

Gesichtserkennung

1 Überblick

Bei der biometrischen Gesichtserkennung wird über eine Kamera das Gesicht einer Person aufgenommen und mit einem oder mehreren zuvor gespeicherten Gesichtsbildern verglichen. Dabei wird zunächst das Bild z.B. in einem PC digitalisiert. Die Erkennungssoftware lokalisiert sodann das Gesicht und berechnet seine charakteristischen Eigenschaften. Das Ergebnis dieser Berechnung, das sog. Template, wird mit den Templates der gespeicherten Gesichtsbilder verglichen. Dies gilt nur dann nicht, wenn als Referenzbild das Originalbild verwendet wird, das für den Erkennungsvorgang gegen ein aktuelles Originalbild verglichen wird.

Es gibt unterschiedliche Ansätze der Gesichtserkennung, wobei alle gewisse Schlüsselemente verwenden. Bei den meisten Verfahren der Gesichtserkennung werden die charakteristischen Merkmale der Gesichtszüge anhand eines digitalisierten Bildes bestimmt. Verwendet werden vor allem solche Merkmale des Gesichts, die sich aufgrund der Mimik nicht ständig verändern, also obere Kanten der Augenhöhlen, die Gebiete um die Wangenknochen und die Seitenpartien des Mundes. Grundsätzlich erfolgt ein Vergleich der charakteristischen Gesichtsmerkmale mit den entsprechenden Referenzmerkmalen mittels klassischer Bildverarbeitungs- und Bildanalyseverfahren, wie etwa nach Lokalisierung der Augen die Berechnung der Gesichtsmerkmale anhand eines Gitternetzes, das über das Gesicht gelegt wird. Die Templategröße beträgt bis zu 1300 Bytes. Eine Sondergruppe der biometrischen Gesichtserkennung ist das sog. Eigenface-Verfahren, das vor allem im Bereich der Personenidentifikation verwendet wird. Schließlich existieren erste (Forschungs-)Ansätze einer 3D-Gesichtserkennung.

2 Geschichte

Im Gegensatz zur automatisierten Fingerbildererkennung ist die IT-gestützte Gesichtserkennung eine vergleichsweise junge Wissenschaft. Während erstere auf die mehr als 100 Jahre alten Erkenntnissen der Daktyloskopie zurückgreifen kann, ist die Geschichte der biometrischen Gesichtserkennungsalgorithmen gerade einmal etwas mehr als 10 Jahre alt. Dabei wurde der Stand der Entwicklungen bereits frühzeitig und inzwischen relativ häufig durch Tests überprüft.

So testete zunächst das US-amerikanische "Department of Defense" (DoD) in seinem "Face Recognition Technology program" (FERET) 1994, 1995 und 1996 die damals noch sehr jungen Gesichtserkennungssysteme. In den Jahren 2000, 2002 erfolgte dann eine Neuauflage und Erweiterung der Tests unter dem Namen "Facial Recognition Vendor Test" (FRVT). Im vergangenen Jahr startete das amerikanische "National Institute for Standardization" (NIST) den "Facial Recognition Grand Challenge" (FRGC) der wiederum die Testszenarien des FRVT erweitert. In Deutschland werden Gesichtserkennungsalgorithmen seit dem Jahr 2002 durch das BSI im Rahmen der Projekte BIOFACE und BioP I und II untersucht. Abgesehen davon wurden Gesichtserkennungssysteme in den vergangenen Jahren weltweit in zahlreichen Praxiserprobungen auf ihre Leistungsfähigkeit geprüft.

Inzwischen haben viele Systeme den Schritt von der technischen Versuchsversion in den echten Einsatz geschafft: Sie werden vorzugsweise als Zutrittskontrollsysteme für

Firmenmitarbeiter und Ausstellungsbesucher, sowie als Zugangüberwachung in Spielkasinos verwendet. Besonders vorteilhaft im Vergleich zu anderen biometrischen Verfahren ist der Einsatz von Gesichtserkennungssystemen vor allem deshalb, weil das Gesicht mit wenigen Ausnahmen öffentlich "zugänglich" und einfach zu fotografieren ist, wobei man allerdings in Kauf nehmen muss, dass sich das Objekt "Gesicht" dabei frei im Raum bewegen kann und die Bildverarbeitung dadurch schwieriger ist, als beispielsweise bei der Finger- oder Iriserkennung.

3 Allgemeine Vorgehensweise

Ähnlich der meisten anderen biometrischen Verfahren ist die Gesichtserkennung bzw. der Gesichtsbildvergleich in mehrere Arbeitsschritte unterteilt. Dabei kann man die drei Hauptarbeitsschritte "Template erzeugen", "Referenzdatensatz erzeugen" und "Gesichtsbilder vergleichen" unterscheiden.

3.1 Template erzeugen

Um den Vergleich zweier Gesichtsbilder möglichst einfach und schnell zu halten, werden zunächst die Merkmale eines Gesichts ermittelt und in Form eines Merkmalsdatensatzes, dem sogenannten "Template", gespeichert.

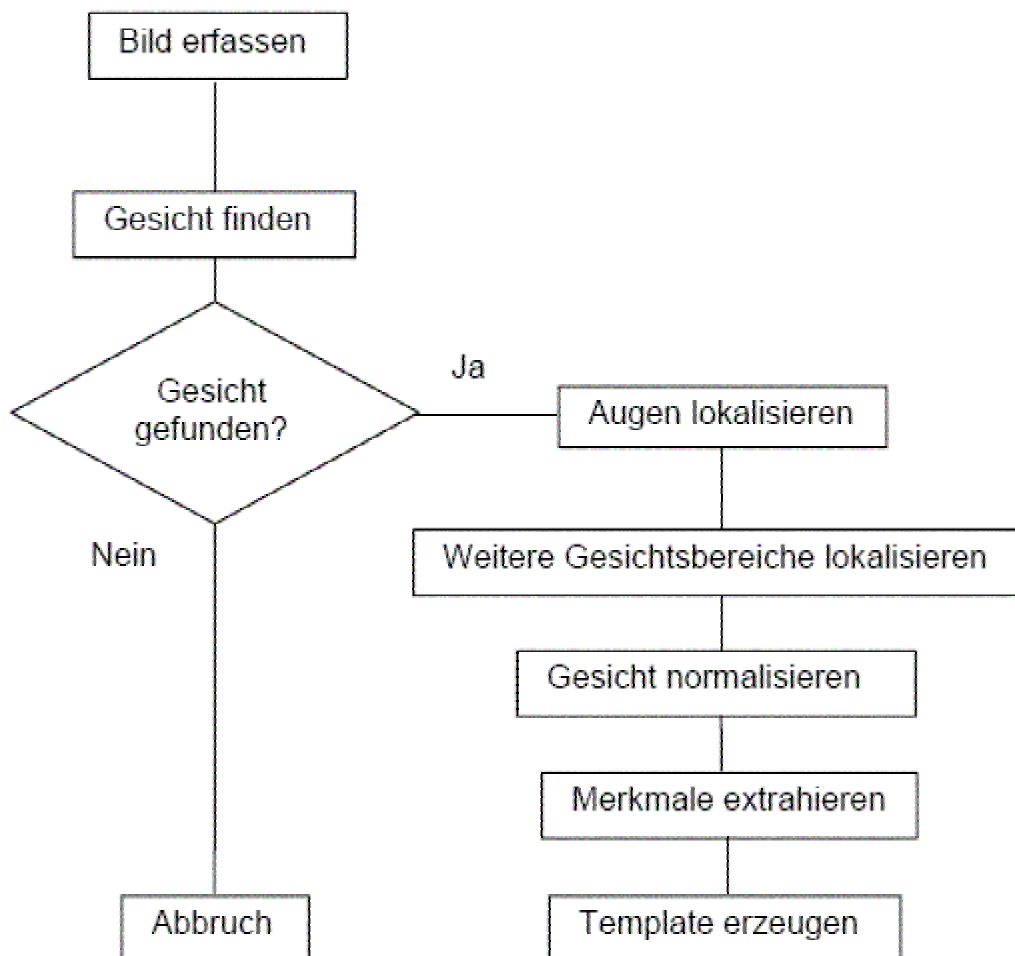


Abbildung 1 zeigt die Erzeugung des Templates¹

¹ Quelle: © Bundesamt für Sicherheit in der Informationstechnik

Die Arbeitsschritte beinhalten folgende Aktionen:

1. "Bild erfassen":
Das Bild des Gesichts einer Person wird mittels einer Kamera im aktuellen Umfeld aufgenommen ("Life-Bild") oder in Form eines Scans eines bereits vorhandenen Bildes der Person z.B. vom Zutrittsausweis o.ä.
2. "Gesicht finden":
Der Bildbereich wird daraufhin untersucht, ob sich eine gesichtsähnliche Form vor dem Bildhintergrund abhebt.
3. "Gesicht gefunden?":
Wenn kein Gesicht (bzw. kein gesichtsähnlicher Bereich) gefunden werden konnte, wird die weitere Verarbeitung des Bildes abgebrochen: Das Bild ist für die Gesichtserkennung unbrauchbar.
4. "Augen lokalisieren":
Die Augen sind typischerweise im Vergleich zu Haut extrem dunkle Punkte in der oberen Hälfte des Bildes.
5. "Weitere Gesichtsbereiche lokalisieren":
Von der Augenposition aus werden alle weiteren typischen Punkte des Gesichts (Nase, Mund-/Kinnpartie, äußerer Rand des Gesichts) gesucht. Das bedeutet: Wird die Augenposition nicht korrekt gefunden (z.B. bedingt durch eine extreme Seitenneigung des Gesichts oder Überschattung durch Hutkrempe o.ä.), scheitert mitunter die Extraktion der Gesichtsmarkmal.
6. "Gesicht normalisieren":
Das Gesichtsbild wird durch Drehung und/oder Streckung bzw. Stauchung auf nahezu einheitliche Maße gebracht, so dass die Augenposition auf allen zu verarbeitenden Gesichtsbildern nahezu im gleichen Bildbereich liegt.
7. "Merkmale extrahieren":
Weitere Gesichtsmarkmal werden gesucht und je nach verwendetem Verfahren vermessen und ggf. zueinander in Bezug gesetzt.
8. "Template erzeugen":
Mittels einer mathematischen Formel werden die Merkmalsdaten codiert und komprimiert, so dass schließlich ein Merkmalsdatensatz von 1000 bis 1300 Byte entsteht.

3.2 Referenzdatensatz erzeugen

Im Bereich der Personenidentifikation wird das Template eines ein Life-Bildes (s.o.) immer mit den Templates aus einer Referenzdatenbank verglichen. Die Referenzdatenbank entsteht durch das sog. "Einlernen" ("Enrolment") von Gesichtsbildern in das Gesichtserkennungssystem:

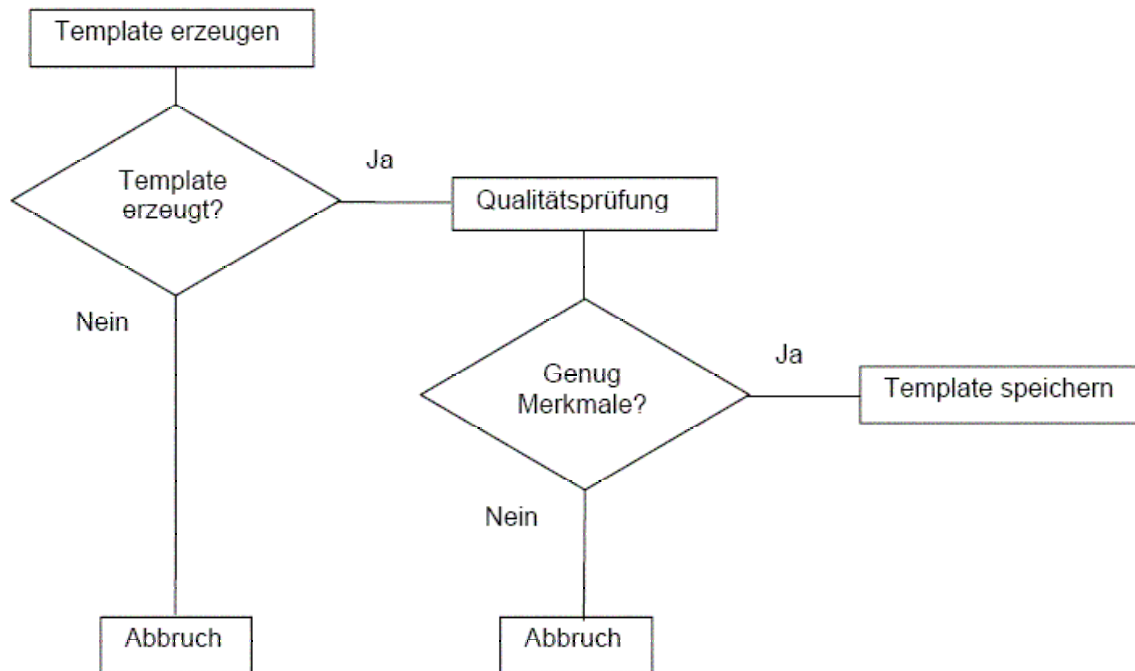


Abbildung 2 zeigt die Fortsetzung der Erzeugung des Templates²

1. "Template erzeugen":
2. "Template erzeugt":
Konnte kein Template erzeugt werden, wird der Einlernprozess abgebrochen: Das Bild ist als Referenzbild unbrauchbar.
3. "Qualitätsprüfung"/"Genug Merkmale?":
Um bei einem späteren Identifikationsvorgang ein qualitativ verlässliches Vergleichsergebnis erreichen zu können (d.h. das System soll zweifelsfrei entscheiden können, ob ein anderes Gesichtsbild dem zum Referenztemplate gehörenden Gesichtsbild gleich, oder nicht), muss das Referenzbild so viele Gesichtsmerkmale aufweisen wie möglich, damit das Template so viele Informationen wie möglich über die Eigenschaft des Gesichts besitzt. Die Toleranzschwelle, ab wann ein Bild als ausreichend erkannt wird, wird von System zu System unterschiedlich eingestellt.
4. "Template speichern":
Das Template wird als Datensatz in der Datenbank o.ä. gespeichert.

Je nach Verfahren werden zur Erstellung des Referenztemplates auch mehrere Gesichtsbilder einer Person verwendet (z.B. indem aus einer Videosequenz mehrere Einzelbilder verwendet werden), um z.B. veränderte Kopfhaltungen, Mundöffnungen o.ä. zu berücksichtigen.

² Quelle: © Bundesamt für Sicherheit in der Informationstechnik

3.3 Vergleich

Um die Ähnlichkeit der Templates zweier Gesichtsbilder zu bestimmen, werden sie byteweise mittels eines mathematischen Algorithmus (z.B. in Form von Vektoroperationen) kombiniert. Die Höhe des mathematischen Resultats repräsentiert den Grad der Ähnlichkeit der Templates. Wenn das Resultat innerhalb gewisser Toleranzgrenzen liegt, werden die beiden Templates, und damit die ihnen zugrunde liegenden Gesichtsbilder identisch eingestuft.

4 Verwendete Verfahren

Den verschiedenen Gesichtserkennungssysteme liegen unterschiedliche Verfahren aus dem Bereich der Mustererkennung bzw. deren Kombinationen zugrunde:

4.1 Template Matching

Beim Template-Matching werden Gesichtsregionen vorgegeben (z.B. Augenpartie, Nasenpartie, Mundpartie) die im zu verarbeitenden Gesichtsbild gesucht werden. Daraufhin werden die grafischen Merkmale dieser Bereiche mathematisch mit denen der entsprechenden Bereichen des Referenzbildes kombiniert und daraus eine Ähnlichkeit der Bereiche berechnet.

4.2 Elastische Graphen

Mittels einer grafischen Analyse des Bildes werden markante Stellen (sog. "Knoten") im Gesicht (Augen, Nasenspitze, Kinnschuppe, Haaransatz, Schläfen, etc.) gesucht und über Linien zu einem Gittermodell verbunden. Mithilfe des Gittermodells eines normierten Gesichts wird das zu verarbeitende Gesichtsbild in eine Standardposition (frontale Ansicht) gedreht (Normalisierung). Beim Vergleich zweier Gesichtsbilder wird durch Verschiebung und/oder Streckung/Stauchung der Gittermodelle der Bilder versucht, diese aufeinander abzubilden. Die verbleibenden Unterschiede in der Lage gleichbedeutender Knoten der beiden Bilder (z.B. Nasenspitze, Augenpositionen) sind dabei das Maß für die Ähnlichkeit der beiden Gesichter.

4.3 Geometrische Merkmale

Ähnlich zum Verfahren der Elastischen Graphen werden die markanten Stellen des Gesichts ermittelt. Ihre Lage, d.h. ihre jeweilige relative Position zueinander bildet dabei einen mathematischen Vektor. Zwischen diesem und dem des jeweiligen Referenzbildes kann dann im Vergleichsvorgang der Abstand, d.h. die Ähnlichkeit, bestimmt werden.

4.4 Eigenfaces

Die Grundlage dieses Verfahrens ist eine Sammlung von sog. "Basisgesichtern" die so kombiniert werden, dass sie dem zu verarbeitende Bild (Referenz- oder Life-Bild) so ähnlich wie möglich sind. Die Koeffizienten, die zu der passendsten Kombination geführt haben, bilden schließlich das Template.

5 Ausblick

Die bisherigen Verfahren beruhen auf einem in der Regel zweidimensionalen Gesichtsbild. Dies führt dann zu Problemen beim Erkennungsvorgang, wenn die Position des Gesichts zur Kamera von einer Norm (z.B. Frontal) abweicht, da der Grad der Abweichung gar nicht oder nur annähernd ermittelt werden kann.

Mittels spezieller Kameras und Aufnahmetechniken soll in zukünftigen Generationen daher ein dreidimensionales Gesichtsbild aufgenommen werden, so dass die Orientierung des Gesicht zweifelsfrei festgestellt werden und das Gesicht in eine normierte Position gedreht werden kann. Damit soll ausgeschlossen werden, dass signifikante Unterschiede beim Vergleich zweier Gesichtsbilder nicht lediglich aufgrund unterschiedlicher Orientierung des Kopfes ein und derselben Person entstehen.

Das BSI führt derzeit einen Test von 3D-Gesichtserkennungssystemen durch, der voraussichtlich Ende 2005 abgeschlossen sein wird.

Weitere Aussagen zur Erkennungsleistung und Sicherheit von marktverfügbaren Gesichtserkennungsverfahren lassen sich dem öffentlichen Abschlussbericht des Projekts BioP II entnehmen, der demnächst auf der BSI-Webseite zum download zur Verfügung stehen wird.