International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:2, 2012

Assessment of Time-Lapse in Visible and Thermal Face Recognition

Sajad Farokhi, Siti Mariyam Shamsuddin, Jan Flusser, Usman Ullah Sheikh

Abstract**—**Although face recognition seems as an easy task for human, automatic face recognition is a much more challenging task due to variations in time, illumination and pose. In this paper, the influence of time-lapse on visible and thermal images is examined. Orthogonal moment invariants are used as a feature extractor to analyze the effect of time-lapse on thermal and visible images and the results are compared with conventional Principal Component Analysis (PCA). A new triangle square ratio criterion is employed instead of Euclidean distance to enhance the performance of nearest neighbor classifier. The results of this study indicate that the ideal feature vectors can be represented with high discrimination power due to the global characteristic of orthogonal moment invariants. Moreover, the effect of time-lapse has been decreasing and enhancing the accuracy of face recognition considerably in comparison with PCA. Furthermore, our experimental results based on moment invariant and triangle square ratio criterion show that the proposed approach achieves on average 13.6% higher in recognition rate than PCA.

Keywords**—**Infrared Face recognition, Time-lapse, Zernike moment invariants

I. Introduction

HUMAN face recognition is a biometric approach which is employed to recognize or verify the identity of a living person based on his/her physiological characteristics by means of automatic methods. It plays an important role in many application areas such as security systems, authentication, intelligent machines and surveillance. Despite considerable progress and some practical successes, face recognition is still a challenging task in the field of computer vision and pattern recognition. The wide-range variations of human face, due to time-lapse, expression, pose and illumination, lead to highly complicated distribution which decreases the accuracy of recognition greatly. Face recognition based on visible spectrum is gaining acceptance as a superior biometric in face recognition systems due to high resolution of acquired samples which are formed due to reflectance. Because of three-dimensional structure of face, external light and angle of incident of light play a significant role in identifying the face in different applications [1]-[2].

Sajad Farokhi and Siti Mariyam Shamsuddin are with the Faculty of Computer Science and Information System, Universiti Teknologi Malaysia, Johor, Malaysia (e-mail: fsajad2@live.utm.my, mariyam@utm.my).

Jan Flusser is with the Institute of Information Theory and Automation, Academy of Science of the Czech Republic, Prague, Czech Republic (e-mail: flusser@utia.cas.sz).

Usman Ullah Sheikh is with the Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor, Malaysia (e-mail: usman@fke.utm.my).

On the other hand, face recognition based on thermal infrared spectrum has received much more attention, because they are entirely free from the influence of external variable light and they can be employed even in total darkness [3]. However, thermal images have their own deficiencies such as, physical activity, stress, time-lapse and health conditions [4]-[5]. Experimental results in [6] show that the accuracy of IR face systems degrades sharply when time-lapse occurs, i.e., there is delay between acquisition of testing and training samples. The time interval can be few weeks, months or even a year. In [7] the comparative study between visible and infrared imagery for face recognition is conducted. It is shown that temperature variation in thermal images affects the accuracy of system as the external light do on visible images. In [8] thermal face recognition over time, has been studied and the effect of time-lapse has been investigated and reported. It has been indicated that the accuracy of face recognition in thermal domain degrades dramatically in comparison with visible images and fusion of both spectrums, could be the best one. A novel method for reducing the effect of time lapse is reported in [9]. Its idea is based on using block-PCA as a feature extractor. In spite of good results in all aforementioned techniques which have used PCA as a superior feature extractor, we will show in section II that the applicability of PCA is limited. Thus to overcome the shortcomings of aforementioned methods, a new research based on moment invariants [10] and nearest neighbor classifier is conducted to study the effect of time-lapse on both modalities and enhance the performance of face recognition system. To the best of our knowledge such a moment based study for time-lapse investigation has not been reported yet.The reminder of paper is organized as follows: Section II introduces principal component analysis. Zernike moments are studied in section III. Feature selection is expressed in section IV. Nearest neighbor classifier is discussed in section V. Proposed system and experimental results are given in section VI. Performance analysis is expressed in section VII. Final conclusion is presented in section VIII.

II. Principal Component Analysis

Principal component analysis (PCA) is one of the most prevalent and successful methods which has been employed widely in the field of computer vision and pattern recognition. It was introduced by Sirovich and Kirby [11] to represent images of human faces. In 1991, the famous Eigenfaces method was presented by Turk [12] and Petland and implemented in face recognition tasks. Its main procedure is to decompose face images into a small set of feature vectors called eigenfaces(Fig. 1) and compare the position of new face images with those of known faces. Although PCA represents an efficient, simple and accurate face recognition model, there are several weaknesses in this concept as follows:

1) It may have small discriminability.
2) It needs a lot of samples in training phase which leads to large computational load.
3) It is very sensitive to facial expression and noise.

Fig. 1 Eigenfaces of a sample in UND database

III. Zernike Moment Invariants

A. Introduction and Definition

Based on the theory of orthogonal polynomials, Teague introduced Zernike moments (ZMs) for image analysis to construct rotation moment invariants [13]. The application of ZM as feature extractor has been a lot of interest in face recognition due to rotation invariance property that makes it insensitive to noise [14]-[15]. In contrary to PCA, they have shown good invariance to facial expressions [16]-[17]. Moreover, it is not necessary to re-train all train samples, when a new image is added to training set. The kernel of Zernike moments is a set of orthogonal Zernike polynomials defined on a unit circle in Cartesian coordinates. Basically, the following generalized expression is employed to obtain rotation invariant moments:

$$F_{pq} = \iint f(r,\theta)\, g_{pq}(r)\, e^{\hat{j} q\theta} r\,dr\,d\theta, \quad \hat{j} = \sqrt{-1} \tag{1}$$

F_{pq} is the pq-order moment, g_{pq} is a function of radial variable, and r, q are integer parameters[18]. It is not difficult to show that the value of $|F_{pq}|$ is invariant to rotation [19]. For a continuous image function $f(x,y)$, Zernike moment of order p with repetition q , is given as follows:

$$Z_{pq} = \frac{p+1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} V_{pq}^{*}(r,\theta)\, f(r,\theta)\, r\,dr\,d\theta\,, \quad |r| \le 1 \tag{2}$$

The symbols * is the sign of complex conjugate and V_{pq} denotes Zernike polynomial of order p and repetition q and is defined as follows:

$$V_{pq}(r,\theta) = R_{pq}(r)\, e^{\hat{j} q\theta} \tag{3}$$

The real-valued radial polynomial R_{pq}, is expressed as follows:

$$R_{pq} = \sum_{k=0}^{\frac{p-|q|}{2}} (-1)^k \frac{(p-k)!}{k!\left(\frac{p+|q|}{2}-k\right)!\left(\frac{p-|q|}{2}-k\right)!} r^{p-2k} \tag{4}$$

Where $p \le q$ *and* p-$|q|$ is even [18]. The coefficients of R_{pq} up to the 10^{th} degree are presented in Fig. 2.

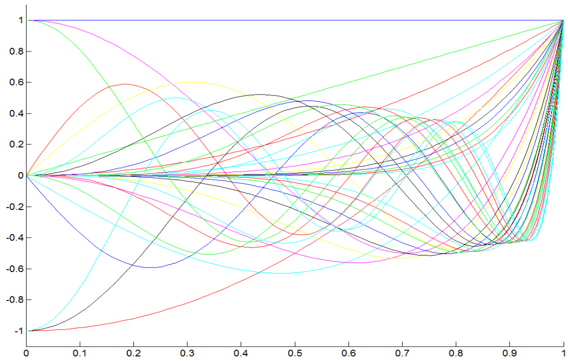

Fig. 2 Zernike Polynomials up to 10th degree

B. Discrete Approximation of Zernike Moments

Since Zernike moments are defined over polar coordinate inside a circle, their calculation needs a linear transformation of the image coordinates to an appropriate domain inside a unit circle which is written as:

$$x_i = \frac{\sqrt{2}}{N-1} i - \frac{1}{\sqrt{2}},\; y_j = \frac{\sqrt{2}}{N-1} j + \frac{1}{\sqrt{2}}$$
$$r_{ij} = \sqrt{x_i^2 + y_j^2},\; \theta_{ij} = \tan^{-1}\left(\frac{y_i}{x_i}\right) \tag{5}$$

In a nutshell, the discrete approximation of continuous Zernike moments is written as follows:

$$Z_{pq} = \lambda_Z(p,R,C) \sum_{i=0}^{R-1} \sum_{j=0}^{C-1} R_{pq}(r_{ij})\, e^{-\hat{j} q\theta_{ij}} f(i,j) \tag{6}$$

Where N=max(R,C) and $\lambda_Z(p,R,C) = \dfrac{2(p+1)}{\pi(R-1)(C-1)}$ (7)

IV. Feature Selection

The feature selection, also called Feature Subset Selection (FSS) has been widely used in pattern recognition systems to reduce dimensionality and enhance the performance, speed and accuracy of classifiers [20]-[21]. As a matter of fact, the aim of feature selection is to select a number of evaluated features from extracted feature set which gives minimum classification error. The evaluation of features is performed by an objective function. The general procedure of feature selection is shown in Fig. 3. Sequential Forward Selection (SFS) is the simplest greedy search algorithm which is based on bottom-up search procedure. The procedure is started by an empty set and selected features which are evaluated by objective function are added respectively to the empty set based on their mean square error which is calculated by objective function. Hence an enhanced subset with high discrimination power can be selected by SFS.

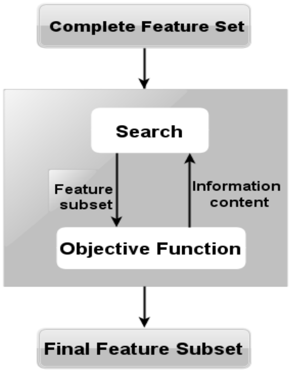

Fig. 3 General procedure of feature selection [22]

V. Nearest Neighbor Classifier

The nearest neighbor (NN) classifier is one of the most popular, powerful and simple classifiers which has been employed in many applications [23]. This classifier can be implemented by some different criterions such as cosine and Euclidean distance. Euclidean distance is the most prevalent one which has been used extensively as a superior criterion. Despite its much strength, it ignores the correlation which plays an important role for measuring the similarity of two feature vectors. On the other hand, cosine criterion considers correlation but ignores the distance between two feature vectors. As a result, triangle square ratio is used to compensate the deficiencies of aforementioned criterions and consider both distance and correlation between two feature vectors. This criterion is introduced in [24] for classification of facial images in wavelet domain. Let V_1 and V_2 be two vectors and θ be the include angle of them, the triangle square ratio is defined as:

$$TSR(V_1,V_2)=\frac{\| V_1-V_2 \|_2^2}{\| V_1 \|_2^2+\| V_2 \|_2^2}=1-\frac{2\| V_1 \|_2.\| V2 \|_2}{\| V_1 \|_2^2+\| V_2 \|_2^2}\cos(\theta) \quad (8)$$

As proven in [24], $\mathrm{TSR}(V_1,V_2)\to 0$ if and only if $\| V_1\|_2 \to \| V_2\|_2$ and $\theta \to 0$, which shows the correlation between V_1 and V_2 should approach 1. As a result, the TSR measures the similarity of V_1 and V_2 based on argument and modulus of each vector. Based on our results in section VI, the superiority and efficiency of TRS in comparison with Euclidean distance is obvious.

VI. Proposed System and Experimental Results

The proposed face recognition system is composed of four main parts as shown in Fig. 4. The most important objectives of this study are as follows:

1) Evaluate the performance of orthogonal moment invariants in time-lapse face recognition and compare the results with PCA.
2) Evaluate the performance of nearest neighbor classifier with new criterion and compare the results with Euclidean distance as a traditional criterion for NN classifier.

This section consists of five subsections as follows: in section *A*, data collection is studied, in section *B*, preprocessing stage is described, section *C* and *D* present two different experiments.

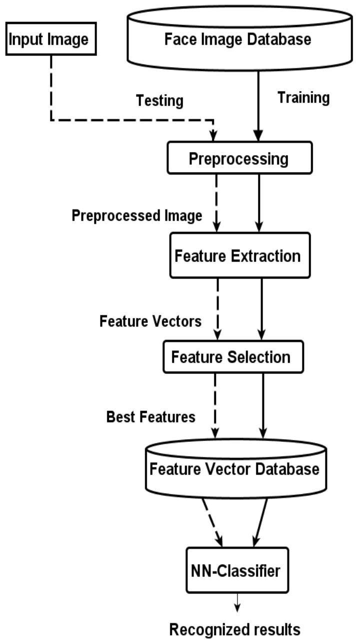

Fig. 4 The general procedure of face recognition system

A. Data Collection

The Notre–Dame Database [4]-[25] time-lapse face database (Collection X1) is used in our experiments which consists of 2292 IR frontal face images and 2292 visible frontal face images from 82 human subjects captured from 2002-2004 with a Merlin long-wavelength infrared camera. Due to IR's opaqueness to glass, all images have been taken without glass. In our experiments, images of 20 subjects from the UND face database were selected. Each subject attended 10 acquisition sessions for each season. So 10 images are used for training and 10 images are used for testing. There is no overlap between testing and training set. While the training set contains no facial expressions or time-lapse, the test set is composed of several images containing variations in time, facial expressions and head rotation. Tolerance for head rotation is utmost 15 degrees. Some sample images of a single subject in visible and infrared imagery are provided in Fig. 5.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:2, 2012

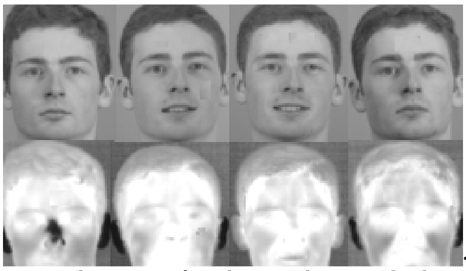

Fig. 5 Sample images of a subject in the UND database which were taken over the span of several months

B. Preprocessing

Generally, image preprocessing is used for normalization of image before feature extraction. Acquired facial images normally include background, hair, clothing, etc, which can deteriorate the recognition performance of the system. If the whole image is taken into consideration for feature extraction, the performance and accuracy of the system may be affected and decreased. This impact can be minimized by normalization. The procedure for normalization is as follows:

1) Face detection: First step in our preprocessing procedure is to locate the face in input images. Thus Viola-Jones [26] algorithm has been employed which minimizes computation time while achieving high detection accuracy.
2) Image registration: Image registration is one of the most important procedures which is done to align two images of the same person (visible and infrared). It is a prerequisite step for thermal and visible images before image fusion [27]. Typically, visible images, and thermal images are considered as reference. This step is performed manually and finally, all of images are resized to 320×240 pixels.
3) Masking: Masking is done to remove some parts of facial images such as background, hair, clothing, etc. This is to make sure that the face recognition system is not affected by redundant features.
4) Histogram equalization: lighting and sensor differences may reduce the performance of the system. Thus histogram equalization is used for normalizing the image histogram and decreasing image variation.

C. Experiment I

In the first experiment, principal component analysis is employed and the results are considered. Table I shows the mean recognition rate for three different modalities of images.

TABLE I
MEAN TOP-MATH RECOGNITION PERFORMANCE FOR TIME LAPSE EXPERIMENTS WITH PCA

Modality	Visible	Infrared	Fusion
PCA	76.23	61.92	83.24

D. Experiment II

In the second experiment Zernike rotation invariants are used to compensate head rotation which is not considered in registration step. Hence Zernike moment $|Z_{pq}|$ up to order 20 is calculated by (6), for visible, infrared and fused images which resulted in feature vectors of size 227. Mahalanobis distance is chosen as an objective function and the best feature vectors of size 62 are selected by SFS. Finally, classification accuracy is measured by different criterions of NN classifier and results are noted. The mean recognition rate is reported in Table II.

TABLE II
MEAN TOP-MATH RECOGNITION PERFORMANCE FOR TIME LAPSE EXPERIMENTS WITH PCA AND ZERNIKE MOMENTS

Modality	Visible	Infrared	Fusion
ZMs with NN(Euclidian)	87.24	73.23	92.37
ZMs with NN(TSR)	91.35	76.93	95.23

VII. PERFORMANCE ANALYSIS

Two main experiments are conducted to compare and evaluate the performance of moment invariants and TSR. Zernike moments and PCA are used for feature extraction. Then, for all modalities, final results for Zernike moments with TSR and PCA are plotted in Fig. 6. Some important results based on Table I, Table II and Fig. 6 can be concluded as follows:

1) As other studies have shown [6]-[28], there is exist no consistent trend for both visible and thermal images in this database.
2) The fluctuation of Fig. 6, based on Zernike moment invariants, is more consistent than that of PCA which shows that orthogonal moment invariants has better stability due to their salient and discriminatory characteristic on Notre–Dame database in comparison with PCA. Some samples which are misclassified by PCA while recognized by Zernike moments are depicted in Fig. 7.
3) TSR as a sophisticated criterion for NN classifier performs better than the Euclidean distance. Hence the advantage of using TSR as a new criterion for NN classifier is obvious. This striking asset can be expressed by the inherent characteristic of TSR which considers both correlation and similarity. Some samples which are recognized by TSR while misclassified by Euclidean distance are shown in Fig. 8.
4) The effectiveness and good performance of orthogonal moment invariants as a dominant feature for face recognition, in comparison with PCA is noticeable. Orthogonal moments not only need smaller training set, but also represent good discrimination power in comparison with PCA. This can be interpreted by invariance and global characteristic of Zernike moments which almost leads to invariance of features regards to time-lapse.
5) As many previous studies have expressed [4]-[29]-[30] fusion of visible and thermal, strengthens the performance

of system considerably. However, one has to take into account that registration of visible and thermal images needs a lot of time and can considerably affect the speed and accuracy of system.

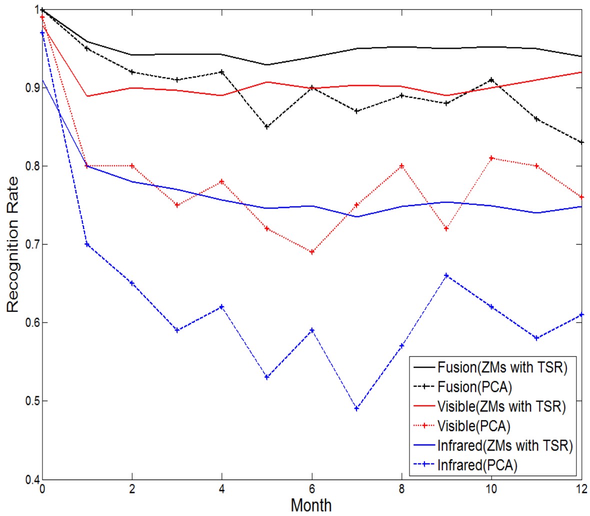

Fig. 6 Mean recognition results of ZMs and PCA for visible, infrared and fusion as a function of months elapsed between training and testing


Fig. 7 Sample of images which are misclassified by PCA whereas recognized by Zernike Moments


Fig. 8 Sample of images which are misclassified by Euclidian distance whereas recognized by TSR

VIII. Conclusion

In this paper, a holistic method based on orthogonal moments is presented to compensate time-lapse which can reduce the accuracy of system greatly. The Notre–Dame database images are used for conducting all experiments. The method uses the values of Zernike moment invariants with lower dimension instead of whole image. Some of these salient features are chosen by means of sequential forward selection algorithm with Mahalanobis distance as an objective function. A modified criterion for nearest neighbor classifier which considers both distance and correlation is employed to measure the accuracy of system. Finally, comparison analysis between PCA and moment invariants with two different criterions is conducted and the results are analyzed. Experimental results show the efficiency of TRS in comparison with Euclidian distance. Moreover it is deduced that due to inherently global characteristic of orthogonal moment invariants; they represent high accuracy in time-lapse scenarios and can be chosen as a suitable feature in face recognition algorithms. With regards to previous analysis and recent results by researchers on time-lapse recognition, we also certify that the use of thermal images for biometric identification as a superior modality can be a challenging task. This is due to large variation of thermal images which can be created by a small variation in temperature. Undoubtedly, fusion of visible and thermal can be the best solution to overcome the deficiencies of both modalities, however higher computational complexity and problematic registration procedure of image fusion should be considered.

Acknowledgment

This work is supported by Ministry of Higher Education (MOHE) and *Universiti Teknologi Malaysia (UTM)* under Research University Grant (VOT Q.J130000.7128.00H71). Authors would like to thank the *Soft Computing Research Group (SCRG)* for the support and incisive comments in making this study a success, *Institute of Information Theory and Automation (UTIA)* for providing MATLAB codes and *the University of Notre Dame (UND)* for providing UND biometric database to carry out this experiment.

References

[1] Socolinsky, Selinger and Neuheisel, *Face recognition with visible and thermal infrared imagery*. Computer Vision and Image Understanding, 2003. **91**(1-2): p. 72-114.

[2] Störring, Andersen and Granum, *Physics-based modelling of human skin colour under mixed illuminants*. Robotics and Autonomous Systems, 2001. **35**(3-4): p. 131-142.

[3] Kong, Heo, Abidi, Paik, and Abidi, *Recent advances in visual and infrared face recognition--a review*. Computer Vision and Image Understanding, 2005. **97**(1): p. 103-135.

[4] Xin Chen , Patrick J. Flynn and Bowyer. *Visible-light and Infrared Face Recognition*. in *The proceedings of Workshop on Multimodal User Authentication*. 2003. Santa Barbara, CA USA.

[5] Buddharaju, Pavlidis, Tsiamyrtzis, and Bazakos, *Physiology-Based Face Recognition in the Thermal Infrared Spectrum*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007. **29**(4): p. 613-626.

[6] Chen, Flynn and Bowyer, *IR and visible light face recognition*. Computer Vision and Image Understanding, 2005. **99**(3): p. 332-358.

[7] Joseph Wilder, P. Jonathon Phillips, Cunhong Jiang, and Wiener. *Comparison of visible and infra-red imagery for face recognition*. in *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*. 1996.

[8] Xin Chen, Patrick J. Flynn and Bowyer. *PCA-based face recognition in infrared imagery: baseline and comparative studies*. in *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*. 2003.

[9] Yu, Zhihua, Zhijun, Jucheng, Shiqian, and Feng. *Time-Lapse Data Oriented Infrared Face Recognition Method Using Block-PCA*. in *2010 International Conference on Multimedia Technology (ICMT)*. 2010.

[10] Jan Flusser, *Moment Invariants in Image Analysis*. Proceedings of World Academy of Science, Engineering and Technology, 2006. **1307-6884**: p. 196-201.

[11] Jian, Zhang, Frangi, and Jing-yu, *Two-dimensional PCA: a new approach to appearance-based face representation and recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004. **26**(1): p. 131-137.

[12] M. Turk and Pentland, *Eigenfaces for Recognition*. Journal of Cognitive Neuroscience, 1991. **3,No 1**: p. 71-86.

[13] Jan Flusser, Tomas Suk and Zitova, *Moments and Moment Invariants in Pattern Recognition*. First ed. 2009, chichester: wiley. 291.

[14] Zahran, Abbas, Dessouky, Ashour, and Sharshar. *High performance face recognition using PCA and ZM on fused LWIR and VISIBLE images on the wavelet domain*. in *International Conference on Computer Engineering & Systems*. 2009.

[15] Lajevardi seyed mehdi and Hussain, *Higher order orthogonal moments for invariant facial expression recognition.* Digital Signal Processing, 2010. **20**(6): p. 1771-1779.
[16] Nabatchian, Makaremi, Abdel-Raheem, and Ahmadi. *Pseudo-Zernike Moment Invariants for Recognition of Faces Using Different Classifiers in FERET Database*. in *Third International Conference on Convergence and Hybrid Information Technology*. 2008.
[17] Zhi and Ruan. *A Comparative Study on Region-Based Moments for Facial Expression Recognition*. in *Congress on Image and Signal Processing*. 2008.
[18] Ali Broumandnia and Shanbehzadeh, *Fast Zernike wavelet moments for Farsi character recognition.* Image and Vision Computing, 2007. **25**(5): p. 717-726.
[19] Shen and Ip, *Discriminative wavelet shape descriptors for recognition of 2-D patterns.* Pattern Recognition, 1999. **32**(2): p. 151-165.
[20] A. Marcano-Cedeño, J. Quintanilla-Domínguez, Cortina-Januchs, and Andina. *Feature selection using Sequential Forward Selection and classification applying Artificial Metaplasticity Neural Network*. in *36th IEEE Annual Conference on Industrial Electronics Society*. 2010.
[21] Durga Prasad Muni, Nikhil R. Pal and Das, *Genetic programming for simultaneous feature selection and classifier design.* IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2006. **36**(1): p. 106-117.
[22] Gutierrez-Osuna, *a lecture on sequential feature selection*. 2008, A&M University: Texas. p. 1-15.
[23] Nor'aini, Raveendran and Selvanathan. *Human Face Recognition using Zernike moments and Nearest Neighbor classifier*. in *4th Student Conference on Research and Development*. 2006.
[24] Chao-Chun Liu, Dao-Qing Dai and Yan, *Local discriminant wavelet packet coordinates for face recognition.* Journal of Machine Learning Research, May 2007. **Vol. 8**: p. 1165-1195.
[25] Flynn, Bowyer and Phillips, *Assessment of Time Dependency in Face Recognition: An Initial Study Audio- and Video-Based Biometric Person Authentication*, J. Kittler and M. Nixon, Editors. 2003, Springer Berlin / Heidelberg. p. 1057-1057.
[26] Viola and Jones. *Rapid Object Detection Using a Boosted Cascade of Simple Features*. in *Computer Vision and Pattern Recognition*. 2001.
[27] Hrkać, Kalafatić and Krapac, *Infrared-Visual Image Registration Based on Corners and Hausdorff Distance Image Analysis*, B. Ersbøll and K. Pedersen, Editors. 2007, Springer Berlin / Heidelberg. p. 383-392.
[28] Socolinsky and Selinger. *Thermal face recognition over time*. in *Proceedings of the 17th International Conference on Pattern Recognition*. 2004.
[29] Sang-ki Kim, Hyobin Lee, Sunjin Yu, and Lee. *Robust Face Recognition by Fusion of Visual and Infrared Cues*. in *1ST IEEE Conference on Industrial Electronics and Applications*. 2006.
[30] Buddharaju and Pavlidis, *Multispectral Face Recognition: Fusion of Visual Imagery with Physiological Information*, in *Face Biometrics for Personal Identification*, R. Hammoud, B. Abidi, and M. Abidi, Editors. 2007, Springer Berlin Heidelberg. p. 91-108.