IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 4, No 4, July 2010
ISSN (Online): 1694-0784
ISSN (Print): 1694-0814

Nose Tip Region Detection in 3D Facial Model across Large Pose Variation and Facial Expression

Anuar L.H., Mashohor S., Mokhtar M. and Wan Adnan W.A.

Multimedia Systems Laboratory, Department of Computer and Communication Systems, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia

Abstract

Detecting nose tip location has become an important task in face analysis. However, for a 3D face model with presence of large rotation variation, detecting nose tip location is certainly a challenging task. In this paper, we propose a method to detect nose tip region in large rotation variation based on the geometrical shape of a nose. Nose region has always been considered as the most protuberant part of a face. Based on convex points of face surface, we use morphological approach to obtain nose tip region candidates consist of highest point density. For each point of each region candidate, a signature is generated and evaluated with trained nose tip tolerance band for matching purpose. The region that contains the point which scores the most is chosen as the final nose tip region. This method can handle large rotation variation, facial expression, combination of all rotations (yaw, pitch and roll) and large non-facial outliers. Combination of two databases has been used; UPMFace and GavabDB as training data set and test data set. The experimental results show that 95.19% nose tip region over 1300 3D face models were correctly detected.

Key words: *Nose Tip Region Detection, Morphology, 3D Face Model, Point Signature, Tolerance Band.*

1. Introduction

Human face recognition has received wide attention for the past few years. Recent development of technology in data acquisition has enabled a 3D face to be captured and analyzed. Extensive studies have shown that 3D face data offers better performance of face recognition in disadvantageous conditions, such as head pose rotation, illumination and facial expression. Methods associated with 3D face recognition such as face registration, face modeling and facial features extraction that have been proposed are largely based on the geometry shape of a face which greatly relies on appropriate 3D surface descriptors and accurate facial landmark locations [1].

Compared to other facial landmarks, nose offers few advantages. Due to the distinct shape and symmetrical property of a nose, it is frequently used as a key feature point in 3D faces representation. For example, finding nose facilitates the search for other landmarks such as eyes and mouth corners in order to employ robust facial feature extraction [2]. Unlike nose, other features can change significantly due to facial expression, e.g., closed eyes and open mouth. In addition, the characteristics of the nose which indicates the center of the face and always pointing frontal are found useful for head pose estimation and face registration.

Even though geometry feature-based method seems to be more straight-forward in representing the whole face, it highly depends on the geometrical characteristics of the face. Any changes of pose or a presence of occlusion can cause severe lost of information. This explained by Ayyagari [3] that during acquisition process, it is possible that some parts of the face will be unobservable from any given position, either due to occlusion, or limitations in the sensor's field of view. Heuristically, nose is always assumed as the nearest point to the camera, and therefore the highest value in z-axis. Although it can largely reduce the complexity of an algorithm, in case of large scale of variation and rotation, this assumption does not always hold.

Alternatively, a search for nose tip can be based on the protuberant parts of a face. Although it cannot give exact location of nose tip, it can reduce the searching space quite effectively. This is followed by one or more other steps to finalize a nose tip. In [4], nose tip candidates were first obtained based on protuberant points. Since only frontal faces were considered, nose tips were obtained by symmetry calculation, which was carried out based on the direction comparison of normal vectors. Similar idea was adopted by Xu [5], who considered nose tip as the highest local point and having peaked cap-like shape. Nose tip were finalized using Support Vector Machine (SMV) to classify between nose-tip and non-nose-tip points.

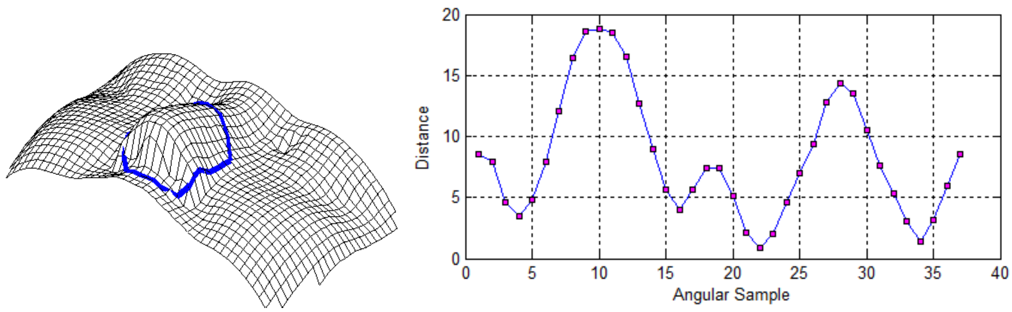


Fig. 1 Orientation shape of a nose tip.

Sun and Yin [6] took the advantage of the symmetrical property of a nose. Based on curvature of facial range image, two clusters of inner eye corners were obtained. With aid of facial reference plane, nose tip can be determined by putting a plane along eyes line and find the point that has the maximum distance from plane. However, occlusion around the eyes, for example glasses or hair, can affect the robustness of this method. Breitenstein [7] introduced a scheme of nose detection that is robust to large pose variation. Shape signature was computed for every point of input image and the corresponding pose hypotheses were generated in parallel. To select nose tip, error function is used to compare the input range image to the pre-computed pose images of an average face model.

In [8], Lu and Jain proposed a feature extractor based on the directional maximum to estimate nose tip location and pose angle simultaneously. At each quantized pose angle in the original coordinate system, the point with the maximum projection value along corresponding pose direction is selected as the nose tip candidate and therefore the directional maximum of that pose angle. A nose profile model represented by subspace is used to select the nose tip and the associated pose angle as the pose estimation result. Same technique was employed later in [9] to find nose tip candidates. However, a nose tip is finalized differently, where the nose check curve of each candidate is extracted and compared with that of trained nose tip check curve for similarity.

Despite of having significantly different nose structure from one person to another, the orientation shape around nose tip is however preserved, even with the presence of rotation. This shape is considered unique, if measured by appropriate radius can be represented by four peaks of nose ridge, nose wings (left and right) and the area between nose and lips which referred as above lips as shown in Fig. 1. Based on this observation, we proposed a new method to locate nose tip region of 3D face model in large variation rotation. Region candidates are gained by studying the geometrical features of a face followed by morphological operations. Nose tip region then finalized using point signatures matching.

The remainder of this paper is organized as follows. Section 2 describes the framework of the proposed method. The performance of the proposed method is presented and discussed in section 3. Finally, we give our conclusion in section 4.

2. Nose Tip Region Detection

This section describes the nose tip region algorithm in four stages as briefly illustrated by Fig. 2.

2.1 Convex point classification

In our proposed method framework, point signature is the key step in verifying the nose tip. However, it is computationally expensive to calculate signature for every point on a face. Therefore, it is always practical to perform elimination steps, giving fewer number of nose tip region candidate. Here, we consider a nose tip as a region rather than one exact point and any part of a face is treated as potential nose tip region. From geometry point of view, a face can be broadly represented as convex, concave or flat points to describe the curviness of a face. Inspired by this, a searching of nose tip candidate can be done by learning the convexity of a face. This method will not give exact location of the nose tip region, but it will effectively reduce the searching space and computational time.

For each point P_0, it can be identified as convex point or others by making use of the known dot product of neighboring unit vectors, P_i and its surface normal, N_p:

$$\cos\theta_i = \overrightarrow{P_iP_0} - \overrightarrow{N_pP_0} \qquad (1)$$

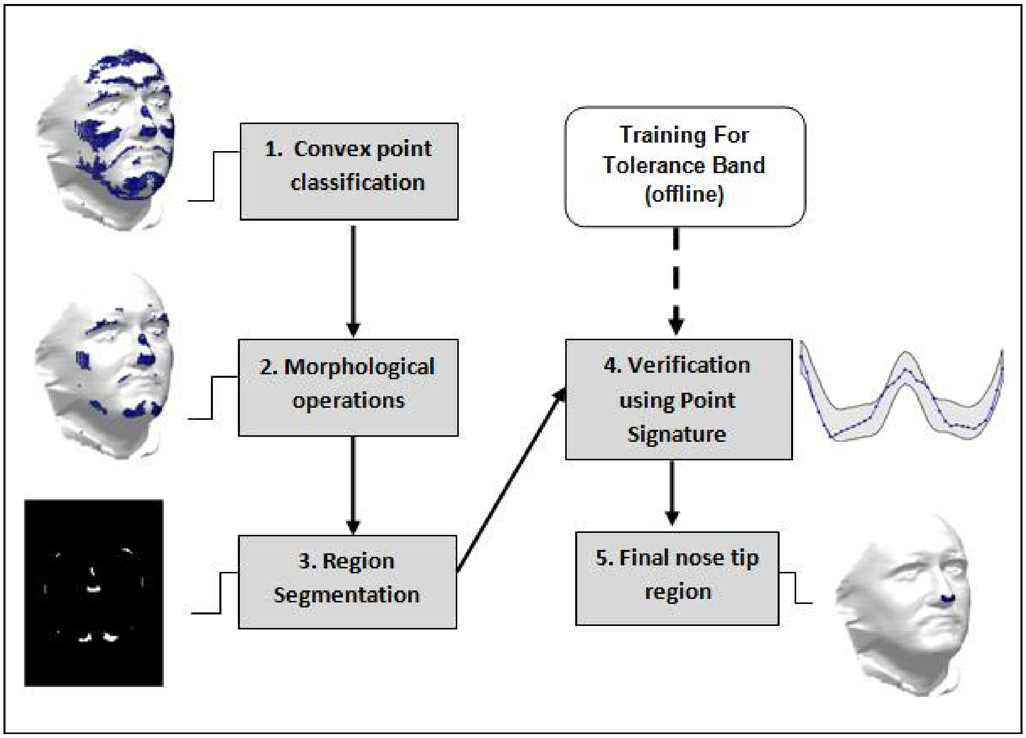


Fig. 2 Framework of the proposed nose tip region detection.

where all vectors magnitude normalized to 1. We define the neighboring points of P_0 as the 8-connected points surrounding the origin. If P_0 is convex, then we have θ more than 90 and $\cos\theta$ always being negative. In theory, surface normal is perpendicular to its surface vectors, meaning that θ and $\cos\theta$ should be exactly 90 and 0 respectively. However, in our experiments, we found out that the calculated $\cos\theta$ is slightly smaller and bigger than 0. Though the difference is very small, it is good enough for us to differentiate between convex surface and others.

During convex classification, apart from nose, convex points are most likely to lie on the chin, jaw line, cheek bone, eyebrows, hair and shirt collar. Thus, to eliminate these non-nose points, we calculate the sum of theta for each P_i, and 30 percent of points of the highest sum are selected representing most protuberant points on face. We chose 30 percent of points because the test data set used contain large non-facial outliers thus any number lower than that can cause nose-tip not fully selected on certain faces.

2.2 Search for candidate regions

Some people have a wide nose tip or bulbous nose, producing flat area on middle of nose tip. Thus, this area will not be picked up during convex classification, causing a hole in the middle of nose tip region. As the solution to this problem, we propose to use a mathematical morphology approach known as closing. Closing is a combination of expanding (dilation) and shrinking (erosion) operations, widely used in image processing to fill gaps exist within image data. In order to do this, we treat the face as binary data of set A, where convex points denoted as 1 and non-convex points denoted as 0. Dilation is performed first followed by erosion by using the same 8-connected structuring element B

$$A \cdot B = (A \oplus B) \ominus B \qquad (2)$$

where $\oplus$ and $\ominus$ denote dilation and erosion respectively. Knowing that nose tip region is among highest point density regions, we can narrow down the searching space by getting rid of regions with lower point density. The morphology operation used specifically for this purpose is erosion which applied twice to produce less number of regions.

At this point, the output image is still in binary form of convex and non-convex points. To extract meaningful regions, segmentation is compulsory. Each region is partitioned and labeled with different number and the nose tip region is finalized using point signatures.

2.3 Point signature

Point signature, was first introduced as a new presentation of free-form 3D object recognition [10], describes the structural neighborhood of a point on a face represented by sets of distance profiles. It is later used in [11-12] for face recognition. Being invariant to rotation and translation, registration can be accomplished by matching the signatures of data points of sensed surface to the signatures of data points representing the reference surface. Although point signature generally known as face recognition algorithm, we use it specifically to recognize a nose tip. This is due to its ability to efficiently describe and represent the unique shape of a nose tip.

The definition of point signature is summarized here based on [10]. For a given point *p*, we place a sphere of radius *r*, centered at *p*. The intersection of the sphere with the object surface is a 3D space curve *C*, whose orientation can be defined by a normal vector, n_1, a "reference" vector n_2, and the vector cross-product of n_1 and n_2. n_1 is defined as the unit normal vector of a plane fitted through the space curve *C*. A new plane *P´* is defined by translating the fitted plane to the point p in a direction parallel to n_1. The perpendicular projection of *C* to *P´* forms a new planar curve *C´* with the projection distance of points on *C´* forming a signed distance profile. This followed by angular sampling of every point of *C* by a clockwise rotation angle about n_1 from the reference direction n_2. This distance profile may now be represented by a discrete set of values $d(\theta_i)$ for $i=1,\ldots,n_\theta$, $0< \theta <360$ degrees where n_θ is the number of samples, as the signature at point *p*. In our experiments, we sample the signed distance for every 10 degree thus giving the number of sample, $n_\theta = 36$. Due to the simple representation of one-dimensional point signature, matching of point signature is efficient and fast.

2.4 Signature Matching

The main purpose of signature matching is to verify the nose tip region by finding the point which has the most similar signature to a nose tip signature. Due to this, we concentrate more on signature shape rather than distance error such as used for recognition. Matching is established by shifting the signature based on the highest positive angular sample, *n2*. We employ simple signature matching by comparing signatures directly to the trained nose tip tolerance band where the total score of angular samples that falls within the tolerance band is counted.

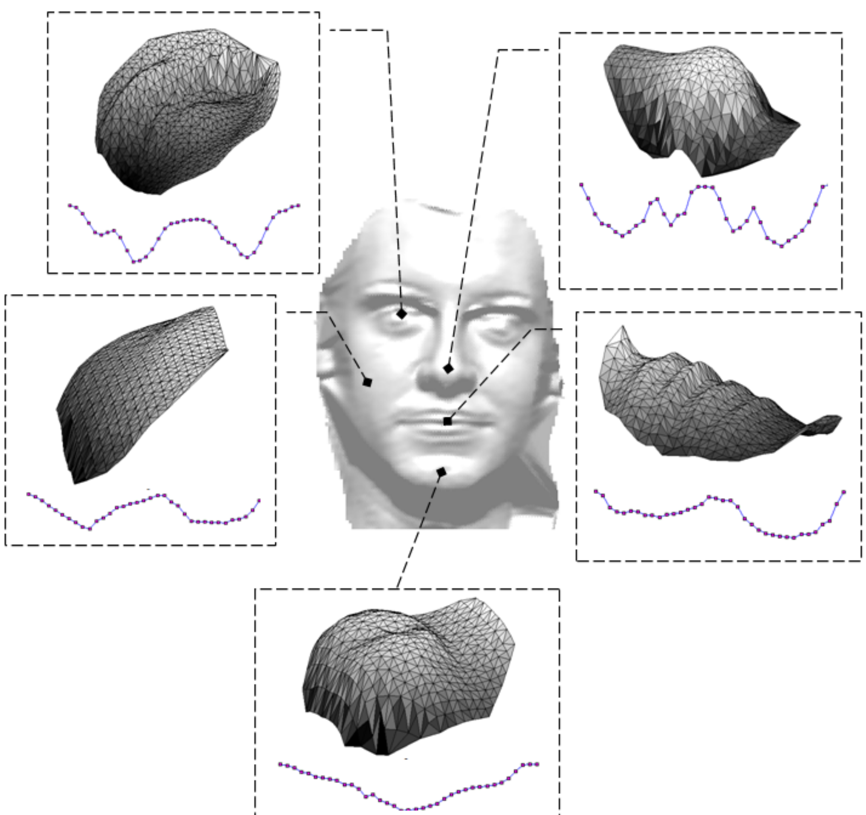


Fig. 4 Point signatures of different parts of a face after normalization.

In order to make signatures comparable with the tolerance band, every single test signature needed to be normalized between 0 to 20 units as shown in Figure 4. This will uniformly distribute the signature and make it more likeable match with the tolerance band.

Given a trained tolerance band, ε_{tol} and signature from candidate region, d_m, we wish to determine total angular samples within tolerance band, T_m:

$$T_m = \sum_{i=1}^{n_\theta} d_m(\theta_i) \le \varepsilon_{tol} \, \forall i = 1, \dots, n_\theta \qquad (3)$$

To obtain the tolerance band, ε_{tol}, we have conducted a training session using 70 frontal 3D face models combined from two databases, 3DUPMFace and GavaDB. For each point of segmented regions, point signature is generated and compared to the tolerance band. The signature that obtains highest score is noted as the maximum point. The next step is to find the region which the maximum point belongs to. The region is then extracted as the final nose tip region.

3. Experimental Setup and Results

3.1 3D Face Pre-Processing

In order to evaluate the developed algorithms, the database used should have enough variations to test its robustness under facial variation. In this work, we have selected GavaDB [13] and UPMFace as our 3D face database. GavabDB is a public 3D face database well known for its variation of facial expressions, consists of 427 Caucasian faces corresponding to 61 persons of 45 male and 16 female. Large number of faces contains spikes and most of them have non-facial outliers such as, hair, neck, shoulder, shirt collar and even a hand. UPMFace is a 3D face database which was collected in University Putra Malaysia using 3D laser scanner, consisting of 9 faces of Malay and Chinese individual. All of these models have been cleaned and the outliers were trimmed, except for one female model wearing head scarf which we kept as the striking feature of the database.

For training and testing purposes, we have combined 61 face models from GavaDB and 9 face model from UPMFace, giving us 70 face models altogether. All selected models are frontal with some of them contains minor yaw, pitch or roll rotation. These two databases consist of 3D models of different resolution range and size, thus pre-processing is necessary to set all input data in the same standard and resolution. Since both databases contain face models with large facial outliers, we cannot scale the models to fit specific number of vertices. Instead, we scale models based on the height of the models' face. In order to do this, we have manually calculated the scaling factor based on the original height and the desired height, which is 100 pixels. This process will not change the resolution of the data, but only change the coordinates of vertices. Therefore, we need to grid back the models into a standard resolution by fitting a mesh grid of 1 pixel onto each face models.

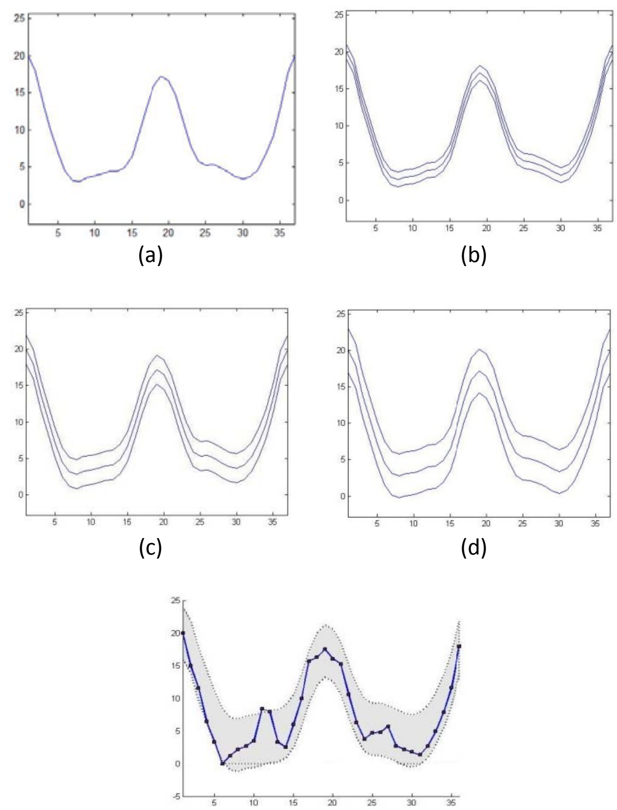
To be able to test our algorithm with respect to large rotation angle, we need a sufficient data of rotated 3D face models. Since both database only provide limited arbitrary head pose variations, it require us to synthetically resample frontal faces into rotated faces in step of 30˚ from -90˚ to 90˚ for yaw (y-axis) rotation and in step of 10˚ from -30˚ to 30˚ for both pitch (x-axis) and roll (z-axis) rotation. This gives us 420 face models for each axis and total number of 1260 face models. On top that, we also construct 20 models with combination of all yaw, pitch and roll rotations. The rotation angles used are randomly selected ranging within the permitted rotations ranges. 20 face models depicting variety of facial expression are also provided for testing. All face models were converted and saved into Wavefront OBJ file format.

3.2 Training for Tolerance Band

The main objective of training session is to model the tolerance band required for signature matching of nose tip region candidates. The tolerance band is constructed based on the orientation shape of nose tip which is significant and unique compared to other parts of a face. The training data set combines 61 models from GavaDB and 9 models from 3DUPMFace giving us 70 face models altogether. All face models in this data set have been re-sampled into a mesh grid form and scaled to the face height of 100 pixels. To initiate the training, nose tip point for all 70 models in frontal view are manually located, and their coordinates are recorded in an external file. Point signatures for all nose tip points are generated and also recorded in other external file. These point signatures will be used as the raw data for training.

To evaluate the shape of nose tip, *k*-fold cross validation method is used where *k* is the number of fold. The advantage of this method is that it matter less how the data gets divided. Every data gets to be in a test set exactly once, and gets to be in training set *k* -1 times. The variance of the resulting estimation is reduced as *k* is increased. Based on [14], the most suitable *k* value is 10, thus in this training, the training set is divided into 10 folds and the method is repeated 10 times. Having 70 models, for every experiment, 7 models are used as the test data set and the remaining 63 models as the training data set. For example, for the first experiment, Model 1 to Model 7 are chosen as the test data set and Model 8 to Model 70 becoming training data set, and so on.

According on [9], better acceptance or rejection of candidate signature is achieved by considering a varying and dynamic band of tolerance. Therefore, to investigate relationship between detection rates with the width of tolerance band, we experiment our nose tip detection algorithm using tolerance band with additional width of error value, as shown in Fig. 5. The training conducted is described in detail in Table 1 showing different accuracy result for different set of experiments for different error value. In each experiment, the average signature is calculated based on the training set.



(a) (b) (c) (d) (e)

Fig. 5 (a) Average signature of training set. Tolerance band with error value (b) E = 1 (c) E = 2 (d) E = 3. (e) Sample of a signature within tolerance band with error value E = 3.

Then we span the average signature outwards based on current error value, and test our nose detection algorithm onto the test data set. This is done to all error values. Accuracy percentage will be calculated for each error value for each experiment. From the result, both error value E=3 and E=4 obtain highest accuracy. Hence error value E=3 will be used to construct the tolerance band used in further experiments for 3D face model with facial expression, yaw, pitch, roll, and the combination of three rotation.

Table 1: Result of training experiments for each tolerance band error value.

Experiment	Test Set	Accuracy for given tolerance band error value, E (%)			
		1	2	3	4
1	1-7	85.71	85.71	85.71	85.71
2	8-14	85.71	100	100	100
3	15-21	100	100	100	100
4	22-28	85.71	85.71	85.71	85.71
5	29-35	100	100	100	100
6	36-42	100	100	100	100
7	43-49	100	100	100	100
8	50-56	85.71	85.71	100	100
9	57-63	100	100	100	100
10	64-70	85.71	85.71	85.71	85.71
Average Accuracy (%)		**94.29**	**94.29**	**95.71**	**95.71**

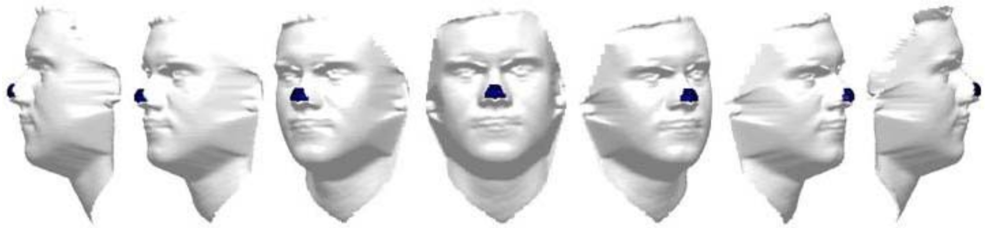
3.3 Detection Results

To test the efficiency of our algorithm towards variety of facial variations, we have conducted separate experiments for different category of variations: yaw, pitch, roll, combination of rotations and facial expression. As for comparison, we also experiment with frontal faces. The percentage of correct detection for each category is tabulated in Table 2.

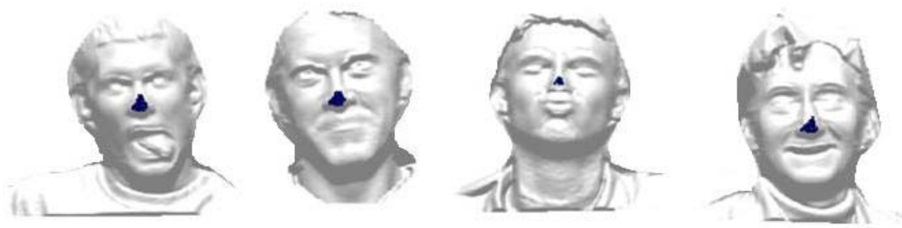
Table 2: Result of nose tip region detection for each variation.

Variation	Number of samples	Number of error detection	Correct detection rate (%)
Yaw	420	21	95.0000
Pitch	420	17	95.9524
Roll	420	21	94.5238
Combination	20	0	100.0000
Facial Expression	20	2	90.0000
Total/Average	**1300**	**64**	**95.0952**
Frontal	70	4	94.2857

During convex classification, apart from nose, chin, jaw line, eyebrows, hair and shirt collar are among regions with high points density. However, our point signature matching manages to identify the nose region quite accurately. From the results, faces with pitch rotation obtained the highest accuracy percentage compared to yaw and roll. The reason behind this is the fact that rotation around x-axis changes the symmetry of the face the least, thus the nose structure does not change so dramatically. Oppose to pitch, rotation around y-axis and z-axis changes the symmetry of the face the most thus obtained lower accuracy percentage.

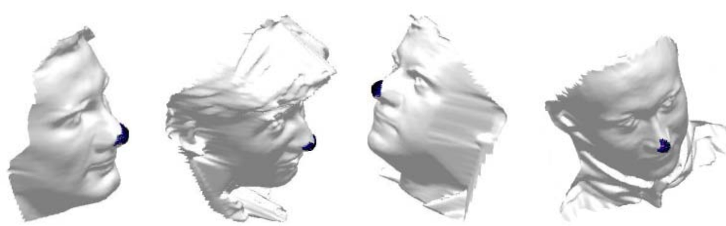


(a)



(b)

(c)



(d)

(e)

Fig. 6 Nose tip detection results: (a) Results for yaw rotation angle for -90˚, -60˚, -30˚, frontal, 30˚, 60˚ and 90˚ accordingly. (b) Results for faces depicting facial expression. (c) Results for faces with large non-facial outliers. Note that the second and last face is incorrectly detected due to protuberant outliers. (d) Results for combination of rotations. (e) Results for faces with occlusions; partial hand and hair.

We also tested our algorithm with face models with combination of rotations and facial expressions as to represent cases that are nearer to the real-time scenarios. Note that some rotations applied to the models are quite dramatic as shown in Fig. 6. As the result, from all 20 face models in both categories, we have only two false detections for facial expression and none for combination of rotations. The experimental results are illustrated in Fig. 6. The dark region in face indicates the detected nose tip region.

The success of this method lies on two main parameter. The first parameter is the appropriate value of sphere radius *r* used for point signature. From our experiments, the results are satisfying for r =18, which is able to handle the intersection of the sphere with both sides of nose wings. The second parameter is the percentage used to select regions with highest point density during convex classification. We have found that the lower percentage used, the higher the chance of the nose region get selected.

Since this method highly dependent on the richness of information of nose information, a complete and high resolution 3D face models are essential. A face with partially missing nose cannot be tolerated. Therefore, for real time application, we suggest the use of suitable acquisition and reconstruction method to obtain the 3D face models from various angles.

4. Conclusions

In this paper, we demonstrate new framework to detect nose tip region in a 3D face model by using local geometry curvature and point signature. The proposed method is fully automated, invariant to large rotation variation and robust enough for challenging face image. A simple convex classification is utilized and point signature manages to verify nose tip region quite accurately. The experiments conducted on challenging face databases demonstrate the effectiveness of our method.

Acknowledgments

This work is supported by research grant Science Fund Malaysia (5450083). The authors would like to acknowledge the use of GavaDB face database by Moreno and S'anchez [13] and 3D laser scanner by Institute of Advanced Technology (ITMA), Universiti Putra Malaysia.

References

[1] Berk Gökberk, M. Okan Irfanoglu, and Lale Akarun, "3D Shape-based Face Representation and Feature Extraction for Face Recognition", Image Vision Computing, Volume 24, No. 8, 2006, pp. 857-869.

[2] W. Zhao, R. Chellappa, A. Rosenfeld, and P.J. Phillips, "Face Recognition: A Literature Survey", ACM Computing Surveys, Volume 35, No. 4, 2003, pp. 399-458.

[3] Ayyagari, V.R. Boughorbel, F. Koschan, A. and Abidi, M.A. , "A New Method for Automatic 3D Face Registration", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2005, pp. 20-26.

[4] Wenhao Zhu, Yanyun Wang, and Baogang Wei, "Robust Nose Detection-based Feature Extraction on both Normal and Abnormal 3D Faces", 8th IEEE International Conference on Computer and Information Technology (CIT 2008), 2008, pp. 312-316.

[5] Chenghua Xu, Yunhong Wang, and Tieniu Tan Long Quan, "Robust Nose Detection in 3D Facial Data Using Local Characteristic", International Conference on Image Processing (ICIP 2004), 2004, pp. 1995-1998.

[6] Yi Sun, and Lijun Yin, "Robust Automatic Pose Estimation of 3D Facial Model", 19th International Conference on Pattern Recognition (ICPR 2008), 2008**,** pp.1-4.

[7] Michael D. Breitenstein, Daniel Kuettel, Thibaut Weise, Luc van Gool, and Hanspeter Pfister, "Real-time Face Pose Estimation From Single Range Images," IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008), 2008, pp.1-8.

[8] Xiaoguang Lu, and Anil K. Jain, "Automatic Feature Extraction for Multiview 3D Face Recognition". Proceeding of the 7th International Conference on Automatic Face and Gesture Recognition (FGR'06), pp. 585-590.

[9] Yang Jun, Liao Zhi-Wu, Li Xiao-Ning, Wu Zhen-Dong, A Method for Robust Nose Tip Location across Pose Variety in 3D Face Data, International Conference on Informatics in Control, Automation and Robotics (CAR 09), pp. 114-117, 2009

[10] Chua, C.S., and Jarvis, R. "Point Signatures: A New Representation for 3D Object Recognition", International Journal of Computer Vision, Vol. 25, No. 1, 1997, pp. 63-85.

[11] Chua, C.S., Han, F., Ho, and Y.K., "3D Human Face Recognition using Point Signature", Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition (FG'00) , 2000, pp. 233–238.

[12] I.Mpiperis, S.Malasiotis, and M.G.Strintzis, "3D Face Recognition by Point Signatures and Iso-contours", Proceedings of the Fourth conference on IASTED International Conference: Signal Processing, Pattern Recognition, and Applications, 2007, pp. 328-332.

[13] A. Moreno, and A. S'anchez, "GavabDB: A 3D Face Database" Proceedings of the 2nd COST275 Workshop on Biometrics on the Internet, 2004, pp. 75-80, URL: http://gavab.escet.urjc.es.

[14] I.H. Witten, E. Frank, Data Mining: Practical Machine Learning Tools and Technique with Java Implementations, ACM SIGMOD Record, Vol. 31, No. 1, pp. 76-77, 2002.

Laili Hayati Anuar received her Bsc. degree of Computer Science from University Technology Malaysia in 2004. She is currently a master student at Department of Computer and Communication Systems Engineering, Universiti Putra Malaysia. Her research interest including image processing and biometric.

Syamsiah Mashohor received her B.Eng degree from the Department of Computer and Communication Systems Engineering, Universiti Putra Malaysia (UPM) in 2002. Then, she continued her study for PhD at Department of Electronics and Electrical Engineering, University of Edinburgh, UK and graduated in 2006. From 2002 until 2006, she was a tutor at Department of Computer and Communication Systems Engineering, UPM and later she has been appointed as a lecturer. Her field of specialization is artificial intelligence especially Genetic Algorithms and she also involves with wide area of interest in image processing. She has taught a range of computing topics and is particularly interested in teaching programming and image processing.

Makhfudzah Mokhtar received the BEng. (Hons.) degree from the Universiti Kebangsaan Malaysia, Malaysia, in 2001 and the PhD. Degree in Electronic Systems Engineering from University of Essex, United Kingdom in 2007. She is currently a senior lecturer in the Faculty of Engineering, Universiti Putra Malaysia, Malaysia. Her research interests lie mainly in the channel coding and optical communication.

Wan Azizun Wan Adnan received her BSc in Mathematics from Southampton University, United Kingdom in 1984. She pursued her Master Degree and PhD at University of Malaya, Malaysia. Currently, she is a senior lecturer at Department of Computer and Communication Systems Engineering, Universiti Putra Malaysia. Her current research interests are software development, information security and biometrics.