

Implementation of Nose and Face Detections in Depth Image

Heung-jun Kim¹, Dong-seok Lee¹, Soon-kak Kwon^{1,*}

Abstract

In this paper, we propose a method which detects the nose and face of certain human by using the depth image. The proposed method has advantages of the low computational complexity and the high accuracy even in dark environment. Also, the detection accuracy of nose and face does not change in various postures. The proposed method first locates the locally protruding part from the depth image of the human body captured through the depth camera, and then confirms the nose through the depth characteristic of the nose and surrounding pixels. After finding the correct pixel of the nose, we determine the region of interest centered on the nose. In this case, the size of the region of interest is variable depending on the depth value of the nose. Then, face region can be found by performing binarization using the depth histogram in the region of interest. The proposed method can detect the nose and the face accurately regardless of the pose or the illumination of the captured area.

Key Words: Nose Detection, Face Detection, Depth Image, Depth Camera.

I. INTRODUCTION

The face detection is one of the most important issues in the image processing. The face detection is not only used for the face recognition, but also used for various applications such as the face auto-focusing and the posture recognition in digital cameras. However, there are problems that the accuracy of existing face detection methods for the color image is drastically reduced when the captured area is dark.

With the advancement of depth sensor technology, the depth image can be acquired accurately [1]. As the result, the depth camera has become widespread. Compared with the color image, the face recognition based on the depth image has the advantage of regardless of the illumination or the pose. Various studies based on the depth image including the face detection, the face element detection, the gender recognition, and the identity recognition are researched.

The studies related to the face detection and the face recognition need the face alignment processing and the face normalization processing [2-4]. At this time, the nose is located at the center of the face, so it becomes

the reference point for the face alignment. The nose is also commonly used to normalize the depth of the face [5, 6, 7]. Gordon [8] used curvature information to detect the nose. However, this method is suitable for clean 3D data and does not work with image including noise. Werghi [9] researched the method of the nose detection and the frontal face extraction by estimating mesh quality. He measured and assessed the mesh surface quality by extracting groups of triangular facets. The nose tip is detected by identifying the single triangular facet using the cascaded filtering framework mostly based on simple topological rules. Bagchi [10] used the HK classification to detect the nose tip and the eye corner which is based on curvature analysis of the entire facial surface.

In this paper, we propose the method of the detection of the nose and the face for the depth image. We detect the nose point using the structural features of the face and the features of the neighboring depth pixels around the nose point in the face. The nose is protruding from the face, so the nose point has the smallest depth locally. Also, the width of the face is within the certain range so we can detect the nose by

Manuscript received March 31, 2017; Revised April 05, 2017; Accepted April 09, 2017. (ID No. JMIS-2017-0005)

Corresponding Author (*): Soon-kak Kwon, Address: (47340) Eomgang-ro 176, Busanjin-gu, Busan, Korea, Tel: +82-51-890-1727, E-mail: skkwon@deu.ac.kr.

¹ Dept. of Computer Software Engineering, Dongeui University, {14177,skkwon}@deu.ac.kr

checking whether the detected point is inside the face. When the nose is detected, we find the face area and perform the binarization using the depth of the nose in order to detect the face area.

II. Nose and Face Detections in Depth Image

In this paper, we propose the method to detect the nose from the body from the depth image and to detect the face area. Figure 1 shows the flow chart of the proposed method.

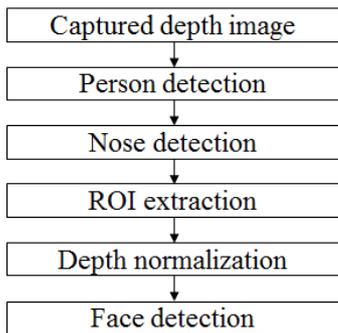


Fig. 1. Flowchart of detecting nose and face.

After capturing the area containing the body by the depth camera, the nose is detected using the relative distance feature of the nose in the face. After that, we set the rectangle whose center is the nose to the region of interest using the position of the nose and the distance of the nose.

We normalize the depth in the region of interest and we perform binarization in order to detect the face.

2.1. Method of detecting nose in depth image

First, we capture the depth image including the human body. We perform the binarization in order to separate the body from the background. In the body in the depth image, the nose is usually the closest distance from the depth camera so the depth of the nose point has the minimum value. Fig. 2 shows the characteristics of the depth of the face including the nose. In Figure 2, the nose has the lowest depth in the depth image of the face.

However, the other body part, such as lips, eyes, or chin, can have the lowest depth depending on the body's pose. Figure 3 and Figure 4 shows this case.

In order to avoid such erroneous detection cases, it is necessary to compare between the depth of the nose point and depths of surrounding pixels. Table 1 shows depth features of the nose which is distinguished from surrounding pixels. Depth features of the nose can be used to detect the nose point correctly and to avoid from the false detection.

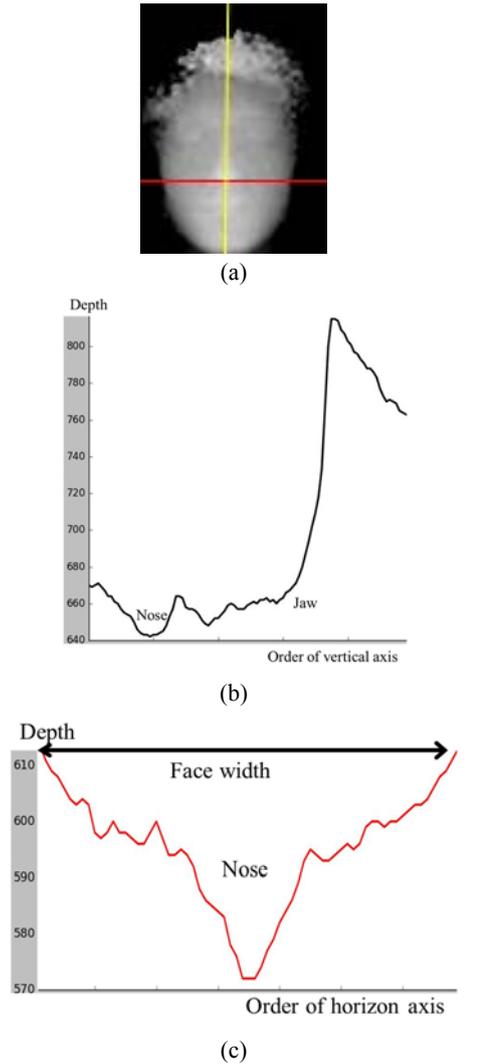


Fig. 2. Flowchart of proposed method of detecting nose and face: (a) Sample depth image of face, (b) Depth graph of face in vertical direction, and (c) Depth graph of face in horizon direction.

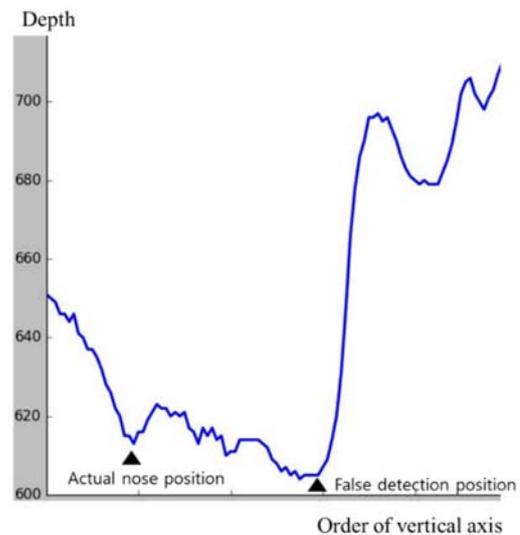


Fig. 3. False detection when chin is lifted.

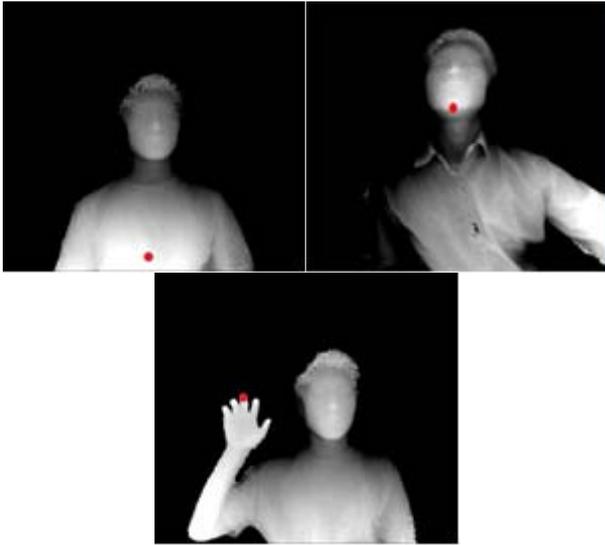


Fig. 4. Samples of false detection case in detecting nose by minimum depth.

Table 1. Features of nose in the depth image.

Classification	Feature
Adjacent pixel features	<ul style="list-style-type: none"> The depth value of the upper, lower, left, and right pixels of the nose end point is larger than the nose end point. There is a large difference in depth between the nose end point and the lower end. The jaw located below the end of the nose is at the boundary of the face and neck, so the depth changes greatly.
Facial structural features	<ul style="list-style-type: none"> The nose end point is located at the center of the width of the face The upper and both sides of the face are background areas. The bottom of the face is the object area. The nose end point is located in the center of the face width. The width of a typical human face ranges from 13cm to 22cm.

The depth decreases continuously as it gets closer to the nose. In horizon consecutive $2N+1$ pixels centered on the nose, N pixels continuous in each the left direction and right direction in one pixel are searched for pixels having the depth larger than that of the immediately preceding pixel as equation (1). In case of vertical consecutive pixels, it is also satisfactory. So we find the pixel which is satisfactory equation (1) as horizontal direction and vertical direction.

$$\begin{aligned}
 p_k &> p_{k+1}, k = i - N, \dots, i - 1 \\
 p_k &< p_{k+1}, k = i + 1, \dots, i + N
 \end{aligned}
 \tag{1}$$

The locally protruding point can be found by Equation (1). However, the depth of the nose point is significantly lower in depths of face. So we compare

the depth to surrounding pixels. We compare the depth of the found pixel to depths of eight points around the pixel located at the distance of $M(M > N)$ from the each found candidate pixel to up, down, left, and right sides to check whether each of the comparison points has a depth value larger than p_i , which is a candidate pixel. This processing is shown by Figure 5.

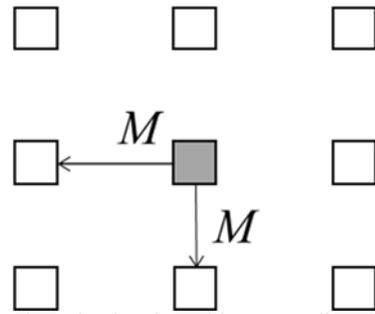


Fig. 5. Comparison depth value with surrounding pixels.

After that, we can find two boundary points $p_{left} \equiv (x_l, y_l)$ and $p_{right} \equiv (x_r, y_r)$ by drawing the horizontal line including the pixel of the nose point. We can obtain actual width of the face by Equation (2) which is the relation between the image coordinate including the depth and the 3D real coordinate [11]. In Equation (2), f is the focal length of the camera as the factor of the depth camera, and x_c, y_c, z_c are the coordinate values in the real world coordinate system. z_c is equal to the depth value.

$$x_c = \frac{d_c}{f} x_v, y_c = \frac{d_c}{f} y_v \tag{2}$$

The actual width of the face w_{face} can be obtained as Equation (3). In Equation (3), d_{left} is the depth of p_{left} , and d_{right} is the depth of p_{right} .

$$w_{face} = \sqrt{\frac{x_l - x_r}{f}^2 + \frac{y_l - y_r}{f}^2 + (d_l + d_r)^2} \tag{3}$$

Since the actual human face width is 13 cm to 22 cm, it can be seen that the candidate pixel is the nose point unless w_{face} is outside the range.

This method can detect the nose accurately even in the posture that is not the frontal face because the face is almost similar to the sphere at the boundary around the nose.

2.2. Method of detecting face using detected nose point

In order to detect the face region using nose detection, a rectangular region of interest is first set. In this case, the width of the face is generally the largest in the nose portion, or almost similar to the face width. The

boundaries of the left and right sides of the rectangle of the area of interest are defined by giving horizontal lines around the nose, L lines from the point where the boundary intersects with the vertical line. However, in the binary image, the neck just below the face is included. At this time, it can be used that the depth of the person's neck suddenly increases. Figure 6. It can be considered that it corresponds to the jaw portion in the face when the variation of the depth value is equal to or more than T_c . Using this, the lower boundary of the region of interest is defined as the jaw, and L pixels are provided in the jaw. As shown in Figure 7.

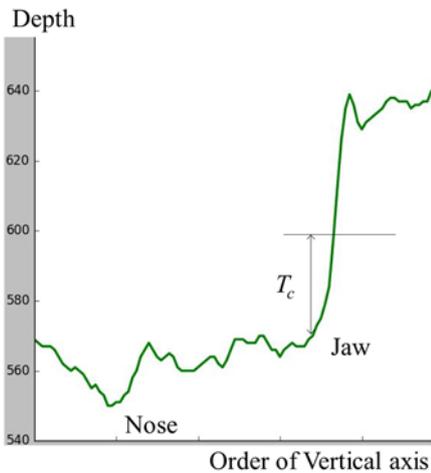


Fig. 6. Detecting bottom boundary of ROI.

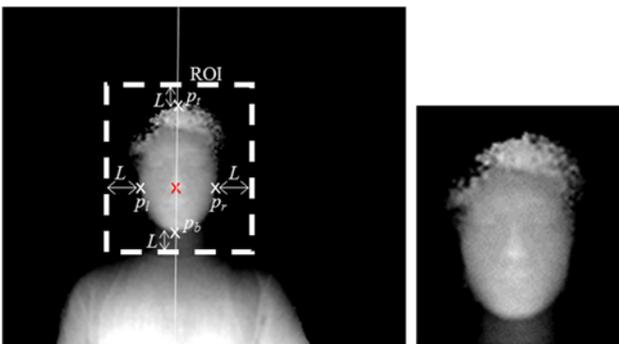


Fig. 7. Detecting bottom boundary of ROI.

Not only the face but also part of the neck area is included in the area of interest. Normalize the range of depth values in the region of interest to remove these areas of interest. To normalize the depth value, the maximum depth value and the non-zero minimum depth value in the region of interest are obtained, and the range is normalized to $[1, P_{max}]$. The pixel whose original depth value is 0 is kept it. The histogram of the normalized depth values in the region of interest is then obtained. In this case, the histogram of ROI with face and neck is shown in Figure 8.

Implementation of Nose Detection and Face Detection in Depth Image

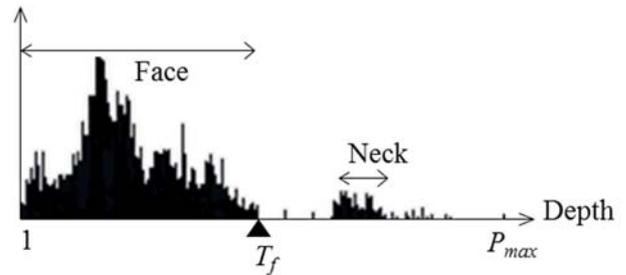


Fig. 8. Histogram of the ROI area including face and neck.

At this time, the characteristic of the histogram of the ROI can be found that the neck part and the face part are separated. By using the feature of the histogram, the threshold value T_f when binarization is performed is obtained. In this case, the depth value in the nose region is 1, which is the minimum value in the ROI, and thereafter, the depth value is continuously distributed according to the depth value of the face. Therefore, we first set the threshold T_f to the point where the depth distribution value is less than ϵ , which is the enough small value. We perform the binarization with a threshold value T_f for the corresponding ROI, then the face region can be obtained as follows.

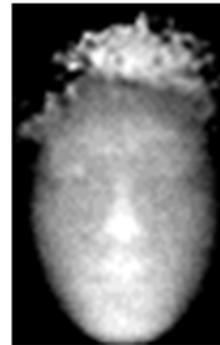


Fig. 9. Result of face detection using proposed method.

III. SIMULATION RESULTS

3.1. Simulation Environment

For examining the performance of the proposed method, we used Kinect v2 as the depth camera. The camera has the following property: the depth image resolution is 512×424 ; the method of obtain the depth is TOF.

We performed the simulation by capturing the face from various angles as Figure 10 (a) and we also performed the simulation by capturing various face as Figure 10 (b).

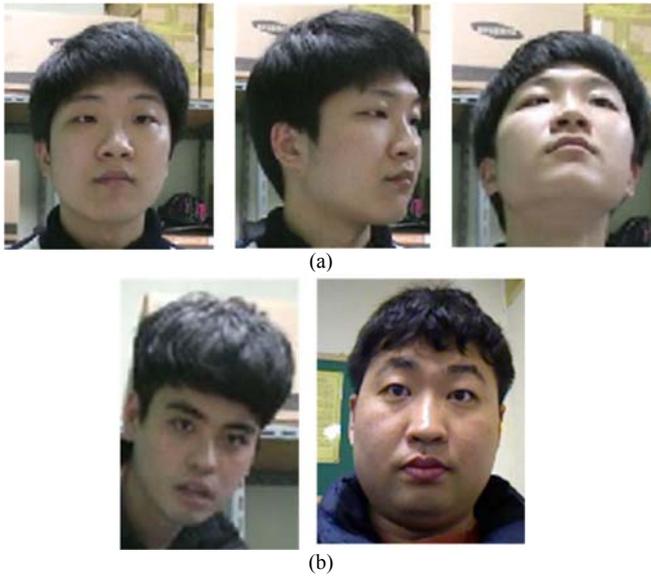


Fig. 10. Sample faces for simulation: (a) capturing various angles and (b) capturing various people's faces

We also performed face detection and comparison using existing color. We used the face detection algorithm included in the OpenCV library as a face detection method using color.

3.2. Simulation Results

The results of the experiment with different face angles are shown as Tables 2 and Table 3.

Table 2. Detection accuracy by horizon angles.

Horizon angle of face (degree)	Accuracy of Detect Nose (%)
-30	94
-15	100
0	100
15	96
30	94

Table 3. Detection accuracy by vertical angles.

Vertical angle of face (degree)	Accuracy of Detect Nose (%)
-30	86
-15	94
0	100
15	96
30	92

The results shows that the nose is correctly detection when the angle of view is close to the front. However, when the captured angle is tilted much from the front of the face, it can be seen that there is the case of misdetection.

The results of measuring the accuracy by changing the distance between the depth camera and the body are

shown by Table 4. At this time, the simulation is performed within the range of about 1.5m ~ 4m due to the characteristics of the depth camera.

Table 4. Detection accuracy according to distance.

Distance between body and camera (m)	Accuracy of Detect Nose (%)
1.5	100
2	100
2.5	100
3	96
3.5	88
4	80

The result shows that the detection accuracy is almost accurate up to 3m depending on the distance. However, it is confirmed that the detection accuracy is slightly lower at a distance of 3.5 m or more

The accuracy of facial detection is compared with the existing color method as shown in Table 5. In this case, simulation is carried out for the case where the illumination is bright and the case where the illumination is dark.

Table 5. Comparison of detection accuracy with conventional method.

Brightness of illumination	Accuracy of Detect Face	
	Conventional Method for color image	Proposed method
Bright	98	96
Dark	21	94

The result shows that the accuracy of the method using the existing color image is higher when the illumination is bright. This is because the measurement accuracy of the depth camera is still lower than the accuracy of the color camera. However, when the illumination is dark, the color image can hardly measure the face, but the depth image can be detected accurately without illumination.

IV. CONCLUSION

In this paper, we propose a face detection method using depth camera only. After separating the body from the background, the position of the end of the nose was detected using the feature of the depth pixel of the nose with respect to the body. We then set the region of interest using the end point of the nose and normalize the depth value of the region of interest. The normalized depth was binarized by separating the neck and face parts by histogram analysis. Through this, I found the face area accurately in the depth image.

In this paper, we propose a face detection method that separates background and object using depth difference. Since the nose end point is located at the center of the human face, the face area can be detected by identifying the nose end point. The method proposed in this paper can perform face detection even in dark environment compared with existing face detection algorithm, and can perform stable nose detection and face detection without being affected by body or face posture.

The nose detection method proposed in this paper is useful not only for face detection but also for face sorting for future face recognition. In existing color images, there is a limited method of correcting faces taken from the side face to face. However, in the depth image, face alignment can be performed by rotating the face using the depth characteristics. In this paper, we propose a face detection method which can reduce the amount of computation as well as simplify and accurate face recognition.

Acknowledgement

This research was supported by The Leading Human Resource Training Program of Regional New industry through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT and future Planning(No. 2016909955).

REFERENCES

- [1] G. Pan, X. Zhang, Y. Wang, Z. Hu, X. Zheng, and Z. Wu, "Establishing point correspondence of 3d faces via sparse facial deformable model," *IEEE Transactions on Image Processing*, vol. 22, no. 11, pp. 4170-4181, 2013.
- [2] S. Yan, X. Hou, S. Z. Li, H. Zhang, and Q. Cheng, "Face alignment using view-based direct appearance models," *International Journal of Imaging Systems and Technology*, vol. 13, no. 1, pp. 106-112, 2003.
- [3] A.S. Mian, M. Bennamoun, and R. Owens, "2D-3D Hybrid approach to automatic face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 1927-1943, 2007.
- [4] N. Pears, T. Heseltine, and M. Romero, "From 3D point clouds to pose-normalized depth maps," *International Journal of Computer Vision*, vol. 89, no. 2-3, pp. 152-176, 2010.
- [5] W. Chew, K. Seng, and L. Ang, "Nose tip detection on a three-dimensional face range image invariant to head pose," *Proceeding of the International Multi Conference of Engineers and Computer Scientists*, vol. 1, 2009.
- [6] A. Mian, M. Bennamoun, and R. Owens, "Automatic 3D Face detection, normalization and recognition," *Proceeding of Third International Symposium on 3D Data Processing, Visualization and Transmission*, pp. 735-742, 2006.
- [7] W. Zhu, Y. Wang, and B. Wei, "Nose detection based feature extraction on both normal and abnormal 3D faces," *Proceeding of IEEE International Conference on Computer and Information Technology*, pp. 312-316, 2008.
- [8] W. Chew, K. Seng, and L. Ang, "Nose tip detection on a three-dimensional face range image invariant to head pose," *Proceedings of the International Multi Conference of Engineers and Computer Scientists*, vol. 1, pp. 18-20, 2009.
- [9] N. Werghi, H. Boukadida, Y. Meguebli, and H. Bhaskar, "Nose detection and face extraction from 3d raw facial surface based on mesh quality assessment Assessment," *Proceeding of IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society*, pp. 1161-1166, 2010.
- [10] P. Bagchi, D. Bhattacharjee, M. Nasipuri, and D. K. Basu, "A novel approach to nose-tip and eye corners detection using H-K curvature analysis in case of 3D images" *Third International Conference on Emerging Applications of Information Technology*, pp. 311-315, 2012
- [11] S. K. Kwon and D, S. Lee, "Correction of perspective distortion image using depth information," *Journal of Korea Multimedia Society*, vol. 18, no. 2, pp. 106-112, 2015.

Authors



Heung-jun Kim(14293@deu.ac.kr) received his B.S. degree in Computer Software Engineering from Dongeui University in 2016. He is currently a graduate student in the Department of Computer Software Engineering at Dongeui University. His research interests are in the areas of multimedia signal processing and image recognition.



Dong-seok Lee(ulsan333@gmail.com) received his B.S. and M.S. degrees in Computer Software Engineering from Dongeui University in 2015 and 2017, respectively. He is currently a Ph.D. student in the Department of Computer Software Engineering at Dongeui University. His research interests are in the areas of multimedia signal processing and image recognition.



Soon-kak Kwon(skkwon@deu.ac.kr) received the B.S. degree in Electronic Engineering from Kyungpook National University, in 1990, the M.S. and Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology (KAIST), in 1992 and 1998, respectively. From 1998 to 2000, he was a team manager at Technology Appraisal Center of Korea Technology

Guarantee Fund. Since 2001, he has been a faculty member of Dongeui University, where he is now a professor in the Department of Computer Software Engineering. From 2003 to 2004, he was a visiting professor of the Department of Electrical Engineering in the University of Texas at Arlington. From 2010 to 2011, he was an international visiting research associate in the School of Engineering and Advanced Technology in Massey University. His research interests are in the areas of image processing, video processing, and video transmission.

