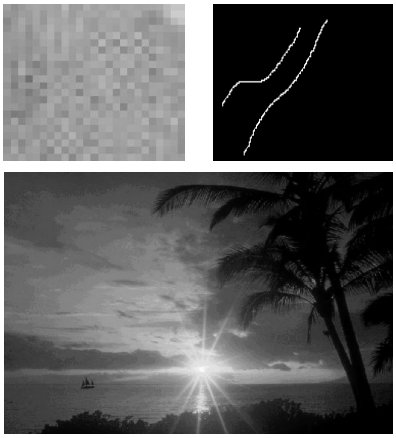
Inhaltsorientierte Such in Bilddatenbanken: Zusammenfassung und Ausblick

Vorlesung Bildverarbeitung 2, Teil 7
Wintersemester 2001/2002
Ullrich Köthe, FB Informatik, Uni Hamburg

Überblick (1)

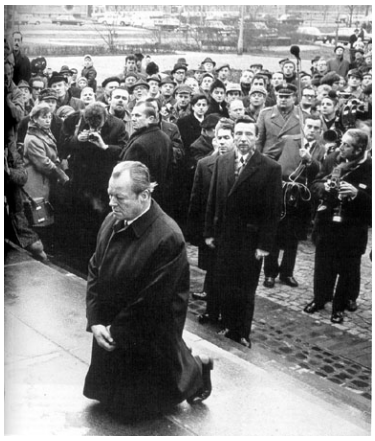
- **Bilder haben viele Interpretationsebenen**



Überblick (2)

- **Visuelle Suche und Schlüsselwörter**
  - High-level Semantik von Schlüsselworten wird von Bildverarbeitung nicht erreicht
  - aber Schlüsselworte sind zu inflexibel für die vielfältigen Interpretationsmöglichkeiten eines Bildes

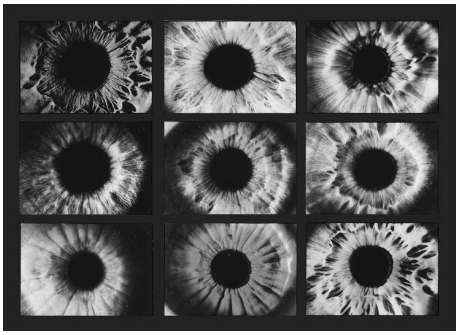
⇨ **Kombination von Schlüsselworten und Bildverarbeitung**

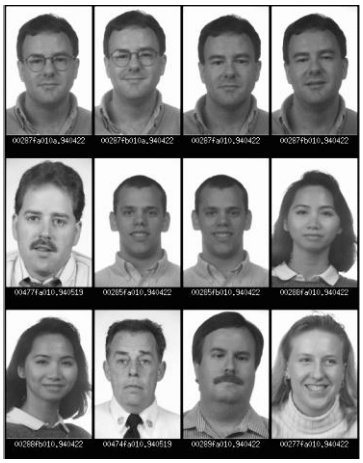


Überblick (3)

Closed world* versus *open world

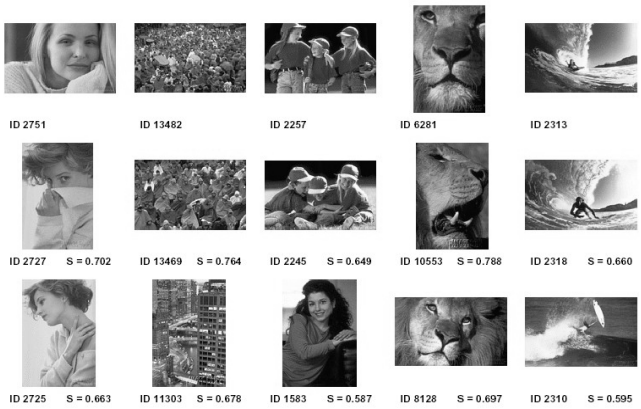
- bei bestimmten Anwendungen gibt es eine eindeutige Semantik und eindeutige Tests der Qualität der Suche





Überblick (4)

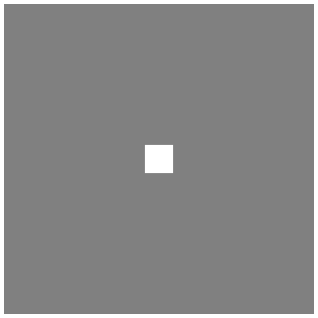
bei open world Anwendungen ist man auf Ähnlich-keitssuche angewiesen, da keine eindeutigen Matches

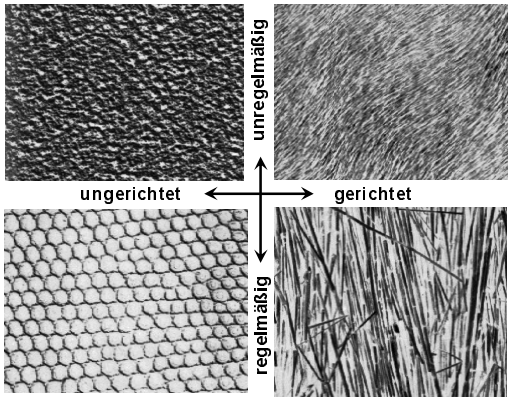


Überblick (5)

- **die Ähnlichkeitsdefinition muß die menschliche Wahrnehmung möglichst genau nachbilden**

L*a*b* Farbraum



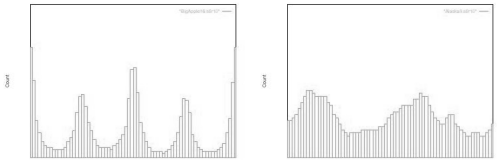


Überblick (5)

globale Bildbeschreibungen für *open world* Probleme

– Histogramme: Farbe, Texturmerkmale, Kantenrichtungen, Regionengrößen, Schlüsselworthäufigkeiten

(a) Richtungshistogramme (b)



	0	2	1	3	
	1	2	3	3	
	2	2	3	0	

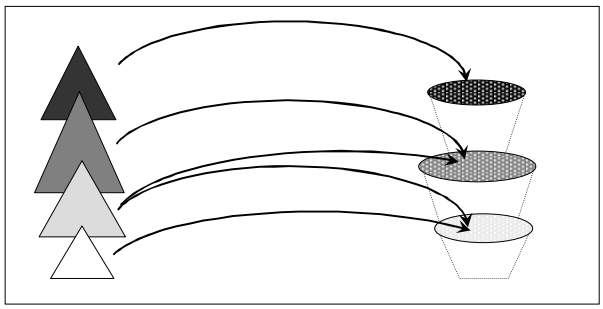
Cooccurence-Matrix

	0	1	2	3	
0	0	0	1	1	
1		0	2	1	
2			1	2	
3				1	

Überblick (6)

- **Vergleich von Histogrammen sollte die Ähnlichkeit der Bins berücksichtigen**
  – *qudratic form distance* und *earth mover distance*

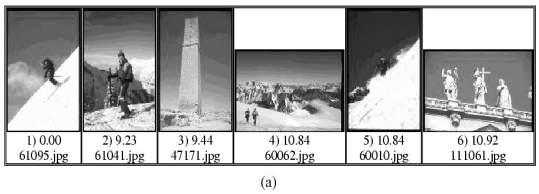
EMD: Histogrammdistanz als Transportproblem ⇨ Quadratische Programmierung



Überblick (7)

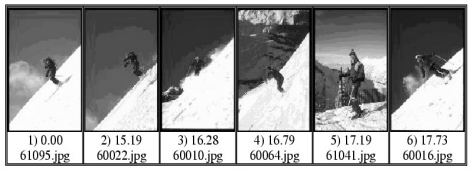
lokalisierte Merkmale:

- kombiniertes Farb- / Ortshistogramm
- Fensterung
- Wavelet-Zerlegung (Haar, Gabor)
- Merkmale für Texturregionen



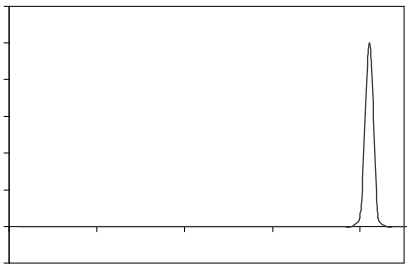
(a)

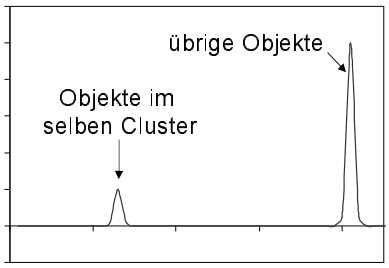
Erfolg begrenzt, da zu wenig Objektbezug bzw. zu instabil



Überblick (8)

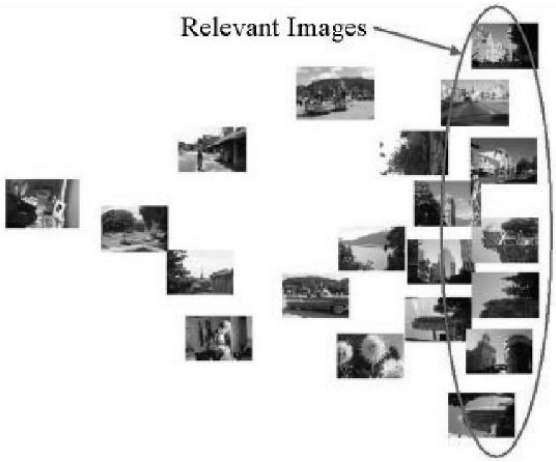
- **Curse of Dimensionality (hochdimensionale Räume):**
  - empty space problem und instabile nächste Nachbarsuche
  - Verhältnis von Standardabweichung zu Erwartungswert des Abstandes zu anderen Punkten konvergiert gegen 0
  - Cluster und irrelevante Dimensionen wirken entgegen





Überblick (9)

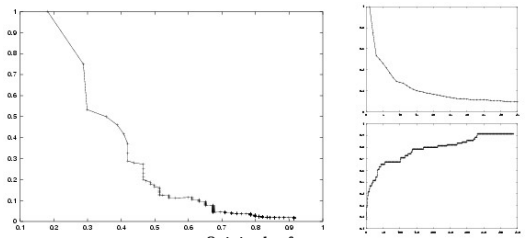
star field display **und Relevanz-Feedback sind effektive Interaktions-techniken**

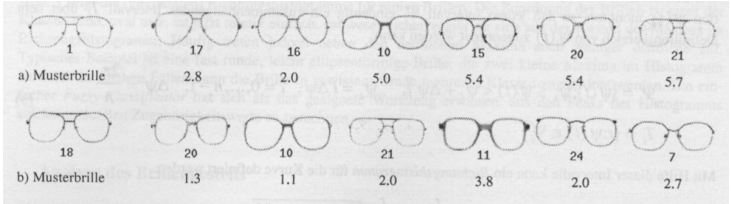


Überblick (10)

Evaluierung:

- Messung von *recall* und *precision*
- fehlen realistische Szenarien
- große Variabilität bei Ähnlichkeitsbeurteilung durch Menschen

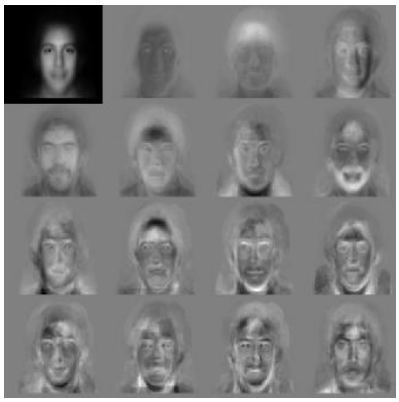




Nützliche Algorithmen (1)

Principle Component Analysis

- Gesichtserkennung („Eigenfaces“)
- Merkmalserkennung, z.B. in Gesichtern („Eigenfeatures“)
- Dimensionsreduktion: Identifikation relevanter Merkmale
- Dimensionsreduktion: Stabilisierung und Beschleunigung der Ähnlichkeitssuche



Projektion der Daten auf die ersten *k* Eigenvektoren der Kovarianzmatrix:

$$C = \frac{1}{M}\sum_{i=1}^{M}\Phi_i\Phi_i^T \qquad C\,\vec{u}_k = \lambda_k\vec{u}_k$$

Nützliche Algorithmen (2)

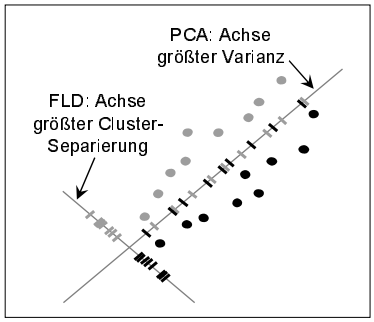
Cluster-erhaltende Dimensionsreduktion

- Fisher‘s Linear Discriminants und Relevant Component Analysis
  - Gesichtserkennung („Fisherfaces“)
  - minimiere *in-class scatter*, maximiere *between-class scatter*

$$S_C = \sum_{i=1}^{C}|m_i|(\vec{\mu}_i - \vec{\mu})(\vec{\mu}_i - \vec{\mu})^T$$

$$S_W = \sum_{i=1}^{C}\sum_{x_k \in m_i}(\vec{x}_k - \vec{\mu}_i)(\vec{x}_k - \vec{\mu}_i)^T$$

$$S_C a_i = \lambda_i S_W a_i$$

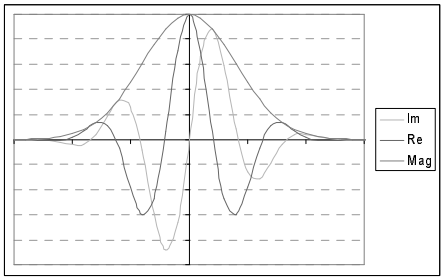


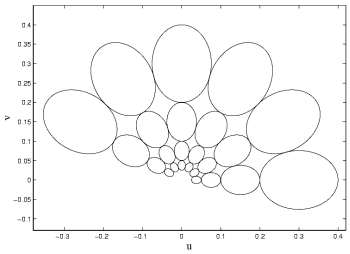
- *Projection Pursuit*: maximiere Clusterungsmaß
- nichtlineare Dimensionsreduktion nicht anwendbar

Nützliche Algorithmen (3)

- **Gaborfilter:**
  - Anwendung einer Filterbank (z.B. 6 Richtungen, 4 Skalen)
  - als Textur- und Kantenmerkmale verwendbar

$$\text{Gabor}_{k,\sigma}(x,y) = \left(\cos(2\pi k\, x) + i\sin(2\pi k\, x)\right) e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)}$$





Nützliche Algorithmen (4)

geometrische Transformationen:

- grauwertbasierte, iterative Bestimmung der Parameter einer affinen Transformation oder Ähnlichkeitstransformation in einer Bildpyramide
- iterative Bestimmung einer Thin Plate Spline Transformation mit Hilfe der „shape context“ Histogramme
  ⇨ Verformungsenergie als Ähnlichkeitsmaß

Ausblick (1)