Gender Classification Based on Binary Haar Cascade

Mustafa E. Yildirim, J. S. Park, J. Song, and B. W. Yoon

***Abstract*—Gender classification is a difficult but also an essential task under the researches of pattern recognition. There are several methods and features used for this task such as face, gait, or full body features. One of the most widely used techniques is Haar cascades. Default Haar features based classifiers can only detect pedestrian, free from gender information. In this paper we aimed to learn the gender of the target pedestrians by Haar cascades that are trained gender specific. We trained the classifier with only male and female images as positive and negative respectively. Once a basic pedestrian detection has been made over whole image, second detection is made in ROI (Region of Interest) which is the first detected rectangle. Even though we implemented this idea for only pedestrians in this step, it can be applied to other binary problems.**

***Index Terms*—Adaboost, gender classification, pattern recognition, pedestrian detection.**

I. INTRODUCTION

In the current technology area, object recognition gains importance and interest. An object recognition system finds objects in the real world from an image of the world, using object models which are known a priori. This task is surprisingly difficult. For humans it is easy and fast to recognize and classify objects.

One of the usage areas of the pattern recognition is gender classification. Gender classification is an essential task in today`s world with various types of applications such as surveillance purposes, medical purposes, monitoring applications, and human-computer interaction. In majority of gender classification studies, face features are used. In real-time conditions, where videos are taken by a Closed Circuit Television(CCTV) system, a capturing face with details much enough to extract features can`t be accurate. Reason of it is that CCTV cameras operating for security are mostly located in quite far distance from people.

At present, data collected from various parts of human body such as finger prints, iris, ear, voice, palm prints, facial images, body image as well as pedestrian gait data from video surveillance equipment have been widely applied in gender recognition technology and great achievements have been obtained [1].

An important aspect of recognition is feature selection and extraction. By using proper features, performance and accuracy of the system can be increased. When we decide what features to use, we should consider the ones that are easy to extract. Generally, features applied to gender recognition task should conform to several criteria: uniqueness, performance, collectability, performance acceptability and circumvention [1], [2].

In this paper, we are using Haar-like features to detect pedestrians and classify their gender. Haar classifier is used for face detection because it can detect the desire image very fast. We created different cascade classifiers by using boosting algorithm for male and female faces separately. A cascaded system is employed for this task.

This paper is organized as fallows. In Section II, we briefly mention about commonly used features, Adaboost algorithm and Haar-like features.

Section III, we explain our work in detail. The last but not least, tests and results are given in Section IV.

II. FEATURE EXTRACTION

A. Overview of Commonly Used Features

Images contain unwanted noises such as light, shadow and occlusion. These effects can result a decrease in the system performance. In order to compensate these handicaps and to have a high accuracy rate, we have to make a proper and effective feature extraction. These features can be global or local depending on color, shape, orientation or texture.

1) Edge feature

Edge is a widely used feature in object recognition. Point and line detection are essential in any segmentation problem, edge detection is the most common approach for detecting meaningful discontinuities in gray level so far. We can define edge as the combination of points which create the boundary between to region. Edge feature is robust against background interference.

2) Haar- like feature

Viola and Jones proposed an algorithm [3], called Haar Classifiers for rapid object detection and then applied to the pedestrian detection. With the simple haar-like features which can be calculated efficiently by using integral images and Adaboost classifiers in a cascade structure, their detector has high detection speed [4]. Experiments showed that object detection using Haar-like features can achieve high accuracy at a considerably low cost. Nowadays, Haar-like features are widely used for pedestrian detection and face recognition because it is very discriminative and very easy to calculate [5], [6].

3) HOG Feature

In [7], Dalal and Triggs proposed HOG algorithm. The basic idea is that local object appearance and shape can often be characterized rather well by the distribution of

Manuscript received November 4, 2013; revised January 26, 2014. This work was supported by the Industrial Strategic technology development program, 10045260, Development of Context Awareness Monitoring and Search System Based on High Definition Multi-video funded By the Ministry of Trade, industry & Energy (MI, Korea).

The authors are with Department of Electronics Engineering, Kyungsung University, Busan, Korea (email: meyeren2@ks.ac.kr, jsong@ks.ac.kr, jsipark@ks.ac.kr, bwyoon@ks.ac.kr).

local intensity gradients or edge directions, even without precise knowledge of the corresponding gradient or edge positions [7].

Histogram of oriented gradients (HOG) features and shapelet features can achieve good performance for pedestrian detection, but they are time consuming [4].

4) Texture feature

Robert M. Haralick proposed the texture feature in 1973. The textural features are based on statistics which summarize the relative frequency distribution (which describes how often one gray tone will appear in a specified spatial relationship to another gray tone on the image) [8]. To extract the texture features, some algorithms are used such as Gabor filter, FFT and wavelet. Fig. 1 shows an example of texture feature extraction.

(a) (b)

Fig. 1. (a) Original SAR image, (b) texture features extracted of the original image.

B. Detector Using Haar-Like Features

Face detection works by scanning up an image at different scales and looking for some simple patterns that identify the presence of a face. The core basis for Haar classifier object detection is the Haar-like features. Two or three adjacent rectangular groups with different contrast values create Haar-like feature. The intensity values of pixels in the white and black groups are separately accumulated. The difference between adjacent groups will give light and dark regions. It is essential to add that the concept of lightness and darkness here is relative. Haar-like features, as shown in Fig. 2 are used to detect any object. By adjusting the size of the rectangles, we can scale up and down. This allows features to be used to detect objects of various sizes.

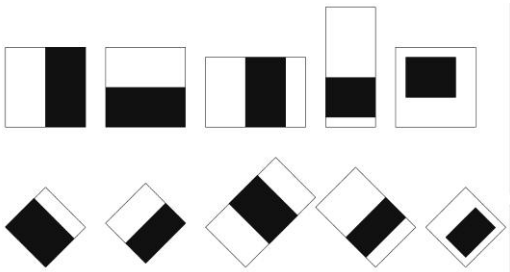


Fig. 2. Commonly used Haar-like features.

C. Adaboost Algorithm

In general, Boosting is a statistical method that uses the original distribution of positive and negative examples to compute simple rules also called weak classifiers and combines them to create a stronger classifier. AdaBoost is most commonly used for binary classification, but with minor modifications it can also deal with multiple classes [9].

In Adaboost algorithm, a weight is assigned for each training sample. These weights indicate the probability of some kind of weak classifier to be selected into the training set [4]. In case of a sample is not classified correctly, during next training set, the probability that will be chosen will increase. Table I shows the boosting algorithms training.

TABLE I: BOOSTING ALGORITHM

Input : $(x_1, y_1), (x_2, y_2)...(x_n, y_n)$

Initialize weights for each element $D_1(i) = 1/n$ where n is the number of examples. For $t = 1,2,..., T$.

1) Train a weak classifier h_t (a simple threshold) using the weights in D_t so that h_t $\{-1,+1\}$.

2) Compute the error associated with the weak classifier:

$$e_t = \sum iD_t(i)\{h_t(x_i) \neq y_i\}$$

3) Compute $\alpha_t = 0.5 \ln\{(1 - e_t)/e_t\}$.

4) Update weight

$D_{t+1} \neq (i) = D_t(i) \exp(-\alpha_t \{h_t(x_i) \neq y_i\})$. [The weights decrease if the element is classified correctly and increase if it is classifies incorrectly.]

5) Normalize the weights $D_{t+1}(i) = D_{t+1}(i) / \sum i D_{t+1}(i)$.

6) After T rounds the final classifier is

$$H(x) = \text{sign}(\sum t\alpha_t h_t(x)).$$

III. PROPOSED ALGORITHM

By using Haar-like features, recognition of gender is very difficult when the capture is done from a far distance. The main motivation of this study comes from this problem. We used a very simple but an effective idea for this task. We applied a cascaded method.

Detecting a female or male face, in an image containing various kinds of objects is difficult and a multi-class problem. Whereas, achieving the same task is easier when our ROI (Region of Interest) is a pedestrian but any other object. In this case, it becomes a binary class problem.

So, our first step was to detect pedestrians only, without the interest of gender recognition. To do this, we used the default pedestrian Haar classifier detector which can be found in OpenCV. Even though this detector cannot guarantee a high accuracy rate, it was enough for us to assist and prove our main study.

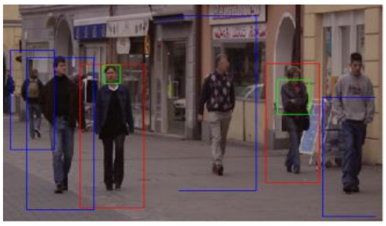
After detection of pedestrian is done, this pedestrian becomes the new ROI. In this ROI, we tried to detect a female face by using the Haar based cascade that we trained.

We trained our own Haar based cascade by using 600 female and 600 male face images. We used frontal face images for training including external features such as make up, hair style, accessories such as earrings and glasses. In a usual XML training, the object we are trying to detect is represented as positive. The negatives can be any object other than the desired object. Whereas, in our study, we used only male and female images as negative and positive respectively. In other words, in a pedestrian image, when our cascade detects a face, it means a female. However, when it does not detect, it means that person is male. This is a simple but efficient approach to gender recognition problem.

IV. Test and Results

In training step we used images from the INRIA person dataset and CVC-CER-01(Computer Vision Center) pedestrian dataset [10]. Training and all tests are done in a 2.60GHz computer by using Microsoft Visual Studio 2010 and OpenCV2.2.

In Fig. 3, blue and red rectangles show pedestrians detections, green ones show face detections. A pedestrian is shown by blue rectangle if face detector cannot find a female face. If it is shown by a red rectangle, it means a female face is detected inside of it. For example in Fig. 3(a), all males are shown by blue rectangle whereas females are shown by red and faces are by green. We can see high accuracy on images (a) to (f) but on (g) and (h), there is one true negative on each.

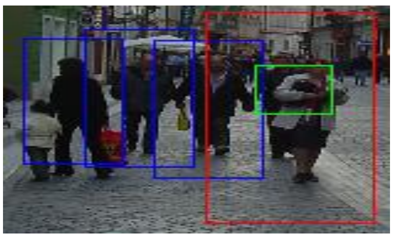


(a)

(b)

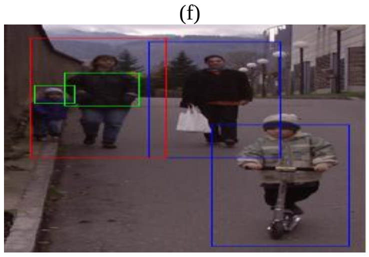
(c)

(d)

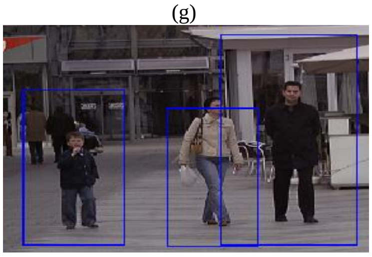


(e)

(f)



(g)



(h)

Fig. 3. Captured results of proposed algorithm.

V. Conclusion

We explained a new approach to gender recognition problem. This study was the first trial of this idea. Although the results are not perfect, they are promising to future studies. Our next step is to create a better Haar cascade and use this method for other multiclass problems.

References

[1] J. Gou, L.Gao, P. Hou, and C. Hu, "Gender recognition based on multiple scale textural feature," presented at the 5th International Congress on Image and Signal Processing, Sichuan, China, October 16-18, 2012.

[2] S. M. E. Hossain and G. Chetty, "Next generation biometric identity verification based on face- gait biometrics," presented at the International Conference on Biomedical Engineering and Technology, Kuala Lumpur, Malaysia, June 17-19, 2011.

[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 511-518.

[4] G. R. Rakate, S. R. Borhade, P. S. Jadhav, and M. S. Shah, "Advanced pedestrian detection system using combination of Haar-like features, adaboost algorithm and edgelet-shapelet," presented at the IEEE Conference on Computational Intelligence and Computing Research, Coimbatore, India, December 18-20, 2012.

[5] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," in *Proc. IEEE International Conference on Image Processing ICIP*, 2002, vol.1, pp. 900-903.

[6] Q. Chen, N. Georganas, and E. Petriu, "Real-time vision-based hand gesture recognition using haar-like features," in *Proc. Instrumentation and Measurement Technology Conference*, 2007, pp. 1-6.

[7] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, vol.1, pp. 886-893.

[8] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textual feature for image classification," *IEEE Trans. On Systems, Man Cybernetics*, vol. 3, no. 3, pp. 610-621, Nov. 1973.

[9] V. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, pp. 137–154, 2004.

[10] D. Geronimo, A. D. Sappa, A. Lopez, and D. Ponsa, "Adaptive image sampling and windows classification for on-board pedestrian detection," in *Proc. the International Conference on Computer Vision Systems*, Bielefeld, Germany, March, 2007.

Mustafa E. Yildirim received his B.S. degree on electrical engineering from Bahcesehir University, Istanbul, Turkey in 2008, M.S. degree on the same major from Kyungsung University, Busan, South Korea in 2010.

Currently, he is a Ph.D candidate on electronics engineering, at Kyungsung University. His research interests are pattern recognition, object tracking, signal processing.

Jang S. Park received the B.S., the M.S. and the Ph.D. degrees in the electronics Eng. from Pusan National University, Korea in 1992, 1994 and 1999, respectively. He joined the Kyungsung University in Korea as professor in the Electronics Engineering Department since March 2011.

Jong K. Song received the B.S., the M.S. and the Ph.D. degrees in the electronics Eng. from Pusan National University, Korea in 1989. He received the M.S. and the Ph.D. degree in the Electrical and Electronics Eng. From KAIST in 1994 and 1999, respectively. He was a researcher at the SK Telecom from 1995 to 1997. He joined the Kyungsung University in Korea as professor in the Electronics Engineering Department since March 1997.

Byung W. Yoon received the B.S, M.S. and Ph.D degrees in electronic engineering from Busan National University in 1987, 1989, 1992 respectively.

He was a researcher at the ETRI from 1992 to 1995. He joined the Kyungsung University in Korea as professor in the Electronics Engineering Department since March 1995. His research interests include the areas of VLSI design, underwater communication.