Human Skin Segmentation Improved by Texture Energy Under Superpixels

Anderson Santos(✉) and Helio Pedrini

Institute of Computing, University of Campinas, Campinas, SP 13083-852, Brazil
acarlos@liv.ic.unicamp.br

Abstract. Several applications demand the segmentation of images in skin and non-skin regions, such as face recognition, hand gesture detection, nudity recognition, among others. Human skin detection is still a challenging task and, although color attribute is a very important clue, it usually generates high rate of false positives. This work proposes and analyzes a skin segmentation method improved by texture energy. Experimental results on a challenging public data set demonstrate significant improvement of the proposed skin segmentation method over color-based state-of-the-art approaches.

Keywords: Skin segmentation · Texture energy · Superpixels

1 Introduction

Several applications require the detection of human skin regions in digital images, such as gesture analysis [17], face detection [4,7], nudity detection [12]. Skin detection can be considered as a binary classification problem, where pixels are assigned to belong to skin or non-skin class.

Human skin segmentation is challenging since it is sensitive to camera properties, illumination conditions, individual appearance such as age, gender and ethnicity, among other factors.

In this work, we propose the use of texture energy to reduce the false positives found by color-based methods. Skin color detection is combined with a skin texture probability to generate a final skin probability map. Experiments conducted on large and challenging data set demonstrate that the proposed method is capable of improving the skin color segmentation approaches available in the literature.

The text is organized as follows. Section 2 briefly reviews some works related to skin detection. Section 3 presents the proposed skin segmentation method based on texture probability. Experiments conducted on public data sets are discussed in Section 4. Finally, Section 5 concludes the paper with final remarks and some directions for future work.

2 Background

Research on skin segmentation is vast, such that the most common is to classify the pixel skin color individually in some color space. Fixed rules over one or

A. Pardo and J. Kittler (Eds.): CIARP 2015, LNCS 9423, pp. 35–42, 2015.
DOI: 10.1007/978-3-319-25751-8_5

more color spaces are the simplest form to define a pixel as skin or non-skin. Sobottka et al. [14] limit the skin to a subsection of the HSV color space. Kovac el al. [7] opt for the RGB space, however, with rules concerning the minimum and maximum values of the channels and their differences. Cheddad et al. [2] propose a transformation of RGB color to a 1-dimensional error signal, where simple low and high threshold values define the skin.

Another approach is to fit a parametric model for the distribution of skin and non-skin color. The most common is to use a single Gaussian [16] or a Gaussian Mixture Model [18]. Jones et al. [5] propose to model both skin ($P(c|skin)$) and non-skin ($P(c|\neg skin)$ in order to define the probability of a pixel as skin given its color (c) to be

$$P(skin|c) = \frac{P(c|skin)P(skin)}{P(c|skin)P(skin) + P(c|\neg skin)P(\neg skin)} \tag{1}$$

as stated by the Bayes' rule.

For a Gaussian Mixture Model, the skin or non-skin prior probability is established as

$$P(c|class) = \sum_{i=1}^{N} w_i \frac{1}{(2\pi)^{\frac{3}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(c-\mu_i)^{\tau} \Sigma_i^{-1} (c-\mu_i)} \tag{2}$$

where c is an RGB color vector, *class* can be skin or non-skin and the contribution of the i-th Gaussian is determined by a scalar weight w_i, mean vector μ_i, and diagonal covariance matrix Σ_i.

The model can also be achieved in a non-parametric way by histogram density [5], where the prior probabilities are calculated as

$$P(c|class) = \frac{H_{class}(c)}{\sum_{i=1}^{N} (H_{class}(i))} \tag{3}$$

As expected, the color attribute can individually be ambiguous in skin and non-skin regions, referred to as skin-like regions. Thus, these methods usually achieve high rate of false positives. There are some approaches to improving pixel-based methods by adapting the model to particular characteristics of the image through the analysis of the entire image [7,11] or just a part of it, such as detecting a face [3] or precise skin blobs [13]. Region-based methods, such as texture [9], can be applied as a second step and also spatial analysis in the form of interactive segmentation [6,15].

3 Proposed Method

We propose a method for reducing the rate of false positives in skin detection caused by skin-like color. Law's texture energy measure [8] is employed in the

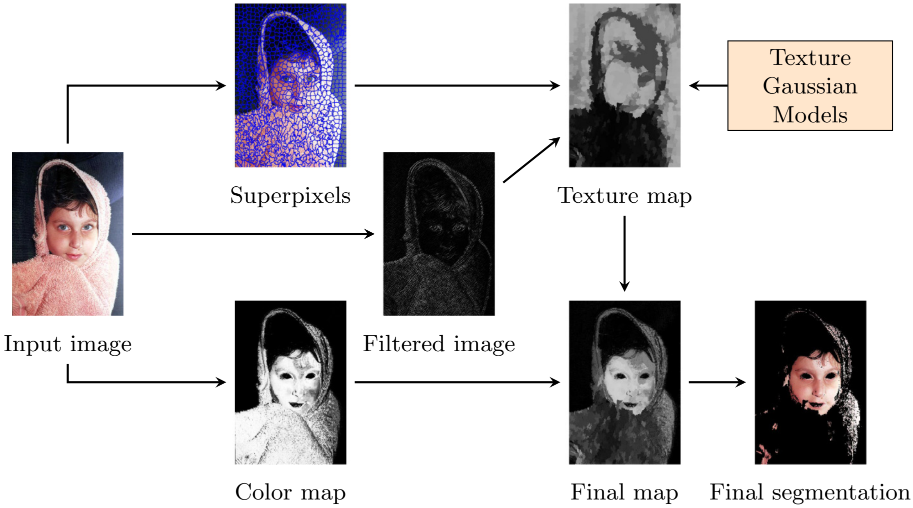

Fig. 1. Main stages of the proposed skin detection method.

process, which works on the response of the intensity image to a special filter mask. The main steps of our skin detection method are illustrated in the flowchart of Figure 1.

The filters defined by Law [8] are built by the product of two vectors obtained from a fixed set of 1D masks designed to detect edges, spots, ripple, among others. A filter is named according to the purpose of the vectors from which it was produced and its size. For example, an E5S5 mask is a 5×5 mask produced by the product of a 1D edge mask and a 1D spot mask. In the following section, we explore the choice of the filter for the proposed method.

To allow the calculation of energy over a region and prevent that the same region covers both skin and non-skin, we use the Simple Linear Iterative Clustering (SLIC) [1] technique for segmenting the image into superpixels (atomic regions, formed by groups of perceptually meaningful pixels). Thus, we calculate the mean energy of each superpixel in the training and test sets.

The goal of the training stage is to obtain two Gaussian models, one for skin and another for non-skin texture energy measures. The images are submitted to superpixels over segmentation and convoluted with a spatial filter. The texture energy is computed for each superpixel, such that mean and standard deviation are extracted for each class (skin and non-skin), forming the two Gaussian models. Algorithm 1 summarizes the training stage.

In the test stage, once the energies of an image have been computed through the same pipeline as in the training step, the skin and non-skin probability densities for each superpixel are obtained. Then, the skin probability given the texture energy is computed in a similar manner to Equation 1, as stated in

Algorithm 1. Proposed Skin Texture Training.

input : List of images L, list of ground-truth pixels G, filter mask f, size of superpixels sp_size.
output: Gaussian Models

$X_{skin} \leftarrow \emptyset$
$X_{\neg skin} \leftarrow \emptyset$
for $I \in L$ **do**
  $SP_{list} \leftarrow SLIC(I, sp\_size)$ /* SP_{list} holds the superpixels coordinates */
  $I_{gray} \leftarrow rgb2gray(I)$
  $I_f \leftarrow I_{gray} * f$ /* where $*$ denotes a convolution */
  for $x \in SPlist$ **do**
    $E_\Phi \leftarrow \frac{\sum I_f(x)^2}{length(SP_{list})}$
    if $I(x) \in G$ **then**
      $X_{skin} \leftarrow \{E_\Phi\}$
    else
      $X_{\neg skin} \leftarrow \{E_\Phi\}$
Compute μ and σ for X_{skin} and $X_{\neg skin}$.
return μ_{skin}, $\mu_{\neg skin}$, σ_{skin}, $\sigma_{\neg skin}$

Equation 4

$$P(skin|E_\Phi) = \frac{f(E_\Phi, \mu_{skin}, \sigma_{skin})}{f(E_\Phi, \mu_{skin}, \sigma_{skin}) + f(E_\Phi, \mu_{\neg skin}, \sigma_{\neg skin})} \tag{4}$$

where E_Φ is the energy measure and $f(E_\Phi, \mu_{class}, \sigma_{class})$ is the Gaussian probability density function for the texture energy.

As texture in a face can vary from the rest of the body, the skin probability in the region close to the nose, around the eyes and mouth will be very low. Thus, it is necessary to apply a heuristic to avoid this type of problem. In our work, we perform a postprocessing mechanism, where areas with low probabilities, surrounded by high probabilities, are filled with the mean of these surroundings high probabilities. Finally, the result of this process constitutes the skin texture probability map.

The texture probability map (T_{map}) is combined with a color probability map (C_{map}) through an AND operation, as shown in Equation 5

$$F_{map} = \sqrt{C_{map} \cdot T_{map}} \tag{5}$$

producing the final skin probability map F_{map}.

The color probability map can be calculated from any color skin detector, even binary output methods that produce only probability 0 or 1. At the final stage, the framework outputs a skin map. Thus, the final segmentation can be performed by a simple threshold or a more sophisticated strategy.

4 Experiments

Experiments were conducted on two different data sets to evaluate the proposed methodology. For training, we used 8963 non-skin images and 4666 skin images from the Compaq database [5], which contains images acquired from the Internet in a diverse variety of settings. For evaluation and comparison purposes, we used the ECU database [10] that was divided into 1000 images for validation and 3000 images for test. This ensures a diversity in terms of background scenes, lighting conditions, and skin types. Both data sets contain a manually labeled ground-truth.

The performance of the skin detection method was assessed through a number of metrics: true positive rate (η_{tp} - percentage of skin correctly classified as skin); false positive rate (δ_{fp} - percentage of non-skin classified as skin); F_{score} (harmonic mean between η_{prec} and η_{tp}) and detection error ($\delta_{min} = (1 - \eta_{tp}) + \delta_{fp}$). Additionally, ROC (receiver operating characteristics) curves are computed.

In order to select the filter, we used four 1D vectors:

$$\begin{array}{lll} \text{L5} & \text{(Level)} & = [\ -1 \quad 4 \quad 6 \quad 4 \quad 1\ ] \\ \text{E5} & \text{(Edge)} & = [\ -1 \quad -2 \quad 0 \quad 2 \quad 1\ ] \\ \text{S5} & \text{(Spot)} & = [\ -1 \quad 0 \quad 2 \quad 0 \quad 1\ ] \\ \text{R5} & \text{(Ripple)} & = [\ 1 \quad -4 \quad 6 \quad -4 \quad 1\ ] \end{array}$$

which generates sixteen 5×5 filters. Each one is convolved with the image; some filters will be just transposed of others, so they are combined to produce only one filter with their mean, resulting in nine features.

From the validation set, we evaluate the gain provided by each feature in relation to the color individually. The filter E5S5/S5E5 produced the best results.

In order to evaluate the proposed method, we selected three widely used skin detectors with different approaches: Cheddad's rule [2] (rule based), Histogram Model [5] (non-parametric) and Gaussian Mixture Model (GMM) [5] (parametric). The Histogram Model was built with 64 bins per channel in the RGB space. For the Gaussian Mixture, we used the 16 kernels trained in the original paper with the same database as used here.

Figure 2 shows comparative ROC curves between the original skin detector and our improvement. It is possible to observe that the proposed method always achieves superior results.

Table 1 shows the result values when considering the closest point to the optimum point (0, 100%) in the ROC curve. For Cheddad's rule, which is a binary method, the tables present isolated point values.

For a more detailed comparison, we provide true positive rate values for a 10% false positive rate in Table 2. In other words, this represents how much of true skin is possible to detect since there is only 10% tolerance for skin-like. In case of the original Cheddad method, we perform a linear approximation preserving the same ratio between η_{tp} and δ_{fp}.

It is worth mentioning that our method always results in higher true positive rates with a considerable advantage over the original approaches.

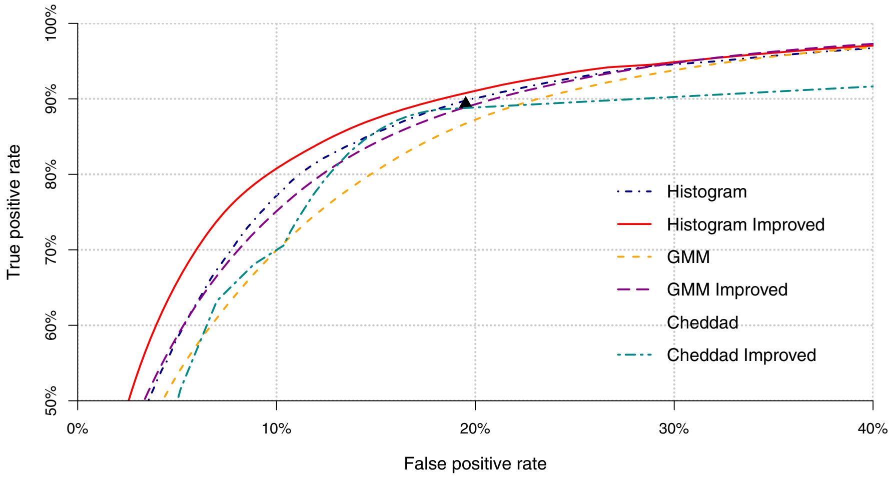

Fig. 2. ROC curve illustrating the results on test data set for the original method and the improvement through our method.

Table 1. Detection results for different methods (ECU data set).

Method	Original				Improved			
	η_{tp} (%)	δ_{fp} (%)	F_{score} (%)	δ_{min} (%)	η_{tp} (%)	δ_{fp} (%)	F_{score} (%)	δ_{min} (%)
Cheddad	89.33	19.51	64.78	30.18	87.32	16.22	67.38	28.90
Gaussian Mixture	87.55	20.30	63.09	32.76	87.37	17.78	65.64	30.41
Histogram Model	87.21	16.54	66.95	29.33	86.96	14.55	**69.17**	**27.59**

Table 2. True positive rates for a fixed value of false positive rate (ECU data set).

Method	$\eta_{tp}(\%)$, $\delta_{fp} = 10\%$	
	Original	Improved
Cheddad	46	71
Gaussian Mixture	70	75
Histogram Model	77	**81**

Figure 3 shows some examples of final segmentation in the tested data set. The first column presents the original image, the second one shows the ground-truth, whereas the remaining columns show the segmentation result for each original method on the left and its corresponding improved segmentation results on the right.

Fig. 3. Examples of skin regions detected through different methods.

5 Conclusions

This work describes a new technique for reducing the high number of false positives in color-based skin segmentation. It can be applied in conjunction with any skin segmentation method, while not adding more sensitive parameters to perform the skin classification.

Experiments conducted on a well-known large test set demonstrated that the proposed technique can provide a significant improvement on the results obtained with different color-based skin detection approaches. Furthermore, the use of texture for skin segmentation is very difficult; although human skin has a distinguishable texture, this can only be noticeable on high resolution images. Age, amount of body hair, expression and pose constitute obstacles in finding a structural pattern for skin.

As directions for future work, we intend to expand the method through multi-resolution analysis to take the variety in terms of image quality and size into account. We also plan to propose a better filter, designed specifically for the skin detection problem.

Acknowledgments. The authors are grateful to FAPESP, CNPq and CAPES for their financial support.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. IEEE Transactions on Pattern Analysis and Machine Intelligence **34**(11), 2274–2282 (2012)
2. Cheddad, A., Condell, J., Curran, K., Mc Kevitt, P.: A Skin Tone Detection Algorithm for an Adaptive Approach to Steganography. Signal Processing **89**(12), 2465–2478 (2009)

3. Fritsch, J., Lang, S., Kleinehagenbrock, M., Fink, G.A., Sagerer, G.: Improving adaptive skin color segmentation by incorporating results from face detection. In: 11th IEEE International Workshop on Robot and Human Interactive Communication, pp. 337–343 (2002)
4. Hsu, R.L., Abdel-Mottaleb, M., Jain, A.K.: Face Detection in Color Images. IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(5), 696–706 (2002)
5. Jones, M.J., Rehg, J.M.: Statistical Color Models with Application to Skin Detection. International Journal of Computer Vision **46**(1), 81–96 (2002)
6. Kawulok, M.: Fast propagation-based skin regions segmentation in color images. In: 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, pp. 1–7 (2013)
7. Kovac, J., Peer, P., Solina, F.: Human skin color clustering for face detection. In: International Conference on Computer as a Tool (Eurocon), vol. 2, pp. 144–148. IEEE, September 2003
8. Laws, K.I.: Rapid texture identification. In: 24th Annual Technical Symposium, pp. 376–381. International Society for Optics and Photonics (1980)
9. Ng, P., Pun, C.M.: Skin color segmentation by texture feature extraction and k-means clustering. In: Third International Conference on Computational Intelligence, Communication Systems and Networks, pp. 213–218. IEEE (2011)
10. Phung, S.L., Bouzerdoum, A., Chai Sr, D.: Skin Segmentation using Color Pixel Classification: Analysis and Comparison. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(1), 148–154 (2005)
11. Phung, S.L., Chai, D., Bouzerdoum, A.: Adaptive skin segmentation in color images. In: International Conference on Multimedia and Expo., vol. 3, pp. III–173 (2003)
12. Platzer, C., Stuetz, M., Lindorfer, M.: Skin sheriff: a machine learning solution for detecting explicit images. In: 2nd International Workshop on Security and Forensics in Communication Systems, pp. 45–56. ACM, New York (2014)
13. Santos, A., Pedrini, H.: A self-adaptation method for human skin segmentation based on seed growing. In: 10th International Conference on Computer Vision Theory and Applications, Berlin, Germany, pp. 455–462, March 2015
14. Sobottka, K., Pitas, I.: A Novel Method for Automatic Face Segmentation, Facial Feature Sxtraction and Tracking. Signal Processing: Image Communication **12**(3), 263–281 (1998)
15. Ruiz-del Solar, J., Verschae, R.: Skin detection using neighborhood information. In: Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 463–468. IEEE (2004)
16. Subban, R., Mishra, R.: Human skin segmentation in color images using Gaussian color model. In: Thampi, S.M., Abraham, A., Pal, S.K., Rodriguez, J.M.C. (eds.) Recent Advances in Intelligent Informatics. AISC, vol. 235, pp. 13–21. Springer, Heidelberg (2014)
17. Wachs, J.P., Kölsch, M., Stern, H., Edan, Y.: Vision-based hand-gesture applications. Communications of the ACM **54**(2), 60–71 (2011)
18. Yang, M.H., Ahuja, N.: Gaussian mixture model for human skin color and its application in image and video databases. In: SPIE: Storage and Retrieval for Image and Video Databases VII, vol. 3656, pp. 458–466 (1999)