
Hauptseminar

Thema 20: **Vergleich von Gesichtserkennungsverfahren**

Vergleich von zwei erweiterten LDA -basierten
Gesichtserkennungsverfahren zur Lösung des Small-Sample-Size-
Problems

Technische Universität Ilmenau
Studiengang: Medientechnologie

Sommersemester 2005

Bearbeiter: Peter Hamm
Matrikelnummer: 30945

Betreuer: Dipl.-Ing. Peter Dunker
Betreuender Professor: Prof. Dr.-Ing. Karlheinz Brandenburg

INHALTSVERZEICHNIS

1. Einleitung.....	3
2. Gesichtserkennung.....	3
2.1 Überblick aktueller Gesichtserkennungsverfahren.....	4
3. Gesichtererkennung mittels Hauptachsentransformation.....	5
3.2 Eigenfaces-Methode.....	7
3.2.1 PCA.....	7
3.2.2 Eigenfaces - Algorithmus.....	9
3.3 Fisherfaces-Methode.....	10
3.3.1 LDA.....	11
3.3.2 Fisherfaces - Algorithmus.....	12
3.4. Klassifizierung.....	12
3.5. Small Sample Size Problem.....	13
4. Erweiterte Methoden.....	14
4.1 A Maximum Uncertainty LDA.....	14
4.2. Direct Fractional LDA.....	15
5. Ergebnisse.....	16
6. Zusammenfassung.....	19
7. Literaturverzeichnis.....	21
8. Abbildungsverzeichnis.....	212

1. Einleitung

Die folgende Hausarbeit ist im Rahmen des Hauptseminars „Digitale Medientechnik“ im Sommersemester 2005 an der TU Ilmenau entstanden und befasst sich mit dem Gebiet der Gesichtserkennungsverfahren, die auf statistischen Methoden der Hauptachsentransformation beruhen.

Ziel dieser Transformation ist es, die Effektivität der Gesichtserkennung gegenüber anderen Verfahren insbesondere bei Helligkeitsschwankungen und Positionsänderungen des Gesichts im Bild zu erhöhen. Die grundsätzlichen Methoden werden in der Literatur als Eigenfaces- und Fisherfaces-Methoden bezeichnet und arbeiten mit der so genannten *Principle Component Analysis* (PCA) bzw. der *Linear Discriminant Analysis* (LDA).

Hauptgegenstand dieser Hausarbeit ist die Erläuterung zweier fortgeschrittenen, auf der LDA basierenden Verfahren, die das so genannte *Small-Sample-Size-Problem* lösen und die Eigenfaces- bzw. Fisherfaces-Methoden in der Performanz übertreffen.

Der ausführlichen Erläuterung dieser zwei Verfahren steht eine kurze Einführung in das Gebiet der Gesichtserkennung voran.

Damit erhält der Leser einen groben Einblick; auch um die detaillierten Ausführungen zu den oben genannten Hauptachsentransformations-Verfahren entsprechend einordnen zu können.

2. Gesichtserkennung

Die automatisierte Gesichtserkennung ist ein Gebiet der Bildverarbeitung, die in der Praxis große Verbreitung findet. Zumeist handelt es sich dabei um Systeme, die versuchen, anhand der Aufnahme eines Gesichtes eine Person zu identifizieren, indem das Bild mit einem vorhandenem Datenpool abgeglichen wird.

Automatisierte Gesichtserkennung, gleich welcher Methode sie sich bedient, kann durch Prozessmodule im Diagramm 2.1 beschrieben werden.

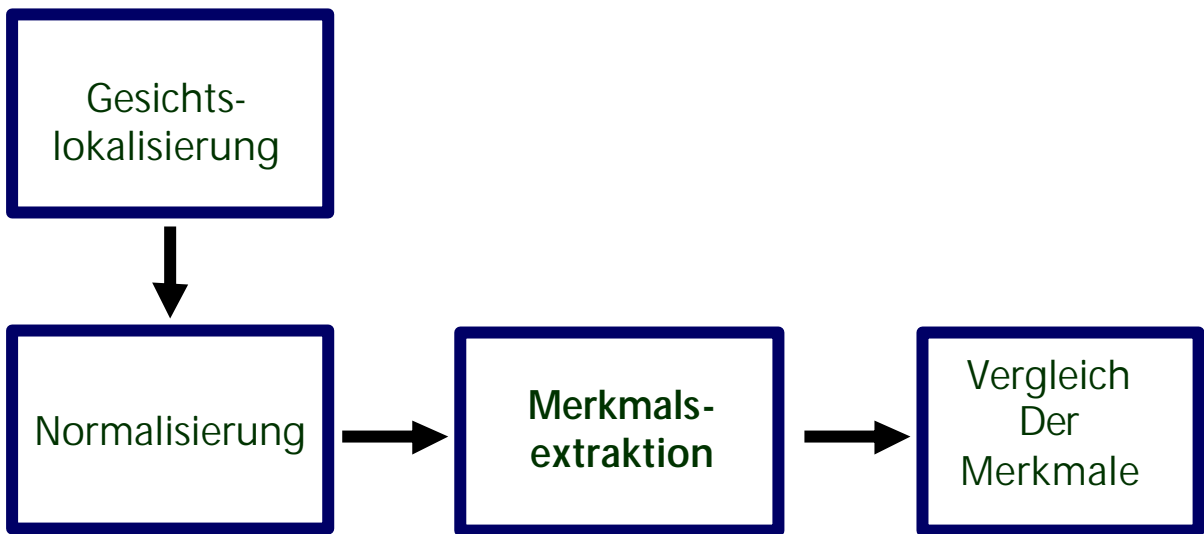


Diagramm 2.1

Es muss zu allererst eine Lokalisierung des Gesichts vorgenommen werden, wenn man davon ausgeht, dass sich neben dem Gesicht auch, für die Gesichtserkennung unwichtige Hintergrundinformationen im Bild befinden. Ist das Gesicht detektiert, wird das Bild auf die entsprechende Größe zugeschnitten. Dieser Schritt ist, wie auch ein eventueller Ausgleich der Helligkeit oder das Rotieren des Gesichtes, dem Normalisierungsprozess zuzuordnen.

Die Frage ist nun, welche Merkmale des Gesichtes verwendet werden (Merkmalsextraktion), anhand derer die Klassifikation (Vergleich der Merkmale) bzw. die Identifikation der Person durchgeführt wird.

2.1 Überblick aktueller Gesichtserkennungsverfahren

Bei dem **Active Shape Model** (Abb. 2.1a) Verfahren sind die Gesichtskonturen die entscheidenden Merkmale zum Vergleich der Gesichter. Konturen werden durch sog. Konturpunkte repräsentiert, die miteinander zu einem Polygonenzug verbunden sind.



Dabei wird dem Gesicht iterativ ein Konturmodell angepasst.

Abb. 2.1a

Ein anderer Ansatz vergleicht Gesichtsausschnitte mit besonderen biometrischen Merkmalen. Die **Template Matching** Methode (Abb. 2.1b) benutzt hierbei die Bereiche der Augenpartie, der Nase und des Mundes.



Abb. 2.1b

Ein Vergleich erfolgt über die Suche nach Ähnlichkeiten zwischen diesen sog. Templates. Da der Vergleich auf aufgrund der Pixelintensität vollzogen wird, ist das Verfahren besonders anfällig für Helligkeits- und Kontraständerungen.

Weiterhin existieren Verfahren, welche die **Gesichtsproportionen** als Merkmale zugrunde legen. Größe und Abstand der Augen als auch die Größe des Mundes und der Nase sind unveränderliche Attribute des menschlichen Gesichts und damit geeignet, als Vergleichsmerkmale herangezogen zu werden.

Bei den genannten Verfahren sind immer nur bestimmte Teile der Gesichtes von Interesse für die Extraktion der Merkmale gewesen.

Einen völlig anderen Weg beschreiten die Verfahren, die Gegenstand dieser Arbeit sind. Sie nehmen eine Analyse des gesamten Gesichtes vor, wobei die Vergleichsmerkmale über die sog. Hauptachsentransformation ermittelt werden.

Es wird sich weiterhin zeigen, dass diese Verfahren eine Reihe von Vorteilen gegenüber den anderen Methoden besitzen.

3. Gesichtserkennung mittels Hauptachsentransformation

Ein Bild wird zu allererst durch die Intensitätswerte der Pixel, bei Schwarz-Weiß-Aufnahmen durch deren Grauwerte, definiert.

Die Gesichtserkennungsverfahren, die auf der Hauptachsentransformation beruhen, benutzen diese Werte der Bildpixel nicht direkt, um eine Klassifikation der Gesichter zu ermöglichen. Vielmehr projizieren sie ein Bild mit Hilfe einer Transformationsmatrix in einen sog. Unterraum. Die neu berechneten Werte, die nicht mehr direkt einem bestimmten Pixel im Bild zugeordnet werden können, sind nunmehr die eigentlichen Merkmale, die zum Vergleich herangezogen werden.

3.1 Einführung – Bildtransformation

Welche Vorteile für die Klassifikation von Gesichtern ergeben sich nun aber aus einer Projektion eines Bildes in einen Unterraum?

Man stelle sich ein Bild mit den Maßen von jeweils 256 Pixel für Höhe und Breite vor. Dieses Bild setzt sich somit aus 65.536 Pixel zusammen und - würde als Vektor repräsentiert - folglich einen Punkt in einem 65,536-dimensionalen Raum darstellen. Der Aufwand für die Berechnung zur Klassifikation, insbesondere bei Korrelationsbetrachtungen, wäre äußerst hoch. Eine Reduktion der Bilddimension würde den Rechenaufwand auf ein erträgliches Maß minimieren.

Ein zweites generelles Problem wird in den Abbildungen 3.1a und 3.1b verdeutlicht.

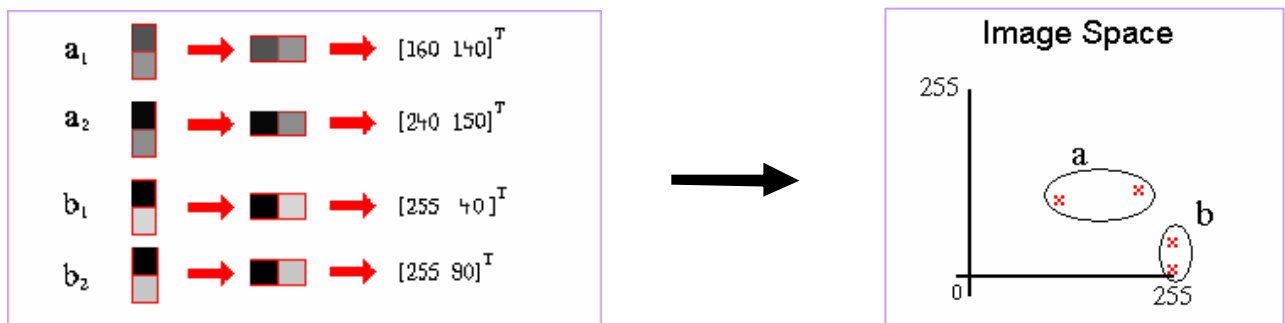


Abb. 3.1a

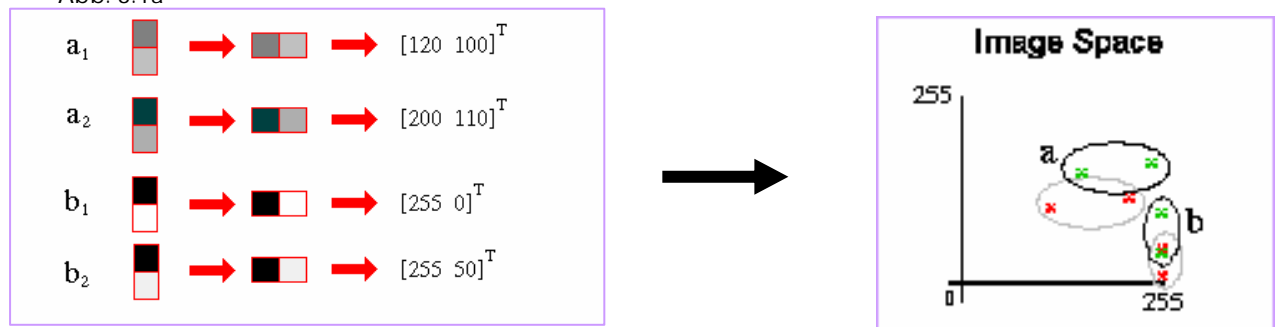


Abb. 3.1b

Sie zeigen die Auswirkungen von Helligkeitsschwankungen auf die Position von Pixel im sog. Bildraum (*Image Space*).

Der Anschaulichkeit halber bestehen die „Bilder“ aus lediglich zwei Pixel, so dass sie in einem zweidimensionalen Raum abgebildet werden können. In Abbildung 3.1a sind zwei Klassen (zwei Personen) mit jeweils zwei Bildern (*Samples*) als Bildvektoren und den korrespondierenden numerischen Grauwerten dargestellt. Die Samples einer Klasse haben eine

geringe Varianz gegenüber der Zwischenklassenvarianz. Die Samples sind schließlich im *Image Space* Diagramm, dessen Achsen die Grauwerte bedeuten, eingetragen. Beide Klassen lassen sich noch sehr gut von einander trennen.

Abbildung 3.1b zeigt die gleichen Grauwertbilder nach einer Helligkeitsänderung um 40 Grauwertstufen. Man sieht, dass sich die Klassencluster durch die Helligkeitsänderungen im Koordinatensystem verschieben und sich beinahe überschneiden. Eine Trennung bzw. Klassifizierung ist nun nicht mehr ohne weiteres möglich.

Beide angesprochenen Probleme - die Dimensionsgröße der Samples und die erschwerte Klassifikation bei Helligkeitsschwankungen - können durch eine geschickte Transformation des Bilddatensatzes auf ein hohes Maß minimiert werden:

$$y_k = W^T x_k \quad [f. 3.1]$$

Das Bild der Gesichtsklasse k wird durch die Transformationsmatrix W in den Bildunterraum y projiziert.

3.2 Eigenfaces-Methode

Ein grundsätzliches Gesichtserkennungsverfahren, das mit einer linearen Transformation arbeitet, ist die so genannte Eigenfaces-Methode. Sie wurde von Pentland und Turk [1] 1991 beschrieben, geht allerdings auf die Bildtransformations-Untersuchungen von Sirovich und Kirby [2] aus dem Jahre 1987 zurück.

Die Eigenfaces-Methode bedient sich der *Principle Component Analysis*, kurz *PCA*, um eine geeignete Transformation des Bilddatensatzes in einen Unterraum zu erreichen.

3.2.1 PCA

Die PCA, oft auch als *Karhunen-Loeve-Transformation* bezeichnet, extrahiert die relevanten Information eines Gesichtes über die Suche nach der Variation in einem gesammelten

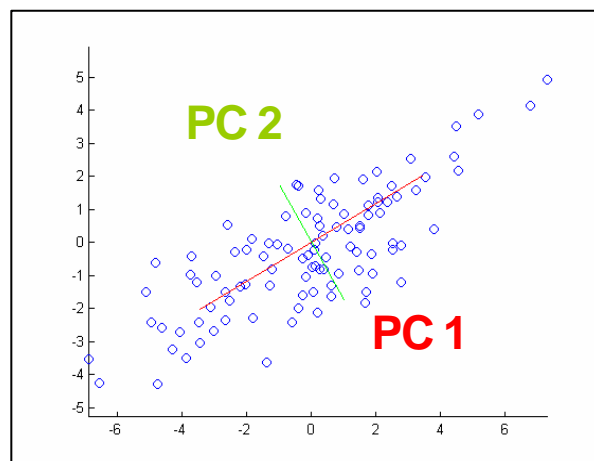


Abbildung 3.2.1.1a

Bilddatensatz. Die Transformation lässt sich auch als eine Drehung der Hauptachsen verstehen, so dass die neuen Koordinatenachsen der Richtungen der größten Varianz entsprechen. Diese Richtungsvektoren stellen die gesuchten *Principle Components* dar.

Mathematisch betrachtet ergeben sich die *Principle Components* (PCs) aus der Berechnung der Eigenvektoren der Kovarianzmatrix des Bilddatensatzes.

Nicht jeder Eigenvektor trägt jedoch gleich zur Repräsentation der Bilder bei. Die Eigenvektoren mit den größten korrespondierenden Eigenwerten beinhalten auch die größte Varianz.

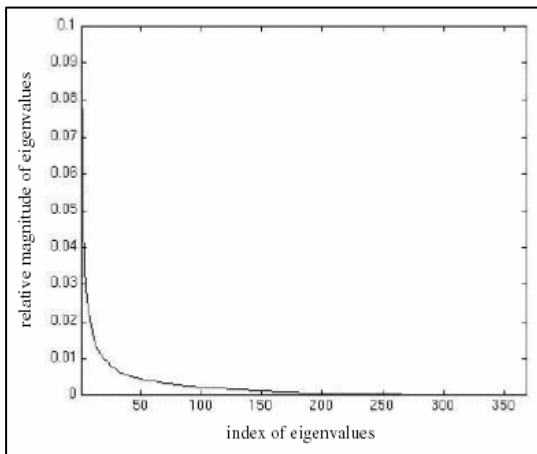


Abbildung 3.2.1.1b

Abbildung 3.2.1.1b verdeutlicht die Abnahme des Informationsgehaltes („*relative magnitude of eigenvalues*“) bezüglich des Eigenwertindex. Eigenvektoren, die mit kleinen Eigenwerten zusammenhängen, liefern keinen praktischen Anteil mehr für die Beschreibung der Bilder, so dass sie nicht mit in die Berechnung der Transformationsmatrix einfließen. Wieviele Eigenvektoren letztendlich einbezogen werden, hängt in der Praxis stark von Bilddatenbank ab

(Praxisnahe Werte liegen zwischen 10 und 160).

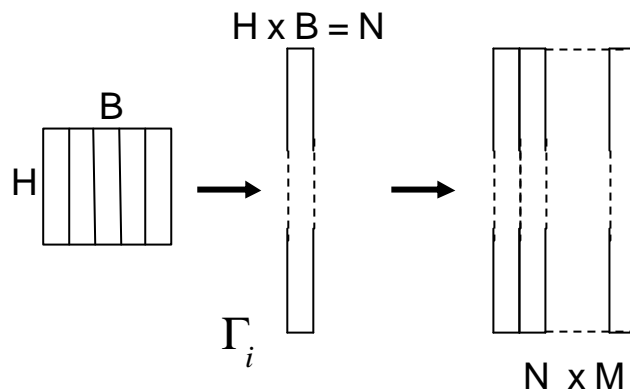
Da die Eigenvektoren dieselbe Dimension wie die Bildvektoren besitzen, lassen sie sich grafisch darstellen. Man erhält die sog. *Eigenfaces* (Abb. 3.2.1c), die der hier besprochenen Methode auch ihren Namen verliehen.



Abb. 3.2.1c

3.2.2 Eigenfaces - Algorithmus

Bei der Berechnung der Eigenfaces bzw. der Transformationsmatrix wird von einer bekannten Bilddatenbank ausgegangen. Die darin enthaltenen Bilder der Gesichter liegen in vektorieller Form vor.



Das heißt, dass ein Bild mit der Höhe H und der Breite B durch den Vektor der Länge $N = H \times B$ repräsentiert wird.

$$\Gamma = [\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M] \quad [\text{f. 3.2.2a}]$$

Γ_1 ist der erste in der Gesamtbildmatrix Γ befindliche Bildvektor mit der Länge N , wobei insgesamt M Bildvektoren existieren.

Für die Berechnung der Kovarianzmatrix dieser Bildmatrix wird das Durchschnittsgesicht (*Image Mean*) gefunden,

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad [\text{f. 3.2.2b}]$$

das zur Abweichung jedes Gesichts vom Durchschnittsgesicht führt:

$$\Phi_i = \Gamma_i - \Psi \quad [\text{f. 3.2.2c}]$$

Die Kovarianzmatrix ergibt sich nun zu:

$$C = \Phi \Phi^T \quad [\text{f. 3.2.2d}]$$

Die jeweiligen Eigenvektoren und die korrespondierenden Eigenwerte erhält man aus der Lösung des Eigensystems:

$$C E_i = \lambda_i E_i \quad [\text{f. 3.2.2e}]$$

Ein Rechenproblem bedeutet jedoch die hohe Dimension der Kovarianzmatrix C von $N \times N$.

Weil sich aus der Kovarianzmatrix C jedoch nur M sinnvolle Eigenvektoren bilden lassen, ist die Berechnung der umgestellten Kovarianzmatrix

$$L = \Phi^T \Phi \quad [\text{f. 3.2.2f}]$$

die nur noch die Dimension $M \times M$ besitzt, möglich.

Es kann mathematisch bewiesen werden, dass die Eigenvektoren von L den Eigenvektoren von C entsprechen.

Das neu zu lösende Eigensystem:

$$L E_i = \lambda_i E_i \quad [\text{f. 3.2.2g}]$$

führt schließlich zu der aus Gleichung [f. 3.1] bekannten Transformationsmatrix:

$$W_{PCA} = \Phi E \quad [\text{f. 3.2.2h}]$$

Die Wahl der Anzahl der verwendeten Eigenvektoren bestimmt die Zahl der Werte (*Features*) im Unterraum, die ein Bild repräsentiert.

3.3 Fisherfaces-Methode

Die sog. Fisherfaces-Methode von Belhumeur, Hespanha und Kriegmann [3] stellt das zweite grundlegende Hauptachsentransformations-Verfahren dar, indem es die *Linear Discriminant Analysis (LDA)** auf den Bereich der Gesichtserkennung überträgt.

Die LDA führt zur Berechnung der Transformationsmatrix mit der Erweiterung, bestehende Klasseninformation (d.h. welche Gesichter zu einer Person gehören) zu verwenden, um nicht nur eine Dimensionsreduktion durchzuführen, sondern auch die Diskriminierung zwischen den Personenklassen zu erhöhen.

Die LDA erzielt im Vergleich zur PCA eine effektivere Trennung der Klassen durch Minimierung der Intra-Klassen-Streuung bei gleichzeitiger Maximierung der Inter-Klassen-Streuung. Mit anderen Worten: Die Samples einer Klasse (d.h. die Bilder einer Person) werden näher zusammengedrückt, während die Klassen selbst von einander wegrücken.

* Die LDA wird oft auch als FLD (Fisher's Linear Discriminant) bezeichnet, da sie auf Klassifikationsverfahren von R. A. Fisher (1936) zurückgehen.

3.3.1 LDA

Die Linear Discriminant Analysis berechnet die optimale Projektion W_{LDA} des Datensatzes aus der Maximierung es sog. *Fisher's Criterion*.

$$W_{LDA} = \underset{W}{\operatorname{arg\,max}} \frac{|W^T S_{BETWEEN} W|}{|W^T S_{WITHIN} W|} \quad [\text{f. 3.3.1a}]$$

Das Fisher-Kriterium ist maximal, wenn der Zähler, welcher die Inter-Klassen-Varianz (engl.: *Between Scatter Matrix* $S_{BETWEEN}$) beinhaltet, größtmöglich und der Nenner mit der Intra-Klassenvarianz (engl.: *Within Scatter Matrix* S_{WITHIN}) kleinstmöglich wird.

Die Spaltenvektoren der Transformationsmatrix lassen sich wie auch bei den Eigenfaces bildlich darstellen. Diese sog. *Fisherfaces* besitzen ein ähnlich geisterhaftes Erscheinungsbild.

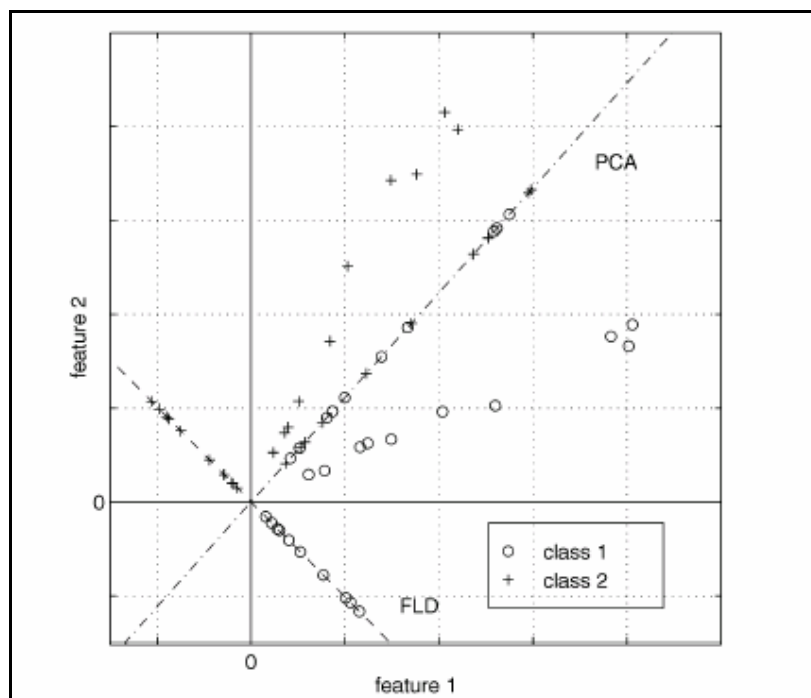


Abb. 3.3.1a

Abbildung 3.3.1a stellt die beiden Transformationen, PCA und LDA, gegenüber. Jeweils zwei Personenklassen, hier mit zwei Features

dargestellt, sind mit 10 Samples vertreten. Zwar findet die PCA die Richtung der größten Varianz, die LDA ermittelt jedoch den Projektionsvektor, der die Klassen optimal trennt und die Samples einer Klasse enger zusammenrückt.

3.3.2 Fisherfaces - Algorithmus

Für die Lösung des Fisher-Kriteriums [f. 3.3.1a] wird die *Within Scatter Matrix*, also die Summe aller Klassenvarianzen, benötigt:

$$S_{WITHIN} = \sum_{j=1}^M p_j \Phi_j \Phi_j^T \quad [f. 3.3.2a]$$

Die Berechnung erfolgt mit einem Trainingsdatensatz der Bilder von insgesamt M verschiedenen Personen bzw. Klassen besitzt. p ist die Wahrscheinlichkeit, mit der ein Bild der Klasse j im Datensatz auftaucht. Ist jede Klasse mit gleich vielen Bilder vertreten, so ist $p = 1/M$ für alle Klassen.

Die Inter-Klassenvarianz gibt die *Between Scatter Matrix* an:

$$S_{BETWEEN} = \sum_{i=1}^K N_i \left[(\Phi_i - \Phi_{global}) (\Phi_i - \Phi_{global})^T \right] \quad [f. 3.3.2b]$$

Die *Between Scatter Matrix* berechnet sich aus der Summe der Abweichungen des Durchschnitts einer Klasse Φ_i (*Class Mean*) vom Gesamtdurchschnitt des Bilddatensatzes Φ_{global} (*Global Mean*). N ist die Anzahl der Bilder in Klasse i .

Die Transformationsmatrix ergibt sich aus der Lösung des Eigensystems des Fisher-Kriterium:

$$\left(S_{WITHIN}^{-1} S_{BETWEEN} \right) E = E \Lambda \quad [f. 3.3.2c]$$

3.4. Klassifizierung

Vor der eigentlich Klassifizierung wurde bei beiden Methoden, Eigenfaces und Fisherfaces ein Trainingsdatensatz von vorhanden Bildern verwendet, um die Transformationsmatrix zu finden.

Um ein unbekanntes Bild eines Gesichts (*Query Picture*) einer Person im Trainingsdatensatz zuordnen zu können, werden folgende Schritte durchgeführt:

1) Transformation des Bilddatensatzes in Unterraum

$$\mathbf{y} = \mathbf{W}^T \Phi \quad [\text{f. 3.4a}]$$

2) Transformation des Query-Bildes

$$\mathbf{y}_{query} = \mathbf{W}^T \Phi_{query} \quad [\text{f. 3.4b}]$$

3) Berechnung der Euklidischen Distanz

$$\mathbf{e}_i^2 = \left\| \mathbf{y}_i - \mathbf{y}_{query} \right\|^2 \quad [\text{f. 3.4c}]$$

Im dritten Schritt wird die Distanz des Feature-Vektors \mathbf{y}_{query} des Query-Bildes zu allen Vektoren des Unterraum \mathbf{y} berechnet.

Ein unbekanntes Gesicht wird derjenigen Personenklasse im Bilddatensatz zugeordnet, bei der die Distanz am kleinsten ist.

Des Weiteren wird ein Schwellenwert bestimmt, der die maximal erlaubte Distanz angibt, die ein Gesicht im Datensatz besitzen darf.

Liegt die Distanz eines Query-Bildes über diesem Schwellenwert, wird von einer sich nicht in der Datenbank befindenden Person ausgegangen.

3.5. Small Sample Size Problem

In der automatisierten Gesichtserkennung begegnet man einem speziellen Problem, das unter dem Begriff *Small Sample Size Problem* zusammengefasst wird.

Im Allgemeinen ist von einer Person nur eine geringe Anzahl von Bildern im Trainingsdatensatz vorhanden. Dadurch kann es passieren, dass diese wenigen Samples die Personklasse schlecht bzw. unzureichend repräsentiert.

Neben der schlechten/unzureichenden Repräsentation einer Klasse ergibt sich auch ein mathematisches Problem bei der Implementation der Gesichtserkennungsalgorithmen: da die Anzahl der Feature N eines

Samples die Anzahl der Samples M im Trainingsatz um ein vielfaches übersteigt ($N \gg M$), kann die Kovarianzmatrix singular werden. Eine singuläre Matrix besitzt Spalten oder Zeilen, die komplett aus Nullwerten bestehen, so dass sich zum einen die Determinante zu Null ergibt, zum anderem die Bildung der Inversen unmöglich ist.

Betrachtet man das Fisher-Kriterium [f. 3.3.1a], lässt sich folgendes feststellen:

Der Nenner des Fisher-Kriteriums ist, mathematisch gesehen, die Berechnung der Determinante der *Within scatter matrix*. Ist diese Intra-Klassenvarianzmatrix singular, wird der Nenner zu Null und die Berechnung der Arguments unmöglich.

Eine einfache Lösung des *Small Sample Size Problem* zeigt Belhumeur [3]. Er stellt der LDA eine Dimensionsreduktion der Features durch eine PCA voran. Damit wird die Anzahl der Features je Sample von N auf c Features (mit: $M < c < (N - M)$) verringert.

Der Schritt sichert, dass die Kovarianzmatrix nicht singular ist, hat aber den Nachteil, dass mit der vorausgehenden PCA auch ein Verlust an Information einhergeht, der letztendlich zu einer Verschlechterung der Klassifikation führt.

4. Erweiterte Methoden

Die im Folgenden besprochenen auf der LDA basierenden Methoden lösen das Singularitätsproblem, ohne eine verlustbringende Dimensionsreduktion mittels PCA anzuwenden.

4.1 A Maximum Uncertainty LDA

Der *Maximum Uncertainty LDA*-Ansatz [5] stabilisiert die *Within Scatter Matrix* durch die Modifikation des Fisher-Kriteriums.

Die Verfasser Thomasz und Gillies lösen das Singularitätsproblem der *Within Scatter Matrix*, indem sie ihr eine gewichtete Einheitsmatrix I hinzuaddieren:

$$\hat{S}_{WITHIN} = (I - g)S_{WITHIN} + g\Lambda I, \quad [f. 4.1a]$$

wobei der sog. *Shrinkage*-Parameter g einen Wert zwischen 0 und 1 annimmt und experimentell ermittelt wird.

Die Einheitsmatrix I wird zusätzlich zu g auch mit dem Mittelwert der Eigenwerte der Within scatter matrix multipliziert:

$$\Lambda = \frac{1}{n} \sum_{i=1}^n \mathbf{I}_i . \quad [\text{f. 4.1b}]$$

Das modifizierte Fisher-Kriterium ergibt sich nun zu:

$$W_{NLDA} = \underset{W}{\arg \max} \frac{\left| W^T S_{BETWEEN} W \right|}{\left| W^T \hat{S}_{WITHIN} W \right|} . \quad [\text{f. 4.1c}]$$

Die Transformationsmatrix der Methode von Thomasz und Gillies (NLDA steht für: *the New LDA*) wird wiederum durch die aus [f. 3.3.2c] bekannten Lösung des Eigensystem s gefunden.

4.2. Direct Fractional LDA

Die DF-LDA-Methode (*Direct Fractional Linear Discriminant Analysis*) von Lu, Plataniotis und Venetdanoppolos [6] befasst sich ebenfalls mit der Lösung des Singularitätsproblems.

Der eigentliche Schwerpunkt ihrer Arbeit liegt jedoch auf der Weiterentwicklung des LDA-Algorithmus, um eine die Klassifizierung zu verbessern.

Die Berechnung der Transformationsmatrix über den DF-LDA-Algorithmus ist in drei Schritte unterteilt:

Im **ersten Schritt** wird eine, die Diskriminierung verbessernde, modifizierte *Between Scatter Matrix* [f. 3.3.2b] berechnet:

$$\hat{S}_{BETWEEN} = \sum_{i=1}^K \Phi_i \Phi_i^T$$

$$\text{mit : } \Phi_i = \frac{M_i}{M} \sum_{j=1}^K \sqrt{\left(\|\Psi_i - \Psi_j\| \right)^{-2p}} (\Psi_i - \Psi_j) \quad [\text{f. 4.2a}]$$

Die Gewichtsfunktion $\sqrt{\left(\|\Psi_i - \Psi_j\|\right)^{-2p}}$ bewertet die Klassen, indem sie dicht zusammenliegende Klassen mit auseinanderrückt. Der Faktor p in der Potenz wurde bei den Tests mit 2 bis 8 gewählt.

Im **zweiten Schritt** wird der *Direct LDA*-Schritt (D-LDA) mit dem neu formulierten Fisher-Kriterium vollzogen:

$$W_{D-LDA} = \arg \max_W \frac{|W^T \hat{S}_{BETWEEN} W|}{|W^T (S_{WITHIN} + \hat{S}_{BETWEEN}) W|} \quad [f. 4.2b]$$

Das Singularitätsproblem der *Within Scatter Matrix* im Nenner des ursprünglichen Arguments [f. 3.3.1a] wird dadurch gelöst, dass die *Within Scatter Matrix* mit der in Schritt 1 [f. 4.2a] berechneten $\hat{S}_{BETWEEN}$ addiert wird.

Der mit der Projektionsmatrix W_{D-LDA} transformierte Datensatz wird im **dritten Schritt** weiteren Reduktionsschritten unterzogen.

Die dafür verwendete Transformationsmatrix ergibt sich aus der gewöhnlichen *Between Scatter Matrix* [f. 3.3.2b] zu:

$$W_{DF-LDA} = \arg \max_W |W^T S_{BETWEEN} W| \quad [f. 4.2c]$$

Die Dimensionsreduktion erfolgt über $r > 1$ Schritten, wobei W_{DF-LDA} nach jeder Transformation neu berechnet wird, weil sich auch die Kovarianzen ändern. Die Anzahl r der Reduktionsschritte wurde bei den Tests mit 20 angegeben.

5. Ergebnisse

An dieser Stelle muss darauf hingewiesen werden, dass es bei den Standardmethoden Eigenfaces und Fisherfaces nicht möglich ist, eine allgemein gültige Zusammenfassung von experimentellen Ergebnissen vorzustellen.

Zum einen erzielen die Algorithmen bei unterschiedlichen Testbilddatenbanken auch teils sehr unterschiedliche Ergebnisse (das siehe Kapitel 6).

Zum anderen präsentieren die verschiedenen wissenschaftlichen Arbeiten [3, 7 -10] trotz gleicher verwendeter Methoden und Bilddatenbanken

unterschiedliche Ergebnisse. Die Ursache dafür könnten geringe Abweichungen in der Implementierung der Algorithmen oder auch eine Vorverarbeitung der Bilder sein. Dies ist jedoch in den Arbeiten nicht explizit erläutert.

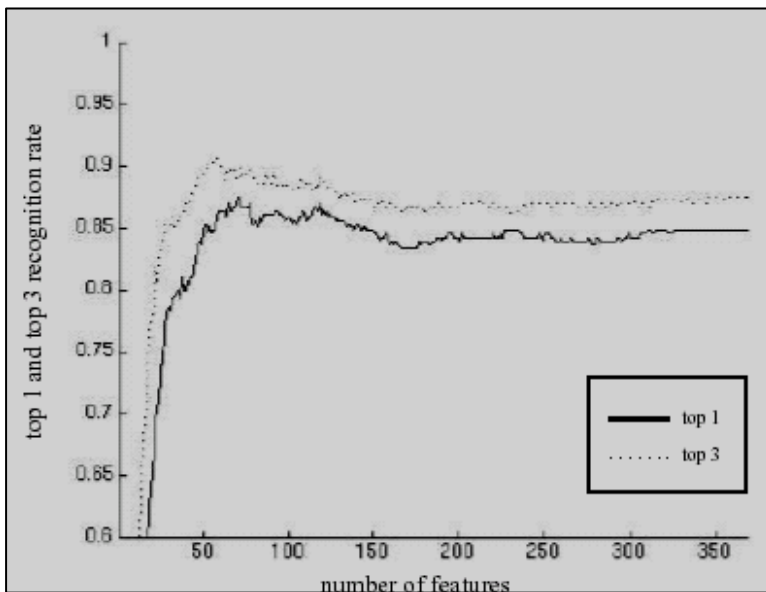


Abb. 5a

auch die Kurve, die einen Treffer unter den ersten drei gefundenen Zuordnungen widerspiegelt („Top 3“- Graph) steigen steil bis zu einem Erkennungsratenmaximum bei etwa 70 genommen PCs an. Danach bringt eine Erhöhung der Anzahl der PCs keine Klassifikationsgewinn.

Dennoch lassen sich generelle allgemeingültige Feststellungen treffen:

Wie Abbildung 5a verbildlicht, hat die Anzahl der *Principle Components* (im Bild bezeichnet als „number of features“) Auswirkung auf das Klassifikationsergebnis. Sowohl die Kurve der exakten Klassifikationen („Top 1“- Kurve), als

Ein zweite Erkenntnis ist, dass die Fisherfaces-Methode mit der LDA generell bessere Klassifikationsergebnisse erzielt als die Eigenfaces-Methode mit der *Principle Component Analysis* (siehe Tabelle 5b aus [7]).

Projektionsmethode	Trainingsbilder / Klasse	PCs	Erkennungsrate
PCA	6	56	87,90%
LDA	6	14	91,50%
PCA	4	46	87,30%
LDA	4	14	90,30%
PCA	2	26	82,70%
LDA	2	14	81,50%

Tabelle 5b

Des Weiteren spielt auch die Anzahl der in die Trainingsdatenbank einfließenden Bilder eine Rolle. In der Tabelle 5b steigt die Erkennungsrate mit der Anzahl der Trainingsbilder. Die gleiche Erkenntnis gewinnt man aus der Abbildung 5c.

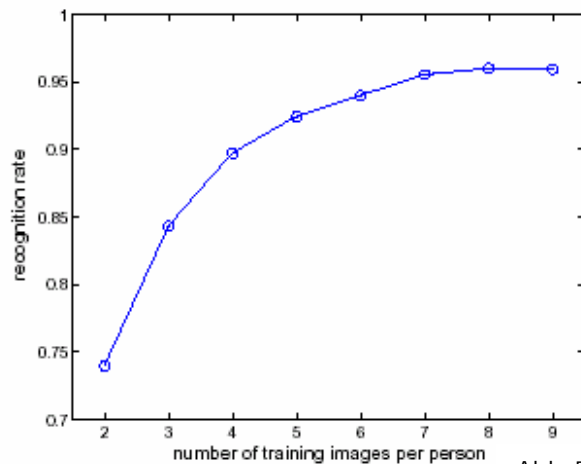


Abb. 5c

Ein Mehr an Trainingsbildern, die den statistischen Methoden zur Berechnung der Transformationsmatrix zur Verfügung stehen, führen zu einer besseren Repräsentation einer

Personenklassen und damit zu besseren Klassifikationsergebnissen.

A Maximum Uncertainty LDA (NLDA)

Für die in Kapitel 4 erläuterten erweiterten Methoden bestand die Aufgabe letztendlich darin, die Erkennungsperformanz der Fisherfaces-Methode zu übertreffen.

Die NLDA [5] erreicht durch die Modifizierung des Fisher-Kriteriums eine insgesamt bessere Erkennungsrate als das Fisherfaces-Verfahren (Tab. 5c). Die gewählte Anzahl der *Principle Components* führte in den Tests jeweils zum besten Klassifikationsergebnis.

Projektionsmethode	Bildtestdatenbank	PCs	Erkennungsrate
Fisherfaces (PCA / LDA)	ORL	PCA: 60 / LDA: 39	94,90%
NLDA	ORL	39	95,80%
Fisherfaces (PCA / LDA)	FERET	PCA: 200 / LDA: 20	91,50%
NLDA	FERET	10	95,40%

Tabelle 5c

Direct Fractional LDA (DF-LDA)

Die Ergebnisse der DF-LDA (Tab. 5d) beweisen, dass auch diese Methode bessere Klassifikationsergebnisse gegenüber der Fisherfaces-Methode erzielt. Die Werte der Erkennungsrate sind die Durchschnittswerte der Ergebnisse mit 5 bis 25 (bei der ORL-Bilddatenbank) bzw. 3 bis 12 (UMIST) verwendeten *Principle Components*.

Die DF-LDA wurde bei Leistungstests mit 5 (ORL) bzw. 8 (UMIST) Trainingsbildern „gefüttert“ und führte jeweils 20 *Fractional Steps* aus.

Projektionsmethode	Bildtestdatenbank	PCs	Erkennungsrate
Fisherfaces (PCA / LDA)	ORL	5 bis 25	38,51%
DF-LDA	ORL	5 bis 25	80,03%
Fisherfaces (PCA / LDA)	UMIST	3 bis 12	47,68%
DF-LDA	UMIST	3 bis 12	79,82%

Tabelle 5d

Die beste Performanz lieferte die Methode bei der ORL-Datenbank mit 22 PCs und einer Fehlererkennungsrate von lediglich 4 %.

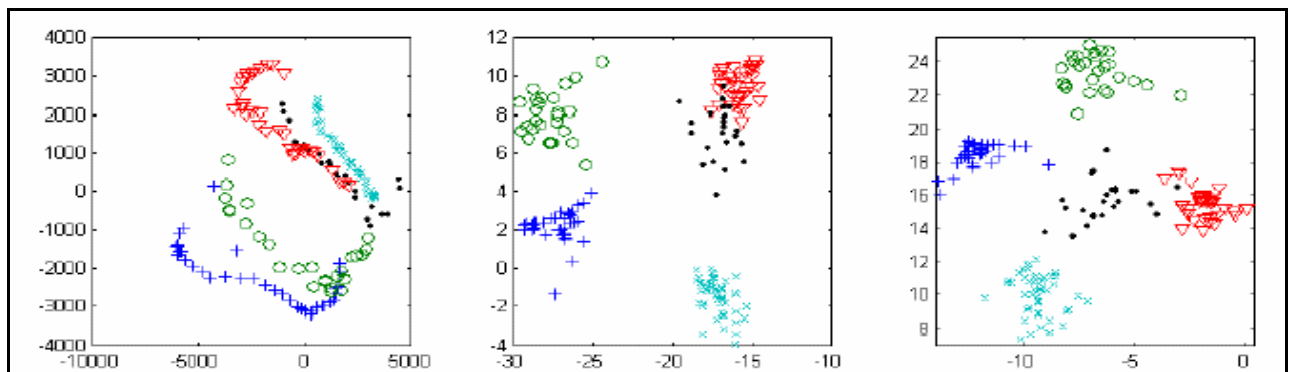


Abbildung 5e

Die Abbildung 5e verdeutlicht noch einmal auf anschauliche Weise die Effizienz der DF-LDA. Das linke Bild zeigt die Features von fünf Personenklassen nach der Transformation mittels PCA. Das mittlere Bild vermittelt die besser trennbaren Klassencluster nach dem Direct Step LDA. Das rechte Bild schließlich zeigt die Projektion mittels DF-LDA. Jede Personenklasse lässt sich sehr gut räumlich von den anderen trennen.

6. Zusammenfassung

Die in dieser Arbeit vorgestellten Gesichtserkennungsverfahren sind so genannte Hauptachsentransformationsverfahren und führen eine Analyse des gesamten Gesichts mit Hilfe statistischer Methoden durch. Die Projektion der vektorisierten Bilder in den Unterraum (*Feature Space*) hat den Zweck, zum einen die Objektdimension zu reduzieren und damit

den Rechenaufwand zu minimieren und zum anderen die Diskriminierung der Objektklassen zu erhöhen. Die Hauptachsentransformationsverfahren erweisen sich gegenüber anderer Gesichtserkennungsverfahren (*Template Matching, Active Shape Mode*) robuster hinsichtlich Helligkeitsschwankungen, Positionsänderungen des Kopfes, Mimik und Bildauflösungsschwankungen.

Es zeigt sich, dass das Fisherfaces-Verfahren, welches die *Linear Discriminant Analysis* (LDA) zugrunde legt, der Eigenfaces-Methode (PCA) überlegen ist. Die LDA führt durch die bessere Trennung der Objektklassen besonders bei Helligkeitsschwankungen und hoher Varianz der Gesichtsmimik zu geringeren Fehlerraten.

Die erweiterten, auf der LDA basierenden Verfahren, NLDA und DF-LDA, verbessern die Erkennungsraten um ein weiteres, indem sie das Fisher-Kriterium modifizierten und die Klassendiskriminierung erhöhen.

Die einzelnen Methoden erzielen bei unterschiedlichen Bilddatenbanken auch unterschiedlich gute Klassifikationsergebnisse, so dass ein direkter Vergleich der Methoden oft nicht möglich ist.



Abb. 6

Die Yale-Bilddatenbank (Abbildung 6) legt den Schwerpunkt auf starke Helligkeitsschwankungen, während die FERET-Datenbank Personen mit unterschiedlichen Kopfpositionen und variierender Mimik beinhaltet.

7. Literaturverzeichnis

- [1] **Turk**, Metthew, **Pentland**, Alex, *Eigenfaces for Face Recognition*, Journal of Cognitive Neuroscience Vol. 3, Nr.1, Seiten 71-86, 1991
- [2] **Sirovich**, L., **Kirby**, M., *Low-dimensional procedure for the characterization of human faces*, Journal of Optical Society of America , 4(3), Seiten 519-524, 1987
- [3] **Belhumeru**, Peter., **Hespanha**, João., **Kriegmann**, David *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*, IEEE Transactions on Pattern Analysis and Machine Intelligence , vol. 19, no. 7, Seiten 711-720, 1997
- [4] **Skurichina**, Martina., *Stabilizing Weak Classifiers - Regularization and Combining Techniques in Discriminant Analysis* , Seiten 199-201,
- [5] **Thomasz**, Carlos, **Gillies**, Duncan, *A Maximum Uncertainty LDA-based Approach For Limited Sample Size Problems – with Application to Face Recognition*
- [6] **Lu**, Juwei, **Plananiotis**, K. N., **Venetsanopoulos**, A. N., *Face Recognition Using LDA Based Algorithms*, IEEE Transactions on Neural Networks, 2002
- [7] **Navarette**, Pablo, **Riuz-del-Solar**, Javier, *Analysis and Comparison of Eigenspace-based Face Recognition Approaches* , 2001
- [8] **Yang**, Jian, **Yang**, Jing-Yu, *Combined Fisherfaces Framework* , Image and Vision Computing 21, Seiten 1037-1044, 2003
- [9] **Liu**, Chenjun, **Wechsler**, Harry, *Enhanced Fisher Linear Discriminant Models For Face Recognition*, 14th International Conference on Pattern Recognition, ICPR, 1998

[10] **Zhuang, X., Dai, Dao-Qing**, *Inverse Fisher discriminate criteria for small sample size problem and its application to face recognition*, <http://www.paper.edu.cn>, 2004

8. Abbildungsverzeichnis

Abb. 2.1a: www.informatik.uni-ulm.de/ni/Lehre/WS03/HSBiometrie/ausarbeitungen/GoldsteinRockmann.pdf, Juni 2005

Abb. 3.1a: vision.ai.uiuc.edu/~wanghc, **Wang, Hengchung**: Facial Recognition as a Pattern Recognition Problem, Juli 2005

Abb. 3.1.1a / Abb. 6:

Belhumeru, Peter., Hespanha, João., Kriegmann, David *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, Seiten 711-720, 1997

Abb. 3.2.1.1a:

<http://www.eng.man.ac.uk/mech/merg/Research/datafusion.org.uk/pca.htm>, Juni 2005

Abb. 3.2.1.1b / Abb. 5a

Liu, Chenjun, Wechsler, Harry, *Enhanced Fisher Linear Discriminant Models For Face Recognition*, 14th International Conference on Pattern Recognition, ICPR, 1998

Abb. 3.2.1c / Abb. 2.1b:

<http://www.markus-hofmann.de/>, Juni 2005

Abb. 6c: **Zhuang, X., Dai, Dao-Qing**, *Inverse Fisher discriminate criteria for small sample size problem and its application to face recognition*, <http://www.paper.edu.cn>, 2004

Abb. 5e: **Lu, Juwei, Plananiotis, K. N., Venetsanopoulos, A. N.**, *Face Recognition Using LDA Based Algorithms*, IEEE Transactions on Neural Networks, 2002