



# Gesichtsdetektion

Institut für Informatik  
Mustererkennung und Bioinformatik

Angewandte Bildverarbeitung, SS 2007



# Überblick

Detektion  
von  
Gesichtern

Überblick

- 1 Techniken im Überblick
- 2 Principal Component Analysis
- 3 Eigenfaces
- 4 Realtime Detektion mit Viola-Jones



# Techniken zur Gesichtsdetektion

## Detektion von Gesichtern

### Techniken

PCA

Eigenfaces

Viola-Jones

- Modell-basiert:  
 explizite Modellierung der Gesichtsgeometrie
- Eigenfaces:  
 automatische Extraktion von statistisch relevanten  
 Merkmalen von Gesichtern
- Filter-Ansätze:  
 Suche nach spezifischen Merkmalen im Bild  
 (Hautfarbe, ovale Formen, Nasenlöcher, Augen, . . . )

# Eigenfaces - Motivation

Detektion  
 von  
 Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

- 1 Berechnung von *geeigneten* Gesichts-Merkmalen aus Mustervektoren
- 2 Training eines Klassifikators, z.B. NN-Klassifikator
- 3 Zuordnung unbekannter Merkmalsvektoren zu einer Klasse

⇒ offene Frage: Wie findet man geeignete Merkmale zur Klassifikation?

- für Bilder als Mustervektoren bewährt:

**Hauptkomponentenanalyse (PCA)**

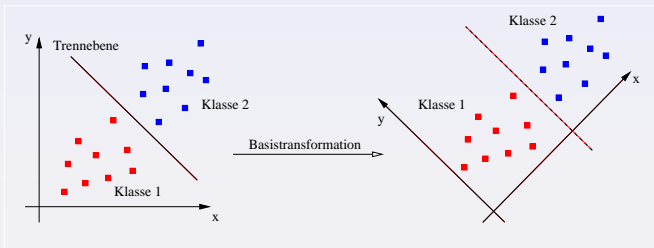


# Principal Component Analysis

”Ermögliche Klassifikation durch Transformation der Mustervektoren in einen niedriger-dimensionalen Unterraum, in dem der Hauptteil der Datenvariation liegt.”

- Annahme dabei:

Variation in den Daten entspricht einem hohen Informationsgehalt!





# Algorithmus - I

- gegeben seien mittelwert-freie Mustervektoren  $\vec{x}^\alpha, \alpha = 1 \dots N, \vec{x}^\alpha \in R^d$ :

$$\frac{1}{N} \sum_{\alpha=1}^N \vec{x}^\alpha = 0$$

- Algorithmus:

- 1 berechne die Autokorrelationsmatrix  $C^{xx}$  der Datenmenge:

$$C_{ij}^{xx} = \frac{1}{N} \sum_{\alpha=1}^N \vec{x}_i^\alpha \vec{x}_j^\alpha$$

- 2 berechne die Eigenwerte  $\lambda_i$  und die -vektoren  $\hat{v}_i$  von  $C^{xx}$ :

$$C^{xx} \cdot \hat{v}_i = \lambda_i \cdot \hat{v}_i \quad (\text{Eigenwertgleichung})$$

⇒ die Eigenvektoren bilden eine Orthonormal-Basis des  $R^d$



# Algorithmus - II

## Detektion von Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

- jeder Datenvektor  $\vec{x}^\alpha$  besitzt die Eigenvektorzerlegung

$$\vec{x}^\alpha = \sum_{i=1}^d t_i^\alpha \cdot \hat{v}_i \quad \iff \quad t_j^\alpha = \vec{x}^\alpha \cdot \hat{v}_j$$

- die  $t_i^\alpha$  sind zentriert und paarweise unkorreliert
- die Eigenwerte  $\lambda_j$  liefern die Varianz in den  $t_j^\alpha$ :

$$\frac{1}{N} \sum_{\alpha=1}^N t_i^\alpha t_j^\alpha = \lambda_j \cdot \delta_{ij} \quad ,$$



# Interpretation

## Detektion von Gesichtern

### Techniken

### PCA

### Eigenfaces

### Viola-Jones

- die Eigenvektorzerlegung beschreibt jeden Vektor  $\vec{x}^\alpha$  durch einen neuen Parametervektor (Merkmalsvektor!)

$$\vec{t}^\alpha = (t_1^\alpha, t_2^\alpha, \dots, t_d^\alpha)^T$$

- $t_j^\alpha$  ergibt sich durch lineare Transformation aus  $\vec{x}^\alpha$ :

$$t_j^\alpha = \vec{x}^\alpha \cdot \hat{v}_j$$

- die Eigenwerte  $\lambda_i$  liefern die Varianzen in den einzelnen  $t_i^\alpha$

⇒ **Dimensionsreduktion** durch Auswahl einer Teilmenge der Basisvektoren bei der Transformation



# Dimensionsreduktion

- die Eigenwerte seien absteigend sortiert:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$$

- Abbruch der Eigenvektorzerlegung nach dem k-ten Term liefert Approximation  $\tilde{\mathbf{x}}^\alpha$  für  $\mathbf{x}^\alpha$ :

$$\tilde{\mathbf{x}}^\alpha = \sum_{j=1}^k t_j^\alpha \cdot \hat{\mathbf{v}}_j$$

- Approximationsfehler  $\delta\tilde{\mathbf{x}}^\alpha$ :

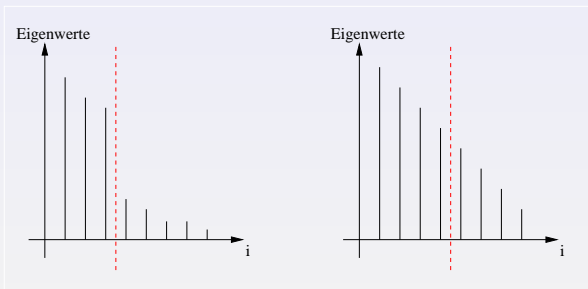
$$\delta\tilde{\mathbf{x}}^\alpha = \mathbf{x}^\alpha - \tilde{\mathbf{x}}^\alpha = \sum_{j=k+1}^d t_j^\alpha \cdot \hat{\mathbf{v}}_j$$

⇒ der mittlere Approximationsfehler ist gleich der Summe unberücksichtigter Eigenwerte!



# Fazit - I

- Mitnahme der  $k$  größten Eigenvektoren führt zu Minimierung des mittleren Approximationsfehlers
- Dimensionsreduktion auch bekannt als **Karhunen-Loeve-Entwicklung**
- Offene Frage: Wie wählt man  $k$  geschickt?!  
 ⇒ anhand der Eigenwertverteilung von  $C^{XX}$





# Fazit - II

## Detektion von Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

- Eigenwertanalyse gibt Aufschluß über *intrinsische Datendimensionalität*
- PCA macht keine Aussage über semantischen Gehalt  
 ⇒ Achtung bei starkem Rauschen in den Daten!!!
- der entstehende (niedrig-dimensionale) Datenraum beschreibt die Mustercharakteristik optimal bei gewählter Dimension  $k$



## Grundidee:

- Repräsentation der gesicht-spezifischen Merkmale von Bildern in einem geeigneten Unterraum
- Klassifikation eines unbekanntem Musters durch Auswertung seiner Projektion in den Unterraum

## Der Klassifikator unterscheidet zwei Modi:

- 1 Systeminitialisierung:  
 Training des Klassifikators auf einer Trainingsmenge
- 2 Arbeitsphase:  
 Klassifikation unbekannter Muster  
 (mit optionalem Update)

# Phase I - Training

Detektion  
 von  
 Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

- gegeben eine Menge von  $M$  Trainingsmustern  
 $\vec{x}_\alpha \in R^{N^2}, \alpha = 1 \dots M$
- fasse  $N \times N$ -dimensionales Bild als  $N^2$ -dimensionalen Vektor auf
- berechne "Facespace" durch Auswahl von  $L$  Eigenvektoren der
- berechne Merkmalsvektoren  $\vec{\omega}_\alpha$  der Trainingsvektoren  $\vec{x}_\alpha$  als Repräsentanten der einzelnen Klassen  $\Omega_j$

Berechnung der Eigenvektoren:

$$C^{xx} = \frac{1}{M} \sum_{\alpha=1}^M \vec{x}_\alpha (\vec{x}_\alpha)^T = \frac{1}{M} A A^T \text{ mit } A = [\vec{x}_1 \dots \vec{x}_M]$$

# Phase I - Training

- Beispiel-Eigenfaces:



Detektion  
 von  
 Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones



## Phase II - Klassifikation

Detektion  
 von  
 Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

Gegeben ein unbekanntes Muster  $\vec{x}$ .

- berechne die Eigenvektorzerlegung des Eingabemusters (Projektion in den "Facespace")

$$t_k = \hat{v}_k \cdot (\vec{x} - \bar{x}) \quad \text{mit} \quad \bar{x} = \frac{1}{M} \sum_{\alpha=1}^M \vec{x}_\alpha$$

- klassifiziere über Distanzen zu den Repräsentanten:

$$\epsilon_i = \|\vec{t} - \vec{\omega}_i\|^2 < \theta_1 \quad \text{mit} \quad \vec{\omega}_i \text{ Repräsentant der Klasse } \Omega_i$$

- Problem: auch Nicht-Gesichter werden unter Umständen auf Merkmalsvektoren nahe den Repräsentanten abgebildet!

Rückweisungskriterium:

$$\psi = \|\vec{x} - \tilde{x}\|^2 > \theta_2 \quad \text{mit} \quad \tilde{x} = \sum_{i=1}^L t_i^{\vec{x}} \hat{v}_i$$

Damit resultieren drei Fälle, die zu unterscheiden sind:

- a)  $\psi > \theta_2$ : Rückweisung
- b)  $\psi < \theta_2, \epsilon > \theta_1$ : unbekanntes Gesicht
- c)  $\psi < \theta_2, \epsilon < \theta_1$ : bekanntes Gesicht

⇒ Hauptproblem: Aufwand!





# Grundidee: Merkmalsdetektion

## Detektion von Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

- Detektion mit "einfachen" Merkmalen, die Gesichter von Nicht-Gesichtern unterscheiden können
  - ⇒ Merkmale oftmals flexibler als eine explizite Wissensmodellierung
  - ⇒ Merkmale schneller als pixelweise Detektion
- Hauptziele:
  - schnell in Real-Time, z.B. 15 Bilder pro Sekunde
  - robust
    - ⇒ wenig *false positives*, aber hohe Erkennungsraten

# Viola-Jones-Detektor



## Detektion von Gesichtern

Techniken

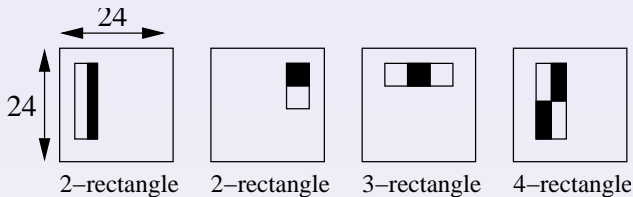
PCA

Eigenfaces

Viola-Jones

- allgemein zur schnellen Detektion von Objekten
- entwickelt von Paul Viola & Michael Jones (2001)
- drei Grundprinzipien:
  - ① *Integral Images* für eine schnelle Merkmalsberechnung
  - ② automatisches Lernen aussagekräftiger Merkmale
  - ③ Kaskade von Klassifikatoren (insgesamt 32)
    - ⇒ Bild wird schrittweise auf interessante Regionen begrenzt (Regions of Interest - ROIs)

- "schnell heisst, Merkmale dürfen nicht zu komplex sein"
- Viola-Jones: *rectangle features*  
 (motiviert durch Haar-Basisfunktionen und Wavelets)



- "Faltung" des Bildes mit den Merkmalen:  
 Summe der Pixel in den schwarzen Regionen minus  
 Summe der Pixel in weißen Regionen
- Größe des Operatorfensters: 24x24

# Integral Images

Detektion  
 von  
 Gesichtern

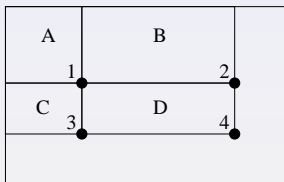
Techniken  
 PCA  
 Eigenfaces  
 Viola-Jones

- erlauben eine schnelle Berechnung der Merkmale  
 (... und können auch selbst schnell berechnet werden)

- Definition:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

- Merkmale erfordern Berechnung von Differenzen  
 zwischen Rechtecken  
 ⇒ konstanter Aufwand für jedes Rechteck



$$\begin{aligned}
 v(2) &= \sum A + \sum B \\
 v(3) &= \sum A + \sum C \\
 v(4) &= \sum A + \sum B + \sum C + \sum D \\
 \Rightarrow \sum D &= v(4) + v(1) - (v(2) + v(3))
 \end{aligned}$$

# Lernen von Merkmalen



## Detektion von Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

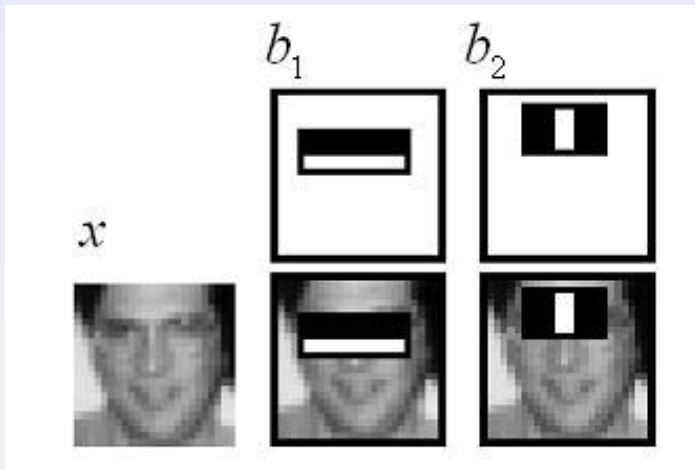
- es gibt insgesamt 45.396 (!) Merkmale
- Anwendung zumeist noch in verschiedenen Skalierungen  
 ⇒ Aufwand noch viel zu hoch!
- Ansatz:  
 Verwendung von wenigen, aber aussagekräftigen Merkmalen
- automatische Auswahl durch *AdaBoost*

- Kombination von einfachen ("dummen") Klassifikatoren zu einem guten
- Training durch Anwendung auf Sequenz von Klassifikationsproblemen:
  - ① Suche nach bestem Klassifikator aus einer gegebenen Menge durch Training & Test (Schwellwert-Klassifikation)
  - ② Neugewichtung der Trainingsbeispiele  
 ⇒ falsch klassifizierte im folgenden "wichtiger"
  - ③ Suche nach nächstem Klassifikator
  - ④ ...
- Klassifikationsfehler geht exponentiell gegen Null
- Viola-Jones: jedes Feature ist ein Klassifikator

# Lernergebnisse

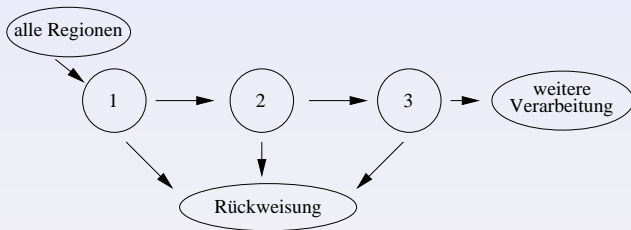
Detektion  
 von  
 Gesichtern

- erster und zweiter Klassifikator:



# Klassifikationskaskade

- in einem Bild sind die meisten Teilregionen uninteressant  
 ⇒ einfache Merkmale am Anfang,  
 komplexe nur für wenige Regionen



- finales System: 32 Schichten mit 4297 Merkmalen





# Abschluss-Bemerkungen: Praktische Aspekte

## Detektion von Gesichtern

Techniken

PCA

Eigenfaces

Viola-Jones

- Trainingsmenge:  
 mehrere 1.000 Gesichter und Nicht-Gesichter  
 ⇒ Trainingszeit ursprünglich mehrere Wochen (!)  
 ⇒ OpenCV enthält vortrainierte Kaskade
- Normierung der Varianz in den Bildern  
 ⇒ Kompensation von veränderter Beleuchtung
- explizite Behandlung von Mehrfach-Detektionen  
 (z.B. Positionsmittelung)