

# Ähnlichkeit im Kontext der Bildsuche

Vorlesung Bildinformationssysteme, Teil 3  
Sommersemester 2005  
Ullrich Köthe, FB Informatik, Uni Hamburg

## Ähnlichkeit

---

- **Motivation:**
  - keine klar definierte Semantik  $\Rightarrow$  Approximation des „korrekten“ Ergebnisses  $\Rightarrow$  Präsentation mehrerer Möglichkeiten
- **was ist Ähnlichkeit?**
  - monoton nicht-wachsende Funktionen  $S \in [0, 1]$
  - Umrechnung in Distanzen: monoton nicht-fallende Funktionen
  - mathematische Axiome (Ähnlichkeitsstrukturen bzw. Metriken):
    - nützliche Eigenschaften für Berechnungen
    - Zoo der Distanzen
  - psychologische Experimente:
    - Ermittlung des menschlichen Ähnlichkeitsempfindens
    - nicht immer metrisch
    - Ziel: Modellierung der Nutzererwartungen an Ähnlichkeitssuche

## Axiome der Metrik bzw. Ähnlichkeit

---

- **Metrik: Distanzfunktion  $d(x, y) \geq 0$** 
  - $d(x, x) = d(y, y) = 0$       konstante Selbstähnlichkeit
  - $d(x, y) \geq d(x, x)$       Minimalität
  - $d(x, y) = d(y, x)$       Symmetrie
  - $d(x, z) \leq d(x, y) + d(y, z)$       Dreiecksungleichung
- **Ähnlichkeitsfunktion  $s(x, y) \in [0, 1]$** 
  - $s(x, x) = s(y, y) = 1$       konstante Selbstähnlichkeit
  - $s(x, x) \geq s(x, y)$
  - $s(x, y) = s(y, x)$       Symmetrie
  - $s(x, z) \geq s(x, y) + s(y, z) - 1$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Eigenschaften

---

- **Metrische Axiome erlauben:**
  - Definition der Grenzwertbildung
 
$$\lim_{i \rightarrow \infty} x_i = x \Leftrightarrow \lim_{i \rightarrow \infty} d(x_i, x) = 0$$
  - Umrechnung zwischen Ähnlichkeit und Distanz:
 
$$s(x, y) = g(d(x, y)) \text{ ist Ähnlichkeitsfunktion} \Leftrightarrow$$

$$d(x, y) \text{ ist Metrik und}$$

$$\forall p \geq 0: g(p) \geq 0, g(0) = 1, g'(p) \leq 0, g''(p) \geq 0$$

z.B.:  $d = 1 - s, d = -\log s$
  - Schranke für die Distanzen:  $|d(x, x_1) - d(x, x_2)| \leq d(x_1, x_2)$
  - Definition effizienter Indexstrukturen und Suchverfahren
- **ist nicht zwingend für die Bildsuche**

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Minkowski-Distanzen

---

- **Spezialfall:**

- $x, y, z$   $n$ -dimensionale Vektoren
- Additivität der Distanz entlang von Geraden
- d.h.: Gleichheit bei der Dreiecksungleichung, wenn Punkte auf einer Geraden liegen

$d(x, y) + d(y, z) = d(x, z)$ , falls  $y$  zwischen  $x, z$  auf einer Geraden

$$\Rightarrow d_p(x, y) = \left[ \sum_i |x_i - y_i|^p \right]^{\frac{1}{p}} \quad \text{Minkowski - Distanzen (} p \geq 1 \text{)}$$

$$d_{\text{Manhattan}}(x, y) = \sum_i |x_i - y_i|, \quad d_{\text{Chessboard}}(x, y) = \max_i |x_i - y_i|$$

$$d_{\text{Euklid}}(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Abgeleitete Metriken (1)

---

- **Apply function Operator:**

geg. Metrik  $d$  :  $h(d(x, y))$  ist Metrik, wenn

$$h(0) = 0, \quad h(x) > 0 \quad (x > 0)$$

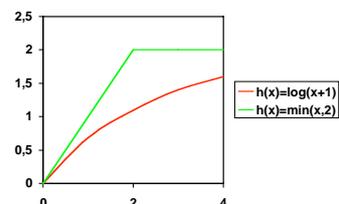
$$h'(x) \geq 0, \quad h''(x) \leq 0$$

oder wenn  $h(x) = \lim_{n \rightarrow \infty} h_n(x)$  wobei alle  $h_n(x)$  obige Bedingungen erfüllen

- **z.B.:**

$$h(x) = \log(x + 1)$$

$$h(x) = \begin{cases} x, & \text{wenn } x < \alpha \\ \alpha, & \text{sonst} \end{cases}$$



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Abgeleitete Metriken (2)

- **Composition Operator**  
geg. Metriken :

$H(d_1(x, y), d_2(x, y))$  ist Metrik, wenn

$$H(0, 0) = 0$$

$$\left. \frac{\partial H(a, b)}{\partial a} \right|_{b=\text{const}} \geq 0, \quad \left. \frac{\partial H(a, b)}{\partial b} \right|_{a=\text{const}} \geq 0 \quad (\forall a, b)$$

Hessesche Matrix  $\begin{pmatrix} H_{aa} & H_{ab} \\ H_{ab} & H_{bb} \end{pmatrix}$  ist negativ semidefinit

oder  $H(a, b) = \lim_{n \rightarrow \infty} H_n(a, b)$  wobei alle  $H_n$  die obigen Bedingungen erfüllen

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Abgeleitete Metriken (3)

- **Beispiele für Kombinationsoperatoren:**

harmonisches, geometrisches, arithmetisches Mittel:

$$H(a, b) = \frac{2ab}{a+b}, \quad H(a, b) = \sqrt{ab}, \quad H(a, b) = \frac{a+b}{2}$$

Minimum, Maximum

$$H(a, b) = \min(a, b), \quad H(a, b) = \max(a, b)$$

gewichtetes Mittel, quasi-lineares Mittel

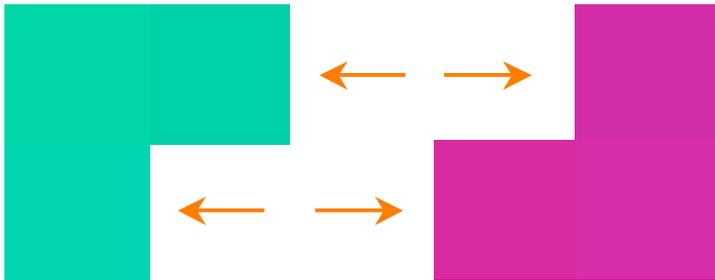
$$H(a, b) = wa + (1-w)b, \quad H(a, b) = (wa^\alpha + (1-w)b^\alpha)^{1/\alpha}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Farbähnlichkeit (1)

- Messung der minimal wahrnehmbaren Farbänderung

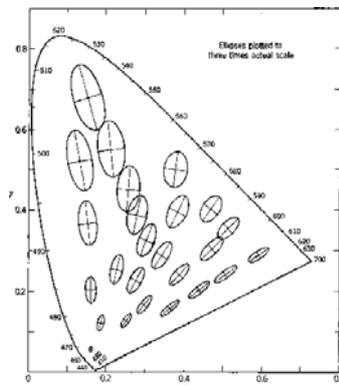
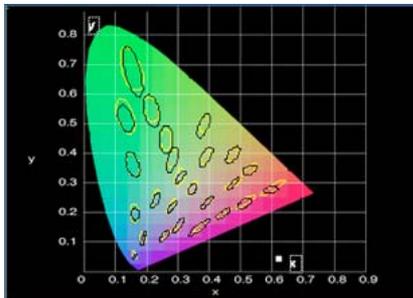


Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Farbähnlichkeit (2)

- Messung für viele Farben und viele Richtungen
- Approximation durch Ellipsoide (McAdams)



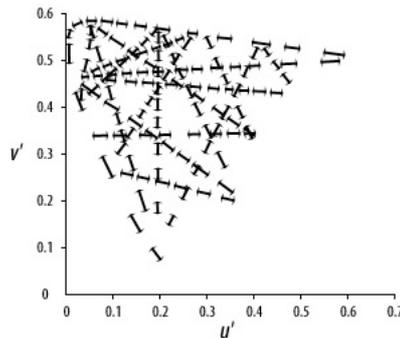
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Farbähnlichkeit (3)

- Transformation des Farbraumes so, daß Ellipsoide zu Kugeln werden
- ⇒ Euklidische Distanz approximiert menschliche Wahrnehmung

$$d_2(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Farbähnlichkeit (4)

- z.B. Verwendung des L\*u\*v\*- oder L\*a\*b\*-Farbraums
  1. Normierung für unterschiedliche Lichtquellen
  2. Berücksichtigung der logarithmischen Rezeptor-Empfindlichkeit (Weber-Fechnersches Gesetz)
  3. affine bzw. nichtlineare Entzerrung des Farbraumes
- ⇒ Verringerung der Verzerrung von 200:1 auf 6:1
- Verwendung des Y'PbPr-Farbraums
  - keine Normierung für unterschiedliche Lichtquellen
  - einfacher zu berechnen
  - fast so gute Ergebnisse wie mit L\*u\*v\* oder L\*a\*b\*

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## L\*u\*v\*- und L\*a\*b\*-Farbraum (1)

---

- **Normierung für unterschiedliche Lichtquellen**

- Transformation auf Wertebereich [0, 1]:  $r = \frac{R}{R_{\max}}, \quad g = \frac{G}{G_{\max}}, \quad b = \frac{B}{B_{\max}}$

- Lichtquellenabhängige Transformation in den XYZ-Farbraum, z.B. Lichtquelle D65 (moderne Farbbildröhre):

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.21267 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix}$$

- Transformation des Weißpunktes (r,g,b) = (1,1,1):

$$(X_n, Y_n, Z_n) = (0.950456, 1.00000, 1.088754)$$

Sommersemester 2005

*Ulrich Köthe: Bildinformationssysteme*

## L\*u\*v\*- und L\*a\*b\*-Farbraum (2)

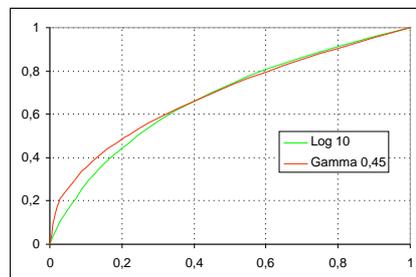
---

- **logarithmische Rezeptorempfindlichkeit: wahrgenommene Helligkeit entspricht Logarithmus der Intensität**

- **Approximation des Logarithmus durch Potenz („Gamma-Korrektur“):**

$$\begin{aligned} \text{gamma}(x) &= x^\gamma, \quad x \in [0,1] \\ &\approx \log_b((b-1)x + 1) \end{aligned}$$

- **L\*u\*v\* und L\*a\*b\*:**  
 $\gamma = 1/3$



Sommersemester 2005

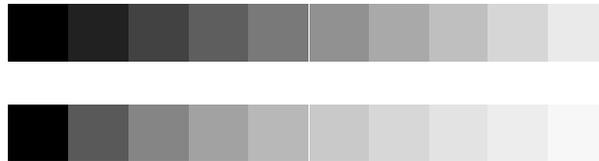
*Ulrich Köthe: Bildinformationssysteme*

## L\*u\*v\*- und L\*a\*b\*-Farbraum (3)

---

- Gamma-Korrektur der Luminanz (mit  $\gamma = 1/3$ )

$$L^* = \begin{cases} 116 \left( \frac{Y}{Y_n} \right)^{1/3} - 16, & \text{if } 0.008856 < \frac{Y}{Y_n} \quad (L^* \geq 8) \\ 903.3 \frac{Y}{Y_n}, & \text{otherwise} \quad (L^* \leq 8) \end{cases}$$



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## L\*u\*v\*- und L\*a\*b\*-Farbraum (4)

---

- XYZ  $\Leftrightarrow$  L\*u\*v\* und L\*a\*b\* mit Referenzweißpunkt ( $X_n, Y_n, Z_n$ )

$$u^? = \frac{4X}{X + 15Y + 3Z}, \quad v^? = \frac{9Y}{X + 15Y + 3Z}$$

$$u^* = 13L^* (u^? - u^?_n), \quad v^* = 13L^* (v^? - v^?_n)$$

$$a^* = 500 \left[ \left( \frac{X}{X_n} \right)^{1/3} - \left( \frac{Y}{Y_n} \right)^{1/3} \right], \quad b^* = 200 \left[ \left( \frac{Y}{Y_n} \right)^{1/3} - \left( \frac{Z}{Z_n} \right)^{1/3} \right]$$

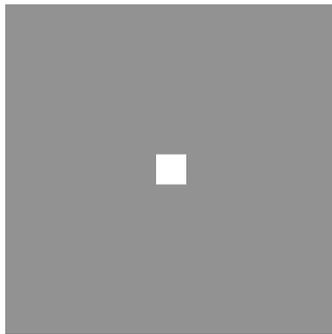
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

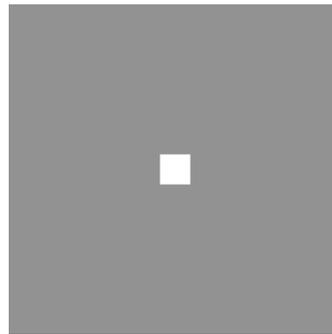
## L\*u\*v\*- und L\*a\*b\*-Farbraum (5)

---

L\*a\*b\* ⇨ RGB



L\*a\*b\* ⇨ R'G'B'



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## RGB, R'G'B' und Y'PbPr (1)

---

- RGB kodiert reine Intensität ⇨ keine Berücksichtigung der logarithmischen Empfindlichkeit  
⇨ R'G'B' als gamma-korrigiertes RGB ( $\gamma = 0.45$ )

$$R' = R_{\max} r^{0.45}, \quad G' = G_{\max} g^{0.45}, \quad B' = B_{\max} b^{0.45}$$

- Y'PbPr als affine Entzerrung von R'G'B':

$$\begin{pmatrix} Y' \\ Pb \\ Pr \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687358916 & 0.3312641084 & 0.5 \\ 0.5 & 0.4186875892 & 0.0813124108 \end{pmatrix} \begin{pmatrix} r' \\ g' \\ b' \end{pmatrix}$$

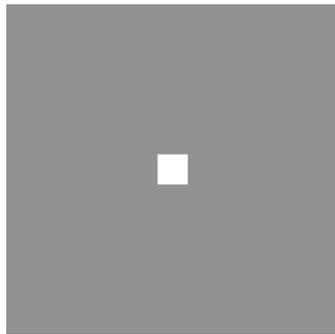
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

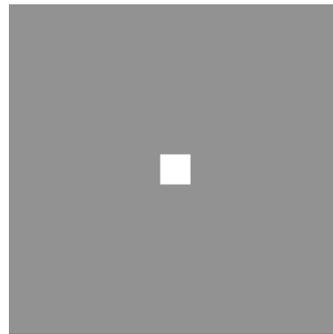
## RGB, R'G'B' und Y'PbPr (2)

---

Y'PbPr  $\Rightarrow$  RGB



Y'PbPr  $\Rightarrow$  R'G'B'



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Probleme mit uniformen Farbräumen

---

- Daten nicht gekennzeichnet: RGB oder R'G'B' ?
- Entzerrung auf der Basis kleiner Differenzen: gilt Transformation auch für große Differenzen ?
- Entzerrung erzeugt nur näherungsweise uniformen Farbraum
- Abhängigkeit der Farbwahrnehmung von Lichtquelle unberücksichtigt
- nur Paare von Farben berücksichtigt: Kontexteffekte (z.B. Farbkonstanz) vernachlässigt
- trotzdem: zur Zeit nichts besseres

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Histogramme und Signaturen

- Ähnlichkeitsfunktion bisher vergleicht einzelne Farben, aber wir wollen ganze Bilder (Mengen von Farben) vergleichen
- Beschreibung von Mengen durch
  - Histogramme:
    - feste Quantisierung des Merkmalsraums („Bins“)
    - Anzahl der Werte pro Bin
  - Signaturen:
    - adaptive Quantisierung des Merkmalsraums („Cluster“)
    - Clusterzentrum und Anzahl der Werte pro Cluster
  - manchmal „Pseudo-Signaturen“ (z.B. Haar-Methode):
    - berücksichtige nur signifikant gefüllte Bins

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Bin-weiser Vergleich (1)

- Bin-weiser Vergleich: einfach und schnell
- z.B. Minkowski-Distanzen (Metriken):

$$D(k,l) = \sum_i |k_i - l_i|, \quad D(k,l) = \sqrt{\sum_i (k_i - l_i)^2}, \quad D(k,l) = \max_i |k_i - l_i|$$

- Histogramm-Schnitt  
(keine Metrik, erlaubt  
partielles Matching):

$$D(k,l) = 1 - \frac{\sum_i \min(k_i, l_i)}{\sum_i k_i}$$

- Jeffrey-Divergenz  
(keine Metrik,  
symmetrisch,  
robust gegen Rauschen und  
Quantisierung)

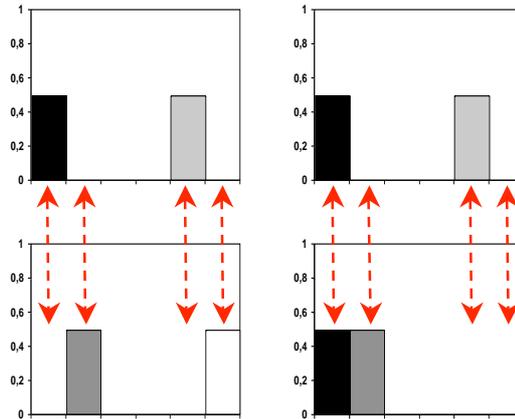
$$D(k,l) = \sum_i \left( k_i \log \frac{k_i}{m_i} + l_i \log \frac{l_i}{m_i} \right)$$
$$m_i = \frac{k_i + l_i}{2}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Bin-weiser Vergleich (2)

- Bin-weise Vergleiche entsprechen oft nicht der Wahrnehmung
- Bsp.: die linken Histogramme werden als ähnlich wahrgenommen, aber ihre Distanz ist größer als die der rechten

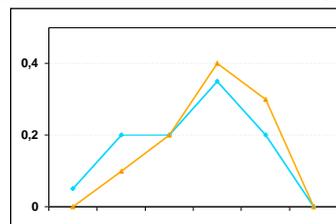


Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Berücksichtigung der Bin-Ähnlichkeit (1)

- Vergleich auf Basis der Summen-Histogramme  $\hat{k}_i = \sum_{j=0}^i k_j$ 
  - Metriken
  - nur für geordnete Daten (1D)
  - Ordnung impliziert Ähnlichkeit
  - funktioniert für adaptiv quantisierte und kontinuierliche Daten

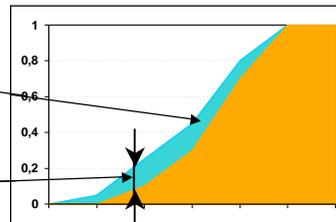


Match-Distanz:

$$D(k, l) = \sum_i |\hat{k}_i - \hat{l}_i|$$

Kolmogorov-Smirnov-Statistik:

$$D(k, l) = \max_i |\hat{k}_i - \hat{l}_i|$$



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Berücksichtigung der Bin-Ähnlichkeit (2)

---

- **Mahalanobis-Distanz (*quadratic form distance*):**
  - Ähnlichkeitsmatrix auf Grund einer Basisdistanz für die Bins
  - ⇒ „Glättung des Histogramms“
  - Ähnlichkeitsmatrix ist symmetrisch (anti-symmetrische Anteile heben sich heraus)
  - Ähnlichkeitsmatrix positiv definit (alle Eigenwerte positiv)
  - ⇒ *quadratic form distance* ist Metrik

$$D(\mathbf{k}, \mathbf{l}) = \sqrt{(\bar{\mathbf{k}} - \bar{\mathbf{l}})^T \mathbf{A} (\bar{\mathbf{k}} - \bar{\mathbf{l}})}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Berücksichtigung der Bin-Ähnlichkeit (2)

---

- **Metrische Eigenschaften der Mahalanobisdistanz**
  - Eigenwertzerlegung einer positiv definiten symmetrischen Matrix
  - $\mathbf{A} = \mathbf{D}^T \mathbf{B} \mathbf{D}$ ,  $\mathbf{B}$  ... positive Diagonalmatrix
  - Diagonalisierung der Distanz ⇒ Euklidische Distanz in linear transformiertem Raum
  - $\tilde{\mathbf{k}} = \sqrt{\mathbf{B}} \mathbf{D} \bar{\mathbf{k}}$ ,  $(\bar{\mathbf{k}} - \bar{\mathbf{l}})^T \mathbf{A} (\bar{\mathbf{k}} - \bar{\mathbf{l}}) = (\tilde{\mathbf{k}} - \tilde{\mathbf{l}})^T (\tilde{\mathbf{k}} - \tilde{\mathbf{l}})$
  - $\mathbf{B} = \sqrt{\mathbf{B}} \sqrt{\mathbf{B}}$
  - diagonal dominante Matrix mit positiven Einträgen ist stets positiv definit, z.B. für kleines  $\sigma$ :

$$\mathbf{A}_{ij} = e^{-\frac{d(\mathbf{k}_i, \mathbf{l}_j)}{\sigma}}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Berücksichtigung der Bin-Ähnlichkeit (3)

- Falls  $A$  nicht positiv definit: sei  $d(P_i, P_j)$  eine  $L_1$ - oder  $L_2$ -Norm, dann
 
$$(\bar{k} - \bar{l})^T H(\bar{k} - \bar{l}) \leq 0, \quad \text{wenn } h_{ij} = d(P_i, P_j)$$
 und  $\sum_i (k_i - l_i) = 0$  (Histogramme gleicher Masse)
- wenn  $A$  geeignet aus  $H$  abgeleitet wird, ist *quadratic form distance* zumindest nicht-negativ (Hafner et al. 95 - keine Aussage zur Dreiecksungleichung), z.B.:

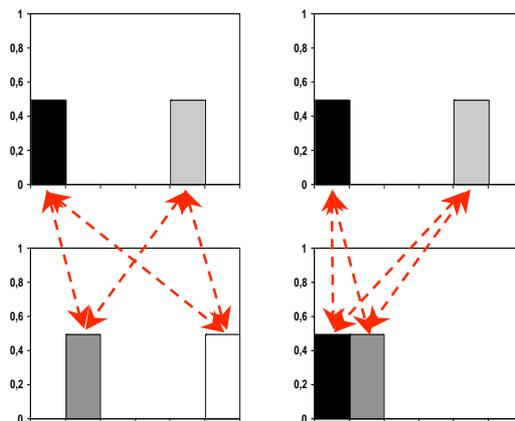
$$A_{ij} = 1 - \frac{d(k_i, l_j)}{\max_{i,j} d(k_i, l_j)}, \quad \text{vermutlich auch } A_{ij} = e^{-\frac{d(k_i, l_j)}{\sigma}}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Berücksichtigung der Bin-Ähnlichkeit (3)

- Mahalanobis-Distanz entspricht der Wahrnehmung besser:**  
linke Histogramme haben geringere Distanz, wie gewünscht



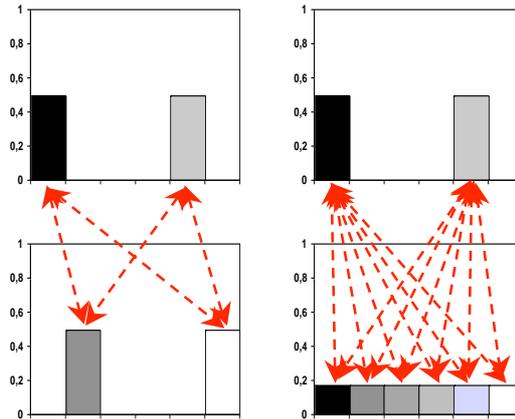
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Berücksichtigung der Bin-Ähnlichkeit (4)

---

- gibt auch Fälle, wo Mahalanobis-Distanz nicht der Wahrnehmung entspricht: die linken Histogramme sind ähnlicher, aber ihre Distanz ist größer als rechts



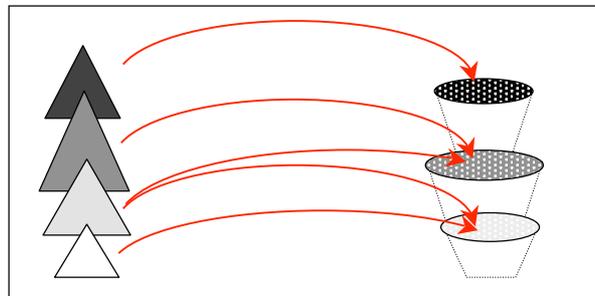
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Vergleich von Histogrammen: Earth Mover Distance (1)

---

- Rubner 99
- Interpretiere Distanz-Berechnung als Transportproblem: Bestimme minimale Kosten, um das eine Histogramm in das andere zu transformieren



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Earth Mover Distance (2)

---

- formale Definition auf Basis der Grunddistanz  $d(k,l)$ :

Voraussetzung:  $\sum_i k_i = \sum_j l_j = 1$

$\text{Work}(k,l,F) = \sum_i \sum_j d(k_i, l_j) f_{ij}$ , wobei

$F = [f_{ij}]$ ,  $f_{ij} \geq 0$  (von  $k_i$  nach  $l_j$  transportierte Masse),

$\sum_j f_{ij} = k_i$ ,  $\sum_i f_{ij} = l_j$ ,  $\sum_i \sum_j f_{ij} = 1$

$\text{EMD}(k,l) = \min_F \text{Work}(k,l,F)$

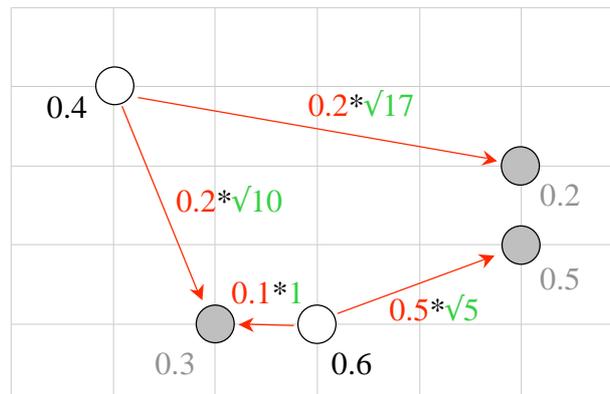
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Earth Mover Distance (3)

---

- Beispiel:  $\text{EMD} = 2.675$



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Earth Mover Distance (4)

### Falls Signaturen $k$ und $l$ unterschiedliche Masse haben

- füge bei der kleineren (o.B.d.A.  $k$ ) ein Hilfs-Bin hinzu

$$\hat{k} = \{k, k_{\text{slack}}\}, \quad \sum_i \hat{k}_i = 1$$

- normiere EMD (andernfalls werden kleine Signaturen bevorzugt):

$$\text{EMD}(k, l) = \frac{\min_F \text{Work}(k, l, F)}{\sum_i k_i}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Earth Mover Distance (5)

### • Metrische Eigenschaften: EMD ist Metrik, wenn die Basisdistanz eine Metrik ist

- Symmetrie, Nicht-Negativität trivial
- Dreiecksungleichung: betrachte 3 Histogramme  $k, l, m$  mit optimalen Verschiebungen  $\underline{kl}, \underline{lm}, \underline{km}$  sowie  $\underline{klm} = \underline{kl} \circ \underline{lm}$

$$\begin{aligned} \text{EMD}(k, m) &= \sum_{i,j} \underline{km}_{ij} d(k_i, m_j) \\ &\leq \sum_{i,r,j} \underline{klm}_{irj} d(k_i, m_j) \\ &\leq \sum_{i,r,j} \underline{klm}_{irj} (d(k_i, l_r) + d(l_r, m_j)) \\ &= \sum_{i,r} \underline{kl}_{ir} d(k_i, l_r) + \sum_{r,j} \underline{lm}_{rj} d(l_r, m_j) \\ &= \text{EMD}(k, l) + \text{EMD}(l, m) \end{aligned}$$

Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Earth Mover Distance (6)

---

- **Berechnung:**

1. Bestimmung einer erlaubten Lösung des Transportproblems (Nord-West Algorithmus, Russel's Methode)
2. Optimierung der Rohlösung mit spezialisiertem Simplexverfahren  
⇒ Komplexität (mit optimalen Algorithmen):  $O(n^3 \log(n))$

	$l_1$	$l_2$	$l_3$	$l_4$	
$k_1$	5	10			0
$k_2$		5	15	5	0
$k_3$				5	0
	0	0	0	5	

Sommersemester 2005

Ulrich Köthe: Bildinformationssysteme

## Earth Mover Distance (7)

---

- **Eigenschaften:**

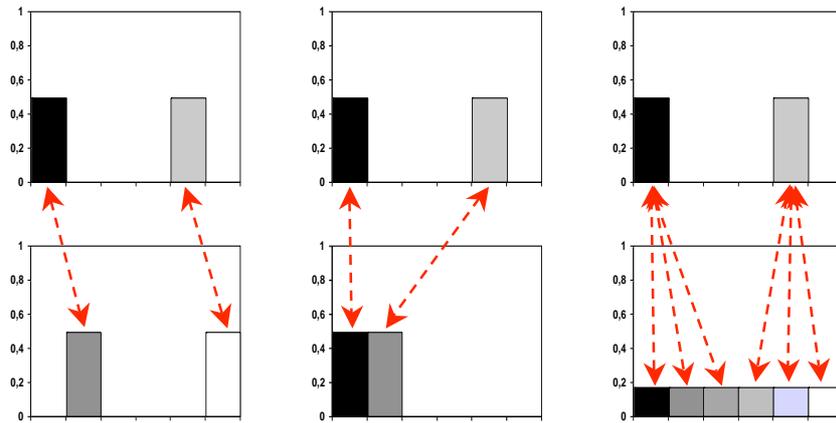
- Anwendbar auf Signaturen und Histogramme
- Matching/Transport gute Approximation für Ähnlichkeitswahrnehmung, löst die Probleme, die bei anderen Distanzen aufgetreten waren
- Unterstützung partieller Matches (Signaturen ungleicher Masse, z.B. bei Verdeckung)
- relativ effiziente Berechnung (da  $n$  relativ klein)
- Metrik

Sommersemester 2005

Ulrich Köthe: Bildinformationssysteme

## Earth Mover Distance (8)

---



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Earth Mover Distance (9)

---

- **Optimierungen – Spezialfälle**

- für 1D Histogramme entspricht die EMD der Matching-Distanz
- untere Schranke:

$$\mathbf{EMD(k,l) \geq d(\bar{k}, \bar{l}), \quad \bar{k}, \bar{l} \dots \text{Schwerpunkte der Signaturen}$$

- robuste EMD, falls  $(1 - \gamma)$  Ausreißer:

$$\sum_{i,j} f_{ij} = \gamma < 1$$

- robuste EMD, falls Transport um maximal  $\tau$  Einheiten möglich:

$$\tau\text{-EMD(k,l)} = \max_F \sum_{i,j} f_{ij} [d(k_i, l_j) \leq \tau]$$

$$\text{mit } [d(k_i, l_j) \leq \tau] = \begin{cases} 1, & \text{falls } d(k_i, l_j) \leq \tau \\ 0, & \text{sonst} \end{cases}$$

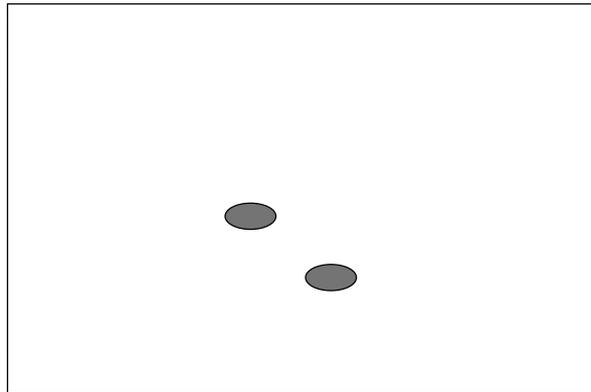
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Psychologische Erforschung der Ähnlichkeitswahrnehmung (1)

---

- Ähnlichkeitswahrnehmung ist abhängig vom Kontext

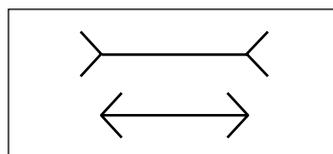


Sommersemester 2005

Ulrich Köthe: Bildinformationssysteme

## Psychologische Erforschung der Ähnlichkeitswahrnehmung (2)

---



Sommersemester 2005

Ulrich Köthe: Bildinformationssysteme

## Psychologische Erforschung der Ähnlichkeitswahrnehmung (3)

---

- Ähnlichkeitswahrnehmung ist abhängig vom Blickwinkel

δημισιονος verwirreub  
απειριτωελ coufnsiug



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Ähnlichkeitsmodell von Thurstone-Shepard

---

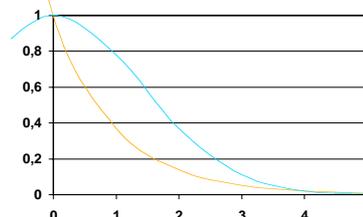
- VP lernen Antwort  $R_i$  für Stimulus  $S_i$ 
  - Messung der Verwechslungswahrscheinlichkeit und Berechnung der „Generalisierungsmatrix“

$$p_{ij} = P[R_i | S_j], \quad g_{ij} = \sqrt{\frac{p_{ij} p_{ji}}{p_{ii} p_{jj}}}$$

- Anpassung des „Generalisierungsmodells“ an die Daten

$$g_{ij} = g(d(S_i, S_j))$$

- Distanz als Minkowski-Metrik
- Sättigung des Abstandes
- exponentielle Generalisierungsfunktion (vgl. Weber-Fechner!)



Sommersemester 2005

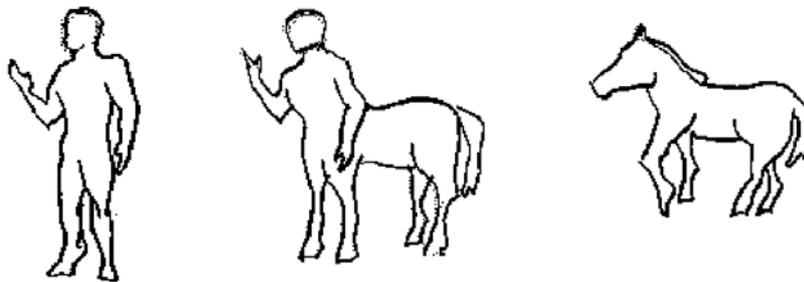
Ullrich Köthe: Bildinformationssysteme

## Tversky's Experimente (1)

---

- **Ähnlichkeitswahrnehmung ist nicht metrisch**

Verletzung der Dreiecksungleichung: Kuba ähnelt der Sowjetunion, Jamaica ähnelt Kuba, aber Jamaica ähnelt der Sowjetunion nicht



Sommersemester 2005

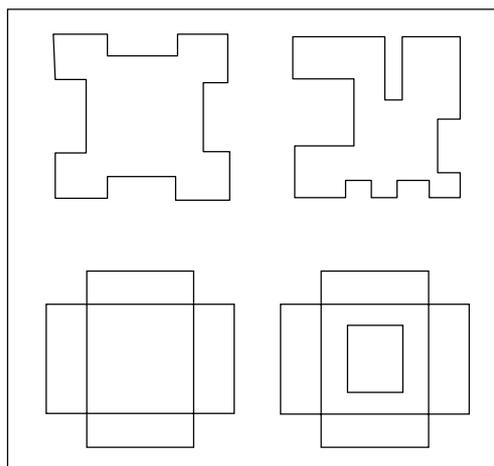
Ullrich Köthe: Bildinformationssysteme

## Tversky's Experimente (2)

---

- **Ähnlichkeitswahrnehmung ist nicht metrisch**

– Verletzung der Symmetrie:  
VP bevorzugen  
„rechtes Muster ähnelt linkem“  
gegenüber „linkes Muster ähnelt rechtem“



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Ähnlichkeitsmodell von Krumhansl

- Ähnlichkeit hängt nicht nur von Distanz der Stimuli ab, sondern auch von der Dichte der Stimuli in deren Nähe:

$$\theta(x, y) = d(x, y) + \alpha h(x) + \beta h(y), \quad h(x) - \text{Dichte}$$

- höhere Dichte = typischere Stimuli
- keine konstante Selbstähnlichkeit: typische Stimuli (hohe Dichte) haben geringe Selbstähnlichkeit (werden leicht verwechselt)

$$\theta(x, x) = (\alpha + \beta) h(x)$$

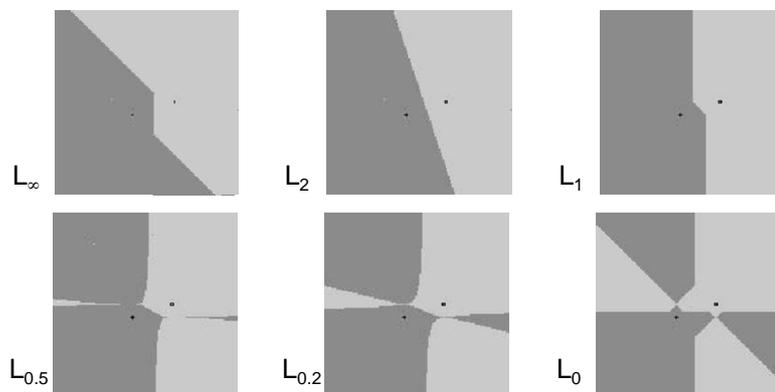
- wenn  $\alpha < \beta$ : untypisches ist typischem ähnlicher als umgekehrt (Verletzung der Symmetrie - vgl. Tversky)
- Dreiecksungleichung gilt

Sommersemester 2005

Ulrich Köthe: Bildinformationssysteme

## Nicht-Metrische Distanzen nach Jacobs et al. (1)

Voronoi-Diagramme für metrische und nicht-metrische Distanzen (Minkowski-Distanzen mit  $p = \infty, 2, 1, 0.5, 0.2, 0$ )



Sommersemester 2005

Ulrich Köthe: Bildinformationssysteme

## Nicht-Metrische Distanzen nach Jacobs et al. (2)

---

- Robuste Vergleiche führen oft auf nicht-metrische Distanzen (robust geg. Ausreißer, Verdeckung)
- Beispiel: Hausdorff-Distanz

$$d(x, y) = \max\left(\max_{i \in x} \min_{j \in y} \|i - j\|, \max_{j \in x} \min_{i \in y} \|j - i\|\right)$$

- robuste Hausdorff-Distanz: Ersetzung der max-Operation durch 60%-Quantil

$$d(x, y) = \max\left(\text{med}_{60} \min_{i \in x} \|i - j\|, \text{med}_{60} \min_{j \in x} \|j - i\|\right)$$

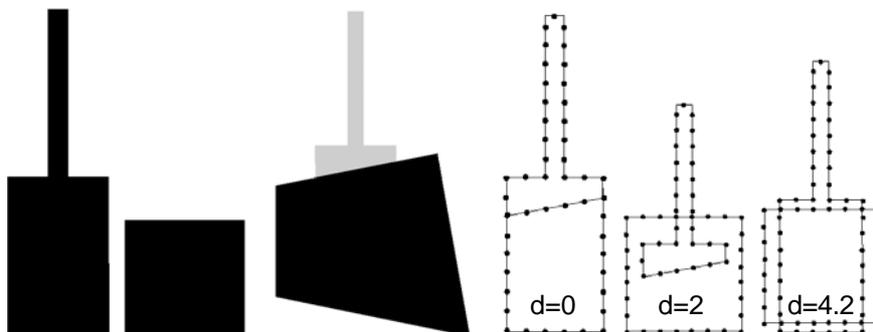
Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Nicht-Metrische Distanzen nach Jacobs et al. (3)

---

- Verwendung zu Erkennung von verdeckten Objekten
- Dreiecksungleichung gilt nicht



Sommersemester 2005

Ullrich Köthe: Bildinformationssysteme

## Zusammenfassung und Bewertung

---

- **existiert ein riesiger Zoo von Distanzen/Ähnlichkeitsfunktionen**
  - unterschiedliche Komplexität der Berechnung
  - geeignete Auswahl für jede Anwendung
  - experimenteller Vergleich: siehe später
- **viele offene Fragen**
  - metrische oder nicht-metrische Distanzen?
  - allgemeine Eigenschaften der nicht-metrischen Distanzen
  - Modellierung des menschlichen Ähnlichkeitsempfindens
  - Kombination verschiedener Distanzen (siehe später)
- **in der Praxis oft Verwendung von heuristischen Distanzen**