International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.5 (2015), pp.373-382
http://dx.doi.org/10.14257/ijmue.2015.10.5.35

Nose Tip Detection in 3D Face Image Based On Maximum Intensity Algorithm

[1]Raed Majeed, [1]Zou Beiji and [1,2]Hiyam Hatem

[1]*School of Information Science and Engineering, Central South University, Changsha, 410083, China*
[2] *Department of Computer Science, College of Sciences, Baghdad University, Iraq*
raed.m.muttasher@gmail.com, bjzou@vip.163.com, hiamhatim2005@yahoo.com

Abstract

In order to increase the ability to track face movements with large head rotations, a 3D shape model is used in the system. In this paper, we present a robust nose tip detection method in 3D facial image that handling facial expression and hair occlusion. The 3D face smoothed by weighted median filter, the holes are filled by linear interpolation during the re-sampling phase and the 3D Gaussian filter is used to remove noise. Since the database, mesh contains unimportant parts like neck, shoulder, clothes and hair that can also change the overall appearance of a face. We propose a 3D mask to cut the face and crop useful part, which helps us to achieve sufficient accuracy for noise detection.

Nose localization is one of the most significant tasks of any facial classification system, compared to other facial landmarks. We proceed to detect the point N from the front image following the assumption that the relevant point has the highest value of the Z axis.

In this framework, the face model is determined from a frontal 3D face image. In this experiment, the performance rate is improved from 90.1% to 98.3%. As indicated by the experimental result, the proposed binary mask with maximum intensity method provides a significant improvement in performance of nose tip detection, also it success in different facial expression.

Keywords: *Denoising, 3D Mask, Nose Tip Detection, 3D facial image, 3D Face Detection*

1. Introduction

Face Recognition (FR) biometric have gained great importance ever since terrorist threats imposed weakness among the implemented security systems. Other biometrics i.e., Fingerprints or iris recognition is not trustworthy in such situations, whereas FR is considered as a fine compromise [1]. Of all the biometric such as fingerprint and iris, face recognition is the most important advantage of brass is that, it can catch at a distance and in a covert manner, face recognition remains one of the most active research topics in pattern recognition. It has become increasingly important owing to rapid advances in image capture devices, like surveillance cameras, camera in mobile phones, the availability of huge amounts of face images on the web, and increased demands for higher security [2]. Three-dimensional face recognition is a sensory system of face recognition methods in which the three-dimensional geometry of the human face is applied. It has been shown that 3D face recognition methods can reach significantly higher accuracy than their 2D rivals that it holds the potential to attain better accuracy than its 2D counterpart by measuring the geometry of rigid features on the face. Gabor filters were applied to extract features, and the senators from the filter response vectors were then learned by K-means clustering. Drira *et al.* [3] proposed partial biometric. That is, they studied the contribution of only the nose region in 3D face recognition using a similar level curve based approach. They used elastic matching for curves comparison. Colbry *et al.* [4] Propose to use two different strategies for nose tip detection under frontal and non-frontal poses. Under a frontal pose, they take the nose tip as the point of the face scan that is closest to the viewer. They did not justify the use of their six criteria. The detection rates for profile views under neutral expressions and smiles are only 82.7%.Chew et al. [5]. Also, use the effective energy condition to select nose tip candidates, but circumvent the SVM-training stage by trying to find mouth-and-eyes regions in the input image, which is not always easy. The nose tip detection rates were moderate as reported in their paper 93% for front faces. Heuristically, the nose is always assumed as the nearest point to the camera, and therefore the highest value in the z-axis. Gordon [6] used the curvature information to detect the nose; their methods are suitable for the clean 3D data and would not work in the case that there are holes around the nose.

The data from laser scanners usually contain much noise and holes around eyes, opening mouth, hair, clothes and so on. To our best knowledge, there seems no robust work on nose tip detection by now. Therefore, in our work, we preferred to estimate the location of the whole nose area instead of localizing individual landmark points. The limitation of this method was that the illumination along the nose is sensitive to the lighting condition. Therefore, in this work, we have extended the median filtering with Gaussian filter and used it to smooth the surface of a 3D face and that helps us to localize the nose tip in high accuracy. For the nose tip localization, we have used the maximum intensity concept as the tool for the selection process, after applying the proposed technique, we obtained formidably good results as is discussed below.

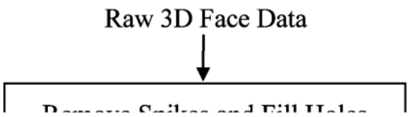


Figure 1. Block Diagram of the Nose Tip Detection

2. 3D Face Preprocessing

When developing a 3D face recognition system, one has to understand what information is provided from the camera or from the dataset, what format it is presented in and what imperfections are likely to exist. The raw data obtained from even the most accurate scanners are imperfect as it contains spikes, holes and noise. Preprocessing stages are tailored to the form and quality of this raw data. We discuss all of the stages in the following subsections. Figure 2 shows the sample of 3D face scan before preprocessing steps.

Figure 2. Sample of the Raw 3D Face Scans

2.1. Data Actuation

The acquisition technique used for capturing 3D data is especially important when collecting facial expressions, as the equipment used can affect the level of imposition on the subject, thereby changing their behavior significantly. A variety of devices and techniques have been employed previously for 3D facial expression data acquisition, including the use of single image reconstruction, structured light technologies. Automatic 3D face reconstruction from a single facial image of low resolution is an emerging research topic in computer vision [7]. Such methods have great potential in facial behavior research since recordings can be made in unconstrained environments with conventional 2D cameras, so that the subject under investigation has minimal awareness of the recording setup. However, these methods still result in errors in the reconstructed mesh which mean that the accuracy at high resolutions is not adequate for detection of subtle expressions and facial action units, and so far there are no extensive experiments on such data that have been reported.

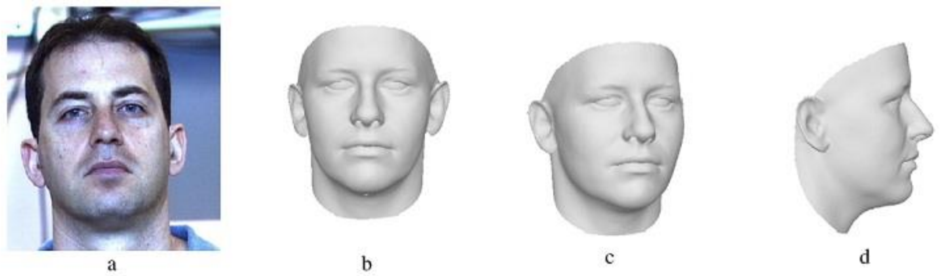


Figure 3. Example of the Morphable Model Fitted to a Single Image; (a) Single Image of Subject. (b)–(d) The Fitted Model in Frontal, Nearly Frontal and Profile Poses

2.2. Denoising and Filling Holes

Small holes in the surface resulting from missing polygons (typically due to failure in finding a stereo correspondence) can easily be filled by computing an average central point and connecting the mesh. However, larger holes occurring on one side of the nose, because the subjects were facing to the left or right, or missing areas around the eyes and forehead due to hair or glasses require a different approach.

Often, some areas that are not visible to either the camera or the projected light cannot be acquired. Similarly, dark regions such as eyebrows and mustaches, or self-occlusion or open mouth, which do not reflect sufficient projected light, are not sensed by the 3D camera. Both can have large areas of missing data, which are oftentimes consulted to as missing parts. They are identified in the input mesh by locating boundary edges, linking them together into loops, and then triangulating the resulting loops [8]. The holes are filled by linear interpolation during the re-sampling phase. The data points of each slice are interpolated at uniform intervals to fill in any holes.

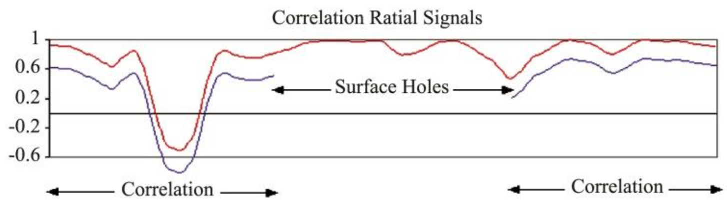


Figure 4. Correlation Partial Signal

Removal of surface noise is particularly significant as a preprocessing step in some methods of descent of the different properties of the surface, such as normal and curvatures.

To the best of our knowledge, GavabDB is the most expression-rich and noise-prone 3D face dataset currently available to the populace. In the case of faces, the eyes, nose tip and teeth are three main regions where spikes are likely to occur. The eye lens sometimes forms a real image in front of the face causing a positive spike. Likewise, the specular reflection from the eye forms an image of the laser behind the eye causing a negative spike. Shiny teeth seem to be bulging out in 3D scans and a small spike can sometimes form on top of the nose tip. Glossy facial makeup or oily skin can also cause spikes at other regions of the face. So we used median filter, it is uncomplicated and very effective instruments for noise suppressing [9].

One simple approach to filtering spikes is to test a small neighborhood of each detail in the mesh or range image and replace its depth (Z-coordinate value) by the median of this modest neighborhood. This is a standard median filter that, although effective, can attenuate fine surface detail. In our present technique we have extended the concept of 2D weighted median filtering technique to 3D face images. The present technique performs filtering of a 3D dataset using the weighted median implementation of the mesh median filtering.

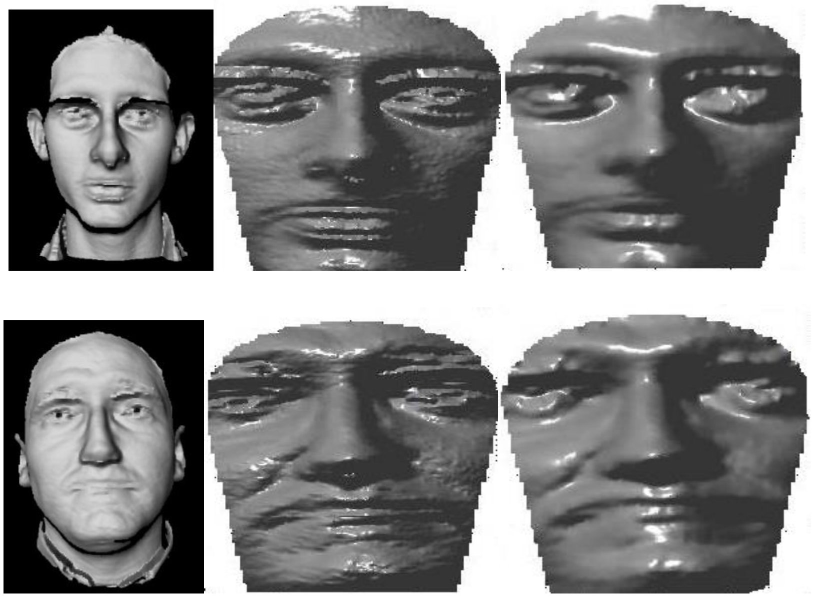
The weighted median filter is a modification of the simple median filter. Weighted Median (WM) filters are the filters that have the robustness and edge preserving capability of the classical median filter and resemble linear FIR filters. In addition, weighted median filters belong to the broad class of nonlinear filters called stack filters. This enables the use of the tools developed for the latter class in characterizing and analyzing the behavior of Weighted Median filters in noise attenuation capability. After eliminating spikes, we used a 3d Gaussian filter on our mesh data to smooth surface while preserving the geometric features of the model [10].

2.3. Face Cropping

Automatic face detection is considered to be the second essential requirement for preprocessing steps. Since the faces are non rigid and have a high degree of variability in location, color, and pose, several facial features that are uncommon to other pattern detection issues make facial detection more complex. Images acquired with a 3D sensor usually contain a larger area than just the face area, and it is often desirable to crop this extraneous data as early as possible.

Since the database, mesh contains unimportant parts like neck, shoulder, clothes and hair that can also change the overall appearance of a face. Filtering operations are sometimes performed only in a small part of an image known as a region of interest (ROI), which can be specified by defining a (usually binary) mask that delimits the portion of the image in which the operation will take place. Image masking is the process of extracting such a sub-image (or ROI) from a larger image for further processing.

We propose a binary mask to cut the face and crop useful part, afterwards, a binary mask is obtained by thresholding the difference map. The binary occlusion mask is then post-processed by morphological dilation and connected component analysis operations, which helps us to achieve sufficient accuracy for noise detection, the mask is developed to crop the facial regions. In order to reduce the size of the extracted features, the cropped region is sampled at uniform (x, y) intervals and only the seed points are kept. Figure 5 will show mesh data after cropping.



(a) (b) (c)

Figure 5. An Illustration of Facial Scan Preprocessing

As shown in figure 5 of facial preprocessing: filling gaps, cropping and filtering. (a) Shows two raw facial scans with some obvious holes, spikes and undefended points. (b) Shows the two respective cropped mesh data. (c) The surfaces have been smoothed to remove the spikes, and interpolated to fill in the holes.

We devised binary mask to crop the facial region. As shown in this figure the results of the decomposition of our binary mask are commensurate with the semi-rigid, the rigid and the non- rigid regions respectively. In order to reduce the size of the extracted features, the cropped regions are sampled at uniform (x, y) intervals (2mm in our case) and only the seed points are kept. There are two advantages of using binary masks. First, complicated algorithms for landmark feature point detection are not required, which results in a reduced computational cost. In addition, the cropped regions of different 3D facial scans will always contain the same number of vertices, which results in feature vectors of the same dimension.

3. Detection of the Nose Tip Point

Faces have a common general form with prominent local structures such as eyes, nose, mouth, chin, etc. Nose localization is one of the most significant tasks of any facial classification system [11]. Compared to other facial landmarks, the nose offers a few advantages, due to the distinct shape and symmetrical properties of a nose; it is frequently used as a key feature point in the 3D face representation. For example, finding nose facilitates the search for other landmarks such as eyes and mouth corners. Unlike nose, other features can change significantly due to facial expression.

The ridge of the nose has been used successfully in previous studies. The assumption is that by registering closely the bridge of the nose other areas that also remain unchanged under a specific expression would also be closely registered in the final result. In most images, the nose is the closest part of the face in 3D scanner; it has the highest value in depth between all points of the face. In this step, we introduce nose tip point N detection method based on both geometry characteristic and curve analysis. We proceed to detect the point N from the front image following the assumption that the relevant point has the highest value of the Z axis figure 6 As well as to verify the location of this point we extract horizontal (X-Z) and transverse (Y-Z) curve of point N.

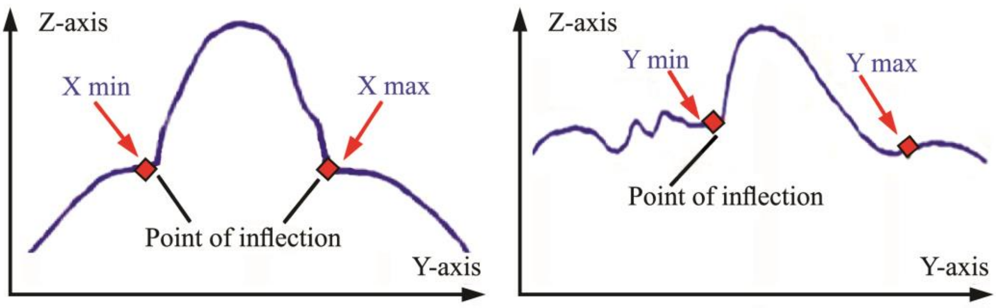


Figure 6. The Nose Tip in Horizontal and Vertical Curves

This section identifies the procedure to find the nose tip using the proposed Maximum Intensity method, after potential smoothing, the algorithm returns the nose points in all frontal cases. For detecting nose-tip the ellipse, we have used a maximum intensity technique to pull the nose-tip correctly.

The facial region avoids the effects of the hair and the bias of the face cropping. In addition, the spike removal, holes fill, and smoothing steps help accurately determine the nose tip. As shown in Figure 6, the nose tips have been labeled on the facial surface, and accordingly, the local regions are constructed based on these points. The maximum intensity algorithm used for our purpose is given:-

---

Algorithm: Find Maximum Intensity

Step 1: -Set max to 0
Step 2: - Run loop for X from 1 to the width of the image
Step 3: - Run loop for Y from 1 to the height of the image
Step 4: - Set VAL to sum (image (X-1: X+1, Y-1: Y+1))
Step 5: - if VAL > max.
Step 6: - Set VAL = VAL2
Step 7: - End if
Step 8: - End loop for X
Step 9: - End loop for Y

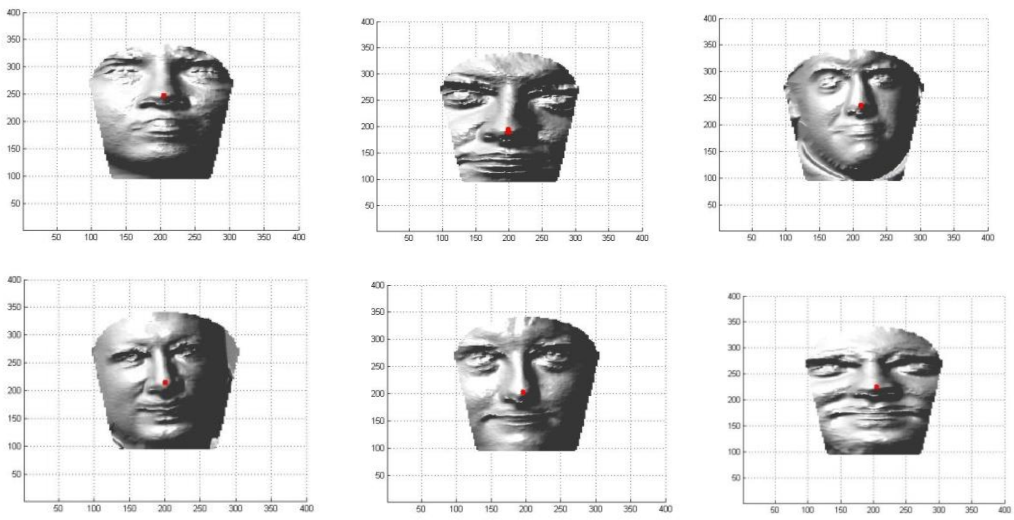


Figure 7. Some Successful Nose Tip Detection Examples; Red Crosses Correspond to our Benchmark nose Tip Positions

4. Experimental Results And Discussion

A GavabDB database is used in this scheme, it consists of Minolta Vi-700 laser range scans from 549 3D scans of facial surfaces corresponding to 61 different individuals (45 male and 16 female) [12]. All individuals are Caucasian and theirs age is between 18 and 40. Each subject was scanned 9 times for different poses and expressions, namely six neutral expression scans and three scans with an expression.

The neutral scans include two different frontal scans, one scan while looking up, one scan while looking down, one scan from the right side, and one from the left side. The expression scans include many variations with respect to the poses of each individual. Such facial expressions are: smile, laugh and a random gesture chosen by the individual. Figure 8 shown all there nine different images (views) captured per each individual.

Figure 8. Examples of All 3D Scans of the Same Subject from GAVAB Dataset

The testing data has been performed on the GavabDB database itself and it incorporates a comparison of the present algorithm using preprocessing as compared to the algorithm without preprocessing. The data set contains 549 3D scans of facial surfaces, which 305 faces were in frontal posing. In this part, the first table would prove the termination in a large set of databases after applying, or as compared to the algorithm without preprocessing and the second table would prove the result after applying preprocessing on the set of faces for each segment. Table 1 presents the performance rate of nose tip detection using the method without and with preprocessing of 3D mesh data in frontal pose.

In this experiment, we utilize the proposed binary mask to crop the face region, an input facial scan would be preprocessed to obtain highest performance rate. The performance rate is improved from 90.1% to 98.3%. As indicated by the experimental result, the proposed binary mask with maximum intensity method provides a significant improvement in performance of 3D face nose tip detection.

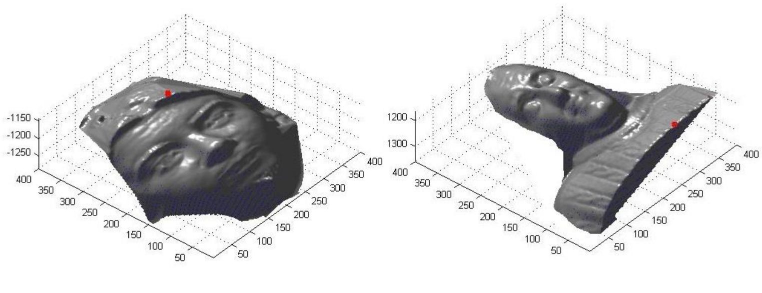


Figure 9. Some False Detection before Preprocessing

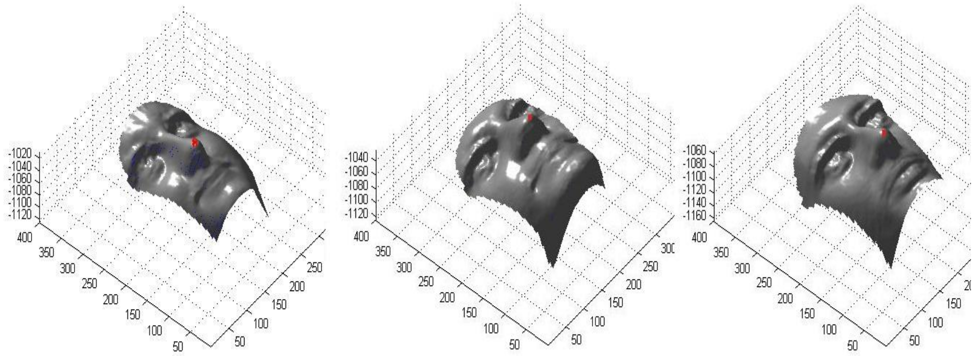


Figure 10. Some Successful Noise Tip Detection

Table 1. The Performance Rate of Nose Tip Detection

Our method	No. of sample	No. of Successful	No. of error	Performance rate
Without preprocessing	305	275	30	90.1%
After preprocessing	305	300	5	**98.3%**

We compare the proposed 3D nose tip detection with some others recently reported in the literature. Table 2 presents the comparison of performance rate with previous works. In order to locate nose tip point that corresponds to highest point in face, we extracted its transverse and horizontal curves to ensure its location. Although we have noticed that only the frontal image provided right positive detection of nose tip point, on the other side only nine scans of profile gave a positive detection.

Table 2. Comparison of Performance Rate with Previous Works

Author, year	Database	Correct rate
Colbry et al., 2005 [4]	FRAV3D	82.7%
Slater et al. ,2007[13]	FRGC v2.0	98%
Chew et al,2009 [5]	FRAV3D	93%
Anuar et al, 2010 [11]	GavabDB & UPMFace	94.2%
Our method	**GavabDB**	**98.3%**

Table 2 provides an exhaustive summary of results obtained using GavabDB; our method outperforms the majority of other approaches in terms of the nose detection rate. Although our method works well on common faces with a range of pose variations within 35 degrees, also it success in different face expression.

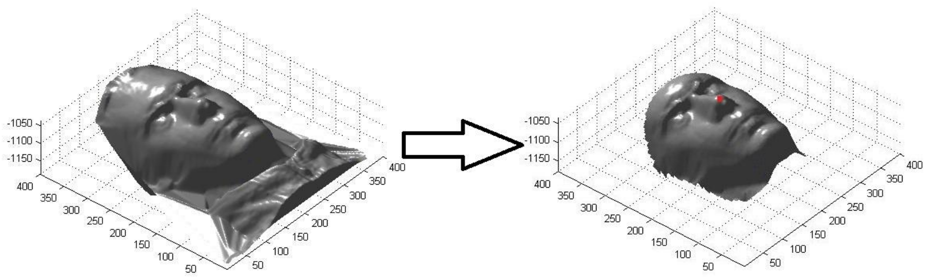


Figure 11. The Final Result of Preprocessing and Nose Tip Detection

5. Conclusion

In this paper, we propose a novel scheme to locate the nose tip in 3D facial data using the high intensity point in the GavabDB 3D database. A facial scan was cut and crop face region by using the proposed binary mask, that increased accuracy of detection the nose tip. The proposed method is fully automated, robust to noisy and incomplete input data, immune to the translation and suitable to the different resolution.

Our tests have that the proposed algorithm can achieve best performance rate after preprocessing, In particular, the performance rate was increased from 90.1% to 98.3%, and the experimental results fully show the excellent performance of our method.

In the future, we will locate the other features based on the position of the nose, and it is necessary to consider another statistic scheme for detection. In the case of GavabDB, the nose tip was manually annotated for non-frontal and occluded faces. In the future, we hope to develop automatic nose tip detection methods for non-frontal views and faces that have undergone occlusion.

References

[1] A. Azeem, M. Sharif, M. Raza, and M. Murtaza, "A Survey: Face Recognition Techniques under Partial Occlusion", The International Arab Journal of Information Technology, vol. 11, no. 1, (**2014**), January.
[2] S. Li and A. Jain, "Handbook of Face Recognition", Springer, Verlag, London Limited, (**2011**).
[3] H. Drira, B. Ben Amor, A. Srivastava, and M. Daoudi,"A Riemannian analysis of 3d nose shapes for partial human biometrics", In IEEE International Conference on Computer Vision, (**2009**), pp. 2050–2057.
[4] D. Colbry, G. Stockman and A. Jain, "Detection of anchor points for 3D face verification", in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 03, **(2005),** p. 118.
[5] W. Chew, K. Seng and L. Ang, "Nose tip detection on a three-dimensional face range image invariant to head pose", In: Proceedings of the International Multi Conference of Engineers and Computer Scientists, vol. 1, **(2009).**
[6] G. Gordon, "Face Recognition Based on Depth and Curvature Features," Proc. IEEE Conf. Computer Vision and Pattern Recognition, **(1992),** pp. 808-810.
[7] S. Wang and S. Lai, "Reconstructing 3d face model with associated expression deformation from a single face image via constructing a low-dimensional expression deformation manifold", IEEE Trans. Pattern Anal. Mach. Intell., vol. 3, no. 10, **(2011)**, pp. 2115–2121.
[8] H. Drira, B. Ben Amor, A. Srivastava, M. Daoudi and R. Slama, "3D Face Recognition under Expressions, Occlusions, and Pose Variations", IEEE Transactions on Pattern Analysis and Machine.vol. vol. 35, no. 9, **(2013),** pp. 2270-2283.
[9] O. Marques, "Practical Image And Video Processing Using Matlab", USA, (**2011**).
[10] S. Fleishman, I. Drori and Cohen O., D., "Bilateral mesh denoising", ACM TRANSACTIONS ON GRAPHICS, vol. 22, no. 3, **(2003),** pp. 950–953.
[11] L. Anuar, S. Mashohor, M. Mokhtar, and W. Adnan W.A." Nose tip region detection in 3d facial model across large pose variation and facial expression", International Journal of Computer Science, vol. vol. 7, no. 4**, (2010),** pp. 1-9.
[12] A. Moreno and A. Sanchez, "GavabDB: a 3D face database", in: Proc. 2nd COST Workshop on Biometrics, on the Internet, Vigo, Spain, (**2004**), pp. 75–80.

[13] A. Slater, Y. Hu and N. Boston, “Multi scale integration in variants for facial landmark detection in 2.5D data”, in: pro. IEEE International Workshop on Multimedia Signal Processing (MMSP007), **(2007)**, pp. 175–178.

Authors

Raed Majeed Muttasher, he received his B.Sc. in Computer Science from Baghdad University, colleges of science, computer department in 2004. He received the Master Degree in applied computer technology from Wuhan University, school of computer in 2011, P.R. China. He is currently working toward his Ph.D. degree at Central South University, School Of Information Science And Engineering, P.R. China. His research interests include 3D Object Recognition, 3D Modeling, Pattern Recognition, and Image Processing.

Zou Beiji, he holds a Ph.D. in Computer science and technology from Hunan University, China. He is currently a Professor in School of Information Science and Engineering of Central South University, China. He worked at Zhejiang University, China as a visiting scholar from Jul.1999 to Jun. 2000 and at Tsinghua University, China as a post-doctor from Jan. 2002 to Nov. 2003 respectively, engaged in research about Human Facial Expressional Recognition and Animation. He worked in the research field of Multimedia Technology at Griffith University in Australia from Dec. of 2003 to Dec. of 2004.Interest area of research: Digital Image Processing, Computer Graphics, Multimedia technology, Software Engineering.

Hiyam Hatem, he received the BS degree from the department of computer science, Baghdad University, Iraq, 2003.The master degree from Huazhong University of science and technology Wuhan, China, in 2010.Currently, she is a PhD student in School of information science and Engineering at Central south university, Changsha, China. Her research interests include face processing and recognition, object detection, pattern recognition, computer vision, and biometrics.

