

Regelbasierte Kantenerkennung zur schnellen kantenbasierten Segmentierung der Glottis in Hochgeschwindigkeitsvideos

Ralph Maschotta, Simon Boymann und Ulrich Hoppe

Institut für Biomedizinische Technik und Informatik,
Technische Universität Ilmenau, 98684 Ilmenau

Zusammenfassung. Kantenerkennungsalgorithmen werden für die Lösung verschiedenster Aufgabenstellungen in verschiedenen Gebieten der Bildverarbeitung verwendet. Dabei wird meist mit Hilfe der ersten oder zweiten Ableitung die Position des steilsten Anstieges einer Kante gesucht. Für die im Folgenden betrachtete Aufgabenstellung der Segmentierung der Glottis in Hochgeschwindigkeitsvideos wurde ein anderes Vorgehen gewählt. Das Bild wird mit eindimensionalen Gradientenfiltern unterschiedlicher Länge gefaltet und die Ergebnisse mit Hilfe eines deterministischen endlichen Automaten derart verknüpft, dass der Beginn und das Ende einer Kante erkannt wird. Zur Beschreibung dieser Kante wird ein linearer Verlauf angenommen und der Beginn, die Länge sowie die Höhe der Kante in einem Merkmalsvektor abgespeichert.

1 Einleitung

Um endoskopische Hochgeschwindigkeitsaufnahmen von Stimmlippenbewegungen zu analysieren, ist die Ermittlung aussagefähiger Merkmale notwendig. Es werden Merkmale gesucht, die auch die örtlichen Unterschiede berücksichtigen. Ausgangspunkt hierfür ist die möglichst genaue Segmentierung der Stimmlippen. Der funktionell interessierende Teil ist die einfacher zu segmentierende Stimmritze (Glottis). Für die Segmentierung der Glottis wurde ein kantenbasiertes Verfahren entwickelt, bei dem die Glottis ohne Nutzerinteraktion segmentiert wird [1]. Im Folgenden soll die hierfür verwendete Kantenerkennung vorgestellt werden. Dabei ist nicht die Position des maximalen Anstieges, sondern die Position des Anfangs und des Endes einer Kante von Interesse. Die Anforderungen, die sich hieraus für die Kantenerkennung ableiten lassen, sind sehr hoch. Zum einen sollen die bei dieser Aufgabenstellung häufig vorkommenden langen Kantenverläufe korrekt erkannt werden. Zum anderen soll deren Position exakt ermittelt werden. Da die Bilder eine sehr geringe Auflösung haben, ist eine möglichst pixelgenaue Positionierung wünschenswert. Um den durch die hohe Aufnahmefrequenz bedingten immensen Datenaufwand zu reduzieren, ist eine geringe Anzahl von Kantenpunkten notwendig. Diese sollten jedoch die nötigen Bildinformationen repräsentieren. Verbreitete Kantenerkennungsalgorithmen, wie der nach Sobel oder Canny können lange Kanten nur dann als Ganzes erkennen, wenn das Bild

vorher geglättet wird oder die Filterlänge sehr groß gewählt wird. Für die exakte Positionsbestimmung ist allerdings ein möglichst kurzer Filter notwendig. Im Folgenden soll ein Kantenalgorithmus vorgestellt werden, bei dem einfache Gradientenfilter unterschiedlicher Länge mit Hilfe eines deterministischen endlichen Automaten so miteinander kombiniert werden, dass sowohl lange Kanten erkannt werden, als auch deren Position genau bestimmt werden kann. Im Gegensatz zu anderen Verfahren wird hierbei nicht die Position des maximalen Anstiegs, sondern der Beginn und das Ende der Kante detektiert.

2 Methoden

Für die Entwicklung der Segmentierung wurden Aufnahmen von 18 weiblichen Probanden mit normaler Stimmfunktion im Alter zwischen 19 und 23 Jahren verwendet. Die Bilder wurden mit einer endoskopischen Hochgeschwindigkeitskamera mit einer Bildrate von 3704 Bildern pro Sekunde und einer Auflösung von 128x64 8-Bit-Grauwertbildpunkten aufgenommen [2]. Aufnahmebedingt liegt die Glottis so im Bild, dass waagerechte Kanten für die Segmentierung ausreichend sind. Aus diesem Grund werden ausschließlich eindimensionale Filter zeilenweise auf das Bild angewendet.

Für die Segmentierung ist eine möglichst genaue Bestimmung des Anfangs und des Endes der Kante erforderlich. Dazu wird ein Gradientenfilter mit der Länge von zwei Pixeln verwendet (1). Als Aufpunkt wird dabei das erste Pixel verwendet.

$$h(n) = [1, -1] \quad (1)$$

Solche Filter sind sehr rauschanfällig, was die Verwendung eines weiteren Filters erforderlich macht. Dazu wird das Originalbild mit verschiedenen Binomial-Filtern unterschiedlicher Länge geglättet. Auf den Ergebnisbildern wird jeweils ein Gradientenfilter angewendet. Als Aufpunkt wird auch hier das erste Pixel verwendet (2). Die Ergebnisse der Filterungen werden addiert. Aufgrund der Linearität der Faltung können die unterschiedlich langen Filterkombinationen zu einem Gesamtfilter zusammengefasst werden (3). Fehlende Filterkoeffizienten werden durch Nullen aufgefüllt. Die Länge dieses zweiten Filters (3) hängt von der Auflösung des Bildes ab.

$$\begin{aligned} t(n) &= [1/2, 1/2] \\ h_0(n) &= h(n) \\ h_1(n) &= t(n) * h(n) \\ h_1(n) &= t(n) * t(n) * h(n) \\ &\vdots \\ h_i(n) &= (t_1(n) * \dots * t_m(n)) * h(n) \end{aligned} \quad (2)$$

$$h_g(n) = \sum_{i=0}^m h_i(n) \quad (3)$$

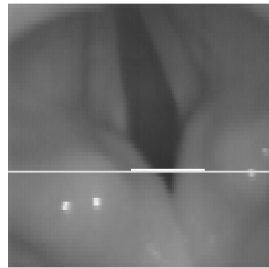
Die Ergebnisse der Faltung werden nun mit Hilfe eines deterministischen endlichen Automaten (siehe Abb. 1b) verknüpft. Der Bereich in dem der Gradient von h (1) ungleich Null ist und das gleiche Vorzeichen besitzt, wird als Kantenbereich definiert. Der Automat besitzt die drei Zustände: kein Kantenbereich, Kantenbereich mit positiven Gradienten (Vorderkante (VK)) und Kantenbereich mit negativen Gradienten (Hinterkante (HK)). Der Beginn des Kantenbereiches wird ausschließlich durch den Filter h bestimmt (in Abb. 1b $F0=h$). Rauschen führt zu einer häufigen Unterbrechung des Kantenbereichs. Daher wird der Filter h_g (3) verwendet, um die Rauscheinflüsse auf einem langen Anstieg zu mildern. Der Automat wird so erweitert, dass das Ende des Kantenbereiches durch eine Oder-Kombination zwischen h und h_g bestimmt wird (siehe Abb. 1b $F0=h$ und $F1=h_g$). Der Filter $F1$ muss so konstruiert werden, dass er das Ende einer Kante korrekt detektiert. Die Länge von $F1$ bestimmt somit, wie lang ein Plateau innerhalb eines Anstiegs sein darf, damit der Kantenbereich auf einer langen Kante nicht unterbrochen wird. Für die Vermeidung von zu vielen Kantenbereichen, kann zusätzlich die Empfindlichkeit des Automaten durch die Schwellen (S_0 und S_1) verändert werden. Für diesen Anwendungsfall werden die Schwellen auf Null gesetzt.

Für die ermittelten Kanten wird zur Vereinfachung ein linearer Kantenverlauf angenommen und für jede ermittelte Kante, die Position des Kantenbeginns, die Kantenlänge und die Gesamtkantenhöhe in einer Merkmalsliste abgespeichert. Für weitere Analysen oder zur nichtlinearen Beschreibung der Kante kann zusätzlich die Position und Höhe des stärksten Anstiegs, sowie der Winkel der Kante gespeichert werden. Um zu viele Kantenpunkte zu vermeiden, werden nur die Kantenpunkte in die Merkmalsliste übertragen, deren Gesamtkantenhöhe eine Schwelle überschreitet. Für einfache Anwendungen reicht ein Filter, mit entsprechenden Eigenschaften aus (wobei in Abb. 1b $F0 = F1$). Auch Standardfilter können als Grundlage für den Automaten verwendet werden.

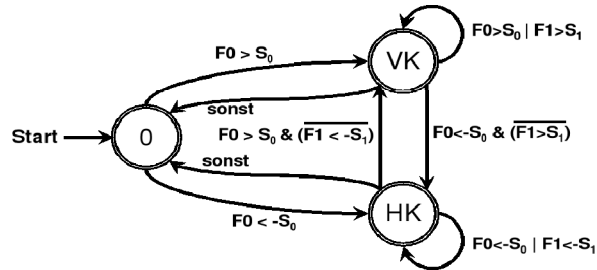
3 Ergebnisse

In Abb. 1d sind die Ergebnisse der Filterung mit $F0$ und $F1$ sowie die ermittelten Kantenbereiche von 1a dargestellt. Der Filter $F1$ hat eine Länge von 6 Pixeln. Wie dargestellt, werden die langen Kanten beim Übergang zwischen Stimmlippe und Glottis gut erfasst. Dabei wird der Beginn und das Ende der Kanten gut erkannt. Kanten mit einer Kantenlänge von einem Pixel werden ebenso gut erkannt. Rauschen wirkt sich stark auf das Ergebnis der Kantenerkennung aus. Auf langen Anstiegen wird das Ergebnis durch den langen Filter $F1$ nicht so stark von Rauschen beeinflusst. Diese Kanten besitzen eine höhere Gesamtkantenhöhe als durch Rauschen bedingte Kanten, wodurch diese sehr gut voneinander getrennt werden können. Salz- und Pfeffer-Rauschen wirkt sich sehr stark auf das Ergebnis aus. Diese Kanten sind jedoch sehr kurz, wodurch auch diese von anderen Kanten unterschieden werden können. Die zusätzlichen Informationen, welche die Kantenerkennung liefert, wie zum Beispiel die Länge oder die Gesamtkantenhöhe, werden für die Segmentierung bzw. Ausreisserbehandlung

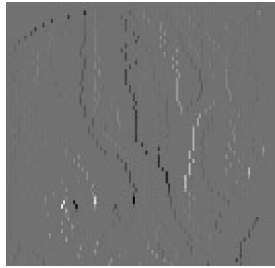
Abb. 1. a) Originalbild; b) deterministischer endlicher Automat. VK (Vorderkante)-Positiver Kantenbereich, HK (Hinterkante) negativer Kantenbereich, F0 und F1 Filterergebnisse, S_0 und S_1 Schwellen; c) Ergebnisbild. Schwarz negativer Kantenbeginn und weiss positiver Kantenbeginn; d) Ergebnisverlauf des im Originalbild markierten Zeilenausschnitts. Positiver Kantenbereich +20, negativer Kantenbereich -20



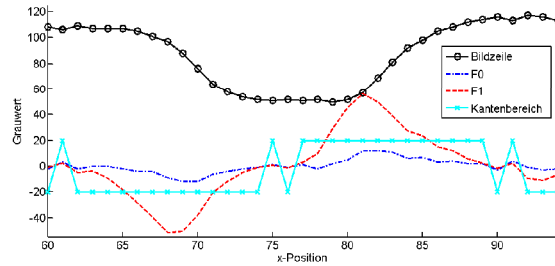
(a)



(b)



(c)



(d)

verwendet. Pro Anstieg wird nur ein Merkmalsvektor abgespeichert, so dass der Speicherbedarf bei rauscharmen Bildern sehr stark sinkt.

4 Zusammenfassung und Diskussion

Die vorgestellte Kantenerkennung basiert auf Gradientenfiltern unterschiedlicher Länge. Die Ergebnisse werden mit Hilfe des vorgestellten deterministischen endlichen Automaten zusammengefasst. Durch den ersten Filter wird die Genauigkeit des Kantenbeginns bestimmt, der zweite Filter bestimmt das Abbruchkriterium innerhalb des Kantenbereichs und somit das Ende des Kantenbereichs. Die für diese Aufgabenstellung angepasste Kombination aus einem einfachen Gradientenfilter und der Summation aus Binomial- und Gradientenfilter liefert gute Ergebnisse und kann auch für andere Aufgabenstellungen verwendet werden.

Für Anwendungen, bei denen sowohl senkrechte als auch waagerechte Kanten von Bedeutung sind, wird die Kantenerkennung in den entsprechenden Richtun-

gen ausgeführt. Die Ergebnislisten werden aneinandergehängt. Dadurch können in der Merkmalsliste an einer Position mehrere Kantenpunkte vorhanden sein. Bei Rotationsoperationen auf der Merkmalsliste ist die Rotationsabhängigkeit der Kantenerkennung zu beachten.

Die vorgestellte Kantenerkennung reagiert empfindlich auf Rauschen. Lange und große Anstiege sind davon nicht so stark betroffen, da $F1$ einen starken Glättungsfilter beinhaltet. Durch nichtlineare Filterung des Ausgangsbildes, kann der Rauscheinfluss gemindert werden. Jedoch verschlechtert sich dadurch die hohe Genauigkeit der Positionierung vor allem bei sehr kurzen Kanten. Auch ohne vorherige Filterung beeinflusst Rauschen die genaue Bestimmung des Kantenbereiches. In Abb. 1d ist ein daraus resultierendes Problem zu erkennen: Wenn sich vor einem großen Anstieg durch Rauschen, ein kleiner Peak befindet, wechselt an dieser Stelle der Automat in den entsprechenden Kantenbereichszustand. Da der lange Filter den folgenden Anstieg bereits erkennt, wird der Kantenbereich nicht mehr unterbrochen, auch wenn anschließend der Wert sogar unter den Anfangswert sinkt. Dies führt zu einer unbeabsichtigten Verlängerung des Kantenbereiches. In Abb. 1d beginnt der Kantenbereich an Position 77, er müsste allerdings erst an Position 79 beginnen. Ist der Peak groß genug, werden die Kanten korrekt getrennt. Somit ist dieses Problem nicht gravierend. Um dieses Problem dennoch zu beheben, kann man für $F0$ und $F1$ die entsprechenden Schwellen S_0 und S_1 erhöhen. Durch diese Schwellen wird jedoch die Genauigkeit der Position des Anfangs bzw. des Endes einer Kante mit steigendem Gradienten, verschlechtert. Eine andere Möglichkeit dieses Problem zu beheben, besteht darin, den Anfang des Kantenbereiches an die Stelle zu verschieben, an der innerhalb des erkannten Kantenbereiches das tatsächliche Maximum bzw. Minimum liegt.

Die für die Aufgabenstellung wichtigen langen Kanten werden in den untersuchten Bildern gut erkannt. Die berechnete Länge und Höhe der Kante nutzt das Segmentierungsverfahren aus. Die vorgestellte Kantenerkennung bildet somit eine gute Grundlage für die kantenbasierte Segmentierung, die ohne Nutzerinteraktion die Glottis mit 200 Bildern pro Sekunde segmentiert. Der Vergleich mit einem Standardsegmentierungs-Verfahren erreichte im Mittel einen guten Tanimoto-Koeffizienten von 0,68.

Literaturverzeichnis

1. Maschotta R, Boymann S, Hoppe U. Kantenbasierte Segmentierung der Glottis in Hochgeschwindigkeitsvideos von Stimmlippenschwingungen. In: Procs. Biomedizinische Technik (BMT) 2004; 2004. p. 918–919.
2. Hoppe U, Rosanowski F, Lohscheller J. Visualisierung und Interpretation von Stimmlippenschwingungen. In: Procs BVM; 2003. p. 140–243.