Erschienen in: »Mustererkennung 1998« - Tagungsband zum
20. DAGM-Symposium, S. 385-392, 29. 9. - 1. 10. 1998, Stuttgart

Konturbasierte Personenlokalisation mittels dreidimensionaler neuronaler Felder und steuerbarer Filter

Ulf-Dietrich Braumann, Andrea Corradini*,
Hans-Joachim Böhme und Horst-Michael Groß**

Fachgebiet Neuroinformatik, Technische Universität Ilmenau
PF 10 05 65, 98684 Ilmenau
{ulf,andreac,hans,homi}@informatik.tu-ilmenau.de
http://cortex.informatik.tu-ilmenau.de/

Zusammenfassung In dieser Arbeit wird ein Verfahren zur Lokalisation von Personen innerhalb unpräparierter visueller Szenen anhand typischer Konturen vorgestellt. Es wird sich dabei auf die äußere Kontur frontal ausgerichteter Personen im Bereich von Kopf und Schultern bezogen (Kopf-Schulter-Partie). Diese Kontur wird approximiert durch ein *räumlich verteiltes Arrangement steuerbarer (steerable) Filter*, das auf einer Anzahl von pyramidal abgestuften Auflösungen eines Bildes angewendet wird.

Unter den Bedingungen realer Umwelten führt der Einsatz dieses Filterarrangements allerdings häufig zu einem unscharfen und teilweise artefaktbehafteten Ergebnis, so daß ein Kriterium zur Selektion als Bedingung für eine gute Lokalisationsleistung benötigt wird. Daher werden zur Bildung einer optimalen Lokalisationsentscheidung *dreidimensionale neuronale Felder* verwendet, innerhalb derer eine dynamische Musterbildung mit dem Ziel der Selektion *einer* Person erfolgt. Mittels einfacher homogener feldinterner Wechselwirkungsmechanismen können robuste Lokalisationsentscheidungen (Richtung und Entfernung) gefunden werden. Die Aktivitäten der Feldneuronen können als interne Zustände angesehen werden, die eine permanente Lokalisation (und damit eine visuelle Verfolgung der Person) wesentlich vereinfachen.

1 Einleitung

Für ein visuell basiertes Gesten-Erkennungssystem, das auf unserer mobilen und autonom navigierenden Roboterplattform *MILVA* eingesetzt wird (siehe dazu auch [3]), hat die Lokalisation eines Nutzers essentielle Bedeutung. Insbesondere spielt die sichere visuelle Lokalisation des Kopfes dieses Nutzers eine wichtige

* gefördert durch ein TMR-Stipendium der Europäischen Union: Marie Curie Research Training Grant # ERB FMBI CT 97 2613

** gefördert durch das Thüringer Ministerium für Wissenschaft, Forschung und Kultur (TMWFK) im Rahmen des Projektes *GESTIK*

Rolle für alle weiteren gestenbezogenen Analysen. In diesem Kontext einer gestenbasierten Mensch-Maschine-Interaktion behandelt diese Arbeit die visuelle Lokalisation von Personen unter realen Umweltbedingungen. Dabei ist davon auszugehen, daß eine solche Lokalisationsaufgabe insbesondere dann relevant wird, wenn eher größere Entfernungen zwischen Roboter und Zielperson vorliegen. Daher rührt die Forderung her, daß die verwendeten Merkmale mit der zur Verfügung stehenden Ortsauflösung erfaßbar sein müssen, weswegen Details wie etwa Gesichtsstrukturen nur teilweise geeignet sind, sofern mit relativ großen Operationsentfernungen zu rechnen ist. Damit erweitert das hier vorgestellte Verfahren unsere bisherigen methodischen Ansätze (siehe [2] zur Lokalisation von Personen, besonders hinsichtlich des nutzbaren Entfernungsbereichs.

Wir sind der Ansicht, daß insbesondere unter Realwelt-Bedingungen die äußere Kontur einer Person die geeignetste Invariante darstellt. Unser einfaches Konturmodell besteht aus einem Arrangement orientierter Filter, das eine stückweise Approximation des Umrisses der Kopf-Schulter-Partie von frontal zur Kamera ausgerichteten Personen vornimmt. Die relevanten Positionen der steuerbaren Filter im Arrangement wurden basierend auf einem Satz von Portraits bestimmt. Durch die Anwendung dieses Filterarrangements auf mehrere Auflösungsstufen eines Bildes kann die Lokalisation eine Entfernungsbestimmung der Person einschließen.

Das zentrale Problem, die für die Personenlokalisation auffälligste Bildposition zu finden, wird mit einem *dreidimensionalen dynamischen neuronalen Feld* gelöst, das in einem iterativen Selektionsprozeß eine abgegrenzte Region aktiver Neuronen (Aktivitätsblob) herausbildet (dynamische Musterbildung, siehe [1]).

2 Arrangements steuerbarer Filter

2.1 Motivation und verwandte Arbeiten

Unser Ansatz beruht auf einer Beschreibung der äußeren Kontur von Kopf und Schultern. Demgegenüber stellen Oren et al. in [9] einen interessanten verwandten Ansatz vor, der sich unter Verwendung sog. *Haar Wavelet Templates* auf den kompletten Umriß von Personen (Fußgänger) bezieht. Der gemeinsame Aspekt beider Ansätze besteht darin, daß räumlich verteilte orientierte Filter genutzt werden, mit denen jeweils ein Orientierungsmaß innerhalb einer begrenzten Region (lokale Orientierung) bestimmt wird. Interessante Bezüge bestehen auch zu Arbeiten von Meller [8], der u. a. mittels verschiedener gitterartig angeordneter Gaborfilter bestimmte Teilmerkmale auf Personen extrahiert, deren lokale Relationen zueinander zur Personenlokalisation verwendet werden.

Die Idee des konturbasierten Ansatzes zur Bildung eines Auffälligkeitsmaßes fußt sowohl auf physiologischen Erkenntnissen als auch auf psychophysischen Effekten. Der visuelle Kortex besteht in verschiedenen Abschnitten aus Zellen, die orientierte rezeptive Felder aufweisen. Bezogen auf eine bestimmte retinale Position kann damit ein breiter Frequenzbereich abgedeckt werden. Eine Vielzahl von Untersuchungen hat gezeigt, daß das Profil der rezeptiven Felder von einfachen

Zellen im primären visuellen Kortex von Säugern durch bestimmte zweidimensionale mathematische Funktionen modelliert werden kann. *Gaborfunktionen*, mit denen Jones und Palmer [6] arbeiten, sowie *Gaußfunktionen* (einschl. Ableitungen niederer Ordnung), die Koenderink und van Doorn verwenden [7], erscheinen insbesondere zur Modellierung typischer Profile visueller rezeptiver Felder als geeignet. Derartige lokale Operatoren ermöglichen es, eine spektrale Zerlegung der visuellen Information vorzunehmen.

Psychophysische Aspekte des konturbasierten Lokalisationsansatzes, wie bspw. *gute Fortsetzung* und *Symmetrie*, die beide in den Gestaltgesetzen wiederzufinden sind, beschreiben offensichtlich Effekte, die aufgrund von Gruppierungsmechanismen zustande kommen können.

Da wir mittels des Filterarrangements eine Modellierung einer bestimmten Kontur vornehmen wollen, indem diese stückweise mit speziell orientierten Filtern approximiert wird, benötigen wir dafür unter Umständen eine Vielzahl verschieden orientierter Filter. Da das sehr bei diskreter Realisierung sehr berechnungsaufwendig wäre, verwenden wir dafür steuerbare Filter [5].

2.2 Bestimmung des Konturverlaufs

In unseren bisherigen Arbeiten [4] wurde ein heuristisch definiertes Konturmodell eingesetzt. Es war manuell entworfen und beschränkte sich auf vier Orientierungen. Es ist allerdings zu erwarten, daß unter Verwendung von mehr Orientierungen ein präziseres Modell bestimmt werden kann. Das macht es jedoch erforderlich, das Konturmodell eng an *reales Bildmaterial* anzulehnen.

Steuerbare Filter haben die angenehme Eigenschaft, daß eine initial feste Anzahl von Faltungen ausreicht, beliebige Orientierungsinformation aus einem Bild zu extrahieren. Sie erlauben, eine deutlich größere Anzahl von Orientierungen zu schätzen und dabei auf zusätzliche Faltungen zu verzichten.

Unser Datenmaterial besteht aus unter identischen Bedingungen vor einem homogenen Hintergrund aufgenommenen Portraits von insgesamt zehn Personen. Jede Person wurde direkt frontal sowie jeweils $(+10°)$ und $(-10°)$ davon abweichend aufgenommen. Um ein symmetrisches Konturmodell zu erhalten, wurde aus jedem dieser Bilder noch eine vertikal gespiegelte Kopie erzeugt, so daß insgesamt 60 Bilder zur Verfügung standen. Diese Grauwertbilder im Format 256×256 wurden entsprechend tiefpaßgefiltert und auf eine Größe von 16×16 unterabgetastet. Anschließend wurden mit einem Sobel-Operator die Kanten dieser Bilder hervorgehoben und schließlich über alle diese Kantenbilder gemittelt. Mit einer einfachen Schwellenoperation erfolgte durch eine Binarisierung die Bestimmung eines repräsentativen Konturverlaufs, der im Mittel einer Kontur der Kopf-Schulter-Partie der Personen aus dem Datensatz entspricht. Die optimale Schwellenhöhe wurde bestimmt, indem sie in einem automatisierten Verfahren solange angehoben wurde, bis durch die Binarisierung nur noch *ein einzelnes, lückenlos zusammenhängendes, binäres Kontursegment* entstand.

Damit liegt der interessierende Konturverlauf in Form einer 16×16-Binärmatrix $\mathbf{\Lambda}$ vor. Was noch offen ist, ist die lokale Orientierung der einzelnen Konturabschnitte. Wir bestimmen sie mittels steuerbarer Filter (s. u.), die als Faltungsoperatoren auf die Binärmatrix $\mathbf{\Lambda}$ angewendet werden. Somit wird korrespondierend zu jedem Element $\lambda_{i,j}$ ein dominanter Orientierungswinkel bestimmt.

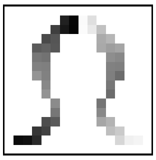

Abb. 1. Der ermittelte Konturverlauf $\mathbf{\Lambda}$: die Orientierungswinkel sind in Grauwerten kodiert (0°: schwarz; 90°: mittelgrau; 180°: weiß). Man beachte, daß im Scheitelbereich Übergänge zwischen 180° und 0° auftreten. Die Kontur ist symmetrisch, da die Originalbilder jeweils zusätzlich gespiegelt vorlagen.

2.3 Die Anwendung steuerbarer Filter

Es wurde bereits darauf verwiesen, daß die Orientierungswinkel entlang der Kontur mittels steuerbarer Filter bestimmt werden. Diese steuerbaren Filter werden hierfür *ohne* Vorzugsrichtung eingesetzt und aus dem Ergebnis wird in einem Schritt die dominante lokale Orientierung bestimmt. Diese Maximierung könnte bspw. auch unter Nutzung von Gaborfiltern vorgenommen werden, jedoch wäre dies nur mit einer Vielzahl von diskreten Faltungen möglich, da für jede beliebige Richtung ein entsprechendes Quadraturpaar an Gaborfiltern benötigt würde. Je feiner man die Richtungsauflösung wählt, desto höher wird der Rechenaufwand. Daher ist es notwendig, einen akzeptablen Kompromiß zwischen Genauigkeit der Orientierungsbestimmung und der Anzahl der Filter und damit der Faltungen zu finden.

In dieser Arbeit gehen wir mit dem Einsatz von *steuerbaren Filtern* (siehe dazu z. B. die Arbeiten von Freeman und Adelson [5]) einen viel einfacheren Weg. Mit ihnen wird eine flexible und gleichzeitig effektive Orientierungsfilterung möglich, indem ein bestimmter Satz von wenigen, verschieden orientierten *Basisfiltern* eingesetzt wird, anhand derer eine Interpolation in beliebiger Richtung erfolgen kann. Daher sind steuerbare Filter insgesamt effizienter als nicht steuerbare, wie etwa Gaborfilter.

Im allgemeinen gilt eine Funktion $f(\cdot)$ als steuerbar, wenn die folgenden zwei Bedingungen erfüllt sind. Erstens muß die Menge der Basisfilter aus insgesamt M von $f(\cdot)$ gewonnenen, jedoch rotierten Kopien $f^{\alpha_1}(\cdot)\ldots f^{\alpha_M}(\cdot)$ mit festen Rotationswinkeln $\alpha_1 \ldots \alpha_M$ bestehen. Zweitens muß eine weitere, um einen beliebigen Winkel ϑ rotierte Kopie $f^{\vartheta}(\cdot)$ ebenfalls aus der Überlagerung der Funktionen der Basisfilter multipliziert mit Interpolationsfunktionen $k_j(\vartheta)$ zu bilden sein als

$$f^{\vartheta}(\cdot) = \sum_{j=1}^{M} k_j(\vartheta) f^{\alpha_j}(\cdot) \qquad (1)$$

Für unsere Arbeiten verwenden wir als Basisfilter ein Quadraturpaar aus der zweiten Richtungsableitung einer Gaußfunktion und einer Approximation ihrer Hilberttransformierten mittels eines Polynoms dritter Ordnung, das eine Gaußfunktion moduliert. Ausgehend vom Steuerbarkeitstheorem (siehe [5]) sind diese beiden Funktionen jeweils steuerbar. Man benötigt dafür $M = 7$ Basisfunktionen. Um nun die stückweise Orientierung entlang unserer betrachteten Kontur zu messen, verwenden wir die phasenlageunabhängige Summe der Quadrate beider Filterergebnisse. Dieses quadrierte Filterergebnis repräsentiert als eine Funktion der Filterorientierung am Punkt (x, y) eine *orientierte Energie* $E^{(x,y)}(\vartheta)$. Aus Gründen der jeweiligen Symmetrie der Funktionen des Filterpaars ist die Energie an jedem Punkt periodisch mit der Periode π.

Um eine genaue Schätzung der *dominanten* lokalen Orientierung zu erhalten, könnte man punktweise nun eine Maximierung der orientierten Energie vornehmen, indem man durch stetige Variation von ϑ für $\vartheta_{MAX}^{(x,y)} = \arg\max\{E^{(x,y)}(\vartheta) \mid \vartheta \in [0, \pi)\}$ eine optimale Lösung findet. Allerdings können wir auf eine iterative Suche des Maximums verzichten, da für dieses Problem von Freeman und Adelson in [5] bereits eine analytische Lösung zur Energiemaximierung zur Verfügung gestellt wurde. Im weiteren bezeichnen wir die Matrix, die alle zu einem Bildausschnitt korrespondierenden Orientierungswinkel $\vartheta_{MAX}^{(x,y)}$ enthält, als $\mathbf{\Theta}$.

Zusammenfassend kann gesagt werden, daß anders als beim Einsatz von Gaborfiltern bei Verwendung von steuerbaren Filtern keine zusätzlichen Faltungen im Anschluß an eine einmalige Faltung mit allen Basisfiltern erforderlich ist. Außerdem haben wir uns für diese steuerbaren Filter entschieden, da alle Basisfilter kartesisch separierbar sind, was den Rechenaufwand weiter erheblich reduziert.

3 Dynamische Neuronale Felder zur Lokalisation

3.1 Bestimmung der Eingabe des Neuronalen Feldes

Der vorangegangene Abschnitt beschreibt die Theorie und die Anwendung von steuerbaren Filtern. Mittels dieser Filter berechnen wir sowohl die Matrix $\mathbf{\Lambda}$, die den typischen Konturverlauf der Kopf-Schulter-Partie beschreibt, als auch die Matrix $\mathbf{\Theta}$, die zu dem jeweiligen Bildausschnitt korrespondiert, in dem eine Person zu detektieren ist und die die dominanten lokalen Orientierungswerte enthält.

Im folgenden wird ein Maß für das Vorhandensein der *visuellen Auffälligkeitskomponente Kopf-Schulter-Partie*, die durch den Faltungskern $\mathbf{\Lambda}$ repräsentiert wird, innerhalb der Matrix $\mathbf{\Theta}$ gebildet. Dafür nutzen wir ein Matchingverfahren, das auf einem *Ähnlichkeitsmaß* $m^{(x,y)}$ basiert. Wegen der π-Periodizität des Filterergebnisses der steuerbaren Filter einerseits, und um andererseits geeignet die „Ähnlichkeit" zweier Elemente von $\mathbf{\Lambda}$ und $\mathbf{\Theta}$ zu bestimmen, benötigt die Ähnlichkeitsfunktion dieselbe Periodizität.

$$m^{(x,y)} = \frac{\displaystyle\sum_{i=0}^{I-1}\sum_{\substack{j=0\\ \lambda_{i,j}\neq 0}}^{J-1} \frac{1}{2}\left[\cos\left(2\left|\lambda_{i,j} - \vartheta_{MAX}^{(x+i-\frac{I}{2},y+j-\frac{J}{2})}\right|\right) + 1\right]}{\operatorname{card}\left(\operatorname{supp}\left(\mathbf{\Lambda}\right)\right)} \tag{2}$$

Hierin bezeichnet $\lambda_{i,j}$ ein Element aus $\mathbf{\Lambda}$ mit den Indizes (i, j) sowie $\vartheta_{MAX}^{(x+i-\frac{I}{2},y+j-\frac{J}{2})}$ ein Element aus $\mathbf{\Theta}$ an der Stelle $(x+i-\frac{I}{2}, y+j-\frac{J}{2})$. $I = J = 16$ sind die Dimensionen der Matrix $\mathbf{\Lambda}$. Die Normierung auf die Kardinalität des Supports von $\mathbf{\Lambda}$ (der Support einer Matrix selektiert nur diejenigen Elemente daraus, die von Null verschieden sind) sichert $m^{(x,y)} \in [0,1]$ für die nachfolgenden Verarbeitungsschritte.

3.2 Das dreidimensionale nichtlineare dynamische Feld

Eine Lokalisation kann umso sicherer erfolgen, je stärker es abgegrenzte Häufungen von hohen Werten des Ähnlichkeitsmaßes gibt. Um eine akzeptable Lokalisation von Personen erzielen zu können, reicht aber die Bildung des Ähnlichkeitsmaßes allein nicht aus, sondern für eine definierte, sichere Lokalisationsentscheidung bedarf es eines *Selektionsmechanismus*. Dynamische neuronale Felder eignen sich gut für eine dynamische Selektion. Sie können mit einfachen homogenen feldinternen Wechselwirkungsmechanismen selektive einzelne, eng abgegrenzte Gruppen aktiver Neuronen herausbilden (sog. Blobs), deren Position im Feld als Lokalisationsergebnis genutzt werden kann.

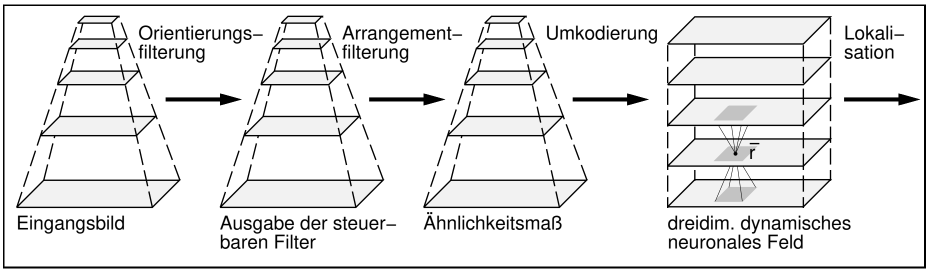

Abb. 2. Die Verarbeitungsstufen zur Personenlokalisation: Die Repräsentation eines Bildes in mehreren Auflösungsstufen wird ebenenweise mit einem Satz steuerbarer Filter verarbeitet. Mit dem anschließenden Einsatz des Filterarrangements wird ein Ähnlichkeitsmaß gebildet, was (nach Umkodierung) als Stimulus in ein dreidimensionales Feld dynamischer Neuronen geleitet wird. Der resultierende Blob (lokal abgegrenztes Muster aktiver Neuronen) wird zur Lokalisierung einer Person genutzt.

Für unsere Arbeiten haben wir ein derartiges Feld aus einer Anzahl von Schichten aufgebaut, da wir insgesamt fünf Auflösungsstufen verarbeiten (siehe Abb. 2). Somit wird ein dreidimensionales Feld aufgespannt, innerhalb dessen die Selektion abläuft. Daher ist es möglich, die Lokalisation auch hinsichtlich verschiedener Distanzen vorzunehmen. Das dynamische neuronale Feld F kann als rückgekoppeltes nichtlineares dynamisches System beschrieben werden. Ziel der Selektion ist die Ausbildung eines *einzelnen* lokalen Blobs aktiver Neuronen, die sich bei Annäherung an einen Gleichgewichtszustand gegenüber den übrigen behaupten können. Die folgende Gleichung beschreibt das System:

$$\tau \frac{d}{dt} z(\boldsymbol{r},t) = -z(\boldsymbol{r},t) - \mathrm{c}_h h(t) + \mathrm{c}_i x(\boldsymbol{r},t) + \mathrm{c}_l \int_N w(\boldsymbol{r}-\boldsymbol{r}\,')y(\boldsymbol{r}\,',t)d^3\boldsymbol{r}\,' \qquad (3)$$

Hierbei steht $\boldsymbol{r}$ für eine dreidimensionale Koordinate einer Neuronenposition im Feld, $z(\boldsymbol{r},t)$ ist die Aktivierung des Neurons $\boldsymbol{r}$ zur Zeit t, $y(\boldsymbol{r},t)$ ist die Ausgabeaktivität dieses Neurons, die mittels einer Sigmoidfunktion bestimmt wird,

$x(\boldsymbol{r}, t)$ bezeichnet den externen Input (korrespondierend zum umkodierten Ähnlichkeitsmaß $m^{\boldsymbol{r}}$, siehe Gleichung 2), $h(t)$ ist eine feldglobale Inhibition zur Zeit t, die alle Neuronenaktivitäten im gesamten Feld $F \subseteq \mathbb{R}^3$ aufsummiert. $w(\boldsymbol{r} - \boldsymbol{r}')$ ist eine Funktion vom Mexikanerhut-Typ. Sie gewichtet die laterale Aktivierung eines Neurons $\boldsymbol{r}$ aus seiner Nachbarschaft $N \subseteq \mathbb{R}^3$. Für ein ausgewähltes $\boldsymbol{r}$ ist N in Abb. 2 (rechts) dunkel markiert. Die Konstanten c_h, c_l und c_i sind Parameter des Systems.

Wie auch in Abb. 2 dargestellt ist, ist zu beachten, daß innerhalb des Feldes die lokalen Korrespondenzen zwischen den Ebenen nicht verletzt werden. Deshalb nehmen wir eine Umkodierung in eine quaderförmige Struktur vor. Ein Nebeneffekt davon ist folgender: je gröber die Auflösungsebene in der Pyramide wird, desto räumlich unschärfer wird das Ähnlichkeitsmaß auf der Kubusschicht und desto ausgedehnter treten dementsprechend lokale Häufungen als Stimuli des dynamischen Feldes auf. Allerdings wurde dieser Effekt von uns nicht gesondert kompensiert. Es war lediglich zu beobachten, daß in den Schichten des Feldes, deren Input aus groben Auflösungsebenen entstammt, einige Schritte mehr zur Herausbildung eines Blobs bzw. Bestandteil eines Blobs benötigt werden, so daß die Anzahl der insgesamt benötigten Iterationsschritte bis zum Erreichen eines Gleichgewichts dadurch etwas erhöht werden kann.

3.3 Ergebnisse

Ein Ergebnis zur Personenlokalisation wir in Abb. 3 illustriert. Die rechte Spalte darin zeigt den Zustand aller fünf Schichten des dynamischen neuronalen Feldes in einer Momentaufnahme, zu der die Aktivitätsänderung des Neurons mit der höchsten Aktivität kleiner als 1% wurde. Im Mittel benötigt das System 11 Iterationszyklen bei Anwendung eines Eulerverfahrens zur numerischen Lösung des Differentialgleichungssystems. Die Ausdehnung des Blobs ist nicht nur auf eine Schicht des Feldes beschränkt. Um eine präzisere Entfernungsbestimmung vornehmen zu können, wäre es auch möglich, die genaue Lage des Blobzentrums zwischen zwei Schichten durch Interpolation zu präzisieren, was im Rahmen dieser Arbeit jedoch nicht realisiert wurde.

Die von uns vorgestellten Ergebnisse sind beispielhaft, wobei wir uns aus Platzgründen auf ein Beispiel beschränken mußten. Eine generelle klare quantitative Bewertung des Verfahrens erscheint uns nicht unproblematisch, da für Art und Komplexität von Person und jeweiligem Hintergrund ein geeignetes Bewertungsmaß benötigt wird. Der hier vorgestellte Ansatz erweist sich aus unserer Sicht als eine gute Lösung für das Problem der Lokalisation von Personen, insbesondere im Hinblick auf seine Toleranz gegenüber Änderungen von Umgebungsbedingungen. Seine Neuheit besteht im Einsatz eines dreidimensionalen dynamischen neuronalen Feldes, das einen wesentlichen Beitrag zur Robustheit des Verfahrens erbringt.

Danksagung Die Autoren danken Anja Brakensiek und Markus Krabbes vom FG Neuroinformatik der TU Ilmenau sowie Heiko Kempe vom Zentrum für Bild- und Signalverarbeitung e. V. Ilmenau für theoretische und praktische Anregungen.

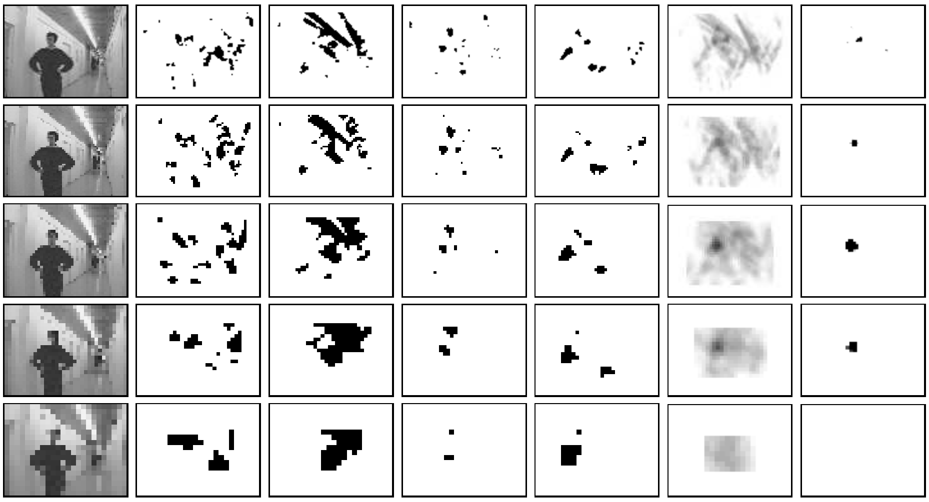

Abb. 3. Lokalisationsergebnis im Innenbereich: Die Lokalisation der Person erfolgt nicht scharf auf einer der Pyramidenebenen, der entstehende räumliche Blob (Spalte ganz rechts) ist am stärksten auf der mittleren der fünf Ebenen ausgebildet. In jeder Zeile sind alle Zwischenergebnisse und Ergebnisse für eine der fünf Auflösungsstufen (Abstand $1/\sqrt{2}$) dargestellt. Die sieben Spalten enthalten folgendes: Input, Ergebnis der Orientierungsfilterung mit steuerbaren Filtern (hier aus Darstellungsgründen beschränkt auf vier Richtungsintervalle zentriert um 0°, 45°, 90° und 135°), das Ergebnis der Anwendung des Filterarrangements (Ähnlichkeitsmaß) und schließlich das Ergebnis der Selektion innerhalb des dreidimensionalen Feldes dynamischer Neuronen.

Literatur

1. Amari, S.: Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields. Biological Cybernetics **27** (1977) 77–87
2. Brakensiek, A., Braumann, U.-D., Böhme, H.-J. und Gross, H.-M.: Farb- und strukturbasierte neuronale Verfahren zur Lokalisierung von Gesichtern in Real-World-Szenen. Mustererk. 1997 – Tagungsband 19. DAGM-Symp. (1997) 113–120
3. Boehme, H.-J., Brakensiek, A., Braumann, U.-D., Krabbes, M. und Gross, H.-M.: Neural Networks for Gesture-based Remote Control of a Mobile Robot. Proc. IEEE Intern. Joint Conf. on Neural Networks (1998) Bd. 1, 372–377
4. Corradini, A., Braumann, U.-D., Brakensiek, A., Krabbes, M., Boehme, H.-J. und Gross, H.-M.: Visual Gesture Localization with Dynamic Neural Fields: Towards a Gesture Recognition System. Proc. X Italian Workshop on Neural Nets (1998)
5. Freeman, W. T. und Adelson, E. H.: The Design and Use of Steerable Filters. IEEE Transactions on Pattern Analysis and Machine Intelligence **13** (1991) 891–906
6. Jones, J. P. und Palmer, L. A.: An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex. Journal of Neurophysiology **58** (1987) 1233–1258
7. Koenderink, J. J. und van Doorn, A. J.: Receptive Field Families. Biological Cybernetics **63** (1990) 291–297
8. Meller, S.: Segmentierungsfreie Personendetektion. Diplomarbeit, Universität Erlangen-Nürnberg (1997)
9. Oren, M., Papageorgiou, C., Sinha, P., Osuna, E. und Poggio, T.: Pedestrian Detection Using Wavelet Templates. Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (1997) 193–199