ICSP2012 Proceedings

A Fast Face Detection Method for JPEG Image

TIAN Qing, ZHAO Shiwei
The First Research Institute of Ministry of Public Security, Beijing, China
tianqing@263.net; zsw.zhao@gmail.com

Abstract—**Face detection is a key technology of face recognition, in this paper, a fast face detection method based on JPEG image is proposed. According to analyzing the feature of face image, the combination of reconstruction of low resolution image in the compressed domain and face detection in the pixel domain method is proposed. For JPEG image, we propose the reconstruction method of low resolution image by using few DCT coefficients. We compare our method with conventional face detection in the pixel domain. The testing results show our method has high face detection precision while greatly improve the detection speed.**

Keywords- compressed domai, face detection; face recognition

I. INTRODUCTION

Face recognition technology is playing an increasingly important role in the social field of public safety. This technology has been successfully applied to exit and entry administration of public security business areas. With the increase of the face image data, how to quickly detect face in images has become an urgent need to address the problem.

In recent years, many face detection methods have been proposed. These methods can be generally classified into two categories: in the pixel domain and in the compressed domain. Due to the face recognition generally require high resolution, therefore, these face images are usually stored and transmitted in JPEG. Complete decompression should be the first step for the face detection methods in the pixel domain, which not only results in computation and storage burden on the system, but also greatly increases the processing time required.

Face detection methods in the compressed domain perform on the fully compressed or partly decompressed data, saving time for decoding. However, in order to detect multi-scale faces, scaling images is required in compressed domain. Although the fast image scaling method in compressed domain, face detection speed is affected. And because of block encoding in JPEG, the accuracy of face detection methods is reduced.

Considering the drawbacks in above schemes, a novel fast face detection method is presented in this paper. This method combines the advantages of both face detection methods in compressed domain and in pixel domain. Firstly, the low-resolution images are reconstructed in the compressed domain. Afterward, faces are detected in the pixel domain in the low-resolution images. Reconstructing low-resolution images in the compressed domain is a key of this method, and its detailed derivation.

The remainder of this paper is organized as follows. The method is introduced in detail in Sec.II. In Sec.III, the experimental results and analysis are described. Conclusions of the method are given in Section IV.

II. OUR METHOD

A. The Scheme of Our Method

The purpose of the face detection in the compressed domain is to improve the detection speed, but the scaling and block encoding of compressed images, the detection speed is low, besides, the accuracy of face detection is degreased. Recently, progress both in speed and precision are get in pixel domain face detection[8, 9]. The most important advantage for face detection in the compressed domain is no need for Discrete Cosine Transform (DCT), which can save the time for decoding. By combining the feature in the compressed domain with the face detection in the pixel domain, we propose a fast face detection method of JPEG image. It recovers a low resolution image stream without DCT and using the face detection method in the pixel domain.

The proposed method shows in Figure 1:

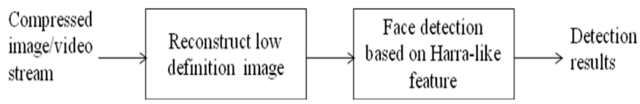

Figure 1. Flowchart of fast face detection

As shows in Figure 1. , the proposed method is divided into 2 steps: first step is to reconstruct the low resolution image; second step is to do face detection by using pixel domain method. The reasons to detect face in low resolution image is as following: 1) For face recognition application, normally high resolution images are required, there are no less than 60-pixel between two eyes; for these kinds of images, even decrease the image resolution, the face area is still big enough for face detection. 2) Reconstruction speed can be improved by recovering low resolution stream, which can avoid DCT. 3) After finish reconstruction low resolution image, the robustness and high speed pixel domain face detection can be adopted.

B. Fast low-resolution image reconstruction

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

Normally, the size of the images for face recognition is big, in the paper, only reconstruction of 1/8×1/8 and 1/4×1/4 are studied.

JPEG using YCbCr color space, first the image is divided into 8×8 image block, then DCT, quantification and entropy coding to finish the image compression. It is unstructured DCT data after entropy coding, which is difficult to be used. So in the paper, the low resolution image reconstruction is based on the DCT coefficients after entropy decoding.

Fast 1/8×1/8 low-resolution image reconstruction

For 8 × 8 image block, the average value of the whole image block can be treated as the pixel value of the 1/8×1/8 low-resolution image. After DCT, DC coefficients include the energy of the image, so DC coefficients can be used for 1/8×1/8 low-resolution image.

Fast 1/4×1/4 low-resolution image reconstruction

Suppose x is matrix of original 8×8 image block, xi,j (0≦i,j ≦7) is the gray value of (i,j), S is a 8×8 sparse matrix, the value of upper left corner is S00,S01,S10,S11, the other element of the matrix is 0, as show in (1).

$$\boldsymbol{S}=\begin{bmatrix} S_{00} & S_{01} & 0 & 0 & 0 & 0 & 0 & 0 \\ S_{10} & S_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{1}$$

The method to reconstruct 1/4×1/4 low resolution image shows in (2), the value of upper left corner 2×2 pixels is the low resolution image.

$$\mathrm{S} = \mathrm{HxH^T} \tag{2}$$

Here, transformation matrix H is a sparse matrix as shows in (3):

$$\boldsymbol{H}=\begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3}$$

For 8×8 image block, two dimensional DCT and inverse DCT is (4) and (5) respectively.

$$\boldsymbol{X} = \boldsymbol{D}\boldsymbol{x}\boldsymbol{D}^T \tag{4}$$

$$\boldsymbol{x} = \boldsymbol{D}^T\boldsymbol{X}\boldsymbol{D} \tag{5}$$

Here X denotes a coefficients matrix, which is DCT output of a 8×8 pixel matrix x, the elements of the matrix is Xi,j, D is transformation matrix as shows in (6), here $\mathrm{C}_\mathrm{N}^{\tau} = \cos\left(\frac{\tau\pi}{\mathrm{N}}\right)$.

$$\boldsymbol{D}=\frac{1}{2}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ C_{16}^{1} & C_{16}^{3} & C_{16}^{5} & C_{16}^{7} & C_{16}^{9} & C_{16}^{11} & C_{16}^{13} & C_{16}^{15} \\ C_{16}^{2} & C_{16}^{6} & C_{16}^{10} & C_{16}^{14} & C_{16}^{18} & C_{16}^{22} & C_{16}^{26} & C_{16}^{30} \\ C_{16}^{3} & C_{16}^{9} & C_{16}^{15} & C_{16}^{21} & C_{16}^{27} & C_{16}^{33} & C_{16}^{39} & C_{16}^{45} \\ C_{16}^{4} & C_{16}^{12} & C_{16}^{20} & C_{16}^{28} & C_{16}^{36} & C_{16}^{44} & C_{16}^{52} & C_{16}^{60} \\ C_{16}^{5} & C_{16}^{15} & C_{16}^{25} & C_{16}^{35} & C_{16}^{45} & C_{16}^{55} & C_{16}^{65} & C_{16}^{75} \\ C_{16}^{6} & C_{16}^{18} & C_{16}^{30} & C_{16}^{42} & C_{16}^{54} & C_{16}^{66} & C_{16}^{78} & C_{16}^{90} \\ C_{16}^{7} & C_{16}^{21} & C_{16}^{35} & C_{16}^{49} & C_{16}^{63} & C_{16}^{77} & C_{16}^{91} & C_{16}^{105} \end{bmatrix} \tag{6}$$

Put (5) into (2), we can get:

$$\boldsymbol{S} = \boldsymbol{H}\boldsymbol{D}^T\boldsymbol{X}\boldsymbol{D}\boldsymbol{H}^T = \boldsymbol{H}\boldsymbol{D}^T\boldsymbol{X}\left(\boldsymbol{H}\boldsymbol{D}^T\right)^T \tag{7}$$

Here we define$\mathrm{C} = \mathrm{HD^T}$, (7) will be:

$$\boldsymbol{S} = \boldsymbol{C}\boldsymbol{X}\boldsymbol{C}^T \tag{8}$$

By calculation, C is a sparse matrix as shows in (9):

$$\boldsymbol{C}=\frac{1}{2}\begin{bmatrix} 1/\sqrt{2} & C_4^1C_8^1C_{16}^1 & 0 & C_4^3C_8^3C_{16}^3 & 0 & C_4^5C_8^5C_{16}^5 & 0 & C_4^7C_8^7C_{16}^7 \\ 1/\sqrt{2} & -C_4^1C_8^1C_{16}^1 & 0 & -C_4^3C_8^3C_{16}^3 & 0 & -C_4^5C_8^5C_{16}^5 & 0 & -C_4^7C_8^7C_{16}^7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{9}$$

When the image block is compressed, after DCT and quantification, most high frequency component become 0, to simplify the calculation, we only reserve the first 4 DCT coefficients after zigzag sorting, then (7) will be:

$$\boldsymbol{S} = \boldsymbol{C}\boldsymbol{X}'\boldsymbol{C}^T \tag{10}$$

Here X is the DCT coefficient after setting the high frequency component to 0, as shows in (11):

$$\boldsymbol{X}'=\begin{bmatrix} X_{0,0} & X_{0,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ X_{1,0} & X_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{11}$$

Put (9) and (11) into (10), keep the 4 element upper left corner to form a 2×2 pixel matrix S1/4×1/4, the matrix is the 1/4×1/4 low-resolution image of original image, the calculation method is (12):

$$S_{1/4\times1/4}=\begin{bmatrix}1/\sqrt{8} & \frac{1}{2}C_4^1C_8^1C_{16}^1\\ 1/\sqrt{8} & -\frac{1}{2}C_4^1C_8^1C_{16}^1\end{bmatrix}\begin{bmatrix}X_{0,0} & X_{0,1}\\ X_{1,0} & X_{1,1}\end{bmatrix}\begin{bmatrix}1/\sqrt{8} & 1/\sqrt{8}\\ \frac{1}{2}C_4^1C_8^1C_{16}^1 & -\frac{1}{2}C_4^1C_8^1C_{16}^1\end{bmatrix} \quad (12)$$

Assume

$$\begin{cases} w_0 = 1/\sqrt{8} \\ w_1 = \frac{1}{2}C_4^1C_8^1C_{16}^1 \end{cases} \quad (13)$$

then (12) can be expressed as:

$$\begin{bmatrix}s_{0,0} & s_{0,1}\\ s_{1,0} & s_{1,1}\end{bmatrix}=\begin{bmatrix}w_0 & w_1\\ w_0 & -w_1\end{bmatrix}\begin{bmatrix}X_{0,0} & X_{0,1}\\ X_{1,0} & X_{1,1}\end{bmatrix}\begin{bmatrix}w_0 & w_0\\ w_1 & -w_1\end{bmatrix} \quad (14)$$

After expansion, (14) becomes:

$$\begin{cases} S_{0,0}=(w_0w_0X_{0,0}+w_0w_1X_{0,1})+ \\ \qquad (w_0w_1X_{1,0}+w_1w_1X_{1,1}); \\ S_{0,1}=(w_0w_0X_{0,0}-w_0w_1X_{0,0})+ \\ \qquad (w_0w_1X_{1,0}-w_1w_1X_{1,1}); \\ S_{1,0}=(w_0w_0X_{0,0}+w_0w_1X_{0,1})- \\ \qquad (w_0w_1X_{1,0}+w_1w_1X_{1,1}); \\ S_{1,1}=(w_0w_0X_{0,0}-w_0w_1X_{0,1})- \\ \qquad (w_0w_1X_{1,0}-w_1w_1X_{1,1}); \end{cases} \quad (15)$$

In the decompressing of image, DCT coefficients after entropy decoding need inverse quantization before substituted into (15). In order to further reduce the calculation complex, we can combine the quantization coefficients with product factor of DCT coefficients. Assume the first 4 quantization coefficients for image compress are Q0,0, Q0,1, Q1,0, Q1,1, DCT coefficient before inverse quantization is $X_{i,j}^{Q}$, according to (15), we can get the method to generate 1/4×1/4 low-resolution image based on the DCT coefficients from entropy decoding, as shows in Fig. 2.

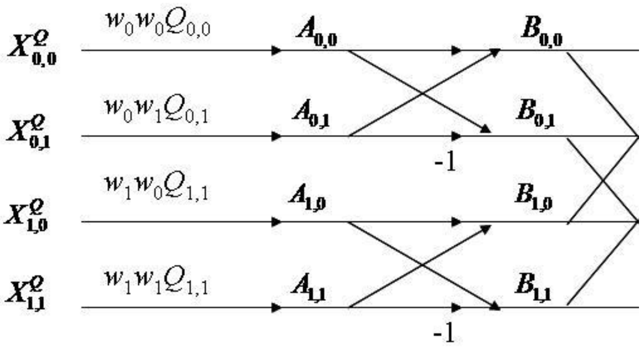

Figure 2. Fast reconstruction method of 1/4×1/4 low-resolution image

According to Figure 2. , using the our method to reconstruction 1/4×1/4 low-resolution image from DCT stream, for each pixel of original image, only 0.0625 multiplication and 0.125 addition are required.

C. Face detection in low-resolution image

After the low-resolution image is reconstruction by the DCT coefficients, we can use pixel domain detection method for face detection, so as to increase the detection speed and guarantee the detection precision. Basically, any pixel domain detection method can be used. In Reference [9], AdaBoost face detection method has better performance both in speed and precision, in this paper, AdaBoost method is used for face detection after low-resolution image reconstruction.

III. THE EXPERIMENTAL RESULTS AND ANALYSIS

Based on the face recognition application background, we choose 1000 images with 1600 × 1200 for experiment. Detection rate (correctly detection face/total face) and false detection rate (false detection face/total detected face) are two key parameters for algorithm evaluation. The average detection time for each image is used for detection speed evaluation.

The experimental results show in Table 1. Here LR_FD8, LR_FD4 are doing the face detection in the mentioned 1/8× 1/8 and 1/4×1/4 low-resolution image respectively. Reference [9] is used as the pixel domain face detection method.

TABLE I. EXPERIMENTAL RESULTS COMPARISONS OF THE PURPOSED METHOD WITH REFERENCE[9]

Detection Algorithm	Detection rate	False detection rate	Average detection time(ms)
LR_FD8	86.3%	0.4%	88
LR_FD4	98.6%	3.6%	287
Reference [9]	100%	70%	4185

Note: hardware platform for experiment is PC with Intel Core i7-2720QM (2.2GHZ).

For detection rate, we can find LR-FD4 has the similar performance with Reference [9], while the false detection rate is much lower. The main reasons are: 1) for face recognition application, since the face area is big, decree the image resolution has little affection to face detection. 2) Decreasing the resolution of image reduce the parts which look like face. From the detection ratio and false detection rate, we can get a conclusion: decreasing the image resolution for face recognition can adapt to the fast face detection requirement.

For average detection time, we can find, by comparing conventional pixel domain face detection, our method has much higher detection speed. By using the 1/4×1/4 low-resolution image, it is 14.56 times faster; by using 1/8×1/8 low-resolution image, it is 47.56 times faster. One reason is the high efficiency method to reconstruct the low-resolution image from compress stream, the other reason is low-resolution image has few windows to be detected, which further improve the detection speed.

In order to have a visualized understanding, Figure 3. shows part of LR_FD4 detection results (the detection results is marked in red block).

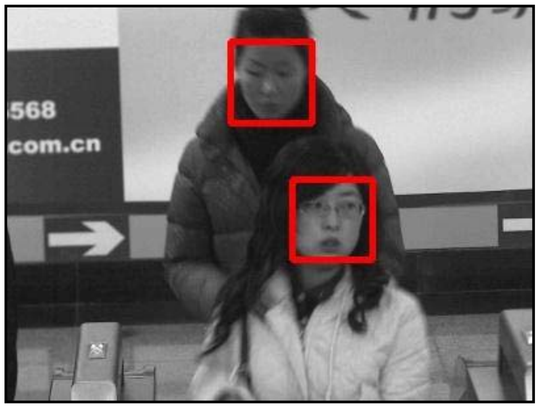

Figure 3. Some face detection experiment results

CONCLUSION

In this paper, a novel fast face detection method in JPEG is proposed. Firstly fewer DCT coefficients are adopted to reconstruct low-resolution image. And then faces are detected in low-resolution image. Experimental results show that, compared with the method in pixel domain, the proposed method achieves higher detection speed and closely detection accuracy.

REFERENCES

[1] M. H. Yang, D. J. Kriegman, N. Ahuja. Detecting faces in images: a survey. IEEE Transactions on Pattern Analysis and Machine intelligence, 2002,24(1): 34-58

[2] 刘党辉, 沈兰荪, Kin-Man Lam. 人脸检测研究进展. 计算机工程与应用, 2003, 28: 5-9.

[3] 李晓华. 小波压缩域图像检索技术的初步研究. 北京工业大学博士论文. 2004.

[4] H. Wang, S. F. Change. A highly efficient system for automatic face region detection in mpeg video. IEEE Transactions on Circuits and Systems for Video Technology, 1997,7(4): 615-628.

[5] T. S. Chua, Y. Zhao, M. S. Kankanhalli. Detection of human faces in a compressed domain for video stratification. The Visual Computer, 2002,18(2): 121-133.

[6] H. Luo, A. Eleftberiadis. On face detection in the compressed domain. In Proceedings of the Eight ACM International Conference on Multimedia, 2000: 285-294.

[7] 李晓光, 李晓华, 沈兰荪. 一种基于多级梯度能量特征的 DCT 压缩域人脸检测算法. 电子学报, 2005,33(12): 2170-2173.

[8] P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, USA, 2001: 511-518.

[9] L. Rainer, K. Alexander, P. Vadim. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In Proceedings of the 25th Pattern Recognition Symposium, Magdeburg, Germany, 2003: 297-304.