Structured-Cut: A Max-margin Feature Selection Framework for Video Segmentation

Nikhil Santosh Naikal

CS 294-69 Project Presentation.
December 5, 2011.
Motivation

- Interactive video segmentation very useful for variety of applications.
  - Examples: Rotoscoping, Compositing, etc.
- State-of-the-art methods combine local and global classifiers [1].

Goal:

Use multiple image features to improve foreground/background separation.

Interactive Image Segmentation

Given marked Foreground/Background regions,

Assign binary label to all other pixels.
Interactive Image Segmentation

Markov Random Fields (MRF) formulation:

\[ E(s) = \sum_i \Psi_u(y_i) + \lambda \sum_{i,j \in N} \Psi_p(y_i, y_j). \]

- **s**: Image segmentation.
- **y**: Image pixel nodes.
- **\( \Psi_u \)**: Unary potential.
- **\( \Psi_p \)**: Pairwise potential.

Can be minimized using Graph Cuts [1]

\[ s^* = \arg \min_s E(s) \]

Using Multiple Features

- MRF framework general enough to incorporate multiple features (color, texture, shape, etc.)

\[ E(s) = \sum_i \Psi_u + \sum_{i,j \in N} \Psi_p = w^T \Theta(s) \]

- Composite unary potential \( \Psi_u = \lambda_{\text{color}} \Psi^{\text{color}}_u + \lambda_{\text{texture}} \Psi^{\text{texture}}_u + \ldots \),
- Composite pairwise potential \( \Psi_p = \mu_{\text{color}} \Psi^{\text{color}}_p + \mu_{\text{texture}} \Psi^{\text{texture}}_p + \ldots \),
- Feature weights \( w = [\lambda_{\text{color}} \lambda_{\text{texture}} \ldots \mu_{\text{color}} \mu_{\text{texture}} \ldots]^T \).
Constraints on Energy

\[ E(s_{\text{incorrect}}) > E(s_{\text{gt}}) \Rightarrow w^T \Theta(s_{\text{incorrect}}) > w^T \Theta(s_{\text{gt}}). \]

Exponentially large number of incorrect segmentations possible for an image.
Learning Weights


Given: parameters $w$
Estimate: segmentation

Graph Cuts

Given: positive example of segmentation
Estimate: parameters $w$

Structural SVMs
Example 1

Minimizing composite energy using graph cuts

\[ \Psi_u : \text{composite} = 2.8 \times \text{color} + 0.9 \times \text{texture} + 0.1 \times \text{blur} \]

\[ \Psi_p : \text{composite} = 0.2 \times \text{color} + 0.0 \times \text{texture} + 1.0 \times \text{blur} \]
Minimizing composite energy using graph cuts

\[
\Psi_u : \text{composite} = 0.9 \times \text{color} + 4.1 \times \text{texture} - 2.6 \times \text{blur}
\]

\[
\Psi_p : \text{composite} = 0.8 \times \text{color} + 0.0 \times \text{texture} + 1.8 \times \text{blur}
\]
Extension to Video

- Foreground/background models don't change drastically over subsequent frames.
- Local feature weights propagated to future frames.
Preliminary Results

frame 1  frame 2  frame 3  frame 10
Conclusion

- Color not always good discriminant for foreground/background segmentation.
- Incorporating multiple features (color, texture, blur, etc.) improves segmentation.
- Presented approach to weight multiple features using max-margin formulation.
- Demonstrated effectiveness of approach for image segmentation.
- Extending approach to video segmentation showed promise.

Questions?