Objektivierung der Beurteilung von Gesichtsasymmetrien durch Bildanalyse

U. Ahlrichs[a], D. Paulus[a], S. Wolf[b]

Lehrstuhl für Mustererkennung (Informatik 5)[a] (*LME*)
und
Klinik und Poliklinik für Hals–, Nasen– und Ohrenkranke[b] (*HNO*)
Universität Erlangen–Nürnberg
Martensstr. 3, D–91058 Erlangen, Germany
email:ahlrichs@informatik.uni-erlangen.de

Erscheint in den Proceedings zum
Workshop Bildverarbeitung in der Medizin

Aachen, November 1996

Objektivierung der Beurteilung von Gesichtsasymmetrien durch Bildanalyse

U. Ahlrichs[a], D. Paulus[a], S. Wolf[b]

Lehrstuhl für Mustererkennung (Informatik 5)[a] (*LME*)
und
Klinik und Poliklinik für Hals–, Nasen– und Ohrenkranke[b] (*HNO*)
Universität Erlangen–Nürnberg
Martensstr. 3, D–91058 Erlangen, Germany
email:ahlrichs@informatik.uni-erlangen.de

Eine objektive Beurteilung von Gesichtsasymmetrien, die durch Störungen in der Funktion der Gesichtsmuskulatur hervorgerufen werden, ist für Verlaufsbeobachtungen und Therapieentscheidungen wünschenswert. In diesem Beitrag stellen wir zwei Systeme vor, die anhand von Videoaufzeichnungen eine computergestützte Bewertung von Nasendeformitäten beziehungsweise Fazialisparesen — eingeschränkt auf die Mundregion — ermöglichen. Die Nasendeformitäten werden aufgrund einer vertikalen Symmetrieachse bewertet, die durch das Gesicht gelegt wird. Für die Beurteilung der Fazialisparesen wird zunächst anhand der Mundkontur, die mit Hilfe von verformbaren Schablonen extrahiert wird, ein Merkmalvektor bestimmt. Dieser wird anschließend einer von neun Klassen zugeordnet.

Keywords: Computergestützte Gesichtsanalyse, Fazialisparese, Nasendeformitäten, Segmentierung, Klassifikation

1 Einleitung

Das Gesicht gilt als “Spiegel der Seele”. Störungen der Funktion der Gesichtsmuskeln sind auf den ersten Blick erkennbar und entsprechend beunruhigend für den Betroffenen. Hervorgerufen werden können solche Störungen beispielsweise durch eine Lähmung des Fazialisnerven (Fazialisparese). Auch das Phänomen von Schiefnasen und Nasendeformitäten läßt sich teilweise auf Störungen in der Muskulatur zurückführen.

Eine Fazialisparese äußert sich in einer sichtbaren Asymmetrie des Gesichts, teilweise während sich das Gesicht in Ruhe befindet, vermehrt jedoch bei der Ausführung mimischer Bewegungen. Typische Symptome sind z.B. das Drehen des Augapfels nach oben beim Schluß der Augen oder das Herunterhängen eines Mundwinkels.

Die Fazialisparese ist die häufigste isoliert auftretende Lähmung überhaupt. In der Klinik und Poliklinik für Hals–, Nasen– und Ohrenkranke (HNO) der Universität Erlangen–Nürnberg werden pro Jahr über 100 Patienten mit neu auftretenden Lähmungen beobachtet. Hierzu bietet [11] einen Überblick.

In der Diagnostik von Lähmungen und Nasendeformitäten ist man bisher auf eine visuelle Beurteilung angewiesen, die naturgemäß bei jedem Untersucher verschieden ist. Eine objektive Quantifizierung ist jedoch für Verlaufsbeobachtungen einschließlich weiterer Therapieentscheidungen wichtig. Diese Bewertung ist Ziel der computergestützten Analyse von Gesichtsbildern.

Als Rahmenbedingungen wird dabei vorgegeben, daß das Verfahren kostengünstig sowie für den Arzt leicht handbar und auswertbar ist. Ebenso soll es für den Patienten durch den Verzicht auf Hilfsmittel möglichst angenehm gestaltet werden.

In diesem Beitrag stellen wir zwei computergestützte Systeme vor, die mit Mitteln der Bildanalyse den Arzt in seiner Diagnose unterstützen. Dazu betrachten wir zunächst die konventionellen Untersuchungsmethoden (Abs. 2). Danach stellen wir innerhalb der Bildanalyse (Abs. 3) ein Segmentierungsverfahren für asymmetrische Münder vor (Abs. 3.1), auf dessen Ergebnissen die in Abschnitt 4 beschriebene Klassifikation von Fazialisparesen basiert, sowie ein segmentierungsfreies Verfahren für Nasendeformitäten (Abs. 3.2). Abschließend geben wir eine Zusammenfassung und einen Ausblick.

2 Konventionelle Untersuchungsmethoden

Die Bandbreite der auf dem visuellen Eindruck des Untersuchers beruhenden Beurteilungsschemata reicht von ausschließlich subjektiven Bewertungen, die das Gesicht im Ganzen betrachten [4], bis zu Bewertungen anhand von Distanzmessungen zwischen festgelegten topographischen Punkten, die zu einer Graduierung für das gesamte Gesicht kombiniert werden. Zur letztgenannten Gruppe gehört das Indizierungssystem von Stennert [8], das aus zwei verschiedenen Untersuchungslisten besteht. Mit dem zehn Untersuchungspunkte umfassenden Lähmungsindex wird ein Vergleich zwischen gesunder und erkrankter Gesichtshälfte in Ruhe und bei der Ausführung bestimmter mimischer Bewegungen wie dem Mundspitzen oder Zähnezeigen vorgenommen. Die Graduierung erfolgt durch die Beantwortung von zehn Ja/Nein Fragen. Von den Untersuchungspunkten entfallen dabei fünf auf die Mundregion — begründet durch ihre große funktionelle Bedeutung —, vier auf die Augenregion und ein Punkt auf die Stirnregion. Ebenso befassen sich sechs der zehn Untersuchungspunkte aufgrund der höheren Information über die Schädigung des Nervens mit dem Zustand der aktiven Bewegung.

Bewegungserscheinungen, die auf eine fehlerhafte Heilung zurückzuführen sind — wie zum Beispiel eine gleichzeitige Bewegung von Mundwinkel und Lidern —, werden mit dem zweiten Index, dem Defektheilungsindex, klassifiziert.

Durch die Festlegung der topographischen Punkte durch den Untersucher ist aber auch dieses Indizierungssystem einem gewissen Maß an Subjektivität unterworfen.

3 Bildanalyse

Neben den im letzten Abschnitt erläuterten Graduierungsschemata sind in der Literatur auch Systeme bekannt, in denen eine computergestützte Bewertung vorgenommen wird. So wird in [9] ein System zur Quantifizierung von Gesichtsasymmetrien vorgestellt, in dem mit Hilfe von vier Kameras die Positionsänderungen von sechs im Gesicht angebrachten Markern während der Durchführung vorgegebener Bewegungen bestimmt werden. In [5] wird auf Hilfsmittel wie Marker verzichtet. Stattdessen werden Videoaufzeichnungen von Patienten ausgewertet, die die Patienten bei der Ausführung von Gesichtsbewegungen zeigen. Nach Digitalisierung der Videoaufzeichnung wird jedes Bild der Bildfolge vom ersten Bild der Bildfolge subtrahiert, das das Gesicht des Patienten in Ruhe zeigt. Mit diesen Differenzbildern lassen sich Asymmetrien

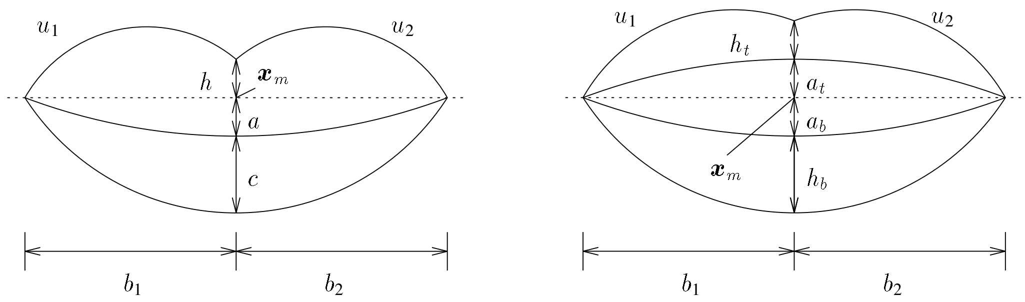

Bild 1: Schablonen für einen geschlossenen und einen geöffneten Mund [12]

anhand von Unterschieden in den beiden Gesichtshälften bestimmen. Jedoch wird auch hier die Symmetrieachse interaktiv durch den Untersucher festgelegt.

Im Gegensatz dazu sind die im folgenden beschriebenen Verfahren unabhängig von etwaigen Eingriffen des Untersuchers. Sie basieren auf Videosequenzen, die in der HNO mit einer handelsüblichen Videokamera aufgezeichnet wurden. Diese zeigen Kopf–Schulter–Szenen von Patienten, die vor einem homogenen, schwarzen Hintergrund sitzend die im Stennertschen Lähmungsindex aufgeführten mimischen Bewegungen ausführen.

3.1 Analyse von Gesichtsregionen bei einer Fazialisparese

In [1] wird zur Analyse der Lähmungen zunächst nur die Mundregion betrachtet, da dieser Region innerhalb des Gesichts die größte funktionelle Bedeutung zukommt (vgl. Abs. 2). Die Berechnung der Merkmale zur Beurteilung der Asymmetrie basiert dort auf der segmentierten Kontur des Mundes. Diese Segmentierung des Mundes gliedert sich in zwei Schritte. Zunächst erfolgt eine grobe Bestimmung der Mundregion. Dazu werden die linke und rechte Gesichtsgrenze durch Suchen in einem eindimensionalen Feld bestimmt, das durch Addition der Grauwerte der einzelnen Bildspalten (vertikale Projektion) entsteht. Unter Zuhilfenahme von antropometrischen Standards läßt sich die Trennungslinie zwischen Ober– und Unterlippe mit einer horizontalen Projektion und einer heuristischen Suche wie bei den Gesichtsgrenzen ermitteln. Bei einer Auswertung des Verfahrens auf 51 Bildfolgen wurde die Mundregion in 90 % der Fälle richtig gefunden und in 88 % der Fälle in der Bildsequenz richtig verfolgt.

Dieser erste Schritt dient der Verkleinerung des Suchraums für die verformbaren Schablonen, mit denen im zweiten Schritt die Kontur des Mundes segmentiert wird. Der Mund wird dabei durch eine parametrisierte Schablone beschrieben, wobei die Lippenkonturen durch Parabeln approximiert werden (vgl. Bild 1). Damit kann a–priori Wissen über die erwartete Form des Mundes dazu verwendet werden, fehlende Bildinformation wie z.B. Kanten bei mangelndem Kontrast zwischen Lippen und Haut zu kompensieren. Um eine optimale Übereinstimmung zwischen Schablone und Mund im Bild zu erreichen, müssen die Parameterwerte entsprechend angepaßt werden. Hierzu wird eine Energiefunktion definiert, mit der die Übereinstimmung zwischen den Mund charakterisierenden Kanten, Intensitätmaxima und –minima im Bild und Schablone bewertet wird. Neben diesen sogenannten externen Energien fließen in die Gesamtenergie Energieterme ein, die die Stabilität der Schablonen garantieren. Die Energiefunktion muß minimiert werden, wozu in [1] stochastische Optimierungsverfahren [3] herangezogen werden.

Während in Experimenten gezeigt werden konnte, daß die in [12] vorgeschlagene Schablone für einen geschlossenen Mund auch zur Segmentierung asymmetrischer Münder geeignet ist (vgl. Bild 2), ließ sich dies für geöffnete Münder nicht bestätigen. Deshalb wurden die beiden Parabeln zur Beschreibung der Mundöffnung wie bei der Oberlippe jeweils in zwei Parabeln geteilt, so daß die Mundhälften unabhängig voneinander beschrieben werden können. Das Ergebnis für diese Schablone ist in Bild 2 dargestellt. Theoretisch ließe sich dieses Ergebnis noch verbessern, wenn man das gleiche Vorgehen für die Parabel der Unterlippe wiederholt. Jedoch wirkte sich die dadurch erhöhte Anzahl an Parametern negativ auf den Aufwand und die Stabilität der Optimierungsverfahren aus, so daß darauf verzichtet wurde.

Da die Beurteilung von Asymmetrien gerade bei der Ausführung von mimischen Bewegungen in der Praxis am wichtigsten ist (vgl. Abs. 2), genügt es nicht, nur Einzelbildauswertungen vorzunehmen. Eine Möglichkeit bestände in der Verwendung bewegungsbasierter Verfahren. Durch die Verwendung von im Gesicht angebrachten Markern lassen sich z.B. Bewegungsvektoren schätzen und aus ihnen direkt Merkmale ableiten [2]. Anfängliche Experimente zeigten jedoch, daß dazu die zeitliche Abtastrate zu gering war, wodurch sich große Unterschiede zwischen den Bildern der Bildfolge zu Beginn und am Ende der Bewegung ergaben. Die verformbaren Schablonen erwiesen sich gegenüber diesen großen Unterschieden als robuster und wurden deshalb auch für die Bildfolgen verwendet.

3.2 Qualitative Symmetriebeurteilung der Nase

Ziel der Arbeiten in [10] war die Entwicklung eines einfachen Systems zur Ermittlung qualitativer Maße für Gesichtssymmetrien, die zur Beurteilung von Schiefnasen herangezogen werden sollen. Dazu wurde ein Streifenmuster auf das Gesicht projiziert. Im Bild wurden horizontale Linien detektiert, deren Symmetrieachse heuristisch bestimmt wurde. Der Aufbau ist unkalibriert, da das Ziel nicht in einer hochgenauen Vermessung, sondern in einer qualitativen Beurteilung besteht, und zudem ein niedriger Preis und Aufwand an Technik gewünscht war. Ein Beispiel zeigt Bild 3.

Zur Streifenerkennung wurde ein modifizierter Kantendetektor aus [6] verwendet, in dem nur die Masken zur Detektion annähernd horizontaler Kanten ausgewählt wurden. Für jede horizontale Linie wird der Punkt bestimmt, der die Linie — bildlich gesprochen — in der Balance hält. Diese Punkte werden wiederum verbunden und außerdem durch eine Ausgleichsgerade

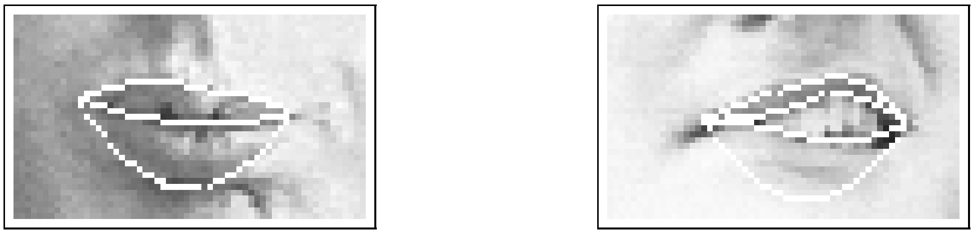

Bild 2: Segmentierungsergebnis für einen geschlossenen und einen geöffneten Mund

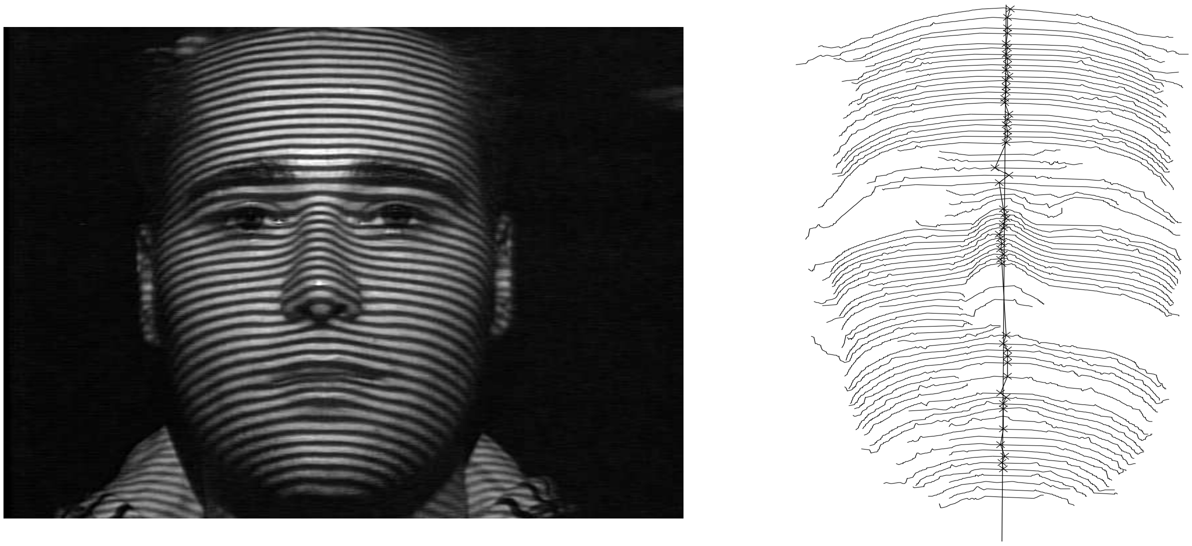

Bild 3: Bilder einer gesunden Personen. Links das Grauwertbild mit projiziertem Streifenmuster, rechts das Segmentierungsergebnis mit Ausgleichsgeraden

approximiert, deren Steigung als Merkmal herangezogen wird.

Das Verfahren wurde mit zehn Patienten und zahlreichen gesunden Testpersonen durchgeführt. Die Symmetriemaße unterschieden sich in den betrachteten Fällen deutlich. Für eine weitergehende Beurteilung des Nutzens dieses Verfahrens ist die Stichprobe noch nicht ausreichend.

4 Klassifikation

Während die zur Analyse der Nasendeformitäten bestimmten Merkmale noch nicht für eine Klassifikation verwendet wurden, wird in [1] ein erster Versuch für die Graduierung der Fazialisparesen auf der Basis des Stennertschen Lähmungsindex vorgestellt. Die dort für den Mundwinkeltiefstand, das Mundspitzen und das Zähnezeigen angegebenen quantitativen Abmessungen können nach Segmentierung entsprechender topographischer Punkte mit Hilfe der verformbaren Schablonen direkt in Merkmalwerte umgesetzt werden. Als Merkmale werden das Verhältnis des Abstandes zwischen Filtrum und Mundwinkel zwischen beiden Mundhälften, die Position der Mundwinkel und für das Zähnezeigen das Verhältnis des Flächeninhaltes in den beiden Mundhälften herangezogen, wobei jeweils der Mittelwert über die Bildfolge gebildet wird.

Anhand dieser Merkmale wird mit Hilfe eines Minimumabstandsklassifikators [7] die Asymmetrie durch Zuordnung des Merkmalvektors in eine von neun Klassen bewertet. Diese Anzahl ergibt sich aus den 2^3 kombinatorisch möglichen Klassen bei drei Untersuchungspunkten sowie einer Rückweisungsklasse. Die Erkennungsraten für die einzelnen Untersuchungspunkte lagen bei einer Teststichprobe mit 28 Patienten mit 55 % für den Mundwinkeltiefstand, 72 % für das Mundspitzen und 61 % für das Zähnezeigen deutlich über der Erkennungsrate für die Gesamtdiagnose von 35 %. Dies ist auf das Training des Klassifikators zurückzuführen, bei dem die Untersuchungspunkte unabhängig voneinander trainiert wurden, da nicht für alle möglichen

Klassen Patienten in der willkürlich aus der HNO–Gesamtstichprobe ausgewählten Trainingsstichprobe enthalten waren. Diese war disjunkt zur Teststichprobe und umfaßte 23 Patienten.

5 Zusammenfassung und Ausblick

In diesem Beitrag haben wir zwei Systeme vorgestellt, eines für die computergestützte Beurteilung von Nasendeformitäten und eines für die Bewertung von Fazialisparesen. In letzterem wurde zunächst zur Vereinfachung der Aufgabe nur die Mundregion betrachtet. Während die Nasendeformitäten mit Hilfe einer vertikalen Symmetrieachse beurteilt werden, erfolgt bei den Fazialisparesen eine Segmentierung des Mundes mit verformbaren Schablonen. Die anhand dieser Kontur extrahierten Merkmale werden mit einem Minimumabstandsklassifikator einer Klasse zugeordnet, die das Ausmaß der Lähmung beschreibt.

In Zukunft soll diese Bewertung auf das gesamte Gesicht ausgedehnt werden. Zunächst ist jedoch insbesondere bei der Klassifikation der Fazialisparesen für die Validierung der Eignung der gewählten Merkmale eine repräsentative Stichprobe notwendig, die das Training von Gesamtdiagnosen erlaubt. Das System soll sich dann nicht nur auf die Unterstützung des Arztes bei der Diagnose beschränken, sondern auch in der Therapie eingesetzt werden.

Literatur

1. U. Ahlrichs: *Beurteilung von Asymmetrie der Mundpartie*, Diplomarbeit, Lehrstuhl für Mustererkennung (Informatik 5), Universität Erlangen–Nürnberg, Erlangen, 1995.
2. K. Finn, A. Montgomery: *Automatic Optically-Based Recognition Of Speech*, *Pattern Recognition Letters*, Bd. 8, Nr. 3, March 1988, S. 159–164.
3. F. Gallwitz: *Lokalisation von 3D–Objekten in Grauwertbildern*, Diplomarbeit, Lehrstuhl für Mustererkennung (Informatik 5), Universität Erlangen–Nürnberg, 1994.
4. J. W. House, D. E. Brackmann: *Facial Nerve Grading System*, *Otolaryngological Head und Neck Surgery*, Bd. 93, 1985, S. 146–147.
5. J. Neely, J. Cheung, M. Wood, J. Byers: *Computerized Quantitative Dynamic Analysis of Facial Motion in the Paralyzed and Synkinetic Face*, *The American Journal of Otology*, Bd. 13, 1992, S. 97–107.
6. R. Nevatia, R. Babu: *Line Feature Extraction and Description*, *Computer Vision, Graphics and Image Processing (CVGIP)*, Bd. 13, 1980, S. 257–269.
7. H. Niemann: *Klassifikation von Mustern*, Springer, Berlin, 1983.
8. E. Stennert, C. Limberg, K. Frentrup: *Parese- und Defektheilungs-Index*, *HNO*, Bd. 25, 1977, S. 238–245.
9. E. Stüssi, S. Handschin, M. Frey: *Quantifizierung von Gesichtsasymmetrien, Eine Methode zur Objektivierung von Beeinträchtigungen der Gesichtsmotorik; eine Pilotstudie*, *Biomedizinische Technik*, Bd. 37, Nr. 1-2, Januar-Februar 1992, S. 14–19.
10. K. Wildermuth: *Optische Vermessung von Gesichtsbildern*, Diplomarbeit, Lehrstuhl für Mustererkennung (Informatik 5), Universität Erlangen–Nürnberg, Erlangen, 1993.
11. S. Wolf, M. Müller, W. Schneider, C. Haid, M. Wigand: *Facial Nerve Function after Transtemporal Removal of Acoustic Neurinomas: Results, Time Course or Function and Rehabilitation*, in M. Samii (Hrsg.): *Skull Base Surgery*, Hannover, 1992, S. 894–897.
12. A. L. Yuille, P. W. Hallinan, D. C. Cohen: *Feature Extraction from Faces Using Deformable Templates*, *International Journal of Computer Vision*, Bd. 8, 1992, S. 99–111.