

2.3 Ähnlichkeitsmodelle für Bilder

Inhaltsbezogene Suche in Bilddatenbanken

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

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

Merkmale von Bildern

- Farbe
 - Farbhistogramme (QBIC) [HSE+ 95]
- Textur
 - Beschaffenheit von Bildsegmenten (z.B. Holzmaserung, Kieselsteine, Karomuster)
 - Evaluierung verschiedener Distanzfunktionen [PBRT 99]
- Formen (Konturen)
 - Algebraische Moment-Invarianten [TC 91] [FBF+ 94]
 - Pixelbasierte Ähnlichkeitsmodelle [WJ 96] [AKS 98]
 - Morphologisches Ähnlichkeitsmodell [KSF+ 98]

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!

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

2.3.1 Farbhistogramme und Quadratische Formen (QBIC)

[HSE+ 95] Hafner J., Sawhney H. S., Equitz W., Flickner M., Niblack W.: *Efficient Color Histogram Indexing for Quadratic Form Distance Functions*. PAMI 17(7), 1995, 729-736.

Farbhistogramme

- 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 $4 \times 4 \times 4 = 64$ Farben oder $8 \times 8 \times 8 = 512$ 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.
 - Beispiel (64 Repräsentanten):

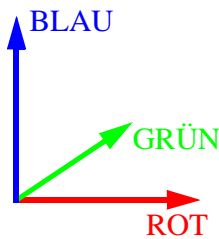


Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

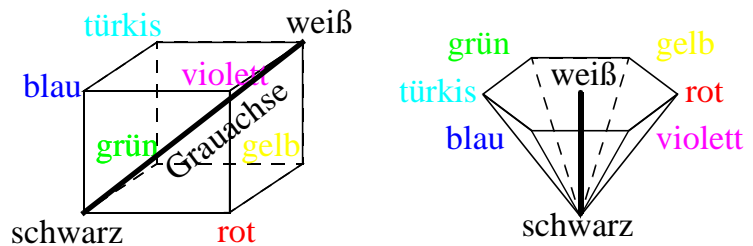
Farbmodelle

Technische Modelle (RGB, CMY) und anschauliche Modelle (HSV, HLS)

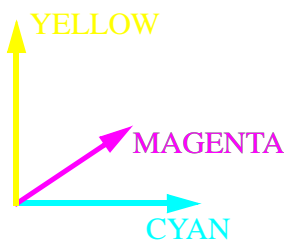
RGB-Modell
(Bildschirm, additiv)



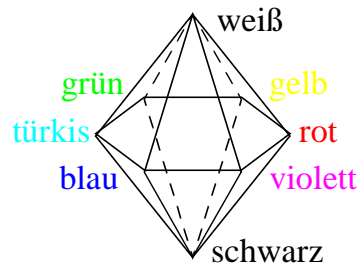
HSV-Modell: Hue, Saturation, Value
(Farbton, Sättigung, Helligkeit)



CMY-Modell
(Drucker, subtraktiv)



HLS-Modell: Hue, Luminance, Saturation
(Farbton, Leuchtkraft, Sättigung)

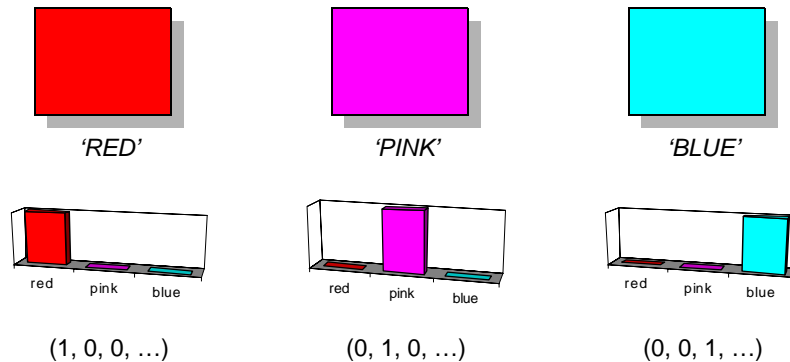


Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

Distanzfunktionen

- Beispiel euklidische Distanz: Seien H^P und H^Q die Farbhistogramme der Bilder P und Q .

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



Es gilt: $D(\text{'RED'}, \text{'PINK'}) = \sqrt{2}$, $D(\text{'RED'}, \text{'BLUE'}) = \sqrt{2}$, $D(\text{'PINK'}, \text{'BLUE'}) = \sqrt{2}$

- Problem der euklidischen Distanz
 - Im Beispiel haben alle Paare von Bildern denselben Abstandswert $\sqrt{2}$.
 - Der Sachverhalt, dass rot ähnlicher zu lila als zu blau ist, wird nicht beachtet.
 - Hintergrund: Querbezüge zwischen Dimensionen werden grundsätzlich negiert.

Quadratische Formen als Distanzfunktionen

- Definition. Sei A eine Ähnlichkeitsmatrix, dann gilt:

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

- 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:

$$D_{A'}(\text{'RED'}, \text{'PINK'}) = \sqrt{0,2}, \quad D_{A'}(\text{'RED'}, \text{'BLUE'}) = \sqrt{2}, \quad D_{A'}(\text{'PINK'}, \text{'BLUE'}) = \sqrt{2}$$

- Beispiele für Ähnlichkeitsmatrizen (vgl. [HSE+ 95]; d_{ij} ist der Abstand der Bins i und j):
 - $a_{ij} = (1 - d_{ij} / d_{\max})$
 - $a_{ij} = \exp(-\sigma (d_{ij} / d_{\max})^2)$ (für $\sigma \rightarrow \infty$ erhält man die Einheitsmatrix)
 - QBIC verwendet eine aus Ergebnissen der Perzeptionsforschung abgeleitete Matrix.

Eigenschaften von Ähnlichkeitsmatrizen

- Symmetrie

Wir dürfen annehmen, dass Ähnlichkeitsmatrizen immer symmetrisch sind, denn:

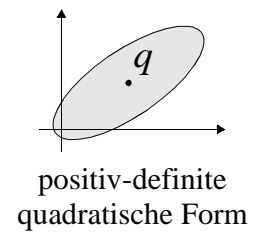
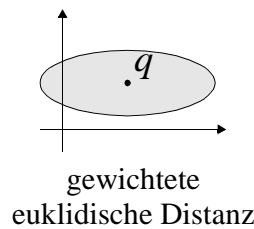
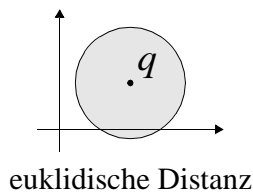
Lemma. Zu jeder Matrix A' gibt es eine symmetrische Matrix $A = (A' + A'^T)/2$ mit

$$D_A(P, Q) = D_{A'}(P, Q).$$

Beweis. Sei $\Delta = H^P - H^Q$, dann gilt:

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

- Geometrie von ϵ -Anfragen



Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

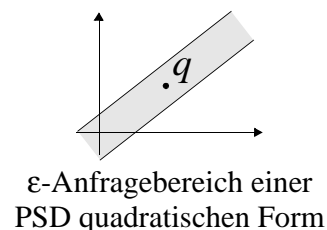
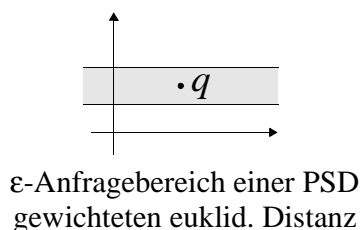
- Positiv definite Matrizen (PD)

- Ähnlichkeitsmatrizen müssen positiv definit sein.
- Definition aus der linearen Algebra: A ist PD 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.

- Positiv semidefinite Matrizen (PSD)

- Für bestimmte Anwendungen kann man sich semi-definite Matrizen vorstellen.
- Definition: A ist PSD gdw. $x A x^T \geq 0$ für alle x .
- 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 $\lambda \in \mathbb{R}$.



Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

Anfragebearbeitung für Quadratische Formen

- 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 $O(n \cdot d^2)$. 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 $W = \text{diag}(w_1, \dots, w_n)$ sowie eine orthogonale Matrix V , d.h. $V V^T = V^T V = Id$, sodass gilt:

$$A = V W V^T$$

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

$$D_A(p, q) = \sqrt{(p - q) \cdot V W V^T \cdot (p - q)^T} = \sqrt{(pV - qV) \cdot W \cdot (pV - qV)^T} = D_W(pV, qV)$$

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

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

- 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 $D_W = D_A$:

$$D_{W,r}(pV, qV) = \sqrt{\sum_{i=1}^r w_i (pV_i - qV_i)^2} \leq \sqrt{\sum_{i=1}^d w_i (pV_i - qV_i)^2} = D_W(pV, qV)$$

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

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

2.3.2 Texturen in Bildern

Die Textur beschreibt die Beschaffenheit von Bildsegmenten (dargestellte Oberflächen).

Texturmodell in QBIC

[FBF+ 94] Faloutsos C., Barber R., Flickner M., Hafner J., et al.: *Efficient and Effective Querying by Image Content*. Journal of Intelligent Information Systems 3, 231-262, 1994.

- **Gerichtetheit**, Orientiertheit (Directionality)
 - Vorhandensein von (Vorzugs-)Richtungen
 - Beispiel: Mauerfugen versus Kieselsteine
 - aus Verteilung der Gradientenrichtungen in den Bildern
- **Kontrast** (Contrast)
 - Lebendigkeit (Unruhe) eines Musters
 - Beispiel: weiße Wand versus Sand
 - Berechnung aus der Varianz im Grauwert histogramm
- **Granularität** (Coarseness)
 - Größenordnung der Textur
 - Beispiele: Sand vs. Kieselsteine; feine vs. grobe Mauer
 - Berechnung durch über das Bild verschobene Fenster unterschiedlicher Größe



Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

2.3.3 Evaluierung verschiedener Distanzfunktionen für Farben und Texturen

[PBRT 99] J. Puzicha, J.M. Buhmann, Y. Rubner, C. Tomasi: *Empirical Evaluation of Dissimilarity Measures for Color and Texture*. Int. Conf. on Computer Vision 1999, 1165-1173.

Repräsentation von Bildinhalten (Farben, Texturen) durch Histogramme

- Charakterisierung
 - Histogramme sind eine diskrete Beschreibung der Verteilung von Featurewerten.
 - Die Diskretisierung basiert auf einer Vorgabe von n Prototypen $\{c_i \mid i = 1, \dots, n\}$.
 - Die resultierenden Histogramme haben n Dimensionen (Bins).
- Formale Definition
 - Für jedes Pixel $I(x)$ wird der nächstgelegene Prototyp c_j bestimmt (Voronoi-Modell).
 - Der Wert $h_I(i)$ ist die Anzahl der dem Prototypen c_i zugeordneten Pixel im Bild I :
$$h_I(i) = \left| \{x \mid \|I(x) - c_i\| < \|I(x) - c_j\|, j = 1, \dots, n\} \right|$$
- Reguläres vs. adaptives Binning
 - Problem: Wieviele Bins sollen verwendet werden?
 - Reguläres Binning: Für alle Bilder werden dieselben Prototypen $\{c_i\}$ verwendet.

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

— Adaptive Binning: Die Menge der Prototypen $\{c_i\}$ wird jeweils dem Bild angepaßt.

- Randhistogramme

— Für mehrdimensionale Features kann man die Randverteilungen darstellen, d.h. die Verteilung pro Dimension. Für eine Dimension r gilt dann:

$$h_I^r(i) = |\{x \mid t_{i-1}^r < I^r(x) \leq t_i^r\}|$$

wobei dem Bin i in der Dimension r das Featureintervall $(t_{i-1}^r, t_i^r]$ zugeordnet ist.

— Nachteil: Informationen über Korrelationen der Dimensionen gehen verloren.

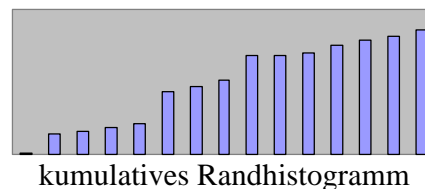
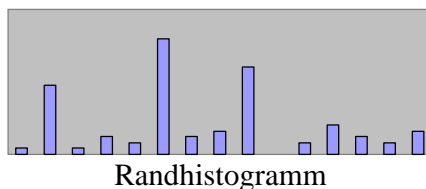
— Vorteil: Geringere Anzahl von Bins als bei Betrachtung des gesamten Raumes (Oberfläche statt Volumen).

- Kumulative Randhistogramme

— Die Einträge in den Bins werden mit aufsteigendem i aufsummiert:

$$H_I^r(i) = |\{x \mid I^r(x) \leq t_i^r\}|$$

— Vorteil: Größere Robustheit gegenüber Ausreißern.



Distanzfunktionen als (Un-)Ähnlichkeitsmaße

Im folgenden bezeichnet $D(I, J)$ die Distanz der beiden Bilder I und J , und $D^r(I, J)$ die Distanz beschränkt auf die Randverteilung in der Dimension r .

- **Heuristische Histogramm-Distanzen**

Gebräuchlichste Form von Distanzfunktionen

— Minkowski-Distanz L_p

$$D(I, J) = \left(\sum_i |h_I(i) - h_J(i)|^p \right)^{1/p}$$

Beispiele: euklidische Distanz (L_2), Summendistanz (L_1), Maximumsdistanz (L_∞)

— Weighted-Mean-Variance (WMV)

$$D^r(I, J) = \frac{|\mu_r(I) - \mu_r(J)|}{|\sigma(\mu_r)|} + \frac{|\sigma_r(I) - \sigma_r(J)|}{|\sigma(\sigma_r)|}$$

wobei $\mu_r(I)$, $\mu_r(J)$ die Mittelwerte und $\sigma_r(I)$, $\sigma_r(J)$ die Standardabweichungen bezeichnen.

- **Nicht-parametrische Teststatistiken**

Gemessen wird die Wahrscheinlichkeit für die Hypothese, daß die beiden Verteilungen I und J empirische Beobachtungen derselben zugrundeliegenden wahren Verteilung sind.

— Kolmogorov-Smirnov (KS)

Maximale Diskrepanz zwischen den kumulativen Randverteilungen:

$$D^r(I, J) = \max_i |H_I^r(i) - H_J^r(i)|$$

— Statistiken vom Cramer/von Mises-Typ (CvM)

Ebenfalls über die kumulativen Randverteilungen definiert:

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

$$D^r(I, J) = \sum_i (H_I^r(i) - H_J^r(i))^2$$

— χ^2 -Statistik

Sei $h'(i) = (h_I(i) + h_J(i))/2$ der gemeinsame Schätzwert der beiden Verteilungen:

$$D^r(I, J) = \sum_i \frac{(h_I(i) - h'(i))^2}{h'(i)}$$

- **Informationstheoretische Divergenzen**

Gemessen wird, wie kompakt eine Verteilung I unter Verwendung der Verteilung J als Code codiert werden kann (vgl. Entropie).

— Kullback-Leibler-Divergenz (KL)

$$D(I, J) = \sum_i h_I(i) \cdot \log \frac{h_I(i)}{h_J(i)}$$

— Jeffrey-Divergenz (JD) oder Jensen-Shannon-Divergenz

Im Gegensatz zu KL ist JD symmetrisch und numerisch stabiler:

$$D(I, J) = \sum_i h_I(i) \cdot \log \frac{h_I(i)}{h'(i)} + h_J(i) \cdot \log \frac{h_J(i)}{h'(i)}$$

- **Grunddistanzbasierte Maße**

- Basis ist eine Grunddistanz zwischen den einzelnen Features (Bins), die durch Expertenwissen aus der Anwendung gedeckt sind (Bsp. Perzeptionsforschung bzgl. Farben und Texturen in Bildern).

- Quadratische Formen (QF)

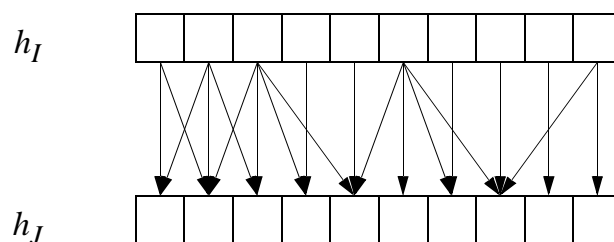
Die Einträge a_{ij} der Ähnlichkeitsmatrix $A = [a_{ij}]$ beschreiben die Ähnlichkeit (Korrelation) der Features (Bins) i und j (z.B. $a_{ij} = e^{-\sigma(d_{ij}/d_{\max})^2}$):

$$D(I, J) = \sqrt{(h_I - h_J) \cdot A \cdot (h_I - h_J)^T}$$

Um eine Metrik zu erhalten, muß die Matrix A positiv definit sein.

- Earth Movers Distance (EMD)

Idee: Minimale Kosten, um eine Verteilung I in eine Verteilung J umzuwandeln. D.h. Minimale Summe aller Einzelkosten, um eine (Feature-)Einheit von einem Bin zu einem anderen zu verschieben.



Berechnung durch lineare Optimierung als Lösung eines Transportproblems:

$$D(I, J) = \frac{\sum_{i,j} g_{ij} d_{ij}}{\sum_{i,j} g_{ij}}$$

wobei d_{ij} die Unähnlichkeit der Bins i und j beschreibt und g_{ij} den optimalen Fluß zwischen den Verteilungen I und J , sodass die Gesamtkosten $\sum_{i,j} g_{ij} d_{ij}$ bezüglich folgender Nebenbedingungen minimal sind:

- (i) $\sum_i g_{ij} \leq h_J(j)$ Häufe auf $h_J(j)$ nicht mehr als möglich.
- (ii) $\sum_j g_{ij} \leq h_I(i)$ Ziehe von $h_I(i)$ nicht mehr als möglich ab.
- (iii) $\sum_{i,j} g_{ij} = \min(\sum_i h_I(i), \sum_j h_J(j))$ Bewege eine der Verteilungen vollständig.

- EMD ist eine Metrik, falls die Grunddistanz d_{ij} eine Metrik bildet und die beiden Verteilungen die gleiche Masse enthalten (vgl. Bedingung iii).
- EMD unterstützt adaptives Binning.

Eigenschaften

	L_p	WMV	KS/ CvM	χ^2	KL	JD	QF	EMD
Symmetrisch	+	+	+	+	o	+	+	+
Dreiecksungleichung	+	+	+	o	o	o	+*	+*
Berechnungskomplexität	mittel	niedrig	mittel	mittel	mittel	mittel	hoch	hoch
Grunddistanzbasiert	o	o	+	o	o	o	+	+
Individuelles Binning	o	+	o	o	o	o	o	+
Mehrdimensional	+	+	o	+	+	+	+	+
Partielle Ähnlichkeit	o**	o	o	o	o	o	o	+
Nicht-parametrisch	+	o	+	+	+	+	***	+

* QF und EMD sind Metriken, falls die obengenannten Bedingungen erfüllt sind.

** nur in einer speziellen Form (“Histogramm-Schnitt”).

*** z.B. aber auch mit Parameter σ .

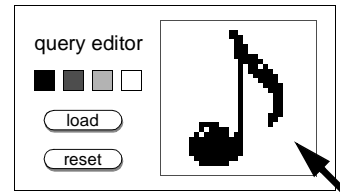
Experimentelle Ergebnisse

- Die Ähnlichkeitsmaße eignen sich für verschiedene Anwendungen unterschiedlich gut.
- EMD ist am teuersten auszuwerten, liefert aber oft die besten Ergebnisse.

2.3.4 Pixelbasiertes Modell für Formen in Bildern

[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

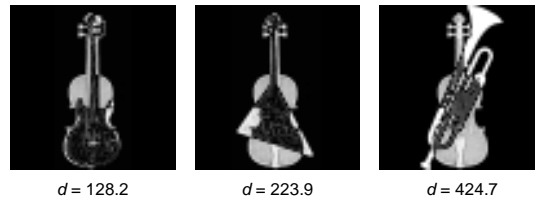


- Konzept der Differenzbilder

Anfrage- und Ergebnisbilder (64x64 Pixel):

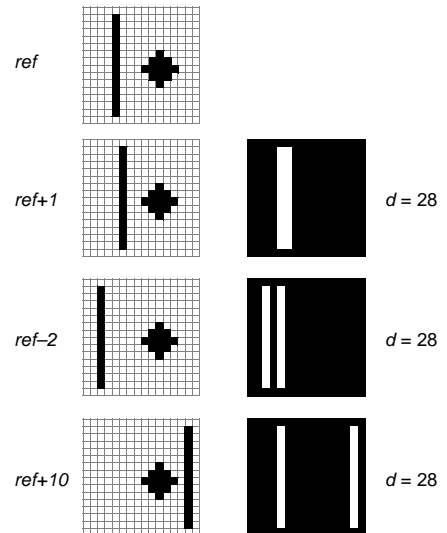


Differenzbilder und euklidische Distanz:



Probleme mit der euklidischen Distanz

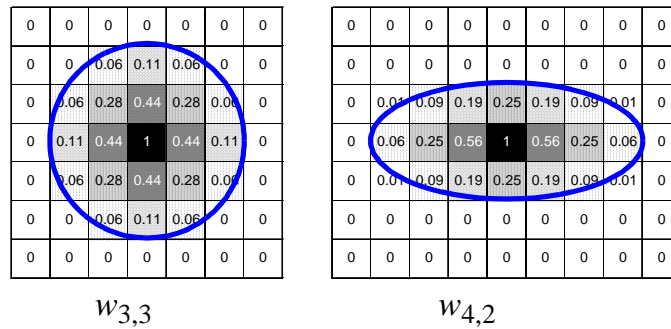
- Beispiel "ref"
 - Referenzbild: Balken links von einem Punkt.
 - Bei den Vergleichsbildern ist der Balken um +1, -2 bzw. +10 Pixel horizontal verschoben.
 - Euklidischer Abstand zu ref bleibt jedoch immer derselbe.
- Problembeschreibung
 - Leichte Verschiebungen sind von starken Veränderungen nicht unterscheidbar.
 - Invarianz gegenüber globalen Translationen stellt keine Lösung dar.
 - Erwünscht ist die Robustheit gegenüber kleinen, lokalen Veränderungen.



Umgebungsbasierte Distanzfunktion

- Lösungsidee: Betrachte die Nachbarschaften der Pixel
 - Statt nur die direkt übereinanderliegenden Pixel wie im Differenzbild zu betrachten, werden nun auch benachbarte Pixel zur Ähnlichkeitsbewertung herangezogen.

- Beispiel für unterschiedliche Gewichtungen benachbarter Pixel



- 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.

- Formale Definition

- Lokale Nachbarschaften

Zu jedem Pixel p zweier Bilder F, G werden die Pixel p' in der Umgebung betrachtet:

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

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:

$$\begin{aligned} d_w(F, G)^2 &= \sum_{p \text{ Pixel}} (F(p) - G(p)) \cdot d_w(F, G)|_p = \\ &= \sum_{p \text{ Pixel}} (F(p) - G(p)) \cdot \sum_{p' \text{ Pixel}} w(p - p') \cdot (F(p') - G(p')) = \\ &= \sum_{p \text{ Pixel}} \sum_{p' \text{ Pixel}} (F(p) - G(p)) \cdot w(p - p') \cdot (F(p') - G(p')) = \\ &= (F - G) \cdot W \cdot (F - G)^T \end{aligned}$$

- Beobachtung:

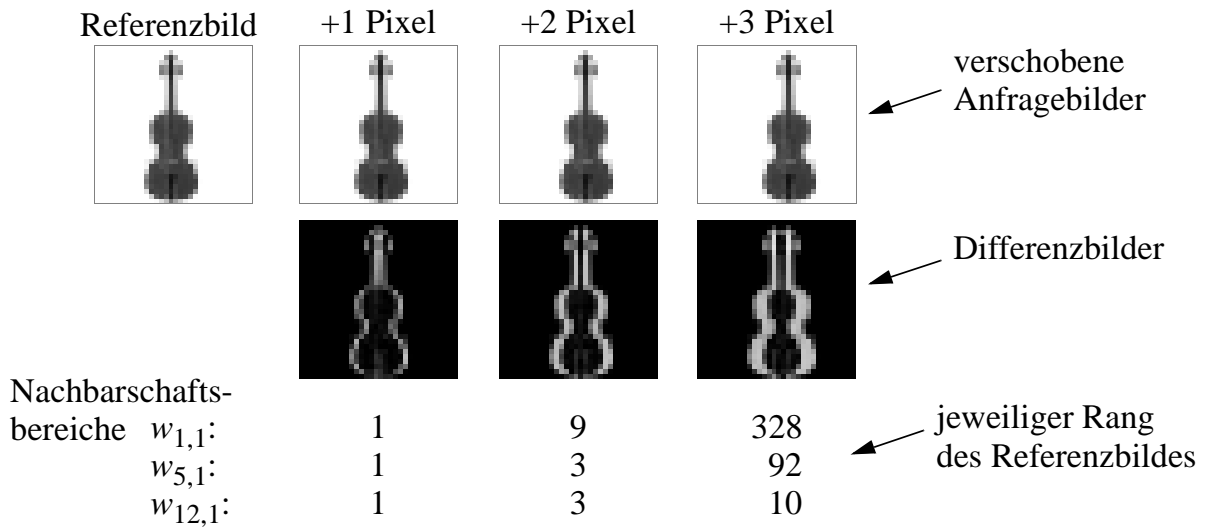
- Die Distanzfunktion $d_w(F, G)$ ist eine quadratische Form.
- Die Ähnlichkeitsmatrix W enthält für jedes Pixelpaar p, p' den Eintrag $w(p - p')$.

- 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 $d_C(F(p), G(p))$ bzgl. einer Farbdistanzfunktion d_C .

Experimentelle Ergebnisse

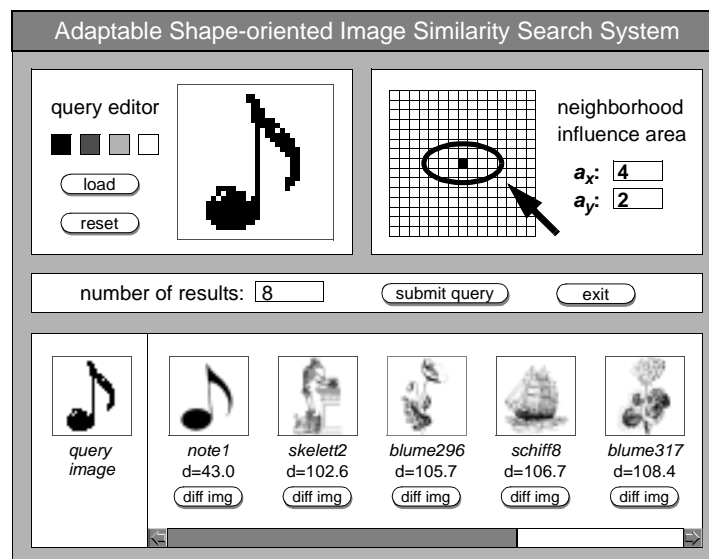
- Testdatenbank mit 10.000 Clip Arts der Auflösung $32 \times 32 = 1.024D$



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

Beispiel für eine Benutzeroberfläche

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



2.3.5 Morphologisches Modell für Formen in Bildern

[KSF+ 98] Korn F., Sidiropoulos N., Faloutsos C., Siegel E., Protopapas Z.: *Fast and Effective Retrieval of Medical Tumor Shapes*. IEEE Transactions on Knowledge and Data Engineering (TKDE) 10(6), 1998, 889-904.

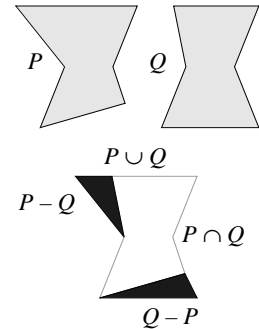
- Motivation
 - Betrachtung von Formen unabhängig von Lage und Orientierung im Bild.
 - Oberflächenbeschaffenheit (rauhe oder glatte Kontur) bzgl. unterschiedlicher Größenmaßstäbe (feingliedrige bzw. grobgliedrige Oberflächenstruktur) wichtig.

Distanz von Formen

- Symmetrische Flächendifferenz zweier Formen P und Q :

$$d(P, Q) = |P - Q| + |Q - P| = |P \cup Q| - |P \cap Q|$$
- Translations- und rotationsinvariante Distanz d^* zweier Formen:
 Sei \mathfrak{R} die Menge aller euklidischen Transformationen (Translationen und Rotationen), dann gilt für zwei Formen P, Q :

$$d^*(P, Q) = \inf_{R \in \mathfrak{R}} d(P, R(Q))$$



Die Infimumsbildung über alle Translationen und Rotationen bewirkt die bestmögliche aller Überlagerungen der beiden Formen P und Q .

- Problem: Rauheit von Oberflächenstrukturen wird nicht adäquat berücksichtigt.

Operationen der mathematischen Morphologie

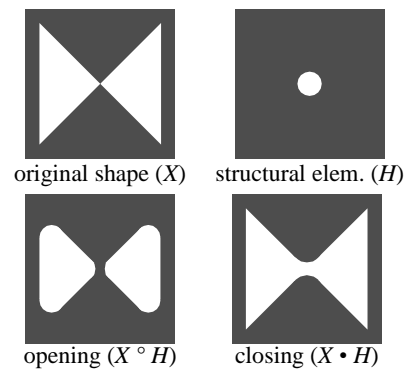
- Hier: Opening und Closing
 Ein strukturelles Element H (z.B. Kreis) wird wie folgt auf eine Form X angewandt:

Opening $X \circ H$:

Zu $X \circ H$ gehören diejenigen Teile von X , die H abdecken kann, ohne X zu verlassen.

Closing $X \bullet H$:

Zu $X \bullet H$ gehört alles, was H nicht abdecken kann, ohne in X einzudringen.



Quelle: [KSF+ 96]

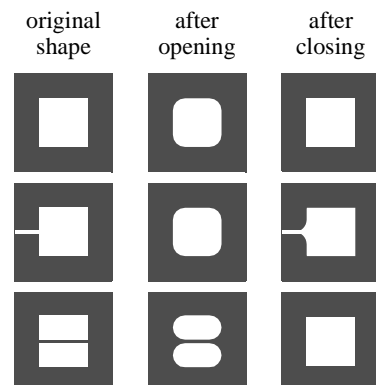
- Auswirkungen
 Opening und Closing haben unterschiedliche Auswirkungen auf die Oberfläche einer Form:

- Opening:

- feine Ausbuchtungen werden entfernt.
- feine Einbuchtungen werden verstärkt.

- Closing:

- feine Ausbuchtungen werden verstärkt.
- feine Einbuchtungen werden entfernt.



Quelle: [KSF+ 96]

Morphologiebasierte Distanz von Formen

- Vektoren von morphologisch veränderten Formen

Die Oberflächenstruktur einer Form wird einbezogen, indem für ein Element H in verschiedenen Größen m ($-M \leq m \leq M$) die entsprechenden Openings ($m > 0$) und Closings ($m < 0$) gebildet und in einem $(2 \cdot M + 1)$ -dim. Vektor $f^H(X)$ zusammengefaßt werden:

$$f^H(X) = \underbrace{(f_{-M}^H(X), \dots, f_{-1}^H(X))}_{M \text{ openings}} \underbrace{(f_0^H(X))}_{= X} \underbrace{(f_1^H(X), \dots, f_M^H(X))}_{M \text{ closings}} \quad \text{wobei } f_m^H(X) = \begin{cases} X \circ mH & m > 0 \\ X & m = 0 \\ X \bullet mH & m < 0 \end{cases}$$

- *Morphologische Distanz*

Für zwei Vektoren $f^H(X_1)$ und $f^H(X_2)$ werden komponentenweise die Distanzen d^* berechnet, über die dann eine p -Norm gebildet wird (Summe, euklidisch, Maximum):

$$d_{\text{morph}, p}^H(X_1, X_2) = \left(\sum_{m=-M}^M |d^*(f_m^H(X_1), f_m^H(X_2))|^p \right)^{1/p}$$

- Die morphologische Distanz $d_{\text{morph}, p}^H$ ist eine Metrik.

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

Granulometrische Distanz

- Probleme der morphologischen Distanz

- Exakte Objektrepräsentation muß im Hauptspeicher vorliegen.
- Berechnung ist sehr teuer (insbesondere Infimumsbildung für d^* -Berechnungen).
- Erwünscht ist ein einfacheres Distanzmaß, das sich für einen Filterschritt eignet.
- Idee: Granulometrisches Formenspektrum.

- Größenverteilung von Formen

Die Größenverteilung einer Form X bezüglich eines strukturellen Elements H ist definiert als der $(2 \cdot M + 1)$ -dimensionale Vektor $sd(X)$ ("size distribution"):

$$sd(X) = (|f_{-M}^H(X)|, \dots, |f_{-1}^H(X)|, |f_0^H(X)|, |f_1^H(X)|, \dots, |f_M^H(X)|)$$

- Invarianzeigenschaft: Die Größenverteilung beinhaltet nur Flächenmaßzahlen und ist deswegen eine translations- und rotationsinvariante Darstellung von Formen.
- Die Größenverteilung kann als kumulatives Formenspektrum angesehen werden: Die Differenzen der $|f_m^H(X)|$ beschreiben das Ausmaß der geometrischen Details, die beim nächstgrößeren Opening bzw. Closing entfernt bzw. hinzugefügt werden.

Skript Multimedia-Datenbanksysteme · Modelle der Datenexploration

- *Granulometrische Distanz*

Für ein strukturelles Element H wird eine p -Norm über dem $(2 \cdot M + 1)$ -dimensionalen Differenzvektoren der Größenverteilungen gebildet:

$$\delta_{\text{gran}, p}^H(X_1, X_2) = \|sd(X_1) - sd(X_2)\|_p$$

- Lemma

Die granulometrische Distanz ist eine untere Schranke der morphologischen Distanz:

$$\delta_{\text{gran}, p}^H(X_1, X_2) \leq d_{\text{morph}, p}^H(X_1, X_2)$$

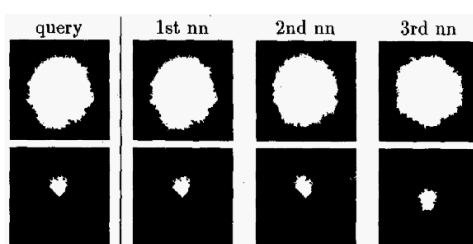
Beweis. Es gilt $d^*(X_1, X_2) \geq ||X_1| - |X_2||$, wobei die Gleichheit nur erreicht wird, wenn X_1 und X_2 durch eine rigide Transformation $R \in \mathfrak{R}$ zur Deckung gebracht werden. Dann gilt auch für alle m : $d^*(f_m(X_1), f_m(X_2)) \geq ||f_m(X_1)| - |f_m(X_2)||$ und insbesondere:

$$\left(\sum_m |d^*(f_m(X_1), f_m(X_2))|^p \right)^{1/p} \geq \left(\sum_m ||f_m(X_1)| - |f_m(X_2)||^p \right)^{1/p} \quad \text{q.e.d.}$$

Bemerkung. Das Lemma gewährleistet die Vollständigkeit einer mehrstufigen Anfragebearbeitung für die morphologische Distanz unter Verwendung der granulometrischen Distanz im Filterschritt.

Architektur der Anfragebearbeitung

- Datenhaltung
 - Die exakten Formen werden im Sekundärspeicher gehalten.
 - Die Größenverteilungen (Vektoren) werden in einem mehrdimensionalen Index verwaltet.
- Anfragebearbeitung
 - Berechne die Größenverteilung $sd(Q)$ der Anfrageform Q und starte mit diesem Vektor eine Anfrage auf dem Index bezüglich der granulometrischen Distanz.
 - Hole für die gewonnenen Kandidaten die exakte Darstellung vom Sekundärspeicher und berechne die morphologische Distanz (ca. 12.69 Sekunden im Mittel pro Paar).
- Experimentelles Ergebnis zweier Ähnlichkeitsanfragen



Quelle: [KSF+ 96]

Referenzen

- [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.
- [CG 99] S. Cohen, L. Guibas: *The Earth Mover's Distance under Transformation Sets*. Proc. Int. Conf. on Computer Vision (ICCV) 1999.
- [FBF+ 94] Faloutsos C., Barber R., Flickner M., Hafner J., Niblack W., Petkovic D., Equitz W.: *Efficient and Effective Querying by Image Content*. Journal of Intelligent Information Systems 3, 231-262, 1994.
- [FSN+ 95] Flickner M., Sawhney H., Niblack W., Ashley J., Huang Q., Dom B., Gorkani M., Hafner J., Lee D., Petkovic D., Steele D., Yanker P.: *Query by Image and Video Content: The QBIC System*. IEEE Computer, 23-32, Sept. 1995.
- [HSE+ 95] Hafner J., Sawhney H. S., Equitz W., Flickner M., Niblack W.: *Efficient Color Histogram Indexing for Quadratic Form Distance Functions*. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 17(7), 1995, 729-736.
- [KSF+ 96] Korn F., Sidiropoulos N., Faloutsos C., Siegel E., Protopapas Z.: *Fast Nearest Neighbor Search in Medical Image Databases*. Proc. Int. Conf. on Very Large Data Bases (VLDB) 1996, 215-226.
- [KSF+ 98] Korn F., Sidiropoulos N., Faloutsos C., Siegel E., Protopapas Z.: *Fast and Effective Retrieval of Medical Tumor Shapes*. IEEE Transactions on Knowledge and Data Engineering (TKDE) 10(6), 1998, 889-904.
- [MM 96] Manjunath B. S., Ma W. Y.: *Texture Features for Browsing and Retrieval of Image Data*. IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI) 18(8): 837-842, 1996.
- [NBE+ 93] Niblack W., Barber R., Equitz W., Flickner M., Glasmann E., Petkovic D., Yanker P., Faloutsos C., Taubin G.: *The QBIC Project: Querying Images by Content Using Color, Texture, and Shape*. SPIE 1993 Int. Symposium on Electronic Imaging: Science and Technology Conference 1908, Storage and Retrieval for Image and Video Databases, San Jose, CA, 1993.
- [PBRT 99] J. Puzicha, J.M. Buhmann, Y. Rubner, C. Tomasi: *Empirical Evaluation of Dissimilarity Measures for Color and Texture*. Int. Conf. on Computer Vision (ICCV) 1999, 1165-1173.
- [RGT 97] Y. Rubner, L. Guibas, C. Tomasi: *The Earth Mover's Distance, Multi-Dimensional Scaling, and Color-Based Image Retrieval*. Proc. ARPA Image Understanding Workshop, 661-668, 1997.
- [RTG 98] Y. Rubner, C. Tomasi, L. J. Guibas: *A Metric for Distributions with Applications to Image Databases*. Proc. 6th IEEE Int. Conf. on Computer Vision (ICCV) 1998, 59-66.
- [TC 91] Taubin G., Cooper D. B.: *Recognition and Positioning of Rigid Objects Using Algebraic Moment Invariants*. in *Geometric Methods in Computer Vision*, SPIE 1570, 1991, 175-186.
- [TMY 78] Tamura H., More S., Yamawaki T.: *Texture Features Corresponding to Visual Perception*. IEEE Trans. on Systems, Man, and Cybernetics (SMC) 8(6): 460-473, 1978.
- [WJ 96] White D.A., Jain R.: *Similarity Indexing with the SS-tree*. Proc. IEEE Int. Conf. on Data Engineering (ICDE) 1996, 516-523.