Quality in Face and Iris Research Ensemble (Q-FIRE)

P. A. Johnson, P. Lopez-Meyer, N. Sazonova, F. Hua, and S. Schuckers

Abstract—**Identification of individuals using biometric information has found great success in many security and law enforcement applications. Up until the present time, most research in the field has been focused on ideal conditions and most available databases are constructed in these ideal conditions. There has been a growing interest in the perfection of these technologies at a distance and in less than ideal conditions, i.e. low lighting, out-of-focus blur, off angles, etc. This paper presents a dataset consisting of face and iris videos obtained at distances of 5 to 25 feet and in conditions of varying quality. The purpose of this database is to set a standard for quality measurement in face and iris data and to provide a means for analyzing biometric systems in less than ideal conditions. The structure of the dataset as well as a quantified metric for quality measurement based on a 25 subject subset of the dataset is presented.**

I. Introduction

FACIAL and iris recognition are two commonly used non-contact biometrics. These two modalities have both been well established in controlled environments, where ideal lighting, resolution, and gaze angle are well controlled. With tasks of semi-cooperative or un-cooperative acquisition, it is difficult if not impossible to regulate these ideal conditions. Various applications in the security and law enforcement fields would benefit greatly from improvements in facial and iris recognition performance in non-ideal conditions.

Recent work in facial recognition technology is moving in the direction of long range surveillance and wide area monitoring, such as the long range facial recognition system for maritime monitoring proposed by [1]. In these areas, image quality assessment is necessary to ensure adequate system performance in these non-ideal scenarios. In long range, high magnified imaging, distortion due to magnification blur becomes a factor as demonstrated by [2]. Additional work by Medioni et al. [3] is focused on long range identification using 3-D face modeling and studying head rotation. Messer et al. [4] presented an extended M2VTS database labeled XM2VTSDB, which incorporates speech and face videos with varying gaze angle to be publicly available.

Iris recognition is an area of research, which has found great success due to its statistical uniqueness among different people [5]. Quality assessment of iris data is a growing area of research, reflected by algorithms such as the Laplacian of Gaussian operator [6] and Daugman's 2-D Fourier analysis method [5] for image blur assessment. The need for specialized acquisition systems for capture of low blur iris images is discussed by Grabowski et al [7]. It is stated by Kalka et al. [8] that there is a need for a publicly available iris dataset including a broad range of quality factors, such as defocus blur, motion blur, off-angle, occlusion, specular reflection, lighting, and resolution for iris recognition system assessment. Available datasets composed of images of varying quality include the Iris Challenge Evaluation (ICE) [9], the Multiple Biometrics Grand Challenge (MBGC) [10], and UBIRIS, a visible spectrum noisy iris database [11].

This paper presents the Quality in Face and Iris Research Ensemble (Q-FIRE) dataset of face and iris videos obtained at distances of 5 to 25 feet and of varying quality. The quality levels of the videos are of high, medium, and low illumination, varying out-of-focus blur, off-angle gaze, motion blur, as well as blinking and multiple faces. The dataset currently consists of 195 subjects, each with two visits. The main contribution of this dataset is to be a standard for the evaluation of face and iris recognition quality and recognition algorithms for non-ideal environments. Multimodal datasets, similar to the one presented here are summarized in Table I. The novel approach to this dataset is video at a distance for both iris and face as well as that the quality of the data in terms of illumination, blur, gaze angle, occlusion, and motion is clearly defined in terms of a metric controlled at acquisition and is consistent throughout the dataset. In this paper, the actual quality metrics will be measured from a randomly selected 25 subject subset due to the fact that manual processing is required at the present time and it was not feasible to use the entire dataset. The dataset is currently being used by NIST IREX II: Iris Quality Calibration and Evaluation (IQCE) and is scheduled to be released at the completion of IREX II [12].

II. Baseline Dataset

The dataset consists of two parts. The first part is a set of baseline images collected using the MBARK system, consisting of an OKI IRISPASS EQ5016A iris camera and an Olympus C-5060 camera for face images. The iris images are 640×480 pixels with approximately 200 pixels across the iris and the face images are 2592×1944 pixels with approximately 400 pixels between the eyes. There are 2 left and 2 right images for iris and 10 images for face per visit. When a subject is wearing glasses, the number of face images is doubled, so that there are 10 with glasses and 10 without glasses. The background for the face images is controlled with 18% reflectance gray color. This part of the dataset establishes a comparison between a constrained dataset from a commercially available system and an unconstrained dataset.

This material is based upon work supported by the National Science Foundation under Grant Nos. 0325333 & 0325640, the Department of Homeland Security, and the Center for Identification Technology Research.

The authors are with Clarkson University, Potsdam NY, U.S.A (e-mail: johnsopa@clarkson.edu; lopezmp@clarkson.edu; nsazonova@clarkson.edu; sschucke@clarkson.edu).

III. Q-FIRE Dataset

The Q-FIRE dataset is set of face and iris videos spanning 5 to 25 feet and with various quantifiable degradations in quality. Where possible, quality ranges from high to low were chosen based on published criteria. There are a total of 28 face videos and 27 iris videos for each of 2 sessions per subject. Each session totals nearly 16 gigabytes.

TABLE I
OVERVIEW OF MULTIMODAL DATASETS

Dataset	Number of Subjects	Number of Sessions	Modalities	Data Type
XM2VTS [4]	295	4	Face, Voice	Video Audio
BiosecurID [13]	400	4	Face, Iris, Fingerprint, Voice, Hand, Hand Writing, Signature, Keystroke	Image Audio
BioSec [14]	250	2	Face, Iris, Voice	Image Audio
BIOMET [15]	91	3	Face, Fingerprint, Voice, Hand, Signature	Image, Audio
MCYT [16]	330	2	Fingerprint, Hand Writing	Image

Videos were captured at five different distances from the cameras; 5, 7, 11, 15, and 25 feet. Each video is six seconds in length, captured at a frame rate of ~25 frames per second. The videos that have variations in illumination to characterize contrast as well as the videos for motion blur were collected simultaneously for face and iris. The methods for the collection of all face and iris videos in this dataset are described below.

A. Face Data

Face videos are captured using a Canon VIXIA HFS100 HD camcorder, 24 Mbps, 1920×1080, 30p progressive camcorder.

Resolution The main distances of interest for face data are 5, 15, and 25 feet, representing high, medium, and low resolution respectively. The standard for determining the resolution of a face image is to measure the number of pixels between the centers of the eyes [17]. For this dataset, the target value ranges for resolution are 350 to 800 pixels for high, 120 to 300 pixels for medium, and 70 to 100 pixels for low resolution, as specified by NIST [17].

Illumination Three different levels of face contrast are achieved by three different illumination settings. For high light, the overhead fluorescent lights are on in addition to two incandescent 500 Watt studio lamps, arranged in two-point balance lighting, each at $30° - 40°$, diffused with umbrellas, and positioned 5 feet from the subject. The medium lighting condition is met by turning off the incandescent lighting and leaving the overhead fluorescent lighting on. Low lighting is then achieved by turning off the overhead lighting and using only a single incandescent 500 Watt studio lamp with diffusing umbrella, positioned at $90°$ to form a shadow on one side of the face. The target ranges for face contrast measurements are determined by measuring the width of the histogram of the face image and are greater than 128 gray levels for high light, less than 128 gray levels for medium light, and shadowing on one side of the face to create a lower quality face image for low light.

Out of Focus Blur In order to create face videos of varying out-of-focus blur, videos were captured at 5, 15, and 25 feet while adjusting the focal plane of the camcorder across the full range over the 6 seconds of the video. A modulation transfer function (MTF) chart is placed beside the subject's head to allow for the quantification of the level of blur in each frame of the video.

Face Angles Variations in subject gaze angle are achieved using fixed markings at each distance for the subject to look at. The different gaze angles are straight, left and right at $70°$, and up and down at $30°$. The subjects are asked to turn their head, keeping their eyes in line with their face, and hold each position for approximately one second. Face gaze angle videos are captured at 5, 15, and 25 feet.

Motion Blur Motion blur videos are captured at 7 and 15 feet and consist of two different speeds of walking. The pace is set by a metronome for the subjects to step in sync with. The two different speeds are 66 beats per minute and 120 beats per minute, which corresponds to approximately 1.5 miles/hour and 2.75 miles/hour respectively. The camcorder is first set to manual focus while the subject is standing in position (7 or 15 feet) in order to set the focal plane and avoid any effects of the autofocus on the motion blur in the videos.

Multiple Face To study the implications of facial recognition in a crowded environment, a video with multiple faces present is available. In this dataset there are videos for each subject with two additional individuals present in the scene, where the additional subjects take eight different positions around the subject of interest to simulate a crowded environment. These videos are also captured at 5, 15, and 25 feet.

B. Iris Data

Iris videos are captured with a Dalsa 4M30 infrared camera with a Tamron AF 70-300mm 1:4.5-5.6 LD DI lens. A Sigma APO 300-800mm F5.6 EX DG HSM lens is used at distances of 15 and 25 feet.

Resolution The main distances of interest for iris data are 5, 7, and 11 feet, representing high, medium, and low resolution respectively. The standard for determining resolution of an iris image is measuring the number of pixels across the diameter of the iris. The target ranges for iris resolution are 280 to 300 pixels for high, 200 to 220 pixels for medium, and 100 to 120 pixels for low. The high and low resolution ranges are those suggested by Daugman [4].

Illumination Iris contrast is achieved by three different illumination settings much like face contrast. The lighting used for the iris collection is made up of 10 LED based infrared light arrays (5 Watts each) mounted on a movable portal, positioned 2 feet from the subject. For the high lighting condition, 8 LED arrays are on, with 6 LED arrays for medium light and 4 LED arrays for low light. For 15 and 25 feet, 10, 8, and 6 LED arrays are used for high, medium, and low light respectively. The measurements for iris contrast are taken at the average gray level difference across the sclera-iris and iris-pupil boundaries. The targets for these contrast measurements are summarized in Table II.

TABLE II
IRIS CONTRAST MEASUREMENT TARGETS IN NUMBER OF GRAY LEVELS ACROSS BOUNDARY

Contrast	Sclera-Iris Boundary	Iris-Pupil Boundary
High	70	60
Medium	55	45
Low	40	30

Out of Focus Blur Variation in out-of-focus blur was obtained by turning the focus ring of the lens during capture of iris video. The focal plane is adjusted over a sufficient range from out of focus to in focus to back out of focus creating frames of varying blur to be selected from. This data is obtained in conjunction with the contrast data to save collection time and hard drive space. To quantify the amount of blur in each frame, a measure known as LoGscore can be implemented which is the average of the output of a Laplacian of Gaussian filter and can be used as the relative sharpness of each of the frames in a video. The Laplacian of Gaussian operator can be used to assess iris image quality in terms of focus as described by [6].

The effect of out-of-focus blur on iris recognition performance was studied from images captured with out-of-focus blur produced at acquisition. To quantify blur, a method was introduced to quantify real out-of-focus blur and compare it to the gold standard of the modulation transfer function (MTF) of a black/white standard chart. The proposed sharpness measure is based on the Laplacian of Gaussian (LoG) operator using non-segmented iris images from a video sequence with varying focus. This measure offers a good approximation of the gold standard MTF. The effect of the 9 blur levels on iris recognition performance was examined. The results have shown that even for very blurry images (sharpness below 40%) there is sufficient information to achieve EER of 3-8.5%.

Eye Angles Videos of variations in eye gaze angle are collected in similar manner to the face gaze angle videos. The subject is asked to look straight, left and right at 30°, and up and down at 30°, identified by fixed markers. For this case, unlike face, the subjects are asked to turn only their eyes, keeping their heads fixed straight ahead. These videos are collected at 5, 7, and 11 feet.

Motion Blur Motion blur videos are collected for iris at the same time as for face. All 10 LED arrays are on for the collection of the walking videos and the exposure time is reduced. This allows for the best tradeoff between illumination and motion blur.

Eye Occlusion A video representing iris occlusion is collected at all distances at the high lighting condition. The subjects are asked to blink repeatedly during the 6 seconds of video capture. Frames are then available at various amounts of eyelid closure.

TABLE III
DESCRIPTION OF FULL DATASET

	Distance (ft)	Description
Iris		
Out-of-focus blur	5, 7, 11, 15, 25	Full range of blur
Illumination	5, 7, 11, 15, 25	Low, Med, High
Angles	5, 7, 11	Straight, Left, Right, Up, Down
Occlusion	5, 7, 11, 15, 25	6 seconds of blinking
Motion blur	7, 15	Slow/Fast Walking
Face		
Out-of-focus blur	5, 15, 25	Full range of blur
Illumination	5, 7, 11, 15, 25	Low, Med, High
Angles	5, 15, 25	Straight, Left, Right, Up, Down
Occlusion	5, 15, 25	Multiple Faces
Motion blur	7, 15	Slow/Fast Walking

IV. RESULTS

Verification of the target parameters in this dataset are demonstrated using a 25 subject subset. Many of the quality factors require manual measurements, and so an analysis of the entire dataset is not feasible at this point in time. Automated methods for quality measurement are currently being developed to allow for analysis of the entire dataset. The full dataset consists of 62% male subjects and 38% female subjects. The eye color distribution is 49% brown, 27% blue, 18% hazel, and 8% green. All eye colors were self-reported by the subjects. The percentage of subjects over 30 years of age is 22%. The percentage of subjects wearing glasses is 21% and the percentage wearing contact lenses is 12%. Table IV summarizes the demographics of the 25 subjects. Example frames from the face videos are presented in Figures 1-5 and examples of frames from iris videos presented in Figures 6-10.

TABLE IV
DEMOGRAPHICS FOR 25 SUBJECT SUBSET: SUBJECT DISTRIBUTIONS BY EYE COLOR

Gender	Total	Brown	Blue	Hazel	Green
Male	15	5	7	2	1
Female	10	6	3	1	0

Two measures that were verified in this data are the resolution and contrast of the face and iris video frames. To analyze this, frames were extracted from the face and iris videos portraying variations in illumination.

For face, the segmentation of the face frames as well as the measurement of the number of pixel between the eyes for resolution was conducted using the FACEIT software. The contrast was then measured by determining the width of the histogram of the face image using MATLAB. The average results for resolution using the 25 subject subset are presented in Figure 11, where it is shown that the resolution targets are met at 5, 15, and 25 for high, medium, and low respectively. The results for contrast measurements, averaged for the 25 subjects are presented in Figure 12. It is observed that the contrast values for the low lighting condition are higher than the other two conditions, which is to be expected given the shadow on half of the face. The medium lighting condition at 5 feet created slightly higher contrast values than expected, but the remainder of the results for high and low contrast fall above and below a histogram width of 128 gray levels.

For analysis of the iris videos, the most in-focus frame was selected from each video using the relative LoGscore measurement in the neighborhood of the iris being analyzed. The number of pixels across the iris was measured manually and the contrast was determined by calculating the changes in average gray level value across the sclera-iris and iris-pupil boundaries. All processing of iris data was conducted in MATLAB. The resolution results for the iris data, averaged over the 25 subjects are presented in Figure 13, where resolution measurements at 5, 7, and 11 feet are shown to closely match the target ranges. The iris contrast results, averaged for the 25 subjects are presented in Figures 14 and 15, as well as being presented in Table V. The iris contrast results are shown to be consistent across all distances and to match well to the targets. Table VI shows the standard deviations for the iris resolution and contrast measurements. The standard deviation for the contrast measurements is observed to be quite high. This is a result of a variation in iris contrast across different eye colors.

(a) (b) (c)

Fig. 1. Example images from illumination videos; high (a), low (b), and shadow (c).

Fig. 2. Example images from out-of-focus blur video.

Fig. 3. Example images from off-angle gaze video.

TABLE V
RESULTS FOR IRIS BOUNDARY REGION CONTRAST AND RESOLUTION MEASUREMENT AVERAGES

Illumination	Boundary	5 Feet	7 Feet	11 Feet
High	Sclera-Iris	69	70	68
	Iris-Pupil	55	56	52
Medium	Sclera-Iris	56	55	50
	Iris-Pupil	43	43	38
Low	Sclera-Iris	42	40	38
	Iris-Pupil	30	30	30
Resolution		285	205	125

TABLE VI
STANDARD DEVIATION RESULTS FOR IRIS BOUNDARY REGION CONTRAST AND RESOLUTION MEASUREMENTS

Illumination	Boundary	5 Feet	7 Feet	11 Feet
High	Sclera-Iris	24	25	28
	Iris-Pupil	19	16	19
Medium	Sclera-Iris	22	19	19
	Iris-Pupil	16	15	15
Low	Sclera-Iris	14	14	16
	Iris-Pupil	10	10	11
Resolution		21	13	8

(a) (b) (c) (d)

Fig. 4. Example images from motion blur videos; 7 feet slow (a), 7 feet fast (b), 15 feet slow (c), and 15 feet fast (d).

Fig. 5. Example images from multiple faces video.

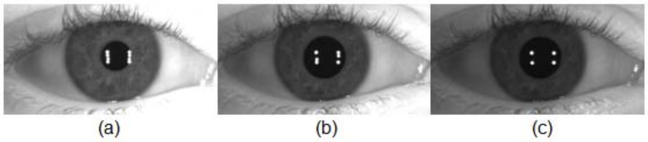

Fig. 6. Example images from iris illumination videos; high contrast (a), medium contrast (b), and low contrast (c)

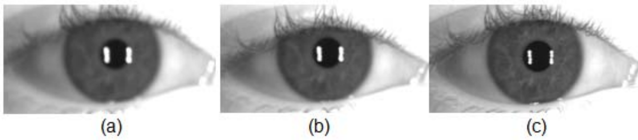

Fig. 7. Example images from iris out-of-focus blur video; high blur (a), medium blur (b), and low blur (c).

Fig. 8. Example images from iris motion blur; slow waling at 7 feet (a), fast walking at 7 feet (b), slow walking at 15 feet (c), fast walking at 15 feet (d).

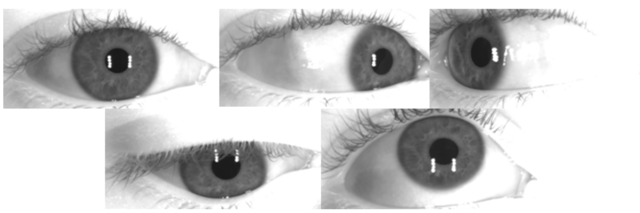

Fig. 9. Example images from iris angle video.

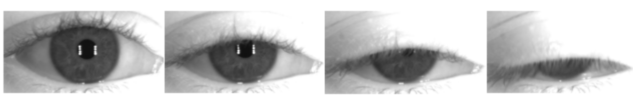

Fig. 10. Example images of blinking video.

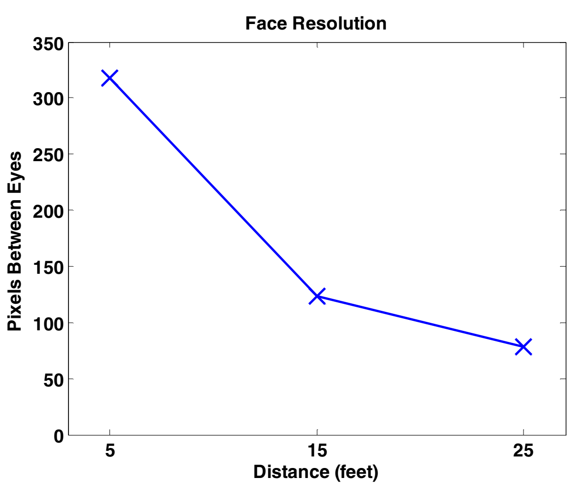

Fig. 11. Face resolution measurements averaged over 25 subjects for 5, 15, and 25 feet.

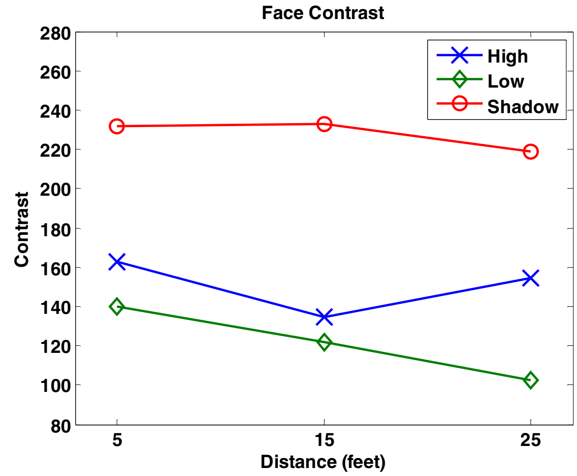

Fig. 12. Face contrast measurements averaged over 25 subjects for 5, 15, and 25 feet.

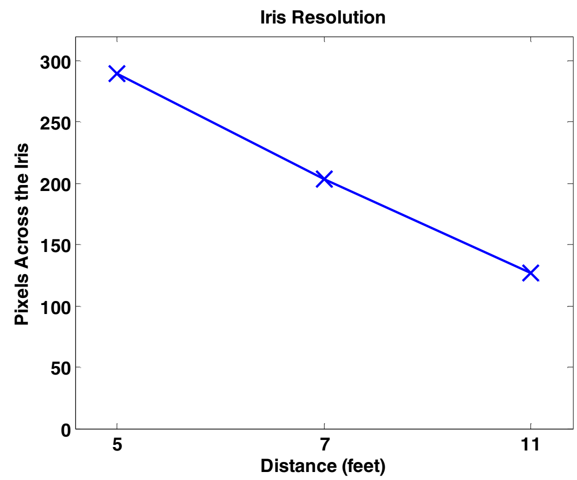

Fig. 13. Iris resolution measurements averaged over 25 subjects for 5, 7, and 11 feet.

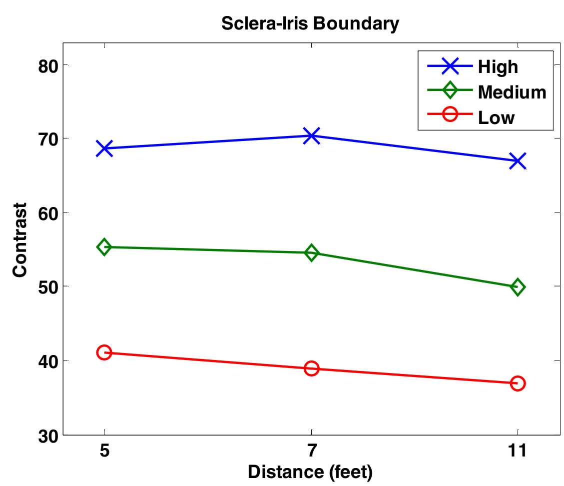

Fig. 14. Iris contrast in sclera-iris boundary averaged over 25 subjects for 5, 15, and 25 feet.

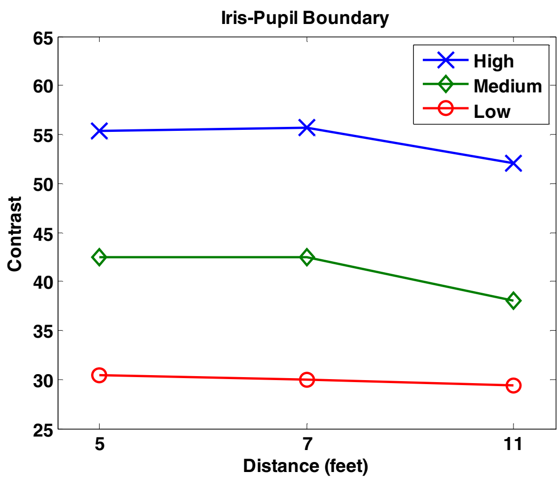

Fig. 15. Iris contrast in iris-pupil boundary averaged over 25 subjects for 5, 15, and 25 feet.

V. Future Work

This database could be used for study of fusion of face and iris in unconstrained environments at a distance. For example, presently the iris portion of the dataset is being used to assess measurement of iris quality in the NIST IREX II: Iris Quality Calibration and Evaluation [12]. By controlling quality at acquisition in a measureable way, independent measures of iris image quality can be assessed. Ultimately, knowledge of quality can inform "smart" fusion of face and iris images, particularly when video collection of images is performed which can require massive data processing.

The next step in analysis of this database is to extend measurement of quality factors over the whole set. Corresponding match scores will then be computed, comparing the distance images to baseline. The link between quality and performance can then be assessed.

A future plan of this study is to extend the dataset by having around 100 of the subjects that participated in the current data collection return to form a supplementary dataset of non-cooperative biometrics at a distance. This new set would consist of face and iris videos obtained in a less constrained environment, i.e. longer standoff distance, outdoor collection, and having the subjects perform unconstrained tasks.

VI. Conclusion

Quality assessment of face and iris data is an important factor in the field of biometric research. This paper presents a face and iris video dataset captured at a distance incorporating multiple quality degradations in face and iris data. This 6.24 terabyte dataset will serve to evaluate face and iris recognition algorithms in terms of performance based on variations in quality.

References

[1] Xiaokun Li, Genshe Chen, Qiang Ji, and E. Blasch, "A non-cooperative long-range biometric system for maritime surveillance," *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, 2008, pp. 1-4.

[2] Y. Yao, B.R. Abidi, N.D. Kalka, N.A. Schmid, and M.A. Abidi, "Improving long range and high magnification face recognition: Database acquisition, evaluation, and enhancement," *Computer Vision and Image Understanding*, vol. 111, Aug. 2008, pp. 111-125.

[3] G. Medioni, J. Choi, Cheng-Hao Kuo, and D. Fidaleo, "Identifying Noncooperative Subjects at a Distance Using Face Images and Inferred Three-Dimensional Face Models," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 39, 2009, pp. 12-24.

[4] K. Messer, J. Matas, J. Kittler, J. Lüttin, and G. Maitre, "XM2VTSDB: The Extended M2VTS Database," *In Second International Conference on Audio and Video-based Biometric Person Authentication*, 1999, pp. 72-77.

[5] J. Daugman, "How iris recognition works," *Image Processing. 2002. Proceedings. 2002 International Conference on*, 2002, pp. I-33-I-36 vol.1.

[6] J. Wan, X. He, and P. Shi, "An Iris Image Quality Assessment Method Based on Laplacian of Gaussian Operation," May. 2007, pp. 248-251.

[7] K. Grabowski, W. Sankowski, M. Zubert, and M. Napieralska, "Focus Assessment Issues in Iris Image Acquisition System," *Mixed Design of Integrated Circuits and Systems, 2007. MIXDES '07. 14th International Conference on*, 2007, pp. 628-631.

[8] N.D. Kalka, J. Zuo, N.A. Schmid, and B. Cukic, "Image quality assessment for iris biometric," *Biometric Technology for Human Identification III*, P.J. Flynn and S. Pankanti, Eds., Orlando (Kissimmee), FL, USA: SPIE, 2006, pp. 62020D-11.

[9] "Iris Challenge Evaluation," *National Institute of Standards and Technology*, http://iris.nist.gov/ICE.

[10] "Multiple Biometrics Grand Challenge," *National Institute of Standards and Technology*.

[11] H. Proença and L.A. Alexandre, "UBIRIS : A noisy iris image database," vol. 3617, 2005, pp. 970-977.

[12] "IREX Iris Exchange," *National Institute of Standards and Technology*, http://iris.nist.gov/IREXII.

[13] J. Fierrez, J. Galbally, J. Ortega-Garcia, M.R. Freire, F. Alonso-Fernandez, D. Ramos, D.T. Toledano, J. Gonzalez-Rodriguez, J.A. Siguenza, J. Garrido-Salas, E. Anguiano, G. Gonzalez-de-Rivera, R. Ribalda, M. Faundez-Zanuy, J.A. Ortega, V. Cardeñoso-Payo, A. Viloria, C.E. Vivaracho, Q.I. Moro, J.J. Igarza, J. Sanchez, I. Hernaez, C. Orrite-Uruñuela, F. Martinez-Contreras, and J.J. Gracia-Roche, "BiosecurID: a multimodal biometric database," *Pattern Analysis and Applications*, vol. 13, 2009, pp. 235-246.

[14] J. Fierrez, J. Ortega-Garcia, D. Torre Toledano, and J. Gonzalez-Rodriguez, "Biosec baseline corpus: A multimodal biometric database," *Pattern Recognition*, vol. 40, Apr. 2007, pp. 1389-1392.

[15] S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J.L.L. Jardins, J. Lunter, Y. Ni, and D. Petrovska-Delacrétaz, "BIOMET: a multimodal person authentication database including face, voice, fingerprint, hand and signature modalities," *Proceedings of the 4th international conference on Audio- and video-based biometric person authentication*, Guildford, UK: Springer-Verlag, 2003, pp. 845-853.

[16] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J. Igarza, C. Vivaracho, D. Escudero, and Q. Moro, "MCYT baseline corpus: a bimodal biometric database," *Vision, Image and Signal Processing, IEE Proceedings* -, vol. 150, 2003, pp. 395-401.

[17] "NIST-ITL 1-2007 Data Format for the Interchange of Fingerprint, Facial, & Other Biometric Information – Part 1," *National Institute of Standards and Technology*, May. 2007, http://fingerprint.nist.gov/standard/Approved-Std-20070427.pdf.