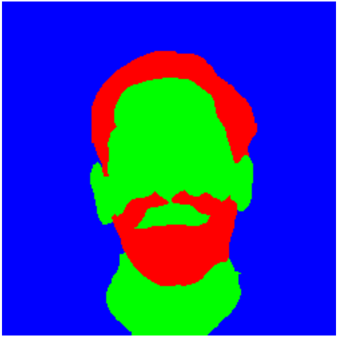
Conditional Random Fields (CRFs) for Facial Segmentation

Karan Goel

November 25, 2015

- Task: Segment faces into hair, skin and background
- Challenging due to varying poses, degree of hair

(a) Original Image



(b) Desired Segmentation

Figure: An example

Conditional Random Fields

- Discriminative modeling in a graphical structure
- Characterized by

$$p(Y|X) = \frac{1}{Z} \exp \left\{ \underbrace{\sum_{k} \lambda_k f_k(y_k, X)}_{\text{unary features}} + \underbrace{\sum_{k_1,k_2} \lambda_{k_1,k_2} f_{k_1,k_2}(y_{k_1}, y_{k_2}, X)}_{\text{binary features}} \right\}$$

- X: observed nodes (pixels)
- Y: hidden nodes (segmentation labels $= 0/1/2$)
- λ: weights, learned using training data
- f_k: node features; likelihood of node label given its pixel value
- f_{k_1,k_2}: edge features, smoothing of label values

Approach

- Implemented CRF approach of Huang et al. [2008]
- Dataset: Labeled Faces in the Wild (Huang et al. [2007])
  - Contains superpixel information – over-segmented images
  - Ground truth information
- Graph Construction
  - 4-way connectivity for hidden layer
  - Each hidden label connected to its pixel as well

Model Details

- Node Features
  - Color: 64 bin color histogram in LAB space
  - Texture: 64 textons
  - Position: Quadrant wise intersection with super-pixel
- Edge Features
  - Color: Euclidean distance in LAB space between mean colors of neighbor superpixels
  - Texture: χ^2 distance between texture histograms,

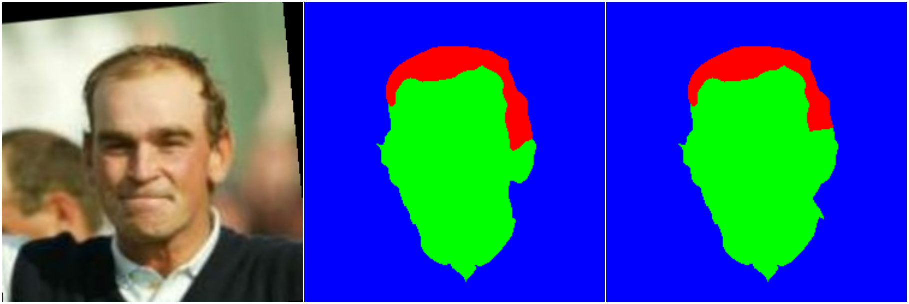
$$\chi^2 = \frac{1}{2} \sum_{i=1}^{64} \frac{(h_1(i) - h_2(i))^2}{h_1(i) + h_2(i)}$$

  - Position: Quadrant wise intersection with super-pixel

Model Learning

- Parameter (λ) Learning: Maximize the conditional log-likelihood of the training data
  - Loopy graph: approximate Z by Bethe approximation with belief propagation to determine marginal probability of y_i
  - Optimized using L-BFGS search
- Inference (segmentation estimation): Maximize marginal probability of y_i in the posterior distribution
  - Implemented using loopy belief propagation with max-marginals

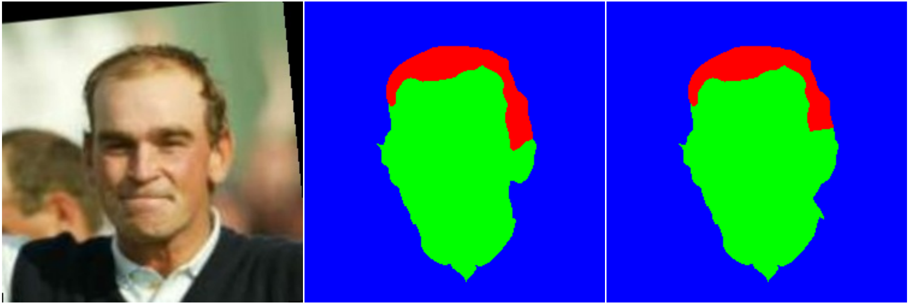
Results



(a) Original Image (b) Ground Truth (c) Segmented Image

Figure: Success Case for CRF-Based Segmentation

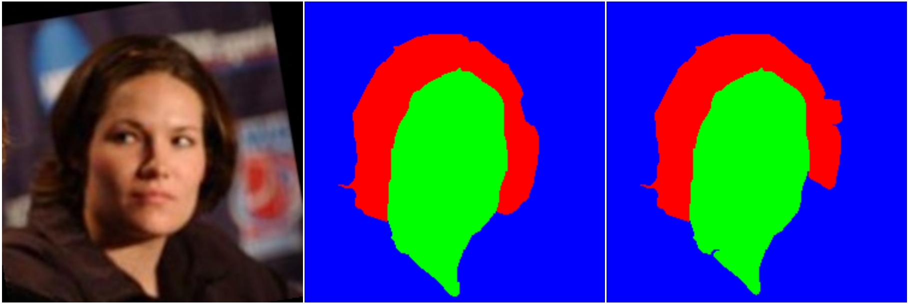
Results



(a) Original Image (b) Ground Truth (c) Segmented Image

Figure: Success Case for CRF-Based Segmentation – Multiple Heads

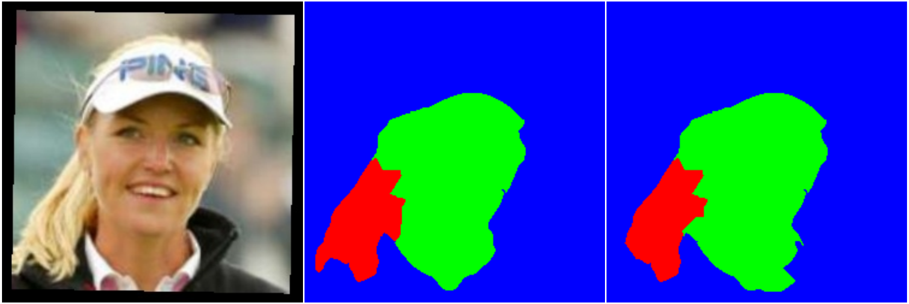
Results



(a) Original Image (b) Ground Truth (c) Segmented Image

Figure: Success Case for CRF-Based Segmentation – Blurred Boundary

Results



(a) Original Image (b) Ground Truth (c) Segmented Image

Figure: Success Case for CRF-Based Segmentation – Hat Occlusion

Results

(a) Original Image (b) Ground Truth (c) Segmented Image

Figure: Failure Case for CRF-Based Segmentation – Comparable Head/Hair Size

Results

(a) Original Image (b) Ground Truth (c) Segmented Image

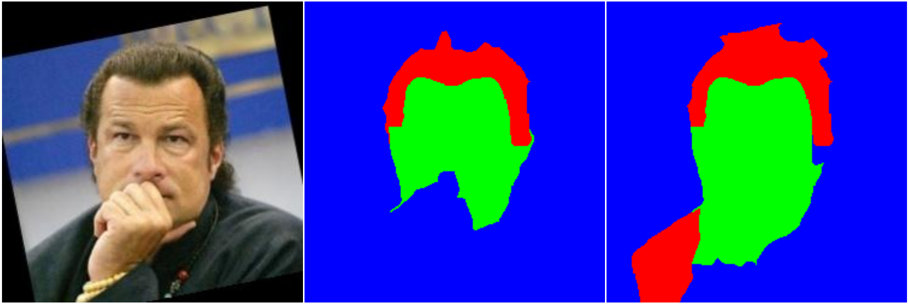
Figure: Failure Case for CRF-Based Segmentation – Many People

Results

(a) Original Image (b) Ground Truth (c) Segmented Image

Figure: Failure Case for CRF-Based Segmentation – Many People

Results



(a) Original Image (b) Ground Truth (c) Segmented Image

Figure: Failure Case for CRF-Based Segmentation – Occlusion

Observations

- Results indicate lack of global information – Kae et al. [2013] explore the use of global shape information using Restricted Boltzmann Machines (RBMs).
- Can make use of part-based models to improve the segmentation.
- Dataset dependence: positional features rely on centrality of face. Can be improved by considering relative spatial information in CRF formulation independent of the location of the face.

References

Gary B Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical report, Technical Report 07-49, University of Massachusetts, Amherst, 2007.

Gary B Huang, Manjunath Narayana, and Erik Learned-Miller. Towards unconstrained face recognition. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pages 1–8. IEEE, 2008.

Andrew Kae, Kihyuk Sohn, Honglak Lee, and Erik Learned-Miller. Augmenting crfs with boltzmann machine shape priors for image labeling. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2019–2026. IEEE, 2013.

Thank You