

Statistisches Formenwissen in Variationsansätzen zur Bildsegmentierung

Daniel Cremers

Lehrstuhl für Bildverarbeitung, Mustererkennung und Computer Grafik
Fakultät für Mathematik und Informatik, Universität Mannheim
<http://www.cs.ucla.edu/~cremers>

Abstract: In der vorliegenden Arbeit werden Bildsegmentierungsverfahren entwickelt, die es ermöglichen, gelerntes Wissen über die Silhouetten bekannter Objekte in den Segmentierungsprozess zu integrieren. Das statistisch repräsentierte Formwissen führt zu deutlich besserer Segmentierung vertrauter Objekte in Inputbildern, die durch Rauschen, teilweise Verdeckungen und störende Hintergrundstrukturen korrumpiert sind.

Es wird ein Überblick über existierende Variationsansätze zur Bildsegmentierung gegeben. Anschließend werden die *Diffusion Snakes* vorgestellt, eine Symbiose zweier etablierter Ansätze, in der flächenbasierte Segmentierung mit einer splinebasierten Konturrepräsentation kombiniert wird.

Es werden statistische Formmodelle verschiedener Komplexität eingeführt. Auf der Grundlage der Kernmethoden wird ein nichtlineares statistisches Formmodell entwickelt. Dieses Modell erlaubt es, mehrere dreidimensionale Objekte durch die Silhouetten verschiedener zweidimensionaler Ansichten zu kodieren und – trotz teilweiser Verdeckungen – über längere Videosequenzen zu verfolgen und sehr präzise zu segmentieren. Ein neues Verfahren intrinsischer Registrierung garantiert ein Formenwissen, welches invariant ist gegenüber Verschiebung, Drehung und Skalierung der entsprechenden Objekte.

Im letzten Teil dieser Arbeit wird eine Modifikation des Datenterms der Kostenfunktion vorgeschlagen, die es ermöglicht, Objekte nicht aufgrund ihres Aussehens zu segmentieren, sondern aufgrund ihrer relativen Bewegung in einer gegebenen Videosequenz. Experimentelle Resultate belegen, daß sich bewegte Objekte auch dann noch präzise segmentieren und über Videosequenzen verfolgen lassen, wenn sich sowohl Objekt als auch Hintergrund bewegen und wenn sich Objekt und Hintergrund in ihrer Helligkeitsstruktur nicht unterscheiden.

Der vorliegende Text liefert einen Abriß der Ergebnisse der Dissertation. Eine ausführlichere Darstellung findet sich in [Cr02].

1 Einleitung

Bildsegmentierung ist ein zentrales Probleme in den Bereichen der Bildverarbeitung und des Maschinensehens. Mit Segmentierung bezeichnet man die Zerlegung eines gegebenen Bildes in sinnvolle Bereiche. In der vorliegenden Arbeit behandeln wir die spezielle Aufgabe, ein Objekt unseres Interesses vom Hintergrund zu trennen. In dieser Formulierung ist das Problem der Segmentierung eng mit dem der Objekterkennung verknüpft.



Abbildung 1: Ein Suchbild: Dalmatinerhund auf einer Blätterwiese.¹ Der Dalmatiner befindet sich rechts der Mitte, kehrt dem Betrachter den Rücken, mit Kopf in der Bildmitte. Für die Segmentierung eines solchen Bildes berücksichtigen menschliche Betrachter sowohl die Helligkeitswerte des Bildes als auch vorher gelerntes Weltwissen.

Eine Kernthese dieser Arbeit ist es, daß Segmentierung und Objekterkennung nicht unabhängig voneinander gelöst werden können. Dies läßt sich anhand von Abbildung 1 verdeutlichen, die einen Dalmatinerhund auf einer Laubwiese zeigt. Durch Grobrasterung und Binarisierung kann der Bildbereich des Hundes kaum von dem des Hintergrundes unterschieden werden. Das Wissen, daß das gesuchte Objekt ein Dalmatinerhund ist, hilft bei der Segmentierung ungemein. Umgekehrt läßt sich das Objekt viel leichter als Dalmatiner identifizieren, wenn man weiß, welche Bildbereiche zum Objekt gehören.

Dieses Zusammenspiel zwischen der Information des Inputbildes und vorher gelerntem Wissen in Segmentierungsprozessen zu modellieren, ist das Thema der vorliegenden Arbeit. Abbildung 2 verdeutlicht die dieser Arbeit zugrunde liegende Problemstellung. Gegeben seien ein Bild, welches ein Objekt unseres Interesses enthält, und eine Anzahl von Silhouetten, die verschiedenen Ansichten dieses Objektes entsprechen. Die Aufgabe ist es nun, Inputbild und Silhouetten in möglichst optimaler Weise in einem Segmentierungsprozess zu integrieren.

2 Helligkeitsbasierte Segmentierung: Die Diffusion-Snakes

Im Folgenden bezeichnen wir mit $\Omega \subset \mathbb{R}^2$ die Bildebene und mit $I : \Omega \rightarrow \mathbb{R}$ das Inputbild, also eine Funktion, die jedem Punkt in der Bildebene einen Helligkeitswert zuordnet. Aufgabe der Bildsegmentierung ist es, eine Kontur $C \subset \Omega$ zu finden, die die Bildebene in sinnvolle Bereiche zerlegt.

¹Das Bild des Dalmatiners wird R. C. James zugeschrieben.



Abbildung 2: Ziel dieser Arbeit ist die Formulierung von Segmentierungsprozessen, die sowohl die Bildinformation (links) als auch vorher gelernte Konturen (rechts) optimal ausnutzen.

In den letzten Jahrzehnten sind viele Verfahren zur Bildsegmentierung vorgeschlagen worden. Besonders attraktiv sind unter diesen die Variationsmethoden, d.h. Verfahren, bei denen die Segmentierung durch Optimierung einer geeigneten Kostenfunktion bestimmt wird. Eine ausführliche Diskussion der Vorteile dieser Ansätze findet sich in [MS95]. Die Variationsansätze lassen sich in kantenbasierte und flächenbasierte Verfahren unterteilen.

Kantenbasierte Verfahren zielen darauf ab, die Kontur C derart in die Bildebene Ω zu legen, daß sie entlang der größten Helligkeitssprünge der Bildfunktion I verläuft. Ein sehr prominenter Vertreter sind die *Snakes* [KWT88]: Ausgehend von einer initialen parametrischen Kontur $C : [0, 1] \rightarrow \Omega$, wird diese derart deformiert, daß die Energie

$$E_{snakes}(C) = \int \alpha \left(\frac{dC(s)}{ds} \right)^2 + \beta \left(\frac{d^2C(s)}{ds^2} \right)^2 + |\nabla I(C(S))|^2 ds, \quad (1)$$

minimiert wird. Die ersten beiden Terme dieser Kostenfunktion, gewichtet mit positiven Parametern α und β , beschreiben die innere Energie der Kontur, nämlich ihre Elastizität und ihre Steifheit. Der letzte Term garantiert, daß die Kontur an Stellen mit möglichst großen Helligkeitssprüngen (großem Gradienten $|\nabla I|$) zu liegen kommt.

Ein Repräsentant der flächenbasierten Ansätze ist das Mumford-Shah Funktional [MS89]:

$$E_{MS}(C, u) = \int_{\Omega} (I - u)^2 dx + \lambda^2 \int_{\Omega - C} |\nabla u|^2 dx + \nu |C|. \quad (2)$$

Durch gleichzeitige Minimierung dieser Kostenfunktion bezüglich C und u wird das Inputbild I durch eine stückweise glatte Funktion $u : \Omega \rightarrow \mathbb{R}$ approximiert – das heißt durch eine Funktion, die möglichst nah an I liegt und überall dort glatt ist (einen kleinen Gradienten $|\nabla u|$ hat), wo keine Kontur C liegt. Der letzte Term in Gleichung (2) garantiert, daß die zu findende Kontur eine möglichst kleine Länge $|C|$ hat.

Kantenbasierte Verfahren leiden bei der Anwendung auf reale Bilder daran, daß Bildfunktionen üblicherweise viele Grauwertsprünge besitzen, sodaß die zu minimierende Energie entsprechend viele lokale Minima besitzt. In Anwendungen äußert sich dieses Phänomen darin, daß die Kontur hinreichend nah am interessierenden Objekt initialisiert werden muß, damit ein lokales Minimierungsverfahren zur gewünschten Lösung konvergiert.

Die Probleme der Rauschunterdrückung und der Konturlokalisierung sind in flächenbasierten Verfahren viel klarer getrennt. Entsprechend lassen sich auch sehr verrauschte Bilder derart segmentieren, daß Kanten und Ecken von Objekten erhalten bleiben. Konturen konvergieren typischerweise über wesentlich größere räumliche Distanzen, sodaß die Initialisierung nicht sehr präzise gewählt werden muß. Aufgrund einer klaren probabilistischen Interpretation lassen flächenbasierte Verfahren elegante Verallgemeinerungen des Segmentierungskriteriums zu. Eine derartige Verallgemeinerung zur Segmentierung bewegter Objekte in Videosequenzen wird im letzten Teil dieser Arbeit vorgeschlagen. Die Arbeit von Mumford und Shah ist eher theoretischer Natur, eine konkrete Konturrepräsentation wird nicht vorgeschlagen.

Die in dieser Arbeit entwickelten *Diffusion Snakes* [CTWS02] sind eine Symbiose der beiden oben vorgestellten Verfahren. Das Kostenfunktional besteht aus dem Datenterm des Mumford-Shah Funktionals und dem Elastizitätsterm der Snakes:

$$E_{DS}(C, u) = \int_{\Omega} (I - u)^2 dx + \lambda^2 \int_{\Omega - C} |\nabla u|^2 dx + \nu \int \left(\frac{dC(s)}{ds} \right)^2 ds. \quad (3)$$

Die Kontur C wird als geschlossene Splinekurve implementiert. Es wird gezeigt, daß der Elastizitätsterm – im Gegensatz zum Längenterm des Mumford-Shah Funktionals (2) – dafür sorgt, daß die Splinekontrollpunkte eine äquidistante Verteilung anstreben, sodaß die mit expliziten Konturen verbundenen numerischen Instabilitäten verhindert werden.

Im Grenzfall $\lambda \rightarrow \infty$ ergeben sich die sogenannten *Simplified Diffusion Snakes*. Diese ermöglichen Echtzeitimplementierungen flächenbasierter Segmentierung: Die Konturevolution entspricht einer Aktualisierung weniger Kontrollpunkte und die Approximation u läßt sich alternierend durch eine einfache Mittelwertbildung bestimmen.

Der Längenterm des Mumford-Shah Funktionals (und der korrespondierende Elastizitätsterm der Diffusion Snakes) entspricht einer Apriori-Wahrscheinlichkeit im Raum aller Konturen: Konturen kürzerer Länge sind apriori wahrscheinlicher. Hat man zusätzliche Information in Form von Trainingsformen vertrauter Objekte, so läßt sich diese rein geometrische Apriori-Wahrscheinlichkeit durch statistische Abstandsmaße ersetzen, die eine Ähnlichkeit zu den Trainingsformen berücksichtigen.

3 Repräsentation und Lernen von Formen

Wir repräsentieren eine Menge von Trainingsformen durch Splinekurven – siehe Abbildung 3. Wir normieren die Kontrollpolygone bezüglich Verschiebung, Drehung, Skalierung [Go91] und zyklischer Permutation – siehe Abbildung 4. Das statistische Lernen besteht darin, die den Kontrollpunktvektoren zugrundeliegende Verteilung zu schätzen.

Eine einfache parametrische Verteilung ist die multivariate Gaußverteilung \mathcal{P}_{gauss} , die einem sogenannten Mahalanobisabstand E_{gauss} entspricht:

$$\mathcal{P}_{gauss}(z) \propto \exp(-E_{gauss}(z)) \quad \text{mit} \quad E_{gauss}(z) = (z - z_o)^t \Sigma^{-1} (z - z_o). \quad (4)$$

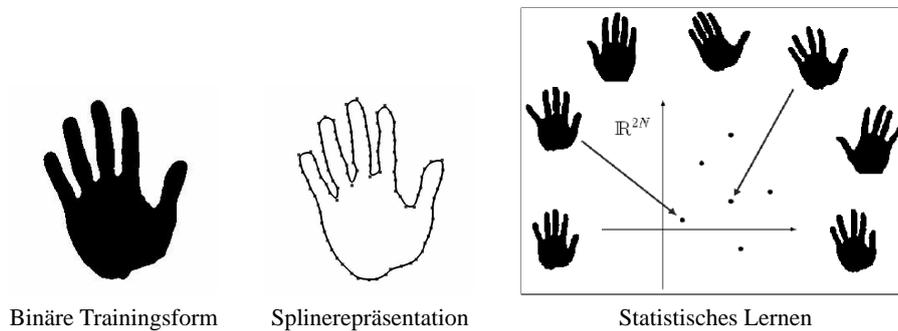


Abbildung 3: Schematische Darstellung des statistischen Lernens von Formen. Formen werden durch Splines repräsentiert, jede Form entspricht einem Vektor von Kontrollpunkten. Das statistische Lernen besteht darin, die den Kontrollpunktvektoren der Trainingsformen zugrundeliegende Verteilung zu schätzen.

Hierbei ist $z \in \mathbb{R}^{2N}$ der Koordinatenvektor der Splinekontrollpunkte – in unseren Implementierungen verwenden wir typischerweise $N = 100$ Kontrollpunkte. Die Trainingsvektoren werden durch den mittleren Vektor z_0 und die Kovarianzmatrix Σ kodiert, die aus den Trainingsformen berechnet werden. Hierbei modelliert Σ die Ausdehnung der Verteilung in die verschiedenen Raumrichtungen.

Aus dieser statistischen Energie läßt sich ein Formähnlichkeitsmaß bilden, welches invariant gegenüber Verschiebung, Drehung und Skalierung einer gegebenen Kontur ist:

$$E_{shape}(z) = E_{gauss}(\hat{z}(z)). \quad (5)$$

Hierbei steht $\hat{z}(z)$ für die Registrierung der gegebenen Form z mit den Trainingsformen – siehe Abbildung 4.

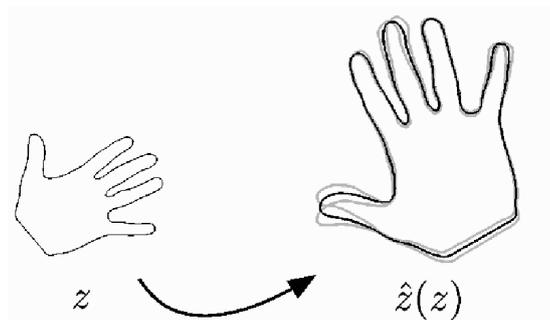
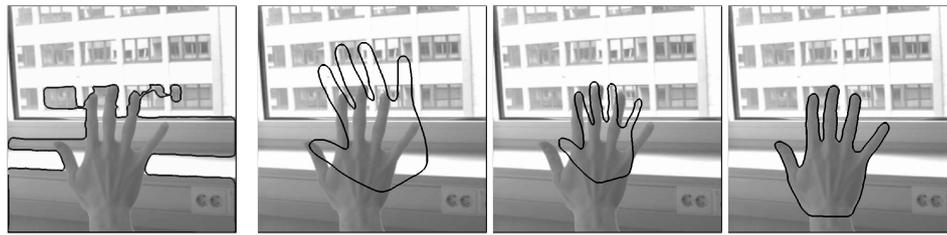


Abbildung 4: Schematische Darstellung der Registrierung einer Form z mit den Trainingsformen bezüglich Verschiebung, Drehung und Skalierung. Durch eine intrinsische Registrierung wird das Formähnlichkeitsmaß (5) invariant gegenüber entsprechenden Transformationen der Form z .



Nur hellkeitsbasiert

Konturevolution mit statistischem Ähnlichkeitsmaß

Abbildung 5: Einfluß des statistischen Konturwissens. Durch die Integration des statistischen Konturähnlichkeitsmasses (5) in den Segmentierungsprozess (6) wird die evolvierende Kontur auf einen Unterraum vertrauter Formen eingeschränkt (siehe die drei Bilder rechts). Dadurch wird der Segmentierungsprozess robust gegen störende Hintergrundstrukturen, Verdeckungen und Rauschen. Im Gegensatz zur Segmentierung ohne Konturwissen (links), ermöglicht das statistische Formenwissen die korrekte Segmentierung (rechts).

4 Wissensgetriebene Segmentierung

Einen Segmentierungsprozeß, der sowohl die Helligkeitsinformation des Inputbildes als auch die gelernten Konturen berücksichtigt, modelliert man dadurch, daß man Bildinformation (3) und Konturenwissen (5) in einer gemeinsamen Kostenfunktion integriert:

$$E_{\text{gesamt}}(C, u) = (1 - \alpha) E_{DS}(C, u) + \alpha E_{\text{shape}}(C). \quad (6)$$

Der Parameter $\alpha \in [0, 1]$ gewichtet den Einfluß des vorher gelernten Wissens gegenüber dem der Bildinformation. Für $\alpha = 0$ ergibt sich ein rein datengetriebener Segmentierungsprozeß, und für $\alpha \rightarrow 1$ wird das gelernte Wissen in zunehmendem Maße berücksichtigt. Diese Kombination von Daten und Wissen in einem Variationsansatz ist äquivalent zu der aus der Statistik bekannten Bayes'schen Inferenzmethode.

Abbildung 5 zeigt einen Vergleich: Für dasselbe Inputbild einer Hand vor einem strukturierten Hintergrund ergibt der Segmentierungsprozeß ohne Wissen ($\alpha = 0$) eine Segmentierung, die helle und dunkle Bereiche trennt (links). Schaltet man das statistische Ähnlichkeitsmaß hinzu ($\alpha > 0$) ergibt sich ein Segmentierungsprozeß, in dem die evolvierende Kontur auf einen Unterraum vertrauter Formen eingeschränkt ist – siehe Abbildung 5, rechte drei Bilder: Das Konturenwissen ermöglicht die gewünschte Segmentierung.

5 Nichtlineare Statistik: Gaußverteilung im Merkmalsraum

Geht man zu komplexeren Trainingsmengen über, die aus verschiedenen Formen bestehen – zum Beispiel Silhouetten eines dreidimensionalen Objektes (vgl. Abbildung 2) – dann bricht die Annahme eines Gauß'schen Modells zwangsläufig zusammen. Eine naheliegende Modellerweiterung ist eine Verteilung, die aus mehreren Gaußverteilungen besteht, denn jede Verteilung läßt sich beliebig gut durch Summen von Gaußverteilungen approximieren.

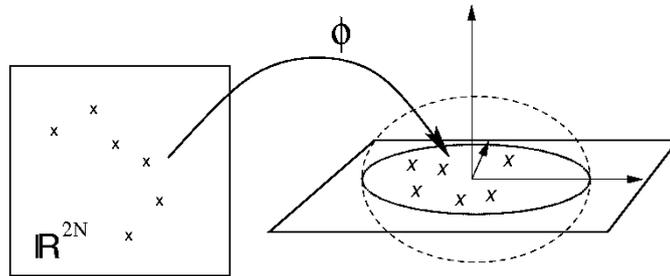


Abbildung 6: Schematische Darstellung der Transformation ϕ der Trainingsformen in einen i.a. höherdimensionalen Raum und der Annahme einer Gaußverteilung der transformierten Daten.

In dieser Arbeit wird ein alternatives Modell vorgeschlagen [CKS02], das auf den aus dem Bereich des Maschinenlernens bekannten Kernmethoden [Va98, SSM98] aufbaut. Von der oben diskutierten Multimodalverteilung unterscheidet es sich in folgenden Punkten:

- Im Gegensatz zur numerisch aufwändigen iterativen Optimierung der multimodalen Gaußverteilung, ergibt sich das hier vorgeschlagene Modell als analytische Funktion der Trainingsdaten.
- Statt einer festen Anzahl von Gaußverteilungen enthält das Modell einen Skalensparameter, der für einen gegebenen Datensatz die Zahl der Cluster der geschätzten Verteilung induziert. Dieser Parameter läßt sich als Granularität auffassen, er bestimmt, auf welcher Skala zwei Datenpunkte als ähnlich anerkannt werden.
- Das hier vorgeschlagene Modell entspricht einer einfachen Gaußverteilung in einem geeigneten Merkmalsraum. Im Originalraum entspricht es einer Verallgemeinerung des klassischen Parzenschätzers.

Die Modellidee läßt sich anhand von Abbildung 6 verdeutlichen. Wir nehmen an, daß die Trainingsformen nach einer geeigneten nichtlinearen Transformation ϕ in einen typischerweise höherdimensionalen Raum F durch eine Gaußverteilung in F approximiert werden können. Diese Annahme entspricht einem Mahalanobisabstand in F :

$$E(z) = (\phi(z) - \phi_0)^t \Sigma_\phi^{-1} (\phi(z) - \phi_0), \quad (7)$$

Analog zu (4) bezeichnen ϕ_0 und Σ_ϕ den Mittelwert und die Kovarianzmatrix der transformierten Trainingsdaten. Formal unterscheidet sich das Modell (7) vom bisherigen in (4) nur durch die zwischengeschaltete Abbildung ϕ . Die entsprechenden Verteilungen im Ursprungsraum der Trainingsformen unterscheiden sich jedoch grundlegend – das nichtlineare Modell ist geeignet, beliebige Datenstrukturen zu approximieren.

Wie ist die Abbildung ϕ zu wählen? Einsetzen von Mittelwert und Kovarianzmatrix in (7) zeigt, daß die Abbildung ϕ ausschließlich in Form von Skalarprodukten zweier transformierter Punkte auftritt – eine Eigenschaft, die alle Kernmethoden besitzen. Statt die Abbildung ϕ explizit zu spezifizieren, ist es daher hinreichend, für zwei gegebene Punkte $z_1, z_2 \in \mathbb{R}_n$ das Skalarprodukt der transformierten Punkte $k(z_1, z_2) = \langle \phi(z_1), \phi(z_2) \rangle$ festzulegen. Wir modellieren dies mit einem Gaußkern $k(z_1, z_2) \propto \exp(-|z_1 - z_2|^2/2\sigma^2)$.

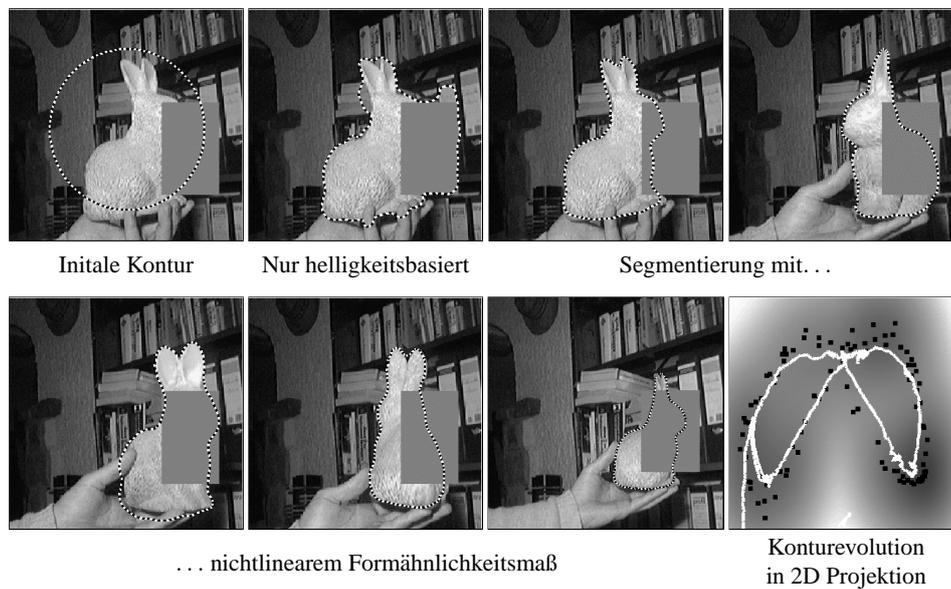


Abbildung 7: Anwendung des nichtlinearen Konturähnlichkeitsmaßes zum Verfolgen einer Hasenskulptur in einem Video. Zur Erschwerung ist das helle Rechteck als eine künstliche Verdeckung in das Video eingeführt worden. Durch Hinzuschalten des Konturähnlichkeitsmaßes (ab dem dritten Bild) wird die Segmentierung des Hasen deutlich verbessert – störende Hintergrundstrukturen werden ignoriert und der fehlende Teil der Silhouette wird aufgrund des Konturähnlichkeitsmaßes rekonstruiert. Das vorletzte Bild zeigt, daß das Ähnlichkeitsmaß invariant gegenüber der Größe der Figur ist – das Gleiche gilt für Drehung und Verschiebung der Figur. Eine Projektion auf einen zweidimensionalen Unterraum (Bild unten rechts) zeigt, daß das Konturähnlichkeitsmaß garantiert, daß die evolvierende Kontur (weiße Spur) in der Nähe der gelernten Silhouetten (schwarze Quadrate) gehalten wird.

Eine Begründung dieser Wahl der Kernfunktion k und Korrespondenzen zu Methoden der klassischen Statistik und zur Kern PCA [SSM98] finden sich in [Cr02].

Integriert man die Energie (7) analog zu (5) als Formähnlichkeitsmaß in den Segmentierungsansatz (6), so lassen sich nun sehr vielfältige Formen in unüberwachter Weise kodieren und segmentieren. Abbildung 7 zeigt eine Anwendung dieser Methode zum Verfolgen einer Hasenskulptur in einer komplexen Bildsequenz. Vorab sind 100 Silhouetten des Hasen aus entsprechenden Binärbildern generiert worden. Aus diesen läßt sich mit Formel (7) eine nichtlineare Formähnlichkeitsmaß berechnen. Der Hase läßt sich nun über die gesamte Sequenz dadurch verfolgen, daß man die Gesamtenergie (6) für jedes Bild der Sequenz minimiert. Einige segmentierte Bilder der Sequenz sind in Abbildung 7 dargestellt.¹ Eine Projektion der 100 Trainingsformen (schwarze Quadrate) und der aktuellen Form (weiße Linie) auf einen zweidimensionalen Unterraum ist in Abbildung 7, unten rechts, dargestellt. Diese macht anschaulich, wie das nichtlineare Ähnlichkeitsmaß die evolvierende Kontur in der Nähe der vertrauten Formen hält.

¹Die gesamte Sequenz findet man unter http://www.cs.ucla.edu/~cremers/Nonlinear/rabbit_small.mpeg.

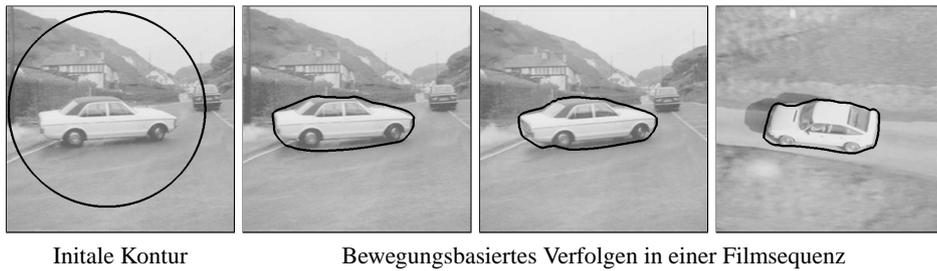


Abbildung 8: Motion Competition zum Verfolgen von bewegten Objekten in Bildsequenzen. In dieser Sequenz aus der Serie Avengers bewegen sich Autos und Hintergrund (i.e. Kamera).

6 Bewegungsbasierte Segmentierung: Motion Competition

Der Datenterm des Mumford-Shah Funktionals (und der der Diffusion Snakes) läßt sich derart modifizieren, daß man Objekte nicht mehr aufgrund ihrer relativen Helligkeit sondern aufgrund ihrer relativen Bewegung in einer Bildsequenz segmentieren kann. Durch Optimierung entsprechender Kostenfunktionale [CS03a] lassen sich die beiden Probleme der Segmentierung und der Bewegungsschätzung simultan lösen. Die Minimierung der Kostenfunktion induziert eine Art “Wettkampf” benachbarter Regionen um die separierende Kontur – daher bezeichnen wir das Modell als *Motion Competition*. Eine Verallgemeinerung zur Segmentierung multipler bewegter Objekte findet sich in [CS03b].

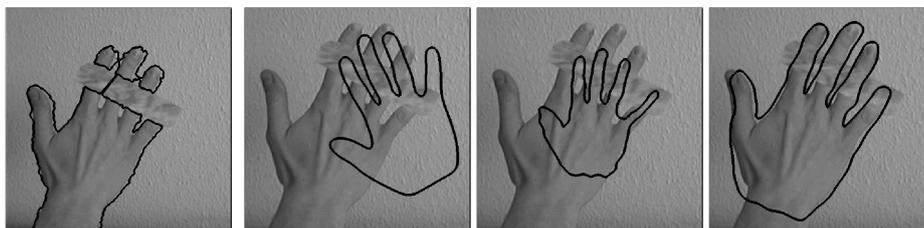
Kombiniert man das entsprechende Funktional mit einem statistischen Formähnlichkeitsmaß, so lassen sich vertraute Objekte auch dann noch segmentieren, wenn ihre Bewegungsinformation durch Rauschen oder Verdeckungen gestört ist – siehe Abbildung 9.

7 Zusammenfassung

In dieser Arbeit wurden Methoden entwickelt, vorher gelerntes Wissen über Konturen vertrauter Objekte in Prozesse der Bildsegmentierung zu integrieren. Zu diesem Zweck sind neue Segmentierungsverfahren entwickelt worden, die eine Integration von Konturenwissen in flächenbasierte Segmentierung erlauben. Aufbauend auf Entwicklungen im Bereich des Maschinenlernens sind nichtlineare statistische Modelle vorgeschlagen worden, die eine präzise statistische Kodierung auch komplexerer Formen ermöglichen. Eine Erweiterung des Datenterms der Kostenfunktion erlaubt zudem eine Segmentierung von Objekten aufgrund ihrer relativen Bewegung in einer Videosequenz. Experimentelle Resultate zeigen, daß das statistische Konturenwissen den entsprechenden Segmentierungsprozeß robust macht gegen Rauschen, Hintergrundstörungen und teilweise Verdeckungen der interessierenden Objekte.

Literatur

- [CKS02] Cremers, D., Kohlberger, T., und Schnörr, C.: Nonlinear shape statistics in Mumford-Shah based segmentation. In: Heyden, A. u. a. (Hrsg.), *Europ. Conf. on Comp. Vis.* volume 2351 of *LNCS*. S. 93–108. Copenhagen. May 2002. Springer, Berlin.



Nur bewegungsbasiert

Konturevolution mit statistischem Apriori-Wissen

Abbildung 9: Einfluß statistischen Konturwissens in der Bewegungssegmentierung. Eine Hand bewegt sich hinter einem statischen Balken. Ohne Konturwissen werden Bildbereiche segmentiert, die sich gleichartig bewegen (links). Das Konturenwissen jedoch erlaubt die Rekonstruktion der fehlenden Konturstücke und damit die gewünschte Segmentierung (rechte Bilder).

- [Cr02] Cremers, D.: *Statistical Shape Knowledge in Variational Image Segmentation*. PhD thesis. Department of Mathematics and Computer Science, University of Mannheim, Germany. 2002. http://www.cs.ucla.edu/~cremers/Publications/cremers_dissertation.pdf.
- [CS03a] Cremers, D. und Schnörr, C.: Statistical shape knowledge in variational motion segmentation. *Image and Vision Computing*. 21(1):77–86. 2003.
- [CS03b] Cremers, D. und Soatto, S.: Variational space-time motion segmentation. In: Triggs, B. und Zisserman, A. (Hrsg.), *Int. Conf. on Computer Vision*. Nice. Oct. 2003. To appear.
- [CTWS02] Cremers, D., Tischhäuser, F., Weickert, J., und Schnörr, C.: Diffusion Snakes: Introducing statistical shape knowledge into the Mumford–Shah functional. *Int. J. of Computer Vision*. 50(3):295–313. 2002.
- [Go91] Goodall, C.: Procrustes methods in the statistical analysis of shape. *J. Roy. Statist. Soc., Ser. B*. 53(2):285–339. 1991.
- [KWT88] Kass, M., Witkin, A., und Terzopoulos, D.: Snakes: Active contour models. *Int. J. of Computer Vision*. 1(4):321–331. 1988.
- [MS89] Mumford, D. und Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.* 42:577–685. 1989.
- [MS95] Morel, J.-M. und Solimini, S.: *Variational Methods in Image Segmentation*. Birkhäuser. Boston. 1995.
- [SSM98] Schölkopf, B., Smola, A., und Müller, K.-R.: Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*. 10:1299–1319. 1998.
- [Va98] Vapnik, V. N.: *Statistical Learning Theory*. J. Wiley & Sons. 1998.

Daniel Cremers studierte von 1992-1997 in Heidelberg Physik und Mathematik, das Jahr 1994/95 verbrachte er als Fulbright Stipendiat in Stony Brook (New York). 1997 legte er sein Diplom mit Auszeichnung in der Arbeitsgruppe von Franz Wegner und Andreas Mielke am Institut für theoretische Physik der Universität Heidelberg ab. Während seiner Studienzzeit wurde er von der Studienstiftung des deutschen Volkes gefördert. 1998/99 war er für ein Jahr wissenschaftlicher Angestellter am Innovationskolleg Theoretische Biologie der Humboldt Universität zu Berlin. In der Arbeitsgruppe von Andreas Herz befaßte er sich mit Feldmodellen neuronaler Aktivität im visuellen System. Von 1999-2002 forschte er am Lehrstuhl für Bildverarbeitung und Mustererkennung der Universität Mannheim unter Betreuung von Christoph Schnörr und Joachim Weickert und promovierte (summa cum laude) mit der hier skizzierten Arbeit. Seit November 2002 ist er an der University of California in Los Angeles tätig. Zu seinen Hobbys zählen Fremdsprachen, Reisen und Sport.