

Merkmalshistogramme für die Suche in Bilddatenbanken

Sven Siggelkow

sven.siggelkow@gmx.de

Abstract: Invariante Merkmale bleiben unverändert, wenn die Daten mit einer vorgegebenen Transformationklasse verändern. Diese Eigenschaft kann in Anwendungen von Vorteil sein, bei denen die Orientierung oder absolute Position von Objekten irrelevant sind, z.B. bei der Suche nach ähnlichen Bildern. Wir verwenden hier Merkmale, welche sich durch Integration über die Transformationsgruppe der Euklidischen Bewegungen bilden lassen. Deren Berechnungskomplexität ist aber zu hoch, zumal es sich um der angestrebten Applikation um Bilddaten oder gar Volumendaten handelt. Anstatt die Merkmale vollständig zu berechnen, wird daher eine Monte-Carlo Schätzung des Integrals durchgeführt, so daß eine konstante Komplexität erreicht werden kann, ohne daß das Merkmal zu stark von der ursprünglichen Berechnung abweichen würde. Um die Aussagekraft der Merkmale zu erhöhen werden mehrdimensionale Merkmals-histogramme statt der ursprünglichen Merkmale konstruiert. Durch eine gewichtete Zuweisung der Werte zu den Histogrammcontainern wird zudem die Unstetigkeit der Histogrammfunktion beseitigt. Schließlich werden die Methoden anhand von zwei Applikationen demonstriert: einem System zur allgemeinen Bildsuche, sowie einem Programm zur automatischen Suche von Briefmarken anhand von eingescannten Vorlagen.

1 Einleitung

Für viele Anwendungen ist die Objektlage und -orientierung uninteressant. Z.B. ist eine typische Fragestellung bei der Bildsuche, diejenigen Bilder zu finden, die ein bestimmtes Objekt enthalten, egal in welcher Position. Ein menschlicher Betrachter kann diese Aufgabenstellung auf einen Blick lösen, solange die Anzahl der Bilder im Rahmen bleibt. Bei größeren Datenmengen ist es jedoch hilfreich, ihm eine bereits vorselektierte Auswahl zu präsentieren. Die Vorauswahl geschieht dabei automatisiert anhand objektiver Bildähnlichkeiten.

In dieser Arbeit wird die Ähnlichkeit dahingehend definiert, daß bestimmte Objekttransformationen zugelassen werden, die keinen Einfluß auf die Bildähnlichkeit haben sollen. In [BS01] ist ein generelles Verfahren zur Bildung von Invarianten beschrieben. Hier beschränken wir uns auf Grauwertbildmerkmale für die Transformationsgruppe der Translationen und Rotationen aus [SM95] und erweitern diese auf Farbbilder. Zusätzlich zur Invarianz gegenüber globalen Translationen und Rotationen haben sich jene Merkmale auch als robust gegenüber unabhängigen Bewegungen unterschiedlicher Objekte, Gelenkbewegungen sowie Objektdeformationen erwiesen. Ein Vorteil dieser Invarianten ist, daß sie

keinerlei Vorverarbeitungsschritte benötigen wie Segmentierung, Kantenextraktion o.ä., sondern direkt auf den originalen Bilddaten berechnet werden können.

In Abschnitt 2 fassen wir kurz die Bildung der oben angesprochenen Invarianten zusammen. Zusätzlich erweitern wir diese zu invarianten Merkmalhistogrammen, welche für die Bildsuche besser geeignet sind. Im darauf folgenden Abschnitt zeigen wir auf, wie man die Berechnungskomplexität durch eine Monte-Carlo Schätzung auf einen konstanten Aufwand reduzieren kann. In Abschnitt 4 demonstrieren wir den Nutzen der eingeführten Merkmalhistogramme für die Anwendung in Bildsuchsystemen. Schließlich fassen wir die Ergebnisse noch einmal im letzten Abschnitt zusammen.

2 Konstruktion invarianter Merkmale

In diesem Abschnitt fassen wir das generelle Vorgehen zur Bildung von Invarianten zusammen und erweitern diese zu mehrdimensionalen, gewichteten Merkmalhistogrammen.

2.1 Globale Invarianten durch Mittelung über die Transformationsgruppe

Ein Typ von Invarianten, der bereits 1897 von Hurwitz erwähnt wurde [Hu97], sind Integralinvarianten:

Es liege ein endliche Gruppe von discreten Substitutionen vor, die sich auf die Variablen $x_1, x_2, \dots, x_n$ beziehen. Dann giebt es ein einfaches Verfahren, um alle Invarianten dieser Gruppe herzustellen. Man wende nämlich auf eine beliebige Function $f(x_1, x_2, \dots, x_n)$ die sämtlichen Substitutionen der Gruppe an und bilde sodann die Summe aus allen so entstehenden Functionen.

Übertragen auf Bilder $\mathbf{X}$ erhält man somit

$$F(\mathbf{X}) = \frac{1}{|G|} \int_G f(g\mathbf{X}) dg, \quad \text{bzw.} \quad F(\mathbf{X}) := \frac{1}{|G|} \sum_G f(g\mathbf{X}). \quad (1)$$

für eine endliche Gruppe G .

Für den Fall einer euklidischen Bewegung (Rotation und Translation)

$$(g\mathbf{X})(\mathbf{n}) = \mathbf{X}(\mathbf{n}'), \quad \text{mit} \quad \mathbf{n}' = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \mathbf{n} + \mathbf{t}. \quad (2)$$

ergibt sich

$$F(\mathbf{X}) = \frac{1}{2\pi N_0 N_1} \int_{t_0=0}^{N_0} \int_{t_1=0}^{N_1} \int_{\varphi=0}^{2\pi} f(g(t_0, t_1, \varphi)\mathbf{X}) d\varphi dt_1 dt_0, \quad (3)$$

bzw. wenn man Rotationen und Translationen nur in einer endlichen Anzahl von Stufen zuläßt

$$F(\mathbf{X}) = \frac{1}{RN_0N_1} \sum_{t_0=0}^{N_0-1} \sum_{t_1=0}^{N_1-1} \sum_{r=0}^{R-1} f\left(g\left(t_0, t_1, \varphi = 2\pi \frac{r}{R}\right) \mathbf{X}\right). \quad (4)$$

Eine genauere Betrachtung dieser Gleichung zeigt, daß sich die Merkmale für Kernfunktionen f mit endlichem Definitionsbereich folgendermaßen berechnen lassen: Zunächst werden für jede Bildposition die Werte von f in sämtlichen Rotationen ermittelt. Alle diese Werte werden anschließend gemittelt. Dies ist in Abbildung 1 dargestellt.

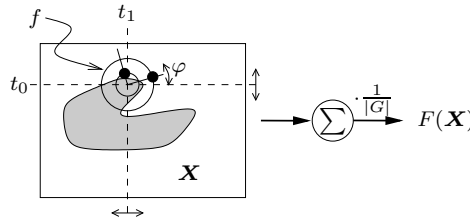


Abbildung 1: Berechnungsschema für die invariante Integration im Fall einer Euklidischen Bewegung.

2.2 Invariante Merkmalhistogramme

Aufgrund der globalen Mittelung der oben dargestellten Merkmale ergeben sich unterschiedliche Probleme bei deren Anwendung zur Bildsuche: Bei größeren Objektverdeckungen oder variierendem Hintergrund ergeben sich starke Abweichungen der Merkmale. Es ist daher sinnvoll, bei Anwendungen, die solche Szenarien einschließen, mehr Informationen über lokale Merkmale zu erhalten als dies bei der Mittelung der Fall ist.

Histogramme sind für diese Aufgabenstellung besser geeignet. Sie haben sich bereits in ihrer einfachsten Form als sehr erfolgreich erwiesen: In [SB91] werden Objekte anhand ihrer Farbhistogramme erkannt. Ein Nachteil dieser Histogramme ist, daß jegliche Informationen über die Nachbarschaft eines Pixels verlorengeht.

Daher verbinden wir die Histogramme mit den oben dargestellten invarianten Merkmalen und konstruieren invariante Merkmalhistogramme. Offensichtlich bleibt die Invarianz erhalten, wenn die Summen in Gleichung (4) durch eine Histogrammbildung ersetzt werden:

$$h \left(\left\{ f \left(g \left(t_0, t_1, \varphi = 2\pi \frac{r}{R} \right) \mathbf{X} \right) \mid \begin{array}{l} t_0 = 0, 1, \dots, N_0 - 1, \\ t_1 = 0, 1, \dots, N_1 - 1, \\ r = 0, 1, \dots, R - 1 \end{array} \right\} \right). \quad (5)$$

Somit ist die Empfindlichkeit der invarianten Merkmale gegenüber Objektverdeckungen

oder Hintergrundänderungen abgemildert, da in solch einem Fall ein Großteil des Histogramms erhalten bleibt, während sich nur der Beitrag des veränderten Bereichs ändert.

Zudem kann man durch Verwendung von unterschiedlichen Kernfunktionen oder von mehreren Farbenen mehrdimensionale Histogramme bilden und so sehr aussagekräftige Merkmale erhalten.

Der Vollständigkeit halber sei erwähnt, daß es bei manchen Applikationen sinnvoll ist, gewichtete Histogramme zu verwenden, d.h. Histogramme, bei denen Einzelwerte nicht einem einzigen Histogrammcontainer zugeordnet werden, sondern allen umliegenden Containern mit einem Gewicht, das sich aus der Lage relativ zu diesen Containern berechnet. Damit läßt sich die Unstetigkeit des Histogrammoperators beseitigen. Diese hat sich z.B. als problematisch erwiesen, wenn Bilder große homogene Flächen aufweisen, die unter anderen Beleuchtungsbedingungen komplett in einen anderen Histogrammcontainer abgebildet werden.

3 Schnelle Merkmalberechnung

Die oben genannten Merkmale sind offensichtlich von linearer Komplexität bezüglich der Bildfläche. Für eine interaktive Bildsuche ist dies noch zu langsam. Daher sollen die Merkmale durch eine Monte-Carlo Integration ermittelt werden, statt sie deterministisch zu berechnen. Zunächst soll dies für die Originalmerkmale (also nicht die Merkmals-histogramme) dargestellt werden: Mit Hilfe der Monte-Carlo Integration können L -dimensionale Integrale über $\mathcal{Z}$ vom folgenden Typ berechnet werden:

$$I = \frac{1}{|\mathcal{Z}|} \int_{\mathcal{Z}} f(z) dz, \quad \text{mit } |\mathcal{Z}| = \int_{\mathcal{Z}} dz \quad \text{und } z = (z_0, z_1, \dots, z_{L-1})^T. \quad (6)$$

Wenn $\mathbf{Z}^{(0)}, \mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(N-1)}$ mit $\mathbf{Z}^{(n)} = (Z_0^{(n)}, Z_1^{(n)}, \dots, Z_{L-1}^{(n)})^T$, $n = 0, 1, \dots, N-1$ unabhängige, im Integrationsbereich $\mathcal{Z}$ gleichverteilte Zufallsvektoren sind, dann sind $Y_n = f(\mathbf{Z}^{(n)})$ unabhängige Zufallsvariablen mit identischer Verteilung und Mittelwert $E(Y_n) = I$, und aufgrund des schwachen Gesetzes großer Zahlen gilt:

$$\lim_{N \rightarrow \infty} P \left(\left| \frac{1}{N} \sum_{n=0}^{N-1} Y_n - I \right| \geq \varepsilon \right) = 0 \quad \forall \varepsilon > 0. \quad (7)$$

Daher kann das Integral I mit folgendem Algorithmus approximiert werden:

- Man generiere ein Set von N zufällig ausgewählten $\mathbf{z}^{(n)}$, $n = 0, 1, \dots, N-1$ die in $\mathcal{Z}$ gleichverteilt sind.
- Für alle $n = 0, 1, \dots, N-1$ bestimme man $y_n = f(\mathbf{z}^{(n)})$.
- Schließlich approximiere man I durch $\bar{y} = \frac{1}{N} \sum_{n=0}^{N-1} y_n$.

Übertragen auf unsere Problemstellung ist $\mathbf{z}$ ein Parametervektor, d.h. die Kernfunktion f wird auf dem Bild $\mathbf{X}$ für eine Menge von zufälligen Parametervektoren $\mathbf{z} = (t_0, t_1, \varphi)^T$ berechnet wie in Abbildung 2 dargestellt. Abschließend werden diese Ergebnisse gemittelt.

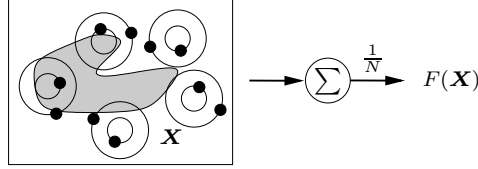


Abbildung 2: Illustration der Merkmalschätzung.

Natürlich erhält man mit diesem Vorgehen eine Abweichung $\bar{y} - I$ vom Originalwert. Diese läßt sich jedoch folgendermaßen abschätzen: Nach dem schwachen Gesetz großer Zahlen und dem zentralen Grenzwertsatz ist $\bar{Y}$ für große N normalverteilt mit Mittelwert bzw. Varianz

$$E(\bar{Y}) = I, \quad V(\bar{Y}) = \frac{V(Y_n)}{N}. \quad (8)$$

Die Anzahl N von Ziehungen für eine angestrebte Genauigkeit ε mit Sicherheit $1 - \delta$ kann daher durch Aufintegration der Normalverteilung von $-\varepsilon$ bis ε bestimmt werden. Damit ergibt sich

$$1 - \Phi\left(\frac{\varepsilon}{\sqrt{V(\bar{Y})}}\right) \leq \frac{\delta}{2} \quad (9)$$

und mit Gleichung (8)

$$N \geq \left(\frac{\Phi^{-1}\left(1 - \frac{\delta}{2}\right)}{\varepsilon}\right)^2 \cdot V(Y_n). \quad (10)$$

Ein konkretes Beispiel: Seien die Werte der Kernfunktion $f(g\mathbf{X})$ im Bereich $[0, f^{\max}]$ und somit $Y_n \in [0, f^{\max}]$. Als gewünschte Toleranz werden folgende Werte vorgegeben: $\varepsilon = 0.01 f^{\max}$ (1% des Maximalwertes) und $\delta = 5\%$. Damit erhält man als Abschätzung für die Varianz $V(Y_n)$:

$$V(Y_n) = \underbrace{E(Y_n^2)}_{\leq (f^{\max})^2} - \underbrace{E^2(Y_n)}_{\geq 0} \leq (f^{\max})^2. \quad (11)$$

Somit erhält man $N \geq (\Phi^{-1}(0.975)/0.01)^2 = 38416$, *unabhängig* von der Bildgröße oder Größe der Kernfunktion. Experimente auf einer Datenbank mit 438 Bildern (QBIC-Daten von IBM [FSN⁺95]) und 100 Schätzungen des Merkmalwertes haben bestätigt, daß die oben festgesetzten Vorgaben eingehalten werden.

Für Merkmalshistogramme gestaltet sich die Herleitung der Fehlergrenze etwas komplizierter. Außerdem ist der Fehler abhängig vom verwendeten Histogramm-Abstandsmaß. In [Si02] ist der zu erwartende Fehler z.B. für die Summe der absoluten Differenzen L_1 hergeleitet und ergibt sich zu

$$N \geq \left(\frac{\left(\Phi^{-1}(1 - \delta) + \sqrt{\frac{2M}{\pi-2}} \right) \cdot \sqrt{(1 - 2/\pi)(1 - 1/M)}}{\varepsilon} \right)^2, \quad (12)$$

worin M die Anzahl der Histogrammcontainer bezeichnet.

Experimente in [Si02] bestätigen, daß der praktische Fehler durch die theoretischen Formeln gut vorhergesagt wird. Für große Histogramme wird der Fehler allerdings stark überschätzt. Das liegt darin begründet, daß große mehrdimensionale Merkmalshistogramme typischerweise spärlich besetzt sind, die Fehlerabschätzung aber den ungünstigsten Fall einer Gleichverteilung berücksichtigt.

4 Bildsuche mit Hilfe invarianter Merkmalshistogramme

Basierend auf den dargestellten Methoden wurde SIMBA (Search Images By Appearance) implementiert [SSB01, SB01], ein Bildsuchsystem, das auf dem MPEG-7 Testdatensatz von fast 2500 Bildern der Größe 376×256 arbeitet.

Der Benutzer kann Merkmalshistogramme kombinieren mit monomialer Kernfunktion sowie relationaler Kernfunktion. Die monomialen Merkmalshistogramme werden gebildet durch Anwendung der Kernfunktion $f(\mathbf{X}) = (\mathbf{X}(4, 0) \cdot \mathbf{X}(0, 8))^{1/2}$ auf jede Farbebene im RGB -Farbraum und durch Berechnung eines multidimensionalen Histogramms dieser lokalen Ergebnisse.

Die relationalen Merkmalshistogramme basieren nur auf den Grauwerten. Als Kernfunktion kommt folgender Ausdruck zur Anwendung [Sc01]:

$$f(\mathbf{X}) = \text{rel}(\mathbf{X}(0, 0) - \mathbf{X}(0, 4)) \quad \text{mit} \quad \text{rel}(x) = \begin{cases} 1 & x < -c \\ \frac{1}{2c}(c - x) & -c \leq x \leq c \\ 0 & c < x \end{cases} \quad (13)$$

und $c = 25$ (bei einem Gesamtwertebereich von $[0, 255]$). Zunächst wird für jeden Bildpunkt in einer runden Nachbarschaft ein gewichtetes Histogramm mit $M = 3$ Klassen berechnet. Bildlich werden in diesem die Bildpunkte sortiert nach „dunkler“, „ungefähr gleichhell“ und „heller“. Anschließend werden aus diesen Histogrammen ein dreidimensionales Gesamthistogramm aller Punkte berechnet. Bildlich gesprochen erfolgt somit eine Klassifikation des Bildes nach stark texturierten und weniger stark texturierten Bildbereichen.

Der Benutzer kann bei der Bildsuche die Ergebnisse der monomialen und der relationalen Merkmale gewichten, so daß die Bedeutung von Farbe oder Textur für die Bildsuche variiert werden kann.

Für den Histogrammvergleich verwenden wir einen Histogrammschnitt $\sum_{m=0}^{M-1} \min(h_m, q_m)$. Dieser ist eng verwandt mit der L_1 -Norm der Histogrammdifferenz, so daß die oben durchgeführte Fehlerabschätzung herangezogen werden kann. Außerdem ermöglicht der Histogrammschnitt eine Suche mittels Bildausschnitten.

Für quantitative Messungen bzgl. der Aussagekraft der Merkmale sei auf [Si02] verwiesen. Wir beschränken uns hier darauf, den Gewinn durch Verwendung von Kombinationen unterschiedlicher Merkmale anhand eines Beispiels zu veranschaulichen. Abbildung 3 zeigt eine Bildsuche mit zwei unterschiedlichen Merkmalssets. Im ersten Teil werden ausschließlich monomiale Merkmale verwendet. Im zweiten Teil werden diese mit relationalen Merkmalen kombiniert, wodurch das Ergebnis stark verbessert werden kann.

In einer zweiten Anwendung wurden die Merkmale für eine Suche in einer Datenbank von ca. 12000 Briefmarkenbildern angepaßt. Da Briefmarken leicht zu segmentieren sind, wurde hier zusätzlich eine Segmentierung vorgenommen, um Fläche und Seitenverhältnisse bestimmen zu können. Zudem können die Merkmalhistogramme dadurch die Lage der Bildpunkte in der Marke mit berücksichtigen.

Mit dem entwickelten Programm MICHELscope lassen sich Marken mit Hilfe eines Scanners identifizieren, ohne lange in Briefmarkenkatalogen suchen zu müssen (Abbildung 4).

Zudem ermöglicht das Programm, Marken mit ähnlichem Motiv aufzufinden. So weisen Briefmarkenserien häufig ähnliche Motivcharakteristika auf. Oft werden auch existierende Motive zu einem späteren Zeitpunkt noch einmal leicht variiert aufgegriffen. Solche Marken können leicht mit Hilfe des Programms aufgefunden werden, während sie im Briefmarkenkatalog an weit auseinander liegenden Stellen zu suchen wären. Abbildung 5 zeigt solch ein Beispiel.

5 Zusammenfassung

In der hier kurz beschriebenen Doktorarbeit wurden histogrammbasierte Techniken zur inhaltsbasierten Bildsuche entwickelt und verbessert. Die Hauptbeiträge bestehen in der Konstruktion robuster, invarianter Merkmalhistogramme. Diese wurden basierend auf invarianten Merkmalen von [SM95] konstruiert, so daß eine vollautomatische Merkmalextraktion ermöglicht wurde. Der zweite Beitrag bestand in der Konstruktion gewichteter Histogramme, um die Unstetigkeit der normalen Histogrammabbildung zu beseitigen. Dieser Punkt wurde hier aus Platzgründen nur sehr kurz dargestellt. Schließlich wurde die Berechnung der Merkmalhistogramme erheblich beschleunigt, indem eine Monte-Carlo Schätzung eingesetzt wurde. Hierfür wurde der zu erwartende Fehler theoretisch abgeschätzt und in praktischen Experimenten bestätigt.

Die Praxistauglichkeit der Verfahren wurde in zwei Anwendungen zur Bildsuche demonstriert: SIMBA, einem Prototypen zur allgemeinen Bildsuche, und MICHELscope, einem Programm zur Suche nach Briefmarken.

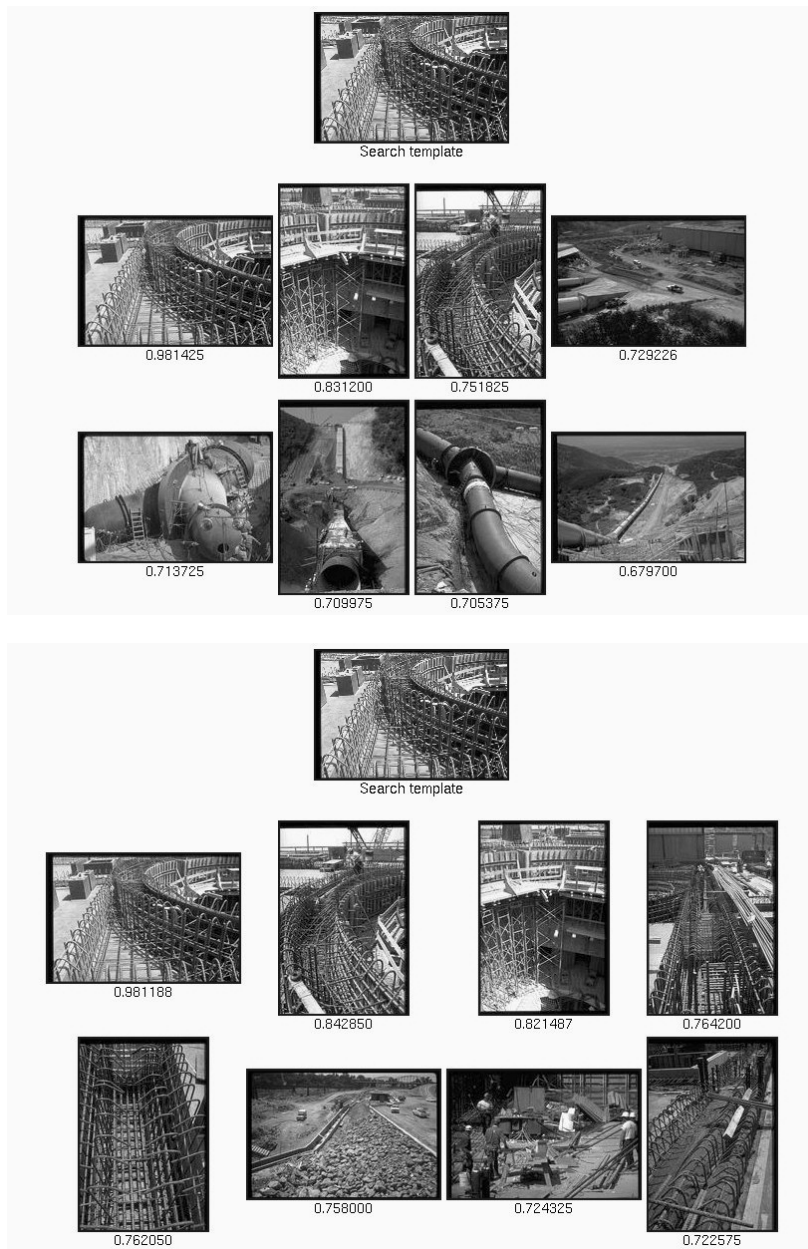


Abbildung 3: Beispielsuche mit monomialen Merkmalen (oben) bzw. einer Kombination aus monomialen und relationalen Merkmalen (unten).



Abbildung 4: Suche mit einer eingescannten Briefmarke.



Abbildung 5: Suche nach Marken mit ähnlichem Motiv.

Literatur

- [BS01] Burkhardt, H. und Siggelkow, S.: Invariant features in pattern recognition – fundamentals and applications. In: Kotropoulos, C. und Pitas, I. (Hrsg.), *Nonlinear Model-Based Image/Video Processing and Analysis*. S. 269–307. John Wiley & Sons. 2001.
- [FSN⁺95] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., und Yanker, P.: Query by image and video content: the QBIC system. *IEEE Computer*. 28(9):23–32. 1995.
- [Hu97] Hurwitz, A.: Über die Erzeugung der Invarianten durch Integration. In: *Nachrichten von der Kgl. Gesellschaft der Wissenschaften zu Göttingen – Mathematisch-Physikalische Klasse*. S. 71–89. 1897.
- [SB91] Swain, M. J. und Ballard, D. H.: Color indexing. *International Journal of Computer Vision*. 7(1):11–32. 1991.
- [SB01] Siggelkow, S. und Burkhardt, H.: Fast invariant feature extraction for image retrieval. In: Veltkamp, R. C., Burkhardt, H., und Kriegel, H.-P. (Hrsg.), *State-of-the-Art in Content-Based Image and Video Retrieval*. S. 43–68. Kluwer. 2001.
- [Sc01] Schael, M.: Texture defect detection using invariant textural features. In: Radig, B. und Florczyk, S. (Hrsg.), *Pattern Recognition, DAGM, LNCS 2191*. S. 17–24. München, Germany. September 2001.
- [Si02] Siggelkow, S.: *Feature Histograms for Content-Based Image Retrieval*. Dissertation, Albert-Ludwigs-Universität Freiburg. 2002.
- [SM95] Schulz-Mirbach, H.: *Anwendung von Invarianzprinzipien zur Merkmalgewinnung in der Mustererkennung*. Fortschritt-Berichte VDI, Reihe 10. VDI-Verlag. 1995.
- [SSB01] Siggelkow, S., Schael, M., und Burkhardt, H.: SIMBA – Search Images By Appearance. In: Radig, B. und Florczyk, S. (Hrsg.), *Pattern Recognition, DAGM, LNCS 2191*. S. 9–16. München, Germany. September 2001.

Sven Siggelkow studierte von 1990 bis 1996 Elektrotechnik mit der Fachrichtung Technische Informatik an der TU Hamburg-Harburg. Seine Diplomarbeit wurde mit dem Preis der Olympus Europa Stiftung „Wissenschaft fürs Leben“ ausgezeichnet. 1995 arbeitete er für einige Monate im Applikationslabor von Philips Semiconductors, wo er zwei Patente verfaßte. Von 1997 bis 2001 war er wissenschaftlicher Angestellter, zunächst an der TU Hamburg-Harburg, später an der Albert-Ludwigs-Universität Freiburg, wo er 2002 zum Dr.-Ing. promovierte. Seine Beiträge zur Tagung der DAGM wurden 1999 und 2001 ausgezeichnet. Außerdem entwickelte er während dieser Zeit nebenberuflich Software für den Schwaneberger-Verlag München. Seit 2001 arbeitet er bei der sci-worx GmbH.