Continuous and Discrete Optimization Methods in Computer Vision

Daniel Cremers
Department of Computer Science
University of Bonn

Oxford, August 16 2007
Segmentation by Energy Minimization

Given an image $I : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, minimize the energy

$$E(C, \mu_1, \mu_2) = \int_{R_1} |I - \mu_1|^2 \, dx + \int_{R_2} |I - \mu_2|^2 \, dx + \nu |C|$$

boundary length

Blake, Zisserman '87, Mumford, Shah '89, Ising '27, Potts '59, Geman, Geman '84
Continuous Versus Discrete Optimization

\[ C = \{ x \in \Omega \mid \phi(x) = 0 \}, \]

\[ \phi : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R} \]

\[ \frac{\partial \phi}{\partial t} = \ldots \]

Osher, Sethian, J. of Comp. Phys. ’88

\[ \text{Min. cut} \leftrightarrow \text{Max. flow} \]

Ford, Fulkerson ’59
Greig et al. ’89
Boykov, Kolmogorov ’02
Level Set Formulation of Mumford-Shah

Chan & Vese ’99
Level Set Segmentation in Feature Space

efficient coarse-to-fine scheme

*Brox, Weickert ’04, ’06*
Overview

Motion segmentation
Obstacle segmentation
Prior shape knowledge
Shape priors in global optimization
Convex 3D reconstruction
Markerless human motion capture
Overview

Motion segmentation

Obstacle segmentation

Prior shape knowledge

Shape priors in global optimization

3D reconstruction: a convex formulation

Markerless human motion capture

_Cremers CVPR ’03, Cremers & Soatto IJCV ’05_
Motion
Motion-based Image Segmentation

1. assumption: Intensity of moving points is preserved.

\[ I(x(t), t) = \text{const.} \quad \frac{d}{dt} I(x(t), t) = \nabla I^\top v + I_t = 0 \]

Problem underconstrained (aperture problem).

2. assumption: Velocity in each region is a Gaussian-distributed random variable

\[ \forall x \in R_i : \ v(x) = v_i + \eta, \quad \text{where } \eta \sim \mathcal{N}(0, \sigma^2) \]

Minimize

\[ E(\{v_i\}, C) = \sum_i \int_{R_i} \frac{|v_i^\top \nabla I + I_t|^2}{|\nabla I|^2} \ dx \ + \ \nu |C| \]

**Cremers CVPR ’03, Cremers & Soatto IJCV ’05**
Motion Competition

\[ E(\{v_i\}, C) = \sum_i \int_{R_i} \frac{|v_i^\top \nabla I + I_t|^2}{|\nabla I|^2} \, dx + \nu |C| \]

Motion Estimation

\[ v_i = \arg \min_v \begin{pmatrix} v \end{pmatrix}^\top M_i \begin{pmatrix} v \end{pmatrix} \]

with \( M_i = \int_{R_i} \frac{\nabla_3 I \nabla_3 I^\top}{|\nabla_3 I|^2} \, dx \)

Segmentation

\[ \frac{\partial C}{\partial t} = (e_j - e_k) n + \nu \kappa n \]

with \( e_i = \frac{|v_i^\top \nabla I + I_t|^2}{|\nabla I|^2} \)
Motion Competition via Level Sets

Cremers, Yuille, ’04
Motion Competition via Level Sets

What is moving?
Motion Competition via Level Sets

Cremers, Yuille, '04
Motion Competition via Level Sets

_Cremers, Soatto, IJCV '05:_ Piecewise Parametric Motion
Motion Competition via Graph Cuts

piecewise constant  piecewise affine

*Schoenemann & Cremers, DAGM ’06*

up to 30 fps
Overview

Motion segmentation

Obstacle segmentation

Prior shape knowledge

Shape priors in global optimization

3D reconstruction: a convex formulation

Markerless human motion capture

*Wedel, Schoenemann, Brox, Cremers DAGM ’07*
WarpCut: Obstacle Segmentation

Segmentation of obstacles in monocular video
WarpCut: Obstacle Segmentation

Local scale changes depend on distance from the camera.
WarpCut: Obstacle Segmentation

Competing hypotheses: Obstacle or road?
WarpCut: Obstacle Segmentation

Wedel, Schoenemann, Brox, Cremers DAGM ’07
Overview

Motion segmentation

Obstacle segmentation

Prior shape knowledge

Shape priors in global optimization

3D reconstruction: a convex formulation

Markerless human motion capture
Statistical Priors for Image Segmentation

Cremers, Kohlberger, Schnörr, ECCV 2002
Statistical Learning of Familiar Shapes

\[ \mathcal{P}(\phi) \propto \frac{1}{N} \sum_{i=1}^{N} \exp \left( -\frac{1}{2\sigma^2} d^2(\phi, \phi_i) \right) \]

*Cremers, Osher, Soatto, IJCV '06*  
*Rosenblatt '56, Parzen '62*
Statistical Learning of Implicit shapes

Without statistical prior

With statistical prior

*Cremers, Osher, Soatto, IJCV ’06*

Sequence data courtesy of Alessandro Bissacco.
Statistical Priors for Image Segmentation

Purely geometric prior

Nonparametric prior

*Cremers, Osher, Soatto, IJCV ’06*
Dynamical Models for Implicit Shapes

Training sequence
Dynamical Models for Implicit Shapes

1. Dimensionality reduction by PCA \((\text{Leventon et al. '00, Tsai et al. '01})\):

\[
\phi_i(x) \approx \phi_0(x) + \sum_{j=1}^{m} \alpha_{ij} \psi_j(x)
\]

\[
\alpha_{ij} = \int (\phi_i - \phi_0) \psi_j \, dx \quad \text{mean } \alpha_i = (\alpha_{i1}, \ldots, \alpha_{im})
\]

2. Autoregressive model for the shape vectors:

\[
\alpha_t = \mu + \sum_{i=1}^{k} A_i \alpha_{t-i} + \eta
\]

3. Synthesized sequence of shape vectors and embedding functions:

mean \quad \text{transition matrices} \quad \text{Gaussian noise}

\[
\phi_t(x) = \phi_0(x) + \alpha_t^\top \psi(x)
\]

\textit{Cremers, IEEE Trans. on PAMI '06}
Statistically Sampled Embedding Functions

Cremers, IEEE Trans. on PAMI ’06
Dynamical Models for Implicit Shapes

1. Low-dim. representation via PCA \( (\text{Leventon et al. '00, Tsai et al. '01}) \):

\[
\phi_i(x) \approx \phi_0(x) + \sum_{j=1}^{m} \alpha_{ij} \psi_j(x)
\]

\[
\alpha_{ij} = \int (\phi_i - \phi_0) \psi_j \, dx \quad \alpha_i = (\alpha_{i1}, \ldots, \alpha_{im})
\]

2. Autoregressive model for the shape coefficients:

\[
\alpha_t = \mu + \sum_{i=1}^{k} A_i \alpha_{t-i} + \eta
\]

3. Synthesize shape vectors and embedding surfaces:

\[
\phi_t(x) = \phi_0(x) + \alpha_t^\top \psi(x)
\]
Dynamical Models for Implicit Shapes

1. Low-dim. representation via PCA \( \text{(Leventon et al. '00, Tsai et al. '01):} \)

\[
\phi_i(x) \approx \phi_0(x) + \sum_{j=1}^{m} \alpha_{ij} \psi_j(x)
\]

\[
\alpha_{ij} = \int (\phi_i - \phi_0) \psi_j \, dx \hspace{1cm} \alpha_i = (\alpha_{i1}, \ldots, \alpha_{im})
\]

2. Autoregressive model for the shape coefficients:

\[
\alpha_t = \mu + \sum_{i=1}^{k} A_i \alpha_{t-i} + \eta
\]

3. Synthesize shape vectors and embedding surfaces:

\[
\phi_t(x) = \phi_0(T_{\theta_t} x) + \alpha_l^\top \psi(T_{\theta_t} x)
\]
Dynamical Priors for Level Set Segmentation

Bayesian Aposteriori Maximization:

\[ \hat{\alpha}_t, \hat{\theta}_t = \arg \max_{\alpha_t, \theta_t} \mathcal{P}(\alpha_t, \theta_t \mid I_t, \hat{\alpha}_{1:t-1}, \hat{\theta}_{1:t-1}) \]

\[ = \arg \max_{\alpha_t, \theta_t} \mathcal{P}(I_t \mid \alpha_t, \theta_t) \mathcal{P}(\alpha_t \mid \hat{\alpha}_{1:t-1}) \]

\[ E(\alpha_t, \theta_t) = -\log \mathcal{P} = E_{\text{dat}}(\alpha_t, \theta_t, I_t) + E_{\text{dyn}}(\alpha_t, \hat{\alpha}_{1:t-1}) \]

\[ E_{\text{dyn}} = \frac{1}{2} (\alpha_t - \nu)^\top C^{-1} (\alpha_t - \nu), \quad \nu \equiv \mu + \sum_{i=1}^{k} A_i \hat{\alpha}_{t-i} \]

Optimization by gradient descent:

\[ \frac{d\alpha_t}{d\tau} = -\frac{\partial E}{\partial \alpha_t}, \quad \frac{d\theta_t}{d\tau} = -\frac{\partial E}{\partial \theta_t} \]
Dynamical Priors for Level Set Segmentation

Input sequence with 90% uniform noise
Dynamical Priors for Level Set Segmentation

Cremers, IEEE Trans. on PAMI ’06

Daniel Cremers
Continuous and Discrete Optimization in Computer Vision
Dynamical Priors for Level Set Segmentation

Pure deformation model

*Cremers, IEEE Trans. on PAMI ’06*
Dynamical Priors for Level Set Segmentation

Model of joint deformation and transformation

Cremers, IEEE Trans. on PAMI ’06
Overview

Motion segmentation
Obstacle segmentation
Prior shape knowledge
Shape priors in global optimization
3D reconstruction: a convex formulation
Markerless human motion capture

Schoenemann & Cremers ICCV ‘07
Limitations of Previous Approaches

- Local optimization
- Simple geometric distances
- Little generalization
- Requires large training sets

- Point correspondence,
- Missing parts,
- Stretching/shrinking,…
- Combinatorial problem

Goal: Combinatorial solution for joint segmentation and alignment
Each node of the graph represents an image pixel and a point on the template. All template points $S$ are allowed to stretch over at most $K$ image pixels.
The Product Space of Image and Template

Each node of the graph represents an image pixel and a point on the template.

All template points $S$ are allowed to stretch over at most $K$ image pixels.

A cycle in this graph defines an assignment of template points to image pixels.
Segmentation & Alignment by Energy Min.

\[
\min_{C,m} \frac{\int_0^l g(C(s)) ds}{l(C')} + \lambda \frac{\int_0^l |\alpha_c(s) - \alpha_s(m(s))|^2 ds}{l(C')} + \nu \frac{\int_0^l \Psi(m'(s)) ds}{l(C')}
\]

- **Data term**
- **Alignment with template**
- **Stretching cost**

\[
g(x) = \frac{1}{1 + |\nabla I(x)|}
\]

Favors strong gradients.

\[
\Psi(m') = \begin{cases} 
  m' - 1, & \text{if } m' \geq 1 \\
  \frac{1}{m'} - 1, & \text{otherwise}
\end{cases}
\]

Favors minimal stretching / shrinking.
Segmentation with a Single Template

Template

Segmentation results

Optimization by Minimum Ratio Cycles on GPU: ~15 sec / frame
Rotational Invariance

No rotational invariance  \hspace{1cm}  Rotational invariance
Tracking

Input sequence

Globally optimal segmentations
Overview

Motion segmentation
Obstacle segmentation
Prior shape knowledge
Shape priors in global optimization
3D reconstruction: a convex formulation
Markerless human motion capture

Kolev, Klodt, Brox, Esedoglu & Cremers  EMMCVPR ‘07
3D Reconstruction
Probabilistic Formulation

\[ P_{\text{obj}}(x), P_{\text{bck}}(x) \]
Probabilistic Formulation

Kolev, Brox, Cremers  DAGM ’06
Continuous Convex Formulation

Introduce binary variable: \( u : \Omega \subset \mathbb{R}^3 \rightarrow \{0, 1\} \)

\[
E(\phi) = \int f(x) \, u \, d^3x + \int g(x) (1-u) \, d^3x + \nu \int \rho(x) |\nabla u| \, d^3x
\]

Relax binary constraint: \( u(x) \in [0, 1] \)

- Convex functional optimized over a convex set.
- Solution of binary-valued problem by thresholding.
- Global minima by gradient descent (or SOR).

*Kolev, Klodt, Brox, Esedoglu & Cremers EMMCVPR ’07*
Multiview 3D Reconstruction

Kolev, Klodt, Brox, Esedoglu & Cremers EMMCVPR ‘07
## Comparison

<table>
<thead>
<tr>
<th>Level sets</th>
<th>Convex formulation</th>
<th>Graph cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td>Continuous</td>
<td>Discrete</td>
</tr>
<tr>
<td>Local optima</td>
<td>Global optima*</td>
<td>Global optima*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Memory limitations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Metrification errors</td>
</tr>
</tbody>
</table>

* for certain cost functionals only
Metrification Errors in Discrete Energies

Synthetic data:

Discrete graph cut optimization

Continuous convex optimization

Kolev, Klodt, Brox, Esedoglu & Cremers EMMCVPR ‘07
Overview

Motion segmentation

Obstacle segmentation

Prior shape knowledge

Shape priors in global optimization

3D reconstruction: a convex formulation

Markerless human motion capture

*Brox et al. DAGM 2005*

*Brox, Rosenhahn, Cremers & Seidel, ECCV ’06*
Human Motion Tracking

\[ \begin{bmatrix} R \mid t \end{bmatrix} \]

\[ \phi_1 \ldots \phi_n \]
Human Motion Tracking

Brox et al. DAGM 2005
Brox, Rosenhahn, Cremers & Seidel, ECCV ’06
Coupled Segmentation and 3D Tracking

Brox et al. DAGM 2005
Brox, Rosenhahn, Cremers & Seidel, ECCV '06
3D Tracking with Statistical Priors

T. Brox, B. Rosenhahn, U. Kersting, D. Cremers, DAGM ’06
Human Motion Tracking

T. Brox, B. Rosenhahn, U. Kersting, D. Cremers, DAGM ’06
Summary

Motion competition

Obstacle segmentation

Dynamical shape priors

3D reconstruction

Human motion tracking
Announcement: IPAM Workshop

Boykov, Cremers, Darbon, Ishikawa, Kolmogorov, Osher:

IPAM Workshop on “Graph Cuts and Related Discrete or Continuous Optimization Problems”, Feb. 25-29 2008

https://www.ipam.ucla.edu/programs/gc2008/