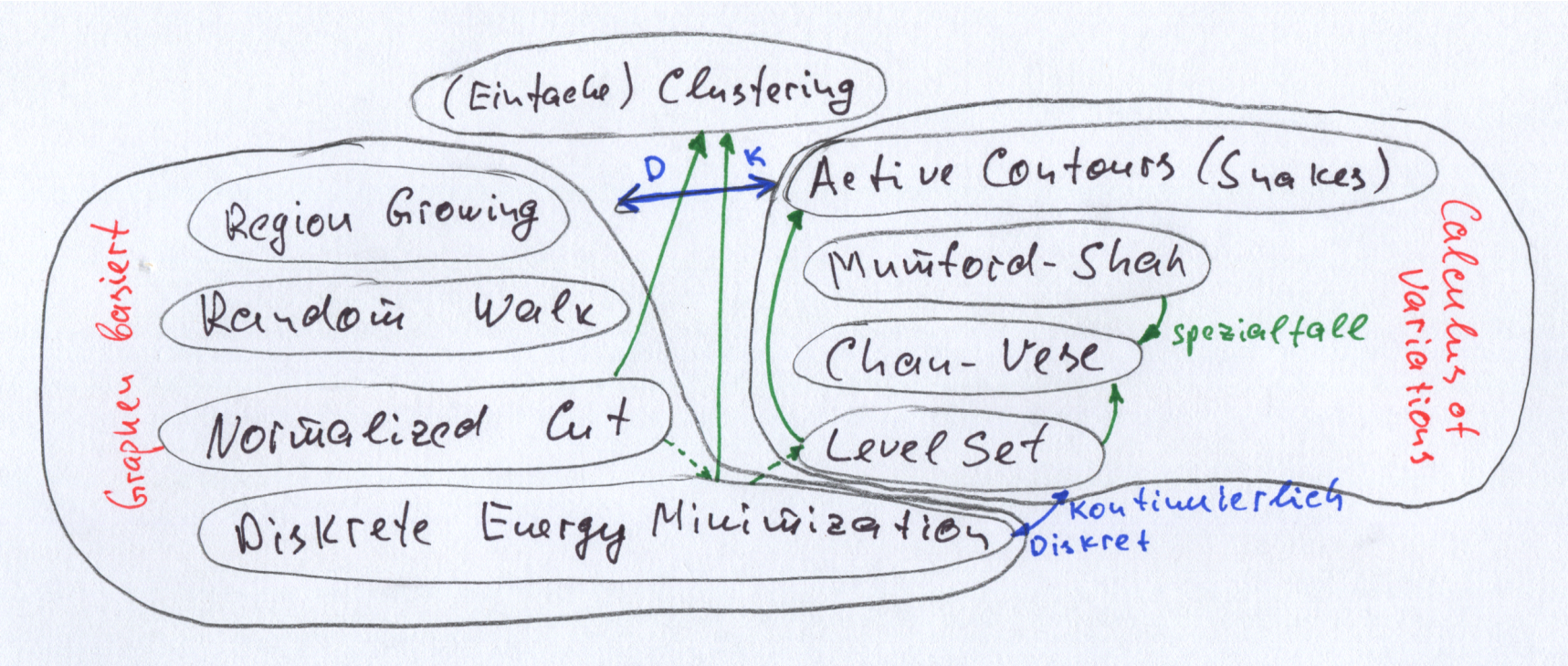
Computer Vision: Segmentierung I

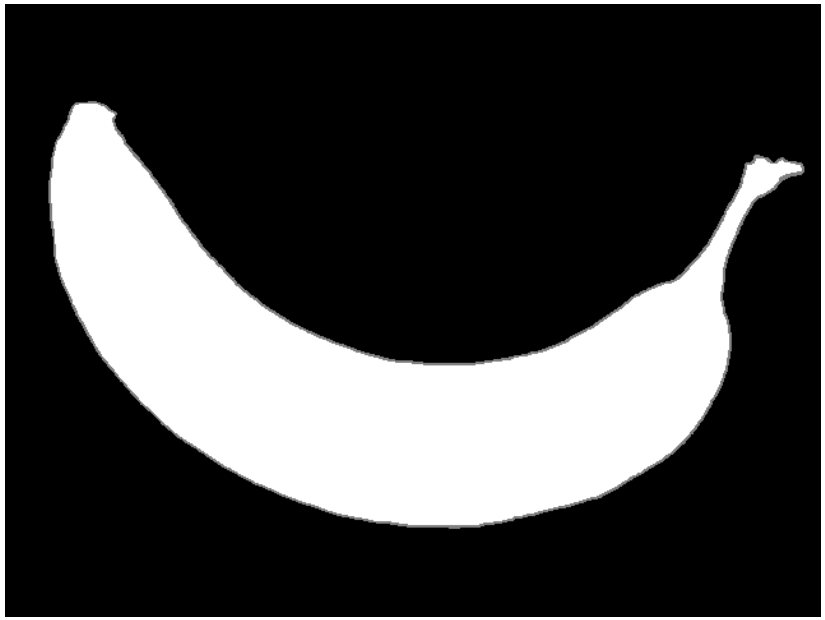
D. Schlesinger – TUD/INF/KI/IS

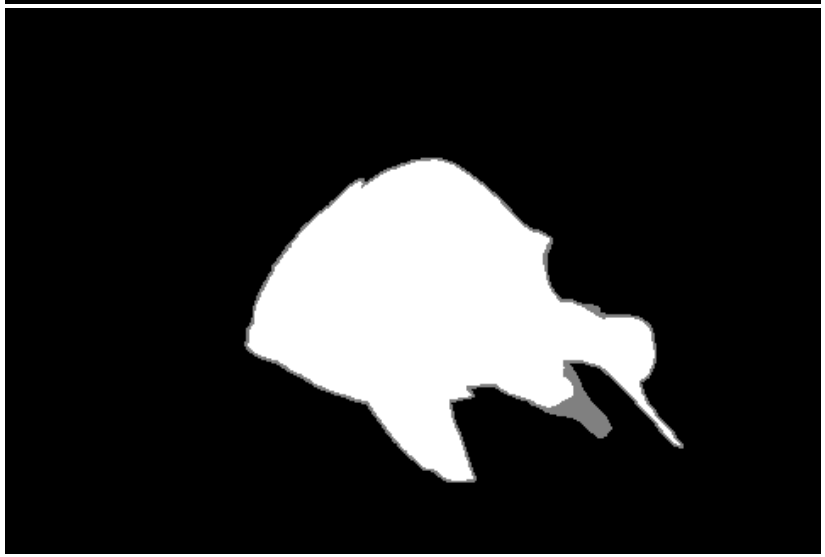
Der Plan:



Segmentierung

Man hat das ...





Man braucht das.

Segmentierung

Man hat nun das ...

Was braucht man?
⇒ ein schlecht gestelltes Problem
(d.h. es gibt keine einheitliche Definition).

Normalerweise soll ein „Objekt“ segmentiert werden.

Was wird dabei benötigt?

- Farbmodelle
- A-priori Wissen über Form
  ↑ alles mit zugelassenen Transformationen
- evtl. Beziehungen zu anderen Objekten,
  Lage, Geometrie, Topologie etc.

(Einfache) Clusterung

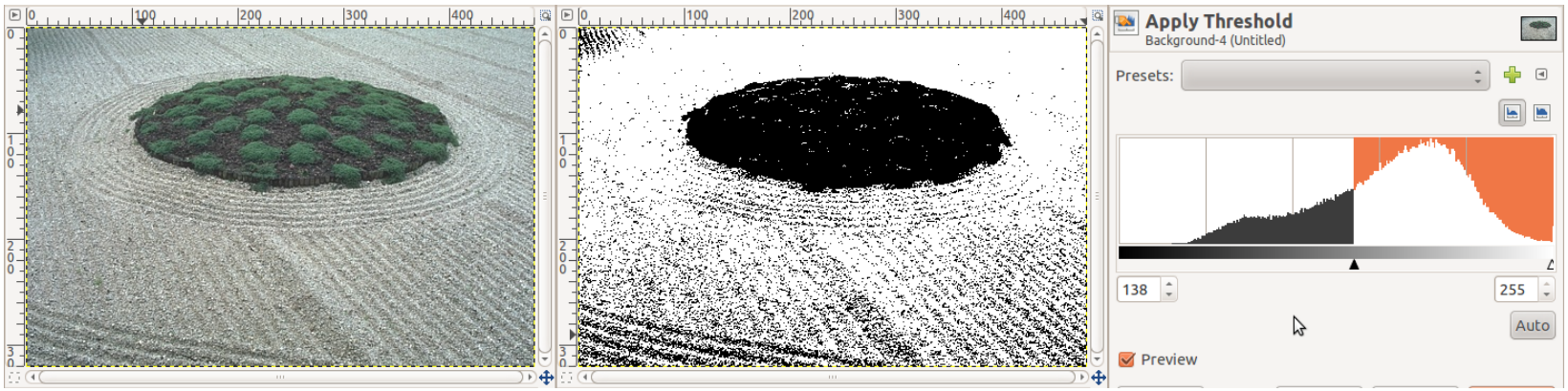
Idee:

- Für jeden Pixel werden „Merkmale“ berechnet,
- Die so entstehende Menge wird geclustert

Beispiel:

- Zwei Clusters (Segmente), Grauwertbilder, Merkmal – Grauwert
- Geclustert wird nach dem K-Means Prinzip (mithilfe der Histogramme)

  $\Rightarrow$ Binarisierung mit einem Schwellwert

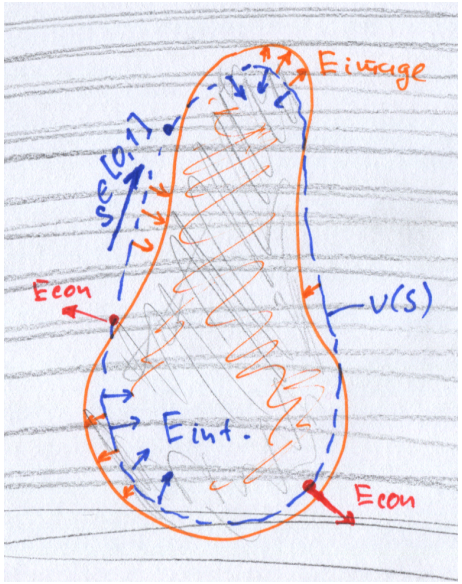


Erweiterungen:

- Glättung vorher/nachher
- Kompliziertere lokale Merkmale
- Andere Clusterung Methoden/Abstände
- ...

Aktive Konturen (Snakes)

[Kass, Witkin, Terzopoulos, 1988]



Eine Segmentierung ist es nicht ganz

Eine Kontur ist **parametrisch** beschrieben:

$v(s) = \Big(x(s), y(s)\Big),\ s \in [0, 1]$
(geschlossen, nicht geschlossen)
– parametrisch, mit Slpines usw.

Seine Energie ist:

$$E(v) = \int_0^1 \Big[E_{int}\Big(v(s)\Big) + E_{image}\Big(v(s)\Big) + E_{con}\Big(v(s)\Big)\Big]\, ds$$

Interne Energie:

$$E_{int}\Big(v(s)\Big) = \Big(\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2\Big)/2$$

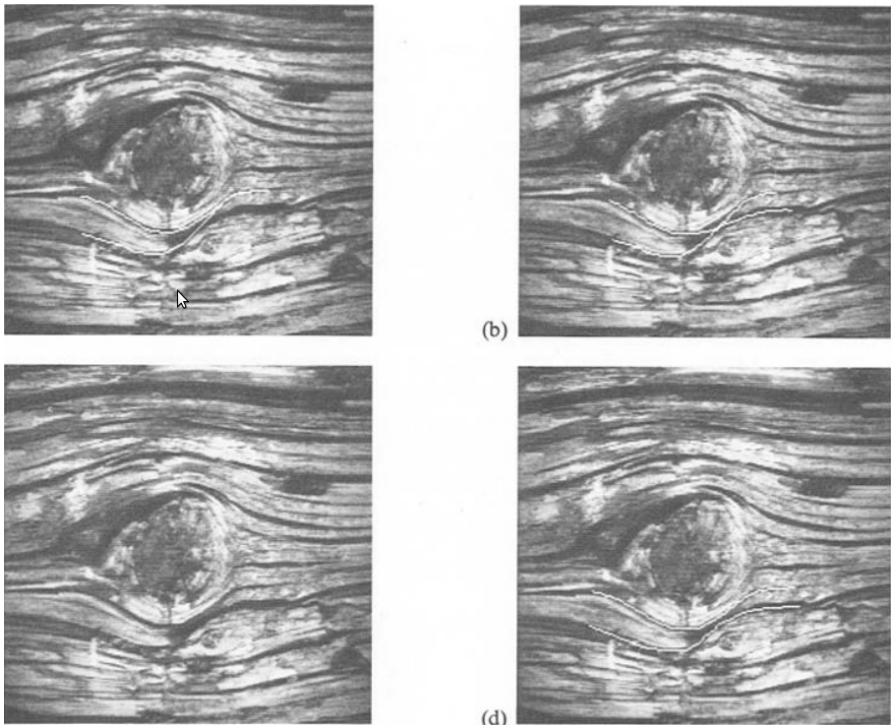
mit v_s und v_{ss} die erste bzw. die zweite Ableitungen,
$\alpha(s)$ und $\beta(s)$ sind positionspezifische Gewichte – repräsentieren die gewünschte Form
z.B. $\beta(s^*) = 0$ bedeutet „Ecke an der Stelle s^* ist möglich“

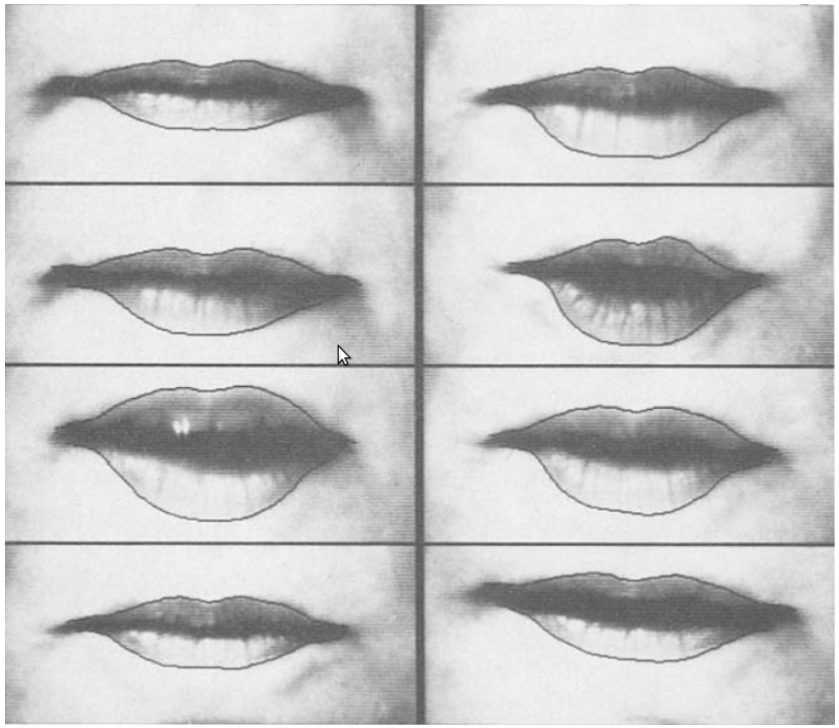
$E_{image}\Big(v(s)\Big)$ ist der Datenterm – $I\Big(x(s), y(s)\Big)$ (Intensität), $\Big|\nabla I\Big(x(s), y(s)\Big)\Big|^2$ (Kanten) ...

$E_{con}\Big(v(s)\Big)$ sind „zusätzliche Kräfte“ (z.B. vom Nutzer explizit angegeben).

Aktive Konturen (Snakes)

Paar Beispiele:





Andrew Blake and Michael Isard: Active Contours

`http://www.robots.ox.ac.uk/~contours/`

Region Growing

Man hat bereits eine Region (diskret – eine Teilmenge der Pixel) und lässt sie wachsen:

- Für die aktuelle Region werden bestimmte Charakteristika berechnet, wie z.B. den mittleren Grauwert, Streuung, andere Farbcharakteristika (Statistik der Kanten Ecken usw.), Größe, Ausdehnung, Umfang (Geometrische Merkmale) ...
- Pixel zu der aktuellen Region benachbarten Pixel werden betrachtet
- Abhängig von den lokalen Pixelmerkmalen und den Charakteristika der Region wird entschieden, ob der Pixel zu der Region gehört.

Ergänzungen:

- Mehrere Regionen
- Split, Merge ...

Ähnlichkeiten zu Aktiven Konturen:

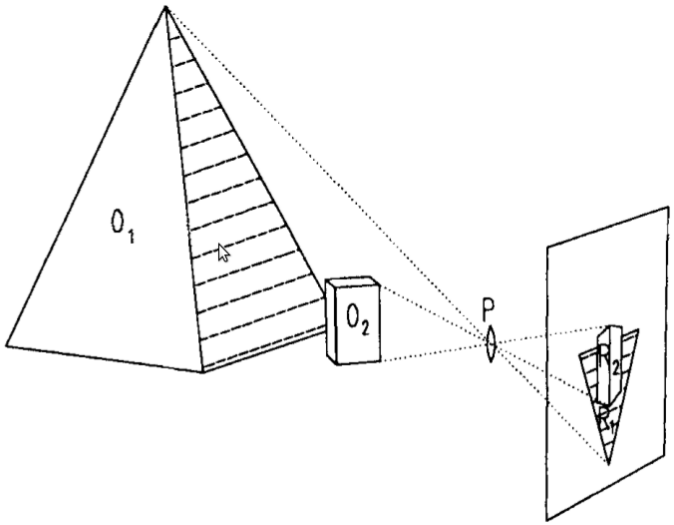
- eine lokale Vorgehensweise – Evaluierung einer Kontur
- eine Initialisierung ist notwendig

Unterschiede:

- Topologie: RG – automatisch, AK – problematisch
- Metrische Fehler (durch Diskretisierung) bei RG
- Globale Charakteristika sind bei AK schwierig zu berücksichtigen
- Modell bei RG ist meist nicht explizit definiert

Mumford-Shah Funktional

[Mumford, Shah, 1989]



Motivation:

Das betrachtete 2-dimensionale Bild ist eine Projektion einer 3D-Szene

Lichtstrahlen treffen Objekte → Bildfunktion innerhalb der Projektion eines Objektes ist glatt

Objektgrenzen in 3D verursachen Diskontinuitäten der Bildfunktion

⇒ Bildfunktion soll stückweise glatt sein

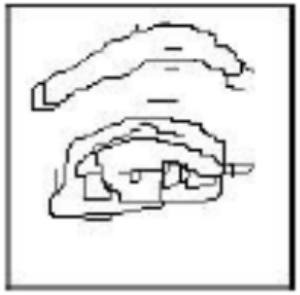
$R \subset \mathbb{R}^2$ – Bilddomäne, Γ – Diskontinuitäten (Segmentgrenzen)
$g : R \to \mathbb{R}$ – (die zu approximierende) Bildfunktion
$f : R \to \mathbb{R}$ – stückweise glatte Approximation

$$E(f,\Gamma) = \mu^2 \int_R (f-g)^2 \, dxdy + \int_{R-\Gamma} \|\nabla f\|^2 \, dxdy + \nu|\Gamma| \to \min_{f,\Gamma}$$

Die Approximation soll

- mit der Bildfunktion möglichst übereinstimmen
- glatt sein
- möglichst wenig Grenzen haben

Mumford-Shah Funktional



← Beispiele für unterschiedliche μ und ν

Problem der Topologie wieder!!!

Dies ist keine Segmentierung im Sinne der Partitionierung des Bildes auf disjunkte Gebiete.

Die gefundenen Kanten sind nicht unbedingt Segmentgrenzen

Mumford-Shah Funktional

Spezialfall: Stückweise **konstante** Approximation – „Cartoon-Limit“

$$E(f,\Gamma) = \mu^2 \int_R (f-g)^2\,dxdy + \int_{R-\Gamma} \|\nabla f\|^2\,dxdy + \nu|\Gamma| \to \min_{f,\Gamma}$$

geht bei $\mu, \nu \to 0$ gegen

$$E(a,\Gamma) = \sum_i \int_{R_i} (g-a_i)^2\,dxdy + \nu|\Gamma| \to \min_{a,\Gamma}$$

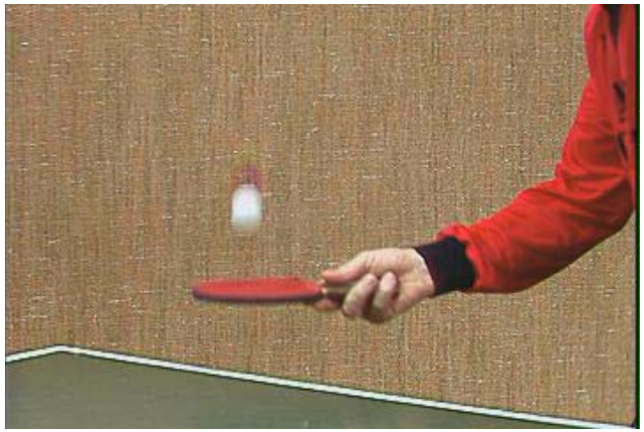
Die Lösung bezüglich a_i bei der fixierten Partitionierung Γ ist:

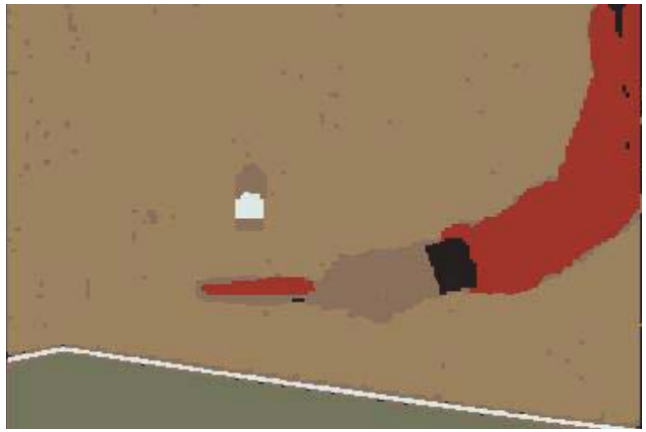
$$a_i = \frac{1}{|R_i|} \int_{R_i} g\ dxdy$$

– das Bild wird in die Regionen zerlegt (partitioniert),
– Farbwert innerhalb einer Region ist konstant

Jetzt ist die Topologie in Ordnung: die Kanten dürfen nicht mehr „willkürlich“ enden, denn das ist offensichtlich energetisch ungünstig (wegen $|\Gamma|$).

Chan-Vese Modell





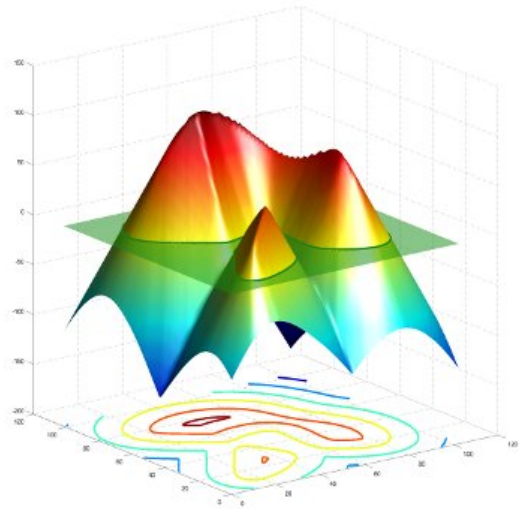
Bei Cartoon-Limit wird die Anzahl der Segmente nicht festgelegt.
Diese Anzahl ergibt sich automatisch.

Moch spezieller: nur zwei Regionen [Chan, Vese, 2001]:

$$E(c_1, c_2, C) = \mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(C)$$
$$+\lambda_1 \int_{inside(C)} |u_0(x,y) - c_1|^2 \, dxdy$$
$$+\lambda_2 \int_{outside(C)} |u_0(x,y) - c_2|^2 \, dxdy$$

Kann effizient mit **Level Set** Methoden gelöst werden.

Level Set Methoden



[Osher, Sethian, 1988 ?]

Idee: eine Kontur wird durch Nullstelle einer **Levelset** Funktion $\phi : \mathbb{R}^2 \to \mathbb{R}$ beschrieben, d.h.

$$C = \{(x, y) \in \Omega : \phi(x, y) = 0\},$$
$$inside(C) = \{(x, y) \in \Omega : \phi(x, y) > 0\},$$
$$outside(C) = \{(x, y) \in \Omega : \phi(x, y) < 0\}$$

Heaviside Funktion $H(z) = \mathbb{1}(z > 0)$ („Step“), Dirac Funktion $\delta(z) = \nabla H(z)$ („Delta“)

Mit diesen Hilfsfunktionen lassen sich alle Bestandteile wie folgt schreiben:

$$\text{Length}(C) = \int_\Omega \delta\Big(\phi(x, y)\Big)\, dxdy, \qquad \text{Area}(C) = \int_\Omega H\Big(\phi(x, y)\Big)\, dxdy$$

$$\int_{inside(C)} bla(x, y)\, dxdy = \int_\Omega H\Big(\phi(x, y)\Big)\, bla(x, y)\, dxdy \quad \text{usw.}$$

Andere Charakteristika der Kontur (wie z.B. die Krümmung) lässt sich auch relativ einfach ausdrücken.

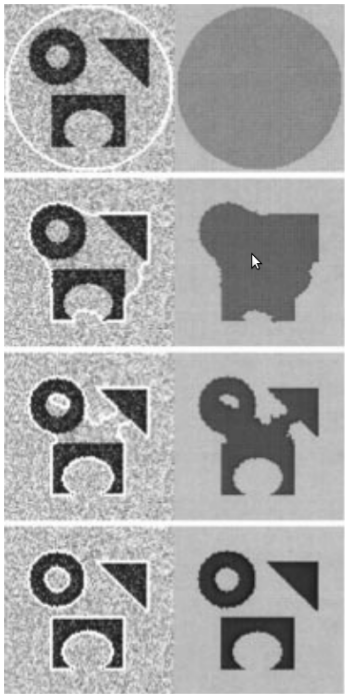
Integrale über bestimmte Gebiete werden durch Integrale über den gesamten Ω ersetzt.

Level Set Methoden

Weiter:

- Nichtdifferenzierbare H und δ werden durch etwas differenzierbares ersetzt
- entweder
  - Gateau Ableitungen auf Null setzen → Euler-Lagrange Gleuchungen
  - Diskretisieren → Gleichungssystem
  - Mit iterativen Methoden lösen

  oder
  - Gateau Ableitungen → Gragienten
  - Gradientabstieg



Die Topologie ist korrekt

Die Topologie kann sich während der Evaluation der Kurve ändern

Weniger sensitiv zur Initialisierung

Vieles lässt sich einheitlich und bequem ausdrücken

Nur mühsam für mehr als zwei Segmente übertragbar

Statistische Interpretation schwierig
⇒ Probleme mit dem Lernen (betrifft alle kontinuierliche Ansätze)