RESEARCH ARTICLE OPEN ACCESS

Efficient Performance Evaluation for Robust Eye Localization System

Krupa Jariwala[1], Dr. Mrs. Upena Dalal[2]
Department of Computer Science and Engineering
Department of Electronics Engineering
SVNIT, Sura
Gujarat – India

ABSTRACT

Accurate eye localization is very important and fundamental to all eye tracking applications. This task is challenging due to variation in appearance, size, shape, facial expression and lighting condition of an eye. Numerous methods have been developed to meet these challenges in past few years. To evaluate the performance of these methods precisely and to make a fair comparison of the accuracy of these methods, it is necessary to categorize them and evaluate them using standard Error Measure Metrics and standard Benchmark Database. We have assessed system architecture of eye localization model, various error measure metrics and benchmark databases to derive their appropriateness in evaluating performance of various methods.

Keywords:- Gaze Tracking, Eye localization**.**

I. INTRODUCTION

Unarguably, the eyes are the most prominent feature of human face. Eye localization is fundamental to all eyes tracking applications. The task of eye localization is different from other applications such as eye detection, eye tracking, gaze estimation and blink detection.

The eye detection confirms the existence of eyes in the given image. Where as eye localization is performed after eye detection that identifies eye position with an error margin of few pixels. The eye tracking is the step after eye localization where time and successive frames are also taken into account. Gaze estimation will find person's focus of attention by analyzing the pupil position in the eye pit. Blink detection is to estimate person's physical state (eg. sleeping, active) by measuring the instances between open and close eyes. However fundamental to all these application is eye localization and the performance of all these applications will improve with robust eye localization.

The task of eye localization is difficult due to variations in appearance, illumination, and environmental condition and motion characteristics of an eye and may lead to designing a complex system. Following are the situations that have highest influence on eye variations which consequence to eye localization algorithms:

Facial Expression: Some facial expression such as laughing or crying may cause eyes to partially or fully close and may deform the shape of an eye.
Occlusion: In real life scenarios, eyes are frequently occluded by hair, sunglasses, and specs and make the localization difficult.

Pose: The appearance of the eyes will be different when a face has frontal, profile, upside down pose. The eyes may also be occluded in a profile face.

Lighting condition: The indoor and outdoor lighting condition, shadow (e.g walking thru the street with tall trees and buildings) and reflection will create significant variations.
Imaging quality: The real world circumstances, poor image quality (noise and blur) changes the eye appearance significantly.

II. EYE LOCALIZATION AND SYSTEM ARCHITECTURE

Eye localization methods are derived based on various features available from eyes. Following are the main categories available for classification of eye localization techniques for developing an effective eye localization model.

Eye characteristics Measurement: These methods use eye features such as shape, strong intensity contrast etc. as a measure and perform eye localization.

Learning statistical appearance model: These methods derive an eye model from large training images. These training images should have possible eye feature variations. The eye features are then extracted from that derived eye model.

Exploiting structural information: These methods uses the geometrical features of an eye and other facial features to derive information. This information may be integrated to statistical eye model to improve performance.

These methods usually divide the task and employ a stage wise implementation of eye localization system. The first stage should be to localize face region. Once the face is detected, the eyes can then be localized from the face window in second stage. Fasel[1] proposed face localization and facial feature extraction model through Hybrid facial expression analysis systems by combining several facial expression analysis methods. Inspired by Fasel[1], Song[2], derived a system architecture for eye localization to indicate a generalized pipeline for an eye localization task into three components as (1) Face processing (2) Eye Feature extraction and representation (3) Eye Localization[2].

Face Processing: The main goal of Face Preprocessing stage is to narrow down the search region to simplify the task of feature extraction and remove noise. Narrowing down the search region will improve the efficiency of eye localization. The Face processing is done by locating and segmenting faces from mixed backgrounds. A modern face detector such as Voila jones[3] can do this task efficiently. The output of a face detector is a face region possibly with variation in shape, scale, rotation and lighting. Therefore these variations needs to be removed and search regions may also be narrowed down to right and left eye prior to second stage of eye feature extraction.

Eye Feature Extraction and Representation: The search for eye region is carried out based on characteristic, appearance or structure of an eye. This approach works well in most cases, except when the location of the eye is occluded or affected with pose variations. In such cases, the search space may be expanded to include all possible images or different templates with different scales to improve the search with appropriate eye model.

Eye localization: Once the eyes are detected from the face region, using appropriate eye model, a post processing is then done to further evaluate the most appropriate eye locations. The eye coordinates are then generated as an output, which is a final result of chosen eye localization model.

III. PERFORMANCE EVALUATION

To evaluate eye localization model precisely and make a fair judgment among various methods, ideally the performance should be measured on benchmark database using standard error measure metrics, and follow a standard evaluation protocol. However, in reality the algorithms are evaluated in different ways with variations in error measure metrics, databases, training samples, testing samples, etc., which makes it difficult to fairly compare eye localization results. We have reviewed the Error Measure Metric and use of Benchmark Databases for efficient performance evaluation of various eye localization methods.

A. Error Measure Metrics for Performance Evaluation

The design of the ground truth: The ground truth location of an eye is defined as a representation of the agreed correct result of the ideal eye localization method [2]. It is the base for all performance comparisons among the methods to be evaluated; hence the design of the ground truth is very crucial. The distance between the predicted position and ground truth location of the eye then measures the localization accuracy of the proposed method.

Normalized error measurement: Once the ground truth information is available, the next step is to measure the normalized error, which is accuracy measure for the estimated eye centers. It indicates the error obtained by the worst of both eye estimations. The most commonly used measurement is the normalized eye localization error proposed by Jesorsky[4] which is defined in terms of the eye center positions according to

$$e \leq \frac{1}{d} \max\left(e_l, e_r\right) \tag{1}$$

where d is the ground-truth position and e_l and e_r are the Euclidean distances between the detected eye centers and the ground-truths[2]. While analyzing the performance of various approaches for eye localization, this measure has the following characteristics:

1. $e \leq 0.25$ ≈ Within eye center and the eye corners.
2. $e \leq 0.10$ ≈ Within diameter of the iris.
3. $e \leq 0.05$ ≈ Within diameter of the pupil.

Performance of $e \leq 0.25$ indicates eye detection and the estimated center might be located within the eye center and eye corner. This may be appropriate for applications that use the overall eye position such as face matching, where comparing the performance for $e \leq 0.25$ is appropriate. To indicate accurate eye center detection inside the pupil, performance higher than $e \leq 0.25$ is necessary. The approach that is used for eye tracking must produce good results for $e \leq 0.05$ to indicate precise eye center detection. Hence, for eye tracking applications a performance for $e \leq 0.05$ is essential.

Many times the normalized error measurement is used, where the error degree is provided by the measures $e_{better} \leq min(e_l , e_r)/d$ and $e_{avg} \leq (e_l + e_r)/2d$ to give an upper bound and an averaged error. For the face under out of plane rotation, the two eyes distance cannot reflect the actual face scale, and the normalized error measurement may be biased. In such case more general evaluation measures such as mean and variance could also be used [2].

W*orse Eye Characteristic (WEC)* curve is also analysed for performance comparison for different values of the normalized error e. The area under the WEC can be used for this purpose[6]. The worse eye characteristic (WEC) curve is similar to the well-known *Receiver Operator characteristic (ROC)* curve which is a graphical plot to illustrates the performance of a system. It is created by plotting the true positive rate against the false positive rate at various threshold values[5]. Additionally, the rank of various methods are also evaluated for comparison. The rank of the method is inversely proportional to the area under the WEC[6].

B. Benchmark Database for Performance Evaluation

There are many standard face databases available for evaluation. However, most of them are not principally targeting eye localization problem. They are collected under well-controlled laboratory conditions with normal lighting, neutral expression and high image quality. Many databases contain remarkable eye pattern variations and distributed with ground truth information, hence used by well-known methods to evaluate accuracy of localization. We have examined databases containing visual face images and providing related ground truth information for eye and gaze detection for our purpose. Inspired by [17], we have derived table 1 that shows the main characteristics of some popular database and table 2 that summarizes the main properties of these benchmark databases. Figure 1 shows example images from some popular database.

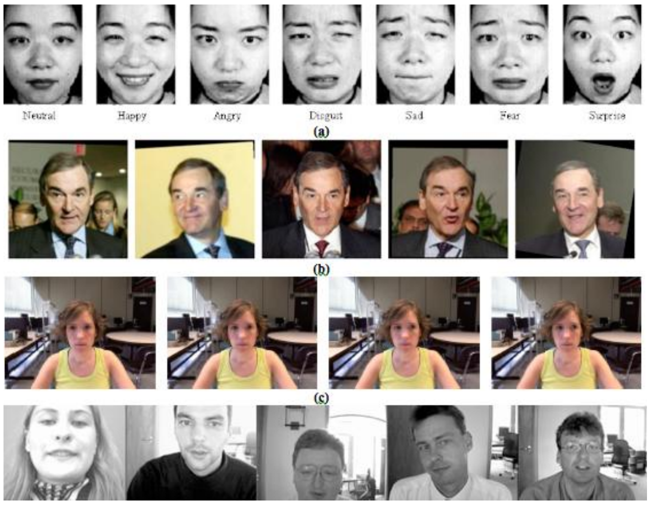

Fig 1 : Example images from (a) JAFEE[9] (b) LFW[12] (c) Gi4E[15] (d) BioID [16] database

From Table 1 and 2, it is observed that most of the databases are designed for face recognition and not eye localization as they have the ground truth information for pose, eye position, expression label, identification of subjects but they do not have ground truth information of eye coordinates which is very crucial for eye localization. The BioID[16] dataset is found to be most appropriate and much more challenging for eye localization for following reasons :

1. The database consists of 1521 grey level images of 23 different subjects with ground truth information for eye localization (eye coordinates) for every subject (The left and right eye centers are marked and provided with the images).
2. There is variation in the background, position, pose of the subjects and illumination conditions, which are comparable to outdoor scenes and real time scenario.

3. There is variation in the eye appearance as some subjects are wearing glasses, have curled hair near to the eye centers, the eyes are closed, head is turned away from the camera or strongly affected by shadows, or eyes are completely hidden by strong reflections on the glasses.
4. The image quality and the image size (384 x 286) is approximately equal to the quality of a low-resolution webcam.
5. The performance on database drops about 7% compared to that on the relatively simple databases [2].

This indicates that complex unrestrained conditions are present as a challenge in BioID database and a method will perform better that provides superior solution to these complex scenarios.

IV. CONCLUSION

The changes in geometric, photometric and motion characteristics of an eye require design of complicated eye localization method. For fair assessment and impartial evaluation of various eye localization methods, we examined various Error Measure Metrics and Benchmark Database on which test may be carried out. We found that the ground truth information for eye localization is necessary for correct error measurement. However, most face databases provide ground truth information for face recognition, which is not suitable for eye localization task. Moreover, these face images have variation in pose and expression, which is not essential condition for eye localization. Considering this fact, BioID database is found to be most suitable and challenging database for eye localization task, containing variation in illumination, background and face size and providing a real test scenario. The normalized eye localization error proposed by Jesorsky[4] is found to be suitable and widely used error measure metrics.

REFERENCES

[1]. B. Fasel, J. Luettin, Automatic facial expression analysis: a survey, Pattern Recognition 36 (1), 259–275, 2003.

[2]. Fengyi Song, XiaoyangTan, Songcan Chen, Zhi-Hua Zhou, “A literature survey on robust and efficient eye localization in real-life scenarios, Journal of Pattern Recognition, Elsevier, (46), 3157–3173, 2013.

[3]. Viola, P. & Jones, M, “Robust Real-Time Face Detection” International Journal of Computer Vision 57(2), 137-154, 2004

[4]. Jesorsky O, Kirchberg K, and Frischholz R, “Robust face detection using the Hausdorff distance” In Proceedings of the 3rd AVBPA, LNCS, Springer pages 90–95, 2001.

[5]. http://en.wikipedia.org/wiki/Receiver_operating_characteristic

[6]. Fabian Timm and Erhardt Barth, "Accurate Eye Centre Localization By Means of Gradients". In Proceedings of the Int. Conference on Computer Theory and Applications (VISAPP), volume 1, pages 125-130, Algarve, Portugal, 2011.

[7]. www.itl.nist.gov/iad/humanid/feret/feret_master.html

[8]. www.face.nist.gov/frgc/

[9]. www.kasrl.org/jaffe.html

[10]. www.vasc.ri.cmu.edu/idb/html/face/

[11]. www.vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html

[12]. www.vis-www.cs.umass.edu/lfw/

[13]. www.jdl.ac.cn/peal/

[14]. S. Asteriadis, D. Soufleros, K. Karpouzis, S. Kollias, “A Natural Head Pose and Eye Gaze Dataset”, International Conference on Multimodal Interfaces (ICMI 2009)

[15]. V. Ponz, A. Villanueva, and R. Cabeza, “Dataset for the evaluation of eye detector for gaze estimation,” ACM Conference Ubiquitous Computing 2012, pp. 681-684.

[16]. www.bioid.com/About/BioID-Face-Database

[17]. http://www.ecse.rpi.edu/~cvrl/database/database.html

TABLE 1 : Main characteristics of some popular database

Database	Source	Characteristics	Purpose
FERET - The Facial Recognition Technology Database[7][17]	Sponsored by the Defence Advanced Research Products Agency (DARPA).	This database contains 7 categories of face pose (Frontal, quarter-left, quarter-right, half-left, half-right, full-left, full-right). Face images have moderate expression change under controlled conditions.	To evaluate the face recognition technology, face pose estimation and general evaluation of eye detection.
FRGC - Face Recognition Grand Challenge Database[8][17]	Jointly sponsored by several government agencies to improve face recognition technology.	This database has face images with variation in lighting condition. The images have higher resolution and frontal view; hence most methods perform well on this database.	To evaluate face recognition and eye detection.
JAFFE - The Japanese Female Facial Expression Database[9]	Developed at Psychology Department in Kyushu University	This database has 7 facial expressions (Angry, Disgust, Fear, Happy, Sad, Surprise & Neutral) of each person. The expression variations are mostly with eyes open, hence suitable for those methods that measure the appearance or intensity characteristics of eyes.	To evaluate facial expression.
CMU Face (Frontal and Profile) Database [10][17]	Developed at The Robotics Institute of Carnegie Mellon University.	This database combines images collected at CMU and MIT. It includes frontal and profile face with various facial expression and lighting condition.	To evaluate face detection, eye detection and facial feature detection.
Yale Face Database B [11][17]	Constructed by Yale University.	This database has 9 various face pose with neutral facial expression, 64 lighting conditions with 1 ambient illumination	To evaluate face recognition, face pose estimation, and eye detection.
LFW (Labelled Faces in the Wild) Database [12][17]	Built by University of Massachusetts, Amherst.	The images of the faces in this database are collected from the web. There is large variation in clothing, pose and background in LFW. Each face in LFW has been labelled with the name of the person pictured.	To evaluate unconstrained face recognition.
CAS-PEAL Face Database[13][17]	Developed by Chinese Academy of Science.	This database has variation of 21 face pose angles (vertical: up, middle, and down and horizontal: left to right) 6 facial expressions (neutral, eye closing, frown, smile, surprise, and mouth open) and 15 lighting conditions. Moreover objects are wearing 3 kinds of glasses and 3 kinds of caps.	To evaluate the face recognition, face pose estimation, facial expression recognition and eye detection
HPEG (Head Pose and Eye Gaze) Dataset [14]	Developed by S. Asteriadis[14] at National Technical University of Athens.	This dataset contains video sequences for 10 subjects where the subjects move their eyeball looking at the camera with frontal and rotated head pose. It contains the yaw and pitch rotation information for the head.	To evaluate head pose and gaze estimation.
Gi4E dataset[15]	Developed by Gaze Interaction Group at Public University of Navarre in Spain.	This dataset images are captured in indoor environments with different backgrounds and illumination conditions. There are 13 images for every subject where in 12 images the user gazes at different points in the screen and last image is with closed eye.	To evaluate iris center and eye corner detection
BioID Database [16]	Constructed by Human Scan company	This database has large variation in illumination, background and face size. The eye positions have been set manually and provided with the images for accuracy calculation. It provides real time scenario hence it is much challenging database for eye localization.	To evaluate face detection, eye detection and localization.

TABLE 2 : Properties of Benchmark databases for eye localization.

Database / Properties	No. of Subject	No. of images	Grey/ Colour	Resolution	Face pose	Facial expression	Illumination	Ground Truth Information
FERET [7][17]	1199	14051	8-bit grey	256*384	7 pose	Moderate change	Controlled	- Positions of eyes, nose, and mouth. - Subject identification.
FRGC [8][17]	Training set - 222 Validation set - 466	Training set-12776 Validation set-4007*8	Colour	1704*2272 or 1200*1600	Frontal view	Neutral & smiling	Controlled & Uncontrolled	- Positions of eyes, nose, and mouth. - Subject identification.
JAFFE[9]	60	213	Grey	256*256	Frontal	7 Facial Expression	Controlled	-Label of expression such as angry, Happy, Sad, Surprise, Disgust, Fear and Neutral.
CMU [10][17]	N/A	Frontal Face-169 Profile face-202	8-bit grey	N/A	Frontal & Profile	Various	Various	- For each frontal view face: Positions of eyes, nose tip, mouth corners, and mouth centre. - For each Profile Face: Positions of eye corner, eye, nose, nose tip, mouth corner, mouth centre, chin, earlobe, and ear tip.
Yale B [11][17]	10	5760	Grey	640*480	9 Poses	Neutral	64 lighting & 1 ambient illumination	- Subject identification. - Face pose. - Illumination positions. - Coordinates of eyes and mouth (frontal view). - Coordinates of face centre (other views).
LFW[12][17]	5749	13233	Colour	250*250	Various	Various	Various	- Subject identification.
CAS-PEAL [13][17]	1040	30900	8-bit grey	360*480	21 pose angle	6 Facial Expression	15 lighting condition	- Positions of eyes. - Subject identification. - Face pose angles, - Facial expression labels. - Illumination position.
HPEG[14]	10	Video of 2 sets (10 recordings each)	Colour	640*480	Various	Neutral	Controlled	- Head pose associated with ground truth for Yaw and Pitch angles (in degrees). - eye gaze directionality. - Timestamps.
Gi4E[15]	103	1339	Colour	800*600	12 gaze point on screen	Neutral	Various	-Iris centre and eye corner points.
BioID[16]	23	1521	Grey	384*286	Frontal	Various	Various	- Coordinates of 20 feature points such as (right and left eye pupil, eyebrow, nose, mouth, chin, lip etc.)