



TECHNISCHE UNIVERSITÄT
BERGAKADEMIE FREIBERG

Die Ressourcenuniversität. Seit 1765.

Fakultät für Maschinenbau, Verfahrens- und Energietechnik
Institut für Automatisierungstechnik

Seminararbeit

Moderne Methoden der Bildverarbeitung am Beispiel Gesichtserkennung

Anwendung von Informations- und Automatisierungssystemen

Samuel Dressel und Hannes Krumbiegel

27. Oktober 2019

Betreuer/1. Korrektor:
Prof. Andreas Rehkopf

Inhaltsverzeichnis

1	Einleitung	3
2	Grundlagen	3
3	Gesichtserkennung (Face Detection)	4
3.1	Genetische Algorithmen	4
3.2	Neuronale Netzwerke	4
3.3	Viola-Jones-Algorithmus	4
3.3.1	Vorverarbeitung	4
3.3.2	Haar-like-Features	5
3.3.3	Integralbilder	5
3.3.4	Feature-Auswahl	6
3.3.5	Kaskade	7
3.3.6	Nachverarbeitung	7
3.3.7	Verwendung des Algorithmus unter Nutzung von OpenCV	7
4	Gesichtsidentifizierung (Face Recognition)	8
4.1	Methoden der Gesichtserkennung	9
4.1.1	Hauptkomponentenanalyse (PCA) mit Eigengesichtern	9
4.1.2	Deep Learning	10
4.1.3	Dreidimensionale Gesichtsidentifizierung	11
4.1.4	Gesichtstexturanalyse	11
4.1.5	Wärmebildverfahren	11
4.2	Anwendung	12
4.3	Kritik und Probleme	12
5	Zusammenfassung und Ausblick	13
	Literatur	15

1 Einleitung

Wenn man von Automatisierungssystemen redet, so erstreckt sich der Anwendungsbereich von diesen Systemen längst nicht nur auf den industriell-technischen Sektor. Vielmehr kommen diese Systeme gerade im Zug der Digitalisierung auch in völlig neuen Anwendungsgebieten zum Tragen. Dazu gehören auch Systeme, die es ermöglichen, Gesichter zu erkennen und zu identifizieren. Der Einsatzspektrum dieser Technik reicht dabei vom automatischen Toilettenpapierspender bis hin zum bargeldlosen Bezahlen mittels Gesichts-ID. Grade in den letzten Jahren wurden hier neue Methoden und Technologien entwickelt, welche die Verwendung solcher Systeme immer effizienter und kostengünstiger machen. Dazu gehören vor allem neue Forschungserkenntnisse im Bereich des maschinellen Lernens und die steigende Rechenleistung von Computertechnik. Dies erlaubt es, unterschiedlichste Prozesse in der Industrie und Dienstleistung zu optimieren und wichtige Ressourcen einzusparen.

In dieser Arbeit soll ein grober Überblick über verschiedene Techniken und Methoden der Gesichtserkennung und -identifizierung gegeben werden. Dabei wird in Kapitel 2 zunächst auf einige wichtige Grundlagen eingegangen; in Kapitel 3 und 4 jeweils die zwei Teilbereiche Gesichtserkennung und Gesichtsidentifizierung näher beleuchtet und in Kapitel 5 eine Zusammenfassung gegeben.

2 Grundlagen

Einführend zum besseren Verständnis des Themenkomplexes ist es wichtig, einige wesentliche Begriffe zu definieren und außerdem voneinander abzugrenzen. Dabei gilt es, zwischen der Lokalisation eines Gesichtes an sich und der Zuordnung des Gesichts zu einer bestimmten Person zu unterscheiden. Im ersten Fall wird geprüft, ob und wo ein Gesicht zu sehen ist, im zweiten, um wen es sich handelt [1]. Wir sprechen somit zum einen von der Gesichtserkennung und zum anderen von der Gesichtsidentifizierung. Im englischen Sprachgebrauch unterteilt man den Teil der Gesichtsidentifizierung zusätzlich noch in eine menschliche Erkennung (*face perception*) und in eine maschinelle Erkennung (*face recognition*). In dieser Arbeit wird sich jedoch lediglich auf die maschinelle Gesichtsidentifizierung beschränkt.

3 Gesichtserkennung (Face Detection)

Verschiedene Methoden können verwendet werden, um zu entscheiden, ob ein Bild ein Gesicht enthält. Die Möglichkeiten beinhalten u. a. neuronale Netze, genetische Algorithmen oder speziell für das Problem entwickelte Lösungsansätze. Die Gesichtserkennung ist ein rechnerisch einfacheres Problem als die in Kapitel 4 behandelte Gesichtsidentifizierung; allerdings besteht beispielsweise bei Fotokameras der Wunsch, auf wenig leistungsfähigen Prozessoren die Aufgabe in Echtzeit durchzuführen, um einen schnellen Fokus auf die Gesichter zu ermöglichen. Dieser Anwendungsfall erfordert einen möglichst effizienten Algorithmus.

Der folgende Abschnitt soll einen kurzen Abriss über die verschiedenen Möglichkeiten der Gesichtserkennung geben und exemplarisch einen besonders effizienten Vertreter, den 2001 von Viola und Jones veröffentlichten Viola-Jones-Algorithmus, im Detail vorstellen [2].

3.1 Genetische Algorithmen

Wong, Lam und Sui entwickelten einen genetischen Algorithmus, der mögliche Gesichtskandidaten als Chromosomen auffasst und die Evolution imitiert, um ein Gesicht zu finden [3]. Die Fitness eines Chromosoms beschreibt, wie sehr es einem Gesicht ähnelt. Chromosomen mit einer hohen Fitness behält der Algorithmus („Selektion“) und modifiziert sie durch von der Natur inspirierten Rekombinations- und Mutationsoperatoren. Diese Schritte wiederholt er mehrere Generationen lang, um letztendlich die besten Gesichtskandidaten zu erhalten.

3.2 Neuronale Netzwerke

Neuronale Netzwerke, darunter insbesondere Convolutional Neural Networks, eignen sich gut für viele Aufgaben des maschinellen Sehens. Neben Objekterkennung und Gesichtsidentifizierung befindet sich darunter auch Gesichtserkennung. Valliant et al. entwickelten schon 1994 einen Algorithmus, der Gesichter mithilfe von neuronalen Netzwerken erkennt [4].

3.3 Viola-Jones-Algorithmus

Der Viola-Jones-Algorithmus ist ein häufig zur Gesichtserkennung eingesetzter Algorithmus. Er verarbeitet in Graustufen umgewandelte Bildausschnitte und nutzt sogenannte *Haar-like-Features*, schwarz-weiße Rechtekanordnungen, um in den Ausschnitten Gesichter zu identifizieren. Diesen Prozess optimiert er mithilfe von verschiedenen Techniken, z. B. den des Integralbilds und einer Attention-Kaskade. Das Resultat ist ein effizienter Algorithmus, der sich auch von schwachen Rechnern in Echtzeit berechnen lässt.

3.3.1 Vorverarbeitung

Der Viola-Jones-Algorithmus operiert ausschließlich auf der Grundlage von Graustufenbildern, weshalb der er im ersten Schritt Farbbilder in diese umwandelt.

Der zweite Schritt dient dazu, die Geschwindigkeit des Algorithmus zu erhöhen. Ein Bild, das weniger Pixel enthält, kann schneller verarbeitet werden, weshalb das Herunterskalieren des Bildes die Performance erhöht. Sehr kleine Gesichter können durch die Reduktion der Auflösung verloren gehen, aber in Anwendungsfällen, wo dies akzeptabel ist, ist der Schritt sinnvoll.

Der dritte Schritt besteht darin, iterativ alle möglichen Bildausschnitte auszuwählen, die ein Gesicht enthalten könnten und demzufolge von einem Gesichtserkennungs-Unteralgorithmus darauf überprüft werden sollen. Es ist beispielsweise möglich, dass das gegebene Bild eine

einzigste Großaufnahme eines Gesichts ist, der gesuchte Bildausschnitt betrifft hier das ganze Bild. In anderen Fällen enthält ein Bild viele kleine Gesichter: Um diese zu finden, muss der Algorithmus alle Bildausschnitte, die so groß sind wie die gesuchten Gesichter, überprüfen. Letztendlich überprüft der Algorithmus alle Bildausschnitte vorher festgelegter Größen, die von sehr groß bis klein variieren.

3.3.2 Haar-like-Features

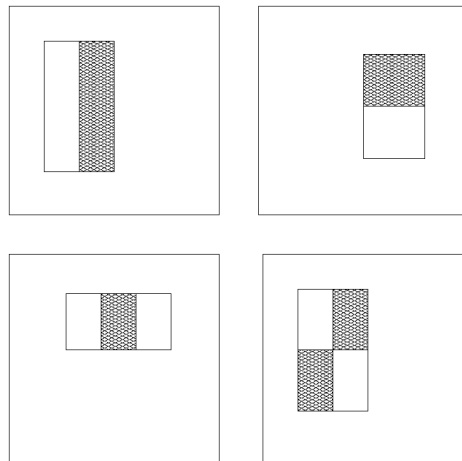


Abb. 1: Mehrere mögliche Haar-like-Features. Quelle: [2]

Ein Haar-like-Feature besteht aus mehreren schwarzen oder weißen Rechtecken und hat das Ziel, markante dunkle bzw. helle Teile eines Gesichts abstrakt abzubilden. Der Wert eines solchen Features beschreibt, wie gut es auf einen Teil eines Graustufenbildes passt.

Ein Feature, das aus einem schwarzen Rechteck über einem weißen Rechteck besteht, wird einen hohen Wert haben, wenn das schwarze Rechteck über die Augenpartie und das weiße über Nase und Wangen gelegt wird (siehe Abbildung 2). Bei Teilen des Bildes, die kein Gesicht enthalten, ist dies selten der Fall. Bei einem Bildausschnitt, bei dem viele - für die Gesichtserkennung ausgewählte - Features hohe Werte aufweisen, handelt es sich wahrscheinlich um ein Gesicht.

Der Wert eines solchen Features berechnet sich aus der Summe der Helligkeitswerte der Pixel unter dem schwarzen Rechteck(en), von denen die Summe der Helligkeitswerte unter den weißen Rechteck(en) abgezogen wird. Ein weißes Pixel hat einen Helligkeitswert von 0, ein schwarzes Pixel entspricht dem Wert 255. Ein Feature erreicht dementsprechend seinen Maximalwert, wenn unter seinen weißen Rechtecken nur weiße Pixel und unter seinen schwarzen Rechtecken nur schwarze Pixel liegen.

3.3.3 Integralbilder

Während des Viola-Jones-Algorithmus müssen die Werte von sehr vielen Haar-like-Features berechnet werden. Eine einfache Implementierung für die Berechnung des Werts eines weißen oder schwarzen Rechtecks bestünde aus einer doppelten For-Schleife, die zwischen den Grenzen der x- bzw. y-Achse iteriert und alle Pixelwerte aufaddiert. Ein Rechteck der Dimension 10x10 würde somit 99 Additionen erfordern.

Viola und Jones entwickelten in [2] das Konzept des Integralbilds, das es ermöglicht, die Anzahl der Rechenoperationen stark zu reduzieren. Das Integralbild ist eine Matrix, die an der

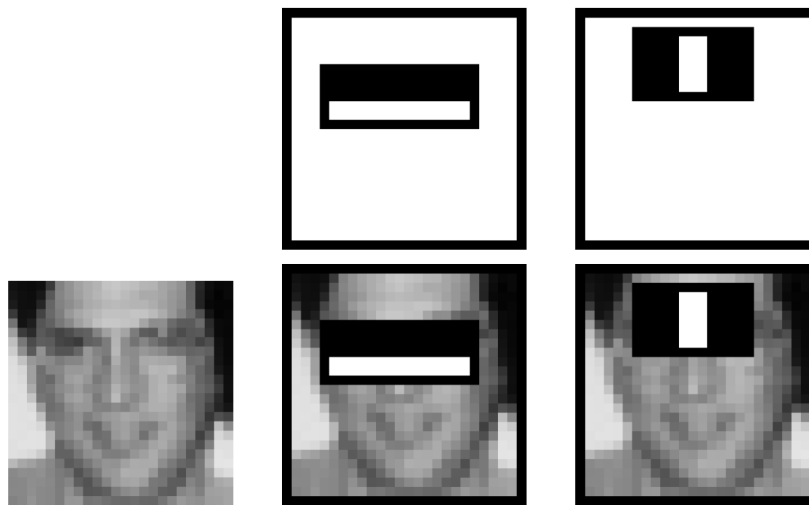


Abb. 2: Zwei Haar-like-Features, die ein Gesicht sehr gut beschreiben. Quelle: [2]

Position jedes Pixels des Bildes nicht den Wert des Pixels, sondern die Summe aller Pixelwerte speichert, die links oberhalb des Pixels liegen. Abbildung 3 verdeutlicht dies: Der Eintrag an Position 1 speichert z. B. die Summe der Pixel der Fläche A.

Das Integralbild ermöglicht es, die Summe der Pixel in einem beliebigen Rechteck nur durch 3 Rechenoperationen zu ermitteln. Die Abbildung verdeutlicht dies: Der Eintrag an Position 4 speichert die Summe der Pixel der Flächen $A + B + C + D$. Wenn wir die Summe der Pixel der Fläche D ermitteln möchten, können wir von dem Wert die Einträge 2 (entspricht d. Summe von $A + B$) und 3 (entspricht d. Summe von $A + C$) subtrahieren. Zuletzt müssen wir lediglich den Flächeninhalt von A - gespeichert in Matrixeintrag 1 - wieder addieren, da er im vorherigen Schritt doppelt subtrahiert wurde.

Der Algorithmus berechnet das Integralbild einmal pro Bild und kann es dann für jedes Haar-like-Feature in jedem Bildausschnitt wiederverwenden.

3.3.4 Feature-Auswahl

Viola und Jones verwenden in ihrer Implementierung Bildausschnitte der Größe 24×24 Pixel [2]. Schon bei dieser relative geringen Auflösung existieren über 160000 verschiedene mögliche Features. Mithilfe einer Gesichtsdatenbank kann ein Brute-Force-Algorithmus bestimmen, welche dieser Features Gesichter am besten von anderen Objekten unterscheiden können. Die Features mit den höchsten Klassifizierungsgenauigkeiten zeigt Abbildung 2. Dennoch haben auch sie eine Fehlerrate von über 10 Prozent. Um hohe Genauigkeiten zu erreichen, müssen mehrere Features kombiniert werden.

Ein einfacher Ansatz wäre es, für die Klassifizierung die Werte der n besten Features für jeden Bildausschnitt zu berechnen. Dies führt jedoch nicht zu optimalen Ergebnissen, da z. B. die Kombination zweier sehr ähnlicher Features, die in den gleichen Fällen ein vorhandenes Gesicht nicht erkennen, zu keiner Verbesserung führt. Viola und Jones haben sich hier für den AdaBoost-Algorithmus für die Featureauswahl entschieden. Dieser wählt Features sukzessive aus und bevorzugt dabei Features, die Gesichter korrekt klassifizieren, welche von den bisher ausgewählten Features falsch klassifiziert werden.

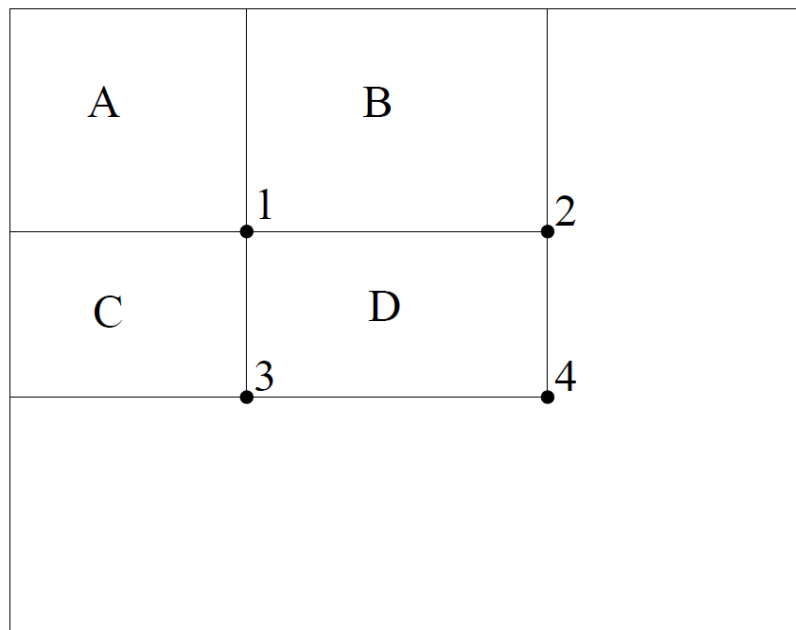


Abb. 3: Eine Illustration eines Integralbildes. Quelle: [2]

3.3.5 Kaskade

Von allen Bildausschnitten, die der Algorithmus auf ein vorhandenes Gesicht prüft, enthalten die meisten keine Gesichter. Um endgültig zu mit hoher Genauigkeit verifizieren, ob ein Ausschnitt ein Gesicht enthält, benötigt der Algorithmus Tausende Features.

Damit er nicht alle Features berechnen muss, um offensichtliche Ausschnitte ohne Gesicht auszusortieren, verwenden Viola und Jones eine sogenannte *Kaskade* [2]. Dieser Ansatz verbessert die Performance erheblich und wird in neueren Arbeiten auch bei der Gesichtserkennung ohne Haar-like-Features verwendet [5].

Eine Kaskade besteht aus mehreren Stufen: Jeder Bildausschnitt muss alle Stufen klassifizieren, um als Gesicht erkannt zu werden. Wenn eine Stufe erkennt, dass ein Ausschnitt sicher kein Gesicht ist, wird dieser aussortiert. Viola und Jones setzten in ihrer Implementierung eine erste Stufe ein, die nur aus zwei Features besteht und 100 Prozent aller Gesichter durchlässt, aber in der Lage ist, 60 Prozent aller Nicht-Gesichter auszusortieren.

3.3.6 Nachverarbeitung

Falls das Bild ein Gesicht enthält, werden einige der Bildausschnitte die Kaskade vollständig durchlaufen und als Gesicht erkannt. Häufig erkennt der Algorithmus mehrere sich stark überschneidende Ausschnitte als zwei separate Gesichter. Der letzte Schritt besteht darin, diese wieder zusammenzuführen.

3.3.7 Verwendung des Algorithmus unter Nutzung von OpenCV

Es ist nicht einfach, einen effizienten Viola-Jones-Algorithmus zu implementieren. Dazu sollte der Programmierer nicht nur den Algorithmus verstehen, sondern auch Technologien wie SIMD oder CUDA. Für Entwickler, die lediglich an den Resultaten des Algorithmus interessiert sind, bietet sich die Verwendung einer schon vorhandenen Implementierung an.

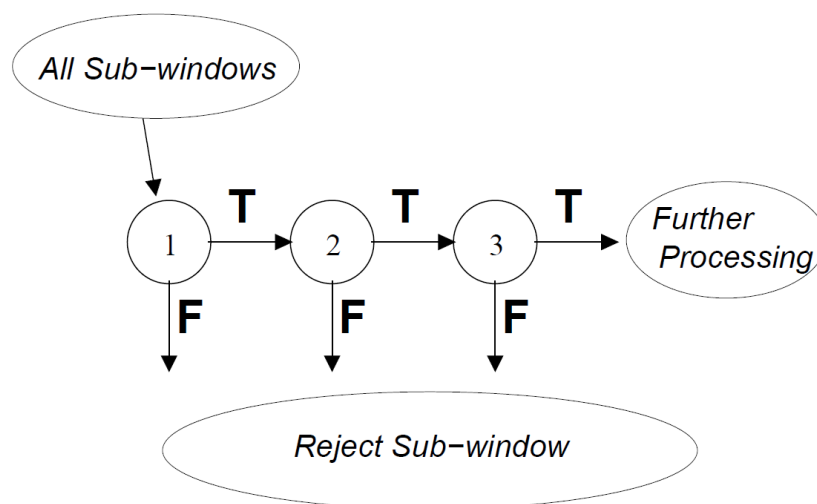


Abb. 4: Eine sogenannte Kaskade, ein Entscheidungsbaum, der Bildausschnitte ohne Gesicht schnell aussortieren soll. Quelle: [2]

Eine Bibliothek, die den Algorithmus implementiert, ist OpenCV [6]. Optional bietet sie CUDA-Unterstützung an, wobei der Großteil des Algorithmus auf der Grafikkarte und nicht auf dem CPU abläuft.

OpenCV erfordert nur wenige Schritte für die Ausführung des Viola-Jones-Algorithmus. Die Bibliothek stellt vorgefertigte Kaskaden in der Form von XML-Dateien zur Verfügung und implementiert den Algorithmus in nur einer Funktion. Der folgende Codeblock zeigt in leicht gekürzter Form die Anwendung der Bibliothek.

```

CascadeClassifier face_cascade; //Initialisiert ueber XML-Datei
Mat picture; //Bild, in dem Gesichter erkannt werden sollen. (Initialisiert in
    nicht gezeigtem Code.)
Mat picture_gray;

cvtColor(picture, picture_gray, COLOR_BGR2GRAY); //Wandle Bild in Graustufen um
std::vector<Rect> faces;
face_cascade.detectMultiScale(frame_gray, faces); //Fuehrt den
    Viola-Jones-Algorithmus aus und speichert die Resultate in der Variable "faces"
  
```

4 Gesichtsidifizierung (Face Recognition)

Das Gebiet der Gesichtsidifizierung umfasst alle Methoden und Techniken, mit der es möglich ist, basierend auf visuellen Material Menschen eindeutig zu identifizieren. Die Gesichtsidifizierung baut dabei auf der Gesichtserkennung auf. Insbesondere in den letzten Jahren wurden nach und nach zahlreichen Algorithmen und Methodiken entwickelt, um Gesichter maschinell zu identifizieren. Insbesondere im Hinblick auf die Automatisierung der Gesichtsidifizierung stellt erst das Jahr 1973 den Startpunkt dar, als Takeo Kanade in seiner Dissertation das erste automatisierte System zur Gesichtsidifizierung vorstellte [7]. Jedoch konnte grade aufgrund der begrenzten Rechenleistung zu dieser Zeit dieses System noch längst nicht effizient arbeiten. Ein wesentlicher Fortschritt wurde erst im Jahr 1990 durch die Forscher Kirby und Sirovich erzielt, welche die Hauptkomponentenanalyse (Karhunen-Loève-Transformation) verwendeten [8]. Hierbei wird eine Vielzahl von generierten Variablen auf

eine geringe Zahl aussagekräftiger Parameter reduziert. Dadurch wird gleichzeitig auch die Komplexität und der Rechenaufwand verringert. Weiterhin stellte die Arbeit von Matthew Turk und Alex Pentland einen weiteren Meilenstein dar, welche sogenannte Eigengesichter (Eigenvektoren) zur Gesichtsidifizierung nutzten [9]. Diese Techniken werden auch noch in den derzeitigen State-of-the-Art-Methoden genutzt.

Im Folgenden soll auf eine Reihe dieser Methoden und Herangehensweisen eingegangen, einige Anwendungsbeispiele genannt und die Problematiken der Gesichtsidifizierung thematisiert werden.

4.1 Methoden der Gesichtserkennung

Grundlegend können alle Methoden der Gesichtsidifizierung in zwei wesentliche Schritte eingeteilt werden. Der erste Schritt ist die Extraktion und Selektion von Merkmalen aus dem visuellen Input; der zweite Schritt stellt die Klassifizierung der Merkmale dar [10]. Desweiteren können Techniken zur Gesichtsidifizierung in die Art der verwendeten Merkmale unterteilt werden. Diese können entweder geometrische Features (Form der Augen, Position der Nase, ...) oder fotometrische Features sein (Texturen, Farbwerte, ...).

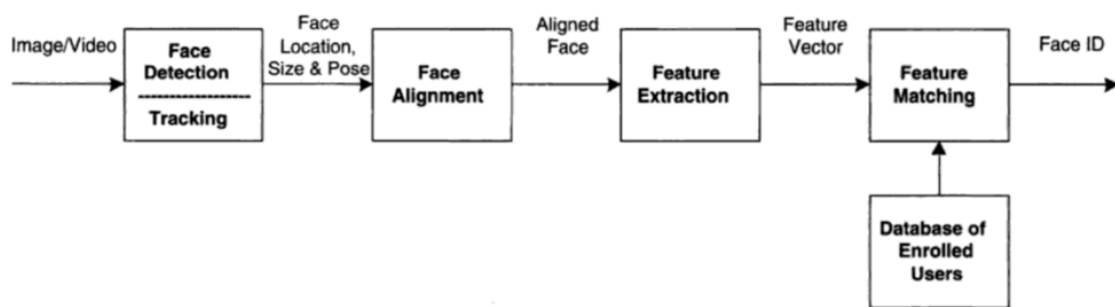


Abb. 5: Verallgemeinerte Darstellung des Gesichtsidifizierungsprozessen als Prozessbild [11]

Die obenstehende Abbildung 5 gibt einen grafischen Überblick über den Prozess der Gesichtsidifizierung. Zunächst wird durch Gesichtserkennung nach einem Gesicht in dem vorhandenen Video- bzw. Bildmaterial gesucht (*Face Detection*). Dann wird das Bild durch bestimmte Bildmanipulationsmethoden in eine zur weiteren Verarbeitung nutzbare Form gebracht (*Face Alignment*). Dies umfasst unter anderem Transformationen (Drehung, Neigung, ...) und die Veränderung von Helligkeits- und Kontrastwerten. In dem folgenden Schritt werden dann die benötigten Features für die jeweilige Gesichtsidifizierungsmethode extrahiert, selektiert und klassifiziert (*Feature Extraction*). Diese werden dann mit einer vorhandenen Datenbank abgeglichen und führen letztendlich zum Identifizierungsergebnis (*Feature Matching*).

4.1.1 Hauptkomponentenanalyse (PCA) mit Eigengesichtern

Das erste Verfahren, welches im Rahmen dieser Ausarbeitung näher betrachtet werden soll, ist die Hauptkomponentenanalyse (Principal Component Analysis, PCA) mit Eigengesichtern. Die Idee der Gesichtsidifizierung mit PCA ist es, vorhandenes Bildmaterial so zu reduzieren, dass es sich für eine Auswertung eignet, ohne dabei die wesentlichen Merkmalsinformationen

zu verlieren [12]. Man spricht hier von einer Dimensionsreduzierung. Die Hauptkomponentenanalyse an sich beinhaltet noch nicht den kompletten Prozess der Gesichtsidentifizierung, sondern stellt vielmehr ein Tool dar, welches von finalen Methoden im ersten Schritt der Selektion genutzt wird. Dies ist vor allem im Hinblick auf die Effizienz dieses Prozesses von großer Wichtigkeit. Um die Funktionsweise der Hauptkomponentenanalyse mit Eigengesich-

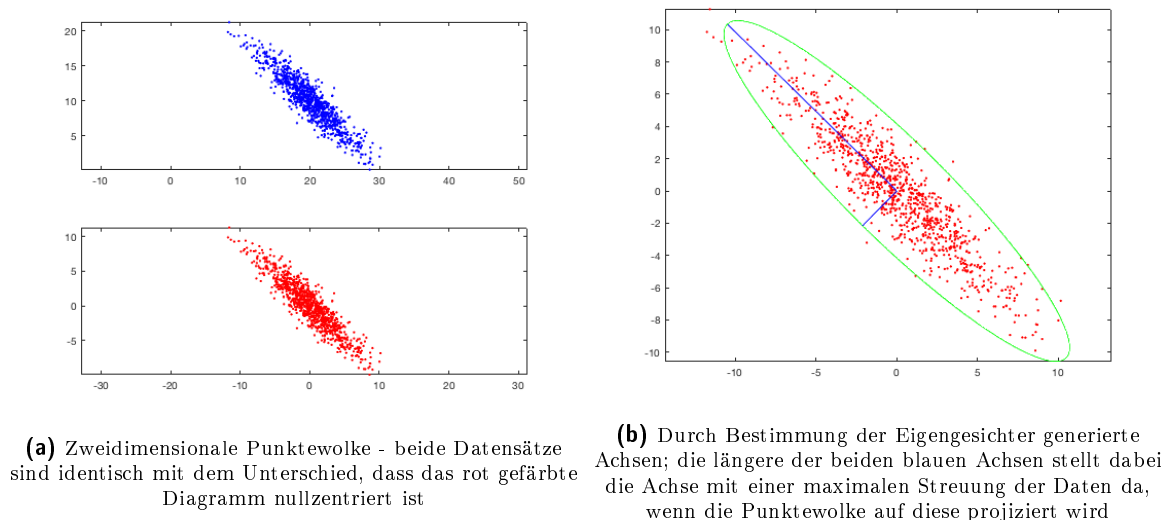


Abb. 6: Reduzierung der Dimension einer zweidimensionalen Punktwolke

tern zu verstehen, kann man sich die einzelnen Merkmale als zweidimensionale Punktwolke vorstellen (Abbildung 6a). Eine Dimensionsreduzierung wird den Datensatz in diesen Fall von einer zweidimensionalen Ebene in eine eine eindimensionale Ebene transformieren [13]. Wie auch bei allen anderen Algorithmen zur Dimensionsreduzierung geschieht dies im Fall der Hauptkomponentenanalyse durch das Finden einer Hyperebene, auf welche die einzelnen Punkte projiziert werden können. Bei der Hauptkomponentenanalyse wird die Hyperebene so gewählt, dass bei einer Projektion alle Punkte eine maximale Streuung erzielen. Es wird sozusagen die Achse der maximalen Varianz gesucht. Im Fall der Punktwolke in Abbildung 6 ist dies die längere blaue Linie. Um diese Linie zu finden, kommen nun die sogenannten Eigengesichter zum Tragen. Diese Eigengesichter sind nichts anderes als die Eigenvektoren der Kovarianzmatrix der gegebenen Datenmenge. Da in dem Fall der Hauptkomponentenanalyse die Achsen der maximalen Varianz generiert werden, bleiben die wichtigsten Informationen aus der Menge von Merkmalen erhalten. Für den zweiten Schritt der Klassifizierung im Rahmen der Gesichtsidentifizierung wirkt sich eine große Streuung zusätzlich positiv aus.

4.1.2 Deep Learning

Die meisten Merkmale, die bei einer Gesichtsidentifizierung durch einen Menschen offensichtlich sind, sind es für einen Computer nicht. Die einfachste Methode, um gewisse Merkmale eines Gesichtes für den Computer als „wichtig“ zu markieren, ist eine eigenständige Abwägung des Computers in Hinblick auf die Wichtigkeit eines Merkmals. Dafür werden meist Methoden des maschinellen Lernens (Deep Learnings) genutzt. Die Methode des maschinellen Lernens kann für eine Vielzahl von Merkmalsformen genutzt werden wie unter anderem Wärmesignaturen, Gesichtsform oder Gesichtstextur [11].

4.1.3 Dreidimensionale Gesichtsidifizierung

Die dreidimensionale Methode der Gesichtsidifizierung nutzt ein gewisses räumliches Sensorsystem, um Informationen zur Form des Kopfes zu generieren [14]. Diese Informationen werden dann genutzt, um bestimmte Merkmale wie die Form und Kontur von Augen, Nase und Mund zur Identifizierung zu benutzen. Der große Vorteil dieser Technik ist, dass die Identifizierungsvorgang auch bei schlechten Lichtverhältnissen durchgeführt werden kann und einen größeren Spielraum hinsichtlich des Aufnahmewinkels besitzt. Der Nachteil dieser Methode ist die Fehleranfälligkeit von durch Mimik gezeigten Emotionen, welche die Form der einzelnen Gesichtsmerkmale verzerren.

4.1.4 Gesichtstexturanalyse

Ein viel genutzter Ansatz zur Gesichtsidifizierung ist die sogenannte Gesichtstexturanalyse (*Skin texture analysis*). Diese Methode extrahiert und nutzt die visuellen Details der Gesichtshaut [15]. Dabei wird ein aufgenommenes Foto zunächst in einzelne Blöcke aufgeteilt. Diese Blöcke werden dann mit diversen Algorithmen in einen mathematischen Raum abgebildet. Diese extrahierten Daten können dann mit einem bestehenden Datensatz verglichen werden. Bei einer hinreichend guten Auflösung kann ein solches Gesichtsidifizierungssystem selbst winzige Poren und Hautlinien erkennen kann. Tests mit diesen Systemen konnten sogar eineiige Zwillinge voneinander unterscheiden, was mit herkömmlichen Systemen meist nicht erreicht werden konnte. Besonders effektiv ist diese Methode der Gesichtsidifizierung in Verbindung mit anderen Methoden der Gesichtsidifizierung.

4.1.5 Wärmebildverfahren

Einen ganz anderen Ansatz als bisher vorgestellte Methoden zur Gesichtsidifizierung bilden die Wärmebildverfahren [16]. Hier kommt der Input für den Identifikationsprozess von Wärmebildkameras. Dabei nehmen diese Kameras nur die Form und die Wärmesignatur des Kopfes auf und ignorieren dabei Accessoires wie Brillen, Kopfbedeckungen oder Makeup. Der Vorteil des Einsatzes einer Wärmebildkamera ist die Tatsache, dass die Bildaufnahme auch bei schlechten Lichtverhältnissen stattfinden kann; Bildaufnahmen bei Nacht inbegriffen. Der Nachteil des Einsatzes von Thermalbildern ist der limitierte Vergleichsdatsatz, der letztendlich für eine Identifikation notwendig ist. 2018 wurde von Forschern des U.S. Army Research Laboratory



Abb. 7: Thermalbild eines Gesichts [16]

eine Technik entwickelt, die einen Vergleich von Wärmebildern mit Bildern einer herkömmlichen Kamera vergleichen kann. Umgesetzt wird dies wiederum durch die Verwendung von

maschinellern Lernen. Dabei wird hier zunächst ein einzelnes Thermalbild synthetisiert indem verschiedene einzelne Gesichtsregionen und darin enthaltene Gesichtsdetails analysiert. Danach kann dieses Bild wiederum mit vorhandenen konventionellen Fotos innerhalb einer Personendatenbank verglichen werden [17].

4.2 Anwendung

Mit den derzeitigen vorhandenen Methoden und einer ausreichend hohen Rechenleistung lässt sich die automatisierte Gesichtsidentifizierung vielfältig nutzen und wird auch mehr und mehr in verschiedensten Anwendungsgebieten eingesetzt. Ein erstes großes Anwendungsgebiet ist sicherlich die Nutzung zu Überwachungszwecken und in Verifizierung von Personen in rechtstaatlichem Kontext. Dies geschieht gerade in Staaten wie China, Russland oder der USA. Als Beispiel sei hier das „Next Generation Identification System“ des FBI genannt, welches mit einer Genauigkeit von 85 % Personen identifizieren kann [18]. Weiterhin setzt China hier unter Verwendung von Künstlicher Intelligenz neue Maßstäbe im Hinblick auf eine totale Überwachung [19]. Beispielsweise werden Bürger, die bei Rot über die Ampel gehen, durch Kameras identifiziert und durch Bildschirme öffentlich als Verkehrssünder angeprangert. Hier werden auch die negativen Aspekte des Einsatzes deutlich.

Weiteren Einsatz findet die Gesichtsidentifizierung auch privaten und finanziellen Sektor. Hier reicht die Einsatzreichweite von dem Entsperren von Türen und Smartphones bis hin zum Bezahlen mithilfe der Gesichts-ID. Einen Meilenstein erzielte Facebook mit der Entwicklung von „Deep Face“, welches die Personenidentifizierung mit einer Genauigkeit von 97 % erlaubt [20]. Um diese Genauigkeit zu erzielen, wurde ein neurales Netz mit neun Schichten und 120 Millionen Gewichtungen benutzt und mithilfe von vier Millionen Facebook-Profilbildern trainiert.

4.3 Kritik und Probleme

Die Thematik der Gesichtsidentifizierung im Vergleich zu einer lediglichen Gesichtserkennung im Sinne der Unterscheidung von anderen Objekten bringt eine Vielzahl von neuen Problemen mit sich. Und dies längst nicht nur auf der technischen Seite, sondern auch im Hinblick auf ethische Aspekte. Betrachtet man die auftretenden Probleme der technischen Umsetzung, so können unterschiedliche Herausforderungen durch die Verwendung von unterschiedlichen Methoden und Technologien gemeistert werden. Jedoch werden bestimmte Schwierigkeiten noch längst nicht vollständig gelöst. Klassische Methoden, welche ihren Input aus klassischem Bildmaterial beziehen, haben mit einer Reihe ganz unterschiedlicher Probleme zu kämpfen [21]:

- Viele zweidimensionale Methoden liefern nur verwendbare Ergebnisse unter ausreichend guten Lichtverhältnissen. Auch ein großer Kontrast bzw. Variation der Beleuchtung kann die Effizienz dieser Methoden erheblich einschränken.
- Die Verdeckung insbesondere der oberen Gesichtsareale durch eventuelle Gadgets wie Brillen, Hüte oder Kapuzen verringert die Genauigkeit oder kann eine Gesichtsidentifizierung nahezu unmöglich machen. Abhilfe schafft hier die Verwendung von Wärmesignaturen zur Gesichtsidentifizierung.
- Klassische Verfahren, die auf die Textur bzw. die Form des Gesichts zur Identifizierung angewiesen sind, funktionieren nur, wenn ein ausreichend optimaler Blickwinkel auf das Gesicht gegeben ist.

- Bei der Identifizierung von Menschen muss der natürliche Alterungsprozess und damit die Veränderung des Gesichts als ein negativer Faktor beachtet werden. Dies erfordert entweder eine aktuelle Vergleichsdatenbank oder das Verwenden einer hinreichend guten KI, um den Alterungsprozess des Gesichts abbilden zu können.
- Ein weiteres Problem im Hinblick auf eine fotometrische Gesichtserkennung sind Emotionen. Diese können ähnlich wie Verdeckungen durch eine damit einhergehende Gesichtsverzerrung die Identifizierung wesentlich ungenauer machen.

Neben den technologischen Herausforderungen der Gesichtsidentifizierung ergeben sich auch eine Reihe von ethischen Problemen bzw. Kontroversen. Viele Menschenrechtsorganisationen und private Kampagnen zeigen immer wieder Zweifel und äußern Kritik, wenn es um den Einsatz dieser Technologien im Überwachungsbereich geht. Diese Bedenken reichen bis hin zum Gegenstand einer der „totalen Überwachungsgesellschaft“, bei dem Staaten und andere Autoritäten die Fähigkeit haben, die Überwachung ihrer Bürger umfassend auszuweiten. Die Verwendung von Identifizierungstools bietet außerdem Möglichkeiten des Missbrauchs im Hinblick auf die digitale Welt (Social Profiling [22]).

5 Zusammenfassung und Ausblick

Die in dieser Arbeit vorgestellten Methoden und Techniken der Gesichtserkennung- und -identifizierung bieten für eine Vielzahl von Problemen in unterschiedlichen Anwendungsbereichen hinreichend gute Lösungen. Die Anwendungsbereiche gehen dabei längst über die klassische Verwendung im Überwachungssektor hinaus. Vielmehr können durch neueste Methoden auch Bereiche wie der Forschungs-, Finanz- und Industriesektor bedient werden. In jedem Fall ist im Hinblick auf die Forschung noch längst nicht die Grenze des Möglichen erreicht. Mit der Entwicklung neuer Technologien in diesem Bereich werden gewisse Nachteile von bestehenden Ansätzen ausgeglichen, jedoch entstehen damit auch gleichzeitig neue Fragen gerade was den ethischen Aspekt von der Gesichtsidentifizierung angeht. Hier gilt es, einen Einsatz immer Hinblick auf die Einschränkung der persönlichen Freiheit abzuwägen. Im Großen und Ganzen überwiegt trotzdem der Vorteil der Nutzung von Systemen zur Gesichtserkennung- und -identifizierung: durch die Verwendung derselben können Zeit und Ressourcen gespart werden. Dafür stehen für verschiedene Einsatzzwecke unterschiedlichen Methoden bereit. Es bleibt nach wie vor spannend, wie die Entwicklung in diesem Bereich in den nächsten Jahren voranschreiten wird.

Literatur

- [1] Gesichtserkennung. <https://de.wikipedia.org/wiki/Gesichtserkennung>. Zuletzt besucht: 23.10.2019.
- [2] Paul Viola; Michael Jones. Rapid object detection using a boosted cascade of simple features, 2001.
- [3] Kwok-Wai Wong, Kin-Man Lam, and Wan-Chi Siu. An efficient algorithm for human face detection and facial feature extraction under different conditions. *Pattern Recognition*, 34(10):1993 – 2004, 2001.
- [4] R. Vaillant. Original approach for the localisation of objects in images. In *Vision, Image, and Signal Processing*, 1994.
- [5] Haoxiang Li, Zhe Lin, Xiaohui Shen, Jonathan Brandt, and Gang Hua. A convolutional neural network cascade for face detection. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
- [6] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [7] Takeo Kanade. *Picture processing system by computer complex and recognition of human faces*. PhD thesis, Department of Information Science, Kyoto University, 1973.
- [8] M. Kirby; L. Sirovich. Application of the karhunen-loueve procedure for the characterization of human faces. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1990.
- [9] M. Turk; A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 1991.
- [10] Facial recognition system. https://en.wikipedia.org/wiki/Facial_recognition_system. Zuletzt besucht: 23.10.2019.
- [11] Oleksii Kharkovyna. An Intro to Deep Learning for Face Recognition. <https://towardsdatascience.com/an-intro-to-deep-learning-for-face-recognition-aa8dfbbc51fb>, June 2019. Zuletzt besucht: 22.10.2019.
- [12] Ana-Maria Sevcenco; Wu-Sheng Lu. Extendend 2-D PCA for Face Recognition: Analysis, Algorithms and Performance Enhancement. In *Face Recognition*. Nova Science Publishers, Inc, 2012.
- [13] Mohit Deshpande. Face Recognition with Eigenfaces. <https://pythonmachinelearning.pro/face-recognition-with-eigenfaces>. Zuletzt besucht: 23.10.2019.
- [14] Song Zhou; Sheng Xiao. 3D face recognition: a survey. *Human-centric Computing and Information Sciences*, 2018.
- [15] A. Sparavigna; R. Marazzato. An image-processing analysis of skin textures. *Skin Research and Technology*, 2010.

- [16] Mrinal Kanti Bhowmik; Kankan Saha; Sharmistha Majumder; Goutam Majumder; Ashim Saha; Aniruddha Nath Sarma; Debotosh Bhattacharjee; Dipak Kumar Basu; Mita Nasipuri. Thermal Infrared Face Recognition – A Biometric Identification Technique for Robust Security system.
- [17] Army develops face recognition technology that works in the dark. <https://www.arl.army.mil/www/default.cfm?article=3199>, April 2018. Zuletzt besucht: 22.10.2019.
- [18] Next Generation Identification (NGI). <https://www.fbi.gov/services/cjis/fingerprints-and-other-biometrics/ngi>. Zuletzt besucht: 23.10.2019.
- [19] Greg Walton. *China's Golden Shield: Corporations and the Development of Surveillance Technology in the People's Republic of China*. Rights and Democracy, 2001.
- [20] Yaniv Taigman; Ming Yang; Marc Aurelio Ranzato; Lior Wolf. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [21] Asit Kumar Datta; Madhura Datta; Pradipta Kumar Banerjee. *Face Detection and Recognition - Theory and Practice*. CRC Press - Taylor & Francis Group, 2016.
- [22] Muhammad Bilal; Abdullah Gani; Muhammad Ikram Ullah Lali; Mohsen Marjani; Nadia Malik. Social Profiling: A Review, Taxonomy, and Challenges. *Cyberpsychology, Behavior and Social Networking*, Vol. 22, No. 7, 2019.