Optimal fusion scheme selection framework based on genetic algorithms, for multimodal face recognition

W. Ben Soltana[1] M. Ardabilian[1] L. Chen[1] C. Ben Amar[2]

[1] LIRIS (Laboratoire d'InfoRmatique en Image et Systèmes d'information)

Université de Lyon, CNRS
Ecole Centrale de Lyon, LIRIS, UMR5205, F-69134, France

{wael.ben-soltana, Mohsen.Ardabilian, Liming.Chen}@ ec-lyon.fr

[2] REGIM (REsearch Group on Intelligent Machines)

University of Sfax
National Engineering School of Sfax, 3038, Sfax, Tunisia

chokri.benamar@ieee.org

Abstract

In this paper, we consider the problem of feature selection and classifier fusion and discuss how they should be reflected in the fusion system architecture. We employed the genetic algorithm with a novel coding to search the worst performing fusion strategy. The proposed algorithm tunes itself between feature and matching score levels, and improves the final performance over the original on two levels, and as a fusion method, it not only contains fusion strategy to combine the most relevant features so as to achieve adequate and optimized results, but also has the extensive ability to select the most discriminative features and their appropriate classifiers. Sparse Representation Classifier (SRC) and Nearest Neighbor classifier with euclidean distance, mahalanobis distance, cosine distance and correlation distance are exploited to calculate all the similarity measures. Experiments are provided on the FRGC database and show that the proposed method produces significantly better results than the baseline fusion methods.

Keywords

Genetic algorithm, fusion strategy, feature level, score matching level, classifier selection, classifier fusion, Sparse Representation Classifier.

1 Introduction

In recent years, there are many research works and studies of multiple classifier systems. It has been frequently demonstrated that combing classifiers can offer significant classification performance improvement for a number of non-trivial pattern recognition problems [1].

Fusion strategies can be roughly classified into three main categories: fusion at an early stage, fusion at a later stage and hybrid fusion. Many systems that integrate information at an early stage are believed to be more effective than those that perform integration at a later stage. Therefore, while it is relatively more difficult to achieve in practice [2], fusion at early stage has drawn more attention in recent years. There exist two types of early fusion: fusion at data level (for example 3D image [3] or 3D/2D image [4]), and fusion at feature level [2]. In fact, at the feature level the concatenated feature vectors may contain noisy or redundant data, thus leading to decreased performances of the classifier [5]. In this case, feature selection procedure is an important step. It is essentially an optimization problem that involves searching within the space of possible feature subsets to find one subset that is optimal (or near-optimal) with respect to a certain criterion. Several search strategies have been put forward and can be classified into three categories: optimal, heuristic, and randomized. Exhaustive search is the most straightforward approach to optimal feature selection and it is guaranteed to find the optimal subset. However, since the number of possible subsets grows exponentially, exhaustive search becomes not feasible and impractical even for moderate feature numbers. The only optimal feature selection method, which avoids the exhaustive search, is based on the branch and bound algorithm [5, 6]. Best individual features, sequential forward selection (SFS) and sequential backward selection (SBS) [5] are three well-known heuristic suboptimal feature selection schemes. Combining SFS and SBS gives birth to plus l-take away r feature selection. A generalization of the plus l-take away r method is twofold: Sequential forward floating search [5, 7] and sequential backward floating search [5, 8], where l and r are determined automatically and updated dynamically. Evolutionary algorithms [1] are random search algorithms. Among them, genetic algorithms (GAs) include a subset of evolutionary algorithms focusing on the application of selection, mutation, and recombination to a population of competing problem solutions. Obviously, GAs are prime candidates for random probabilistic search algorithms within the context of feature selection.

In fusion at later stage, there are three fusion sub-levels: score match level [9], rank level [10] and decision level [11]. Kittler and al. [12] presented and developed a common theoretical framework for these combining classifiers. At the first level, similarity scores generated by classifiers are combined by various techniques [13], for example, Sum Rule, Product Rule, etc. Gabrys and al. [14] developed a weighted soft combiners based on genetic algorithms. At the second level, sorted lists computed by classifiers are merged based on different approaches such as Borda Count and Logistic Regression [15]. At the third level, all the candidates of the classifiers are fused by adopting several methods [16], i.e., Majority Vote. The last category contains intermediate fusion schemes, such as serial fusion and multilevel fusion. The main motivation of the serial architecture [16] is to filter out the most similar K classes using a simple classifier and then to feed these K classes into a more complex and powerful second classifier. On the other side, there are few works that describe multilevel fusion. In [17], fusion is introduced in both feature level and confidence level.

In this work, we develop a common framework for selecting features, classifiers and combining the latest; we confirm that many existing schemes can be considered as special cases of our generic fusion scheme. We show that our fusion method is able to obtain a global sub-optimal solution while lessening the complexity of calculation. Other contributions of this paper are: the use of genetic algorithm with a novel coding strategy and sequential backward floating search for effective feature selection; at the same time an optimal fusion strategy scheme is generated. Furthermore, an appropriate classifier for each feature type was determined automatically.

The remainder of this paper is organized as follows: a framework for feature and score matching levels fusion is introduced in section2, and section 3 presents experimental results. Section 4 concludes the paper.

2 Optimal fusion scheme selection framework

The proposed framework is shown in Fig. 1. It is based on genetic algorithm, using a novel coding technique, to search the optimal fusion scheme.

2.1 Framework Overview

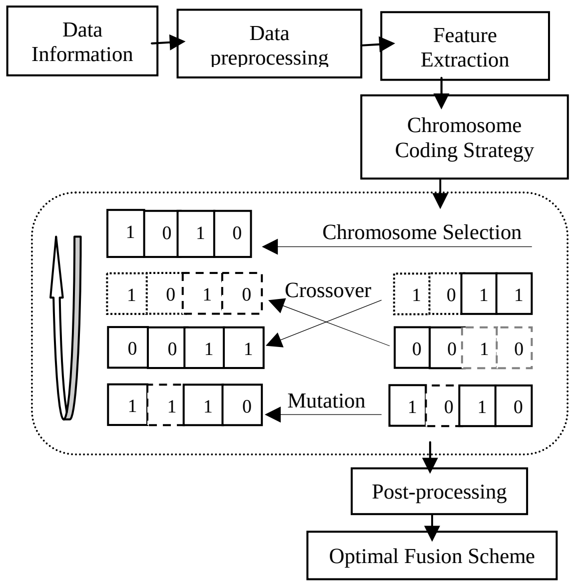

Figure 1 - *Algorithm Overview*

The proposed framework consists of four steps. In the first step, Data Preprocessing, the data validity and integrity are checked and noisy data is rectified. The second step consists to features extraction. For measurement cost and classification accuracy, Linear Discriminant Analysis (LDA) is used to reduce the dimensionality for each feature. The third step lead to: 1) finding one subset of features that is optimal with respect to the corresponding fusion scheme and 2) determining, automatically, an appropriate classifier for each feature type. So, all features are coded to form individual “chromosomes” according to the model described in the section 2.3. Furthermore, these chromosomes are used by a genetic algorithm [18] to encode the trial solution for the current problem. Iterative selection, crossover, and mutation were used to make evolve a new population. At each new generation, a new set of chromosomes is produced, using the fittest genes of the previous generation, for a better solution. Assessment of the satisfactory degree of this solution, encoded as individuals, is reflected in the fitness. In fact, fitness corresponds to performance rate of each fusion strategy represented by individual chromosome. This fitness is calculated according to eq. (2) taking into account different classifiers. Figure.2 illustrates this process. Also, the individuals with higher fitness have a high probability of being selected and producing offspring. The crossover operator produces better offspring by exchanging the characteristics of the parents. This enables the most efficient characteristics to be concentrated in the same individual. The mutation operator randomly changes the genetic representation of an individual and tends to inhibit the possibility of converging to a local optimum, rather than the global optimum. The evolution is carried out until a desired solution is arrived, or a pre-specified number of iterations are completed. The final solution with higher fitness represents the final fusion strategy.

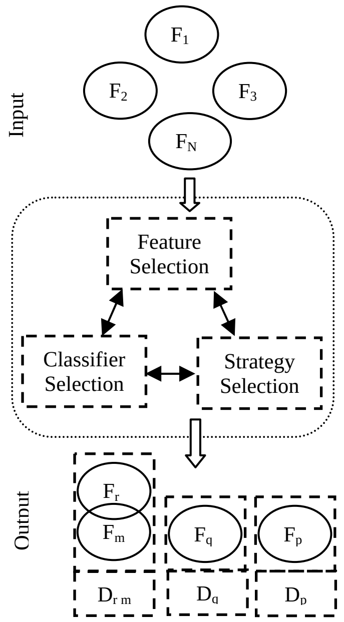

Figure 2 - *Genetic algorithm optimization for feature, classifier and strategy selection*

The last step can generate an optimal final fusion strategy. In fact, SBFS [5] is used to select the best features in new concatenated features (See Section 2.3). Furthermore, if the final strategy represents a matching score level strategy, we assign new different weights following eq.(3) to ameliorate the performance rate.

2.2 Performance Rate

The performance rate is calculated for each chromosome. So, all features output different scores. Min-Max Normalization [2] is used to map the matching scores to the range of [0, 1]. At the strategy selection stage, for each *g* in the gallery set, we compute a similarity score $S^{g,f}$ by classifier using each feature *f* with the probe. All these similarities $S^{g,f}$ are then sorted in a descending order. We assign to each score $S^{g,f}$ a weight $w^{g,f}$ which is a function of its ordered position $p^{g,f}$. Specifically, the weight $w^{g,f}$ is defined as:

$$w^{g,f} = f(p) \propto \ln\,(N_g/p^{g,f}). \tag{1}$$

where N_g is the number of the subjects in the gallery.
The matching score, in the strategy selection stage, between the *g* in the gallery and the probe is:

$$S(g) = \sum_{f\,\in features} w^{g,f}\,.\,S^{g,f} \tag{2}$$

This weighting strategy gives more importance to the scores ranked at the first positions and aims to discard wrong matching of each feature in test by assigning a lower weight to its corresponding similarity with a gallery sample.
At the post-preprocessing stage, we use another Genetic Algorithm to assign a weight $P^{g,f}$ to scores of a particular feature. The final matching score between g in the gallery and the probe is:

$$S(g) = \sum_{f\,\in features} P^{g,f}.\,w^{g,f}\,.\,S^{g,f} \tag{3}$$

The probe face is recognized as the one in the gallery which obtained the highest final score according to (3).

2.3 Feature subset and Strategy Selection

We propose a novel coding strategy to select simultaneously the efficient feature, the best classifier and the optimal fusion scheme. This coding strategy consists to divide the chromosome into two parts: Part A and Part B (See Figure.3). Given N features, Part A has *N* gene positions that correspond to each feature, and represented with integer values: 1 implies that the feature is active and used in feature level fusion, 0 implies that the feature is active and used in score level fusion, and -1 implies that the feature is inactive. Part B codes the fusion model that depends on the number N_F of active features at feature level fusion. In this model, we generate all possible combinations. However, we can't create a strategy that contains a single feature and we consider that combinations obtained by permutation are equivalent. Part B is also composed of two parts P1and P2: P1 refers to the model M and P2 associates the features in this model.

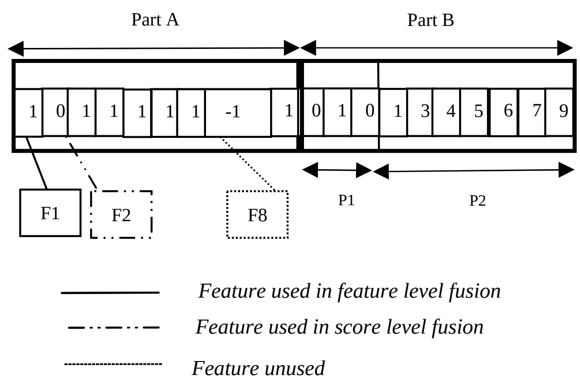

Figure 3 - *Example of chromosome coding strategy*

An example of this representation is illustrated in Figure 3. With a Part A as 1011111-11, we can generate 4 models M_i, with *i* in {1..4}: M_1=(7, 0), M_2= (2, 5), M_3= (3, 4), M_4=(2, 2, 3). The number of the selected model is represented in the chromosome by its binary code: the model M_2 is selected and represented by (010) and two vectors V1, V2 are created by concatenation, V1= [F1, F3] and V2= [F4, F5, F6, F7, F9]. The fusion strategy corresponds to a score matching level with V1, V2, and F2. The fitness of this strategy is calculated based on performance rate described in previous section. Stochastic universal sampling [19] is used to select best chromosomes “strategies”. Uniform crossover is used only on Part A and random mutation may occur on Part A or Part B of chromosome. Stopping criteria chosen for problem solving is selected from these conditions: 1) either the maximum number of iterations over the terminal number max of generations, 2) the best fitness value beyond the value of fitness limits.

3 Experimental Results

The proposed algorithm is tested in a face recognition application, where the objective is to find an optimal subset of features and their adequate fusion strategy.

3.1 Database, Experiment Settings and Feature Extracted

The FRGC [20] database was chosen for our experiments. Each face data consists of one 3D face model and its registered 2D color image. The 3D face models are first cropped with a sphere of radius of 80 *mm* centered at the nose tip and preprocessed with techniques in [21]. In order to avoid the impact of registration errors in our analysis, we used a manual registration method, namely Region based Iterative Closet Point (R-ICP) [22] for 3D face model. . As 2D color texture information is densely registered to its corresponding 3D face data, the previously cropped and registered 3D face model has also its 2D texture counterpart. The positions of both two eye inner corners are further used for rotation normalization. Finally, all the 2D color faces are converted to the gray-level, and resized to 80×92 pixels.

FRGC v1.0 dataset is used for estimating LDA parameters while FRGC v2.0 is utilized for training and test; 116 subjects having each 4 face models were selected from FRGC v1.0 to train subspace based approaches such as estimating LDA parameters. One face scan (3D+2D) with a neutral expression was selected from each subject to make a gallery of 410 subjects (gallery database) and 3541 face scans were treated as probes and were separated to build training database (2332 face scan) and test database (1209 face scan). The test database is divided into two subsets according to their expression labels. The first subset contains face scans with the neutral expression (713 probes), and the other one with face scans producing non-neutral expressions (496 probes). In the second step, we have used 2D and 3D features. Geometric features include normal (Nor Vec), binormal (BiN Vec), tangent vector (Tang Vec) [23] and curvature. Four categories of curvature-based features are extracted. The first two types rely on main directions corresponding to maximum (Max Curv) and minimum (Min Curv) curvatures [24]. The last two are their derivatives, i.e., the mean (Mean Curv) and Gaussian (Gauss Curv) curvatures. We further investigated another type of 3D feature based on the anthropometric (Anthr Mes) approach which advocates extracting a signature from some anthropometric points considered the most relevant. Three different features are extracted from the 2D texture images. The first one is the simple pixel-based method that encodes grayscale intensity (Intensity) values into a vector. The second one is a non-parametric feature namely Local Binary Patterns (LBP) [25]. The most important properties of LBP are the tolerance against the monotonic illumination changes and its computational simplicity. The third feature is extracted by Gabor filters (Gabor) [26] which are spatially localized and selective to spatial orientations and scales. Five different frequencies and eight equally spaced orientations are utilized to generate Gabor kernels.

For evaluating the proposed approach, experiments were designed in identification task with training and test stages. The training stage outputs the optimal fusion strategy. The same gallery database is used in all stages. In training stage, we achieve one experiment using gallery database and training database. The training database contains neutral and non neutral expressions. In test stage, three experiments were carried out: Neutral vs. Neutral, Neutral vs. Non-Neutral, and Neutral vs. All. In Neutral vs. Neutral and Neutral vs. Non-Neutral, only the neutral and non-neutral probe subsets were used.

3.2 Classifiers

Two classifiers, Sparse Representation Classifier and Nearest Neighbor Classifier were used in our experiments.

Sparse Representation Classifier: Sparse representation for signal classification (SRSC) is proposed in [27]. SRSC incorporates reconstruction properties, discriminative power and sparsity for robust classification. In [28], a general classification approach for (image-based) object recognition is proposed based on a sparse representation computed by L1-minimization. The method based on sparse representation can often achieve high performance based on a data dictionary [29].

Nearest Neighbor Classifier: Euclidean distance (4), Mahalanobis distance (5), cosine distance (6) and correlation distance (7) are introduced, and their performances are also compared in our experiments. A description of each of these metrics can be found below:

$$d(x,y)^2 = (x-y)(x-y)', \qquad (4)$$

$$d(x,y)^2 = (x-y)C^{-1}(x-y)', \qquad (5)$$

$$d(x,y) = 1 - \frac{(x\,y')}{(x'x)^{1/2}(y'y)^{1/2}}\ , \qquad (6)$$

$$d(x,y) = 1 - \frac{(x-x_M)(y-y_M)'}{\left((x-x_M)(x-x_M)'\right)^{1/2}\left((y-y_M)(y-y_M)'\right)^{1/2}}. \quad (7)$$

where x and y are two rows vectors to compare, C is the covariance matrix, x_M is the mean value of x, and y_M is the mean value of y.

3.3 Results and Analysis

First, LDA is applied to reduce dimensionality of all features in FRGC v2 database. In order to use the genetic algorithm in training stage, we define some parameters. Part A of chromosome is organized as follows: {Tang Vec, BiN Vec, Nor Vec, Gauss Curv, Max Curv, Mean Curv, Min Curv, LBP, Gabor, Anthr Mes, Intensity}. Five similarity measure of each feature was computed with SRC classifier (SRC), 1-Nearest Neighbor with euclidean distance measure (1-NN-ED), 1-Nearest Neighbor with mahalanobis distance measure (1-NN-MD), 1-Nearest Neighbor with cosine distance measure (1-NN-CosD), 1-Nearest Neighbor with correlation distance measure (1-NN-CorrD). The selection algorithm used a population of 50 chromosomes. The mutation rate was set to 0.1 and the GA was stopped after 100 generations for experiment.

Table 1. Rank-one recognition rate of individual type of feature and classifiers selected by the GA for the best fusion strategy

Classifier	Features			
SRC	Gabor	81.97	Bin Vec	73.12
	LBP	77.01	Anthr Mes	58.97
	Intensity	53.35		
1-NN-CorrD	Tang Vec	82.22		
	Mean-Gauss	74.16		

The final fusion strategy generated by training stage is coded as follows. Part A: 0,0,-1,1,-1,1,-1,0,0,0,0, Part B: [P1: 001, P2: 4, 6]. It consists firstly to concatenate {Mean Curv, Gauss Curv} in vector V_1. Secondly, we use this optimal subset {V_1, Tang Vec, BiN Vec, LBP, Gabor, Anthr Mes, Intensity} in score level fusion. The post processing step is used to optimize the final fusion strategy. Firstly, SBFS is used to select the best feature in the new concatenated vector V1. Secondly, new weight processing eq. (3) is used. The final recognition rate is 95.67% in training stage. In test stage, we apply the final fusion strategy. In this case, the final recognition rate is 97.27% using Neutral vs All experiment. Table 2 compares the proposed fusion strategy with other fusion approaches (simple sum rule (baseline method), Gökberk and al. [7], Mian and al.[2]). The performance of each feature and classifiers selected by the GA is displayed in Table1. Others experiments (Neutral vs Neutral and Neutral vs Non Neutral) are presented in Table 3. In all experiments, our method improves rank-one recognition accuracy as compared with other methods in three aspects: selecting the most discriminative features, selecting appropriate classifier and proposing an optimized fusion strategy.

Table 2. Identification Results Using Neutral vs All

Method	Training Phase	Test Phase
Simple Sum Rule	_	94.13%
Gökberk and al. [7]	_	95.28%
Mian and al.[2]	_	94.71%
Our approach	95.67%	97.27%

Table 3. Identification Results (Rank-one) Using Neutral vs Neutral (N-N), Neutral vs NonNeutral (N-nonN)

Method	N-nonN	N- N
Simple Sum Rule	91.33%	97.76%
Gökberk and al. [7]	92.75%	97.88%
Mian and al.[2]	90.73%	97.76%
Our approach	95.16%	99.16%

4 Conclusions and Future Works

In this paper, we developed a common framework for optimal fusion strategy selection. The proposed framework, based on a genetic algorithm and a novel associated coding strategy, generates automatically a subset of best features, an appropriate classifier for each feature, and an optimal fusion strategy scheme. Experiments are provided on the FRGC database and show that the proposed method produces significantly better results.

In future works, we can integrate other features and classifiers to improve the potential of our method. We intend also to extend this fusion scheme in order to generate the best model for each application. We plan as well to analyze the impact of the quality of information on fusion strategy.

Acknowledgement

This work was partially carried out within the French FAR3D project supported by ANR under the grant ANR-07-SESU-004 FAR3D.

References

[1] Dymitr Ruta, Bogdan Gabrys. Classifier selection for majority voting. *J. Inf. Fusion*, 6 (1), 63–81, (2005).

[2] Ajmal S. Mian, Mohammed Bennamoun, Robyn Owens. Keypoint detection and local feature matching for textured 3D face Recognition. International Journal of Computer Vision, 1-12, 2008.

[3] Theodoros Papatheodorou, Daniel Rueckert. Evaluation of Automatic 4D Face Recognition Using

Surface and Texture Registration. *Automatic Face and Gesture Recognition*, 321 – 326, May 2004.

[4] Gede Putra Kusuma and Chin-Seng Chua. Image Level Fusion Method for Multimodal 2D + 3D Face Recognition. *ICIAR* 2008, 984-992.

[5] Anil K. Jain, Robert P. W. Duin, Jianchang Mao, Statistical Pattern Recognition: A Review. IEEE Trans. Pattern Anal. Mach. Intell, 22(1), 4-37, 2000.

[6] Patrenahalli M. Narendra and Keinosuke Fukunaga. A Branch and Bound Algorithm for Feature Subset Selection. *IEEE Transactions on Computers*, vol. C-26, issue 9, 917-922, Sept 1977.

[7] Berk Gökberk , Helin Dutagacı, Lale Akarun, Bülent Sankur. Representation plurality and fusion for 3D face recognition. *IEEE Trans. on Systems Man and Cybernetics-Part B*, 38(1), 155–173, 2008.

[8] Ververidis, D., Kotropoulos, C.: Fast and accurate sequential floating forward feature selection with the Bayes classifier applied to speech emotion recognition. *Signal Processing* 88(12), 2956-2970, 2008.

[9] Jamie Cook, Mark Cox, Vinod Chandran, Sridha Sridharan. Robust 3D face recognition from expression categorization. *International Conference on Biometrics*, 271-280, 2007.

[10] Berk Gökberk, Albert Ali Salah and Lale Akarun. Rank-based decision fusion for 3D shape-based face recognition. *International Conference on Audio- and Video-Based Biometric Person Authentication*, 1019-1028, 2005.

[11] Timothy Faltemier, Kevin Bowyer, Patrick Flynn**.** 3D face recognition with region committee voting. *International Symposium on 3D Data Processing, Visualization, and Transmission*, 2006.

[12] Josef Kittler, Mohamad Hatef, Robert P.W. Duin, Jiri Matas. On Combining Classifiers. *IEEE Trans. Pattern Anal. Mach. Intell*, 226-239, 20(3) 1998.

[13] Afzal Godil, Sandy Ressler and Patrick Grother. Face recognition using 3D facial shape and color map information: comparison and combination. *Biometric Technology for Human Identification*, *SPIE*, 5404, 351–361, 2005.

[14] Bogdan Gabrys Dymitr Ruta. Genetic algorithms in classifier fusion. *Appl. Soft Comput.* 6(4), 337-347 (2006).

[15] Md. Maruf Monwar , Marina Gavrilova. FES: a system for combining face, ear, and signature biometrics using rank level fusion. *International Conference on Information Technology: New Generations*, 922-927, 2008.

[16] Berk Gökberk and Lale Akarun. Comparative analysis of decision level fusion algorithms for 3D face recognition. *ICPR*, 2006.

[17] Congcong Li, Guangda Su, Yan Shang, Yingchun Li, and Yan Xiang. Face Recognition Based on Pose-Variant Image Synthesis and Multi-level Multi-feature Fusion. *AMFG 2007*, 261-275, 2007.

[18] Mohammad Sedaaghi, Constantine Kotropoulos, Dimitrios Ververidis. Improving speech emotion recognition using adaptive genetic algorithms. *Proc. EUSIPCO*, Polland, 2007.

[19] James E. Baker. Reducing Bias and Inefficiency in the Selection Algorithm. *Proceedings of the Second International Conference on Genetic Algorithms and their Application*, 14-21, 1987.

[20] P. Jonathon Phillips, Patrick J. Flynn, W. Todd Scruggs, Kevin W. Bowyer, Jin Chang, Kevin J. Hoffman, Joe Marques, Jaesik Min, William J. Worek. Overview of the Face Recognition Grand Challenge, *CVPR*, 947-954, 2005.

[21] Przemyslaw Szeptycki, Mohsen Ardabilian, and Liming Chen. A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking. *International Conference on Biometrics: Theory, Applications and Systems*, 2009.

[22] Boulbaba Ben Amor, Mohsen Ardabilian, Liming Chen. New experiments on ICP-based 3D face recognition and authentication. *ICPR*, 2006.

[23] Steven W. Zucker. Differential geometry from the frenet point of view: boundary detection, stereo, texture and color. *Handbook of Mathematical Models in Computer Vision*, N. Paragios, Y. Chen, and O. Faugeras, eds., Springer, 2005.

[24] Hiromi T. Tanaka, Masaki Ikeda and Hisako Chiaki. Curvature-based face surface recognition using spherical correlation. *IEEE International Conference on Automatic Face and Gesture Recognition*, 372–377, 1998.

[25] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face Recognition with Local Binary Patterns. *ECCV*, 2004.

[26] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger , Christopher von derMalsburg**.** Face recognition by elastic bunch graph matching. *IEEE Trans. on PAMI*, 1997.

[27] Ke Huang and Selin Aviyente. Sparse representation for signal classification. *International Conference on Neural Information Processing Systems*, 609-616, 2006.

[28] John Wright, Allen Y. Yang, Arvind Ganesh, S. Shankar Sastry, and Yi Ma. Robust face recognition via sparse representation. *IEEE Trans. on PAMI*, 2009.

[29] Zongbo Xie, Jiuchao Feng. KFCE: a dictionary generation algorithm for sparse representation. *Signal Processing*, 2009.