

Unsupervised Tattoo Segmentation Combining Bottom-Up and Top-Down Cues

Josef D. Allen^{1a,b}, Nan Zhao^b, Jiangbo Yuan^b, Xiuwen Liu^b

^aOak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN, U.S 37831;

^bDepartment of Computer Science, Florida State University, Love Building, Tallahassee, FL, U.S 32306

ABSTRACT

Tattoo segmentation is challenging due to the complexity and large variance in tattoo structures. We have developed a segmentation algorithm for finding tattoos in an image. Our basic idea is split-merge: split each tattoo image into clusters through a bottom-up process, learn to merge the clusters containing skin and then distinguish tattoo from the other skin via top-down prior in the image itself. Tattoo segmentation with unknown number of clusters is transferred to a figure-ground segmentation. We have applied our segmentation algorithm on a tattoo dataset and the results have shown that our tattoo segmentation system is efficient and suitable for further tattoo classification and retrieval purposes.

Keywords: Tattoo segmentation, Tattoo, Soft Biometrics, Top-Down, Bottom-Up, Figure-Ground

1. INTRODUCTION

Scars, Marks and Tattoos, (SMT), are useful clues for criminal identification and personal identification in criminal convictions and medical forensics respectively. Although canonical biometric identifiers such as fingerprint, DNA and iris are still the primary assured way for identification, SMT can supplement; these identifiers can be considered *soft-biometrics*. Hence, a large amount of tattoo images have also been taken from victims, suspects and incarcerated personnel for identification in law enforcement¹. Collectively, these canonical and soft biometric markers are collected, maintained and analyzed by national security systems like the Integrated Automated Fingerprint Identification System (IAFIS) for retrieval purposes².

Manual tattoo searches over a large datasets are very time-consuming and inefficient. Several Content-Based Image Retrieval (CBIR) systems have been proposed for tattoo matching and retrieval^{1, 3, 4}. The performance of these systems is sensitive to tattoo segmentation, which is a preprocessing step removing varying background. Our goal in this paper is to design a semantic-level segmentation system to automatically mark tattoo regions for CBIR systems. Among a myriad of other things this type of semantic level segmentation will enable us to categorize the gangs amongst themselves by understanding the uniqueness of the intra gang affiliations, “i.e. sets”. It will also enable us to autonomously classify the Tattoo/Ink Artist and how his/her style varies from set-to-set. Finally, we would like our work to be able to categorize the database into what we would call a “friend or foe” type of way such that migratory elements of individual gang/set populations are better understood from views of both the continental United States (CONUS) and the outside of Contiguous United States (OCONUS).

¹ Further author information: (Send correspondence to Josef.D.Allen.)
Josef.D.Allen.: E-mail: allenjd@ornl.gov, Telephone: 1 321 759 0345
Xiuwen Liu: E-mail: liux@cs.fsu.edu, Telephone: 1 850 644 0050
Nan Zhao: E-mail: jolen217@gmail.com, Telephone: 1 850 345 1975
Jiangbo Yuan: E-mail: yuanjiangbo@gmail.com, Telephone: 1 850 320 1279

The objective of tattoo segmentation systems is to automatically extract regions solely containing component(s) of a tattoo(s) in an image. Tattoo segmentation requires that each extracted region have a semantic component, *contain meaningful and actionable data*, of the tattoo. This is more than traditional image segmentation⁵⁻⁸, which only requires that each segmented region be homogeneous and are often called *oversegmentation*. In contrast, the tattoo regions are usually not homogeneous in color space. In fact, tattoo segmentation can be considered as a case of *figure-ground semantic* (object) segmentation, i.e. distinguishing the semantically meaningful object(s), tattoos, from the background to include skin.

Semantic segmentation is a more challenging topic that has been widely discussed in recent years⁹⁻¹². Bottom-up segmentation is sensitive to intra-object variance. Various forms of top-down cues are usually combined with bottom-up cues for the purpose of obtaining semantic meaningful results. For example, Schitman et. al.¹² found that a group of patch sets (one for each class and labels are known) can help label the homogeneous regions obtained from bottom-up over segmentation. Each set contains patches sampled from one class in the labeled training image and the cost of the assignment to each class is computed for each over segmented region. A graph-cut optimization based on these costs is used to find a globally optimal segmentation. Carreira et al. applied multi-scale binary segmentation on an image using parametric min-cuts technique⁹. Then a feature-based regressor is trained to rank the pool of segmentation results to predict each result's likelihood of being an object. Such regressor's is learned from the statistic distribution of a large number of features (related to graph, region and Gestalt properties) among a set of annotated images. To defeat manually labeling the training set, a large number of images containing the same object were simultaneously segmented in another work¹⁰. It is under this assumption that the common parts of an object will appear frequently while the effect from varied background will diminish. In such an approach, superpixels and interest points are re-organized as mid-level over segmentation results and visual words representations from bottom-up and top-down priors of a hybrid graph model respectively.



Figure 1. Three groups of tattoo examples from three gangs. The major patterns for these three groups (from left to right) are number four, a cross with a posture of hand for prayers nearby and rabbit head respectively.

The key to semantic segmentation is how to incorporate top-down priors. However, obtaining top-down priors for tattoo is very challenging due to large variances in tattoo appearance, shapes and spatial connectedness. Different gangs/sets have different types of tattoos with their own symbolic system. It could be certain number, an animal or a combination of several meaningful components (see Figure. 1). Despite the fact that the number of gangs/sets is limited, the appearance of the tattoo components (such as letters, numbers and animals) in an individual gang still varies from person to person. As an illustration, consider tattoo examples in Figure. 1. The image on the left, with four smaller tattoo images, shows the variance of tattoo in its appearance. Even though number four is the major pattern for all these four tattoos, the appearance (such as texture and color) of each number four is significantly different. The image in the middle illustrates the variance of tattoo shape. In such tattoos, different fonts were used for the letters on the paper and those beside the animals. Letters in a single tattoo may even have different shapes. A considerable spatial variance of tattoo is shown via the image on the right. A tattoo may not be a collection of spatially related components, such as rabbit head and letters in this image. In such a case, the neighborhood of components is difficult to be involved in segmentation model as spatial constraints for region aggregation, a strategy that can finally obtain the object as a whole. Suffering from these shortcomings, it is difficult to segment tattoos from the background directly through any prior on location or shape of tattoo as a whole and thus this problem is rarely discussed from the work we have read so far.^{1, 4, 5} Jain et al. proposed an image retrieval system for tattoo images via SIFT-feature-based image matching.⁴ However, better performance was reported when their automatic tattoo segmentation step was replaced by manual segmentation, which is really time-consuming on the other hand. Acton and Rossi proposed a segmentation approach based on active contour

and vector field convolution.⁵ Unfortunately, the contour initialization is difficult since the structure of a tattoo is complicated.

We propose a tattoo segmentation system combining both bottom-up and top-down priors. We make the assumptions that each component in a tattoo is arbitrarily distributed in spatial space of a tattoo image. Moreover, we consider skin and tattoo as a whole and deal with a figure-ground segmentation for both skin and tattoo as the foreground. Figure-ground segmentation is a recognition process that needs to figure out the aimed object(s). Thus top-down priors should be involved for this step. Yet, due to the intra-variance of skin, tattoo and background, these top-down priors should be learned from the image being segmented, rather than from other labeled images. After obtaining regions with skin and tattoo, another figure-ground segmentation distinguishes the tattoo from skin. Similar to the segmentation in the previous step, top-down priors learn from the image. This two-stage process agrees with the hierarchy and adaptiveness of human visual system in visual scene understanding¹³.

The rest of this paper is organized as follows: Section 2 presents the proposed system for tattoo segmentation in details and experimental results are shown in section 3. Section 4 discusses some potential of proposed system and a conclusion is given at the end.

2. ALGORITHM DESCRIPTION

Regarding the complexity and intra-class variance of tattoo, our main idea is to transfer the tattoo segmentation into skin detection followed by a figure-ground segmentation. The outline of our algorithm is depicted in Fig. 2.

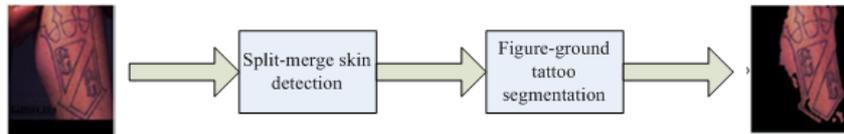


Figure 2. The outline of tattoo segmentation.

Initially a clustering technique is used on the range domain (color space) to separate the tattoo image into oversegmented regions in an bottom-up manner. This is a very important step in that regions containing both skin and tattoo are much more non-homogeneous than these oversegmented regions. Based on top-down cues learned from the image itself, a region merging step is introduced to group skin regions together. Through this split-merge process, skin and tattoo are distinguished from the background. Finally, K-Means algorithm is applied for figure-ground segmentation, where now the tattoo is the foreground and the skin is the background.

2.1 Split-merge skin detection

For skin detection, an image is usually regarded as a group of feature vectors. Each pixel corresponds to a feature vector in multi-dimensional feature space. The statistic properties of the histograms or the distributions built on these dimensions are widely discussed^{14, 15}. To make our system clear, we simply begin by applying an initial clustering process on the gray-scale distribution $f(I)$ of the image I so as to obtain an initial segmentation. Here $f(I)$ evaluates the density estimate covering the range of intensity in the gray-scale form of I . Since pixels from the same cluster are more likely to have the similar intensity, we segment the tattoo images according to the local minima of distribution $f(I)$. In that sense, the pixels with intensity values between either two closest local minima are labeled as a cluster. To this end, other clustering algorithms like watershed¹⁶ and meanshift¹⁷ could be used as alternatives.

After applying the histogram-based clustering, $d-1$ local minima in $f(I)$ indicates d clusters. Obviously, such initial segmentation may suffer from under-segmentation due to background with similar intensity and over-segmentation due to illuminant variation on the skin. However, the following points can help solving such problems:

1) Initialization

Since only weak prior information in detecting tattoos can be obtained in a given image, the segmentation system needs to be initialized, where an initial model of skin and tattoo needs to be estimated. In this paper, based on the observation that the center region in almost all the tattoo images in the dataset contains skin, we use a center in focus initialization, considering the statistic properties of the center region as prior knowledge for tattoo and skin detection

(Note: not only tattoo) Clearly this initialization step is not universal and this can be replaced if some prior information about the skin color is available.

2) Connected structure

Although skin with tattoo may be split into several regions (clusters) due to shading and intra-class variance of tattoo components, these regions are more likely to be connected to each other in spatial space. If the main pattern of tattoo is not in the center of an image, it may still be merged since its background in the cluster (the skin) may be adjacent to the pure-skin region in the center. Therefore, it is reasonable to use the neighborhood of clusters in spatial domain as a top-down cue for merging potential skin regions.

Following these two points above, an m -by- n sample patch p_0 in the center of I is sampled from I for obtaining prior knowledge (empirically, m and n is half of the height and width of I correspondingly). How to obtain the seeds for learning the top-down prior of the objects is widely discussed in figure-ground segmentation^{9, 10}. In the work of Joao et al.⁹, foreground seeds are randomly picked up from a group of pixels uniformly distributed over the spatial space for several times. The background seeds are those on the boundary of an image. Liu et al.¹⁰ use interest points obtained from a large amount of image containing the same object as the prior for such object. Regarding our initialization, clusters covering the major area of p_0 are most likely to be skin and tattoo. Thus in our work, a region sampled from the image is regarded as the seeds containing top-down priors. In detail, the number of overlapped pixels between p_0 and each cluster c_i are sorted as $\{(n_j, i) | n_j > n_{j+1} \text{ and } i, j = 1 \dots d\}$. The first k clusters with the largest number of overlap pixels are labeled as potential foreground (skin and tattoo) under the constraint that their overlap ratio $\frac{\sum_{i=1}^k n_i}{\sum_{i=1}^d n_i}$ exceeds a threshold t (typically 75%). In most cases, such potential foreground may contain arbitrary number of regions merged by clusters. However, only the region with overlap in the sample patch is labeled as foreground due to the connected structure of skin. Therefore, only one region with sample patch inside is segmented as skin. If either skin or tattoo dominates the sample patch, regions belonging to the other may be excluded from the segmented region. Thus an operator filling the holes inside such region should be applied as a post-process of skin detection.

2.2 Figure-ground tattoo segmentation

To this point, we transferred a problem of tattoo segmentation with unknown number of clusters to a skin-tattoo binary segmentation. In this section, skin pixels should be distinguished from pixels belonging to tattoo. In that sense, skin pixel in this section indicates merely the skin pixel that is not covered by tattoo. Since we already know the number of potential clusters now, a k -means algorithm ($k = 2$) can be applied on the RGB color space of the foreground (skin). Now the issue is which cluster should be tattoo. If distinction between tattoo and skin is needed, the pixels on the contour of the skin region are more likely to be skin pixels rather than pixels in tattoo. Because, otherwise, skin are fully covered by tattoo and distinguishing tattoo from skin is thus unnecessary. Following this rule, the cluster with more pixels on the contour of the foreground is labeled as the skin and the other the tattoo. If the structure of tattoo is preferred rather than the whole region containing tattoo in application, an alternative way of marking the tattoo is to apply a ridge or edge detector^{18, 19} on the skin region. This is reasonable since tattoo is a kind of man-made painting with clear boundaries while skin regions are more textureless in contrast.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Experimental results

We have tested our proposed algorithm over a dataset of 256 tattoo images collected from the GANGINK.COM tattoo database². Each tattoo is unique in the sense that no two images are of the same tattoo, and under different views and/or illuminant conditions. Figure 3 shows the segmentation results for a small sample set. It can be shown that the algorithm can separate the tattoo regions from most of the backgrounds. Although in some examples certain small tattoo regions are further eliminated. This is not a big problem for the application of tattoo classification purpose since these regions are small in contrast to the whole tattoo and the main patterns of the tattoo are well preserved and unique.

²<http://www.gangink.com/>



Figure 3. Segmentation results. The first and third row shows the original tattoo images; the second and fourth row shows the results of segmentation using K-Means in the figure-ground segmentation step.

Figure 4 shows the accuracy distribution of the proposed algorithm. Here the segmentation accuracy is the most widely used evaluation metric defined as following:

$$accuracy = \frac{|S_f| + |S_b|}{|S|} \quad (1)$$

where S_f and S_b are correctly assigned foreground and background pixels correspondingly and S is the image. $|X|$ means the number of data (pixels) in the set X . Since bad segmentation may receive a good “accuracy”. If tattoo is small, as suggested by Liu et al.¹¹, our algorithm was also evaluated under a popular measure for information retrieval, called F-measure:

$$F_\alpha = \frac{(1 + \alpha) \cdot R_r \cdot R_p}{\alpha \cdot R_p + R_r} \quad (2)$$

where α is a balance parameter for precision

$$R_p = \frac{|A \cap B|}{|A|} \quad (3)$$

and recall

$$R_r = \frac{|A \cap B|}{|B|}, \quad (4)$$

where A indicates the man-made ground truth segmentation and B the result of proposed segmentation algorithm. α is usually set as 2. Its distribution is shown in Fig. 5. More details of the accuracy and the F measure of proposed algorithm are shown in Table 1.

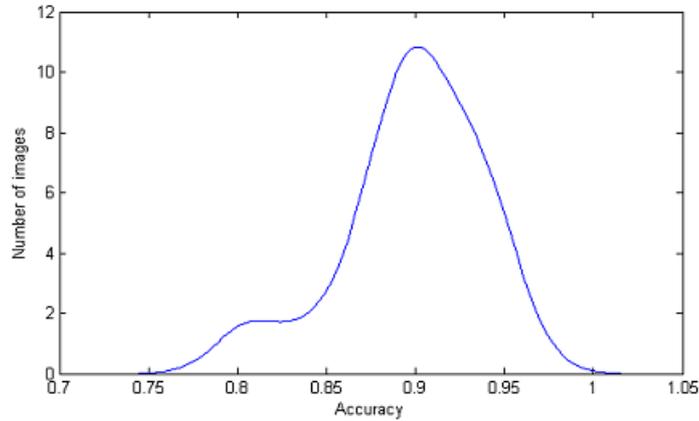


Figure 4. Accuracy of the proposed algorithm. The x-axis is the accuracy and the y axis is the number of images involved in each accuracy.

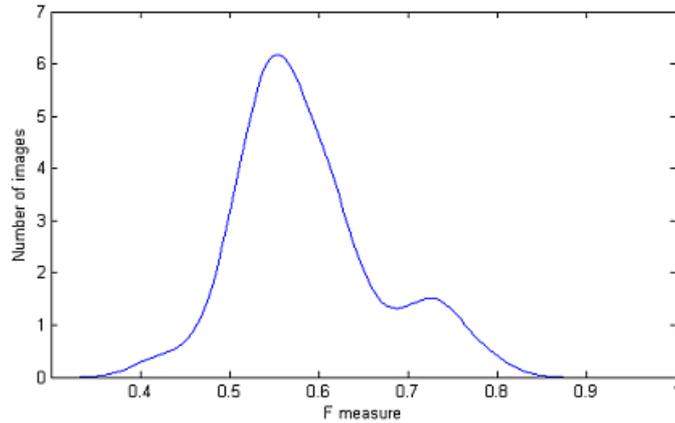


Figure 5. F measure of the proposed algorithm ($\alpha = 2$). The x-axis is the accuracy and the y axis is the number of images involved in each F measure.

Table 1. More details of the accuracy and the F measure of proposed algorithm

	Min	Max	Mean	Variance
Accuracy	0.7989	0.9614	0.8983	0.0014
F measure	0.4204	0.7854	0.5866	0.0061

3.2 Discussion

Even though a tattoo usually contains several components with arbitrary shapes, the typical distinctive components between gang tattoos are often unique embedded patterns in the tattoo. They usually contain semantic level patterns with more describable structures, e.g number four and rabbit head in (see Figure. 1). Unlike a tattoo as a whole, these types of patterns have good connectedness and can thus give us clues on the tattoo artist and particular “set” that

the person may originate from. In general, we call them *tattoo features*. In other words, tattoo features are the primary connected components in tattoos with special structure contributing to tattoo matching and retrieval.

To achieve intrinsic tattoo segmentation, tattoos may be further separated into different connected components. As mentioned in the end of section 2, we can represent the shape via a ridge detector¹⁸. As the local extreme pixels along the largest surface curvature, ridges can be extracted via computing the eigenvalue of the Hessian matrix. After applying the ridge detector, a set of connected components based images can be collected, where each connected ridge can be regarded as a finer component. In other words, each tattoo is ‘cropped’ into finer components with clearer shape structure. Since tattoo features are most distinguishable and thus obvious in tattoo, they are among the components with the best connectedness in most cases. Therefore, it is reasonable to label the first k connected components with the best connectedness as potential tattoo features for further analysis.

Figure 6 shows the shape structure as well as its components for some of the segmentation results in Fig. 3. It indicates that, even though structure for the tattoo as a whole is difficult to be represented, the way of tattoo classification based on shape analysis of primary components is worth trying. Based on these unit-connected components, the shape-like features and the ink style features can be modeled separately for each part of tattoos. This is of great importance as these aforementioned components give us additional information about the region, country, set, etc of a potential. Those connected components based features build up the database. For ink style features, it will be gradient-like features (measuring the sharpness of the lines) and kernel-density estimation of the color distributions. For shape-like features, there are several of choices such as spectral histogram features, shape context and other features that are more sensitive to shapes. For shape context, we can borrow the method from the previous working done by researchers from the University of California at Berkeley²⁰. This kind of shape recognition has been applied well to recognize silhouettes, trademarks and handwritten digits. Since the similar properties between tattoos and those types of data, the shape context features are intuitively a strong potential method for tattoo. Given a new query for a tattoo image, the steps above can be repeated and generate a set of features based on connected components. Through similarity measurement, the query tattoo can be labeled as one of the known classes in the database.

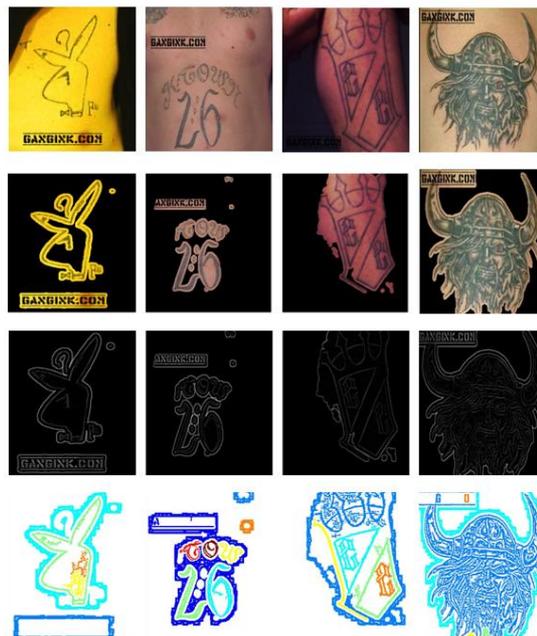


Figure 6. Results shown are connected components of tattoo. The first row shows four original tattoo images; the second row shows the results of initial segmentation; the third row shows the further segmentation results from the ridge-based descriptor; the last row shows different connected components in different colors.

4. CONCLUSIONS AND FUTURE WORK

To help tattoo image retrieval, we have proposed a prototype tattoo segmentation system. With the growing use of tattoos for identification purpose in law enforcement, our segmentation system will be of great value for tattoo image retrieval systems. Despite the simplified design of the algorithm, its process is quite general for tattoo segmentation. The preliminary segmentation results based on a relatively small tattoo dataset are promising, closing to 90% in average accuracy. However, the complicated structure of tattoo and varied background require a more robust and sophisticated segmentation algorithm for improving the performance of tattoo image retrieval. We have also shown that tattoo features (principle connected components of tattoo) obtained via segmentation have great potential in distinguishing tattoos during the retrieval. Our future work addresses (i) enlarging the tattoo dataset, (ii) generating more manually segmented results for evaluation purpose, (iii) developing a more sophisticated segmentation algorithm for tattoo retrieval, (iv) finding a way to model tattoo features, and (v) building an image retrieval system based on tattoo features.

ACKNOWLEDGMENTS

The authors would like to thank Tony West, Amanda West & Chris Hall of Emtec Federal and Dr. John Schmitt, President of ID Solutions, Inc. for their support, encouragement and assistance in preparing this manuscript.

REFERENCES

- [1] Jain, A. K., Lee, J.-E., and Jin, R., "Tattoo-id: automatic tattoo image retrieval for suspect and victim identification," Proc. of the multimedia 8th Pacific Rim conference on Advances in multimedia information processing, PCM'07, 256-265, Springer-Verlag, Berlin, Heidelberg (2007).
- [2] www.biometriccoe.gov/doc/IAFIS_0808_One-Page825.pdf
- [3] Jain, A. K., Lee, J. E., Jim, R., and Gregg, N., "Content-based image retrieval: an application to tattoo images," Proc. Int. Conf. on Image Processing, 2745-2748 (2009).
- [4] Acton, S. T. and Rossi, A., "Matching and retrieval of tattoo images: Active contour cbir and glocal image features," Proc. of the 2008 IEEE Southwest Symposium on Image Analysis and Interpretation, 21-24, IEEE Computer Society, Washington, DC, USA (2008).
- [5] Comaniciu, D. and Meer, P., "Mean shift: A robust approach toward feature space analysis," IEEE Trans. on Pattern Analysis and Machine Intelligence 24(5), 603-619 (2002).
- [6] Vincent, L. and Soille, P., "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," IEEE Trans. on Pattern Analysis and Machine Intelligence 13(6), 583-596 (1991).
- [7] Boycov, Y. Y. and Jolly, M. P., "Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images," Proc. of the 8'th International Conference on Computer Vision (ICCV 2001), 105-112 (2001).
- [8] Shi, J. and Malik, J., "Normalized cuts and image segmentation," IEEE Trans. on Pattern Analysis and Machine Intelligence 8(22), 888-905 (2000).
- [9] Joao, C. and Cristian, S., "Constrained parametric min-cuts for automatic object segmentation," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 3241-3248 (2010).
- [10] Liu, G., Lin, Z., You, Y., and Tang, X., "Unsupervised object segmentation with a hybrid graph model (hgm)," IEEE Trans. on Pattern Analysis and Machine Intelligence 32(5), 910-924 (2010).
- [11] Zhu, Z., Wang, Y., and Jiang, G., "Statistical image modeling for semantic segmentation," IEEE Trans. on Consumer Electronics 56(2), 777-782 (2010).

- [12] Schnitman, Y., Caspi, Y., Cohen-Or, D., and Lischinski, D., "Inducing semantic segmentation from an example," Proc. of the 7th Asian Conference on Computer Vision (ACCV 2006) 3852, 373-384 (2006).
- [13] Sharon, E., Galun, M., Sharon, D., Basri, R., and Brandt, A., "Hierarchy and adaptivity in segmenting visual scenes," Nature 442(7104), 810-813 (2006).
- [14] Jones, M. J. and Rehg, J. M., "Statistical color models with application to skin detection," Int. Journal of Computer Vision 46(1), 81-96 (2002).
- [15] Vezhnevets, V., Sazonov, V., and Andreeva, A., "A survey on pixel-based skin color detection techniques," GraphiCon03, 85-92 (2003).
- [16] Aldiol, A., Torres, L., and Delp, E. J., "An unsupervised color image segmentation algorithm for face detection applications," Proc. of the IEEE Conference on Image Processing 2, 681-684 (2001).
- [17] Lu, J., Gu, Q., Plataniotis, K. N., and Wang, J., "A comparative study of skin-color models," Proc. of the IEEE Conference on Image Analysis and Recognition 3656, 729-736 (2005).
- [18] Lindeberg, T., "Edge detection and ridge detection with automatic scale selection," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 465-470 (1996).
- [19] Canny, J. F., "A computational approach to edge detection," IEEE Trans. on Pattern Analysis and Machine Intelligence 8(6), 679-698 (1986).
- [20] Belongie, S., Malik, J., and Puzicha, J., "Shape matching and object recognition using shape contexts," IEEE Trans. on Pattern Analysis and Machine Intelligence 24, 509-522 (April 2002).