

Gesichtserkennung

Seminar
Biometrische Identifikationsverfahren
SoSe2004

Yasushi Tomii
Steffen Scheer

1. Einleitung

Bei der Gesichtserkennung werden die charakteristischen Merkmale eines menschlichen Gesichtes verwendet um Personen zu identifizieren bzw. zu verifizieren.

Erste Versuche charakteristische Gesichtsmarkmale mit Hilfe von Zahlenserien zu abstrahieren wurden schon im 19. Jahrhundert durchgeführt. Sir Galton, ein engl. Mathematiker war einer der ersten, der sich mit dieser Problematik auseinandersetzte. Daraus entwickelte sich im letzten Jahrhundert eine Einzeldisziplin: die Physiognomie. Die systematische Untersuchung von Gesichtern. In den 70er Jahren wurde die automatische Gesichtserkennung mittels computergestützter Systeme, Ziel verschiedener Forschungsprojekte. Zu erwähnen ist hier die Veröffentlichung „Picture processing by computer complex and recognition of human faces“ von 1973. Durch den Gesichtserkennungswettbewerb „FERET“ (Face Recognition Technology) der US Army im Jahr 1993, hat sich die Gesichtserkennung enorm weiterentwickelt. Heute steht dabei die Identifikation und Authentisierung von Personen anhand von Körpermerkmalen, die sogenannte Biometrie im Vordergrund. Innerhalb der Biometrie ist die Gesichtserkennung mit Sequenzen als Einzeldisziplin, dem sogenannten „Tracking“ entstanden. Dabei findet die Gesichtserkennung in allen Bereichen in denen eine Gesichtsreproduktion erforderlich ist Anwendung, und hat sich zum wichtigsten bzw. populärsten Teilbereiche der Biometrie entwickelt.

2. Gesichtserkennungssysteme

Moderne Gesichtserkennungssysteme können heutzutage aus reinen Softwarelösungen bestehen, die Bilder durch bestehende Überwachungskameras verarbeiten, bis hin zu vollintegrierten Systemen mit Kamera, Terminal, Back End Geräten, und eigener Verarbeitungssoftware, die für diese Hardware optimiert wurde.

In der Regel sind diese Systeme so entwickelt, daß sie Bilder aus verschiedenen Quellen verarbeiten können. Dazu gehören unter anderem Fotos, Webcams und Überwachungskameras.

Ein Gesichtserkennungssystem besteht in der

Regel aus 2 Komponenten:

Gesichtslokalisierungs-Engine:

Verfahren um Gesichter in einem Bild zu erkennen und zu extrahieren.

Gesichtserkennungs-Engine

Verfahren um Gesichter mit gespeicherten Gesichtstemplates zu vergleichen.

Beide Komponenten bilden eine Einheit und befinden sich meist auf dem selben Geräte bzw. PC. Klassentypen für Gesichtserkennung sind das Alter, das Geschlecht, Gemütszustand und ethnologische Zugehörigkeit. Die Unterschiede werden dabei in Binär, Grau und Farbwertbildern gespeichert.

Einige der Gründe warum die Gesichtserkennung heute so populär ist, sind unter anderem die einfache Integration in schon bestehende Überwachungssysteme, die hohe Akzeptanz, denn an Kameras hat man sich schon gewöhnt, und moderne Systeme können die Identifikation „on the fly“ durchführen. Das bedeutet man braucht lediglich an einer Kamera vorbeizugehen und nicht für eine gewisse Zeit vor dieser stehen um identifiziert zu werden. Auch das gesteigerte Sicherheitsbedürfnis trägt dazu bei das diese Systeme eine immer höhere Akzeptanz, und auch Anwendung finden.

3. Automatisierte Gesichtserkennung

Bei der automatisierten Gesichtserkennung unterscheiden wir zwischen der Identifikation und der Verifikation von Personen.

1:N Identifikation

Das Bild einer Person wird mit einem Datenpool von Gesichtern abgeglichen, und es wird versucht die Person eindeutig zu identifizieren.



Schema der 1:N Identifikation

1:1 Verifikation

Überprüfung ob es sich bei einer Person, wirklich um diese handelt.



Schema der 1:1 Verifikation

Die Gesichtserkennung wird hauptsächlich in Systemen für die 1:N Identifikation eingesetzt. Sehr häufig werden diese Systeme in Kombination mit ID Karten Systemen, in Buchungsstationen oder im Bereich der Überwachung eingesetzt. Dabei bietet sich die Integration dieser Systeme in schon bestehende Umgebungen mit Kameras bzw. Bildaufzeichnungsgeräten an.

Die Gesichtserkennung wird selbstverständlich auch zur 1:1 Verifikation eingesetzt, meist um physischen bzw. logischen Zugang zu autorisieren.

Allerdings gibt es nur sehr wenig wirkliche Umsetzungen bzw. kommerzielle Systeme.

Generell kann man davon sprechen das 1: N Identifikationssysteme, meist in einer kontrollierten und konsistenten Umgebung eingesetzt werden. Der Benutzer steht in der Regel in einem festen Abstand vor der Kamera, mit gleichbleibendem Licht und Hintergrund.

In 1:1 Identifikationssystemen kommen noch Variablen, wie Änderungen bei den Lichtverhältnissen oder Aufnahmewinkel hinzu. Außerdem werden 1:1 System meist nicht beaufsichtigt und müssen somit wesentlich zuverlässiger Arbeiten.

Die wesentlichen Arbeitsschritte der Gesichtserkennung sind folgende:

1. Aufnahme und Lokalisierung des Gesichts
2. Normierung, Modifikation
3. Merkmalsextraktion
4. Abgleich

4. Arbeitsschritte

4.1 Bildaufnahme

Gesichtserkennungssysteme können Bilder aus fast jeder statischen bzw. Videokamera verarbeiten, vorausgesetzt die Qualität und Auflösung ist ausreichend. Im idealsten Fall werden die Bilder durch eine hochauflösende Kamera aufgenommen, es herrschen moderate Lichtverhältnisse und die Person steht direkt vor der Kamera.

Das sogenannte „High Quality Enrollment“, die Bildaufnahme in einer hohen Qualität ist essentiell, da diese Aufnahmen die einzige Grundlage für den späteren Authentifizierungsvorgang sind. Es gibt Systeme denen digitale Bilder mit einer Höhe von 30 Pixel ausreicht. In

der Regel arbeiten die Systeme aber besser wenn sie ein Bild größer als 100x100 Pixel haben

4.2 Bildnormierung

Nachdem das Bild aufgenommen wurde, folgt die Bildverarbeitung. Bilder werden beschnitten und reduziert auf die wesentlichen Bildinformationen, dazu gehört auch die Konvertierung in das Schwarz/Weiß Format. Anschließend werden die Gesichtsaufnahmen normalisiert, so dass sie einheitlich werden in Bezug auf Aufnahmewinkel und Abstand. Um dies zu tun, wird zunächst versucht, charakteristische Gesichtmerkmale zu identifizieren. z.B. die Mitte der Augen. Von diesem Referenzrahmen aus wird das Bild dann gedreht und vergrößert um den minimal benötigten Bildausschnitt zu ermitteln.



Bildaufnahme und Normierung

4.3 Merkmalsextraktion

Wenn das aufgenommene Bild einmal standardisiert wurde, kann der Lokalisierungsprozess der charakteristischen Unterschiede von Gesichtmerkmalen beginnen. Es werden charakteristische Stellen im Gesicht identifiziert, die sich über einen längeren Zeitraum nicht/kaum verändern. Dazu gehören der Bereich um die Wangenknochen, Mundwinkel, Nasenumriss und Augenringe. Zusätzlich werden die Positionen der Hauptmerkmale wie Nase, Augen und Mund zueinander bestimmt. Dabei werden Bereiche, die sehr wahrscheinlich Änderungen unterliegen, z.B. Stirn und Haaransatz ausgespart. Aus diesen Daten der fast unveränderten charakteristischen Stellen im Gesicht wird ein Musterbild des Gesichts reproduziert, das sogenannte Template.



Bestimmung der charakteristischen Merkmale eines Gesichtes

4.4 Matching

Nachdem ein Musterbild hergestellt wurde, wird es mit vorgegebenen Musterbilder aus der Datenbank verglichen, um zu identifizieren bzw. verifizieren. Dieser Vorgang wird als Matching bezeichnet.

Bei der Gesichtererkennung kann generell zwischen zwei verschiedenen Ansätzen unterschieden werden:

1. merkmalsbasierte Gesichtererkennung
Einzelne Merkmale (Features) werden aus dem Gesicht extrahiert, und die Gesichter anhand dieser Merkmale klassifiziert. Beispiele dafür sind das Elastic Bunch Graph Matching oder die „Gesichtserkennung anhand geometrischer Merkmale“.
2. holistischer Ansatz:
Das komplette Gesicht wird betrachtet und zum Matching verwendet. Beispiele dafür sind das Template Matching, die Eigenface-Methode oder die Fourier-Transformation.

In der Praxis kommen oft Kombinationen aus beiden Ansätzen vor.

5. Matching Verfahren

5.1 Template Matching

Das Template Matching wird sehr häufig für die Gesichtserkennung verwendet. Das Verfahren versucht dabei die Ähnlichkeit zwischen einem Bild und einer vorgegebenen Gesichtsmaske zu berechnen. Die Maske besteht aus mehreren markanten Gesichtsausschnitten, den sogenannten Templates. Für die Identifikation wird das gegebene Gesicht mit allen in der Datenbank gespeicherten Gesichtern, genauer gesagt mit den generierten Templates, verglichen. Wenn n Templates für die Gesichtserkennung verwendet werden, erhält man einen n dimensional

Vektor, der die Ähnlichkeit jedes Templates zum getesteten Bild der Datenbank angibt. Die Ähnlichkeit kann zum Beispiel durch die Summe der Fehlerquadrate bestimmt werden. Probleme dieses Verfahrens liegen darin, dass die Maske bei möglichst vielen unterschiedlichen Personen passen muss, außerdem muss die Maske möglichst unabhängig gegenüber Helligkeits- und Kontrastveränderungen sein. Zusätzlich ist das Verfahren sehr rechenaufwändig, da immer mehrere Templates pro Gesicht verglichen werden müssen.



Bild mit 4 Templates für Augen, Nase, Mund und das ganze Gesicht

5.2 Matching geometrischer Merkmale

Bei diesem Verfahren werden geometrische Merkmale eines Gesichts aus dem Bild extrahiert und als Zahlenwerte in einem Vektor gespeichert. Aus n Merkmalen wird ein n dimensionaler Vektor gebildet, der das gegebene Gesicht beschreibt. Dazu müssen wieder Merkmale gewählt werden die sich über einen längeren Zeitraum nicht verändern, außerdem ist darauf zu achten dass kleine Änderungen im Bild, nur kleine Änderungen im Merkmalsvektor verursachen. Um die Merkmalsextraktion zu erleichtern/ beschleunigen werden Gesetzmäßigkeiten, sogenannte Constraints, eines Gesichts ausgenutzt. Zum Beispiel: Augen liegen nebeneinander, Nase liegt unter den Augen. Eine Möglichkeit um Merkmale aus einem Gesicht zu extrahieren, besteht darin die Grauwerte einer Spalte bzw. einer Zeile aufzusummieren, eine Zweite wäre ein Gradientenoperator einzusetzen. Dieser bestimmt die vertikalen und horizontalen Kanten im Bild. Dadurch entsteht ein Umriss des Gesichtes.



Vertikales Aufsummieren

Vertikales und horizontales Kantenbild

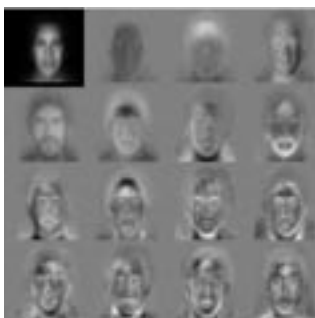
Probleme entstehen bei diesem Verfahren, wenn Gesichter unter verschiedenen Winkeln aufgenommen werden. In diesem Fall muss zuerst eine Bildnormierung durchgeführt werden.

5.3 Eigenface Verfahren

Die Methode der *Eigenfaces* basiert auf dem Verfahren der Karhunen-Loeve-Transformation (KLT), auch Principal-Component-Analysis (PCA) genannt. Grundlage dieses Verfahrens ist die Idee, daß eine Sammlung von Bildern vollständig wiederhergestellt werden kann, indem man die Gemeinsamkeit aller Bilder bestimmt und dann lediglich die Abweichung für jedes Bild speichert.

Dazu werden zweidimensionalen Gesichtsbilder als Vektoren in einem multidimensionalen Raum betrachtet. Da Bilder von Gesichtern eine Menge von Eigenschaften gemeinsam haben, werden sie nicht zufällig in diesem Raum verteilt sein, sondern sammeln sich in einem bestimmten Bereich. Dieser Bereich lässt sich als Subraum mit einer wesentlich geringeren Anzahl an Dimensionen darstellen. Dazu wird aus allen Gesichtsbildern ein Durchschnittsgesicht bestimmt und dieses wird benutzt um die Eigenvektoren, den Unterschied vom Durchschnittsgesicht, jedes Gesichtes zu errechnen. Diese Eigenvektoren bezeichnet man als Eigenfaces und spiegeln die markanten Merkmale der Gesichter wieder. Anhand dieser Vektoren wird dann der Untervektorraum, auch Facespace genannt gebildet in dem alle Bilder liegen.

Ein Gesicht, muss in diesem Untervektorraum liegen, und kann somit als Linearkombination dieser Eigenvektoren dargestellt werden. Dazu wird der Beitrag(Gewichte) der einzelnen *Eigenfaces* zum Eingabebild berechnet. Dieser Vektor wird dann benutzt um festzustellen ob ein Gesicht erkannt ist oder nicht, indem man den Abstand zu den bekannten Gesichtern bestimmt.



Visualisierung von Eigenfaces

5.4 Schwächen von Gesichtserkennungssystemen

Im Vergleich zu anderen biometrischen Erkennungssystemen ist die Treffergenauigkeit beim Gesichtserkennungssystem niedriger. D.h. die Fehlerwahrscheinlichkeit, die False Acceptance Rate (FAR) sowie die False Rejection Rate (FRR), sind wesentlich höher. Die Gründe für die niedrige Treffergenauigkeit sind folgende.

Die Umgebung bzw. Lichtverhältnisse, Aufnahmewinkel und Abstand bei der Bildaufnahme haben unmittelbaren Einfluss auf die Treffergenauigkeit (FRR). Das Gesichtserkennungssystem ist besonders lichtempfindlich. Die Abweichungen gegenüber einer optimale Aufnahmeposition, sollte nicht mehr als 15 Grad in der Horizontale oder Vertikalen betragen. Ein unterschiedlicher Abstand zur Kamera kann die Treffergenauigkeit zusätzlich beeinflussen. Die Änderungen in physiologischen Merkmalen reduzieren auch die Treffergenauigkeit. Die Änderung des Haarstils, Schnur- oder Kinnbart, Make-Up oder eine Brille erhöhen besonders die FRR.

Eine weitere Schwäche ist die sehr hohe Gefahr des Missbrauchs. Die Gesichtsdatabanken können leicht für andere Zwecke verwendet werden. Beispielsweise wurde beim Super Bowl 2001 in Florida ein Gesichtserkennungssystem im Stadion eingeführt und alle Besucher wurden heimlich mit einer Datenbank der Polizei abgeglichen. Ein weiterer Fall ist ein Gesichtserkennungssystem in Uganda. Dieses wurde vergeblich gegen Wahlbetrugs eingebaut, wurde jedoch von der diktatorischen Regierung genutzt, um die Aufsicht über die Wähler zu führen.