Volume 4, Issue 5, May 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Object Identification in Real Time Systems

Harsha Singh[1]
Computer Science and Engg,
PDM college of Engg for women,MDU
Bahadurgarh, Haryana, India

Ms.smita[2]
A.P in CSE Department,
PDM college of Engg for women,MDU
Bahadurgarh,Haryana, India

Abstract-Recent studies of object identification shows that it plays a vital role in a wide range of application that intersect with our lives: automatic driving, driver assistance system, human–computer interaction ,smart homes and many more. The advancement of machine perception systems which detect eyes and analyze eye behavior will also enable development of new approaches . We derive optimal inference algorithms for finding eyes within this framework.The approach represent the image as a collage of patches of arbitrary size, out of which few contain the object of interest and few are background. These models are learned using boosting methods.This paper describe a framework capable of processing images rapidly and achieving high detection rates.This framework is a three-stage process and approach searches the entire image plane in each frame, making it robust, unpredictable motion. The system is robust to changes in illumination and system can simultaneously track the eyes and blinks of multiple individuals. The proposed system is specialized in detecting eye in a particular pose: upright frontal. The upright frontal view is needed to indicate fatigue in car drivers. This method is very effective under controlled illumination conditions. Finally we represent the development of perceptive system which may help in understanding human brain.

Keywords :face detection ,eye detection, blink detection.

I. INTRODUCTION

Traditional research in the area of developmental psychology , face perception is one of the most important step for understanding how human mind develops from infancy to adulthood.In general , face processing and eye detection are the landmark in many influential research area. The researches has postulated their need in eye detection and gaze processing modules.In recent years , automatic detection of faces and facial behavior has been examined flourishingly. In realistic environment , good eye detection have tremendous impact on face perception technology [3]. As we use our eye in expressing our thoughts and eye movement provide conceivable information for activity recognition. A good eye detection can be achieved by following ways:1) proper registration . In a recent evaluation face recognition system proposed that large no.of failures occurred in outdoor conditions due to poor registration and alignment of facial features.2)FACS(facial action coding system) developed by Ekman and Frisen in 1978 which is considered to be a comprehensive standard for coding facial behavior. The system describes 15 action units of eye behavior[2]. This system reflects that eye behavior is very rich and informative. The action unit 45 is most important eye related behavior i.e blink.Blinks are known to reflect the emotional state of the person. Blink detection is most important in several fields like neurology, psychology and physiology.

Code	Descriptor	Muscles Involved	Example
AU5	Upper Lid Raiser	Levator Palpebrae Superioris	
AU6	Cheek Raiser	Orbicularis Oculi, Pars Orbitalis	
AU7	Lid Tightener	Orbicularis Oculi, Pars Palebralis	
AU41	Lid Droop	Relaxation of Levator Palpebrae Superioris	
AU42	Slit	Orbicularis Oculi	
AU43	Eyes Closed	Relaxation of Levator Palpebrae Superioris; Orbicularis Oculi, pars Palpebralis	
AU44	Squint	Orbicularis Oculi, pars Palpebralis	
AU45	Blink	Relaxation of Levator Palpebrae Superioris; Orbicularis Oculi, pars Palpebralis	
AU46	Wink	Relaxation of Levator Palpebrae Superioris; Orbicularis Oculi, pars Palpebralis	
AU61	Eyes Turn Left	Lateral and Medial Rectus	
AU62	Eyes Turn right	Lateral and Medial Rectus	
AU63	Eyes Up	Superior Rectus	
AU64	Eyes Down	Inferious Rectus	
AU65	Walleye	Lateral Rectus	
AU66	Crosseye	Medial Rectus	

II. Object identification in real time systems

We evaluate an algorithm that performs optimal inference under the assumption of generative model [1].The real time system utilizes two types of eye detectors: The first type is a face detector which begins with total uncertainty about the possible position of eyes on the image plane. Its function is to narrow down the uncertainty about the location of the eyes while operating in a very wide background conditions and variety of illuminations. The second type of detector is eye detector, works on the output of the face detector used in first type. By doing this, we can assume a restricted context and attain high location accuracy. Once the very likely eye location is selected after this, image patch containing the location of eyes is obtainted. The image patch is then passed to a blink detection for analysis.The procedure is shown by the help of flowchart in Fig.4

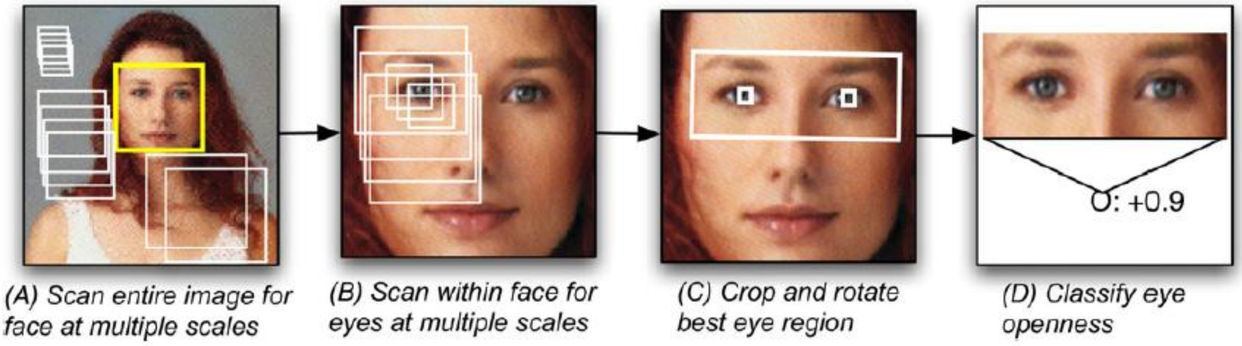

Fig. 1. Flowchart for face, eye, and blink detection.

The system described above performs operation on each frame and treats each as independent of the previous frames.This system is useful for static images as for video. The purpose of treating each video frame separately allows the system to simultaneously code eye location and behavior on multiple faces that may come in and out of the scene at random times

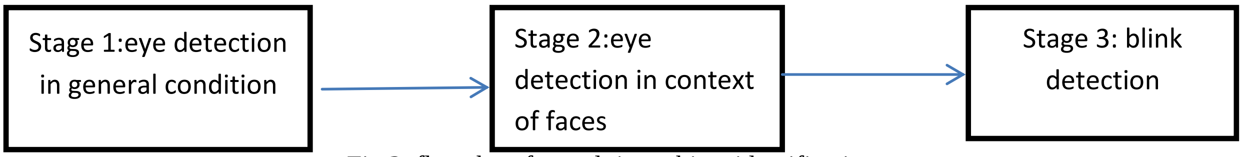

Fig 2 .flowchart for real time object identification

Stage I: Eye detection in general background conditions

As mentioned above the first component is the inference process,which locates regions of the image patch that contain faces, and thus eyes. This module works under very general background and illumination conditions and greatly reduce the feasible locations of eyes on the image plane. This module makes no prior assumptions about the location of the face. The generally image are processed using approach similar to the multiscale search [6], a single binary classifier is used to classify face vs. non-face for patches of fixed size (20 ×20 pixels), then that classifier is used to classify all possible patches in the image. Faces larger than the original size were found by repeating the search in copies of the image scaled to smaller sizes (thus, a 20 × 20 pixel face in a 1/4 size copy of the image corresponds to an 80 × 80 pixel face in the corresponding

location in the original). We developed a likelihood-ratio model using a dataset of Web images provided by Compaq Research Laboratories. This dataset carries 5000 images containing frontal upright faces taken under a variety of illumination conditions, facial expressions, facial hair, eyeglasses, hats, etc. Faces were cropped and scaled to 24 × 24 pixels square. The negative examples collected from web were sampled from a dataset of 8000 images and known not to contain faces. Likewise, these images contained a wide variety of natural indoor and outdoor scenes, illustrations , etc., with varying image quality. The merits of this Web dataset is that it includes more variability than most other closed databases.

Due to the multi-scale search, about 1 billion total patches are possible in these 8000 images. The initial negative examples for training, 10,000 square patches, of arbitrary size and at arbitrary locations in the images, were sampled from this dataset. Patches were then scaled down to 24 × 24 pixels. The set of negative samples changes during training thanks to the bootstrap round , so ultimately all 1 billion possible patches were used at some time during training (see figure 3).

Figure 3: examples of faces and non faces used in training face detector

The likelihood-ratio model was trained using the GentleBoost,sequentially chooses wavelets from a large pool and combines them to minimize a $\chi 2$ error function. Based on Viola and Jones(2001)[4]; shakhnarovich et al.(2002)[5] we choose the pool of wavelets and consists of Haar-like wavelets. The main purpose for their use is that by taking the sum of two ,three or four pixels their output can be calculated very easily shown in fig 4 . We can do this by adding a center-surround type wavelets and corresponding mirror image wavelets that are sensitive to patches symmetric about vertical axis (Fig. 5).

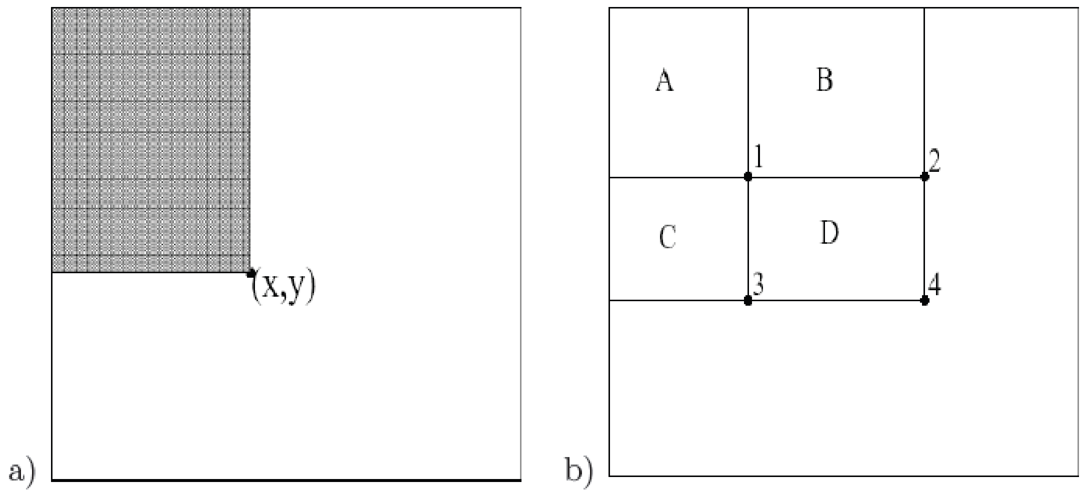

Figure 4: The Integral Image (after Viola & Jones, 2000).: (a) The value of the pixel at (x, y) is the sum of all the pixels above and to the left. (b) The sum of the pixels within rectangle D can be computed as $4 + 1 - (2 + 3)$.

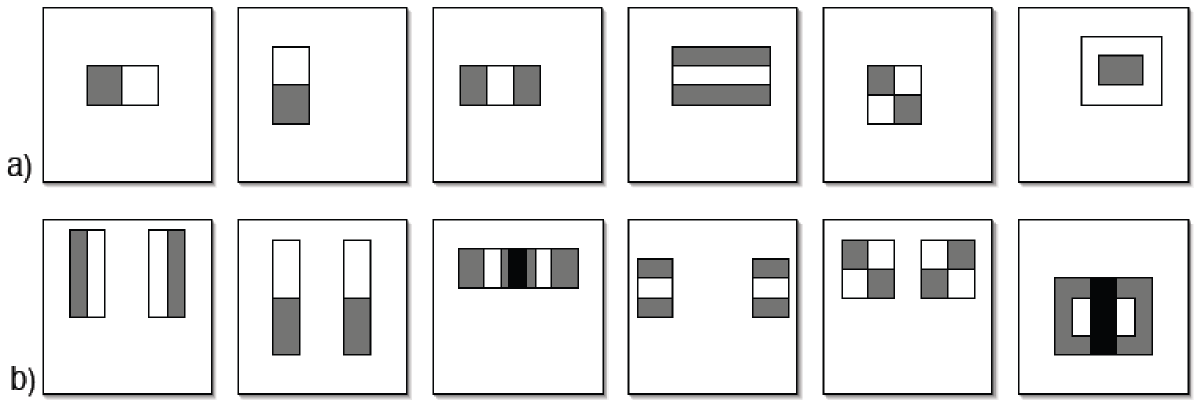

Figure 5: Each wavelet is computed by taking the difference of the sums of the pixels in the white boxes and grey boxes. (a) Wavelets types include those in (Viola and Jones, 2001), as well as a center-surround type wavelet. (b) During the refinement step, the same wavelet types superimposed on their reflection about the Y axis are also possible.

It is very computationally expensive to perform an exhaustive search over all these wavelets—in a 24×24 pixel window, there are over 170,000 possible wavelets of this type. To speed up training, we break the wavelet selection step into two stages (see Fig. 5). First, at each round of boosting, we take a random sample of 5% of the possible wavelets. For each wavelet we find the tuning curve that minimizes the loss function ρ if that particular wavelet were added to the pool of already chosen wavelets(shown in fig 6).In step two, we refine the selection by finding the best performing single-wavelet classifier from a new set of wavelets generated by shifting and scaling the best wavelet by two pixels in each direction, as well as composite wavelets made by reflecting each shifted and scaled wavelet horizontally about the center and superimposing it on the original. Using the chosen classifier as the weak learner for this round ofboosting, the weights over the examples are then adjusted using to the GentleBoost rule. This wavelet selection process is then repeated with the new weights, and the procedure continues until the performance of the system on a validation set no longer decreases.

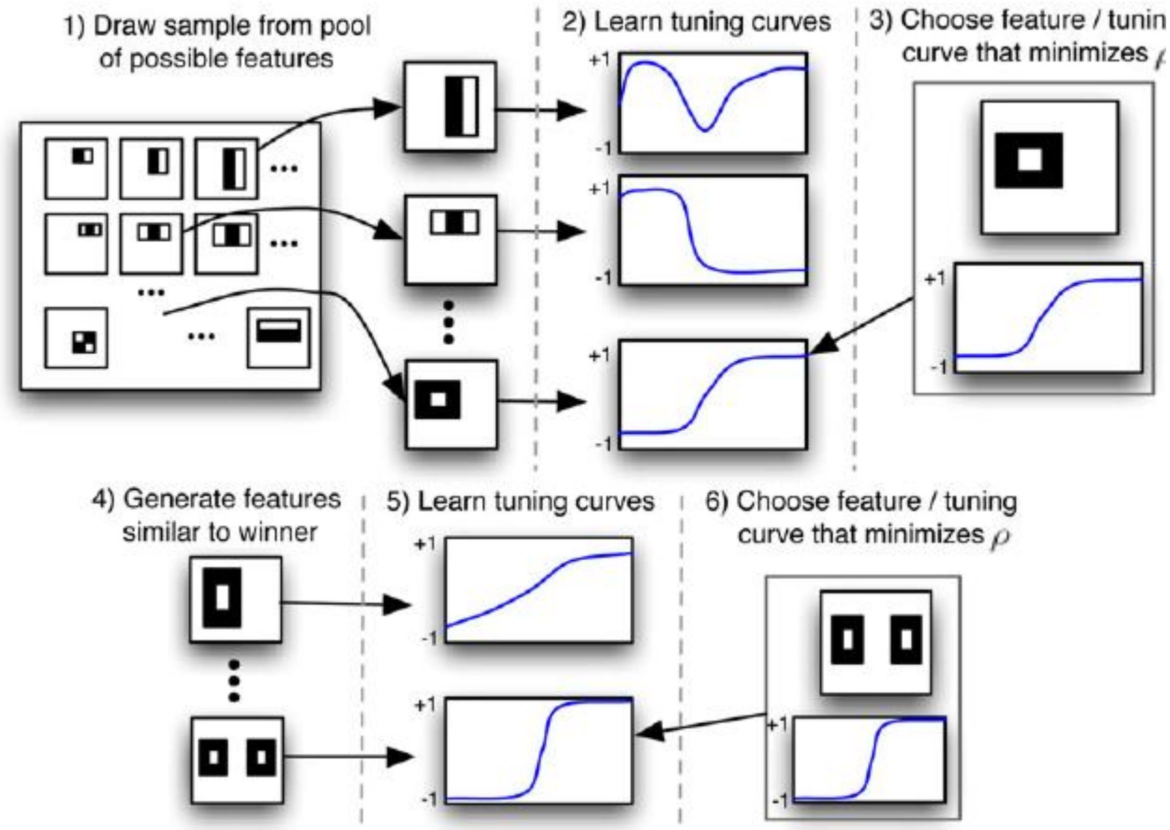

Fig 6 .flowchart for real time object identification

The inference algorithm calls for likelihood ratio models at multiple scales. Likelihood ratios are obtained by linearly scaling for larger image patches, the patches reduced to 24 × 24 pixels and then passed to likelihood ratio model trained on that specified scale. Selecting of Haar-like wavelets for the higher level image representation, this interpolation step can be completed in constant time if the scale factor is an integer. Following Viola and Jones (2001)[4], rather than training a "monolithic" classifier which evaluates all its wavelets before it makes a decision, we divided the classifier into a sequence of smaller classifiers which can make an early decision to abort further processing on a patch if its likelihood-ratio falls below a minimum threshold.Performance on the CMU-MIT dataset (a standard, public data set for benchmarking frontal face detection systems) is comparable to Viola and Jones (2001)[4]. While CMU-MIT contains wide variability in the images due to illumination, occlusions, and differences in image quality, the performance incontrolled environments, such as in the BioID dataset (used later in this study), containing faces that are frontal, focused and well lit, with simple background, is often close to 100% hit rate with few, if any, false alarms. We made the source code for this stage available at http://kolmogorov.sourceforge.net.

Stage II: Eye Detection in the Context of Faces

The first stage in the eye detection system focus on finding common regions of the image plane that contain high probability to contain eyes. The output of the systemis very reliable as well as resistant to false alarms. The system lacks in specifying the precise location of the eyes. To overcome the drawback of the first stage ,we introduce the second stage which detects eye in context of faces. The technique used by this stage is same as used by the previous stage;all patches in the sub region of the face are classified as eye vs non-eye but restricted in both location and scale.

The major advantage of second stage is high accuracy provided. The second stage operates on the regions selected by the first stage. The data received from CMU-MIT face database and Compaq face database was used for training, by including all positive examples of eyes within location 24× 24 pixel patch at canonical scale. To perform positive training samples, take example images by cropping the images in a way that distance from the centre of eye to the left and upper edges of the window are equivalent to a fixed ratio r of the distance from the eyes to the source face. After scaling at a measurement of 24×24 pixels. Assume d to be the distance between the eyes, offset parameter t and scale parameter q, then we formulate an equation $r = q\,(d+td)$.In case of high resolution ,q will give small results with small receptive field. Similarly in case of low resolution, q will give large results with large receptive field. Shifting of t with the location of eye with respect to centre of the patch.

In contextual cascade approach, the background image contain very little information ,when we are within a face then we should select t and q in a way that it maximizes the no.of pixels generated by the face in positive example patch. This approach also allow us to create a criteria for choosing non eye examples .The criteria is based on our prior belief where we take eye location π as normal distribution along with mean of parameters and standard deviation of true eye. From the training set ,scaling with respect to window is done by the face detector. Using the above criteria, we develop two training examples for each eye and six negative training examples ,where we randomly select the negative examples from the set of patches. After collecting data , training eye detector uses GentleBoost technique as this technique achieve excellent performance.figure 7 shows example wavelets and their tunning curves for one of the best eye detector.

In our system , the stage 1 makes no prior assumptions about no.of faces in the image plane.The goal of the stage 1 is face detection.In the stage 2 where we assume one patch describing the left eye and one patch describing the right eye.Case1: To maximize the probability of selecting the correct rendering patch, the optimal inference process should choose a patch which maximize log posterior ratio.Case2:To minimize , take expected squared distance from eye optimal inference by computing the mean of posterior distribution.

Stage 3: Blink detection

Blink detection is similar to face detection and eye detection. Boosted classifer are used for blink detection. On a single patch per image ,binary classification is performed as compared to the previous stages search across multiple patches is performed. As compared to the previous stages , we create 44×24 pixel patch containing location of eye , rotation and scaling is done using simple linear interpolation.Training data was collected from web with the help of eye detector containing 120 each open and closed eye images. After collecting the data cropping and region rotation around the eye to an upright frontal view.GentleBoost is used to select the wavelet and tunning curves for this this task

III. Conclusion

One of the major advantage of generative model is that they give attention to explicit conditions in which inference algorithm give optimal results. In this paper , approach was used to detect eye and eye blink in natural conditions using visible spectrum cameras. During eye detection, we faced several problems as eye blink and eye behavior are human's physiological characteristics. Firstly, we found difficulty in analyzing eye behavior without eye localization. Moreover, it is difficult to develop detectors which can work under general condition and achieve high level of accuracy. This paper has proposed a system specialized in detecting eye in upright frontal pose.

In future, we will put emphasis on other face position for eye blink detection. This system could work under all condition (like general and illumination condition).

References

[1] Javier R.Movellan,Bret fortenberry & Ian Fasel (MPLab TR 2003.03) Machine Perception Laboratory Institute for Neural Computation University of California San Diego.
[2] Ekman, P., Friesen, W., 1978. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto,CA.
[3] J. Phillips, DARPA Symposium on Human ID, Washington, DC, 2003.
[4] Viola, P., Jones, M., 2001. Robust real-time object detection. Tech. Rep. CRL 20001/01, Cambridge Research Laboratory.
[5] Shakhnarovich, G., Viola, P., Moghaddam, B., 2002. A unified learning framework for real-time face detection and classification. International Conference on Automatic Face and Gesture Recognition.
[6] H. Rowley, S. Baluja, T. Kanade, Neural network-based face detection, IEEE Trans. on Pattern Analysis and Machine Intelligence (1998).
[7] Johnson, M. H., 2001. The developmental and neural basis of face recognition: Comment and speculation. Infant and Child Development .
[8] Ji, Q., Yang, X., 2002. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. Real-Time Imaging.
[9] Ji, Q., Yang, X., 2001. Real time visual cues extraction for monitoring driver vigilance. Second International Workshop on Computer Vision Systems.
[10] Farah, M. J., Wilson, K. D., Drain, M., Tanaka, J. N., 1988. What is special about face perception? Psychological Review.
[11] Freund, Y., Schapire, R., 1999. A short introduction to boosting. URL citeseer.nj.nec.com/freund99short.html

