Effect of Pretreatment methods on Face Detection in Video Images

Beibei Zeng$^{1,2}$, Ruian Liu$^{1,2,*}$, Nan Liu$^{1,2}$ and Ninghao Yin$^{1,2}$

$^1$Tianjin Key laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
$^2$College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

*Corresponding author e-mail: ruianliu@sina.com

Abstract. At present, face detection technology in video has attracted more and more attention, and its application areas are also very extensive, such as video surveillance, image detection, human-computer interaction, etc. However, the real-time and accuracy of face detection in the video environment are key issues that needs to be solved. In this paper, we preprocess differently frame images captured from video stream, then the AdaBoost face detection method is used to detect, along with tracking of Optical flow method, observe the influence of preprocessing methods on face detection. Thereby different pretreatment methods can be selected according to different detection environments to improve the accuracy of face detection.

1. Introduction
From the 1990s, face detection has become the focus of face recognition research. Most of the face detection research focuses on the relatively simple static background at early stages. However, in recent years, face detection technology under the video environment has begun to become a research trend due to the strong demand for identity identification in virtual games and the numerous demands in the business domain. Unlike face detection in static images, face detection in video faces many problems and challenges, such as complex background, external interference, etc. Pretreatment is the first and key step in face detection. It affects the implementation of face detection algorithms. In this paper, different methods of preprocessing the captured video frame images are used to observe the effect of face detection. Therefore, different pretreatment methods can be selected according to different detection environments to improve the accuracy of face detection.

2. Video Frame Image Acquisition and Preprocessing Theory

2.1 Video Frame Image Acquisition
This article adopts OpenCV+vs2015 for programming, 2165 frame images are extracted from the video through the OpenCV function, including 1902 single images, 39 images of two people, 23 images of three people, 45 images of nine people, 32 images of the image of 11 people, 124 images of non-face images, and the image format of 1432*704 pixels.

2.2 Preprocessing Theory
The purpose of image preprocessing is to weaken the noise in the image, improve the quality of the image to be detected, and provide the effective information for the next image detection.
2.2.1 Conversion of RGB color image and gray image. First, the image is grayscale processed to reduce the computation amount of subsequent processing and face detection. Then, a noise reduction process is performed on the converted grayscale image to reduce the interference of noise, such as Gaussian noise, salt and pepper noise, etc., on the subsequent face detection in the video frame image. Finally, the histogram equalization process is performed so that the contrast between the images is more obvious, interference caused by the over- or under-illuminated image contrast significantly is reduced.

2.2.2 Conversion of RGB Color Images and HSV Spatial Images. Because the HSV color space can more intuitively express the brightness and darkness of the color space and the vividness of the overall image, we can transform the RGB color space into an HSV space which is closer to human visual perception. After extracting the V and S components of HSV color space, histogram equalization is processed and finally transformed to RGB color space.

2.2.3 Conversion of RGB color image and YCBCR space image. YCBCR is a kind of color space and is usually used for continuous processing of images in movies. YCBCR space can be obtained through linear transformation of RGB color space, in which the component Y represents the brightness, and the CB and CR are the blue and red density offset components. In the YCBCR space, without considering the influence of component Y, it’s directly mapped to CBCR space, showing two dimensional independent distribution, which is convenient to compute.

3. Face detection method

Face detection in video is usually accompanied by complex illumination and other adverse environment, and the method based on statistics can be well adapted to the face detection of complex background images. Therefore, this article uses Adaboost face detection based on statistical methods. The commonly used face detection methods are Adaboost face detection algorithm based on Haar features and the characteristics of LBP (Local Binary Pattern).

4. Optical flow tracking

There are motion blur, occlusion, morphological changes, and illumination changes in the video, it is not possible to achieve good results only using image detection methods. Because the materials studied are video, video has space-time consistency, sports field information, which can be used to improve the detection effect. For the input video, detection is performed for each frame. Considering the motion field information between two consecutive frames, the detection results and the tracking information of the adjacent frames can be used to eliminate those parts that are easily jittered. Taking Fig.1 as an example, figure 1 is the result of adjacent frame image detection using Haar-adaboost detection method for face detection under HSV space preprocessing, in which a lot of misdiagnosis appears in the t+1 frame and the t+2 frame image. The optical flow tracking method can track human faces, which is a good solution to this kind of misdiagnosis.

The result of the optical flow tracking method is shown in Fig.2. The red dot in the figure is the position of the feature point of the previous frame in the current frame. It can be seen that the offset of the adjacent frame is very small and the human face has been detected in previous frame, so the false
detection portion in the current frame can be directly eliminated to improved the face detection effect in the video image.

**Figure 2.** Optical flow tracking results

5. The influence of different preprocessing methods on face detection

In this paper, two face detection methods that Haar-adaboost and LBP-Adaboost are tested, and the effect comparison is shown in Fig.3. It shows that the false detection rate of the LBP-Adaboost method is lower than that of the Haar-adaboost method, but the Leakage rate is increased. This is because the LBP model is extracted from the local texture features, and the calculation speed is fast, which can lead to a certain leakage, but this method can also eliminate the influence of light on face detection to a certain extent, thus reducing the error detection, so we can analyze the two methods respectively.

**Figure 3.** Contrast diagram of detection effect

5.1 Influence of grayscale processing on Haar-adaboost and LBP-Adaboost

After grayscale processing, it is obvious that the error rate of the detection results of the two detection methods have been reduced, which is because the effect of light is reduced after the grayscale treatment. The contrast effect of Haar-adaboost and LBP-Adaboost detection before and after grayscale processing is shown in Fig.4.

**Figure 4.** Contrast diagram of grayscale treatment effect

In order to further improve the speed of detection, we can use geometric normalization for it, because geometric normalization only changes the size of the picture and does not change the quality of the picture, we take the haar-adaboost method to detect. The video frame image pixel is 1432*704, and the pixel is reduced to 706*352 after geometric normalization. The detection time is reduced by
half, and the detection effect is unchanged for the images of one person, two people and even nine people, but the images of eleven people are not detected. The experimental results are shown in Fig.5.

![Figure5. Contrast diagram of geometric normalization effect](image1)

This is because the geometric normalization itself is a process of scaling the image. When it is reduced to a certain proportion, it will affect the features of the face in the image so that the face can not be detected. Therefore, geometric normalization is only applicable to single or small number of person images, which is very unsuitable for multi-person image.

5.2 Influence of HSV spatial processing on Haar-adaboost and LBP-Adaboost

Through experiments, we can see that the preprocessing of color space conversion in the dimly lit environment has certain improvements for both detection methods. The detection effect for Haar-adaboost is to improve positive inspection, while for LBP-Adaboost it is to reduce false detection. This is because the HSV color space is more intuitive than RGB to express the bright and dark colors of the color space and the brilliance of the overall image, and it is also closer to the human visual perception. The contrast effect of Haar-adaboost and LBP-Adaboost detection before and after the HSV space processing is shown in Fig.6.

![Figure6. Contrast diagram of the effect of HSV space processing](image2)

Moreover, the color space conversion preprocessing is more effective than the grayscale preprocessing method for face detection in the dimly lit environment, and the false detection rate is significantly reduced, thus we can know this color space conversion preprocessing method is better for a dimly lit imaging environment. The effect of grayscale and HSV space preprocessing on Haar-adaboost and LBP-Adaboost detection is shown in Fig.7.
5.3 Influence of YCRBR spatial processing on Haar-adaboost and LBP-Adaboost

After processing in the YCRBR space, the two detection methods have higher detection rates compared with the other two preprocessing methods in multi-person images. But for the Haar-adaboost face detection method, the detection rate is improved still accompanied by false detection, while face detection method does not increase the false detection for LBP-Adaboost. This is due to that the images of the images collected in this paper are from different races. They have different complexion, and the characteristics of complexion presents a good clustering and statistical distribution regularity in the YCBCR spatial distribution. Therefore, this pretreatment method is suitable for people with different skin color and race. Fig.8 shows the contrast between Haar-adaboost and LBP-Adaboost before and after YCRBR spatial processing.

6. conclusion

In this paper, the video frame image obtained from the OpenCV platform is used to experiment and analyze the process of face detection preprocessing. Due to the different face detection environments and different detection methods, we adopt different preprocessing flow, and take the optical flow tracing method to eliminate the jitters according to the video characteristics, then improve the accuracy of face detection.

References
[1] B. Heisele, T. Poggio, and M. Pontil. Face Detection in Still Gray Images[J]. Artificial Intelligence Laboratory. MIT, 2000.
[2] Kin-Pong F, Wai-Chee Fu A, Yu C. Haar Wavelets for efficient Similarity Search of i. Time-Series: With and Without Time Warping[J]. IEEE Transactions on Knowledge
ii. &Data Engineering, 2003, 15(3): 686-705.

[3] Zhang L, Chu R, Xiang S. Face Detection Based on Multi-Block LBP Representation[J]. Lecture Notes in Computer Science, 2007, 4642: 11-18.

[4] Xiaoyu Hou, Hongbo Wang. Design and Implementation of Face Detection and Recognition System in Video[J]. Beijing University of Posts and Telecommunications. 2017, 1, 17

[5] Zhichao Huang, Peng zhang, Huarong Zhao, Wenming Zhao. Research on Face Detection Algorithm Based on Skin Color Information in Intelligent Video Surveillance[J]. Modern Electronics Technique. 2018. 4. 1