MVA2005 IAPR Conference on Machine VIsion Applications, May 16-18, 2005 Tsukuba Science City, Japan
13-15

Detecting Faces in Low-resolution Images

Shinji HAYASHI
Tokyo Institute of Technology
R2-52, 4259 Nagatsuta, Midori-ku,
Yokohama, 226-8503 JAPAN
hayashi@isl.titech.ac.jp

Osamu HASEGHAWA
Tokyo Institute of Technology
R2-52, 4259 Nagatsuta, Midori-ku,
Yokohama, 226-8503 JAPAN
hasegawa@isl.titech.ac.jp

Abstract

Face detection is one of the hot research topics in Computer Vision, and greatly progressed in past decade. However, as far as we know, face detection in low-resolution images has not been studied (most system detects faces bigger than 20×20 or 24×24 pixels). A conventional AdaBoost based face detection method (Viola & Jones) can detect merely 32% of faces in 1/4 resolution MIT+CMU frontal face test set. In this paper, we propose a new face detection method for low-resolution images by combined use of two classifiers: one classifier detects faces and the other detects upper-bodies. These classifiers are applied to magnified low-resolution images. The combination of classifiers is realized by using a neural network. As the result, our method achieved 83% of the face detection rate for the 1/4 resolution MIT+CMU test set.

1 Introduction

In recent years, many methods for detecting faces in general scenes are proposed [1, 2, 3, 4]. Those methods work efficiently under various conditions such as illumination fluctuation and containing multiple face directions. However, to date, detection of faces in low-resolution images has not been explicitly studied as far as we know.

In this paper, we propose a new method of face detection for low-resolution images. We consider that the proposed method can be applied to many applications such as face detection in surveillance images that often include distant face images.

2 Conventional method

Recently, many methods for face detection are proposed. Especially, AdaBoost based face classifier by Viola [4] is widely used in face detection research because of its speed and accuracy. AdaBoost based face classifier is thought as one of the standard methods for face detection. Therefore, we use AdaBoost based face classifier for our research. In this section, we show the result of AdaBoost based face classifier's application to low-resolution images.

To evaluate detection rate of conventional method for low-resolution images, we trained and applied

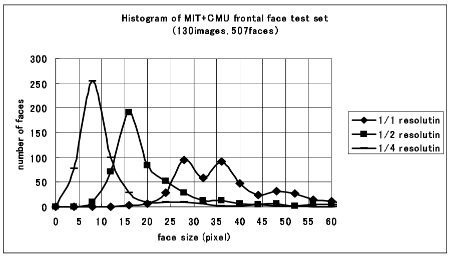

Figure 1: Histogram of face size in MIT+CMU face set

AdaBoost-based classifiers to three kind of resolution of MIT+CMU frontal face test set (130 images including 507 faces): those sets are original size images, 1/2 resolution images, and 1/4 resolution images. The 1/2 and 1/4 resolution images are made by scaling down of original resolution images by the 'bicubic' method.

In the experiment, we firstly trained three classifiers by using three sets of face images (24×24,12×12,and 6×6 pixels). Then we applied (1) 24×24 classifier to original images, (2) 12×12 classifier to 1/2 resolution images, and (3) 6×6 classifier to 1/4 resolution images respectively. Here, we would like to emphasize that we cannot apply the 24×24 classifier to 1/4 resolution images because it cannot detect 6×6 pixel faces in the images. Fig.1 is the histogram of the size of faces contained in three resolution levels of MIT+CMU frontal face test set. [1]

Size of training data is the minimum size of detectable face because in face detection process, input image pyramid is made by scaling down [2]. Furthermore, faces bigger than 24×24 pixel are hardly contained in 1/4 resolution MIT+CMU frontal face test set, so it is meaningless to apply 24×24 face classifier to 1/4 resolution set.

Fig.2 shows averaged faces of the three resolution training sets (24×24,12×12,6×6 pixels).

Fig.3 shows ROC curves for the three kinds of resolution images obtained by using conventional AdaBoost method. At the point of 100 false positives, face detection rate falls from 89% to 32% as resolution

[1]The size of a face is defined as 2.4 times of the interval of both eyes.

24×24 12×12 6×6

Figure 2: Averaged faces used for learning of classifiers

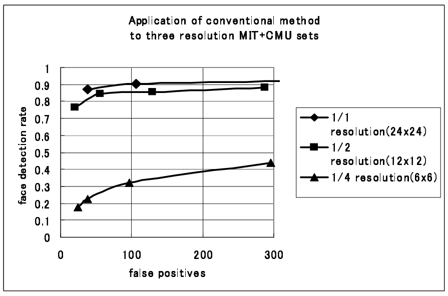

Figure 3: ROC curves obtained by applying conventional AdaBoost method to MIT+CMU frontal face test set

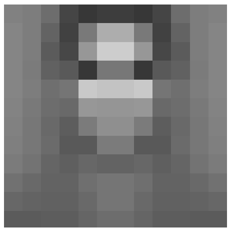

Figure 4: 12×12 pixel averaged upper-body images

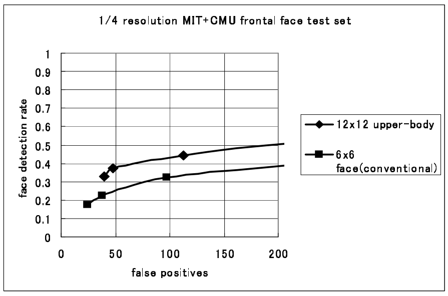

Figure 5: effect of upper-body classifiers

falls from 1/1 to 1/4. This result shows that we cannot obtain sufficient detection rate for low-resolution images by simply applying conventional method.

3 Proposed method

As shown in Section 2, face detection by using conventional method in low-resolution images is difficult. Therefore, we use 1/4 resolution MIT+CMU frontal face test set as evaluation images of detecting faces in low-resolution images.

In this paper, we propose a new face detection method for low-resolution images. Our method consists of three techniques. 1. To use a classifier trained with upper-body images instead of face images. 2. To magnify an input image. 3. To combine two classifiers. One is face classifier, the other is upper-body classifier. We show each details below.

3.1 Using upper-body images

We trained a classifier using 12×12 pixel upper-body images instead of 6×6 pixel face images as training data. The average image is shown in Fig.4. The size of the face in an upper-body image is 6×6 pixel.

The idea of using upper-body classifier is based on Torralba's psychological experiment [5]. Their result indicates that a man can recognize a face in a low-resolution image well when using an upper-body image than simple face image.

Then we applied 12×12 upper-body classifier to 1/4 resolution MIT+CMU frontal face test set. Fig.5 is the result. For comparison, result of 6×6 face classifier applied to 1/4 resolution MIT+CMU frontal face test set is plotted too. (The data of 6×6 face classifier is the same as Fig.3.)

At the point of 100 false positives, 6×6 face classifier detects merely 32% of faces, but 12×12 upper-body classifier detects 42% of faces in the 1/4 resolution MIT+CMU frontal face test set. It can be said that 12×12 upper-body classifier can detect faces well as compared with 6×6 face classifier.

From this result, we thought to use only upper-body classifier. However, by carefully seeing faces of two classifiers detected, it turns out that two classifiers complement each other. Example is Fig.6. (i.e. there are faces that only one classifier detected.) Therefore, we use not only upper-body classifier, but also face classifier. In section 3.3, we will try to combine these two classifiers to improve face detection rate more.

3.2 Magnifying input images

Face detection rate for 1/4 resolution MIT+CMU frontal face test set was improved by using 12×12 upper-body classifier. However, we thought 42% of

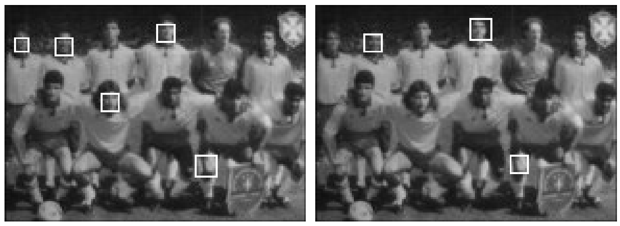

Figure 6: left:6×6 face classifier's result. right:12×12 upper-body classifier's result

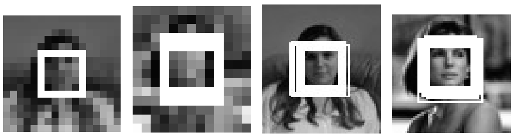

Figure 7: difference of the number of "face coordinates candidates". :left two images contains 6×6 pixel faces, right two images contains 24×24 pixel faces,

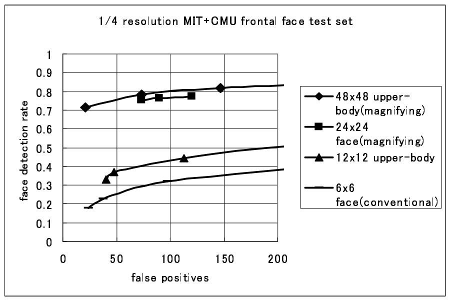

Figure 8: effect of magnifying input images

face detection rate is still low.

In face detection, two or more "face coordinates candidates" usually occurs around one face. This is because a classifier judges as a face, even if position and size changes somewhat. Two or more detection coordinates generated around one face are merged, and, finally turn into one face detection coordinates to one face.

Fig.7 is detected results of 24×24 and 6×6 pixel faces. "Face coordinates candidates" in Fig.7 are not merged. There are more "face coordinates candidates" in 24×24 pixel face than 6×6 pixel face. We counted number of "face coordinates candidates" by applying face classifier to 100 of 24×24 face images and 6×6 pixel face images respectively. For 24×24 pixel face images, average number of "face coordinates candidates" is 20. For 6×6 pixel face images, average number of face coordinates candidates is 2. This difference is the difference of robustness for position and size change. We thought this is one of the reason why face detection rate for 1/4 resolution MIT+CMU frontal face test set is so low.

So we magnify low-resolution input image and detect faces by 24×24 face classifier. We magnified 1/4 resolution MIT+CMU frontal face test set by bicubic to magnify smoothly, and applied 24×24 face classifier and 48×48 upper-body classifier. In magnification, we use 4 as a scaling factor. The result is Fig.8. For comparison, the result before using magnifying is plotted. As for face classifier, face detection rate is improved from 32% to 78% at the point of 100 false positives. As for upper-body classifier, face detection rate is improved from 42% to 80% at the point of 100 false positives.

It turned out that face classifier and upper-body classifier can improve face detection rate by magnifying input images. In next section, we improve face detection rate furthermore, by combine use of two detectors.

3.3 Combination of two classifiers

In this section, face detection rate is improved by combination use of face classifier and upper-body classifier.

Now, since there are two classifiers, two face-likeness are calculated about the image judged to be a face.

Making a final judgment based on this information means determining the domain of face and nonface in the two dimension plane which takes face-likeness which two classifiers output on both axes. In our research, this is realized by a neural network. Face-likeness is defined as below. $h_i(x)$ is a weak leaner and α_i is a weight of the weak learner. k is the number of "face coordinates candidates" and i is the number of weak learners.

$$z_k = \sum_{weak\ learners} \alpha_i h_i(x) \tag{1}$$

Two or more "face coordinates candidates" generated around one face is merged, and one coordinates are made to correspond to one face finally in detection process.

When merging cadidate locations,

$$Z = \sum_k z_k \tag{2}$$

is calculated. This corresponds to a "face coordinates" which was made by merging "face coordinates candidates". This value is thought as "face-likeness".

Now, two detectors are applied independently to an input images and the 2D vector Z is obtained about a image finally detected by merging the result further. Final judgement is made by a neural network whose input is this 2D vector Z.

As for training data of a neural network, about 10000 images (these images contain both faces and nonfaces) with a 2D vector Z were obtained by applying two detectors to 6570 general images containing faces. We trained a neural network to separate these images appropriately. Number of neurons of input layer, hidden layer, and output layer are 2, 8, and 1 respectively.

Here, we summarize briefly procedure of proposed method.

- Magnifying an input low-resolution image by a factor of 4, by using 'bicubic' method.

- Applying two classifiers to the magnified image. The first classifier (trained by 24×24 pixel size face images) is for detecting faces, and the second classifier (trained by 48x48 pixel size upper-body images) is for detecting upper-bodies.

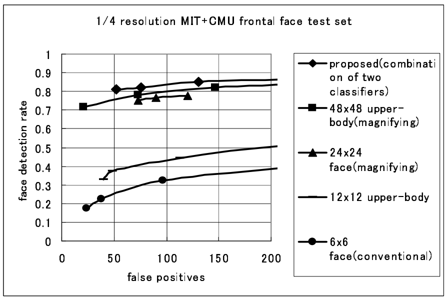

Figure 9: effect of combination of two classifiers

- Combining outputs of two classifiers using a three-layer neural network.
- Determining sizes and positions of faces in the input low-resolution image.

We applied the proposed method to the 1/4 resolution MIT+CMU frontal face test set. Fig.9 shows the experimental result. As shown in the Fig.9, the conventional method detects only 32% faces at 100 false positives, the proposed method can detect 83% faces at the same false positives. We show the result images of 1/4 resolution MIT+CMU frontal face test set.

Fig.10 shows faces detected in 1/4 resolution MIT+CMU frontal face test set by the proposed method. As shown in Fig.10, it is very difficult to find faces if we see only face regions.

Fig.11 shows the results of the experiment. Three images are (from the top to bottom) the original resolution image, 1/4 resolution image detected by conventional method and 1/4 resolution image detected by the proposed method.

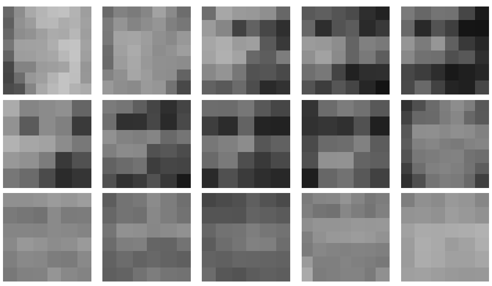

Figure 10: Faces in 1/4 resolution MIT+CMU set detected by proposed method

4 Conclusion and future work

We proposed a new method consists of three techniques for face detection in low-resolution images. And we showed that conventional method can detect only 32% of faces in 1/4 resolution MIT+CMU frontal face test set, but our proposed method can detect 83% of faces in MIT+CMU set.

In this paper, we use 1/4 resolution MIT+CMU frontal face test set as evaluation images. As shown in Fig.1, face of various sizes exists in this set. The effect of proposed method may differ according to the size of a face. Systematic evaluation according to the size of a face is a future work.

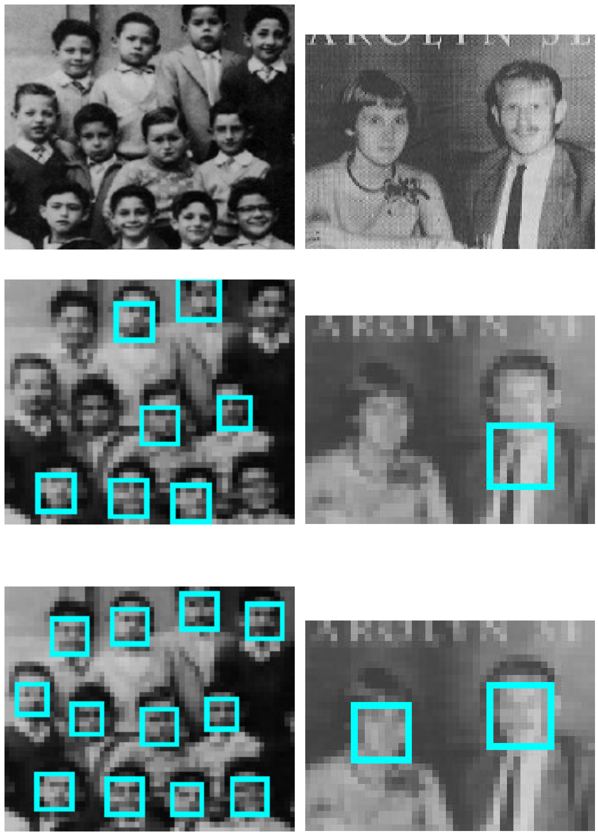

Figure 11: Experimental results (Top: original image, Mid: 1/4 resolution images and detected results by conventional method, Bottom: 1/4 resolution images and detected results by proposed method)

References

[1] M.H. Yang, D. Kriegman and N. Ahuja. “Detecting faces in images.A survey,” PAMI,24(1)34-58,January 2002.

[2] Henry A Rowley, Shumeet Baluja, and Takeo Kanade. “Neural Network-Based Face Detection,” In IEEE PAMI,volume 20, pages 23-38, 1998

[3] H. Schneiderman. “Feature-Centric Evaluation for Efficient Cascaded Object Detection,” CVPR,2004.

[4] P. Viola and M. Jones. “Rapid object detection using a boosted cascade of simple features,” CVPR,2001.

[5] A. Torralba and P. Shina. “Detecting Faces in Impoverished Images,” AI Memo 2001-028, CBCL Memo 208,2001