

# Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection

Peter N. Belhumeur, João P. Hespanha, David J. Kriegman  
IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7,  
July 1997

Marco Kortkamp  
Visuelle Überwachung - Gesichtserkennung

## 1 Vorwort

Das zweite Kapitel behandelt eine kurze Einleitung zum Thema Gesichtserkennung. In der Einleitung wird beschrieben was man unter der Gesichtserkennung versteht und wo sich Schwierigkeiten ergeben.

Im dritten Kapitel folgt der Paper Review. Der Review beginnt mit dem Kernbeitrag des Papers in einem Satz. Danach wird die Relevanz des Papers in Bezug auf das Seminar bzw. den praktischen Einsatz dargestellt. Im Anschluss daran folgen die kurze Zusammenfassung des Inhalts sowie die Diskussion des Papers.

## 2 Einleitung zur Gesichtserkennung

Das Ziel der Gesichtserkennung ist die automatische Zuordnung eines Gesichtsbildes zu einer Person bzw. eine Rückweisung. Hierbei ist die zugewiesene Person dem Gesichtserkennungsverfahren bekannt, d.h. durch eine Datenbank mit Gesichtsbildern und zugehörigen Personen. Grob gesagt geht es also um die Frage, *wer* auf einem Bild zu sehen ist.

Bevor das Gesicht einer Person jedoch erkannt werden kann, ist es in der Regel notwendig, dass Gesicht im Bild zunächst zu finden. Dieser Schritt wird als Gesichtsdetektion bezeichnet. Bei den meisten Verfahren die sich mit der reinen Gesichtserkennung beschäftigen wird angenommen, dass das Gesicht im Bild bereits detektiert wurde.

Man kann drei grundlegende Arten von Gesichtserkennungsszenarien unterscheiden: *Identification* („Wer bin ich?“), *Verification* („Bin ich der, der ich vorgebe zu sein?“) und *Watch-List* („Suchst du nach mir?“). Dabei können bei der *Verification* und *Watch-List* auch Bilder von Personen auftreten, die nicht in der Datenbank sind. Bei der *Identification* ist dies nicht der Fall. Manchmal wird jedoch nicht explizit zwischen den drei Arten unterschieden, da viele grundlegende Probleme und Techniken gleich sind. Für eine genauere Beschreibung der verschiedenen Gesichtserkennungsszenarien siehe [1].

Viele praktische Aufgaben der Gesichtserkennung werden durch die drei grundlegenden Arten von Gesichtserkennungsszenarien abgedeckt. Einige Beispiele sind visuelle Überwachung, Zugangskontrolle, Gesichtsrekonstruktion, Design von Mensch-Maschine-Schnittstellen und Biometrie. Vergleiche [1].

Die Schwierigkeiten der Gesichtserkennung ergeben sich durch die Variationen verschiedener Parameter. Einige dieser Parameter sind Kopfpose, Beleuchtung (Intensität, Lage und Anzahl der Lichtquellen), Alterung, Gesichtsausdrücke, Größe des Bildes, Verkleidungen (Brillen, Hüte etc.), Verdeckungen und schlechte Bildqualität (Unschärfe etc.). Vergleiche [1] und [2].

Neben den Schwierigkeiten bei der Gesichtserkennung ist auch die Evaluation von Verfahren schwierig. Für die Evaluationen in den Publikationen treffen die Wissenschaftler ihre eigenen Entscheidungen bzgl. der verwendeten Datenbank, des verwendeten Evaluierungsverfahrens (leave-one-out, cross-validation etc.) und der verwendeten Bildgröße. Dadurch wird ein Vergleich schwierig. Darüber hinaus ist der direkte Vergleich von Methoden auf der gleichen Datenbank schwierig, da den Methoden unterschiedliche Annahmen zugrunde liegen können. Mehr Details zur Evaluation von Gesichtserkennungsverfahren findet man in [1] und [3].

Das man bei Anwendungen wie der visuellen Überwachung meist Videodaten vorliegen hat, bringt für die Gesichtserkennung einige Vor- und Nachteile. Auf Videodaten werden oft Standbildverfahren auf einzelne Bilder einer Bildsequenz angewendet. Ein Vorteil bei Bildsequenzen kann sein, dass man „günstige“ Bilder für die Gesichtserkennung aus der Sequenz auswählen kann. Ein Nachteil bei Bildsequenzen ist z.B. die schlechte Bildqualität. Einen kurzen Überblick bietet [2].

### **3 Paper Review: Eigenfaces vs. Fisherfaces**

In dem Paper [4] schlagen die Autoren einen (damals) neuen Ansatz zur Gesichtserkennung vor, bei dem durch die Verwendung der Fisher Diskriminanzanalyse eine Unempfindlichkeit gegen relativ große Variationen der Beleuchtung und von Gesichtsausdrücken erzielt werden kann.

#### **3.1 Relanz in Bezug auf Seminar bzw. praktischen Einsatz**

Wie bereits in Kapitel 2 erwähnt, kann man bei der visuellen Überwachung meist von Video-Daten ausgehen. Die im Paper vorgestellten Verfahren - insbesondere Fisherfaces - sind jedoch nicht explizit für Bildsequenzen gedacht. Zu der Zeit, als das Paper entstanden ist, war die Gesichtserkennung in Videos auch erst im aufkommen (vgl. [3]). Allerdings kann man Fisherfaces sowie die anderen Verfahren des Papers auf einzelne Bilder einer Bildsequenz anwenden.

In den praktischen Aufgaben der visuellen Überwachung (z.B. Überwachung öffentlicher Plätze) können wir von unkontrollierten Umgebungen und für die Gesichtserkennung von einer Vielzahl der in Kapitel 2 genannten Variationen ausgehen. Es wird hier umso wichtiger, Methoden zu entwickeln, die gegenüber diesen Variationen robust sind.

Das Fisherface Verfahren ist unter gewissen Annahmen relativ unempfindlich gegenüber starken Variationen der Beleuchtung und von Gesichtsausdrücken. Dies ist der Versuch, mit einem Teil der Variabilität umzugehen. Die Methode ist jedoch nur unter gewissen Annahmen und in halbwegs kontrollierten Umgebungen einsetzbar. Für die Gesichtserkennung in realistischen Umgebungen ist das im Paper vorgestellte Verfahren jedoch zu schwach, da es insgesamt mit zu wenig Variationen klar kommt. Einen kleinen Überblick über einfache und schwere Gesichtserkennungsszenarien findet man in [2].

### 3.2 Zusammenfassung

Jeder Pixel eines  $M \times N$  Bildes wird als eine Koordinate in einem  $M \cdot N$  dimensionalen Vektorraum (Bildraum) betrachtet. D.h. ein Bild ist ein „langer“ Vektor, in dem die einzelnen Intensitätswerte der Bildzeilen untereinander geschrieben sind. Darüber hinaus werden folgende Annahmen getroffen:

- Gesichter sind bereits detektiert (ggfs. normalisiert); d.h. es werden reine Gesichtsbilder verwendet
- die verwendeten Bilder sind Grauwertbilder der gleichen Größe
- Kopfhöhe variiert nicht
- im Grunde: keine anderen Variationen bis auf Beleuchtung und Gesichtsausdrücke

Im Paper werden die folgenden vier Verfahren miteinander verglichen. Zur Klassifikation eines Gesichts verwenden alle Verfahren einen Nächster-Nachbar-Klassifikator (NNK). Dabei wird ein Merkmalsvektor der Klasse zugeordnet, die das nächstgelegene Stichprobenelement besitzt.

**Correlation:** Bei diesem Ansatz werden die Bilder direkt als hochdim. Vektoren zur Klassifikation verwendet. Durch eine vorherige Normierung aller Bilder entsteht eine gewisse Unabhängigkeit von der Intensität der Lichtquelle. Einige Nachteile dieses Ansatzes sind der Rechenaufwand sowie der benötigte Speicherplatz.

**Eigenfaces:** Bei diesem Verfahren wird zur Dimensionsreduktion eine Hauptkomponentenanalyse (Principal Components Analysis) verwendet. Die Hauptachsentransformation lässt sich durch eine Transformationsmatrix angeben, die aus den Eigenvektoren der Kovarianzmatrix gebildet wird. Die Kovarianzmatrix ist definiert als

$$S_T = \sum_{k=1}^N (\underline{x}_k - \underline{\mu})(\underline{x}_k - \underline{\mu})^T$$

wobei  $\{\underline{x}_1, \dots, \underline{x}_N\}$  die Trainingsbilder und  $\underline{\mu}$  das Durchschnittsgesicht (Bsp. siehe Abbildung 1) ist. Die Eigenvektoren von  $S_T$  werden als Eigenfaces bezeichnet (siehe Abbildung 1). Sie haben die gleiche Dimension wie die Bilder. Man verwendet die Eigenvektoren  $\{\underline{w}_1, \dots, \underline{w}_m\}$  der Matrix  $S_T$ , die die  $m$  größten Eigenwerte haben. Für ein Bild  $\underline{x}$  ist die Transformation in den niedrig dimensional Merkmalsraum gegeben durch  $\underline{y} = W\underline{x}$  mit  $W = [\underline{w}_1, \dots, \underline{w}_m]$ . Die PCA maximiert die Varianz



Abbildung 1: Durchschnittsgesicht (oben) [1]; die ersten 7 Eigenfaces (unten) [1]

aller projizierten Trainingsbilder, d.h. sowohl Interklassenvarianz als auch Intraklassenvarianz werden maximiert.

**Linear Subspaces:** Dieser Ansatz verwendet die folgende Beobachtung: Alle Bilder einer Lambertischen Oberfläche<sup>1</sup> liegen in einem 3D linearen Unterraum des hochdim. Bildraums, wenn sie unter einem festen Betrachtungswinkel und variierender Beleuchtung aufgenommen werden. Man verwendet 3 oder mehr Bilder die unter verschiedener Beleuchtung aufgenommen wurden, um eine Basis des 3D Unterraums zu bilden. Ein Gesicht wird der Person zugeordnet, die den nächstliegenden Unterraum hat. Da Gesichter keine Lambertische Oberfläche sind, weichen die Bilder von dem 3D Unterraum ab.

**Fisherfaces:** Die Verwendung der Fisher Diskriminanzanalyse (Fisher's Linear Discriminant) zur Dimensionsreduktion, hat gegenüber der PCA den Vorteil, dass die Intraklassenvarianz minimiert und die Interklassenvarianz maximiert wird. Das Ergebnis der FLD liefert vom Unterscheidungsstandpunkt aus gesehen bessere Ergebnisse, da die Klassifikation vereinfacht wird (siehe Abbildung 3). Die Interklassen-Kovarianzmatrix ist definiert als

$$S_B = \sum_{i=1}^C N_i (\underline{\mu}_i - \underline{\mu})(\underline{\mu}_i - \underline{\mu})^T$$

und die Intraklassen-Kovarianzmatrix als

$$S_W = \sum_{i=1}^C \sum_{x_k \in X_i} (x_k - \underline{\mu}_i)(x_k - \underline{\mu}_i)^T$$

mit  $\underline{\mu}_i$  Durchschnittsgesicht der Klasse  $X_i$  und  $N_i$  Anzahl der Bilder in Klasse  $X_i$ . Die Fisherfaces sind die generalisierten Eigenvektoren von  $S_B$  und  $S_W$ . Sie haben die gleiche Dimension wie die Bilder (siehe Abbildung 2). Man verwendet nur die ersten  $m$  Eigenvektoren, die die  $m$  größten Eigenwerte haben. Für ein Bild  $x$  ist die Transformation in einen niedriger dimensional Merkmalsraum gegeben durch  $\underline{y} = Wx$  wobei  $W = [\underline{w}_1, \dots, \underline{w}_m]$  die Matrix der  $m$  Eigenvektoren ist. Dadurch wird das Verhältnis von  $S_B$  und  $S_W$  maximiert.

Bei der Gesichtserkennung kann das Problem auftreten, dass  $S_W$  singular ist. Dieses Problem kann man z.B. vermeiden, indem man die Bilder mit einer PCA in einen

<sup>1</sup>Eine Lambertische Oberfläche ist eine Oberfläche mit absolut matten Reflexionseigenschaften, d.h. sie die das Lambertische Kosinus-Gesetz befolgt. Die Helligkeit ist somit konstant und unabhängig vom Betrachtungswinkel.



Abbildung 2: die ersten 7 Fisherfaces [1]

niedrig dimensionalen Vektorraum projiziert und dann die FLD anwendet.

Die beschriebenen Verfahren wurden auf zwei Datenbanken getestet. Die eine Datenbank enthält 330 Bilder von 5 Personen und die andere Datenbank enthält 160 Bilder von 16 Personen.

Die Fisherface Methode scheint am besten mit Variationen in der Beleuchtung klarzukommen, wobei die Linear Subspace Methode knapp dahinter liegt. Die Fisherface Methode scheint besser als die anderen Methoden zu sein, wenn sowohl Beleuchtung als auch Gesichtsausdrücke variieren.

### 3.3 Diskussion

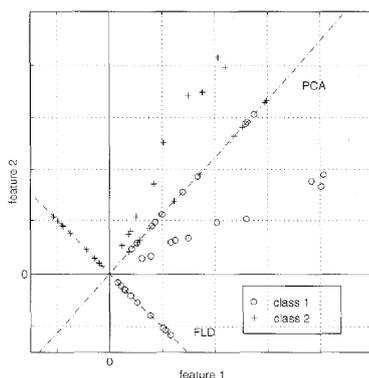


Abbildung 3: PCA vs. FLD [4]

Meiner Meinung nach wird in dem Paper gut erklärt, warum der Einsatz der FLD zur Gesichtserkennung Sinn macht und warum man mit der FLD bessere Ergebnisse erzielt als mit der PCA. Bei der Evaluierung der Verfahren würde ich bemängeln, dass die verwendeten Datenbanken einen zu geringen Umfang haben (insgesamt weniger als 500 Bilder). Weiterhin fehlt mir die Evaluation unter extremen Bedingungen. Im Hinblick auf den praktischen Einsatz müsste meiner Meinung nach die Erweiterbarkeit bzw. die Möglichkeit des Umgangs mit anderen Variationen (z.B. Kopfpose) diskutiert werden.

Als Alternative zu den „traditionellen“ Verfahren mit linearen Transformationen (wie Eigenfaces und Fisherfaces) werden heute eher nicht lineare Transformationen (wie Kernel PCA) verwendet.

Einen guten und halbwegs aktuellen Überblick zur Gesichtserkennung findet man in [1] und [3].

Das Paper ist heute noch immer gut zu lesen und bietet einen „traditionelles“ Verfahren zur Gesichtserkennung. In kontrollierten Umgebungen und unter gewissen Annahmen lassen sich relativ gute Ergebnisse erzielen. Bewegt man sich jedoch weiter in Richtung Praxis und unkontrollierter Umgebungen, sind Fisherfaces zunehmend nicht mehr einsetzbar.

## Literatur

- [1] **Xiaoguang Lu**  
*Image Analysis for Face Recognition*;  
personal notes; 2003
- [2] **Luis Torres**  
*Is There Any Hope for Face Recognition?*;  
Proc. of the 5th International Workshop on Image Analysis for Multimedia  
Interactive Services; 2004
- [3] **W. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips**  
*Face Recognition: A Literature Survey*;  
ACM Computing Surveys; 2003
- [4] **Peter N. Belhumeur, Joao P. Hespanha, David J. Kriegman**  
*Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*;  
IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No.  
7; 1997