11

Recognizing Face Images with Disguise Variations

Richa Singh, Mayank Vatsa and Afzel Noore
Lane Department of Computer Science & Electrical Engineering, West Virginia University
USA

1. Introduction

Automatic face recognition, required in law enforcement applications such as surveillance, border security and forensic investigation, is a process in which an individual is identified or verified based on facial characteristics. Researchers have proposed several algorithms that can effectively recognize individuals in controlled environment with minor variations in pose, expression, and illumination (Zhao et al., 2003), (Li & Jain, 2005), (Wechsler, 2006), (Delac & Grgic, 2007). In recent face recognition test reports such as FRVT 2002, FRGC 2004, and FRGC 2006 (Philips et al., 2006 & Philips et al., 2007), the results show that under normal changes in constrained environment, the performance of existing face recognition systems is greatly enhanced. However, in most real world applications, images may not be of good quality or user may not be cooperative or there may be temporal variations and dissimilarities in facial characteristics that are artificially created using disguise accessories.

Challenges in automatic face recognition can be classified into six categories: illumination, image quality, expression, pose, aging, and disguise. Among these challenges, recognition of faces with disguise is a major challenge and has only been recently addressed by few researchers (Alexander & Smith, 2003), (Ramanathan et al, 2004), (Silva & Rosa, 2003), (Singh et al., 2008). As shown in Fig. 1, the inter-personal and intra-personal characteristics can be modeled using disguise accessories to alter the appearance of an individual, to impersonate another person, or to hide one's identity. For example, a criminal can alter facial features and appearance using makeup tools and accessories to remain elusive from law enforcement. The challenges due to disguise cause change in visual perception, alter actual data, make pertinent facial information disappear, mask features to varying degrees, or introduce extraneous artifacts in the face image. Existing face recognition algorithms may not be able to provide the desired level of security for such cases.

In literature, Ramanathan et al. (Ramanathan et al., 2004) studied facial similarity for several variations including disguise by forming two eigenspaces from two halves of the face, one using the left half and other using the right half. From the test image, optimally illuminated half face is chosen and is projected into the eigenspace. This algorithm has been tested on the AR face database (Martinez & Benavente, 1998) and the National Geographic database (Ramanathan et al., 2004) which consists of variations in smile, glasses, and illumination. An accuracy of around 39% for best two matches is reported on the AR database. Silva and Rosa

Source: Recent Advances in Face Recognition, Book edited by: Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett,
ISBN 978-953-7619-34-3, pp. 236, December 2008, I-Tech, Vienna, Austria

(Silva & Rosa, 2003) proposed using Eigen-eyes to handle several challenges of face recognition including disguise. Using the Yale database (Yale face database), the algorithm was able to achieve an accuracy of around 87.5%. The advantage of the algorithm is that alterations in the facial features excluding the eye region do not affect the accuracy. Pamudurthy et al. (Pamudurthy et al., 2005) proposed a face recognition algorithm which uses dynamic features obtained from skin correlation and the features are matched using nearest neighbor classifier. On a database of 10 individuals, authors reported that this approach gives accurate results. Alexander and Smith (Alexander & Smith, 2005) used PCA based algorithm with Mahalanobis angle as the distance metric. The results show an accuracy of 45.8% on the AR database (Martinez & Benavente, 1998). The limitation of these algorithms is that the performance degrades when important regions such as the eye and mouth are covered. Moreover, the AR and Yale databases do not contain many images with disguise and therefore are not ideal for validating algorithms under comprehensive disguise scenarios.

Fig. 1. Face disguise: use of makeup tools and accessories to alter facial features and appearance of the same individual.

In this chapter, we focus on the problem of recognizing faces that are altered due to variations in disguise, i.e. the ability to recognize individuals when their appearances are intentionally altered to defraud law enforcement and the public using disguises. Specifically, in this research we identify different types of disguise accessories that can be used to alter facial information and analyze their effect on face recognition algorithms. Eight face recognition algorithms, including state-of-the-art algorithms and algorithms that are tailored for disguise, are evaluated using a heterogeneous face database and the face disguise database. Experimental results suggest that the performance of appearance and feature based algorithms is affected when features are altered or hidden whereas texture based algorithms fail to perform with multiple disguises. Further, results indicate that face recognition with variations in disguise is a major challenge and the performance of existing algorithms are not adequate. The next section identifies the types of disguise that can alter facial appearance and features pertinent for face recognition.

2. Types of disguises

The performance of face recognition algorithms can be affected by alteration in appearance, feature and combination of multiple variations. The possible variations of disguise can be classified into the following eight categories depending on their effect on facial appearance and features.

1. **Minimal variations**: Two face images captured at different time instances can have minimal variations in appearance and features. In such cases, face recognition algorithms usually yield correct results.
2. **Variations in hair style**: Hair style can be changed to alter the appearance of a face image or hide facial features. Fig. 2 shows an example of facial variations of an individual due to changes in hair style.

Fig. 2. Face images with variation in hair style

3. **Variations due to beard and moustache**: Facial hair such as beard and moustache can alter facial appearance and features in the lower half of the face, specifically near mouth and chin regions. Fig. 3 shows an example where face images with and without beard and moustache show different appearance.

Fig. 3. Face images with variation in beard and moustache

4. **Variations due to glasses**: Glasses, especially sun-glasses are one of the easiest ways to alter facial appearance. In general, glasses affect upper facial region by hiding the facial features (e.g. eyes and eyebrows). As shown in Fig. 4, structural differences in glasses and opacity of lens can also change the appearance of an individual.

Fig. 4. Face images with variation in eye glasses

5. **Variations due to cap and hat**: In general, use of cap and hat hides hairs and some part of the forehead which are not used by the recognition algorithms. However, as shown in Fig. 5, some specific types of cap and hat (e.g., monkey cap) can hide pertinent facial features, thereby affecting the performance of face recognition algorithms.

Fig. 5. Face images with variation in cap and hat

6. **Variations due to lips, eyebrows and nose**: Makeup tools can be used to alter the shape and size of lips, eyebrows and nose (Fig. 6). These key local features, if altered, can affect the performance of feature based face recognition algorithms.

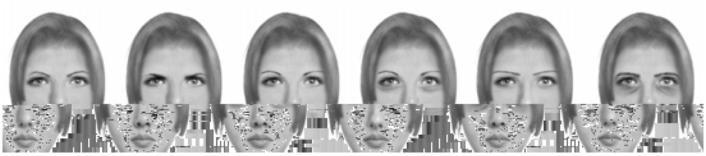

Fig. 6. Face images with variation in lips, eyebrow and nose characteristics

7. **Variations due to aging and wrinkles**: Aging can be natural (due to age progression) and artificial (using makeup tools). In both the cases, aging and wrinkles can severely affect the performance of face recognition algorithms. An example of facial variations due to aging and wrinkles is shown in Fig. 7.

Fig. 7. Face images with aging and wrinkle variations

8. **Multiple variations**: A combination of the above mentioned variations can be used to disguise and defraud law enforcement. Fig. 8 shows an example in which multiple variations are used to alter the appearance and features of an individual.

Fig. 8. Face images of an individual with multiple disguise variations

3. Characteristics of face recognition algorithms used for evaluation

In general, face recognition algorithms can be broadly classified into three classes (1) appearance based algorithms, (2) feature based algorithms, and (3) texture based algorithms. To compare the performance of these algorithms; eight algorithms are selected and are briefly explained below.

1. **Appearance based algorithms**: Three appearance based algorithms are used in experiments that are specifically tailored for recognizing individuals with altered

appearances. These algorithms are: (1) PCA algorithm with Mahalanobis distance (Alexander & Smith, 2005), (2) Half-face based algorithm (Ramanathan et al., 2004), and (3) Eigen-eyes based algorithm (Silva & Rosa, 2003). Additional details of these algorithms are explained in the Introduction section.

2. **Feature based algorithms**: Two feature based algorithms are selected namely, Geometrical Feature (GF) (Cox et al., 1996) and Local Feature Analysis (LFA) (Penev & Atick, 1996).
   Geometrical feature based recognition algorithm (Cox et al., 1996) uses mixture distances of the facial features for matching. This algorithm works on the distance between geometrical features. Facial features such as nose, mouth, eyes, and ears are extracted and their shape information is computed. For matching two images, this shape information is matched using Euclidean distance measure. Hence, the algorithm depends on the correspondence between the facial features and works only in cases when this information is preserved. If the facial features are occluded using accessories such as glasses, beard, moustache, and scarf, then the performance decreases.
   Local feature analysis (Penev & Atick, 1996) is one of the most widely used face recognition algorithms which can accommodate some changes in facial expression. LFA refers to a class of algorithms that extract a set of geometrical metrics and distances from facial images and use these features as the basis for representation and comparison. The recognition performance is dependent on a relatively constant environment and quality of the image.
3. **Texture based algorithms**: In the third class of face recognition algorithms, i.e., texture based, three algorithms are selected namely Independent Gabor Features (IGF) (Liu & Wechsler, 2003), Local Binary Pattern (LBP) (Ahonen et al., 2006), and 2D-log polar Gabor transform and Neural Network (2DLPGNN) (Singh et al., 2008).
   Independent Gabor features (Liu & Wechsler, 2003) extract Gabor features from the face image and then reduces the dimensionality using PCA. The independent Gabor features are obtained from the reduced dimensionality feature vector by applying Independent Component Analysis. These independent Gabor features are classified using Bayes classifier and then matched using the Manhattan distance measure.
   Local binary pattern based face recognition algorithm (Ahonen et al., 2006) extracts textural feature from the face images. In this algorithm, a face image is divided into several regions and weighted LBP features are extracted to generate a feature vector. Matching of two LBP feature vectors is performed using weighted Chi square distance measure based algorithm.
   2D-log polar Gabor and neural network based face recognition algorithm (Singh et al., 2008) extracts phase features from the face images. It uses dynamic neural network architecture to extract the phase features using 2D log polar Gabor transform. The phase features are divided into frames which are matched using hamming distance.

4. Characteristics of face databases and experimental protocol

The experiments are divided into two parts. In the first part, the performance of face recognition algorithms is evaluated on a heterogeneous face database that contains variations due to pose, expression and illumination. This experiment is performed as the baseline experiment. The second experiment is performed to evaluate the effect of disguises on face recognition algorithms. For these experiments, we have created two face databases, (1) heterogeneous face database and (2) face disguise database.

1. **Heterogeneous face database**: For evaluating the performance on a large database with challenging intra-class variations, we combined images from multiple face databases and created a heterogeneous database of 882 subjects. Table 1 lists the databases used and the number of subjects selected from the individual databases. The CMU-AMP database (CMU-AMP face database) contains images with large expression variations while the CMU-PIE dataset (Sim et al., 2003) contains images with variation in pose, illumination and facial expressions. The Equinox database (Equinox face database) has images captured under different illumination conditions with accessories and expressions. The AR face database (Martinez & Benavente, 1998) contains face images with varying illumination and accessories, and the FERET database (Philips et al., 2000) has face images with different variations over a time interval of 3-4 years. The Faces in the Wild database (Huang et al., 2007) contains real world images of celebrities and popular individuals. This database contains images of more than 1600 subjects from which we selected 294 subjects that have at least 6 images. To the best of our knowledge, there is no single database available in the public domain which encompasses all these types of intra-class variations.

Face database	Number of subjects
CMU-AMP	13
CMU-PIE	65
Equinox	90
AR	120
FERET	300
Faces in the Wild	294
Total	882

Table 1. Composition of the heterogeneous face database

2. **Face disguise database**: This database is prepared by the authors. It contains real and synthetic face images from 125 subjects. For every subject, disguise variations are collected based on the eight classes of disguise variations described in Section 2. The database contains real face images of 25 individuals with 15-25 different disguise variations of each individual. Since our goal is to evaluate the performance of the face recognition algorithms on disguise, the database contains frontal face images with less emphasis on variations due to illumination, expression and pose. Fig. 9 shows an example of this database. Further, we used FACES software (Faces software) to generate 4000 frontal face images of 100 subjects with a comprehensive set of variations for disguise. An example of the synthetic face database is shown in Fig. 10. The complete face disguise database is used to broadly evaluate the performance of the proposed algorithm for disguised images.
3. **Experimental protocol**: For both the experiments, the images are partitioned into two sets: (1) the training dataset is used to train the individual face recognition algorithms and (2) the gallery-probe dataset (the test set) is used to evaluate the performance of the recognition algorithms. The training set comprises of randomly selected three images of each subject and the remaining images are used as the test data to evaluate the

algorithms. This train-test partitioning is repeated 20 times (cross validation) and the Receiver Operating Characteristics (ROC) curves are generated by computing the genuine accept rates (GAR) over these trials at different false accept rates (FAR). Furthermore, verification accuracies are reported at 0.01% FAR.

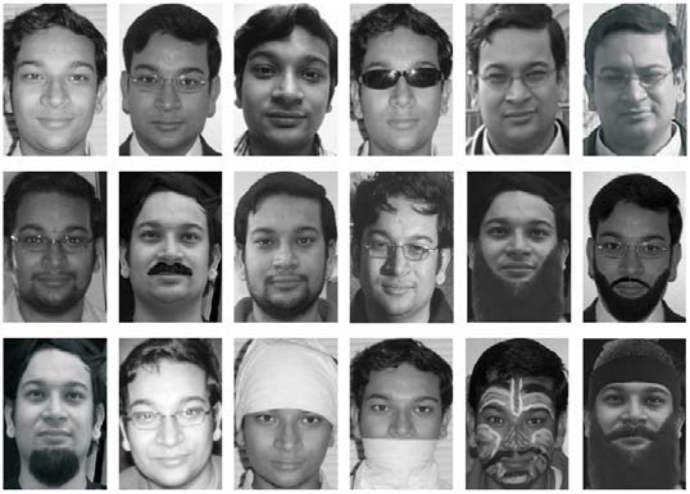

Fig. 9. Sample images from the real face database of the same individual.

5. Performance evaluation

The experiments are divided into two parts: (1) evaluation using the heterogeneous face database, and (2) evaluation using the disguise database. Experimental protocol described in Section 4 is used for training and testing. Training images are used to train the face recognition algorithms and testing images are used for gallery probe matching and evaluation.

5.1 Evaluation using heterogeneous face database

This experiment is conducted to evaluate the effect of three covariates namely pose, expression, and illumination on the performance of face recognition. Fig. 11 shows the ROC plot and Table 2 illustrates the verification accuracies at 0.01% FAR. The key results and their analysis are summarized below:

1. From Table 2, covariate analysis suggests that among the three covariates, variations in pose cause a large reduction in verification accuracy compared to expression and illumination.
2. Results also suggest that the texture based algorithms yield better accuracy compared to appearance and feature based algorithms. This is because pose, expression and illumination variations can cause substantial changes in appearance and spurious/missing features, thereby reducing the verification performance.
3. Among all the algorithms, 2DLPGNN yields the best verification accuracy of 84.2%, which is at least 12% better than other algorithms. The 2DLPGNN algorithm effectively encodes textural features that can handle minor to moderate variations in pose, expression and illumination.

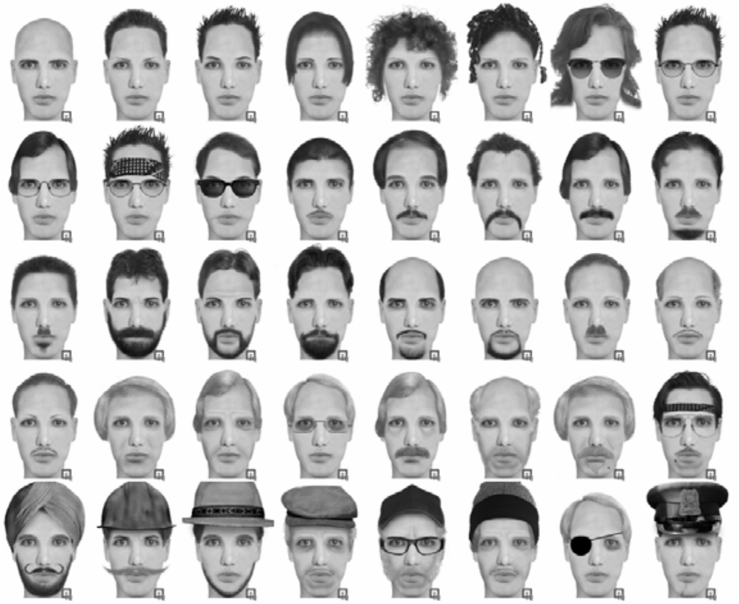

Fig. 10. Sample disguise variations of the same image from the synthetic face database

Effect of disguises on accuracy

In this section, we analyze the performance of face recognition algorithms for each disguise category using the disguise database. The ROC plot in Fig. 12 and Table 3 summarizes the performance of face recognition algorithms. The key results of the experiments are explained below.

1. For most of the disguise variations, appearance based algorithms yield lower verification accuracy because these algorithms use facial appearance to determine the identity, and the makeup tools and accessories significantly alter the facial information. Similarly, feature based algorithms suffer due to feature alterations that are caused by disguise accessories.
2. Texture based algorithms provide significantly better verification accuracy compared to appearance based algorithms. Conversely, these algorithms do not yield good verification accuracy with moderate to large disguise variations.
3. 2DLPGNN algorithm (Singh et al., 2008) yields the best verification performance. However, for the challenging scenarios of multiple disguise variations, the accuracy is only 65.6% but still outperforms other algorithms. This shows that existing algorithms are not efficient enough to handle large degree of disguise variations.
4. Another important comparison is among pose, expression, illumination and multiple disguise variations. From Table 2 and 3, it is quite clear that multiple disguise variations is the most difficult challenge to handle (e.g. 2DLPGNN yields accuracies in the range of 75-86% for pose, expression and illumination whereas for multiple disguise variations, it is only 65%).

These comprehensive experimental results indicate that further research is needed to address high degree of disguise variations.

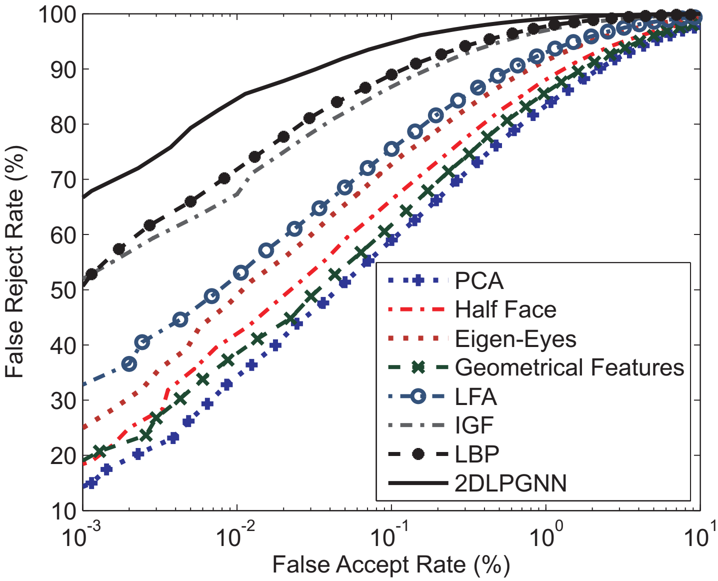

Fig. 11. ROC to evaluate the performance of face recognition algorithms on the heterogeneous face database

Verification accuracy at 0.01% FAR								
Covariates	Appearance based algorithms			Feature based algorithms		Texture based algorithms		
	PCA	Half-face	Eigen-eyes	GF	LFA	IGF	LBP	2D-LPGNN
Pose	31.9	36.6	29.7	35.4	50.1	60.7	73.2	75.3
Expression	35.5	42.1	78.6	38.2	52.3	69.8	71.4	87.6
Illumination	35.3	45.2	45.8	41.7	53.5	68.6	70.9	86.5
Overall	34.4	41.8	49.3	38.9	52.7	67.3	72.1	84.2

Table 2. Verification performance of appearance, feature and texture based face recognition algorithms for different covariates.

6. Conclusion

Currently, many security applications use human observers to recognize the face of an individual. In some applications, face recognition systems are used in conjunction with limited human intervention. For autonomous operation, it is highly desirable that the face recognition systems be able to provide high reliability and accuracy under multifarious conditions, including disguise. However, most of the algorithms are not robust to high

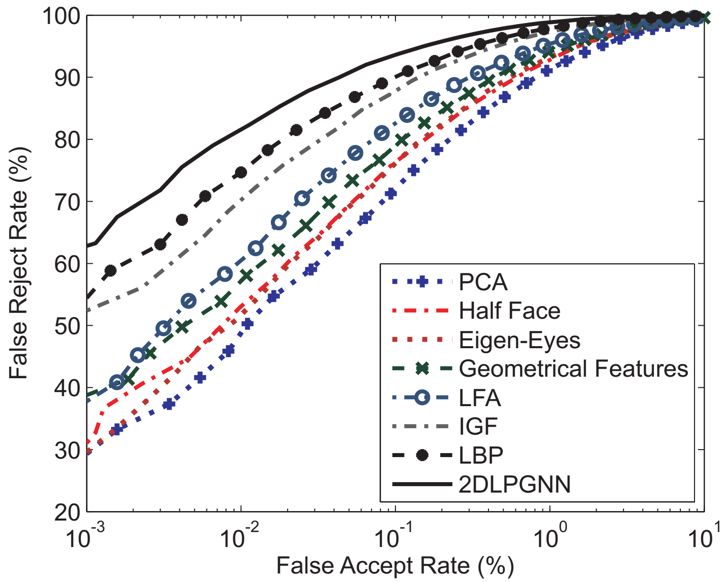

Fig. 12. ROC to evaluate the performance of face recognition algorithms on the face disguise database

Verification accuracy at 0.01% FAR								
Variations	Appearance based algorithms			Feature based algorithms		Texture based algorithms		
	PCA	Half-face	Eigen-eyes	GF	LFA	IGF	LBP	2DLPGNN
Minimal variations	60.3	59.5	63.1	61.3	63.7	74.2	85.5	96.9
Hair	56.8	61.4	57.2	61.9	63.1	73.8	85.1	96.4
Beard and moustache	32.2	34.1	60.5	53.2	54.5	58.7	61.0	77.3
Glasses	41.4	44.6	6.9	52.4	53.8	57.6	62.5	81.9
Cap and hat	55.7	56.8	50.4	58.9	61.4	71.0	80.4	86.3
Lips, nose, and eyebrow	56.3	59.2	47.7	49.1	56.3	70.9	78.6	89.2
Aging and wrinkles	49.6	53.9	41.6	51.8	54.9	55.1	70.3	80.8
Multiple variations	14.7	16.4	30.3	31.6	32.0	32.8	50.3	65.6
Overall	48.2	52.9	51.1	58.0	60.4	70.1	74.7	82.0

Table 3. Verification performance of the appearance, feature and texture based face recognition algorithms for different disguise variations

security applications such as border crossing and terrorist watch list, when an individual attempts to defraud law enforcement by altering his or her physical appearance with disguises. This chapter emphasizes this important aspect of face recognition. It describes different types of disguise variations and experimentally analyzes their effect on face recognition algorithms. The performance of appearance based algorithms, feature based algorithms, and texture based algorithms are compared using the heterogeneous face database and the disguise face database. Experimental results suggest that high degree of disguise variations is more challenging to address compared to variations in pose, expression and illumination. Furthermore, it also suggests that a careful and thorough investigation is required to develop a robust face recognition algorithm that can fulfil the operational needs of real world applications.

7. Acknowledgement

The authors would like to thank NIST, Robotics Institute CMU, CMU AMP Research Lab, Dr. A.R. Martinez, Dr. E.G.L. Miller and Equinox Corporation for granting us access to the face databases used in this research.

8. References

Ahonen, T.; Hadid, A. & Pietikäinen, M. (2006). Face description with local binary patterns: application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 12, pp. 2037-2041

Alexander, J. & Smith, J. (2003). Engineering privacy in public: Confounding face recognition, privacy enhancing technologies. *Proceedings of International Workshop on Privacy Enhancing Technologies*, pp. 88–106

CMU AMP face database: http://amp.ece.cmu.edu/projects/FaceAuthentication/download.htm

Cox, I.J.; Ghosn, J. & Yianilos, P.N. (1996). Feature-based face recognition using mixture-distance. *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 209-216

Delac, K. & Grgic, M. (2007) Face recognition, I-TECH Education and Publishing

Equinox face database: http://www.equinoxsensors.com/products/HID.html

Faces software: http://www.iqbiometrix.com/products_faces_40.html

Huang, G.B.; Ramesh, M.; Berg, T. & Learned-Miller E. (2007). Labeled faces in the wild: A database for studying face recognition in unconstrained environments. University of Massachusetts, Amherst, Technical Report

Liu, C. & Wechsler, H. (2003). Independent component analysis of Gabor features for face recognition. *IEEE Transactions on Neural Networks*, Vol. 14, No. 4, pp. 919-928

Li, S. & Jain, A. (2005) Handbook of face recognition. New York: Springer

Martinez, A. & Benavente, R. (1998). The AR face database. Computer Vision Center, Technical Report.

Pamudurthy, S.; Guan, E.; Mueller, K. & Rafailovich M. (2005). Dynamic approach for face recognition using digital image skin correlation. *Proceedings of Audio- and Video-based Biometric Person Authentication*, pp. 1010–1018

Penev, P. & Atick, J. (1996). Local feature analysis: a general statistical theory for object representation. *Network: Computation in Neural Systems*, Vol. 7, pp. 477-500

Phillips, P.J.; Moon, H.; Rizvi, S. & Rauss, P.J. (2000). The FERET evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1090–1104

Phillips, P.; Flynn, P.; Scruggs, T.; Bowyer, K. & Worek, W. (2006). Preliminary face recognition grand challenge results, *Proceedings of International Conference on Automatic Face and Gesture Recognition*, pp. 15–24

Phillips, P.; Scruggs, W.; O′ Toole, A.; Flynn, P.; Bowyer, K.; Schott, C. & Sharpe, M. (2007). FRVT 2006 and ICE 2006 large-scale results, NIST Technical Report NISTIR 7408

Ramanathan, N.; Chowdhury, A. & Chellappa, R. (2004). Facial similarity across age, disguise, illumination and pose, *Proceedings of International Conference on Image Processing*, Vol. 3, pp. 1999–2002

Silva P. & Rosa, A.S. (2003). Face recognition based on eigeneyes. *Pattern Recognition and Image Analysis*, Vol. 13, No. 2, pp. 335–338

Sim, T.; Baker, S. & Bsat, M. (2003). The CMU pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 12, pp. 1615-1618

Singh, R.; Vatsa, M. & Noore A., (2008). Face recognition with disguise and single gallery images. *Image and Vision Computing*, doi:10.1016/j.imavis.2007.06.010

Wechsler, H. (2006). Reliable Face Recognition Methods: System Design, Implementation and Evaluation. Springer

Yale face database: http://cvc.yale.edu/projects/yalefaces/yalefaces.html

Zhao, W.-Y.; Chellappa, R.; Phillips, P. J. & Rosenfeld A. (2003). Face recognition: A literature survey. *ACM Computing Survey*, Vol. 35, No. 4, pp. 399–458