Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2010, Article ID 345743, 9 pages
doi:10.1155/2010/345743

Research Article

Facial Recognition in Uncontrolled Conditions for Information Security

Qinghan Xiao[1] and Xue-Dong Yang[2]

[1] *Defence Research and Development Canada, Ottawa, 3701 Carling Avenue, Ottawa, ON, Canada K1A 0Z4*
[2] *Department of Computer Science, University of Regina, Regina, SK, Canada S4S 0A2*

Correspondence should be addressed to Qinghan Xiao, qinghan@canada.com

Received 1 December 2009; Accepted 3 February 2010

Academic Editor: Yingzi Du

With the increasing use of computers nowadays, information security is becoming an important issue for private companies and government organizations. Various security technologies have been developed, such as authentication, authorization, and auditing. However, once a user logs on, it is assumed that the system would be controlled by the same person. To address this flaw, we developed a demonstration system that uses facial recognition technology to periodically verify the identity of the user. If the authenticated user's face disappears, the system automatically performs a log-off or screen-lock operation. This paper presents our further efforts in developing image preprocessing algorithms and dealing with angled facial images. The objective is to improve the accuracy of facial recognition under uncontrolled conditions. To compare the results with others, the frontal pose subset of the Face Recognition Technology (FERET) database was used for the test. The experiments showed that the proposed algorithms provided promising results.

1. Introduction

With the growing need to exchange information and share resources, information security has become more important than ever in both the public and private sectors. Although many technologies have been developed to control access to files or resources, to enforce security policies, and to audit network usages, there does not exist a technology that can verify that the user who is using the system is the same person who logged in. Considering heightened security requirements of military organizations to exchange information, Defence Research and Development Canada (DRDC) Ottawa started a research project in 2004 to develop a demonstration system that automatically logs the user off the computer or locks the screen when the authenticated user cannot be identified by examining images of the person sitting in front of the computer [1]. Facial recognition technology has been adopted to monitor the presence of the authenticated user throughout a session. Therefore, only the legitimate user could operate the computer and unauthorized entities have less chance to hijack the session. The objective is to enhance the level of system security by periodically checking the user's identity without disrupting the user's activities.

Various biometric technologies, which measures human physiological or behavioural characteristics, have been proposed for user authentication [2–4]. Physiological biometric traits, such as fingerprints, hand geometry, retina, iris, and facial images, are collected from direct measurements of the human body, while behavioural biometric characteristics, such as signature, keystroke rhythms, gait pattern, and voice recordings, are associated with a specific set of actions of a person. Based on the level of user involvement when capturing the biometric traits, biometrics can be further defined as either active or passive. Passive biometrics do not require the user to actively submit a measurement, while active biometrics need cooperation from the user [2]. For approaches that enable continuous verification of identity but do not interrupt the user's activity, passive biometric technologies, such as keystroke analysis and facial recognition, have shown great potential [5–8]. However, the user alternates between the mouse and the keyboard,

thus rendering monitoring difficult with keystroke rhythms. Recently, *some researchers have investigated the possibility of using* multiple biometric modalities to continuously authenticate the user [9, 10]. It has been demonstrated that a multimodal biometric system provides a higher *level* of *authentication* assurance, but needs more computational resources than a unimodal biometric system. Therefore, we developed a video-based facial presence monitoring demonstration system, which acquires images from a video camera and runs on a Windows-based computer in near-real time [5].

Experiments have been carried out where users were allowed to perform different tasks, such as answering a phone call or drinking a soda, while being able to move freely within their normal working space in front of a camera. A major challenge is that facial images are taken under uncontrolled conditions, such as changes in illumination, pose, facial expression, and so forth. The authors of [11] claimed that "In such uncontrolled conditions, all current commercial and academic face recognition systems fail". This motivated us to conduct further research on new algorithms to improve the performance of accuracy.

The rest of the paper is organized as follows. Section 2 reviews the video-based facial recognition technologies. Section 3 briefly introduces the research background and previous work. Section 4 presents the image preprocessing algorithms. Section 5 deals with multiangle facial image analysis. Section 6 presents performance evaluation and experimental results, and the conclusion and future work are discussed in Section 7.

2. Video-Based Facial Recognition

Video-based facial recognition is a promising technology that allows covert and unobtrusive monitoring of individuals. Generally, video sequences are a collection of sequential static frames, thereby allowing the use of still-image-based techniques. However, in video-based techniques, one can utilize the temporal continuity of the image sequences to enhance robustness of the recognition process.Chellappa and Zhou [12] proposed a system that uses static images as the training data and video sequences as the probe data. They used a state space model to fuse the temporal information contained in the video sequences by tracking the subject identity using kinematics. A computationally efficient sequential importance sampling (SIS) algorithm was developed to estimate the posterior distribution. For identity n at each time instant t, by propagating the joint posterior distribution $p(n_t, \theta_t|_{z_0:t})$ of the motion (denoted by θ_t) and subject information $(z_0 : t = (z_0, z_1, \ldots z_t))$, a marginalization procedure yielded a robust estimate of identity. Evaluation was performed on two databases containing 12 and 30 subjects. The training data consisted of a single frontal face image of each subject and the probe data were videos of each of the subjects walking straight towards the camera. The first database contained images with no variation in illumination or pose, while the second, larger database contained images with large illumination variation and slight variations in pose. With the first database, the proposed system achieved a recognition rate of 100% and interestingly, it was shown that the posterior probability $p(n_t, |_{z_0:t})$ of identity increased with time, whereas the conditional entropy decreased. Using the second database, with large fluctuations of illumination, the system produced an average classification rate of 90.75%.

Chen et al. [13] used the spatio-temporal nature of video sequences to model the motion characteristics of individual faces. For a given subject, they extracted motion flow fields from the video sequences using wavelet transforms. The high dimensional vectors encoding these flow fields were reduced in size by applying a Principal Component Analysis (PCA) followed by a Linear Discriminant Analysis (LDA). Recognition was performed using a nearest neighbour classifier. The training data was collected by recording 28 subjects pronouncing two words in Mandarin. For each subject, nine video sequences were captured under different poses. For the testing data, they used the same sequences, but applied an artificial light source of varying intensity, as the goal of their evaluation was to measure robustness to illumination variations. Face alignment was performed by cropping the faces below the positions of the eyes, which were indicated manually. This method was evaluated against the Fisherface algorithm and exhibited much more stable performance across a wide range of illumination, it alsoachieved a correct classification rate of ~70% versus ~20% for Fisherface on the equivalent test data.

Liu and Chen [14] applied adaptive Hidden Markov Models (HMM) for pose-varying video-based face recognition. All face images were reduced to low-dimensional feature vectors by using PCA. In the training process, an HMM was generated to learn both the statistics of the video sequences and the temporal dynamics of each subject. During the recognition stage, the temporal characteristics of the probe face sequence were analyzed over time by the HMM corresponding to each subject. The likelihood scores provided by the HMMs were compared. Based on maximum likelihood scores, the identity of a face in the video sequence was recognized. Face regions were manually extracted from the images. The test database included 21 subjects. Two sequences were collected for each subject: one contained 322 frames for training and the other had around 400 frames for testing. The recognition performance of the proposed algorithm was compared with an individual PCA method, which showed a 4.0% error rate for HMM versus a 9.9% error rate for PCA.

A key factor in our application is that the system must be able to perform multiple tasks and recognize the user's face in near-real time. Unfortunately there is little emphasis on such a requirement in the existing literature. Therefore, DRDC conducted research on facial recognition algorithms and developed a prototype system that performs periodic verification and allows the user to carry out ordinary tasks. As illustrated in Figure 1, while a user is typing in Microsoft Word and running Internet Explorer, the system automatically performs the facial verification every 30 milliseconds.

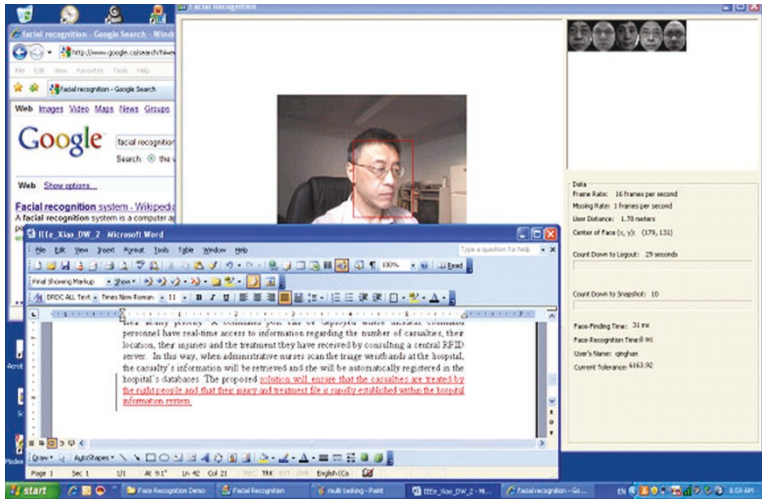


Figure 1: Authentication periodically while a user is working on a Microsoft Word file.

3. Previous Work and System Overview

Traditionally, the authentication process only verifies the identity of a user once at login or sign-on. Afterward, the assumption is that the system remains under the control of the same authenticated user. This authentication mechanism is fairly secure for one-time applications, such as accessing a protected file or withdrawing money from an automatic banking machine. However, there is a security threat if an unauthorized user takes over the session after the legitimate user successfully logged in. A facial presence monitoring system was developed in our previous work to verify the user's identity throughout the entire session [5].

Facial recognition applications can be categorized into three scenarios: under tightly controlled, loosely controlled, and uncontrolled conditions. A tightly controlled facial recognition system operates under strict rules. The users are required to cooperate with the system and the facial images can only be accepted when they satisfy certain conditions, such as a full front view of the face with a neutral expression and both eyes open under a uniform lighting condition without reflection or glare. Unfortunately, most real-world applications cannot satisfy these conditions. In an uncontrolled facial recognition system, the users may be unaware that a system is taking their facial images. Hence, there exist considerable variations in illumination, pose, and expression. Many applications fall into the loosely controlled category. For instance, the facial recognition system may work under uncontrolled illumination and background, but a narrow range of face pose. The system we developed works under uncontrolled conditions. Figure 2 shows the system diagram, and the module functionalities are briefly summarized.

(1) Input. The system captures video images and processes either 24-bit or 32-bit colour images. The captured images are displayed on the computer screen via DirectX in real time.

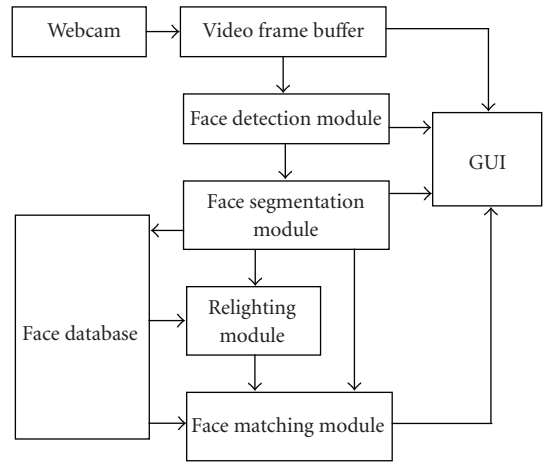


Figure 2: Overall system architecture.

(2) Face Detection. A face detection algorithm subsequently examines the captured images. The locations of *potential human faces* are recorded. Later, when a video image is finally rendered to the monitor, a red rectangle encloses each potential face. Because of the near-real time requirement, this algorithm cannot be expected to perform flawlessly. Therefore, it is anticipated that some detected objects are not actually faces. In order to obtain an accurate result, the system examines more than one image frame to determine if an object is likely a face or not, which is one of the advantages of using video-based facial recognition. When an object presumed to be a face is discovered, the corresponding region in the previous frame is examined. If there was no face found in that area in the previous frame, then the current object is unlikely to be a face. As a result, the object will not be recorded as a potential face. Conversely, if there had been a face present in that region, the likelihood that the current object represents a face is greater. In such a situation, the system assumes that the object is a face.

(3) Face Segmentation. Because the position of a face computed by the face detection module is not accurate, a more precise location is necessary for a good face-matching performance. Since the size of a user's face appearing in a video frame also varies depending on the distance of the user from the web camera, the face image must be normalized to a standard size. There are some features in face images that may change from time to time. For example, the hairstyle can change significantly from one day to another. In order to reduce the effects of such dynamic features, a standard elliptical region with a fixed aspect ratio is used to extract the face region.

(4) Face Matching. Turk and Pentland pioneered the eigenface method [15], which relies on the Karhunen-Loeve (KL) transform or the Principal Component Analysis (PCA). To improve the performance of the eigenface method, it is important to have a good alignment between the live and the stored face images. It means that the nose has to be in the middle, the eyes have to be at a stable vertical position, and the scale of the face images must be normalized. A significant portion of our efforts addressed these issues. An elliptical facial region extracted from a video frame is matched against the facial models stored in a database. Each face image is first converted to a vector. This vector is projected onto eigenfaces through inner product calculations. Each face produces a weight vector. The Euclidean distance between two weight vectors is used to measure the similarity between the two faces. This distance is then mapped to a normalized matching score.

(5) Relighting. This module provides a histogram-based intensity mapping function to normalize the intensity distribution of the segmented face image. It is noted that some areas, such as the eyes, can be very dark due to light direction. It is potentially beneficial to enhance the features in dark regions to improve the recognition performance.

(6) Facial Database. It is assumed that data from up to eight users may be saved in the database. Each user is required to take at least one picture in the user's normal working environment within the normal sitting space and under the normal lighting conditions.

(7) Output. The demonstration system has three main outputs.

(i) *Live video*: the detected face in the scene is surrounded by a red rectangle.

(ii) *Matching results*: the segmented face image from the current test scene is displayed, along with up to five candidate faces from the database in descending priority order.

(iii) *Performance Data*: several performance data are displayed in real time, such as the overall frame rate, the face detection time, the face recognition time, and the best matching score.

In order to evaluate the robustness of the system, we conducted tests in different scenarios that might happen in a real office environment. These tests include detecting multiple users in a scene and recognizing a user answering a phone call or drinking a soda. As shown in Figure 3, the system performed very well and even partially-occluded faces, such as mouth covered by telephone and eyes covered by dark sunglasses, could still be recognized correctly.

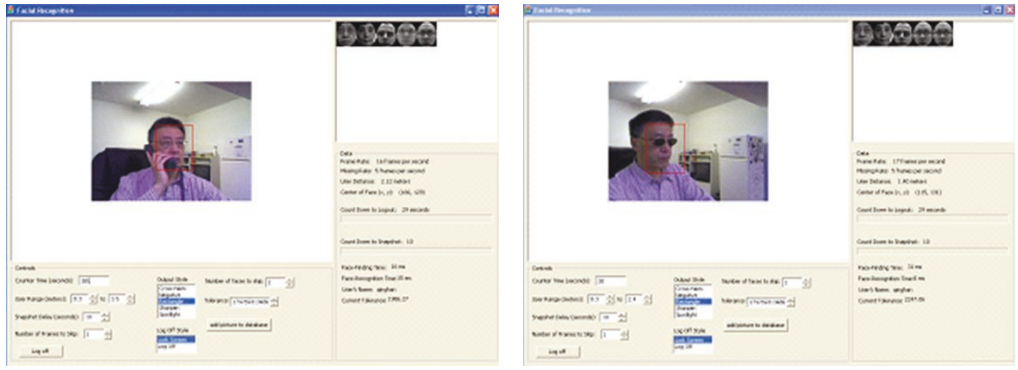
4. Image Preprocessing

A study on image preprocessing algorithms has been carried out to improve the accuracy performance. It focused on the areas that affect the accuracy of facial recognition, which include geometric correction, face alignment, masking and photometric normalization.

4.1. Face and Eye Detection. Each image frame is searched for faces using a fast Viola-Jones [16] face detector. The approach is enhanced over conventional implementations in following aspects. It is invariant to faces of multiple sizes, in-plane (IP) head rotations of $\theta = \pm 45^\circ$, and out-of-plane (OOP) head rotations of $\pm 30^\circ$, as shown in Figure 4. Once the face is located, a virtual bounding box is formed around it and a built-in eye finder is initiated in this region. The eye finder outputs the 2D locations of the left and right irises. To improve operating speed, it is essential to adapt the face search regions based on previous face localization results. This dynamically limits the search area, and as a result, allows this step to operate in near-real time at resolutions up to 1024×768.

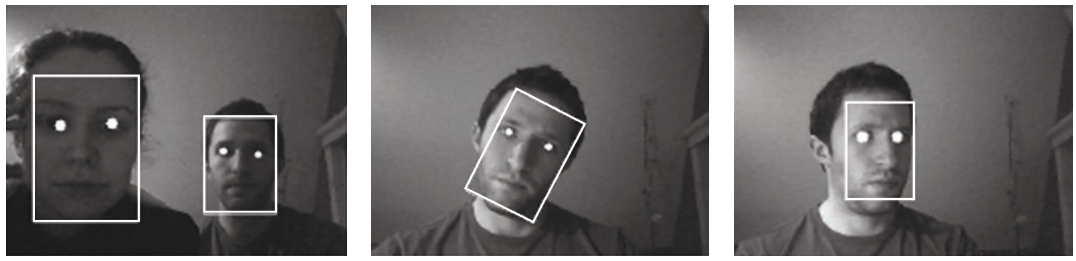
4.2. Face Alignment and Masking. Using the 2D coordinates of the left and right irises, the in-plane rotation angle θ of the face is estimated trigonometrically. The face is then rotated by $-\theta^\circ$, effectively placing both eyes at equal height, as shown in Figure 5. Once normalized for rotation, the face is geometrically warped in order to set the inter-eye distance to 40 pixels, chosen as a reasonable value to ensure effective performance of the subsequent recognition steps. The system then crops the face image to 64×64 pixels, leaving out the top of the head in order to reduce the impact of hair style on the classification. Similarly, a mask is applied to the shoulder and lower-neck areas to avoid the possible influence of different clothing on the face recognition.

4.3. Photometric Normalization. Variable lighting, both in intensity and location, can cause dramatic changes to a person's appearance as a result of shadows and specularity. These issues arise especially in uncontrolled environments, such as outdoors or even in a windowed office. For a facial recognition system using appearance-based method, it is crucial to apply photometric normalization. Here, homomorphic filtering [17] is adopted to mitigate the effects of shadows and specularity on the face. Homomorphic filtering is commonly used to simultaneously normalize brightness across an image and increase contrast. As shown in Figure 6(b), the shadows caused by lateral lighting



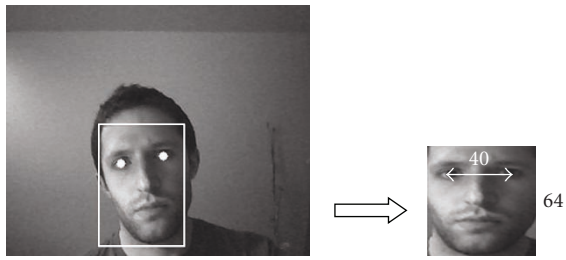
(a) (b)

Figure 3: Recognize partially occluded faces.



(a) Scale (b) IP head rotation (c) OOP head rotation

Figure 4: Face and eye detection under different situations.



(a) (b)

Figure 5: Face alignment and masking to: (a) Original frame, (b) Aligned and masked face ready to be analyzed (dimensions shown for illustrative purposes).

(a) Original frame (b) Filtering result

Figure 6: Homomorphic filtering.

on the right side of the person's face are filtered while preserving important structural details. As this step usually demands many computationally intensive steps, the filter is implemented in a separable fashion by breaking a two-dimensional signal into two one-dimensional signals with a vertical and a horizontal projections. In homomorphic filtering, it is necessary to convolve the input image with a Gaussian filter, which is separable by nature because 2D Gaussians are circular symmetric.

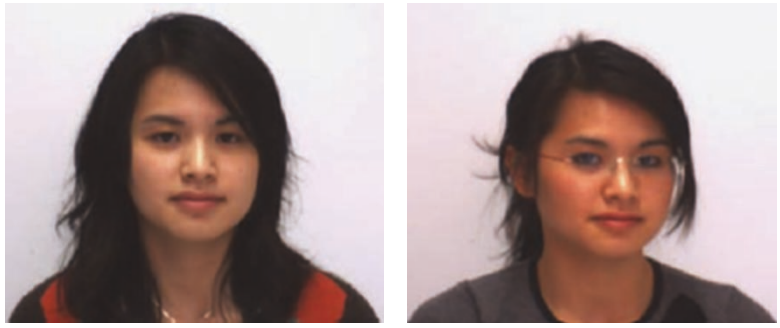
To convolve an image with a separable filter kernel, each row in the image is convolved with the horizontal projection to obtain an intermediate image. Next, we convolve each column of the intermediate image with the vertical projection of the filter. Hence the resulting image is identical to the direct convolution of the original image and the filter kernel. The convolution of an $N \times N$ image with an $M \times M$ filter kernel requires a time proportional to $N^2 \times M^2$. In comparison, convolution in the separable fashion only requires a time proportional to $N^2 \times M$. Therefore, the processing speed is improved to achieve real time operation.

5. MultiAngle Facial Image Analysis

Because there were more side-view face images than front-view images in the captured video stream, a study has been conducted to explore the possibility of using multiangle face images to increase the recognition rate. A database called

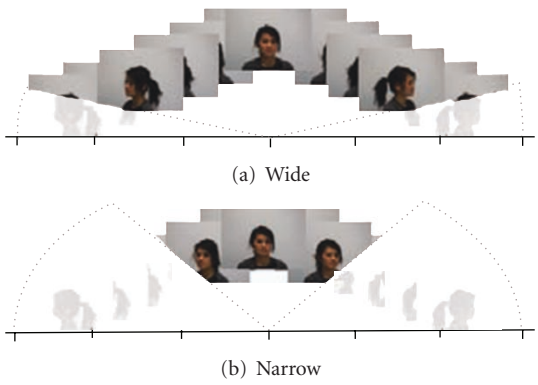
(a) (b) (c)

Figure 7: Database capture conditions: (a) Overhead light with neutral expression, (b) Overhead and frontal light with neutral expression, (c) Overhead and frontal light with varying facial expressions.



(a) First session (b) Second session

Figure 8: An example of repeat subject.



(a) Wide

(b) Narrow

Figure 9: Angular range.

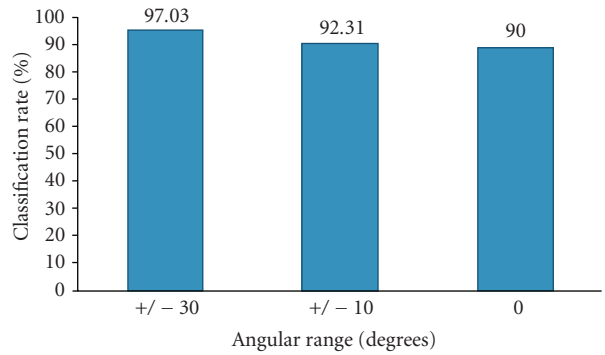


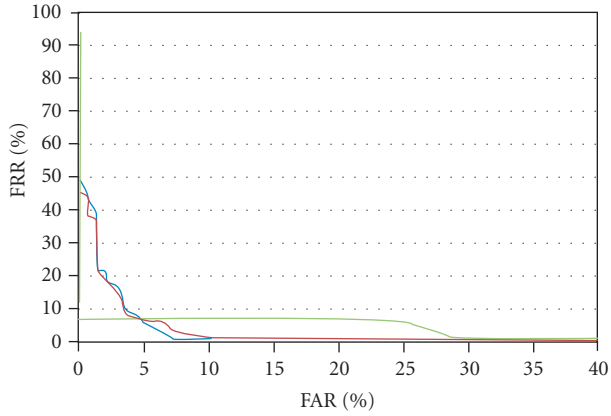
Figure 10: Mean classification rate versus angular range.

the CIM face database was constructed at the Centre for Intelligent Machines (CIM) at McGill University. Currently, it consists of 43 subjects, 19 of which returned for a second recording session. The recording sequences involved the subject rotating through a range of 180°. For each subject, three sequences were recorded under different lighting conditions and facial expressions, as shown in Figure 7.

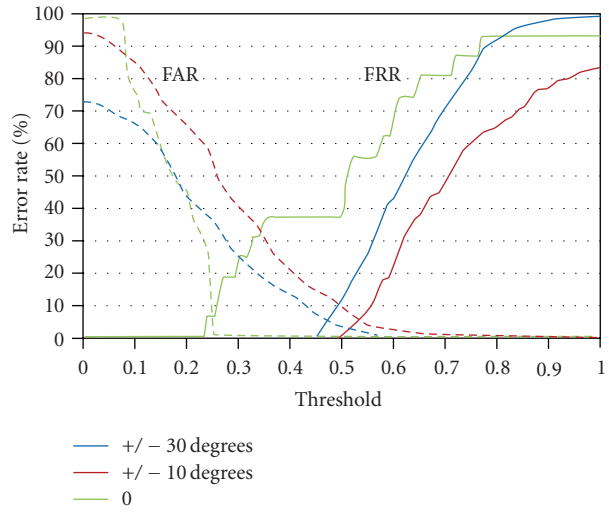
About one-third of the participants were asked to return for a second recording session at least one week after their first session. They were requested to change their facial appearance by growing facial hair, wearing glasses, applying makeup, and so forth, as illustrated in Figure 8. Each sequence was saved as a raw video file and manually cropped at the approximate start and stop times of the participant's rotation. This step allowed us to estimate the rotation angle difference between consecutive frames, which was 2.5° on average.

The benefits of using multiangle images were evaluated by increasing the training data within an angular range of either ±30°, ±10°, or 0°, as illustrated in Figure 9. In addition, an ad-hoc third trial was conducted in which the test set consisted of outdoor images of eight of the subjects.

In the experiments, thirty subjects were randomly selected to construct the training set and twelve subjects were used as imposters to evaluate the false accept rate (FAR).

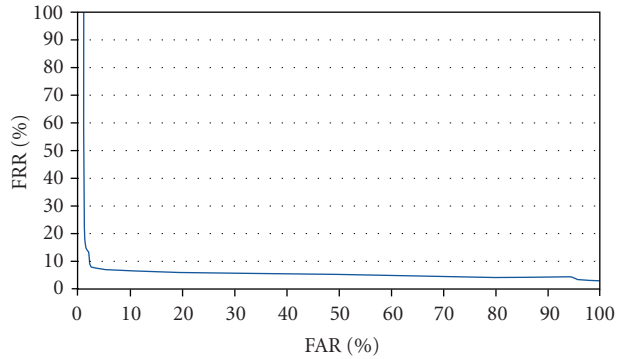


(a) ROC curve

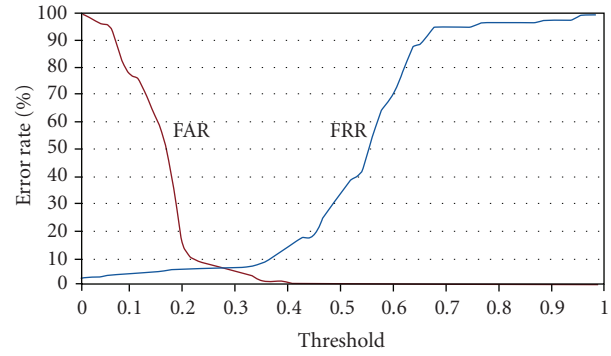


(b) EER

Figure 11: CIM database results.



(a) ROC curve



(b) EER

Figure 12: FERET database results.

In order to compare with other published algorithms, the proposed approach has been tested using the widely-used FERET database. The fb (frontal pose) subset, consisting of 725 subjects, was used in which 580 subjects were selected randomly as the training set and the remaining 145 used as impostors. The performance metrics consisted of the mean classification rate, the false accept rate, and the false recognition rate (FRR). The FAR is the likelihood of incorrectly accepting an impostor, while the FRR is the likelihood of incorrectly rejecting an individual in the training set. As a common rule, the thresholds were adjusted based on the classification confidence values to evaluate the trade-off between FAR and FRR.

6. Experimental Evaluation

Three sets of experiments were conducted on the CIM and FERET [18] face databases. Performance of the classifier is associated with the angular range of training data, as seen in Figure 10. The best performance was achieved on the $\pm 30^\circ$ range, that is, the same range used for the testing sequences. Reducing the range to $\pm 10^\circ$ resulted in a 4.72% drop in performance, while training on the 0° range (i.e., a single frontal image) led to a 7.03% decrease.

It is desirable to evaluate system accuracy by comparing the FAR to the FRR for different choices in angular range of training data. For any given threshold T, we measured the FAR and FRR values. By varying T, we recorded the value sets of FAR and FRR. Plotting the value sets, the result is called a Receiver Operating Characteristic (ROC) curve (Figure 11(a)). ROC curve is a concise graphical summary of the performance of the biometric device [19], and it enables the performance comparisons between different systems under similar conditions or of a single system operating under differing conditions [20]. In the plot, an ROC curve, which lies to the lower left of another curve, has a better accuracy performance. Therefore, in Figure 11(a), the results for $\pm 10^\circ$ and $\pm 30^\circ$ are very similar.

Figure 11(b) presents the FAR and FRR, individually, as functions of threshold value. Since higher threshold values increase FRR and decrease FAR, analysis of these results is only meaningful when it takes into account both FAR and FRR together. The point at which the FAR is equal to the FRR is called the equal error rate (EER). This is another commonly used measure to assess the overall performance for biometric systems. The result in Figure 11(b) shows that error is minimized when training with greater angular range, as expected. Training with an angular range of $\pm 30^\circ$ yields an equal error rate of 4.86%, whereas this increases to 6.80% and 7.42% for angular ranges of $\pm 10^\circ$ and 0°, respectively.

To deal with outdoor scenarios, a preliminary test has been conducted by using static images of eight of the subjects taken outdoors; six of the eight subjects (75%) were successfully classified. Note that the training data for these individuals came exclusively from indoor data. Due to the very limited number of testing samples, these results should not be taken as definitive statements of performance.

Despite very limited opportunity for tuning the algorithms, the mean classification rate obtained on the FERET database was 92.5% (an EER of 7.5%), demonstrating that the presented algorithm is scalable to relatively large databases. This result is only 1.5% lower than the best published classification rate (94%) in the literature for the same database [18]. The FAR versus FRR plots and error rates versus threshold values are shown in Figure 12.

7. Conclusions and Future Work

As networks become larger, more complex, and more distributed, information security has become more critical than it has ever been in the past. Many efforts have been made aiming to accurately authenticate and authorize trusted individuals, and audit their activities. Once a user is successfully logged in, the prevailing technique assumes that the system is controlled by the same person. Focusing on this security challenge, we developed an enhanced authentication method that uses video-based facial recognition technology to monitor the user during the entire session in which the person is using the system. It can automatically lock the screen or log out the user when the authenticated user's face disappears from the vicinity of the computer system for an adjustable time interval.

In order to improve the performance in accuracy, further research has been conducted in developing image preprocessing algorithms and using multiangle facial images in training. The experiments conducted on the CIM and FERET face databases showed promising results. On the FERET database, an EER of 7.5% is obtained which is comparable to the best published EER rate of 6% in the literature. A major advantage of video-based face recognition is that a set of images of the same subject can be captured in a video sequence, while a main problem of video-based face recognition lies in the low images quality in video frames. In order to improve recognition accuracy, an effort has been put into combining front and angle face images.

Uncontrolled face recognition from video is still a challenging task. During our experiments, it became clear that most classification errors were due to instability in the alignment process. Specifically, we noticed that even though the eye detector accurately finds the eyes, the position tends to oscillate in the dark area of the pupils, which causes fluctuations in the computed in-plane head rotation angle. One area for future research is to investigate bootstrapping and integration of spatio-temporal filtering methods in the eye detector to mitigate this issue. We will perform more research on the relationships among front and angle-face images to extract some nose features that cannot be obtained or accurately measured from the front face itself, such as nose slope and depth. Not only will we use newly developed algorithms to improve the facial presence monitoring system, but also we will explore other application areas that will benefit from uncontrolled face recognition. In addition, we need to conduct research to analyze legal and social aspects of monitoring the user's presence at a workplace.

Acknowledgments

The authors would like to thank for the great contribution that Dr. Martin Levine and Dr. Jeremy Cooperstock made under the contract W7714-071076. Without their contribution, they would not have been able to achieve the results presented herein. The authors would also like to thank Dr. Daniel Charlebois for his support and valuable comments throughout the course of this research project.

References

[1] X. D. Yang, P. Kort, and R. Dosselmann, "Automatically log off upon disappearance of facial image," Contract Report CR 2005-051, DRDC, Ottawa, Canada, March 2005.

[2] P. Reid, *Biometrics for Network Security*, Prentice-Hall, Upper Saddle River, NJ, USA, 2003.

[3] J. Chirillo and S. Blaul, *Implementing Biometric Security*, John Wiley & Sons, Indianapolis, Ind, USA, 2003.

[4] R. Manoj, "Biometric security: the making of biometrics era," *InfoSecurity*, pp. 16–22, July 2007.

[5] Q. Xiao and X. D. Yang, "A facial presence monitoring system for information security," in *Proceedings of the IEEE Workshop on Computational Intelligence in Biometrics: Theory, Algorithms, and Applications (CIB '09)*, pp. 69–76, March 2009.

[6] R. Janakiraman, S. Kumar, S. Zhang, and T. Sim, "Using continuous face verification to improve desktop security," in *Proceedings of the 7th IEEE Workshop on Applications of Computer Vision (WACV '07)*, pp. 501–507, January 2007.

[7] B. Rao, *Continuous keystroke biometric system*, M.S. thesis, Media Arts and Technology, University of California, Santa Barbara, Calif, USA, 2005.

[8] R. H. C. Yap, T. Sim, G. X. Y. Kwang, and R. Ramnath, "Physical access protection using continuous authentication," in *Proceedings of the IEEE International Conference on Technologies for Homeland Security (HST '08)*, pp. 510–512, May 2008.

[9] T. Sim, S. Zhang, R. Janakiraman, and S. Kumar, "Continuous verification using multimodal biometrics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 687–700, 2007.

[10] A. Azzini and S. Marrara, "Impostor users discovery using a multimodal biometric continuous authentication fuzzy system," in *Proceedings of the 12th International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES '08)*, vol. 5178 of *Lecture Notes in Computer Science*, pp. 371–378, Springer, September 2008.

[11] S. J. D. Prince, "Latent identity variables: a generative framework for face recognition in uncontrolled conditions," EP/E065872/1, EPSRC, September 2007.

[12] R. Chellappa and S. Zhou, "Face tracking and recognition from videos," in *Handbook of Face Recognition*, S. Z. Li and A. K. Jain, Eds., pp. 169–192, Springer, Berlin, Germany, 2005.

[13] L.-F. Chen, H.-Y. M. Liao, and J.-C. Lin, "Person identification using facial motion," in *Proceedings of the IEEE International*

Conference on Image Processing (ICIP '01), vol. 2, pp. 677–680, October 2001.

[14] X. Liu and T. Chen, "Video-based face recognition using adaptive hidden Markov models," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 340–345, June 2003.

[15] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[16] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[17] D. B. Williams and V. Madisetti, *Digital Signal Processing Handbook*, CRC Press, Boca Raton, Fla, USA, 1999.

[18] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[19] M. Schuckers, "Some statistical aspects of biometric identification device performance," *Stats Magazine*, p. 3, September 2001.

[20] D. Davis, P. Higgins, P. Kormarinski, J. Marques, N. Orlans, and J. Wayman, "State of the art biometrics excellence roadmap: technology assessment: volume 1," Tech. Rep., MITRE Corporation, 2008, http://www.biometriccoe.gov/SABER/index.htm.