Improved skin segmentation for TV image enhancement, using color and texture features

Citation for published version (APA):

Zafarifar, B., Martiniere, A., & With, de, P. H. N. (2010). Improved skin segmentation for TV image enhancement, using color and texture features. In *Proceedings International Conference on Consumer Electronics (ICCE2010)*, 9-13 January 2010, Las Vegas, Nevada (pp. 373-374). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/ICCE.2010.5418755>

DOI:

[10.1109/ICCE.2010.5418755](https://doi.org/10.1109/ICCE.2010.5418755)

Document status and date:

Published: 01/01/2010

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Improved Skin Segmentation for TV Image Enhancement, Using Color and Texture Features

Bahman Zafarifar^{1,2}, Anthony Martinière¹ and Peter H. N. de With^{2,3}, *IEEE Fellow*
¹NXP Semiconductors, ²Eindhoven University of Technology, ³CycloMedia Technology

Abstract—Maintaining natural appearance for human face and skin areas is essential in TV image enhancement. High-end TV video processing chips include skin detectors for this purpose. This paper presents a technique for improved skin detection, using a trainable color feature and a multi-scale texture feature. Evaluation on a database of 173 manually annotated natural images shows that at a true positive rate of 90%, the proposed technique yields a false positive rate of 12%, as compared to 19% and 20% of two color-only skin detectors.

I. INTRODUCTION

Image restoration and image enhancement are standard features in current digital TV sets. Restoration functions such as de-interlacing and noise/artifact reduction aim at regaining lost or corrupted image information. Enhancement functions such as sharpness, contrast and color boosting aim at creating a more pleasant image. Applying these functions to human skin areas may result to unnatural appearance or objectionable artifacts. For example, a TV post-processing chip [3] that we use for case study includes three skin detectors for mitigating the level of contrast, sharpness and saturation boosting, and for skin color enhancement.

Currently, these skin detectors are based on color information only, and judge all skin-colored image areas as skin (false positives), leading to poor enhancement and false color modification in non-skin skin-colored areas. We aim at reducing these false positives, primarily by using a *multi-scale texture analysis*. The current skin detector builds upon our previously developed algorithms for segmentation of sky and grass areas [1][2], and extends it by a flexible histogram-based color detector, and specific processing for improved detection around facial features. We compare the results against two color-based skin detectors, including the best detector available in the mentioned hardware [3].

In the following, Section II describes the algorithm and Section III presents the results and conclusions.

II. SEGMENTATION METHOD

A. Color feature

Skin detectors of existing TV video-enhancement systems (e.g. [3]) are typically based on a simple function defined in a 3D color space (e.g. YUV or HSV). The properties (color center, orientation, size and decrease slope) of these functions can be defined by the user. Such a detector is limited in that, first, it is not fully adaptive to all desired skin tones, and second, determining appropriate parameters involves

This research was sponsored by NWO Casimir program of the Dutch government.

subjective visual assessment of the detection result on multiple test images, which is a manual and non-reproducible process. To address these shortcomings, we employ a flexible and automatically tunable color segmentation, based on the color histogram of skin pixels computed on a set of annotated training images.

A Bayesian skin color classifier based on non-parametric density estimation of skin and non-skin classes is defined by

$$p_{color}(x) = H_{skin}(x) / H_{non-skin}(x) \quad (1)$$

where H_{skin} and $H_{non-skin}$ stand for histogram (i.e. an estimation of class-conditional probability) of skin and non-skin classes, respectively. It is shown in [4][5] that this classifier outperforms many other classifiers based on parametric probability density functions (e.g. Gaussian, piece-wise linear), and classification methods.

We have validated the performance of the Bayesian color classifier in an experiment based on a database of 100 manually annotated skin images captured from several TV sources, containing different skin tones. Using YUV space for the feature vector and a 10-fold cross-validation, our experiment has confirmed the higher performance of the Bayesian method compared to Mixture of 2 Gaussian models. The same holds also when it is compared to the best detector of [3], which is based on 3D function in HSV color space. Fig. 1-right compares the ROC curves of these three skin detectors.

Fig. 1. Left: overview of a Bayesian skin color classification system. Right: Bayesian color feature outperforms other color classification methods.

Despite the higher statistical performance of the Bayesian classifier, we have found that division by estimated class-conditional probability of the non-skin class, $H_{non-skin}(x)$ in (1), causes large variations in the skin segmentation map. This is due to the fact that reliable estimation of the 3-dimensional probability of the non-skin class requires a huge number of annotated training samples. For this reason, we assume uniform probability for the non-skin class. Since the number of training samples is limited, we apply low-pass filtering on the histogram of the skin class in order to increase generalization, and store the filtered result in 32^3 bins.

B. Texture feature

In typical camera zoom-level of TV images, skin can be characterized by its smooth appearance, with slow gradients due to changing light angle. However, reflections and shadows cause sudden changes in skin color components, especially in the luminance channel. Based on this observation, we define the texture feature P_{text} , which yields high values on smooth areas, and is inversely related to

$$T(i, j) = \left[|Y(i, j-1) - 2Y(i, j) + Y(i, j+1)| + |Y(i-1, j) - 2Y(i, j) + Y(i+1, j)| - t_n \right]_0^l, \quad (2)$$

where $Y(i, j)$ is the luminance component at pixel position (i, j) , t_n is a noise threshold and $[x]_0^l$ clips the value of x between 0 and l . This feature gives low response on slow gradients (0 on linear ramps), and clipping to l limits the contribution of the aforementioned sudden luminance changes.

Fig. 2 A multi-scale analysis of color and texture features computes an initial skin segmentation map.

C. Combining features, and segmentation

The color and the texture features are computed at two image scales and combined according to Fig. 2 to produce the *initial skin segmentation*.

Fig. 3 A low-resolution position model is computed by filling low initial skin-segmentation values at facial features, followed by low-pass filtering.

The texture feature causes coarse variations in the initial skin segmentation. Therefore, we compute a smooth low-resolution (45·36 in case of an SD input) version of the initial skin segmentation, denoted by *position model* in Fig. 3. In the position model, low initial skin segmentation values at the location of facial features such as mouth, nose and eyes are filled with the higher skin probabilities in their surrounding (compare the images in Fig. 3-left and middle). This is performed by the downscaling function of Fig. 3, where the maximum of pixel values in a window of 4·4 pixels is assigned to the downsampled output pixel. The result is then filtered by a Gaussian filter to increase spatial coherence.

Fig. 4 Our final skin segmentation combines position and color features at input resolution.

The final skin segmentation map is now computed by linearly upscaling the position model to the input image resolution, and by multiplying this upscaled version with a color probability computed at input image resolution (Fig. 4).

Fig. 5 The proposed method outperforms two color-only skin detectors.

III. RESULTS AND CONCLUSION

We have evaluated the performance of the proposed skin detector versus an existing color detector [3] defined in HSV color space and versus a histogram-based color detector, using a set of 173 annotated natural images containing various skin types and non-skin skin-colored images. The ROC curves in Fig. 5 show that the histogram-based color detector has a similar performance to the existing detector of [3], while the proposed technique yields lower false positive rates at medium to high true-positive rates. When skin detectors of [3] are replaced by the proposed technique, we observe clear improvement in image enhancement results, especially in sharpness perception of detailed areas. The method has limited performance on small faces and heavily textured skin areas (e.g. elderly people). A practical system may therefore require additional measures to address this limitation and to enforce temporal coherence of the skin segmentation result.

- [1] Bahman Zafarifar and Peter H. N. de With, "Blue Sky Detection for Content-based Television Picture Quality Enhancement", *IEEE Intern. Conference on Consumer Electronics*, January 2007, pp. 437-438.
- [2] Bahman Zafarifar and Peter H. N. de With, "Grass Field Detection for TV Picture Quality Enhancement", *IEEE Intern. Conference on Consumer Electronics*, January 2008, pp. 329-330.
- [3] PNX5100 Nexperia video back-end processor, www.nxp.com.
- [4] Phung S.L, Bouzerdoum A. and Chai D., "Skin Segmentation Using Color and Edge Information", *Proc. Int. Symposium on Signal Processing and its Applications*, July 2003, pp. 525- 528.
- [5] Michael J. Jones and James M. Rehg, "Statistical Color Models with Application to Skin Detection", *Int. Journal of Computer Vision*, Vol. 46, Number 1 / Jan. 2002.