10th Symposium on Neural Network Applications in Electrical Engineering, NEUREL-2010
Faculty of Electrical Engineering, University of Belgrade, Serbia, September 23-25, 2010

A New Method for Human Face Recognition using Texture and Depth Information

Atefe Assadi and Alireza Behrad

***Abstract*—The efficiency of a human face recognition system depends on the capability of face recognition in presence of different changes in the appearance of face. One of the main difficulties regarding the face recognition systems is to recognize face in different views and poses. In this paper we propose a new algorithm which utilizes the combination of texture and depth information to overcome the problem of pose variation and illumination change for face recognition. In the proposed algorithm, we first use intensity image to extract efficient key features and find probable face matches in the face database using feature matching algorithm. We have defined some criteria to find the final match based on texture information or leave the decision to second stage. In the second stage the depth information are normalized and used for pose invariant face recognition. We tested the proposed algorithm using a face database with different poses and illumination and compared the results with those of other methods. We obtained the recognition rate of 88.96 percent which shows the considerable enhancement compared to previous methods.**

***Index Terms*—Face recognition, Feature extraction, Image Texture, Depth information**

I. Introduction

FACE recognition using image processing has been a challenging research topic during recent decades that includes different areas of research such as computer vision, artificial intelligence, pattern recognition and so forth. There are several challenging issues in this field such as changes in environment illumination, changes in age and recognizing face in different views and poses to name a few.

Face recognition algorithms are mainly divided into three groups: Feature-based approaches [1]-[4], holistic approaches (based on images) [5]-[9] and hybrid approaches [10],[11].

Feature based approaches extract special features from face images. These features are then used for face matching and face recognition. Different features are used for face recognition such as curvature based features [2], features based on a descriptive statistics of the image and its corresponding face components [3]. In the second group, whole face is considered as input data for face recognition. The techniques such as Eigenfaces [5], Fisherfaces [6] and support vector machine based method [7] are categorized in this group.

The third group is the combination of the first and second groups. This group of algorithms uses face images together with the local features for face recognition. The work of Cheng Zhong and his colleagues [11] which is called RLLGH[1] algorithm can be categorized in this group. In this algorithm the three-dimensional data are first mapped to two-dimensional images. Then features from middle and high frequency bands are extracted by applying Log-Gabor filter and the outputs of the filter are clustered using k-means algorithm. Then by comparing LVC[2] histogram of clustered patches using L1 norm, the face with least difference of histograms is considered as the output of system.

Another criterion to categorize the face recognition algorithm is the type of data that are used for face recognition. Some algorithms use only 2D images or texture information for recognition [3],[5]-[7],[10]. In [1],[2],[4] and [11] 3D face image or depth information is used for face recognition. As it is shown in [12], texture information is more efficient than depth information for face recognition. However texture information is more sensitive to illumination and poses variation and the recognition mostly fails in environment with illumination changes. Some algorithms utilize both depth and texture information to enhance the accuracy of face recognition algorithms [8],[9]. In [9] 3D and 2D face information is used for face recognition. However 3D information are only used for the estimation of face rotation. The rotation compensated 2D images are then used for face recognition.

In this paper we propose a new algorithm which utilizes the combination of texture and depth information to overcome the problem of pose variation and illumination change for face recognition. Our algorithm is a two-stage algorithm. We first use intensity image to extract efficient key features and find probable face matches in the face database using feature matching algorithm. We have defined some criteria to find the final match based on texture information or leave the decision to second stage. In the second stage the depth information are normalized and used for pose invariant face recognition.
We tested the algorithm with different images and results showed the efficiency of the proposed algorithm for face recognition in images with rotation, scale, pose and illumination changes.

The paper is organized as follows. In section 2 the face database is introduced, the proposed method is described in

Alireza Behrad is with the Faculty of Engineering, Electrical Engineering Department of Shahed University Tehran, Iran. (phone: (+98-21)51212036; e-mail: behrad@ shahed.ac.ir).
Atefe Assadi is Msc. Student of Faculty of Engineering, Shahed University Tehran, Iran (e-mail: aasadi@shahed.ac.ir).

[1] Robust Local Log-Gabor Histograms
[2] Learned Visual Codebook

978-1-4244-8819-3/10/$26.00 ©2010 IEEE

section 3, the experimental results are discussed in section 4 and conclusions appear in section 5.

II. Face Database

Currently one of the available perfect face databases is FRAV3D [13] that contain face images for 106 different persons in 16 different poses. The database contains face images with both texture and depth information in VRML files, however the size of texture images is almost 1.5 times of depth images. Fig. 1 shows different poses for a typical person in FRAV3D database. Pose number 1 to 4 are frontal view from the face with closed and open eyes. In pose numbers 5 to 8, the face has different rotation around Y axis. Pose number 9 and 10 has rotation around Z axis and pose number 13 and 14 has rotation around X axis. Pose 11 and 12 are smiling face and the face with open mouth respectively. Pose number 15 and 16 are frontal view from face with different lighting condition.

Fig.1. Different poses for a person in FRAV3D face database

III. Proposed Method for Face Recognition

We used the combination of two-dimensional (texture) information and three-dimensional (depth) information for face recognition. We first use texture information for recognition. We have defined some criteria to check the validity of face recognition using texture information. If the recognition criteria is satisfied the algorithm terminates otherwise we use depth information to recognize the input face. The proposed face recognition algorithm has the following steps:

1) Proper local features are extracted from input texture image and are compared with all local features that are extracted from texture face images in database.

2) Some predefined criteria are checked for matched local features. If the criteria are satisfied, we determine the recognized face; otherwise depending on the matched local features some face images are selected from database for recognition. The selected faces from database are called restricted face database.

3) The depth information is used to calculate and compensate for the 3D rotation matrix and translation vector between input image and each image in the restricted face database.

4) Using proper thresholding algorithm, unnecessary parts of depth images like hair, neck, dress and collar are removed from input image and the restricted face database.

5) Tip of nose is detected in depth images and used to compensate for scale change in depth images.

6) The correlation between the input image and restricted face database are calculated and the final recognized face is determined based on the correlation values.

In the following subsections we discuss about different steps of the proposed algorithm.

A. *Face Recognition using Texture Information*

1) *Feature Extraction and Matching using SIFT Descriptor*

The first step of the proposed algorithm uses texture information and local feature for face recognition. A good local feature should be detectable in different conditions and have less probability of mismatch. We used SIFT descriptor [14] which is a proper method to select key points in the case different scales and rotation. The SIFT algorithm detects point features in image and small patch around it depending on the image scale. To calculate SIFT descriptor, the patch is divided into 16 regions and the histogram of gradient vector is calculated in 8 different directions which results in a descriptor of size 16x8.

To make the descriptor more robust against the illumination change, we apply the histogram equalization algorithm to intensity image before SIFT feature extraction. Fig. 2 shows the results of applying histogram equalization on different poses of a subject

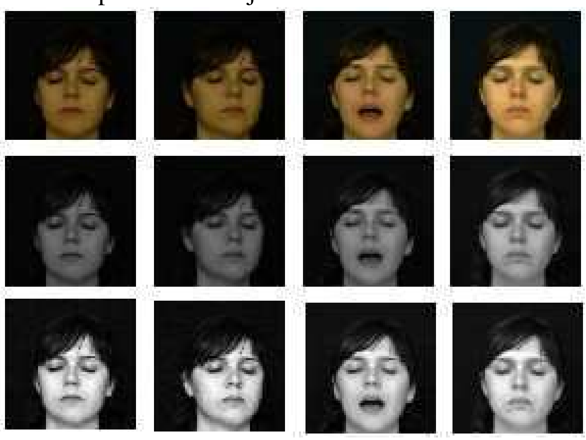

Fig. 2. The result of histogram equalization. Original images (top row), grayscale images (middle row) and images after histogram equalization (bottom row).

2) *Matching Criteria*

After histogram equalization the SIFT features are extracted from database image as well as input image. Then matching points for the feature points of input image are calculated in database images using SIFT feature vector. We used the number of matching points as a criterion for verifying the recognized face. For this purpose we select the two face images in the database with large number of matches and specify the number match with M_1 and M_2 respectively. The image with maximum number of matches is considered as recognized face if the following criteria are satisfied:

$$\begin{cases} M_1 > T_1 \\ M_1 - M_2 > T_2 \end{cases} \quad (1)$$

where T_1 and T_2 are two thresholds that are obtained experimentally. Fig. 3 shows input image and the recognized face using the SIFT features. If the criterion of Eq. 1 doesn't satisfy, we select some images in the database which have enough number of matches as a probable face matches and leave the decision for the next stage of the algorithm.

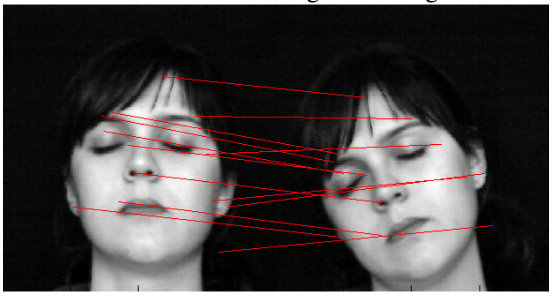

Fig. 3. Results of face recognition using SIFT feature extraction and matching

B. *Face Recognition using Depth Information*

The face images which are not recognized using texture information are transferred to this stage. As mentioned before, in this case, we select some face images in the database as a probable face matches which is called restricted face database (with T_1 threshold).

We simply used the normalized cross correlation algorithm to calculate similarity between input image and the restricted database images. However the normalized cross correlation is not a suitable algorithm for measuring similarity for face images with geometric misalignment. To cope with the problem of geometric misalignment, we estimate and compensate for geometric misalignment including the 3D rotation matrices and translation vectors between input image and the restricted database images.

We used ICP (Iterative Closest Point) technique [15] to calculate the 3D rotation matrix and translation vector between two face images. This technique is widely used for geometric alignment of three-dimensional models. The algorithm has three steps: at first, for each point in the first point set, the closest point in the second point set is located, then the rigid transformation that minimizes the mean square error across all matched points is computed and finally the transformation is applied and the error is recomputed. This process is iterated until convergence to a local minimum. In each iteration, the algorithm selects the closest points as correspondences and calculates the transformation (R,T), by minimizing the distance between point-correspondences, known as closest point.

Starting from the two sets of points, $P=\{p_i\}$, as a reference data of image, and as a test data of image. The least square criterion is defined as follow:

$$e(R,T) = \frac{1}{N}\sum_{i=0}^{N}\left\|(Rp_i + t) - y_i\right\| \quad (2)$$

where p_i and y_i are corresponding points. Many variants of ICP algorithms have been proposed to solve the minimization problem in the iterations. We used SVD based algorithm of Arun et al. [16], to calculate R and T in each iteration. In this method the minimization problem is represented using the following equations.

$$X' = \left\{ x_i - \frac{1}{N_x}\sum_{i=1}^{N_x} x_i \right\} = \{x_i'\} \quad (3)$$

$$P' = \left\{ p_i - \frac{1}{N_p}\sum_{i=1}^{N_p} p_i \right\} = \{p_i'\} \quad (4)$$

$$W = \sum_{i=1}^{N_p} x_i' p_i' \quad (5)$$

Then singular value decomposition of W is calculated as follow:

$$W = U\begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} V^T \quad (6)$$

If the rank of W matrix is 3, the optimal solution of *e(R,t)* is given by:

$$R = UV^T \quad (7)$$

$$t = \frac{1}{N_x}\sum_{i=1}^{N_x} x_i - R\frac{1}{N_p}\sum_{i=1}^{N_p} p_i \quad (8)$$

When the rotation matrix R and the vector T, are calculated they are used to compensate for geometric misalignment. The result of compensation is shown in Fig. 4.

Fig. 4. Database image (left) .Test image (middle). Test image after rotation compensation (right)

As it is obvious from Fig. 4, there may be some scaling mismatch between two face image after compensation for rotation and translation. Therefore it is necessary to use a normalization step to eliminate scale misalignment between images. For scale normalization we first remove background and unnecessary part of face like neck from face images. For this purpose we apply a threshold to remove area with less depth values. To determine the threshold value we used the method proposed by Otsu [17] which determines the threshold value by minimizing intra-class variances.

When the threshold value is calculated, it is used to remove unnecessary part of the face images and normalized them to the same size using the following algorithm:

1) Apply the threshold to remove area with less depth value like neck and hair.

2) Find nose tip as a reference point in the face area. The nose tip is a point with highest depth value. To overcome the noise sensitivity of the algorithm we used the average depth value in a 3x3 windows centered on image pixels.

3) Translate the nose tip to the center of face image.

4) Calculate the bounding box of face area and scale it to a predefined size.

Fig. 5 shows the result of nose tip detection and bounding box extraction. The results of applying threshold and normalization steps are demonstrated in Fig. 6.

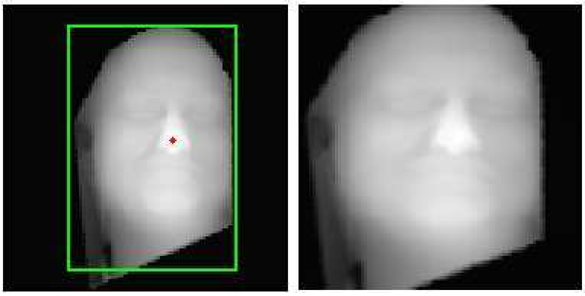

Fig. 5. Result of nose tip detection and bounding box extraction (Left). The output of normalization step (right)

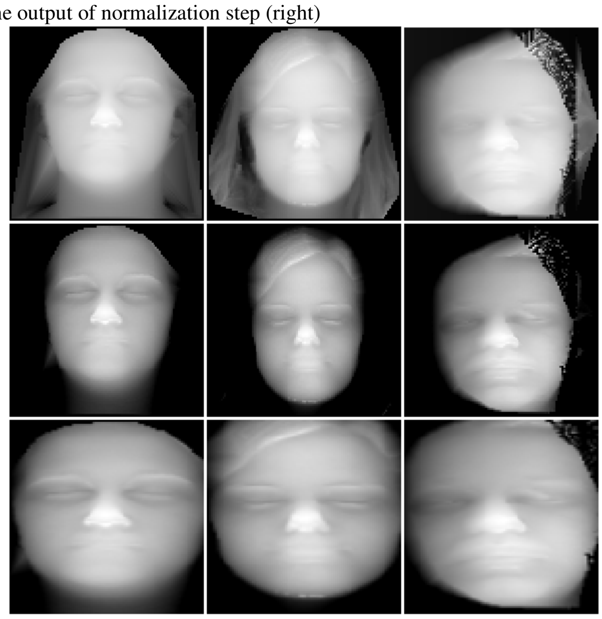

Fig. 6. Results of face image normalization. Original images (top row). Images after applying threshold (middle row). Images after normalization (bottom row)

After normalization stage the normalized cross correlation value between input face image and all images in the restricted database is calculated. The face image in the restricted face database with maximum correlation value is considered as recognized face image

IV. Experimental Results

The proposed algorithm implemented using a MATLAB program and tested using FRAV3D face database. As mentioned before the database contains face images for 106 different persons in 16 different poses. We used one of the poses (frontal view) as database and other 15 poses for tests. Therefore the database contains 106 face images, and test images contain 106x15=1590 samples.

We tested the proposed algorithm with different T_1 and T_2 values. Table I shows the recognition rate (RR) and false alarms (FA) and not recognized ratio (NRR) for the proposed algorithm using different T_1 and T_2 values. The definition of the evaluation parameters in Table I, are as follows:

$$RR = \frac{number\ of\ correctly\ recognized\ faces}{total\ number\ of\ test\ faces} \tag{9}$$

$$FA = \frac{number\ of\ falsy\ recognized\ faces}{total\ number\ of\ test\ images} \tag{10}$$

$$NRR = \frac{number\ of\ non\ recognized\ faces}{total\ number\ of\ test\ images} = 1 - FA - RR \tag{11}$$

TABLE I

Recognition Results for the Proposed Algorithm

$T_1 = 1$						
T_2	2D data			2D+3D data		
	RR	FA	NDR	RR	FA	NDR
0	0.8453	0.0844	0.0703	0.8687	0.0875	0.0438
1	0.8172	0.0312	0.1516	0.8523	0.0359	0.1118
2	0.7891	0.0078	0.2031	0.8521	0.0162	0.1317
3	0.7672	0.0047	0.2281	0.8432	0.0148	0.142
4	0.7438	0.0031	0.2531	0.8419	0.0162	0.1419
6	0.7016	0.0016	0.2968	0.8349	0.0194	0.1457
8	0.6672	0	0.3328	0.8286	0.0215	0.1499
$T_1 = 4$						
T_2	RR	FA	NDR	RR	FA	NDR
0	0.8438	0.0766	0.0796	0.8698	0.0801	0.0501
1	0.8156	0.0312	0.1532	0.8533	0.0362	0.1105
2	0.7891	0.0078	0.2031	0.8526	0.0163	0.1311
3	0.7672	0.0047	0.2281	0.8442	0.015	0.1408
4	0.7438	0.0031	0.2531	0.8428	0.0163	0.1409
6	0.7016	0.0016	0.2968	0.8356	0.0195	0.1449
8	0.6672	0	0.3328	0.8291	0.0216	0.1493
$T_1 = 5$						
T_2	RR	FA	ND	RR	FA	ND
0	0.8359	0.0703	0.0938	0.8708	0.075	0.0542
1	0.8156	0.0297	0.1547	0.8543	0.0349	0.1108
2	0.7891	0.0078	0.2031	0.8553	0.0166	0.1281
3	0.7672	0.0047	0.2281	0.8449	0.0151	0.14
4	0.7438	0.0031	0.2531	0.8431	0.0163	0.1406
6	0.7016	0.0016	0.2968	0.8355	0.0195	0.145
8	0.6672	0	0.3328	0.8297	0.0217	0.1486
$T_1 = 6$						
T_2	RR	FA	NRR	RR	FA	NRR
0	0.8312	0.0531	0.1157	0.8711	0.0584	0.0705
1	0.8125	0.0266	0.1609	0.8545	0.0322	0.1133
2	0.7891	0.0078	0.2031	0.8555	0.0167	0.1278
3	0.7672	0.0047	0.2281	0.8451	0.0151	0.1398
4	0.7438	0.0031	0.2531	0.8433	0.0164	0.1403
6	0.7016	0.0016	0.2968	0.8356	0.0195	0.1449
8	0.6672	0	0.3328	0.8299	0.0217	0.1484
$T_1 = 8$						
T_2	RR	FA	NRR	RR	FA	NRR
0	0.8047	0.0266	0.1687	0.8798	0.0366	0.0836
1	0.7938	0.0172	0.189	0.8625	0.0264	0.1111
2	0.7812	0.0047	0.2141	0.8637	0.0157	0.1206
3	0.7656	0.0031	0.2313	0.8529	0.0147	0.1324
6	0.7016	0.0016	0.2968	0.8413	0.0202	0.1385
8	0.6672	0	0.3328	0.8344	0.0223	0.1433
$T_1 = 10$						
T_2	RR	FA	NRR	RR	FA	NRR
0	0.7719	0.0141	0.214	0.8818	0.0147	0.0894
1	0.7703	0.0109	0.2188	0.8626	0.0123	0.1142
2	0.7672	0.0047	0.2281	0.8638	0.0129	0.1186
3	0.7547	0.0031	0.2422	0.8529	0.0131	0.1309
6	0.7016	0.0016	0.2968	0.8419	0.0187	0.1378
8	0.6672	0	0.3328	0.8350	0.0224	0.1426
$T_1 = 12$						
T_2	RR	FA	NRR	RR	FA	NRR
0	0.7516	0.0047	0.2437	0.8896	0.0184	0.0873
3	0.7438	0.0031	0.2531	0.8703	0.0169	0.1097
6	0.7	0.0016	0.2984	0.8511	0.0201	0.1272
8	0.6656	0	0.3344	0.8402	0.0233	0.1365

As it is shown in Table I, the best results for the proposed algorithm is obtained for T_1=12 and T_1-T_2=0. To compare the results of the proposed algorithm with those of other methods we also implemented the face recognition algorithm using 2DPCA [4] and RLLGH [11] algorithms. Fig. 7 shows the recognition rate using FRAV3D database for different implemented algorithms. For comparison, the results of recognition using proposed algorithm by using only texture information is also shown in this figure. As it is shown in this figure the results of the proposed algorithm is promising. As it is shown in this figure the use of 3D information enhanced the RR about 5 percent, however as it is obvious in Table I the FA is also decreased about 7 percent.

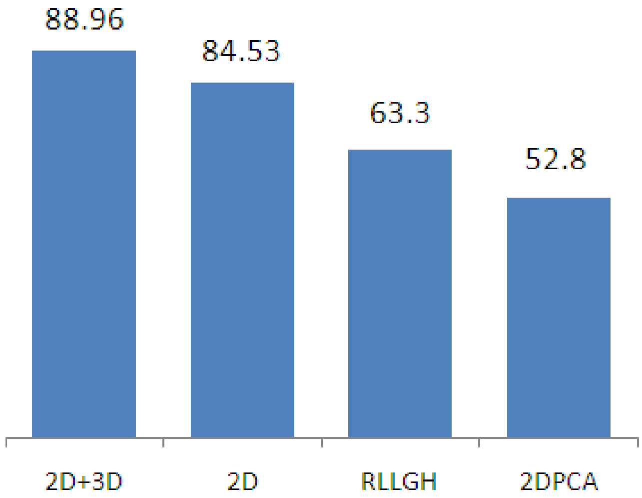

Fig. 7. Recognition rate for different face recognition algorithms. From left to right: 1- Proposed algorithm, 2 Proposed algorithm using only texture data, 3- RLLGH[11] and 4- 2DPCA[4]

V. Conclusion

In this paper we proposed a new algorithm for face recognition. We used both texture and depth information for recognition. The proposed algorithm is a two-stage algorithm. In first stage the texture information is used for recognition and the second stage uses depth information. We tested the proposed algorithm with FRAV3D database and result showed that the proposed algorithm has promising results.

References

[1] G. Gunlu. and H.S. Bilge, "Feature extraction and discriminating feature selection for 3D face recognition" , *in Proc. 24th International Symposium on Computer and Information Sciences*, pp. 44 – 49, Sept 2009.

[2] Mohammad H. Mahoor and Mohamed Abdel-Mottaleb,"3D Face Recognition Based On 3d Ridge Lines In Range Data" *in Proc. IEEE International Conference on Image Processing(ICIP 2007)*, pp. IV-509 -IV-512, 2007.

[3] R. Kam-art, T. Raicharoen and V. Khera," Face recognition using feature extraction based on descriptive statistics of a face image", *IEEE Machine Learning and Cybernetics, International Conference on*,vol.1, pp:193 – 197, 2009.

[4] Amr Rashad and Alaa Hamdy, "3D face recognition using 2DPCA ," *IJCSNS International Journal of Computer Science and Network Security*, vol. 9 no.12, December 2009.

[5] Y.L.Yu Zhujie ,"Face recognition With Eigenfaces," *IEEE International Conference on Industrial Technology* , pp. 434 – 438, Dec. 1994.

[6] P. Belhumeur, J. Hespanha and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.

[7] M. Savvides, R. J. Heo, S. Park, C. Xie, and B.V.K. Vijayakumar," Partial & Holistic Face Recognition on FRGC-II data using Support Vector Machine," *in Proc. Computer Vision and Pattern Recognition Workshop (CVPRW '06)*, pp.48-48, June 2006.

[8] Minghua Zhao, Zhisheng You, Yonggang Zhao and Zhifang Liu. "Face Recognition Method Combining 3D Face Model with 2D Recognition," *in Proc. Fourth International Conference on Image and Graphics (ICIG2007)*, pp.655-660, 2007.

[9] A. Bossaghzade, M. Doostari and A. Behrad, " Face recognition using non frontal face image", *in Proc. 17th Iranian Conference Electrical Engineering*, pp. 196-201,2009. Published in Farsi.

[10] Chen Pan and Feilong Cao, "Face Image Recognition Combining Holistic and Local Features", *in Proc. The 6th International Symposium on Neural Networks*, pp. 407- 415, 2009.

[11] Cheng Zhong, Zhenan Sun, Tieniu Tan and Zhaofeng He,"Robust 3D Face Recognition in Uncontrolled Environments", *in Proc. IEEE Conference on Computer Vision and Pattern Recognition.(CVPR 2008)*, pp.1-8, 2008.

[12] Andrea F. Abate, Michele Nappi, Daniel Riccio and Gabriele Sabatino, "2D and 3D face recognition: A survey" *Pattern Recognition Letters*, vol. 28, Issue 14, pp.1885-1906, 2007.

[13] http://www.frav.es/research/facerecognition/FRAV3D/

[14] David G. Lowe, "Distinctive Image Feature from Scale Invariant Features," *International Journal of computer Vition*, vol. 60, no. 2, pp.91-110, 2004.

[15] P. J. Besl and N. D. McKay, "A method of registration of 3-D shapes," *IEEE Trnsactions on Pattern Analysis and Machine Intelligence*, vol.14, no.14, pp. 239–256, Feb 1992.

[16] K. S. Arun, T.S. Huang and S.D. Blostein, "Least-Squares Fitting of Two 3-D Point Sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no.5, pp. 698-700, September 1987.

[17] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Transactions on systems Man and Cybernetics*, vol.8, no.1, pp.62-67, 1979.