Face Detection

ABSTRACT

Augmentation in computer technology has made feasible to induce new video processing applications in field of biometric recognition. Applications implicate face detection and face recognition incorporated to investigation systems, gesture analysis etc. Face detection is extensively used in interactive user interfaces and plays a very important role in the field of computer vision. In order to build a fully automated system that can analyze the information in face image, there is a need for robust and efficient face detection algorithms. In this paper a system is proposed for human face detection in images acquired under various illumination conditions. The features established on Gabor filters extracted from local image are applied to be the input of the Haar like classifier. In the end, the experiments show high performance in both accuracy and speed of the developed system.

INTRODUCTION

Face detection is the primary stage in the direction of automated face recognition. Potential applications of face detection and extraction are in interactive user interfaces, advertising industry, entertainment services, and video coding. This research is mainly interested in the face detection issue, which means how to find, based on visual information, the entire occurrences of faces in spite of who the person is. Face detection is likely one of the most challenging problems in computer vision and no solution has been achieved with performance similar to humans each in precision and speed.

Face detection can be viewed as a two-class classification problem in which an image region is classified as either a face or non-face [1]. There are over more than 150 approaches for face detection. These approaches are classified into four different categories, they are Knowledge-based, Feature invariant, Template matching, and Appearance-based [2]. In the following, a brief review of these four categories is given. (1) Knowledge based methods are rule-based methods, which encode human knowledge about what a face is. For example, in our human mind, symmetric eyes, ears, nose and mouth are key face feature. Developing these methods in different situations is sometimes difficult because not all states are countable. (2) Feature invariant methods are reforming methods with intention to find robustious structural features which are invariant to pose, lighting, etc. This method is one of the most important methods for face detection. (3) Template matching methods compute the correlation between patterns of a face and an input image in order to detection. In these methods, the correlation of several patterns of face in different poses and the input images are stored to be a criterion for face validation. (4) Appearance-based methods use models learned from training sets to represent the variability of facial appearance. Actually in these methods the templates are learned from face image samples. Generally, appearance-based methods utilize statistical analysis and machine learning to find characteristics related to face or non-face images [3].

Nowadays the field of face detection has made significant progress in the real world applications. In particular, the work by Viola and Jones has made face detection feasible in digital cameras and photo organization software [4]. In this paper, an effective real time face detection system based on Viola Jones approach is presented. The rest of this article is organized as follows. Section 2 gives the related works on face detection. Section 3 we will give brief overviews on feature extraction for our face detection task using Gabor feature extraction and Haar-like features in our proposed system. Section 4 gives experiments and results of proposed system. Section 5 discusses the conclusions of our work.

II. Related works

Prior attempts of face detection research have been concentrated on correlation or template matching, matched filtering, subspace methods, deformable templates, etc. Neural network and convolution neural network are used for face detection of static images respectively. These methods have good detection performance, but they suffer from high computational complexity [5]. C. Ding et al. proposed quick face detection based on skin color and pattern plate. This method can do well in detecting the frontal face. However the above work fails to detect face under the illumination changes, occlusion and faces with arbitrary poses [7]. Stan Z.Li et al. proposed Recursive Nonparametric Discriminant Analysis (RNDA) to extract facial features. The Adaboost classifier was utilized to detect faces with arbitrary poses. Due to facial skin color negligence some wrong objects are detected as a face [8]. ManminderSingh et al. demonstrated that face images with the same pose, under different illumination conditions, form a demonstrated that face images with the same pose, under different illumination conditions, form a convex cone, the illumination cone [9]. H.Nanda et al. have introduced a face detecting method of template matching, this method is easy to achieve, but it has low detection rate for both pose-varied face and decorations [6]. Ke Yan et al. proposed new color local texture features including color local Gabor wavelets and color local binary for face recognition and that could provide excellent recognition rates, but this study was limited to extract features from fixed color-component configuration [10].

III. Proposed Methodology

The proposed method of face detection in this paper is derived on multi-phase architecture. As shown in Figure the setup has two phases: first phase is pre-processing stage where we use Viola and Jones algorithm to detect faces from real time web camera or still images selected from hard disk drive, and the result of this phase will be the candidate images which will be either face or non- face

Fig 1: Block diagram of Proposed system

Then, these candidate images will be the input for the second phase which consists of three sub-phases; Gabor filters for feature extraction and haar like feature extraction for face detection. So first sub-phase of the second phase will be Gabor filter feature extraction, in which the feature will be extracted from the candidate images by applying the Gabor filters wavelet. After that, the produced feature vector is used as an input to the second sub-phase Haar like feature extraction, to detect the face. Viola and Jones method needs first to transform the images from RGB to grey scale. Contrast and illumination adjustment operations will be achieved to have good results. In some cases, facial images are often corrupted by various types of noise; so they will be processed with the proper filters for noise removal restoration. The output of this stage will be the cropped candidate faces or non faces from the source image. In our research we will adopt the implementation of Viola and Jones using built in Matlab libraries. Viola and Jones face detection method consists of a group of algorithms for real time face detection. It became one of the most important and robust face detection techniques. In this technique, the processing steps and the operations are carried based on the integral image technique which allows rapid computation and fast response. In general Viola and Jones method is appropriate for our proposed technique, and was chosen in this paper for its robustness, speed of detection [11].

A. Gabor Feature Extraction

Human beings can recognize faces since childhood. It is an easy process for our brains; we can detect faces and recognize people even if they are wearing glasses or having moustaches [12]. So, the process of detecting faces for human being is trivial and very easy, but it is not as easy for machines to do so. The main challenge of face detection is how to extract features from an image in order to decide whether the image represents a face or not [13]. We can define the process of extracting useful and valuable information from an image as a feature extracting process, having in mind that the extracted information must be useful, valuable and representative, so that it can be passed to other stages of the face detection technique [14]. Also, the process of feature extraction must not put a burden on the system so that it consumes long time to be calculated.

In our proposed technique, we will use Gabor wavelet transform to extract certain local features of the image. Gabor wavelet transform works in both resolution and orientation properties and optimal for measuring local spatial frequencies. The Gabor representation of an image is the convolution of the image with the Gabor filter. Based on the Gabor representations, a feature vector is formed [15]. Due to the above reasons Gabor wavelet transform has been used in many different image analysis and processing applications.

Main advantage of this filter is that its representation of frequency and orientation is the same as in the visual perception of humans. This makes the filter suitable for texture representation and discrimination. A two dimension Gabor filter is a Gaussian function modulated by a sinusoidal wave [16]. The main goal behind using this two dimensional Gabor wavelet filter is to get the main feature vectors that will be the input to the next stage of our proposed technique, in order to get better characterizations of the visual content of facial images. The impulse response of the Gabor filter can be defined mathematically by multiplying the harmonic function of a sinusoidal wave with the Gaussian function. The impulse response then will have both a real and imaginary components which can be calculated using equations (1) (2) and (3).

The Complex Part:

$$g(x,y;\lambda,\theta,\Psi,\sigma,\gamma) = exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right) exp\left(i\left(2\pi\frac{x'}{\lambda}+\Psi\right)\right) \quad ----- (1)$$

The Real Part:

$$g(x,y;\lambda,\theta,\Psi,\sigma,\gamma) = exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right) cos\left(2\pi\frac{x'}{\lambda}+\Psi\right) \quad ----- (2)$$

The Imaginary Part:

$$g(x,y;\lambda,\theta,\Psi,\sigma,\gamma) = exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right) sin\left(2\pi\frac{x'}{\lambda}+\Psi\right) \quad ----- (3)$$

Where:

λ:Wave length of the Sinusoidal wave.

θ:The orientation angle of the parallel stripes angle of the Gabor function.

Ψ:Phase offset

σ:Sigma of the Gaussian envelope

γ:Spatial aspect ratio

$$x' = x\ \cos\theta + \mathrm{y}\sin\theta$$

$$y' = -x\ \sin\theta + \mathrm{y}\cos\theta$$

B. Haar like features

A cascade of classifiers is used for face detection on an image. Classifiers analyze the same window of the initial image sequentially. If the first classifier returns true, the second classifier will analyze a current window etc. If the classifier returns false, the cascade stops analyzing the current window and turns to the next window. The next window is shifted to one pixel right or down according to the current window. The classifier is built on Haar-like features [17]. The traditional Haar-like feature set used, which is composed of four basic types of rectangle features. The features can locate at any area within a scanning image window. A feature is evaluated by the difference between the summed pixel values of dark rectangular and that of light rectangular. The rectangular feature can represent the local appearance of object [18]. Feature A and feature B in Figure 1 is designed to measures the difference in intensity between two adjacent rectangular regions. For example, feature B can measures the difference in intensity between the region of the eyes and a region across the upper cheeks. The most important characteristic of the Viola-Jones framework is that the Haar like features can be computed rapidly at all scales in constant time by using the integral image. The Haar-like features of a gray-level image is intrinsically the gray contrast of local areas [19]. The color contrast to skin color or the difference of gray gradient can be regard as another form of gray image, this kind of single valued feature represented in the form of gray image [20]. In this way, the Haar like features are not only extended in their forms but also in their physical natures, which add additional domain-knowledge to the learning framework and which is impossible to learn from gray-level image. All features represented in the form of gray like image can be used in the Viola-Jones framework, as they can be converted to Haar like features efficiently after computing their integral images.

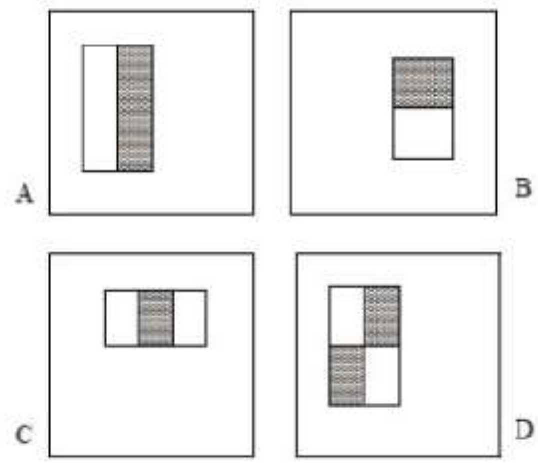

Fig 2 : Examples of Rectangle features

An integral image has been used to accelerate feature sum calculation (1). The integral image is an array with the same size as a source image. An integral image element value is a sum of all pixels positioned left and higher than a current element in the source image. For example, element 1 (fig.3) is a sum of all pixels in area A. Eqn (4) is used to calculate the sum of pixels of any rectangle area in the image.

$$sum = pt_4 - pt_3 - pt_2 + pt_1 \quad - - - - - \ (4)$$

Coefficient indexes indicate corresponding area angles (see fig.3). A classifier window has a fixed size. Thus, the classifier can detect face with a comparable size. There are two approaches for searching a face with an arbitrary size: a classifier scaling and an image scaling.

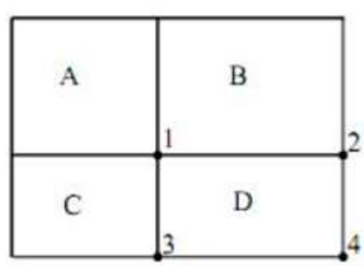

Fig 3 : Example for calculating sum

IV. Experiments and Results

We evaluate the performance of the face detector on FDDB database [21]. We also provide an analysis of the proposed method, report the face detection speed, and report unconstrained face detection performances under illumination variations, pose variations, occlusion, and blur, respectively. The FDDB dataset [21] covers challenging scenarios for face detection. Images in FDDB comes from the Faces in the Wild dataset , which is a large collection of Internet images collected from the Yahoo News. It contains 2,845 images with a total of 5,171 faces, with a wide range of challenging scenarios including arbitrary pose, occlusions, different lightings, expressions, low resolutions, and out-of-focus faces [8]. All faces in the database have been annotated with elliptical regions. We compared our method with other algorithms in FDDB database and the results are shown in Table, Which shows are proposed work outperforms other methods.

VI. Conclusion

We presented an effective and robust method for detecting faces in images under various illumination conditions based on Haar features. Gabor feature extraction technique is incorporated into a conventional face detector to alleviate illumination variation problem. It is improved the face detection rate and reduce the processing time and the size of the bounding box more accurately attached to the detected face. Face detection in the video sequences is an important research branch of face recognition. In future we will investigate face recognition from video sequences under various illumination conditions.

References

1. Ç. Sofuogglu and A. T. Cemgil, "Change-point detection via switching Kalman filters," *2017 25th Signal Processing and Communications Applications Conference (SIU)*, Antalya, 2017, pp. 1-4. doi: 10.1109/SIU.2017.7960439
2. A. Wu, W. S. Zheng and J. H. Lai, "Robust Depth-Based Person Re-Identification," in *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2588-2603, June 2017. doi: 10.1109/TIP.2017.2675201
3. F. Bowen, J. Hu and E. Y. Du, "A Multistage Approach for Image Registration," in *IEEE Transactions on Cybernetics*, vol. 46, no. 9, pp. 2119-2131, Sept. 2016. doi: 10.1109/TCYB.2015.2465394
4. G. Deore, R. Bodhula, V. Udpikar and V. More, "Study of masked face detection approach in video analytics," *2016 Conference on Advances in Signal Processing (CASP)*, Pune, 2016, pp. 196-200. doi: 10.1109/CASP.2016.7746164
5. Paul Viola, Michael J Jones. Robust real-time face detection[J]. Internation Jounrnal of Computer Vision, 2004, 57(2):137-154.
6. Harsh Nanda, LarryDavis. Probabilistic Template Based PedestrianDetection in Infrared Videos[D]. University of Maryland, College Park, MD-20742.
7. C. Ding, C. Xu and D. Tao, "Multi-Task Pose-Invariant Face Recognition," in *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 980-993, March 2015. doi: 10.1109/TIP.2015.2390959
8. S. Liao, A. K. Jain and S. Z. Li, "A Fast and Accurate Unconstrained Face Detector," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 211-223, Feb. 1 2016. doi: 10.1109/TPAMI.2015.2448075
9. ManminderSingh A.S.AroraDr. "Varying Illumination and Pose Conditions in Face Recognition" in Procedia Computer Science Volume 85, 2016, Pages 691-695 ELSEVIER
10. Ke Yan, Youbin Chen and D. Zhang, "Gabor Surface Feature for face recognition," *The First Asian Conference on Pattern Recognition*, Beijing, 2011, pp. 288-292. doi: 10.1109/ACPR.2011.6166553

11. M. Da'san, A. Alqudah and O. Debeir, "Face detection using Viola and Jones method and neural networks," *2015 International Conference on Information and Communication Technology Research (ICTRC)*, Abu Dhabi, 2015, pp. 40-43.
12. S. Caharel, M. Ramon and B. Rossion, "Face Familiarity Decisions Take 200 msec in the Human Brain: Electrophysiological Evidence from a Go/No-go Speeded Task," in *Journal of Cognitive Neuroscience*, vol. 26, no. 1, pp. 81-95, Jan. 2014.
13. Yi-Chong Zeng, "Automatic extraction of useful scenario information for dramatic videos," *2013 9th International Conference on Information, Communications & Signal Processing*, Tainan, 2013, pp. 1-5. doi: 10.1109/ICICS.2013.6782967
14. K. Zhang, Q. Liu, Y. Wu and M. H. Yang, "Robust Visual Tracking via Convolutional Networks Without Training," in *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1779-1792, April 2016. doi: 10.1109/TIP.2016.2531283
15. J. Geng, J. Fan, H. Wang, X. Ma, B. Li and F. Chen, "High-Resolution SAR Image Classification via Deep Convolutional Autoencoders," in *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2351-2355, Nov. 2015.
16. P. Liu and D. Liu, "Filter-based compounded delay estimation with application to strain imaging," in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 58, no. 10, pp. 2078-2095, October 2011.
17. P. Wang, C. Shen, N. Barnes and H. Zheng, "Fast and Robust Object Detection Using Asymmetric Totally Corrective Boosting," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 1, pp. 33-46, Jan. 2012
18. Jean Paul, Niyoyita Zhao Hui Tang, Jin Ping Liu "Multi-view Face Detection Using Six Segmented Rectangular Features" The Sixth International Symposium on Neural Networks, Springer, Berlin, Heidelberg 2009
19. S. Liao, A. K. Jain and S. Z. Li, "A Fast and Accurate Unconstrained Face Detector," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 211-223, Feb. 1 2016. doi: 10.1109/TPAMI.2015.2448075
20. W. ZHANG, X. ZHAO, J. M. Morvan and L. Chen, "Improving Shadow Suppression for Illumination Robust Face Recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, no. 99, pp. 1-1.
21. V. Jain and E. Learned-Miller, "FDDB: A benchmark for face detection in unconstrained settings," University of Massachusetts, Amherst, Tech. Rep. UM-CS-2010-009, 2010