

AUTOMATISCHE BESTIMMUNG DER TRAJEKTORIEN OHNE MARKER ZUR BEWEGUNGSERKENNUNG VON PERSONEN

D. Meyer und H. Niemann

Universität Erlangen-Nürnberg
Lehrstuhl für Mustererkennung (Informatik 5)
Martensstr. 3, D-91058 Erlangen
Tel.: +49-9131-857775, FAX: +49-9131-303811
email: demeyer@infomatik.uni-erlangen.de

ZUSAMMENFASSUNG

Die Bedeutung der Erkennung, Verfolgung und Bewegungsanalyse von Menschen hat in den letzten Jahren zugenommen. Vor allem im medizinischen Bereich wird der Gang von Menschen meist mit Hilfe von Videoaufnahmen der Patienten analysiert. Dabei werden an bestimmten Stellen des Körpers Marker angebracht, die anhand der Bilder automatisch verfolgt werden. In diesem Artikel werden verschiedene Verfahren vorgestellt und ein Ansatz erstellt, die Trajektorien ohne Marker zu verfolgen. Dazu wird das Vektorverschiebungsfeld mit Hilfe von Monotonieoperatoren berechnet. Da sich die einzelnen Körperteile in verschiedene Richtungen bewegen, können sie aus dem berechneten Feld detektiert werden.

Keywords: Ganganalyse, Verschiebungsvektorfelder, Trajektorien

1. MOTIVATION

Anwendungen der Ganganalyse finden sich im medizinischen Bereich bei der Diagnose, aber auch in der Rehabilitation und im Sport. So ist es möglich, bestimmte anomale oder krankhafte Veränderungen z.B. der Knochen oder der Muskeln zu erkennen [1, 2].

Die Bewegungsanalyse geschieht anhand von Videoaufnahmen von gehenden Menschen. Diese Bildsequenzen werden automatisch ausgewertet. Die Ansätze zur Erkennung kann man im groben in zwei Gruppen unterteilen, die Segmentierung der Bilder und Erkennung mit Hilfe eines Menschenmodells zum einen und die direkte Nutzung der Bewegungsinformation zum anderen.

Die einzelnen Bilder können an sich benutzt und segmentiert werden. Dabei wird der menschliche Körper durch geometrische Körper modelliert. Der Mensch im ganzen oder nur bestimmte Teile werden auf den einzelnen Bildern erkannt und die Merkmale anhand dieser abgeleitet. Durch Verfolgen der einzelnen Kör-

perteile in Bildsequenzen wird auf die Bewegung geschlossen.

Eine Erkennung der Bewegungen oder Tätigkeiten von Menschen ohne den Körper geometrisch zu modellieren und diese Formen mit Hilfe von Kantenextraktion zu erkennen, geschieht mit Hilfe der Information der Bewegung, die man z.B. mit dem optischen Fluß oder mit Differenzbildern bei statischer Kamera erhält. Eine andere Beschreibung besteht aus den Orten, die bestimmte Punkte des Körpers durchlaufen, den Trajektorien.

In der klinischen Praxis ist es üblich, diese anhand von Markern zu bestimmen, die an der Person befestigt und auf den Bildern detektiert werden. Die Kurven werden dann von Medizinern ausgewertet. Dieses aufwendige Verfahren motiviert einen Ansatz zur automatischen Ganganalyse, der sowohl auf Marker verzichtet als auch die Kurven automatisch auswertet. Ebenso ist es möglich, anhand der Trajektorien zwischen verschiedenen Gangarten, z.B. Gehen, Laufen, Hüpfen zu unterscheiden.

2. BISHERIGE ARBEITEN UND ANSÄTZE

2.1. Modellierung und Erkennung des Menschen

Der menschliche Körper wird meist aus einfachen geometrischen Formen modelliert, die Unterschiede ergeben sich aus der Anzahl und Zusammensetzung dieser. So wird in [3] ein Modell aus Kegelstümpfen bzw. Zylindern, je einen für Arme, Beine und Rumpf, benutzt.

[4] benutzt ebenfalls ein Modell, in dem Menschen aus Zylindern modelliert werden. Im Gegensatz zu [3] werden die Gliedmaßen mit je drei Segmenten modelliert, so daß die Bewegungen schon genauer nachvollzogen werden können. Die Größe des Menschen wird als bekannt vorausgesetzt und dadurch die 3D-Position bestimmt. Die Kanten, die bei der Segmentierung des Bildes erkannt werden, werden mit denen

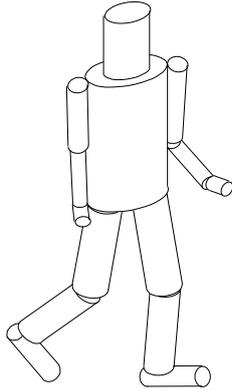


Bild 1: Beispiel zur Modellierung eines Menschen aus Zylindern

des projizierten Modells verglichen und zugeordnet, wobei deren Länge, der Abstand vom Kantenmittelpunkt im Grauwertbild zum Modell und der Winkel zwischen Grauwert- und Modellkante betrachtet werden. Verdeckte Kanten werden entfernt. Die Lage der Zylinder des Modells, das im nächsten Bild projiziert wird, wird mit Hilfe von diskreten linearen Kalman-Filtern geschätzt [4].

In [5] wird der Körper des Menschen aus 17 Segmenten und 14 Gelenken zusammengesetzt betrachtet. Zunächst wird der Kopf detektiert, dann abwärts die anderen Körperteile bestimmt, wobei bestimmte Einschränkungen aus der Anatomie gemacht werden. Die einzelnen Winkelbeschleunigungen werden aufsummiert und die Lösung mit der minimalen Summe als Kostenfunktion mit Hilfe des A*-Algorithmus bestimmt. Dieses entspricht einer möglichst gleichmäßigen Bewegung.

[6] modelliert den Körper aus einer Zusammensetzung von Zylindern, Ellipsoiden und Kugeln, mit insgesamt 22 Freiheitsgraden, drei für die Positionierung im Raum, drei für den Rumpf, je vier für die Arme und Beine.

[7] bestimmt die Kanten, die bei der Segmentierung des aktuellen Bildes mit dem des Differenzbildes zwischen aktuellem und vorhergehendem übereinstimmen. Diese Kanten werden weiterhin durch ihre Lage zueinander (innen/außen) unterschieden. Außerdem fließen in die Berechnung die Kanten früherer Bilder mit ein, damit bewegungslose Teile nicht verlorengehen. Aus diesen Konturen werden Paare von parallelen Linien extrahiert und mit Hilfe des Längen-/Abstandsverhältnisses den einzelnen Körperteilen zugeordnet. Die erkannten Teile sind Beine, Kopf, Arme und Rumpf.

[8] gewinnt ein Modell des Menschen aus mehreren aufeinanderfolgenden Bildern, indem zunächst

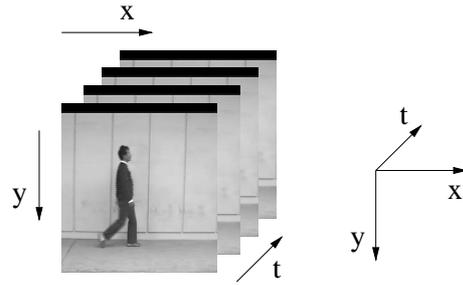


Bild 2: Bezeichnung der Achsen in den Bildern und im xyt -Raum

der gesamten Körper als ein Teil betrachtet wird. Wenn sich einzelne Teile bewegen, wird das Modell in zwei zusammenhängende Teile zerlegt, deren jeweilige Rotation zueinander bestimmt wird. Nach mehreren Iterationen sind sämtliche Einzelstücke, die den verschiedenen Körperteilen entsprechen, bestimmt. Deren Form und Lage können im 3D-Raum aus Aufnahmen verschiedener Richtungen bestimmt werden.

[9] modelliert aus der Bildsequenz einen dreidimensionalen Körper im xyt -Raum. Die Bezeichnung der Achsen ist in Bild 2 dargestellt. Sich bewegende Körper erzeugen in der xt -Ebene ein Muster, das mit Hilfe von Snakes verfolgt wird. In verschiedenen Höhen werden diese Schnitte angefertigt. Die Translation ergibt sich aus dem Schnittbild in Kopfhöhe.

2.2. Erkennung ohne geometrisches Menschenmodell

Es gibt auch verschiedenen Ansätze, die Bewegungsinformation direkt zur Erkennung zu benutzen.

[10, 11] benutzen die Information aus der Bewegung. Sie beschreiben die Bildfolge als Körper im Orts-Zeit-Bereich, den sie in xyt -Zellen normalisierter Größe einteilen, 32×32 pixel im Ortsbereich, $1/6$ Periodendauer in zeitlicher Richtung. Zur Normalisierung wird demnach die Periodizität der Bewegungsabläufe benötigt. In jeder einzelnen Zelle werden lokale Eigenschaften berechnet. Dieses sind z.B. der normale optische Fluß, d.h. der optische Fluß projiziert auf die Gradientenrichtung des Grauwertbildes, oder der summierte Fluß in der Hauptbewegungsrichtung, die sich aus einem Histogramm der Richtungen des normalen optischen Flusses gewichtet mit dem Betrag ergibt. Mit Hilfe eines dreidimensionalen Schablonenvergleichs werden die Aktionen Gehen, Rennen, Schaukeln und Springen klassifiziert.

[12] benutzt einen statistischen Ansatz, um Geste zu erkennen. Nach einer örtlichen Quantisierung extrahiert er eine große Anzahl von Merkmalsvektoren, die sich aus den Grauwerten der Zeilen bzw. Spalten nach einer Mittelwertbildung ergeben. Nach einer

Vektorquantisierung klassifiziert er auf zwei verschiedene Arten, mit Hilfe eines neuronalen Netzes und eines HMM.

[13] entwickelt zwei verschiedene Arten von Bildern aus der Bildsequenz mit dem Ziel, verschiedene Aktionen wie Setzen und gymnastische Übungen zu erkennen. Zum einen wird ein binäres Bewegungsbild bestimmt. Dieses ist ein Bild, bei dem aufgrund des optischen Flusses und einem Schwellwert festgelegt wird, ob und wo es Bewegung gibt. Es wird eine Hypothese aufgestellt, welche Aktion erkennbar sein könnte. Diese Hypothese wird mit Hilfe eines Bildes der Bewegungsvorgeschichte getestet. In diesem Bild ist anhand des Grauwertes die Zeit angegeben, zu der in den einzelnen Bereichen Bewegung detektiert werden konnte. Daraus wird die endgültige Aktion bestimmt.

2.3. Bewegungserkennung mit Trajektorien

Bewegungserkennung erfolgt oft mit am Patienten angebrachten Markern, die verfolgt werden, und der sich daraus ergebenden Trajektorien.

[14] versucht die Bewegungserkennung mit den Trajektorien verschiedener Punkte des Körpers. Betrachtet werden die Kurven, die von den Knien, Fußgelenken, Hüften, Handgelenken, Ellbogen und Schultern beschrieben werden. Die Kurven werden zum einen manuell aus Bildfolgen bestimmt, zum anderen durch LEDs, die an diesen Punkten befestigt werden und detektiert werden. Als Modell wird jeweils eine komplette Periode gespeichert. Da die Kurve im xyt -Raum betrachtet wird, ist eine Voraussetzung, daß die Bewegung senkrecht zur Kameraachse stattfindet.

Die Krümmung einer Trajektorie wird berechnet durch

$$\kappa = \frac{\sqrt{(\det(\mathbf{A}))^2 + (\det(\mathbf{B}))^2 + (\det(\mathbf{C}))^2}}{((x')^2 + (y')^2 + (t')^2)^{\frac{3}{2}}}$$

mit

$$\mathbf{A} = \begin{pmatrix} y' & t' \\ y'' & t'' \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} t' & x' \\ t'' & x'' \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} x' & y' \\ x'' & y'' \end{pmatrix}.$$

Hierbei werden diskrete Näherungen der Ableitungen benutzt, z.B. $x'(t) = x(t) - x(t-1)$, $x''(t) = x'(t) - x'(t-1)$ und bei konstanter Bildrate $t' = 1$, $t'' = 0$. Zunächst wird aus der unbekanntem Anzahl der aufgenommenen Perioden der Trajektorie mit Hilfe der Fouriertransformation ein voller Zyklus extrahiert. Danach wird die Trajektorie mit dem vorhandenen Modell verglichen und zugeordnet.

Ein weiteres Verfahren, das zu ihrer Berechnung verwendet wird, ohne Marker zu benutzen, ist das von

[15] beschriebene. Der Anfangspunkt wird im ersten Bild einer Bildfolge interaktiv festgelegt und mit Hilfe von adaptiven Templates verfolgt.

3. ERSTELLUNG EINES VERSCHIEBUNGSVEKTORFELDES MIT HILFE VON MONOTONIEOPERATOREN

Die Monotonieoperatoren nach [16] dienen zur Berechnung eines Verschiebungsvektorfeldes aus Bildsequenzen. Durch die Verschiebung punktförmiger Bildmerkmale, der lokalen Grauwertverteilungen, in zwei aufeinanderfolgenden Bildern wird Bewegungsinformation extrahiert. Der Vorteil dieses Verfahrens ist, daß die verschiedenen Bewegungsrichtungen der einzelnen Körperteile erhalten bleiben und nicht verschmiert werden. Die anschließende Zuordnung der Verschiebungsvektoren zu den Objekten kann durch Segmentierung des Verschiebungsvektorfeldes geschehen, oder dadurch, daß die Vektoren aus den Segmentierungsergebnissen des Grauwertbild den erkannten Objekten zugeordnet werden.

Beim Monotonieoperator handelt es sich um ein merkmalsorientiertes Verfahren. Implementiert wurde hier das Verfahren von [17]. Um die Merkmale zu extrahieren, werden acht Punkte innerhalb einer $(2L+1) \times (2L+1)$ Maske betrachtet. L bezeichnet die Abgriffweite. Die Anzahl der Punkte mit kleinerem Grauwert als der Mittelpunkt bestimmt die Klasse. So steht eine Klasse m_0 , also kein Wert mit kleinerem Grauwert, für ein lokales Minimum, während eine Klasse m_8 ein lokales Maximum beinhaltet. Im weiteren Verfahren werden nur diese beiden Arten von Klassen weiter verfolgt, alle anderen Punkte werden einer dritten Klasse zugeordnet. Vor den Monotonieoperator werden ein Tief- und ein Hochpaß geschaltet. Bei dieser Bandpaßfilterung um die Mittenfrequenz $\nu = 1/\lambda$ entstehen nach [16] Flecken maximaler Länge, wenn der Monotonieoperator die Abgriffweite L hat. Diese Kombination von Mittenfrequenz und Abgriffweite ergibt die stabilsten Flecken, das heißt, die daraus resultierenden Flecken lassen sich gut über verschiedene Bilder hinweg verfolgen. Stabilitätskriterien in dieser Beziehung sind möglichst große kreisrunde Fleckmerkmale, damit sich der Schwerpunkt gut bestimmen läßt.

Die Zuordnung in zwei aufeinanderfolgenden Aufnahmen geschieht mit Hilfe von zwei Attributen, dem Flecktyp (Maximum und Minimum) und dem Schwerpunkt. Die Lage wird bereits durch die Bandpaßfilterung eingeschränkt. Bei einer Bandpaßfilterung um die Mittenfrequenz $\nu = 1/\lambda$ entstehen nach [16] Flecken des Durchmessers $L = \lambda/2$. Da sich also innerhalb eines Radius von L nur ein Fleck desselben Typs befinden kann, wird auch nur in diesem Be-

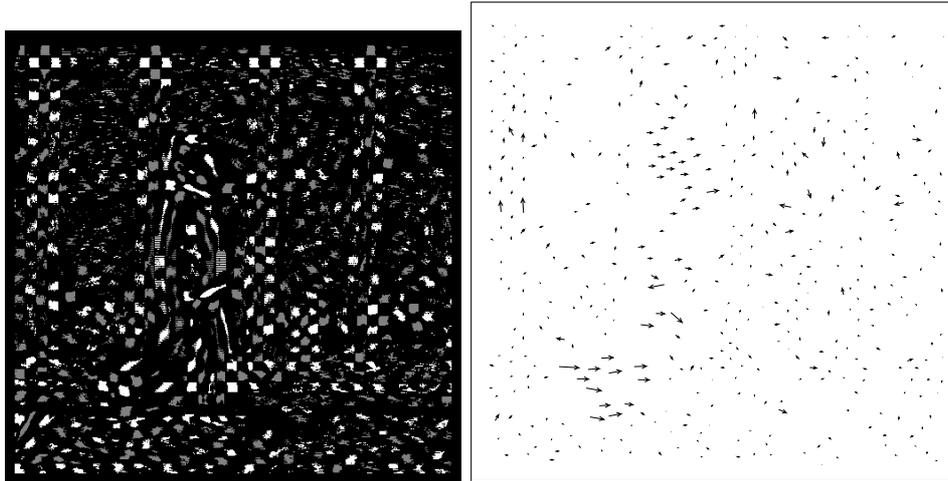


Bild 3: Fleckenmerkmale erzeugt durch den Monotonieoperator und das daraus berechnete Verschiebungsvektorfeld

reich gesucht, um Fehlzuordnungen zu vermeiden. Bei schnellen Bewegungen kann dieses allerdings zu Problemen führen. Deshalb wird die hierarchische Bewegungsschätzung eingeführt. Zunächst wird bei einer groben Auflösung mit größerer Abgriffweite und größeren Fleckenmerkmalen die Verschiebung geschätzt. Durch Mittelung verschiedener Verschiebungen, der Verschiebungskompensation, wird der Suchbereich in der nächsten Auflösungsstufe bereits um den Mittelwert verschoben. So lassen sich auch größere Verschiebungen mit höherer Genauigkeit bestimmen.

4. EXPERIMENTE

In ersten Versuchen wurde das Verschiebungsvektorfeld zweier Bilder einer Sequenz mit Hilfe von Monotonieoperatoren berechnet. Diese Bildfolge, siehe Bild 2, wurde einer Datenbank der University of Southern California, Los Angeles, entnommen.

Man kann erkennen (Bild 3), daß es Bereiche gibt, in denen die Vektoren in die gleiche Richtung zeigen. Deutlich ist zu sehen, daß im Kopf- bzw. Oberkörperbereich und am Fuß eine Häufung auftritt. Der zweite Fuß ist nicht erkennbar, da er sich zu diesem Zeitpunkt nicht bewegt.

Es gibt außerdem die Möglichkeit, direkt mit dem bei [16] beschriebenen und von [17] implementierten Verfahren die Trajektorien zu bestimmen. Dabei werden die aus einer Bildfolge extrahierten Verschiebungsvektorfelder miteinander verbunden. Bei einer Verschiebung im Raum im Sinne einer Trajektorie stimmt der Anfangspunkt eines Vektors mit dem Endpunkt aus dem des vorherigen Bildes überein.

Das Ergebnis ist in Bild 4 zu sehen. Es ist zu erkennen, daß sich die Merkmale grundsätzlich über



Bild 4: Trajektorien aus einer Bildfolge

eine Bildfolge hinweg verfolgen lassen, vor allem im Kopf- und Rumpfbereich. Hier werden nur die Linien dargestellt, die sich mindestens über die Hälfte der Bilder hinweg verfolgen lassen. Darum und weil die Füße teilweise verdeckt sind, kann man die für diese Punkte charakteristischen Kurven nicht erkennen. Außerdem fehlt eine genaue Bestimmung charakteristischer Punkte, da die Fleckenmerkmale nicht bei allen Bildfolgen an den Enden bzw. den gleichen Stellen der Gliedmaßen liegen. Auch in den Bereichen, die im Vektorverschiebungsfeld nicht gut erkennbar sind, wie im unteren Rumpfbereich, werden Trajektorien erstellt.

In zwei anderen Sequenzen eines gehenden Menschen wurden die Trajektorien von Knie und Ferse

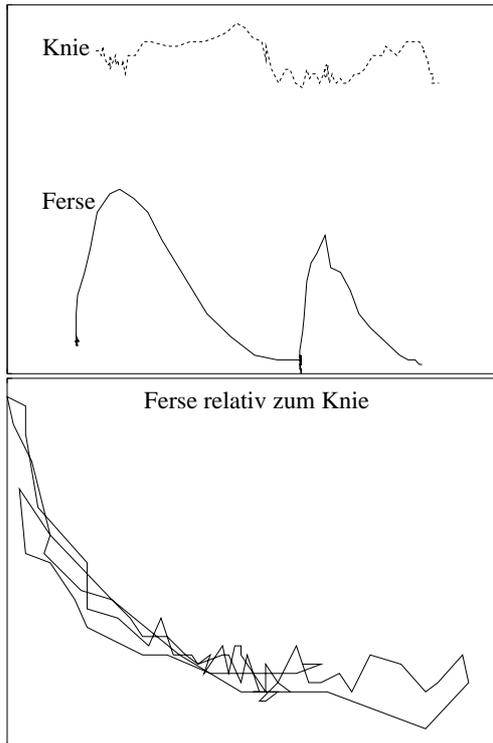


Bild 5: Kurven des linken Knies/Ferse und die Relativbewegung der Ferse zum Knie

manuell bestimmt. Die sich ergebenden Kurven zeigen ähnliche Verläufe, wobei allerdings anzumerken ist, daß eine davon mit nicht ganz zwei Schritten zu kurz ist, um genaue Aussagen zu machen.

In Bild 5 sind oben die Kurven von Knie und Ferse dargestellt, bzw. unten deren Differenz, also die relative Bewegung, was auf die Bewegung des Kniegelenkes schließen läßt. Es sind drei Schritte dargestellt, zwei rechte und ein linker, daher wird die Kurve eineinhalb mal durchlaufen. Der Kreisbogen ist aber auch bei dieser kurzen Zeit schon gut zu erkennen.

5. BEWERTUNG

Der Gang des Menschen kann durch die Bewegungen der einzelnen Gliedmaßen analysiert werden. Es gibt verschiedene Parameter, die dafür von Bedeutung sind. Dazu gehören die Winkelbewegungen der Gelenke, vor allem in Hüft-, Knie- und Fußbereich, die Frequenz der Schritte, evtl. die Armbewegungen und die Symmetrie der Bewegung, die vorhanden sein sollte.

Die Gesamtaufgabe der Ganganalyse läßt sich in drei große Bereiche unterteilen. Im ersten Teil muß der laufende Mensch detektiert und erkannt werden. Darauf folgt die Zuordnung der einzelnen Körperteile

zu den erkannten Teilen. Im dritten Teil wird anhand der Bewegung der Gang analysiert und klassifiziert.

Da angenommen wird, daß der laufende Mensch das einzige sich bewegende Objekt im Bild ist, kann bei statischer Kamera eine grobe Einschränkung der Region mit Hilfe eines Differenzbildes geschehen. Eine bessere Lösung, die auch schon einige Aussagen über die Bewegungen einzelner Körperteile macht, ist die Bestimmung des Verschiebungsvektorfeldes mit Hilfe der Monotonieoperatoren. Es zeigt sich, daß es möglich ist, die verschiedenen Bewegungsrichtungen zu unterscheiden. Zumindest lassen sich Bereiche extrahieren, in denen Häufungen des Verschiebungsvektorfeldes auftreten, die die gleiche Richtung haben. Im Bereich des Oberkörpers und Kopfes sind diese besonders deutlich, aber auch im Bereich der Beine zu erkennen. Des weiteren wurde versucht, anhand des bei [16] beschriebenen Verfahrens Trajektorien zu extrahieren.

Ziel ist es nun, die Bewegung der einzelnen Körperteile zu extrahieren. Aussagekräftige Punkte sind die Füße, die Arme, der Rumpf, die Hüften und die Knie. Kopf und Beine lassen sich aus dem Modell gut erkennen, wobei der Kopf zwar die Detektion erleichtern kann, aber über den Gang wenig aussagt. Wenn ein Teil gefunden ist, kann man durch die Information, die man über die Lage der einzelnen Körperteile zueinander aus der Anatomie hat, die Position der anderen abschätzen. Während für die Arme die Lage der Hände relativ wichtig erscheint, ist am Rumpf der Schwerpunkt interessant. Dieser Punkt beschreibt beim normalen Gang eines Menschen sowohl in der transversalen als auch in der sagittalen Ebene eine Sinuskurve [2]. Um Winkelbewegungen der Beine insbesondere der Oberschenkel zu beschreiben, müssen Punkte an der Hüfte verfolgt werden.

Die Kurven der verfolgten Punkte am Körper sagen einiges über die Art der Bewegung aus. Die Form, Periode, Geschwindigkeit und Beschleunigung sind charakteristisch für einzelne Bewegungstypen. Wenn man die Differenz der Kurven von benachbarten Gelenken bildet, wird die relative Bewegung berechnet. Es ergeben sich Kreisbögen, die Länge macht Aussagen über die Beweglichkeit des Gelenkes, der Abstand der Punkte über die Geschwindigkeit. Die Trajektorien von drei Punkten können die Winkelbewegungen des Gelenkes beschreiben.

Die Frequenz der Schritte läßt sich ebenfalls durch das Verfolgen eines Punktes beschreiben. Sie ergibt sich entweder aus einer Fouriertransformation z.B. der Krümmung der Trajektorie im Zeitbereich, oder einfach dadurch, wie viele Abtastpunkte zwischen dem Erreichen der Umkehrpunkte liegen. Welche Merkmale der Bahnen im einzelnen zur Klassifikation heran-

gezogen werden, muß durch Versuche geklärt werden.

6. WEITERES VORGEHEN

In weiteren Arbeiten wird zunächst versucht, eine zuverlässige Segmentierung zu finden, die es erlaubt, die einzelnen Körperteile zu extrahieren. Die Initialisierung hierfür kann durch das Verschiebungsvektorfeld geschehen. Durch diese Segmentierung sollten auch Punkte festgelegt werden, die als Trajektorien verfolgt werden, z.B. die Ecken oder Mittelpunkte der Kopfflächen. Die Berechnung des Verschiebungsvektorfeldes in diesen Punkten geschieht dann nach dem hier vorgestellten Verfahren.

Zunächst wird versucht, auch bei den weiteren Sequenzen, die aufgenommen wurden, die Trajektorien zu bestimmen und Merkmale zu bestimmen, mit denen der Gang analysiert und klassifiziert werden kann. Um Aussagen über die Bewegungen zu machen, muß außerdem ausreichendes Bildmaterial gesammelt werden. Deshalb werden die Untersuchungen sich zuerst einmal auf die verschiedenen Gangarten beschränken.

7. LITERATUR

- [1] O. Schmid, H. Hirschfelder, D. Hohmann: *Biomechanical Gait Evaluation of Preoperative Lower Extremity Tumorpatients*, in B. Girod, H. Niemann, H.-P. Seidel (Hrsg.): *3D Image Analysis and Synthesis '96*, Infix, Sankt Augustin, November 1996, S. 245–247.
- [2] J. Rose, J. Gamble (Hrsg.): *Human Walking*, Williams and Wilkins, Baltimore, 1994.
- [3] K. Akita: *Image Sequence Analysis of Real World Human Motion*, *Pattern Recognition*, Bd. 17, Nr. 1, 1984, S. 73–83.
- [4] K. Rohr: *Towards Model-Based Recognition of Human Movements in Image Sequences*, *Computer Vision Graphics and Image Processing*, Bd. 59, Nr. 1, 1994, S. 94–115.
- [5] Z. Chen, H.-J. Lee: *Knowledge-Guided Visual Perception of 3-D Human Gait from a Single Image Sequence*, *IEEE Transactions on Systems, Man and Cybernetics*, Bd. 22, Nr. 2, 1992, S. 336–342.
- [6] D. Gavrilu, L. Davis: *Tracking of Humans in Action: A 3-D Model-Based Approach*, in *ARPA Image Understanding Workshop*, Palm Springs, 1996.
- [7] M. Leung: *First Sight: A Human Body Outline Labeling System*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Bd. 17, Nr. 4, 1995, S. 359–377.
- [8] I. Kakadiaris, D. Metaxas: *3D Human Body Model Acquisition from Multiple Views*, in *Proceedings of the 5th International Conference on Computer Vision (ICCV)*, IEEE Computer Society Press, Boston, Juni 1995, S. 618–623.
- [9] S. Niyogi, E. Adelson: *Analysing and Recognizing Walking Figures in XYT*, in *IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, Washington, 1994, S. 469–474.
- [10] R. Polana, R. Nelson: *Recognizing Activities*, in *IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, Washington, 1994, S. 815–818.
- [11] R. Polana, R. Nelson: *Nonparametric Recognition of Nonrigid Motion*, Technical report, Department of Computer Science, University of Rochester, 1995.
- [12] G. Rigoll, A. Kosmala, M. Schuster: *A New Approach to Video Sequence Recognition Based on Statistical Methods*, in *Proceedings of the International Conference on Image Processing (ICIP)*, Bd. 3, IEEE Computer Society Press, Lausanne, Schweiz, September 1996, S. 839–842.
- [13] J. Davis: *Appearance-Based Motion Recognition of Human Actions*, TR No. 387, M.I.T. Media Lab Perceptual Computing Group, Massachusetts, 1996.
- [14] P. Tsai, M. Shah, K. Keiter, T. Kasparis: *Cyclic Motion Detection for Motion Based Recognition*, *Pattern Recognition*, Bd. 27, Nr. 12, 1994, S. 1591–1603.
- [15] D. N. Bauer, R. Frischholz: *Three Dimensional Automatic Motion Analysis*, in *ISATA 28th Symposium, Proceedings on Robotics, Motion and Machine Vision in the Automotive Industries*, 1995.
- [16] D. Koller: *Detektion, Verfolgung und Klassifikation bewegter Objekte in monokularen Bildfolgen am Beispiel von Straßenverkehrsszenen*, Bd. 13 von *Dissertationen zur künstlichen Intelligenz*, Infix, St. Augustin, 1992.
- [17] N. Schneider: *Merkmalsgestützte Bestimmung von Verschiebungsvektorfeldern*, Studienarbeit, Lehrstuhl für Mustererkennung (Informatik 5), Universität Erlangen–Nürnberg, Erlangen, 1995.