Video Object Segmentation via Shape Optimization

Ganesh Sundaramoorthi
Assistant Professor
King Abdullah University of Science and Technology (KAUST)
Object Segmentation from Images

label each pixel of image corresponding to objects
Some Motivating Applications

Robotics
Some Motivating Applications

Robotics

Entertainment

3D reconstruction
Some Motivating Applications

Robotics

Entertainment

3D reconstruction

Medicine
Some Motivating Applications

Robotics

Entertainment

3D reconstruction

Medicine

Geology & Other Sciences
Dominant Approach

Discriminating Image Intensity Statistics

Example:

minimize for $R_{in}$ (object segmentation)

$$E(R_{in}) = -D(p_{in}, p_{out})^2,$$

$(D$ distance between pdf’s $)$

$p_{in} = $ color histogram inside $R_{in}$

$p_{out} = $ color histogram inside $R_{out}$

e.g., Geman & Geman 1984, Mumford & Shah 1989, Caselles et al., 1995, Chan & Vese 2001, Kolmogorov et al., 2001, Bresson et al., 2007, etc
Dominant Approach

Discriminating Image Intensity Statistics

Example:

\[
E(R_{in}) = -D(p_{in}, p_{out})^2, \\
(D \text{ distance between pdf’s })
\]

\[
p_{in} = \text{color histogram inside } R_{in} \\
p_{out} = \text{color histogram inside } R_{out}
\]

minimize for \( R_{in} \) (object segmentation)

Further Constraints Needed to Segment Objects

e.g., Geman & Geman 1984, Mumford & Shah 1989, Caselles et al., 1995, Chan & Vese 2001, Kolmogorov et al., 2001, Bresson et al., 2007, etc
Using Changing Optical Pattern via Video

- J. J. Gibson (1978): from changing pattern on ambient optic array of moving animal, animal can directly perceive properties of the environment for survival

- biological visual systems: motion plays an important role
Segmentation of Motion from Video

Main observation: induced motion of pixels is similar for distinct objects

Segment image into regions of similar motion
Segmentation of Motion from Video

Main observation: induced motion of pixels is similar for distinct objects

Segment image into regions of similar motion

Step 1: Assume segmentation of object in frame $t$, estimate motion
Challenges in Motion Estimation

variation from image formation

viewpoint occlusions

illumination quantization and noise

“What makes [vision] hard is not that any of the above transformations are that hard to detect and decode in isolation, but rather that all of them tend to coexist, and then the decoding becomes hard.”

- David Mumford, “Pattern Theory: A Unifying Perspective,” 1994
Our Model: Incorporating Occlusions

- **Occlusion**: part of object that goes out of view between frames
- **Disocclusion**: part of object that comes into view

Yang & Sundaramoorthi, “Modeling Self-Occlusions in Dynamic Shape & Appearance Tracking,” ICCV 2013
Our Model: Incorporating Occlusions

**Occlusion**: part of object that goes out of view between frames

**Disocclusion**: part of object that comes into view

**Dynamic Model:**

\[
R_{t+1} = w_t(R_t \setminus O_t) \cup D_{t+1}
\]

\[
a_{t+1}(x) = \begin{cases} 
  a_t(w_t^{-1}(x)) + \eta_t(x) & x \in w_t(R_t \setminus O_t) \\
  a_{t+1}^d(x) + \eta_t(x) & x \in D_{t+1}
\end{cases}
\]

Besides occlusion/dis-occlusion, this model is **Brightness-Constancy Plus Noise**

**Notes**:

- \(R\) - region of interest
- \(w\) - general diffeomorphism (defined only on \(R \setminus O\))
- \(a\) - radiance function defined on \(R\)

---

Yang & Sundaramoorthi, “Modeling Self-Occlusions in Dynamic Shape & Appearance Tracking,”
ICCV 2013
Our Model: Incorporating Occlusions

Occlusion: part of object that goes out of view between frames
Disocclusion: part of object that comes into view

Dynamic Model:

\[ R_{t+1} = w_t(R_t \setminus O_t) \cup D_{t+1} \]

\[ a_{t+1}(x) = \begin{cases} 
  a_t(w_t^{-1}(x)) + \eta_t(x) & x \in w_t(R_t \setminus O_t) \\
  a_{t+1}^d(x) + \eta_t(x) & x \in D_{t+1}
\end{cases} \]

Besides occlusion/dis-occlusion, this model is **Brightness-Constancy Plus Noise**

Goal: Determine segmentation in next frame, \( w(R \setminus O) \cup D \)

Yang & Sundaramoorthi, “Modeling Self-Occlusions in Dynamic Shape & Appearance Tracking,” ICCV 2013
Deformation / Motion Estimation

un-occluded region, $R \setminus O$

occlusion, $O$

non-rigid deformation, $w$

warp of un-occluded region, $w(R \setminus O)$

disocclusion, $D$
Deformation / Motion Estimation

- $w$ is defined on un-occluded region $R \setminus O$, but $O$ is unknown
Deformation / Motion Estimation

- $w$ is defined on un-occluded region $R \setminus O$, but $O$ is unknown
- $O$ is subset of $R$ where $w$ is not defined, but $w$ is unknown
Deformation / Motion Estimation

- \(w\) is defined on un-occluded region \(R \setminus O\), but \(O\) is unknown
- \(O\) is subset of \(R\) where \(w\) is not defined, but \(w\) is unknown

\(w\) and \(O\) are coupled: should be estimated jointly
Deformation / Motion Estimation

- \( w \) is defined on un-occluded region \( R \setminus O \), but \( O \) is unknown
- \( O \) is subset of \( R \) where \( w \) is not defined, but \( w \) is unknown
  \( w \) and \( O \) are coupled: should be estimated jointly

Joint Optimization Problem:

\[
E(O, w | I, a, R) = \int_{R \setminus O} (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) + \beta_o \text{Reg}(O)
\]

- **1st term**: penalizes \( w \) that doesn’t transform radiance between frames
- **2nd term**: regularization on \( w \) to resolve aperture ambiguity
- **3rd term**: regularization on \( O \) corresponds to a prior on \( O \)
Need for New Optimization Tools

First consider the sub-problem (no occlusion):

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) \]
Need for New Optimization Tools

First consider the sub-problem (no occlusion):

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) \]

• **Non-convex** energy
  • space of warps, \( w \), is not a linear space
Need for New Optimization Tools

First consider the sub-problem (no occlusion):

\[
E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w)
\]

- **Non-convex** energy
  - space of warps, \( w \), is not a linear space

- **Energy regularization** (e.g., Horn & Schunck 1981, Sun & Black 2010) linearizes by assuming infinitesimal warps:

\[
E(v|I, a, R) = \int_R (I(x) + \nabla I(x) \cdot v(x) - a(x))^2 dx \\
+ \alpha \int_R |\nabla v(x)|^2 dx
\]

\( w(x) = x + v(x) \), where \( v(x) \) is small
Need for New Optimization Tools

First consider the sub-problem (no occlusion):

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) \]

• **Non-convex** energy
  • space of warps, \( w \), is not a linear space

• **Energy regularization** (e.g., Horn & Schunck 1981, Sun & Black 2010) linearizes by assuming infinitesimal warps:

\[
E(v|I, a, R) = \int_R (I(x) + \nabla I(x) \cdot v(x) - a(x))^2 dx \\
+ \alpha \int_R |\nabla v(x)|^2 dx
\]

\[ w(x) = x + v(x), \text{ where } v(x) \text{ is small} \]

• Global optimum by solving PDE:

\[
\begin{cases}
-\alpha \Delta v(x) + \nabla I(x) \nabla I(x)^T v(x) = -(I(x) - a(x))\nabla I(x) & x \in R \\
\nabla v(x) \cdot N = 0 & x \in \partial R
\end{cases}
\]
Need for New Optimization Tools

Non-infinitesimal deformations: Iterate

\[ R \]

\[ \phi_0(x) \]

\[ x = \phi_0(x) \]

\[ R_\tau = \phi_\tau(R) \]

\[ \phi_\tau(x) \]

\[ v_\tau(\phi_\tau(x)) \]
Need for New Optimization Tools

Non-infinitesimal deformations: Iterate

1. Set forward / backward maps to identity $\phi_0(x) = \phi_0^{-1}(x) = x$
Need for New Optimization Tools

Non-infinitesimal deformations: Iterate

1. Set forward / backward maps to identity
   \[ \phi_0(x) = \phi_0^{-1}(x) = x \]

2. Compute infinitesimal deformation
   \[ v_{\tau} = \arg \min_v E(v|a_{\tau}, I, R_{\tau}) \]
1. Set forward / backward maps to identity 
\[ \phi_0(x) = \phi_0^{-1}(x) = x \]

2. Compute infinitesimal deformation 
\[ v_\tau = \arg \min_v E(v | a_\tau, I, R_\tau) \]

3. Update forward map 
\[ \partial_\tau \phi_\tau = v_\tau(\phi_\tau), \quad R_\tau = \phi_\tau(R) \]
Need for New Optimization Tools

Non-infinitesimal deformations: Iterate

1. Set forward / backward maps to identity
   \[ \phi_0(x) = \phi_0^{-1}(x) = x \]

2. Compute infinitesimal deformation
   \[ v_\tau = \arg \min_v E(v|a_\tau, I, R_\tau) \]

3. Update forward map
   \[ \partial_\tau \phi_\tau = v_\tau(\phi_\tau), \quad R_\tau = \phi_\tau(R) \]

4. Update backward map
   \[ \partial_\tau \phi_\tau^{-1} = -\nabla \phi_\tau^{-1} \cdot v_\tau, \text{ in } R_\tau \]
Need for New Optimization Tools

Non-infinitesimal deformations: Iterate

1. Set forward / backward maps to identity
   \( \phi_0(x) = \phi_0^{-1}(x) = x \)
2. Compute infinitesimal deformation
   \( v_\tau = \arg \min_v E(v|a_\tau, I, R_\tau) \)
3. Update forward map
   \( \partial_\tau \phi_\tau = v_\tau(\phi_\tau), \quad R_\tau = \phi_\tau(R) \)
4. Update backward map
   \( \partial_\tau \phi_\tau^{-1} = -\nabla \phi_\tau^{-1} \cdot v_\tau, \quad \text{in } R_\tau \)
5. Update region radiance
   \( a_\tau = a \circ \phi_\tau^{-1} \)
Need for New Optimization Tools

Non-infinitesimal deformations: Iterate

1. Set forward / backward maps to identity
   \( \phi_0(x) = \phi_0^{-1}(x) = x \)

2. Compute infinitesimal deformation
   \( v_\tau = \arg \min_v E(v|a_\tau, I, R_\tau) \)

3. Update forward map
   \( \partial_\tau \phi_\tau = v_\tau(\phi_\tau), \quad R_\tau = \phi_\tau(R) \)

4. Update backward map
   \( \partial_\tau \phi_\tau^{-1} = -\nabla \phi_\tau^{-1} \cdot v_\tau, \text{ in } R_\tau \)

5. Update region radiance
   \( a_\tau = a \circ \phi_\tau^{-1} \)

6. Repeat 2-5 until deformation is zero
Need for New Optimization Tools

Given template $\alpha$ small $\alpha$ medium $\alpha$ large
Need for New Optimization Tools

Given template $\alpha$ small $\alpha$ medium $\alpha$ large
Need for New Optimization Tools

Given template  \( \alpha \) small  \( \alpha \) medium  \( \alpha \) large
Need for New Optimization Tools

Given template

\( \alpha \) small

\( \alpha \) medium

\( \alpha \) large
Need for New Optimization Tools

\( \alpha \) small

favors fine deformations; trapped by fine-scale details

\( \alpha \) medium

not trapped by fine-scale details; only captures coarse motion

\( \alpha \) large
Need for New Optimization Tools

\( \alpha \) small

\( \alpha \) medium

\( \alpha \) large

- favors fine deformations; trapped by fine-scale details

- not trapped by fine-scale details; only captures coarse motion

- No “optimal” choice for parameter
Need for New Optimization Tools

- $\alpha$ small
  - favors fine deformations; trapped by fine-scale details
- $\alpha$ medium
  - not trapped by fine-scale details; only captures coarse motion
- $\alpha$ large

- No “optimal” choice for parameter
- Ad-Hoc “Fix”: 
Need for New Optimization Tools

- $\alpha$ small: favors fine deformations; trapped by fine-scale details
- $\alpha$ medium: not trapped by fine-scale details; only captures coarse motion
- $\alpha$ large: not trapped by fine-scale details; only captures coarse motion

• No “optimal” choice for parameter
• Ad-Hoc “Fix”:
  1. Start with alpha large, iterate until convergence
Need for New Optimization Tools

- $\alpha$ small
- $\alpha$ medium
- $\alpha$ large

- favors fine deformations; trapped by fine-scale details
- not trapped by fine-scale details; only captures coarse motion

• **No “optimal” choice for parameter**

• **Ad-Hoc “Fix”**:  
  1. Start with alpha large, iterate until convergence  
  2. Continue to lower alpha, iterate and repeat
Need for New Optimization Tools

$\alpha$ small

$\alpha$ medium

$\alpha$ large

favors fine deformations; trapped by fine-scale details

not trapped by fine-scale details; only captures coarse motion

- No “optimal” choice for parameter
- Ad-Hoc “Fix”:
  1. Start with alpha large, iterate until convergence
  2. Continue to lower alpha, iterate and repeat
- Let’s do this in a convenient and principled manner by constructing the right Shape Optimization tools
Optimization on a Riemannian Manifold

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) \]

Yang & Sundaramoorthi, “Shape Tracking With Occlusions Using Region-Based Sobolev Descent,”
TPAMI 2015
Optimization on a Riemannian Manifold

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) \]

Different Approach:

- Remove regularization in the energy
Optimization on a Riemannian Manifold

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) \]

Different Approach:

- Remove regularization in the energy
- Construct optimization scheme to prefer certain perturbations of warps over others

Yang & Sundaramoorthi, “Shape Tracking With Occlusions Using Region-Based Sobolev Descent,” TPAMI 2015
Optimization on a Riemannian Manifold

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 dx + \alpha \text{Reg}(w) \]

Different Approach:

- Remove regularization in the energy
- Construct optimization scheme to prefer certain perturbations of warps over others
- Avoid linearization

Yang & Sundaramoorthi, “Shape Tracking With Occlusions Using Region-Based Sobolev Descent,” TPAMI 2015
Optimization on a Riemannian Manifold

\[ E(w|I, a, R) = \int_R (I(w(x)) - a(x))^2 \, dx + \alpha \text{Reg}(w) \]

Different Approach:

- Remove regularization in the energy
- Construct optimization scheme to **prefer certain perturbations of warps** over others
- Avoid linearization
- Use gradient descent on space of warps, a **Riemannian manifold**, choosing the **Shape Metric** that results in favorable properties of the optimizing flow

Yang & Sundaramoorthi, “Shape Tracking With Occlusions Using Region-Based Sobolev Descent,” TPAMI 2015
Optimization on a Riemannian Manifold

Key Point: Gradient Depends on *Shape Metric*
Optimization on a Riemannian Manifold

**Key Point:** Gradient Depends on *Shape Metric*

Gradient is *most efficient* deformation, i.e., maximizes

\[
\frac{\text{change in energy in deforming warp by } h}{\text{cost (measured by metric) of deforming by } h}
\]

over all possible deformations $h$ of the warp.
**Optimization on a Riemannian Manifold**

**Key Point:** Gradient Depends on *Shape Metric*

Gradient is *most efficient* deformation, i.e., maximizes

\[
\text{change in energy in deforming warp by } h \\
\text{cost (measured by metric) of deforming by } h
\]

over all possible deformations \( h \) of the warp

\[
w(R) \\
(\text{region})
\]

\[
\begin{align*}
&h(x) \\
x + \varepsilon h(x)
\end{align*}
\]

\[
(w + \varepsilon h)(R) \\
(\text{warp/region deformed by } h)
\]
Optimization on a Riemannian Manifold

**Key Point:** Gradient Depends on *Shape Metric*

Gradient is *most efficient* deformation, i.e., maximizes

\[
\frac{\text{change in energy in deforming warp by } h}{\text{cost (measured by metric) of deforming by } h}
\]

over all possible deformations \( h \) of the warp

The diagram shows a region \( w(R) \) being deformed by a transformation \( h \) from point \( x \) to \( x + \varepsilon h(x) \). The metric \( \| h \|_w \) measures the amount of deformation.
Choosing the Shape Metric

\[ M = \{ w : R \to \Omega \mid w : R \to w(R) \text{ is a diffeomorphism } \} \]
Choosing the Shape Metric

\[ M = \{ w : R \rightarrow \Omega \mid w : R \rightarrow w(R) \text{ is a diffeomorphism} \} \]

Conventional metric:

\[ \| h \|^2_{L^2} = \int_{w(R)} |h(x)|^2 \, dx \]
Choosing the Shape Metric

\[ M = \{w : R \rightarrow \Omega \mid w : R \rightarrow w(R) \text{ is a diffeomorphism} \} \]

Conventional metric:
\[ \| h \|_{L^2}^2 = \int_{w(R)} |h(x)|^2 dx \]

Region-Based Sobolev metric:
\[ \| h \|_{Sob}^2 = \int_{w(R)} |\nabla h(x)|^2 dx \]

Sobolev metrics for curves (Younes 1998, Klassen et al., 2004, Michor & Mumford 2007, Sundaramoorthi et al., 2007, 2011, Charpiat et al., 2007) for images (e.g., Trouve 1999, Beg et al. 2005), for regions (Wirth et al., 2011, Zolesio 2009)
Choosing the Shape Metric

\[ M = \{ w : R \to \Omega \mid w : R \to w(R) \text{ is a diffeomorphism} \} \]

Conventional metric:

\[ \| h \|^2_{L^2} = \int_{w(R)} |h(x)|^2 dx \]

Region-Based Sobolev metric:

\[ \| h \|^2_{Sob} = \int_{w(R)} |\nabla h(x)|^2 dx \]

Why use the Sobolev Metric?

Sobolev metrics for curves (Younes 1998, Klassen et al., 2004, Michor & Mumford 2007, Sundaramoorthi et al., 2007, 2011, Charpiat et al., 2007) for images (e.g., Trouve 1999, Beg et al. 2005), for regions (Wirth et al., 2011, Zolesio 2009)
Choosing the Shape Metric

\[ M = \{ w : R \rightarrow \Omega \mid w : R \rightarrow w(R) \text{ is a diffeomorphism} \} \]

Conventional metric:

\[ \| h \|_{L^2}^2 = \int_{w(R)} |h(x)|^2 \, dx \]

Region-Based Sobolev metric:

\[ \| h \|_{Sob}^2 = \int_{w(R)} |\nabla h(x)|^2 \, dx \]

Why use the Sobolev Metric?

- **smooth deformations** (e.g., coarse motions) are **favored** for gradients

Sobolev metrics for **curves** (Younes 1998, Klassen et al., 2004, Michor & Mumford 2007, Sundaramoorthi et al., 2007, 2011, Charpiat et al., 2007) for **images** (e.g., Trouve 1999, Beg et al. 2005), for **regions** (Wirth et al., 2011, Zolesio 2009)
Choosing the Shape Metric

\[ M = \{ w : R \rightarrow \Omega \mid w : R \rightarrow w(R) \text{ is a diffeomorphism} \} \]

Conventional metric:

\[ \| h \|_{L^2}^2 = \int_{w(R)} |h(x)|^2 \, dx \]

Region-Based Sobolev metric:

\[ \| h \|_{Sob}^2 = \int_{w(R)} |\nabla h(x)|^2 \, dx \]

Why use the Sobolev Metric?

• **smooth deformations** (e.g., coarse motions) are **favored** for gradients

• **non-smooth (fine) deformations** can be gradients, but only if coarse deformations cannot minimize energy

Sobolev metrics for **curves** (Younes 1998, Klassen et al., 2004, Michor & Mumford 2007, Sundaramoorthy et al., 2007, 2011, Charpiat et al., 2007) for **images** (e.g., Trouve 1999, Beg et al. 2005), for **regions** (Wirth et al., 2011, Zolesio 2009)
Sobolev Gradient Descent
Sobolev Gradient Descent
Sobolev Gradient Descent

**Automatic** Coarse-to-Fine Evolution
Sobolev Gradient Descent

Automatic Coarse-to-Fine Evolution
Sobolev Gradient Descent

Automatic Coarse-to-Fine Evolution

color code

deformation
Sobolev Gradient Descent

Automatic Coarse-to-Fine Evolution

color code

defformation
Sobolev Gradient Descent

Automatic Coarse-to-Fine Evolution

color code

decoration
Sobolev Gradient Descent

**Automatic** Coarse-to-Fine Evolution

color code
deforation
Sobolev Gradient Descent

Automatic Coarse-to-Fine Evolution

color code
deforation
Sobolev Gradient Descent

**Automatic Coarse-to-Fine Evolution**

color code

def ormation
Sobolev Gradient Descent

**Automatic** Coarse-to-Fine Evolution

No Parameter
Comparison of Final Results

Conventional Tools

$\alpha$ small
Comparison of Final Results

Conventional Tools

\(\alpha\) small \hspace{1cm} \(\alpha\) medium
Comparison of Final Results

Conventional Tools

\( \alpha \) small

\( \alpha \) medium

\( \alpha \) large
Comparison of Final Results

Conventional Tools

$\alpha$ small  $\alpha$ medium  $\alpha$ large

Sobolev
Computing the Sobolev Gradient

Gradient is defined by the property:

\[ dE(w) \cdot h = \left. \frac{d}{d\varepsilon} E(w + \varepsilon h) \right|_{\varepsilon=0} = \langle \nabla_w E, h \rangle_w \] for all \( h \in T_w M \)

- directional derivative of \( E \) in \( h \)
- metric (i.e., inner product)
Computing the Sobolev Gradient

Gradient is defined by the property:

\[
\frac{d}{d\varepsilon} E(w + \varepsilon h) \bigg|_{\varepsilon=0} = \langle \nabla_w E, h \rangle_w \quad \text{for all } h \in T_w M
\]

directional derivative of \( E \) in \( h \)

metric (i.e., inner product)

Sobolev Gradient, \( G \), satisfies Poisson PDE:

\[
\begin{align*}
-\Delta G(x) &= -(I(x) - a(w^{-1}(x))) \det (\nabla w^{-1}(x)) \\
&\quad + \text{avg} \left[ (I(x) - a(w^{-1}(x))) \det (\nabla w^{-1}(x)) \right] \quad x \in w(R) \\
\nabla G(x) \cdot N &= 0 \quad x \in \partial w(R) \\
\text{avg}(G) &= 0
\end{align*}
\]
Computing the Sobolev Gradient

Gradient is defined by the property:

\[ dE(w) \cdot h = \frac{d}{d\varepsilon} E(w + \varepsilon h) \bigg|_{\varepsilon=0} = \langle \nabla_w E, h \rangle_w \text{ for all } h \in T_w M \]

directional derivative of \( E \) in \( h \)

metric (i.e., inner product)

Sobolev Gradient, \( G \), satisfies Poisson PDE:

\[
\begin{align*}
-\Delta G(x) &= -(I(x) - a(w^{-1}(x))) \det (\nabla w^{-1}(x)) \\
&\quad + \text{avg } [(I(x) - a(w^{-1}(x))) \det (\nabla w^{-1}(x))] \\
\nabla G(x) \cdot N &= 0 \\
\text{avg}(G) &= 0
\end{align*}
\]

\( x \in w(R) \)

\( x \in \partial w(R) \)

Simple to implement
Computing the Sobolev Gradient

Gradient is defined by the property:

\[
dE(w) \cdot h = \frac{d}{d\varepsilon} E(w + \varepsilon h) \bigg|_{\varepsilon=0} = \langle \nabla_w E, h \rangle_w \text{ for all } h \in T_w M
\]

directional derivative of \( E \) in \( h \)

metric (i.e., inner product)

Sobolev Gradient, \( G \), satisfies Poisson PDE:

\[
\begin{cases}
-\Delta G(x) = - (I(x) - a(w^{-1}(x))) \det (\nabla w^{-1}(x)) \\
+ \text{avg} \left[ (I(x) - a(w^{-1}(x))) \det (\nabla w^{-1}(x)) \right] \\
\nabla G(x) \cdot N = 0 \\
\text{avg}(G) = 0
\end{cases}
\]

\( x \in w(R) \)

\( x \in \partial w(R) \)

**Simple** to implement

**Same** computational cost as conventional approach (Horn & Schunck)
Sobolev Gradient Flow

1. Initialize warps to identity
   \[ \phi_0(x) = \phi_0^{-1}(x) = x \]

2. **Compute Sobolev Gradient**
   \[ G_\tau = \nabla_{\text{Sob}} E(\phi_\tau) \]
   \[ \partial_\tau \phi^{-1}_\tau = \nabla \phi^{-1}_\tau \cdot G_\tau, \text{ in } R_\tau \]

3. Update backward warp
   \[ \partial_\tau \phi_{\tau} = -G_{\tau}(\phi_{\tau}), \text{ in } R \]

4. Update forward map in gradient direction

5. Repeat 2-4 until convergence
Joint Optimization With Occlusion

Alternating optimization of

\[ E(O, w; I, a, R) = \int_{R \setminus O} (I(w(x)) - a(x))^2 dx + \int_O \beta_o dx \]
Joint Optimization With Occlusion

Alternating optimization of

\[ E(O, w; I, a, R) = \int_{R\setminus O} (I(w(x)) - a(x))^2 dx + \int_O \beta_0 dx \]

\( w \) defined on all of region \( R \)

un-occluded region, \( R\setminus O \)

occlusion, \( O \)
Joint Optimization With Occlusion

Alternating optimization of

\[ E(O, w; I, a, R) = \int_{R \setminus O} (I(w(x)) - a(x))^2 dx + \int_{O} \beta_0 dx \]

Optimization in occlusion is then easy given \( w \):

pixel \( x \) adds to the energy

\[
\begin{cases}
(I(w(x)) - a(x))^2 & \text{if } x \text{ assigned to } R \setminus O \\
\beta_0 & \text{if } x \text{ assigned to } O
\end{cases}
\]
Joint Optimization With Occlusion

Alternating optimization of

\[ E(O, w; I, a, R) = \int_{R \setminus O} (I(w(x)) - a(x))^2 dx + \int_O \beta_0 dx \]

Optimization in occlusion is then easy given \( w \):

- Pixel \( x \) adds to the energy
  
  \[
  \begin{cases} 
  (I(w(x)) - a(x))^2 & \text{if } x \text{ assigned to } R \setminus O \\
  \beta_0 & \text{if } x \text{ assigned to } O 
  \end{cases}
  \]

Since \( w \) is defined on all of \( R \), a global optimum for \( O \) given \( w \) is

\[ O = \{ x \in R : \beta_0 < (I(w(x)) - a(x))^2 \} \]
Final Joint Optimization

1. Initialize warp to identity and occlusion to empty set
   \( \phi_0(x) = \phi_0^{-1}(x) = x, \ x \in R, \ O_0 = \emptyset \)

2. Compute Sobolev gradient
   \( G_\tau = \nabla_{Sob}E(\phi_\tau | O_\tau, R_\tau) \)

3. Update backward warp
   \( \partial_\tau \phi_\tau^{-1} = \nabla \phi_\tau^{-1} \cdot G_\tau, \ \text{in} \ R_\tau \)

4. Update forward warp
   \( \partial_\tau \phi_\tau = -G_\tau(\phi_\tau), \ \text{in} \ R \)

5. Update estimate of occlusion
   \( O_\tau = \{ x \in R_\tau : (I(x) - a(\phi_\tau^{-1}(x)))^2 > \beta_0 \} \)

6. Repeat 2-5 until convergence
Simulation of Joint Deformation/Occlusion Estimation

Radiance, $a$

Region $R$ in Image to Match

Infinitesimal Deformation (Sobolev Gradient)

Occlusion Likelihood Map
Simulation of Joint Deformation/Occlusion Estimation

Radiance, $a$

Region $R$ in Image to Match

Infinitesimal Deformation (Sobolev Gradient)

Occlusion Likelihood Map
Result of Occlusion/Deformation Estimation

Region and Radiance in Frame t

Warped Region in Frame t+1
Result of Occlusion/Deformation Estimation

Region and Radiance in Frame t

Warped Region in Frame t+1

Warped Occlusion Determined (White)
Result of Occlusion/Deformation Estimation

Region and Radiance in Frame t

Warped Region in Frame t+1

Warped Occlusion Determined (White)

Warped Occlusion Removed
Un-occluded Region Determined
Result of Occlusion/Deformation Estimation

Region and Radiance in Frame $t$

Warped Region in Frame $t+1$

Warped Occlusion Determined (White)

Warped Occlusion Removed
Un-occluded Region Determined

Disocclusion must be determined
Disocclusion: Grouping by Motion

\[ I_t \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \frac{w_1}{R_1} \quad \frac{w_2}{R_2} \quad I_{t+1}

assign pixels to regions by minimum residual of warp:

\[
E_{seg}(\{R_i\}_{i=1}^N) = \sum_{i=1}^N \int_{R_i} \rho(I_{t+1}(w_i|R_i(x)) - I_t(x)) \, dx,
\]

\[
\rho(x) = \begin{cases} 
  x^2 & x^2 \leq \beta_o \\
  \beta_o & x^2 > \beta_o 
\end{cases}
\]

\(w_i\) - Sobolev warp (computed as before)

assign pixels to regions by minimum residual of warp:

\[ E_{seg}(\{R_i\}_{i=1}^{N}) = \sum_{i=1}^{N} \int_{R_i} \rho(I_{t+1}(w_i|R_i(x)) - I_t(x)) \, dx, \quad \rho(x) = \begin{cases} x^2 & x^2 \leq \beta_0 \\ \beta_0 & x^2 > \beta_0 \end{cases} \]

\( w_i \) - Sobolev warp (computed as before)

optimize in regions by Sobolev gradient flow

Ambiguities in Grouping by Motion

Problems

- motion does not exist in occluded region
- motion cannot be determined in constant regions

Ambiguities in Grouping by Motion

Problems

- motion does not exist in occluded region
- motion cannot be determined in constant regions

Solution: Complementary Motion and Appearance

Ambiguities in Grouping by Motion

Problems

- motion *does not* exist in occluded region
- motion *cannot be* determined in constant regions

Solution: *Complementary Motion and Appearance*

- use *backward motion* for region occluded at $t+1$ and not occluded at $t$

Ambiguities in Grouping by Motion

Problems

- motion *does not* exist in occluded region
- motion *cannot be* determined in constant regions

Solution: *Complementary Motion and Appearance*

- use *backward motion* for region occluded at $t+1$ and not occluded at $t$
- use *appearance cues* for
  - regions disoccluded at $t$ and occluded at $t+1$
  - constant regions

Ambiguities in Grouping by Motion

Problems
- motion *does not* exist in occluded region
- motion *cannot be* determined in constant regions

Solution: *Complementary Motion and Appearance*
- use *backward motion* for region occluded at $t+1$ and not occluded at $t$
- use *appearance cues* for
  - regions disoccluded at $t$ and occluded at $t+1$
  - constant regions

**Energy Formulation**

energy:

$$E_{seg} \left( \{ R_i \}_{i=1}^N \right) = \sum_{i=1}^{N} \int_{R_i} f_i(x) dx$$

complementary motion & appearance:

$$f_i(x) = (1 - maf(x)) \text{Res}_i(x) - maf(x) \log p_{i,x} (I_t(x))$$

local histogram

motion ambiguity function:

$$maf(x) = \begin{cases} 
1 & \text{min}_i \text{Res}_i(x) > \beta_0 \text{ or } \sigma(I_{Bx}) < \varepsilon \\
0 & \text{otherwise} 
\end{cases}$$

Sample Disocclusion Detection

motion only
Sample Disocclusion Detection

motion only
Sample Disocclusion Detection

motion only

complementary motion and appearance
Sample Disocclusion Detection

motion only

complementary motion and appearance
Experiment: Occlusion/Disocclusion Modeling is Crucial

**Without** Occlusion / Disocclusion Modeling
Experiment: Occlusion/Disocclusion Modeling is Crucial

**Without** Occlusion / Disocclusion Modeling

**With** Occlusion and Disocclusion Modeled
Experiment: Going Around an Object
Experiment: Going Around an Object
Experiment: Going Around an Object
Comparison to State-of-the-Art
Comparison to State-of-the-Art
Comparison to State-of-the-Art
Comparison to State-of-the-Art
Comparison to State-of-the-Art

Original Video

Proposed Method

Adobe After Effects 2012

Scribble Tracker
Comparison to State-of-the-Art

Original Video

Proposed Method

Adobe After Effects 2012

Scribble Tracker
Comparison to State-of-the-Art

from Freiburg / Berkeley Motion Segmentation Dataset (2014)
Comparison to State-of-the-Art

from Freiburg / Berkeley Motion Segmentation Dataset (2014)
Comparison to State-of-the-Art

from Freiburg / Berkeley Motion Segmentation Dataset (2014)
Comparison to State-of-the-Art

from Freiburg / Berkeley Motion Segmentation Dataset (2014)
Comparison to State-of-the-Art

from Freiburg / Berkeley Motion Segmentation Dataset (2014)
Comparison to State-of-the-Art

from Freiburg / Berkeley Motion Segmentation Dataset (2014)
Quantitative Comparison

Freiburg / Berkeley Motion Segmentation Dataset (Ochs, Malik, Brox (2014))

<table>
<thead>
<tr>
<th></th>
<th>Training set (29 sequences)</th>
<th>Test set (30 sequences)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>[16]</td>
<td>79.17</td>
<td>47.55</td>
</tr>
<tr>
<td>[24]</td>
<td>81.50</td>
<td>63.23</td>
</tr>
<tr>
<td>[1]</td>
<td>87.20</td>
<td>59.60</td>
</tr>
<tr>
<td>[31]</td>
<td>85.00</td>
<td>67.99</td>
</tr>
<tr>
<td>[31]-NC</td>
<td>83.00</td>
<td>70.10</td>
</tr>
<tr>
<td>ours</td>
<td><strong>89.53</strong></td>
<td><strong>70.74</strong></td>
</tr>
</tbody>
</table>

P-precision, R-recall, F-measure, N-objects detected
(higher is more accurate to ground truth)

[16] - Grundman et al., CVPR 2010
[24] - Ochs et al., PAMI 2014
[1] - Ayvaci et al., PAMI 2012
[31] - Taylor et al., CVPR 2015

Summary

• Segmentation by low-level information (relative motion of camera and object) suggests promising direction for objects
Summary

- Segmentation by low-level information (relative motion of camera and object) suggests promising direction for objects.

- Occlusions / dis-occlusions posed a challenge.
Summary

- Segmentation by low-level information (relative motion of camera and object) suggests promising direction for objects.
- Occlusions / dis-occlusions posed a challenge.
- Occlusions, warp treated as a joint problem.
Summary

- Segmentation by low-level information (relative motion of camera and object) suggests promising direction for objects.
- Occlusions / dis-occlusions posed a challenge.
- Occlusions, warp treated as a joint problem.
- New Region-Based Sobolev Shape Metric yielded inherent coarse-to-fine optimization scheme: beneficial for motion estimation.
Summary

- Segmentation by low-level information (relative motion of camera and object) suggests promising direction for objects.

- Occlusions / dis-occlusions posed a challenge.

- Occlusions, warp treated as a joint problem.

- New Region-Based Sobolev Shape Metric yielded inherent coarse-to-fine optimization scheme: beneficial for motion estimation.

- Dis-occlusions solved by complementary motion & appearance grouping.