

Verfahren zur Fluchtpunktschätzung

Studienarbeit im Studiengang Computervisualistik

vorgelegt von

Timo Krusche

Betreuer: Prof. Dr.-Ing. Dietrich Paulus, Institut für Computervisualistik,
Fachbereich Informatik
Erstgutachter: Prof. Dr.-Ing. Dietrich Paulus, Institut für Computervisualistik,
Fachbereich Informatik
Zweitgutachter: Herr Clemens Schmidt, externer Doktorand

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	Projekt Ornamente	7
1.1.1	Vorverarbeitung	8
1.1.2	”Region of Interest”	9
1.1.3	Analyse und Klassifikation	9
1.2	Motivation und Vorhandenes	9
1.3	Gliederung der Arbeit	12
2	Stand der Wissenschaft	15
2.1	Perspektivische Verzerrung und Fluchtpunkte	15
2.2	Kaskadierte Hough Transform	18
2.3	”Triangle Area Minimization”	21
2.4	Weitere Verfahren zur Fluchtpunktfindung	22
2.5	Homographie und Referenzpunkte	23
3	Eigener Ansatz	27
3.1	Warum ist die Entzerrung so wichtig?	27
3.2	Problemstellung und Zielsetzungen	30

3.2.1	Umstrukturierung des Vorhandenen	30
3.2.2	”Level of Detail”	33
4	Experimente und Ergebnisse	39
4.1	Erläuterung des Tests	39
4.2	Ergebnisse und Verifikation	41
5	Zusammenfassung	45

Kapitel 1

Einleitung

1.1 Projekt Ornamente

Im Juli 2004 wurde an der Universität Koblenz-Landau das Forschungsprojekt "Ornamente¹ - Anwendungen der Kunstwissenschaft" ins Leben gerufen. Unter der Leitung von Prof. Dr. Dietrich Paulus (Institut für Computervisualistik) und Prof. Dr. Ludwig Tavernier (Institut für Kunstwissenschaft) verbindet dieses interdisziplinäre Projekt die Disziplinen der Informatik und der Kunstwissenschaft.

Es wird versucht, Methoden der Informatik zu entwickeln, um das kunstwissenschaftliche Forschungsgebiet der Ornamentik zu untersuchen. Es soll eine digitale Diathek aufgebaut werden, die es ermöglicht, per Webinterface eine beliebige Abbildung von kunsthistorischem Interesse hochzuladen, das Abgebildete automatisch auf Herkunft, Stil und Zeitepoche zu analysieren und einzukategorisieren. Momentan liegt das Hauptaugenmerk auf der Analyse von Abbildungen historischer Gebäude.

Kunstvoll gestaltete Gebäudefassaden sind meist entlang der Propyläen, also oberhalb von Säulengängen, mit Dekorationen versehen, die charakteristische Merkmale für die kunsthistorische Einordnung enthalten. Diese Ornamente sollen automatisch gefunden,

¹Homepage Projekt Ornamente

<http://www.uni-koblenz.de/FB4/Institutes/ICV/AGPaulus/Researches/Arts>

erkannt und klassifiziert werden.

Der hierfür entworfene Versuchsaufbau, in Bild 1.1 schematisch dargestellt, basiert auf einer digitalen Bilddatenbank, die momentan mehr als 1500 Abbildungen beinhaltet. Dabei handelt es sich um eingescannte Abbildungen aus den Bänden der "Propyläen der Kunstgeschichte" (verschiedene Bücher; zu finden in der Universitätsbibliothek Koblenz-Landau) und eingescannten Dias, die in der Diathek des Instituts für Kunstgeschichte der Universität Koblenz-Landau gesammelt wurden.

Die Bilddatenbank stellt keineswegs besonders bearbeitetes oder speziell aufgenommenes Bildmaterial bereit. Vielmehr sind hier gescannte Fotos gespeichert, wie sie jeder Hobbyfotograf schießen und digitalisieren könnte - teilweise gedreht eingescannte Abbildungen; Dias, deren Rahmen zu sehen sind; durch Artefakte wie Flusen, Staub und Kratzer verunreinigte Bilder.

Um gute Ergebnisse der Analyse zu gewährleisten muss jeder Verarbeitungsschritt zuverlässige Teilergebnisse liefern. Wird ein Bild der Datenbank analysiert, durchläuft es das Gesamtsystem, das sich in folgende Arbeitsschritte gliedert:

1.1.1 Vorverarbeitung

Clipping - Dieses Thema bearbeitete Andreas Thun u.A. in seiner Diplomarbeit [Thu05]. Ist auf einem Bild der Diarahmen oder der weiße Hintergrund einer sehr dunklen, eingescannten Abbildung zu sehen, ergeben sich überflüssige harte Kanten im Bild, die eine Liniendetektion in darauffolgenden Arbeitsschritten verfälschen würden. Die für die Analyse störenden und unbedeutenden Ränder werden abgeschnitten und das Bild freigestellt. Gleichzeitig werden in einem beliebigen Winkel schräg eingescannte Bilder horizontal ausgerichtet.

Artefakte - Kratzer, Flusen und Staub, wie auch Rauschen im Bild, gelten als Störfaktoren während der Analyse. Die eigentlichen Bildinformationen sind dadurch verfälscht. Die Bilder müssen gesäubert werden. Hierzu ist eine Studienarbeit ausgeschrieben (Stand 10.10.06). Nähere Informationen sind der Webseite der AGAS² zu entnehmen.

²Arbeitsgruppe Aktives Sehen: www.uni-koblenz.de/FB4/Institutes/ICV/AGPaulus

Perspektivische Entzerrung - Diese Studienarbeit der Fluchtpunktdetektion ist diesem Arbeitsschritt zuzuordnen. Mit Hilfe detektierter Fluchtpunkte lassen sich Referentpunkte ermitteln, über die eine Homographiematrix H aufgestellt wird. Transformiert man das verzerrte Bild mit dieser Matrix H , erhält man ein entzerrtes Bild, ohne Verlust der geometrischen Eigenschaften des Abgebildeten. Die orthogonal projizierte, entzerrte Darstellung der Bilder ist für die weitere Analyse auf Ornamente nötig.

1.1.2 "Region of Interest"

Um nicht in Bereichen nach Ornamenten zu untersuchen, in denen keine zu sehen sind, werden interessante, ornamentale Bereiche vorab eingegrenzt. Mit Hilfe des von Timo Dickscheid implementierten SIFT-Operators [Dic05] definierte Stephan Steneberg die "Region of Interest" (kurz: ROI) in seiner Studienarbeit [Ste06]. Die ROI wird als kleineres, separates Bild weiterverarbeitet, was die Performanz späterer Arbeitsschritte verbessert.

1.1.3 Analyse und Klassifikation

Die Klassifikation geschieht entweder durch einen direkten Mustervergleich (vgl. NC-CF [PH03]), wobei das Muster aus der Knowledge Base entnommen und, in Größe und Schräglage variierend, mit dem Bild verglichen wird. Die Merkmalsuche mit Hilfe des SIFT-Operators stellt eine zweite Variante dar, die das Bild mehrmals analysiert und nach jedem Durchgang die Parameter des SIFTs optimiert. Ein dritte Variante von Matthias Dennhardt [Den05] sucht Symmetrien. Jede Variante zielt darauf ab, einen Merkmalsvektor c zu erstellen, der die gefundenen Ornamentmuster repräsentiert und zu jedem Bild in der Bilddatenbank abgelegt wird.

1.2 Motivation und Vorhandenes

Die Fluchtpunktdetektion ist der Vorverarbeitung zugeordnet. Die Eingabebilder sind bereits von möglichen Rändern entfernt und ihre Schräglage korrigiert. In der Bilddatenbank

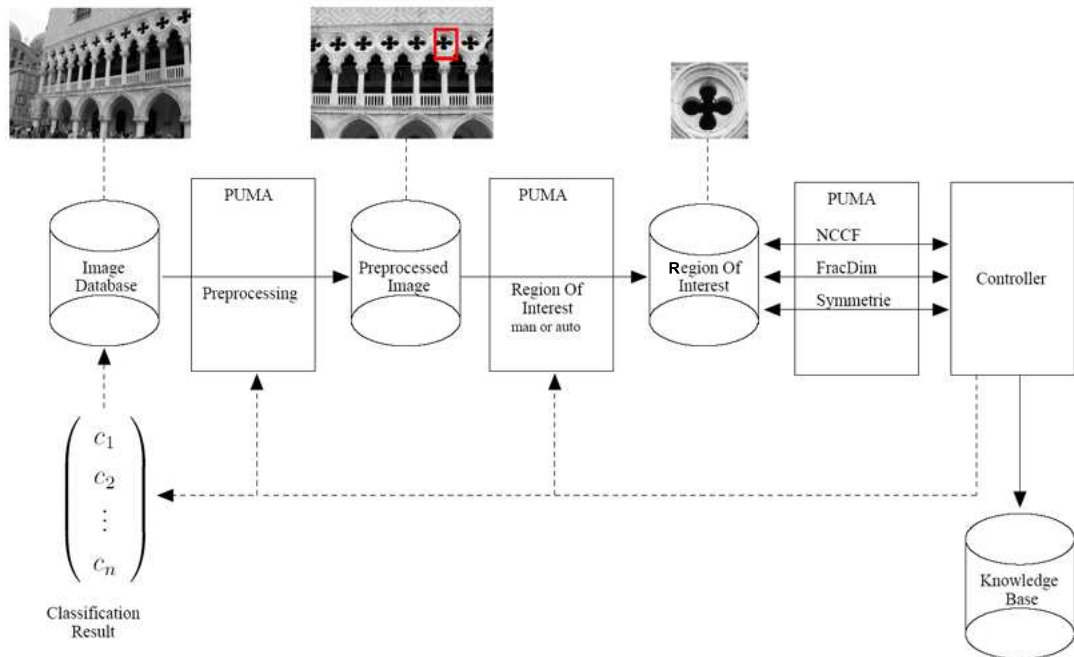


Bild 1.1: System zur automatischen Ornamentklassifikation. Bilder werden vorverarbeitet und müssen perspektivisch entzerrt werden, um Ornamente detektieren zu können. Diese werden anschließend klassifiziert. Das Ergebnis wird in einer Datenbank gespeichert. *Quelle:* [Ste06]

sind Skizzen zu verschiedenen Ornamenten zu finden, die nicht perspektivisch verzerrt sind, alle anderen sind dies jedoch. Eine nicht verzerrte Darstellung ist für die darauf folgende ROI-Bestimmung und die Analyseverfahren notwendig.

Den Algorithmus zur Entzerrung setzte Andreas Thun [Thu05] mit dem Modul `PerpsCorr` um. Er berechnet eine projektive Transformationsmatrix H , die jeden Punkt x' eines verzerrten Bildes in den korrespondierenden Punkt x im Idealbild, die entzerrte Darstellung, überführt. Diese Matrix H wird *Homographiematrix* genannt, der Vorgang selbst *Homographie*¹ - "Die Abbildung einer Ebene in eine andere" [Thu05]. Um H aufzustellen, sind

¹Homographie, aus dem Griechischen: homos - gleich; graphien - zeichnen, schreiben

vier Referenzpunkte zu wählen. Diese werden anhand der Fluchtpunkte² ermittelt.

Soll eine abgebildete Gebäudefront untersucht werden, muss diese orthogonal projiziert dargestellt sein. Zur Berechnung der Homographmatrix muss die Ebene definiert werden, in der die zu entzerrende Gebäudefront liegt. Nehmen wir an, es handelt sich um ein rechteckiges Gebäude mit mehreren Fenster mit parallelen Ober- und Unterkanten. Ist während der Ablichtung die Projektionsfläche nicht parallel zur Gebäudefront ausgerichtet, entsteht die perspektivische Verzerrung. Die parallelen Kanten konvergieren in ihrer Abbildung und schneiden sich einem Fluchtpunkt. Durch die Projektion eines dreidimensionalen, kubischen Objekts auf eine zweidimensionale Projektionsfläche können drei Fluchtpunkte in der Abbildung entstehen. Die Ebene in der eine Seitenfläche dieses Objekts liegt, wird anhand von zwei Fluchtpunkten, die sich durch die perspektivische Verzerrung des Objekts ergeben, ermittelt. Die vertikalen Kanten der Seitenfläche ergeben einen Fluchtpunkt, die horizontalen Kanten einen zweiten. Diese Kanten müssen detektiert und in vertikalen und horizontalen Geraden gruppiert werden. So können die Fluchtpunkte ermittelt und mit ihrer Hilfe Referenzpunkte der zu entzerrenden Ebene bestimmt werden.

Sind die Referenzpunkte der Ebene bekannt, können sie beim Ausführen von `PerpsCorr` mit angegeben werden. Sollte dies nicht der Fall sein, wird das Modul `HoughTransform` [Thu05] aufgerufen, um die Fluchtpunkte und die sich daraus ergebenden Referenzpunkte automatisch zu ermitteln. Das Eingabebild wird in ein Kantenbild (vgl. Sobel [PH03]) umgewandelt und per `Hough` transformiert. Alle Pixel im Kantenbild mit gleicher Orientierung und gleichem Abstand zum Ursprung werden aufsummiert, wie in [PH03] zu `Hough` beschrieben. Im `Hough`raum können so durch Histogrammanalysen konvergierende Kanten im Bild gefunden werden. Ausgehend von diesen Kanten wird per "Triangle Area Minimization" (kurz: TAM) 2.3 ihr zugehöriger Fluchtpunkt geschätzt.

Mit dem vorhandenem Verfahren `HoughTransform` ist es möglich die Ebene der dominanten Gebäudeseite zu entzerren.

Aus diesem Sachverhalt heraus entstand die Aufgabenstellung für dieser Studienarbeit:

1. Nachdem dominante Gebäudeseiten bereits entzerrt werden können, soll das Beste-

²Im Fluchtpunkt einer perspektivischen Abbildung schneiden sich alle Geraden, die im Original parallel zueinander verlaufen.

hende um die Möglichkeit erweitert werden, zusätzlich eine kleinere, zweite Gebäudefront zu entzerren. Dies stellt die Hauptaufgabe dieser Arbeit dar.

2. Verfahren zur Fluchtpunktschätzung sollen auf ihre Vor- und Nachteile überprüft werden, um eine Möglichkeit für das "Entzerren von mehreren Gebäudeseiten" zu finden.
3. Die Programme, die im Zuge dieser Arbeit entstehen, müssen wie das Gesamtsystem in die PUMA³-Umgebung eingegliedert und in objektorientiertem C++ implementiert sein.
4. Der modulare Aufbau des Systems soll ausgenutzt werden, so dass Kantenextraktion und Fluchtpunktbestimmung unabhängig voneinander eingesetzt werden können. Spätere Erweiterungen oder Änderungen dieser Teilprozesse zur Entzerrung können so leicht in das System eingebaut werden.
5. Eine eigene Idee mit dem Namen "Level of Detail" wurde zur Lösung des Hauptproblems der zweiten, zu entzerrenden Gebäudefront entwickelt. Sie ist als Erweiterung des vorhandenen Moduls `HoughTransform` anzusehen.

1.3 Gliederung der Arbeit

Die Entstehung einer perspektivischen Verzerrung bei der Projektion eines Objekts auf die Bildebene während der Ablichtung wird in Kapitel 2 beschrieben. Die Definition für Fluchtpunkte wird gegeben und erklärt, wie diese als charakteristisches Merkmal der vorliegenden Perspektive dienen. Ferner werden verschiedene Verfahren zur Fluchtpunktfindung vorgestellt und ihre Funktionsweisen erläutert. Zuletzt wird die Homographie zur perspektivischen Entzerrung angesprochen.

Eine genauere Betrachtung der Problemstellung und Lösungsansätze dieser Studienarbeit findet sich in Kapitel 3. Hierbei wird auf bereits bestehende Prozesse und das daran ansetzende eigene Verfahren names "Level of Detail" eingegangen.

³Programmierungsumgebung für die Bildverarbeitung und Musteranalyse

<http://www.uni-koblenz.de/~puma>

Kapitel 4 erläutert den Test, der zur Überprüfung auf korrekt detektierte Fluchtpunkte gewählt wurde und zeigt einige Beispielbilder.

Kapitel 5 weist auf weiterhin bestehende Probleme hin und schließt diese Studienarbeit ab.

Kapitel 2

Stand der Wissenschaft

Dieses Kapitel liefert unerlässliche Informationen zum Verständnis über die Entstehung der perspektivischen Verzerrung und die Definition eines Fluchtpunktes. Verschiedene Verfahren zur Fluchtpunkt-detektion und die Homographie zur perspektivischen Entzerrung werden an dieser Stelle beschrieben.

2.1 Perspektivische Verzerrung und Fluchtpunkte

Wird ein Objekt abgelichtet, wird es durch das Kameraobjektiv auf die Bildebene projiziert. Abhängig von der Ausrichtung der Bildebene in Bezug auf das Objekt entsteht die perspektivische Verzerrung. Bild 2.1 zeigt die Abbildung einer kubischen CD-Hülle aus drei verschiedenen Blickrichtungen. In der Welt verläuft die Oberkante des roten Balkens parallel zur Tischkante, wie es im mittleren Bild zu sehen ist. Hier ist das Kameraobjektiv vertikal und horizontal auf den Mittelpunkt des Kreises ausgerichtet. Die Bildebene liegt somit koplanar zur Frontseite der CD-Hülle, was dazu führt, dass das Cover bildparallel und ohne perspektivische Verzerrung abgelichtet wird. In dieser Perspektive sind die gegenüberliegenden Raumkanten der Hülle parallel, die Kreise kreisrund und die rechteckige Hülle auch als Rechteck zu sehen. Die beiden äußeren Abbildungen weisen eine perspektivische Verzerrung auf, da die Bildebene zum Zeitpunkt der Aufnahme nicht koplanar zur CD-Hülle lag. Die Hülle ist nicht mehr als Rechteck abgebildet, die Kreise sind elliptisch

und der rote Balken verjüngt sich nach außen hin und erzeugen so einen Tiefeneffekt.



Bild 2.1: Verzerrung einer CD-Hülle durch verschiedene Blickwinkel. *mitte*: Koplanare Abbildung mit quadratischer Hülle und gleichbleibende Höhe des roten Balkens *außen*: Man beachte die Verjüngung des roten Balkens und die elliptische Form der Kreise

Die Entstehung der Verzerrung wird klar ersichtlich, betrachtet man den Lastkraftwagen in Bild 2.2, aufgenommen in unterschiedlicher Distanz zur Kamera. Je weiter er von der Kamera entfernt ist, desto kleiner erscheint er. Genauso verhält es sich mit den Seitenkanten der CD-Hülle. Trotz gleicher Größe wird die entferntere Kante kleiner abgebildet.

Die Ausrichtung der Bildebene in Bezug auf das Abgelichtete während der Aufnahme ist somit entscheidend für die perspektivische Verzerrung. Jede Veränderung des Kamerastandortes oder des Blickwinkels zeigt die selbe Szene in einer anderen Perspektive. Für die Entzerrung eines Bildes liegen in der Regel keinerlei Informationen über den räumlichen Zusammenhang zwischen Bildebene und dem Abgelichteten vor, aufgrund derer man im Vorfeld geometrisch berechnen könnte, welche Perspektive und somit Verzerrung vorliegt. Die Informationen, die für die Entzerrung notwendig sind, müssen aus dem Bild selbst entnommen werden.

Jede veränderte Perspektive würde Ober- und Unterkante der CD-Hülle in einem anderen Winkel zueinander zeigen. Verlängert man diese beiden Kanten zu Geraden, schneiden sie sich im Fluchtpunkt. Nach der Definition für Fluchtpunkte gilt: "Im Fluchtpunkt einer perspektivischen Abbildung schneiden sich alle Geraden, die im Original parallel zueinander verlaufen."¹

¹Quelle: www.wikipedia.org



Bild 2.2: Lastkraftwagen in unterschiedlicher Distanz zur Bildebene abgelichtet

Als spezifisches Merkmal der vorliegenden Perspektive können somit die Fluchtpunkte herangezogen werden, mit Hilfe derer die Homographiematrix H bestimmt wird.

Zu ergänzen ist, dass ein Fluchtpunkt auch im Unendlichen liegen kann. Denn die parallelen Seitenkanten einer Gebäudefront, die es zu entzerren gilt, können auch in ihrer Abbildung parallel verlaufen. Die Ebene dieser Front wird dennoch über die Fluchtpunkte ihrer Seitenkanten bestimmt, der in diesem Fall im Unendliche liegt.

Ferner kann keine Vorhersage über die Position der Fluchtpunkte getroffen werden. Sie können innerhalb wie außerhalb des Bildes liegen und vom Mittelpunkt des Bildes aus gesehen in allen Richtungen über 360 Grad verteilt vorhanden sein. Zur Verdeutlichung betrachte man die möglichen Perspektiven.

Bild 2.3 zeigt drei Perspektiven. Die mittlere und rechte Abbildung zeigen Vogel- und Froschperspektive mit jeweils drei Fluchtpunkten. Links ist die *Über-Eck-Perspektive* zu sehen, wie sie in der darstellenden Geometrie verwendet wird. In dieser konstruierten Darstellung wird angenommen, dass alle vertikalen Geraden im Lot stehen, wodurch sie sich

sehr stark an die korrekte Projektion eines dreidimensionalen Körpers auf eine zweidimensionale Bildebene annähert. Bei einer korrekten Projektion durch Fotografieren ist dies jedoch nur der Fall, wenn sich der Augpunkt genau auf halber Höhe des Objektes befindet und der Sehstrahl orthogonal auf das Objekt trifft. Weicht der Augpunkt oder Sehstrahl davon ab, entsteht wiederum eine Verzerrung und die Seitenkanten verjüngen sich nach oben oder nach unten hin. In dieser Arbeit wird davon ausgegangen, dass diese Perspektive ein Sonderfall ist, der gleichzeitig Vogel- und Froschperspektive zugeordnet werden muss. Der Fluchtpunkt ober- oder unterhalb des Objekts liegt somit bei parallelen Seitenkanten im Unendlichen, wird aber als Fluchtpunkt behandelt.

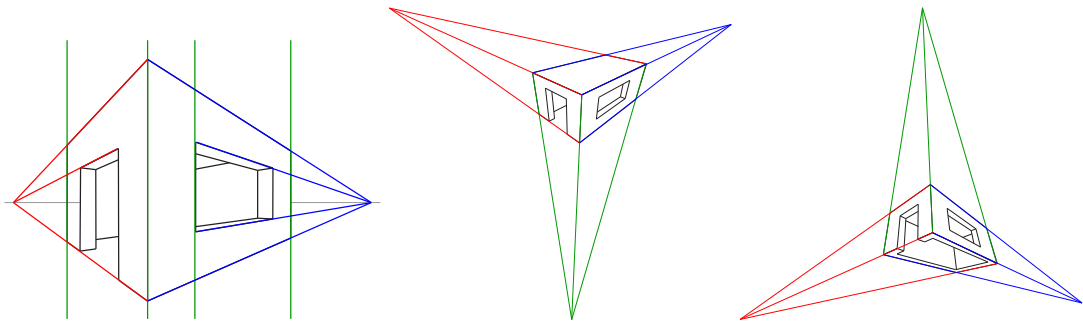


Bild 2.3: Über-Eck-, Vogel- und Froschperspektive

Die Fluchtpunkte werden anhand der im Bild enthaltenen Kanten ermittelt und dienen als Grundlage für die Berechnung der Homographiematrix H . Mit dieser wird das Bild transformiert, und somit entzerrt. Im Folgenden werden verschiedene Varianten zur Fluchtpunktdetektion vorgestellt und die Homographie erläutert.

2.2 Kaskadierte Hough Transform

Diese Variante zur Fluchtpunktbestimmung von Tuytelaars et al. [TGPM98] ist eine Erweiterung der Hough Transformation von P.V.C. Hough [HA62], die 1962 patentiert wurde. Der Grundgedanke besteht darin, charakteristische Merkmale eines Bildes in parametrisierter Form darzustellen, um so die Suche nach diesen Merkmalen zu beschleunigen.

Eine genauere Beschreibung der Standard Hough Transformation für Geraden und Kreise gibt Paulus [PH03].

Wie im Beispiel der CD-Hülle beschrieben, schneiden sich die konvergierenden Seitenkanten im Fluchtpunkt. Um die Fluchtpunkte anhand von Geraden im Bild zu finden, wird vorausgesetzt, dass das Eingabebild als Kantenbild vorliegt. Der gewählte Kantendetektor ist hierfür nicht entscheidend. Im Kantenbild muss lediglich für jeden Punkt die Stärke und die Richtung der Kante an diesem Punkt im Kantenbild vorliegen.

Die Standard Hough Transformation parametrisiert Geraden durch ihre Steigung α , also der im Kantenbild eingetragenen Steigung für den jeweiligen Punkt auf der Geraden und dem Abstand d dieses Punktes zum Ursprung. So wird eine Gerade im Bild auf einen Punkt im Parameterraum abgebildet. Vice versa gilt dies bei der Standard Hough Transformation jedoch nicht. Im Gegensatz dazu zielt die kaskadierende Hough Transformation (kurz: KHT) auf die Symmetrie zwischen Bildraum und Parameterraum ab.

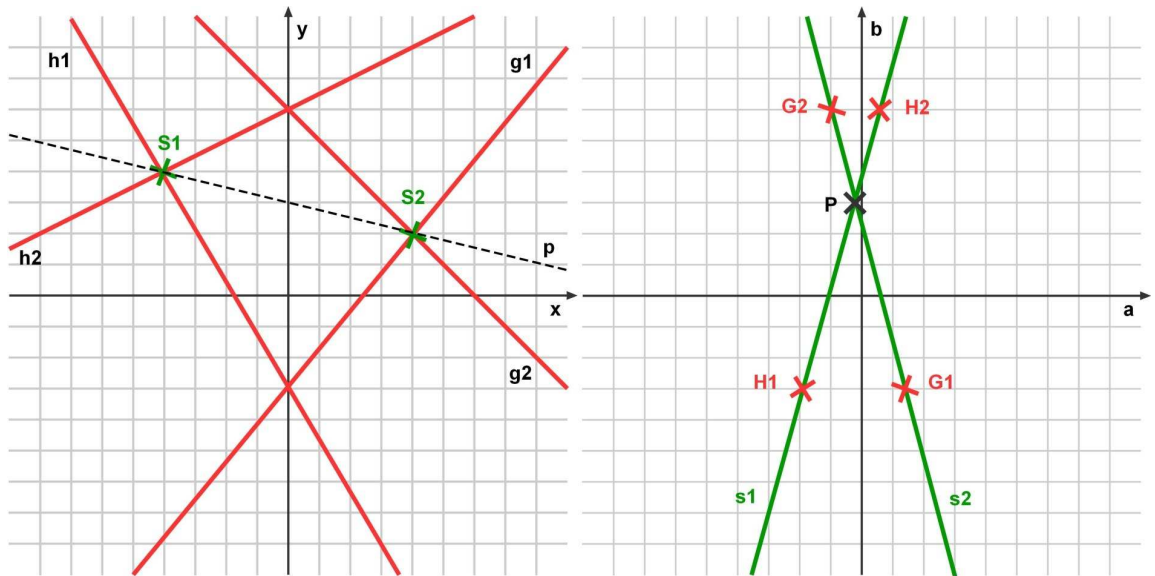
Bei der KHT geschieht die Parametrisierung einer Geraden durch die Parameter (a,b) nach der Gleichung

$$ax + b + y = 0 . \quad (2.1)$$

Durch diese (a,b) -Parametrisierung wird ein Punkt im (x,y) -Kantenbild auf eine Gerade im (a,b) -Parameterraum abgebildet. Ein Punkt im (a,b) -Parameterraum korrespondiert gleichzeitig zu einer Geraden im (x,y) -Kantenbild. Die Parameter a und b verhalten sich zum Bildraum (x,y) genau wie x und y zum Parameterraum (a,b) . Diese Symmetrie ermöglicht, dass das Ergebnis der ersten Transformation erneut auf gleicher Weise transformiert werden kann.

Zur Verdeutlichung zeigt Bild 2.2 zwei Geradenpaare h_1, h_2 mit Schnittpunkt S_1 und g_1, g_2 mit Schnittpunkt S_2 . So wird nach der ersten Transformation im Parameterraum S_1 als Gerade s_1 repräsentiert. Diese verläuft durch die zu den Geraden h_1 und h_2 korrespondierenden Punkte H_1 und H_2 . Gleichermaßen verläuft die Gerade s_2 durch G_1 und G_2 . Im Parameterraum ist der Schnittpunkt P von s_1 und s_2 gleichbedeutend, wie die Gerade p durch die beiden Schnittpunkte S_1 und S_2 , die im Ausgangsbild als gestrichelte Linie dargestellt ist.

Durch die erste Transformation können Linien des Ausgangsbildes im Parameterraum er-

Bild 2.4: (x,y) -Bildraum und (a,b) -Parameterraum

kannt werden. Durch die zweite Transformation, auf den Parameterraum angewandt, werden Linien kollinear Punkte gefunden, die die Schnittpunkte von konvergierenden Geraden im Ausgangsbild widerspiegeln. Eine weitere Transformation des Ergebnisses der zweiten Transformation ermöglicht den Zugriff auf kollineare Schnittpunkte des Ausgangsbildes.

Bei der (a,b) -Parametrisierung nach Gleichung 2.1 können a und b allerdings unendlich große Werte annehmen. Um dies zu umgehen, wird der unbeschränkte (a,b) -Raum in drei beschränkte "Subspaces", wie in Bild 2.2 gezeigt, aufgeteilt.

Um die Symmetrie zwischen Bildraum und Parameterraum zu gewährleisten, wird der Bildraum entsprechend dem Parameterraum unterteilt. Es werden wieder drei "Subspaces" mit den Koordinaten (x,y) , $(1/x,y/x)$ und $(1/y,x/y)$ erstellt. Der Bildraum, zuerst beschränkt durch die Größe des Bildes an sich, ist nun als unbegrenzter Raum dargestellt, wodurch Fluchtpunkte, die im Unendlichen liegen, auf einfache Weise in einem der äußeren "Subspaces" abgebildet werden.

Diese Variante der Fluchtpunktdetektion ermöglicht das Auffinden von Fluchtpunkten in

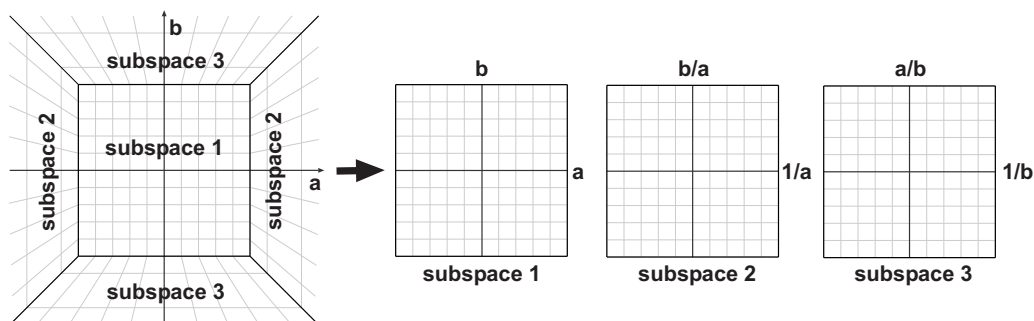


Bild 2.5: Unterteilung des Parameterraums und des Bildes in "Subspaces"

sehr kurzer Zeit, ist allerdings sehr speicherintensiv. Die passende Größe der Subräume muss gefunden werden. Gleichzeitig muss jedes Ergebnis einer Transformation entsprechend den zu untersuchenden Merkmalen gefiltert werden. Nach Tuytelaars stellt dies das Hauptproblem dieses Ansatzes dar.

2.3 "Triangle Area Minimization"

Eine Variante von Bräuer-Burchardt und Voss [BBV00] nutzt zur Fluchtpunktbestimmung wieder die Geraden im Bild. Sind die auf einen Fluchtpunkt $\mathbf{v} = (x_v, y_v, 1)^T$ zulaufenden Geraden s_i gefunden, bilden die Endpunkte p_i und q_i von s_i mit \mathbf{v} ein Dreieck, dessen Fläche A_i es zu minimieren gilt.

Für die Fläche $A_i = 0$ gilt für alle s_i

$$A_i = \frac{1}{2} \begin{pmatrix} x_{ip} & y_{ip} & 1 \\ x_{iq} & y_{iq} & 1 \\ x_v & y_v & 1 \end{pmatrix} = 0 \quad (2.2)$$

Diese Gleichung wird umgeformt zu:

$$x_{ip}y_{iq} - x_{ip}y_v - x_{iq}y_{ip} + x_{iq}y_v + x_vy_{ip} - x_vy_{iq} = 0 \quad (2.3)$$

und anschließend in die Matrixschreibweise umgeformt.

$$(y_{ip} - y_{iq}, x_{iq} - x_{ip}, x_{ip}y_{iq} - x_{iq}y_{ip}) \begin{pmatrix} x_v \\ y_v \\ 1 \end{pmatrix} = 0 \quad (2.4)$$

Sind mehr als 2 Geraden gefunden, also $n > 2$, die auf einen gemeinsamen Fluchtpunkt v_i hinzeigen, lässt sich ein überbestimmtes lineares Gleichungssystem der Form $\mathbf{A}\mathbf{v} = 0$ aufstellen.

$$\begin{pmatrix} y_{1p} - y_{1q} & x_{1q} - x_{1p} & x_{1p}y_{1q} - x_{1q}y_{1p} \\ y_{2p} - y_{2q} & x_{2q} - x_{2p} & x_{2p}y_{2q} - x_{2q}y_{2p} \\ \dots & \dots & \dots \\ y_{np} - y_{nq} & x_{nq} - x_{np} & x_{np}y_{nq} - x_{nq}y_{np} \end{pmatrix} \begin{pmatrix} x_v \\ y_v \\ 1 \end{pmatrix} = 0 \quad (2.5)$$

Liegt die Gleichung in dieser Form vor, kann \mathbf{v} anhand der Singularwertzerlegung (SVD: Singular Value Decomposition), wie in [TV98] von Trucco und Verri beschrieben², bestimmt werden.

2.4 Weitere Verfahren zur Fluchtpunktfindung

Neben der TAM und der KHT gibt es noch weitere Ansätze. Die Mittelwertbildung nach Liebowitz und Zissermann [LZ98] berechnet zu allen Geradensegmenten s_i , die einen gemeinsamen Fluchtpunkt v_i besitzen, paarweise ihre Schnittpunkte (x_i, y_i) . Durch Rauschen im Bild und abhängig von der Genauigkeit der Kantendetektion liegen diese Schnittpunkte, die sich im Idealfall alle in einem Fluchtpunkt treffen, nicht genau in einem Punkt. Für den Fluchtpunkt wird der Mittelwert aller Schnittpunkte nach den Gleichungen

$$x_v = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.6)$$

$$y_v = \frac{1}{n} \sum_{i=1}^n y_i \quad (2.7)$$

gemittelt.

²SVD nach Trucco, Verri in [TV98] Appendix A.6 S.322ff

Anstatt das Eingabebild per Standard Hough Transformation zu transformieren, um im Parameterraum die Fluchtpunkte über die dort enthaltenen Maxima zu bestimmen, stellt Rother in [Rot00] eine Variante vor, bei der die Geraden im Bild auf eine Gaußsche Kugel im Zentrum der Kameraöffnung abgebildet werden. Die Geraden werden als Kreislinien dieser Kugel beschrieben. Der Schnittpunkt zweier Kreislinien stellt so den Fluchtpunkt dar. Vorteil hierbei ist, dass Fluchtpunkte, die im Unendlichen liegen, dennoch leicht ermittelt werden können.

Einen weiteren Ansatz, der Fluchtpunkte im Unendlichen erkennt und gleichzeitig die Anzahl falscher Alarme minimiert, entwickelte Almansa et al. [ADV03]. Als falschen Alarm wird ein als Fluchtpunkt erkannter Schnittpunkt zweier Geraden bezeichnet, wobei die Geraden sich nur zufällig schneiden, aber nicht zur selben Gebäudefront gehören. Auch hier werden Geraden auf die Gaußsche Kugel im Zentrum des Kameraobjektivs abgebildet. Diese ist allerdings tesseliert und die Fluchtpunkte werden aufgrund der Wahrscheinlichkeit, dass der Fluchtpunkt in einer der Kacheln dieser Tessellierung liegt, bestimmt.

Nachdem einige Varianten, Fluchtpunkte zu bestimmen, vorgestellt wurden, stellt sich die Frage, was mit den gewonnenen Informationen über die Beschaffenheit der Verzerrung geschehen muss, um das Ausgangsbild zu entzerren.

2.5 Homographie und Referenzpunkte

Mit Hilfe der gefundenen Fluchtpunkte werden nun Referenzpunkte der zu entzerrenden Ebene ermittelt. Die Referenzpunkte stellen die Grundlage zur Berechnung der Homographiematrix H dar.

Homographie bezeichnet die Abbildung einer Ebene auf eine andere Ebene. Ziel ist es, über die Referenzpunkte die Ebene im Bild zu bestimmen, in der die zu entzerrende Gebäudefront liegt. Diese verzerrte Ebene wird mit der Entzerrungsmatrix H in eine unverzerrte Ebene überführt. Im verzerrten Bild beschreiben die Referenzpunkte ein beliebiges Viereck, die nach der Entzerrung ein Rechteck bilden. Das entzerrte Bild wird Idealbild genannt.

Bild 2.6 zeigt den Idealfall, in dem die Eckpunkte der Gebäudefront als Referenzpunk-

te, rot eingezeichnet, gewählt wurden. Die hierfür angewandte projektive Transformation wird durch die Homographiematrix \mathbf{H} beschrieben, in diesem konkreten Fall:

$$\mathbf{H} = \begin{pmatrix} 0.122 & 0.0442 & 186 \\ -0.823 & 1.92 & 220 \\ -0.00131 & 0.000112 & 1.3 \end{pmatrix} \quad (2.8)$$

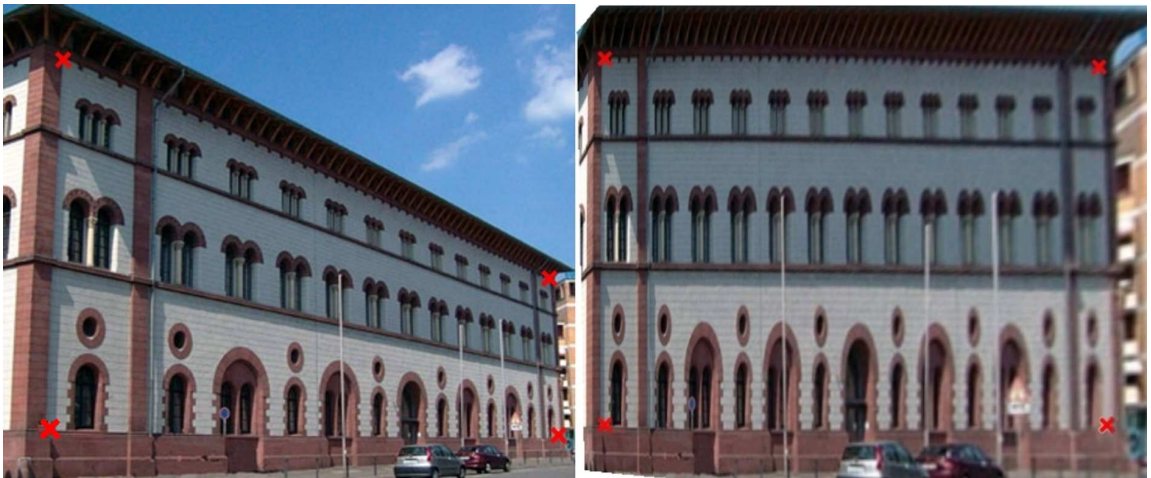


Bild 2.6: Verzerrtes Bild und Idealbild - Referenzpunkte bilden nach der Transformation ein Rechteck (*Quelle*: Bilddatenbank Ornamente)

Punkt x' im Idealbild korrespondiert mit Punkt x im verzerrten Bild (beide in homogenen Koordinaten angegeben) und wird nach der Gleichung $x' = \mathbf{H}x$, und somit durch

$$\begin{pmatrix} x' \\ y' \\ k' \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2.9)$$

transformiert.

Zerlegt man diese Gleichung in ihre Komponenten, so erhält man

$$\begin{aligned} x' &= h_{11}x + h_{12}y + h_{13} \\ y' &= h_{21}x + h_{22}y + h_{23} \\ k' &= h_{31}x + h_{32}y + h_{33}. \end{aligned} \quad (2.10)$$

Gelten muss, dass $k' = 1$ ist, was zu

$$\begin{aligned}x' &= \frac{x'}{k'} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \\y' &= \frac{y'}{k'} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \\x' &= \frac{k'}{k'} = 1\end{aligned}\tag{2.11}$$

führt, und für x' und y' die nach 0 aufgelösten Gleichungen

$$\begin{aligned}h_{11}x + h_{12}y + h_{13} - h_{31}x'x + h_{32}x'y + h_{33}x' &= 0 \\h_{21}x + h_{22}y + h_{23} - h_{31}y'x + h_{32}y'y + h_{33}y' &= 0\end{aligned}\tag{2.12}$$

ergibt. Matrix \mathbf{H} entspricht dem Spaltenvektor \mathbf{h}

$$\mathbf{h} = (h_{11} \ h_{12} \ h_{13} \ h_{21} \ h_{22} \ h_{23} \ h_{31} \ h_{32} \ h_{33})^T .\tag{2.13}$$

Die Gleichung 2.12 kann in Matrixschreibweise dargestellt werden:

$$\begin{pmatrix}x & y & 1 & 0 & 0 & 0 & -x'x & -x'y & -x' \\0 & 0 & 0 & x & y & 1 & -y'x & -y'y & -y'\end{pmatrix} \mathbf{h} = 0 .\tag{2.14}$$

Eine projektive Homographiematrix hat acht Freiheitsgrade. Somit werden auch acht Gleichungen benötigt, um das Gleichungssystem lösen zu können. Ein Referenzpunkt liefert zwei Gleichungen. Somit müssen vier Referenzpunkte gewählt werden um die Lösungsmatrix \mathbf{A} nach $\mathbf{A}\mathbf{h} = 0$ aufstellen zu können.

$$\begin{pmatrix}x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1x_1 & -x'_1y_1 & -x'_1 \\0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1x_1 & -y'_1y_1 & -y'_1 \\x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2x_2 & -x'_2y_2 & -x'_2 \\0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2x_2 & -y'_2y_2 & -y'_2 \\x_3 & y_3 & 1 & 0 & 0 & 0 & -x'_3x_3 & -x'_3y_3 & -x'_3 \\0 & 0 & 0 & x_3 & y_3 & 1 & -y'_3x_3 & -y'_3y_3 & -y'_3 \\x_4 & y_4 & 1 & 0 & 0 & 0 & -x'_4x_4 & -x'_4y_4 & -x'_4 \\0 & 0 & 0 & x_4 & y_4 & 1 & -y'_4x_4 & -y'_4y_4 & -y'_4\end{pmatrix} \mathbf{h} = 0\tag{2.15}$$

In der Gleichung 2.15 ist \mathbf{h} als Spaltenvektor enthalten und entspricht dem Nullraum oder Kern von Matrix \mathbf{A} . Mittels angesprochener Singulärwertzerlegung (SVD)³ kann der Nullraum einer Matrix berechnet werden. Die gesuchte Homographiematrix \mathbf{H} ist somit bestimmt.

³Trucco, Verri [TV98], Appendix A.6 S.322ff

Kapitel 3

Eigener Ansatz

Nachdem die Fluchtpunktfindung im vorherigen Kapitel allgemein beschrieben wurde, bezieht sich dieses Kapitel speziell auf die durch diese Arbeit entstandenen Erweiterungen des Projekts Ornamente. Zu Beginn wird dargelegt, warum in diesem Zusammenhang die Entzerrung von so großer Bedeutung ist. Ferner werden Problemstellungen der Implementation, angestrebte Ziele und ihre Lösungsansätze erklärt.

3.1 Warum ist die Entzerrung so wichtig?

Das Projekt Ornamente zielt auf die Kategorisierung von historischen Gebäuden nach Herkunft, Stil und Zeitepoche ab. Die in den Abbildungen der Datenbank zu sehenden Ornamente werden hierfür herangezogen. Diese müssen gefunden und klassifiziert werden, um so den gesuchten Klassifikationsvektor c zu gewinnen. Die Abbildungen sind jedoch in keiner Weise für die Analyse nach Ornamenten vorbereitet.

Nach Palmowska [Pal04] ist ein "Ornament ein einzelnes, zweidimensionales Verzierungs-motiv, welches zusammen mit anderen Ornamenten als Vorlage für eine Dekoration dient". Es muss klar zwischen der Dekoration, die aus vielen sich wiederholenden Ornamenten besteht, und dem Ornament an sich als grundlegendes Muster der Dekoration, unterschieden werden. Bild 3.1 zeigt links das Ornament namens Vierpass und rechts eine

Dekoration, die Vierpässe enthält.

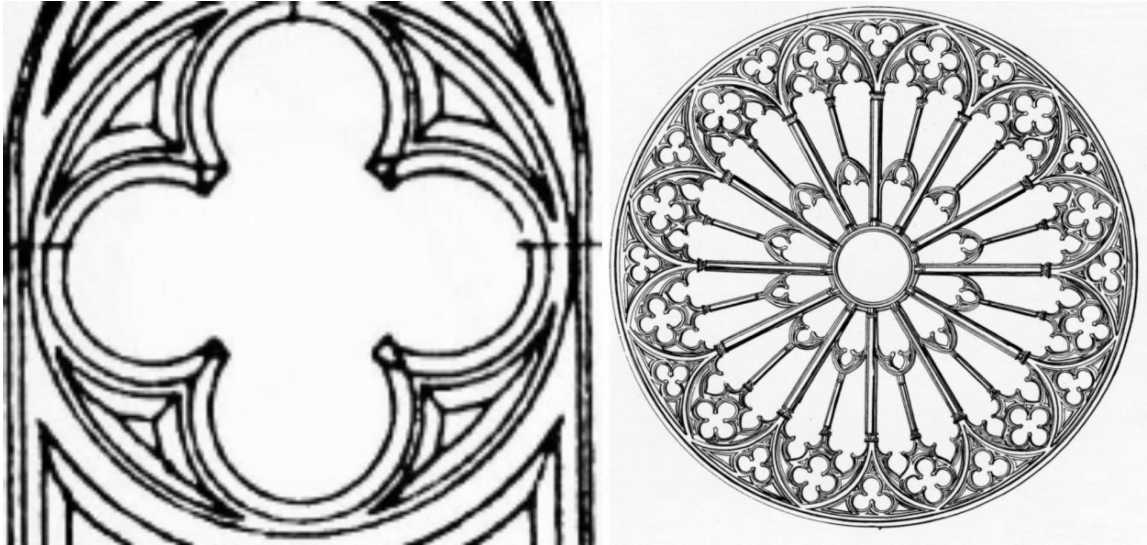


Bild 3.1: *links*: Ornament Vierpass *rechts*: Dekoration mit Vierpassen (*Quelle: Bilddatenbank Ornamente*)

Während Ornamente in beinahe allen Dekorationen von Vasen, Säulen und anderen Kunstobjekten zu finden sind, bezieht sich unsere Betrachtung auf Ornamente von kunsthistorischen Gebäuden. Diese sind meist oberhalb von Fenstern und Türen oder entlang von Säulengängen vorhanden. Das charakteristische Merkmal dieser Dekorationen ist die Wiederholung der darin enthaltenen Ornamente. In Bild 3.2 ist zu sehen, dass die an einer Gebäudefassade angebrachten Ornamente durch die Perspektive verzerrt werden, wie auch die Gebäudefront an sich. Die vertikalen Seitenkanten der Front sind kleiner abgelichtet, je weiter sie von der Kamera entfernt sind. Dadurch verjüngt sich eine Gebäudeseite zu ihrem Fluchtpunkt hin, wie es sich auch für alle daran angebrachten Dekorationen und somit Ornamenten verhält.

Wird ein Bild auf die darin enthaltenen Ornamente hin analysiert, werden bei dem Vergleich mittels NCCF alle vordefinierte Muster der Ornamente, nach denen gesucht werden soll, mit dem Bild verglichen. Im Falle eines Vierpasses wird versucht, den Mittelpunkt und die Größe von allen abgebildeten Vierpassen im Bild zu bestimmen. Gleichzeitig auch

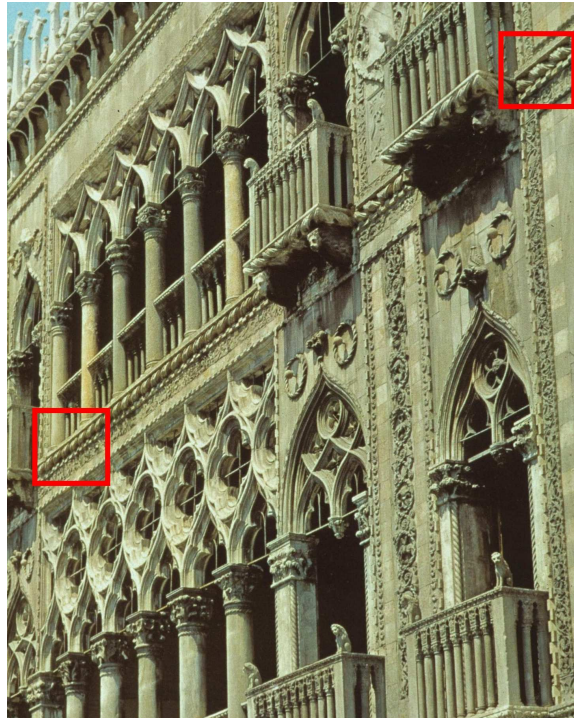


Bild 3.2: Die Ornamente sind gleich der Gebäudefront verzerrt. Je weiter sie entfernt von der Kamera sind, desto kleiner erscheinen sie. *Quelle:* Bilddatenbank Ornamente

ihre Anzahl. Variiert die Größe des Vierpasses werden mehrere unterschiedliche Vierpässe gefunden, die jedoch ein und dem selben Ornament zuzuordnen sind. Das Ergebnis der Anzahl ist somit fehlerhaft, wenn versucht wird, Ornamente in einem verzerrten Bild zu finden. Sind die Vierpässe nach der Entzerrung in gleicher Größe gezeigt, ist diese Fehlerquelle beseitigt.

Nicht nur die Größe, sondern auch die Abstände der Mittelpunkte sind im verzerrten Bild unterschiedlich. In der Realität sind die Abstände jedoch gleich groß. Um nicht das gesamte Bild untersuchen zu müssen, zielt Steneberg [Ste06] darauf ab, das Bild in ornamentale Bereiche einzuteilen. Diese für die Analyse interessanteren Bereiche werden "Region of Interest" (*kurz:* ROI) genannt. Die Suche nach Ornamenten wird auf die ROIs beschränkt, was die Performanz des Gesamtsystems erheblich erhöht. Steneberg bestimmt die ROIs

mit Hilfe des von Dickscheid implementierten SIFT-Operators [Dic05], der für gleiche Ornamente gleiche Merkmale extrahiert. Über diese Merkmale werden die Mittelpunkte der Ornamente gefunden und deren Abstände zueinander gemessen. Wiederholen sich diese Abstände mehrmals in einer Bildzeile, ist dies ein Indiz für einen ornamentalen Bereich. Für die Bestimmung der ROI ist somit von entscheidender Rolle, dass die Abstände in der Abbildung des Gebäudes den Abständen in der Realität entsprechen.

Die Gebäudefront muss somit zur Analyse orthogonal projiziert vorliegen. Wie in Kapitel 2 beschrieben, müssen hierfür die Fluchtpunkte gefunden werden, wodurch die vorliegende Perspektive beschrieben wird. So können die Referenzpunkte bestimmt werden, die der Homographiematrix H zu Grunde liegen. Mit dieser Matrix H wird das gesamte Bild transformiert, was dazu führt, dass die Ebene, die durch die Referenzpunkte definiert wurde, entzerrt vorliegt. Die Gebäudefront in dieser Ebene ist daraufhin orthogonal projiziert dargestellt. In der entzerrten Gebäudefront finden sich jedoch der Realität entsprechend die selben geometrische Verhältnisse. So liegt die in der Realität rechteckige Front auch in ihrer entzerrten Abbildung rechteckig vor. Die Abstände der Mittelpunkte entsprechen denen des Originals in der realen Welt. Die ROIs können dadurch bestimmt und das Bild weiter analysiert werden.

3.2 Problemstellung und Zielsetzungen

3.2.1 Umstrukturierung des Vorhandenen

Die Hauptaufgabe zur Vorbereitung des Bildes auf die ROI-Bestimmung ist die Entzerrung, die von Thun [Thu05] bereits realisiert wurde. Für die Entzerrung arbeiten die implementierten Module `PerspCorr` und `HoughTransform` eng verbunden. Einstiegspunkt in den Prozess ist die Klasse `PerspCorr`, die mit dem ausführbaren Programm `undisPersp` angesprochen wird. Das Eingabebild kann diesem Prozess auf zwei Arten übergeben werden. Einerseits mit bereits vordefinierten Referenzpunkten.

```
perspcorr Input.ppm Output.ppm -x1 -y1 -x2 -y2 -x3 -y3 -x4 -y4
```

Hierbei ist der Input das verzerrte und der Output das entzerrte Bild. Die Koordinaten der Referenzpunkte müssen als Integerwerte und in der Reihenfolge

1. Oben Links	2. Oben Rechts
4. Unten Links	3. Unten Rechts

angegeben sein. Andererseits kann der Prozess mit

```
perspcorr Input.ppm Output.ppm
```

ohne Angabe der Referenzpunkte gestartet werden.

Sind keine Referenzpunkte mit angegeben oder Referenzpunkten mit falschem Umlaufsinn, ruft `PerspCorr` automatisch `HoughTransform` auf. Dieses Modul ermittelt die Fluchtpunkte über eine Histogrammanalyse der Kantenrichtungen mit anschließender Standard Hough Transformation und bestimmt die daraus resultierenden Referenzpunkte. Diese werden an `PerspCorr` zurückgegeben, die Homographiematrix H berechnet und umgehend auf das Eingabebild angewandt. Das Bild wird zur Weiterverarbeitung entzerrt gespeichert.

Thun implementierte Algorithmen zur Flucht- und Referenzpunktbestimmung, zur Berechnung der Homographiematrix und liefert das gesamte Rahmenwerk der Entzerrung. Allerdings ist es nicht möglich, das Modul `HoughTransform`, das Fluchtpunkte detektiert, für sich alleine aufzurufen. Der Zugriff auf gefundenen Fluchtpunkte ist nicht möglich. In diese Studienarbeit soll genau dieser Teilprozess behandelt und erweitert werden. Um die bereits bestehenden Algorithmen der Entzerrung aus `PerspCorr` weiterhin zu nutzen, bedarf es der Umstrukturierung der bestehenden Programme. Der Vorgang der Entzerrung muß separat von der Fluchtpunktfindung gestartet werden können. Die gefundenen Fluchtpunkte sollen bei der Initiierung der neuen Klasse `PerspUndistortion` übergeben werden.

Bild 3.3 und Bild 3.4 zeigen den Programmablauf vor und nach der Umstrukturierung.

Die Vorgänge zur Fluchtpunkt detektion, die zuvor `HoughTransform` durchführte, werden jetzt von der Klasse `VanishingPointDet` übernommen. Die Bestimmung der Referenzpunkte, die Erstellung der Homographiematrix H und das Speichern des entzerrten

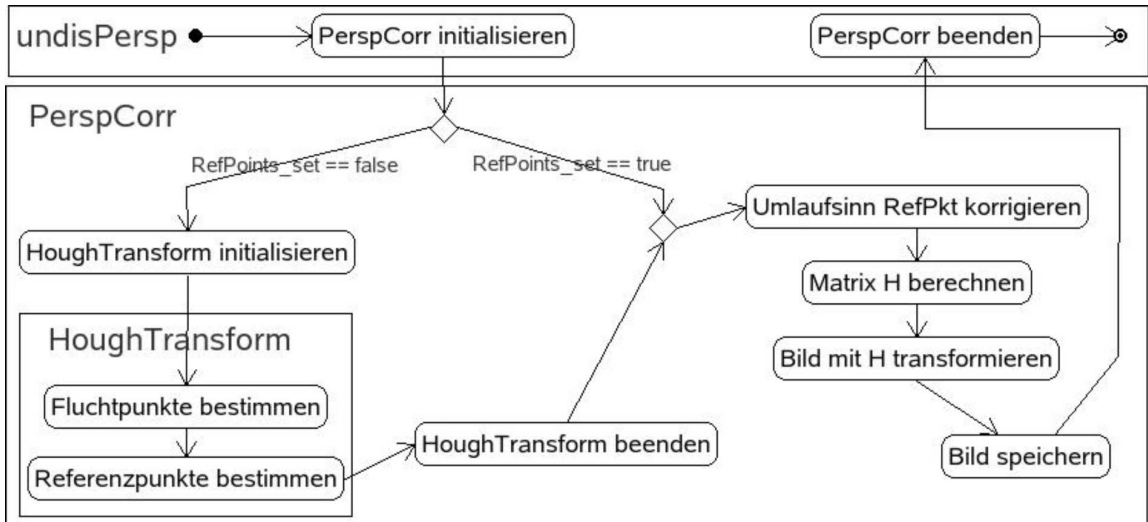


Bild 3.3: Vor Umstrukturierung: HoughTransform als Teil von PerspCorr

Bildes vollzieht sich in PerspUndistortion. Hierbei wurden die Algorithmen aus PerspCorr übernommen und lediglich zur Trennung der beiden Prozesse minimal angepasst.

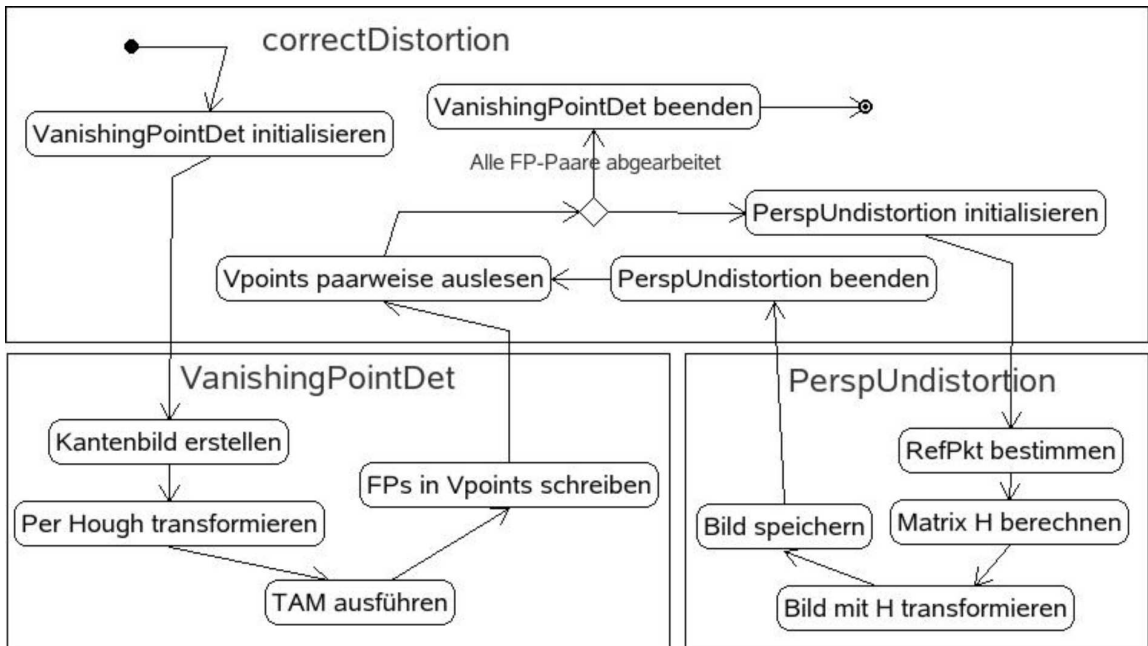


Bild 3.4: Nach Umstrukturierung: VanishingPointDet getrennt von PerspUndistortion

Wird `correctDistortion` eingesetzt, um ein Bild zu entzerren, werden zuerst mehrere Fluchtpunkte von `VanishingPointDet` ermittelt und in einem klasseninternen Array `Vpoints` gespeichert. Die public-Methode `getVpoint(int n)` ermöglicht das Auslesen eines gefundenen Fluchtpunktes an Stelle n . Paarweise werden nun jeweils zwei Fluchtpunkte ausgelesen und während der Initialisierung an `PerspUndistortion` übergeben.

Zu dieser Umstrukturierung kam es, da als Hauptziel der Studienarbeit das Entzerren von zwei oder mehreren Gebäudeseiten genannt wurde. Während `HoughTransform` lediglich zwei Fluchtpunkte detektiert und so maximal eine Gebäudefront entzerren kann, bietet `VanishingPointDet` die Möglichkeit, mehrere Fluchtpunkte zu verarbeiten, um mit ihrer Hilfe weitere Gebäudefronten zu bestimmen und zu entzerren.

Darüber hinaus können durch die Trennung der beiden Prozesse die detektierten Fluchtpunkte ausgelesen und zu jedem Bild der Datenbank abgespeichert werden. Sollte ein Bild ein wiederholtes Mal analysiert werden, erübrigt sich die Suche nach den Fluchtpunkten.

Der entscheidende Vorteil der Trennung liegt jedoch darin, dass das Verfahren zur Fluchtpunktfindung jederzeit austauschbar ist, ohne dass dies den Prozess der Entzerrung beeinflusst.

3.2.2 "Level of Detail"

"Level of Detail" (*kurz:* LOD) bezeichnet den Lösungsansatz, um mehrere Gebäudefronten automatisch zu entzerren. Thun realisierte die Entzerrung der größten Gebäudeseite, die daraufhin nach Ornamenten analysiert werden kann. In diese Analyse sollen Dekorationen mit einbezogen werden, die auf kleineren Gebäudefronten oder Nebengebäuden angebracht sind. Hierfür müssen auch diese Ebenen bestimmt und entzerrt werden.

Entgegen dem Titel dieser Studienarbeit stellt LOD kein eigenständiges Verfahren zur Fluchtpunktschätzung dar. Vielmehr ist es eine selbst entwickelte Erweiterung des bestehenden Verfahrens der Entzerrung. Sie zielt darauf ab, die Hauptaufgabe Entzerrung von mehreren Gebäudeseiten umzusetzen. Im Folgenden werden die Probleme, die es dabei zu bewältigen gilt, erklärt und die Idee der LOD-Variante erläutert.

Eine Gebäudefront kann durch das bestehende Verfahren von Thun bereits entzerrt werden. Sollen nun weitere Fronten entzerrt werden, liegt es nahe, das existierende Verfahren genauer zu betrachten. Es wird deutlich gemacht, warum die LOD-Erweiterung gewählt wurde und nicht ein weiteres Verfahren zur Fluchtpunktfindung umgesetzt wurde.

Grundlegend gilt, dass die Ebene, die entzerrt werden soll, über ihre Fluchtpunkte bestimmt wird. Als Verfahren zur Fluchtpunktfindung kam bisher die TAM zum Einsatz, die in Kapitel 2 beschrieben wurde. Hierfür sind lediglich mehr als zwei konvergierende Geraden von Nöten, um rechnerisch ihren Fluchtpunkt zu bestimmen. Dieser Fluchtpunkt wird der Ebene zugeordnet, in der die Geraden liegen. Allerdings kann keine Aussage darüber gemacht werden, ob die Geraden auch wirklich innerhalb der betrachteten Ebene liegen. Entscheidend für ein korrektes Ergebnis der Fluchtpunktfindung bezüglich dieser Ebene ist somit nicht das Verfahren an sich, sondern die Wahl der richtigen Geraden.

Für die Entzerrung einer zweidimensionalen Ebene ist ein Fluchtpunkt pro Dimension nötig. Sind diese für beide Dimensionen bekannt, werden die Referenzpunkte der Ebene bestimmt. Hierfür werden zwei Geraden durch den Mittelpunkt des Bildes und jeweils einen Fluchtpunkt gelegt und in den Fluchtpunkten um $\pm 10^\circ$ rotiert. Die Schnittpunkte dieser rotierten Geraden liegen in derselben Ebene, wie die Geraden, die auf den Fluchtpunkt zulaufen. Diese Punkte werden als Referenzpunkte der Ebene genutzt.

Zu jeder verzerrten Ebene wird die Homographiematrix berechnet und das gesamte Bild damit transformiert. Die entsprechenden Referenzpunkte, über die die Matrix bestimmt wird, müssen für jede dieser Ebene neu gewählt werden. Da dies wiederum über ihre beide Fluchtpunkte geschieht, können anhand der Referenzpunkte nicht entschieden werden, ob damit die richtigen Ebene im Bild bestimmt wurde. Hierfür müssen wieder die den Fluchtpunkten zu Grunde liegenden Geraden betrachtet werden.

Für die Gesamtanalyse eines Bildes und Einkategorisierung des Abgelichteten ist entscheidend, dass überhaupt Ornamente detektiert werden. Es ist ausreichend, wenn nur eine Gebäudefront entzerrt vorliegt. In diesem Fall wird angenommen, dass das abgebildete Gebäude den größten Bildbereich einnimmt und eine Front sehr groß zu sehen ist. Auf diese Weise können die Kanten dieser Front mit Hilfe einer Histogrammanalyse der Kantenrichtungen erfolgen.

Das Histogramm wird erstellt, indem alle Pixel des Kantenbildes mit gleicher Richtung gezählt werden. Dies erfolgt von 1° bis 180° , da nicht zu entscheiden ist, ob ein Kante mit 90° nach oben oder mit -90° nach unten zeigt. Der Fluchtpunkt auf den sie zuläuft ist der selbe. Die beiden Maxima des Histogramms geben die Hauptrichtungen α_1 und α_2 im Bild an. Diese beziehen sich auf die Kanten der dominantesten Gebäudeseite, da sie den größten Bildbereich einnimmt. Ihre Kanten sind am häufigsten im Bild zu finden und ergeben ein Maximum im Histogramm. Das Bild wird anschließend per Standard Hough transformiert. Der Houghraum wird über die Kantenrichtung α und dem Abstand d der Kante zum Ursprung erstellt. In einer Spalte α sind alle Kanten mit gleicher Richtung eingetragen.

Für jede Hauptrichtung α_1 und α_2 wird eine Kandidatenlisten K_i erstellt. In der Spalte α sind alle Kanten dieser Richtung eingetragen. Die Spalten im Bereich $(\alpha_i - 10, \alpha_i + 10)$ werden ausgelesen und in die Kandidatenliste K_i aufgenommen. K_i enthält alle konvergierenden Kanten, die auf einen gemeinsamen Fluchtpunkt v_i hinzeigen. Übergibt man K_i der TAM wird v_i bestimmt.

Für die Fluchtpunkte sind die Kanten K_1 und K_2 entscheidend und nicht das Verfahren, mit dessen Hilfe die Fluchtpunkte bestimmt werden. Werden die Kanten anhand der Histogrammanalyse und darauffolgender Houghtransformation interpretiert, können wie im vorhandenen Verfahren nur die am häufigsten vorkommenden Kanten betrachtet werden. Es liegt keine Information über den Zusammenhang der detektierten Kanten und denen im Bild tatsächlich vorhandenen Ebenen vor. Die Ebene eine kleiner Gebäudefront kann durch der Betrachtung des Histogramms nicht eindeutig bestimmt werden.

Im Histogramm der Kantenstärke wird die Anzahl der Pixel mit gleicher Kantenrichtung festgehalten und im Houghraum zusätzlich ihre Kantenstärke aufsummiert. Daraus ergeben sich zwei Hauptprobleme:

1. Eine kleine Gebäudefront kann nicht aus dem Histogramm ausgelesen werden. Die beiden Hauptrichtungen entsprechen dem globalen Maximum und dem größten lokalen Maximum. Die Anzahl der Kanten einer kleineren Front ist geringer als die Anzahl der Kanten der dominanten Gebäudefront. Die Informationen der Kleineren werden von denen der Größeren überlagert.

2. Die Ebene der größten Front kann nicht bestimmt werden, sollten harte Kanten im Bild zu sehen sein, die nicht dieser Front angehören. Die Kantenstärken werden im Houghraum mit aussummiert. Wieder werden die Informationen die entscheidenden Informationen zur Gebäudefront von denen der unbedeutenden Kante überlagert.

Um diese beiden Hauptproblemen zu umgehen, kann der zu entzerrende Bereich des Bildes von Hand eingegrenzt werden. Diese Funktion wurde bereits im Webinterface¹ zum Gesamtsystem des Ornamente Projekts zur Verfügung gestellt. Auf diese Weise können die Bereiche mit störenden harten Kanten, die nicht Teil der zu analysierenden Gebäudefront sind, ausgegrenzt werden. Ferner können kleiner Gebäudefronten gezielt noch vor der Entzerrung zur Analyse ausgewählt werden.

Diese Lösung bedarf allerdings die Steuerung von Hand für jedes Bild, das analysiert werden soll. Die Analyse von mehreren Bilder wird dadurch aufwendig und mühsam. Eine automatisierte Analyse ohne Eingriff des Benutzers ist anzustreben.

Mit dem existierenden Verfahren kann keine Aussage über die Anzahl der im Bild zu sehenden Gebäudefronten oder deren Lage gemacht werden. Alleine die detektierten Kanten stehen für die Entzerrung zu Verfügung. Eine besser an das Problem "Entzerrung von mehreren Gebäudefronten" angepasste Linienextraktion könnte dies erreichen. Das Problem kann jedoch nicht durch den Einsatz eines anderen Verfahrens zur Fluchtpunktfindung gelöst werden.

Aus diesem Grund wurde davon abgesehen, die KHT zu implementieren. Es entstand die LOD-Variante. Ihr Grundgedanke besteht darin, das Histogramm der Kantenrichtungen detaillierter zu betrachten, um so mehrere Maxima zu bestimmen. Über diese können weitere Fluchtpunkte ermittelt werden, die im günstigsten Fall nichtdominante Gebäudefronten beschreiben. Diese können daraufhin mit den Algorithmen des bereits implementierten Entzerrungsverfahrens entzerrt werden.

Bisher wurde im Histogramm das globale Maximum und Minimum, daraufhin zwischen diesen beiden das lokale Maximum gesucht. Globales und lokales Maximum ergaben die

¹Webinterface des Ornamente Projekts

happy.uni-koblenz.de/puma/ornamente/experiment/

beiden Hauptrichtungen. In die Kandidatenlisten wurden um $\pm 10^\circ$ von der Hauptrichtung abweichende Kanten übernommen.

Die LOD unterteilt das Histogramm in mehrere Teilbereiche. Ihre Größe muss so gewählt sein, dass sie addiert 180° ergeben, anderenfalls bricht das Programm ab. Abhängig von dieser Größe entstehen n Teilbereiche. In jedem wird das globale Maximum bestimmt und in einem Array gespeichert. Dieser wird abhängig von der zu jedem gefundenen Maximum abgespeicherten Kantenstärke sortiert.

Alle Kanten mit der den Maximum entsprechenden Orientierung werden wieder in eine Kandidatenliste geschrieben. Zusätzlich alle Kanten, die von dieser Orientierung um $\pm 3^\circ$ abweichen. Zu jeder Liste wird ein Fluchtpunkt ermittelt. Hierfür werden die Kandidatenlisten der TAM übergeben.

So werden n -viele Fluchtpunkte durch die Klasse `VanishingPointDet` ermittelt. Um weiterhin den Zugriff auf diese Fluchtpunkte zu gewährleisten, werden sie im klasseninternen Array `Vpoints` abgespeichert, der durch die public-Methode `getVpoint(int n)` ausgelesen werden kann.

Es wird deutlich, dass die Umstrukturierung des bestehenden Verfahrens bereits auf die Anforderungen der LOD-Erweiterung angepasst wurde. Zwei der durch die LOD bestimmten Fluchtpunkte entsprechen denen der dominantesten Gebäudeseite. Die LOD liefert somit die selben Ergebnisse, wie es auch das alte Verfahren tat. Es werden jedoch mehr Fluchtpunkte gefunden, abhängig von Aufteilung des Histogramms in die Teilbereiche und somit der Detailstufe.

Kapitel 4

Experimente und Ergebnisse

4.1 Erläuterung des Tests

In diesem Kapitel wird das Erarbeitete anhand von Bildern, die der Bilddatenbank entnommen sind, getestet. Zusätzlich wird eine eingescannte Grafik auf die korrekte Detektion der Fluchtpunkte in einem künstlichen Bild überprüft. Diese Grafik wurde manuell freigestellt. Die restlichen Testbilder wurden randentfernt.

Für diesen Test wurden alle Bilder auf die gleiche Größe von 945 Pixel in der Breite bei einer Auflösung von 150dpi skaliert.

Zu den Testbildern liegen keine Informationen über die tatsächlichen Fluchtpunkte vor. Ein Vergleich zwischen diesen und den Fluchtpunkten, die von `VanishingPointDet` berechnet werden, kann nicht erfolgen.

Aus diesem Grund sind in den folgenden Abbildungen die vier am stärksten vertretenen Kantenrichtungen eingezeichnet. Die ermittelten Kantenrichtungen sind von entscheidender Bedeutung für die Fluchtpunktschätzung. Sie wurden aus den Teilbereichen des Histogramms als Maximum ausgelesen. Aufgrund dieser Richtungen werden die Fluchtpunkte bestimmt.

Anhand der eingezeichneten Richtungen kann sehr schnell überprüft werden, ob diese Kantenrichtung vermehrt im Bild zu sehen ist, oder eine falsche Kantenrichtung gewählt

wurde. Dieser rein visuelle Test wurde gewählt, da zuverlässige Referenzwerte, weder für Fluchtpunkte noch für entscheidende Kantenrichtungen vorhanden sind.

Zu beachten ist, dass für die korrekte Entzerrung einer Gebäudeseite stets zwei Fluchtpunkte gefunden werden müssen. Somit müssen zwei der eingezeichneten Linien den markanten Kanten im Bild entsprechen, die gleichzeitig einer Gebäudefront zuzuordnen sind.

Ideal wäre, wenn die Neigungen zweier Linien eindeutig den Kantenrichtungen einer Gebäudeseite entsprechen. Gleichzeitig müssen die beiden anderen eingezeichneten Richtungen aufgrund der Kanten einer weiteren Front entstanden sein. In diesem Fall wäre das entzerren von mehr als nur einer Gebäudefront realisiert.

Die Klasse `VanishingPointDet` wurde für diesen Test direkt angesprochen, und die gefundenen Fluchtpunkte bestimmt. Neben den Koordinaten dieser Punkte sind gleichzeitig die Werte der in den Histogrammen gefundenen Maxima aufgelistet.

4.2 Ergebnisse und Verifikation

Wie in Bild 4.1 zu sehen ist, wird eine Kantenrichtung korrekt ausgewählt. Eine ist jedoch für die Entzerrung nicht ausreichend. Das entzerrte Bild überstieg die Größenbeschränkung von 20 000 Pixel und die Entzerrung brach ab. Trotz der vier gefundenen Fluchtpunkte konnte keine Ebene eindeutig bestimmt werden.

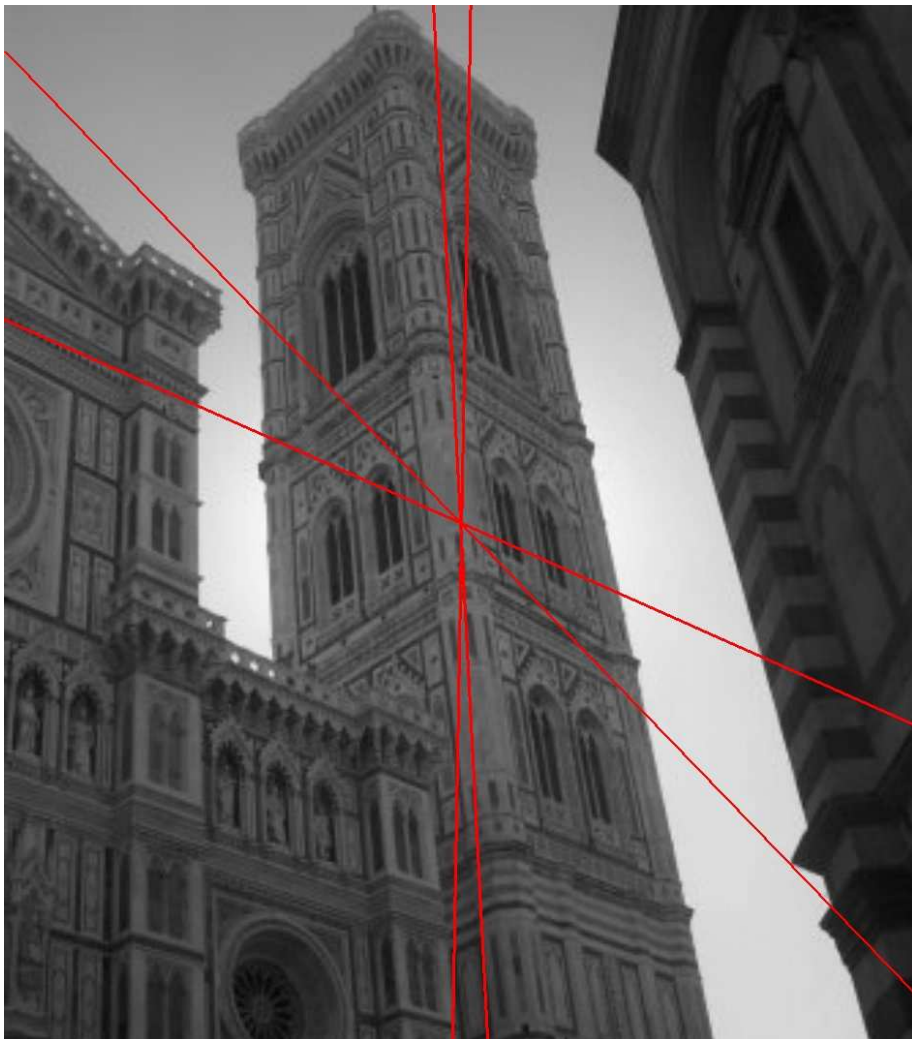


Bild 4.1: Ein Kantenrichtung wurde richtig detektiert

Nr.	MaxRichtung	MaxWert	Fluchtpunkt
1	177	1219	(-16246/14783)
2	1	540	(22867/963)
3	136	41	(14/143)
4	114	16	(-979/-1955)

Tabelle 4.1: Maxima und Fluchtpunkte zu Bild 4.1

In Bild 4.2 wurde nicht eine Hauptrichtung korrekt gewählt. Die Entzerrung anhand der gefundenen Fluchtpunkte führte zu keinem für die Weiterverarbeitung brauchbaren Ergebnis.



Bild 4.2: Keine entscheidende Kantenrichtung erkannt

Nr.	MaxRichtung	MaxWert	Fluchtpunkt
1	5	2597	(685/380)
2	179	1507	(1869/327)
3	83	726	(-18803/-23766)
4	32	456	(7632/-4326)

Tabelle 4.2: Maxima und Fluchtpunkte zu Bild 4.2

Bild 4.3 zeigt das eingescannte Werbeplakat des *Kompetenzzentrums für Gestaltung, Fertigung und Kommunikation der Handwerkskammer Koblenz*. Diese künstlich erstellte Abbildung beinhaltet weniger Kanten, dafür größere und einheitlichere Farbflächen. Darüber hinaus wurde der Hintergrund komplett entfernt.

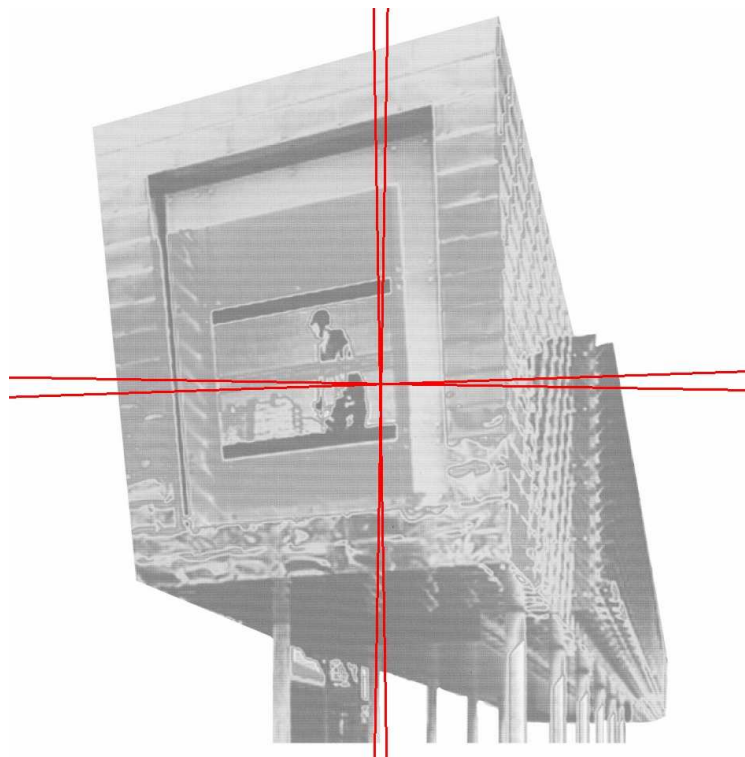


Bild 4.3: Falsche Kantenrichtungen

Nr.	MaxRichtung	MaxWert	Fluchtpunkt
1	88	244	(404/1924)
2	91	243	(388/2790)
3	1	191	(689/901)
4	179	187	(616/902)

Tabelle 4.3: Maxima und Fluchtpunkte zu Bild 4.3

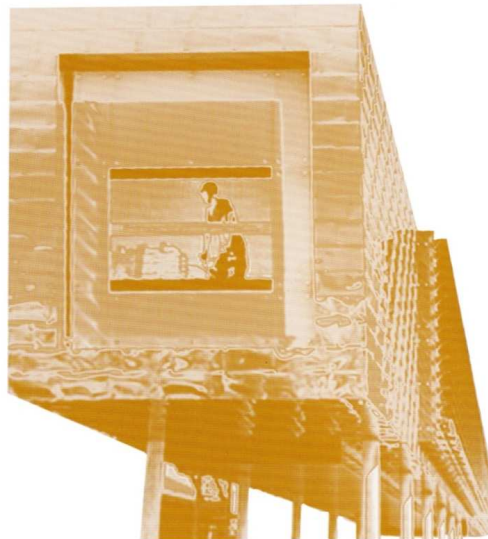


Bild 4.4: Entzerrte Darstellung von Hand generiert

Störende Kanten sollten dabei eliminiert werden, um die Kanten entlang der Gebäude-seiten klarer hervorzuheben. Dennoch konnte das Programm diese nicht bestimmen. Wie Bild 4.3 entzerrt hätte aussehen sollen, zeigt Bild 4.4.

Kapitel 5

Zusammenfassung

Für diese Studienarbeit können zwei Schwerpunkte genannt werden. Einerseits sollten verschiedene Verfahren zur Fluchtpunktschätzung aus Wissenschaft und Forschung eingänglich untersucht und erörtert werden. Dies im Hinblick auf ein detaillierteres Analyseverfahren, das die Möglichkeit bietet, mehrere Gebäudeseiten automatisiert entzerren zu können. Andererseits sollten sich die gewünschten Verbesserungen in das bereits vorhandene Gesamtsystem des Projekts Ornamente eingliedern, um so das Endergebnis der Klassifizierung von Ornamenten zu verbessern.

Daraus entstanden die in Kapitel 1 genannten Hauptaufgaben. Neben dem TAM-Verfahren, das im vorhandenen Teilprozess der Entzerrung bereits zum Einsatz kam, wurde in Kapitel 2 das Verfahren KHT nach Tuytelaars beschrieben. Ansätze der KHT waren im Bestehenden zu erkennen, wie sich während der anfänglichen Einarbeitung in das Themengebiet Fluchtpunktfindung und dem Gesamtsystem der Ornamentklassifizierung herausstellte. Allerdings waren einige Aspekte, wie sie von Tuytelaars et al. in [TGPM98] zur KHT beschrieben sind, nicht enthalten.

Der erste Lösungsansatz zur Entzerrung von mehreren Gebäudeseiten bestand darin, die KHT unabhängig von allen Prozessen des Gesamtsystems zu implementieren, um so die Genauigkeit der Fluchtpunkt-detektion zu erhöhen. Mit dieser detaillierteren Fluchtpunktfindung sollte das bereits bestehende Modul der Entzerrung zu besseren Ergebnissen führen.

Um die Entzerrung für sich alleine nutzen zu können, musste sie vorerst von der vorhandenen Fluchtpunktschätzung isoliert werden. Während der in Kapitel 3 beschriebenen Umstrukturierung und Trennung der beiden Prozesse wurde das eigentliche Problem der Verarbeitung von mehreren Gebäudeseiten erkannt.

Nicht die Fluchtpunkte und die Verfahren für ihre Detektion sind ausschlaggebend, weitere Ebenen im Bild erkennen zu können. Vielmehr verhindert dies der fehlende Rückschluss von extrahierten Kanten auf die Lage, Größe und Anzahl der im Bild vorhandenen Gebäudeseiten.

Wären hierzu Informationen bekannt, könnten, wie auch für ornamentale Bereiche, ROIs festgelegt werden, die mit einer hohen Wahrscheinlichkeit eine abgegrenzte Gebäudeseite beinhalten. Um diese daraufhin zu entzerren, kann das jetzt isolierte Programm zur Entzerrung genutzt werden.

Die KHT umzusetzen, wurde als Lösungsweg verworfen und der eigene Lösungsansatz "Level of Detail" aus Kapitel 3 wurde entwickelt. Die entstandenen Programme wurden wie gefordert in PUMA, der "Programmierungsumgebung für die Musteranalyse" eingebunden.

Wie die Test aus Kapitel 4 jedoch zeigen, konnte damit keine Verbesserung erzielt werden.

Literaturverzeichnis

- [ADV03] Andres Almansa, Agnes Desolneux, and Sebastien Vamech. Vanishing point detection without any a priori information. *IEEE Trans. on PAMI*, page 502 507, 2003.
- [BBV00] Christian Bräuer-Burchardt and Klaus Voss. Robust vanishing point determination in noisy images. *International Conference on Pattern Recognition*, page 1559 1562, 2000.
- [Den05] Matthias Dennhardt. Symmetriesuche in ornamentbildern. Studienarbeit, Universität Koblenz-Landau, 2005.
- [Dic05] Timo Dickscheid. Automatische referenzpunktverfeinerung in panoramabildern mittels sift-operator. Studienarbeit, Universität Koblenz-Landau, 9 2005.
- [HA62] P. V. C. Hough and A. Arbor. Method and means for recognizing complex patterns. Technical Report US Patent 3069 654, US Patent, 1962.
- [LZ98] D. Liebowitz and Andrew Zisserman. Metric rectification for perspective images of planes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 482–488, 6 1998. Liebowitz98.
- [Pal04] Marta Palmowska. Bildanalyse von ornamentik. Studienarbeit, Universität Koblenz-Landau, 7 2004.
- [PH03] Dietrich Paulus and Joachim Hornegger. *Applied pattern recognition: A practical introduction to image and speech processing in C++*. Advanced Studies in Computer Science. Vieweg, Braunschweig, 4 edition, 2003.

- [Rot00] Carsten Rother. A new approach for vanishing point detection in architectural environments. In *British Machine Vision Conference (BMVC)*, pages 382–391, 2000.
- [Ste06] Stephan Steneberg. Detektion ornamentaler bereiche in architektonischen bildern. Studienarbeit, Universität Koblenz-Landau, Campus Koblenz, Fachbereich 4 Informatik, Institut für Computervisualistik, 6 2006.
- [TGPM98] Tinne Tuytelaars, Luc J. Van Gool, Marc Proesmans, and Theo Moons. A cascaded hough transform as an aid in aerial image interpretation. In *ICCV*, pages 67–72, 1998.
- [Thu05] Andreas Thun. *Perspektivische Entzerrung gescannter Dia-Positive*. Diplomarbeit, Universität Koblenz-Landau, 3 2005.
- [TV98] E. Trucco and A. Verri. *Introductory Techniques for 3-D Computer VISION*. Prentice Hall, 1998.