

Asiatic skin color segmentation using an adaptive algorithm in changing luminance environment

Chutisant Kerdvibulvech

Faculty of Information Technology, Rangsit University, Pathumthani 12000, Thailand
E-mail: dr-chutisant@hotmail.com

Submitted 29 February 2012; accepted in final form 15 June 2012

Abstract

Skin color of Asiatic people is a unique color. Therefore, Asiatic skin color segmentation is not trivial, particularly in non-static background. This paper proposes a novel algorithm to segment an area of Asiatic skin color effectively. By using the on-line adaptation of color probabilities, the algorithm proposes two main benefits: it can cope with luminance changes very well, and also it can be processed in real-time. Bayes' theorem and Bayesian classifier are employed to compute the probability of skin color of Asiatic people. Representative results from the experiments are presented to show the efficiency of the proposed system. The system presented can be further used to develop the real-time vision-based applications in many challenging environments.

Keywords: *Asiatic skin, color segmentation, Bayes theorem, Bayesian, adaptive algorithm, luminance*

1. Introduction

Skin color segmentation is still very challenging in the field of computer vision, especially skin color of Asiatic people. Research about color segmentation and a model of color is a dynamic topic because it can be used to further development of many applications in real-life. A model of color relates to the selection of the color space: color spaces efficiently separating the chrominance from the luminance components of color are typically considered preferable (Sigal et al., 2000). Many color spaces have been proposed including RGB (Jebara & Pentland, 1997), normalized RGB (Kim et al., 1998), YCrCb (Chai & Ngan, 1998), etc. The comprehensive survey of detecting skin color in images different color spaces provides an interesting overview of color detection (Kakumanu et al., 2007). In this survey, it gives a review of skin modeling and classification methods based on color information in the visual spectrum, but it does not yet talk about Asiatic skin color segmentation.

Previous methods determined the proper threshold values of each color model, e.g. HSV (Hue, Saturation and brightness Value) (Gonzalez & Woods, 2002). Another recent method used (Cabrol et al., 2005) color region segmentation followed by a color classification and region. After it was segmented, it was used for RoboCup application, i.e. four-legged league or an industrial conveyor wheeled robot. A three-stage face detection approach using

self-organizing Takagi–Sugeno-type fuzzy network with support vector (SOTFN-SV) learning is proposed (Chia-Feng & Shen-Jie, 2008). More recently, a method for skin color segmentation on color photos is presented (Iraji & Yavari, 2011) using Fuzzy YCrCb Color Space with the Mamdani Inference.

Nevertheless, by using these conventional algorithms, it is not able to deal with luminance changes robustly. An essential aspect of any skin color segmentation algorithm is the robustness against luminance variations because changing lighting conditions markedly affects the skin color distribution in an image.

This paper proposes an algorithm to segment an area of Asiatic skin color using a Bayesian classifier that is bootstrapped with a small set of training data and refined through an off-line iterative training procedure (Argyros & Lourakis, 2004). This paper focuses on Asiatic skin color because the Asiatic skin color is not easily segmented. This is because the color is not obviously recognized and is not similar to other skin colors such as black and white. This makes it difficult to segment robustly. Also, luminance changes make it more challenging. This paper proposes an algorithm to solve these problems.

Initially, the proposed algorithm calculates the color probabilities of being Asiatic skin color. The learning process is composed of two phases.

In the first phase, the color probability is obtained from a small number of training images

during an off-line pre-process. The color representation used in this process is YUV 4:2:2 (Jack, 2004). This color space contains a luminance component (Y) and two color components (UV). The main advantage for using the YUV space is that the luminance and the color information are independent. Thus, it is easy to separate the chrominance from the luminance components of color. As Asiatic skin tones differ mostly in chrominance and less in intensity, by employing only chrominance-dependent components of color, one can achieve some degree of robustness to changes in luminance. The probability map in the U and V axis in color model is depicted in Figure 1.

In the second phase, the probability is gradually updated automatically and adaptively from the additional training data images. However, the Y-component of this representation is not employed for two reasons. Firstly, the Y-component corresponds to the luminance of an image pixel. By omitting this component, the developed classifier becomes less sensitive to luminance changes. Secondly, compared to a 3D color representation (YUV), a 2D color representation (UV) is lower in dimensions and, therefore, less demanding in terms of memory storage and processing costs. This disregard of the luminance value has also been shown to be useful in detection and tracking of faces (Hua et al., 2002) and color night vision (Shi et al., 2007).

The proposed method stipulates that the adapting process can be disabled as soon as the achieved training is deemed sufficient. Therefore, when the on-line Asiatic skin color adaptation is started to learn, the assumption is that there is enough skin area in the image. As soon as the on-line adapting process is deemed sufficient (i.e. the Asiatic skin color probability converges to a proper value), the adapting process is stopped manually. In this way, after finishing the on-line learning process, although the Asiatic skin area disappears from the scene, it does not affect the skin color probability.

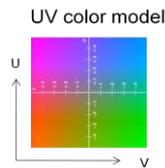


Figure 1 U and V axis in color model used to calculate probability in the proposed algorithm (Jack, 2004)

Therefore, this algorithm is able to get accurate color probability of the Asiatic skin from only a small set of manually prepared training images. This is because the additional skin region

does not need to be segmented manually. Also, because of the adaptive learning, it can be used robustly with changing luminance during the on-line operation.

2. Proposed algorithm

In the proposed algorithm, the color probabilities of being Asiatic skin color adaptively are calculated. The learning process is composed of two phases, as displayed in Figure 2.

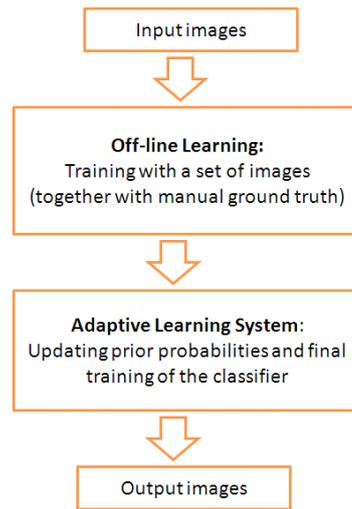


Figure 2 System flowchart

2.1 Off-line learning

During an off-line phase, a small set of training input images is selected, on which a human operator manually delineates skin-colored regions. Figure 3 shows a set of training input images.

Following this, assuming that image pixels with coordinates (x,y) have color values $c = c(x,y)$, training data are used to calculate:

(i) The prior probability $P(a)$ of having Asiatic skin color a in an image. This is the ratio of the Asiatic skin-colored pixels in the training set to the total number of pixels of whole training images.

(ii) The prior probability $P(c)$ of the occurrence of each color in an image. This is computed as the ratio of the number of occurrences of each color c to the total number of image points in the training set.

(iii) The conditional probability $P(c/a)$ of an Asiatic skin being color c . This is defined as the ratio of the number of occurrences of a color c within the skin-colored areas to the number of skin-colored image points in the training set.

By employing Bayes' theorem, the probability $P(a/c)$ of a color c being an Asiatic skin color can be computed by using

$$P(a | c) = \frac{P(c | a)P(a)}{P(c | a)P(a) + P(c | m)P(m)} \quad (1)$$

in which m is the complement of a . This equation determines the probability of a certain image pixel being Asiatic skin-colored using a lookup table indexed with the pixel's color. The prior probability $P(c)$ of the occurrence of each color in an image is employed by using

$$P(c) = P(c | a)P(a) + P(c | m)P(m). \quad (2)$$

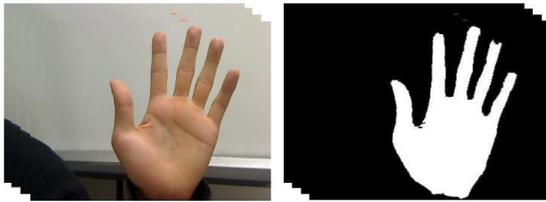


Figure 3 Some examples of training input images during off-line learning process have been done manually.

The resultant probability map thresholds are then set to be threshold T_{max} and threshold T_{min} , where all pixels with probability $P(a|c) > T_{max}$ are considered as being Asiatic skin-colored—these pixels constitute seeds of potential Asiatic skin-colored blobs—and image pixels with probabilities $P(a|c) > T_{min}$ where $T_{min} < T_{max}$ are the neighbors of Asiatic skin-colored image pixels being recursively added to each color blob. The rationale behind this region growing operation is that an image pixel with relatively low probability of being Asiatic skin-color should be considered as a neighbour of an image pixel with high probability of being Asiatic skin-color. Smaller and larger threshold values cause the false Asiatic skin detection. For example, if the threshold value T_{max} chosen is too big, any pixels that constitute the seeds of potential Asiatic skin blobs cannot be detected. Therefore, the values for T_{max} and T_{min} should be determined by test experiments (this experiment uses 0.5 and 0.15, respectively). In other words, these values came from the experimental results. A standard connected component labelling algorithm (i.e. depth-first search) is then responsible for assigning different labels to the image pixels of different blobs.

Size filtering on the derived connected components is also performed to eliminate small isolated blobs that are attributed to noise and do not correspond to the Asiatic skin-colored regions of

interest. Hence, connected components that consist of less than the threshold size are assumed to be noise and then rejected from further consideration. Each of the remaining connected components corresponds to an Asiatic skin-colored blob. In this step, the biggest region as an Asiatic skin-colored blob is selected.

2.2 Adaptive learning system

In Section 2.1 (offline), training with a set of images (together with manual ground truth) is described. In Section 2.2 (online), updating prior probabilities and final training of the classifier are explained.

Training is an off-line procedure that does not affect the on-line performance of the tracker. Nevertheless, the compilation of a sufficiently representative training set is a time-consuming and labor-intensive process. To cope with this problem, an adaptive training procedure has been developed. Training is performed on a small set of seed images for which a human provides ground truth by defining Asiatic skin-colored regions. Following this, detection together with hysteresis thresholding is used to continuously update the prior probabilities $P(a)$, $P(c)$ and $P(c/a)$ based on a larger image data set. The updated prior probabilities are used to classify pixels of these images into Asiatic skin-colored and non-Asiatic skin-colored ones. The final training of the classifier is then performed based on the training set resulting from user editing. This process for adapting the prior probabilities $P(a)$, $P(c)$ and $P(c/a)$ can either be disabled as soon as the achieved training is deemed sufficient for the purposes of the tracker, or continue as more input images are fed to the system.

The success of the color detection depends crucially on whether or not the luminance conditions during the on-line operation of the detector are similar to those during the acquisition of the training data set. Despite the fact that using the UV color representation model has certain luminance independent characteristics, the Asiatic skin color detector may produce poor results if the luminance conditions during on-line operation are considerably different to those used in the training set. Thus, a means of adapting the representation of Asiatic skin-colored image pixels according to the recent history of detected colored pixels is required.

To solve this problem, Asiatic skin color detection maintains two sets of prior probabilities (Zabulis et al, 2009). The first set consists of $P(a)$, $P(c)$, $P(c/a)$ that have been computed off-line from the training set. The second is made up of $P_w(a)$, $P_w(c)$, $P_w(c | a)$, corresponding to the $P(a)$, $P(c)$, $P(c/a)$ that the system gathers during the W most

recent frames respectively. Obviously, the second set better reflects the “recent” appearance of Asiatic skin-colored Asiatic skins and is therefore better adapted to the current luminance conditions. Asiatic skin color detection is then performed based on the following moving average formula:

$$P_A(a | c) = \gamma P(a | c) + ((1 - \gamma) P_W(a | c)). \quad (3)$$

where $P_A(a | c)$ represents the adapted probability of a color c being an Asiatic skin color. $P(a/c)$ and $P_W(a | c)$ are both given by Equation (1) but involve prior probabilities that have been computed from the whole training set [for $P(a/c)$] and from the detection results in the last W frames [for $P_W(a | c)$]. γ is a sensitivity parameter that controls the influence of the training set in the detection process ($0 \leq \gamma \leq 1$). If $\gamma = 1$, then the Asiatic skin color detection takes into account only the training set (35 images in the off-line training set), and no adaptation takes place; if γ is close to zero, then the Asiatic skin color detection becomes very reactive, relying strongly on the recent past for deriving a model of the immediate future. W is the number of history frames. If W value is too high, the length of history frames will be too long; if W value is set too low, the history for adaptation will be too short. In this implementation, $\gamma = 0.8$ and $W = 5$ are empirical conditions which gave good results in the tests that have been performed.

Thus, the Asiatic skin color probability can be determined adaptively. By using on-line adaptation of Asiatic skin color probabilities, the classifier is easily able to cope with considerable luminance changes, and also it is able to segment the Asiatic skin color even in the case of a dynamic background.

3. Experimental results

In this section, representative results from the experiment are shown. Figure 4 provides representative snapshots of the experiment. In the experimental results, the training set used is collected from one Asiatic person. The reported experiment is based on a sequence that has been acquired with USB camera at a resolution of 320x240 pixels. This process is done in real time and on-line using a Intel (R) Core (TM)2 Duo CPU P8600 laptop computer running MS Windows at 2.4 GHz. To explain the characteristics of the training set, note that the training set (35 images in the off-line training set) was collected from a different room, so that the luminance is obviously different from the experimental room. So the luminance changes make it quite challenging. Also note that the person in the training set is not the Asiatic person in the test

sequence. The left window depicts the input images, while the right window represents the output images.

In the initial stage (frame 17), when the experiment starts, the Asiatic skin color probability does not converge to a proper value. In other words, the color probability is scattering. So the segmented output cannot be achieved satisfactorily because it uses only from the off-line data set which the lighting is extremely different. At frames 22, 32 and 42, after performing the adaptive learning process, the Asiatic skin color probability gradually converges to a proper value: the result becomes better. Later at frames 52, 62 and 75, the result can be achieved robustly. Note that the adapting process is stopped manually in frame 52.

It is important to note that the lighting used to test between off-line and on-line is obviously different. However, it can be observed that the segmented result of Asiatic skin can be still determined without effects of different light sources in each representative frame. This is because a Bayesian classifier and an on-line adaptation of color probabilities are utilized to deal with this. For comparison with previous methods, to the best of our knowledge, there is no previous method focusing on Asiatic skin color segmentation. Therefore, it is extremely difficult to compare directly with the results of other methods. However, demonstration video of additional result can be found online at the following website:

<http://www.rsu.ac.th/ric/chutisant/RJAS/result.avi>

4. Conclusion and future works

This paper has presented an adaptive learning system which is performed by using a Bayesian classifier and the online adaptation of Asiatic skin color probabilities. This algorithm effectively deals with any luminance changes in different areas. Furthermore, the computation time of the method is real-time. The result of the proposed method is the segmenting of Asiatic skin to be applied for a human-computer interaction system for virtual reality and augmented reality environments. Note that this paper focuses on detecting one connected contour representing Asiatic skin tone.

While the results from this paper are encouraging, there are some possible issues. Even though some successful results have been produced, the current system has been tested with a limited number of Asiatic people. More specifically, if a skin color of Asiatic people has totally different tone to the training data set or if the color of the background is closed to the color of the skin, the system may not segment perfectly. To achieve robustness, this could be considered in the future and

a larger sample population of Asiatic people can be examined. Moreover, if the size of Asiatic skin color during online adaptation is significantly different from the size of Asiatic skin color used during offline training, the proposed probability approach may not cope with it very well. Also, this algorithm could be further utilized to a real-life augmented reality application as well. The future work will be undertaken to refine these problems.



Frame 17 (the adaptive learning process is started)



Frame 22



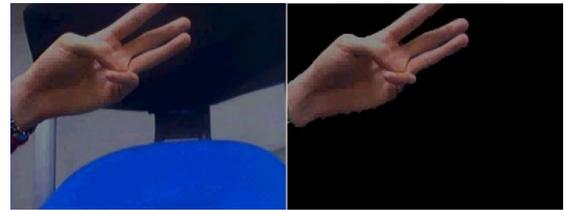
Frame 32



Frame 42



Frame 52 (the adapting process is stopped in this frame)



Frame 62



Frame 75

Figure 4 Asiatic skin segmentation based on the color probability: from off-line learning to adaptive learning

5. References

- Argyros, A. A. & Lourakis, M. I. A. (2004). Real time Tracking of Multiple Skin-Colored Objects with a Possibly Moving Camera. *European Conference on Computer Vision*.
- Cabrol, A. D., Bonnin, P., Costis, T., Hugel, V., & Blazevic, P. (2005). A New Video Rate Region Color Segmentation and Classification for Sony Legged RoboCup Application. *Lecture Notes in Computer Science. RoboCup*.
- Chai, D. & Ngan, K. N. (1998). Locating facial region of a head-and-shoulders color image. *3rd IEEE International Conference on Automatic Face and Gesture Recognition (FG'98)*.
- Chia-Feng, J. & Shen-Jie, S. (2008). Using self-organizing fuzzy network with support vector learning for face detection in color images. *Neurocomputing*, 71(16-18), pp. 3409-3420.
- Gonzalez, R. & Woods, R. E. (2002). *Digital Image Processing*. 2nd ed, Prentice Hall Press., ISBN 0-201-18075-8.
- Hua, R. C. K., Silva, L. C. D. & Vadakkepat, P., (2002). Detection and Tracking of Faces in Real-Time Environments. *International Conference on Imaging Science, Systems, and Technology*.

- Iraji, M. S. & Yavari, A. (2011). Skin Color Segmentation in Fuzzy YCBCR Color Space with the Mamdani Inference, *American Journal of Scientific Research*, ISSN 1450-223, X Issue, pp.131-137.
- Jack, K. (2004). Video demystified, Elsevier science, 4th edition.
- Jebara, T. S. & Pentland, A. (1997). Parameterized structure from motion for 3D adaptive feedback tracking of faces, *IEEE Conference on Computer Vision and Pattern Recognition*.
- Kakumanu, K., Makrogiannis, S., & Bourbakis, N. (2007), A survey of skin color modelling and detection methods. *Pattern Recognition* 40(3), pp.1106-1122.
- Kim, S. H., Kim, N. K., Ahn, S. C., & Kim, H. G. (1998). Object oriented face detection using range and color information. 3rd *IEEE International Conference on Automatic Face and Gesture Recognition*.
- Sigal, L., Sclaroff, S. & Athitsos, V. (2000). Estimation and prediction of evolving color distributions for skin segmentation under varying illumination. *IEEE International Conference on Computer Vision and Pattern Recognition*. Volume 2, ISBN 0-7695-0662-3, pp. 152-159.
- Shi, S., Wang, L., Jin, W., & Zhao, Y. (2007). Color night vision based on color transfer in YUV color space. *International Symposium on Photoelectronic Detection and Imaging*, pp. 66230B-66230B-12.
- Zabulis, X., Baltzakis, H. & Argyros, A.A. (2009). Vision-based Hand Gesture Recognition for Human-Computer Interaction, In "*The Universal Access Handbook*", Lawrence Erlbaum Associates, Inc. (LEA), Series on "*Human Factors and Ergonomics*", ISBN: 978-0-8058-6280-5.