Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 46 (2015) 1778 – 1785
Procedia Computer Science

International Conference on Information and Communication Technologies (ICICT 2014)

Face Detection, Registration and Feature Localization Experiments with RGB-D Face database

Bodhi S.R.[a,*], Naveen S.[b]

[a]PG Scholar, Department of ECE, LBSITW, Poojappura, Trivandrum-695012, India
[b]Assistant Professor, Department of ECE, LBSITW, Poojappura, Trivandrum-695012, India

Abstract

RGB-D Face database is a multimodal database based on Kinect. This paper focuses on pre-processing techniques for depth data that involve face detection, registration and facial feature localization which are very critical for face authentication and identification systems. Faces are detected in the range data, assisted by detection of face on corresponding texture data using Viola-Jones detector. Validated faces using a depth PCA classifier are registered using conventional ICP algorithm. Finally Nose tip and eye corners are localized on registered faces.

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

Keywords: Viola-Jones detector;2.5D data; ICP;PCA face/non-face classifier;curvature analysis

1. Introduction

Most of the algorithms using facial data were based on using two dimensional data of face or simply face images. Images can be affected by changeable factors like illumination, pose and expression etc. These problems associated with 2D systems can be avoided by the use of depth data. It is possible to incorporate depth information along with two dimensional data or through purely three dimensional approaches. Both the multimodal and 3D systems require proper pre-processing of depth data can improve the system performance and necessary steps to be included in pre-

* Corresponding author. Tel.:+0-808-965-6130
*E-mail address:*sreebodhi@gmail.com

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)
doi:10.1016/j.procs.2015.02.132

processing of facial data depends on nature of data and type of application. Some of the important steps include detection of faces in complicated scenes, registration of facial data to a reference model and extraction of relevant facial features from registered data that can improve the performance of the system. Development of low cost, easily available depth sensors like Kinect made revolutionary change in 3D data based works, as it makes real time depth data acquisition easier. Face detection in the input data can be considered as the first stage in any practical face analysis system. In literature, 3D data is found to be used in different ways for face detection. Some researchers used depth data to improve the performance and reliability of 2D face detection system. Robotic vision can be improved by using depth cues[1,2]. Face detection literature is incomplete without mentioning Viola and Jones. They proposed a real time object detection frame work that can be used for solving face detection problem[3]. It is based on extraction of Haar like features from face and cascading several week classifiers to form a strong classifier. Adaboost algorithm is proposed for selecting relevant features or weak classifiers. They also made use of integral images to make algorithm faster. Even though training is slow, better detection rate can be achieved. Implementation is available in Open CV library. Most practical system follows their approach. Even many recent works are based on viola-Jones detection algorithm. Basic Viola-Jones algorithm is adapted in[4] to work with range images acquired using TOF camera. Researchers state that better detection rate could be achieved using both 2D and 3D data than training detectors using either data individually. In[5], researchers tried to improve the basic Viola-Jones detection method and present a real time scale invariant 3D face detection algorithm by using an orthogonal projection technique for obtaining range images. A robust real time face detection framework based on cascading two level classifiers is proposed in[6] in which windows detected as face by first level 2D classifier is verified by second level cascaded classifiers. A heuristic classifier combined with LDA classifier constitute second level which uses depth data generated using stereo camera and finally reject the non face regions so that considerable change in false positive rate is achieved. LDA classifier is trained with a number of face and non face samples. An efficient algorithm for registering 3D shapes is proposed by Besl and McKay[7]. It is a general purpose, representation independent method for accurate and computationally efficient registration of 3D shapes. Iterative Closest Point (ICP) algorithm is applicable for 3D face registration and many researchers follow this technique for face registration. Different modified ICP algorithms are also proposed later by incorporating modifications on different stages of basic ICP algorithm. Classification and comparison of different ICP variants are discussed in[8]. In some works ICP is included as a fine registration approach after performing a coarse registration using some other methods. Gordan explains how face can be represented by features based on shape and curvature of surface in[9]. Many researchers got inspired by this work and focus on the curvature analysis for facial feature extraction from the range images of face. In[10], a face detection system based on curvature analysis on range images is proposed. Most of the publically available 3D data bases are made especially for recognition purpose. As faces are already detected automatically or manually, face detection is not that much challenging if deals with such databases. Only face localization is to be performed prior to face recognition. So[11] presents a public database for research purposes where faces are not detected in both color and depth images.

In this paper, frame work for detecting 3D faces from the RGB-D Face data samples is explained in section 2. Section 3 explains how detected faces can be registered efficiently using ICP algorithm and facial feature localization is given in section 4. Results obtained on experiments with RGB-D database are discussed in section 5.

2. 3D Face detection assisted by 2D

Face detection is an important step to be done in practical systems especially when the input depth data is acquired from complex scenes. Here RGB-D face database is used which consists of 2D and 3D data of scenes where face detection and localization is challenging as compared to other publically available databases which mostly contain face data only. Simultaneous capturing of texture and depth data by Kinect enables 3D face detection assisted by 2D face detector. Proposed approach uses Viola-Jones detector which effectively detects faces in 2D scene. Corresponding candidate regions are selected in the depth image or range data. In RGB-D face database, texture and range data were simultaneously captured with 1280x960 and 640x480 resolutions respectively. So to establish correspondence between the two modalities, texture data has to be first down-sampled to the size of range data. Still correspondence problem exists since position of same element in the depth data has a shift towards left compared to its corresponding texture data. It may be due to the difference in the position of sensors in the Kinect. So left shifted bounding boxes are considered in the depth data corresponding to the bounding boxes detected by Viola-Jones detector in the texture data. Fig. 1 shows detected faces in some samples from RGB-D face database. One of the drawbacks of Viola-Jones detector is high false positive rate even though detection rate is better. All the bounding boxes given by detector on texture data need not be faces. Corresponding candidate regions selected in the range data include some non face regions also as shown in last column of Fig. 1. A depth PCA face or non-face classifier can be cascaded to reject such false faces. Depth PCA classifier makes decision based on the reconstruction error of faces. A set of 2.5D faces (range data) are used to form the reduced subspace of original face space, called as eigen space. Basis vectors of subspace are obtained from original face vectors. These are the eigen vectors corresponding to the largest eigen values of covariance matrix of face vectors. Test sample is projected into the sub space and it is re-projected in the original face space. Reconstruction error, the error between reconstructed face and original face is considered as the measure of “faceness”. Each candidate region obtained in the range data of scene with the help of viola- Jones detector is reconstructed as described above and classified as actual face if the reconstruction error is below certain threshold.

Fig. 1. 3D face detection assisted by 2D face detector. Top row shows bounding boxes detected by Viola-Jones detector in texture data and bounding boxes selected in corresponding range data is shown in bottom row.

3. 3D Face registration

3D face registration is a key step to improve the performance of face identification and authentication systems. Registration is done to align sample faces to a reference face. In real time systems, the subjects need not be cooperative with data acquisition. So input faces can be of different poses. Registration brings all faces to a common coordinate system. Reference face is usually selected as a face with frontal pose. Reference face selection is also critical as it can make considerable change in the performance of the registration algorithms. The pre-processing stage for face analysis systems should consider the pose correction stage to ensure better performance of the system as well as to make the system adaptable for real time applications. ICP is an effective algorithm for registering 3D shapes. The goal of the ICP algorithm is to find registration parameters such that the error between the sample face corrected using those parameters and reference face is minimum. It is an iterative procedure. After a number of iterations the sample face gets aligned with the reference face. During each iteration closest point estimation, error minimization and updating sample face are done. Matching process is the large time consuming part which involves finding out closest points on the reference shape for each point on the sample face. After corresponding points are estimated, transformation parameters are obtained such that the mean square error between the corresponding points of sample and reference faces is minimized. The sample data is getting updated using those parameters and it will be the sample data for next iteration.

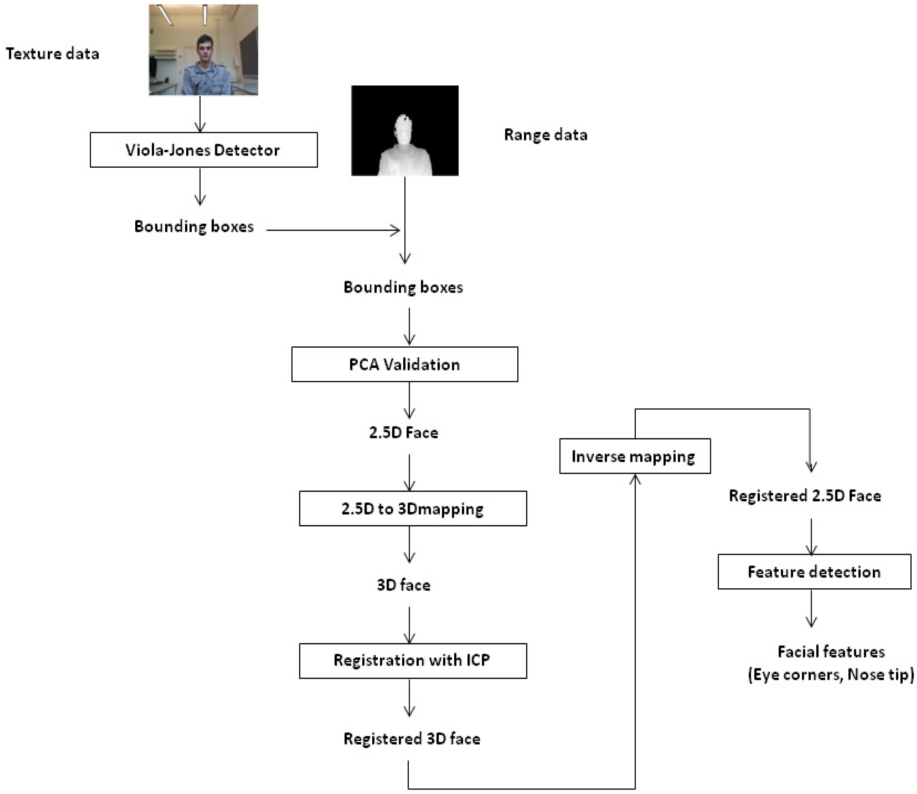

Fig. 2. Flowchart of proposed algorithm

Mathematically ICP can be explained as follows

Let $X = \{\vec{x}_i\}$ be the cloud data of reference face and $P = \{\vec{p}_i\}$ be the sample data to be corrected.

Let the distance between a vector $\vec{p}$ and X can be obtained as

$$d(\vec{p}, X) = \min_{i \in \{1,2,\ldots,N_X\}} d(\vec{p}, \vec{x}_i) \tag{1}$$

Where,

$$d(\vec{p}, \vec{x}_i) = || \vec{P} - \vec{x}_i || \tag{2}$$

Then closest point or corresponding point $\vec{x}_j$ of $\vec{p}$ on X satisfies the condition

$$d(\vec{p}, X) = d(\vec{p}, \vec{x}_j) \tag{3}$$

Once the corresponding points on the reference face for entire points on the input sample are estimated as described above, registration vector is to be obtained such that the mean square objective function is minimized and is defined as

$$f(\vec{q}) = \frac{1}{N_p} \sum_{i=1}^{N_p} || \vec{x}_i - R(\vec{q}_R)\, \vec{p}_i - \vec{q}_T ||^2 \tag{4}$$

Where $\vec{q}_R$ and $\vec{q}_T$ are part of registration vector corresponding to rotation and translation. $R(\vec{q}_R)$, Rotation matrix corresponding to $\vec{q}_R$. Updated sample data is given by

$$P' = \{\vec{p}_i'\} \tag{5}$$

$$\vec{p}_i' = R(\vec{q}_R)\,\vec{p}_i + \vec{q}_T \tag{6}$$

These steps are repeated in each iteration. Sample data converges to the reference face after a number of iterations. Computational complexity of registration algorithm can be reduced by limiting the size of input cloud data using a down sampling filter.

4. Facial feature Localization

Curvature analysis is an effective tool for detection of eye corners which are regions of higher curvature. Curvature analysis isolates regions of higher curvature like eye corners. Forehead, jaw lines and nose are also having high curvature values. So locating eye corners alone is not an easy task. Here search for eye corners is done according to the constrain that eye corners are always above nose region. So preliminary search is done for nose tip location. Then region of interest to search for eye corners is selected with reference to the nose tip.

4.1. Nose tip localization

Nose tip is characterized by its highest depth value compared to other facial components. It can be simply detected on 2.5D faces using conventional method. A window is used to search location having maximum intensity value in 2.5D representation of face. Accuracy can be improved by considering intensity of neighbouring values.

4.2. Eye Corner localization

Curvature analysis is done on range image of face and involves 3 main steps- range image smoothing, curvature map generation and curvature thresholding. Gaussian smoothing is done on the range data of face first. In the second step two kind of curvature maps Gaussian (H) and Mean (K) curvature maps are generated which can give curvature values of facial components.

If the surface S be represented by twice differentiable real valued function $f : U \to R$, defined on an open set $U \subseteq R^2$:

$$S = \{(x, y, z) \mid (x, y) \in U; z \in R; f(x, y) = z\} \tag{7}$$

Then two curvature values can be obtained using first and second order derivatives of *f* for every point $(x, y, f(x, y))$ using the equations

$$H(x,y) = \frac{(1+f_y^2)f_{xx} - 2f_xf_yf_{xy} + (1+f_x^2)f_{yy}}{2(1+f_x^2+f_y^2)^{\frac{3}{2}}} \tag{8}$$

$$K(x,y) = \frac{f_{xx}f_{yy} - f_{xy}^2}{(1+f_x^2+f_y^2)^2} \tag{9}$$

Finally desired high curvature regions are isolated by selecting proper thresholds T_h and T_k for H and K maps respectively.

$$H(x,y) \geq T_h$$
$$K(x,y) \geq T_k \quad (10)$$

5. Experimental results

All experiments were performed with RGB-D face database on MATLAB 2013a. Face detection is done with the cascade object detector of computer vision toolbox based on the Viola-Jones algorithm. A multimodal approach is followed as described in section 2 to isolate face data alone from the background. False 2.5D faces detected are eliminated with the help of PCA based face/ non-face classifier thus improving reliability of detection frame work. PCA face or non-face classifier makes decision based on the reconstruction error. An optimum threshold value for reconstruction error that discriminates a candidate region into face or non-face is to be obtained. An experiment is conducted with a set of face and non-face samples to analyze accuracy of PCA classifier for face or non-face classification as well as to discuss how the threshold value for two class separation problem is selected. Results are added in table 1.

Error distribution of faces and non faces are separately plotted as shown in Fig. 3. Optimum threshold value is determined based on these two distributions. From the histograms, it is clear that error is below 40 for most of face samples and greater than 40 for most of non faces. Hence choosing a threshold around the value of 40 can give better separation between faces and non-faces in this particular experiment.

Table 1. PCA face/ non-face classifier results

	Samples	Detected as face	Detected as non-face	Detection Rate	Error Rate
Face	125	113	12	90.4	9.6
Non-face	100	91	9	91	9

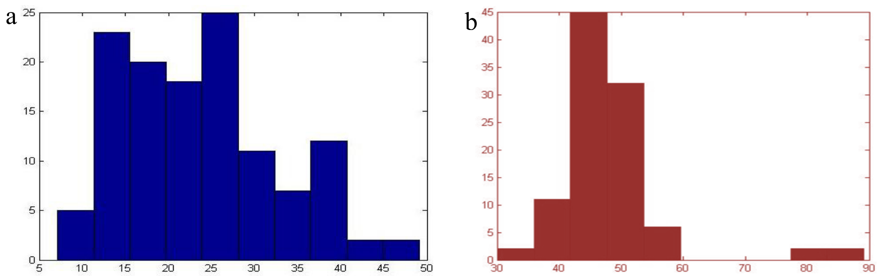

Fig. 3. (a) Error distribution of faces; (b) Error distribution of non-faces

Effect of cascading PCA classifier is analyzed with another experiment considering 150 samples from RGB-D face database. Experimental results in table 2 shows the effect of cascading face/ non face classifier to eliminate false faces. In this particular experiment all non faces could be eliminated by the second classifier through the selection of proper threshold. Thus finally true 2.5D faces can be obtained. Sometimes 2.5D to 3D mapping is required especially if registration of detected faces is involved as discussed in the pre-processing framework. Stair case effect is one of the artefacts associated with 3D surfaces reconstructed from 2.5D data of RGB-D face database as reported in [5]. Reconstruction from 2.5D face, smoothed using wiener filtering is found to be resolving this problem to some extent. Effect of wiener filtering is shown in Fig. 4.

Table 2. Face detection and validation results

Total no. of input samples	150
Total bounding boxes returned by Viola-Jones detector	159
Detected non faces (False Positives)	9
Validated faces by PCA classifier	150
Rejected non faces by PCA classifier	9

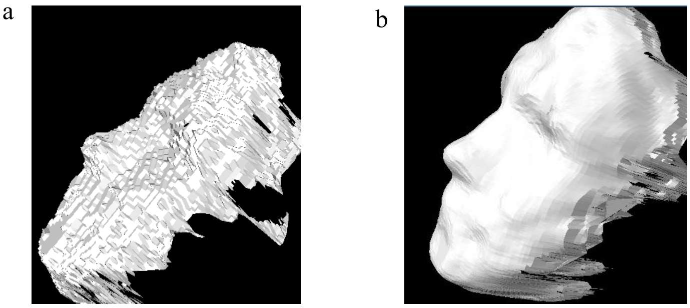

Fig. 4. Face surfaces reconstructed from 2.5D face (a) not filtered; (b) filtered using wiener filter

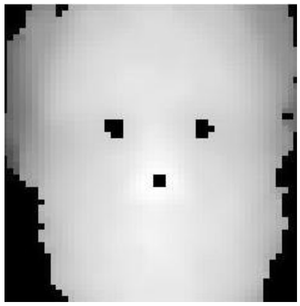

Fig. 5. Automatically localized facial features

A database including more pose variations is artificially created from 3D faces with frontal pose for registration experiment. Suppose P_o is original point cloud representation of face with frontal pose, the transformed face representation P will be

$$P = R * P_o \quad (11)$$

Where, R being rotation matrix corresponding to required rotation about a particular axis. Usually rotations about X and Y axes are considered as these are the mostly occurring cases in practical systems. 3D face data for rotation of different angles about all 3 axes are considered here. Further experiments are done with this database.

All 3D faces registered to frontal pose using a variant of ICP algorithm. Down sampling operation is done on the input 3D faces to ICP so that size of data and thus computational complexity can be reduced. Face data with size17061x3 is reduced to 2133x3 when down sampled by a factor of 8. Detection of nose tips and eye corners are done on range data using algorithms discussed in section IV. Automatically detected nose tip and eye corners on a 2.5D face from RGB-D face database is shown in Fig. 5. Performance of feature detection algorithms is subjectively evaluated. Observation results on registration and feature detection are added in table 3.

Table 3. Registration and feature localization results

No. of test samples	Registration		Nose tip		Eye corners	
	No. of correctly registered	*Registration rate*	*No. of correctly detected*	*Detection rate*	*No. of correctly detected*	*Detection rate*
624	620	99.35	576	92.31	566	90.71

6. Conclusion

In this paper, face detection, registration and facial feature localization on depth data from RGB-D face database are discussed. Detection of face data alone from a complicated scene is challenging task. Widely used Viola-Jones 2D detector is used to assist 3D detection and better face detection performance is achieved. By cascading depth PCA classifier for validation, reliability of detection frame work can be improved. Well known ICP algorithm is found to be effectively registering almost all faces even if input faces to ICP are size limited by down sampling. Facial feature extraction is also relevant for some systems. Experimental results show that consistent performance in all stages of proposed frame work can be achieved even though depth data from RGB-D face is of low quality compared to other databases acquired using 3D scanners.

References

1. W. Burgin, C. Pantofaru, W. D. Smart. Using depth information to improve face detection. in Proceedings of the 6th international conference on Human-robot interaction, USA, 2011.
2. Vo Duc My, Andreas Masselli, Andreas Zell. Real Time Face Detection using Geometric Constraints, Navigation and Depth-based Skin Segmentation on Mobile Robots. IEEE 2012.
3. P. Viola, and M. Jones. Rapid object detection using a boosted cascade of simple features. in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 511-518, 2001.
4. Böhme M., Haker M., Riemer K., Martinetz T., Barth E. Face detection using a time-of-flight camera. Lecture Notes in Computer Science vol. 5742, 167–176.
5. Pamplona Segundo M., et al. Orthogonal projection images for 3D face detection. Pattern Recognition Lett. (2013)
6. Mahmood Rahat, Masoom Nazari, Akram Bafandehkar, Saeed Shiry Ghidary. Improving 2D Boosted Classifiers Using Depth LDA Classifier for Robust Face Detection. IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 2, May 2012
7. P. Besel, N. Mckay. A method for registration of 3D-shapes. IEEE trans. Pattern analysis and Machine intelligence, 14(2):239-256, 1992.
8. Szymon Rusinkiewicz,Marc Levoy. Efficient Variants of the ICP Algorithm. Third International Conference on 3D Digital Imaging and Modeling(3DIM),June 2001
9. G.G. Gordon. Face recognition based on depth maps and surface curvature. SPIE Geom. Methods Comput. Vision 1570 (1991)234–274.
10. A. Colombo, C. Cusano, and R. Schettini. 3D face detection using curvature analysis. Pattern Recognition, vol. 39, no. 3,2006.
11. Hg,R., Jasek P., Rofidal, C., Nasrollahi, K., Moeslund, T., Tranchet, G. An RGB-D database using Microsoft's Kinect for windows for face detection. In: Proceedings of the Eighth International Conference on Signal Image Technology and Internet Based Systems, pp. 42–46.
12. P. Bagchi, D. Bhattacharjee, M. Nasipuri, D. K. Basu. A novel approach for registration of 3D face images. proc of Int. Conf on Advances in Engineering, Science and Management(ICAESM), March 30-31, 2012
13. http://www.vap.aau.dk/rgb-d-face-database