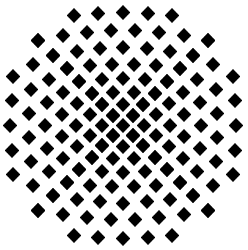
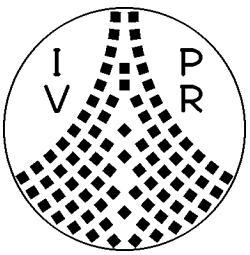


Universität Stuttgart
Fakultät Informatik



Institut für parallele und verteilte
Höchstleistungsrechnersysteme
Breitwiesenstraße 20-22
D-70565 Stuttgart

Statistische Objekterkennung durch nichtparametrische Hypothesenverifizierung

N. Oswald und P. Levi

1997/17

Oktober 1997

CR: I.2.10, I.3.5, I.4.7, I.5.1

Zusammenfassung

In diesem Bericht wird ein Verfahren vorgestellt, mit dem die Ähnlichkeit zweier komplexer 3D-Objekte an Hand nichtparametrischer statistischer Tests festgestellt werden kann. Dabei werden Objekte zunächst in einen normierten skalierungs- und translationsinvarianten Merkmalsraum abgebildet und anschließend unter Anwendung von Verfahren, die auf dem Vergleich von Verteilungsfunktionen basieren, mit verschiedenen Modellansichten verglichen. Zur Generierung eines Bewertungsmaßes für die Ähnlichkeit von Objekten verwenden wir die uni- und bivariaten Teststatistiken von Kolmogoroff und Smirnow sowie die von Cramer und von von Mises. Wie die Experimente zeigen, läßt sich mit diesem schnellen Verfahren besonders bei den bivariaten Tests eine hohe Signifikanz erzielen. Durch Gruppierung der Merkmale ist eine weitere Effizienzsteigerung bei der Erkennung der Objekte möglich.

1 Einleitung

Eine zentrale Aufgabe in der Robotik, sei es in Verkehrs-, Fertigungs- oder Dienstleistungsszenarien, ist das schnelle Erkennen von Objekten in einer Szene, in der die Objekte an beliebiger Position und in beliebiger Orientierung sichtbar sind. Bei diesen Objekten handelt es sich meist um hochkomplexe Gebilde, deren Identifizierung hohe Anforderungen an die Bildverarbeitung stellt. Um komplexe Objekte zu erkennen, reduziert man üblicherweise die zu verarbeitende Information, indem zur Segmentierung Aufmerksamkeitserreger wie Bewegung oder Farbe benutzt werden. Die Qualität der Objekterkennung hängt dann stark von der Qualität der Segmentierung ab. Um auch zusätzliche Einflüße wie projektive Verzerrungen oder Beleuchtungsschwankungen weitestgehend kompensieren zu können, müssen Objekterkennungsverfahren vor allem robust sein.

In [12] wurden zwei schnelle Matchingverfahren vorgestellt, die auf einem translations- und skalierungsinvarianten Konfigurationsraum für ikonische Merkmale basieren. Durch Aufsummieren der Differenzen korrespondierender Raster wurden bei großen Merkmalsmengen Energiewerte bestimmt, die allerdings nicht normiert waren, so daß sich die anschließende Klassifikation als schwierig gestaltete. Dies führte zu der Überlegung, die Verteilung der Punktanordnungen im Konfigurationsraum als zufällig aufzufassen und mit mathematisch fundierten statistischen Methoden zu analysieren. Um Unabhängigkeit zu bestimmten Verteilungsformen zu gewährleisten, haben wir nichtparametrische Verfahren verwendet, die auch unter relativ geringen Anforderungen an der den Daten zugrundeliegenden Grundgesamtheit angewendet werden können. Diese Verfahren sind selbst im Falle normalverteilter Grundgesamtheiten nur unwesentlich weniger effizient als parametrische Verfahren [8]. Die Qualität einzelner Testverfahren wird im Schrifttum sehr unterschiedlich bewertet [10]. In unseren Anwendungen hat sich gezeigt, daß insbesondere Testverfahren geeignet sind - im Gegensatz beispielsweise zu Rangordnungsverfahren -, die auf dem Vergleich von Verteilungsfunktionen basieren. Aus den potentiellen Testverfahren haben wir das bekannte Testkriterium der maximalen Differenz von Kolmogoroff und Smirnow sowie das der integrierten quadrierten Differenz von Cramer und von Mises ausgewählt.

Die statistische Objekterkennung ist bislang nicht sehr verbreitet. Sie besteht nach [16] erstens aus dem Erstellen einer formalen Beschreibung für ein Objekt und zweitens dem Klassifizieren des daraus resultierenden Musters. Statistische Methoden werden bisher hauptsächlich zur Klassifikation eingesetzt (siehe z.B. [11]), während das Erstellen von Objektbeschreibungen mangels formaler Modelle meist der Intuition des Designers unterliegt [16]. Eine formale Objektbeschreibung, die auf der statistischen Modellierung von Objekten beruht, wird in [9] vorgestellt. Dieser wahrscheinlichkeitsbasierte Ansatz zur Erkennung und Lokalisierung von Objekten basiert jedoch auf parametrischen statistischen Methoden. Auch [6] verwendet eine feste Verteilungsfunktion. Die statistische Objektidentifikation erfolgt dort, indem die Ereignisverteilung der Daten im Hough-Raum mit einem Anpassungstest auf eine geometrische Verteilung hin untersucht werden. In [15] wird ein statistisches Modell vorgestellt, mit dem korrekte und nicht korrekte Objekthypothesen unterschieden werden können, womit die Fehlerwahrscheinlichkeit bei der Wahl von Schwellwerten reduziert wird.

2 Erstellen einer formalen Objektbeschreibung

Zum Erstellen einer formalen Objektbeschreibung werden aus einem Grauwertbild, welches vorwiegend das gesuchte Objekt enthält, zunächst m relevante ikonische Merkmale M_i ($i \in [1, m]$) wie Grauwerte, Kanten oder Ecken bestimmt. Die Merkmale werden durch geometrische Attribute beschrieben, die die Positionen der Merkmale im Objektraum O angeben. Anschließend erfolgt eine Normierung der Merkmale bzgl. ihrer geometrischen Positionen durch Transformation vom n-dimensionalen Objektraum in den in [12] vorgestellten translations- und skalierungsinvarianten Konfigurationsraum κ. Faßt man die entstandene Anordnung der Punktmerkmale in κ als zufällig auf, so können mit Hilfe nichtparametrischer Verfahren aus dem Matching eines Objekts mit vorhandenen Modellobjekten statistische Merkmale extrahiert und anschließend klassifiziert werden.

2.1 Repräsentation von Merkmalen im Konfigurationsraum

Der Konfigurationsraum ist ein normierter kontinuierlicher Raum mit dem Wertebereich $[0, 1]$ für jede Koordinatenachse x_t ($t \in [1, n]$). Bei Verwendung punktförmiger Merkmale (z.B. Kantenpixel oder Grauwerte) entspricht die Dimension des Konfigurationsraumes der des Objektraumes O. Die den Konfigurationsraum aufspannenden Merkmalsvektoren $\vec{B}_i = (b_{1,i}, \ldots, b_{n,i})^T$ sind entsprechend der geometrischen Position der Merkmale M_i mit $i \in [1, m]$ im Objektraum O relativ zu den übrigen Merkmalen M_j ($j \in [1, m] \quad \forall j \neq i$) angeordnet. Der Konfigurationsraum entspricht somit einem Quantilisierungsraum. Die Abbildung eines Merkmals M_i von O nach κ mit $\kappa \subseteq [0, 1]^n$ erfolgt durch Bestimmung der Werte $b_{t,i}$ für jede Koordinatenachse x_t:

$$b_{t,i} = \frac{1}{m-1} \sum_{j=1}^{m-1} f_{t,ij} \quad \text{mit} \quad f_{t,ij} = \begin{cases} 1 & \text{falls } o_{t,i} \leq o_{t,j} + \varepsilon_t \quad \forall j \neq i \\ 0 & \text{sonst} \end{cases} \tag{1}$$

Der Merkmalvektors $\vec{B}_i$ eines transformierten Merkmals i im Konfigurationsraum besteht aus n Elementen $b_{t,i}$, die aus dem Vergleich der Koordinaten $o_{t,i}$ des Merkmals M_i mit allen übrigen Merkmalen in O bestimmt werden, wobei der Wert ε_t eine Überlappungstoleranz definiert. Der so erstellte Konfigurationsraum ist, wie schon in [12] gezeigt, sowohl translations- als auch skalierungsinvariant und wird im weiteren Verlauf die Basis für die angewandten statistischen Methoden darstellen. Der Einfluß der Skalierung auf den Konfigurationsraum wird an den verschiedenen Auflösungen eines kantengefilterten mobilen Roboters in Abb. 2-1 deutlich. Die Pixelreduktion sorgt zwar für eine Ausdünnung der Punktanordnung, die Objektform wird jedoch beibehalten.

Abbildung 2-1: Skalierung des Konfigurationsraumes mit 100%, 80%, 60%, 40%, 30% und 20%

2.2 Extraktion statistischer Merkmale

Die statistische Testtheorie basiert auf der Vorgehensweise, über eine Stichprobe eine Vermutung in Form einer Nullhypothese auszudrücken und diese unter Auswahl einer geeigneten Testfunktion gegen eine Alternativhypothese unter Angabe von Gütemaßen zu überprüfen. Die Frage nach der Verteilung einer solchen Stichprobenfunktion gehört dabei zu den Grundproblemen in der mathematischen Statistik [5]. Für wenige Sonderfälle, in denen normalverteilte Grundgesamtheiten angenommen werden, gibt es verhältnismäßig einfache Methoden zur Bestimmung der exakten Verteilung der Stichprobenfunktion. Ansonsten ist die Bestimmung der exakten Verteilung sehr aufwendig, so daß man stattdessen die Grenzverteilung der Stichprobenfunktion für $m \to \infty$ bestimmt, ohne dabei Annahmen an eine Verteilungsfunktion zu treffen.

2.2.1 Bestimmen eines Bewertungsmaßes für statistische Tests

Betrachtet man die Modelldaten als gegeben und vollständig, so geht es beim Einstichprobenfall darum, ein zu analysierendes Objekt, dessen Repräsentation im Konfigurationsraum als Stichprobe aufgefaßt wird, an ein Modell anzupassen, indem auf den kontinuierlichen Punktmerkmalen des Konfigurationsraumes uni- oder multivariate Tests angewandt werden. Als Nullhypothese H_0 für die verschiedenen Tests wird die Vermutung aufgestellt, daß die Bild- und Modelldaten aus der gleichen Grundgesamtheit, als Alternativhypothese H_A, daß sie aus verschiedenen Grundgesamtheiten entstammen. Ziel bei diesen dualen Entscheidungen ist nun einerseits, eine hohe Sicherheit für H_0 zu erzielen, so daß eine fälschliche Ablehnung von H_0 vermieden wird und α, der Fehler 1. Art, gegen Null geht. Andererseits möchte man eine hohe Trennschärfe zwischen H_0 und H_A erreichen, so daß der Fehler 2. Art (β), die irrtümliche Beibehaltung von H_0, gegen Null geht. Einen derart idealen Test kann man jedoch nicht erhalten [5].

Um von dem dualen Entscheidungskonzept statistischer Tests hin zu einem Bewertungsmaß für die Ähnlichkeit von Objekten zu gelangen, wird zunächst für den Vergleich zweier Stichproben aus κ_1 und κ_2 das nominelle Signifikanzniveau α_k (der sogenannte p-Wert [7]) des angewandten Tests ermittelt, bei dem dieser gerade abgelehnt worden wäre. Je größer α_k, desto größer ist die Wahrscheinlichkeit, daß H_0 fälschlicherweise abgelehnt wurde, gleichzeitig sinkt aber auch der Fehler 2. Art. Diese kritische Irrtumswahrscheinlichkeit α_k wird im folgenden als Bewertungsmaß für das Matching verwendet. Das bedeutet, falls der Wert α_k kleiner als ein vorgegebenes α ist, wird der jeweils angewandte Test zu einer Ablehnung von H_0 führen.

2.2.2 Der uni- und bivariate Test von Kolmogoroff und Smirnow

Der Test von Kolmogoroff und Smirnow (KS-Test) ermittelt den Unterschied zweier Verteilungsfunktionen, indem er die zentrale Tendenz, Streuung und Schiefe der Verteilungen erfaßt. Voraussetzungen für den KS-Test sind unabhängige Daten, die ein kardinales Meßniveau besitzen und eine stetige Verteilungsfunktion besitzen. Diese Voraussetzungen sind gewährleistet zum einen durch die Quantilisierungseigenschaft von κ, womit der Abstand zweier Merkmale $\vec{B}_i$ bestimmbar ist, zum anderen durch die aus den $\vec{B}_i$ generierte empirische Verteilungsfunktion $F(\vec{B}_i)$ mit

$$F(\vec{B}_i) = \begin{cases} 0 & \text{für } \vec{B}_i \le \vec{B}_1 \\ \frac{k}{m} & \text{für } \vec{B}_k < \vec{B}_i \le \vec{B}_{k+1} \quad (k=1,2,\ldots,m) \\ 1 & \text{für } \vec{B}_i > \vec{B}_m \end{cases} \tag{2}$$

die nicht abnehmend und linksseitig stetig ist. Als Prüfgröße dient die größte Ordinatendifferenz D zwischen den empirischen Verteilungsfunktionen von Objekt $F^O(\vec{B}_i)$ und Referenzmodell $F^R(\vec{B}_i)$:

$$D = \sup_i \left| F^O(\vec{B}_i) - F^R(\vec{B}_i) \right| \tag{3}$$

Für einfache Stichproben vom Umfang m ergibt sich für die Operationscharakteristik $Q(\lambda_{KS})$ nach dem Satz von Kolmogoroff [5] die folgende Grenzverteilung:

$$\lim_{m \to \infty} Q_m(\lambda_{KS}) = Q(\lambda_{KS}) = \begin{cases} \sum_{j=-\infty}^{\infty} (-1)^j \cdot e^{-2j^2\lambda_{KS}^2} & \text{für } \lambda_{KS} > 0 \\ 0 & \text{sonst} \end{cases} \tag{4}$$

Nach [14] läßt sich das benötigte λ_{KS} im Einstichprobenfall für den univariaten KS-Test durch folgende Formel näherungsweise bestimmen:

$$\lambda_{KS} = \left(\sqrt{m} + 0,12 + \frac{0,11}{\sqrt{m}} \right) D \tag{5}$$

λ_{KS} errechnet sich in Abhängigkeit vom Stichprobenumfang m und D, wobei sich der Wert m aus der Anzahl der Merkmale des Objekts m^O und des Referenzmodells m^R wie in (6) zusammensetzt.

$$m = \frac{m^O \cdot m^R}{m^O + m^R} \tag{6}$$

Die empirische Verteilungsfunktion $F^O(\vec{B}_i)$ konvergiert nach dem Satz von Gliwenko [5] für $m \to \infty$ mit Wahrscheinlichkeit 1 zur tatsächlichen Verteilungsfunktion $F^R(\vec{B}_i)$, falls die $\vec{B}_i$ von κ gemäß $F^R(\vec{B}_i)$ verteilt sind. Als Bewertungsmaß verwenden wir die kritische Irrtumswahrscheinlichkeit $\alpha_k = 1 - Q(\lambda_{KS})$. Liegen nun gleiche Grundgesamtheiten vor, werden sowohl λ_{KS} als auch $Q(\lambda_{KS})$ mit zunehmendem m immer kleiner, α_k wird dagegen immer größer. Hohe Werte von α_k bedeuten also, daß eine Ähnlichkeit zwischen Objekt und Modellansicht vorliegt.

Einen univariaten KS-Test kann man auf einen zweidimensionalen Konfigurationsraum anwenden, indem man κ in S_1 vertikale und S_2 horizontale Streifen zerlegt und die größte Ordinatendifferenz D für eine Koordinatenachse x_t nach (3) bestimmt. Der KS-Test wird dann in jedem Streifen $S_{1,s}$ bzw. $S_{2,s}$ durchgeführt, so daß sich α_k gemäß (7) aus dem arithmetischen Mittel der Streifenergebnisse ergibt. Für den Fall eines dreidimensionalen κ verfährt man analog.

$$\alpha_k = \frac{1}{S_1 + S_2} \left(\sum_{s=1}^{S_1} \alpha_k^{S_{1,s}} + \sum_{s=1}^{S_2} \alpha_k^{S_{2,s}} \right) \tag{7}$$

Der bivariate Test von Kolmogoroff und Smirnow (KS2D-Test) läßt sich auf einen zweidimensiona-

len Konfigurationsraum dagegen direkt anwenden. D berechnet sich nach (3), diesmal jedoch aus dem Maximum über beide Koordinatenachsen x_1 und x_2. Voraussetzung für den KS2D-Test ist, daß die beiden Verteilungsfunktionen $F^O(\vec{B}_i)$ und $F^R(\vec{B}_i)$ ähnliche Korrelationskoeffizienten r besitzen. Für die Zufallsvariable λ_{KS2D} verwenden wir die in [14] angegebene Näherungformel:

$$\lambda_{KS2D} = D \frac{\sqrt{m}}{1 + \sqrt{1 - r^2}\left(0,25 - \frac{0,75}{\sqrt{m}}\right)} \tag{8}$$

Die Operationscharakteristik $Q(\lambda_{KS2D})$ bzw. das Bewertungsmaß α_k berechnet sich durch Einsetzen von λ_{KS2D} in (4). Im Falle korrelierter Daten reduziert sich der KS2D-Test auf den KS-Test. Bei dreidimensionalen Eingabedaten kann der dreidimensionale κ entlang einer beliebigen Koordinatenachse x_t in Ebenen zerlegt werden. Anschließend wird in jeder dieser Ebenen der KS2D-Test durchgeführt und das Bewertungsmaß α_k z.B. wie in (7) aus dem arithmetischen Mittel aller Ebenenergebnisse bestimmt.

2.2.3 Der uni- und bivariate Test von Cramer und von Mises

Der Test von Cramer und von Mises (CM-Test) basiert ebenfalls auf dem Vergleich zweier Verteilungsfunktionen. Statt der größten Differenz wie im KS-Test wird die integrierte quadrierte Differenz zweier Verteilungsfunktionen berechnet. Für den CM-Test gelten die gleichen Voraussetzungen wie für den KS-Test. In Anlehnung an [2] verwenden wir für den univariaten CM-Test bezogen auf den Konfigurationsraum die folgende Teststatistik für eine Koordinatenachse x_t:

$$\lambda_{CM} = m \int_0^1 \left(F^O(\vec{B}_i) - F^R(\vec{B}_i)\right)^2 d\vec{B}_i \tag{9}$$

Um ein Bewertungsmaß α_k zu erhalten, wurde aus den Quantilen für λ_{CM}, die man beispielsweise aus [3] entnimmt, die folgende Bewertungsfunktion für α_k approximiert [1]:

$$\alpha_k = \begin{cases} 1 - 1244,68\lambda_{CM}^3 + 16,90\lambda_{CM}^2 & 0 \le \lambda_{CM} < 0,05 \\ e^{4,29\lambda_{CM}^2 - 9,05\lambda_{CM} + 0,32} & 0,05 \le \lambda_{CM} < 1,0 \\ 0 & \text{sonst} \end{cases} \tag{10}$$

Auch beim bivariaten Cramer und von Mises Test [4] wird die Verteilungsfunktion der Merkmale $\vec{B}_i$ im Falle eines zweidimensionalen κ direkt bestimmt. Die Teststatistik λ_{CM2D} hat folgende Form:

$$\lambda_{CM2D} = m \int_0^1 \int_0^1 \left(F^O(\vec{B}_i) - F^R(\vec{B}_i)\right)^2 d\vec{B}_i \tag{11}$$

Das Bewertungsmaß α_k erhält man, indem λ_{CM2D} in (10) eingesetzt wird.

3 Klassifikation

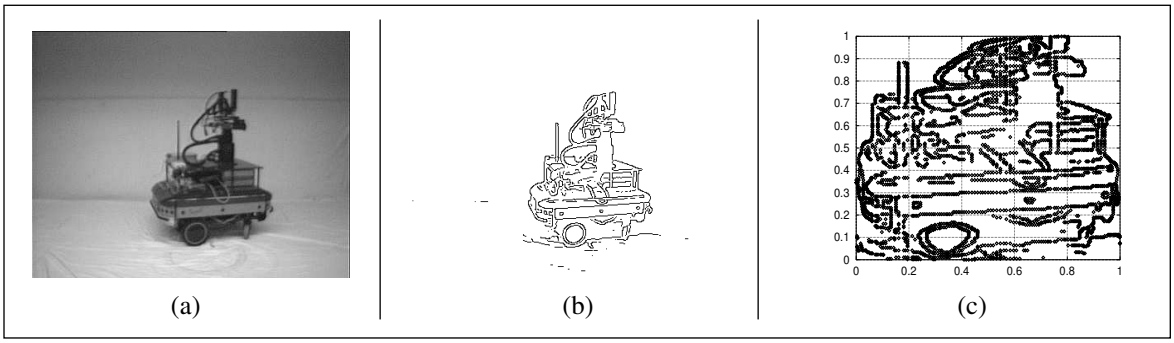
Jeder Vergleich zweier Stichproben durch eines der oben beschriebenen Testverfahren liefert ein Bewertungsmaß α_k. Gleichzeitig erfolgt durch jeden Test bzgl. eines Niveaus α eine Klassifikation

nach H_0 oder H_A. Bezogen auf das Bewertungsmaß α_k bedeutet das, falls α_k kleiner als ein vorgegebenes α ist, führt der verwandte Test zu einer Annahme von H_A. Diese Art der Klassifizierung ist jedoch für eine Objekterkennungsaufgabe nicht ausreichend. Stattdessen wird nun die Klassifikation der α_k so durchgeführt, daß einerseits eine Zuordnung zu einer Objektklasse erfolgen kann, andererseits die jeweilige Ansicht eines Objekts bestimmbar wird. Dazu wird die ansichtsbasierte Modelldatenbank als eine hierarchische baumartige Struktur aufgefaßt, die j verschiedene Modelle R_j^M als Blätter enthält, von denen jedes l_j definierte Ansichten R_{j,l_j}^A besitzt.

Nach jedem Test existiert für jedes Modell ein Merkmalsvektor $\vec{v}_j$ mit $\vec{v}_j = (P(R_{j,1}^A), P(R_{j,2}^A), \ldots, P(R_{j,l_j}^A))^T$, der die Bewertungen $\alpha_k = P(R_{j,l_j}^A)$ für alle R_{j,l_j}^A enthält. Die $\vec{v}_j$ spannen somit einen statistischen Merkmalsraum auf. Die Hypothesenwahrscheinlichkeit $H(R_j^M)$ für ein Modell R_j^M ergibt sich aus der besten Bewertung $\max_{l_j} P(R_{j,l_j}^A)$. Damit entspricht diese Art der Klassifikation dem Verwenden von Diskriminanzgeraden. Eine Aussage über die Trennschärfe der Modelle bzw. Aspekte ergibt sich aus der Modell- und der Aspektsignifikanz. Die Modellsignifikanz $SG(R_{j_1}^M)$ berechnet sich aus der relativen Differenz der beiden besten Modellbewertungen $H(R_{j_1}^M)$ und $H(R_{j_2}^M)$, die Aspektsignifikanz $SG(R_{j,l_1}^A)$ ergibt sich aus der relativen Differenz der beiden besten Aspektbewertungen $H(R_{j,l_1}^A)$ und $H(R_{j,l_2}^A)$. Ein großer Wert $SG(R_{j_1}^M)$ bedeutet, daß es sich mit hoher Wahrscheinlichkeit um R_j^M handelt. Die Ähnlichkeit eines analysierten Objekts zu einer bestimmter Ansicht R_{j,l_j}^A wird durch einen hohen Wert $SG(R_{j,l_1}^A)$ ausgedrückt.

4 Experimentelle Ergebnisse

In den Experimenten verwenden wir als Eingabe Grauwertbilder, die überwiegend das zu identifizierende Objekt enthalten. Mit Hilfe des Canny-Operators werden Punktmerkmale extrahiert und in den Konfigurationsraum transformiert (Abb. 4-2). Das Matching mit dem Eingabeobjekt erfolgt durch Vergleich mit Modellansichten aus einer ansichtsbasierten Modelldatenbank, in der sich definierte Ansichten von Objekten befinden, die zu einer festen Kameraposition in 15°-Schritten aufgenommen wurden (vgl. Abb. 4-3). Somit liegen für jedes Objekt pro Entfernungsposition 24 Ansichten vor.



(a) (b) (c)

Abbildung 4-2: Grauwertbild (a), canny-gefiltertes Bild (b) und der resultierende Konfigurationsraum (c)

Zum Zwecke einer effizienten Berechnung der Differenzen der Verteilungsfunktionen wird der Konfigurationsraum diskretisiert und werden die Merkmale gruppiert. Bei einer sehr feinen Diskretisierung - mit maximal einem Merkmal pro Raster - bildet man die empirische Verteilungsfunktion

quasi direkt aus dem Konfigurationsraum. Eine grobe Diskretisierung hat, wie die Experimente zei-

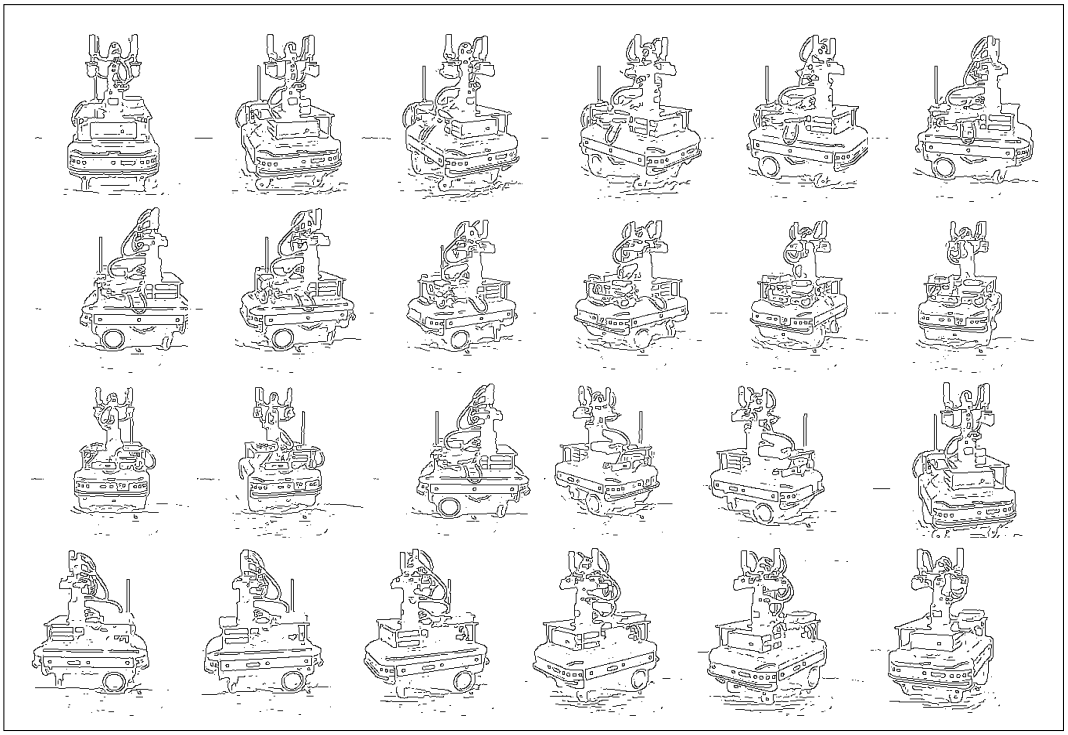


Abbildung 4-3: Modellansichten des mobilen Roboters *Athos* in 15^{o}-Schritten

gen, zur Folge, daß nur geringfügige Signifikanzverluste gegenüber einer feinen Diskretisierung auftreten [1], gleichzeitig aber der quadratische Rechenaufwand sinkt. Aus Abb. 4-4 (a+b) wird am Beispiel des KS2D-Tests deutlich, daß bereits mit einer 10×10-Rasterung eine hohe Trennschärfe erreicht wird, da die Bewertungen sehr eng um die korrekte Ansicht von 60^{o} liegen.

Der Einfluß der Skalierung auf den Konfigurationsraum wird an den verschiedenen Auflösungen eines mobilen Roboters in Abb. 2-1 deutlich. Versuche haben gezeigt, daß trotz der theoretischen Eigenschaft der Skalierungsinvarianz aufgrund der Quantisierung eine Mindestanzahl an Kantenpixel, mindestens 30%, sichtbar sein muß. Zur Bestimmung der Orientierungsgenauigkeit wurden Aufnahmen eines Greifarms in 5^{o}-Schritten mit Modelldaten, die in 30^{o}-Schritten aufgenommen wurden, verglichen. Abb. 4-4 (c+d) zeigt, daß bei einer Abweichung von bis zu 15^{o} noch eine korrekte Zuordnung erfolgen kann. Die Aspekttoleranz für eine grobe Objektunterscheidung liegt damit bei ca. 30^{o}. Eine feinere Unterscheidung der Ansichten eines Objektes bedarf dagegen bei der Modellierung einer sehr hohen Meßgenauigkeit, wobei eine genauere Auflösung als 5^{o} in der Regel nicht erzielt werden kann. Die Ansicht eines Objektes hängt nicht nur von seiner Orientierung ab, sondern auch von der Beobachtungshöhe bzw. -entfernung. Die Robustheit gegenüber perspektivischen Verzerrungen aufgrund der Beobachtungshöhe wurde bei etwa 8^{o} ausgemacht.

Das Bewertungsmaß (Aspektbewertung) für eine korrekte Zuordnung ist in einer idealen Szene typischerweise hoch und sehr trennscharf, so daß man eine hohe Modellsignifikanz erhält. Für den Versuch in Abb. 4-4 (e), in dem der Roboter *Aramis* aus etwa 225^{o} mit den Modellen von *Athos*, *Aramis*

und *Bagger* verglichen wurde, beträgt die Modellsignifikanz ca. 95%. Die Aspektsignfikanz liegt

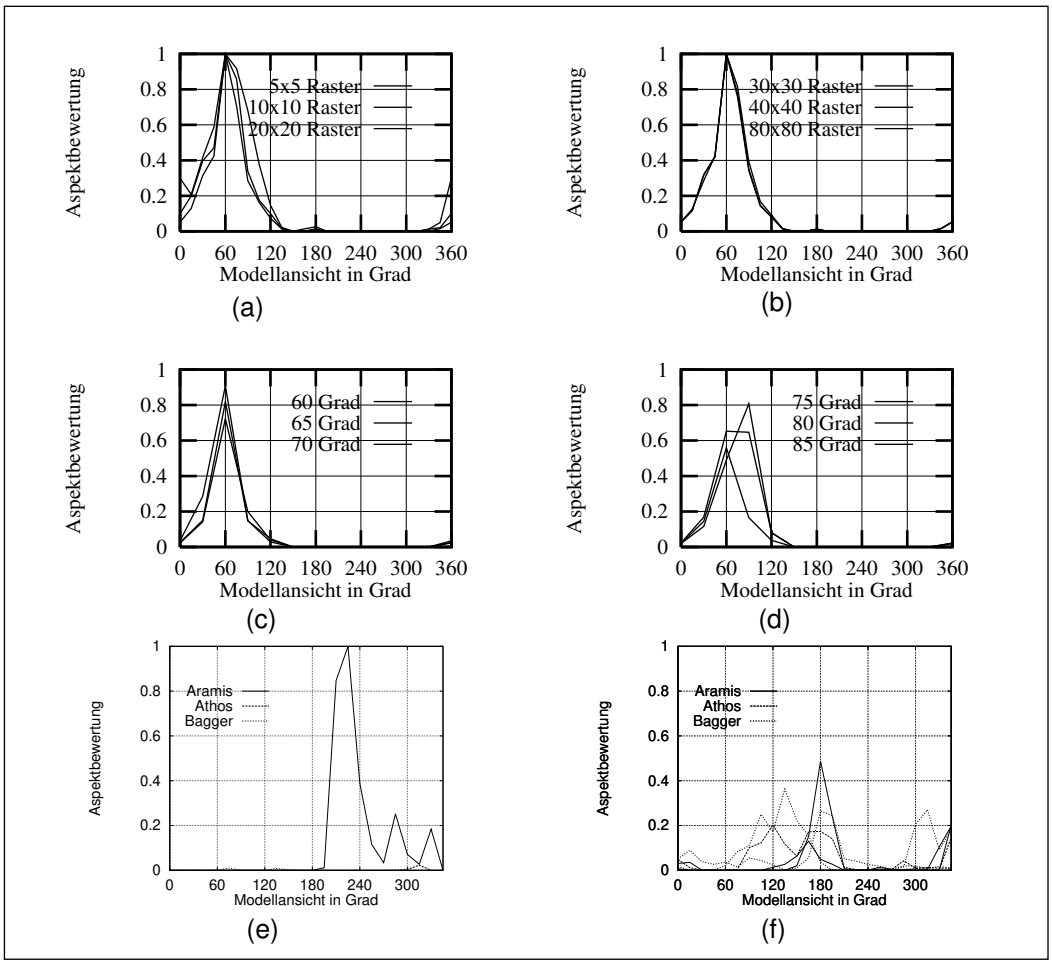


Abbildung 4-4: Einfluß der Diskretisierung auf die Aspektbewertung beim KS2D-Test (a + b), Aspekttoleranz bzgl. der Objektorientierung beim CM2D-Tests (c+d), Aspektbewertungen bei bekannten (e) und unbekannten Objekten (f) mit KS2D-Test

hier bei etwa 15%. Beim Vergleich mit unbekannten Objekten erhält man nur sehr schwache Aspektbewertungen, wie etwa in Abb. 4-2 (f), wo eine *Person* mit den drei Modellen aus Abb. 4-4 (e) verglichen wurde.

In realen Szenen mit heterogenem Hintergrund müssen die Verfahren robust gegenüber Verzeichnungen des Hintergrundes sein. Um dies zu testen, wurden für die Experimente sogenannte Regions of Interest (RoI) als Eingabedaten verwendet, die durch Ausnutzen der Objektbewegung in Form einer konvexen Hülle - wie in [13] beschrieben - segmentiert wurden. Durch dieses nur grobe Umschließen der Objekte kann die zu analysierende RoI Störungen in Form fehlender oder zusätzlicher Informationen beinhalten, die dann in den Konfigurationsraum abgebildet werden. Wie die Bildfolge in Abb. 4-5 zeigt, erhält man mit beiden bivariaten Tests trotz der vorhandenen Störung korrekte Ergebnisse, die Modellsignifikanz ist jedoch deutlich niedriger. Die univariaten Tests schneiden hier deutlich schwächer ab. Man erkennt an diesen Bewertungen außerdem, daß bei der Identifizierung mit ansichtsbasierten Modellen vor allem bei Symmetrien Entscheidungsprobleme

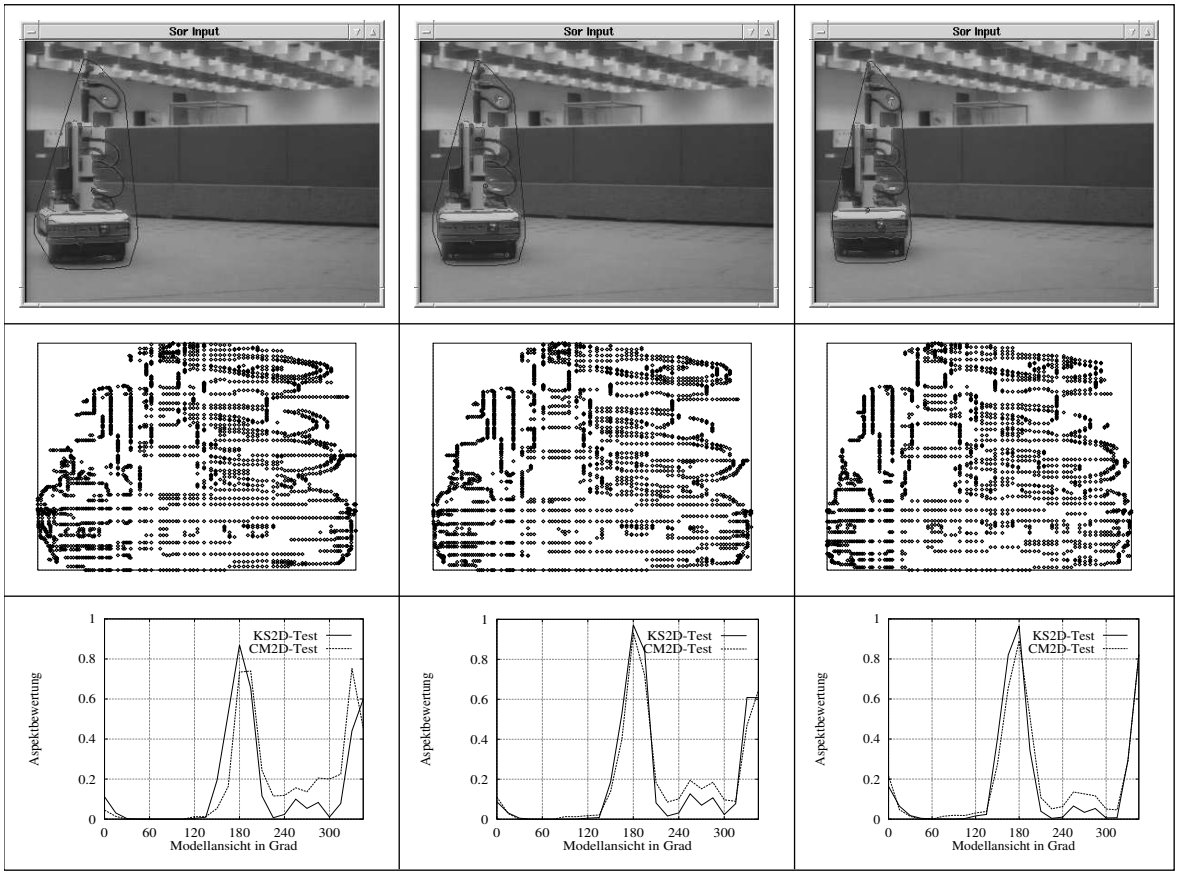


Abbildung 4-5: Bildfolge aus einer realen Szene, Konfigurationsräume aus den konvexen Hüllen und Vergleich der Bewertungen mit KS2D- und CM2D-Test

auftreten können. Solche und andere Klassifikationsprobleme können beispielsweise durch die in [13] vorgestellte fortlaufende Betrachtung bzw. kooperative Objekterkennung gelöst werden.

Insgesamt schneiden von den vorgestellten Tests die bivariaten am besten ab, da sie signifikantere Bewertungen durch eine höhere Trennschärfe auch für reale Szenen liefern. Die Laufzeiten - gemessen auf einer SPARCstation 10 - werden bei diesen Tests praktisch nur durch die Anzahl der Diskretisierungsraster beeinflußt, dann jedoch quadratisch. Bei einem 10×10-Raster benötigen sowohl KS2D- als auch CM2D-Test fast konstant 100ms, sind damit also nahezu unabhängig von der Anzahl der verwendeten Modellansichten. Die univariaten Tests werden außer durch die Anzahl der Streifen auch durch die Anzahl der Kantenpixel beeinflußt. So liegt die Laufzeit bei je 10 Streifen und ca. 3300 Kantenpixeln bei etwa 400ms, bei ca. 7200 Kantenpixel schon bei 1,2s.

5 Zusammenfassung und Ausblick

Es wurde eine Methode vorgestellt, mit der die Ähnlichkeit zweier Objekte an Hand statistischer Merkmale bestimmt werden kann. Ausgehend von der Merkmalsrepräsentation im Konfigurationsraum wird aus dem Unterschied der empirischen Verteilungsfunktionen von Objekt und Modell mit Hilfe der nichtparametrischen Tests von Kolmogoroff und Smirnow und von Cramer und von Mises ein Bewertungsmaß bestimmt.

Die statistischen Verfahren ermöglichen, wie sich an Experimenten gezeigt hat, in idealen Szenen ein absolut robustes Matching, da projektive Verzerrungen oder fehlende Kantenpixel durch die Eigenschaften des Konfigurationsraumes bis zu einem gewissen Grad kompensiert werden. Dies gilt auch für reale Szenen, sofern die Segmentierung nicht zuviele Störungen beinhaltet. Von den vorgestellten Tests schneiden die bivariaten insgesamt am besten ab, da sie meist eine signifikantere Bewertung mit einer höheren Trennschärfe liefern. Zudem sind die bivariaten Verfahren durch eine grobe Diskretisierung des Konfigurationsraumes sehr schnell, so daß die Indizierung in diesem Fall eine untergeordnete Rolle spielt. Experimente zur Bestimmung der Aspekttoleranz haben gezeigt, daß - je nach Aufgabenstellung - eine Modellierung der Ansichten bzgl. der Orientierung in 30^0-Schritten und bzgl. der Beobachtungshöhe in 8^o-Schritten ausreichen kann.

Gegenwärtig erweitern wir die Objekterkennung dahingehend, daß sie einerseits robuster gegenüber segmentierten Eingabedaten niedriger Qualität wird, andererseits, daß sie den Prozeß der Segmentierung zur Steigerung der Segmentierungsqualität unterstützt.

Literatur

1. Biedert, R. *Robuste und schnelle Identifizierung komplexer 3D-Objekte durch statistische Hypothesenverifizierung.* IPVR Diplomarbeit Nr. 1401, Universität Stutgart (1996).
2. Büning, H. und Trenkler, G. *Nichtparametrische statistische Methoden*. De Gruyter (1994).
3. Conover, W.J. *Practical Nonparametric Statistics*. Wiley & Sons (1971).
4. Csörgö, M. In P.R. Krishnaiah und P.K.Sen (Hrsg.). *Handbook of Statistics Vol. 4*, North-Holland (1984) 444-449.
5. Fisz, M. *Wahrscheinlichkeitsrechnung und mathematische Statistik*. VEB Deutscher Verlag der Wissenschaften (1989).
6. Grimson, W.E.L. *Object Recognition by Computer: The Role of Constraints*. MIT Press (1990).
7. Hartung, J. *Statistik: Lehr-und Handbuch der angewandten Statistik*. Oldenbourg Verlag (1995).
8. Hollander, M. und Wolfe, D.A. *Nonparametric Statistical Methods*. Wiley & Sons (1973).
9. Hornegger, J. *Statistische Modellierung, Klassifikation und Lokalisation von Objekten.* Shaker (1996).
10. Mardia, K.V. *Tests of Univariate and Multivariate Normality.* In P.R. Krishnaiah (Hrsg.). *Handbook of Statistics Vol. 1*, North-Holland (1980) 279-320.
11. Niemann, H. *Klassifikation von Mustern.* Springer (1983).
12. Oswald, N., Gerl, S. und Biedert, R. *Konfigurationsbasiertes Verfahren zur schnellen Identifikation komplexer Objekte*, Mustererkennung 1996, 18. DAGM-Symposium, Springer Verlag, S. 187-195.
13. Oswald, N. und Levi, P. *Cooperative Vision in a Multi-Agent Architecture*. In *Image Analysis and Processing 1997*, Lecture Notes in Computer Science 1310, Springer Verlag (1997) 709-716.
14. Press, W.H et al. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press (1992).
15. Sarachik, K.B. *An Analysis of the Effect of Gaussian Error in Object Recognition.* MIT, Technical Report AI-94 (1994).
16. Sonka, M. et al. *Image Processing, Analysis and Machine Vision*. Chapman & Hall (1993).