Robust Real Time based Face Recognition using Support Vector Machine & Histogram Equalization Algorithm

Dr. V. S. Manjula

Associate Professor & Head, Department of Information System & Network Engineering
St. Joseph University College of Engineering & Technology, Dar-Es-Salaam, Tanzania, East Africa

Abstract: ***A face recognition algorithm based Support Vector Machine & Histogram Equalization Methods. These methods allow standardizing the faces illumination reducing in such way that the variations for further features extraction; which are extracted using the image phase spectrum of the histogram equalized image together with the principal components analysis. This algorithm operates by mapping training set into a high-dimensional feature space, and separate positive and negative samples. In statistical learning theory, for some classes of well behaved data, the choice of the maximum margin hyper plane will lead to maximal generalization when predicting the classification. The input dataset is divided into training and testing dataset and experiments are performed by varying dataset size. The effect of performing image intensity normalization, histogram equalization, and input scaling are observed. Proposed scheme allows a reduction of the amount of data without much information loss. Evaluation results show that the proposed feature extraction scheme, when used together with the support vector machine (SVM), provides a recognition rate higher than 97% and a verification error lower than 0.003%.***

Keywords: Histogram Equalization, Preprocessing Techniques, Fast Fourier Transform, Principal Component Analysis, Support Vector Machine.

1. Introduction

Face recognition is one of the most traditional research and development applications of using an image in face detection and tracking are focused on video images or in still images. This has received significant attention, because of its wide range of commercial and law enforcement applications. Even though current machine recognition systems have reached a certain level of *maturity*, their success is limited by the conditions imposed by many real applications. For example, recognition of face images acquired in an outdoor environment with changes in illumination and/or pose remains a largely unsolved problem. The structure of the face, a direct lighting source can cast strong shadows that accentuate or diminish certain facial features. The illumination problem is basically the variability of an object's appearance from one image to the next with slight changes in lighting conditions and viewpoint. This often results in large changes in the object's appearance.

One of the challenges of automatic gender classification is to account for the effects of pose, illumination and background clutter. Practical systems have to be robust enough to take these issues into consideration. Most of the work in gender classification assumes that the frontal views of faces, which are pre-aligned and free of distracting background clutters, are available. It provides a framework that is free of these assumptions and can classify faces by first automatically detecting, then localizing and finally extracting features from arbitrary viewpoints. But in this paper, only the facial images with full frontal views are considered.

2. Backgrounds and Prior Work

2.1 Motivation

Image preprocessing techniques represent an essential part of face recognition systems, which has a great impact on the performance and robustness of the recognition procedure. Amongst the number of techniques already presented in the literature, histogram equalization has emerged as the dominant preprocessing technique and is regularly used for the task of face recognition. With the property of increasing the global contrast of the facial image while simultaneously compensating for the illumination conditions present at the image acquisition stage, it represents a useful preprocessing step, which can ensure enhanced and more robust recognition performance. Even though, more elaborate normalization techniques, such as the multistage retime technique, isotropic and anisotropic smoothing, have been introduced to field of face recognition, they have been found to be more of a complement than a real substitute for histogram equalization. However, by closer examining the characteristics of histogram equalization, one can quickly discover that it represents only a specific case of a more general concept of histogram remapping techniques which may have similar characteristics as histogram equalization.

2.2 Support Vector Machine

Support vector machine is a very effective method for general purpose pattern recognition. SVM is particularly a good tool to classify a set of points which belong to two or more classes. SVMs are based on statistical learning theory and try to find the biggest margin to separate different classes. SVM embed data into a high dimensional feature space. This method uses the hyper plane that separates the largest possible fraction of points of the same class on the same side, while it maximizes the distance of either class from the hyper-plane. Hence there is only the inner product involved in SVM, learning and predicting is much faster than a multilayer neural network. We introduced the SVM as a supervised learning algorithm. This algorithm operates by mapping training set into a high-dimensional feature space, and separate positive and negative samples. In statistical learning theory, some classes of well behaved data, the choice of the maximum margin hyper plane will lead to maximal generalization when predicting the classification of examples.

The basic concept of SVM is as follows: Assume we want to fit a function g(x) through a set of samples S={(x1,y1),(x2,y2),…,(xn,yn)}. x ϵ Rn and y ϵ {- 1,+1}. y is the label of two different classes. Φ(x) maps input data to a high dimensional feature space as follows. The above equation can be combined into one equation: These points lie on the hyper-plane H1: *w* TΦ(x) + b = 1 and H2: *w* TΦ(x) + b = -1, The margin between these two hyper-plane is: d = 2/||w|| , shows the points lie one the H1 and H2 are support vectors.

To tackle these issues, we present in this paper an empirical assessment of the concept of histogram remapping with the following target distributions: the uniform, the normal, the lognormal and the exponential distribution. The given image databases and conclude that similar or even better recognition results that those ensured by histogram equalization can be achieved when other non-uniform distribution are considered for the histogram remapping. This enhanced performance, however, comes at a price, as the non-uniform distributions rely on some parameters which have to be trained or selected appropriately to achieve the optimal performance.

Face recognition under surveillance, where a person should be recognized without intentional, cooperative effort. Extrinsic factors, including eyeglasses, hairstyle, expression, posture, and environmental lighting, which make distributions of face data highly complex should be minimized for reliable face recognition. Among several extrinsic factors, problems with uncontrolled environmental lighting are the topmost issue to solve for reliable face-based biometric applications in practice.

However, most current face recognition systems, academic and commercial, are based on face images captured in the Visible Light spectrum; they are compromised in accuracy by changes in environmental illumination, even for co-operative user applications indoors. In an in-depth study on the influence of illumination changes on face recognition, several distance measures and several local image operators, including Gabor filters, local directive filters, and edge maps, which were considered to be relatively insensitive to illumination changes for face recognition. Several conclusions are made there: 1) Lighting conditions, and especially light angle, drastically change the appearance of a face. 2) When comparing unprocessed images, the changes between the images of a person under different illumination conditions are larger than those between the images of two people under the same illumination. 3) All of the local filters under study are insufficient by themselves to overcome variations due to changes in illumination direction. The influence of illumination is also shown in the recent Face Recognition Vendor Test. The issues of face detection and tracking have long been studied in the fields of computer vision and pattern recognition. The widespread interests devoted to such research and development, are due in part to increasingly growing performance ratio of computing power and related hardware, and beyond that, due to potential important applications in surveillance, human-computer interaction, retrieval among others.

2.3 Grayscale Values

Grayscale preprocessing was the most imprecise method. If grayscale testing performance data (averaged from a set of 10 runs with different datasets) for neural networks of 5, 10, 15, 20 and 25 hidden neurons. The PCA minimum information percentage that allows the best testing performance of the given input image.

Figure 1: Gray Scale Conversion to Equalize Image

3. Experiments and Discussion

In this paper, we propose a real-time robust method for eye tracking under variable lighting conditions and face orientations, based on combining the appearance-based methods and the active IR illumination approach. Combing the respective strengths of different complementary techniques and over-coming their shortcomings, the proposed method uses an active infrared illumination to brighten subject's faces to produce the bright pupil erect. The bright pupil erect and the appearance of eyes are utilized simultaneously for eyes detection and tracking. The latest technologies in pattern classification recognition (Support Vector Machine) and in object tracking are employed for pupil detection and tracking based on eyes appearance.

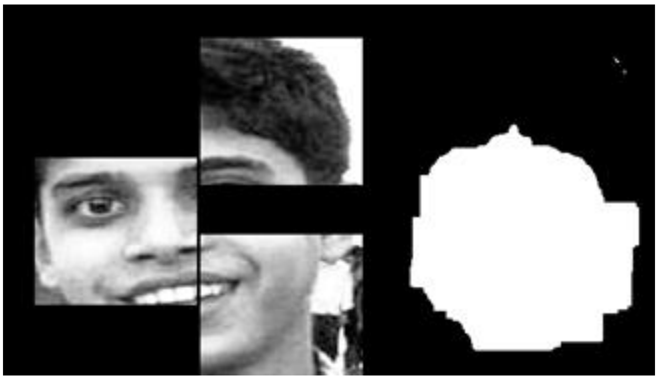


Figure 2: Face detection using Neural Network

n this paper, Our method consists of three parts: eye detection, nose detection and eye tracking. Eye detection is accomplished by simultaneously utilizing the bright/dark pupil erect under active illumination and the eye appearance pattern under ambient illumination via the Support Vector Machine. Eye tracking is composed of two major modules. In case Kalman eye tracker fails due to either weak pupil intensity or the absence of the bright pupils, eye tracking based on the mean shift is activated and to continue tracking the eyes. Eye tracking returns to the Kalman filtering tracker as soon as the bright pupils reappear since eye tracking using bright pupils is much more robust than the mean shift tracker. The two trackers alternate, complementing each other and overcoming their limitations.

3.1 Filters & Features

Image features are called *Rectangle Features* and are reminiscent of rectangle feature; there are three types of rectangle filters. The value of a *two-rectangle filter* is the difference between the sum of the pixels within two rectangular regions. The regions have the same size and shape and are horizontally or vertically adjacent a *three-rectangle filter* computes the sum within two outside rectangles subtracted from the sum in a center rectangle

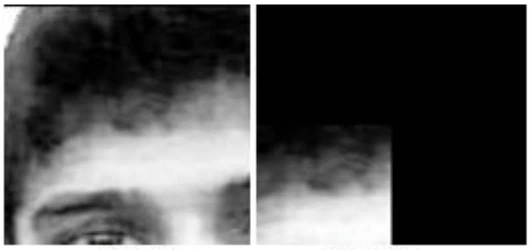
Figure 3: Face detection using Neural Network

Finally a *four rectangle filter* computes the difference between diagonal pairs of rectangles. Given that the base resolution of the classifier is 24 by 24 pixels, the exhaustive set of rectangle filters is quite large, 190,800, which is roughly O(N4). The actual number is a smaller since filters must fit within the classification window.

3.2 Tracking

The analysis of surveillance video presents different problems and opportunities than the static analysis of mugshots or passport photos. As mentioned earlier surveillance video is low quality, has poor lighting, and may never capture a completely "frontal" facial image. These flaws are somewhat offset by the large number of face images available for any individual. Given 15 frames per second, it is not unreasonable to assume that 15 to 30 usable images of each individual will be available. Though no single image might be of high quality, the collection of images is more likely to yield confident predictions of gender or race. This sort of temporal integration requires that detected faces be tracked across time. In this section we will briefly describe a new form of face tracking which has distinct advantages over previous approaches.

Most tracking algorithms implicitly assume that a complete "brute force" search of an image for the target of interest is prohibitively expensive. As a result location evidence from the previous frames is used to focus the search in subsequent images. Given strong prior assumptions about the movement of the target, each new frame can be processed very quickly. Another key component of tracking algorithms is disambiguation: given the locations of several targets in one image and their location in the next image, the tracking algorithm can act to disambiguate the identities of the objects. Once again a prior model for target motion is required. One form of ambiguity is the absence of evidence for a particular tracked target. Prior knowledge can be used to compute target location as a combination of weak image evidence and strong prior expectations.



Face 1 Face 2

Figure 4: Detected Face

The final property common to almost all tracking algorithms is the necessity of initialization. One either assumes a simple process is sufficient for initial detection or initialization is done by hand. In the domain of face tracking, the appearance of a real-time algorithm can radically change many of the classical motivations for tracking. The tracking process no longer needs to "focus" the search for faces. The tracking process no longer requires initialization, since new faces will be automatically detected in each frame.

3.3 Face Deduction Experiment

1) Original images, 2) horizontal and vertical derivates filtered images, 3) laplacian filtered images, 4) horizontal and vertical derivates filtered images joined to the laplacian, and 5) horizontal and vertical derivates filtered images joined to the grayscale. These operations over the original image are part of the preprocessing step in the whole face detection process.

3.4 Face Recognition Experiment

The appearance based methods detect eyes based on their photometric appearance. These methods usually need to collect a large amount of training data, representing the eyes of divergent subjects, under divergent face orientations, and under divergent illumination conditions. These data are used to train a classifier such as a neural network or the Support Vector Machine and detection is achieved via classification. Proposed several improvements on the neural network based eye detector. The trained neural network eye detector can detect rotated or scaled eyes under divergent lighting conditions. But it is trained for the frontal view face image only. Feature based methods explore the characteristics of the eyes to identify some distinctive features around the eyes. At the upper-left corner is an ideal example of eyes which can be easily detected, with a "bright pupil" detector an appearance-based detector. Secular reflection on eyeglasses is the most serious problem. Eyelid occlusion, which happens among people in some ethnic groups and senior people and eye closing due to blinking are among other problems. Eye detection in these situations cannot be done by using a simple eye detector.

4. Discussion

We are taking input image of same subject but different expression and variations in facial expression (open/closed eyes, smiling/non-smiling), and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an up-right, frontal position, with tolerance for some tilting and rotation of up to about 20 degrees. There is some variation in scale of up to about 10%. The images are grayscale with a resolution of given input image.

Figure 5: Rectangle features extraction using

5. Neural Network

The first step in facial feature detection is detecting the face. This requires analyzing the entire image. The second step is using the isolated face(s) to detect each feature. The result is shown in Figure 5. Since each the portion of the image used to detect a feature is much smaller than that of the whole image, detection of all three facial features takes less time on average than detecting the face itself. The mouth detection has a lower rate due to the minimum size required for detection. By changing the height and width parameter to more accurately represent the dimensions of the mouth and retraining the classifier the accuracy should increase the accuracy to that of Rectangle.

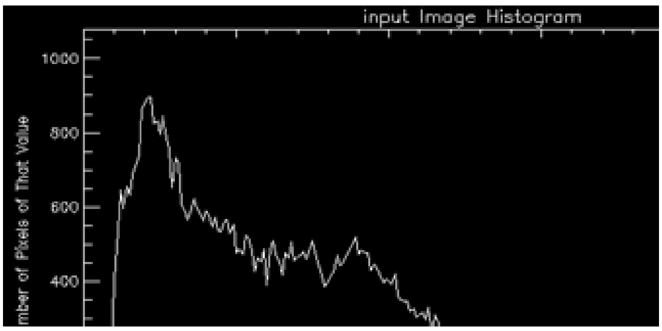


Figure 6: Input Image Histogram

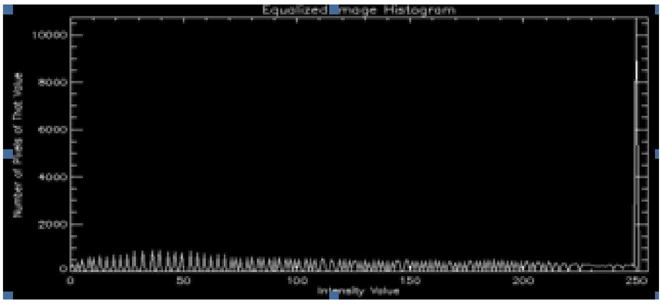


Figure 7: Histogram for Equalize image

Since a frame rate of 5 frames per second was achieved in facial detection only by using a much faster processor, regionalization provides a tremendous increase in efficiency in facial feature detection.

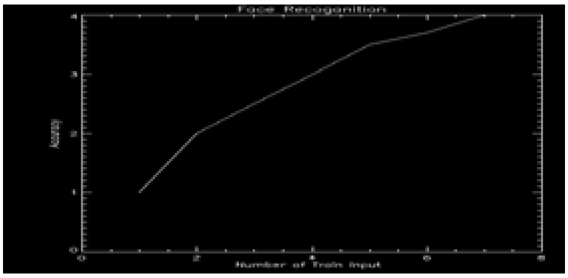


Figure 8: Result of input face image

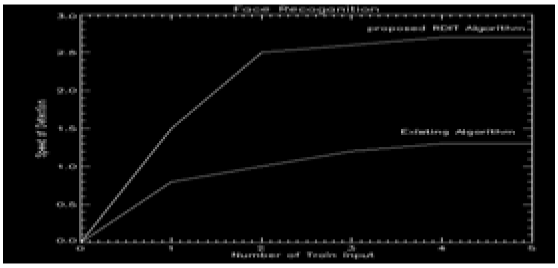


Figure 9: Fitness Measure using Face image

Regionalization also greatly increased the accuracy of the detection. All false positives were eliminated, giving a detection rate of around 95% for the eyes and nose. Figure (5) contains an eye and a part of the nose. This indicates that the rect- angular regions are not necessary to use a single facial feature as basic unit and the other features. The locations of most rectangles are around eyes and nose. This is consistent with the intuition that these facial features are critical in face recognition. For any application, but especially for an integrated real-time system, the time it takes to classify a new image is of critical importance. The low error rates shown in the anecdotal experiments in Figures 2 and 1 are not a typical. These images are actually of much higher quality than many examples from our web test set. In a number of applications where image quality is higher than web data, lower error rates can be assumed. A real-time system provides an additional opportunity for reduction in classification error, the integration of classifications across time.

6. Conclusion

In this paper, we propose a real-time face recognizing and tracking system which can detect face, then recognize and track it. Our method performs well regardless of whether the faces is in the complex background, moving fast or partly covered, and the hybrid algorithm, PCA & SVM further enhance the accuracy of face recognition. By experimental results, we have demonstrated that the proposed method dramatically improves the robustness and accuracy of the real-time face recognizing and tracking system.

References

[1] **D. Comaniciu, V. Ramesh(2000), "***Robust Detection and Tracking of Human Faces with an Active Camera",* IEEE Int'l Workshop on Visual Surveillance, Dublin, Ireland, 11-18.

[2] **X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang,** *"Face Recognition Using Laplacianfaces," IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[3] **Feng, Guo Can; Yuen, Pong;** "Multi-cues eye detection on gray intensity image*", Pattern Recognition*, *34*(5), 2001, 1033-1046.

[4] **Yang, M. H.; Kriegman, D. J.; and Ahuja N.** *"Detecting Faces in Images: A Survey",* IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1): 34-58, 2002

[5] D. Comaniciu, V. Ramesh (2000), *"Mean Shift and Optimal Prediction for Efficient Object Tracking"*, IEEE Int'l Conf. Image Processing, Vancouver, Canada, Vol. 3: 70-73.

[6] **G. Dai and C. L. Zhou,** *"Face Recognition Using Support Vector Machines with the Robust Feature",* Proc. of the 2003 IEEE International Workshop on Robot and Human interactive Communication, 2003.

[7] **Huang, J., S. Gutta, and H. Wechsler (1996),** "Detection of Human Faces Using Decision Trees", *2nd International Conferenceon Automatic Face and Gesture Recognition (ICAFGR)*, Killington, VT.

[8] **M. Pantic and L. J. M. Rothkrantz,** *"Automatic analysis of facial expressions",* The state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1424–1445, 2000.

[9] **P. Viola and M. Jones,"** *Robust real-time object detection",* Technical Report 2001/01, Compaq Cambridge Research Lab, 2001.

[10] **C.W.Hsu and C.J.Lin (2002),** *"A simple decomposition method for support vector machines",* machine Learning, 46: 291-314.

[11] **Haddadnia, J.; Faez, K.; and Moallem, P.** "*Neural Network Based Face Recognition with Moment Invariants",* In Proceedings of the IEEE International Conference on Image Processing, 1: 1018-1021. Thessaloniki Greece, 2001