Closed World Anwendungen von CBR

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

Closed World Anwendungen

- **Anfragen haben eine eindeutige Lösung**
- ***Ground truth* leicht zu definieren**
- **Ähnlichkeit mit klassischer Objekterkennung, aber wesentlich größere Datenbestände**
- **Beispiele:**
  - Biometrie (Identifizierung, Authentizierung)
    Erkennung von Gesicht, Iris, Retina, Hand, Fingerabdrücken
  - einfache Bildklassen
    Bildmarken (Logos) und andere Symbole

Gesichtserkennung: Die *eigenface* Methode (1)

- **Grundidee: identifiziere ein Gesicht durch nächste Nachbar-Suche in der Datenbank, aber:**
  - zu langsam
  - empfindlich gegen zufällige Aspekte (z.B. Beleuchtung)
- **Lösung: Projektion der Gesichter in einen niedrig-dimensionalen Raum**
  - optimale Basis: Hauptkomponentenanalyse
  - schnell (10 - 100 Dimensionen entsprechend Invarianzeigensch.
  - robust: geeignete Auswahl der Eigenvektoren

  ⇨ Eigenvektoren = *eigenfaces* (Turk and Pentland 1991)

Gesichtserkennung: Die *eigenface* Methode (2)

Vorverarbeitung:

- **Transformation der Bilder auf Einheitsgröße $N * N$**
- **Bestimmung des „durchschnittlichen Gesichts“ und des „Differenzgesichts“**

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i \qquad \Phi_i = \Gamma_i - \Psi$$

- **Kovarianzmatrix ($N^2 * N^2$), Eigenvektoren:**

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

Gesichtserkennung: Die *eigenface* Methode (3)

- **aber: wenn $M < N^2$, nur M Eigenwerte $\neq 0$**

⇨ **Optimierung:**

$$\frac{1}{M}\sum_{i=1}^{M}\Phi_i\Phi_i^T =: AA^T, \quad A = \left[\Phi_1 \quad \Phi_2 \quad \ldots \quad \Phi_M\right]$$

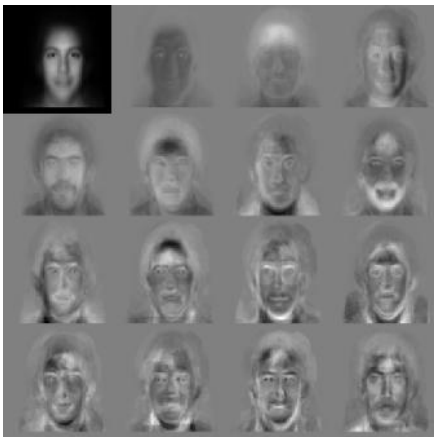
$$A^T A\,\vec{v}_k = \mu_k \vec{v}_k$$

$$AA^T A\,\vec{v}_k = \mu_k A\vec{v}_k, \quad \Rightarrow \quad A\vec{v}_k \text{ sind EV von } C$$

$$\vec{u}_l = \sum_{i=1}^{M} \vec{v}_{lk}\Phi_k$$

Gesichtserkennung: Die *eigenface* Methode (4)

- **typische Eigenfaces:**
  - Lernmenge: 128 Gesichter
  - *face space*: 20 Eigenvektoren, Bandpaßcharakteristik
  - erste EV weggelassen, da sie die allgemeine Experimentiersituation codieren (Beleuchtung, Vordergrund/Hintergrund etc.) - Tiefpaßcharakter



Gesichtserkennung: Die *eigenface* Methode (5)

- **Suche: Projektion auf die *eigenfaces* und Bestimmung des nächsten Nachbarn:**

$$\omega_l = \vec{u}_l^T (\Gamma - \Psi), \quad \Omega = [\omega_1 \quad \ldots \quad \omega_{M'}] \quad \text{Projektion}$$

$$\Phi_r = \sum_{l=1}^{M'} \omega_l \vec{u}_l \quad \text{Rekonstruktion}$$

$$\varepsilon_r = \|\Phi - \Phi_r\| \quad \text{Rekonstruktionsfehler}$$

$$\varepsilon_k = \|\Omega - \Omega_k\| \quad \text{Differenz zu anderen Gesichtern}$$

- **Bewertung: ε_k und ε_r klein: gefunden**
  ε_k groß, ε_r klein: unbekanntes Gesicht
  ε_r groß: kein Gesicht

Gesichtserkennung: Die *eigenface* Methode (6)

- **Resultate:**
  - Datenbank mit 7562 Bildern, ca. 3000 Personen
  - Erkennungsrate: 95%
    - Fehler: bester *match* ist andere Person
    - abhängig von Rasse (Schwarze: 95%, Weiße: 90%, Asiaten: 80%) ⇨ bei Asiaten merkmalsbasierte Erkennung erforderlich
  - Zugangskontrolle: 1.5% *false rejections*, 0.01% *false acceptances*
- **Weiterentwicklungen:**
  - Robustheit gegenüber Blickrichtung, Beleuchtung
  - Fisher's Discriminant Analysis: *eigenfaces,* die Ähnlichkeit in jeder Klasse maximieren, zwischen den Klassen minimieren
  - Bayes-Klassifikator: Analyse der Wahrscheinlichkeitsverteilungen

Gesichtserkennung: Die *antiface* Methode (1)

- ***eigenface*: Erkennung bei *hoher* Übereinstimmung**
- ***antiface*: Erkennung bei *niedriger* Übereinstimmung**
- **Motivation:** typische Bilder sind glatt: mittlere Inhomogenität (z.B. Gradientenbetrag) ist exponentialverteilt:

$$S(I) = \sum_{u,v} (u^2 + v^2)\,\mathrm{DCT[I](u,v)}, \qquad \text{mittlere Inhomogenität}$$

$$\Pr(I) \propto \exp(-S(I)), \qquad \text{Wahrscheinlichkeit}$$

⇨ Erwartungswert des Skalarprodukts zweier zufälliger, glatter, normierter Bilder ist groß

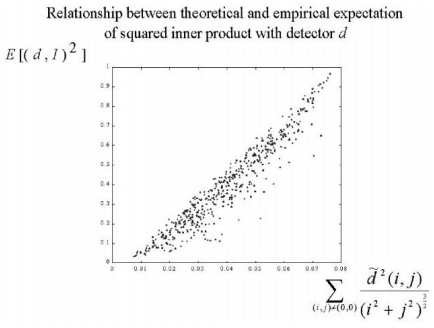
$$\mathrm{E}\left[\langle J, I\rangle^2\right] = \int \langle J, I\rangle^2 \Pr(I)\, dI \propto \sum_{u,v} \frac{\mathrm{DCT}[J](u,v)}{\left(u^2+v^2\right)^{3/2}} >> 0$$

Gesichtserkennung: Die *antiface* Methode (2)

- **experimentelle Überprüfung: Vergleich von 5000 Zufallsbildern gemäß Pr() mit 15000 natürlichen Bildern (20x20)**

  linearer Zusammenhang zwischen Erwartungswert des Skalarprodukts und Inhomogenität gemäß Frequenzanalyse

  Keren et al. 2001



Gesichtserkennung: Die *antiface* Methode (3)

- **Folgerung: wähle Basisvektoren so, daß**
  - Skalarprodukt mit den gewünschten Zielstrukturen klein

$$\forall i : \max_{t \in T} \left| \langle d_i, t \rangle \right| \rightarrow \text{Minimum}$$

  - Skalarprodukt mit zufälliger Struktur groß (glatte Basisvektoren)

$$\mathrm{S}(d_i) = \sum_{u,v} \left(u^2 + v^2\right) \mathrm{DCT}[d_i](u,v) \rightarrow \text{Minimum}$$

  - sie unabhängig sind, d.h. die Wahrscheinlichkeit kleiner Skalarprodukte mit einem Bild faktorisiert für alle Basisvektoren:

$$p\left(\langle d_1, I \rangle < \varepsilon, \langle d_2, I \rangle < \varepsilon\right) = p\left(\langle d_1, I \rangle < \varepsilon\right) p\left(\langle d_2, I \rangle < \varepsilon\right) \quad \Leftrightarrow$$

$$0 = \int \langle I, d_i \rangle \langle I, d_j \rangle \Pr(I)\, dI = \sum_{u,v} \frac{\mathrm{DCT}[d_i](u,v)\,\mathrm{DCT}[d_j](u,v)}{\left(u^2 + v^2\right)^{3/2}}$$

Gesichtserkennung: Die *antiface* Methode (4)

- **Bestimmung der Basisvektoren:**
  - wähle Minimalwert für Skalarprodukt mit Zielstrukturen, z.B. $M = 10^{-5}$
  - Bestimme ersten Basisvektor: minimiere (Nelder-Mead Verfahren)

$$\max_{t \in T} \left| \langle d_1, t \rangle \right| + \lambda S(d_1)$$

    und wähle λ so daß $\max_{t \in T} \left| \langle d_1, t \rangle \right| = M$
  - Bestimme weitere Basisvektoren ebenso, mit der zusätzlichen Bedingung:

$$\sum_{u,v} \frac{\mathrm{DCT}[d_i](u,v)\,\mathrm{DCT}[d_j](u,v)}{\left(u^2 + v^2\right)^{3/2}} = 0$$

  - schnellere Approximation: minimiere $\sum_{t \in T} \langle d_i, t \rangle^2 + \lambda S(d_i)$ (lineares Gleichungssystem)

Gesichtserkennung: Die *antiface* Methode (5)

Suchalgorithmus:

1. Falte Bild mit d_1
2. für *i*=1...*K*:
   wenn im Punkt (x,y) gilt: $\forall k < i : \langle d_k, I \rangle < \varepsilon$
   berechne Skalarprodukt des Fensters um (x,y) mit d_i
   sonst: Abbruch

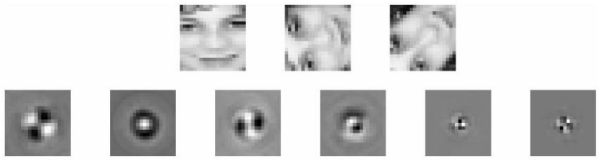
⇨ **Fenster, wo alle Skalarprodukte klein sind, enthalten die gesuchte Struktur**

Ergebnisse:

- deutlich geringere Zahl an Basisvektoren als *eigenface*-Methode
- Robustheit gegen Rotation: *ef.*: 12 Vektoren, *af.*: 3 Vektoren
- Robustheit gegen affine Transformationen: *ef.*: 145, *af.*: 4

Gesichtserkennung: Die *antiface* Methode (6)

- **Beispiel: Vektoren f. rotationsinvariante Gesichtserk.**



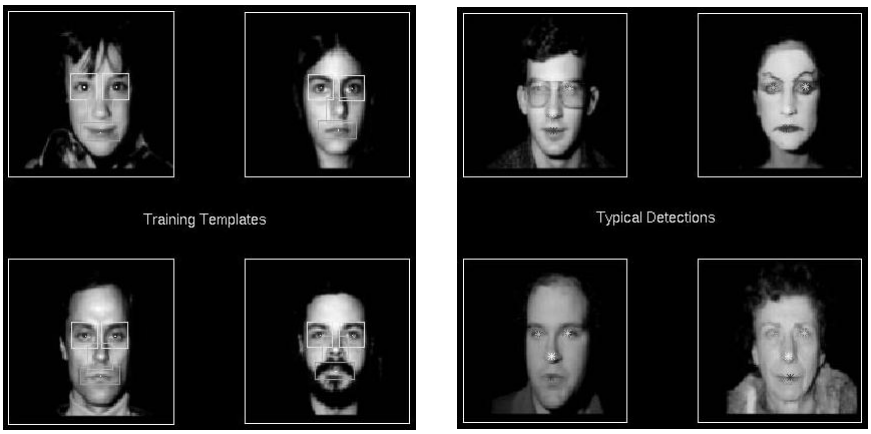
- **Suchbeispiele: (Auto)**

Gesichtserkennung: Die merkmalsbasierte Methode (1)

- **Höhere Robustheit:**
  - zufällige Veränderungen: Blickwinkel, Beleuchtung, Gesichtsausdruck, teilweise Verdeckung
  - systematische Veränderungen: Haare, Bart, Brille, Alter
- **Gesichtsdetektion**
- **Merkmalsgewinnung: Signatur salienter Punkte**
  - *eigenface* oder *antiface*-Methoden angewendet auf Merkmale (Augen, Nase, Mund, Mundwinkel)
  - Anwendung einer Filterbank (z.B. Gabor)
- **geometrische und photometrische Normierung**
- **Beschreibung der Gesichter durch Signatur der Merkmale und deren relativer Lage**

Gesichtserkennung: Die merkmalsbasierte Methode (2)

Training und Erkennung mit lokalen *eigenfaces* und Bayes-Klassifikator

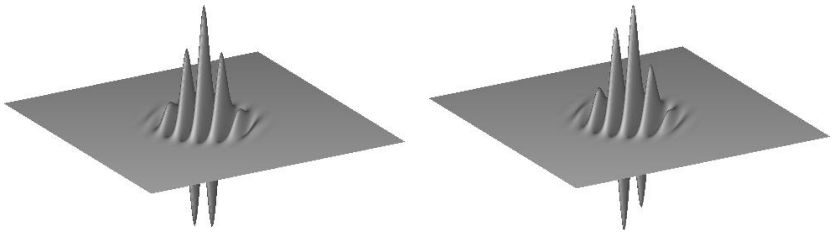


Gesichtserkennung: Die merkmalsbasierte Methode (3)

- **Erkennung mit Gabor-Jets an ausgewählten Punkten**

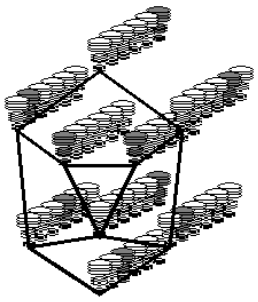
$$\mathrm{Gabor}_{\vec{k},\sigma}(\vec{x}) = \frac{k^2}{\sigma^2} e^{-\frac{k^2 x^2}{2\sigma^2}} \left(e^{i\vec{k}\vec{x}} - e^{-\frac{\sigma^2}{2}} \right), \quad \left\| \mathrm{Gabor}_{\vec{k},\sigma}(\vec{x}) \right\|^2 \sim k^2$$

Realteil: **Imaginärteil**

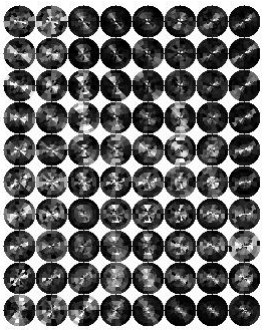


Gesichtserkennung: Die merkmalsbasierte Methode (4)

- **Bunch Graph: Gabor-Jets an ausgewählten Punkten (manuelles Markieren in ca. 70 Gesichtern)**



face bunch graph



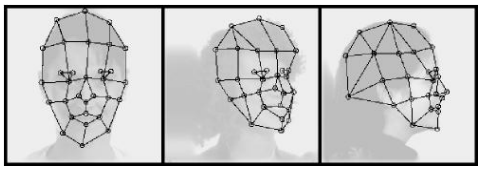
Gesichtserkennung: Die merkmalsbasierte Methode (5)

- **Elastic Bunch Graph Matching:**
  - Vergleich vom Gabor-Jets (in Polarkoordinaten):

$$S_m(I,I') = \frac{\sum_i a_i a_i'}{\sqrt{\sum_i a_i^2 \sum_i a'^2_i}} \qquad S_\varphi(I,J) = \frac{\sum_i a_i a_i' \cos(\varphi_i - \varphi_i' - \vec{d}\vec{k})}{\sqrt{\sum_i a_i^2 \sum_i a'^2_i}}$$

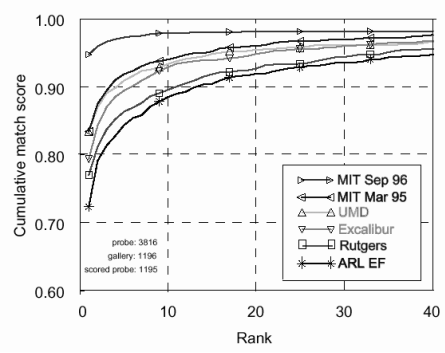
  - günstigster Graph (grober Blickwinkel), günstigster Jet in jedem Knoten, Verschiebung der Knoten (feiner Blickwinkel)

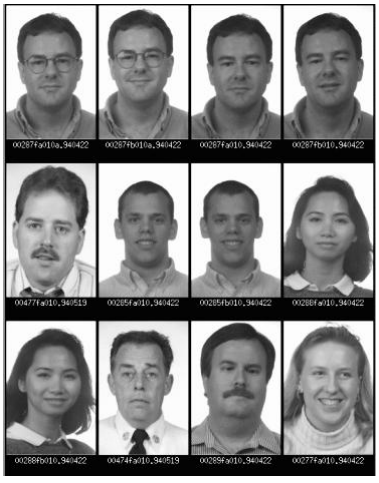
$$S(I,G) = \sum_n \max_j \left(S(J_n^I, J_n^j)\right) + \lambda \sum_n \text{contraction}_n$$



Gesichtserkennung: Bewertung

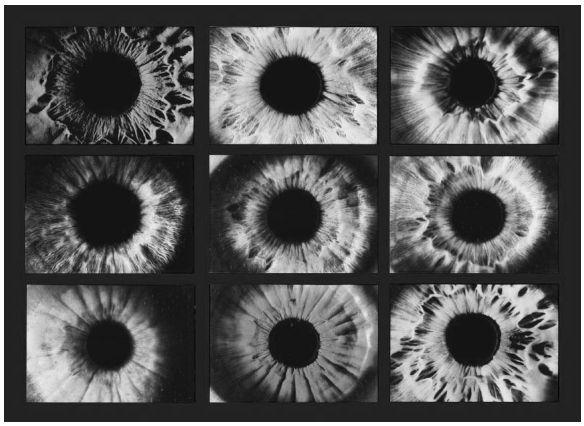
- **FERET-Datenbank der US-Army (7500 Bilder)**





Iriserkennung (1)

- **hohe Individualität der Iris ⇨ Eignung für Biometrie**
  (Vgl. mit anderen Verfahren später)
- Iris der Augen einer Person oder von Zwillingen unterschieden sich wie bei Fremden
- kaum Änderungen im Lauf des Lebens
- Fälschungssicher



Iriserkennung (2)

- **Beispiel Korrelationsverfahren (Wildes at al. 94):**
  - Multiskalenzerlegung (Laplace-Pyramide)
  - bestmögliche starre Registrierung mit jedem Beispielbild (Spezialhardware)
  - Ähnlichkeitsfunktion auf jedem Band (Median von normierten 8x8 Korrelationen)
  - Fishers Diskriminanten-Analyse zur Klassifikation
  - Validierung: 10 Berechtigte, 40 Unberechtigte (je 10 Bilder) sowie je 2 Bilder des „falschen“ Auges der berechtigten

  ⇨ 100% korrekt bei 520 Bildern

Iriserkennung (3)
Starre Registrierung

- Annahme: *J* entspricht *I* bis auf Translation, Rotation, Skalierung:

$$I(\vec{x}) = J(\vec{x} - \vec{u}(\vec{x}))$$

- Fehlermaß:

$$E(\vec{u}) = \sum_{\vec{x}} \left(I(\vec{x}) - J(\vec{x} - \vec{u}(\vec{x}))\right)^2$$

- Linearisierung: (Taylorreihe)

$$E(\vec{u}_i + \delta\vec{u}_i) = \sum_{\vec{x}} \left(I(\vec{x}) - J(\vec{x} - \vec{u}_i(\vec{x})) + (\nabla I)^T \delta\vec{u}_i\right)^2$$

- Starres Bewegungsmodell: (mit gleichmäßiger Skalierung)

$$\vec{u}(\vec{x}) = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} a_3 & a_4 \\ -a_4 & a_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\vec{u}(\vec{x}) = \mathbf{X}\vec{a}, \quad \vec{a} = (a_1, a_2, a_3, a_4)^T$$

$$\mathbf{X} = \begin{pmatrix} 1 & 0 & x & y \\ 0 & 1 & y & -x \end{pmatrix}$$

Iriserkennung (4)
Starre Registrierung

- Einsetzen:

$$\vec{u}_i + \delta\vec{u}_i = \vec{u}_i + \mathbf{X}\delta\vec{a}_i$$

$$E(\vec{u}_i + \delta\vec{u}_i) = \sum_{\vec{x}} \left(I(\vec{x}) - J(\vec{x} - \vec{u}_i(\vec{x})) + (\nabla I)^T \mathbf{X}\delta\vec{a}_i\right)^2$$

- Fehlerminimierung durch iteratives Gaußsches Fehlerquadratverfahren: ableiten und 0 setzen:

$$\left[\sum_{\vec{x}} \mathbf{X}^T (\nabla I)(\nabla I)^T \mathbf{X}\right] \delta\vec{a}_i = -\sum_{\vec{x}} \mathbf{X}^T (\nabla I)\left(I(\vec{x}) - J(\vec{x} - \vec{u}_i(\vec{x}))\right)$$

- nach $\delta\vec{a}_i$ auflösen (lineares Gleichungssystem), 4 Iterationen auf jeder Pyramidenstufe

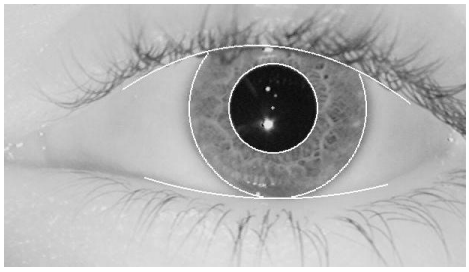
Iriserkennung (5)
Verfahren von Daugman

- **Daugman, PAMI 11/93, patentiert 1994**
  - höchste Zuverlässigkeit (alle anderen Verfahren verdrängt)
  - kommerzielle Systeme von zahlreichen Firmen (u.a. Iridian Technologies, Panasonic, OKI, IrisAccess, Diebold ATMs)
    - Zugangskontrolle, Passkomtrolle
    - Informationssicherheit (Iris login, e-commerce)
    - Ticketless Air Travel
    - Geldautomaten ohne Karte und PIN
- **Algorithmus:**
  - Lokalisierung der Iris (Infrarotbilder, 700-900 nm)
  - Anpassung eines normierten Polarkoordinatensystems
  - Berechnung des IrisCodes mit Hilfe von polaren Gabor-Filtern
  - Matching: Minimale Hamming-Distanz (7 Rotationen)

Iriserkennung (6)
Verfahren von Daugman

- **Lokalisierung der Iris:**
  Suche Kreise mit maximaler Kantenstärke entlang der Kreislinie (coarse-to-fine)

$$\max_{r,x_0,y_0}\left|G_\sigma(r) * \frac{\partial}{\partial r}\oint_{S_r(x_0,y_0)}\frac{I(x,y)}{2\pi r}ds\right|$$



- **Diskretisierung:** Kreisintegral als Summe über diskrete Winkel, Ableitung: Faltung mit finiten Differenzen der Gaußfunktion

$$\max_{n\Delta r,x_0,y_0}\left|\frac{1}{\Delta r}\sum_k\left\{\begin{array}{l}\left(G_\sigma((n-k)\Delta r)-G_\sigma((n-k-1)\Delta r)\right)\times \\ \qquad\sum_m I(k\Delta r\cos(m\Delta\theta)+x_0, k\Delta r\sin(m\Delta\theta)+y_0)\end{array}\right\}\right|$$

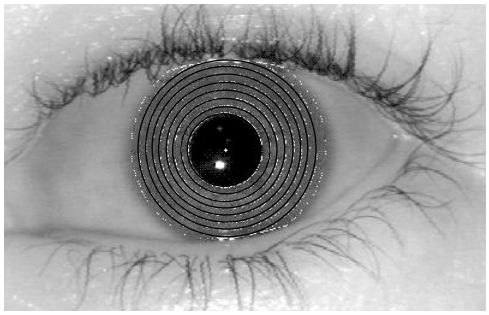
Iriserkennung (7)
Verfahren von Daugman

- **Koordinaten der beiden Kreislinien:**

$$\text{innen}: \quad x_i(\theta) = r_i \cos(\theta) + x_{ci}, \quad y_i(\theta) = r_i \sin(\theta) + y_{ci}$$
$$\text{außen}: \quad x_o(\theta) = r_o \cos(\theta) + x_{co}, \quad y_o(\theta) = r_o \sin(\theta) + y_{co}$$

- **normierte Quasi-Polarkoordinaten:**
  (Kreise nicht konzentrisch, bis 15% Abweichung!)

$$\hat{r} \in [0,1] \quad \text{normierter Radius}$$
$$x(\hat{r},\theta) = (1-\hat{r})x_i(\theta) + \hat{r}\,x_o(\theta)$$
$$y(\hat{r},\theta) = (1-\hat{r})y_i(\theta) + \hat{r}\,y_o(\theta)$$

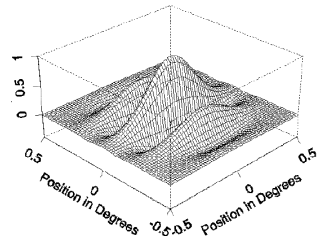


Iriserkennung (8)
Verfahren von Daugman

- **Gaborfilter in Polarkoordinaten**

$$\text{Gabor}(\hat{r},\theta;r_0,\theta_0,\alpha,\beta) = e^{-i\omega(\theta-\theta_0)} e^{-(r-r_0)^2/\alpha^2} e^{-(\theta-\theta_0)/\beta^2}$$

- **komplexwertige Filterantwort codiert lokale Energie und Phase eines orientierten Frequenzkanales**



Iriserkennung (9)
Verfahren von Daugman

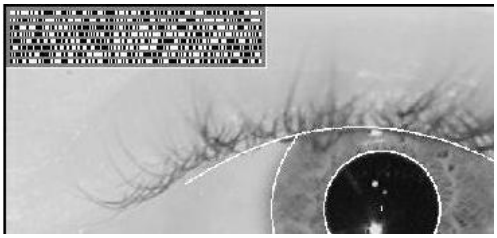
- **IrisCode: Vorzeichen der Filterantworten für verschiedene Orte, Skalen und Frequenzen:**

$$h_{\mathrm{Re}}(P_k) = \mathrm{sig}\left\{\mathrm{Re}\left(\mathrm{Gabor}(\hat{r},\theta;P) * I(x(\hat{r},\theta), y(\hat{r},\theta))\right)\Big|_{P=P_k}\right\}$$

$$h_{\mathrm{Im}}(P_k) = \mathrm{sig}\left\{\mathrm{Im}\left(\mathrm{Gabor}(\hat{r},\theta;P) * I(x(\hat{r},\theta), y(\hat{r},\theta))\right)\Big|_{P=P_k}\right\}$$

$$\mathrm{sig}(\lambda) = \begin{cases} 1 & \text{wenn } \lambda \geq 0 \\ 0 & \text{sonst} \end{cases}$$

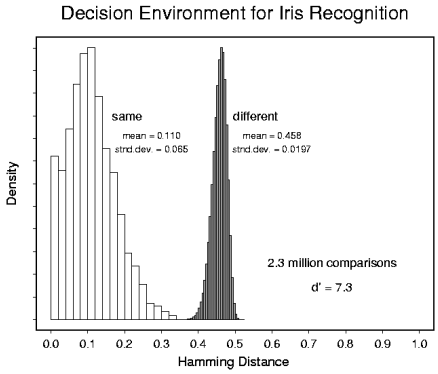
$$P_k = \hat{r}_{0k}, \theta_{0k}, \alpha_k, \beta_k, \omega_k$$



- k = 1...2048 ⇨ 512 byte pro Iris

Iriserkennung (10)
Verfahren von Daugman

- **Matching: Hamming-Distanz (Anteil ungleicher Bits)**
- wenn reiner Zufall:
  - 0 bei gleichen IrisCodes
  - 0.5 bei zufälligen
- 2.3 Mio. Vergleiche (mit Rotation): annähernd Binomialverteilung
- ungleiche Iriden:
  - 244 Freiheitsgrade
  - (p = 0.45, σ = 0.02)
- gleiche Iriden:
  - 40 Freiheitsgrade
  - (p = 0.11, σ = 0.07)

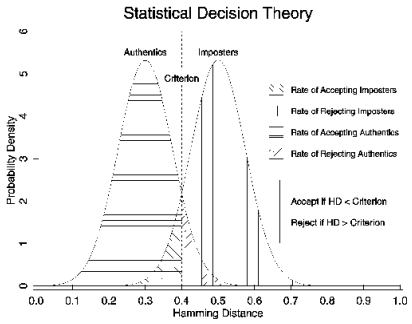


Iriserkennung (11)
Verfahren von Daugman

- Binomialverteilung mit *N* Freiheitsgraden: Erwartungswert der Anzahl „1“ in binärer Zufallsfolge, wenn Wahrscheinlichkeit einer „1“ = p

$$f(x) = \frac{N!}{m!(N-m)!} p^m (1-p)^{N-m}, \quad \sigma = \sqrt{Np(1-p)}$$

- minimaler Fehler:
  maximum likelihood decision
  = Schwellwert bei Überschneidung
  ⇨ 0.34: p(error) ~ 10^{-5}
- konservativerer Wert wenn höhere Kosten bei *false admission*
  ⇨ 0.25: p(*false admission*) ~ 10^{-14}
  p(*false rejection*) ~ 10^{-2}
- **Outlier ?**



Iriserkennung (12)
Verfahren von Daugman

- **Vergleich mit anderen Verfahren: deutlich höhere Erkennungssicherheit**

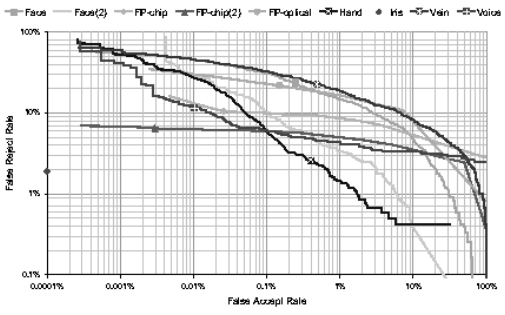


Figure 5. Detection error trade-off: FAR vs FRR

Überprüfung vs. Erkennung

- **Wahrscheinlichkeit p_1 für korrekte Überprüfung der Identität**
- **Wahrscheinlichkeit p_N für korrekte Erkennung aus *N* Personen**

$$p_N = p_1^N$$

- **z.B. p_1 = 99.9%** ⇨ **p_N = 82%** (***N*** **= 200)**
  p_N = 14% (***N*** **= 2000)**
  p_N = 0.005% (***N*** **= 10000)**
- **z.B. $p_1 = (1\text{-}10^{-5})\%$** ⇨ **$p_N$ = 99.8%** (***N*** **= 200)**
  p_N = 98% (***N*** **= 2000)**
  p_N = 90% (***N*** **= 10000)**
- **mit $(1\text{-}p_N) \sim N(1\text{-}p_1)$** ⇨ *identifier* muß für gleiche *false accept* Rate *N*-mal besser sein als *verifier*

Kombination mehrerer Verfahren

- **Benutze 2 unabhängige Biometrieverfahren *i*=1,2**
  - Fehlerraten p_i(*false admission*), p_i(*false rejection*)
  - Regel: akzeptiere Person, wenn beide Verfahren akzeptieren

$p_{combined}(FA) = p_1(FA)\, p_2(FA)$

$p_{combined}(FR) = 1 - (1\text{-}p_1(FR))(1\text{-}p_2(FR))$
$= p_1(FR) + p_2(FR) - p_1(FR)\, p_2(FR)$

Beispiel: $p_1(FA) = p_2(FA) = p_1(FR) = p_2(FR) = 0.001$
⇨ $p_{combined}(FA) = 0.000001$, $p_{combined}(FR) = 0.001999$
$p_2(FA) = p_2(FR) = 0.01$
⇨ $p_{combined}(FA) = 0.00001$, $p_{combined}(FR) = 0.01099$
$p_2(FA) = 0.2$, $p_2(FR) = 0.001$
⇨ $p_{combined}(FA) = 0.0002$, $p_{combined}(FR) = 0.001999$

Zusammenfassung: *Closed World* Verfahren

- spezielle statistische Test auf speziellen Beschreibungen
- hohe Genauigkeit möglich und beweisbar
- Biometrie: Bequemlichkeit vs. Genauigkeit
- Erkennung hat wesentlich höhere Anforderungen als Verifikation
- Kombination von Verfahren hat nur begrenzten Nutzen, oftmals Auswahl des optimalen Verfahrens besser
- ***trademarks***: später (da meist Ähnlichkeitssuche)