Body Orientation Estimation Using Template Matching Based on Shape and Skin Color Information

Hsin-Chun Tsai, Jhing-Fa Wang, and Hung-Chieh Tai
Department of Electrical Engineering
National Cheng Kung University
Tainan
tsaihcmail@gmail.com, wangjf@csie.ncku.edu.tw, shakerathing@gmail.com

***Abstract*—In real world, body orientation is a problem of person identification. The variation of orientation makes features hard to be extracted in a uniform form. If the orientation is estimated, we can select suitable features based on orientation. It may improve the performance of person identification. For the reason, we propose a body orientation estimation method in this paper. First, body region in images is extracted by background subtraction technology. Template matching is adopted to estimate the body orientation. Because the head and trunk may face the different orientations, we construct the templates with the images of head and shoulder region. In addition to contours, shapes and skin color areas are reserved to provide more orientation information for templates. In the experiments, the average accuracy for detecting frontal, left, right, and back orientation is higher than 97.8%. It explains that shape and skin color information is credible to describe the orientations. Besides, the body orientation estimation scheme works in real-time.**

I. INTRODUCTION

In some situations such as surveillance system, the features in a uniform form are hard to be received, because people who pass the camera seldom face camera straightly unless they are intentional. Therefore, the accuracy of person identification is influenced by body orientations. A solution is to collect all features in different orientations. However, the features in different orientation are often very different. It is difficult to find the common components to train the classifiers, so the accuracy may be low. Another solution for this question is to detect the orientation of body. The images in near orientations are gathered and used to train respective classifiers. This way reduces the influence of body orientation, and let person identification retain the original performance. Besides, as a preprocessing algorithm, the orientation estimation method should not cost much computation time. Thus, we propose a body orientation estimation method for these purposes.

Before estimating the orientation, the body region has to be extracted. The result has a great effect upon the body orientation estimation. Segmentation method [3] is a reliable way. It segments an image into many sub-blocks based on similar color firstly. A larger sub-block which consists of some neighbor sub-blocks is classified by pre-trained shape classifiers. If the classifiers are trained with body shape in different orientations, body region extraction and body orientation estimation can be implemented at the same time [2]. An advantage of the segmentation method is that background images are unnecessary. However, the method costs much time. Another method for getting body region is background subtraction techniques. Possible color ranges of each pixel are recorded as a background model based on pre-prepared background images. The pixels of foreground are often out of the color range. The images including subjects are subtracted with the background model, so the foreground stands out. Furthermore, the method needs fewer computing time than segmentation.

After receiving the body region, body orientation can be estimated. Sometimes, face direction is the same as body orientation. In [5], the face area and mouth location are extracted by color information, and they are used to estimate face direction. However, if the distance between subjects and the camera becomes larger, the relationship of face area and mouth location cannot be extracted precisely. The authors of [1] separate face images into different directions as well as different sizes, and estimate face direction with SVM algorithm. Besides, eigenspace technique [4] is adopted in face direction estimation. But in some situations, face direction and trunk orientation are not same. If images including faces and trunks are used to find subspace, the clothes may influence the recognition performance. In the paper [7], whole body region is vertically divided into some blocks. The feet block is used to estimate the orientation based on the orientation from heel to shoe toes. In [6], head and shoulder contour model for different orientations is built to estimate body orientation. Observing the images in [1], we find that the face images in small size are too blurred to recognize identities. But the accuracy of orientation estimation is still high. We refer the situation to color and shapes which provide credible information for orientations, and adopt template matching for orientation estimation.

The remaining of this paper is organized as follows. In section II, we take an overview of the proposed body orientation estimation method and introduce the adopted background subtraction technology as well as skin color range in color space. Section III explains the procedure of the template generation. Then, the body orientation estimation based on template matching is presented in section IV. Finally, section V shows the experiments and section VI concludes this paper.

II. Overview

A. Proposed Framework

The framework of the body orientation estimation is shown in figure 1. The body region in the captured image is extracted by background subtraction. Skin region extraction is employed in marking the skin area of body region and catching the shape and skin color information. The head and shoulder region (HS region) is set to be region of interest (ROI). The ROI is compared with the pre-built templates. If the similarity between the ROI and the most matching template is high enough, the body orientation which the template represents will be considered as the result.

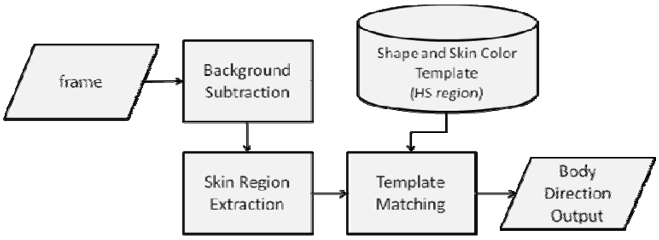


Figure 1. Proposed Framework

B. Background Subtraction

We adopt the background subtraction technology provided in [8]. The background model is built as a codebook. According to the prepared background images, the color variation of each pixel is recorded based on range-adjustable boxes. The built background scope is limited by the range of boxes. YUV or HSV color space is adopted to construct the background model in [8], because variation is along to the brightness axis. In order to avoid considering shadow as foreground, we choose HSV color space and give V a large range for variation. Besides, some criterions for HSV which are presented in reference [9] are employed in our background subtraction method. When brightness is not high enough, the hue and saturation are confused. When saturation is very low, hue is unreliable. In addition, hue is in the range of 0-360. It is necessary for background subtraction to notice that the hue located near 0 and 360 is the same. Figure 2 shows an example of the background subtraction.

C. Relationship of Color Space and Skin

Skin color is important information in person detection and recognition field. Face region and hands can be detected or confirmed by skin color [10], [11], [12]. Researchers construct skin color model in difference color space such as RGB, YCbCr, HIS, and HSV. Reference [11] presents the distribution of skin color in YCbCr and HSV. It works efficiently. In [12], comparison of the performance is displayed. The highest performance is obtained from YCbCr among the traditional color spaces. For this reason, we refer YCbCr parameters [11] to skin color equation as follows.

$$
\begin{aligned}
&if\,(Y>128)\quad \theta_1=-2+\frac{256-Y}{16};\quad \theta_2=20-\frac{256-Y}{16};\\
&\qquad \theta_3=6;\quad \theta_4=-8\\
&if\,(Y\le 128)\quad \theta_1=6;\quad \theta_2=12;\quad \theta_3=2+\frac{Y}{32};\\
&\qquad \theta_4=-16+\frac{Y}{16}\\
&Cr\ge -2(Cb+24);\quad Cr\ge -(Cb+17);\\
&Cr\ge -4(Cb+32);\quad Cr\ge 2.5(Cb+\theta_1);\\
&Cr\ge \theta_3;\quad Cr\ge 0.5(\theta_4-Cb);\\
&Cr\le \frac{220-Cb}{6};\quad Cr\le \frac{4}{3}(\theta_2-Cb)
\end{aligned}
\tag{1}
$$

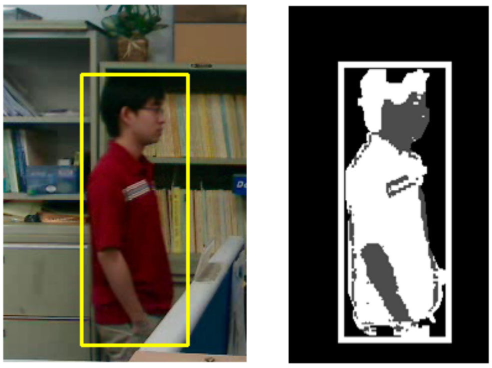
III. Generating Template with Shape and Skin Color

A. Skin Region Extraction

Traditional method for orientation estimation is to build the model with contours. However, it is only efficient in classifying frontal and profile orientation, because the width of body is significantly different in these orientations. Apart from contours, we add skin color and shape information into orientation estimation method. When a person faces forwards, the skin area in head region must be the largest. The smallest skin area in head region exists when a person faces backwards. Moreover, profile can be identified with the concentration of skin area. If a person faces left side, the skin area in the left-half side of head region is larger than that in right-half side. The rest estimation for other body orientations may be deduced by analogy. After skin area and contour are extracted, the body shape, average hair style, location of collar as well as facial features, these kinds of shape information are received at the same time. The information is credible for body orientation estimation.

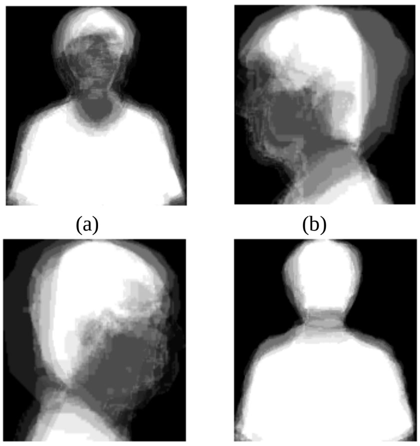
B. Template with Shape and Skin Color Information

Head and trunk may face different orientations. Therefore, the body orientation refers to both head and trunk for a decision. Because shoulder orientation is equal to trunk orientation, we build templates with the head and shoulder region (HS region). The images in database are divided according to orientations of the HS region. After background subtraction and skin region extraction we receive processed images including shape and skin information. Till now, the color of background region is set black by background subtraction; the color of skin region is painted with blue, and the color of the rest region is white. Take figure 2 as an example. The upper bound of the rectangle is calvaria, and it



(a) (b)

Figure 2. An example of background subtraction and skin filter. (a) Original image. (b) Processed image.



(a) (b) (c) (d)

Figure 3. (a) Frontal template. (b) Left template. (c) Right template. (d) Back template.

is also the upper bound of HS region. The width of the rectangle which is equal to the thickness of body is set as the width of HS region. Empirically, the length which is the product of the width of rectangle and 1.15 is calculated as the height of the HS region. The values can contain the region from calvaria to shoulder. The HS region of these processed images which belong to the same orientation are converted to gray level and normalized to a fixed size. Finally, a template is generated by averaging these HS region images. Figure 3 shows the templates in four different orientations. These templates are built with 12 images which are provided equally by 3 males and 1 female. The regions in gray are the skin areas.

IV. Body Orientation Estimation Based on Template Matching

When body region is extracted, the background is set black. The skin area and remaining area in body region are painted blue and white separately. After the body region is converted to gray level, the skin area is gray. Then, the HS region is extracted and normalized as the same size of templates.

The procedure of body orientation estimation based on template matching is formulated in equation (2) and (3). If there are N orientations for estimation, N templates shall be prepared. The differences between HS region and templates are obtained by equation (2):

$$d^n = \sum_{j=0}^{H-1}\sum_{i=0}^{W-1}(I_{i,j} - T^n_{i,j}), \quad \forall n \in [1...N] \tag{2}$$

Where $I_{i,j}$ represents the normalized HS region image, and $T^n_{i,j}$ means the nth template. The width and height of templates are notated with W and H. Therefore, the difference between HS region and the nth template which is represented as d^n is calculated. Next, the differences are estimated by equation (3) for body orientation.

$$r = \begin{cases} n\,, & if \quad d^n < d^m - \tau, \quad \forall m,n \in [1...N] : m \neq n \\ unsure\,, & otherwise \end{cases} \tag{3}$$

The symbol d^m is the difference between HS region and the m^{th} template. The τ is a threshold. When d^n is smaller than the all results of d^m minus τ, it means the orientation of the n^{th} template is a more likely result than the other templates. The result which is notated as r is the orientation of the n^{th} template.

V. Experiments

A. Experimental Setup

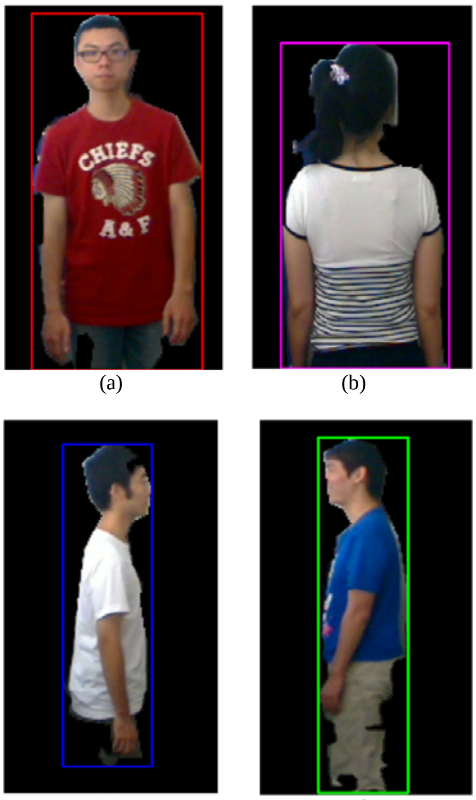
We prepare a self-fabricating video database for test, with 4 males and 1 female individually. The videos are captured in the environment with stabilized light source. Besides, the width and height of each frame are 640 and 480 pixels. In each video, the subject sets out from 4.0 meters away from the camera, and goes straight toward the camera. When the subject reaches the place which is 1.5 meters away from the camera, he or she moves backwards to the starting point and keeps facing the camera. Each subject is asked to repeat the process two times. Then, the subject turns left at 4.0 meters away from camera and goes toward the camera with side step. When the subject reaches the place which is 1.5 meters away from camera, he or she is asked to turn around to face the opposite orientation, and backs with side step. Therefore, we can test the performance of body orientation estimation in different distances. On the other hand, the templates for these four orientations are built in advance with pre-prepared images. There are 12 images which are provided by 4 people equally used to build a template. Three people who provide images to build templates participate in testing video database.

B. *Experimental Results*

In order to evaluate the performance of the proposed body orientation estimation method, we present the experimental result in different orientations and distances. The accuracy is calculated based on the ratio of correct result numbers and believable result numbers. In other words, the equivocal answers are ignored. Figure 4 presents the result images. The subjects in these images are framed with different colors according to body orientations. The experimental results are shown in table I. The highest average accuracy, 100%, appears in frontal orientation. The average accuracy of the other orientations is about 97.8%-98.5%. The result describes that skin color and shapes provide credible information to estimate body orientation. Besides, the variation of distance does not influence the accuracy obviously. At last, the method can work in real-time.

Table I
Experimental Results

	Frontal	Left	Right	Back
Subject 1	100%	92.9%	100%	100%
Subject 2	100%	100%	100%	89%
Subject 3	100%	100%	100%	100%
Subject 4	100%	100%	100%	100%
Subject 5	100%	100%	89.9%	100%



(a) (b) (c) (d)

Figure 4. Result images in different orientations are painted with different colors. (a) Frontal orientation – red. (b) Back orientation – purple. (c) Right orientation – blue. (d) Left orientation – green.

VI. CONCLUSIONS

In this paper, we propose a body orientation estimation method. The body region is caught by background subtraction. The skin filter based on YCbCr color space reserves shape and skin color information. According to the information, templates are built in different orientations. The testing images are compared with the templates by template matching technology. The most matching pair provides the estimated orientation.

In the experiment, four body orientations are tested to evaluate the performance of the proposed method. The average accuracy in each orientation is higher than 97.8%, and the best one reaches 100%. The result explains that the skin color and shapes provide credible information for body orientation estimation. Besides, the proposed framework can work in real-time. It means that the body orientation estimation can be applied as a preprocessing for person identification.

REFERENCES

[1] J. Hirata, M. Morimoto, and K. Fujii, "Estimating Face Direction from Low Resolution Images," World Automation Congress, 2008.

[2] X. B. Cao, H. Qiao, and J. Keane, "A Low-Cost Pedestrian-Detection System with a Single Optical Camera," IEEE Transactions on Intelligent Transportation Systems, vol. 9. No. 1, March 2008, pp.58-67.

[3] E. Borenstein, E. Sharon and S. Ullman, "Combining Top-down and Bottom-up Segmentation," IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop, 2004.

[4] J. Okubo, B. Sugandi, H. Kim, J. K. Tan, and S. Ishikawa "Face Direction Estimation Based on Eigenspace Technique," International Conference on Control, Automation and Systems, Korea, 2008.

[5] X. P. Chen, Q. Q. Yang, H. H. Liao, W. P. Sun, and S. S. Yu, "Real-Time Face Pose Estimation in Video Sequence," 2nd International Workshop on Education Technology and Computer Science, 2010.

[6] M. W. Lee, and R. Nevatia, "Body Part Detection for Human Pose Estimation and Tracking," IEEE Workshop on Motion and Video Computing, 2007.

[7] A. Genco, and E. Varrica, "Human Figure Recognition and the Estimation of its Direction," 7th International Workshop on Computer Architecture for Machine Perception, 2005.

[8] "OpenCV," Software available at http://sourceforge.net/projects /opencvlibrary/

[9] K. Ohba, Y. Sato, and K. Ikeuchi, "Appearance-based Visual Learning and Object Recognition with Illumination Invariance," Machine Vision and Applications, 2000, pp. 189-196.

[10] C. Garcia, and G. Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis," IEEE Transactions on Multimedia, Vol. 1, No. 3. September 1999.

[11] R. Lukac, and K. N. Plataniotis, "Color Image Processing Methods and Applications," 2007, pp. 294-307.

[12] S. K. Singh, D. S. Chauhan, M. Vatsa, and R. Singh, "A Robust Skin Color Based Face Detection Algorithm," Tamkang Journal of Science and Engineering, Vol. 6, No. 4, 2003, pp. 227-234.