Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Pattern Recognition ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

Rapid and brief communication

A recall or precision oriented skin classifier using binary combining strategies

Francesca Gasparini*, Silvia Corchs, Raimondo Schettini

DISCO (Dipartimento di Informatica, Sistemistica e Comunicazione), Università degli Studi di Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy

Received 12 April 2005; accepted 18 April 2005

Abstract

Skin detection is a preliminary step in several applications, and many different methods are available in the literature. We show that the performance of explicit skin cluster classifiers can be enhanced by preprocessing the images with a white balance algorithm. Different combining strategies are then applied to these binary classifiers to further improve their performance in terms of recall and/or precision. Experimental results on a large and heterogeneous image database are presented.

Keywords: Skin color classifier; Combining; White balancing

1. Introduction

The detection of skin regions in color images is a preliminary step in several applications of pattern recognition, and many different methods for discriminating between skin pixels and non-skin pixels are available in the literature. The simplest, and often applied, is to build what is called an “explicit skin cluster” classifier which expressly defines the boundaries of the skin cluster in certain color spaces [1–4], through simple decision rules. This type of binary method is very popular as it is easy to implement and does not require a training phase. Skin detection, just like any other color-based feature computation, may not, however, be completely reliable, especially when the images to process are collected from many different sources (the web, for example). In fact, acquisition conditions and imaging devices not known a priori, or not carefully controlled, can introduce significant color distortion. To alleviate this problem we preprocess the images with a white balance algorithm.

To quantify the performance of these skin detection methods we use recall and precision scores. Classification results are assigned as true positive (TP), false positive (FP) and false negative (FN). Recall is defined as the ratio between the number of skin pixels correctly classified and the total number of actual skin pixels ($TP/(TP+FN)$), while precision is defined as the ratio between the number of skin pixels correctly classified and the total number of pixels labeled as skin pixels by the skin detection method considered ($TP/(TP+FP)$). It is tempting to believe that a good classifier should have high recall and high precision, but typically, as recall increases, precision decreases. Consequently, the classifier will be chosen to offer high recall or high precision according to application demands. Keeping this in mind, we propose different combining strategies for binary classifiers to improve the results in terms of recall or precision.

The effectiveness of these strategies is evaluated on a large and heterogeneous image database.

* Corresponding author. Tel.: +39 02 64487856; fax: +39 02 64487839.
E-mail addresses: gasparini@disco.unimib.it (F. Gasparini), schettini@disco.unimib.it (R. Schettini).

0031-3203/$30.00 © 2005 Published by Elsevier Ltd on behalf of Pattern Recognition Society.
doi:10.1016/j.patcog.2005.04.007

2. Method

We have implemented several skin classifiers which label skin and non-skin pixels using a piecewise linear decision boundaries. We consider here only the five of these algorithms that have proved more precise, (C_j), $j = 1, \ldots, 5$; they are referred to by the name of the color space adopted: $C_1 =$ HSI [1], $C_2 =$ RGB [2], $C_3 =$ YCbCr [3], $C_4 =$ HSV1 [4] and $C_5 =$ HSV2 [4]. The details of their implementation can be found in the referenced papers.

Four methods of color balancing were tested: gray world (GW), white point (WPt), white patch (WP), and our self-tunable color balancing (STCB) [5]. The gray world algorithm assumes that, given an image of sufficiently varied colors, the average surface color in a scene is gray. The white point and white patch algorithms assume that there is always a white point, or white region respectively, in the scene. The white point algorithm determines the reference white as the maximum R, maximum G and maximum B found in the image, while the white patch takes as reference white the average of the region with the higher value of luminance (here the top 5 percent of the luminance range). Our self-tunable color balancing is based on preliminary image statistics for color distribution in the CIELAB color space. It can be considered a weighted mixture of the white patch and gray world procedures, which permits the solution of cases where one, or both of the two assumptions are not valid.

The combination of classifiers has been widely used to improve the performance achieved by a single classifier. Among the possible variations of this idea, in the case of binary classifiers, we have considered the sum rule (Eq. (1)), the product rule (unanimity vote, Eq. (2)) and the majority vote (Eq. (3)):

$$C_{sum} = \bigcup_{i=1}^{5} C_i, \tag{1}$$

$$C_{product} = \bigcap_{i=1}^{5} C_i, \tag{2}$$

$$C_{majority} = \left(\sum_{i=1}^{5} C_i \right) \geqslant 3. \tag{3}$$

We also propose here a fourth combination rule, C_{SCNS} (skin corrected by a non-skin), exploiting the fact that binary classifiers produce both a skin and a non-skin map. The final skin map (Eq. (4)) is obtained from a preliminary skin map (produced here by C_1, the algorithm with the greatest recall) corrected by a non-skin map, NS, obtained as the intersection of the non-skin maps of all the remaining algorithms (Eq. (5)):

$$C_{SCNS} = C_1 \cap (1 - \mathrm{NS}), \tag{4}$$

$$\mathrm{NS} = \bigcap_{j=2}^{4} (1 - C_j). \tag{5}$$

3. Experimental results

All the experiments reported here have been performed on 2000 images taken from the Compaq skin database containing 22.669.739 skin pixels and 149.119.846 non-skin pixels.

The performance of the five algorithms in terms of recall versus precision is presented in Fig. 1a, while in the sequence of Figs. 1b–f, the performance of each algorithm before and after the application of the four white balance methods is shown. We can see that the different methods may be either more precision- or more recall-oriented (Fig. 1a). The qualification of "best method" is therefore application-dependent. The performance of the five methods considered always improves in terms of precision after the application of any of the white balance algorithms, with the exception of HSV1 with GW (Fig. 1e). In fact, GW does not seem to be an efficient white balance algorithm; as an increase in precision does not compensate the loss in terms of recall. STCB, instead, always gives the best performance, increasing precision and preserving recall.

The performance of the combining strategies in terms of recall versus precision is presented in Fig. 2. In Fig. 2a, their results are compared with those of the original five skin classifiers. The unanimity rule is the most precision-oriented, but presents a significantly reduced recall. The majority vote rule achieves a high recall with significantly high precision, showing a good trade-off between correct and incorrect classifications. The sum rule is recall-oriented, but shows a significant reduction in precision. The SCNS method is also recall-oriented, but obtains a slight increase in precision compared with the single classifiers with similar recall. Preserving the true positives was in fact our objective in choosing the method with the highest recall to generate the skin map to be corrected by the non-skin map.

In Fig. 2b, the results of the combining strategies are compared, before and after white balancing with the STCB method. As in the case of the original skin classifiers (Fig. 1b–f), the performance of all the combining rules increases in terms of precision when images are preprocessed with STCB.

All these experimental results show that the application of a reliable and efficient white balance algorithm, such as STCB, together with a combining strategy, significantly enhances the overall performance of skin detection.

4. Conclusions

We have evaluated the performance of various skin classifiers preprocessed with different white balance algorithms, and investigated different combinations of these classifiers.

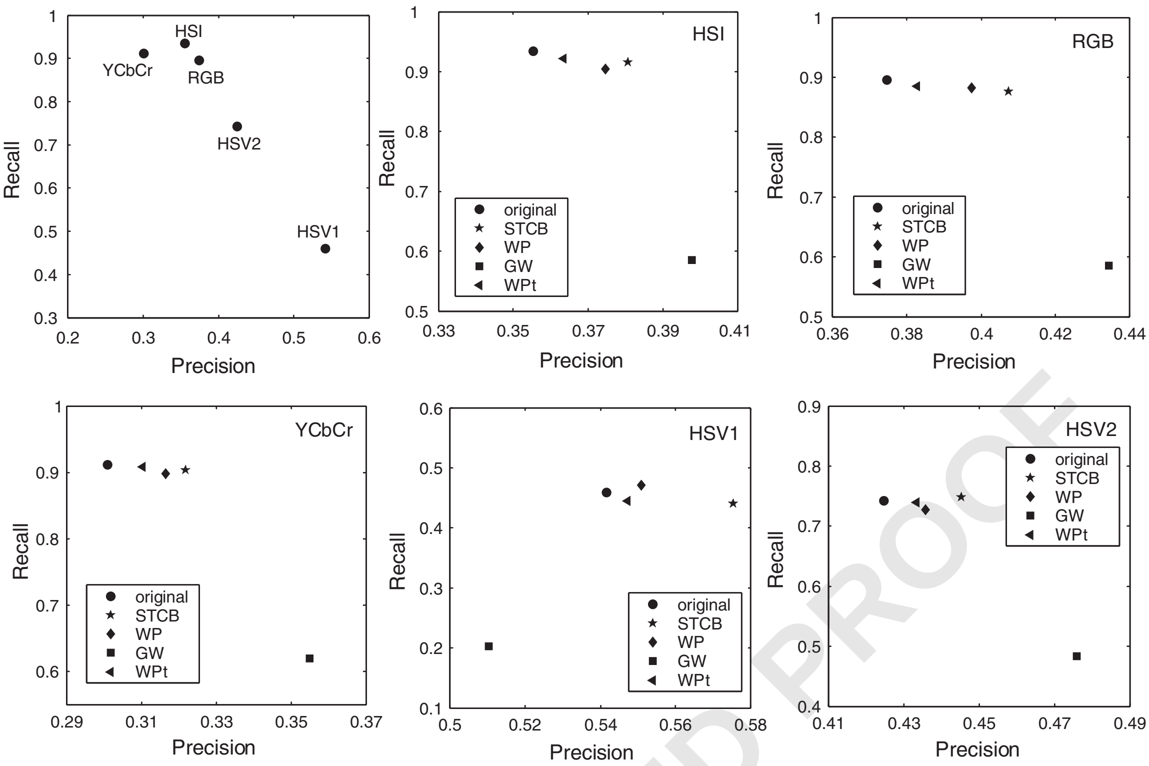

Fig. 1. Comparison of the performance (recall versus precision) of the five skin classifiers (a) and of each original classifier before and after the application of the four white balance methods (b–f). Note that the axes of (b)–(f) are scaled to guarantee equal metrics for cross comparisons.

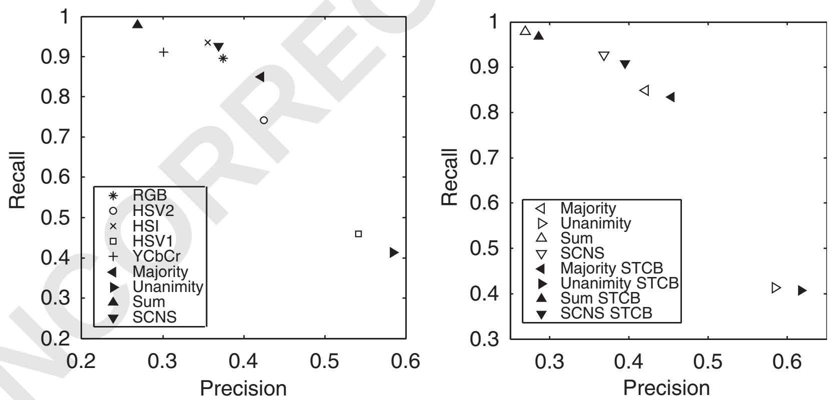

Fig. 2. (a) Comparison of the performance (recall versus precision) of the combining strategies with respect to the original five skin classifiers. (b) Comparison of the performance of all the combining strategies before (blank triangles) and after white balancing with the STCB method (black triangles).

Since recall and precision show a complementary behavior, as in the single methods, the choice of combination rule, in terms of recall or precision, is application-dependent.

We plan in the future to evaluate new combining strategies for non-binary skin classifiers, such as parametric or histogram-based models.

References

[1] I.-S. Hsieh, K.-C. Fan, C. Lin, A statistic approach to the detection of human faces in colour nature scene, Pattern Recogn. 35 (2002) 1583–1596.

[2] J. Kovac, P. Peer, F. Solina, 2D versus 3D colour space face detection, 4th EURASIP Conference on Video/Image Processing and Multimedia Communications, Croatia, 2003, pp. 449–454.

[3] D. Chai, K.N. Ngan, Face segmentation using skin colour map in videophone applications, IEEE Trans. Circuits Syst. Video Technol. 9 (4) (1999) 551–564.

[4] C. Garcia, G. Tziritas, Face detection using quantized skin colour regions merging and wavelet packet analysis, IEEE Trans. Multimedia 1 (1999) 264–277.

[5] F. Gasparini, R. Schettini, Colour balancing of digital photos using simple image statistics, Pattern Recogn. 37 (2004) 1201–1217.