Ludwig Maximilians Universität München
Institut für Informatik
Lehr- und Forschungseinheit für Datenbanksysteme

Skript zur Vorlesung

Datenbanksysteme II

Sommersemester 2006

Kapitel 7: Ähnlichkeitsmodelle für Bilder

Vorlesung: Christian Böhm
Übungen: Elke Achtert, Peter Kunath, Alexey Pryakhin

http://www.dbs.informatik.uni-muenchen.de/Lehre/DBSII

Inhalt

1. Einführung
2. Farben
3. Formen

Inhalt

1. Einführung

2. Farben

3. Formen

Inhaltsbezogene Suche

- **Suche über Standardattribute**
  - Primärschlüssel (z.B. Dateiname): Keine “Suche”, da Identifikator bereits bekannt.
  - Sekundäre Merkmale (Kontextinformationen) wie Entstehungsdatum, Entstehungsort, Rechteinhaber sind nur begrenzt hilfreich.

Inhaltsbezogene Suche

- **Suche über Schlüsselwörter**
  - Manuelle Verschlagwortung bedeutet großen Aufwand.
  - Schlagwörter müssen normiert sein (Abhilfe durch Dictionaries möglich).
  - Schlagwörter decken immer nur bestimmte ausgewählte Aspekte ab (z.B. abgebildete Gegenstände, Indoor/Outdoor/…-Klassifikation)
  - Schlagwortsuche versagt, wenn betrachteter Aspekt nicht als Schlagwort aufgenommen wurde (z.B. "Suche alle Bilder mit hohem Grünanteil am unteren Rand").
- **Suche über den eigentlichen Bildinhalt**
  - Konzept: Inhalt aus der internen Bildrepräsentation (Pixel) ableiten.
  - Aufwand und Probleme der manuellen Verschlagwortung entfallen.
  - Möglichkeiten: Farben, Texturen, Formen

Systeme zur Inhaltsbasierten Suche

- *QBIC*: Query By Image (and Video) Content. IBM Almaden Research Center
- *ImageMiner*. Technologie-Zentrum Informatik, Uni Bremen
- *VisualSeek*. Center for Telecom Research, Columbia Univ., NY
- *MARS: Multimedia Analysis and Retrieval System*. U. Illinois/Urbana-Champaign
- *Surfimage*. INRIA Recquencourt, France
- … und viele mehr!

Merkmale von Bildern

- **Farbe**
  - Farbhistogramme (QBIC) [HSE+ 95]
- **Form** (Konturen)
  - Algebraische Moment-Invarianten [TC 91] [FBF+ 94]
  - Pixelbasierte Ähnlichkeitsmodelle [WJ 96] [AKS 98]
  - Morphologisches Ähnlichkeitsmodell [KSF+ 98]

Inhalt

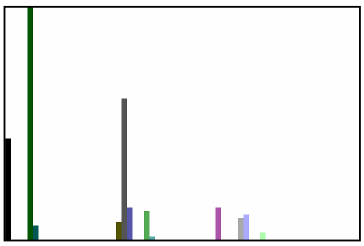
1. Einführung
2. Farbe
3. Form

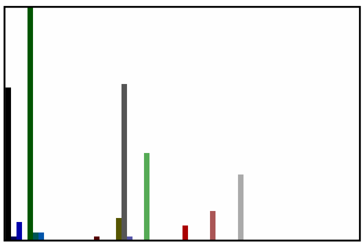
Farbhistogramme (1)

- Repräsentation der Farbverteilung in einem Bild (auf Pixelbasis)
- Definition der Farbhistogramme
  - Farbraum festlegen (z.B. RGB, HSV, HLS, Munsell, …)
  - Menge von Repräsentanten im Farbraum auswählen (*sample points*)
  - z.B. Gitter im Farbraum mit 4x4x4 = 64 Farben oder 8x8x8 = 512 Farben
- Berechnung der Farbhistogramme
  - Für jedes Pixel, erhöhe den Zähler des nächstgelegenen Repräsentanten um eins.
  - Evtl. Normierung, um Histogramm von der Bildgröße unabhängig zu machen.

Farbhistogramme (2)

Beispiel für Farbhistogramme (64 Repräsentanten):

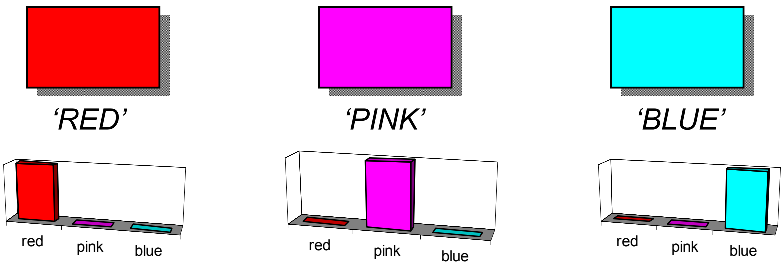



Quadratische Form (1)

Beispiel **Euklidische Distanz**:

- Seien H^P und H^Q die Farbhistogramme der Bilder P und Q.

$$\mathrm{D}(P,Q) = \sqrt{(H^P - H^Q)\cdot(H^P - H^Q)^T}$$



(1, 0, 0, …) (0, 1, 0, …) (0, 0, 1, …)

$$D('RED','PINK') = \sqrt{2}, \quad D('RED','BLUE') = \sqrt{2}, \quad D('PINK','BLUE') = \sqrt{2}$$

- Problem: Sachverhalt, dass rot ähnlicher zu pink als zu blau ist, wird nicht beachtet
- Hintergrund: Querbezüge zwischen den Dimensionen werden grundsätzlich negiert

Datenbanksysteme II
Kapitel 7: Ähnlichkeitsmodelle für Bilder

Quadratische Form (2)

Abhilfe: Quadratische Formen als Distanzfunktionen

- **Definition**: Sei A eine Ähnlichkeitsmatrix, dann gilt:

$$\mathrm{D}_A(P,Q) = \sqrt{(H^P - H^Q)\cdot A \cdot (H^P - H^Q)^T} = \sqrt{\sum_i \sum_j a_{ij}(H_i^P - H_i^Q)(H_j^P - H_j^Q)}$$

- Die Einträge a_{ij} einer **Ähnlichkeitsmatrix** $A = [a_{ij}]$ beschreiben die Ähnlichkeit der Dimensionen i und j in den Vektoren (Bins i und j in den Histogrammen)

$$A = \begin{bmatrix} 1 & & & \\ & \dots & & \\ & a_{ij} & \dots & \\ & & & 1 \end{bmatrix}$$

- Im obigen Beispiel erhalten wir für die Matrix $A' = \begin{bmatrix} 1{,}0 & 0{,}9 & 0{,}0 \\ 0{,}9 & 1{,}0 & 0{,}0 \\ 0{,}0 & 0{,}0 & 1{,}0 \end{bmatrix}$ die Abstandswerte:

$$D('RED','PINK') = \sqrt{0.2}, \quad D('RED','BLUE') = \sqrt{2}, \quad D('PINK','BLUE') = \sqrt{2}$$

Beispiele für Ähnlichkeitsmatrizen

Im folgenden ist d_{ij} der Abstand der Bins i und j (vgl. [HSE+ 95])

- $a_{ij} = (1 - d_{ij} / d_{max})$
- $a_{ij} = \exp(- \sigma (d_{ij} / d_{max})^2)$
  (für $\sigma \to \infty$ erhält man die Einheitsmatrix)
- QBIC verwendet eine aus Ergebnissen der Perzeptionsforschung abgeleitete Matrix

Eigenschaften von Ähnlichkeitsmatrizen (1)

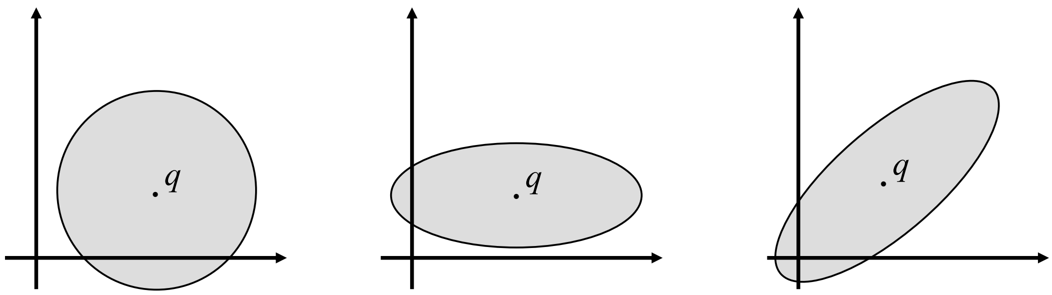
- **Symmetrie**
  Wir dürfen annehmen, dass Ähnlichkeitsmatrizen immer symmetrisch sind, denn:
  - **Lemma**: zu jeder Matrix A' gibt es eine symmetrische Matrix $A = (A' + A'^{T})/2$, so dass gilt: $D_A(P, Q) = D_{A'}(P, Q)$
  - **Beweis**: Sei $\Delta = H^P - H^Q$, dann gilt:

$$D_A(P,Q) = \sqrt{\Delta \cdot A \cdot \Delta^T} = \sqrt{\Delta \cdot \frac{A' + A'^T}{2} \cdot \Delta^T} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{a_{ij} + a_{ji}}{2} \right) \Delta_i \Delta_j}$$

$$= \sqrt{\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \Delta_i \Delta_j + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ji} \Delta_i \Delta_j} = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \Delta_i \Delta_j} = D_{A'}(P,Q)$$

Eigenschaften von Ähnlichkeitsmatrizen (2)

- **Geometrie von ε-Anfragen**



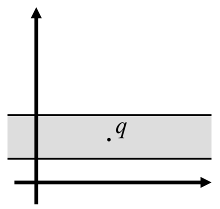
euklidische Distanz | gewichtete euklidische Distanz | positiv-definite quadratische Form

Eigenschaften von Ähnlichkeitsmatrizen (3)

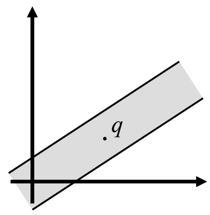
- **Positiv definite Matrizen (PD)**
  - Ähnlichkeitsmatrizen müssen positiv definit sein
  - **Definition** (aus der linearen Algebra):
    A ist positiv definit gdw. $x\,A\,x^T > 0$ für alle $x \neq 0$
  - D.h. Matrix ist PD gdw. Distanzfunktion ist PD (nötig für Metrik!)
  - Test, ob eine Matrix PD ist z.B. durch Berechnung der Cholesky-Zerlegung

Eigenschaften von Ähnlichkeitsmatrizen (4)

- **Positiv semidefinite Matrizen (PSD)**
  - Für bestimmte Anwendungen kann man sich semi-definite Matrizen vorstellen
  - **Definition**: A ist positiv semi-definit gdw. $x\ A\ x^T \geq 0$ für alle $x \neq 0$
  - D.h. auch für $x \neq 0$ (Histogramme: $H^P \neq H^Q$) kann der Distanzwert verschwinden
  - Geometrische Deutung für ε-Anfragen: der Anfragebereich ist unbeschränkt
  - **Lemma**: Falls $x\ A\ x^T = 0$ für ein $x \neq 0$, dann auch $\lambda x\ A\ \lambda x^T = 0$ für alle $\lambda \in IR$



ε-Anfragebereich einer PSD gewichteten euklid. Distanz



ε-Anfragebereich einer PSD quadratischen Form

Anfragebearbeitung für Quadratische Formen (1)

Problem:

- Die Auswertung einer quadratischen Form in d Dimensionen benötigt $O(d^2)$ viele arithmetische Operationen
- Die Laufzeit einer sequentiellen Auswertung für eine Datenbank mit n Objekten ist $O(n \cdot d^2)$.

$\Rightarrow$ Gesucht sind schnellere Verfahren.

Diagonalisierung der Ähnlichkeitsmatrix

- Idee: Quadratische Form in gewichteten euklidischen Abstand überführen.
- Grundlage: Jede PD-Matrix A läßt sich diagonalisieren.
- Das bedeutet: Es gibt eine Diagonalmatrix $W = \text{diag}(w_1, \ldots, w_n)$ sowie eine orthonormale Matrix V, d.h. $V\ V^T = V^T V = Id$, so dass gilt: $A = V\ W\ V^T$

Anfragebearbeitung für Quadratische Formen (2)

- Damit gilt für alle Vektoren (Histogramme) p und q:

$$\mathrm{D}_A(p,q) = \sqrt{(p-q)\cdot VWV^T \cdot (p-q)^T}$$
$$= \sqrt{(pV-qV)\cdot W \cdot (pV-qV)^T} = \mathrm{D}_W(pV,qV)$$

- D_A ist also äquivalent zum gewichteten euklidischen Abstand D_W, nachdem alle beteiligten Vektoren der Basistransformation V unterworfen wurden.
- Die Gewichte $w_1, \ldots, w_n$ sind die *Eigenwerte* der Matrix A, die zugehörigen Spalten in V sind die *Eigenvektoren* von A.

Anfragebearbeitung für Quadratische Formen (3)

- **Dimensionsreduktion**
  - Da A positiv definit ist, sind alle Eigenwerte w_i positiv, d.h. $w_i > 0$, und ein “Abschneiden” der d-dimensionalen Vektoren pV und qV auf $r < d$ Dimensionen liefert eine garantierte untere Schranke für $\mathrm{D}_W = \mathrm{D}_A$:

$$\mathrm{D}_{W,r}(pV,qV) = \sqrt{\sum_{i=1}^{r} w_i (pV_i - qV_i)^2} \le \sqrt{\sum_{i=1}^{d} w_i (pV_i - qV_i)^2} = \mathrm{D}_W(pV,qV)$$

  - Die untere Schranke-Eigenschaft stellt die Vollständigkeit der Anfragebearbeitung sicher.
  - Beobachtung: Die Transformation hängt von der Ähnlichkeitsmatrix A ab!

Anfragebearbeitung für Quadratische Formen (4)

- **Anpassbarkeit durch den Benutzer**
  - Ähnlichkeit hat einen stark subjektiven Charakter.
  - Ein Benutzer ist mit der vorgegebenen Matrix möglicherweise nicht zufrieden.
  - Neuaufbau eines Indexes zur effizienten Datenbanksuche bzgl. D_A ist sehr teuer; erwünscht sind deshalb flexiblere Techniken, die eine Modifikation der Ähnlichkeitsmatrix zur Anfragezeit ermöglichen.

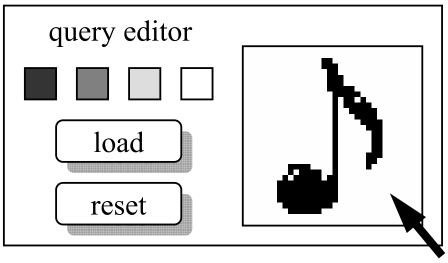
Inhalt

1. Einführung
2. Farbe
3. Form

Pixelbasiertes Modell (1)

([AKS 98] Ankerst M., Kriegel H.-P., Seidl T.: A Multistep *Approach for Shape Similarity Search in Image Databases*. IEEE Transactions on Knowledge and Data Engineering (TKDE) 10(6), 1998, 996-1004.)

- **Anwendungen für formbasierte Ähnlichkeitssuche**
  - Query By Sketch (z.B. mausgesteuerter Editor)

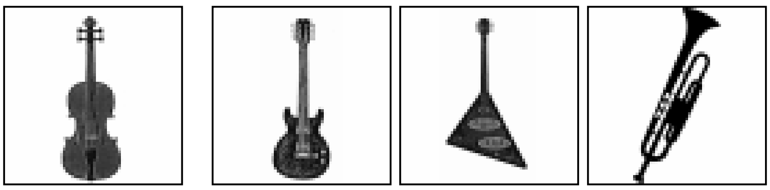


  - Vorgegebene Bilder
    - Grafikarchive
    - Patentrecherche
    - Medizinbilder

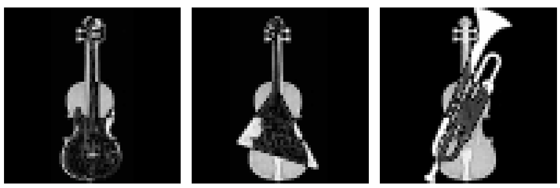
Pixelbasiertes Modell (2)

- **Konzept der Differenzbilder**

Anfrage- und Ergebnisbilder (64×64 Pixel)



Differenzbilder und eukidische Distanz



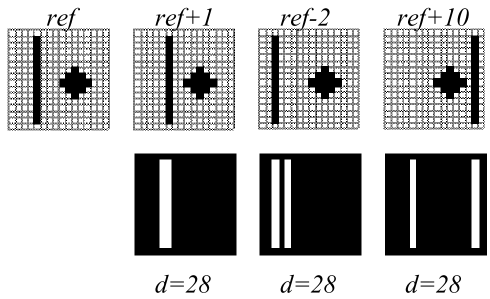
d = 128.2 d = 223.9 d = 424.7

Datenbanksysteme II
Kapitel 7: Ähnlichkeitsmodelle für Bilder

Umgebungsbasierte Distanzfunktion (1)

Probleme der Euklidischen Distanz

- Beispiel Referenzbild: Balken links von einem Punkt.
  - Bei den Vergleichsbildern ist der Balken um +1, –2 bzw. +10 Pixel horizontal verschoben.
  - Euklidischer Abstand zu *ref* bleibt jedoch immer derselbe.



- Problembeschreibung
  - Leichte Verschiebungen sind von starken Veränderungen nicht unterscheidbar.
  - Invarianz gegenüber globalen Translationen stellt keine Lösung dar.
  - Erwünscht ist die Robustheit gegenüber kleinen, lokalen Veränderungen.

Datenbanksysteme II
Kapitel 7: Ähnlichkeitsmodelle für Bilder

Umgebungsbasierte Distanzfunktion (2)

Lösungsidee: Betrachte die Nachbarschaften der Pixel

- Statt nur die direkt übereinander liegenden Pixel im Differenzbild zu betrachten, werden nun auch benachbarte Pixel zur Ähnlichkeitsbewertung herangezogen.
- Beispiel für unterschiedliche Gewichtungen benachbarter Pixel

0	0	0	0	0	0	0
0	0	0.06	0.11	0.06	0	0
0	0.06	0.28	0.44	0.28	0.06	0
0	0.11	0.44	1	0.44	0.11	0
0	0.06	0.28	0.44	0.28	0.06	0
0	0	0.06	0.11	0.06	0	0
0	0	0	0	0	0	0

$w_{3,3}$

0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0.01	0.09	0.19	0.25	0.19	0.09	0.01	0
0	0.06	0.25	0.56	1	0.56	0.25	0.06	0
0	0.01	0.09	0.19	0.25	0.19	0.09	0.01	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

$w_{4,2}$

  - Die Gewichtung kann für alle Pixel gleich gewählt werden, sie kann aber auch über die Bildfläche hinweg variieren.
  - Durch die Gewichtung $w_{1,1}$ wird der euklidische Abstand beschrieben.

Umgebungsbasierte Distanzfunktion (3)

Formale Definition

- Lokale Nachbarschaften: zu jedem Pixel *p* zweier Bilder *F*, *G* werden die Pixel *p'* in der Umgebung betrachtet:

  $$\boxed{\mathrm{d}_W(F,G)|_p = \sum\nolimits_{p' \text{ Pixel}} w(p-p')\cdot(F(p')-G(p'))}$$

  Im obigen Modell sind nur wenige Gewichte $w(\Delta) = w(p - p')$ ungleich Null.
- Gesamtdistanz: die mit dem lokalen Umgebungsabstand gewichteten Pixeldifferenzen werden über alle Pixel der gesamten Bildfläche addiert:

$$\mathrm{d}_W(F,G)^2 = \sum_{p \text{ Pixel}} (F(p)-G(p))\cdot \mathrm{d}_w(F,G)|_p =$$

$$= \sum_{p \text{ Pixel}} (F(p)-G(p))\cdot \sum_{p' \text{ Pixel}} w(p-p')\cdot(F(p')-G(p')) =$$

$$= \sum_{p \text{ Pixel}} \sum_{p' \text{ Pixel}} (F(p)-G(p))\cdot w(p-p')\cdot(F(p')-G(p')) =$$

$$= (F-G)\cdot W\cdot(F-G)^T$$

Umgebungsbasierte Distanzfunktion (4)

Beobachtung:

- Die Distanzfunktion $\mathrm{d}_w(F, G)$ ist eine quadratische Form.
- Die Ähnlichkeitsmatrix *W* enthält für jedes Pixelpaar *p*, *p'* den Eintrag $w(p - p')$.

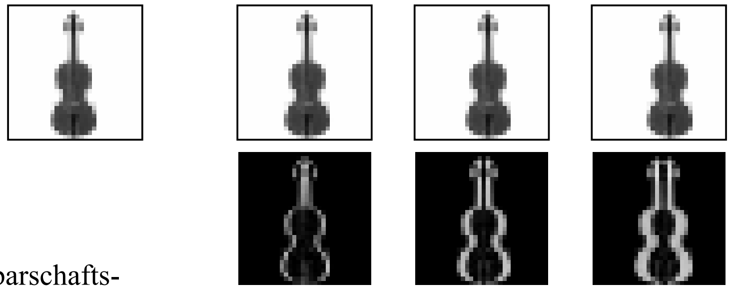
Übertragung auf Farbbilder

- Bisher werden Pixeldifferenzen $F(p) - G(p)$ bezüglich der Grauwerte benutzt.
- Bei Farbbildern berechnet man nicht Grauwertdifferenzen $F(p) - G(p)$, sondern Farbabstandswerte $\mathrm{d}_C(F(p), G(p))$ bzgl. einer Farbdistanzfunktion d_C.

Experimentelle Ergebnisse

- Testdatenbank mit 10.000 Clip Arts der Auflösung 32×32 = 1.024D

Referenzbild +1 Pixel +2 Pixel +3 Pixel



verschobene Anfragebilder

Differenzbilder

Nachbarschafts-bereiche	+1 Pixel	+2 Pixel	+3 Pixel
$w_{1,1}$	1	9	328
$w_{5,1}$	1	3	92
$w_{12,1}$	1	3	10

jeweiliger Rang des Referenzbildes

- Beobachtung: Für den Nachbarschaftsbereich $w_{12,1}$ wird die um drei Pixel nach rechts verschobene Violine unter die “Top-Ten” eingeordnet.

Datenbanksysteme II
Kapitel 7: Ähnlichkeitsmodelle für Bilder

Beispiel für eine Benutzeroberfläche

- Anfragespezifikation: Sketch Editor, Ellipsoid Editor, Parameter k.
- Ergebnisausgabe: Anfragebild, Ergebnisbilder, Distanzwerte, Differenzbilder