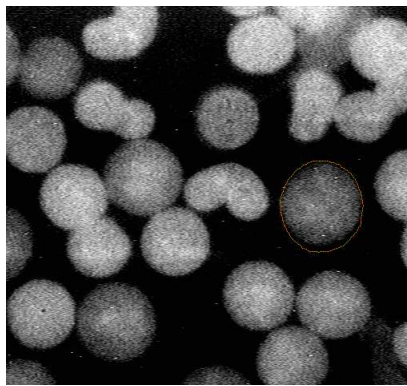


Kantenextraktion mit Snakes

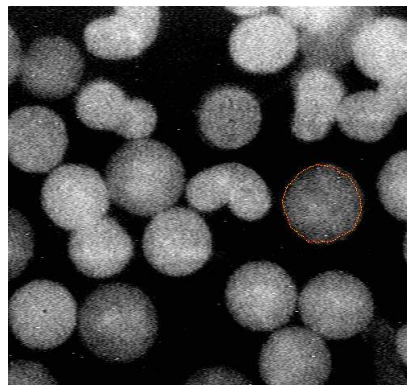
Seminar Bildsegmentierung und Computervision
WS2005/2006 Universität Ulm

Christian Hägele

1 Einführung



(a) Ausgangssnake



(b) stabiler Zustand der Snake

Abbildung 1: Einleitendes Beispiel von [4]

In der *Kantenextraktion* geht es darum markante Kanten in einem Bild zu finden, um so einzelne Objekte bestimmen zu können. Mit einer *Snake* wird dabei versucht *genau ein* Objekt in einem Bild über seine Kanten zu erfassen. Das steht im Gegensatz zu den meisten anderen Methoden der Bildsegmentierung, die versuchen das gesamte Bild sinnvoll zu segmentieren.

Eine Snake ist dabei nichts anderes als eine explizite Parameterdarstellung einer *geschlossenen Kurve*.

Die Snake muss *nahe* an dem gesuchten Objekt *initialisiert* werden (vgl Abb 1). Ist die Initialisierung nicht nah genug am gesuchten Objekt, findet die Snake oft nicht die Kanten des gesuchten Objekts sondern irgendwelche anderen Kanten im Bild. Die Snake zieht sich dann schrittweise zusammen um die Kanten des Objekts zu finden, bis sie einen stabilen Zustand erreicht.

1.1 Definitionen

Graubilder

Im Folgenden werden nur Graustufenbilder betrachtet.

$u(x, y)$ ist der Grauwert des Bildes am Punkt (x, y) . $u(x, y)$ wird die *Graustufenfunktion* des Bildes genannt.

Kanten

Eine Kante im Bild wird über ihren Kontrast beschrieben. Wo der Kontrast die größte Veränderung bzw. die größten Sprünge macht liegt auch die Kante. Sprünge bzw Veränderungen werden bekanntlich über die Ableitung beschrieben. Deshalb wird hier die 2. Ableitung der Graustufenfunktion des Bildes betrachtet.

Definition einer Kante nach *Hildreth und Marr*:

$$\Delta u(x, y) = \frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y) \quad (1)$$

An der Nullstelle dieses Laplace-Operators befindet sich die Kante.

Kurven

Wie bereits erwähnt, ist eine Snake eine explizite Parameterdarstellung einer geschlossenen Kurve. Deshalb werden zunächst einige Grundlagen, die im Folgenden gebraucht werden, definiert:

- $\gamma(s)$ Kurve $\gamma : [a, b] \mapsto \mathbb{R}^2$ (Bildebene)
- s ist die Euklidische Parametrisierung, d.h. $|\gamma'| = 1$
Dies setzt man nur voraus, um sich Rechen- und Schreibearbeit zu sparen. Es wird auch nur im ersten, theoretischen Teil verlangt. Später wird auch der allgemeine Fall einer nicht euklidischen Parametrisierung betrachtet.
- $L(\gamma)$ ist die Länge der Kurve γ
Weil $L(\gamma) = \int_a^b |\gamma'(s)| ds$ ist, sieht man, dass im Falle der Euklidischen Parametrisierung die Länge der Kurve gerade die Länge des Intervalls $[a, b]$ ist.
- $\vec{n}(s) = \gamma'(s)^\perp$ wobei $v = (x, y)$ $v^\perp = (-y, x)$
- $Du = \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \end{pmatrix}$
Also ist Du nichts anderes als der *Gradient der Graustufenfunktion*. In praktischen Anwendungen muss man diesen Gradienten aber auf einem *diskreten* Bild berechnen. Dafür gibt es unterschiedliche Verfahren, z.B. 3×3 Differenzen. Auf die verschiedenen Verfahren wird hier aber nicht näher eingegangen.

1.2 Allgemeines Vorgehen zur Kantenbestimmung mit Snakes

Es wird zunächst eine Ausgangssnake nahe des gesuchten Objekts vorgegeben. Diese Näherung sollte möglichst gut sein und kann in Anwendungsfällen beispielsweise durch

ein Polygonzug realisiert werden. Diese Ausgangssnake hat sehr großen Einfluss auf das Endergebnis und ist deshalb bereits ein sehr wichtiger Faktor der Segmentierung. Damit die Kurve jetzt die Kante “findet“, muss man den Gradient der Graustufenfunktion entlang der Kurve betrachten. Die Kurve wird dann in Richtung des größten *Kontrastes* “gezogen“. Das geschieht schrittweise so lange, bis die Kurve an einer Stelle ankommt, an der bereits maximaler Kontrast herrscht. Also ein *lokales Extremum* ist. An dieser Stelle befindet sich dann auch eine *Hildreth-Marr Kante*, was später noch genauer gezeigt wird.

1.3 Verschiedenen Modelle zur Bestimmung von Snakes

Für das beschriebene Vorgehen gibt es einige verschiedene Modelle, von denen ich nur wenige vorstellen werde. Aber alle diese Modelle gehen auf ein erstes Snake Modell zurück. Auf das *Kass-Witkin-Terzapoulos* Snake Modell von 1988.

Kass-Witkin-Terzapoulos Modell

$$\int_0^{L(\gamma)} g(|D\gamma(s)|) ds + C \int_0^{L(\gamma)} (a + |\gamma''(s)|) ds \rightarrow \min \quad (2)$$

Wie man sieht, besteht das zu minimierende Funktional aus zwei Summanden. Der erste Summand gibt den Kontrast entlang der Kurve an. Der zweite Summand die “Glattheit“ der Kurve. Es wird also eine Snake gesucht, die so glatt wie möglich ist und gleichzeitig den stärksten Kontrast hat.

Der zweite Summand wird auch *innere Energie* genannt. Er beinhaltet zwei Parameter C und a . Wobei man unter der “Glattheit“ auch eine bestimmte Form der Snake verstehen kann. Dort könnte man also Vorwissen über das Aussehen des gesuchten Objekts einbringen.

Der erste Summand wird auch *äußere Energie* genannt. Er beinhaltet nur einen Parameter. Die Funktion g . g wird *Kontrastfunktion* genannt und ist in diesem Modell *fallend*. Es fällt auf, dass nur die äußere Energie Informationen aus dem Bild, also aus $u(x, y)$ beinhaltet. Deshalb ist man auf die Idee gekommen die innere Energie einfach wegzulassen. Das einfache Streichen des zweiten Summanden hat aber nicht zum Erfolg geführt, da die Snake immer zu kurz wurde, was daran lag, dass über die Länge der Snake integriert wird. Deshalb ist man auf das folgende Modell gekommen.

Fua-Leclerc-Modell

Es wird der *maximale durchschnittliche* Kontrast gesucht:

$$\frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(|D\gamma(s)|) ds \rightarrow \max \quad (3)$$

Hier wurde noch ein Faktor $\frac{1}{L(\gamma)}$ vor das Integral gesetzt, um das zu starke Verkürzen der Kurve zu verhindern. In diesem Modell ist g irgendeine *wachsende*, positive und gerade Funktion. Deshalb wird das Funktional in diesem Fall auch maximiert. Die innere Energie bzw. das Aussehen der Kurve wird nichtmehr explizit betrachtet. Deshalb

fallen auch zwei Parameter aus dem Modell heraus und als einziger Parameter bleibt die Kontrastfunktion g übrig, auf die später noch genauer eingegangen wird.

Eine weitere Verbesserung von (3) wurde von *Montanari und Martelli* gemacht. Anstatt nur $|Du(\gamma(s))|$ wird der *Kontrast senkrecht zur Kurve* betrachtet:

$$u_n(s) = \frac{\partial u}{\partial \vec{n}}(s) = Du(\gamma(s)) \bullet \vec{n}(s)$$

So hängt der Wert von $u_n(s)$ sowohl von der Größe von $|Du|$, als auch von der *Richtung* des Gradienten ab. Es wird *nur* der Kontrast senkrecht zur Kurve betrachtet. Das ist auch sinnvoll, da das auch der Kontrast ist, der für die Bewegung der Snake am Wichtigsten ist. Deshalb wird im Folgenden auch nur noch dieser Term verwendet. Also ergibt sich als verbessertes Funktional, das maximiert werden muss:

$$\frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(u_n(s)) \, ds \quad (4)$$

Grundsätzlich muss man bemerken, dass die globalen Extrema der Modelle irrelevant sind, da sie immer in einem Punkt mit dem größten Kontrast zusammenfallen würden. Deshalb wird hier nur nach lokalen Extrema gesucht. Diese lokalen Minima hängen aber noch stark von der Wahl der Kontrastfunktion g und insbesondere der Ausgangssnake ab.

2 Entwicklung der Kurve in Richtung des maximalen Kontrast

Nachdem man jetzt ein Modell gegeben hat, das den Kontrast oder die Energie entlang der Kurve angibt, geht es darum die Kurve in die Richtung des maximalen Kontrastes zu bewegen. Dazu muss man aber die *Kontraständerung* bzw. die *Ableitung des Energiefunktionals* betrachten.

2.1 Ableiten des Kontrast Funktionals

Da das Energiefunktional aber von der Kurve $\gamma(s)$, also von einer Funktion, abhängt, ist die Bestimmung der Ableitung nicht trivial. Man benützt dafür eine *Gateaux Ableitung*.

Gateaux Ableitung

Wir betrachten zunächst die Kimmel-Bruckstein Energie:

$$E(\gamma) = \int_0^{L(\gamma)} g(u_n(s)) \, ds \quad (5)$$

Das ist die Energie (4) ohne den Faktor $\frac{1}{L(\gamma)}$ davor. Das macht man um sich das ganze in einem ersten Schritt zu vereinfachen. Um diese Energie abzuleiten verwenden wir die *Gateaux Ableitung*.

Bei einer *Gateaux Ableitung* wird in die ‘‘Richtung‘‘ einer *Funktion*, in diesem Fall $\varepsilon : [0, L(\gamma)] \mapsto \mathbb{R}^2$ abgeleitet:

$$\nabla_{\varepsilon} E(\gamma) = \lim_{\lambda \rightarrow 0} \frac{E(\gamma + \lambda\varepsilon) - E(\gamma)}{\lambda} = \frac{d}{d\lambda} E(\gamma + \lambda\varepsilon) \Big|_{\lambda=0}$$

Ableitung der Kimmel-Bruckstein Energie

Man hat jetzt eine Ableitung von (5), die aber noch von der Funktion ε abhängt. Deshalb betrachtet man im nächsten Schritt nur die Ableitung in *Normalenrichtung* \vec{n} . Vereinfacht ergibt sich dann als Ableitung von (5), die nichtmehr von ε abhängt:

$$\nabla E(\gamma) = (g'(u_n))'(Du^{\perp}(\gamma) \bullet \vec{n})\vec{n} + g'(u_n)\Delta u(\gamma)\vec{n} - h(u_n)\gamma'' \quad (6)$$

wobei

$$h(t) = g(t) - tg'(t) \quad \text{und} \quad \Delta u(x, y) = \frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y)$$

Wie genau man auf diese Darstellung kommt will ich nicht weiter erläutern. Dazu sei auf [1] verwiesen.

Man kann diese Ableitung für Anwendungsfälle noch weiter vereinfachen. Zunächst setzt man die Kontrastfunktion $g(t) = |t|$. Die Funktion ist wachsend auf \mathbb{R}^+ , positiv und gerade. Erfüllt also alle Voraussetzungen und ist eine Standardform von $g(t)$. Durch diese Wahl ist $h(t) = g(t) - tg'(t) = 0$. Dadurch vereinfacht sich (6) zu:

$$\nabla E(\gamma) = ((g'(u_n))'(Du^{\perp}(\gamma) \bullet \vec{n})\vec{n} + g'(u_n)\Delta u(\gamma)\vec{n} \quad (7)$$

Im nächsten Schritt macht man eine weitere, meistens erfüllte Annahme. Wenn u_n konstantes Vorzeichen auf γ hat, ergibt sich $g'(u_n) = \text{sign}(u_n) = \pm 1$ und deshalb $(g'(u_n))' = 0$. Die Annahme vom konstanten Vorzeichen entlang γ besagt nichts anderes, als dass die Kurve komplett außerhalb oder komplett innerhalb des Objekts verläuft. Sieh also nur ausdehnt oder nur zusammenzieht, was bei einer sinnvollen Initialisierung meistens gegeben ist. (7) vereinfacht sich dadurch zu:

$$\nabla E(\gamma) = \text{sign}(u_n)\Delta u(\gamma)\vec{n} \quad (8)$$

In (8) sieht man jetzt gut, dass eine Hildreth-Marr Kante (vgl. (1)) ein stabiler Zustand der Snake ist, da dort $\nabla E(\gamma) = 0$ ist.

Damit hat man gezeigt, dass dieses Snake Modell auch wirklich in der Lage ist Hildreth-Marr Kanten zu finden. Das heißt, wenn die Snake an einer Kante angekommen ist bewegt sie sich nicht weiter.

Ableitung der Fua-Leclerc Energie

In diesem Modell wird eine geschlossene Kurve gesucht deren *durchschnittlicher Kontrast* maximal wird:

$$F(\gamma) = \frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(u_n(s)) ds = \frac{E(\gamma)}{L(\gamma)}$$

Das ist das eigentliche Snake Modell, wie in der Einleitung beschrieben. Auch hier macht man genau das gleiche wie zuvor bei der Kimmel-Bruckstein Energie.

Wie zuvor bildet man die Gateaux Ableitung in "Richtung" ε . Wegen $F(\gamma) = \frac{E(\gamma)}{L(\gamma)}$ ergibt sich zunächst nach der Quotientenregel:

$$\nabla_\varepsilon F(\gamma) = \frac{1}{L(\gamma)} (\nabla_\varepsilon E(\gamma) - F(\gamma) \nabla_\varepsilon L(\gamma))$$

Durch Einsetzen der bereits bekannten Ableitung von $E(\gamma)$ und Brechen von $\nabla_\varepsilon L(\gamma)$ bekommt man $\nabla_\varepsilon F(\gamma)$.

Mit den gleichen Vereinfachungen wie zuvor, also nur die Veränderung in Normalenrichtung betrachten, $g(t) = |t|$ und u_n ändert sein Vorzeichen entlang der Kurve nicht, ergibt sich folgendes:

$$\nabla F(\gamma) = \text{sign}(u_n) \Delta u(\gamma) \vec{n} + F(\gamma) \gamma'' \quad (9)$$

Wie man in (9) sieht, spielt auch γ'' , also die "Glattheit" der Kurve, wieder eine Rolle in diesem Modell.

Numerisches Modell

Nachdem wir jetzt im theoretischen Fall die Ableitung des Energie-Funktional bestimmt haben, betrachten wir jetzt das gleiche mit einer allgemeinen, nicht Euklidischen Parametrisierung $\gamma(p) : [a, b] \mapsto \mathbb{R}^2$. Die Energie, die maximiert werden soll, sieht dann folgendermaßen aus:

$$F(\gamma) = \frac{\int_a^b g(Du(\gamma(p)) \bullet \frac{\gamma'(p)^\perp}{|\gamma'(p)|}) |\gamma'(p)| dp}{\int_a^b |\gamma'(p)| dp} \quad (10)$$

Wenn man sich (10) genau anschaut, ist es nichts anderes als das bereits bekannte Fua-Leclerc Modell (4). Nur ist diesmal keine Euklidische Parametrisierung vorausgesetzt. Das ganze wird etwas komplizierter und umfangreicher, weshalb im ersten, theoretischen Teil auch eine Euklidische Parametrisierung vorausgesetzt war.

Für praktische Anwendungsbeispiele muss man aber noch einige Dinge beachten. Zunächst ist die Snake $\gamma(s)$ in der Praxis nicht als schöne Kurve, sondern als Polygonzug gegeben. Es wird also die Snake als Polygonzug M_1, \dots, M_n betrachtet. Daraus ergibt sich dann:

$$L = \sum_i \|\Delta_i\| \quad \text{mit} \quad \Delta_i = M_{i+1} - M_i$$

Jetzt kann man damit die diskrete Energie $F = \frac{E}{L}$ so schreiben:

$$E = \sum_i g(t_i) \|\Delta_i\| \quad L = \sum_i \|\Delta_i\|$$

wobei

$$t_i = \omega_i \frac{\Delta_i^\perp}{\|\Delta_i\|} \quad \omega_i = Du(\Omega_i) \quad \Omega_i = \frac{M_i + M_{i+1}}{2}$$

Also ergibt sich als Ableitung nach M_k

$$\nabla_{M_k} F = \frac{1}{L} (\nabla_{M_k} E - F \nabla_{M_k} L)$$

Da F von dem Polygonzug $(M_i)_i$ abhängt, muss auch nach jedem einzelnen M_k abgeleitet werden. So ist auch die letzte Gleichung zu verstehen. Es darf auch nicht vergessen werden, dass die Graustufenfunktion $u(x, y)$ jetzt eine diskrete Funktion ist und deshalb Du auch eine "diskrete Ableitung" darstellt, die über *diskrete Differenzen* berechnet wird.

Vorgehensweise

Eine grobe, allgemeine Vorgehensweise zum iterativen Berechnen bzw Verschieben von Snakes sieht dann folgendermaßen aus:

Zunächst wird die Snake auf Bogenlänge parametrisiert. Das bringt Stabilität in das Schema und Singularitäten treten nicht so schnell auf.

In zweiten Schritt, der *Gradientenentwicklung*, passiert die eigentliche Verschiebung. Sei $(M_i^n)_i$ der Polygonzug der die Snake im n -ten Schritt darstellt und $(\tilde{M}_i^n)_i$ die reparametrisierte Version nach Schritt 1, dann setzt man:

$$M_i^{n+1} = \tilde{M}_i^n + \delta \nabla_{\tilde{M}_i^n} F \quad \text{wobei } \delta \text{ die Schrittweite ist}$$

Es wird also der alte Punkt in Richtung des stärksten Kontrastes verschoben. Die Schrittweite δ bestimmt wie schnell die Verschiebung vonstatten geht, also wie viele Schritte insgesamt benötigt werden. Die Wahl von δ hängt hauptsächlich von der Größe des Bildes ab und wird meistens auf 1, 3 oder 5 gesetzt.

3 Wahl der Kontrastfunktion g

In diesem Teil geht es um die Wahl der Kontrastfunktion g . Es wird gezeigt, dass diese Funktion sehr großen Einfluss auf das Ergebnis der Snake hat. Die Energie $F(\gamma)$ wird erhöht, wenn man die Kurve in kontrastreichen Gegenden verlängert oder die Kurve in kontrastarmen Gebieten verkürzt. Wenn die Funktion g stärker wächst, ist es einfacher Bereiche mit niedrigem Kontrast abzuschneiden, was oft zu Fehlern führt. Deshalb sollte man eher eine schwach steigende Funktion g wählen. Folgende Beispiele sollen das verdeutlichen. Abb.2 zeigt ein künstlich erzeugtes Bild. Im Hintergrund verläuft die Helligkeit linear ansteigend von links nach rechts. Ganz links schwarz, ganz rechts weiß. Darauf ist ein weißes Quadrat abgebildet. Die Snake wird in Abb.2(a) genau an den Kanten des Quadrats initialisiert. Wie man in Abb.2(b) und 2(c) sehen kann werden bei schlechter Wahl von $g(t)$ unter Umständen große Fehler gemacht. Es läßt sich aus solchen experimentellen Überlegungen folgende Aussage machen. Hat man

$$F(\gamma) = \frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(u_n(\gamma(s))) ds \quad \text{mit} \quad g(t) = |t|^\alpha \quad (11)$$

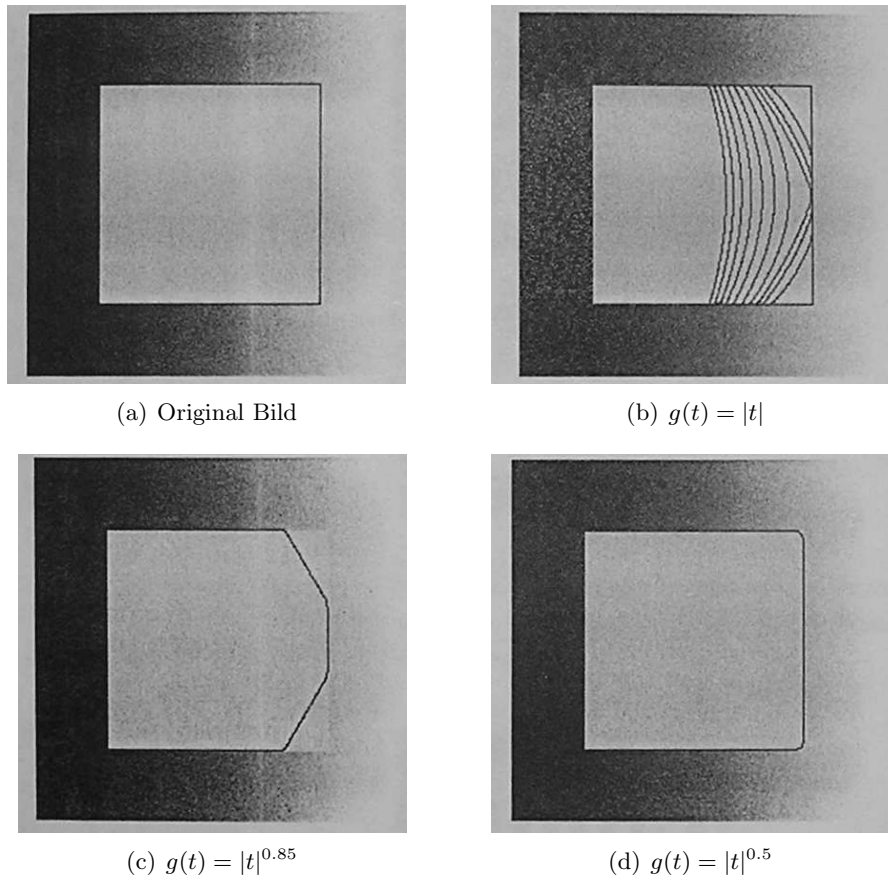


Abbildung 2: Verschiedene Kontrastfunktionen $g(t)$

gegeben. Dann wird bei großem α die Snake in Bereichen mit schwachem Kontrast gekürzt. Wobei $\alpha = 1$ bereits groß ist. Deshalb ist ein kleineres α meistens besser. Die Snakes sind meistens stabiler und treuer an den Kanten. Allerdings läßt sich kein optimales α für alle Arten von Bilder bestimmen. Auch ein zu kleines α führt zu Fehlern. Es werden dann womöglich Kanten erkannt, die nicht zum Objekt, sondern zum allgemeinen Bildrauschen oder Hintergrund gehören.

4 Fazit

Snakes funktionieren, im Gegensatz zu vielen anderen Verfahren der Kantenextraktion, nicht vollautomatisch. Man muss einige Parameter an das Problem anpassen.

Das Erstellen des Polygonzuges am Anfang muss von Hand erledigt werden. Diese Ausgangssnake hat auch sehr starken Einfluss auf das Endergebnis und muss möglichst nahe am Objekt initialisiert werden um brauchbare Ergebnisse zu liefern.

Die Kontrastfunktion g muss als Parameter an das Bild angepasst gesetzt werden. Man

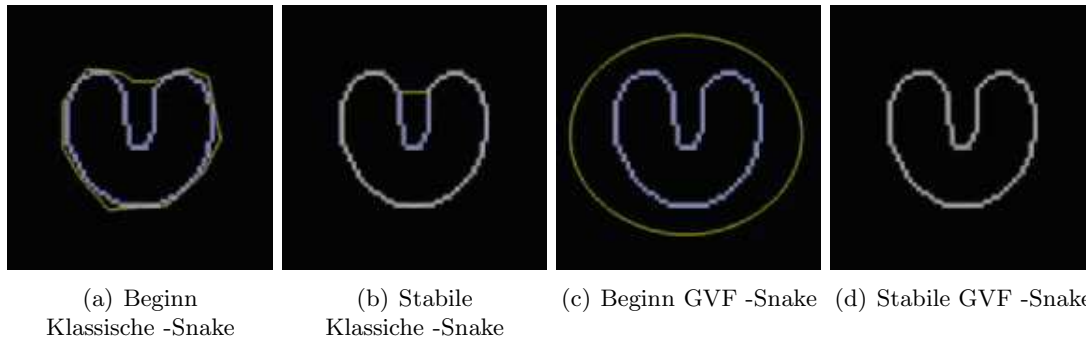


Abbildung 3: Klassische- vs. GVF-Snakes

hat gesehen, dass $g(t) = |t|^\alpha$ mit kleinem α meistens die besten Ergebnisse liefert. Einfluss auf das Ergebnis haben auch die Gradientenschrittweite δ und die Art, wie man die diskrete Ableitung Du praktisch berechnet. Diese Parameter hängen hauptsächlich von der Bildgröße ab.

Im Vergleich mit anderen Methoden zur Kantenextraktion und Bildsegmentierung bietet sich das Verfahren mit Snakes nur an, wenn man genau ein Objekt erfassen will. Darin besteht der Vorteil von Snakes, da sie immer ein geschlossenes Objekt als Ergebnis liefern. Allerdings muss man viele Parameter, insbesondere die Ausgangssnake, vorher auf das Problem angepasst setzen. Will man das gesamte Bild sinnvoll segmentieren, oder viele Objekte erfassen, sind andere, automatisierte Verfahren der Bildsegmentierung meist besser.

Die hier vorgestellten Snakes haben noch einige weitere Nachteile. Sie können sich nicht an weit entfernte Objekte "rantasten", da nur der Kontrast in der nächsten Umgebung der Snake betrachtet wird. Außerdem können sie sich auch nicht in Wölbungen oder Einbuchtungen ausbreiten (vgl. Abb.3(b)). Eine mögliche Lösung dieser Probleme sind *Gradient-Vector-Flow-Snake* (GVF-Snakes).

GVF-Snakes

Wie man in Abb.3(a) und 3(b) sieht kann sich eine klassische Snake nicht in solche Einbuchtungen wie bei diesem künstlich erzeugten U ausbreiten, obwohl eine sehr gute Näherung als Initialisierung angegeben wurde. Dieses Problem können *Gradient-Vector-Flow-Snakes* allerdings lösen, wie man in Abb.3(d) sieht. Außerdem muss die Ausgangssnake nicht so gut initialisiert werden und findet trotzdem noch weiter entfernte Kanten wie man in Abb.3(c) sieht.

GVF-Snakes lösen die Probleme, weil sie nicht nur den Kontrast in der nächsten Umgebung, sondern das "Kontrastvektoren-Feld" auf dem gesamten Bild betrachten. Solch ein *Gradient Vector Flow Field* ist in Abb.4 zu sehen. Ich möchte hier nicht weiter auf GVF-Snakes eingehen. Das wäre die Aufgabe einer eigenen Seminararbeit. Auf der Internetseite [2] stehen sehr viele Informationen zu diesem Thema bereit. Jedem, der sich

näher zu diesem Thema informieren will, sei diese Seite sehr empfohlen.

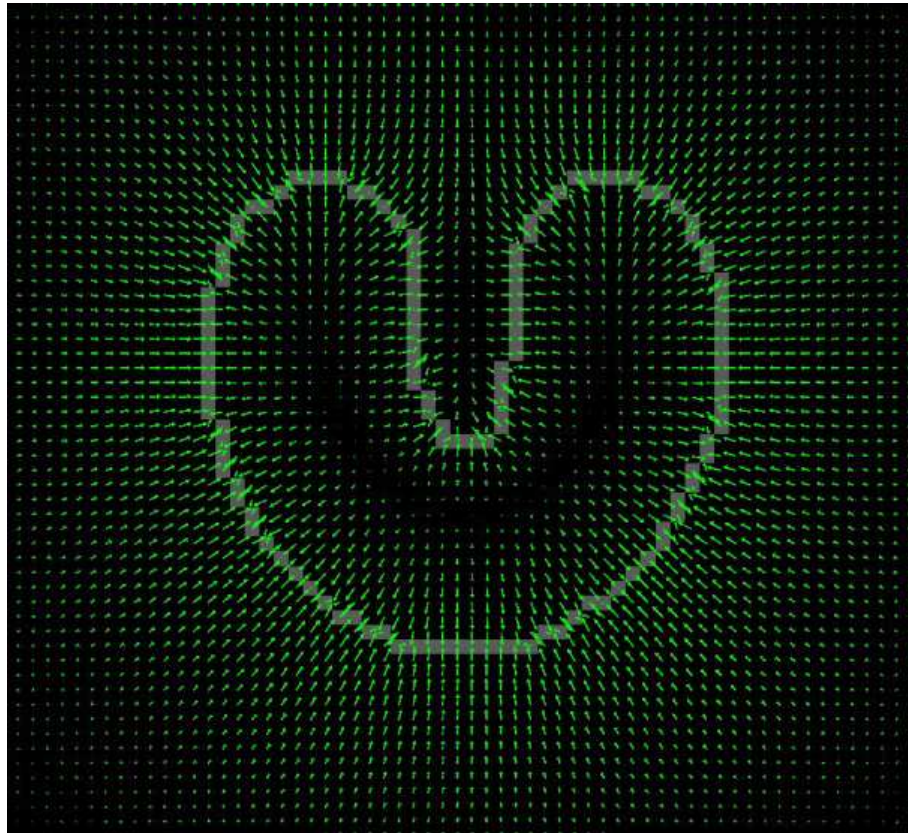


Abbildung 4: Gradient Vector Flow Field

Literatur und URLs

- [1] S. Osher, N. Paragios, eds.
Geometric level set methods in imaging, vision, and graphics
Kap II, §5
Springer, 2003
- [2] GVF-Snakes
<http://iac1.ece.jhu.edu/projects/gvf/>
- [3] Beispiele zu verschiedenen Snake Modellen
<http://iac1.ece.jhu.edu/~chenyang/research/levset/movie>
- [4] Java-Applet zum experimentieren mit Snakes
<http://www.markschulze.net/snakes/>