

# Lokale invariante Objektlokalisierung mittels Gaborfiltern

Dipl. Ing. (FH) **R. Heintz**, Karlsruhe; Prof. Dr.-Ing. **G. Schäfer**, Karlsruhe

## **Kurzfassung:**

Die Erkennung von Objekten gehört zu den Grundaufgaben der digitalen Bildverarbeitung, z.B. in der Robotvision oder der Automatisierungstechnik. Gängige Verfahren sind meist spezialisiert auf bestimmte Umgebungsbedingungen und damit erst nach erheblichen Änderungen auf andere Problemstellungen anpassbar. Eine besondere Rolle bei der Anpassung spielen dabei Drehung, Skalierung und Translation der Objekte. Verfahren, die bereits in ihrer Struktur die Erkennung dieser Transformationen mit einbeziehen, sind im Grundsatz flexibler für Erkennungsprobleme einsetzbar.

Diese Arbeit beschäftigt sich mit der Auswertung eines einfachen, auf Gaborfiltern basierenden Merkmalsraumes. In diesem Merkmalsraum manifestieren sich Rotation und Skalierung als einfache Verschiebungen und sind effizient auswertbar.

## **1. Einführung**

Bei Objekterkennungsaufgaben sind Objekte aus einer vorgegebenen Szene herauszuarbeiten. Liegt ein vorhersagbarer Hintergrund vor, kann eine Segmentierung in Objekt und Hintergrund erfolgen. Das Objekt lässt sich im Anschluss mittels invarianter Merkmale [1] oder Merkmalsräume [2] vergleichen.

Bei komplexen Bildszenen liegen keine oder nur begrenzte Informationen über den Hintergrund vor. Daher können zur Objektlokalisierung nur Objektinformationen verwendet werden. Zumeist kann das Objekt unterschiedliche Repräsentationen annehmen, wodurch die Anzahl der Erscheinungsformen und damit die Anzahl der Suchmuster steigt.

Viele Verfahren basieren auf einem schnellen varianten Modellvergleich. Um Rotation und Skalierung zu erfassen, wird eine Modelldatenbank mit Modellen der möglichen Ansichten des Objektes angelegt. Zur Abdeckung von Rotation und einen gewissen Skalierungsbereich sind zumeist mehrere tausend Modelle nötig, wodurch der Vorteil der schnellen Modellvergleiches zunichte gemacht wird.

Der hier vorgestellte Ansatz bestimmt anhand lokaler, gaborfilterbasierenden Merkmalsräumen, die Position bestimmter Umgebungen und daraus wiederum das Objekt.

## 2. Motivation

Bereits 1962 entdeckten Hubel und Wiesel [3], dass die Filterantworten von Gaborfilter, den Antworten von einfachen Zellen im primären visuellen Kortex des Menschen entsprechen. Zudem ist die zweidimensionale Gaborfunktion die einzige Funktion, welche maximales kombiniertes Auflösungsvermögen im Orts- und Ortsfrequenzbereich besitzt. Der vorgestellte Ansatz führt eine lokale Frequenzanalyse durch, daher ist eine gute Auflösung in Orts- und Ortsfrequenzbereich wichtig [9]. Diese Eigenschaften und der erfolgreiche Einsatz von Gaborfiltern in der Texturanalyse [4] sowie der Gesichtserkennung [5], führte zur Entscheidung Gaborfilter als Basis des Merkmalsraumes zu verwenden.

Zur gleichzeitigen Verarbeitung der Ergebnisse mehrerer Gaborfilterungen führten Kyrki, Kamarainen, Kälviäinen [6] eine Merkmalsmatrize ein. Diese durch Variation der Orientierung und der Frequenz des Gaborfilters entstehende Merkmalsmatrize, im Weiteren als Jetmatrize bezeichnet, diente als Grundlage der Untersuchungen. Da die Parameter für die Orientierung und der Frequenz des Gaborfilters kontinuierlich sind, kann man diese Jetmatrize als Abtastung eines Merkmalsraumes (Jetraum) ansehen.

## 3. Der Merkmalsraum

Zur Erkennung lokaler Umgebungen trotz Rotation und Skalierung, wurde der Ansatz von [6] genauer untersucht. Für die Umgebung jedes Bildpunktes ist ein zweidimensionaler Merkmalsraum zu erzeugen. Skalierung und Rotation spannen den Merkmalsraum auf, dabei soll die Translation keine Auswirkungen auf den Merkmalsraum haben. Die Achsen sind derart zu normieren, dass Skalierung bzw. Rotation zu einer Verschiebung des Merkmalsraumes entlang der zugeordneten Dimension führt.

Ausgehend von einem Bild  $E_1$ , lässt sich durch Translation um  $\bar{x}_B$ , Rotation mit dem Winkel  $\alpha$  um den Drehpunkt  $\bar{x}_0$ , und einer Skalierung mit dem Faktor  $a$  das Bild  $E_x$  erzeugen.

$$(3.1) \quad E_x(\bar{x}) = E_1\left(a \cdot \left(B \cdot ((\bar{x} - \bar{x}_B) - \bar{x}_0) + \bar{x}_0\right)\right) \quad \text{mit} \quad B = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

Eine zweidimensionale Faltung wird mittels des Zusammenhangs

$$(3.2) \quad A_x(\bar{x}, P_1, P_2) = E_x(\bar{x}) ** h(\bar{x}, P_1, P_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_x(\bar{x}) \cdot h(\bar{x} - \bar{a}, P_1, P_2) \cdot d\bar{a}$$

Unter Annahme, dass der Merkmalsraum durch Faltungsfilter erzeugt werden soll, ist folgende allgemeine Forderung an die Faltungsantwort stellen.

$$(3.3) \quad A_x(\bar{x}, P_1, P_2) = A_1\left(a \cdot \left(B \cdot ((\bar{x} - \bar{x}_B) - \bar{x}_0) + \bar{x}_0\right), P_1 - f(a), P_2 - g(\alpha)\right)$$

Das Filterergebnis  $A_1$  muss der gleichen Transformation wie  $E_1$  unterworfen sein, da die im Bild  $E_1$  enthaltenen Umgebungen, ebenfalls dieser Transformation unterworfen sind. Durch Verschiebung der Filterparameter in Abhängigkeit von Rotation und Skalierung ist die Filterantwort  $A_1$  an  $A_x$  anpassbar. Mittels den Funktionen  $f(a)$  und  $g(\alpha)$  lassen sich Rotation und Skalierung so anpassen, dass sie im Merkmalsraum in Verschiebungen übergehen.

Zur weiteren Erläuterung wird der Gaborfilter mathematisch beschrieben:

$$(3.4) \quad g_k(x, y, \sigma_x, \sigma_y, \omega_k, \theta_n) = \frac{1}{\sigma_x \sigma_y} \cdot e^{-\left(\frac{1}{2} \left( \frac{\bar{x}^2}{\sigma_x^2} + \frac{\bar{y}^2}{\sigma_y^2} \right)\right)} \cdot e^{(j\omega_k \cdot (\bar{x} + \bar{y}))} \quad \text{mit} \quad \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} \cos(\theta_n) & \sin(\theta_n) \\ -\sin(\theta_n) & \cos(\theta_n) \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

$$(3.5) \quad g_w(x, y, \omega_k, \theta_n) = \frac{\omega_k^2}{\pi^2} \cdot e^{-\left(\frac{(x^2 + y^2) \cdot \omega_k^2}{2 \cdot \pi^2}\right)} \cdot e^{(j\omega_k \cdot (\bar{x} + \bar{y}))} \quad \text{mit} \quad \sigma_x = \sigma_y = \sigma = \frac{\pi}{\omega_k}$$

Mittels der Abhängigkeit von  $\sigma$  und  $\omega_k$  wird aus dem allgemeinen Gaborfilter (3.4) ein skalierbarer Gaborfilter (3.5), bei dem die Anzahl der Wellen unter der Gaußglocke für jede Frequenz konstant bleibt. Über den Winkel  $\theta_n$  lässt sich die Orientierung und über die Frequenz  $\omega_k$  kann der Mittelfrequenz des Filterkernes festlegen.

In [8] wurde nachgewiesen, dass die Forderung (3.3), von Gaborfiltern erfüllt wird. Für Gaborfilter ergibt sich folgender Zusammenhang.

$$(3.6) \quad A_x(\bar{x}, \omega_k, \theta_n) = A_1\left(a \cdot \left(B \cdot \left(\left(\bar{x} - \bar{x}_B\right) - \bar{x}_0\right) + \bar{x}_0\right), \omega_k - r^a, \theta_n - \alpha\right) \quad \text{mit} \quad r \in \mathbf{R}_+$$

Um einen Wert im Merkmalsraum zu ermitteln, ist eine Filterung nötig. Für eine kontinuierliche Darstellung des Merkmalsraumes wären daher unendlich viele Filterungen nötig. Da dies technisch nicht möglich ist, muss der Merkmalsraum durch Abtastwerte beschrieben werden. Zur Reduzierung des Rechenaufwandes wird eine möglichst kleine Anzahl von Abtastwerten angestrebt. Für die Verteilung der Abtastwerte wurde auf die in [6] vorgeschlagenen Werte zurückgegriffen.

$$(3.7) \quad \omega_k = \omega_0 \cdot 2^{-\frac{k}{2}} \quad \text{mit} \quad k = \{0 \dots K-1\} \quad \text{und} \quad \theta_n = h \cdot \frac{\pi}{N} \quad \text{mit} \quad h = \{0 \dots N-1\}$$

Zur Erreichung einer optimalen Frequenzabdeckung, wird bei Gaborfiltern meist ein exponentielle Zusammenhang zwischen  $\omega_k$  und  $k$  verwendet, da dadurch eine optimale Frequenzabdeckung erreicht wird. Zur Erzeugung des Merkmalsraumes ist dieser exponentielle Zusammenhang zwingend notwendig, da ansonsten die Skalierung nicht in eine Verschiebung übergeht (3.6).

Führt man Faltungen mit obigen Parametern durch, erhält man einen diskreten Merkmalsraum, welche in Anlehnung an [WFKM97] im Weiteren als Jetmatrize  $j(x,y)$  bezeichnet wird.

$$(3.8) \quad j(x,y) = \begin{bmatrix} g_w(x,y,\omega_0,\theta_0) & g_w(x,y,\omega_0,\theta_1) & \cdots & g_w(x,y,\omega_0,\theta_{N-1}) \\ g_w(x,y,\omega_1,\theta_0) & g_w(x,y,\omega_1,\theta_1) & \cdots & g_w(x,y,\omega_1,\theta_{N-1}) \\ \vdots & \vdots & \ddots & \vdots \\ g_w(x,y,\omega_{K-1},\theta_0) & g_w(x,y,\omega_{K-1},\theta_1) & \cdots & g_w(x,y,\omega_{K-1},\theta_{N-1}) \end{bmatrix}$$

Bild1 zeigt einen Überblick über den Ablauf der Merkmalsextraktion. Durch Filterung mit den Gaborfiltern entstehen Merkmalsbilder. Diese Merkmalsbilder besitzen die gleiche Größe wie das Eingangsbild. Somit erhält man für jeden Bildpunkt des Eingangsbildes eine Jetmatrize. Die Jetmatrize besteht aus komplexen Zahlen, zur Vereinfachung wurden in Bild1 die Beträge der Jetwerte dargestellt.

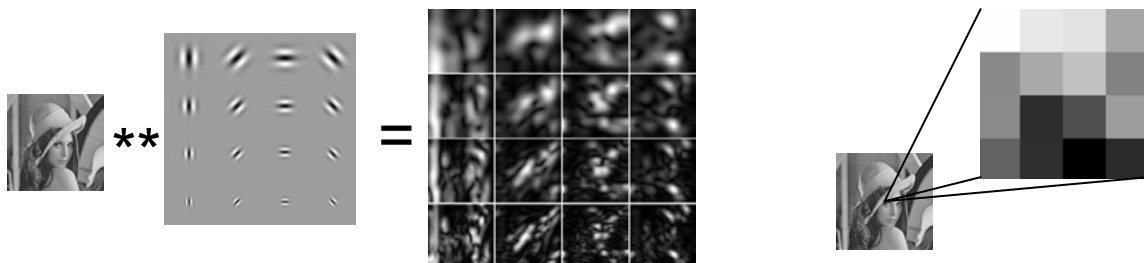


Bild1: Ablauf der Merkmalsextraktion

Gaborfilter sind sensitiv für konstante Eingangssignale, dies führt nach [7] dazu, dass schwache Muster in hellen Umgebungen schlechter als in dunklen Umgebungen identifizierbar sind. Durch Kompensation des Mittelwertes, ergibt sich ein mittelwertfreier Gaborfilter, welcher alle Eigenschaften eines Wavelet erfüllt und somit als Gaborwavelet bezeichnet wird. In [6] wurde bewiesen, dass beim Gaborfilter Bildrotation und Bildskalierung durch Parameteranpassung kompensieren lassen. In [8] wurde gezeigt, dass für Gaborfilter, aber nicht für Gaborwavelets gilt, daher sind Gaborwavelets für die Erzeugung des Merkmalsraumes nicht geeignet.

#### 4. Untersuchung des Merkmalsraums

Zunächst wurde der zulässige Frequenzbereich festgelegt. Unter idealen Bedingungen ist es offensichtlich, dass bei Bilddrehung und gleichzeitiger Drehung des Filterkernes die Filterantwort konstant bleibt. Verwendet man mehrere zueinander rotierte Filterkerne, sind deren Filterantworten bei Drehung des Bildes, um den Winkel der Rotation der Filterkerne, zueinander verschoben. Kompensiert man diese Drehung kann man die Filterantworten vergleichen. Durch die Diskretisierung der Jetposition entsteht ein Fehler, welcher unter realen Bedingungen ebenfalls auftritt. Der Fehler wirkt sich auf die Filterantworten aus, daher

kann durch Betrachtung der Auswirkungen dieses Fehlers, der Parameterbereich von  $\omega_k$  eingeschränkt werden.

Um den Fehler bei verschiedenen Frequenzen zu ermitteln und somit den möglichen Wertebereich des von  $\omega_k$  abzuschätzen zu können, wurde eine Fehlerbetrachtung durchgeführt. Zuerst wurde für jede Bildrotation die Standardabweichung und daraus der prozentuale Fehler der überlagerten Filterergebnisse ermittelt. Dies liefert jedoch nur eine Aussage für eine Umgebung, daher wurde der Mittelwert der prozentualen Fehler mehrerer Punkte bestimmt. Es zeigte sich, dass der Fehler für  $\omega_k < 1$  bei ca. 5% liegt. Ab einer Frequenz  $\omega_k = 1$  steigt der Fehler stark an, wodurch Filter mit größerem  $\omega_k$  für die Erkennung nicht geeignet sind.

Da sich mit sinkendem  $\omega_k$  der Fehler eher reduziert, lässt sich mittels dieses Fehlermaßes keine untere Schwelle von  $\omega_k$  festlegen. Der Filterkern besitzt eine unendliche Ausdehnung, jedoch fällt es exponentiell ab. Durch den schnellen Abfall kann eine endliche Filtergrenze definiert werden. Diese Filtergrenze steigt mit sinkendem  $\omega_k$  an. Legt man die Filtergrenze bei einem prozentualen Abfall auf 10% gezogen auf das Maximum, ergibt sich für  $\omega_k = 0.125$  ein Radius von 54 Pixel. Es wird eine vom Hintergrund nahezu unabhängige Objekterkennung angestrebt, daher sollte sich der Filterkern zum größten Teil innerhalb des Objektes befinden, da ansonsten zuviel hintergrundabhängige Information aufgenommen wird. Die maximale Filterausdehnung ist somit von den Randbedingungen der Objektsuche abhängig, bewegt sich aber erfahrungsgemäß um den Bereich  $\omega_k = 0.1$ .

Wie bereits erwähnt, führen die Bildrotation und -skalierung zu Verschiebungen im Jetraum. Über den kontinuierlichen Jetraum liegen jedoch nur ortsdiskrete Informationen in Form der Jetmatrize vor. Um beliebige Verschiebungen im Jetraum anhand der Jetmatrize erkennen zu können, muss der Jetraum anhand der Jetmatrize interpolierbar sein. Somit lässt sich über die Interpolierbarkeit die Anzahl der benötigten Abtastwerte im Jetraum und somit die Schrittweite der Jetmatrize festlegen.

Zur Bestimmung der Interpolierbarkeit wurde eine Testreihe durchgeführt. Dazu wurde der Jetraum jeweils fein abgetastet und danach mit einer interpolierten Grobabtastung verglichen. Der betrachtete Jetraum wurde anhand der zuvor gewonnenen Kenntnisse auf den Bereich von  $\omega_k = 0.125 \dots 1$  festgelegt.

Die Abtastung des Jetraums wurde aufgrund folgender Formel durchgeführt.

$$(5.5) \quad \omega_k = 2^{-\frac{4 \cdot k}{M}} \quad \text{mit } k = \{0, 1, \dots, M-1\} \quad \text{und} \quad \theta_n = n \cdot \frac{\pi}{M} \quad \text{mit } h = \{0 \dots M-1\} \quad M > 1$$

Der Parameters  $M$  gibt die Anzahl der Abtastwerte in Richtung von  $\omega_k$  und  $\theta_h$  an. Mittels der Formel wird erreicht, dass  $\omega_k$  einen  $M$  unabhängigen Wertebereich von  $0.125 \dots 1$  besitzt.

Die Feinabtastung erfolgte mittels 1000 Abtastwerten für die Grobabtastung wurden  $M$  von 2 bis 16 variiert und mittels komplexer bikubischer Interpolation auf 1000 Abtastwerte interpoliert. Es zeigte sich, dass schon bei kleinen Jetmatrizen sehr gute Interpolationsergebnisse erzielbar waren.

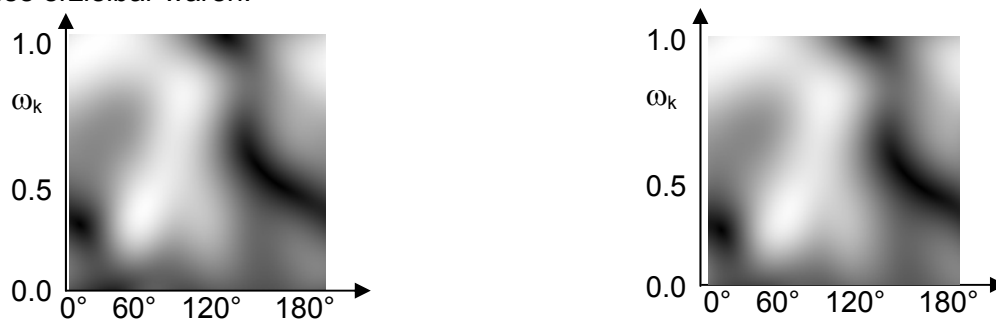


Bild 2: Berechnete Jetmatrize im Vergleich mit interpolierter Jetmatrize des Bildes Lena 256x256 bei dem zufällig gewählten Pixel  $x:152$   $y:174$  (nur Beträge dargestellt)

In Bild 2 wurde eine  $7 \times 7$  Jetmatrize auf die Größe der Referenzmatrize interpoliert und der Referenzmatrize gegenübergestellt. Wie an der Ähnlichkeit der Ergebnisse erkennbar, ist der Jetraum mit nur wenigen Stützpunkten interpolierbar.

Zur Bestimmung der nötigen Stützpunkte, wurden aus 20 Bildern 1000 Stichproben von Jetmatrizen entnommen und mittels dem oben beschriebenen Verfahren untersucht. Der Vergleich zwischen Feinabtastung und interpolierter Grobabtastung geschah mittels kleinster Fehlerquadrate. Bei einer Jetdimension unter 6 Stützpunkten, stieg der Fehler stark an. Daher wurde diese Anzahl an Stützpunkten als Untergrenze angesehen.

## 6. Auswertung der Jetmatrize

Auf Grundlage der bestimmten Parameterbereiche, wurde ein Objekterkennungssystem basierend auf einem Jetvergleich entwickelt.

Um Umgebungen vergleichen zu können, muss ein Ähnlichkeitsmaß bestimmbar sein. Zum Vergleich der Jetmatrizen ist eine Aufhebung der Verschiebung der Jetmatrizen nötig. Dazu wurde das Leistungsdichtespektrum der Jetmatrizen verwendet, da dieses verschiebungsinvariant ist. Die Ähnlichkeit zweier Jetmatrizen erfolgt durch Bestimmung des mittleren quadratischen Fehlers der Leistungsspektren.

Die durch die Bildskalierung entstehende Verschiebung des Jetraums bewirkt, dass unbekannte Information in die Jetmatrize geschoben wird, wodurch die Erkennungsfähigkeit des Verfahrens sinkt. Durch Vergrößerung der Jetmatrize des Referenzbildes kann, wie in Bild 3 gezeigt, die unbekannte Information reduziert werden.

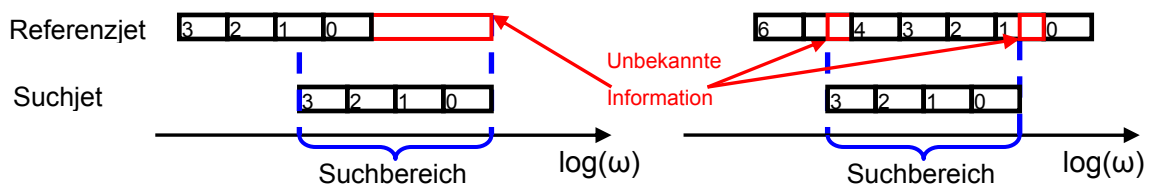


Bild 3: Fehler bei Skalierung des Suchbildes und deren Beseitigung

Bild 4 gibt einen Überblick über die Bestimmung des Fehlermaßes. Der Referenzjet wird in Submatrizen unterteilt und davon die Leistungsdichtespektren bestimmt. Ebenso wird vom Suchjet das Leistungsdichtespektrum ermittelt. Danach werden die kleinsten Fehlerquadrate zwischen den Leistungsdichtespektren des Referenzjets und der Suchmatrize bestimmt. Der kleinste Wert der Fehlerquadrate entspricht dem Fehlermaß einer Bildumgebung.

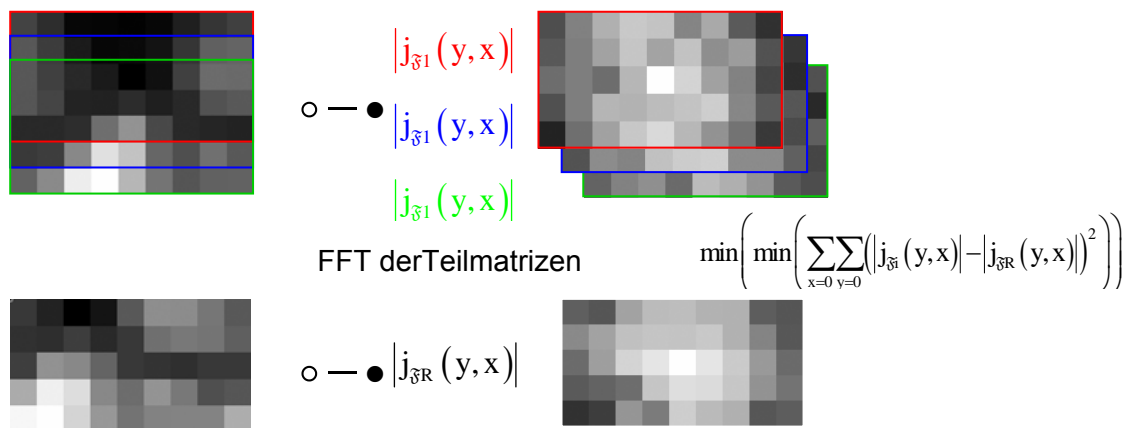


Bild 4: Bestimmung des Fehlermaßes (nur Beträge der Jetmatrizen dargestellt)

Bestimmt man für alle Bildpunkte des Suchbildes dieses Fehlermaß, erhält man eine Aufmerksamkeitskarte in Größe des Suchbildes. Durch Bestimmung des Minimums lässt sich die gesuchte Umgebung lokalisieren. Die Auswertung wurde sowohl mittels komplexer Jetmatrizen als auch Betrags - Jetmatrizen durchgeführt. Es zeigte sich, dass durch die Nutzung der Phaseninformation die Auswertung bessere Ergebnisse lieferte.



Bild 5: Referenzbild, Suchbild, Aufmerksamkeitskarte für Betrags- und komplexe Jetmatrize

In Bild 5 ist ein Beispiel der Punktsuche dargestellt. Eine 9x6 Jetmatrize wurde als Referenzmatrize verwendet und 6x6 Jetmatrizen zur Punktsuche. Nur bei Auswertung der komplexen Jetmatrize wurde die korrekte Position erkannt.

## 7. Zusammenfassung

Auf Basis der vorgestellten Punktsuche ist eine Objektsuche möglich. Mittels der Verschiebungsinformation ist zusätzlich eine Bestimmung von Rotation und Skalierung möglich. Anhand dieser Information kann ein sequentieller Suchalgorithmus aufgebaut werden, da lokale Betrachtungen stattfinden, können durch Anpassung des Suchalgorithmus auch teilverdeckte Objekte erkannt werden. Gegenüber Verfahren die für die unterschiedlichen Ansichten unterschiedliche Modelle erzeugen, lässt sich die Rotation und Skalierung in einem Modell erfassen, wodurch sich die Datenmenge reduziert und eine schnelle Suche erlaubt. Die Auswertung mittels Leistungsspektren besitzt zwei Nachteile. Durch die Betragsbildung bei der Leistungsspektrumsbestimmung geht die Phaseninformation verloren und bei der Bestimmung des Spektrums wird zirkuläre Verschiebung vorausgesetzt, welche jedoch nur in Rotationsrichtung besteht. Daher sind weitere Verfahren zur Bestimmung eines geeigneteren Ähnlichkeitsmaßes zu untersuchen.

## 8. Literaturverzeichnis

- [1] Suk T.; Flusser J.: Combined blur and affine moment invariants and their use in pattern recognition. *Pattern Recognition*, 36 (2003), 12, 2895-2907
- [22] Derrode S.; Ghorbel F.S.: Shape analysis and symmetry detection in gray-level objects using the analytical Fourier-Mellin Representation. *Signal Processing*, 84/1(2004), 25-39
- [3] Hubel, D. H.; Wiesel, T. N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, in *Journal of Physiology* (1962), 160:106 – 154
- [4] Grigorescu, S.E.; Petkov, N.; Kruizinga, P.: Comparison of texture features based on Gabor filters, in: *IEEE Transactions on Image Processing* (2002), 11/10, 142-147
- [5] Messer, K.; Kittler, J.; Sadeghi, M. und andere: Face Authentication Test on the BANCA Database. *17th International Conference on Pattern Recognition* (2004), 523-532
- [6] Kyrki, V.; Kamarainen, J-K.; Kälviäinen, H.: Simple Gabor feature space for invariant object recognition. *Pattern Recognition Letters*, 25/ 3 (2004), 311 - 318
- [7] Fröhlinghaus, T.: Multiscale Gabor Wavelet-based Stereo Disparity Estimation. Dissertation, Friedrich-Wilhelms-Universität Bonn, 2000
- [8] Heintz, R.: Invariantseigenschaften von Gaborfiltern. Interner Bericht 01/05, FH Karlsruhe, 2005
- [9] Burke Hubbard, B.: *Wavelets: Die Mathematik der kleinen Wellen*. Basel: Birkhäuser Verlag, 1997