Computergestützte Detektion charakteristischer Bildpunkte in Gesichtsbildern zur Unterstützung der Phänotypdiagnose*

L. Witta[1], R. Herpers[2], M. Michaelis[2], G. Sommer[3] und S. Stengel-Rutkowski[4]

[1]Lehrstuhl für Mensch-Maschine-Kommunikation, TU-München, 80290 München
[2]GSF-Institut für Medizinische Informatik und Systemforschung, Neuherberg, 85764 Oberschleißheim, E-mail: (herpers,michaeli)@gsf.de
[3]Institut für Informatik, Christian-Albrechts-Universität, 24105 Kiel
[4]Kinderzentrum München, Institut für Soziale Pädiatrie und Jugendmedizin der Universität München, Abteilung Genetik, Heiglhofstr. 63, 81377 München

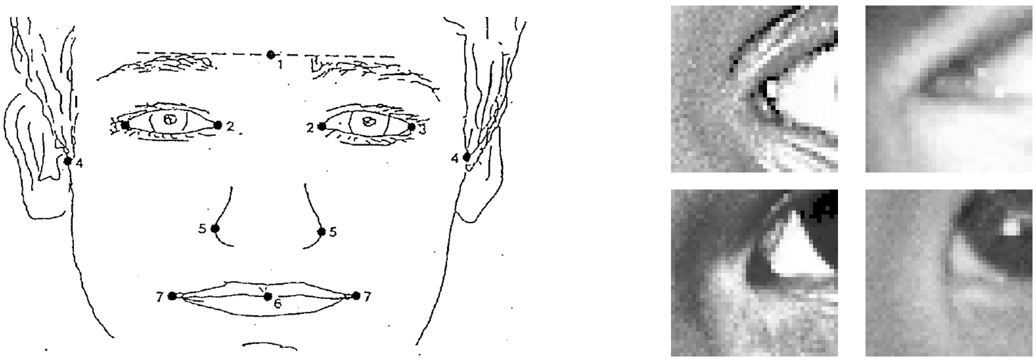
In diesem Beitrag wird ein sequentielles Suchschema vorgestellt, welches automatisch charakteristische Punkte in Bildausschnitten aus Gesichtsbildern detektiert. Die Lokalisierung der Punkte basiert auf einer modellgesteuerten Detektion und Verfolgung der vorhandenen Linien und Kanten. Ein steuerbares Filterschema dient zur flexiblen und schnellen Berechnung der orientierungsselektiven, skalierbaren Kanten- und Linienfilter. Das Suchschema zeichnet sich durch sehr gute Invarianzeigenschaften gegenüber Skalierung, Rotation und Qualität der Bildausschnitte aus. Die so detektierten Bildpunkte dienen zur Ableitung von charakteristischen Gesichtsmaßen und -proportionen, die für eine dysmorphologische Phänotypdiagnose benötigt werden.

Schlüsselwörter: Gesichtsbildverarbeitung, Keypoint-Detektion, Computer Vision

1 Einleitung

Neben der klassischen Gesichtsbilderkennung stellt die Klassifikation oder Typisierung von Gesichtsbildern ein weiteres wichtiges Gebiet in der Gesichtsbildverarbeitung dar. Eine Zuordnung eines Gesichtes zu einer Klasse mit klassenspezifischen Eigenschaften beansprucht in gewisser Hinsicht anspruchsvollere Verarbeitungsmethoden als eine Identifikation, da die interindividuellen Unterschiede zwischen den Gesichtern verschiedener Personen einer Klasse in der Regel wesentlich größer sind, als die intraindividuellen Unterschiede zwischen verschiedenen Bildern derselben Person. Die mit dieser Arbeit verfolgte Anwendung, die Unterstützung der Phänotypbestimmung von Dysmorphiesyndromen, ist gerade eine solche Anwendung, bei der die Gesichtsbilder nicht bestimmten Individuen zugeordnet werden, sondern bestimmten Typenklassen mit charakteristischen Krankheitsbildern. Die Schwierigkeit der Zuordnung zu einer solchen Typklasse liegt in der Auswahl und Berechnung der Merkmale, die signifikant für die Unterscheidung der einzelnen Klassen sind. Ein syndromtypisches Erscheinungsbild setzt sich im Allgemeinen aus der Kombination multipler Dysmorphien zusammen. Dysmorphien sind dabei kleine morphologische Abweichungen vom vorherschenden Phänotyp [1]. Zur Bestimmung dieser Dysmorphien werden geometrische Maße als Merkmale verwendet, die auf der Basis charakteristischer Landmarken definiert sind [1] (Abb. 1a).

*Diese Arbeit wurde von der DFG unter den Aktenzeichen: So 320\1-1 und Ei 322\1-1 gefördert.



(a) (b)

Abbildung 1: (a) Wichtige anatomische Landmarken im frontalen Gesichtsbild. (b) Beispiele von realen inneren Augenecken. Die starke Variabilität der Bilddaten erlaubt keine rein datengetriebene Lokalisation.

Das grundlegende Problem bei einer automatischen Detektion anatomischer Landmarken ist, daß solche markanten Punkte eher durch ihre semantische Bedeutung (Ecke des Mundes oder Auges) als durch eine einheitliche lokale Struktur in den Bilddaten definiert sind (Abb. 1b). Deshalb lassen sich diese Punkte auch nicht mit Verfahren zuverlässig detektieren, die lediglich die Bilddaten in der lokalen Nachbarschaft auswerten, ohne das vorhandene Kontextwissen mit einzubeziehen.

In diesem Beitrag wird ein Verfahren vorgestellt, welches automatisch charakteristische Punkte in Gesichtsbildern detektiert. Die Bildregionen werden bereits vorab durch eine attentive Regionensuche extrahiert [2]. Diese werden anschließend gezielt nach markanten Punkten untersucht. Die Detektion dieser Punkte basiert dabei auf einer sequentiellen Auswertung der Kanten- und Linieninformation. Wesentlich dabei ist, daß der Kantenverfolgung ein Modell der untersuchten Bildregion zugrundeliegt.

2 Steuerbare Filter

Die Grundlage der Merkmalsextraktion bildet ein steuerbares Filterschema, welches Kanten- und Linienfilter einer hohen Orientierungsselektivität in beliebigen Skalen und Orientierungen auf das Bild anwendet. Die Kanten- und Linienfilter $F(x)$ basieren auf der ersten und zweiten Ableitung einer Gaußfunktion (Abb. 2a,b) und werden in ihrer Orientierung und Skala (θ, σ) gesteuert. Der Begriff der Steuerung bezieht sich dabei auf die lineare Rekonstruktion der Filter aus einem festen und kleinen Satz von sogenannten Basisfunktionen $A_k(x)$ (Formel 1 und Abb. 2c,d).

$$F_{\theta,\sigma}(x) \;\approx\; \sum_{k=1}^{N} b_k(\theta,\sigma) A_k(x) \tag{1}$$

Alle Faltungen und Projektionen des Bildes werden ausschließlich mit den Basisfunktionen ausgeführt. Die Kanten- und Liniendetektoren selbst und deren Antworten erhält man durch Superpositionen der Basisfunktionen bzw. deren Antworten. Die Basisfunktionen $A_k(x)$ und die Interpolationsfunktionen $b_k(\theta, \sigma)$ werden durch ein auf der Singulärwertzerlegung basierendes Verfahren berechnet, das auf Perona zurückgeht [3]. Allerdings werden hier Orientierung und Skala simultan gesteuert. Diese Methode wird detailliert in [4] beschrieben.

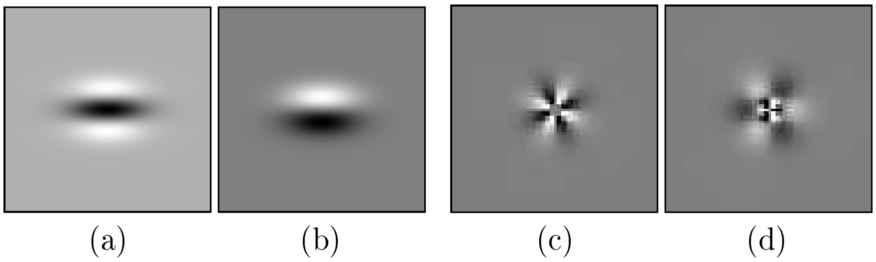


Abbildung 2: (a,b) Linien- und Kantendetektionsfilter (mit einem Seitenverhältnis von 2:1 und einer Faltungskerngröße von 27x27 Pixeln). (c,d) Beispiele für Basisfunktionen zur Steuerung der Filter von (a,b).

Die hier verwendeten Basisfunktionen erlauben eine flexible Anpassung der Approximationsqualität und der Geschwindigkeit. Dies beruht auf folgenden Eigenschaften:
1. Die Basisfunktionen sind orthogonal, d.h. für eine bessere Approximationsqualität können Basisfunktionen hinzugenommen werden, ohne die Interpolationsfunktionen ändern zu müssen. Die neuen Funktionen bzw. Projektionskoeffizienten werden einfach zu den alten addiert.
2. Eine beliebige Anzahl von Basisfunktionen rekonstruiert alle deformierten Filter. Lediglich die Approximationsqualität ist abhängig von der Anzahl der verwendeten Basisfunktionen. Deshalb kann die Geschwindigkeit der Merkmalsgewinnung gesteigert werden, indem anfangs grobe, aber schnelle Approximationen berechnet werden und erst bei Bedarf genauere.

3 Detektion markanter Punkte

3.1 Filtergrundoperationen

Die Detektion der markanten Punkte basiert auf einer modellbasierten Suche und Verfolgung von Kanten- und Linienstrukturen in den Bildregionen. Die Kanten– und Liniendetektion wird durch die folgenden drei Grundoperationen realisiert, welche die eingeführte Steuerbarkeitseigenschaft der verwendeten Filter intensiv nutzen.

- Die erste Filtergrundoperation (BFO1) sucht eine Kante oder Linie bei vorgegebener Orientierung und Skala in einer definierten Region.
- Die zweite Filtergrundoperation (BFO2) bestimmt die Orientierung einer Linie oder Kante an einer vorgegebenen Position durch die maximale Antwort eines rotierten Filters, der bereits bzgl. der Skala an die gesuchte Struktur angepaßt ist.
- Die dritte Filtergrundoperation (BFO3) verfolgt eine Linie oder Kante für einen Schritt. Dazu wird der Filter etwas in die Richtung des Verlaufs der bereits detektierten Struktur verschoben und anschließend senkrecht zu der neuen Position nach einer Kante oder Linie gesucht.

Diese Filtergrundoperationen werden in verschiedenen Orientierungen und Skalen verwendet und zu komplexeren Operationen zusammengesetzt.

3.2 Sequentielles Suchschema

Zur Verdeutlichung des Prinzips der sequentiellen Suchstrategie wird im folgenden die Detektion der Iris in einer Augenregion genauer erklärt (siehe Abb. 3). Jeder Schritt beinhaltet dabei verschiedene Anwendungen der bereits vorgestellten Filtergrundoperationen

(Kap. 3.1) wobei deren Auswahl und Parametrisierung hinsichtlich Filtertyp (Kante, Linie, lokale Energie), Skala und Orientierung bzw. Richtungskorridor von den Ergebnissen vorhergehender Detektionsschritte zusammen mit dem Modellwissen festgelegt wird.

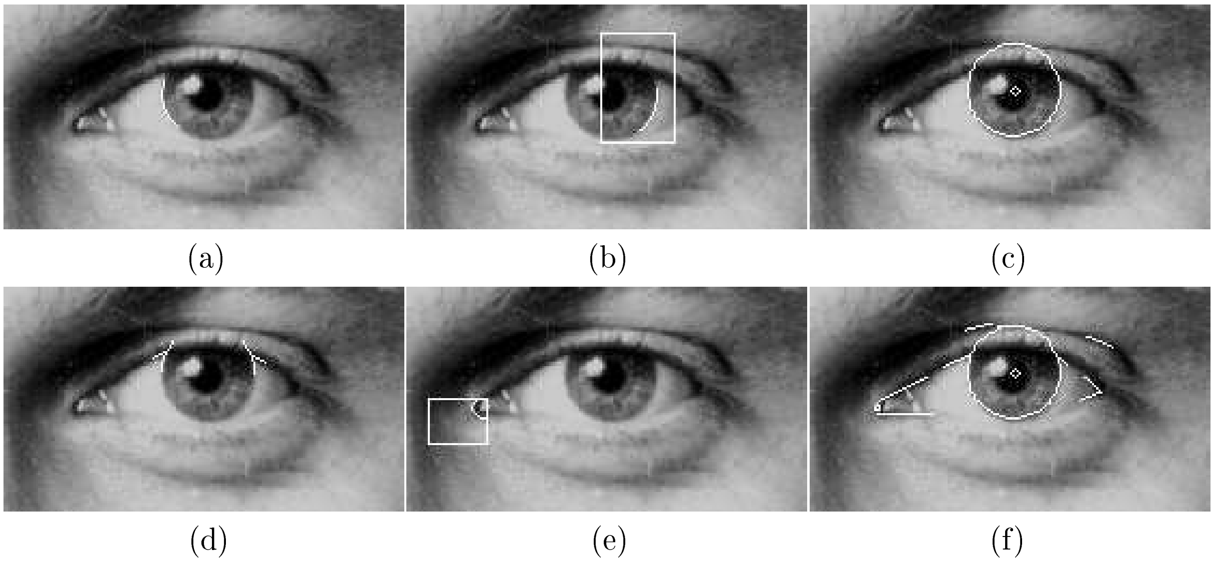


Abbildung 3: Sequentielle Suchstrategie, dargestellt an einem linken Auge. **(a)** Detektion der linken Iriskante. **(b)** Suchbereich für das dazugehörige rechte Kantensegment und Detektion der rechten Iriskante. **(c)** Aus den detektierten Kantensegmenten berechneter Irisradius und Iriszentrum. **(d)** Detektion der Schnittpunkte mit den Augenlidern. **(e)** Suchbereich und Segmentation des inneren Augenwinkels. **(f)** Komplett analysiertes Auge nach Beendigung des sequentiellen Suchschemas. Außer den Koordinaten der markanten Punkte sind zusätzlich noch die Öffnungswinkel der Augenecken, die Orientierung der Augenlider an den Schnittpunkten mit der Iris, Anfangs- End- und Scheitelpunkt der Lidfalte, sowie deren Orientierung an diesen Punkten bestimmt worden.

Die sequentielle Suchstrategie startet mit der Detektion des linken Kantensegments der Iris (Abb. 3a). Dazu wird nach einer markanten Hell/Dunkel-Kante mit nahezu vertikaler Orientierung mit BFO1 gesucht. In diesem ersten Verarbeitungsschritt wird die ganze Augenregion betrachtet, während sich alle nachfolgenden Schritte lediglich auf kleine begrenzte Suchregionen in Abhänigkeit bereits bekannter Informationen beschränken. Das Kantensegment wird in beide Richtungen mit BFO3 solange verfolgt, bis ein signifikanter Wechsel der Filterantwort das Ende der Kante signalisiert. An der Krümmung des detektierten und verfolgten Kantensegments kann nun zwischen einem Kantensegment eines Glanzlichtes und dem des Irisrandes unterschieden werden. In einem ersten Bewertungsschritt wird der Radius und das Zentrum der Iris initial geschätzt. Die bereits abgeleitete Information wird nun zur Definition der Suchregion für das zweite, rechte Kantensegment der Iris verwendet (Abb. 3b). Die Detektion des Kantensegments erfolgt wiederum durch Einsatz von BFO1 und durch Kantenverfolgung mit Hilfe von BFO3. Anschließend wird überprüft, ob beide Kantensegmente zu demselben Kreis gehören, wodurch eine genaue Schätzung des Radius und des Zentrums der Iris ermöglicht wird (Abb. 3c). Anschließend werden sukzessive die Lidkanten unter Einsatz vergleichbarer Vorgehensweisen verfolgt, um die Augenecken und weitere Merkmale in der Augenregion zu detektieren (Abb. 3d-f).

Das sequentielle Suchschema läßt sich flexibel auf andere Gesichtsregionen übertragen, wobei entsprechendes Modellwissen zugrundegelegt werden muß [5] (Abb. 4). Dabei basiert sowohl die Detektion der Nasenflügelkanten, der Mundecken und der Augenbrauenkontur auf Modellwissen, welches aus der Analyse der Augenregion gewonnen wurde.

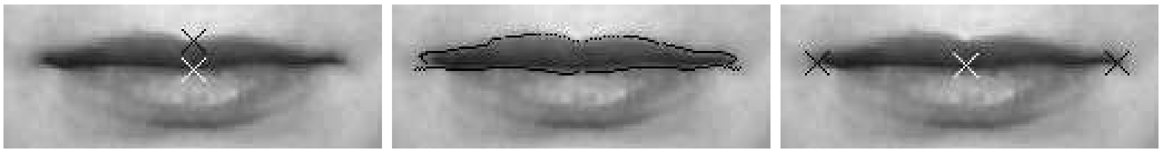
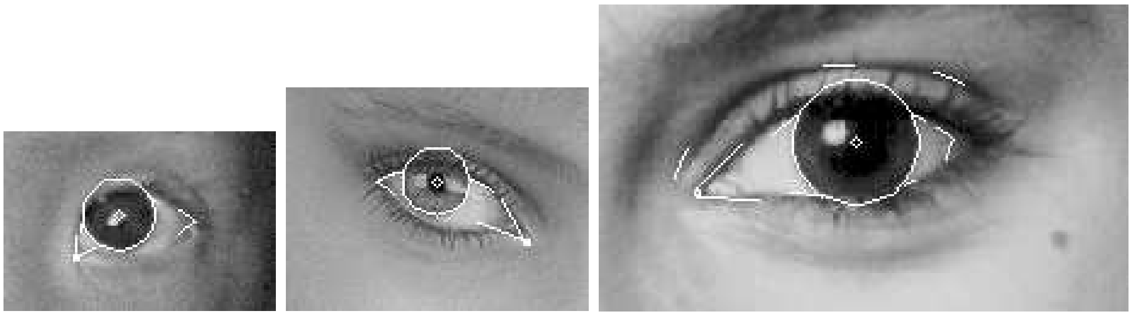


Abbildung 4: Sequentielle Suche in der Mundregion. Nach der initialen Detektion der Oberlippenkante und der Kante der Mundspalte (links) werden diese bis zu ihrem Ende verfolgt (Mitte). Der Schnittpunkte der Kanten definert die Mundecken (rechts).

3.3 Detektionsergebnisse

Das vorgestellte Verfahren wurde an über 100 Gesichtsbildern von verschiedener Individuen getestet. Der gesamte Detektionsfehler für die vollständige Detektion der Iris in Augenregionen betrug dabei 2.35%. In 94.4% der verbleibenden Bildausschnitte konnte die äußere Augenecke erfolgreich detektiert werden, während die innere Augenecke in 91.6% der Fälle richtig gefunden werden konnte. Bei den Mundwinkeln wurde eine Trefferrate von 78,2% erzielt, da diese häufig durch Verschattungen und andere ungünstige Bedingungen schwer erkennbar waren. Die Nasenflügelbreite konnte in 93,4% der untersuchten Fälle richtig bestimmt werden, was eine erfolgreiche Detektion beider Kanten der Nasenflügelränder voraussetzte. Die Detektionsgenauigkeit des vorgestellten Verfahrens liegt bei 1-2 Pixeln.

Das sequentielle Suchschema weist ein hohes Maß an Invarianzeigenschaften gegenüber Grösse der Bildregionen (z.B. Irisdurchmesser zwischen 27 und 104 Pixel)(vgl. Abb. 5b,c), Orientierung (+/- 20 Grad) (Abb. 5b), Rauschen und Qualität der Bildausschnitte auf. Die große Variabilität der erfolgreich untersuchten Bildregionen wird an den Beispielen in Abb. 5 demonstriert.

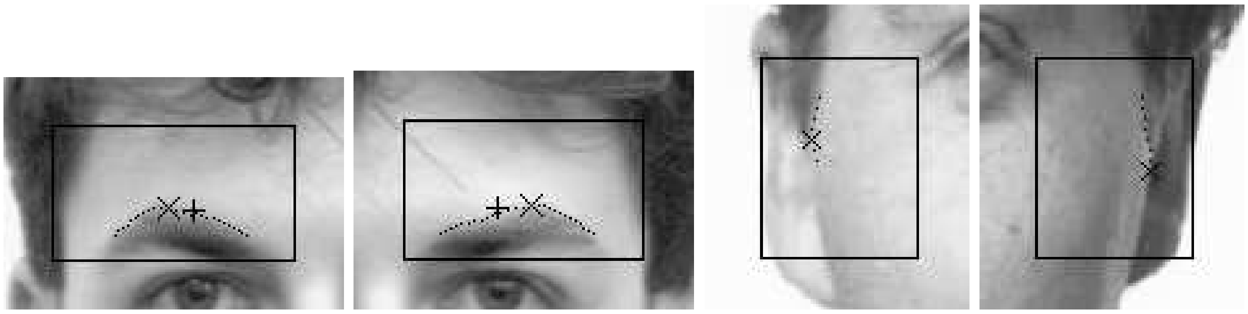


(a) (b) (c)

Abbildung 5: Beispiele von erfolgreich untersuchten Augenregionen. Alle Bilder besitzen den gleichen Maßstab.

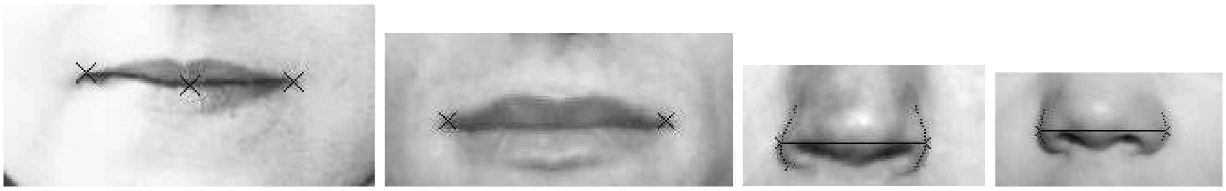
4 Zusammenfassung und Diskussion

Die in dieser Arbeit vorgestellte Vorgehensweise ermöglicht eine robuste und präzise Lokalisierung von markanten Punkten in Gesichtsbildern. Die Robustheit des Verfahrens ist erstens durch die Implementation von Modellwissen in den Detektionsvorgang zu erklären,



(a) (b)

Abbildung 6: Beispiele von erfolgreich detektierten Augenbrauenkanten (a) und Jochbogenpunkten (b).



(a) (b)

Abbildung 7: Beispiele von erfolgreich detektierten Mundecken (a) und Nasenflügelabstand (b).

was die Berücksichtigung des Bildkontextes ermöglicht. Zweitens sind die verwendeten Filter mit hoher Orientierungsselektivität weniger anfällig gegen Störungen des Kanten- oder Linienverlaufes durch Falten, Wimpern, Schatten oder andere unerwartete Strukturen. Das sequentielle Suchschema läßt sich durch weitere Konsistenzüberprüfungen (Verfeinerung der Modelle) weiter optimieren. Darüber hinaus können die Detektionsfehler durch einen sich anschließenden Verifikationsschritt der gefundenen Bildposition reduziert werden [6].

Die Adaptivität des verwendeten Filterschemas stellt eine wesentliche Innovation gegenüber einer auf "einfachen" Kantendetektoren basierenden Kantenverfolgung dar. Das verwendete steuerbare Filterschema trägt wesentlich zur Robustheit und Effizienz des sequentiellen Suchschemas bei.

Literatur

[1] S. Stengel-Rutkowski and P. Schimanek, *Chromosomale und nichtchromosomale Dysmorphiesyndrome*, Enke Verlag, Stuttgart, 1985.

[2] R. Herpers et al., *GAZE: An attentive processing strategy to detect and analyze the prominent facial regions*, Proc. Int. Works. on Auto. Face- and Gesture-Rec., Zurich, 214-220, 1995.

[3] P. Perona, *Steerable-scalable kernels for edge detection and junction analysis*, ECCV'92, G. Sandini (Ed.), LNCS 588, Springer-Verlag, 3-18, 1992.

[4] M. Michaelis, *Low level image processing using steerable filters*, PhD thesis, Christian-Albrechts-Universität, D-24105 Kiel, Germany, 1995.

[5] R. Herpers et al., *Edge and keypoint detection in facial regions*, 2. int. Conf. on Automatic Face and Gesture Recognition, 14.-16.10.96, Killington, (im Druck).

[6] L. Witta et al., *Anwendung eines DCS-Netzwerkes zur Verifikation von charakteristischen Bildpunkten*, in diesem Tagungsband, Workshop "Bildverarbeitung für die Medizin", 1996.