Robust Extraction of Moving Objects Based on Hue and Hue Gradient

Yoo-Joo Choi[1], Je-Sung Lee[1], and We-Duke Cho[2]

[1] Seoul University of Venture and Information, Department of Computer Science and Application, Seoul, Korea
yjchoi@suv.ac.kr, mpegman@paran.com
[2] Ajou University, Center of Excellence for Ubiquitous System, Suwon, Korea
chowd@ajou.ac.kr

Abstract. This paper presents a new method for robustly extracting moving objects in an environment with varying illuminations. The proposed method applies background subtraction scheme based on hue and hue gradient to minimize the effect of the illumination change. First, we train the background images in the HSI color space and build the Gaussian background model with respect to the hue and hue gradient. Next, image subtraction is performed between the trained background image and the current input image based on the Gaussian background model. Finally, the morphological operations are applied to remove the background noise. In this paper, we compare the previous subtraction schemes to our method applied to both the hand and body tracking in order to prove the robustness of the proposed method in sudden illumination changes.

1 Introduction

As the ubiquitous computing paradigms become more widely spread, intuitive human-computer interactions that do not employ cumbersome interaction devices has gained positive attention. In particular, vision-based interfaces have been an active and growing research area for the last decade due to the advantage of a vision-based approach, which promotes easy interaction for the users through natural motion and eliminating the use of any input devices. The vision-based interface offers various technical challenges for motion tracking which requires exact extraction of the human body from the video images and analyzing the variable appearance of human forms in the high dimensionality of the state space.

Robustness is an important feature of the user interface. Robustness of the vision-based interface is generally dependent on the accuracy of the human extraction from the input video images and the exact analysis of the human motion based on the segmented human body images. The segmentation of a human body in the input video images should be insensitive to the changes in the background and illumination. The background subtraction method is usually used to separate the human body region from the background. The numerous background subtraction methods based on different color models have been presented to reduce the effect of the illumination [1,3,4,5,6].

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2007, LNCS 4555, pp. 784–791, 2007.

In this paper, we present a new background subtraction scheme based on hue and hue gradient to robustly extract a moving object in an environment with varying illuminations. The HSI color model is divided into hue, saturation and intensity factors. The hue factor generally represents the unique value of color itself while minimizing the effect of the illumination. The hue gradient image maintains the feature of the background image while reducing the effect of the shadow and illumination changes. Therefore, we propose the background subtraction scheme based on hue and hue gradient. In the proposed method, we train the background images in HSI color space and build the Gaussian background model with respect to hue and hue gradient. Next, background subtraction is performed between the trained background image and the current input image based on the Gaussian background model. Finally, morphological operations are applied to remove the background noise. We prove the robustness of the proposed method by comparing the results of the previous schemes with those of our scheme in the varying illumination environment.

The rest of the paper is organized as follows. We briefly summarize the features of the previous vision-based approaches in Section 2 and address background modeling based on hue and hue gradient in Section 3. Section 4 presents our background subtraction approach. Experimental results are shown in Section 5 and finally Section 6. concludes the paper.

2 Related Work

Vision-based interfaces can be categorized into two classes [2]. One is the marker-based approach and the other is the marker-free approach which generally uses the background subtraction techniques. The marker-based approach is able to extract the motion information by simply tracking the markers. However, this approach requires the user to wear a special suit with many markers and tracking these markers when multiple markers are occluded is prone to errors. In the marker-free approach, the special suit with markers is not necessary and the human body is generally extracted by applying the background subtraction techniques.

Previous background subtraction techniques have exploited the statistical color differences between an input video image and the background reference image. The background reference image is built during a certain period of the initial time [1,3,4, 6]. Hong and Woo [1] exploited the RGB and the normalized RGB color space for the background subtraction to reduce the shadow effect. Elgammal et al. [3] applied the chromaticity coordinates, that is, the normalized RG color space, for detection of the target without detecting its shadow. Thakkar[4] et al. proposed a novel technique for automatically extracting a moving object in gray-scale images captured by a moving camera. They basically exploited the intensity values for comparison of adjacent frame differences. Their approach can miss the motionless object because they use the adjacent frame differences to obtain the moving object area. Lee et al. [5] exploited the YCbCr color model to extract the skin color range among multiple moving objects. Lee [6] used the gray-scale intensity value in the moving object segmentation. He combined the pixel-based frame difference approach and the background difference approach and focused to connect the broken boundary of the moving objects. His approaches can also miss the motionless object.

3 Background Modeling Based on Hue and Hue Gradient

3.1 HSI(Hue-Saturation-Intensity) Color Space

The HSI color model was designed with graphic designers and artists in mind and how they perceive and think colors. The HSI color model consists of the saturation(the pureness of a color), hue(the color in itself) and intensity(the brightness of the color). This color model is represented by the basic triangle coordinate as shown in Fig. 1.

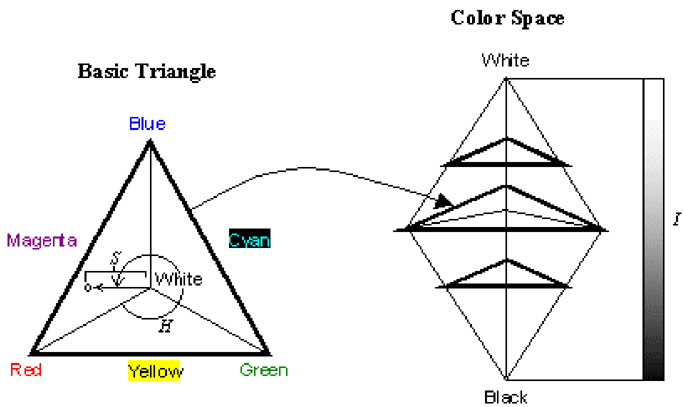


Fig. 1. HSI Color Model

Hue(H) is the angle of the vector over the basic triangle, starting from red (0 degrees). Saturation(S) is the proportional size of the module of the projection of the vector over the basic triangle. Intensity(I) is the distance from the end of the vector to the basic triangle. HSI color is converted from the RGB color space as Eq. 1.

$$H = \begin{Bmatrix} \theta & if & B \le G \\ 360 - \theta & if & B \succ G \end{Bmatrix}$$

$$\theta = \cos^{-1}\left(\frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right) \qquad (1)$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)]$$

$$I = \frac{1}{3}(R+G+B)$$

3.2 Background Modeling

We train the background images in HSI color space for fifty frames and all pixels of the static background scene image are modeled as a Gaussian distribution with respect to the hue and the hue gradient during background training. Each pixel of the background reference image is modeled as follows.

$$< \mu(H_i), \sigma(H_i), \mu(\Delta H_i), \sigma(\Delta H_i) > \quad (2)$$

where, $\mu(H_i)$, and $\sigma(H_i)$ are the vectors of the mean and standard deviation of pixel i's hue channel in HSI color space. $\mu(\Delta H_i)$, and $\sigma(\Delta H_i)$ are vectors of the mean and standard deviation of pixel i's hue gradient. Hue gradient is evaluated as Eq. 3.

$$\Delta H_{(x,y)} = \sqrt{(H_{(x+1,y)} - H_{(x,y)})^2 + (H_{(x,y+1)} - H_{(x,y)})^2} \quad (3)$$

where, $H_{(x,y)}$, $H_{(x+1,y)}$ and $H_{(x,y+1)}$ are the hue channel values in pixels (x, y), (x+1, y) and (x, y+1), respectively. The mean and standard deviation of the hue and hue gradient are gradually updated for the training as Eq. 4 and Eq. 5. Fig. 2 depicts the gradual progress to build the Gaussian background model.

Mean Update:

$$\begin{aligned}
&\mu(H_i(0)) = H_i(0) && t = 0 \\
&\mu(H_i(t)) = (1-\alpha)\mu(H_i(t-1)) + \alpha H_i(t) && t \geq 1 \\
&\mu(\Delta H_i(0)) = \Delta H_i(0) && t = 0 \\
&\mu(\Delta H_i(t)) = (1-\alpha)\mu(\Delta H_i(t-1)) + \alpha \Delta H_i(t) && t \geq 1
\end{aligned} \quad (4)$$

Variance Update:

$$\begin{aligned}
&\sigma^2(H_i(0)) = (H_i(1) - \mu(H_i(0)))^2 && t = 1 \\
&\sigma^2(H_i(t)) = (1-\alpha)\sigma^2(H_i(t-1)) + \alpha(H_i(t) - \mu(H_i(t)))^2 && t \geq 2 \\
&\sigma^2(\Delta H_i(0)) = (\Delta H_i(1) - \mu(\Delta H_i(0)))^2 && t = 1 \\
&\sigma^2(\Delta H_i(t)) = (1-\alpha)\sigma^2(\Delta H_i(t-1)) + \alpha(\Delta H_i(t) - \mu(\Delta H_i(t)))^2 && t \geq 2
\end{aligned} \quad (5)$$

where, t is the frame number and α is the constant for the weighted averaging.

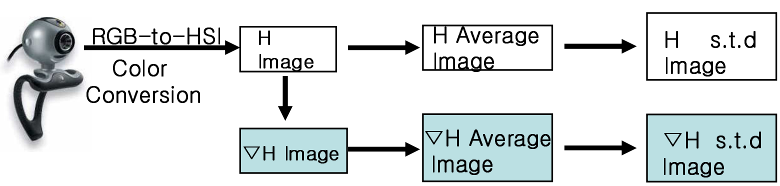


Fig. 2. Gradual Progress to build the Background Model

4 Moving Object Extraction Based on Background Subtraction

The general background subtraction technique is to subtract a current image from the reference background image. The proposed method exploits the hue and hue gradient of the pixel as the cue for the subtraction. Fig. 3 shows the overall procedure of the background modeling and background subtraction for the moving object extraction.

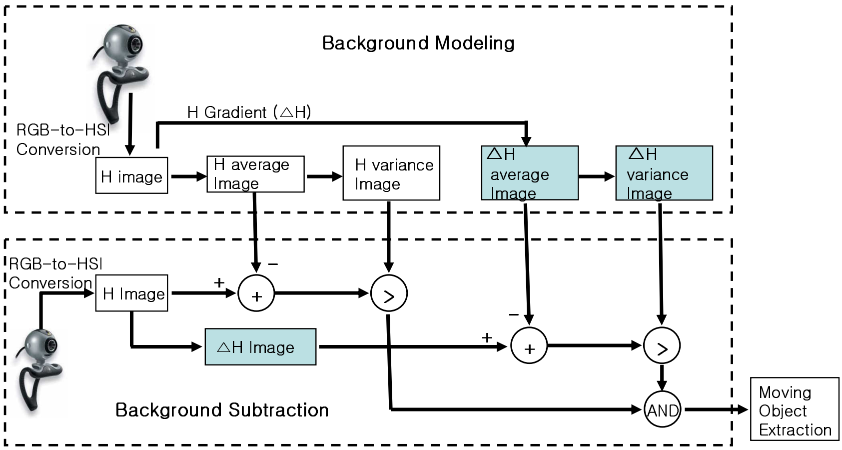


Fig. 3. Overall Procedure for the proposed subtraction

In order to extract the moving object, we compute the hue channel value and the hue gradient for each pixel of the current image frame. We then compare the computed hue and hue gradient values with those of the background model for each pixel as Eq. 6. If the subtraction values are greater than the threshold values which are determined based on the background variance images, those pixels belong to the moving object region.

$$R_i(x) = \begin{cases} 1, & \text{if} \quad \left|H_i(x) - H_{b_i}(x)\right| > \omega_1 \sigma(H_{b_i}(x)) \quad \text{and} \\ & \quad \left|\Delta H_i(x) - \Delta H_{b_i}(x)\right| > \omega_2 \sigma(\Delta H_{b_i}(x)), \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where, $H_i(x)$ and $\Delta H_i(x)$ are the hue and the hue gradient of the pixel *i* in the current frame image, respectively. $H_{b_i}(x), \Delta H_{b_i}(x), \sigma(H_{b_i}(x))$ and $\sigma(\Delta H_{b_i}(x))$ are respectively the averaging hue, the averaging hue gradient, the standard deviation of the hue and hue gradient of the pixel *i* to be saved in the background averaging image and variance image. ω_1 and ω_2 are the constants that determine the threshold ranges. Finally, we apply the morphological opening and closing operations to remove the noise and enhance the quality of the moving object.

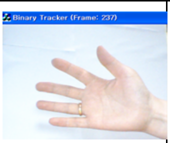
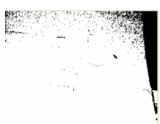
5 Experimental Results

5.1 Hand Tracking

We captured 320 x 240 24bit RGB images from the normal Web Cam called the Logitech QuickCam Chat for the hand tracking. We trained the background scene for fifty image frames with respect to the RGB, normalized RGB, hue and hue gradient.

In order to prove the robustness of the proposed method, we compared the results of the previous methods with those of the proposed method in the sudden illumination change made by switching the auxiliary light off. Table 1 shows the experimental results for the hand tracking. The background training was performed in the bright illumination condition and the first row of Table 1 shows the subtraction results in the same illumination condition as that for the background training. After turning the auxiliary light off, we performed the subtraction schemes without re-training the background scene. The second row of Table 1 shows the subtraction results after the sudden illumination change. The contour of the hand was robustly extracted by applying the proposed subtraction scheme.

Table 1. Comparison of background subtraction schemes for the hand tracking in the sudden illumination change

	Original Image	RGB-based Subtraction	Normalized RGB-based Subtraction	H-based Subtraction	H and H-Gradient Subtractio n
1					
2					

5.2 Body Tracking

The experimental environment for the body tracking was installed as shown in Fig. 4. 640 x 480 frame images were captured in 30 FPS by two Dragonfly2 IEEE 1394 digital cameras. The projector and screen were installed in order to simulate a magic mirror. The moving human body was simultaneously extracted from the two synchronized frame images captured by two stereo cameras for 3-D tracking.

The experiment for the body tracking was performed in the same way as the hand tracking. We made a sudden illumination change by switching the auxiliary light off after the background training and the first body tracking. We then performed the body tracking without the new background modeling. The experimental results are as shown in Table 2. Among the previous subtraction schemes, the hue-based subtraction showed the most robust subtraction result. Table 2 compares the hue-based subtraction with hue and hue gradient-based subtraction. In sudden illumination change, the result of the hue-based subtraction included a lot of background noise, while the proposed subtraction scheme only produced a little amount of the background noise

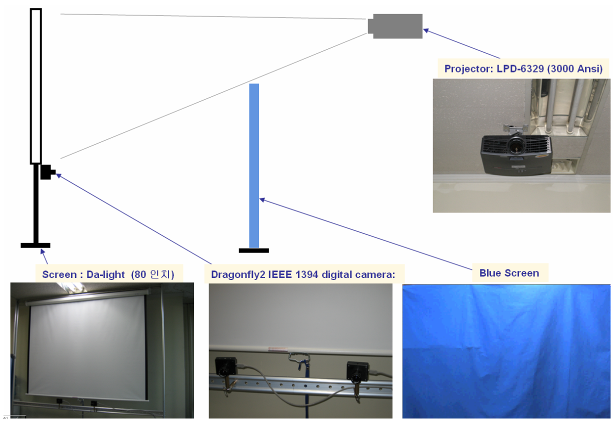
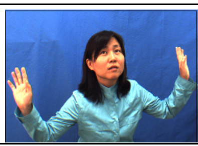
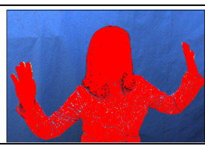
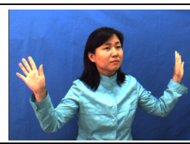
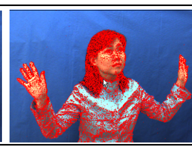
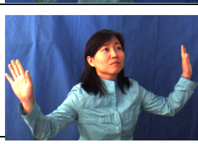
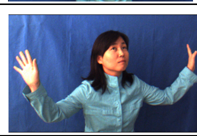
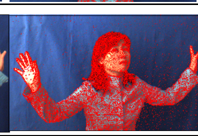


Fig. 4. Experimental environment for the body tracking

Table 2. Comparison of background subtraction schemes for the body tracking in the sudden illumination change

	H-based Subtraction		H and H-Gradient Subtraction	
1				
2				

6 Conclusions

This paper proposed a new background subtraction scheme based on hue and hue gradient. The proposed scheme was applied to the hand tracking and the body tracking to prove the robustness in the illumination change. The robustness of the proposed scheme was compared with those of the previous background subtraction schemes. The proposed scheme showed the most robust results in the hand and the body tracking during the sudden illumination change. In the proposed scheme, the background model is effectively built by gradually updating the averaging and variance images during the background scene training. The effect of the illumination change is minimized by exploiting the hue gradient as the cue of the background subtraction.

For the future work in the area of body tracking, we would like to extract the body feature points and track the body segments based on the extracted feature points in real time. The exact real-time body segment tracking has been an active challenge due to the errors from the segments occlusion and the high computation for the pose estimation. Real-time tracking of the body segments can be applied to many potential applications such as the magic mirror for the virtual dress room and gesture-based game interfaces.

Acknowledgements. This work was supported by the Ubiquitous Computing and Network(UCN) project, the Ministry of Information and Communication (MIC) 21^{st} Century Frontier R&D Program in Korea.

References

1. Hong, D., Woo, W.: A Background Subtraction for a Vision-based User Interface. In: Proc. of PCM03-CD, 1B3.3 (2003)
2. Woo, W., Kim, N., Wong, K., Tadenuma, M.: Sketch on Dynamic Gesture Tracking and Analysis Exploiting Vision-based 3D Interface. In: Proc. of SPIE PW-EI-VCIP'01 4310, 656–666 (2001)
3. Elgammal, A., Duraiswami, R., Harwood, D., Davis, L.S.: Background and Foreground Modeling using Non-parametric Kernel Density Estimation for Visual Surveillance. In: Proc. of the IEEE 90(7), 1151–1163 (2002)
4. takoor, N. , Gao, J.: Automatic Video Object Detection and Recognition with Camera in Motion. In: Proc. of IEEE International Conference on Image Processing (2005)
5. Lee, K.-M., Lee, Y.-M.: Adpative Model-based Multi-object Tracking Robust to Illumination Changes and Overlapping. Journal of the Korea Information Science Society(Software and Applications) 32(5), 449–460 (2005)
6. Lee, H.S.: Moving Object Segmentation using Space-oriented Object Boundary Linking and Background Registration. Journal of the Korea Information Science Socienty(Software and Applications) 32(2), 128–139 (2005)