

Subpixel genaue Bildregistrierung

Studienarbeit im Studiengang Computervisualistik

vorgelegt von

Sabine Achilles

Betreuer: Dipl.-Inf. Detlev Droege, Institut für Computervisualistik, Fachbereich Informatik
Erstgutachter: Dipl.-Inf. Detlev Droege, Institut für Computervisualistik, Fachbereich Informatik
Zweitgutachter:

Koblenz, im Mai 2007

Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Die Richtlinien der Arbeitsgruppe für Studien- und Diplomarbeiten habe ich gelesen und anerkannt, insbesondere die Regelung des Nutzungsrechts

Mit der Einstellung dieser Arbeit in die Bibliothek bin ich einverstanden. ja nein

Der Veröffentlichung dieser Arbeit im Internet stimme ich zu. ja nein

Koblenz, den

Unterschrift

Inhaltsverzeichnis

1	Einleitung	7
1.1	Problemstellung	7
1.2	Überblick eigene Implementierung	9
1.3	Gliederung	10
2	Stand der Wissenschaft	13
3	Erläuterung und Umsetzung der Verfahren	17
3.1	Direkte Methode	18
3.1.1	Überblick	18
3.1.2	Herleitung	20
3.1.3	Implementierung	22
3.2	Merkmalsbasierte Methode	23
3.2.1	Überblick	23
3.2.2	Matching von Merkmalen	25
3.2.3	DLT - Direkte Lineare Transformation	27
3.2.4	RANSAC - RANdom SAmple Concensus	30
4	Experimente und Ergebnisse	33

4.1	Untersuchung des Rotationsverhaltens	33
4.2	Untersuchung des Translationsverhaltens	34
4.3	Untersuchung der Robustheit gegenüber Farb- und Helligkeitsschwankungen	36
4.4	Versuche mit realer Sequenz	38
4.4.1	Registrierung mit Anwendungsbeispiel	40
5	Zusammenfassung	43
	Literaturverzeichnis	46

Kapitel 1

Einleitung

1.1 Problemstellung

In der Bildverarbeitung haben sich in den letzten Jahren immer mehr Anwendungsbereiche entwickelt, in denen eine Serie von Bildern derselben Szene so aneinander ausgerichtet ist, dass ihr Inhalt genau übereinander liegt. Dieses Problem entsteht, weil die Ausgangsbilder meist ohne Wissen ihrer relativen Position aufgenommen wurden, sie zueinander also transliert, rotiert oder sogar skaliert sein könnten. Zusätzlich können noch - abhängig von der zeitlichen Länge der aufgenommenen Bildserie - Probleme durch Beleuchtungsschwankungen und nicht immer existierende bzw. sich bewegende Objekte (wie etwa durch das Bild laufende Menschen) im Bildbereich entstehen.

Eine solche Bildregistrierung wird zum Beispiel auch bei der Erstellung von Superresolution Bildern gebraucht, bei der eine Serie von niedrig aufgelösten Bildern benutzt wird, um daraus ein höher aufgelöstes Bild zu berechnen. Die Bilder der niedrig aufgelösten Bildserie werden dazu leicht versetzt voneinander aufgenommen und müssen mittels Bildregistrierung zu einem Mosaik zusammengefügt werden, welches dann die Auflösung des hoch aufgelösten Bildes besitzt. So können mehr Informationen über die aufgenommene Szene gesammelt und zu einem Superresolution Bild vereint werden. Um ein gutes Superresolution Ergebnis zu erzielen, ist es extrem wichtig, dass subpixel-genaue Informationen über vorliegende Veränderungen der Kameraposition vorliegen, da es sonst sehr schnell zu

Fehlbildern kommen kann. Schon bei einer falschen Berechnung der Translation von 0,5 Pixeln kann es zu einer deutlich erkennbaren Unschärfe im Superresolution Bild kommen.



Bild 1.1: 1. Reihe: Mit einer Kamera aufgenommene Bildserie, 2. Reihe: Bildserie nach der Registrierung

Die verschiedenen Anwendungsbereiche für Bildserien stellen jedoch unterschiedliche Ansprüche an einen Bildregistrierungsalgorithmus. Deshalb kann nicht von einem perfekten Ansatz für das Problem Bildregistrierung gesprochen werden, sondern sollte den Algorithmus stets für den jeweiligen späteren Verwendungszweck optimieren. Für die spätere Verwendung für Superresolution-Algorithmen können somit gewisse Einschränkungen formuliert werden, deren Registrierungsansätze zum Beispiel für eine Registrierung von HDR-Bildern nicht in Frage kämen:

- Die Bilder werden idealerweise direkt an einem Stück aufgenommen. So entstehen also kaum Beleuchtungsschwankungen und es kann zusätzlich darauf geachtet werden, dass die Objekte der Szene statisch sind.
- Die Translationen, Rotationen und Skalierungen liegen in einem minimalen Bereich. Das liegt deshalb im Sinne der Auflösungserhöhung, da auf diese Weise weniger Bereiche mit Informationen von einem oder nur wenigen Bildern entstehen, und die Überlappungsregionen groß gehalten werden können, in denen viele Zusatzinformationen über die Szene vorhanden sind.
- Da die meisten Algorithmen zur Registrierung nicht skalierungs-invariant sind, kann

ggf. noch darauf geachtet werden, dass sich die Position der Kamera von der Entfernung zur Szene hin nicht verändert.

- Auch auf die Rotation könnte theoretisch verzichtet werden, ohne einen Qualitätsverlust in den Superresolution Ergebnissen zu erleiden. Da die meisten Algorithmen jedoch rotationsinvariant sind, sollte sie nur zur Not weggelassen werden.

Minimal wird also ein Algorithmus gebraucht, der subpixel-genau Translationen berechnet und maximal auch mit Rotationen und Skalierung umgehen kann.

Ziel dieser Studienarbeit ist es, verschiedene Bildregistrierungsverfahren zu implementieren und deren Subpixelgenauigkeit zu untersuchen.

1.2 Überblick eigene Implementierung

Die verschiedenen Ansätze zur Registrierung können in zwei grundlegende Bereiche eingeteilt werden: die direkte- und die merkmalsbasierte Bildregistrierung.

In der direkten Registrierung werden alle Pixel des Bildes benutzt, um die Bilder aneinander auszurichten, wohingegen in der merkmalsbasierten Registrierung nur bestimmte Bereiche des Bildes mit wertvollen Informationen benutzt werden.

Im Rahmen dieser Studienarbeit habe ich das direkte Verfahren nach [KPB88] implementiert, das die Differenz der Intensitäten der Bilder iterativ minimiert. Um den Prozess zu beschleunigen, wird eine Gaußpyramidenstruktur benutzt, bei der erst eine Bildversion mit reduzierter Auflösung untersucht wird, und dann die Auflösung Schritt für Schritt wieder erhöht wird.

Als merkmalsbasiertes Verfahren habe ich mich in groben Zügen an den Algorithmus aus [HZ03] gehalten, bei dem zuerst Merkmale aus den Bildern extrahiert, und diese dann aufgrund ihrer Nähe und Ähnlichkeit in der Intensitätsnachbarschaft zu Merkmalspaaren zusammengefügt werden.

Zur Merkmalsdetektion habe ich zum einen den Harris Corner Detector und zum anderen den Tomasi-Kanade Point Detector benutzt, die beide in PUMA (die Bibliothek für

Bildverarbeitung und Musteranalyse der Universität Koblenz) implementiert sind. Für das Matching der Merkmale habe ich die normalisierte Kreuzkorrelation (NCC) implementiert, die die Merkmalspaare mit der größten Korrelation zu einem Merkmalspaar zusammenfügt.

Auf diese recht ungenaue und fehlerbehaftete Menge von Merkmalspaaren wird als nächstes der RANSAC Algorithmus angewandt, der die Paarmenge von groben Fehlern befreit und sie damit in zwei Untermengen aufteilt: *Inlier* und *Outlier*. Inlier sind die Paare, die innerhalb eines Schwellwerts liegen und somit zusammenpassen. Outlier hingegen sind diejenigen Paare, die falsch zusammengefügt wurden.

RANSAC wählt in jedem Schritt 4 zufällige Proben von Merkmalspaaren, aus denen mittels Diskreter Linearer Transformation (DLT) eine Homographie berechnet wird. Auf Basis dieser Homographie, wird die Menge der Inlier und Outlier bestimmt. Dazu wird jedes Merkmalspaar mit dem Symmetric Transfer Error bewertet. Ist der berechnete Fehler unter einem bestimmten Schwellwert, so wird das Merkmalspaar als Inlier gewertet. Ist der Algorithmus terminiert, wird diejenige Probe als Lösung genommen, die die größte Unterstützung durch Inlier hat. Um diese Lösung weiter zu verfeinern, wird anhand der Menge der Inlier noch einmal eine Homographie mit DLT berechnet.

1.3 Gliederung

Zuerst wird in Kapitel 2 ein Überblick über die verschiedenen Verfahren zur Bildregistrierung gegeben, und die von mir in dieser Studienarbeit behandelten Methoden in Zusammenhang gestellt.

In Kapitel 3 folgt eine detaillierte Beschreibung und Herleitung der direkten und der merkmalsbasierten Registrierungstechnik mit verschiedenen Ansätzen zur Fehlerminimierung. Einer allgemeinen Herleitung der direkten Methoden folgt eine eingehende Schilderung des von mir implementierten Verfahrens. Bei der merkmalsbasierten Methode gebe ich zuerst einen allgemeinen Überblick über den schrittweisen Ablauf des Algorithmus, gefolgt von einer genaueren Beschreibung der einzelnen Schritte.

Kapitel 4 ist eine Zusammenfassung der von mir erzielten Ergebnisse mit einem Vergleich

der beiden Methoden unter verschiedenen Voraussetzungen. Dabei wird das Verhalten der beiden Verfahren bei Rotation, Translation, Helligkeitsschwankungen und Farbstichen anhand einer synthetisch erstellten Bildserie untersucht. Zusätzlich werden auch noch Ergebnisse einer realen Bildserie und dessen Anwendungsbeispiel (der Superresolution) gezeigt.

Abschließend wird in Kapitel 5 eine Zusammenfassung dieser Arbeit gegeben.

Kapitel 2

Stand der Wissenschaft

In der Bildverarbeitung existieren viele Ansätze zur Bildregistrierung, die teilweise pixel genau, teilweise subpixel genau arbeiten. In den letzten Jahren wurden eine Vielzahl von Registrierungstechniken entwickelt, die auf verschiedene Problemstellungen spezialisiert sind. Sie sind entstanden, da fast jede Forschungsrichtung andere Anforderungen an einen Registrierungsalgorithmus stellt. Den historisch ersten Überblick über Registrierungsverfahren bildet [KS83], in dem größtenteils Methoden basierend auf Bildkorrelationen behandelt werden. Den umfangreichsten Überblick bietet [GB92], das im Jahre 2003 durch [Zit03] ergänzt wurde. Gute Zusammenfassungen von medizinischen Registrierungstechniken bieten [HBHH01] [LA99] und [MV96].

In der medizinischen Bildverarbeitung wird meist auf die automatische Erkennung von Merkmalen verzichtet, da die Bilder diese Informationen oft nicht hergeben. Stattdessen setzt der Arzt in solchen Fällen Marker an charakteristischen Stellen entweder nach der Aufnahme auf den Bildern (intrinsische Registrierung [MV96]) oder vorher direkt auf dem Körper des Patienten (extrinsische Registrierung [MV96] [SAX⁺94]). Die meisten Methoden verwenden zur Lösungsfindung ein Minimierungsproblem, wie zum Beispiel *Powells Methode* [HHH91], die *Downhill Simplex Methode* [HHHR93], der *Levenberg-Marquardt Algorithmus* [FPM⁺94] oder die *Gradient Descent Methode* [CKMV96]. Diese Verfahren werden oft durch eine mehrstufige Auflösungs- oder eine Skalierungstechnik erweitert, um lokale Minima zu vermeiden. Außerdem ist vielen Methoden gemein, dass

sie mehr als ein Optimierungsverfahren verwenden: ein schnelles, aber ungenaues zur Lösungsapproximation und dann ein langsames zur exakten Lösungsfindung. Viele dieser Optimierungsverfahren werden in [PVTf02] beschrieben und erklärt.

Im Bereich des Rechnersehens und der Mustererkennung sowie der computergestützten Photographie entstehen andersartige Probleme in der Registrierung von Bildern, die auch durch andere Registrierungstechniken gelöst werden müssen. So entstehen hier oft Bilder, die aus unterschiedlichen Blickwinkeln aufgenommen wurden (Stereovision) oder Bildserien, die eine Objektverfolgung benötigen.

Im Allgemeinen kann gesagt werden, dass die meisten Methoden zur Bildregistrierung aus den folgenden Schritten bestehen:

- *Merkmalsdetektion*: manuelle oder automatische Detektion von interessanten Merkmalen (Punkte, Linien oder komplexere Formen).
- *Matching von Merkmalen*: Punktkorrespondenzen werden gebildet und mit Ähnlichkeitsmaßen bewertet.
- *Parameterbestimmung der Transformation*: Anhand der Merkmalspaare werden die Mapping-Parameter bestimmt. Hierzu können verschiedenste Fehlermaße benutzt werden.
- *Transformation und Resampling der Bilder*: die berechneten Transformationen werden angewandt und mögliche subpixel Koordinaten werden interpoliert.

Die Registrierungsverfahren können in zwei große Gebiete unterteilt werden: die *direkten Methoden* und die *merkmalbasierten Methoden*. Während die direkten Methoden große Teile oder sogar das gesamte Bild als Merkmale ansehen und diese dann korrelieren, detektieren merkmalsbasierte Methoden einige herausragende Merkmale im Bild und korrelieren diese. Solche Merkmale können Punkte, Linien oder komplexere Formen sein, von denen angenommen wird, dass ihre Position über eine bestimmte Zeitspanne hinweg unverändert bleibt.

Die am häufigsten genutzten Formen sind Punktmerkmale, also Schnittpunkte von Linien, die auch als Punkte mit hoher Krümmung in der Umgebung verstanden werden kön-

nen. Einen guten Überblick über die verschiedenen Punktdetektoren bieten [Roh01] und [ZWT99]. Der meistgenutzte Detektor ist der *Harris Corner Detector* [HS88] [Nob88], der eine Verbesserung des *Moravec Corner Detector* [Mor80] darstellt und auf der lokalen Autokorrelationsfunktion eines Signals basiert. Aufbauend darauf wurde der *Tomasi-Kanade Corner Detector* entwickelt, in dem der „Cornerness“-Wert verbessert berechnet wird. Mikolajczyk und Schmid haben den Harris Corner Detector auf Skalierungsinvarianz erweitert [MS04], indem sie eine mehrstufig skalierte Darstellung berechnen und dann die Punkte wählen, an denen der Laplace-Operator maximal über alle Skalierungsstufen ist. Ein weiterer guter Merkmalsdetektor ist der *Scale Invariant Feature Transform (SIFT)* [Low99], der auf lokalen 3D Extrema in einer, mit difference-of-Gaussian berechneten, Skalenraumpyramide basiert.

Sind die Merkmale gefunden, werden im nächsten Schritt Merkmalspaare gebildet, die aller Wahrscheinlichkeit nach zusammen gehören. Ein oft genutztes Verfahren ist die *normalisierte Kreuzkorrelation (NCC)* [Lew01], bei der die Intensitätsnachbarschaft der Merkmalspunkte miteinander verglichen, und die Paarung mit der größten berechneten Korrelation als Merkmalspaar angesehen wird. Andere Methoden arbeiten mit komplexeren Merkmalsdeskriptoren, wie z. B. zusätzlich die räumliche Verteilung der umliegenden Merkmale mit einzubeziehen [Mur92]. Auch die SIFT Deskriptoren bestehen aus einer erweiterten Beschreibung der Umgebung.

Das Ergebnis der Transformationsbestimmung verändert das Sequenzbild so, dass es zum Schluss genau über dem Referenzbild liegt und bestenfalls kein Unterschied mehr zu sehen ist. Die Merkmale des einen Bildes sollten dann also so nah wie möglich bei den korrespondierenden Merkmalen des anderen Bildes liegen. Um dies zu erreichen, muss zuerst eine geeignete, also zu der möglichen geometrischen Deformation des Sequenzbildes passende, Mappingfunktion gewählt werden. Die zur Berechnung benötigten Anzahl der Parameter, und damit auch die Anzahl der mindestens erforderlichen Merkmalspaarproben, wird durch die Freiheitsgrade (Degrees Of Freedom - DOF) der Transformation vorgegeben. Mögliche Funktionen sind hierbei Drehstreckung (4 DOF, oft bei Satellitenaufnahmen oder Scannerbildern [CKK⁺96] [SV95]), affine Transformation (6 DOF, Aufnahmen durch Teleobjektiv [IAH95] [CKK⁺96]), planare projektivische Transformation (8 DOF, [OLL04]) oder sogar elastische oder biquadratische Transformationen (12 DOF

[KAI⁺95] [ZMI99]).

Eine mögliche Beschreibung einer Transformation ist die Homographie (welche eine planare projektivische Transformation ist). Diese kann durch Maximum-Likelihood-Schätzer, wie z. B. der *diskreten linearen Transformation*, anhand der berechneten Merkmalspaare approximiert werden. Um die Ergebnisse gegen Outlier (falsch zusammengefügte Merkmale) robust zu machen, wird in [HZ03] zusätzlich die Benutzung des *RAndom SAmple Consensus (RANSAC)* vorgeschlagen. Okuma, Little und Lowe haben in [OLL04] am Beispiel eines NHL Spiels, ein auf RANSAC basierendes Verfahren beschrieben. Eine Verallgemeinerung des RANSAC ist der von Torr und Zisserman in [TZ00] vorgestellte MLE-SAC, der das Ergebnis anhand einer Maximum Likelihood Schätzung anstatt der Anzahl der Inlier bestimmt.

Bei direkten Methoden wiederum werden zu der Deformation passende Gleichungen aufgestellt [AI99], die als Gleichungssystem oder mit Maximum-Likelihood-Schätzern gelöst werden können. Durch ihre generelle Subpixelgenauigkeit arbeiten die direkten Methoden sehr gut mit Superresolution Algorithmen zusammen. Ein Beispiel ist die subpixel genaue Registrierung von Keren, Peleg und Brada [KPB88], in der unter Verwendung von Taylorreihen ein Gleichungssystem mit drei Gleichungen und drei Unbekannten erstellt wird. Auch im Mosaicing [IAH95] [IAB⁺96] [SA96] [Sze94] oder in der Videoaufarbeitung [IAH95] [IAB⁺96] finden direkte Methoden Verwendung.

Kapitel 3

Erläuterung und Umsetzung der Verfahren

Um zwei Bilder aneinander zu registrieren, muss eine Transformation gefunden werden, die jeden Punkt des einen Bildes in einen des anderen Bildes überführt. Es gibt viele verschiedene Arten, eine solche Transformation zu bestimmen, ich habe mich in meiner Studienarbeit mit Methoden aus zwei unterschiedlichen Bereichen beschäftigt, einer *direkten Methode* nach [KPB88] bzw. [IP91] und einer *merkmalsbasierten Methode* nach [HZ03]. Der Unterschied zwischen den beiden Verfahren ist, dass die direkte Methode die gesamten verfügbaren Pixelinformationen benutzt, um eine Transformation zu berechnen, wohingegen die merkmalsbasierte Methode nur herausstechende Informationen aus bestimmten Bereichen verwendet (z. B. Ecken oder Kanten).

Obwohl die Verfahren auch ohne Probleme auf eine Serie von mehreren Bildern erweitert werden können, beziehe ich mich in meinen Ausführungen immer auf zwei Bilder (das *Referenzbild* und das *Sequenzbild*), die zueinander ausgerichtet werden sollen.

3.1 Direkte Methode

3.1.1 Überblick

Grundlage für die Direkten Methoden ist die Minimierung eines Fehlermaßes, das aus Informationen aller verfügbaren Pixel gebildet wird.

Intensitätsbedingung

Ausgangspunkt der meisten Methoden ist der *brightness constancy constraint*, der die Intensitäten zweier Bilder in Zusammenhang stellt. Seien zwei Bilder $I_{x,y}$ und $I'_{x,y}$ gegeben, so ist:

$$I_{x,y} = I'_{x+u_{x,y},y+v_{x,y}} \quad (3.1)$$

mit den Pixelkoordinaten x, y und einer Verschiebung von $u_{x,y}$ und $v_{x,y}$ zwischen den beiden Bildern. Wird nun zusätzlich angenommen, dass diese Verschiebung sehr klein ist, kann folgende Gleichung aufgestellt werden:

$$I_x u + I_y v + I_t = 0 \quad (3.2)$$

mit den jeweils nach x und y abgeleiteten Bildintensitäten I_x und I_y und $I_t = I - I'$. Diese Berechnung muss für jedes Pixel x, y durchgeführt werden. Da die genaue Verschiebung eines Pixels durch zwei Variablen u und v definiert ist, kann diese jedoch nicht allein durch die Bildintensität bestimmt werden. Deshalb wird noch eine zweite Bedingung hinzugenommen, das *globale Bewegungsmodell* (global motion model).

Globales Bewegungsmodell in 2-D

Das meistgenutzte globale Bewegungsmodell ist die affine Transformation, welche eine sehr gute Annäherung an die Verschiebung zweier Bilder darstellt, wenn die Kamera sich

in weiter Entfernung zum aufgenommenen Objekt befindet:

$$\begin{aligned} u_{x,y} &= a_1x + b_1y + c_1 \\ v_{x,y} &= a_2x + b_2y + c_2 \end{aligned} \quad (3.3)$$

Andere Modelle, die in direkten Methoden benutzt werden, sind z. B. das quadratische Bewegungsmodell [BAHH92] [IRP94], das auch kleine Kamerarotationen zulässt, oder die 2-D projektive Transformation (Homographie) [SC94], die alle möglichen Kamerabewegungen abdeckt.

All diese Modelle können in Formel 3.2 integriert werden, aus dem sich dann folgende Gleichung (hier am Beispiel der affinen Transformation) mit sechs Unbekannten ergibt:

$$I_x(a_1x + b_1y + c_1) + I_y(a_2x + b_2y + c_2) + I_t = 0 \quad (3.4)$$

Es reichen also sechs Pixel aus, um eine Verschiebung zweier Bilder zu berechnen. Da dies aber, bedingt durch Rauschen und Pixelfehler, zu ungenauen Ergebnissen führt, werden in der Praxis die Gleichungen von allen betrachteten Pixeln (entweder ein Ausschnitt des Bildes oder auch das gesamte Bild) in die Berechnung mit einbezogen und die Summe der Fehlerquadrate minimiert:

$$E(a_1, b_1, c_1, a_2, b_2, c_2) = \sum (I_x(a_1x + b_1y + c_1) + I_y(a_2x + b_2y + c_2))^2 \quad (3.5)$$

Wie nun zu sehen ist, ist die Gewichtung, zu der jedes Pixel in das Ergebnis einfließt, abhängig von der Struktur des Bildes. So haben z. B. Pixel auf einer horizontalen Kante im Bild einen hohen I_y - Wert zur Folge, wohingegen der I_x - Wert kaum eine Rolle spielt. Liegt ein Pixel in einer homogenen Region, fließt dieser kaum in das Ergebnis mit ein.

Eine direkte Methode braucht somit keine bestimmten Merkmale, anhand denen eine Positionierung bestimmt werden kann. Die zu registrierenden Bilder müssen lediglich eine normalisierte Helligkeit haben und genug Intensitätsunterschiede in x - und y - Richtung aufweisen, die eine ausreichende Gradientenbestimmung zulassen.

Iterative Verfeinerung

Da die oben beschriebene Linearisierung der Intensität in Formel 3.2 nur gut funktioniert, wenn die Verschiebung der Bilder gering ist, wird in der Praxis ein *coarse-to-fine-refinement* benutzt, um mit den häufig doch größeren Abständen zwischen zwei Bildern zurecht zu kommen. Dafür wird eine mehrstufige Gaußpyramidenstruktur benutzt, in der die Bilder jeweils - mit Hilfe einer Gaußmaske - um die Hälfte in der Auflösung reduziert werden, und somit unterschiedlich aufgelöste Versionen der Bilder entstehen (kleine Auflösung an der Spitze und normale Auflösung am Boden der Pyramide). Nun wird das angewandte Verfahren zuerst auf dem am niedrigsten aufgelösten Bild durchgeführt und dann das Ergebnis iterativ zur nächsten Auflösungsstufe propagiert, bis die eigentliche Auflösung erreicht ist.

Somit können direkte Verfahren gut mit einer Verschiebung von bis zu 10% der Bildgröße und einer Rotation von bis zu 10° umgehen. Sollte eine größere Entfernung vorliegen, wird jedoch eine Anfangsschätzung nötig.

3.1.2 Herleitung

Die von mir gewählte Methode basiert auf der Registrierungstechnik von Lucas und Kanade [KL81] und wurde in [KPB88] erstmals vorgestellt. Sie minimiert die Intensitätsdifferenz der zu registrierenden Bilder iterativ, bis ein bestimmter Schwellwert erreicht ist. Dazu wird eines der Bilder in jedem Schritt dem anderen immer weiter angepasst. Diese Transformationen werden durch ein Gleichungssystem mit drei Gleichungen und drei Unbekannten berechnet.

Gegeben sind die zwei Bilder I und I' , deren horizontale Verschiebung a , vertikale Verschiebung b und Rotation Θ durch folgende Gleichung ausgedrückt werden kann:

$$I(x, y) = I'(x \cos(\Theta) - y \sin(\Theta) + a, y \cos(\Theta) - x \sin(\Theta) + b) \quad (3.6)$$

Erweitert man nun $\cos(\Theta)$ und $\sin(\Theta)$ mit den ersten zwei Termen ihrer Taylorreihe, so erhält man:

$$I(x, y) \approx I' \left(x + a - y\Theta - \frac{x\Theta^2}{2}, \quad y + b + x\Theta - \frac{y\Theta^2}{2} \right) \quad (3.7)$$

Wird nun wiederum I' mit dem ersten Term seiner Taylorreihe erweitert, ergibt sich:

$$I(x, y) \approx I'(x, y) + \left(a - y\Theta - \frac{x\Theta^2}{2} \right) \frac{\delta I'}{\delta x} + \left(b + x\Theta - \frac{y\Theta^2}{2} \right) \frac{\delta I'}{\delta y} \quad (3.8)$$

Daraus kann dann eine Fehlerfunktion zwischen I und I' aufgestellt werden:

$$E(a, b, \Theta) = \sum \left[I'(x, y) + \left(a - y\Theta - \frac{x\Theta^2}{2} \right) \frac{\delta I'}{\delta x} + \left(b + x\Theta - \frac{y\Theta^2}{2} \right) \frac{\delta I'}{\delta y} - I(x, y) \right]^2 \quad (3.9)$$

Wird diese Funktion nun minimiert, indem jeweils die Ableitungen nach a , b und Θ gebildet und einige nicht-lineare Terme und kleine Koeffizienten vernachlässigt werden, so erhält man drei Gleichungen mit drei Unbekannten:

$$\begin{aligned} \sum I_x^2 a + \sum I_x I_y b + \sum A I_x \Theta &= \sum I_x I_t \\ \sum I_x I_y a + \sum I_y^2 b + \sum A I_y \Theta &= \sum I_y I_t \\ \sum A I_x^2 a + \sum A I_x b + \sum A^2 \Theta &= \sum A I_t \end{aligned}$$

mit $I_x = \frac{\delta I}{\delta x}$, $I_y = \frac{\delta I}{\delta y}$, $I_t = I' - I$ und $A = xI_x - yI_y$. a , b und Θ werden nun aus diesem Gleichungssystem berechnet, wobei lediglich die Werte auf der rechten Seite des Gleichheitszeichens in jedem Schritt neu berechnet werden müssen, die neun restlichen nur einmal am Anfang, wodurch die Rechenzeit in den Iterationen selber, aber auch bei der Berechnung von mehreren Bildern in einer Bildserie klein gehalten werden kann.

Das Gleichungssystem gilt jedoch nur unter der Annahme, dass Verschiebung und Rotation gering sind. Da aber in der Praxis durchweg größere Unterschiede vorkommen, wird noch die in 3.1.1 beschriebene iterative Verfeinerung angewandt, die dem Algorithmus zusätzliche Robustheit bringt.

3.1.3 Implementierung

Die konkrete Implementierung sieht wie folgt aus:

1. Die Bilder werden in die unterschiedlichen Auflösungsstufen der Gaußpyramide überführt (jeweils Auflösung halbiert, siehe 3.1.1).
2. Beginnend mit der niedrigsten Auflösung bis hin zur höchsten werden folgende Schritte ausgeführt:
 - (a) Ableitungen beider Bilder in x- und y - Richtung werden berechnet (z. B. mit einem Sobelfilter).
 - (b) Die Koeffizienten der drei Gleichungen aus 3.10 werden berechnet.
 - (c) a , b und Θ werden bestimmt, indem das aus den drei Gleichungen entstandene Gleichungssystem gelöst wird (z. B. durch das Invertieren der Koeffizientenmatrix).
 - (d) Nun werden folgende Schritte wiederholt, solange ein bestimmter Schwellwert nicht unterschritten wird:
 - i. Das Referenzbild wird um die berechneten Werte a und b in x- und y - Richtung verschoben und danach um Θ rotiert
 - ii. Die Koeffizientenmatrix wird auf Basis des aktualisierten Referenzbildes neu berechnet.
 - iii. a , b und Θ werden neu berechnet und auf die vorherigen Werte aufaddiert.
 - (e) Die Ergebnisse werden zur nächsten Auflösungsstufe propagiert, indem das Sequenzbild um $2a$ und $2b$ verschoben und um Θ rotiert wird. Wenn Die höchste Auflösungsstufe erreicht ist, werden a , b und Θ als Ergebnis ausgegeben und der Algorithmus terminiert.

3.2 Merkmalsbasierte Methode

3.2.1 Überblick

Die Grundlage der merkmalsbasierten Methoden bildet eine Menge M , die zueinander gehörige Merkmalspaare beschreibt:

$$M(I, I') = \{f, f' | f \in F, f' \in F'\} \quad (3.10)$$

wobei F und F' die Mengen aller extrahierten Merkmale aus beiden Bildern und f, f' deren jeweilige Elemente darstellen. Diese Merkmale werden anfangs in den beiden Bildern detektiert und dann mit Hilfe von Korrelationsmaßen zu Merkmalspaaren zusammengefasst.

Zur Extrahierung von Merkmalen eignen sich Algorithmen wie der Canny-Operator, Tomasi-Kanade Punkt-Detektion oder der Moravec-Operator. Das meistbenutzte Verfahren jedoch ist der Harris Corner Detector. Punkte als Merkmale zu benutzen hat gegenüber komplexeren Formen (wie z. B. Kanten oder Linien) vor allem den Vorteil, dass sie einfach zu detektieren sind und zur späteren Korrelationsbestimmung weniger komplexe Abstandsmaße benötigen.

Geeignete Ähnlichkeitsmaße zur Korrelationsbestimmung sind z. B. die Summe quadrierter/absoluter Differenzen (SSD, SAD) oder auch die normalisierte Kreuzkorrelation (NCC).

Unabhängig von den erkannten Merkmalen kommt es bei der Bestimmung solcher Merkmalspaare oft zu fehlerhaften Zuordnungen und damit zu ungenauen bis falschen Homographien. Da dies auch bei den besten Korrelationsmaßen nie ganz zu vermeiden ist, besteht ein weiterer Schritt in der Parameterbestimmung darin, die falschen Paare (Outlier) mit Hilfe von statistischen Ansätzen auszusortieren. Ein solcher robuster Parameterschätzer ist der RANdom SAmple Concensus (RANSAC), der iterativ zufällige Mengen aus $M(I_1, I_2)$ auswählt, daraus eine Homographie erstellt und mittels geeigneter Fehlermaße dessen Güte feststellt. Die Teilmenge mit der besten Unterstützung, also mit den meisten Merkmalpaaren, die innerhalb eines bestimmten Schwellwertes liegen (Inlier), wird anschließend zur Bestimmung der resultierenden Homographie benutzt.

Um eine noch bessere Schätzung zu erhalten, kann man mit allen berechneten Inliern die Homographie mit Maximum-Likelihood Verfahren wie z. B. dem Levenberg-Marquardt Algorithmus, verfeinern. Auch die Berechnung einer besseren Homographie mit der Diskreten Linearen Transformation (DLT) ist möglich, wenn dazu alle Inlier in die Berechnung mit einfließen.

In [HZ03] wird ein Algorithmus zur Homographieschätzung mit RANSAC im Detail vorgestellt und beschrieben. Zusammengefasst ist der Ablauf des Algorithmus wie folgt:

1. **Merkmale:** In jedem Bild werden Merkmale berechnet (z. B. mit dem Harris Corner Detector).
2. **Matching:** Eine Menge von Merkmalspaaren werden anhand von Ähnlichkeits- oder Distanzmaßen in ihrer Intensitätsnachbarschaft berechnet (z. B. mit der NCC).
3. **Robuste Schätzung mit RANSAC:** Wiederhole für N Proben (N wird adaptiv berechnet)
 - Wähle eine zufällige Probe von 4 Merkmalspaaren und berechne die Homographie H . (z. B. mit DLT)
 - Berechne die Distanz d_{\perp} für jede mutmaßliche Punktkorrespondenz. (z. B. mit symmetric transfer error)
 - Berechne die Anzahl der Inlier für die gilt: $d_{\perp} < t = \sqrt{5.99}\sigma^2$ Pixel.

Wähle H mit der größten Anzahl an Inliern.

4. **Optimierung:** H wird erneut mit allen Merkmalspaaren, die als Inlier klassifiziert wurden, geschätzt, indem die ML Kostenfunktion minimiert wird (z. B. mit dem Levenberg-Marquardt Algorithmus oder DLT)
5. **Guided Matching:** Unter Verwendung von H , Berechne weitere Merkmalspaare.

3.2.2 Matching von Merkmalen

Es stellt sich nun die Frage, wie die extrahierten Merkmale aus den beiden Bildern zu Merkmalspaaren zusammengefügt werden können. Es muss ein Verfahren angewandt werden, das die Zusammengehörigkeit zweier Merkmale beschreibt. Als Merkmalspaar wird dann das Paar deklariert, das den höchsten Zusammengehörigkeitswert besitzt.

Summe quadrierter/absoluter Differenzen (SSD, SAD)

Die Summe quadrierter/absoluter Differenzen sind die zwei einfachsten Arten, Merkmale miteinander zu vergleichen. Nehmen wir an, es wurden die Merkmale f in Bild I an der Position x, y , und die Merkmale f' in Bild I' an der Position $x + i, y + j$ gefunden.

Sei also:

$$\Delta_{x,y} = I_{x,y} - I'_{x+i,y+j} \quad (3.11)$$

das Differenzbild von I und I' , so ist

$$d_{SAD} = \sum_{y=y-\frac{m}{2}}^{\frac{m}{2}} \sum_{x=x-\frac{n}{2}}^{\frac{n}{2}} |\Delta_{x,y}| = \sum_{y=y-\frac{m}{2}}^{\frac{m}{2}} \sum_{x=x-\frac{n}{2}}^{\frac{n}{2}} |I_{x,y} - I'_{x+i,y+j}| \quad (3.12)$$

die Summe absoluter Differenzen zweier Bilder in einem Fenster der Größe $m \times n$.

Analog ist:

$$d_{SSD} = \sum_{y=y-\frac{m}{2}}^{\frac{m}{2}} \sum_{x=x-\frac{n}{2}}^{\frac{n}{2}} \Delta_{x,y}^2 = \sum_{y=y-\frac{m}{2}}^{\frac{m}{2}} \sum_{x=x-\frac{n}{2}}^{\frac{n}{2}} (I_{x,y} - I'_{x+i,y+j})^2 \quad (3.13)$$

die Summe quadrierter Differenzen zweier Bilder in einem Fenster der Größe $m \times n$.

Diese beiden Ähnlichkeitsmaße haben den Vorteil, dass die sehr einfach und effizient zu implementieren sind. Negativ ist jedoch anzumerken, dass wie zum Beispiel durch Bildrauschen entstandene Ausreißer genauso stark in das Ergebnis eingehen, wie „echte“ Fehler, die auf unterschiedliche Merkmale zurückzuführen sind. Außerdem ist dieses Maß nicht invariant gegenüber Kontrast- und Helligkeitsunterschieden der Art: $I' = \alpha I + \beta$. Die Bilder müssen also vorher Farb- und Helligkeitsnormiert sein, damit sinnvolle Ergebnisse entstehen können.

Normalisierte Kreuzkorrelation (NCC)

Die normalisierte Kreuzkorrelation beruht auf dem aus der Wahrscheinlichkeitstheorie stammenden Begriff der Kovarianz:

$$Cov(X, Y) = E[X - E(X), Y - E(Y)] \quad (3.14)$$

wobei X und Y Zufallsvariablen sind. Sollen nun die Ergebnisse auf Werte zwischen -1 und 1 normiert werden, so werden sie durch das Produkt der beiden Standardabweichungen geteilt:

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[X - E(X), Y - E(Y)]}{\sigma_X \sigma_Y} \quad (3.15)$$

Die Korrelation zweier Zufallsvariablen beträgt 1, wenn ein linearer Zusammenhang besteht. 0 wird das Ergebnis, wenn die Variablen unkorreliert sind und -1, wenn ein negativer linearer Zusammenhang besteht. Auch non-Integer Werte zwischen -1 und 1 sind möglich.

Werden nun die Bildintensitäten als Proben von Zufallsvariablen angesehen, so ist:

$$NCC = \frac{1}{\sigma \sigma' N^2} \sum_{(x,y) \in W} (I_{x,y} - \mu)(I'_{x+i,y+j} - \mu') \quad (3.16)$$

wobei die Fensterfunktion W die $N \times N$ Nachbarschaft beschreibt, die zur Berechnung des NCC mit einbezogen wird. Um einen besseren NCC-Wert zu erzielen, hat sich hier die Verwendung einer Gaußmaske als vorteilhaft erwiesen.

Außerdem ist:

$$\mu = \frac{1}{N^2} \sum_{(x,y) \in W} I_{x,y} \quad (3.17)$$

$$\sigma = \frac{1}{N^2} \sqrt{\sum_{(x,y) \in W} (I_{x,y} - \mu)^2} \quad (3.18)$$

Analog sind die Berechnungen für μ' und σ' . Dieses Maß ist im Gegensatz zu den beiden vorherigen invariant gegenüber Kontrast- und Helligkeitsunterschieden.

Um die Rechenzeit der Korrelationen zwischen den Merkmalen zu verringern, sollte der Suchraum, in dem nach Merkmalen im zweiten Bild gesucht wird, nicht auf das ganze Bild ausgebreitet werden, sondern auf eine angemessene Größe beschränkt werden.

3.2.3 DLT - Direkte Lineare Transformation

Ein einfacher linearer Algorithmus zum Bestimmen einer Homographie H ist die *Direkte Lineare Transformation* (DLT). Da eine 2D-Transformation 8, und ein 2D-Merkmalpaar 2 Freiheitsgrade besitzt, reichen vier 2D-Merkmalpaare, $x_i \leftrightarrow x'_i$, aus, um eine Homographie der Form $x'_i = Hx_i$ zu berechnen. Hierbei muss beachtet werden, dass es sich bei x_i und x'_i um homogene Vektoren mit drei Einträgen handelt. Eine andere Schreibweise für H führt zu einer linearen Lösungsfindung: $x'_i \times Hx_i = 0$.

In dieser Herleitung werden folgende Abkürzungen und Annahmen für Vektoren und Matrizen benutzt:

- die Reihe j der Matrix H kann als h^{jT} geschrieben werden
- der Vektor $x'_i = (x'_i, y'_i, w'_i)$

- \mathbf{h} ist ein 9-dimensionaler Vektor, der die Einträge der Matrix \mathbf{H} wie folgt enthält:

$$\mathbf{h} = \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \quad (3.19)$$

Sei $\mathbf{H}\mathbf{x}_i$ geschrieben als:

$$\mathbf{H}\mathbf{x}_i = \begin{pmatrix} \mathbf{h}^{1\top}\mathbf{x}_i \\ \mathbf{h}^{2\top}\mathbf{x}_i \\ \mathbf{h}^{3\top}\mathbf{x}_i \end{pmatrix} \quad (3.20)$$

So ist das Kreuzprodukt von \mathbf{x}'_i und $\mathbf{H}\mathbf{x}_i$:

$$\mathbf{x}'_i \times \mathbf{H}\mathbf{x}_i = \begin{pmatrix} y'_i \mathbf{h}^{3\top}\mathbf{x}_i - w'_i \mathbf{h}^{2\top}\mathbf{x}_i \\ w'_i \mathbf{h}^{1\top}\mathbf{x}_i - x'_i \mathbf{h}^{3\top}\mathbf{x}_i \\ x'_i \mathbf{h}^{2\top}\mathbf{x}_i - y'_i \mathbf{h}^{1\top}\mathbf{x}_i \end{pmatrix} \quad (3.21)$$

Da $\mathbf{h}^{j\top}\mathbf{x}_i = \mathbf{x}_i^\top \mathbf{h}^j$, ergeben sich daraus drei Gleichungen für \mathbf{H} , die als

$$\begin{pmatrix} \mathbf{0}^\top & -w'_i \mathbf{x}_i^\top & y'_i \mathbf{x}_i^\top \\ w'_i \mathbf{x}_i^\top & \mathbf{0}^\top & -x'_i \mathbf{x}_i^\top \\ -y'_i \mathbf{x}_i^\top & x'_i \mathbf{x}_i^\top & \mathbf{0}^\top \end{pmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0}. \quad (3.22)$$

geschrieben werden können und die Form $\mathbf{A}_i \mathbf{h} = \mathbf{0}$ besitzen, wobei \mathbf{A}_i eine 3×9 Matrix vom Rang 8 ist. Der Rang 8 entsteht dadurch, dass dieses Gleichungssystem zwar aus drei Gleichungen besteht, jedoch nur die ersten zwei linear unabhängig voneinander sind. Deshalb kann die dritte Gleichung zum Lösen der Homographie weggelassen werden. Es ergibt sich also:

$$\begin{pmatrix} \mathbf{0}^T & -w'_i \mathbf{x}_i^T & y'_i \mathbf{x}_i^T \\ w'_i \mathbf{x}_i^T & \mathbf{0}^T & -x'_i \mathbf{x}_i^T \end{pmatrix} \begin{pmatrix} h^1 \\ h^2 \\ h^3 \end{pmatrix} = 0. \quad (3.23)$$

Zur besseren Übersicht geschrieben als $\mathbf{A}_i \mathbf{h} = 0$, wobei \mathbf{A}_i nun eine 2×9 Matrix ist.

Für vier Merkmalspaare hat \mathbf{A}_i die Dimension 8×9 und somit den Rang 8 und besitzt damit einen eindimensionalen Nullraum, der auch gleichzeitig das Ergebnis dieser Gleichung ist. Dieser kann gefunden werden, indem eine Singulärwertzerlegung (SVD) auf \mathbf{A}_i angewandt wird. Der Einheitsvektor mit dem kleinsten Eigenwert ist das Ergebnis für \mathbf{h} .

Da das Ergebnis des DLT Algorithmus stark abhängig von dem Koordinatensystem ist, in dem sich die Punkte befinden, wird in [HZ03] empfohlen, vor der Anwendung des DLT eine Normalisierung der Daten durchzuführen. Diese Normalisierung besteht aus einer Translation und einer Skalierung, so dass der Mittelpunkt der Punktmenge im Ursprung (0,0) liegt, und ihre durchschnittliche Entfernung bei $\sqrt{2}$.

Liegt das Zentrum der Punktepaare vor der Normalisierung bei z_x, z_y , und die durchschnittliche Entfernung s , so lautet die Normalisierungsmatrix \mathbf{T} für \mathbf{x} :

$$\mathbf{T} = \begin{pmatrix} \frac{1}{s} & 0 & -\frac{z_x}{s} \\ 0 & \frac{1}{s} & -\frac{z_y}{s} \\ 0 & 0 & 1 \end{pmatrix} \quad (3.24)$$

Analog wird eine Transformationsmatrix \mathbf{T}' für \mathbf{x}' berechnet, die zur Normalisierung der Punkte des zweiten Bildes führt. Auf den neu erhaltenen Punktepaaren $\tilde{\mathbf{x}}_i$ und $\tilde{\mathbf{x}}'_i$ wird nun mit dem DLT Algorithmus, wie oben beschrieben, eine Homographie $\tilde{\mathbf{H}}$ berechnet. Anschließend wird diese Homographie noch denormalisiert, indem $\mathbf{H} = \mathbf{H}'^{-1} \tilde{\mathbf{H}} \mathbf{T}$ berechnet wird.

3.2.4 RANSAC - RANdom SAMple Concensus

Da die bisher gefundenen Merkmalspaare nicht fehlerfrei genug sind, um eine genaue Homographie zu bestimmen, wird eine Schätzung der Inlier durch RANSAC benötigt. Outlier können eine Homographie extrem verfälschen und sollten daher auf alle Fälle so gut es geht aus dieser Berechnung eliminiert werden. Es wird im Zusammenhang mit dem RANSAC Algorithmus von einer *robusten Schätzung* gesprochen, da er sehr gut auch mit einer großen Menge von Outliern umgehen kann.

Die Schätzung verläuft so, dass ausgehend von vier zufällig gewählten Merkmalpaaren eine Homographie berechnet wird, die dann mit Hilfe eines Fehlermaßes auf ihre Güte untersucht wird.

Probenauswahl

Um eine gute Grundlage zur Inlierklassifizierung zu haben, sollten die vier zufällig ausgewählten Merkmalsproben auf ihre Tauglichkeit untersucht werden.

Schlechte Proben sind:

- Punkte, die linear voneinander abhängig sind, da somit keine zuverlässige Homographie durch DLT berechnet werden kann
- Punkte, die zu nah beieinander liegen, da in diesem Fall die berechnete Homographie zu ungenau werden kann, vor allem beim Berechnen des Winkels

Zusätzlich zur Überprüfung auf Kollinearität, sollte die Zufallsprobe also unbedingt daraufhin untersucht werden, ob sich die Punkte in einem bestimmten Radius zueinander befinden. In [OLL04] wird z. B. vorgeschlagen, das Bild in vier Regionen zu unterteilen, sodass jede der vier Punktkorrespondenzen aus einer anderen Region stammt.

Inlierklassifizierung

Die Genauigkeit einer Homographie kann mit verschiedenen Fehlermaßen, wie z. B. dem *Gold Standard Error*, dem *Symmetric Transfer Error*, dem *Projection Error* oder dem *Sampson Error* überprüft werden. Der *Symmetric Transfer Error* ist hierbei die einfachste Methode, die trotzdem noch genau genug ist. Dieses Fehlermaß gibt vor allem gute Ergebnisse, wenn angenommen werden kann, dass Fehler in beiden Bildern existieren. Es wird also die Vorwärts- (\mathbf{H}) und Rückwärtstransformation (\mathbf{H}^{-1}) mit einbezogen und von diesen beiden jeweils der geometrische Fehler $d(\mathbf{x}, \mathbf{H}^{-1}\mathbf{x}')$ und $d(\mathbf{x}', \mathbf{H}\mathbf{x})$ addiert:

$$d_{transfer}^2 = d(\mathbf{x}, \mathbf{H}^{-1}\mathbf{x}')^2 + d(\mathbf{x}', \mathbf{H}\mathbf{x})^2 \quad (3.25)$$

Nun kann jeder Punkt \mathbf{x}_i mit der binären Funktion γ als Inlier oder Outlier klassifiziert werden:

$$\gamma(\mathbf{x}_i) = \begin{cases} 0, & \text{falls } \{d_{transfer}^2\}_i > (\sqrt{5.99} * \sigma)^2 \quad (\text{Outlier}) \\ 1, & \text{sonst} \quad (\text{Inlier}) \end{cases} \quad (3.26)$$

Der Schwellwert $t^2 = 5.99 * \sigma^2$ basiert auf der Wahrscheinlichkeit $\alpha = 0.95$, dass ein Punkt ein Inlier ist und zusätzlich dem angenommenen Modell (in diesem Fall eine Homographie) entspricht [HZ03]. σ ist die Standardabweichung der Summe der symmetrischen Fehler von allen Merkmalspaaren:

$$\sigma = \frac{\sqrt{\sum_{1 \leq i \leq N} (\{d_{transfer}^2\}_i - \mu)^2}}{N - 1} \quad (3.27)$$

mit der Anzahl der Merkmalspaare N . Es ist auch möglich einen konstanten Schätzwert für σ anzugeben, um die Berechnungen zu beschleunigen.

Anzahl der Proben und Terminierung des Algorithmus

Um die Rechenzeit so gering wie möglich zu halten, ist es sinnvoll, nicht jede mögliche Probe auszuprobieren. Die Anzahl der Proben N sollte so gewählt werden, dass mit der Wahrscheinlichkeit p wenigstens eine der Proben der Größe s keine Outlier besitzt. Für diese Wahrscheinlichkeit wird normalerweise $p = 0.99$ gewählt. Sei nun $\epsilon = 1 - w$ die Wahrscheinlichkeit, dass der gewählte Punkt ein Outlier ist, so werden mindestens N Proben der Größe s gebraucht, für die gilt:

$$N = \frac{\log(1 - p)}{\log(1 - (1 - \epsilon)^s)} \quad (3.28)$$

Da in der Praxis die Outlierwahrscheinlichkeit ϵ oft nicht bekannt ist, sollte die Anzahl der zu benutzenden Proben adaptiv bestimmt werden. Initialisiert wird ϵ mit einer Schätzung des schlimmsten Falles (also der Fall in dem die meisten Proben Outlier sind), sinnvollerweise z. B. 1.0. Damit würde dann der anfängliche Wert von N unendlich groß. Wird eine Probe entdeckt, die weniger Outlier besitzt als die vorher gewählte, wird ϵ auf diesen Wert gesetzt und N damit neu berechnet, bis die Anzahl der bereits untersuchten Proben diese Zahl übersteigt:

- Initialisierung: $\epsilon = 1.0$, $N = \infty$ und $\text{SampleCount} = 0$
- solange $N > \text{SampleCount}$:
 - Wähle zufällige Probe und bestimme Anzahl der Inlier
 - Setze $\epsilon = 1 - \frac{\text{Anzahl der Inlier}}{\text{Gesamtanzahl Punkte}}$, falls ϵ kleiner als im vorherigen Schritt
 - Berechne N mit der neuen Inlierwahrscheinlichkeit ϵ und $p = 0.99$
 - Erhöhe SampleCount um 1
- Terminiere Algorithmus

Kapitel 4

Experimente und Ergebnisse

Um die Robustheit beider Verfahren zu testen, wurden verschiedene synthetische Testreihen erstellt. Auf diese Weise wird geprüft, wie sich eine Rotation, Translation und auch verschiedene Helligkeiten und Farbunterschiede auf die Ergebnisse auswirken. Anschließend wird untersucht, wie die beiden Methoden mit einer real aufgenommenen Bildserie umgehen und welche Schwierigkeiten es dabei gibt. Zuletzt wird noch eine Bildserie mit Anwendungsbeispiel vorgestellt. Mit Hilfe der subpixel genauen Registrierung wird so ein Superresolution Bild erstellt.

Als Qualitätsmerkmal wurde das Differenzbild zwischen dem Referenzbild und dem registrierten Sequenzbild erstellt und daraus der RMS-Fehler ¹ (quadratischer Mittelwert) berechnet. Daran kann gut erkannt werden, wie stark die errechnete Registrierung von der tatsächlich vorliegenden Deformation abweicht und wie sich die beiden Verfahren qualitativ unterscheiden.

4.1 Untersuchung des Rotationsverhaltens

Für die Untersuchung des Rotationsverhaltens wurde das Ausgangsbild in 5° - Schritten um den Bildmittelpunkt nach rechts gedreht. Getestet wurden Rotationen zwischen 5° und

¹Root Mean Square: $\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$

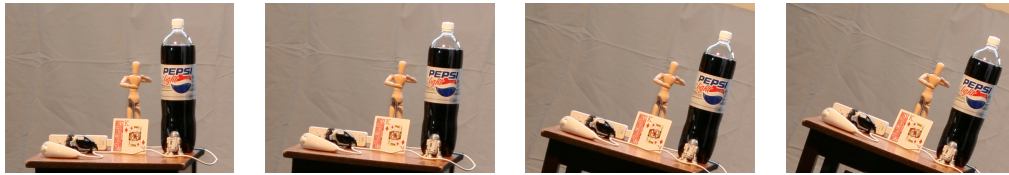


Bild 4.1: 1. Spalte: Referenzbild, 2. Spalte: um 5° , 3. Spalte: um 15° gedreht, 4. Spalte: um 20° gedreht

30° . In 4.1 sind Beispielbilder mit Drehungen um 5° , 15° und 20° abgebildet.

Die Grafik 4.2 zeigt den Vergleich der beiden Verfahren anhand ihrer RMS-Fehler. Beide Methoden erzielen bis 15° recht gute und brauchbare Ergebnisse, wobei die merkmalsbasierte insgesamt etwas besser abschneidet. Ab einer Rotation von 20° jedoch vermerken beide Verfahren Einbußen. Während die merkmalsbasierte nur ein wenig schlechter wird, springt der Fehler der direkten Methode auf das Achtfache und beschreibt somit eine sehr schlechte Registrierung.

Schließlich wird bei einer Rotation von über 20° auch die merkmalsbasierte Methode immer schlechter. Auch wenn diese Registrierungen vom Fehler her viel besser ausfallen, sind die Ergebnisse für weitere Anwendungen nicht mehr zu gebrauchen.

Es ist also zu sehen, dass die merkmalsbasierte Methode viel besser mit stark rotierten Bildern umgehen kann, als die direkte Methode. Tests mit einer Bildserie anderen Inhalts haben ähnliche Ergebnisse erzielt, dort konnte die direkte Methode bis 10° und die merkmalsbasierte bis 25° gute Ergebnisse erzielen. Zusammengefasst haben die Versuche jedoch gezeigt, dass die direkte Methode ab ca. 10 - 15° und die merkmalsbasierte Methode ab 20° an ihre Grenzen stoßen.

4.2 Untersuchung des Translationsverhaltens

Um das Translationsverhalten zu untersuchen, wurde das Referenzbild der Größe 600×444 in 5 Pixel Schritten von 5 bis 100 Pixeln in y - Richtung verschoben. Zusätzlich wurde noch darauf geachtet, dass sich der Inhalt des Bildes nicht zu stark ändert (die Auswir-

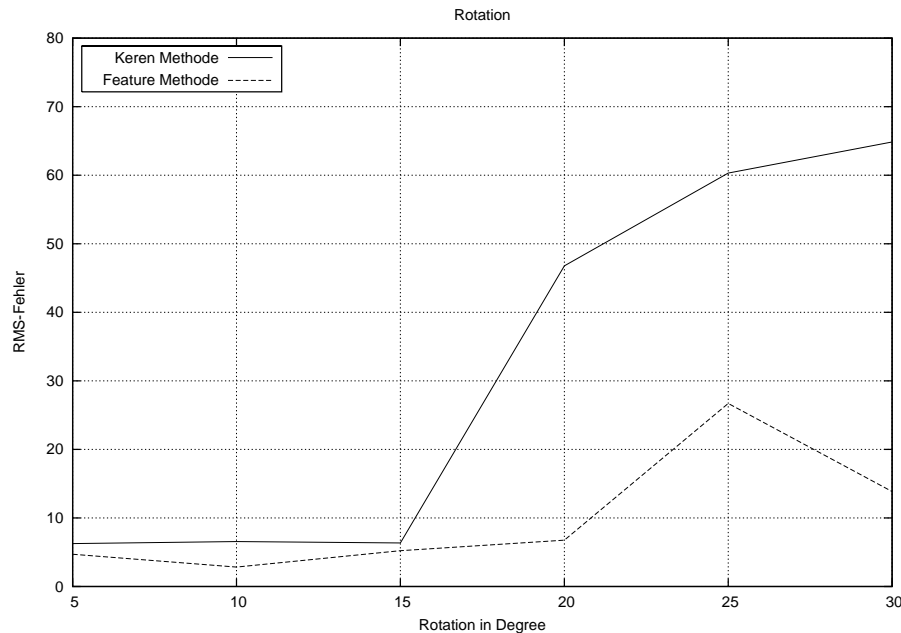


Bild 4.2: Vergleich des Rotationsverhaltens beider Verfahren

kung einer starken Inhaltsveränderung kann in Abbildung 4.8 betrachtet werden).

Wie in Grafik 4.4 zu sehen ist, können beide Methoden gut mit einer Translation bis 60 Pixeln umgehen. Danach ist jedoch bei der direkten Methode schon eine sehr starke und sprunghaft auftretende Fehlregistrierung zu erkennen, wohingegen die der merkmalsbasierten noch keine Fehler aufweist. Des Weiteren ist in der Grafik zu erkennen, dass die direkte Methode fortan einen logarithmischen Verlauf annimmt, die merkmalsbasierte jedoch auch bei einer Verschiebung von 100 Pixeln noch immer keine Fehlregistrierung produziert.

Da der Bildinhalt und die Verschiebung immer abhängig von der Größe des Bildes ist, können diese Zahlen nicht als allgemeingültig gesehen werden, sondern müssen in Verhältnis mit den Dimensionen des Bildes gebracht werden. Bei einer Höhe von 600 Pixeln ist die Grenze der direkten Methode also eine Translation von 10% des Bildes. Weitere Versuche mit der merkmalsbasierten Methode haben gezeigt, dass die Ergebnisse erst bei einer starken Veränderung des Bildinhalts schlechter werden. Bleiben also die Objekte mit

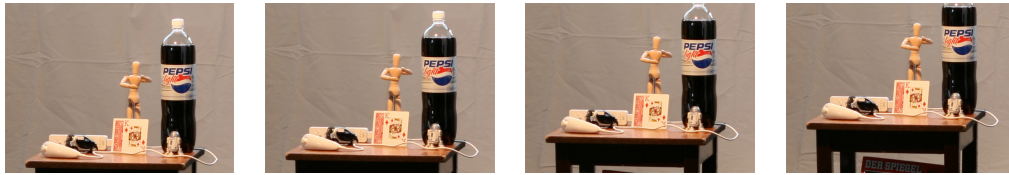


Bild 4.3: 1. Spalte: Referenzbild, 2. Spalte: um 20px, 3. Spalte: um 65px, 4. Spalte: um 100px in y-Richtung erschoben

den interessantesten Merkmalen im Bild, so ist die Registrierung sehr gut, auch wenn die Merkmalspunkte weit verschoben sind. Hierfür ist also keine prozentuale Angabe möglich, weil sogar trotz einer Verschiebung von 50% noch gute Ergebnisse erzielt werden können, wenn der „interessante Inhalt“ im Bild zu sehen ist. Ist dies nicht der Fall, kann auch die merkmalsbasierte Methode schon bei einer 10% - igen Verschiebung Fehlregistrierungen erkennen lassen.

4.3 Untersuchung der Robustheit gegenüber Farb- und Helligkeitsschwankungen

In dieser Versuchsreihe wird untersucht, wie sich die beiden Verfahren gegenüber unterschiedlichen Helligkeiten und Farbstichen in den Bildern der Serie verhalten. Dazu wurde von einem Referenzbild eine Belichtungsreihe erstellt, in der die Helligkeit in jeweils fünf Schritten erhöht und verringert wurde. Außerdem wurden jeweils sieben unterschiedlich starke rot- und blaustichige Bilder des Originals erstellt. Ansonsten gab es im Vergleich zum Referenzbild keinerlei Modifikationen, es wurde also keine Translation und keine Rotation vorgenommen. Beispielbilder aus diesen beiden Reihen sind in den Abbildungen 4.5 und 4.6 zu sehen.

Grafik 4.7 links zeigt den Unterschied zwischen den beiden Verfahren in Hinblick auf Helligkeitsunterschiede. Erwartungsgemäß werden mit der direkten Methode immer schlechter werdende Ergebnisse erzielt, je dunkler oder heller das Sequenzbild wird, da sie auf den Intensitäten des Bildes arbeitet und diese als Ähnlichkeitsmerkmal benutzt. Die merkmals-

4.3. UNTERSUCHUNG DER ROBUSTHEIT GEGENÜBER FARB- UND HELLIGKEITSSCHWANKUNGEN

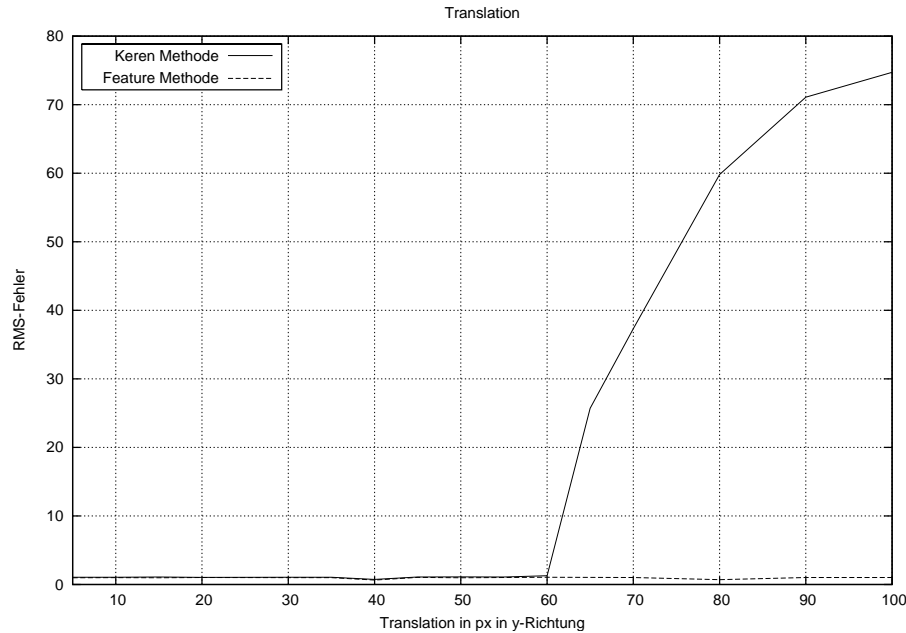


Bild 4.4: Vergleich des Translationsverhaltens beider Verfahren

basierte Methode hat keinerlei Probleme mit unterschiedlichen Belichtungen, der RMS-Fehler bleibt hier konstant unter 1.

Bei der Serie mit den Farbstichen ist in der Grafik 4.7 rechts zu sehen, dass die merkmalsbasierte Methode auch hier keinen relevanten Registrierungsfehler produziert, wohingegen die direkte Methode sofort eine starke Fehlregistrierung erzeugt. Zu bemerken ist außerdem, dass der Blaustich einen größeren Fehler zur Folge hat als der Rotstich. Auch hier liegt die Erklärung wieder in der Art und Weise der Ähnlichkeitsschätzung. Der Blaukanal hat einen stärkeren Einfluß auf die Intensität des Bildes und somit auch auf das Ähnlich-



Bild 4.5: Helligkeitsverlauf von dunkel nach hell, Referenzbild in der Mitte

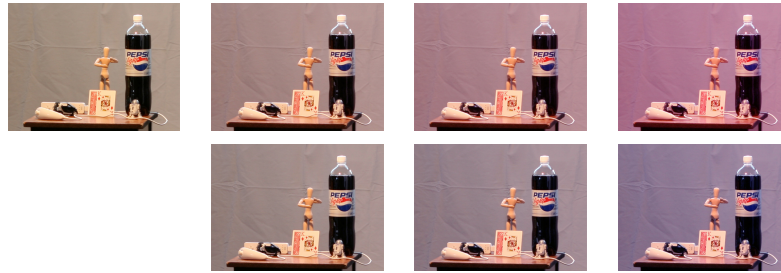


Bild 4.6: Oben links: Referenzbild, oben Spalte 2-3 Farbstichverlauf rot, unten Farbstichverlauf blau

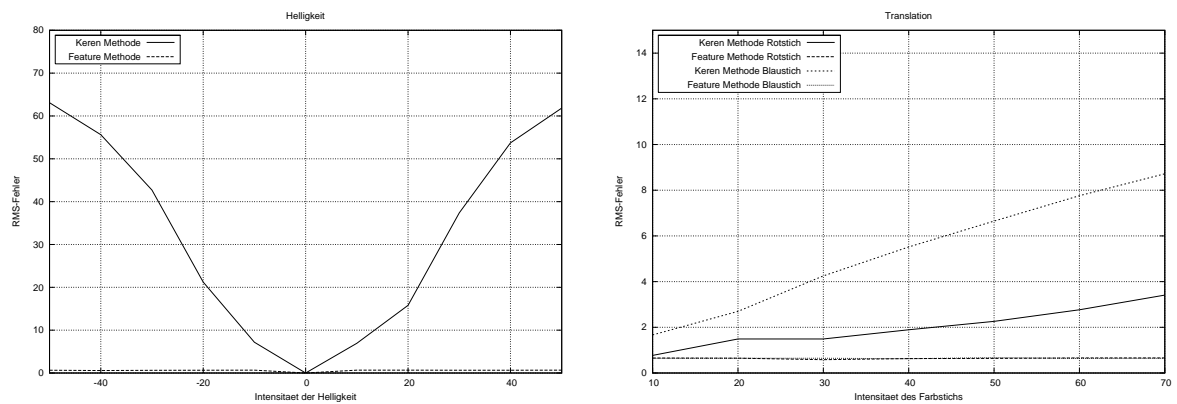


Bild 4.7: Vergleich der Robustheit gegenüber Helligkeitsschwankungen (links) und Farbstichen (rechts)

keitsmaß der direkten Methode.

4.4 Versuche mit realer Sequenz

Der letzte Versuch wurde mit einer real aufgenommenen Bildserie durchgeführt. Dabei wurde bei der ersten Serie weder auf die Art der Deformation, noch auf sich verändernde Bildinhalte und Helligkeitsschwankungen geachtet. In Abbildung 4.8 sind drei der aufgenommenen Bilder und ihre jeweiligen Differenzbilder abgebildet. Zu sehen ist, dass die



Bild 4.8: 1. Spalte: Referenzbild, in den Spalten 2 - 4 in der zweiten Zeile jeweils das Differenzbild der merkmalsbasierten Methode und unten das der direkten Methode, 2. Spalte: Translation in x- und y - Richtung, 3. Spalte: geringe Rotation um ca 5° und leichte Translation, 4. Spalte: starke Translation in y - Richtung und zusätzlich starke Änderung des Bildinhaltes

merkmalsbasierte Methode recht gut mit Translationen und Rotationen umgehen kann, jedoch eine starke Veränderung des Bildinhaltes ein sehr schlechtes Ergebnis zur Folge hat. Ähnliches ist bei der direkten Methode zu beobachten, die jedoch zusätzlich bei einer mittelstarken Translation in x- und y - Richtung eine falsche Registrierung erzielt.

In den Differenzbildern beider Verfahren sind - auch bei den guten Ergebnissen - immer kleine weiße Ränder zu sehen. Der Grund hierfür ist die Bewegung innerhalb der Bilder, die durch den Wind und die sich dadurch bewegenden Bäume entstanden sind.

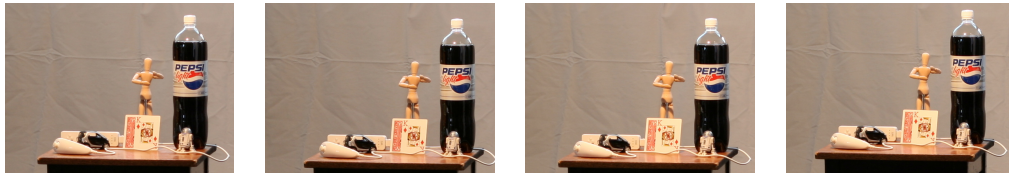


Bild 4.9: Links: Referenzbild, in den Spalten 2 - 4 verschiedene Bilder der Bildserie



Bild 4.10: Links Superresolution mit der direkten, rechts mit der merkmalsbasierten Methode

4.4.1 Registrierung mit Anwendungsbeispiel

In diesem Versuch soll getestet werden, wie gut die beiden Verfahren geeignet sind, um ein Superresolution Bild zu erzeugen. Dazu wurde eine Bildserie mit leichtem Versatz (hauptsächlich bestehend aus Translationen in x - und y -Richtung) aufgenommen, die viele verschiedene Objekte enthält. Außerdem wurde konkret darauf geachtet, dass keine Helligkeits- und Farbschwankungen enthalten sind, damit beide Registrierungsmethoden einwandfrei arbeiten können. Einige Beispielbilder aus dieser Serie sind in Abbildung 4.9 zu sehen.

In Grafik 4.11 sind die RMS-Fehler für jedes Bild abgetragen. Obwohl die merkmalsbasierte Methode in den Tests mit den synthetischen Bildern durchweg besser abgeschnitten hat, ist hier eine Dominanz der direkten Methode zu verzeichnen. Ein möglicher Grund hierfür ist die genaue Beachtung der Stärken und Schwächen der Methoden bei der Bildaufnahme, von der die direkte Methode mehr davon profitieren kann.

Das Superresolution Ergebnis dieses Tests ist in Abbildung 4.10 zu sehen. Aus der Bildserie wurde eine Region of Interest (ROI) ausgeschnitten und mittels des MAP Verfahrens mit

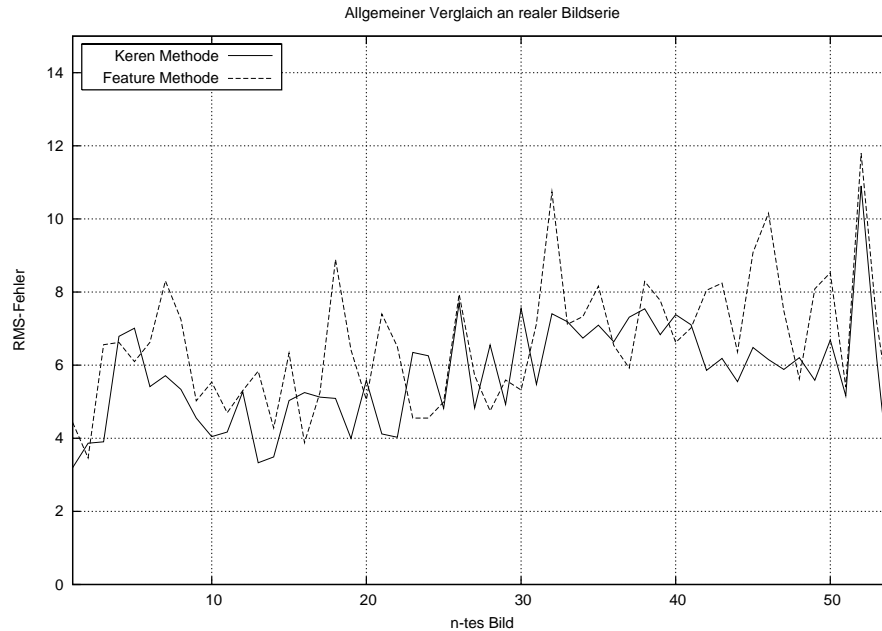


Bild 4.11: Vergleich der Fehler beider Verfahren beim Registrieren einer realen Bildserie

Gaussian-Markov Random Field Prior (GMRF) zweifach vergrößert. Es ist zu erkennen, dass - aufgrund der recht scharfen Ergebnisse - die direkte, sowie auch die merkmalsbasierte Methode eine subpixel genaue Registrierung geliefert haben, die zur weiteren Verwendung durch Superresolution Algorithmen geeignet sind.

Kapitel 5

Zusammenfassung

In dieser Studienarbeit wurde jeweils ein Verfahren aus den zwei Bereichen der Bildregistrierung implementiert und beschrieben. Eine *direkte*- und eine *merkmalsbasierte Methode* wurden verglichen und auf ihre Grenzen hin überprüft. Die implementierten Verfahren funktionieren gut und registrieren beide verschiedenste Bildserien sub-pixelgenau.

In Kapitel 2 wurde ein Überblick über die existierenden Verfahren zur Registrierung von Bildern und deren Anwendungsgebiete gegeben. Des Weiteren beschreibt Kapitel 3 detailliert die von mir gewählten Verfahren und zeigt auch mögliche Variationen in der Implementierung auf.

Bei der direkten Methode ist vor allem die Wahl des Transformationsmodells ausschlaggebend, um ein geeignetes Gleichungssystem zu erstellen, aus dem später die Registrierung berechnet wird. Auch das Einbetten der Methode in eine Gaußpyramidenstruktur hat sich als wichtig herausgestellt, da somit die Genauigkeit gesteigert werden kann.

Da die merkmalsbasierte Methode aus verschiedenen Komponenten aufgebaut ist (siehe 3.2.1) kann jeder einzelne Schritt durch unterschiedliche Verfahren ausgetauscht werden. Die Detektion der Merkmale kann zum Beispiel auch durch Tomasi-Kanade, SIFT oder Moravec ersetzt werden. Auch das Matching der Merkmale kann von anderen Korrelationsmaßen als dem NCC durchgeführt werden, wie zum Beispiel die Summe quadrierter/absoluter Differenzen (SSD, SAD). Weitere Ähnlichkeitsmaße können in [Kub06] nachgelesen werden. Weiter können auch andere robuste Schätzer (z. B. MLESAC) und

wiederum darin andere Fehlermaße (z. B. den Sampson Error) genutzt werden. Letztlich kann die erneute Schätzung der Homographie statt mit dem DLT auch mit dem Levenberg-Marquardt Algorithmus verbessert werden. Auch ist es möglich, durch eine Veränderung der Genauigkeit des Symmetric Transfer Error bei der Inlierbestimmung die Ergebnisse zu beeinflussen.

In der direkten Methode kann die Genauigkeit der Ergebnisse zum einen durch den gewählten Schwellwert und zum anderen durch die Anzahl der Pyramidenstufen beeinflusst werden. Bei der merkmalsbasierten Methode wiederum können unterschiedlich viele Merkmale benutzt werden, die einen unterschiedlich hohen Schwellwert besitzen können. In dieser Studienarbeit konnten nicht alle Variationen der zwei Verfahren behandelt und implementiert werden. Es wäre interessant zu sehen, wie sich die Ergebnisse der merkmalsbasierten Methode verändern, wenn diese Alternativen benutzt würden.

Abschließend wurde in Kapitel 4 gezeigt, dass beide Methoden zu guten Ergebnissen führen, wenn davon ausgegangen werden kann, dass die Verschiebung, sowie die Rotation gering sind. Bei stärkeren Veränderungen jedoch wird die direkte Methode recht ungenau, während die merkmalsbasierte Methode noch gute Ergebnisse erzielt. An ihre Grenzen gerät die merkmalsbasierte Methode erst, wenn sich entweder der Bildinhalt stark ändert, oder die Rotation einen Winkel von 20° überschreitet.

Beide Verfahren arbeiten also sub-pixelgenau, können aber unter verschiedenen Voraussetzungen zu Ungenauigkeiten führen. Werden die Probleme der beiden Methoden beachtet und am Besten bei der Aufnahme oder vor der Registrierung eliminiert werden, so können sehr gute Ergebnisse erzielt werden. Vorteile der merkmalsbasierten Methode ist, dass sie sehr robust gegenüber stärkeren Deformationen, Helligkeits- und Farbschwankungen der Bildserien ist. Von Nachteil ist der vergleichsweise umfangreiche Code, der wiederum bei der direkten Methode sehr minimalistisch gehalten werden kann. Außerdem kann es bei ungünstigen Deformationen vorkommen, dass der Algorithmus sehr lange braucht um zu terminieren (meist im Zusammenhang mit Rotationen). Nachteilig bei der direkten Methode sind die großen Einschränkungen, die vor der Registrierung getroffen werden müssen. Rotation und Translation dürfen nicht zu groß sein, und vor allem wirken sich schon die geringsten Helligkeits- und Farbveränderungen negativ auf das Ergebnis aus. Während diese Einschränkungen für die Verarbeitung von Superresolution Bildern noch

gut zu verkraften sein müssten, ist es unmöglich dieses Verfahren z. B. zum Registrieren von HDR-Bildern zu benutzen.

Literaturverzeichnis

- [AI99] ANANDAN, P. ; IRANI, Michal: About Direct Methods. In: *Workshop on Vision Algorithms*, 1999, S. 267–277
- [BAHH92] BERGEN, James R. ; ANANDAN, P. ; HANNA, Keith J. ; HINGORANI, Rakesh: Hierarchical Model-Based Motion Estimation. In: *ECCV '92: Proceedings of the Second European Conference on Computer Vision*. London, UK : Springer-Verlag, 1992, S. 237–252
- [CKK⁺96] CHEESEMAN, Peter ; KANEFSKY, Bob ; KRAFT, Richard ; STUTZ, John ; HANSON, Robin: Super-Resolved Surface Reconstruction from Multiple Images. Version:1996. citeseer.ist.psu.edu/cheeseman96superresolved.html. In: HEIDBREDER, Glenn R. (Hrsg.): *Maximum Entropy and Bayesian Methods*. Dordrecht, the Netherlands : Kluwer Academic Publishers, 1996, 293–308
- [CKMV96] CHRISTENSEN, G.E. ; KANE, A.A. ; MARSH, J.L. ; VANNIER, M. W.: Synthesis of an Individualized Cranial Atlas with Dysmorphic Shape. In: *mbia 00* (1996), S. 0309
- [FPM⁺94] FRITSCH, Daniel S. ; PIZER, Stephen M. ; MORSE, Bryan S. ; EBERLY, David H. ; LIU, Alan: The multiscale medial axis and its applications in image registration. In: *Pattern Recogn. Lett.* 15 (1994), Nr. 5, S. 445–452
- [GB92] GOTTESFELD BROWN, Lisa: A Survey of Image Registration Techniques. In: *ACM Computing Surveys* 24 (1992), Nr. 4, 325-376. <http://citeseer.ist.psu.edu/gottesfeldbrown92survey.html>

- [HBHH01] HILL, Derek L. G. ; BATCHELOR, Philipp G. ; HOLDEN, Mark ; HAWKES, David J.: Medical image registration. In: *Physics in Medicine and Biology* (2001)
- [HHH91] HAWKES, D.J. ; HARDINGHAM, C. R. ; HILL, D.L.: The use of anatomical knowledge to register 3-D blood vessel data derived from DSA with MR images. In: LOEW, M. H. (Hrsg.): *Proc. SPIE Vol. 1445, p. 348-357, Medical Imaging V: Image Processing, Murray H. Loew; Ed. Bd. 1445, 1991* (Presented at the Society of Photo-Optical Instrume), S. 348–357
- [HHHR93] HILL, Derek L. G. ; HAWKES, David J. ; HARRISON, Neil A. ; RUFF, Cliff F.: A Strategy for Automated Multimodality Image Registration Incorporating Anatomical Knowledge and Imager Characteristics. In: *IPMI '93: Proceedings of the 13th International Conference on Information Processing in Medical Imaging*. London, UK : Springer-Verlag, 1993, S. 182–196
- [HS88] HARRIS, C. ; STEPHENS, M.: A combined corner and edge detector. In: *Fourth Alvey Vision Conference*. Manchester, UK, 1988, 147-151
- [HZ03] HARTLEY, Richard I. ; ZISSERMAN, Andrew: *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003
- [IAB⁺96] IRANI, M. ; ANANDAN, P. ; BERGEN, J. R. ; KUMAR, R. ; HSU, S.: Efficient representations of video sequences and their applications. In: *Signal Processing: Image Communication* (1996). citeseer.ist.psu.edu/irani96efficient.html
- [IAH95] IRANI, M. ; ANANDAN, P. ; HSU, S.: Mosaic based representations of video sequences and their applications. In: *ICCV '95: Proceedings of the Fifth International Conference on Computer Vision*. Washington, DC, USA : IEEE Computer Society, 1995, S. 605
- [IP91] IRANI, Michal ; PELEG, Shmuel: Improving resolution by image registration. In: *Graph. Models Image Process.* 53 (1991), Nr. 3, 231-239. http://www.wisdom.weizmann.ac.il/~irani/PAPERS/SR_CVGIP91.pdf

- [IRP94] IRANI, Michal ; ROUSSO, Benny ; PELEG, Shmuel: Computing occluding and transparent motions. In: *Int. J. Comput. Vision* 12 (1994), Nr. 1, S. 5–16
- [KAI⁺95] KUMAR, Rakesh ; ANANDAN, P. ; IRANI michal ; BERGEN, James R. ; HANNA, Keith J.: Representation of Scenes from Collections of Images. In: *VSR '95: Proceedings of the IEEE Workshop on Representation of Visual Scenes*. Washington, DC, USA : IEEE Computer Society, 1995, S. 10
- [KL81] KANADE, Takeo ; LUCAS, Bruce: An Iterative Image Registration Technique with an Application to Stereo Vision. In: *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, 1981
- [KPB88] KEREN, Danny ; PELEG, Shmuel ; BRADA, Rafi: Image Sequence Enhancement Using Sub-pixel Displacements. In: *Computer Vision and Pattern Recognition, 1988. Proceedings CVPR '88., Computer Society Conference on*, 1988
- [KS83] KASHEF, B. G. ; SAWCHUK, A. A.: A survey of new techniques for image registration and mapping. In: TESCHER, A. G. (Hrsg.): *Applications of Digital Image Processing VI, SPIE Proceedings, Vol. 432. Edited by Andrew G. Tescher. Bellingham, WA: Society for Photo-Optical Instrumentation Engineers, 1983., p.222 Bd. 432, 1983 (Presented at the Society of Photo-Optical Instrume), S. 222–+*
- [Kub06] KUBIAS, Alexander: *Effiziente, adaptive 2-D/3-D-Registrierung von Röntgenbildern und 3-D-Volumen*, Universität Koblenz Landau, Campus Koblenz, Fachbereich 4 Informatik, Institut für Computervisualisitik, Diplomarbeit, 9 2006
- [LA99] LESTER, H. ; ARRIDGE, S. R.: A survey of hierarchical non-linear medical image registration. In: *Pattern Recognition* (1999), S. 129–149
- [Lew01] LEWIS, J. P.: *Fast Normalized Cross-Correlation*. 2001. – Industrial Light and Magic

- [Low99] LOWE, David G.: Object Recognition from Local Scale-Invariant Features. In: *International Conference on Computer Vision*. Corfu Greece, 9 1999, 1150-1157
- [Mor80] MORAVEC, Hans P.: *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover*. Pittsburgh, PA, Carnegie Mellon University, Robotics Institute, Diss., 5 1980. – Available as Stanford AIM-340, CS-80-813 and CMU-RI-TR-3
- [MS04] MIKOLAJCZYK, Krystian ; SCHMID, Cornelia: Scale & Affine Invariant Interest Point Detectors. In: *International Journal of Computer Vision* 60 (2004), Nr. 1, S. 63–86
- [Mur92] MURTAGH, Fionn: A feature-based $O(N^2)$ approach to point pattern matching. In: *Pattern Recognition, 1992. Vol.II. Conference B: Pattern Recognition Methodology and Systems, Proceedings., 11th IAPR International Conference on, 1992*
- [MV96] MAINTZ, J. B. A. ; VIERGEVER, Max A.: *An Overview of Medical Image Registration Methods*. citeseer.ist.psu.edu/255284.html. Version: 1996
- [Nob88] NOBLE, J. A.: Finding Corners. In: *Image and Vision Computing Journal* 6 (1988), 5, Nr. 2, S. 121–128
- [OLL04] OKUMA, Kenji ; LITTLE, James J. ; LOWE, David G.: Automatic rectification of long image sequences. In: *Proceedings of the 6th Asian Conference on Computer Vision, 2004*
- [PVTF02] PRESS, W. ; VETTERLING, W. ; TEUKOLSKY, S. ; FLANNERY, B. P.: *Numerical Recipes in C++: the art of scientific computing*. Cambridge University Press, 2002
- [Roh01] ROHR, Karl: *Landmark-Based Image Analysis: Using Geometric and Intensity Models*. Norwell, MA, USA : Kluwer Academic Publishers, 2001

- [SA96] SAWHNEY, Harpreet S. ; AYER, Serge: Compact Representations of Videos Through Dominant and Multiple Motion Estimation. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (1996), Nr. 8, S. 814–830
- [SAX⁺94] STROTHER, S. C. ; ANDERSON, J. R. ; XU, X. ; LIOW, J. ; BONAR, D. C. ; ROTTENBERG, D. A.: Quantitative comparisons of image registration techniques based on high-resolution MRI of the brain. In: *Journal of Computer Assisted Tomography* 18 (1994), Nr. 6, S. 954–62
- [SC94] SZELISKI, Richard ; COUGHLAN, James: Spline-Based Image Registration. Version:4 1994. citeseer.ist.psu.edu/szeliski94splinebased.html. 1994 (94/1). – Forschungsbericht
- [SV95] SCHUTTE, K. ; VOSSEPOEL, A. M.: Accurate mosaicking of scanned maps, or how to generate a virtual A0 scanner. In: *ASCI'95, Proc. First Annual Conf. of the Advanced School for Computing and Imaging*, 1995
- [Sze94] SZELISKI, Richard: Image Mosaicing for Tele-Reality Applications. In: *Applications of Computer Vision, 1994., Proceedings of the Second IEEE Workshop on*, 1994, 44-53
- [TZ00] TORR, P. H. S. ; ZISSERMAN, A.: MLESAC: a new robust estimator with application to estimating image geometry. In: *Comput. Vis. Image Underst.* 78 (2000), Nr. 1, S. 138–156
- [Zit03] ZITOVA, Barbara: Image registration methods: a survey. In: *Image and Vision Computing* (2003)
- [ZMI99] ZELNIK-MANOR, Lihi ; IRANI, Michal: Multi-Frame Alignment of Planes. In: *cvpr* 01 (1999), S. 1151
- [ZWT99] ZHENG, Zhiqiang ; WANG, Han ; TEOH, Eam K.: Analysis of gray level corner detection. In: *Pattern Recognition Letters* 20 (1999), Nr. 2, S. 149–162