

Extracting Human Body based on Background Estimation in Modified HLS Color Space

Jang-Hee Yoo, Doosung Hwang, Jong-Wook Han, and Ki-Young Moon

Abstract—The ability to recognize humans and their activities by computer vision is a very important task, with many potential application. Study of human motion analysis is related to several research areas of computer vision such as the motion capture, detection, tracking and segmentation of people. In this paper, we describe a segmentation method for extracting human body contour in modified HLS color space. To estimate a background, the modified HLS color space is proposed, and the background features are estimated by using the HLS color components. Here, the large amount of human dataset, which was collected from DV cameras, is pre-processed. The human body and its contour is successfully extracted from the image sequences.

Keywords—Background Subtraction, Human Silhouette Extraction, HLS Color Space, and Object Segmentation

I. INTRODUCTION

RECENT interest in computer vision has emerged which deals with the analysis of image sequences involving humans. This interest is motivated by the various application domains [1]: video surveillance, athletic performance analysis, human robot interaction, computer graphics, and biometrics. The detection and extraction of moving objects from image sequences is a fundamental and important problem in many vision applications. There are two basic methods for detecting moving objects: temporal differencing and background subtraction. Temporal differencing [2] can adapt to dynamic environments, but cannot robustly extract all relevant object pixels. Thus, the method has been mainly used for tracking moving objects. Conversely, background subtraction [3] extracts the most complete representation of an object, but this method is very sensitive to dynamic scene changes due to illumination [4].

However, background subtraction has been successfully applied to many vision systems as a pre-processing phase for object detection and extraction in an image sequence [3][5][6]. The background subtraction based on color or intensity is a commonly used technique to promptly identify foreground elements. Typical problems in background subtraction include

J. H. Yoo is with the Electronics and Telecommunications Research Institute, Daejeon 305-700, S. Korea (e-mail: jhy@etri.re.kr).

D. Hwang is with the Department of Computer Science, Dankook University, Chungnam 330-714, S. Korea (e-mail: dshwang@dnakook.ac.kr).

J. W. Han is with the Electronics and Telecommunications Research Institute, Daejeon 305-700, S. Korea (e-mail: hanjw@etri.re.kr).

K. Y. Moon is with the Electronics and Telecommunications Research Institute, Daejeon 305-700, S. Korea (e-mail: kymoon@etri.re.kr).

foreground objects with similar colors to the background, and shadows or other variable lighting conditions [8]. Especially, background luminance variations are mainly due to noise and illumination change in indoor sequences. In this paper, we describe a human body extraction method based on background estimation using color information in modified HLS space. The moving human body is detected by background subtraction methods. The body region is then determined by analyzing histogram projection profiles, and its location is verified by prior knowledge such as size and shape. In addition, thresholding and morphology is employed to extract the body contour of a detected human body.

II. BACKGROUND ESTIMATION IN COLOR SPACE

In computer vision and graphics, many different color models exist, and each model uses its own 3D coordinate system to identify uniquely individual colors [9]. The RGB (Red, Green, and Blue) space is the most commonly used color space, because it is directly supported by most color displays and scanners. The HSV (Hue, Saturation, and Value) and HLS (Hue, Lightness, and Saturation) color models are transformations of RGB space that can describe colors in terms more natural to an artist [10]. To estimate the background of an image, the H component is defined as hue value in HSV space, and L and S components are defined as Lightness and Saturation values in modified HLS space. The conversion method from RGB color space to modified HLS space is given by

$$m = \min(g, b)$$
$$H = \begin{cases} -1 & \text{undefined} & \text{if } r = m \\ 60 \times (b - r) / (m - r) & \text{if } g = m \\ 60 \times (r - g) / (m - r) & \text{if } b = m \\ H + 360 & \text{if } H < 0 \end{cases} \quad (1)$$

$$L = (m + r) / 2$$

$$S = \begin{cases} 0 & \text{if } m = r \\ (m - r) / (m + r) & \text{if } L \leq 0.5 \\ (m - r) / (2 - m - r) & \text{otherwise} \end{cases}$$

where r , g , and b are defined as normalized RGB ranges from 0 to 1. The modified HLS space is very similar to original HLS space but slightly more robust to noise.

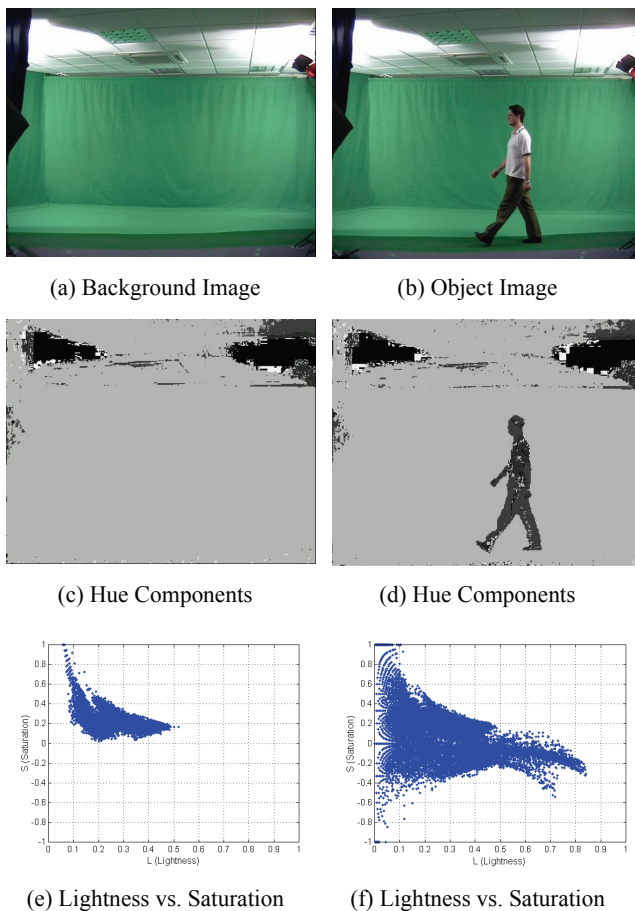


Fig. 1 Color Components for Background Estimation

Here, the chroma-key laboratory is used to allow controlled lighting conditions, and subject is captured against a uniform background. To estimate the background, RGB color space as shown in Figs. 1(a) and (b) is converted to HSV space, and hue component of images is shown in Figs. 1(c) and (d). As can be seen in the figures, hue component is very uniform, thus the mode of hue component is simply calculated by

$$\xi = \text{mode}(h_{x,y}) \quad (2)$$

where $h_{x,y}$ is the hue value at coordinate (x, y) in $m \times n$ image region. In this data, HSV space offered a better estimation of hue component i . The RGB color space is also converted to the modified HLS space. The range of lightness and saturation components in the HLS space is defined as

$$\psi = [\sigma L(i), \sigma S(i)] \quad (3)$$

where σ is a variance of each color component. The background feature (ψ) can be calculated by color clustering method [11] using lightness (L) and saturation (S) components. Figs. 1(e) and (f) show the features of background and object images in rectangular coordinate system.

III. EXTRACTING THE HUMAN BODY SILHOUETTE

Now, the background features can be removed by using Eqns. 2 and 3, namely, the pixel values in HLS space can be re-defined by background estimation as

$$H, L, S = \begin{cases} 0 & \text{if } (p_{k_H} \subset \xi) \wedge (p_{k_{LS}} \subset \psi) \\ p_k & \text{otherwise} \end{cases} \quad (4)$$

where p_{k_H} is H component of pixel k in HSV space, and $p_{k_{LS}}$ is L and S components of pixel k in modified HLS space. Figs. 2(a) and (b) show the results of background components subtraction by using Eqn. 4.

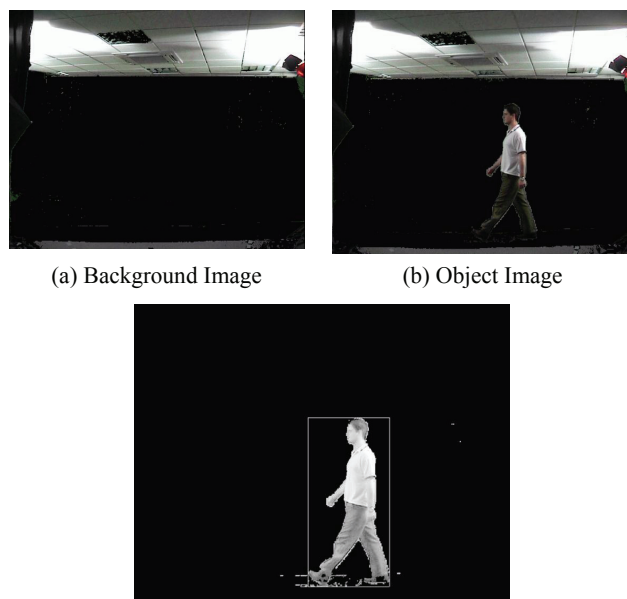


Fig. 2 Background Subtraction

As can be seen in the figures, the object image still has some noise (background components), and the object has also lost some foreground components. Here, to remove this noise and to recover the lost components, noise filtering and histogram logarithm methods are applied to the background subtracted image. The noise filter is defined by

$$g(x,y) = \begin{cases} g(x,y) & \text{if } \left(\sum_{k=1}^1 \sum_{j=1}^1 f(x-j, y-k) \right)_{(j,k) \in \{-j,-k\}} > 2 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where $f(x, y)$ is set to 1 if it is greater than a threshold value T_n , otherwise it is set to 0. Histogram logarithm is defined as

$$p'_k = c \times \log(1 + p_k) \quad (6)$$

where p_k is pixel value at index k , and c is constant. The histogram logarithm increases the dynamic range of grayscale via contrast stretching and is useful to enhance detail in the

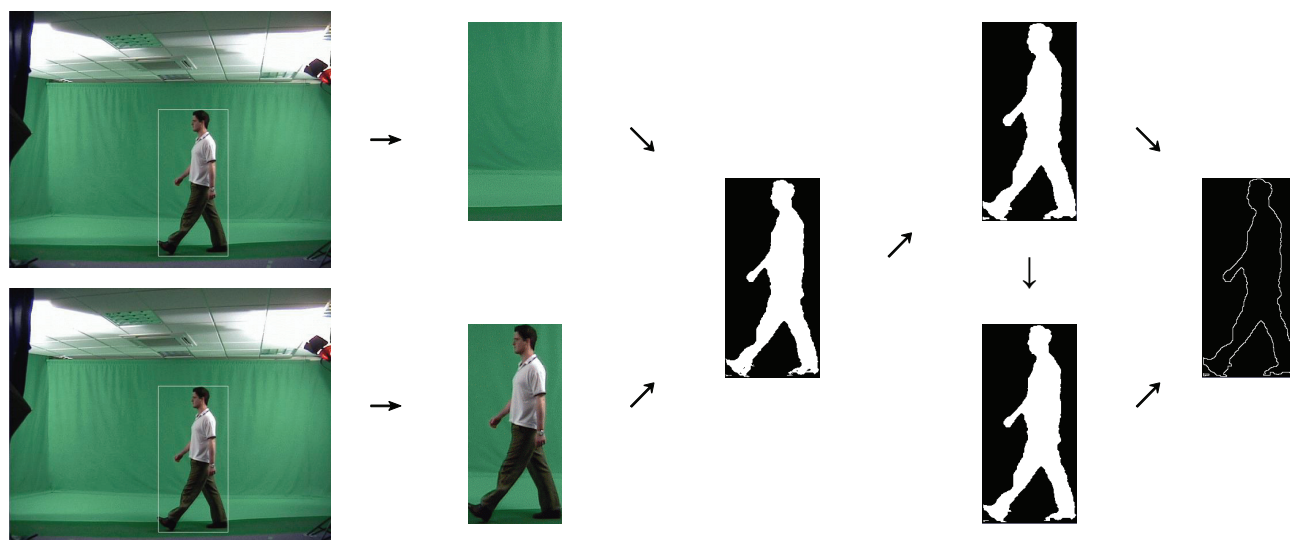


Fig. 3 Extraction of Human Body Contour

Open Science Index, Electrical and Computer Engineering Vol:3, No:11, 2009 publications.waset.org/11683.pdf

darker region of the image. However, the object can be detected by the histogram projection method for grayscale images. Fig. 2(c) shows the detected object by using projection profile in the noise filtered and contrast enhanced object image.

To extract the contour of a detected human body, a thresholding and morphological method is used here. Fig. 3 shows the procedure of thresholding and extracting a human body contour. Thresholding is one of the most important approaches in the field of image segmentation, and choosing a correct threshold is difficult under irregular illumination. However, using the background information in an image can lead to improvement. Accordingly, a thresholding method based on similarity (or dissimilarity) measures between the background and the object image is used. Let $I_b(x, y)$ and $I_o(x, y)$ be the feature (or brightness in grayscale image) of a pixel with coordinate (x, y) in the background image (I_b) and the object image (I_o). Then, the similarity $\Theta(x, y)$ at coordinate (x, y) is computed by

$$\Theta(x, y) = |I_b(x, y) - I_o(x, y)| \quad (7)$$

Similarity values close to zero imply a high probability of being background; conversely, large values of similarity imply high probability as an object. Therefore, the binary image $I_{bi}(x, y)$ is thresholded as

$$I_{bi}(x, y) = \begin{cases} 1 & \text{if } (\Theta(x, y) > \tau) \wedge (I_o(x, y) > \lambda) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Theoretically, thresholding is a very simple image segmentation method, which is very effective and useful for small and low-resolution images, but suffers from difficulty with change in illumination. To improve this method, a more effective algorithm using the probability density of the similarity for determining appropriate values of τ and λ is

required.

In addition, the binary image can have some noise inside the object which is actually a human body part and some noise outside the object. So, morphological filtering is used to remove the noise and to extract the human body contour, by the dilation and erosion. In mathematical terms, the dilation of a set A by a structuring element B is denoted by $A \oplus B$ and is defined as

$$A \oplus B = \{x | B_x \cap A \neq \emptyset\} \quad (9)$$

Finally, the human body contour is determined by arithmetic subtraction between the dilation and erosion images as

$$C = (A \oplus B) - (A \ominus B) \quad (10)$$

If dilation can be said to add pixels to an object or to make it bigger (thickening), then erosion will make an image smaller (thinning). In the simplest case, binary erosion will remove the outer layer of pixels from an object. Therefore, we can obtain an object contour easily just by using the subtraction operation of the dilation and erosion image.

IV. EXPERIMENTAL RESULTS

In experiments, SOTON gait database, which is seven indoor image sequences of each of the 100 different subjects, are used. The detection and extraction of the human body is accomplished by background subtraction and by histogram projection analysis, and thresholding and morphology is then used to extract the contour of a detected human body. The size of the human body in the image sequences is approximately 160x360 pixels in 720x576 image. Fig. 4 show extracted human body contours during one full stride (or two steps) from an image sequence of the SOTON database.

V. CONCLUSIONS

We have described a human body segmentation method based on background estimation using modified HLS color space. To estimate the position of a human body, the histogram projection profiles are analyzed, and the body region is verified by prior knowledge such as size and shape. Also, thresholding method based on similarity measures between the background and the object image is used. The body contour is extracted by subtraction followed by dilation and erosion. However, only the chroma-key laboratory database is used in this work, although the subjects of the SOTON databases were filmed indoors and outdoors. Therefore, the usefulness of the pre-processing methods, which are developed in this study, is not demonstrated in outdoor applications.

REFERENCES

- [1] D. M. Gavrilu, "The Visual Analysis of Human Movement: A Survey," *Computer Vision and Image Understanding*, **73**(1), pp.82-98, Jan. 1999.
- [2] I. I. Koprinska and S. Carrato, "Temporal Video Segmentation: A Survey," *Signal Processing: Image Communication*, **16**(5), pp.477-500, Jan. 2001.
- [3] T. Horprasert, D. Harwood, and L. S. Davis, "A Statistical Approach for Real-Time Robust Background Subtraction and Shadow Detection," in *Proceedings of IEEE International Conference on Computer Vision: Frame-Rate Workshop*, Kerkyra, Greece, Sep. 1999.
- [4] H. Fujiyoshi and A. J. Lipton, "Real-Time Human Motion Analysis by Skeletonization," in *Proceedings of IEEE Workshop on Application of Computer Vision*, pp.15-21, Princeton, NJ, USA, Oct. 1988.
- [5] A. Neri, S. Colonnese, G. Russo, and P. Talone, "Automatic Moving Object and Background Separation," *Signal Processing*, **66**(2), pp.219-232, Apr.1998.
- [6] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pffinder: Real-Time Tracking of the Human Body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**(7), pp.780-785, Jul. 1997.
- [7] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter, "On a Large Sequence-based Human Gait Database," in *Proceedings of Recent Advances in Soft Computing*, pp.66-71, Nottingham, UK, Dec. 2002.
- [8] G. Gordon, T. Darrell, M. Harville, and J. Woodfill, "Background Estimation and Removal based on Range and Color," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Vol.2, pp.2459-2464, Fort Collins, CO., USA, Jun. 1999.
- [9] C. A. Poynton, "Frequently Asked Questions about Color," http://www.poynton.com/notes/colour_and_gamma/ColorFAQ.txt, 1998.
- [10] J. D. Foley, A. Van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics: Principles and Practice*, Addition-Wesley, 1990.
- [11] T. Q. Chen and Y. Lu, "Color Image Segmentation-An Innovative Approach," *Pattern Recognition*, **35**(2), pp.395-405, Feb. 2002.
- [12] J. H. Yoo, D. S. Hwang, K. Y. Moon, and M.S. Nixon, "Automated Human Recognition by Gait using Neural Network," in *Proceedings of IEEE International Workshops on Image Processing Theory, Tool and Applications*, Sousse, pp.362-367, Tunisia, Nov. 2008.

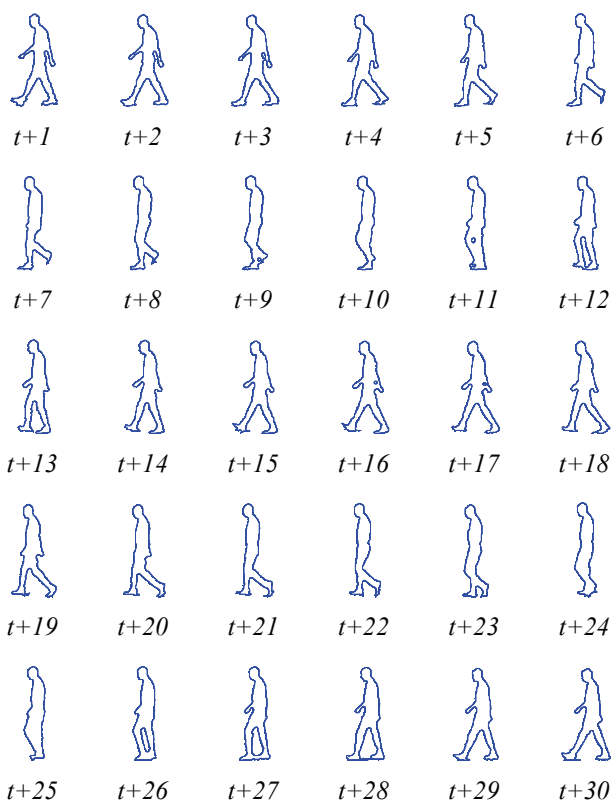


Fig. 4 Extracted Body Contours (Male, Grade A)

The human body contour is extracted with origin of the segmented body region, and the extracted body contours are visually inspected and graded. The average quality levels of the contour data of each subject can be evaluated by comparison with the subject in original images. Here, the quality level is graded as *A* (good), *B* (fair), and *C* (bad) to be 30%, 40%, and 30% of subjects, respectively. The noise, which is related to quality of the contour data, is mostly caused by the shadow and color of shoes at ground. The shoes can appear to change in color due to their reflectance of the walking surface. However, our approach appears to extract the human body contours from image sequences of the SOTON database successfully. In addition, Table I shows the gait recognition results for each group of the different quality levels [12]. As can be seen in the table, the better the quality of results of the pre-processing stage, the better the performance of the recognition rate.

TABLE I
EXPERIMENTAL RESULTS FOR GAIT RECOGNITION

Quality of Gait Silhouette Data (# of Subjects)	# of Feature Vectors		Recognition Rate	
	Train- ing	Test	Training Phase	Recog. Phase
Class A - Good	150	30	100.0%	90.0%
Class B - Fair	150	30	100.0%	83.3%
Class C - Bad	150	30	100.0%	83.3%