Automatische Graduierung von Gesichtsparesen

A. Gebhard[1], D. Paulus[1], B. Suchy[2], I. Fucak,[2], S. Wolf[2] und H. Niemann[1]

[1]Lehrstuhl für Mustererkennung
Friedrich-Alexander-Universität Erlangen-Nürnberg
Martensstraße 3, 91054 Erlangen
Email: {gebhard,frank,paulus,niemann}@informatik.uni-erlangen.de
[2]Klinik und Poliklinik für Hals-, Nasen- und Ohrenkranke (HNO)
Friedrich-Alexander-Universität Erlangen-Nürnberg
Waldstraße 1, 91054 Erlangen
Email: {bernhard.suchy,stefan.wolf}@rzmail.uni-erlangen.de
Gefördert von der DFG, Sonderforschungsbereich 603, TP B3

Abstract. Ein System zur automatischen Graduierung von Gesichtslähmungen wird vorgestellt. Dabei werden die Module zur Verfolgung von Gesichtern und Gesichtsmerkmalen basierend auf Support Vector Machines und der automatische Graduierung vorgestellt. Die Graduierung der Lähmungen orientiert sich am verbreiteten, aber subjektiven House-Brackmann-Index. Ergebnisse, die an einer Stichprobe mit 28 Patienten erzielt wurden, werden präsentiert.

1 Einleitung

Symptome einer einseitigen Gesichtslähmung sind der Funktionsausfall bzw. die Fehlfunktion der Gesichtsmuskulatur durch fehlende oder falsche Innervation. Dadurch bedingt sind Asymmetrien im Gesicht, die während mimischer Bewegungen auftreten. In [1] haben wir ein System vorgestellt, das diese Asymmetrien in statischen Aufnahmen auswertet. Die Information über die Lähmung wurde dabei aus den Öffnungswinkel von Augen und Mund bestimmt, der implizit in der Filterantwort eines rotierenden Keilfilters [8] enthalten ist. Der Vergleich der Signaturen der Filter zwischen linken und rechten äußeren Augenwinkel bzw. Mundwinkel beim Erreichen der extremalen Position von mimischen Bewegungen liefert die Information über Gesichtsasymmetrie, die anschließend zur Detektion von Gesichtslähmungen eingesetzt wird. Ein Experiment mit 58 Patienten (15 vom Grad House I: normal/gesund) hat gezeigt, dass mit einer Detektionsrate von 86% (93% Recall) Gesichtslähmungen detektiert werden konnten. Jedoch traten verschiedene Mängel zutage:

1. Die Analyse des Gesichts setzt eine genaue Lokalisation der äußeren Augen- bzw. Mundwinkel voraus. Eine ausreichend genaue automatische Lokalisation ist nur in ca. der Hälfte aller Fälle gelungen.

2. Da nur statische Aufnahmen des Patientengesichts ausgewertet wurden, wird auf einen Großteil der Information aus der Gesichtsbewegung verzichtet.

3. Es konnten keine Aussagen über Synkinesien (Fehlbewegungen) gemacht werden, da die gesamte Gesichtsasymmetrie als Bewegungsdefizit interpretiert wurde.

Die in dem Beitrag vorgestellten Neu-/Weiterentwicklungen der einzelnen Komponenten erfolgte unter Berücksichtigung diese Mängel. Die Diagnoseunterstützung basiert nun nicht mehr auf der lokalen Orientierungsanalyse der Augen- und Mundwinkel, sondern auf der Bewegungsanalyse der Gesichts während der Ausführung von mimischen Übungen. Dadurch ist es sowohl möglich, Fehl- als auch Falschbewegungen zu erkennen und zu bewerten.

Die Arbeit gliedert sich wie folgt: In Abs. 2 stellen wir die Lokalisation und Verfolgung von Gesichtern und Gesichtsmerkmalen auf Basis von Support Vektor Maschinen vor. Aus Differenzbildern und dem optischen Fluss, deren Berechnung aus Bildfolgen während der Ausführung der mimischen Übungen geschieht, erhalten wir die zur Graduierung notwendigen Information. Darauf wird in Abs. 3 eingegangen. Dem folgt in Abs. 4 die Beschreibung der durchgeführten Experimente, zusammen mit den erzielten Ergebnissen. Abgeschlossen wird der Beitrag mit einer Zusammenfassung in Abs. 5.

2 Lokalisation und Verfolgung von Gesichtern und Gesichtsmerkmalen

Zur Extraktion der benötigten Information aus den Gesichtsbildern werden die Postionen des Gesichts (d. h. der Nasenspitze) und die des „Zwischenaugenpunktes“, dem Punkt, der am obersten Punkt des Nasenrückens zwischen den Augen liegt, bestimmt. Damit ist implizit die Gesichts-, Augen- und Mundregion festgelegt. Zur Lokalisation dieser beiden Punkte (und der damit betrachteten Regionen) setzen wir *Support Vector Machins* (SVM) [7] ein, deren Einsatzmöglichkeit zur Lokalisation von Gesichtern in [5] gezeigt wurde.

Die Lokalisation und Verfolgung von Gesichtern und Gesichtsmerkmalen erfolgt ansichtenbasiert, d. h., ein Eingabebild wird nicht segmentiert, sondern die Bildinformation wird evtl. unteragbetastet und in einem Merkmalsvektor während des Trainings und der Klassifikation zur Verfügung gestellt. Für das Training des Lokalisators wurden Gesichter verschiedener Patienten aus Aufnahmen geschnitten und in zwei verschiedenen Auflösungsstufen (8×8 und 16×16 Pixel) als Trainingsmaterial benutzt. Der Mittelpunkt der ausgeschnittenen Bilder war die Nasenspitze. Aus jedem Gesichtsbild werden noch jeweils zufällig 5 Teilbilder zufällig gewählt, die Trainingsmaterial der Klasse „Nicht-Gesicht“ ergeben. Für die beiden Auflösungsstufen wurden zwei SVMs ϕ_8 und ϕ_{16} trainiert.

Zur Lokalisation des Gesichts wird zunächst das Eingabebild unterabgetastet und in einem 24×18-Bild gespeichert. Die zentralen 16×10 Bildpunkte des Bildes werden mittels des Klassifikators ϕ_{16} getestet. Bildpunkte, die von ϕ_{16} zur Klasse „Gesicht“ zugewiesen werden, werden als Gesichtshypothesen betrachtet und von ϕ_8 in doppelter Bildauflösung (4-fache Anzahl der Bildpunkte) verifiziert. Sind mehrere Bildpunkte auf diese Weise als Gesicht identifiziert worden, so wird die tatsächliche Gesichtsposition der Bildposition zugesprochen, die im transformierten Raum die größte Distanz zur Hyperebenen von ϕ_8 besitzt. Dieser heuristische Ansatz wurde u. a. in gleicher Weise in [6] eingesetzt.

Eine weitere SVM ϕ_a wird anhand von unterabgetasteten Ansichten eines Augenpaares trainiert. Die Größe des Augenbereichs ist 32×12 Bildpunkte. Damit kann in der oberen Gesichtshälfte (d. h. über der Nasenspitze) nach den Augen (bzw. dem Augenzwischenpunkt) gesucht werden.

Zur Verfolgung wird die gefundene Gesichtsposition als Beobachtung einem Kalman-Filter (Modellierung mit konstanter Geschwindigkeit) übergeben, das nach einer Initialisierung die mögliche Gesichtsposition und den damit verbundenen Hypothesenraum einschränkt. In Fig. 1a) wird wird das Ergebnis einer Gesichtmerkmalsverfolgung gezeigt.

3 Graduierung von Gesichtslähmungen

3.1 Extraktion von Information zur Graduierung

Die Graduierung einer Gesichtslähmung basiert auf der Analyse der Bewegung der beiden Gesichtshälften während der Ausführung spezifischer mimischer Übungen. Ziel ist es, durch eine automatische Bewertung der Bewegung eine objektive Graduierung der Lähmung zu bestimmen (s. a. [4]). Der Patient sitzt vor einer Kamera. Sein Gesicht wird vom Modul aus Abs. 2 verfolgt. Der Patient führt die mimischen Übungen „Stirn runzeln“, „Augen schließen“, „Mund spitzen“ und „Zähne zeigen“ aus. Von jeder dieser vier Übungen wird je eine Bildsequenz $\boldsymbol{f}_s, s = 1, ..., 4$ aufgezeichnet.

Zwischen allen aufeinanderfolgenden Bilder der Bildfolgen werden Folgen von Differenzbildern $\boldsymbol{d}_s$ (siehe Fig. 1b und c) und der optische Fluss $\boldsymbol{v}_s$ (nach Lukas-Kanade [3]) bestimmt. Aus diesen Repräsentationen werden für die Augen/Stirn- und die Mund/Nase-Region folgende Informationen extrahiert:

1. Die gesamte Veränderung d_{al}, d_{ar}, d_{ml} und d_{mr} in den linken und rechten Hälften der entsprechenden Regionen A_l, A_r, M_l und M_r (siehe Fig. 1) durch

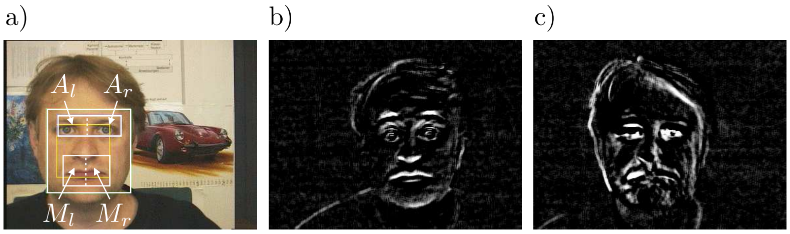

Fig. 1. a) Verfolgung eines Gesichts und der Gesichtsmerkmale; Regionen werden in linke (A_l und M_l) und rechte (A_r und M_r) unterteilt. b) Differenzbild mit symmetrischen Stirnrunzeln. c) Differenzbild mit asymmetrischen Augenschließen

Summation der absoluten Differenzen aus den Differenzbildern $d_{s,j}$ (ähnlich dem Verfahren aus [4]):

$$d_{s,al} = \sum_j \sum_{\boldsymbol{x} \in A_l} |d_{s,j}(\boldsymbol{x})|, \quad d_{s,ar}, d_{s,ml}, d_{s,mr} \text{ analog}$$

2. Die Übereinstimmung der Bewegungen c_a, c_m in der linken und rechten Gesichtshälfte durch Summation über die absoluten Differenzen entsprechender Bildpunkte der Differenzbilder $d_{s,j}$. Der zu einem Bildpunkt $\boldsymbol{x}_1 \in A_l$ entsprechende entsteht durch Spiegelung von $\boldsymbol{x}_1$ an der Trennlinie zwischen A_l und A_r (Notation: $\boldsymbol{x}_1 \| \boldsymbol{x}_2$). Analoges gilt für die Mundregion.

$$c_{s,a} = \sum_j \sum_{\substack{x_1 \in A_l \\ x_2 \in A_r \\ x_1 \| x_2}} |d_{s,j}(\boldsymbol{x}_1) - d_{s,j}(\boldsymbol{x}_2)|, \quad c_{s,m} \text{ analog}$$

3. Die Differenz der Flußrichtung $f_{a,s}$ und $f_{m,s}$ an entsprechenden Bildpositionen aus der linken und rechten Gesichtsregion aus dem Bewegungsrichtungsfeld gewichtet mit dem Wert aus dem Differenzbild.

$$f_{s,a} = \sum_j \sum_{\substack{x_1 \in A_l \\ x_2 \in A_r \\ x_1 \| x_2}} |\boldsymbol{v}_{s,j}(\boldsymbol{x}_1) d_{s,j}(\boldsymbol{x}_1) - \boldsymbol{v}_{s,j}(\boldsymbol{x}_2) d_{s,j}(\boldsymbol{x}_2)|_2, \quad f_{s,m} \text{ analog}$$

Während einer Untersuchung werden so für den Patienten je 16 Parameter für die Augen- und Mundregion bestimmt.

3.2 Graduierungsschema

Ziel bei der Graduierung von Gesichtslähmungen ist die automatische Zuordnung eines Patienten zu den Klassen „Gesund (G1)“, „Leichte Lähmung (G2)“, „Lähmung (G2)“und „Starke Lähmung (G4)“. Dieses Graduierungsschema leitet sich direkt aus dem Lähmungsindizierungsschema nach House-Brackmann[2] ab, wobei Patienten vom Grad House I der Klasse G1, vom Grad House II oder House III der Klasse G2, vom Grad House IV oder House V der Klasse G3 und Patienten vom Grad House VI der Klasse G4 zugewiesen werden. Die automatische Klassifikation erfolgt derzeit anhand eines Nächster-Nachbar-Klassifikators der 32-dimensionalen Merkmalsvektoren unter Einsatz der L_2-Norm.

4 Ergebnisse

Das im Beitrag beschriebene System wurde bereits an der HNO-Klinik Erlangen installiert und an einer Stichprobe von 28 Patienten getestet. Die Klassifikationsergebnisse, die durch die Verfahren aus Abs. 3 erzielt wurden, werden in Tab. 1 gezeigt. Hierbei ist zu bemerken, dass bei einigen Patienten aus G3 eine starke Lähmung entweder in der Augen- oder in der Mundregion vorlag, die der House-Brackmann-Index nicht differenziert beschreibt.

	G1	G2	G3	G4
G1	5	2	0	0
G2	1	3	2	0
G3	1	2	3	2
G4	0	0	3	4

Tab. 1. Ergebnisse der Graduierung von 28 Patienten mit Gesichtslähmungen

5 Zusammenfassung

Im vorliegenden Beitrag wurde ein System zur automatischen Graduierung von Gesichtslähmungen vorgestellt. Damit wird ein Beitrag zur empfohlenen und oft angemahnten Objektivierung der Lähmungsgraduierung gegeben. Im einzelnen wurde bei der Präsentation auf folgende Systemmodule eingegangen: Lokalisation und Verfolgung von Gesichtern und Gesichtsmerkmalen mit Support Vector Machines, Extraktion von Information zur Lähmungsgraduierung aus Gesichtsbildfolgen und der Klassifikation der Lähmung des Patienten. Dieses automatische Graduierungssystem wurde anhand einer Patientenstichprobe (28 Patienten, davon 7 House I/gesund) getestet.

Literaturverzeichnis

1. A. Gebhard, D. Paulus, B. Suchy, S. Wolf, and H. Niemann. System zur Diagnoseunterstützung von Patienten mit Gesichtslähmungen. In *4. Workshop Bildverarbeitung für die Medizin*, pages 249–253. Springer, 2000.
2. John W. House and Derald E. Brackmann. Facial nerve grading system. *Otolaryngological Head und Neck Surgery*, 93:146–147, 1985.
3. B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proceedings of the* 7^{th} *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.
4. J. G. Neely, J. Y. Cheung, M. Wood, and John Byers. Computerized quantitative dynamic analysis of facial motion in the paralyzed and synkinetic face. *The American Journal of Otology*, 13:97–107, 1992.
5. E. Osuna, R. Freund, and F. Girosi. Training support vector machines: an application to face detection. In *Proceedings of Computer Vision and Pattern Recognition*, pages 130–143, Puerto Rico, Juni 1997. IEEE Computer Society Press.
6. B. Schölkopf. *Support Vector Learning*. R. Oldenbourg Verlag, München, 1997.
7. V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, Heidelberg, 1996.
8. W. Yu, K. Daniilidis, and G. Sommer. Rotated wedge averaging method for junction characterization. In *Proceedings of CVPR'98*, pages 390–395, Santa Barbara, California, USA, 1998.