SKIN DETECTION BASED ON MULTI-SEED PROPAGATION IN A MULTI-LAYER GRAPH FOR REGIONAL AND COLOR CONSISTENCY

Insung Hwang, Yoonsik Kim, Nam Ik Cho

Dept. of Electrical and Computer Engineering
Seoul National University, Seoul, Korea

ABSTRACT

We propose a new skin detection method based on multi-seeds propagation in a multi-layer graph representation of an image. Initially, some of nodes in the graph are set to be foreground or background seeds based on a simple Bayesian skin detector, and they are propagated through the graph to find the skin probability in the manner of semi-supervised learning. The graph is designed to consider not only local and global coherence but also to consider the color consistency by constructing a multilayer graph of image and cluster layers. Extensive experiments on several datasets are conducted, which demonstrate that our method outperforms the existing methods in terms of various quantitative measures, such as accuracy, precision, recall and F-measure.

Index Terms— skin detection, multi-layer graph, seed propagation

1. INTRODUCTION

Skin detection is still regarded as a difficult problem due to illumination variations, skin color variations according to race and makeup, and the overlap between skin and non-skin pixels in the color space. We focus on the third problem in this paper by exploiting the spatial relationship of the regions. The proposed method is based on graph representation of an image, where the graph is composed of two layers. One layer consists of image nodes which will be called image layer, where each node is a region of an over-segmented image. The other layer is called cluster layer, where the nodes are the clusters of image nodes according to their color similarity. The nodes in the image layer are connected according to their spatial connectivity, and they are also indirectly linked to the cluster layer according to color similarity regardless of spatial distance. The graph is learned by propagating predefined seeds over the whole nodes of graph [1], where we define two types of seeds in our problem, i.e., skin and non-skin (background) seed. Each kind of seed is separately propagated through the graph and a skin probability map is generated by combining the results of each propagation.

The contributions of the proposed method are summarized as: 1) introducing a method that links the spatial and color information in a unified framework to maintain color and spatial coherence together, 2) exploiting non-skin seeds along with skin seeds to effectively suppress background regions.

This research and paper is supported by

2. RELATED WORKS

The simplest skin detection method may be to set a threshold or hyperplane that discriminates color features of skin in a certain color space [2, 3]. There are also parameteric [4–6] and non-parametric modeling [7, 8] of skin regions, which usually show higher accuracy than the threshold methods. However, all of these methods are still sensitive to illumination changes, ethnic groups and training data. Hence, adaptive skin models have been proposed to overcome this problem. Specifically, Zhu *et al.* proposed a two-stage skin detection method that extracts skin pixels by a classical method at first, and then adaptively learns a Gaussian mixture model [9]. An illumination adaptation method was also proposed in [10], and there are also some methods that find skin region by using human-related features [11, 12]. There is also a fusion method combining a dynamic threshold and a single Gaussian model [13], and a method based on a luminance adaptive color channel [14].

Recently, some researchers have focused on the spatial relation of skin pixels along with the skin colors. For example, a controlled diffusion method was used to transfer the skin probability of pixels to its neighbors in [15]. The method that propagates skin seeds based on the distance transform was proposed by [16], and they also modified their previous method by introducing a texture based seed extraction scheme [17]. Considering the spatial relations have contributed to overcome the problems of overlapping skin and detecting non-skin pixels on the color spaces in a certain extent.

3. PROPOSED METHOD

The proposed method is consisted of 5 steps: preprocessing for image segmentation and initial skin probability computation, graph construction, multi seed extraction, seed propagation, and region to pixel refinement.

3.1. Preprocessing

3.1.1. Bayesian Skin Probability

We employ the Bayesian skin color classifier [7] to extract the initial skin seeds. To be precise, we build two histograms of $64 \times 64 \times 64$ bins referring to [17], each of which respectively represents skin or non-skin histogram from a large training dataset. The histograms are normalized in order to make the sum of each histogram becomes 1, which is given by $P(c|S)$ and $P(c|N)$ for skin and non-skin respectively. Assuming that the presence of skin and non-skin pixels is identical ($P(S) = P(N) = 0.5$), the Bayesian probability comes as:

$$
\begin{aligned}
P(S|c) &= \frac{P(c|S)P(S)}{P(c|S)P(S) + P(c|N)P(N)} \\
&= \frac{P(c|S)}{P(c|S) + P(c|N)}.
\end{aligned} \tag{1}
$$

A look-up table for skin probability is obtained from the above equation, and we generate the initial skin probability map (iSPM).

3.1.2. Image Segmentation

An image is divided into over-segmented regions based on the SLIC algorithm [18]. The set of over-segmented regions are denoted as $V^I = \{v_i^I\}_{i=1}^N$ where N is the number of segments, and each region is described by two types of features: color and initial skin probability of the region denoted as $\mathbf{x}_i$: *Lab mean color vector* obtained by averaging color vectors of all pixels belonging to v_i^I, and p_i: *mean skin probability* obtained by averaging iSPM of all pixels belonging to v_i^I.

3.2. Graph Construction

A graph for our problem is constructed with a set of nodes V and a set of edges E, which is denoted as $G = (V, E)$. We construct a two layer graph (image layer and cluster layer) where each layer has different type of node from the other, denoted as $V = \{V^I, V^C\}$ where V^I is the node in the image layer defined in the above subsection and V^C is the node in the cluster layer. Fig. 1 is the illustration of the proposed multilayer graph, which shows that the first layer is the image layer whose node is over-segmented region and the second layer has the nodes as the cluster of these regions.

As we have two layer graph, there are three types of edges connecting the nodes inside and inter layers, denoted as $E = \{E^I, E^C, E^{VC}\}$, which are the edges connecting the nodes inside the image nodes, inside the cluster nodes and between the image and cluster nodes respectively. We design the graphs for image and cluster layers separately, and then connect them with E^{VC}.

The nodes of the first layer are connected to their neighbors with weights defined as:

$$
w_{ij}^I = \exp(-(|p_i - p_j|^2)/\sigma_p^2), \tag{2}
$$

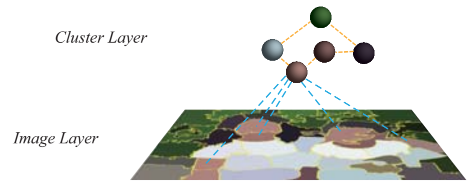

Fig. 1: Illustration of the proposed graph for skin detection. It also shows inter graph connection that the nodes in the image layer are connected with a node in the cluster layer with the highest color similarity (blue dot lines).

where σ_p is a control parameter for intra node affinity. As in [19], the σ_p is designed to be adjusted according to the input image statistics, which is defined as:

$$
\sigma_p = \sqrt{\frac{1}{|E^I|} \sum_{e_{ij} \in E^I} (p_i - p_j)^2}, \tag{3}
$$

where $\left|E^I\right|$ is the number of intra edges. In the proposed method, we link the edges by measuring the similarities of skin probability between them, while other methods including [19, 20] connect the edges according to color affinity. From the skin probability weights of edges, an affinity matrix $\mathbf{W}_I$ is constructed whose (i, j)-th element is the weight between the i-th and j-th node, defined as:

$$
(\mathbf{W}_I)_{i,j} = \begin{cases} w_{ij}^I, & \text{if } i \in Q_j, \\ 0, & \text{otherwise,} \end{cases} \tag{4}
$$

where Q_j is a set of neighbor nodes for v_i^I, which implies that the nodes in the image layer are sparsely connected to only its neighbors.

For the construction of cluster layer graph, we first categorize the image nodes into some groups using K-means clustering algorithm from color feature $\mathbf{x}_i$. We define this cluster as a node in the cluster layer denoted as v_i^C, and define the set of these nodes as $V^C = \{v_i^C\}_{i=1}^K$, where K is the number of clusters. We also define the centroid of the i-th cluster as the average of all color vectors of the image nodes belonging to the cluster, which is denoted as $\mathbf{c}_i$. The affinity for intra cluster graph is defined as the similarity between the centroids of clusters which is written as:

$$
w_{ij}^C = \exp(-\left\|\mathbf{c}_i - \mathbf{c}_j\right\|/\sigma_c), \tag{5}
$$

where σ_c is a control parameter for the similarity, which is a fixed value unlike the image graph case, because there is no sufficient connections between the clusters for computing statistics. The affinity matrix $\mathbf{W}_C$ is constructed such that its (i, j)-th element is the affinity between i-th and j-th node:

$$
(\mathbf{W}_C)_{i,j} = \begin{cases} w_{ij}^C, & \text{if } i \in k\text{-NN}(j) \text{ or } j \in k\text{-NN}(i), \\ 0, & \text{otherwise,} \end{cases} \tag{6}
$$

where the k-NN(j) and k-NN(i) are the k-nearest neighbors for the nodes v_j^C and v_i^C, which means that it is also constructed by linking the nodes sparsely to their neighbors in the Lab color space.

We now build an integrated graph that links the image layer graph and the cluster layer graph in the form of a multi-layer graph. That is, the nodes in the two graphs are connected by the inter-edges. The similarity between an image node and a cluster node is similarly defined as eq. (5), which is written as:

$$w_{ij}^{IC} = \exp(-\|\mathbf{x}_i - \mathbf{c}_j\| / \sigma_c), \tag{7}$$

where σ_c is a color control parameter for the inter cluster similarity. In the proposed method, each image node is connected to only one cluster node and thus the (i,j)-th element of the affinity matrix $\mathbf{W}_{IC}$ is represented as:

$$(\mathbf{W}_{IC})_{i,j} = \begin{cases} w_{ij}^{IC}, & \text{if } i \in C_j, \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

where C_j is a set of image layer nodes included in the cluster, v_j^C. Finally, the overall graph representation using an affinity matrix can be expressed as a block-wise matrix form:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_I & \mathbf{W}_{IC} \\ \mathbf{W}_{CI} & \mathbf{W}_C \end{bmatrix}, \tag{9}$$

where $\mathbf{W}_{CI}$ is the transpose of $\mathbf{W}_{IC}$.

3.3. Skin and Background Seed Selection

For the semi-supervised learning [1], it is needed to select the seeds whose labels are known from a simple preceding classification process or from some assumptions and/or prior knowledge. In the proposed method, we select skin seeds as a set of nodes with high iSPM, and the non-skin seeds as a set of nodes adjacent to image boundaries as well as the ones with low iSPM. Specifically, K-means clustering is conducted to make three groups of nodes, based on the features as a set of skin probability, $\{p_i\}_{i=1}^N$ of all nodes. Then, there are three clusters whose centroids represent the skin probability of clusters. We select the nodes belonging to a cluster that is the most skin-probable as the skin seeds, excluding the ones that are located at image boundaries. In the case of non-skin seeds, among the nodes on the image boundary, we select the ones which belong to the cluster which is the least skin-probable (among the three clusters stated above). The skin and the non-skin seeds are notated as $\hat{\mathbf{y}}_S$ and $\hat{\mathbf{y}}_B$ respectively.

3.4. Skin Detection via Semi-supervised Learning

Referring to [1], a matrix for the semi-supervised learning is defined as:

$$\mathbf{W}_L = (1-\alpha)\left(\mathbf{D} - \alpha\mathbf{W}\right)^{-1}, \tag{10}$$

where α is a regularization control parameter, and $\mathbf{D}$ is the degree matrix of $\mathbf{W}$ where each diagonal entry is a row sum of $\mathbf{W}$. As a result, $\mathbf{W}_L$ becomes a fully connected graph even though we connect the edges sparsely in $\mathbf{W}$.

We conduct the propagation based on the graph ($\mathbf{W}_L$) along with both skin seeds ($\mathbf{y}_S$) and background seeds ($\mathbf{y}_B$). It is necessary that the seed vectors are zero-padded to adjust its size, because the graph is composed of image nodes and cluster nodes, while there are only image nodes for the seed vectors, which is given by $\hat{\mathbf{y}}_S$ and $\hat{\mathbf{y}}_B$ respectively. Then, scores of being a skin node and being a non-skin are computed by propagating both seeds to all the nodes which is simply defined as:

$$\hat{\mathbf{f}}_S = \mathbf{W_L}\hat{\mathbf{y}}_S, \tag{11}$$

$$\hat{\mathbf{f}}_B = \mathbf{W_L}\hat{\mathbf{y}}_B, \tag{12}$$

which are interpreted as the overall affinities to the skin and the background seeds respectively. We define the skin probability using these scores as in [19], which is written as:

$$\hat{\mathbf{p}}_S = (\hat{\mathbf{f}}_S - \hat{\mathbf{f}}_B)./(\hat{\mathbf{f}}_S + \hat{\mathbf{f}}_B), \tag{13}$$

where the $./$ is a element-wise division operator. Then, skin probability for the region, v_i^I, which is the i-th element of $\hat{\mathbf{p}}_S$, is notated as $P(v_i^I)$.

3.5. Pixel-wise Refinement

From the above procedures, we get the SPM for the segments, not for the pixels itself. Since the segmentation algorithm is not perfect, the segmented regions sometimes contain the pixels which are much different from the others in the same segment. To alleviate this problem, we devise a region (over-segment) to pixel similarity measure as a pixel-wise refinement step which plays a role in suppressing the small non-skin pixels included in skin regions. It is defined with color distance between a region and a pixel as

$$P(\mathbf{x}_i^p|v_j^I) = \exp\left(-\|\mathbf{x}_i^p - \mathbf{x}_j\|^2 / \sigma_r^2\right), \tag{14}$$

where $\mathbf{x}_i^p$ is a color vector for the i-th pixel and σ_r is a control parameter for the refinement. Pixel-wise skin probability is computed from the above equation and $P(v_j^I)$ as

$$P(\mathbf{x}_i^p) = P(\mathbf{x}_i^p|v_j^I) \cdot P(v_j^I). \tag{15}$$

4. EXPERIMENTS

4.1. Experiment Setup

The number of over-segmented regions is set to 250, and the K used in K-means clustering for cluster nodes is 50, while the number of nearest neighbors used for connecting the cluster layer is set to 5. We set the control parameters for the cluster affinity (σ_c) and refinement (σ_r^2) to 0.05 and 0.1 respectively. As in [1], the regularization control parameter, α, is set to 0.99. The Bayesian skin classifier was trained from $2,000$

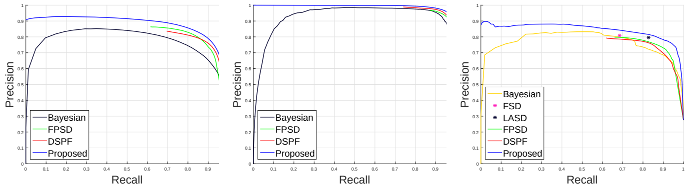

Fig. 2: Comparison of PR curves on ECU, HGR and Pratheepan datasets (from left to right).

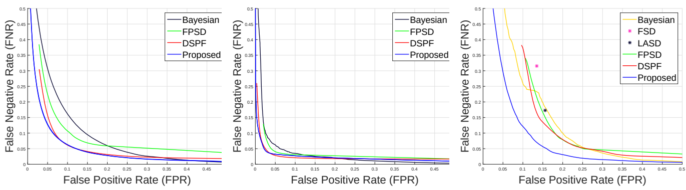

Fig. 3: Comparison of ROC curves on ECU, HGR and Pratheepan datasets (from left to right).

Table 1: Evaluation on ECU and HGR datasets at peak F-measure.

Methods	ECU dataset [21]				HGR dataset [16, 17]			
	Accuracy	*Precision*	*Recall*	*F-measure*	*Accuracy*	*Precision*	*Recall*	*F-measure*
Bayesian [7]	0.8910	0.7292	0.8220	0.7728	0.9598	0.9447	0.9152	0.9297
FPSD [16]	0.9106	0.7948	0.8534	0.8231	0.9610	0.9348	0.9458	0.9403
DSPF [17]	0.9190	0.7713	**0.8864**	0.8249	0.9701	0.9494	0.9437	0.9465
Proposed	**0.9306**	**0.8085**	0.8805	**0.8430**	**0.9735**	**0.9593**	**0.9517**	**0.9555**

images among ECU dataset [21] which consists of 4,000 images. We evaluate the performance of our algorithm on three datasets: ECU [21], HGR [17] and Pratheepan [22]. For the evaluation, we adopt precision-recall curves (PR curves) and receiver operating characteristic curves (ROC curves) for overall performance comparison. In addition, we validate the performance of binarized skin maps with four statistics: *Accuracy*, *Precision*, *Recall* and *F-measure* in a fixed threshold which maximizes the F-measure of each method.

4.2. Comparison with Other Methods

The proposed method is compared with other methods: Bayesian [7], FPSD [16], DSPF [17], FSD [13] and LASD [14]. The overall performance of the compared methods are plotted on PR curves and ROC curves in Fig. 2 and 3 respectively. It shows that our method provides more accurate results in most cases.

Table 2: Evaluation on the Pratheepan dataset [22].

Methods	*Accuracy*	*Precision*	*Recall*	*F-measure*
Bayesian [7]	0.8237	0.6881	0.8972	0.7788
FSD [13]	0.8255	**0.8077**	0.6851	0.7414
LASD [14]	0.8361	0.7954	0.8275	0.8111
FPSD [16]	0.8419	0.7387	0.8991	0.8070
DSPF [17]	0.8521	0.7543	0.8436	0.7964
Proposed	**0.8782**	0.7659	**0.9328**	**0.8412**

Fig. 4: Visual comparison with other methods on the HGR dataset: (from left to right) input, Bayesian, FPSD, DSPF, FSD, LASD, proposed and ground truth images.

In addition, we investigate the quality of binary skin maps by applying a threshold. The threshold for each method is selected such that it maximizes the F-measure, which means that the classification is well-balanced in terms of precision and recall measures. It is shown in Table 1 that our method also shows better performance in most measures on the ECU and the HGR datasets. Our method is also compared with FSD and LASD along with other methods on the Pratheepan dataset as shown in Table. 2. Visual comparison is also presented in Fig. 4, and additional results and binary executables are publicly avaiable at http://ispl.snu.ac.kr/terryoo/skin_detection .

5. CONCLUSION

We have proposed a new skin detection method based on the multi-layer graph representation of an image and multi-seed propagation over the graph. The input image is represented by a two-layer graph, where the nodes in the first layer are the over-segmented regions and the nodes in the second are the clusters of these segments. We design the graph in such a way that the connectivities of nodes represent the color similarity and also the spatial relationship. Experiments show that the proposed method performs better than others even though we do not use texture based seed extraction or facial detectors.

6. ACKNOWLEDGEMENT

This research was supported in part by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the University Information Technology Research Center support program (IITP-2016-R2718-16-0014) supervised by the IITP (Institute for Information & communications Technology Promotion), and in part by the Brain Korea 21 Plus Project in 2016.

7. REFERENCES

[1] Dengyong Zhou, Olivier Bousquet, Thomas Navin Lal, Jason Weston, and Bernhard Schölkopf, "Learning with local and global consistency," in *Advanced in Neural Information Processing Systems*, 2004, pp. 321–328.

[2] Giovani Gomez and E Morales, "Automatic feature construction and a simple rule induction algorithm for skin detection," in *ICML workshop on Machine Learning in Computer Vision*, 2002, pp. 31–38.

[3] Mohammad Abdullah-Al-Wadud, Mohammad Shoyaib, and Oksam Chae, "A skin detection approach based on color distance map," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, pp. 1–10, 2009.

[4] Rein-Lien Hsu, Mohamed Abdel-Mottaleb, and Anil K Jain, "Face detection in color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696–706, 2002.

[5] Ming-Hsuan Yang and Narendra Ahuja, "Gaussian mixture model for human skin color and its applications in image and video databases," in *SPIE: Storage and Retrieval for Image and Video Databases VII*, 1998, vol. 3656, pp. 458–466.

[6] Jae Y Lee and Suk I Yoo, "An elliptical boundary model for skin color detection," in *International Conference on Imaging Science, Systems, and Technology*, 2002.

[7] Michael J Jones and James M Rehg, "Statistical color models with application to skin detection," *International Journal of Computer Vision*, vol. 46, no. 1, pp. 81–96, 2002.

[8] Rehanullah Khan, Allan Hanbury, and Julian Stoettinger, "Skin detection: A random forest approach," in *IEEE International Conference on Image Processing*, 2010, pp. 4613–4616.

[9] Qiang Zhu, Kwang-Ting Cheng, Ching-Tung Wu, and Yi-Leh Wu, "Adaptive learning of an accurate skin-color model," in *IEEE International Conference on Automatic Face and Gesture Recognition*, 2004, pp. 37–42.

[10] Jie Geng, Zhenjiang Miao, and Cencen Zhong, "Skin detection with illumination adaptation in single image," in *IEEE International Conference on Multimedia and Expo*, 2011, pp. 1–6.

[11] Chen-Chiung Hsieh, Dung-Hua Liou, and Meng-Kai Jiang, "Fast enhanced face-based adaptive skin color model," in *International Conference on Image Processing and Pattern Recognition in Industrial Engineering*, 2010, pp. 782027–782027.

[12] Pratheepan Yogarajah, Joan Condell, Kevin Curran, Paul McKevitt, and Abbas Cheddad, "A dynamic threshold approach for skin tone detection in colour images," *International Journal of Biometrics*, vol. 4, no. 1, pp. 38–55, 2011.

[13] Wei Ren Tan, Chee Seng Chan, Pratheepan Yogarajah, and Joan Condell, "A fusion approach for efficient human skin detection," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 138–147, 2012.

[14] Insung Hwang, Sang Hwa Lee, Byungseok Min, and Nam Ik Cho, "Luminance adapted skin color modeling for the robust detection of skin areas," in *IEEE International Conference on Image Processing*, 2013, pp. 2622–2625.

[15] Javier Ruiz-del Solar and Rodrigo Verschae, "Skin detection using neighborhood information," in *IEEE International Conference on Automatic Face and Gesture Recognition*, 2004, pp. 463–468.

[16] Michal Kawulok, "Fast propagation-based skin regions segmentation in color images," in *IEEE International Conference on Workshops Automatic Face and Gesture Recognition*, 2013, pp. 1–7.

[17] Michal Kawulok, Jolanta Kawulok, and Jakub Nalepa, "Spatial-based skin detection using discriminative skin-presence features," *Pattern Recognition Letters*, vol. 41, pp. 3–13, 2014.

[18] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[19] Insung Hwang, Sang Hwa Lee, Jae Sung Park, and Nam Ik Cho, "Saliency detection based on seed propagation in a multilayer graph," *Multimedia Tools and Applications*, pp. 1–19, 2016.

[20] Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan, and Ming-Hsuan Yang, "Saliency detection via graph-based manifold ranking," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3166–3173.

[21] Son Lam Phung, Abdesselam Bouzerdoum, and Douglas Chai, "Skin segmentation using color pixel classification: analysis and comparison," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 148–154, 2005.

[22] Pratheepan Yogarajah, Joan Condell, Kevin Curran, Abbas Cheddad, and Paul McKevitt, "A dynamic threshold approach for skin segmentation in color images," in *IEEE International Conference on Image Processing*, 2010, pp. 2225–2228.