2001

Skin color detection for face localization in human-machine communications

Douglas Chai
Edith Cowan University

Son Lam Phung
Edith Cowan University

Abdesselam Bouzerdoum
Edith Cowan University

10.1109/ISSPA.2001.949848
This is an Author's Accepted Manuscript of: Chai, D. K., Phung, S. L., & Bouzerdoum, A. (2001). Skin color detection for face localization in human-machine communications. Proceedings of 6th Internatrional Symposium on Signal Processing and its Applications. (pp. 343 - 346 vol.1). Malaysia. IEEE. Available here

International Symposium on Signal Processing and its Applications (ISSPA), Kuala Lumpur, Malaysia, 13 – 16 August, 2001. Organized by the Dept. of Microelectronics and Computer Engineering, UTM, Malaysia and Signal Processing Research Centre, QUT, Australia

SKIN COLOR DETECTION FOR FACE LOCALIZATION IN HUMAN-MACHINE COMMUNICATIONS

Douglas Chai, Son Lam Phung, Abdesselam Bouzerdoum

Edith Cowan University
School of Engineering and Mathematics
100 Joondalup Drive, Joondalup 6027, Perth, Australia
Email: {d.chai, s.phung, a.bouzerdoum}@ecu.edu.au

ABSTRACT

This paper presents the proposed user interface design for computers whereby users can navigate in 3D graphics scene and change camera viewpoint via head movement. This human-machine communication relies very much on the performance of its face localization module, which must determine head pose and track head movement. We have employed the skin color detection approach to face localization. The approach is studied and presented in this paper. The experimental results show that our chosen methodology is very effective. Furthermore, we demonstrate that skin color detection approach can cope of the variations of skin color and lighting condition.

1. INTRODUCTION

As computers become more ubiquitous in society, there is a desire for more sophisticated ways for humans to interact with computers. We regularly witness new line of human-computer interface being developed through voice or computer vision. The design of algorithms that allow computers to recognize speech, face and hand gestures, and even emotional state are some of the current challenges that researchers are facing in order to realize the ultimate dream of having natural human-machine communications.

This paper addresses part of these challenges and examines a particular computer vision tool that makes use of human skin color characteristics. The paper is organized as follows. Section 2 explains face localization, and its roles in human-machine communications and also in our proposed user interface for computers. Section 3 discusses a powerful skin-color-based approach to face localization. The performance of this approach is investigated and experimental results are presented in Section 4. Lastly, the concluding remarks can be found in Section 5.

2. FACE LOCALIZATION

Face localization deals with an image processing problem of locating facial region in a given image. It plays an important role in the development of human-machine communications as it is often the first task that needs to be successful achieved before subsequent steps such as face tracking, face recognition, facial expression analysis, etc. can be carried out.

Indeed, face localization plays a vital part in our proposed user interface that we are designing for computer game control and navigation in 3D graphics world. In our setup, a digital camera is used to capture the user's face. Once face localization is achieved, head pose can then be determined and also head movement can be tracked. In a computer game scenario such as Quake or other first person shooter typed games, players can move their characters not only by keyboard and mouse control, but also through the players' head movement. For instance, players can move left or right in the game by moving their head left or right, and squat by moving their head in downward direction. In choosing such metaphor, other similar controls can be applied to forward and backward head movements, tilting of the head and other gestures. Such computer vision based head movement control should be more intuitive and natural to users than keyboard and mouse interfaces. It also has an advantage of not having to wear any devices such as helmets or goggles.

Our proposed interface design relies heavily on face localization capability. The study of face localization is active and it is gaining increasing attention. Among many methodologies that were proposed over the years (for example, see [1]), it is fair to say that the most popular and superior approach to face localization is the use of color information, whereby estimating areas with skin color is often the first essential step of such strategy.

3. SKIN COLOR DETECTION

The popularity and superiority of skin color approach to face localization comes from the fact that it is fast, reliable and effective. Skin color detection involves relatively simple computation and can be implemented in real-time. Its underlying assumptions are minimal as it can cope with varying lighting condition as well as high level of structure and texture of the background scene. Furthermore, background and camera motions do not pose a problem. More importantly, it is very capable of identifying skin color pixels and providing good coverage of all human races.

We can use color information as a feature to identify a person's face in an image because human faces have a special color distribution that differs significantly, although not entirely, from those of the background objects [2]. Previous studies [2,3] have found that pixels belonging to skin region exhibit similar chrominance components within and across different human races. In the YCbCr color space, chrominance components are represented by Cb and Cr values. Thus, skin color model can be derived from these values.

There are a number of ways to model skin color and to use these models for classifying image pixels into skin color and non skin color. Here we briefly describe two methods for YCbCr color space:

- Simple thresholding technique
- Bayes decision rule for minimum cost

In simple thresholding technique, skin color pixels are identified by the presence of a certain set of Cb and Cr values that is narrowly distributed in the YCbCr color space. Let R_{Cr} and R_{Cb} denote the respective ranges of Cr and Cb values that correspond to skin color, and consequently define our skin color model. Pixel X is classified as skin color (ω_1) if both its Cr and Cb values fall inside their respective ranges of R_{Cr} and R_{Cb}. Otherwise, the pixel is classified as non skin color (ω_2). Details of this technique can be found in [3].

In the Bayesian approach, the skin color detection problem becomes finding the class that gives the minimal cost when considering different cost weightings on the classification decisions. In this technique, we first generate the statistics of skin color and non skin color distributions from a set of training data so that the conditional probability density functions of skin color, $p(X|\omega_1)$ and non skin color $p(X|\omega_2)$ can be obtained. Let C_{ij} denotes the cost of deciding $X \in \omega_i$ when $X \in \omega_j$. It represents the cost of correct classification when $i = j$, and false classification when $i \neq j$. The classification is performed using the Bayes decision rule for minimum cost, which can be expressed as

$$\frac{p(X|\omega_1)}{p(X|\omega_2)} > \tau \Rightarrow X \in \omega_1$$

$$\frac{p(X|\omega_1)}{p(X|\omega_2)} < \tau \Rightarrow X \in \omega_2$$

where

$$\tau = \frac{(C_{12} - C_{22})}{(C_{21} - C_{11})} \cdot \frac{p(\omega_2)}{p(\omega_1)}.$$

In the above equations, $p(\omega_i)$ is the *a priori* probability of class ω_i; and τ represents the adjustment threshold. Note that the costs of false classifications are manipulated by C_{12} and C_{21} for false detection and false dismissal, respectively, while the costs of correct classifications (i.e. C_{11} and C_{22}) are typically set to zero. Further description of this approach can be found in [4].

4. RESULTS AND DISCUSSION

To evaluate the robustness of skin color detection against varying lighting conditions, two images as shown in Figure 1 were captured at different angles to the sun.

The images were stored in YCbCr format. The faces were segmented so that the histograms of the facial regions could be calculated. Figure 2 shows the normalized histograms, in percentage form, of the separate Y, Cb and Cr components of the two facial regions. Two observations were made. Firstly, the distribution of Y values was altered. Secondly, the characteristics of the Cb and Cr distributions remain more or less unchanged. This means that skin color model derived from Cb and Cr values is not sensitive to illuminant changes.

Figure 1. Subject facing at different angles to the sun.

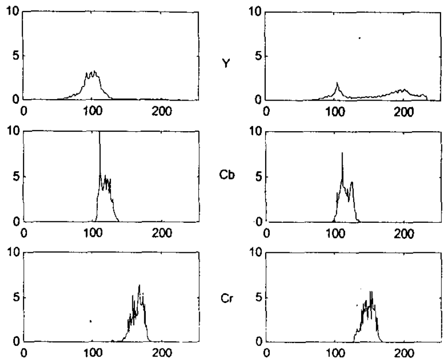

Figure 2. The respective normalized histograms of Y, Cb and Cr components of the facial regions shown in Figure 1.

Two subjects with different skin colors as shown in Figure 3 were used to test the effectiveness of skin color detection against skin color variation. The figure also shows the segmented facial regions. The histograms of Y, Cb and Cr values for the two facial regions were calculated and can be found in Figure 4.

The Cb and Cr histograms show that skin color distributions of the two faces are quite similar even though the color of their skin clearly appears to be different. We have found that the different skin color that we perceived from image cannot be differentiated from the chrominance information of that image region. Therefore, skin color model derived from Cb and Cr values will remain effective regardless of skin color variation. Moreover, the apparent difference in skin color that viewers perceived is mainly due to the darkness or fairness of the skin; these features are characterized by the difference in the brightness of the color, which is governed by Y values.

Both the simple thresholding approach and the Bayesian approach to skin color detection were implemented and tested. Figure 5 shows our program called Face Tracker, which runs in real-time using the simple thresholding approach.

To evaluate the performances of the two skin color detection techniques, testing was carried out on 50 images of different subjects, background complexities, and lighting conditions. The correct classification rates for both the simple thresholding and Bayesian approaches are shown in Figure 6. In term of correct classification, the Bayesian approach has produced the best result.

Figure 3. Subjects of different skin colors.

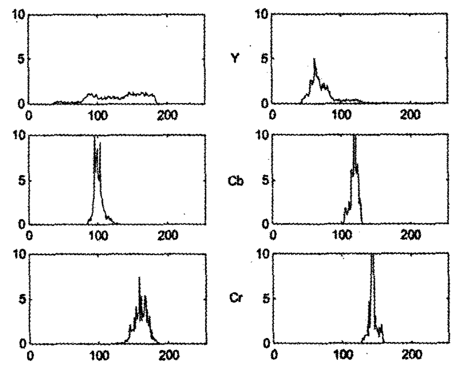

Figure 4. The respective Y, Cb and Cr histograms of the facial regions shown in Figure 3.

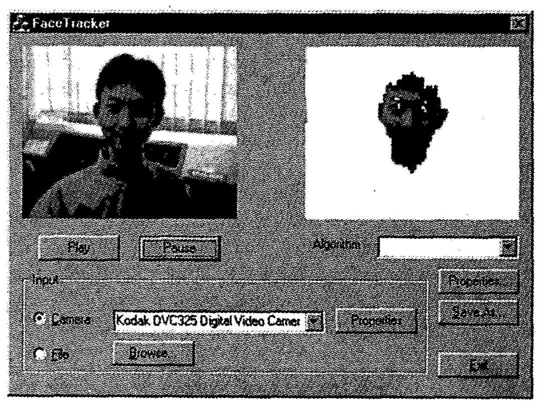

Figure 5. Face tracker program that tracks a person's face in real-time.

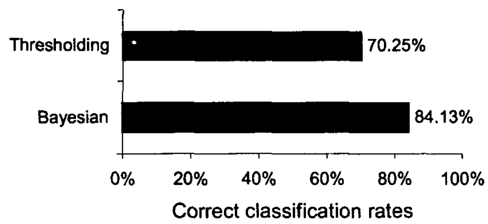

Figure 6. Correct classification rates for simple thresholding and Bayesian approaches to skin color detection.

Table 1. False classification rates.

	False Detection	False Dismissal
Thresholding	28.12%	1.63%
Bayesian	13.40%	2.47%

The false classification rates can be categorized into false detection rates and false dismissal rates as shown in Table 1. A representative result is shown in Figure 7.

Note that post processing tasks can be used to refine the results from the skin color detection by removing skin color like background object and filling in holes resulted from false dismissal in the facial region. Some morphology operation typed post processors can be found in [2].

5. CONCLUSION

In this paper, we have discussed the use of face localization and skin color detection for facilitating human-machine communications. We have explained why the skin color detection approach is so far the most popular and superior approach to face localization. Two skin color detection techniques, namely, the simple thresholding approach and the Bayesian approach were presented. Experimental results were provided to demonstrate the effectiveness of both techniques, and their robustness against varying lighting condition and skin color variation.

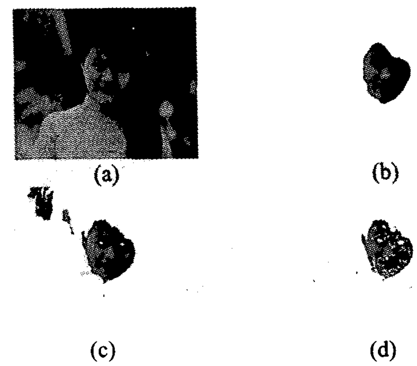

Figure 7. (a) Original image, (b) manually segmented image, (c) result from simple thresholding technique, and (d) result from Bayesian technique.

6. REFERENCES

[1] Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition (FG'98), Nara, Japan, Apr. 1998.

[2] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 4, pp. 551-564, Jun. 1999.

[3] D. Chai and K. N. Ngan, "Locating facial region of a head-and-shoulders color image," in *Third IEEE International Conference on Automatic Face and Gesture Recognition* (*FG'98*), Nara, Japan, pp. 124-129, Apr. 1998.

[4] D. Chai and A. Bouzerdoum, "A Bayesian approach to skin color classification in YCbCr color space," *IEEE Region Ten Conference* (*TENCON'2000*), Kuala Lumpur, Malaysia, vol. II, pp. 421-424, Sep. 2000.