

(12)

Patentschrift

(21) Anmeldenummer: A 1284/2004 (51) Int. Cl.⁷: **G06K 9/00**
 (22) Anmeldetag: 2004-07-28 G06T 7/40, G06K 9/80
 (42) Beginn der Patentdauer: 2005-07-15
 (45) Ausgabetag: 2007-01-15

(56) Entgegenhaltungen:
 SUN Z., BEBIS G., MILLER R.,
 "IMPROVING THE PERFORMANCE OF
 ON-ROAD VEHICLE DETECTION BY
 COMBINING GABOR AND WAVELET
 FEATURES", IN: INTELLIGENT
 TRANSPORTATION SYSTEMS, 5TH
 INTERNATIONAL CONFERENCE,
 SINGAPUR, IEEE 2002, SEITEN 130-135

(73) Patentinhaber:
 X-PIN.COM GMBH
 A-1060 WIEN (AT).
 (72) Erfinder:
 TRAXLER JOHANNES DIPL.ING.
 ZWETTL, NIEDERÖSTERREICH (AT).

(54) VERFAHREN ZUR OBJEKTERKENNUNG

(57) Ein Verfahren zur Objekterkennung, insbesondere im zweiten Schritt auf Basis von in der Mehrzahl gekrümmten Gabor-Wavelets ($\psi_{B,j}$; B_B) erzeugt werden.

in einem ersten Schritt im Objektbild (P_{obj}) die aktuellen Koordinaten ($Im_{obj,i}$) der signifikanten Punkte (Im_i) ermittelt werden, indem jeweils im Umgebungsbereich einer Schätzung ($Im_{est,i}$) der Koordinaten die Koordinaten jenes Punktes mit der maximalen Ähnlichkeit seines Wavelet-Jets (J'_i) zu den Wavelet-Jets ($J_{bib,i,r}$) der signifikanten Punkte ($Im_{man,i,r}$) des Referenzbildes ($P_{bib,r}$) ermittelt werden, und
 in einem zweiten Schritt die Erkennung des Objektes durch Ähnlichkeitsvergleich zwischen den Wavelet-Jets ($J_{obj,i}$) der lokalisierten signifikanten Punkte ($Im_{obj,i}$) des Objektbildes (P_{obj}) und den Wavelet-Jets ($J_{ref,i,s}$) der signifikanten Punkte ($Im_{man,i,s}$) des Referenzbildes ($P_{ref,s}$) erfolgt,

zeichnet sich dadurch aus, daß die Wavelet-Jets ($J_{est,i}$, $J_{bib,i,r}$, J'_i) im ersten Schritt auf Basis von Haar-Wavelets ($\psi_{H,j}$; B_H) und die Wavelet-Jets ($J_{obj,i}$, $J_{ref,i,s}$)

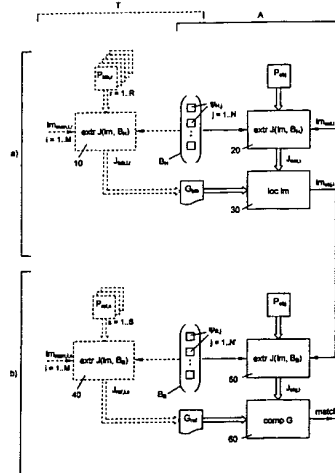


Fig. 1

Die vorliegende Erfindung betrifft ein Verfahren zur Objekterkennung, insbesondere Gesichtserkennung, in einem Objektbild durch Vergleichen von Wavelet-Jets signifikanter Punkte des Objektbildes mit Wavelet-Jets zumindest eines Referenzbildes.

5 Ein derartiges Verfahren ist in der Arbeit von David S. Bolme, „Elastic Bunch Graph Matching“, Thesis MSc, Colorado State University, Sommer 2003, beruhend u.a. auf der Arbeit von Laurenz Wiskott, Jean-Marc Fellous, Norbert Kruger und Christoph von der Malsburg, „Face Recognition by Elastic Bunch Graph Matching“, Tech. Report 96-08, Ruhr-Universität Bochum, April 1996, beschrieben.

10

Bei den bekannten Verfahren wird als Basis für die Erzeugung der Jets ein Satz von geradlinigen Gabor-Wavelets verwendet, und zwar sowohl im ersten Schritt der Lokalisierung der signifikanten Punkte als auch im zweiten Schritt des Ähnlichkeitsvergleiches mit dem Referenzbild. Es hat sich gezeigt, daß praktische Implementierungen dieser Verfahren eine überaus hohe Rechenleistung erfordern und dennoch eine verhältnismäßig hohe Fehlerrate bei der Objekterkennung aufweisen, was einem breiteren industriellen Einsatz, z.B. in der Sicherheitstechnik, entgegensteht.

15

Die Erfindung setzt sich zum Ziel, ein Verfahren der genannten Art so weiterzubilden, daß es eine geringere Fehleranfälligkeit und höhere Robustheit besitzt als die bekannten Lösungen, bei gleichzeitig geringerem Rechenleistungsbedarf. Dieses Ziel wird mit einem Verfahren der einleitend genannten Art erreicht, welches sich gemäß der Erfindung dadurch auszeichnet, daß die Wavelet-Jets im ersten Schritt auf Basis von Haar-Wavelets und im zweiten Schritt auf Basis von in der Mehrzahl gekrümmten Gabor-Wavelets erzeugt werden.

20

Die Erfindung beruht auf der Erkenntnis, daß sowohl die Verarbeitungsgeschwindigkeit als auch die Treffsicherheit der Objekterkennung wesentlich gesteigert werden können, wenn die beiden Schritte des Verfahrens, Lokalisation und Ähnlichkeitsvergleich, auf vollkommen getrennte, jeweils speziell angepaßte Wavelet-Basen abgestellt werden. Der Einsatz von einfachen Haar-Wavelets im Lokalisierungsschritt ermöglicht ein fehlertolerantes, robustes und rasches Auffinden der signifikanten Punkte, während die anschließende Auswertung der aufgefundenen Punkte mit Hilfe von detailreichen gekrümmten Gabor-Wavelets genau und treffsicher erfolgt. Die Fehlerrate des Verfahrens kann damit wesentlich gesenkt werden. Die Jet-Generierung auf Basis von Haar-Wavelets benötigt nur äußerst geringe Rechenleistung, wobei die erzielte Einsparung den zusätzlichen Rechenleistungsbedarf für die gesonderte Jet-Extraktion im zweiten Schritt bei weitem überwiegt. Der Rechenbedarf des gesamten Verfahrens ist damit wesentlich geringer als bei den bekannten Verfahren, und dies bei gleichzeitig höherer Treffsicherheit. Der geringe Rechenleistungsbedarf ermöglicht erstmals praktische Realisierungen mit langsamen, lüfterlosen Prozessoren, was einen industriellen Einsatz in breiterem Umfang, beispielsweise für Sicherheitsschleusen oder Schlösser mit Onboard-Prozessoren, in greifbare Nähe rückt.

25

30

35

40

Eine bevorzugte Ausführungsform der Erfindung zeichnet sich dadurch aus, daß für jeden signifikanten Punkt die Koordinatenermittlung im ersten Schritt durch Ähnlichkeitsvergleich mit einem Bündel von ersten Wavelet-Jets verschiedener Referenzbilder erfolgt. Ein solcher Bündelvergleich ist aus den genannten Arbeiten von Bolme und Wiskott et al. an sich bekannt und verbessert die Lokalisierung der signifikanten Punkte in dem zu untersuchenden Objektbild.

45

Besonders vorteilhaft ist es, wenn gemäß einem weiteren Merkmal der Erfindung als Haar-Wavelet-Basis eine Lienhart-Maydt-Basis verwendet wird, bevorzugt in mehreren Skalierungsschritten, besonders bevorzugt in 24 Skalierungsschritten.

50

Unter einer Lienhart-Maydt-Basis wird in der vorliegenden Beschreibung der von Rainer Lienhardt und Jochen Maydt in der Arbeit „An Extended Set of Haar-Like Features for Rapid Object Detection“, Intel Laboratories, 2002, vorgeschlagene Satz von 14 bzw. 15 Haar-Funktionen zur schnellen Objekterkennung verstanden, welcher u.a. auf Arbeiten von Papageorgiou et al. und

55

Viola et al. zurückgeht.

Bevorzugt wird die Lienhart-Maydt-Basis in mehreren Skalierungsschritten angewandt, um die Genauigkeit der Lokalisierung zu erhöhen, ohne den Rechenleistungsbedarf des gesamten Verfahrens signifikant zu erhöhen. In der Praxis hat sich ein Wert von 24 Skalierungsschritten als besonders günstig erwiesen.

Ein weiteres bevorzugtes Merkmal der Erfindung besteht darin, daß als Gabor-Wavelet-Basis Wavelets der Funktion

$$\psi(x, y, C, \Theta, \lambda, \varphi, \sigma, \gamma) = e^{-\frac{\dot{x}^2 + \gamma^2 \dot{y}^2}{2\sigma}} \cos\left(2\pi \frac{\dot{x}}{\lambda} + \varphi\right)$$

mit

$$\dot{x} = x \cos(\Theta) + y \sin(\Theta) + y^2 C,$$

$$\dot{y} = -x \sin(\Theta) + y \cos(\Theta)$$

und dem Parametersatz

$$C = \{0, \pm 0.04, \pm 0.07, \pm 0.1\},$$

$$\Theta = \left\{0, \frac{\pi}{8}, \frac{2\pi}{8}, \frac{3\pi}{8}, \dots, \frac{7\pi}{8}\right\},$$

$$\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\},$$

$$\varphi = \left\{0, \frac{\pi}{2}\right\},$$

$$\sigma = \lambda,$$

$$\gamma = 1.0,$$

verwendet werden. Mit diesem Parametersatz ergibt sich eine Basis von 800 Gabor-Wavelets, von denen 700 gekrümmt sind. Dies führt zu einer äußerst detailreichen, genauen Analyse der Bildbereiche in der Umgebung der signifikanten Punkte. Aufgrund der gewählten Krümmungen ist die erfindungsgemäße Wavelet-Basis an Gesichtskonturen besonders gut angepaßt und damit für die Gesichtserkennung besonders gut geeignet.

Die Erfindung wird nachstehend anhand eines in den beigeschlossenen Zeichnungen dargestellten Ausführungsbeispiels näher erläutert. In den Zeichnungen zeigt

Fig. 1 ein Flußdiagramm des Verfahrens der Erfindung,

Fig. 2 eine Lienhart-Maydt-Basis von Haar-Wavelets und

Fig. 3 einige beispielhafte Wavelets der Basis von gekrümmten Gabor-Wavelets.

Das in Fig. 1 dargestellte Verfahren baut auf den genannten Arbeiten von Bolme, Wiskott et al. und Lienhart-Maydt auf und für eine genaue Erläuterung der theoretischen Grundlagen des Verfahrens wird auf diese Schriften verwiesen. Ziel des Verfahrens ist es, mittels digitaler Bildverarbeitung ein Objekt, z.B. ein Gesicht, ein Fahrzeug usw., in einem Bild P_{obj} des Objektes zu erkennen, und zwar durch Ähnlichkeitsvergleich mit einem oder mehreren Referenzbildern P_{ref}

des Objektes.

Der Ähnlichkeitsvergleich erfolgt nicht durch Vergleichen des gesamten Bildes P_{obj} mit dem gesamten Referenzbild P_{ref} , sondern durch Vergleichen von einzelnen Bildbereichen rund um
5 signifikante Punkte des Objektes, sog. „Landmarks“ Im_i . Im Falle der Gesichtserkennung sind geeignete signifikante Punkte bzw. Landmarks Im_i beispielsweise das Zentrum der Pupille des linken Auges, das Zentrum der Pupille des rechten Auges, die Nasenspitze, der linke und rechte Mundwinkel usw. Im weiteren werden M Landmarks Im_i ($i = 1..M$) betrachtet.

10 Der Bildbereich P rund um einen Landmark Im_i wird nicht direkt, sondern mittels Wavelet-Transformation in einem Transformationsraum J ausgewertet. Das Ergebnis der Wavelet-Transformation an einem Ort (x, y) wird als Jet J bezeichnet und kann als Vektor

$$J = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_j \\ \vdots \\ c_N \end{pmatrix}$$

dargestellt werden, dessen N komplexe Koeffizienten c_j den Faltungen des Bildes P am Ort
20 (x, y) mit einem Satz von N Wavelets ψ_j entsprechen:

$$c_j = C \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(x, y) \Psi_j(x, y) dx dy.$$

Die Menge aller N Wavelets ψ_j ($j = 1..N$) wird als Wavelet-Basis B der Wavelet-Transformation
30 bezeichnet.

Zurückkommend auf Fig. 1 ist das im Betrieb ablaufende Verfahren zur Objekterkennung als
Ablaufteil A („runtime“) in der rechten Hälfte von Fig. 1 dargestellt und gliedert sich in einen
35 ersten Schritt a), in welchem die Landmarks $Im_{obj,i}$ im Objektbild P_{obj} lokalisiert werden, und einen zweiten Schritt b), in welchem Jets $J_{obj,i}$ an den Landmarks $Im_{obj,i}$ extrahiert und mit Jets $J_{ref,i,s}$ eines oder mehrerer Referenzbilder $P_{ref,s}$ verglichen werden, um die Objekterkennung („match“) durchzuführen.

In einem vorhergehenden Trainingsteil T des Verfahrens werden die für die Schritte a) und b)
40 des Ablaufteiles A erforderlichen Referenzdaten ermittelt und bereitgestellt. Es versteht sich, daß der Trainingsteil T nur ein Mal für eine Vielzahl von Ablaufteilen A durchlaufen werden muß und auch in großem zeitlichen Abstand zu letzterem liegen kann; zum besseren Verständnis werden die Teile T und A jedoch im weiteren gleichzeitig erläutert.

45 Im Trainingsteil T werden für den Schritt a) in einem Satz von R Referenzbildern $P_{bib,r}$ ($r = 1..R$) in jedem einzelnen Referenzbild die Koordinaten $Im_{man,i,r}$ vorgegebener Landmarks Im_i manuell ermittelt, beispielsweise durch manuelles Markieren der Pupillenzentren, der Nasenspitze usw. mit Hilfe eines interaktiven Bildbearbeitungswerkzeuges.

50 Im Block 10 werden aus jedem Referenzbild $P_{bib,r}$ am Ort jedes Landmarks $Im_{man,i,r}$ Jets $J_{bib,i,r}$ generiert, und zwar auf Grundlage einer Wavelet-Basis B_H aus N Haar-Wavelets $\psi_{H,j}$ ($j = 1..N$).

Die Haar-Wavelets $\psi_{H,j}$ können an sich aus jeder beliebigen Haar-Funktion aufgebaut sein.
Bevorzugt wird jedoch als Haar-Wavelets $\psi_{H,j}$ eine Lienhart-Maydt-Basis von 15 bzw. 14 Haar-
55 Wavelets verwendet, welche in Fig. 2 in ihrer Struktur graphisch dargestellt ist und eine normier-

te Ausdehnung von 3 x 3 bis 3 x 5 Pixeln besitzt.

Die Haar-Wavelets $\psi_{H,j}$ der Lienhart-Maydt-Basis B_H von Fig. 2 können in beliebiger Skalierung verwendet werden; bevorzugt werden sie gleichzeitig in mehreren, insbesondere 24 Skalierungsschritten von 1- bis 24-facher Größe eingesetzt, so daß die gesamte Anzahl N an Haar-Wavelets $\psi_{H,j}$ in der Basis B_H $14 \times 24 = 336$ beträgt.

Das Ergebnis der Jet-Extraktion von Block 10 wird als Referenzdatensatz G_{bib} gespeichert und dem Ablaufteil A zur Verfügung. Der Referenzdatensatz G_{bib} kann als Matrix

$$G_{bib} = \begin{pmatrix} J_{1,1} & J_{1,2} & \dots & J_{1,r} & J_{1,R} \\ J_{2,1} & J_{2,2} & & & \\ \vdots & & & & \\ J_{i,1} & & & J_{i,r} & \\ J_{M,1} & & & & J_{M,R} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ b_M \end{pmatrix}$$

geschrieben werden und stellt gleichzeitig einen Spaltenvektor von M Bündeln b_i ($i = 1..M$) von Jets, sog. „bunches“, für jeden Landmark lm_i dar. Der Referenzdatensatz G_{bib} wird daher auch als „Bunch Graph“ bezeichnet.

Im Ablaufteil A sind die genauen Koordinaten (x, y) der gewählten Landmarks lm_i (linkes Pupillenzentrum, rechtes Pupillenzentrum usw.) im Objektbild P_{obj} zunächst nicht bekannt und werden durch Anfangswerte und/oder Schätzungen $lm_{est,i}$ ersetzt. Im Block 20 werden am Ort der geschätzten Landmarks $lm_{est,i}$ Jets $J_{est,i}$ extrahiert, und zwar auf Grundlage derselben Wavelet-Basis B_H , die im Block 10 zur Erzeugung der Referenzdaten verwendet und anhand von Fig. 2 erläutert wurde.

In einem anschließenden Block 30 werden die tatsächlichen bzw. aktuellen Koordinaten (x, y) der Landmarks lm_i des Objektbildes P_{obj} lokalisiert. Dazu werden die Jets $J_{est,i}$ an den geschätzten Landmarks $lm_{est,i}$ des Objektbildes P_{obj} mit den Jets $J_{bib,i,r}$ des Bunch Graphs G_{bib} verglichen.

Sofern mehrere Referenzbilder $P_{bib,r}$ zugrunde gelegt wurden ($R > 1$), d.h. der Bunch Graph G_{bib} aus Bündeln b_i von Jets für jeden Landmark lm_i besteht, wird jeweils der ähnlichste Jet aus einem Bündel b_i für einen bestimmten Landmark lm_i zum Vergleich herangezogen. Dies kann beispielsweise für den ersten Landmark lm_1 (z.B. „linkes Pupillenzentrum“) ein vom dritten Referenzbild $P_{bib,3}$ stammender Jet $J_{1,3}$ sein, für den zweiten Landmark lm_2 (z.B. „rechtes Pupillenzentrum“) ein vom ersten Referenzbild $P_{bib,1}$ stammender Jet $J_{2,1}$ usw., wodurch für den Vergleich das jeweils „passendste“ Auge, die jeweils „passendste“ Nase usw. herangezogen wird.

Der Vergleich im Block 30 kann auf verschiedene in der Technik bekannte Arten durchgeführt werden, sofern daraus die örtliche Abweichung des tatsächlichen Landmarks $lm_{obj,i}$ vom geschätzten Landmark $lm_{est,i}$ ermittelt werden kann. Beispielsweise könnten die Jets von allen Punkten in einer Umgebung des geschätzten Landmarks aus dem Objektbild extrahiert und mit den entsprechenden Referenz-Jets des Referenzbildes auf maximale Ähnlichkeit verglichen werden.

Ein von Bolme vorgeschlagenes, besonders rasch zu berechnendes Verfahren beruht auf einer Schätzung des Jets J' in einem Abstand \vec{d} vom geschätzten Landmark $lm_{est,i}$ durch Anwendung der Näherung, daß sich ein Jet J' in einem geringen Abstand von einem Jet J vorwiegend in seinem Phasenversatz unterscheidet:

In Polarkoordinatendarstellung

$$c_j = (a_j, \varphi_j)$$

kann ein Jet

5

$$J' = \begin{pmatrix} c_1 \\ \vdots \\ c_2 \\ \vdots \\ c_N \end{pmatrix}$$

10

in einem Abstand

15

$$\vec{d} = \begin{pmatrix} dx \\ dy \end{pmatrix}$$

von einem vorgegebenen Jet

20

$$J = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_j \\ \vdots \\ c_N \end{pmatrix}$$

25

30 angenähert werden zu

$$a_j \approx a'_j$$

$$\varphi'_j \approx \varphi_j + \bar{k}_j \cdot \vec{d}$$

35

mit

40

$$\bar{k}_j = \begin{pmatrix} \frac{2\pi \cos \Theta_j}{\lambda_j} \\ \frac{2\pi \sin \Theta_j}{\lambda_j} \end{pmatrix}$$

und

45

Θ_j ... Orientierung des Wavelets ψ_j

λ_j ... Wellenlänge des Wavelets ψ_j .

50

Durch Variieren des Abstandsvektors \vec{d} solange, bis der Umgebungsjet J' maximale Ähnlichkeit mit dem Jet J des gesuchten Landmarks lm_j hat, können der Abstandsvektor \vec{d} und damit die aktuellen Koordinaten $lm_{obj,i}$ des Landmarks lm_i im Objektbild P_{obj} ermittelt werden.

55

Das Variieren des Abstandsvektors \vec{d} kann mit allen in der Technik bekannten Suchalgorithmen durchgeführt werden, z.B. mittels Grid-Sample-, Predictive-Step-, Predictive-Iteration-, Fixed-Local-Search- oder Narrowing-Local-Search-Algorithmen, wie z.B. in der Arbeit von

Bolme erörtert.

Für den Ähnlichkeitsvergleich zweier Jets, wie der Jets J und J', kann ebenfalls jedes in der Technik bekannte Ähnlichkeitsmaß verwendet werden, beispielsweise ein betragsmäßiges Korrelationsmaß

$$S_a(J, J') = \frac{\sum_{j=1}^N a_j a'_j}{\sqrt{\sum_{j=1}^N a_j^2 \sum_{j=1}^N a'^2_j}}$$

oder ein auch den Phasenversatz berücksichtigendes Korrelationsmaß

$$S_\varphi(J, J') = \frac{\sum_{j=1}^N a_j a'_j \cos(\varphi_j - \varphi'_j)}{\sqrt{\sum_{j=1}^N a_j^2 \sum_{j=1}^N a'^2_j}}$$

Wird in letztere Gleichung die obige Phasenabschätzung für den Umgebungsjet J' im Abstand \bar{d} eingesetzt, kann letzteres Ähnlichkeitsmaß auch geschrieben werden als

$$S_\varphi(J, J', \bar{d}) = \frac{\sum_{j=1}^N a_j a'_j \cos(\varphi_j - (\varphi'_j + \bar{k}_j \cdot \bar{d}))}{\sqrt{\sum_{j=1}^N a_j^2 \sum_{j=1}^N a'^2_j}}$$

bzw. durch Taylorreihenentwicklung noch weiter genähert als

$$S_\varphi(J, J', \bar{d}) = \frac{1}{N} \sum_{j=1}^N a_j a'_j \left[1 - 0.5(\varphi_j - \varphi'_j - \bar{k}_j \cdot \bar{d})^2 \right]$$

Durch Variieren des Abstandsvektors \bar{d} mittels der genannten Suchalgorithmen im Hinblick auf eine Maximierung des Ähnlichkeitsmaßes S_φ kann somit der Versatz \bar{d} des geschätzten Ortes $Im_{est,i}$ vom richtigen Ort $Im_{obj,i}$ des Landmarks Im_i und damit der Landmark $Im_{obj,i}$ im Objektbild P_{obj} lokalisiert werden.

Das Ergebnis des Blocks 30 ist somit eine Liste mit allen im Objektbild P_{obj} lokalisierten Landmarks $Im_{obj,i}$, beispielsweise in der Form

$Im_{obj,i}$	x-Position	y-Position
linksaugenmitte	41.047727	49.115525
rechtsaugenmitte	90.913636	49.852737
nasenbrücke	67.162529	38.257454
usw.		

Auf Grundlage der im Schritt a) lokalisierten Landmarks $Im_{obj,i}$ kann nun im Schritt b) der Ähnlichkeitsvergleich des Objektbildes P_{obj} mit einem bestimmten Referenzbild P_{ref} („Verifizierung eines Objektes“) oder mit mehreren Referenzbildern $P_{ref,s}$ ($s = 1..S$) („Identifizierung eines Objektes unter mehreren“) erfolgen.

Es ist zu beachten, daß die in Schritt b) verwendeten Referenzbilder $P_{ref,s}$ dieselben sein können wie die Referenzbilder $P_{bib,r}$ von Schritt a) oder auch nicht. So können beide Sätze von Referenzbildern ident sein, oder aber der erste Satz von Referenzbildern $P_{bib,r}$ ist z.B. eine umfangreiche Bibliothek repräsentativer Gesichtsformen, -farben, Brillen- und Bartaccessoires usw., während der zweite Satz von Referenzbildern $P_{ref,s}$ nur das oder die zu verifizierende(n)

oder identifizierende(n) Objekt(e) umfaßt.

Im Trainingsteil T werden im Block 40 für den Schritt b) aus den Referenzbildern $P_{\text{ref},s}$ an manuell selektierten Landmarks $Im_{\text{man},i,s}$ Jets $J_{\text{ref},i,s}$ extrahiert, u.zw. auf Grundlage einer Wavelet-Basis B_B , welche einen Satz von N' Gabor-Wavelets $\psi_{B,j}$ umfaßt, von denen die Mehrzahl gekrümmt ist.

Unter gekrümmten Gabor-Wavelets („Banana-Wavelets“) werden Wavelets der allgemeinen Gleichung

$$\psi(x, y, C, \Theta, \lambda, \varphi, \sigma, \gamma) = e^{-\frac{\dot{x}^2 + \gamma^2 \dot{y}^2}{2\sigma}} \cos\left(2\pi \frac{\dot{x}}{\lambda} + \varphi\right)$$

mit

$$\dot{x} = x \cos(\Theta) + y \sin(\Theta) + y^2 C$$

$$\dot{y} = -x \sin(\Theta) + y \cos(\Theta)$$

verstanden,

wobei

- C .. Krümmung des Wavelets
- Θ .. Orientierung des Wavelets
- λ .. Wellenlänge der Sinusoidfunktion
- φ .. Phasenverschiebung
- σ .. Radius der einhüllenden Gaußkurve
- γ .. Seiten/Höhen-Verhältnis des Wavelets

ist. Einige beispielhafte Wavelets $\psi_{B,j}$ sind in Fig. 3 gezeigt.

Bevorzugt wird zur Erzeugung der Wavelet-Basis B_B der folgende Parametersatz gewählt:

$$C = \{0, \pm 0.04, \pm 0.07, \pm 0.1\},$$

$$\Theta = \left\{0, \frac{\pi}{8}, \frac{2\pi}{8}, \frac{3\pi}{8}, \dots, \frac{7\pi}{8}\right\},$$

$$\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\},$$

$$\varphi = \left\{0, \frac{\pi}{2}\right\},$$

$$\sigma = \lambda \text{ und}$$

$$\gamma = 1.0.$$

Dies führt zu 700 gekrümmten und 100 geradlinigen Gabor-Wavelets, d.h. $N' = 800$, und damit zu einer äußerst detailreichen Bildanalyse.

Die Jets $J_{\text{ref},i,s}$ an den Landmarks $Im_{\text{man},i,s}$ der Referenzbilder $P_{\text{ref},s}$ werden in einem Referenzda-

tensatz G_{ref} für den Ablaufteil A zur Verfügung gestellt. Der Referenzdatensatz G_{ref} umfaßt für jedes Referenzbild $P_{ref,s}$ einen sogenannten „Face-Graphen“ G_s :

$$G_{ref} = \begin{pmatrix} J_{1,1} & J_{1,2} & \dots & J_{1,s} & J_{1,S} \\ J_{2,1} & J_{2,2} & & & \\ \vdots & & & & \\ J_{i,1} & & & J_{i,s} & \\ J_{M,1} & & & & J_{M,S} \end{pmatrix} = (G_1 \ G_2 \ \dots \ G_s \ G_S)$$

Im Ablaufteil A des Verfahrens wird nun im Block 50 für jeden lokalisierten Landmark $lm_{obj,i}$ des Objektbildes P_{obj} ein Jet $J_{obj,i}$ extrahiert, und zwar auf Grundlage derselben Gabor-Wavelet-Basis B_B , welche im Extraktionsblock 40 des Trainingsteiles T verwendet und anhand von Fig. 3 erläutert wurde.

In einem anschließenden Block 60 wird der Ähnlichkeitsvergleich zwischen den Jets $J_{obj,i}$ des Objektbildes P_{obj} und den Jets $J_{ref,i,s}$ des Referenzdatensatzes G_{ref} durchgeführt. Als Ähnlichkeitsmaße für die Bestimmung der Ähnlichkeit zweier Jets $J_{obj,i}$ und $J_{ref,i,s}$ können beliebige Korrelationsmaße herangezogen werden, bevorzugt das obengenannte Phasenmaß

$$S_{\varphi}(J_{obj}, J_{ref}) = \frac{\sum_{j=1}^{N'} a_{obj,j} a_{ref,j} \cos(\varphi_{obj,j} - \varphi_{ref,j})}{\sqrt{\sum_{j=1}^{N'} a_{obj,j}^2 \sum_{j=1}^N a_{ref,j}^2}}$$

Dabei kann ein Gesamtähnlichkeitsmaß $S_{\varphi,ges}$ zwischen allen Landmarks lm_i des Objektbildes P_{obj} und einem bestimmten Referenzbild $P_{ref,s}$ verwendet werden, beispielsweise durch Mittelwertbildung über die Ähnlichkeitsmaße $S_{\varphi,i,s}$ der Landmarks lm_i bezüglich eines bestimmten Referenzbildes $P_{ref,s}$:

$$S_{\varphi,ges,s} = \frac{1}{M} \sum_{i=1}^M S_{\varphi,i,s}$$

Alternativ oder zusätzlich könnten auch die geometrischen Abweichungen zwischen den Landmarks eines bestimmten Referenzbildes $P_{ref,s}$ und den Landmarks des Objektbildes P_{obj} in das Ähnlichkeitsmaß eingehen.

Wird das Verfahren zur Verifikation eines Objektes, z.B. Gesichts, eingesetzt, d.h. zum Vergleich mit einem einzigen Referenzbild P_{ref} oder mit einem Satz von verschiedenen Referenzbildern $P_{ref,s}$ ein und desselben Objekts, wird als Ausgang des Blocks 60 eine Verifikation („match“) angezeigt, wenn das im Block 60 ausgewertete Ähnlichkeitsmaß einen vorgegebenen Schwellwert überschreitet.

Wird hingegen das Verfahren zur Identifikation eines Objektes unter mehreren möglichen Objekten verwendet, z.B. zur Identifikation des Gesichts eines Benutzers von mehreren möglichen Benutzern, von denen jeweils ein oder auch mehrere Referenzbilder $P_{ref,s}$ vorliegen, wird als Ausgang des Blocks 60 dasjenige Objekt bzw. Gesicht identifiziert („match“), dessen Ähnlichkeitsmaß das aller anderen übersteigt.

Die Erfindung ist nicht auf die dargestellten Ausführungsformen beschränkt, sondern umfaßt alle Varianten und Modifikationen, die in den Rahmen der angeschlossenen Ansprüche fallen.

Patentansprüche:

1. Verfahren zur Objekterkennung, insbesondere Gesichtserkennung, in einem Objektbild (P_{obj}) durch Vergleichen von Wavelet-Jets signifikanter Punkte (Im_i) des Objektbildes mit Wavelet-Jets zumindest eines Referenzbildes ($P_{bib,r}, P_{ref,s}$), wobei
 5 in einem ersten Schritt im Objektbild (P_{obj}) die aktuellen Koordinaten ($Im_{obj,i}$) der signifikanten Punkte (Im_i) ermittelt werden, indem jeweils im Umgebungsbereich einer Schätzung ($Im_{est,i}$) der Koordinaten die Koordinaten jenes Punktes mit der maximalen Ähnlichkeit seines Wavelet-Jets (J'_i) zu den Wavelet-Jets ($J_{bib,i,r}$) der signifikanten Punkte ($Im_{man,i,r}$) des Referenzbildes ($P_{bib,r}$) ermittelt werden, und
 10 in einem zweiten Schritt die Erkennung des Objektes durch Ähnlichkeitsvergleich zwischen den Wavelet-Jets ($J_{obj,i}$) der lokalisierten signifikanten Punkte ($Im_{obj,i}$) des Objektbildes (P_{obj}) und den Wavelet-Jets ($J_{ref,i,s}$) der signifikanten Punkte ($Im_{man,i,s}$) des Referenzbildes ($P_{ref,s}$) erfolgt,
 15 *dadurch gekennzeichnet*, daß die Wavelet-Jets ($J_{est,i}, J_{bib,i,r}, J'_i$) im ersten Schritt auf Basis von Haar-Wavelets ($\psi_{H,j}; B_H$) und die Wavelet-Jets ($J_{obj,i}, J_{ref,i,s}$) im zweiten Schritt auf Basis von in der Mehrzahl gekrümmten Gabor-Wavelets ($\psi_{B,j}; B_B$) erzeugt werden.
2. Verfahren nach Anspruch 1, *dadurch gekennzeichnet*, daß für jeden signifikanten Punkt (Im_i) die Koordinatenermittlung im ersten Schritt durch Ähnlichkeitsvergleich mit einem Bündel (b_i) von ersten Wavelet-Jets ($J_{bib,i,r}$) verschiedener Referenzbilder ($P_{bib,r}$) erfolgt.
3. Verfahren nach Anspruch 1 oder 2, *dadurch gekennzeichnet*, daß als Haar-Wavelet-Basis (B_H) eine Lienhart-Maydt-Basis verwendet wird, bevorzugt in mehreren Skalierungsschritten, besonders bevorzugt in 24 Skalierungsschritten.
4. Verfahren nach einem der Ansprüche 1 bis 3, *dadurch gekennzeichnet*, daß als Gabor-Wavelet-Basis (B_B) Wavelets der Funktion

$$\psi(x, y, C, \Theta, \lambda, \varphi, \sigma, \gamma) = e^{-\frac{\dot{x}^2 + \gamma^2 \dot{y}^2}{2\sigma}} \cos\left(2\pi \frac{\dot{x}}{\lambda} + \varphi\right)$$

mit

$$\begin{aligned}\dot{x} &= x \cos(\Theta) + y \sin(\Theta) + y^2 C, \\ \dot{y} &= -x \sin(\Theta) + y \cos(\Theta)\end{aligned}$$

und dem Parametersatz

$$C = \{0, \pm 0.04, \pm 0.07, \pm 0.1\},$$

$$\Theta = \left\{0, \frac{\pi}{8}, \frac{2\pi}{8}, \frac{3\pi}{8}, \dots, \frac{7\pi}{8}\right\},$$

$$\lambda = \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\},$$

$$\varphi = \left\{0, \frac{\pi}{2}\right\},$$

$$\sigma = \lambda,$$

$$\gamma = 1.0,$$

verwendet werden.

Hiezu 2 Blatt Zeichnungen

5

10

15

20

25

30

35

40

45

50

55

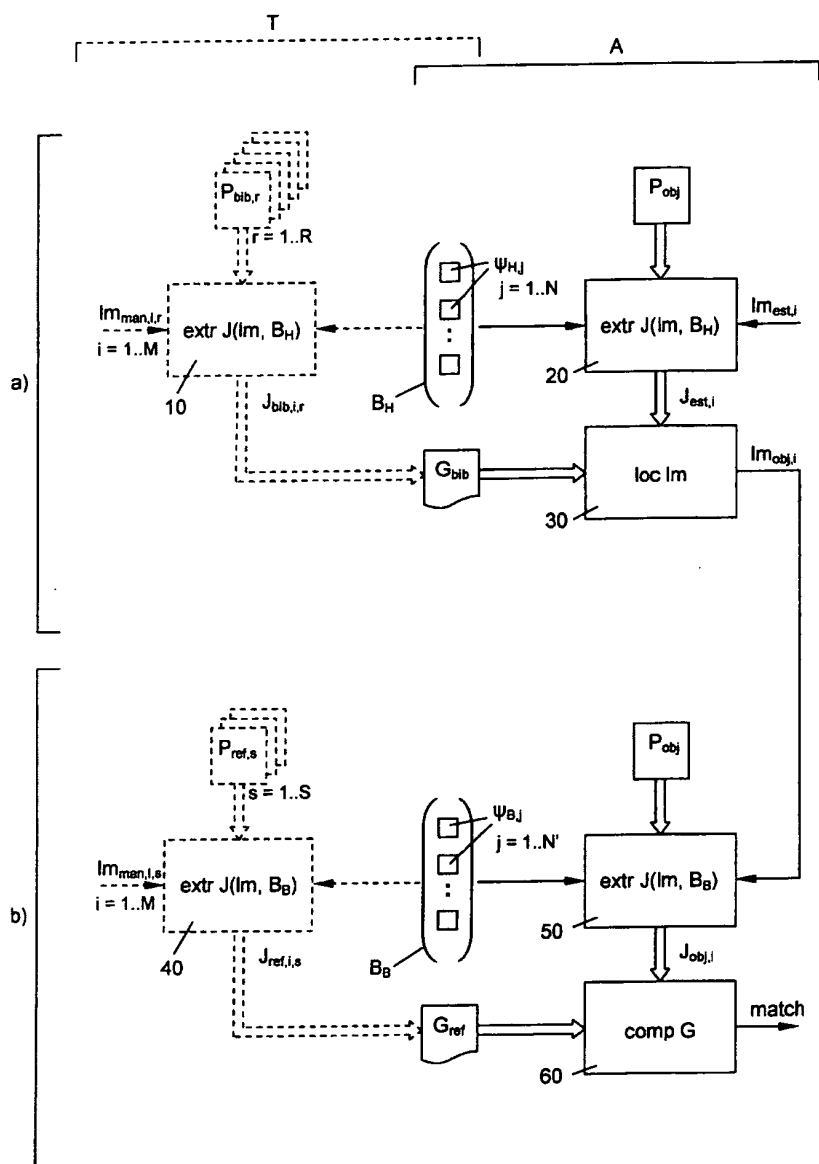


Fig. 1

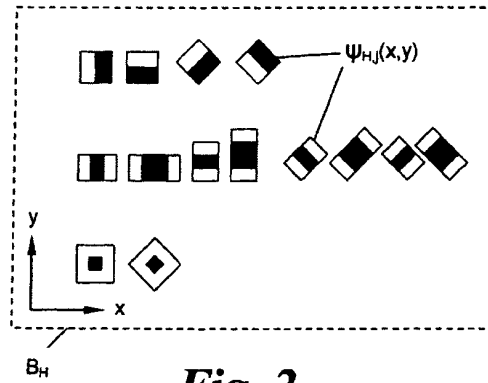


Fig. 2

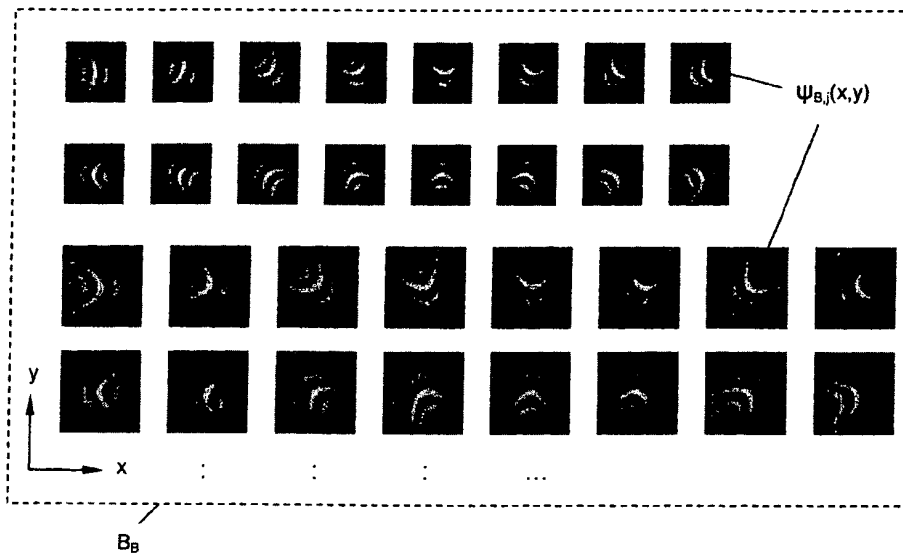


Fig. 3