Family Member Identification from Photo Collections (supplemental materials)

Qieyun Dai[1,2] Peter Carr[2] Leonid Sigal[2] Derek Hoiem[1]

[1]University of Illinois at Urbana-Champaign [2]Disney Research

dai9@illinois.edu {peter.carr, lsigal}@disneyresearch.com dhoiem@uiuc.edu

In this document, we explain how our clothes segmentation is generated, and show results demonstrating how clothes information and role information help with face verification.

1. Clothes Segmentation

To generate clothes segmentation, we first gather clothes prior on the training set, then perform color-based graph cuts on a given test image.

1.1. Clothes Prior Generation

For each training image, we make use of the labeled body parts and body bounding box to define foreground and background regions (see Figure 1.) and build color histograms using pixels inside the foreground and background region by counting. We use L*a*b histograms with 20 bins for the L channel, and 22 bins for the a and b channel each.

For each pixel, we compute

$$P(clothes) = \frac{P(foreground)}{P(foreground) + P(background)}$$

and use its negative log likelihood as the unary potential. We use color difference between adjacent pixels (defined as $\exp(-\alpha\,|(R_i, G_i, B_i) - (R_j, G_j, B_j)|)$) as the pairwise potential and perform graph cuts to generate the clothes region for each training image.

All the generated clothes regions are aligned by resizing the face bounding box to the same size and using the center of the face as the origin. Then we average all the generated clothes region to get the clothes prior (Figure 1).

We also generate a max body bbox prior from the training data. This is done by aligning all the detected faces and choose the mode of annotated body bounding box.

1.2. Test Image Segmentation

For each testing image, we find the regions corresponding to the clothes prior and the maximum body bounding box. Then we build a foreground and background color model (a L*a*b histogram as described in Section 1.1) using pixels inside the clothes prior as foreground seeds and pixels outside the maximum bounding box with a width of 20 pixels as background seeds. To reduce confusion with skin pixels, pixels inside the detected face region are also added as background seeds.

For each pixel, we compute

$$P(clothes) = \alpha_1 \frac{P(foreground)}{P(foreground) + P(background)} + \alpha_2 P(clothes_p rior)$$

and we set $\alpha_1 = 0.7$ and $\alpha_2 = 0.3$, and use its negative log likelihood as unary potential. We use the same pairwise potential as in Section 1.1 and perform graph cuts to generate the clothes region for a testing image. Examples of generated clothes segmentation are shown in Figure 2.

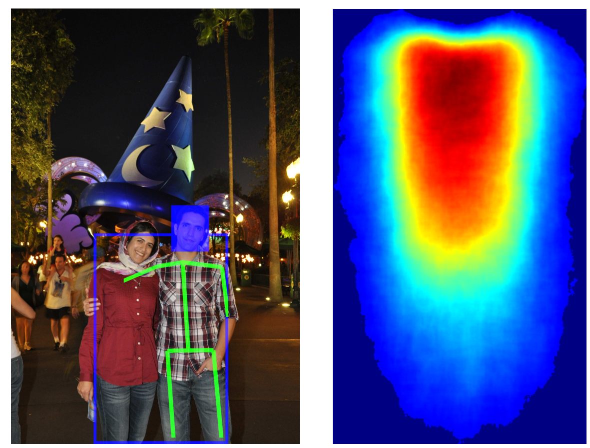

Figure 1. Example images of clothes prior generation from training images and the generated clothes prior. Left: a training image and the regions selected as foreground and background. Foreground regions are shown in green, and background regions are shown in blue. Right: clothes prior averaged over the training set.

2. Face Verification

In this section, we show face verification results on the test set using different feature combinations. We compare classification using all the features (facial feature similarity, clothes similarity, role similarity, verificaiton score(Li

Figure 2. Examples of generated clothes segmentation on the testing set.

Set	Score		All		-Score		-Face		-Clothes		-Role	
	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC
1	0.701	0.824	0.777	**0.936**	**0.815**	0.928	0.780	0.921	0.757	0.917	0.677	0.887
2	0.654	0.762	0.786	0.867	**0.787**	**0.869**	0.756	0.848	0.779	0.858	0.710	0.795
3	0.687	0.700	0.692	**0.771**	0.663	0.753	0.682	0.760	**0.697**	0.769	0.681	0.707
4	**0.679**	0.631	0.593	0.639	0.570	0.626	0.597	0.619	0.584	**0.643**	0.626	0.640
5	0.805	0.842	**0.862**	**0.920**	0.856	0.891	0.848	0.911	0.850	0.915	0.812	0.873
6	0.696	0.724	0.725	**0.797**	0.716	0.766	0.723	0.802	**0.726**	**0.797**	0.683	0.716
7	0.608	0.643	0.668	**0.717**	0.656	0.708	**0.669**	**0.717**	0.655	0.692	0.631	0.678
Avg	0.690	0.732	**0.729**	**0.807**	0.723	0.792	0.722	0.797	0.721	0.799	0.689	0.757

Table 1. Face Verification Results Using different feature combinations. Score: face verification score returned by Li et al. [1], All: classifier trained using all features, *-feature_name*: classifier trained by removing the specific features. Acc: classifier accuracy thresholded at 0.5. AUC: area under the ROC curve.

et al. [1])) and models trained by removing one of the features. For a detailed description of the features, please refer to the main paper. The results are shown in Table 1. It's clear that, by adding role information, the verification performance increased greatly in terms of both accuracy and area under the ROC curve.

References

[1] H. Li, G. Hua, Z. Lin, J. Brandt, and J. Yang. Probabilistic elastic matching for pose variant face verification. In *CVPR*, 2013.