegriffen, die im Rahmen des BMBF-Projektes Mqube⁷ entstanden ist. Die Szene zeigt ein eigens für die Entwicklung eines modellbasierten Objektverfolgers angefertigtes Objekt, das im Folgenden kurz Hantel genannt wird (siehe Abbildung 5.13).

5.2.1 Die Versuchsdaten

Aus den zahlreichen verfügbaren Sequenzen wurden drei kalibrierte Einzelbilder in der Größe von 382×286 Bildpunkten ausgewählt (siehe Abbildung 5.14). Die Hantel selbst hat eine Länge von 220 Millimetern und eine Breite bzw. Höhe von 100 Millimetern und wird in den ausgewählten Bildern mit Aufnahmeabständen zwischen 0,94 und 1,05 Metern abgebildet. Das der Hantel entsprechende Modell setzt sich aus drei triangulierten Quadern zusammen und besteht aus 36 Dreiecken.

Um die oben ausführlich diskutierten Korrelationen zwischen Modell-, Lage- und Kameraparametern so gering wie möglich zu halten, wurde von dem Hantelmodell nur die Länge variiert. Die Tatsache, dass diese Variation mathematisch einfach modellierbar ist, macht eine Modellerzeugung, wie sie in Kapitel 3 beschrieben wurde, unnötig. Stattdessen wurden aus dem Hantelmodell durch Skalierung in einer Dimension zwei Realisierungen erzeugt, eine um 30 Millimeter verkürzt, die andere um 30 Millimeter verlängert. Diese zwei Realisierungen bilden zusammen mit der durch das Dreiecksnetz des Originalmodells gegebenen Topologie das adaptive Modell.

⁷Förderkennzeichen 01 IR A04 des BMBF

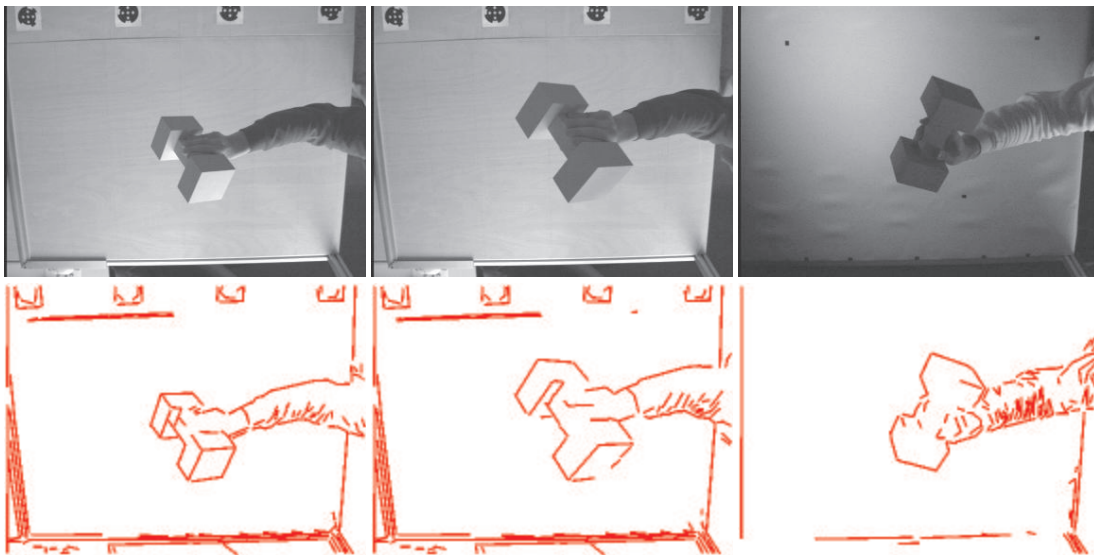


Abbildung 5.14: Die drei Bilder für die Konvergenztests, die im Text (von links nach rechts) mit Bild 1, Bild 2 bzw. Bild 3 bezeichnet sind. Unten sind die extrahierten Bildkantensegmente dargestellt.

5.2.2 Testbeschreibungen

Mit Hilfe dieses adaptiven Modells wurden zwei Konvergenztests durchgeführt. Der erste Test gilt dem Vergleich zwischen

- einer Lageschätzung mit dem exakten, rigiden Modell und
- einer Lageschätzung mit dem adaptiven Modell mit zufällig gewählten Initialwerten für die Hantellänge.

Aus diesem Versuch lässt sich ableiten, welche Fehler in Kauf genommen werden müssen, wenn man, weil die reale Hantellänge unbekannt ist, mit einem adaptiven Modell arbeitet. Der zweite Test gilt dem Vergleich zwischen

- einer Lageschätzung mit einem rigiden Modell in mehreren verschiedenen Längen und
- einer Lageschätzung mit dem adaptiven Modell in mehreren verschiedenen Längen.

Dieser Versuch wiederum gibt Aufschluss darüber, welche Fehler entstehen, wenn man, weil die reale Hantellänge unbekannt ist, mit einem rigiden Modell falscher Länge arbeitet. Zudem lässt sich eine Aussage darüber ableiten, ab welcher Ungenauigkeit des rigiden Modells die Leistung der Lageschätzung durch den Einsatz eines adaptiven Modells verbessert wird.

	Eulerwinkel	Translation x - und y -Achse	Translation z -Achse	Hantellänge
Experiment 1	$\pm 5^\circ$	± 10 mm	± 100 mm	± 30 mm
Experiment 2	$\pm 10^\circ$	± 20 mm	± 200 mm	± 30 mm

Tabelle 5.4: Die Offsets für die Konvergenztests der zwei Experimente zur Lage-schätzung.

Alle Versuche wurden unter Verwendung eines Aufmerksamkeitsbereiches durchgeführt, d. h. vor Versuchsbeginn wurde interaktiv ein grob umrissener Bildbereich angegeben, auf den die Bildkantenextraktion eingeschränkt wurde. Diese Einschränkung der Bildkantenmenge reduziert die benötigte Rechenzeit, hat aber keine Auswirkung auf die erzielten Messergebnisse, da durch diese Aufmerksamkeitsbereiche nur Bildkanten aus dem Randbereich der Bilder ausgeblendet werden, die durch den Filter der Gütefunktion ohnehin nicht berücksichtigt werden.

Für die Konvergenztests wurden zunächst Startpositionen bestimmt, in dem das Modell mehrfach an die Bilder angepasst und anschließend für jedes Bild die entsprechende mittlere Lage als Startposition berechnet wurde. Auf diese Startpositionen wurden dann für jeden Versuch 100 zufällig erzeugte Offsets aufaddiert und als initiale Parameter für jeweils 100 Registrierungen herangezogen. Die Offsets wurden gleichverteilt in den Intervallen entsprechend Tabelle 5.4 erzeugt. Insgesamt wurden für jedes Bild zwei Experimente mit unterschiedlich erzeugten initialen Parametern durchgeführt.

5.2.3 Die Leistungsmessung

Für die Konvergenztests wurde zwischen robustem Verhalten einerseits und Messgenauigkeit andererseits differenziert. Unter der Robustheit ist in diesem Zusammenhang die Anzahl der zurückgewiesenen Messungen zu verstehen, d. h. diejenigen Optimierungsläufe, die an lokalen Extremwerten der Gütefunktion konvergiert sind und automatisch detektiert wurden. Unter der Messgenauigkeit ist die Streuung derjenigen Messdaten zu verstehen, die vom System als korrekt deklariert wurden. Diese Differenzierung ist deshalb wichtig, weil grobe Fehlmessungen nicht zufällig entstehen, sondern durch systematische Strukturen in den Bilddaten, die sich als lokale Extremwerte in der Gütefunktion ausdrücken. Bezieht man die Optimierungsläufe, die in solchen lokalen Extremwerten konvergieren, in die Messgenauigkeit ein, so misst man letztlich die systematischen Störungen im Bild und nicht die Genauigkeit der Lage-schätzung. Die Zurückweisung von Optimierungsläufen, also die Deklaration einer Lageschätzung als Ausreißer, wurde anhand einer Schwelle für den Vertrauenswert der Modellanpassung festgelegt.

Bei den hier durchgeführten Hantelversuchen hat sich gezeigt, dass die zurückgewiesenen Optimierungsläufe immer in denselben, maximal zwei oder drei verschiedenen Lagen konvergiert sind, die sich am erreichten Vertrauenswert der Modellanpassung unterscheiden lassen. In Abbildung 5.15 beispielsweise sind die Vertrauenswerte von

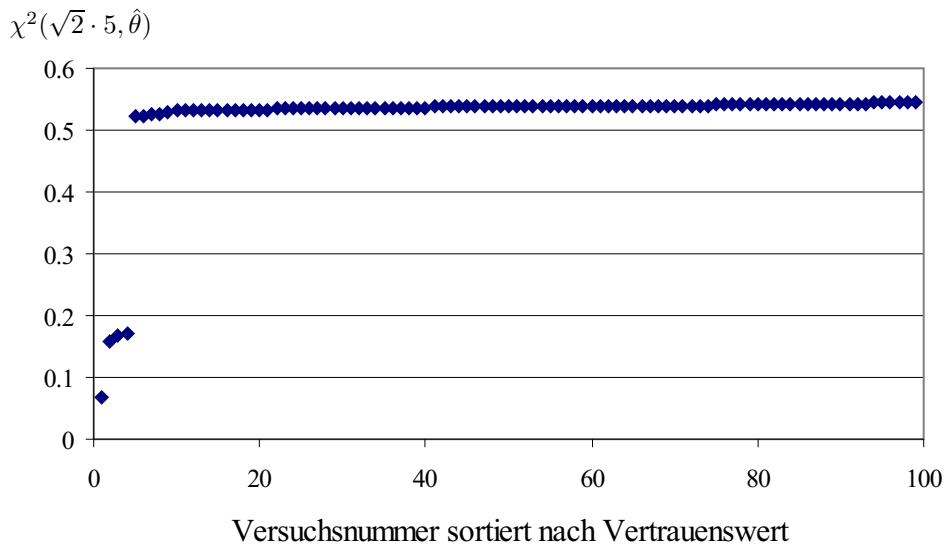


Abbildung 5.15: Die aufsteigend sortierten Vertrauenswerte für Bild 2 und Experiment 2 zur reinen Lageschätzung. Die Vertrauenswerte sind bezüglich der Gesamtlänge der abgebildeten Modellkantensegmente normiert. Da in Bild 2 keine parallelen Bildkantensegmente in der Nähe der Hantel vorhanden sind, können diese Werte als Prozentzahl der Überlappung von Modellkantensegmenten durch Bildkantensegmente interpretiert werden. Die ersten vier Messungen wurden anhand des geringen Vertrauenswertes zurückgewiesen.

100 Testläufen auf Bild 2 dargestellt, von denen die ersten vier in zwei verschiedenen lokalen Extremwerten konvergiert sind.

5.2.4 Die Ergebnisse

Im ersten Test werden Optimierungsläufe mit dem exakten, rigiden Modell und dem adaptiven Modell verglichen. Bei der Optimierung mit dem adaptiven Modell ist zusätzlich die Hantellänge mit zufälligen Werten im Bereich von ± 30 Millimetern initialisiert. Das Ergebnis dieses Tests ist in Tabelle 5.5 dargestellt. Zunächst ist festzustellen, dass die Streuung der Eulerwinkel in Experiment 1 bei einer Lageschätzung mit dem adaptiven Modell gegenüber dem rigiden Modell nur geringfügig erhöht wird. Dasselbe gilt für die translativen Anteile in der Bildebene. Die erhöhte Streuung bei der Schätzung des Aufnahmeabstands (Translation längs der z -Achse) lässt sich auf eine Korrelation zwischen den Modellparametern und dem Aufnahmeabstand zurückführen.

Die Robustheit wird in Experiment 1 durch das adaptive Modell offensichtlich nur geringfügig beeinflusst. Bei den weiter gestreuten Startwerten in Experiment 2 hingegen zeigen sich Messungenauigkeiten, die eher an den einzelnen Bildern festzumachen

		Standardabweichungen						
		Eulerwinkel [Grad]			Translation [mm]			Länge
Anzahl Ausreißer		ω	φ	κ	t_x	t_y	t_z	[mm]
Experiment 1: Lage								
Bild 1	0	0,51	0,30	0,24	0,24	0,26	4,39	-
Bild 2	1	0,72	0,38	0,26	0,42	0,33	3,32	-
Bild 3	3	0,65	0,30	0,25	0,29	0,42	2,99	-
Experiment 1: Lage und Form								
Bild 1	0	0,49	0,53	0,35	0,44	0,43	8,60	2,38
Bild 2	3	0,65	0,41	0,25	0,30	0,32	4,72	1,82
Bild 3	3	1,20	0,95	0,42	0,47	0,65	9,00	2,80
Experiment 2: Lage								
Bild 1	3	0,45	0,31	0,24	0,21	0,34	4,92	-
Bild 2	4	0,70	0,40	0,27	0,40	0,35	3,37	-
Bild 3	21	0,62	0,62	0,24	0,37	0,52	4,55	-
Experiment 2: Lage und Form								
Bild 1	8	0,62	1,02	0,64	0,50	1,39	8,91	3,69
Bild 2	4	0,50	0,45	0,19	0,35	0,27	3,30	1,58
Bild 3	18	1,20	0,88	0,49	0,46	0,69	8,80	2,62

Tabelle 5.5: Die Standardabweichungen der reinen Lageschätzung in Gegenüberstellung mit den Standardabweichungen bei simultaner Lage- und Formschätzung.

sind als an der Verwendung eines adaptiven oder rigidem Modells. Dies zeigt sich auch an der Anzahl der Ausreißer, die für Bild 3 bei Experiment 2 sprunghaft ansteigt. Insgesamt wirkt sich die Verwendung adaptiver Modelle nicht negativ auf die Robustheit aus und die Messgenauigkeiten werden nur geringfügig verschlechtert. Die größten Messgenauigkeiten treten für den Aufnahmeabstand auf. Ein weiterer Schluss kann aus dem Test gezogen werden: Die hier vorgestellten Module zur Modellanpassung sind bei den Hantelbildern in Experiment 2 an die Grenzen gestoßen, innerhalb derer die initialen Parameter liegen müssen, um eine robuste Lageschätzung durchzuführen.

Die entsprechenden Rechenzeiten für einen Pentium IV 1.9 GHz wurden unter Verwendung der Aufmerksamkeitsbereiche ermittelt und sind in Tabelle 5.6 festgehalten. Die Verwendung eines Aufmerksamkeitsbereiches bedeutet z. B. für Bild 2 die Reduktion der Zahl der Bildkanten von 144 auf 33, also auf 23 Prozent. Zum Vergleich wurde für Bild 2 die Rechenzeit für die reine Lageschätzung, also ohne Formparameter, auf allen 144 Bildkanten ermittelt (0,086 Sekunden). Vergleicht man diese Rechenzeit mit der Rechenzeit, die unter Verwendung eines Aufmerksamkeitsbereiches benötigt wurde (0,075 Sekunden), so ergibt sich eine geringe Erhöhung der Rechenzeit um 15 Prozent. Dieser geringe Faktor ist darauf zurückzuführen, dass zwar die Summanden der Gütefunktion für alle Bild-Modellkantenpaare berechnet werden müssen, nicht jedoch

Rechenzeiten [sek]	Lage		Lage und Form	
	Experiment 1	Experiment 2	Experiment 1	Experiment 2
Bild 1	0,07	0,08	2,24	2,33
Bild 2	0,08	0,07	1,92	3,93
Bild 3	0,08	0,07	1,81	1,90

Tabelle 5.6: Die mittleren Rechenzeiten für die Lage- und Formschätzung unter Verwendung eines Aufmerksamkeitsbereiches.

die rechenintensiven Gradienten. Letztere werden nur für Bildkanten berechnet, die einen Beitrag zur Kostenfunktion leisten.

Bei den in Tabelle 5.6 angegebenen Rechenzeiten ist unbedingt zu berücksichtigen, dass das Modell aus nur 36 Dreiecken besteht und entsprechend wenige Modellkantensegmente liefert. Da technische Objekte in der Regel mindestens eine dreistellige Anzahl an Dreiecken besitzen, können die gemessenen Zeiten nicht uneingeschränkt als Echtzeitfähigkeit interpretiert werden. In jedem Fall kann aber festgehalten werden, dass die Rechenzeit für den robusten Schätzer weniger von der Anzahl der Bildkanten als von der Anzahl der Modellkanten beeinflusst wird.

Stellt man die Rechenzeiten der Anpassung rigider Modelle der Anpassung adaptiver Modelle gegenüber, so sind Rechenzeitverlängerungen von mehr als Faktor 50 zu beobachten. Dies ist auf die Verwendung des Passpunktschätzers und dessen Implementierung auf Basis von Schablonenanpassungen zurückzuführen, wie eine Anpassung des adaptiven Modells ohne den Passpunktschätzer auf Bild 2 gezeigt hat: Bei diesem Versuch⁸ wurde eine mittlere Rechenzeit von 0,24 Sekunden ermittelt, also um Faktor 3 mehr als für den entsprechenden Versuch mit dem rigidem Modell.

In dem zweiten Test wurde nur Bild 2 ausgewertet und die Lageparameter wurden entsprechend Experiment 2, also auf die ungenauere Initialisierung gesetzt. Nun wurden sowohl für das rigide als auch für das adaptive Modell 13 Versuche gestartet. Die Versuche unterscheiden sich darin, dass auf die korrekte Hantellänge ein fest gewähltes, also nicht zufällig erzeugtes Offset aufaddiert wurde. Die Versuche wurden mit Offsets zwischen -30 und 30 Millimetern erzeugt, in einer Schrittweite von 5 Millimetern. Der Unterschied zwischen der Lageschätzung mit dem rigiden Modell einerseits und dem adaptiven Modell andererseits besteht nun darin, dass das adaptive Modell die Möglichkeit hat, die falsche Länge zu korrigieren, wohingegen das rigide Modell die beste Position mit dieser falschen Länge finden muss.

Die Ergebnisse sind im Diagramm 5.16 in Form der Standardabweichungen dargestellt. Während sich die Streuungen der geschätzten Parameter für das rigide Modell mit zunehmend falscher Modelllänge drastisch erhöhen, bleiben die Streuungen für das adaptive Modell konstant. In den Abbildungen 5.17 und 5.18 sind die Ergebnisse bezogen auf die Mittelwerte zu sehen. Sie schwanken für rigide Modelle erheblich.

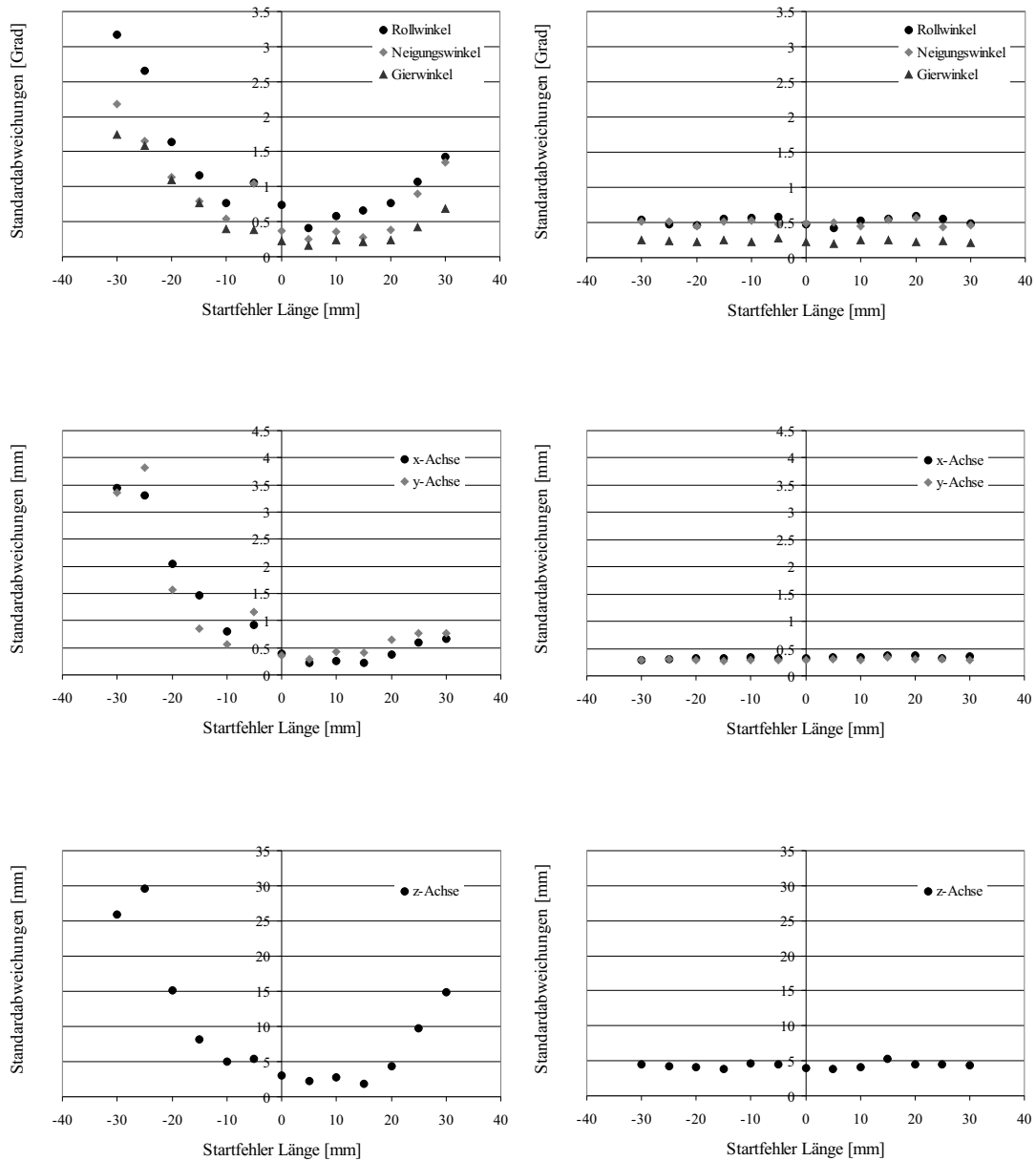
⁸Dieser Versuch zeigte wie alle Versuche mit adaptiven Modellen ohne Passpunktschätzer ein schlechtes Konvergenzverhalten.

An den Diagrammen ist außerdem auffällig, dass die besten Ergebnisse der rigiden Modellanpassung für ein Hantelmodell mit einer Länge von 225 anstatt der realen 220 Millimeter erzielt werden. Diese Beobachtung deckt sich mit der Tatsache, dass bei der Anpassung des adaptiven Modells im ersten Test auf Bild 2 eine mittlere Länge von 224,7 Millimetern gemessen wurde, also ein Fehler von 4,7 Millimetern. Der Versuch wurde für eine synthetische Reproduktion von Bild 2 wiederholt, wobei eine mittlere Länge von 220,26 Millimetern gemessen wurde, was einem Längenfehler von 0,26 Millimetern entspricht.

Daraus muss man schließen, dass der Fehler auf die Modellierung des physikalischen Abbildungsvorgangs zurückzuführen ist. Er kann durch Berücksichtigung von optischen Verzerrungen vermieden werden. Zudem lässt sich feststellen, dass das adaptive Modell bereits ab einem Modellfehler von 5 Millimetern mit der Anpassung rigider Modelle Schritt halten kann und bei Fehlern, die darüber hinausgehen, genauere Messergebnisse liefert.

5.2.5 Zusammenfassung der Lageschätzung

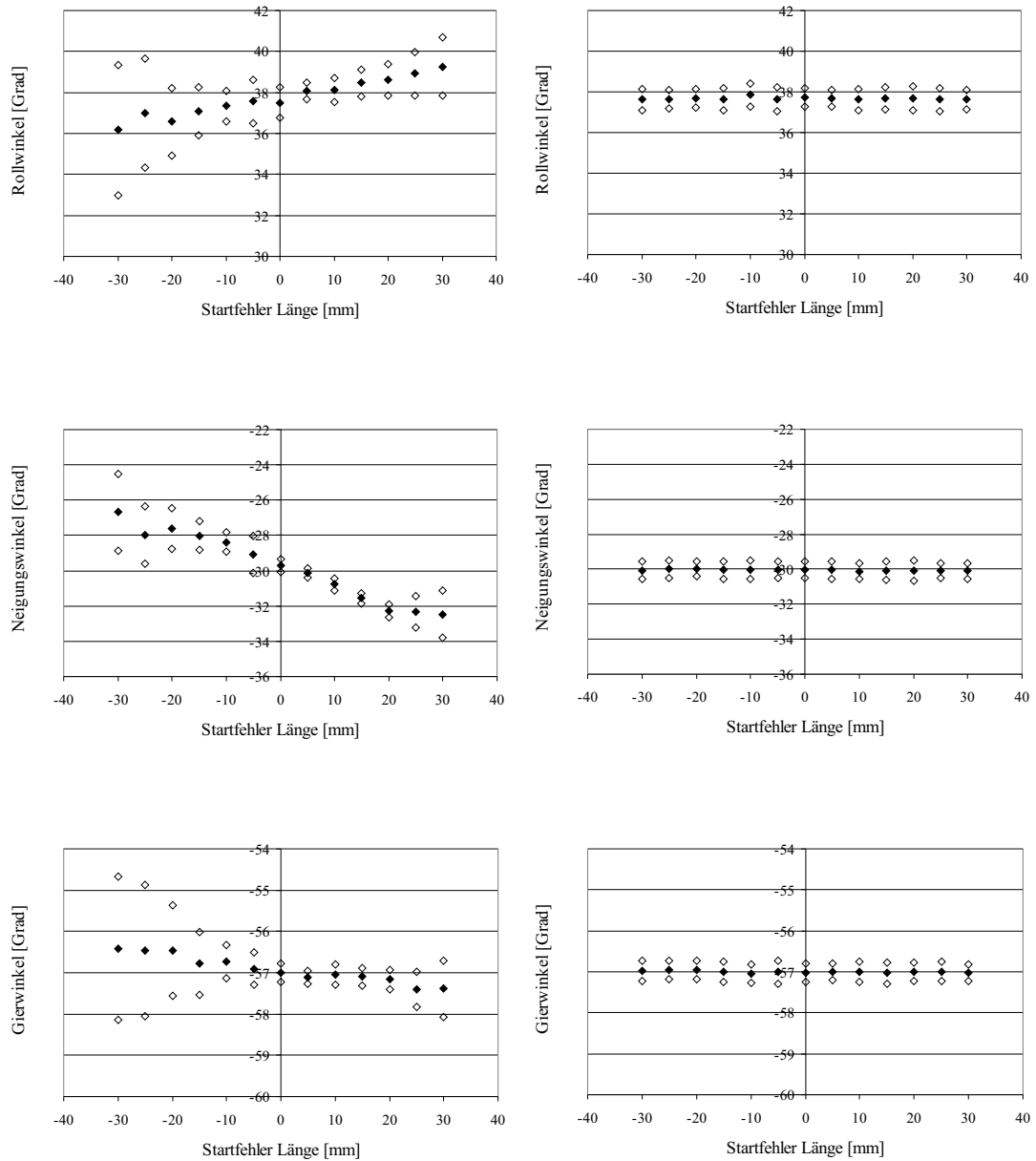
In den oben durchgeführten Experimenten mit der Hantel wird zum einen gezeigt, dass die in dieser Arbeit vorgestellten adaptiven Modelle auch zur Lageschätzung eingesetzt werden können. Seine wahren Stärken offenbart das adaptive Modell bei seiner Anwendung auf Objekte, deren Form nur ungenau bekannt ist. Experimentell wurde gezeigt, dass die Anpassung mit adaptiven Modellen bereits ab Modellfehlern von 5 Millimetern, also 2,3 Prozent der Hantellänge, mit der Anpassung des rigiden Modells gleichzieht. Bei Modellfehlern, die darüber hinaus gehen, werden erheblich genauere Messungen geliefert. Der Preis, der dafür zu zahlen ist, ist die um bis zu Faktor 50 erhöhte Rechenzeit, die auf die rechenintensive Implementierung des Passpunktschätzers zurückzuführen ist.



(a) Rigides Modell

(b) Adaptives Modell

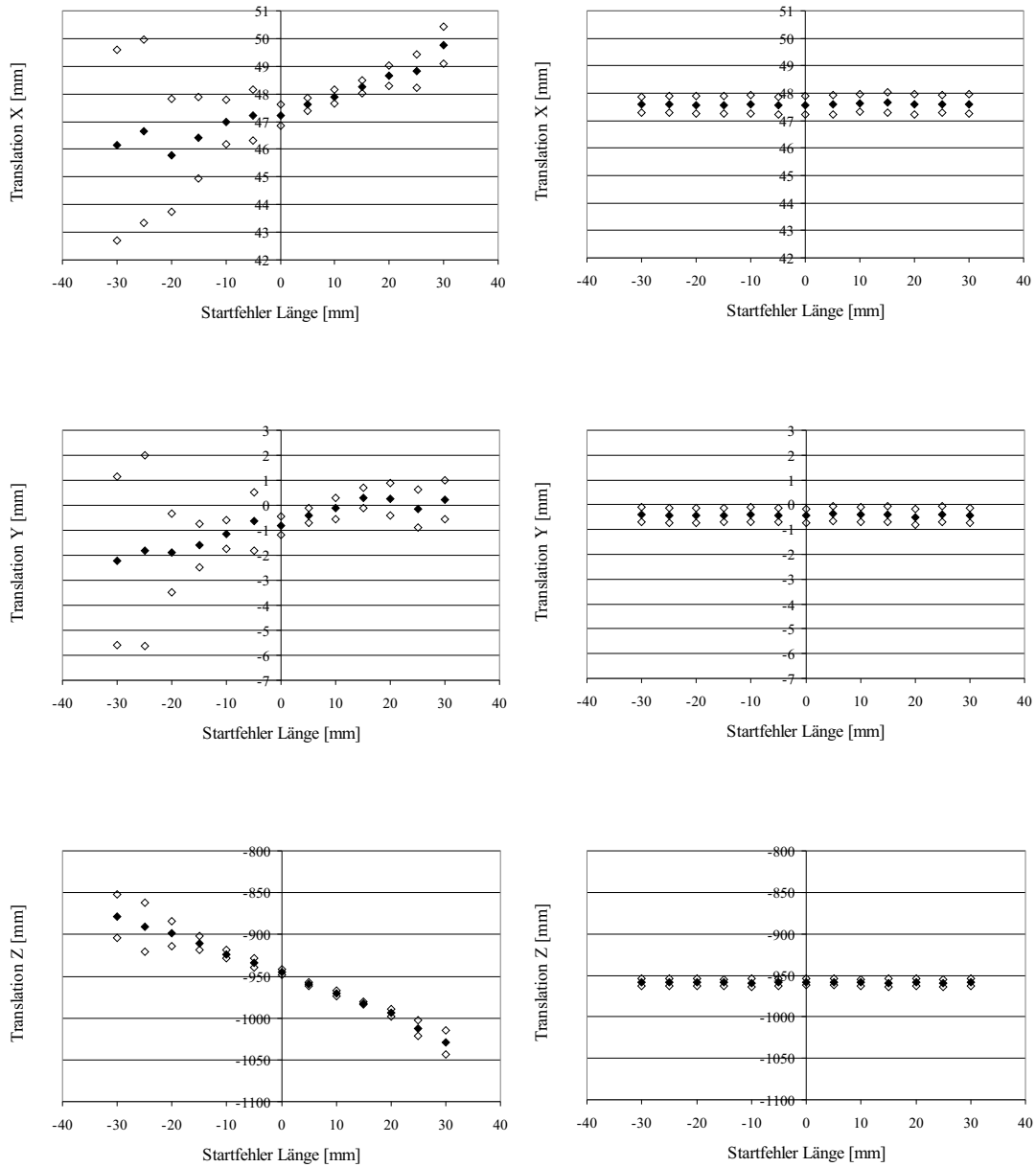
Abbildung 5.16: Vergleich der Standardabweichungen mit (a) rigidem und (b) adaptivem Modell in Abhängigkeit vom jeweiligen Modellstartfehler.



(a) Rigides Modell

(b) Adaptives Modell

Abbildung 5.17: Die Mittelwerte (ausgefüllte Rauten) und die Mittelwerte plus/minus Standardabweichungen (nicht gefüllte Rauten) für das rigide und das adaptive Modell im Vergleich.



(a) Rigides Modell

(b) Adaptives Modell

Abbildung 5.18: Die Mittelwerte und die Mittelwerte (ausgefüllte Rauten) plus/minus Standardabweichungen (nicht gefüllte Rauten) für das rigide und das adaptive Modell im Vergleich.

Kapitel 6

Zusammenfassung und Ausblick

6.1 Zusammenfassung

Adaptive Modelle gewinnen in der Bildauswertung zunehmend an Bedeutung, da sie die Modellierung dynamischer Szenenaktivitäten ermöglichen und Flexibilität bezüglich neu auftretender Objekte schaffen. In der hier vorliegenden Arbeit wird diese Flexibilität eingesetzt, um unvollständige Datenbasen und geometrisch ungenaue Systemeingaben zu kompensieren. Hierfür wird zum einen ein automatischer Konstruktionsprozess für adaptive Modelle vorgestellt, der die Modellierung von technischen Objekten ermöglicht. Zum anderen werden neu entwickelte Module zur Anpassung adaptiver Modelle an Grauwertbilder vorgestellt, die anwendungsspezifisch zusammengestellt werden können.

Eine Errungenschaft der vorliegenden Arbeit ist die automatische datengetriebene Konstruktion adaptiver Modelle. Für diesen Konstruktionsprozess werden handelsübliche 3D-Modelle als Systemeingabe verwendet. Eine Neuerung ist dabei die Kombination oberflächen- und volumenbasierter Verfahren, die auch die Modellierung dünner Objektteile und scharfer Kanten ermöglicht. Letzteres ist für die Anwendung auf technische Objekte von elementarer Bedeutung und wird von keinem anderen Verfahren in dieser Form gezeigt.

Die Elastizität der adaptiven Modelle wird - im Gegensatz zu parametrisierten Flächen und interaktiv erzeugten Modellen - durch die Auswahl der Eingabemodelle definiert. Wenngleich keine Faustregel für die Zusammenstellung dieser Basis existiert, so können Kenntnisse über die Variationen der vorliegenden Objektklasse leicht zu einer geeigneten Auswahl verhelfen. Sind die erforderlichen Modelle nicht vorhanden, so kann die unvollständige Datenbasis durch eine Erweiterung des automatischen Konstruktionsprozesses ausgeglichen werden. Die für diese Erweiterung des adaptiven Modells benötigte Interaktion ist geringfügig und steht in keinem Verhältnis zu dem Aufwand, den die manuelle Konstruktion von generischen Modellen in der Art von [Sullivan et al. '95, Ferryman et al. '95, Koller et al. '93, Zhao & Nevatia '01] mit sich bringt.

Die Module zur Modellanpassung sind auf Basis von Kantenmerkmalen realisiert, wodurch unter anderem die Beleuchtungsinvarianz des Verfahrens erzielt wird. Die Modellanpassung setzt sich aus zwei Schätzern zusammen, wobei der erste Schätzer durch eine robuste, nichtlineare Optimierung realisiert ist und den Kern der Modellanpassung darstellt. Der zweite Schätzer, ein Passpunktschätzer, dient zur Überwindung lokaler Extremwerte.

Eine Demonstration der Leistungsfähigkeit dieser Verfahren ist am Beispiel der Typerkennung von Verkehrsflugzeugen erfolgt. Das hierfür entwickelte System zeigt die größte Stärke des Verfahrens: Aus einer Menge von handelsüblichen Modellen, deren geometrische Fehler jenseits der erforderlichen Genauigkeiten zur Typerkennung liegen, kann es eine Menge von automatisch erzeugten dreidimensionalen Prototypen erzeugen, die trennungswirksame Genauigkeiten besitzen. Auf diese Weise wird eine Datenbasis für die Flugzeugerkennung geschaffen, die die Differenzierung zwischen verschiedenen Flugzeugtypen und deren Varianten ermöglicht. Insbesondere wird mit der Flugzeugtypisierung ein Differenzierungsgrad erreicht, der weit über den anderer Veröffentlichungen, wie z. B. [Sullivan et al. '95, Ferryman et al. '95], hinaus geht.

Eine weitere wichtige Eigenschaft des hier vorgestellten Verfahrens ist die Tatsache, dass das System im laufenden Betrieb - durch Modellanpassung an Beispielfelder - um weitere Prototypen ergänzt werden kann. Vergleicht man diese Eigenschaft mit Verfahren, die auf rigiden Modellen aufbauen [Kölzow '03], so wird ein weiterer Gewinn der Anwendung des hier vorgestellten Verfahrens deutlich, nämlich die automatische Erweiterbarkeit der Datenbasis um neue Instanzen innerhalb der modellierten Objektklasse. Eine manuelle Nachführung neuer Instanzen in der erforderlichen Qualität entfällt.

An Bildern einer Innenbereichsszene wurde zudem gezeigt, dass das entwickelte Verfahren auch zur Lageschätzung einsetzbar ist. In Versuchen wurde demonstriert, dass bereits ab Modellierungsungenauigkeiten von weniger als drei Prozent durch adaptive Modelle genauere Lagemessungen erreichbar sind als mit rigiden Modellen. Damit hat sich das hier entwickelte Modellanpassungsverfahren auch für Anwendungen im Rahmen von Messaufgaben als einsetzbar erwiesen. Anwendungen sind in solchen Bereichen vorstellbar, in denen deformierbare Objektteile beobachtet werden, keine genauen Angaben über die beobachteten Objekte existieren oder aber eine Vielzahl sehr ähnlicher Objekte vorliegt. Echtzeitanwendungen kommen dabei, zumindest bei der derzeitigen Implementierung des Passpunktschätzers, jedoch nicht in Frage.

6.2 Ausblick

Adaptive Modelle stellen inzwischen eine anerkannte Methode in der Bildauswertung dar und die entsprechenden Methoden werden sich in den nächsten Jahren weiter entwickeln. Auch für das hier vorgestellte Verfahren bieten sich eine Reihe von Entwicklungsmöglichkeiten an, die insbesondere die Anwendbarkeit in anderen Kontexten als der Typerkennung ermöglichen.

Multiokulare Kamerasysteme

Ein Punkt, der bereits im Rahmen der Versuche zur Lageschätzung angesprochen wurde, ist die Anwendung des Verfahrens auf multiokulare Kamerasysteme. Durch die Messung von Bildkantensegmenten aus mehreren Blickrichtungen wird die Parameterschätzung robuster, weil dadurch Korrelationen zwischen Form und Lage reduziert werden. Durch eine solche Erweiterung kommt den adaptiven Modellen eine erheblich größere Bedeutung im Bereich der Lageschätzung zu und damit in Anwendungen der Fertigungssteuerung oder im Bereich der Qualitätskontrolle.

Adaptive Modelle mit Materialdifferenzierung

Ebenfalls lohnenswert ist die Erweiterung des Konstruktionsprozesses für die adaptiven Modelle um Materialeigenschaften. Dies ist weniger zur Modellierung der Oberflächentexturen erforderlich als zur Stabilisierung der Modellanpassung. An Heckansichten von Fahrzeugen hat sich beispielsweise gezeigt, dass die Lageschätzung durch die Modellierung des Nummernschildes und der Heckscheibe erheblich an Robustheit gewinnt. Sowohl das Nummernschild als auch die Heckscheibe heben sich jedoch nur durch die Farbwerte und nicht durch die Form von der Karosserie ab. Hinzu kommt, dass Fahrzeuge zunehmend aerodynamischer und damit auch runder werden. Die Form liefert an solchen Rundungen nur wenige markante Modellkanten, wodurch eine robuste Modellanpassung erschwert wird.

Sensorfusion

Eine weitere vielversprechende Richtung ist die Fusion des Systems mit einer anderen Sensorik. Dies kann durch eine Gewichtung der Bildkantensegmente durch Merkmale vorgenommen werden, die aus Daten anderer Sensoren extrahiert wurden. Als Beispiel ist hier die Verwendung von Entfernungsdaten zu nennen, die - wenn sie mit dem Bild referenziert sind - im Bereich der Objektkontur Sprungstellen besitzen und somit zur Hervorhebung der Silhouettenkanten herangezogen werden können. Ebenfalls ist die Verwendung von Wärmebilddaten denkbar, z. B. um diejenigen Bildkanten, die zum Objekt gehören, gegenüber Schattenkanten abzugrenzen. Durch eine solche Sensorfusion ließe sich die Anwendung des Verfahrens auch für nicht kooperative, also visuell getarnte Objekte durchführen.

Anhang A

Die Hauptkomponentenanalyse

Zur Durchführung der Hauptkomponentenanalyse aus Abschnitt (3.1) stehen mehrere Verfahren zur Verfügung. Gegeben sind die zentrierten geometrischen Realisierungen $\Delta \mathbf{g}_n := \mathbf{g}_n - \bar{\mathbf{g}}$ für $n = 1, \dots, N$, die jeweils die Dimension

$$m = 3 \times \text{Anzahl Ecken der Meta-Topologie}$$

besitzen. Zu berechnen sind die Eigenvektoren \mathbf{e}_n und Eigenwerte λ_n der Datenmatrix

$$X = (\Delta \mathbf{g}_1 \dots \Delta \mathbf{g}_N) \in \mathbb{R}^{m \times N}.$$

Die Eigenvektoren und -werte lassen sich aus der Diagonalisierung der Kovarianzmatrix

$$C = \frac{1}{N} X X^T \in \mathbb{R}^{m \times m} \quad (\text{A.1})$$

berechnen. Eine solche Diagonalisierung (siehe z. B. [Jolliffe '86]) hat die Form

$$C = U \Sigma U^T, \quad (\text{A.2})$$

wobei die Matrix U die Form

$$U = (\mathbf{e}_1 \dots \mathbf{e}_N) \in \mathbb{R}^{m \times N}$$

hat, also die gesuchten Eigenvektoren liefert. Die Diagonalelemente der Matrix $\Sigma \in \mathbb{R}^{N \times N}$ entsprechen den Eigenwerten $\lambda_1, \dots, \lambda_N$ von C . Diese Methode ist allerdings im hier vorliegenden Fall nicht praktikabel, da adaptive Modelle durchaus 1 000 Ecken besitzen können und die Kovarianzmatrix folglich eine Dimension von $3\,000 \times 3\,000$ erreichen kann. Typischerweise wird die Dimension des Problems in solchen Fällen, also für $N \ll m$, in der Dimension reduziert, indem die Matrix

$$A = \frac{1}{N} X^T X \in \mathbb{R}^{N \times N}$$

diagonalisiert wird. Bezeichnen $\mathbf{u}_1, \dots, \mathbf{u}_N$ die Eigenvektoren und μ_1, \dots, μ_N die Eigenwerte der Matrix A , so sind die gesuchten Eigenvektoren der Kovarianzmatrix gegeben durch

$$\mathbf{e}_n = \frac{1}{\sqrt{N \mu_n}} X \mathbf{u}_n \quad (\text{A.3})$$

und die Eigenwerte der Kovarianzmatrix entsprechen gerade den μ_1, \dots, μ_N . Einen Beweis für Gleichung (A.3) findet man z. B. in [Blanz '00]. In derselben Arbeit wird jedoch darauf hingewiesen, dass die durch diese Methode berechneten Eigenvektoren weit von der Orthogonalität entfernt sind, ihre Skalarprodukte also erheblich von 0 bzw. 1 abweichen. Blanz vermutet die Ursache in der großen Konditionszahl, also dem großen Verhältnis zwischen dem größten und dem kleinsten Singulärwert der Matrix A . Er propagiert alternativ die Berechnung einer Singulärwertzerlegung der Matrix X , also eine Zerlegung der Form

$$X = U\Gamma V^T$$

mit orthogonalen Matrizen $U \in \mathbb{R}^{m \times N}$ und $V \in \mathbb{R}^{N \times N}$ sowie der Diagonalmatrix $\Gamma \in \mathbb{R}^{N \times N}$. Blanz zeigt, dass die gesuchte Diagonalisierung von C dann gegeben ist durch

$$C = \frac{1}{N}U\Gamma^2U^T,$$

wobei die $U \in \mathbb{R}^{m \times N}$ die größte auftretende Matrix ist, die zudem der Matrix U aus Gleichung (A.2) entspricht, sich also aus den gesuchten Eigenvektoren zusammensetzt. Die Eigenwerte der Kovarianzmatrix lassen sich aus den Diagonalelementen $\gamma_1, \dots, \gamma_N$ der Matrix Γ berechnen und zwar durch

$$\lambda_n = \frac{1}{N}\gamma_n^2.$$

Anhang B

Die Abbildungskette

Die Abbildung vom Objekt- ins Bildkoordinatensystem wird durch die Abbildung

$$\begin{pmatrix} X \\ Y \\ W \end{pmatrix} = \mathcal{P}\mathcal{A} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

realisiert, wobei die Matrix \mathcal{A} die affine Abbildung vom Objekt- ins Kamerakoordinatensystem beschreibt und die Matrix \mathcal{P} die Projektion vom Kamera- ins Bildkoordinatensystem. Die Transformationsmatrix \mathcal{A} wird durch die Eulerwinkel ω , φ und κ sowie drei Translationsparameter t_x , t_y und t_z parametrisiert. Die Rotation ist erklärt durch die Matrix

$$\mathcal{R} = \begin{pmatrix} \cos \kappa \cos \varphi & \cos \kappa \sin \varphi \sin \omega - \sin \kappa \cos \omega & \cos \kappa \sin \varphi \cos \omega + \sin \kappa \sin \omega \\ \sin \kappa \cos \varphi & \sin \kappa \sin \varphi \sin \omega + \cos \kappa \cos \omega & \sin \kappa \sin \varphi \cos \omega - \cos \kappa \sin \omega \\ -\sin \varphi & \cos \varphi \sin \omega & \cos \varphi \cos \omega \end{pmatrix}$$

und die Transformation lautet in homogener Schreibweise

$$\mathcal{A} = \left(\begin{array}{ccc|c} & & & t_x \\ & \mathcal{R} & & t_y \\ & & & t_z \\ \hline 0 & 0 & 0 & 1 \end{array} \right).$$

Die Projektionsmatrix wird im Fall einer zentralperspektivischen Projektion unter der Annahme eines Lochkameramodells definiert. Sie ist also durch den Bildhauptpunkt $(c_x, c_y)^T$, die Brennweite f und die Bildpunktbreite bzw. -höhe d_x bzw. d_y gegeben. Die homogene Projektionsmatrix lautet dann

$$\mathcal{P}_{perspective} = \begin{pmatrix} \frac{f}{d_x} & 0 & c_x & 0 \\ 0 & \frac{f}{d_y} & c_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (\text{B.1})$$

Um die perspektivische Division durchzuführen, muss der homogene Ergebnisvektor nach Anwendung der Projektionsmatrix normiert werden, so dass $W = 1$ gilt. Wird eine orthogonale Projektion durchgeführt, so ist die Projektionsmatrix durch

$$\mathcal{P}_{ortho} = \begin{pmatrix} \frac{1}{d_x} & 0 & 0 & c_x \\ 0 & \frac{1}{d_y} & 0 & c_y \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{B.2})$$

erklärt.

Anhang C

Bilddatenübersicht

Gruppe	Typ und Variante	Länge [m]	Spannweite [m]	Anzahl Bilder	Daten- bank	Modell
Airbus A300	A300-B2/B4	53,6	44,8	19	+	+
	A300-600/600R	54,1	44,8			
Airbus A310	A310-200/300	47,7	43,9	9	+	+
Airbus A318	A318	31,5	34,1	1	+	+
Airbus A320	A320-100/200	37,6	34,1	4	+	+
Airbus A321	A321-100/200	44,5	34,1	9	+	+
Airbus A340-200	A340-200	58,4	60,3	2	+	-
Airbus A340-300	A340-300	63,7	60,3	3	+	+
Boeing 737 klein	737-300	33,4	28,9	8	+	+
	737-400	36,4	28,9	2	+	+
	737-500	31,0	28,9	1	+	+
Boeing 737 groß	737-600	31,2	34,3	0	+	+
	737-700	33,6	34,3	0	+	+
Boeing 707	707-320B	46,6	44,4	0	+	+
Boeing kurz	747SP	56,3	59,6	8	+	+
Boeing 747 lang	747-100/200/300	70,7	59,6	8	+	+
	747-400	70,7	64,4	12	+	+
McDonnell DC 10	DC 10-30	55,0	50,4	1	-	-
McDonnell MD 11	MD 11	61,2	51,8	1	-	-

Tabelle C.1: Die Flugzeugvarianten wurden entsprechend dieser Tabelle gruppiert. Die Spalte „Datenbank“ indiziert, ob für die entsprechende Flugzeugvariante ein Eintrag in der Datenbank vorhanden ist (siehe Abschnitt 5.1.4). Die Spalte „Modell“ zeigt an, ob die Variante als Basismodell vorliegt (siehe Abschnitt 5.1.3). (Quelle technische Angaben: [Figgen et al. '00])

Literaturverzeichnis

- [Aumann & Spitzmüller '92] G. Aumann und K. Spitzmüller. *Computerorientierte Geometrie*, Bd. 89 in *Reihe Informatik*. BI Wissenschaftsverlag, 1992.
- [Berger et al. '00] L. Berger, M. Esswein, N. Heinze, J. Kanthak, A. Korn, W. Krüger, E. Lakatsch, A. Laubenheimer, N. Link, M. Müller, W. Schwerdtmann, G. Sauer, C.-K. Sung, C. Wetzel und G. Zimmermann. *WITMUS - Wissensbasierte, teilautomatische Bildauswertung für die multisensorielle Aufklärung, Abschlußbericht*. Fraunhofer IITB Bericht Nr. 10508, 2000.
- [Blake et al. '01] A. Blake, M. Isard und D. Reynard. Learning to Track the Visual Motion of Contours. *Artificial Intelligence*, 78, S. 101 – 133, 2001.
- [Blanz & Vetter '99] V. Blanz und T. Vetter. A Morphable Model for the Synthesis of 3D Faces. In *Proceedings Computer Graphics, SIGGRAPH'99*, S. 187 – 194, Los Angeles, CA, USA, 1999.
- [Blanz & Vetter '03] V. Blanz und T. Vetter. Face Recognition Based on Fitting a 3D Morphable Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 25, Nr. 9, S. 1063 – 1074, 2003.
- [Blanz '00] V. Blanz. *Automatische Rekonstruktion der dreidimensionalen Form von Gesichtern aus einem Einzelbild*. Dissertation, Fakultät für Physik der Eberhard-Karls-Universität zu Tübingen, 2000.
- [Bloomenthal & Ferguson '95] J. Bloomenthal und K. Ferguson. Polygonization of Non-Manifold Surfaces. In *Proceedings Computer Graphics, SIGGRAPH'95*, S. 309 – 316, Los Angeles, CA, USA, 1995.
- [Bloomenthal '94] J. Bloomenthal. An Implicit Surface Polygonizer. In P. S. Heckbert (Hrsg.), *Graphics Gems IV*, S. 324 – 349. Academic Press, 1994.
- [Bloomenthal '97] J. Bloomenthal (Hrsg.). *Introduction to Implicit Surfaces*. Morgan Kaufman, San Francisco, 1997.
- [Buhmann '01] M. D. Buhmann. *A new class of radial basis functions with compact support*, S. 307 – 318. American Mathematical Society, 2001.

- [Buxton '03] H. Buxton. Learning and Understanding Dynamic Scene Activity: a Review. *Image and Vision Computing*, 21, S. 125 – 136, 2003.
- [Campbell & Flynn '01] R. Campbell und P. Flynn. A Survey of Free-Form Object Representation and Recognition Techniques. *Proceedings Computer Vision and Image Understanding (CVIU)*, 81, Nr. 2, S. 166 – 210, 2001.
- [Carr et al. '01] J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum und T. R. Evans. Reconstruction and Representation of 3D Objects with Radial Basis Functions. In *Proceedings Computer Graphics, SIG-GRAPH 2001*, S. 67 – 76, Los Angeles, CA, USA, 2001.
- [Cohen-Or et al. '98] D. Cohen-Or, D. Levin und A. Solomovici. Three-Dimensional Distance Field Metamorphosis. *ACM Transactions on Graphics*, 17, Nr. 2, S. 116 – 141, 1998.
- [Comaniciu & Meer '97] D. Comaniciu und P. Meer. Robust Analysis of Feature Spaces: Color Image Segmentation. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97)*, S. 750 – 755, 1997.
- [Cootes & Taylor '01] T. F. Cootes und C. J. Taylor. Statistical Models of Appearance for Medical Image Analysis and Computer Vision. In *Proceedings of SPIE Medical Imaging*, S. 236 – 248, 2001.
- [Cootes et al. '00] T. F. Cootes, K. N. Walker und C. J. Taylor. View-Based Active Appearance Models. In *Proceedings of 4th International Conference on Automatic Face and Gesture Recognition*, S. 227 – 232, 2000.
- [Cramer '03] M. Cramer. Erfahrungen mit der direkten Georeferenzierung. *Photogrammetrie Fernerkundung Geoinformation*, 4, S. 267 – 278, 2003.
- [DeCarlo & Metaxas '02] D. DeCarlo und D. Metaxas. Adjusting Shape Parameters Using Model-Based Optical Flow Residuals. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 24, Nr. 6, S. 814–823, 2002.
- [Duchon '76] J. Duchon. Splines Minimizing Rotation-Invariant Semi-Norms in Sobolev Spaces. In A. Dold und B. Eckmann (Hrsg.), *Constructive Theory of Functions of Several Variables, Proceedings of a Conference, Lecture Notes in Mathematics 571*, S. 85 – 100, Oberwolfach, Germany, 1976.
- [Eisenreich & Sube '94] G. Eisenreich und R. Sube. *Dictionary of Mathematics*. Verlag Harri Deutsch, Thun und Frankfurt am Main, 1994.
- [Fautz '03] M. Fautz. *Objekt- und Texturrekonstruktion mit einer robotergeführten Kamera*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH). Shaker Verlag, Aachen, 2003.

- [Felzenszwalb & Huttenlocher '00] P. Felzenszwalb und D. Huttenlocher. Efficient Matching of Pictorial Structures. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2000)*, S. 66 – 73, Los Alamitos, CA, USA, 2000.
- [Ferryman et al. '95] J. Ferryman, A. Worrall, G. Sullivan und K. Baker. A Generic Deformable Model for Vehicle Recognition. In *Proceedings of British Machine Vision Conference*, S. 127 – 136, 1995.
- [Figgen et al. '00] A. Figgen, D. Plath und B. Rothfischer. *Verkehrsflugzeuge*. Bechtermünz Verlag, 2000.
- [Garland & Heckbert '97] M. Garland und P. S. Heckbert. Surface Simplification Using Quadric Error Metrics. *Computer Graphics*, 31, S. 209 – 216, 1997. Annual Conference Series.
- [Gengenbach et al. '98] V. Gengenbach, K. Schäfer, H. H. Nagel, K. Fleischer, H. Leuck, F. Muth, A. Bachem, W. Enkelmann, F. Heimes und M. Tonko. *Sichtgestütztes Andocken von Flugzeugen*, S. 31 – 35. IITB Mitteilungen. Fraunhofer Institut für Informations- und Datenverarbeitung (IITB), 1998.
- [Gengenbach '94] V. Gengenbach. *Einsatz von Rückkopplungen in der Bildauswertung bei einem Hand-Auge-System zur automatischen Montage*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH). Infix Verlag, Sankt Augustin, 1994.
- [Green '80] W. Green. *The Observer's Book of Aircraft*. Frederick Warne, London, 1980.
- [Green '92] W. Green. *Flugzeuge der Welt heute - morgen*. Werner Classen Verlag, Zürich, 1992.
- [Gregory et al. '98] A. Gregory, A. State, M. C. Lin, D. Manocha und M. A. Livingston. Feature-based Surface Decomposition for Correspondence and Morphing between Polyhedra. In *Proceedings Computer Animation '98*, S. 64 – 71, Philadelphia, USA, 1998.
- [Hashimoto & Skalansy '87] M. Hashimoto und J. Skalansy. Multiple-Order Derivatives for Detecting Local Image Characteristics. *Computer Vision, Graphics, and Image Processing*, 39, S. 28 – 55, 1987.
- [He et al. '94] T. He, S. Wang und A. Kaufmann. Wavelet-based Volume Morphing. In R. Bergeron und A. Kaufmann (Hrsg.), *Proceedings of Visualization 1994, IEEE Computer Society Press*, S. 85 – 92, Washington DC, USA, 1994.
- [Heckbert & Garland '97] P. S. Heckbert und M. Garland. Survey of Polygonal Surface Simplification Algorithms. Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA, 1997.

- [Hoppe et al. '93] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald und W. Stuetzle. Mesh Optimization. In *Proceedings Computer Graphics, SIGGRAPH'93*, S. 19 – 26, Anaheim, CA, USA, 1993.
- [Jähne '97] B. Jähne. *Digitale Bildverarbeitung*. Springer Verlag, Berlin, Heidelberg, 4. Auflage, 1997.
- [Jänich '94] K. Jänich. *Topologie*. Springer Verlag, Berlin, Heidelberg, 4. Auflage, 1994.
- [Jolliffe '86] I. J. Jolliffe. *Principal Component Analysis*. Springer Verlag, New York, 1986.
- [Kanai et al. '97] T. Kanai, H. Suzuki und F. Kimura. 3D Geometric Metamorphosis Based on Harmonic Maps. In *Proceedings 5th Pacific Graphics International Conference*, S. 87 – 104, Los Alamitos, CA, USA, 1997.
- [Kanai et al. '00] T. Kanai, H. Suzuki und F. Kimura. Metamorphosis of Arbitrary Triangular Meshes with User-Specified Correspondence. *IEEE Computer Graphics and Applications*, 20, Nr. 2, S. 62 – 75, 2000.
- [Karkanis & Stewart '01] T. Karkanis und A. Stewart. Curvature-Dependent Triangulation of Implicit Surfaces. *IEEE Computer Graphics and Applications*, 21, Nr. 2, S. 60 – 69, 2001.
- [Kecici '01] H. F. Kecici. *Form- und Positionsschätzung durch binokulare Bildfolgenauswertung*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH). Shaker Verlag, Aachen, 2001.
- [Kent et al. '92] J. R. Kent, W. E. Carlson und R. E. Parent. Shape Transformation for Polyhedral Objects. In *Proceedings Computer Graphics, SIGGRAPH'92*, S. 47 – 54, 1992.
- [Klausmann '00] P. Klausmann. *Anwendungsorientierte Leistungsprädiktion und Bewertung von Detektionsverfahren*. Dissertation, Fakultät für Elektrotechnik und Informationstechnik der Universität Karlsruhe (TH). Shaker Verlag, Aachen, 2000.
- [Kobbelt '00] L. Kobbelt. $\sqrt{3}$ -Subdivision. In *Proceedings Computer Graphics, SIGGRAPH 2000*, S. 103 – 112, New Orleans, LA, USA, 2000.
- [Koller et al. '93] D. Koller, D. Daniilidis und H. H. Nagel. Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes. *International Journal on Computer Vision*, 10, Nr. 3, S. 257 – 281, 1993.
- [Kölzow '03] T. Kölzow. *System zur Klassifikation und Lokalisation von 3D-Objekten durch Anpassung vereinheitlichter Merkmale in Bildfolgen*. Dissertation, Technische Fakultät der Universität Bielefeld, 2003.

- [Korn '99] A. Korn. Toward a Symbolic Representation of Intensity Changes in Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10, Nr. 5, S. 610 – 625, 1999.
- [Kostiainen et al. '01] T. Kostiainen, I. Kalliomäki, T. Tamminen und J. Lampinen. 3D Object Recognition Based on Hierarchical Eigen-Shapes and Bayesian Inference. In *Proceedings SPIE Intelligent Robots and Computer Vision XX*, S. 165 – 173, 2001.
- [Kroschel & Heckmann '02] K. Kroschel und M. Heckmann. Lip Parameter Extraction for Speechreading. In *Tagungsband 13. Konferenz Elektronische Sprachsignalverarbeitung (ESSV 2002)*, S. 58 – 65, Dresden, 2002.
- [Krüger '01] W. Krüger. Robust and Efficient Map-to-Image Registration with Line Segments. *Machine Vision and Applications*, 13, Nr. 1, S. 39 – 50, 2001.
- [Laubenheimer & Link '03] A. Laubenheimer und N. Link. Towards Adaptive Models for Classification of Technical Objects. In *Proceedings Vision Modeling and Visualization*, S. 319 – 327, München, 2003.
- [Lazarus & Verroust '97] F. Lazarus und A. Verroust. Metamorphosis of Cylinder-like Objects. *The Journal of Visualization and Computer Animation*, 8, Nr. 3, S. 131 – 146, 1997.
- [Lazarus & Verroust '98] F. Lazarus und A. Verroust. Three-dimensional Metamorphosis: a Survey. *The Visual Computer*, 14, Nr. 8/9, S. 373 – 389, 1998.
- [Lee et al. '99] A. Lee, D. Dobkin, W. Sweldens und P. Schröder. Multiresolution Mesh Morphing. In *Proceedings Computer Graphics, SIGGRAPH'99*, S. 343 – 350, Los Angeles, CA, USA, 1999.
- [Lerios et al. '95] A. Lerios, C. D. Garfinkle und M. Levoy. Feature-Based Volume Metamorphosis. In *Proceedings Computer Graphics, SIGGRAPH'95*, S. 449 – 456, Los Angeles, CA, USA, 1995.
- [Mauch '03] S. Mauch. *Efficient Algorithms for Solving Static Hamilton-Jacobi Equations*. Dissertation. California Institute of Technology, Pasadena, California, 2003.
- [McInerney & Terzopoulos '96] T. McInerney und D. Terzopoulos. Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis*, 1, Nr. 2, S. 91 – 108, 1996.
- [Morse et al. '01] B. S. Morse, T. S. Yoo, P. Rheingans, D. T. Chen und K. R. Subramanian. Interpolating Implicit Surfaces from Scattered Surface Data Using Compactly Supported Radial Basis Functions. In *Proceedings Shape Modeling International - SMI 2001, International Conference on Shape Modelling and Applications*, S. 89 – 98, Genova, Italy, 2001. IEEE Computer Society.

- [Narcowich et al. '99] F. J. Narcowich, R. Schaback und J. D. Ward. Multilevel Interpolation and Approximation. *Applied and Computational Harmonic Analysis*, 7, S. 243–261, 1999.
- [Ohtake & Belyaev '02] Y. Ohtake und A. Belyaev. Dual/Primal Mesh Optimization for Polygonized Implicit Surfaces. In *ACM Solid Modeling Symposium*, S. 171 – 178, Saarbruecken, Germany, 2002.
- [Ohtake et al. '02] Y. Ohtake, A. Belyaev und A. Pasko. Dynamic Mesh Optimization for Polygonized Implicit Surfaces with Sharp Features. *The Visual Computer*, 19, Nr. 2 – 3, S. 115 – 126, 2002.
- [Parent '92] R. E. Parent. Shape Transformation by Boundary Representation Interpolation: a Recursive Approach to Establishing Face Correspondences. *The Journal of Visualisation and Computer Animation*, 3, Nr. 4, S. 219 – 239, 1992.
- [Pasko & Savchenko '95] A. Pasko und V. Savchenko. Construction Functionally Defined Surfaces. In *Proceedings Implicit Surfaces 1995*, S. 97 – 106, Genoble, France, Eurographics Association, 1995.
- [Plagemann '04] C. Plagemann. Ansichts-basierte Erkennung und Lokalisierung von Objekten zur Initialisierung eines Verfolgungsprozesses. Diplomarbeit, Universität Karlsruhe (TH), Fakultät für Informatik, 2004.
- [Pope & Lowe '96] A. Pope und D. Lowe. Learning Appearance Models for Object Recognition. In *Proceedings International Workshop on Object Representation for Computer Vision*, S. 201 – 219, Cambridge, UK, 1996.
- [Press et al. '92] W. H. Press, S. A. Teukolsky, W. T. Vetterling und B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2. Auflage, 1992.
- [Rosenhahn et al. '03] B. Rosenhahn, C. Perwass und G. Sommer. Pose Estimation of Free-Form Surface Models. In *25. Symposium für Mustererkennung, DAGM 2003*, S. 574 – 581, Magdeburg, 2003.
- [Schlosser et al. '03] C. Schlosser, J. Reitberger und S. Hinz. Automatic Car Detection in High Resolution Urban Scenes Based on Adaptive 3D-Modells. In *Proceedings of Remote Sensing and Data Fusion over Urban Areas*, Berlin, 2003.
- [Selinger & Nelson '00] A. Selinger und R. C. Nelson. Improving Appearance-Based Object Recognition in Cluttered Backgrounds. In *Proceedings of the International Conference on Pattern Recognition*, S. 46 – 50, Barcelona, Spain, 2000.
- [Shapiro & Tsukanov '99] V. Shapiro und I. Tsukanov. Implicit Functions with Guaranteed Differential Properties. In *Proceedings of the Fifth ACM Symposium on Solid Modeling and Applications*, S. 852 – 857, 1999.

- [Shelton '00] C. R. Shelton. Morphable Surface Models. *International Journal of Computer Vision (IJCV)*, 38, Nr. 1, S. 75 – 91, 2000.
- [Stegmann '01] M. B. Stegmann. Object Tracking Using Active Appearance Models. In Søren I. Olsen (Hrsg.), *Proceedings 10th Danish Conference on Pattern Recognition and Image Analysis*, Bd. 1, S. 54–60, Copenhagen, Denmark, 2001.
- [Steward '99] C. Steward. Robust Parameter Estimation in Computer Vision. *Siam Review*, 41, Nr. 3, S. 513 – 537, 1999.
- [Sullivan et al. '95] G. Sullivan, A. Worrall und J. Ferryman. Visual Object Recognition Using Deformable Models of Vehicles. In *Context95*, 1995.
- [Terzopoulos & Metaxas '91] D. Terzopoulos und D. Metaxas. Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, Nr. 7, S. 703 – 714, 1991.
- [Theobalt et al. '03] C. Theobalt, J. Carranza, M. Magnor und H.-P. Seidel. A Parallel Framework for Silhouette-based Human Motion Capture. In *Proceedings Vision Modeling and Visualization*, S. 207 –214, München, 2003.
- [Tsai '87] R. Y. Tsai. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. *IEEE Transactions on Robotics and Automation*, 3, Nr. 4, S. 323 – 344, 1987.
- [Turk & O'Brian '99] G. Turk und J. O'Brian. Shape Transformation Using Variational Implicit Functions. In *Proceedings Computer Graphics, SIGGRAPH'99*, S. 335 – 342, Los Angeles, CA, USA, 1999.
- [Turk & Pentland '91] M. Turk und A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3, Nr. 1, S. 71 – 86, 1991.
- [Vasilescu & Terzopoulos '02] M. Vasilescu und D. Terzopoulos. Multilinear Image Analysis for Facial Recognition. In *Proceedings of the International Conference on Pattern Recognition*, S. 511 – 514, 2002.
- [Wachter '97] S. Wachter. *Verfolgung von Personen in monokularen Bildfolgen*. Dissertation, Fakultät für Informatik der Universität Karlsruhe (TH). Vice Versa Verlag, Berlin, 1997.
- [Wendland '95] H. Wendland. *Piecewise Polynomial, Positive Definite and Compactly Supported Radial Functions of Minimal Degree*, S. 389 – 396. J. C. Baltzer AG, Science Publishers, 1995.
- [Zhang & Kambhamettu '00] Y. Zhang und C. Kambhamettu. Robust 3D Head Tracking Under Partial Occlusion. In *Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, S. 176 – 182, 2000.

- [Zhao & Nevatia '01] T. Zhao und R. Nevatia. Car Detection in Low Resolution Aerial Image. In *Proceedings of the International Conference on Computer Vision*, 2001.