

Markus Volkmer

Erkennung mehrerer Objekte durch parallele Suchagenten

Diplomarbeit

Bibliografische Information der Deutschen Nationalbibliothek:

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Dieses Werk sowie alle darin enthaltenen einzelnen Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsschutz zugelassen ist, bedarf der vorherigen Zustimmung des Verlanges. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen, Auswertungen durch Datenbanken und für die Einspeicherung und Verarbeitung in elektronische Systeme. Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen Wiedergabe (einschließlich Mikrokopie) sowie der Auswertung durch Datenbanken oder ähnliche Einrichtungen, vorbehalten.

Copyright © 1998 Diplom.de
ISBN: 9783832428303

Markus Volkmer

Erkennung mehrerer Objekte durch parallele Suchagenten

Markus Volkmer

Erkennung mehrerer Objekte durch parallele Suchagenten

Diplomarbeit
an der Universität Bielefeld
Dezember 1998 Abgabe



Diplomarbeiten Agentur
Dipl. Kfm. Dipl. Hdl. Björn Bedey
Dipl. Wi.-Ing. Martin Haschke
und Guido Meyer GbR

Hermannstal 119 k
22119 Hamburg

agentur@diplom.de
www.diplom.de

ID 2830

Volkmer, Markus: Erkennung mehrerer Objekte durch parallele Suchagenten /

Markus Volkmer - Hamburg: Diplomarbeiten Agentur, 2000

Zugl.: Bielefeld, Universität, Diplom, 1998

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtes.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, daß solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Die Informationen in diesem Werk wurden mit Sorgfalt erarbeitet. Dennoch können Fehler nicht vollständig ausgeschlossen werden, und die Diplomarbeiten Agentur, die Autoren oder Übersetzer übernehmen keine juristische Verantwortung oder irgendeine Haftung für evtl. verbliebene fehlerhafte Angaben und deren Folgen.

Dipl. Kfm. Dipl. Hdl. Björn Bedey, Dipl. Wi.-Ing. Martin Haschke & Guido Meyer GbR
Diplomarbeiten Agentur, <http://www.diplom.de>, Hamburg 2000

Printed in Germany



Diplomarbeiten Agentur

Wissensquellen gewinnbringend nutzen

Qualität, Praxisrelevanz und Aktualität zeichnen unsere Studien aus. Wir bieten Ihnen im Auftrag unserer Autorinnen und Autoren Wirtschaftsstudien und wissenschaftliche Abschlussarbeiten – Dissertationen, Diplomarbeiten, Magisterarbeiten, Staatsexamensarbeiten und Studienarbeiten zum Kauf. Sie wurden an deutschen Universitäten, Fachhochschulen, Akademien oder vergleichbaren Institutionen der Europäischen Union geschrieben. Der Notendurchschnitt liegt bei 1,5.

Wettbewerbsvorteile verschaffen – Vergleichen Sie den Preis unserer Studien mit den Honoraren externer Berater. Um dieses Wissen selbst zusammenzutragen, müssten Sie viel Zeit und Geld aufbringen.

<http://www.diplom.de> bietet Ihnen unser vollständiges Lieferprogramm mit mehreren tausend Studien im Internet. Neben dem Online-Katalog und der Online-Suchmaschine für Ihre Recherche steht Ihnen auch eine Online-Bestellfunktion zur Verfügung. Inhaltliche Zusammenfassungen und Inhaltsverzeichnisse zu jeder Studie sind im Internet einsehbar.

Individueller Service – Gerne senden wir Ihnen auch unseren Papierkatalog zu. Bitte fordern Sie Ihr individuelles Exemplar bei uns an. Für Fragen, Anregungen und individuelle Anfragen stehen wir Ihnen gerne zur Verfügung. Wir freuen uns auf eine gute Zusammenarbeit

Ihr Team der *Diplomarbeiten Agentur*

Dipl. Kfm. Dipl. Hdl. Björn Bedey –
Dipl. Wi.-Ing. Martin Haschke —
und Guido Meyer GbR —————

Hermannstal 119 k —————
22119 Hamburg —————

Fon: 040 / 655 99 20 —————
Fax: 040 / 655 99 222 —————

agentur@diplom.de —————
www.diplom.de —————

Inhaltsverzeichnis

1	Einleitung und Problemstellung	1
1.1	Aufbau der Arbeit	2
1.2	Objekterkennung durch Suchagenten – Motivation	3
1.3	Notation	5
2	Lösungsansätze zur Objekterkennung	7
2.1	Geometrie-basierte Verfahren	8
2.1.1	Erkennung durch Invarianten	8
2.2	Appearance-basierte Verfahren	14
2.2.1	Appearance-based Object Recognition und Eigenraumverfahren	14
2.3	Motivation des gewählten Ansatzes: Erkennung durch Template-Matching	22
2.3.1	Geometrie-basiert versus Appearance-basiert	22
2.3.2	Eigenraumverfahren versus Template-Matching	23
3	Template-Matching und Korrelationstechniken	26
3.1	Template-Matching	27
3.1.1	Modellbasierte Templates	27
3.1.2	Bildbasierte Templates	27
3.2	Ähnlichkeits- und Korrelationsmaße	28
3.2.1	Korrelationsebenen	29
3.2.2	Kreuz-Korrelation und abgeleitete Maße	30
4	Simulated Annealing zur Korrelations-Maximierung	37
4.1	Objekterkennung durch Optimierung	38
4.1.1	Lokale Verfahren	39
4.1.2	Globale Verfahren	41
4.2	Simulated Annealing und (stochastische) Relaxationsverfahren	43
4.2.1	Simulated Annealing	44
4.2.2	Annealing versus Evolutionäre (Genetische) Algorithmen	49
4.2.3	Deterministische Relaxation	50

4.2.4	Fast Simulated Annealing mit Temperatur-parametrisiertem Suchraum	54
5	Parallele Suchagenten zur Lösung des Multiple-Objekte-Problems	57
5.1	Verteilte Lösungen und Parallelisierungsstrategien	58
5.1.1	Daten- versus Aufgaben-Parallelisierung	59
5.1.2	Parallelisierungen des Erkennungsalgorithmus	60
5.2	Parallele Suchagenten	66
5.2.1	Agenten, Agentensysteme und Suchagenten	66
5.2.2	Agenten in der Computer Vision	71
5.2.3	Interface-Agent <i>MatchIt</i>	78
5.2.4	Der Suchagent (<i>Visual-Slave</i>)	80
6	Implementierung	83
6.1	Unterstützende Software und Werkzeuge	84
6.1.1	Die Bildverarbeitungsumgebung Vista	84
6.1.2	Die Parallele Virtuelle Maschine (PVM)	85
6.2	Master and Slave(s) – <i>MatchIt</i> und die Suchagenten	87
6.2.1	<i>MatchIt</i>	88
6.2.2	Slave und Visualslave	95
7	Experimentelle Ergebnisse, Diskussion und Bewertung	101
7.1	Erkennung durch einen Suchagenten	102
7.1.1	Erkennung nicht verdeckter Objekte	102
7.1.2	Erkennung eines teilweise verdeckten Objektes	107
7.1.3	Variation der Objektmenge	108
7.2	Parallele Erkennung durch mehrere Suchagenten	109
7.2.1	Parallele Erkennung – Einzelaufgaben	109
7.2.2	Parallele Erkennung – „Team“	112
7.2.3	Parallele Erkennung von Objektteilen (Sub-Templates)	116
7.3	Korrelation als Matching-Kriterium	118
7.4	Experimentelle Ergebnisse – Fazit	120
8	Zusammenfassung und Ausblick	121
A	Literaturverzeichnis	124
B	Abbildungsverzeichnis	152
C	Tabellenverzeichnis	153
D	Algorithmen	154

1

Einleitung und Problemstellung

Im Rahmen der vorliegenden Diplomarbeit wird ein System zur maschinellen Erkennung mehrerer Objekte (basierend auf Intensitätsbildern) erarbeitet, implementiert und analysiert.

Diese Problemstellung erforderte eine selbständige Einarbeitung in die die Arbeit betreffenden Gebiete und deren Einordnung zur Entwicklung eines Systems. Die in Erwägung gezogenen aktuellen Verfahren werden dazu erläutert und vor dem Hintergrund der Problemstellung und weiteren Randbedingungen betrachtet. Als Folge dieser Betrachtungen ergibt sich die Umsetzung eines homogenen Multi-Agenten-Systems. Die implementierten Agenten realisieren parallel Intensitäts-basiertes Template-Matching und Fast Simulated Annealing mit Temperatur-parametrisiertem Suchraum.

Der konzeptionelle Weg von der Betrachtung von Lösungsmöglichkeiten, über die Entscheidungsfindung, die Integration, bis hin zur Implementierung und Bewertung des Systems, ist daher Teil dieser Arbeit und nicht lediglich die Vorstellung eines umgesetzten Verfahrens.

Die der Arbeit zugrundeliegenden theoretischen Prinzipien werden erörtert. Eine Darstellung der implementatorischen Aspekte und die Analyse der Ergebnisse wird durch eine Bewertung des Systems und seiner problemspezifischen Anwendbarkeit, sowie durch Aufzeigen von Modifikationen und möglichen Erweiterungen ergänzt. Zur Entwicklung zählt auch die Installation unterstützender Software sowie die Implementierung einer graphischen Benutzeroberfläche zur Steuerung des Systems, einer Testumgebung für die relevanten (Bild-)Operationen des Verfahrens und einer Online-Visualisierung.

Es folgt zunächst eine kurze inhaltliche Übersicht über die einzelnen Kapitel. Danach wird die *Erkennung mehrerer Objekte durch parallele Suchagenten*, im Vorgriff auf die anschließende Entwicklung dieses Konzeptes, motivierend vorgestellt. Mit der Festlegung der im Rahmen der Arbeit verwendeten Notation endet dieses Kapitel.

1.1 Aufbau der Arbeit

Der Aufbau der Arbeit entspricht dem verfolgten Weg von der Problemdefinition bis zur Umsetzung und Evaluierung der implementierten Lösung.

Zuerst wird in diesem Kapitel in **Abschnitt 1.2** eine Motivation für den verfolgten Suchagenten-Ansatz vor dem Hintergrund der konkreten Problemstellung gegeben.

Nach einem kurzen **Abschnitt (1.3)** zur Festlegung der Notation werden in **Kapitel 2** die in der Recherche nach Lösungsmöglichkeiten erwogenen Ansätze vorgestellt. Der Entscheidungsweg wird skizziert und der favorisierte Ansatz zur Lösung der Aufgabenstellung „Objekterkennung“ motiviert. Dabei kommen die in den folgenden Kapiteln erläuterten Methoden Template-Matching- und Optimierungstechniken zum Einsatz.

Kapitel 3 erörtert das verwendete Konzept des Template Matching und gibt eine Übersicht zu Ähnlichkeits- und Korrelationsmaßen in Bezug auf die verwendeten bildbasierten Templates. Auf deren Anwendungsgebiete und Eigenschaften wird ebenfalls eingegangen.

Das folgende Kapitel (**Kapitel 4**) stellt das zur Korrelations-Maximierung verwendete Annealing-Verfahren vor und geht auf verwandte und abgewandelte Verfahren ein. Dabei werden Optimierungsprinzipien allgemein und im Kontext der *Computer Vision* betrachtet. Die in den Kapiteln 3 und 4 erläuterten Konzepte und Methoden finden direkten Eingang in das im nächsten Kapitel vorgestellte Agentensystem.

Nachdem somit eine Grundlage für die Erkennung von Objekten vorhanden ist, wird in **Kapitel 5** der Suchagenten-Ansatz eingeführt. Dabei werden, ausgehend von der Problematik der Erkennung mehrerer Objekte, grundlegende Konzepte zum Agentenkonzept und konkrete Strategien zur Lösung des vorliegenden Problems dargestellt. Auch auf verteilte Strategien zur Erkennung einzelner Objekte und die Möglichkeiten der Parallelisierung allgemein wird eingegangen, da hier ein entwicklungsgeschichtlicher und implementierungstechnischer Zusammenhang besteht.

Kapitel 6 geht auf implementatorische Details ein. Die benutzten Software-Pakete und Bibliotheken werden kurz erläutert. Danach wird das Programmpaket bestehend aus den Programmen *MatchIt* und *(Visual-)Slave* im Detail erläutert, das die *Erkennung mehrerer Objekte durch parallele Suchagenten* implementiert.

Die experimentellen Ergebnisse werden in **Kapitel 7** dokumentiert. Hierbei liegt der Schwerpunkt auf der Erkennungsleistung des Systems unter verschiedenen experimentellen Bedingungen und nicht auf der durch parallele Implementierung mögliche Beschleunigung (*Speed-Up*) in der Erkennung mehrerer Objekte. Weiterhin erfolgt hier die Diskussion und Bewertung des Systems und der erzielten Ergebnisse.

Die Arbeit schließt in **Kapitel 8** mit einer rückblickenden Zusammenfassung der Diplomarbeit und einem Ausblick auf mögliche zukünftige Erweiterungen und Anwendungen des vorgestellten Verfahrens.

Das **Literaturverzeichnis** im Anhang wird ergänzt durch Verzeichnisse der Abbildungen, Tabellen und Algorithmen.

1.2 Objekterkennung durch Suchagenten – Motivation

Die Agenten-orientierte Programmierung (Agent Oriented Programming) entwickelte sich als Zweig der Verteilten Künstlichen Intelligenz (Distributed Artificial Intelligence). Ausgehend von Strategien zur verteilten Problemlösung (Distributed Problem Solving) hat sie sich heute zu einem allgemeineren Konzept entwickelt. Darin werden neben der reinen Verteilung von (Teil-)Aufgaben auch erweiterte Kommunikationsstrukturen und Interaktion bis hin zur Definition von sozialen Eigenschaften solcher Agenten verfolgt. Obwohl es keine allgemeingültige Definition von Agenten gibt, versteht man sie als unabhängige, autonom agierende Systeme zur Problemlösung in definierter Umgebung [JSG98]. Agenten arbeiten dabei typischerweise auf unvollständiger und unsicherer Information ihre Umwelt betreffend.

Gerade in der Modellierung der für Menschen so selbstverständlichen Vorgänge der Wahrnehmung und deren Verknüpfung mit dem Ziel intelligenten Agierens liegen (quasi natürlicherweise) Möglichkeiten der Modellierung durch verteilte (unabhängige wie abhängige), „intelligente“ und kommunizierende Einheiten – Agenten.

Die „intelligenten“ Eigenschaften dieser Agenten können dabei implementiert sein oder erst durch ihr Zusammenwirken zutage treten (emergentes intelligentes Verhalten).

Im Rahmen dieser Arbeit kann natürlich kein komplexes System aus unterschiedlichsten, interagierenden Agenten mit komplexer Kommunikationsstruktur umgesetzt werden. Vor dem Hintergrund der konkreten Aufgabenstellung „Objekterkennung“ jedoch kann ein einfaches Agentensystem umgesetzt werden, das zwar in sich geschlossen ist, aber dennoch prinzipielle Möglichkeiten des Agentenkonzepts und einfache Kommunikationsstrukturen in sich trägt.

Das System zur Objekterkennung wurde unter Berücksichtigung der anwendungsspezifischen Gegebenheiten im Rahmen des *Baufix-Szenarios* entwickelt.

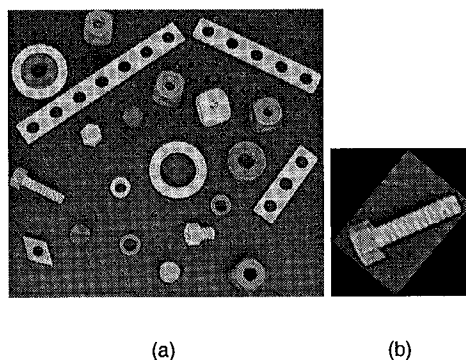


Abbildung 1.1: (a) Objekte aus dem *Baufix-Szenario* und (b) ein Objekt-Template

In diesem Szenario ist von einer begrenzten Objektmenge auszugehen. Die Einzelobjekte

(in Abgrenzung zu Aggregaten) sind oft einfarbig und stellen idealisiert Schrauben, Würfel, Leisten etc. dar. Das System ist jedoch nicht auf dieses Szenario festgelegt und kann prinzipiell beliebige Objekte erkennen.

Die Erkennung erfolgt direkt auf Intensitäts-Einzelbildern (Grauwertbildern) und nutzt keine Farbinformation. Auch auf dem Gradienten eines Bildes (mit bzw. ohne nachfolgende Schwellwertoperation) kann die Erkennung erfolgen. Da keine aufwendige Vorverarbeitung zur Merkmalsgewinnung erfolgt (z.B. Kantendetektion, Generierung von Geraden oder Polygonzügen) und das Verfahren nicht modellbasiert arbeitet, ist eine strikte Differenzierung nach unterschiedlichen Ebenen der Computer Vision („Low-level“, „Mid-level“ und „High-level“) nicht sinnvoll. Innerhalb dieser Ebenen werden Stufen von der (Bild-)Vorverarbeitung bis hin zur Anwendung klassischer KI-Techniken realisiert, um eine Erkennung zu erreichen.

Durch die in dieser Arbeit gewählte Formulierung der Erkennung als Optimierungsproblem und durch die Verwendung einer Intensitäts-basierten Objektdarstellung gehen zumindest die Ebenen „Low-level“ und „Mid-level“ ineinander über.

Weiterhin ist das System zur (parallelen) Erkennung mehrerer Objekte ausgelegt. Der gewählte Suchagenten-Ansatz realisiert aufgabenorientiert paralleles Template-Matching und Korrelationsoptimierung via *Fast Simulated Annealing* und Suchraumkontraktion.

Die einzelnen Suchagenten bekommen ein Template mit der Aufgabe, eine Instanz desselben selbständig im Zielbild finden. Dabei werden neben der Position im Bild rotierte und skalierte Instanzen vom Agenten erzeugt. Jeder einzelne Agent implementiert ein Annealing-Verfahren zur Suche mit dem Ziel der Maximierung eines Korrelationskoeffizienten.

Die Agenten können die gleiche oder unterschiedliche Aufgaben bearbeiten. Ihre Anzahl ist effektiv durch die Prozessor-Last und letztlich durch die Anzahl der zur Verfügung stehenden Rechner im Cluster begrenzt. In der vorliegenden Version besteht nur eine vertikale Kommunikationsstruktur – eine Inter-Agent-Kommunikation ist noch nicht vorgesehen.

Damit entsteht letztlich ein verteiltes System zur Erkennung mehrerer Objekte durch Suchagenten. Diese (hier unabhängigen Agenten) suchen eigenständig nach dem angefragten Objekt und liefern in jedem Fall ein Ergebnis – die Qualität dieser Ergebnisse könnte letztlich durch den Benutzer, den Interface-Agenten oder durch die Suchagenten selbst geeignet bewertet werden. Eine Weiterverarbeitung der Ergebnisse erfolgt im Rahmen dieser Arbeit nicht.

Teil des Systems ist (jeweils) auch eine graphische Benutzeroberfläche mit Funktionen zur Bearbeitung von Bildern, der Generierung von Templates und nicht zuletzt der Kontrolle der Erkennung sowie der Darstellung der Ergebnisse.

Bevor nun in Kapitel 2 alternative Konzepte zur Erkennung vorgestellt werden, erfolgt zunächst die Festlegung einer geeigneten Notation.

1.3 Notation

Skalare Größen

$$s \text{ oder } S$$

werden in Groß- und Kleinschreibung verwendet.

Vektoren

$$\vec{v} \text{ oder } \vec{V}$$

werden in Groß- und Kleinschreibung verwendet und je nach Zweckmäßigkeit im Kontext unter Weglassen des Pfeils. Dabei symbolisiert $\vec{0}$ den Nullvektor.

Das **Skalarprodukt** zweier Vektoren a und b wird durch

$$a^T b \text{ oder auch } a \cdot b,$$

der Betrag

$$|\vec{v}| \text{ oder } |s|$$

eines Vektors oder einer skalaren Größe durch einfache Betragsstriche dargestellt. Die **Norm** wird in Abgrenzung dazu durch $\|\cdot\|$ dargestellt.

Matrizen

$$M \text{ oder } M_{ij}$$

werden entweder Fett (die Matrix), oder einfach mit Indizierung (der Eintrag der Matrix an der Stelle ij) dargestellt.

Doppelte **Indizierung** wird nur bei Matrizen M_{ij} „doppelt-unten“, ansonsten x_i^n „oben-unten“ verwendet. Die **Determinante** einer Matrix wird dargestellt durch

$$|M| \text{ oder } \det M$$

Die **Inverse** bzw. die **Transponierte** einer Matrix:

$$M^{-1} \text{ und } M^T$$

Mengen werden in Großbuchstaben

$$\mathbb{R}, \mathbb{N}, A, \dots$$

und ihre **Kardinalität** wie folgt dargestellt:

$$\text{card}Q = |Q|$$

Für die **Exponentialfunktion** gilt:

$$\exp(\cdot) \equiv e^{(\cdot)}$$

Zustandsübergänge werden durch \rightarrow symbolisiert:

$$f \rightarrow f' \quad \text{mit Zuständen } f, f'.$$

Für die **Summation** gilt

$$\sum_{x,y}(\cdot) \equiv \sum_x \sum_y(\cdot)$$

und wird nur im Kontext stark vereinfacht ohne Summations-Grenzen dargestellt (analog für Integrale).

Unter einem digitalen **Intensitätsbild (Grauwertbild)** wird im folgenden die Abbildung

$$I : M \times N \mapsto [0, 255] \subset \mathbb{N} \cup 0 \quad \text{mit } M, N \in \mathbb{N}$$

verstanden.

Algorithmen werden in *Pseudocode* vereinfacht dargestellt.

2

Lösungsansätze zur Objekterkennung

Verfahren zur Objekterkennung sind ebenso zahlreich wie anwendungsspezifisch, was die Auswahl eines für das konkrete Problem geeigneten Verfahrens erheblich erschwert. In den letzten Jahren haben sich jedoch zwei unterschiedliche Modelle ausgeprägt, die ein wichtiges Entscheidungskriterium darstellen.

Für die gegebene Aufgabenstellung wurde daher

Geometrie-basierte und Appearance-basierte

Modellierung betrachtet.

In *Geometrie-basierten* Verfahren werden Objekte bezüglich der Auswirkungen der Effekte, die durch Gruppen von Transformationen (der Kamera) hervorgerufen werden, beschrieben.

Appearance-basierte Verfahren beschreiben Objekte in Abhängigkeit von Intensitätstransformationen und werden im Rahmen dieser Arbeit ausführlicher beschrieben, da sie mit dem implementierten Verfahren einige Verbindungen aufweisen.

Ein experimenteller Vergleich (in erster Linie hinsichtlich der Verdeckungsproblematik ohne Berücksichtigung variierender Beleuchtungsbedingungen) wird in [MLP⁺96] vorgestellt. Aus der Problemstellung der Behandlung komplexerer Objekte wird ein Vergleich in [BP93] gegeben. In der (noch nicht gedruckten) Publikation [LOKC99] werden Methoden zur Invarianz bezüglich Beleuchtungseinflüssen erörtert.

Im folgenden Abschnitt werden zunächst Aspekte der Geometrie-basierten Ansätze angesprochen und ihre zentralen Eigenschaften genannt. Danach wird entsprechend auf die Appearance-basierten Verfahren eingegangen und eine Auswahl auf der Basis dieser Erörterungen getroffen.

2.1 Geometrie-basierte Verfahren

Im allgemeinen versuchen Geometrie-basierte Ansätze ein (3D-)Objektmodell mit einer Menge von Merkmalen zu „matchen“, die aus dem Bild (der Szene) extrahiert wurden. Diese Merkmale basieren auf lokalen Informationen, wie z.B. Kanten und Ecken, und ermöglichen daher potentiell eine Erkennung unter (teilweiser) Verdeckung. Die Anwendbarkeit ist aber begrenzt auf solche Objekte, die durch einfache geometrische Primitiva beschrieben werden können. Die Extrahierbarkeit solcher Primitiva – und die Modellierbarkeit durch diese – wird dabei vorausgesetzt.

2.1.1 Erkennung durch Invarianten

Geometrische Invarianten sind (skalare) Form-Deskriptoren, die sich unter bestimmten geometrischen (oft projektiven) Transformationen nicht ändern. Allgemein transformiert sich die Invariante einer $I(\mathbf{P})$ einer Konfiguration \mathbf{P} wie folgt

$$I(\mathbf{P}') = |\mathbf{A}|^w I(\mathbf{P}) \quad (2.1)$$

Dabei ist $I(\mathbf{P}')$ die relative Invariante mit Gewicht w und \mathbf{P}' die transformierte Konfiguration. Für $w = 0$ spricht man von einer skalaren Invariante. Für projektive Transformationen ist \mathbf{A} eine nicht-singuläre quadratische Matrix angewandt auf homogene Koordinaten (vgl. ZFMRLP:94).

Geometrische Transformationen

Die zweidimensionalen Transformationen

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2.2)$$

umfassen Skalierung s_x, s_y , Scherung α und Rotation θ

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (2.3)$$

Also für den zweidimensionalen Fall in kartesischen Koordinaten kurz

$$\vec{x}' = \mathbf{A}\vec{x} - \vec{x}_0 \quad (2.4)$$

mit

$$\mathbf{A} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin(\theta + \alpha) & \cos(\theta + \alpha) \end{pmatrix} \quad (2.5)$$

um eine lineare Transformation mit den Parametern $\vec{a} = (x_0, y_0, \theta_0, s_x, s_y, \alpha)^T$ zu beschreiben. Befindet sich der Sensor (die Kamera) nahe am Objekt, so ist perspektivische