People detection in still images based on a skin filter and body part evidence*

Claudio Cavalcanti and Herman Gomes
Universidade Federal de Campina Grande, Departamento de Sistemas e Computação
{csvcc, hmg}@dsc.ufcg.edu.br

Abstract

This paper presents an approach for determining the presence of people in a still image. The approach starts with the segmentation of skin areas in the images using the YCbCr color space and a set of experimentally defined thresholds. The detected skin areas are then grouped into body part candidate regions. Regions that are too small (typically smaller than 5x5) are discarded. A simple heuristic rule, based on the proportions and pixel density of these regions, is then applied to map these regions into body part evidence. The application of this approach to a database of still photographs resulted in the correct classification of 77.2% of the images.

1. Introduction

In this paper, we propose an approach for finding image regions that have a high likelihood of belonging to a person by using some properties of the human skin tone and heuristic rules to refine these regions. Since the human skin tone does not vary sharply, even for people of different races and from different geographic regions, it is possible to find, in a given color space, small intervals that may represent the color of the natural human skin and other intervals with very low likelihood of being human (e.g. blue, green, cyan etc). After finding the regions which may represent skin tone, the next step is to filter out false candidates, such as background objects whose tone interval may be similar to the human one. Jones and Rehg [1] present an analysis on the use of color spaces for skin detection. Figure 1 illustrates the main steps of the proposed approach. This is an initial stage of the work.

Figure 1. *Steps of the proposed approach (skin areas are painted white): top left, original image; top right, output of the skin filter; bottom, body part candidates found*

The main objective was to prove that body proportions and skin colors can be used together to produce fast and approximate results.

2. Color Spaces

Zarit et al [2] presented a performance evaluation of five of the most used color spaces for skin filter design. Overall, their results have shown that the worst and best combination of detection technique and color space had a recognition rate varying from around 70% to 80%.

Other recent works compare the application of skin color segmentation algorithms. Butler et al [3] used Gaussian Mixture Models based on 8x8 blocks. Ruiz-del-Solar et al [4] uses neighborhood information for classifying skin tones and Jayaram et al [5] analyze which combination of several colorspace transformation algorithms are the best ones for skin color detection.

In the present work, we chose to use the YCbCr color space because it favors the identification of skin tone intervals. The consequence is that some comparisons can be avoided. We experimentally determined intervals in the YCbCr color space in such a way that for a number of test images containing people the segmented skin regions were maximized. The resulting intervals are shown on Table 1.

Table 1. Experimentally determined skin intervals in the YCbCr space.

Cb		Cr	
Min	Max	Min	Max
205	250	20	140

It is not possible to completely isolate skin color using information from color space segmentation alone. If the intervals are too restrictive, many skin tones would not be detected, whereas if they are too wide, the chances of false positives dramatically increase. In order to solve this problem, Forsyth and Fleck [6] also uses geometrical features to complement the skin filter.

This work proposes making very simple further processing to verify if the pixels found are skin related or not. The main idea is to consider only the dimensions and density of candidate skin regions. This might be imprecise and lead to some errors, but it is fast to compute.

The algorithm for selecting the candidate skin regions for body parts, considers the set of pixels found as a skin tone. It scans the image starting from its origin and defines boxes that grow around each skin tone found. The step of choosing the correct candidates obey to pre-defined rules which are divided in two categories: dimension and the color density. The dimension score will map the dimension of a candidate region to the chance of being human. It is based on the face proportions. Width / height ratios between 0.6 and 1.0 gives a high score. Medium ratios (1.0 to 2.0) scores less than the first one. Ratios higher than 2.0 do not usually represent standing people, so a negative score is assigned. The density score considers the amount of skin tone pixels within a candidate. A typical human body part typically occupies at least 25% of a region. A score is then assigned and if it is higher than 35% it is doubled. These scores are not, by their own, definitive, but influences on the final combined score. All the candidates are evaluated and the final results are then passed through a threshold in order to make the final decision.

All the scores and thresholds presented above have been manually chosen, based on the subjective evaluation of a set of test images.

3. Experiments and Results

Three different experiments are presented. First two experiments were performed on a set of 167 uncontrolled (different illumination, capture quality, distance to the camera, etc.) images. Each of the first 83 images contains an arbitrary number of people. However, we considered only pictures with people that were standing up. The other 84 images contained no people. We also deliberately included pictures with objects that presented skin-like tones (such as a red car, a dog, etc.). The second experiment improves on the first one by using the YCbCr filter and the body part evidence. Results are presented on Table 2.

Table 2. YCbCr Skin filter plus body part evidence performance. PWP – Photos With People; PWOP – Photos Without People.

	Correct	Incorrect	Correct (%)
PWP	72	11	86.7%
PWOP	57	27	67.8%
Total	129	38	77.2%

The third experiment uses the same algorithm of experiment 2 on a new larger set of uncontrolled images (155 images). It calculates the detection rate of the proposed approach. The classification rate of the skin correctly detected is higher than 90% (Table 3).

Table 3. YCbCr Skin filter plus body part evidence performance on a people-containing set of images. PWP – Photos With People.

	Correct	Incorrect	Correct(%)
PWP	141	14	90.9%

4. Conclusions

The use of thresholding in different color spaces to verify the human presence in images is not sufficiently accurate. Higher-level analysis is needed for achieving higher detection performance. The simplicity and associated low processing costs, however, makes this approach useful as a preliminary discriminating tool. As future work, we plan to compare our results with the approaches discussed in Section 2. We believe that the color space influence in the human detection will play a secondary role when compared to the higher level post-processing algorithms, such as the body part evidence proposed in this paper.

References

[1] M. J. Jones and J. M. Rehg, “Statistical Color Models with Application to Skin Detection”. Int. J. Computer Vision 46(1), pp. 81-96, 2002.

[2] B. D. Zarit, B. J. Super and F. K. H. Quek, “Comparison of Five Color Models in Skin Pixel Classification”. Proc. Int. Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 1999.

[3] D. Butler, S. Sridharan, V. Chandran, “Chromatic colour spaces for skin detection using GMMs”. Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing, pp: 3620-3623, v.4, 2002.

[4] J. Ruiz-del-Solar, R. Verschae, “Skin detection using neighborhood information”. Proc. of 6a IEEE Int. Conf. on Automatic Face and Gesture Recog., pp.463–468, 2004.

[5] S. Jayaram, S. Schmugge, M.C. Shin, L.V. Tsap, “Effect of colorspace transformation, the illuminance component, and color modeling on skin detection”. Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recog., pp: 813-818 V.2, 2004.

[6] D. Forsyth and M. Fleck, “Automatic Detection of Human Nudes”. Int. J. of Computer Vision 32/1, pp: 63-77, 1999.