Silhouette extraction based on histogram of pixel features for automatic diagnosis of golf swing

Yasunori KOYAMA*, Yoshiaki SHIRAI*, Masahiko UEDA†
*Ritsumeikan University, Shiga, Japan
†SRI R&D Ltd, Kobe, Japan

***Abstract*— We propose a novel method of extracting a human silhouette from a outdoor golf swing movie taken by a mobile phone camera for automatic golf swing diagnosis. At every pixel, we make color and brightness histograms during a swing. Each pixel is classified into three categories:background pixel, human pixel, mixture pixel. Histograms of the "background pixel" have only one peak corresponding to the background, histograms of the "human pixel" have dispersed dispersed distribution corresponding to the various points in the human region, and histograms of the "mixture pixel" have a peak and dispersed distribution. Based on this histogram features, we classify each pixel we segment the swing period into background frames and human frames. Then we extract human silhouette at each frame. Experimental results show the validity of our method.**

I. INTRODUCTION

In order to improve the skill of golf, the golfer may take lessons from an instructor. However, there are many golfers who cannot take lessons because of lack of time or the high cost. To reduce the cost and time, a system has been developed which automatically diagnoses the golfer's swing.

Ueda et al. developed "The automatic diagnosis system of golf swing"[1][2]. This system detects features in indoor swing movie and background image taken by the movie camera which was fixed and diagnoses golf swing on the basis of the original diagnosis algorithm that was based on instructor's knowledge. But there are two problems in the system:

1) The system needs the background image for background subtraction.
2) The golf swing movie must be taken indoors where a good illumination condition is availabe.

These two problems are burden for users.

The automatic diagnosis of outdoor golf swing was studied by Mochizuki et al[3] which employes two cameras to observe the golfer from the back and side directions. The initial human region is estimated by the background subtraction method.

This paper describes a method of extracting a human silhouettes from a series of outdoor golf swing movie taken by the user with a mobile phone camera. Because a user takes a movie, only one camera is used, and the background subtraction can not be applied because the camera may move.

The human and the background regions are segmented based on the variation of the pixel value during a swing. At every pixel, we make color and brightness histograms during a swing. From each pixel at each frame, the pixel is classified into three categories: background pixel, human pixel, and mixtures based on histogram features. Those that can not be clearly classified, we classify them based on the histograms of the neighboring classified pixels.

(c) address (d) top

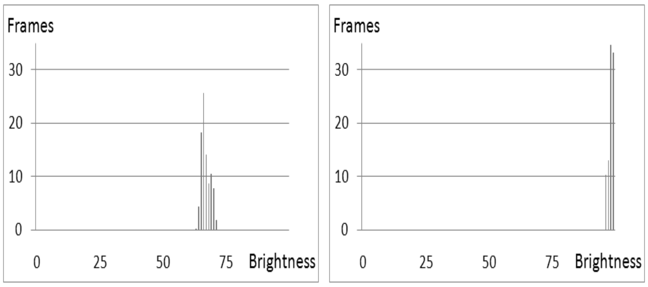
(e) impact (f) dinish

Fig. 1. The part of golf swing movie

II. EXTRACTION OF HUMAN SILHOUETTE

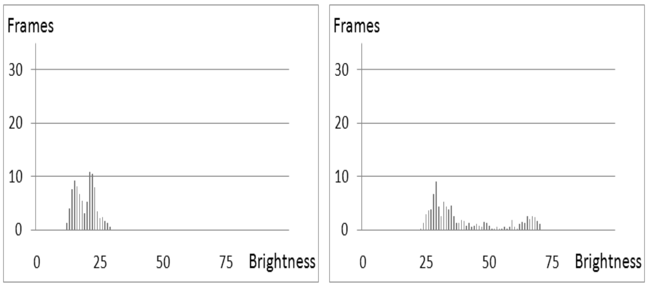
A. Relation Between Pixel Category and the Histogram of the Pixel Properties

During the swing (Fig.1), each pixel in the image may correspond to the background and the human regions. we classify each pixel into three pixel categories: the "background



A B

Fig. 2. Background pixel



C D

Fig. 3. Human pixel

pixel”, the ”human pixel” and the ”mixture pixel”.

- ”background pixel” $\cdots$ only background region during the swing movie.
- ”human pixel” $\cdots$ only human region during the swing movie.
- ”mixure pixel” $\cdots$ both the background and the human region during the swing movie.

Figure.2 shows brightness histograms of the ”background pixel” at A and B in Fig.1. Most of ”background pixel” have only one peak corresponding to the background.

Figure.3 shows brightness histograms of the ”human pixel” at C and D in Fig.1. Most of ”human pixel” have dispersed distribution corresponding to the various points in the human .

Figure.4 shows brightness histograms of the ”mixture pixel” at E and F in Fig.1. Most of ”mixture pixel” have a peak and others due to a combination of background and human regions.

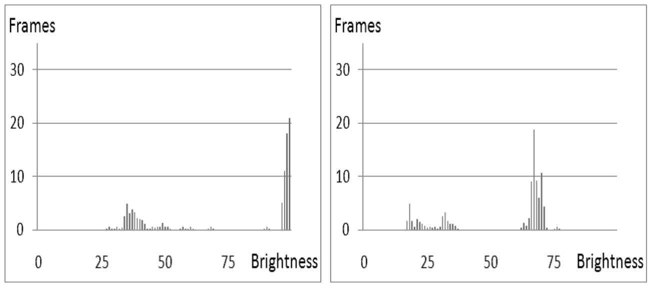
We classify each pixel into those categories based on a histogram of those pixel properties.

B. Pixel Classification

we make color and birghtness histograms at each pixel during the swing. (1) shows a method of making histogaram of brightness ”$\mathbf{h}^B_{x,y}$”, (2) shows a method of making histogram of color ”$\mathbf{h}^C_{x,y}$”.

1)

$$\mathbf{h}^B_{x,y} = (h^B_0, k^B_1, \cdots, k^B_{99})$$



E F

Fig. 4. Mixture pixel

where h^B_i is a frequency of frames that the brightness value has i in the pixel (x, y).

$$i = 99 \times (0.299R + 0.587G + 0.114B)$$

$$0 \le R, G, B \le 1$$

2)

$$\mathbf{h}^C_{x,y} = (h^C_0, \cdots, h^C_{99}, h^C_{100}, \cdots, h^C_{199})$$

where h^C_h (from h^C_0 to h^C_{99}) is a frequency of frames whose the hue has h in the case that saturation s has above 0.15 in the pixel (x, y), and h^C_{i+100} (from h^C_{100} to h^C_{199}) is a frequency of frames whose the brightness has i in the case that saturation s has below 0.15 in the pixel (x, y).

$$s = \frac{(B-R)^2 + (R-G)^2 + (G-B)^2}{2}$$

$$h = \frac{99}{359} \times \begin{cases} \text{if } determin = 0, G > B \\ \quad 90 \\ \text{if } determin = 0, G < B \\ \quad 270 \\ \text{if } determin < 0 \\ \quad \arctan(\sqrt{3}(\frac{G-B}{determin})) + 180 \\ \text{if } determin > 0 \\ \quad (\arctan(\sqrt{3}(\frac{G-B}{determin})) + 360)\%360 \end{cases}$$

$$determin = 2R - G - B$$

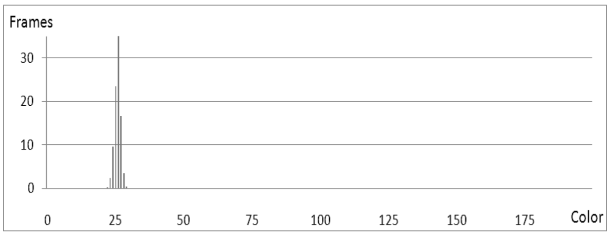
Figure.5 shows color histograms of at A, C and E in Fig.1.

The algorithm to classify each pixel based on histograms is below:

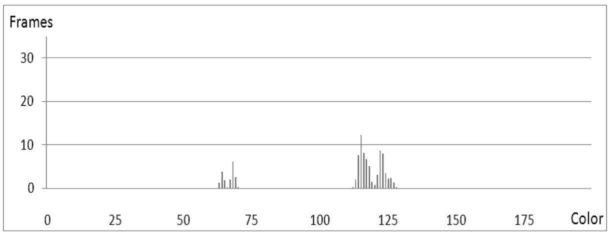
The histogram is classified into two categories:

- background class: The class whose samples more than 60% and are included in interval with lengh 20
- human class: Otherwise

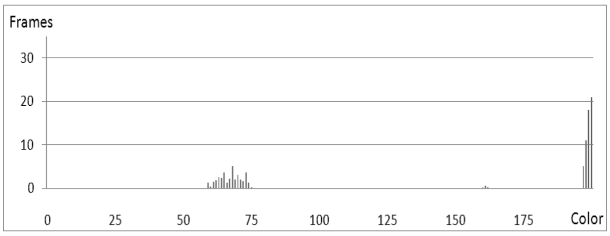
1) Pixels outside of the mask region are classified (Fig.6(a) black pixels).
   The pixel (x, y) whose $\mathbf{h}^B_{x,y}$ and $\mathbf{h}^C_{x,y}$ have only ”background class” is classified into the ”background pixel”.



A: ”background pixel”



C: ”human pixel”



E ”mixture pixel”

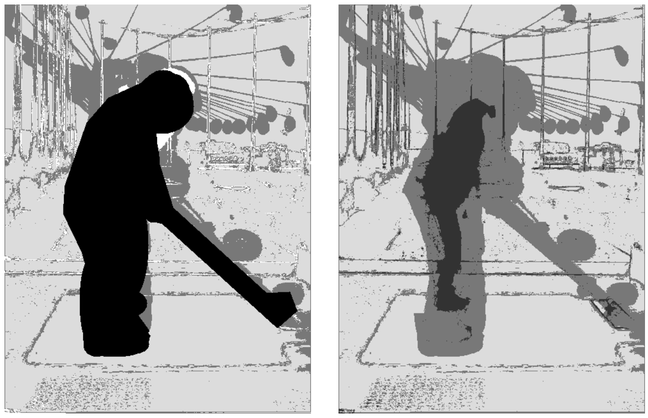
Fig. 5. Color histogram

The pixel (x, y) whose $\mathbf{h}^B_{x,y}$ and $\mathbf{h}^C_{x,y}$ have the ”background class” and the ”human class” is classified into the ”mixture pixel”.
Figure.6(a) shows the ”background pixel” and the ”mixture pixel” besed on (1).

2) We classify the rest pixel (x, y) based on $\mathbf{h}^B_{x+\alpha,y+\beta}$ and $\mathbf{h}^C_{x+\alpha,y+\beta}(-1 \leq \alpha, \beta \leq 1)$. If $\mathbf{h}^B_{x,y}$ and $\mathbf{h}^C_{x,y}$ have the class which is similar to the ”backgraond class” of $\mathbf{h}^B_{x+\alpha,y+\beta}$, $\mathbf{h}^C_{x+\alpha,y+\beta}$, this class is the ”background class” and others are the ”human class” and the pixel (x, y) is the ”mixture pixel”. If $\mathbf{h}^B_{x,y}$ and $\mathbf{h}^C_{x,y}$ which is not similar, all classes are the ”human class” and the pixel (x, y) is the ”human pixel”.

3) We cassify all pixels by ditting (2) with pixels that are not classified.
Figure.6(b) shows ”background pixel”, ”human pixel” and the ”mixture pixel” based on (1),(2).

Based on pixel classification, we extract the human silhouette at each frame.



(a) Classification based on features

(b) Classification based on a neighbor pixel

background pixel:light gray, mixture pixel:gray, human pixel:heavy gray
not clear:white, mask:black

Fig. 6. Pixel classification

Figure.7 shows the human silhouette based on our method. Noises in the background are due to blurring of golf swing movie. We process this human silhouette by the dilation and the erosion and the labeling.

III. CONCLUSION

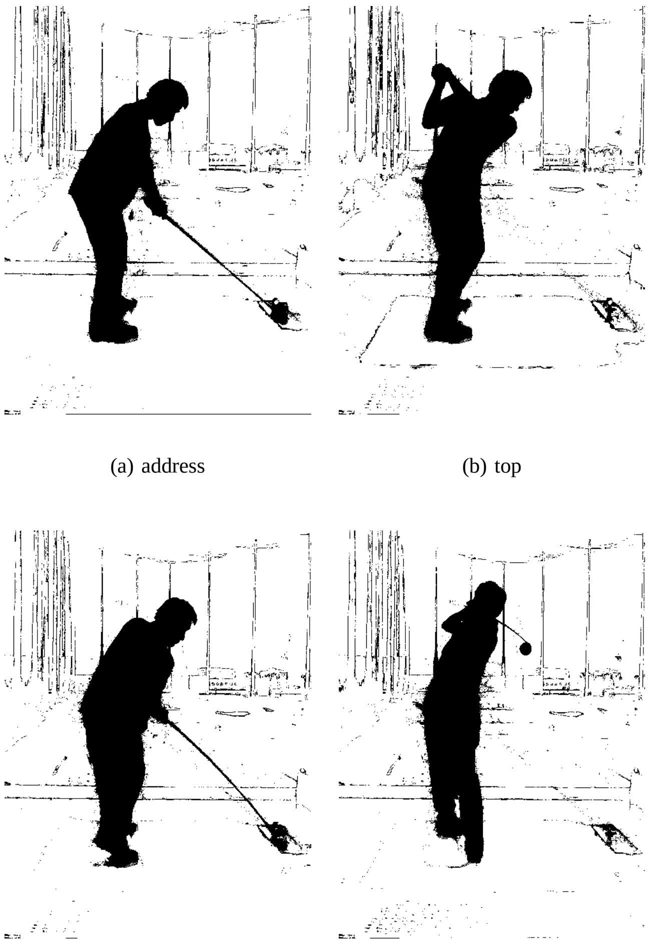
We proposed a method of extracting a human silhouette from a outdoor golf swing movie taken by a mobile phone camera for automatic golf swing diagnosis. Because of no backgraound image, we extracted the human silhouette based on histograms during the swing movie at every pixel. Experimental results Fig.7 shows the validity of our method. We have two future challenges. One is to extract accurate the human silhouette even where the human color resembles the background. The other is to run curve approximation to smooth the human silhouette.

REFERENCES

[1] M. Ueda, N. Shimada, Y. Shirai and M. Oonuki, “Automatic Diagnosis of Golf Swing using a image taken by a TV camera (in Japanese)”, *Dinamic Image Processing and Applications Workshop 2006 Collection of Papers (in Japanese)*, pp.228–232, Mar.2006

[2] M. Ueda, M. Oonuki, Y. Shirai and N. Shimada, “Automatic Diagnosis System of Golf Swing”, *Proc. of The Impact of Technology on Sport 2*, pp.271–276, 2007

[3] M. Mochizuki, M. Ueda, N. Shimada and Y. Shirai, “Feature extraction for Automatic Diagnosis System of Golf Swing (in Japanese)”, *Dinamic Image Processing and Applications Workshop 2007 Collection of Papers (in Japanese)*, pp.11–16, Mar.2007



(a) address (b) top

(c) impact (d) finish

Fig. 7. Human silhouettes