

Optische Zutrittskontrolle mit ZN-Face

Ein System zur automatischen Gesichtserkennung

Dr. Wolfgang Konen, ZN GmbH¹

Die Gesichtserkennung ist ein bemerkenswertes Beispiel für die Leistungsfähigkeit und Robustheit menschlicher visueller Objekterkennung. Wir können Tausende von Gesichtern, die wir im Laufe unseres Lebens erlernen, praktisch "auf einen Blick" wiedererkennen. Dabei ist diese visuelle Leistung sehr robust gegen Störungen und Variationen, wie beispielsweise Mimik, Blickwinkel, Beleuchtung, wechselnden Hintergrund, Größe, Alterung und teilweise Verdeckungen. Es gehört zu den Herausforderungen künstlicher Informationsverarbeitung, diese Erkennungsleistung zumindest in Ansätzen nachzubilden. Dabei zeigt uns die Perfektion, mit der dieses Problem vom menschlichen Gehirn, also einem biologischen neuronalen System, gelöst wird, daß offensichtlich neuronale Verarbeitungsprinzipien sehr wichtig zur Lösung dieser Aufgabe sind.

Bis heute ist es nicht möglich, ein rechnergestütztes System zur Gesichtserkennung zu konstruieren, das der menschlichen Gesichtserkennung in Bezug auf Leistungsfähigkeit und Schnelligkeit ebenbürtig ist. Jedoch wird auf diesem Feld intensive Forschung betrieben, und besonders durch den Einsatz von Methoden der Neuronalen Netze und der massiv-parallelen Informationsverarbeitung, wie sie ja auch im menschlichen visuellen System stattfindet, sind in diesem Gebiet in den letzten Jahren erhebliche Fortschritte erzielt worden, die unter standardisierten Umfeldbedingungen eine zuverlässige Gesichtserkennung erlauben. Durch diese Fortschritte wird es heute erstmalig möglich, die Ergebnisse dieser Forschung in Anwendungen auf dem Markt zu integrieren und dadurch neue und innovative Marktpotentiale zu erschließen.

Zutrittskontrollsystem *ZN-Face*

Eine automatisierte Lösung des Problems der Gesichtserkennung ist von sehr großer praktischer Bedeutung: Beispiele für mögliche Anwendungen sind Zutrittskontroll- und Identifikationssysteme, die durch eine zusätzliche Gesichtserkennung wesentlich höhere

¹Dr. Wolfgang Konen arbeitet am Zentrum für Neuroinformatik GmbH, Bochum, und leitet dort die Produktentwicklung des Systems *ZN-Face*.

Sicherheit erlangen, da ein sehr personenspezifisches Merkmal abgefragt wird. Eine andere denkbare Anwendung ist das automatisierte Durchsuchen von Datenbanken nach ähnlichen Gesichtern oder anderen als Bilder gespeicherten Objekten.

Im Bereich der Zutrittskontrollsysteme hat das Zentrum für Neuroinformatik nun das Zutrittskontrollsystem *ZN-Face* entwickelt, das erstmalig die automatische Gesichtserkennung einsetzt und auf neuartigen neuronalen Verfahren zur Bildverarbeitung basiert. Die Benutzung ist denkbar einfach: Am Zutrittskontrollpunkt tritt die Person vor eine Konsole (s. Abb. 1) mit eingebauter Kamera und identifiziert sich durch Geheimnummer oder Magnetkarte. *ZN-Face* vergleicht das aufgenommene Bild des Gesichtes mit den zu dieser Person gespeicherten Merkmalen. Nur wenn beide Kriterien, Bild und Geheimnummer, übereinstimmen, gewährt *ZN-Face* Zutritt. Damit ist ausgeschlossen, daß z. B. der Diebstahl einer Magnetkarte einem Unberechtigten Zutritt ermöglicht. Der Bildvergleich dauert ca. 3 Sekunden (Pentium-PC), unabhängig von der Größe der Datenbasis.

Das System *ZN-Face* stellt keine besonderen Anforderung an die Raumbeleuchtung, ist jedoch konzipiert für den Einsatz in Innenräumen; im Außenbereich können durch direktes Sonnenlicht zu große Beleuchtungsunterschiede auftreten.

Marktchancen

Das Marktpotential für das Zutrittskontrollsystem *ZN-Face* liegt in all jenen Bereichen, die eine besondere Sicherheit erfordern: Sensible Bereiche in Firmen und Banken, wie beispielsweise deren Rechenzentren; zusätzliche Absicherung von SB-Schließfachanlagen in Banken; Hochsicherheitsbereiche in Kernkraftwerken u.v.a.m.

Für den Sicherheitsmarkt besonders interessant ist, daß mit der Gesichtserkennung *ZN-Face* eine neue Methode der biometrischen Kontrolle verfügbar wird, die für den Benutzer besonders einfach handhabbar ist. Deshalb sind geringere Akzeptanzprobleme zu erwarten als bei anderen biometrischen Methoden, wie Fingerabdruck- oder Augenhintergrund-Erkennung.

Das System *ZN-Face* wurde auf den Messen VISION'94 und SECURITY'94 erstmalig einer breiten Öffentlichkeit als voll funktionsfähiger Prototyp präsentiert. Die Messebesucher hatten dabei die Gelegenheit, *ZN-Face* "live" zu testen: Nach Speicherung eines Referenzbildes konnte man sich beispielsweise davon überzeugen, daß auch bei anderer Mimik eine Wiedererkennung sichergestellt ist. Erstinstallationen des Produktes sind für das Quartal I/95 geplant.

Automatischer Gesichtsvergleich

Wie funktioniert nun der Gesichtsvergleich per Computer im einzelnen? Das hier beschriebene Gesichtserkennungs-System *ZN-Face* zeichnet sich dadurch aus, daß verschiedene für die Gesichtserkennung typischen Teilprobleme, nämlich die Lokalisation des Gesichtes in einer Szene, die Trennung von Vordergrund- (Gesichts-) und Hintergrundmerkmalen und eine mögliche Änderung in der Größe des Gesichtes durch *ein* kohärentes Verfahren behandelt werden können. Das Verfahren ist eine Weiterentwicklung des in [1] beschriebenen Elastic-Graph-Matching-Verfahrens.

Die Weiterentwicklung des Systems *ZN-Face*, das am Zentrum für Neuroinformatik zur Produktreife als Identifikationssystem gebracht wurde, beinhaltet unter anderem aus

folgenden Aspekten:

- Turn-Key-Lösung als Gesamtsystem bestehend aus Aufnahme- und Verifikationskonsole mit integrierter Kamera und Rechereinheit.
- Drastische Beschleunigung des Verfahrens: Während frühere Implementierungen 20 sec Rechenzeit auf einem Parallelrechner mit 20 T800-Transputern erforderten, benötigt die jetzige Implementierung auf einem Pentium-PC nur 3 sec. Damit ist eine wichtige Voraussetzung für die wirtschaftliche Anwendbarkeit des Verfahrens erfüllt.
- Automatische Anpassung an unterschiedliche Objektgrößen: Selbst wenn Gesichter mit bis zu etwa 30% abweichender Größe gegenüber der gespeicherten Vorlage präsentiert werden, ist eine verlässliche Erkennung möglich.
- Verbesserte Erkennungsleistung durch Einsatz eines Neuronalen Netzes bei der Signifikanzbewertung (s.u.).

Beim Verfahren des Elastic-Graph-Matching werden Gesichter durch flexible Graphen oder Gitter (s. Abb. 2) gespeichert, indem charakteristische visuelle Merkmale an den Knoten des Graphen angeheftet sind, während der Graph selber in seiner Struktur den geometrischen Zusammenhang dieser Merkmale speichert. Die visuellen Merkmale an einem Knoten charakterisieren die betreffende Stelle des Bildes inklusive ihrer näheren Umgebung (die etwa bis zu den benachbarten Knoten reicht). Eine solche "Arbeitsteilung", die das Bild in sich überlappende räumliche Regionen aufteilt, ist typisch für neuronale Verarbeitungsmechanismen (in biologischen visuellen Systemen spricht man von *rezeptiven Feldern*).

Die Vorteile einer solchen Datenrepräsentation liegen auf der Hand:

- Durch die Beschränkung auf wenige (hier: 70) Graphknoten anstelle der wesentlich zahlreicheren Bildpixel wird eine enorme Datenkompression erzielt. Ein Beispiel: Wenn an jedem Graphknoten 20 verschiedene Merkmale mit je 1 Byte Auflösung gespeichert sind, so bedeutet dies 1.4 kB zur Speicherung eines Graphen gegenüber 16 kB zur Speicherung eines Grauwertbildes mit 128×128 Pixeln.
- Ein spärlicher Graph kann mit sehr geringerem Rechenaufwand an eine veränderte Geometrie (Größe, Perspektive) angepaßt werden, während dies bei einem Grauwertbild eine umfangreiche Transformation erfordert.
- Der Graph charakterisiert die im Gesicht enthaltene Information ganzheitlich und besitzt damit Redundanz gegenüber lokal begrenzten Veränderung wie z. B. dem Tragen einer Brille oder leicht veränderter Frisur. Das weiter unten beschriebene Matching-Verfahren zwischen Graphen ist nicht auf die Lokalisation bestimmter Schlüsselmerkmale angewiesen und arbeitet in der Regel auch noch dann, wenn einzelne Teilbereiche des Gesichtes verdeckt sind.

Die Datenrepräsentation beschreibt die Speicherung der zu erkennenden Gesichter. Es bleibt das Problem zu lösen, wie ein aktuelles Bild mit der gespeicherten Information abgeglichen wird (Matching). Die Grundidee des Matching-Verfahrens ist wie folgt: Soll ein Gesicht wiedererkannt werden, so wird der betreffende Graph aus dem Speicher geholt und an einer zunächst beliebigen Stelle im Bild plazierte. Die gespeicherten Merkmale jedes Knoten werden mit den aktuellen Merkmalen im Bild verglichen und so ein Maß für die Güte des Matches berechnet. Ein Optimierungsverfahren gestattet es nun, durch schrittweise Veränderungen von Lage, Größe und Form des Graphen, die Match-Güte zu verbessern, bis eine optimale Positionierung gefunden wird. Damit ist das Gesicht im Bild lokalisiert, vom Hintergrund abgetrennt und in seiner Größe bestimmt. Der so gefundene Graph im Bild wird dann in einem zweiten Schritt mit gespeicherten Graphen verglichen und anschließend wird eine Signifikanzbewertung durchgeführt.

Erkennungsleistung

Die optimale Erkennungsleistung eines Identifikationssystems ist standardmäßig dadurch definiert, daß die Summe aus fälschlichen Zurückweisungen (*false rejection rate, FRR*) und fälschlichen Akzeptanzen (*false acceptance rate, FAR*) minimal wird. Verschiedene Anwendungen können jedoch eine unterschiedliche Ausbalancierung beider Effekte erforderlich machen; dies kann in der Konfiguration von *ZN-Face* durch eine einstellbare Akzeptanzschwelle berücksichtigt werden, die beispielsweise die FAR reduziert, in der Regel jedoch auf Kosten einer erhöhten FRR (s. Abb. 3).

Abb. 3 zeigt die Ergebnisse der neuronal trainierten Signifikanzbewertung bei Anwendung auf das Datenmaterial aus [1], das von 87 Personen jeweils 3 unterschiedliche Bilder enthält. Je ein (Abb. 3A) oder zwei (Abb. 3B) dieser 3 Bilder werden in der Datenbasis des Systems gespeichert, das dritte wird zum Test der Wiedererkennung benutzt. Ein Vorteil des Elastic-Graph-Matching-Verfahrens ist, daß es mit wenigen gespeicherten Graphen pro Person – im Extremfall einem einzigen – auskommen kann (Abb. 3A). In diesem Fall erzielt das durch das Neuronale Netz erlernte Signifikanzkriterium bei FAR=0% eine FRR von etwa 9%, eine leichte Verbesserung gegenüber FRR=12% in [1]. Wie in Abb. 3B gezeigt, lassen sich FAR und FRR jedoch noch erheblich verbessern – bei der hier untersuchten Datenbasis sogar auf jeweils 0% – wenn pro Person je ein weiterer Graph, der von einem leicht anderen Bild der Person erzeugt wurde, gespeichert wird (Abb. 3B). Damit wird die Erkennungsrate berechtigter Personen stark erhöht, *ohne* daß Einbußen bei der FAR auftreten.

Abschließend bleibt festzuhalten, daß es mit den Methoden der Neuronalen Netze erstmals möglich wird, ein Identifikationssystem zur visuellen Gesichtserkennung herzustellen, das diese Erkennungsaufgabe mit der für die Anwendung erforderlichen hohen Genauigkeit lösen kann.

– hier beiliegendes Foto einbauen –

Abbildung 1: Das Zutrittskontrollsystem *ZN-Face*: Der Benutzer blickt in den halbdurchlässigen Spiegel der Konsole, hinter dem sich die Kamera befindet. Das aufgenommene Bild wird automatisch mit den gespeicherten Merkmalen verglichen.

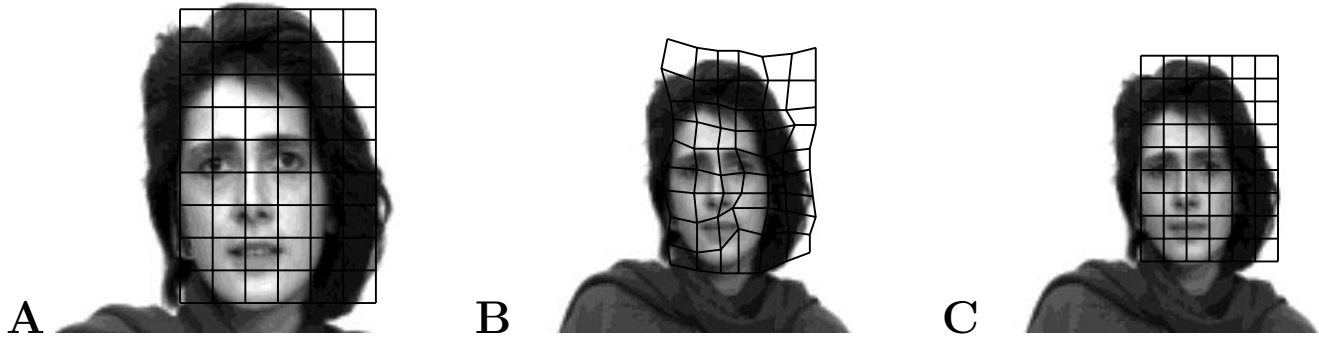


Abbildung 2: Die Datenrepräsentation in *ZN-Face*: Gesichter werden als flexibler Graph mit (hier) 7×10 Knotenpunkten gespeichert (**A,B**). Solche Graphen lassen sich effizient im Bildbereich verschieben oder skalieren (**C**).

Literatur

- [1] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C.v.d. Malsburg, R.P. Wrtz and W. Konen, *Distortion invariant object recognition in the dynamic link architecture*, IEEE Transaction on Computers, 42, p. 300-311, 1993.

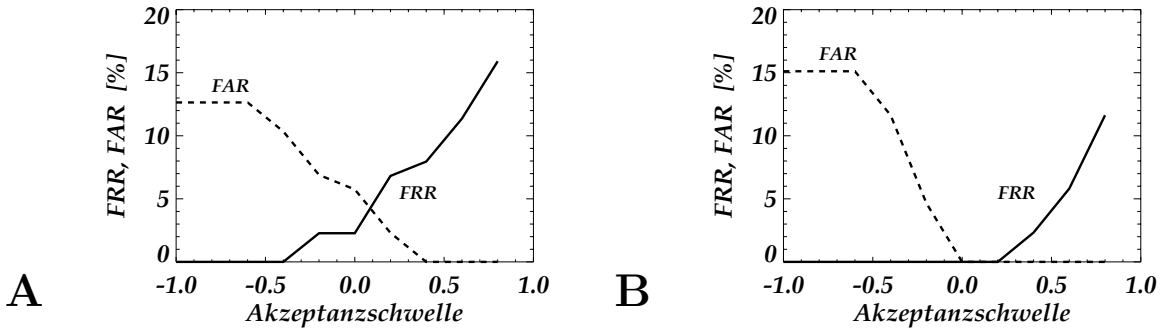


Abbildung 3: Rate der fälschlichen Zurückweisungen (*false rejection rate*, *FRR*) und fälschlichen Akzeptanzen (*false acceptance rate*, *FAR*) bei einer Erkennung aus einer Datenbank von 87 Personen, aufgetragen als Funktion der einstellbaren Akzeptanzschwelle. Die Akzeptanzschwelle 0 entspricht dem durch das Neuronale Netz erlernten Signifikanzkriterium. **(A)** Zu jeder Person ist genau ein Graph abgespeichert: Die Crossover-Rate von FAR und FRR beträgt etwa 3.5% (d. h. der Prozentsatz der korrekt erkannten berechtigten Personen und der korrekt zurückgewiesenen unberechtigten Personen beträgt jeweils 96.5%). **(B)** Zu jeder Person sind die Graphen zweier leicht unterschiedlicher Bilder abgespeichert. Damit wird die Crossover-Rate in diesem Lauf auf 0% gesenkt. (Aufgrund der durch den Algorithmus und das empirische Datenmaterial bedingten leichten Streuungen muß eine konservative Abschätzung allerdings von einer mittleren Crossover-Rate bis zu 1% ausgehen.)