

Prototyp zur automatischen Wiedererkennung von mindestens drei Personen

Deniz Karaayak, Julia Jochum, Pascal Trebert

28.11.2013

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit dem Thema der Gesichtserkennung, im englischen „Face Recognition“. Dabei wird das Gesicht einer Person anhand sichtbarer biometrischer Merkmale, im Bereich des frontalen Kopfes, live von einer Web-Cam analysiert und durch geometrische Anordnung und Textureigenschaften der Oberfläche wieder erkannt. Heutzutage existiert neben dem traditionellen Passwort eine neue Generation von Zugangskontrollen wie Face Recognition, Fingerprint (Fingerabdruck-Scanner) und Venenscanner. Face Recognition wird zum Beispiel zum Entsperren von Mobile Devices benutzt oder in Zugangskontroll-Systemen geschlossener Räumlichkeiten von Unternehmen. Die Grundlagen zu dieser Arbeit basieren im Wesentlichen auf dem ausführlichen Buch [1] *MASTERING OPENCV*

1 Voraussetzung

Aufgrund von Lizenzfreiheit und Betriebssystemunabhängigkeit kommen bei der Umsetzung des System die Programmiersprache Python 2.7.4 mit PyQt, Numpy und der Bildverarbeitungs-Bibliothek PyOpenCV 2.4 zum Einsatz.

Um die Gesichtserkennung ausführen zu können, müssen dem System mindestens drei Personen bekannt sein. Beim Anlegen einer neuen Person, ist es sehr wichtig, dass die ID korrekt ist. Das Programm erkennt nur Personen ohne eine Brille. Um eine gute Erkennung zu starten, wird verlangt, dass mehr als 100 Bilder im Trainingsset angelegt werden.

2 Face detection

Das Gesicht eines Menschen ist eines der wichtigsten biometrischen Merkmale.

Bis zum Jahr 2000 waren die Gesichtserkennungen langsam, sehr unzuverlässig oder beides. Ab 2001 wurde dann die Haarcascade erfunden. Die Haarcascade ist in einer XML Datei hinterlegt, mit Koordinaten von jeweiligen Gesichtspunkt. Mithilfe der Haarcascades werden die Gesichtserkennungen noch verlässlicher dargestellt. Es gibt zu jedem Gesichtsmerkmal ein Haarcascade.

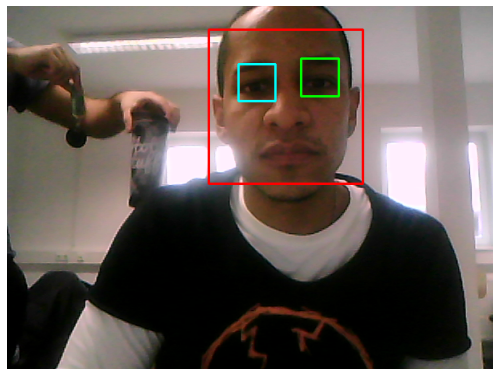


Abbildung 1: Face detection

Die Gesichtserkennung ist in der Lage, Gesichtszüge zu erkennen und alles andere wie Gebäude, Bäume, Gegenstände und Einrichtungen werden ignoriert. Es kann im Allgemeinen als Lokalisierung des Gesichtes betrachtet werden.

Wenn der Benutzer vor der Webcam steht, erkennt das System sofort das Gesicht und die Augen, und zeichnet einen Quadrat um die erkannten Bereiche Gesicht, linkes- und rechtes Auge.

Digitalkameras benutzen Gesichtserkennung, um in einem Motiv das Gesicht zu finden und darauf zu fokussieren. Es verbessert die Bilder im Bereich Personenfotografie. Die Kameras erkennen im Motiv, anhand des Kopfform oder Augenabstand, das Gesicht, und stellt automatisch den Bereich scharf und stellt die Hauttöne richtig dar.

3 Face preprocessing

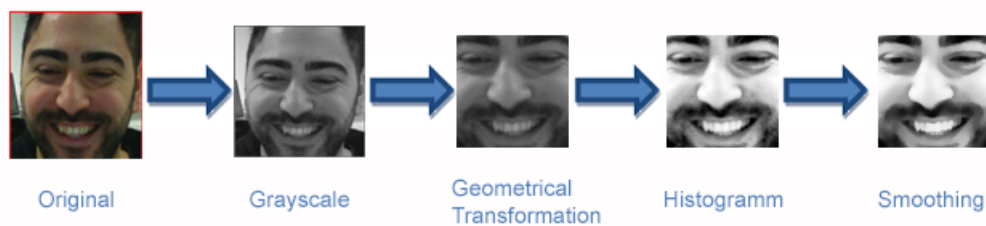


Abbildung 2: Face preprocessing

Gesichter zu vergleichen ist sehr schwierig, weil sich ein Gesicht allein wegen des Lichtverhältnisses unterscheiden kann. Ziel des Face-Preprocessing ist es aufgenommene Bilder aller Gesichter gut vergleichbar abzuspeichern und die Lichteinflüsse zu mildern. Um ein Gesicht rechnerisch vergleichbar zu machen, ist es nötig, einen möglichst hohen Informationsgehalt im Bild zu erhalten und unwichtiges zu entfernen. Demnach besteht der erste Schritt darin, das RGB-Farbbild in ein Graustufenbild zu konvertieren, da Farben sehr von Licht beeinflusst werden.

Nachfolgend wird das Bild affin transformiert. Dabei wird zum einen das Bild ohne Änderung des Seitenverhältnisses auf eine feste Breite skaliert und zum anderen so rotiert, dass die Augen horizontal ausgerichtet sind. Ausserdem wird das Gesichtsbild so geschnitten, dass nur die zentrale Gesichtsmerkmale wie Augen, Nase und Mund bestehen. Stirn, Kinn und Wangen werden ausser Acht gelassen. Anschließend wird zur Reduktion von Lichteinflüssen ein Histogrammausgleich auf die linke- rechte- und den mittleren Teil des Gesichtsbilds angewendet. Dabei wird der mittlere Teil aus dem Histogrammausgleich des linken-, rechten- und gesamt-Bild berechnet. Dieser Art des Histogrammausgleichs erzeugt ein starkes Rauschen und erfordert abschließend einen smoothing Vorgang mit einem Bilateralfilter. Zum Abschluss wird durch ein elliptisches Abschneiden der Seitenränder nicht aussagekräftige Gesichtsmerkmale wie Haare, Bart und Ohren abgeschnitten.

4 Training Set

Die Bilder, die gespeichert werden, besitzen eine Struktur. Es wird für jede einzelne Person, die erkannt werden soll, eine ID angelegt. Die ID wird beim Erzeugen des Bildes, in einem Label-Feld, von Benutzer selbst, angegeben.

Die ID ist dafür da, eine Person fest zuzuordnen. Die Strukturierung beginnt jeweils mit der Ziffer 0, fortlaufend. Jede bekannte Person besitzt dadurch sein eigenes Training-Set.

Was passiert eigentlich, wenn die Person mit den Bildern am nächsten Tag fortfahren möchte?

Der Trainingsset wird ganz einfach fortgesetzt. Wenn die Person mit der "ID = 0" am nächsten Tag seine Bilder fortsetzen möchte, gibt er seine eigene ID ein und die Bilder, die sich in dem Verzeichnis befinden, werden nicht ersetzt, sondern hinzugefügt. Somit erhält man noch mehr Bilder von dieser Person, im Verzeichnis.

Nicht zu vergessen ist, je mehr Bilder im Trainingsset vorhanden, desto präziser kann man das Gesicht erkennen. Jedoch wird hier eine Grenze gesetzt. Es werden 50 Trainingsset-Bilder gespeichert, und diese nochmals als gespiegelt gespeichert werden. So wären dann insgesamt 100 Bilder im Trainingsset vorhanden, was zu einem sehr guten Ergebnis führt.

5 Fisherface

Fisherfaces benutzt Principal Component Analysis (PCA) gekoppelt mit Linear Discriminant Analysis (LDA). Die LDA wurde von Sir R. A. Fisher 1936 erfunden und wurde von Belhumeur, Hespanha und Kriegman [4] für Face Recognition eingesetzt.

Die Principal Component Analysis (PCA) zu Deutsch Hauptkomponentenanalyse findet eine lineare Kombination von Eigenschaften (Hauptkomponenten), welche die Gesamtabweichung in einer Sammlung von Bildern maximiert. Sie fasst eine Sammlung von Bildern in einer Matrix von Eigenvektoren zusammen, die nur aus Hauptkomponenten bestehen, die nicht Null sind.

LDA sucht nach Kombinationen von Eigenschaften die sich innerhalb einer Klasse und zwischen den verschiedenen Klassen befinden. Eine Klasse ist in diesem Fall ein Training-Set beziehungsweise Sammlung von Bildern einer Person. Man erhält zum Schluss (Anzahl Verschiedener Klassen - 1) Eigenvektoren. Zum Beispiel er-

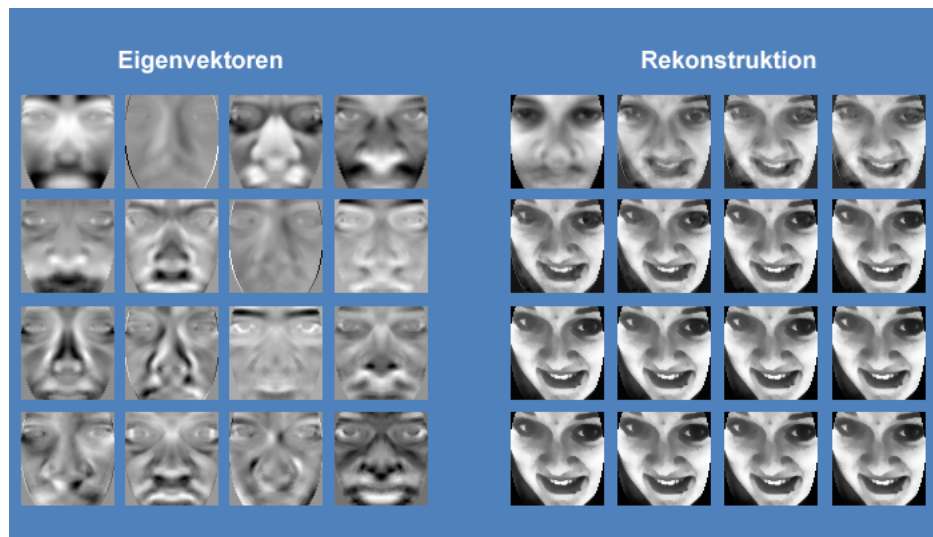


Abbildung 3: Eigenvektoren - die ersten 16 Bilder die entstehen

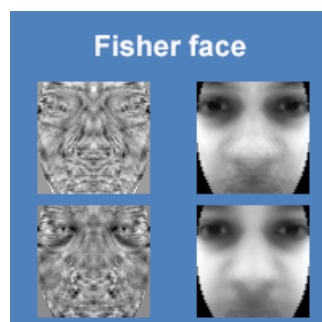


Abbildung 4: Fisher Face

hält man bei drei Training-Sets, zwei Eigenvektoren.

Beim Trainieren des Fisherface-Algorithmus werden die Matrizen von PCA Eigenvektoren und LDA-Eigenvektoren multipliziert und ergeben eine Transformationsmatrix. Anschließend wird jedes Bild aus den Training-Sets in den Fisherface-Unterraum projiziert.

Wenn man wissen möchte zu welcher Klasse ein unbekanntes Gesichtsbild gehört, wird das Bild zuerst in den Fisherface-Unterraum projiziert und anschließend die Distanzen zwischen dem unbekannten Bild und der beim trainieren erlangten Sammlung von Projektionen ermittelt. Die Klasse der Projektion mit der kleinsten Distanz zum unbekannten Bild, ist am wahrscheinlichsten die getestete Person. Die Distanz kann auf verschiedene Weise errechnet werden. Zum Beispiel durch die Euklidische-Distanz, jedoch hat sich die Kosinus-Distanz als präziser erwiesen. [3]

$$\text{Cosinus} := 1 - \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|} \quad (1)$$

Im Gegensatz zu Eigenfaces ist Fisherfaces genauer bei der Gesichtswiedererkennung dafür hat Eigenfaces Vorteile bei Gesichtsrekonstruktion was bei Abbildung 4 zu sehen.

6 Face Recognition

Das Herzstück der Gesichtserkennung, ist die Gesichtswiedererkennung. Es läuft ganz nach dem Motto „Zeig mir dein Gesicht, und ich sag dir, wer du bist“.

Wichtige Basis der Gesichtswiedererkennung sind die Training-Sets da sie essenziell für den Fisherface-Algorithmus sind.

Falls die Person, die erkannt werden soll, in der Datenbank von Training-Sets vorhanden ist, wird das Gesicht erkannt und ihr Name zusammen mit der ID wird angezeigt, wie in Abbildung 5 zu sehen ist.

Ein großes Problem ist, dass bei der Vorhersage, auch eine unbekannte Person, das heißt eine Person die kein Training-Set besitzt, ein Resultat liefert. Das Resultat enthält den Namen und ID, wer am wahrscheinlichsten zu der Person, passen kann. Ein Lösungsansatz ist das unbekannte Gesichtsbild vorher mit einem rekonstruierten Gesichtsbild zu vergleichen, durch Anwendung des L2-Error einen Ähnlichkeitswert ergibt. Eine gute Ähnlichkeit liegt zwischen 0 und 0.5 während eine schlechte zwischen 0.5 und 1. Ist die Ähnlichkeit größer als ein bestimmter Schwellwert, in

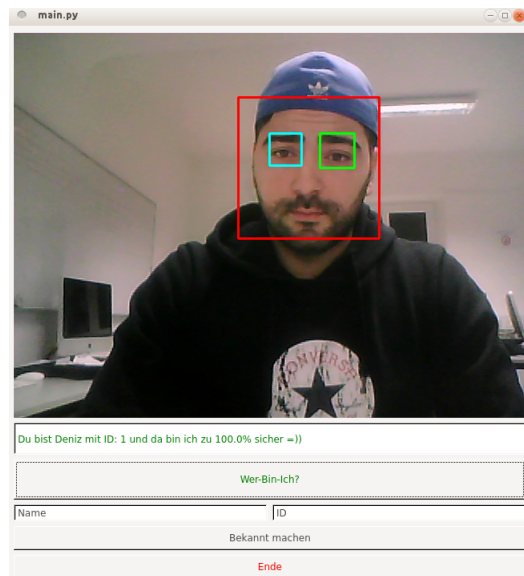


Abbildung 5: Gesichtserkennung

diesem Fall 0.7 wird das unbekannte Gesichtsbild als unbekannte Person angezeigt. Obwohl mit diesem Ansatz und unter Verwendung der Kosinusdistanz nur Werte kleiner als eins als bekannte Person eingestuft werden sollte, ist das Erkennen von unbekannten Personen sehr schwach. Dieser Punkt sollte noch verfeinert werden.

7 Fazit

Es wurden viele verschiedene Ansätze beschrieben, wie die PCA und LDA Verfahren. Fisherface Algorithmus besitzt beide dieser Verfahren, während Eigenface nur das PCA Verfahren benutzt.

Gesichtserkennungen kommen immer mehr in der Vordergrund der Technik. Großkonzerne wie Apple und Samsung, bauen immer mehr auf das Konzept der Gesichtserkennung auf. Nicht nur Mobiltelefone werden mit dieser Technik ausgestattet sondern auch Digitalkameras, um das Gesicht einer Person besser zu fokussieren.

Google vermarktet mit dem Werbespruch “Seien sie kein Fremder, verbinden sie ihre persönlichen Daten mit dem einzigartigen Merkmal, das Sie haben – ihrem Gesicht.”, der wohl ersten Gesichtserkennungs-App für Datenbrillen.

Im Großen und Ganzen, besitzt jeder einzelne Mensch ein Passwort. Ein Passwort, das einmalig ist. Ein Passwort, das nicht kopierbar ist, nämlich das eigene Gesicht!

Literatur

- [1] Lelis Baggio, Daniel...: *Mastering OpenCV with Practical Computer* O'Reilly Verlag, Dezember 2012
- [2] Ranftl, Andreas...: *Digitale Gesichtserkennung* Diplomica@Verlag, Hamburg 2012
- [3] Anonym...: *Learning to Recognize Faces in Realistic Conditions*, 2014
- [4] Belhumeur,P.N...: *Eigenfaces vs. fisherfaces: Recognition using class specific linear projection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997