

Three-dimensional Head Tracking Using Adaptive Local Binary Pattern in Depth Images

Joongrock Kim¹ and Changyong Yoon²

¹Machine Learning Team, Intelligence Lab., Convergence Center, LG Electronics, Seoul, Korea

²Department of Electrical Engineering, Suwon Science College, Hwaseong, Korea



Abstract

Recognition of human motions has become a main area of computer vision due to its potential human-computer interface (HCI) and surveillance. Among those existing recognition techniques for human motions, head detection and tracking is basis for all human motion recognitions. Various approaches have been tried to detect and trace the position of human head in two-dimensional (2D) images precisely. However, it is still a challenging problem because the human appearance is too changeable by pose, and images are affected by illumination change. To enhance the performance of head detection and tracking, the real-time three-dimensional (3D) data acquisition sensors such as time-of-flight and Kinect depth sensor are recently used. In this paper, we propose an effective feature extraction method, called adaptive local binary pattern (ALBP), for depth image based applications. Contrasting to well-known conventional local binary pattern (LBP), the proposed ALBP cannot only extract shape information without texture in depth images, but also is invariant distance change in range images. We apply the proposed ALBP for head detection and tracking in depth images to show its effectiveness and its usefulness.

Keywords: Head detection, Local binary pattern, Feature extraction, Depth images, Kinect sensor

1. Introduction

Recognition of human motions has become a main area of computer vision due to its potential human-computer interface (HCI) and surveillance [1-3]. Among those existing recognition techniques for human motions, head detection and tracking is basis in all human motion recognitions. Many efforts have been devoted to detect and trace the human head using a variety of cue information such as edge [4], texture [5] and shape [6] in two-dimensional (2D) image during the past decades. However, it is still very challenging in case of large pose variations and light changes.

Recently, novel depth sensors which can acquire three-dimensional (3D) information in real-time such as time-of-flight (ToF) camera [7, 8] and Kinect sensor [9-11] have been launched into commercial market. First, ToF camera provide a full-range distance data by measuring the distance from the flight time taken for light to travel from a near infrared emitting source to the objects corresponding with near-infrared image [7, 8]. Second, Kinect sensor is based structured light system consisting of a near infrared camera and a near-infrared

Received: Jun. 10, 2016
Revised : Jun. 17, 2016
Accepted: Jun. 20, 2016

Correspondence to: Changyong Yoon
(cyoon@ssc.ac.kr)
©The Korean Institute of Intelligent Systems

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

pattern emitted source as well as a visual ranged sensor [9-11]. Consequently, 2D image based application can be extended into 3D depth image base application using additional distance information.

There are many feature extraction methods such as local binary pattern (LBP) [12-14], histogram of gradient (HOG) [15, 16] and scale invariant feature transform (SIFT) [17, 18] for 2D image. Especially, well-known LBP is useful to analysis texture information in 2D image. However, it cannot directly apply to depth image, since the image dese not include texture as well as color. Even though many real-time 3D depth sensors have been released into commercial market, there is not effective feature extraction methods for depth image.

In this paper, we propose a pattern extraction method, called adaptive local binary pattern (ALBP), which size is changed by distance information in depth images. In order to show the effectiveness of proposed ALBP, we apply it into head detection and tracking in only range image from Kinect sensor. Since a pixel intensity of depth image represents the distance from camera to the object, we can estimate the size of a head adaptively using distance. After then, using properties of ALBP, we perform head detection and tracking in real-time. Experimental results show the robust performance against distance change as well as various movements of a head at different distances.

The paper is organized as follows. In the next section, we introduce the proposed ALBP and ALBP based head detection and tracking using depth image. Section 3 presents experimental results of our system. Finally, our conclusion is given in Section 4.

2. Proposed Method

In this section, we propose ALBP which is an effective feature extraction method in depth images. Subsequently, we apply the proposed ALBP to head detection and tracking using depth images. Figure 1 shows an overview of the proposed system.

2.1 Adaptive Local Binary Pattern in Depth Images

In this paper, we perform a head detection and tracking by a novel pattern called ALBP which is an effective approach to extract useful feature of head in depth images. Even though proposed ALBP is similar with LBP [12-14] which is a texture descriptor for gray-scale image, ALBP is specialized to detect head like shaped feature in depth image.

ALBP consists of a number of points around a pixel which is fixed number points and whose radius is changed by pixel

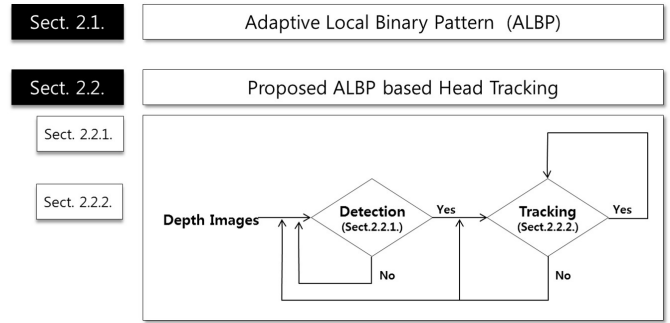


Figure 1. An overview of the proposed system.

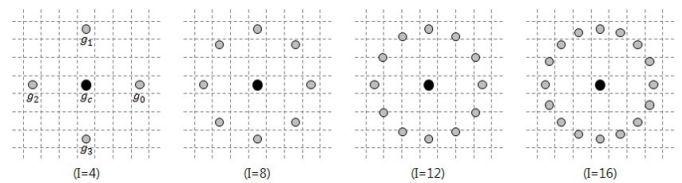


Figure 2. Adaptive local binary pattern with different I, but same r.

intensity of depth image to classify if it is the head shaped or not. Since the pixel intensity of depth image represents a distance from camera to objects, we can adjust the size of ALBP as lager when closer and smaller when far away.

First, we define texture T in a local neighborhood around a center pixel in a depth image as shown in Figure 2.

$$T_r(g_c) = t(g_0, g_1, \dots, g_{I-1}), \quad (1)$$

where g_c is a pixel intensity of center point of ALBP and $g_i (i = 0, \dots, I - 1)$ consisting of I points correspond to the pixel intensity of ALBP on a circle neighborhood of radius $r (r > 0)$ around the center point.

Second, we estimate radius r of ALBP with respect to g_c which is distance information in depth image. Size of head in an image is adaptively changed by distance between a head and a camera. Since a pixel intensity of a depth image represents distance from Kinect sensor to object, we can estimate the size of a head according to pixel information of center point (g_c). In this paper, we estimate a linear regression function, which represents relation between g_c and r :

$$r(g_{ic}) = \beta_0 + \beta_1 g_{ic} + \beta_2 g_{ic}^2 + \dots + \beta_N g_{ic}^N, \quad (2)$$

where g_{ic} is the i th observation of g_c , $r(g_{ic})$ is the radius with respect to g_{ic} , $\beta_k (k = 0, 1, \dots, N)$ is coefficient and N is the order of the regression function.

Cumulated K-observation of g_c and r can be written as

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_K \end{bmatrix} = \begin{pmatrix} 1 & g_{1c} & g_{1c}^2 & \cdots & g_{1c}^N \\ 1 & g_{2c} & g_{2c}^2 & \cdots & g_{2c}^N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & g_{Kc} & g_{Kc}^2 & \cdots & g_{Kc}^N \end{pmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix}. \quad (3)$$

In matrix notation, Eq. (3) is written as

$$R = GB, \quad (4)$$

where R is a column vector $[r_1 \ r_2 \ \dots \ r_K]^T$, B is a column vector $[\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_N]^T$ and G is the matrix between R and β .

The coefficient β can be calculated by the least squares parameter estimation by:

$$B = (G^T G)^{-1} G^T R. \quad (5)$$

Then, we can find the fitted regression function to the head size according to any distance;

$$\hat{r}(g_{ic}) = GB. \quad (6)$$

Third, texture T is subtracted by the pixel value of the center pixel (g_c) from the pixel value of neighborhood g_i ($i = 0, \dots, I - 1$) after being set the radius r of ALBP.

$$T_r(g_c) = t(g_0 - g_c, g_1 - g_c, g_2 - g_c, \dots, g_{I-1} - g_c). \quad (7)$$

Especially, this provide useful information to analysis local texture in depth images since boundary between foreground and background in depth image more distinct than color image.

Fourth, signed difference of (7) can be represented as

$$T_r(g_c) = t(s(g_0 - g_c), s(g_1 - g_c), s(g_2 - g_c), \dots, s(g_{I-1} - g_c)), \quad (8)$$

where

$$s(x) = \begin{cases} 1, & x \geq \text{threshold}, \\ 0, & x < \text{threshold}. \end{cases}$$

Also, we can reformulate (8) as (9) to easy to explain.

$$T_r(g_c) = t(s_0, s_1, s_2, \dots, s_{I-1}), \quad (9)$$

where $s_i = s(g_i - g_c)$.

Finally, we use the number of transitions (bitwise 0/1 or 1/0 changes) in the ALBP to decide if it is head shaped or not. How many transition has is very important information since it is

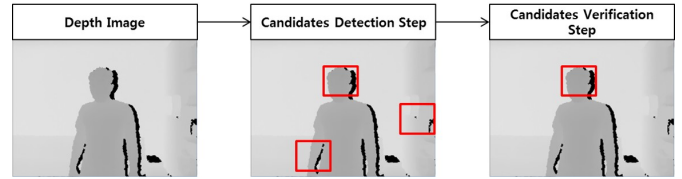


Figure 3. The proposed head detection consists of candidates detection and candidates verification step.

highly related with the texture in gray-scale image, or shape in depth image. The number of transitions can be calculated as

$$\sum_{i=0}^{I-2} p(s_i - s_{i+1}) + p(s_0 - s_{I-1}), \quad (10)$$

where

$$p(x) = \begin{cases} 1, & x = 0, \\ 0, & \text{otherwise.} \end{cases}$$

2.2 Proposed ALBP based 3D Head Tracking

In this section, we apply the proposed ALBP to detect and track the position of head in depth image. First, we detect an initial head's position to be tracked in depth image using ALBP. Based on the detected initial head position, the head tracking is performed to find head's position precisely and fast.

2.2.1 ALBP based head detection

Detection for initial head location can be classified two steps: detection step for position of head candidates and verification step of detected head candidates as shown in Figure 3.

First, to detect positions of head shaped features in depth image, we use two kinds of properties; the number of transition and signed difference of ALBP. The size of ALBP can be estimated by regression function with respect to distance information of depth image. Two size of ALBP are used: $3 \times r(g_{ic})$ for head candidates detection and $0.5 \times r(g_{ic})$ for head candidates verification.

We use a transition property of ALBP in order to detect candidates position of head in all region of depth image. Since a pixel value (g) of depth image represents a real distance information from Kinect sensor, we check the relation between head and neck. Since neck always exists with head, we check being neck using ALBP which size is $3 \times r(g_{ic})$ as shown in Figure 4(a). We assume the distance between a head and neck as 20 cm. Therefore, a point, which is the number of transition of ALBP is "1" and the number of signed difference of "0" is

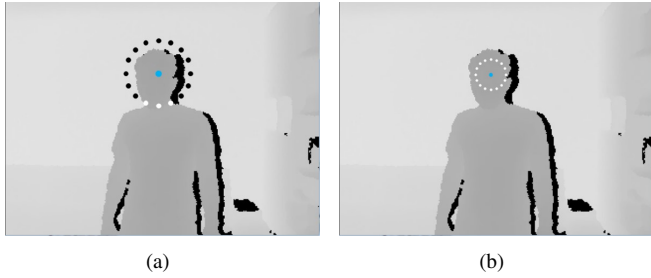


Figure 4. The results of ALBP with assumptions for head detection: (a) head candidates detection and (b) the head candidates verification; white and black points represent 0 and 1 of signs of differences, respectively.

larger than $I/4$, is real head as (11).

$$T_r(g_c) = t(s_0, s_1, s_2, \dots, s_{I-1}) = t(1, 0, 0, 0, \dots, 1, 1), \quad (11)$$

where

$$s(x) = \begin{cases} 1, & x \geq 20 \text{ cm}, \\ 0, & x < 20 \text{ cm}. \end{cases}$$

Second, we need to verify if a detected point by (13) is a real head or not. As shown in Figure 3, we used ALBP which size of ALBP is $0.5 \times r(g_{ic})$ with 10 cm threshold. Since head in depth images has features which there is no transition by threshold value 10 cm in ALBP as shown in Figure 4(b) and all signed difference of is "0."

$$T_r(g_c) = t(s_0, s_1, s_2, \dots, s_{I-1}) = t(0, 0, 0, \dots, 0), \quad (12)$$

where

$$s(x) = \begin{cases} 1, & x \geq 10 \text{ cm}, \\ 0, & x < 10 \text{ cm}. \end{cases}$$

To reduce the processing time, skip mode which does not search all coordinate of a depth image is used in this step. We choose a point which is detected during N-frame continuously is initial detected location of head. After then, tracking step is performed from detected initial location in detection step.

2.2.2 ALBP based head tracking

From the detected head's location in detection step, head tracking is performed to estimate head's position rapidly and precisely. Head tracking step in this paper can be divided into three steps; update of search range, extraction of head features and selection of a point to be tracked as shown in Figure 5.

First, for fast estimation of head locations, we need to set search range, not full search of depth image as shown in Figure

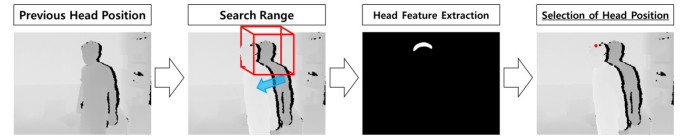


Figure 5. Overview of proposed head tracking.

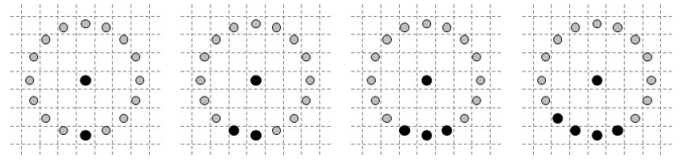


Figure 6. Some example of ALBP for head tracking; the number of transition of all examples is 1 and the number of signed difference of "0" is 1, 2, 3, and 4, respectively (from left to right); white and black points represent 1 and 0 of signs of differences, respectively.

5, according to the regression function with distance information in Section 2.1. In addition, we can set the distance ranges for head tracking since we can use distance information from depth images. In other words, we can set the search ranges X and Y coordinates as well as Z coordinate in depth images as shown in Figure 5. Search ranges of X and Y coordinate are set by 6 times of value of regression function which represents size of width of a head. The distance range of Z coordinates is set as ± 20 cm.

Second, head feature points, which are head-like shaped, should be extracted by $r(g_{ic})$ sized ALBP in search range. We use points having 2 transitions and the number of signed difference of "0" is less than $I/4$ with threshold value 10 cm in ALBP as (13)

$$T_r(g_c) = t(s_0, s_1, s_2, \dots, s_{I-1}) = t(1, 0, 0, 0, \dots, 1, 1), \quad (13)$$

where

$$s(x) = \begin{cases} 1, & x \geq 10 \text{ cm}, \\ 0, & x < 10 \text{ cm}. \end{cases} \quad (14)$$

Some examples of ALBP for head tracking are shown in Figure 6.

Third, we should select a point to be tracked from extracted many feature points. In this study, we choose a tracking point, which is a nearest feature point from a center point of extracted feature points. Since the center point of extracted feature points might be non-head location such as empty locations between fingers, the tracking point should be selected in the extracted feature points on the definite head. The nearest point from the center point is the most suitable point to be tracked since it is the less affected by noise in a depth image. After then, search



Figure 7. Kinect sensor.

region is updated by a point to be tracked and repeat step 1 through 3.

3. Experiments

Before evaluation of proposed methods for head detection and tracking, we need to estimate a regression function, which represents relation between size of a head in depth image and distances. After then, we compare the performance of our proposed approaches with ground truth at difference distances to show the robustness of proposed head detection and tracking algorithms against distance.

In this paper, we use a Kinect sensor [9-11], which provides RGB and depth images of 640×480 at 30 fps as shown in Figure 7. Even though data acquisition is implemented in Open Natural Interaction (OpenNI) [19] which is a library to access natural interaction devices, other works including detection and tracking have been simulated using C in a machine of the configuration 2.93 GHz Intel Core i7 870 and 4 GB of physical memory.

3.1 Estimation of Regression Function

Since size of ALBP is changed according to distance information in depth image and the relation between the size of ALBP and distance is not linear, it is difficult to measure all relations such as a look-up-table. Therefore, we should calculate a regression function to estimate every relation between size of ALBP and distance by the millimeter.

In order to estimate the precise regression function with respect to distance, we assume the size of head as 20 cm and measure the pixel size projected onto depth images over changing the distance from 60 cm to 750 cm at intervals of 20 cm as shown in Figure 8. Since it is hard to extract feature points such as line and corner on the plane in depth image, we select points in RGB image which is perfectly registered with depth image, after then measure the size in depth image.

The result of regression function according to order of the function, 2nd to 5th, is shown in Figure 9. The error which is pixel difference between ground truth and result of regression function is shown in Figure 10.



Figure 8. Sample images to estimate regression function according to distance; left, middle and right images represent when 60 cm, 400 cm and 700 cm, respectively.

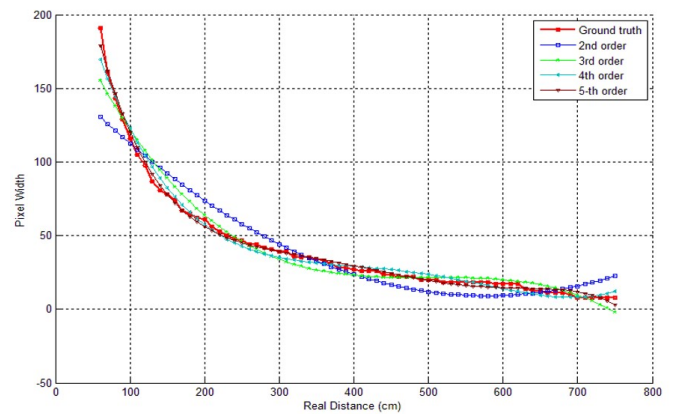


Figure 9. Ground truth and result of regression function according to order.

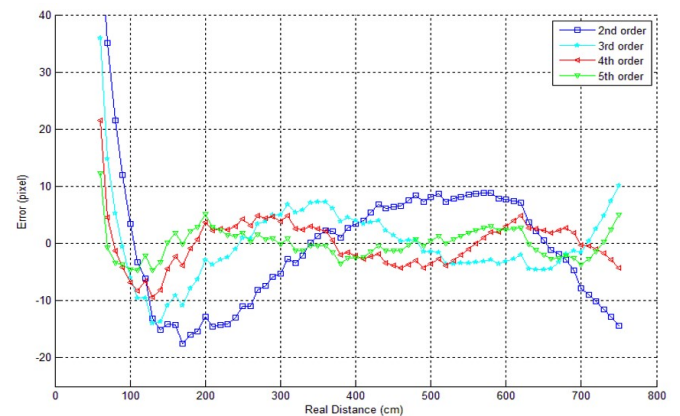


Figure 10. Error between ground truth and result of regression function according to order.

The average error of each regression function can be calculate as

$$\text{Error} = \frac{\sum_{x=60\text{cm}}^{750\text{cm}} |gt(x) - r(x)|}{\text{number of ground truth}} \quad (15)$$

where $gt(x)$ is ground truth and $r(x)$ is estimated regression

Table 1. Average error of each regression function

Case	2nd order	3rd order	4th order	5th order
Error (pixels)	9.0256	5.0102	3.3508	2.0211

functions.

The results of average error of each regression function are shown in Table 1.

As a result of experiments, since the result of 5th order regression function is the most similar with ground truth, we use the 5th order regression function to estimate size of the head in this study.

The estimated regression function used in this study is

$$\hat{r}(x) = 327.9824 - 3.2828 \times x + 0.0152 \times x^2 - 3.5441 \times 10^{-5} \times x^3 + 3.9893 \times 10^{-8} \times x^4 - 1.731 \times 10^{-11} \times x^5$$

where x is the distance (cm).

3.2 Head Detection

Using the estimated regression function, we apply ALBP to detect head's initial position in depth images. Distances in this experiments are set from 1 m to 7 m at intervals of 50 cm. Each detection rate is calculated as an average of 10 times attempts of 20 persons. To evaluate the performance of the proposed head detection method, we use precision and recall based on the classical true positive (TP), false positive (FP) and false negative (FN) rate:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

As a result of experiments, Table 2 shows a robustness of detection performance at distance, 1 m to 4 m, as 100% detection rate. However, the detection rate rapidly decreases over 4.5 m since the size of head is too small to be recognized in depth image.

The size of head in depth image according to distance is shown in Table 3. Since the size of head is less than 10 pixels over 3 m, it is difficult to verify if it is noise or heads in depth images. If all distance range has been used to detect head position, false detections which detect a non-head's position frequently are occurred. Also, Kinect cannot get the depth information

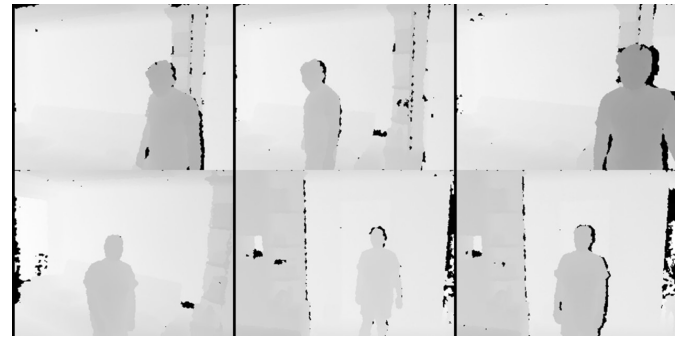


Figure 11. Data set for head detection with various head movements at different distances.

when distance is less than 0.5 m. Therefore, we set the reliable distance for head detection from 0.5 m to 4 m which ignore over 4 m or closer than 0.5 m distance in this paper. However, the resolution of depth image is rapidly increasing as great progress of depth sensor. In addition, the depth noise is decreasing as improving the quality of depth sensor such as time-of-flight. Consequently, the reliable range for head detection should be increased due to the enhancement the performance of depth sensor.

3.3 Head Tracking

To verify the robustness of head tracking against different motions and distance, we made a data set of 10 free movements from 20 persons, which walk around at different distance as Figure 11. We manually select the bounding box of head as ground truth for each frame. The result of head tracking system is in terms of a relation between ground truth and tracked bounding box.

The experimental results of head tracking with respect to X, Y and Z axis at different distance, 1 m to 3 m, are shown in Table 4. We calculate the average error between ground truth and tracking position during head movements of 10 persons. The unit of X and Y axis error is pixels and unit of Z axis is millimeter. The proposed system can be applied in various 3D applications such as surveillance and human computer interaction since it shows the reliable performance with respect to X and Y axis as well as Z axis.

Figure 12 shows the tracking trajectory of head motions with respect to each axis respectively. X and Y axis of each graph represent the number of frames and coordinate of each axis respectively. The red dashed line is the results of proposed tracking using ALBP and blue solid line represent the ground truth, which is manually selected. The results show the robust-

Table 2. Detection performance according to distance

Distance (m)	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5
Precision (%)	100	100	100	100	98.4	95.2	92.1	74.3	53.2	24.1	7.3	0
Recall (%)	100	100	100	100	100	100	91.2	81.4	60.5	37.5	14.2	0

Table 3. The size of head in depth image according to distances

Distance (m)	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7
Head size	32×54	27×42	24×36	18×25	13×20	11×17	8×13	7×12	7×11	6×11	6×10	5×9	5×8

Table 4. Mean square error (MSE) between ground truth (a center of bounding box) and tracked results

Distance (m)	1			2			3		
	X	Y	Z	X	Y	Z	X	Y	Z
MSE (mm)	50.812	47.291	164.257	69.194	61.360	288892	25.693	25.420	195.840

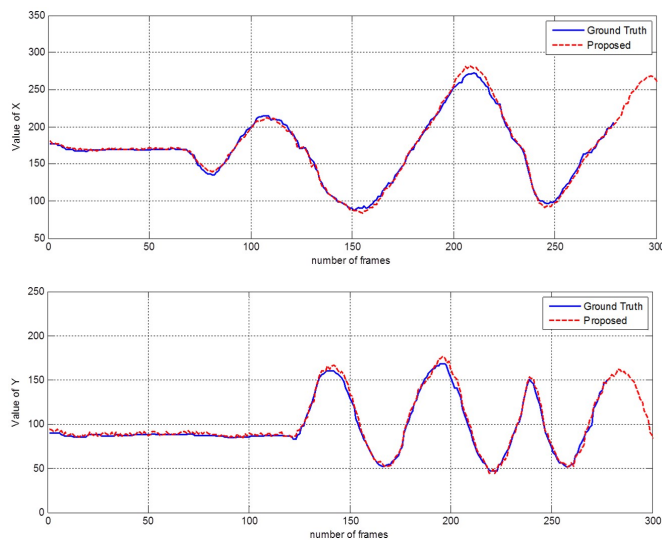


Figure 12. The tracking trajectory of head motions with respect to X and Y axis.

ness against distance as well as various movements of head without tracking failure.

4. Conclusion

In this paper, we propose an efficient feature extraction method, called adaptive local binary pattern, for depth image based applications. Even though conventional local binary pattern is useful to analysis texture information in 2D image, it cannot apply in depth image which does not contain texture information.

In addition, the depth image does not have color information which is the one of the important information to detect object. On the other hand, proposed ALBP can directly extract shape information from depth image without any additional classifiers.

In the paper, we apply the proposed ALBP to head detection and tracking in depth image. As a result of experiments, we set a reliable distance range from 0.5 m to 4 m since the head cannot be verified if it is head or noise over 4 m. In the reliable distance range, we show the effectiveness of proposed ALBP based head detection and tracing in the experiments. Also, ALBP can be applied for any depth image based applications such as human head detection and head tracking.

Conflict of Interest

No potential conflict of interest relevant to this article has been reported.

References

- [1] J. K. Aggarwal and Q. Cai, "Human motion analysis: a review," in *Proceedings of IEEE Nonrigid and Articulated Motion Workshop*, San Juan, Puerto Rico, 1997, pp. 90-102. <http://dx.doi.org/10.1109/NAMW.1997.609859>
- [2] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, et al., "A system for video surveillance and monitoring," Carnegie Mellon University, Pittsburgh, PA, 2000.

- [3] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: real-time surveillance of people and their activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, 2000. <http://dx.doi.org/10.1109/34.868683>
- [4] H. Wang, J. Chen, B. Fang, and S. Dai, "Human detection algorithm based on edge symmetry," in *Robot Intelligence Technology and Applications 3*, J. H. Kim, W. Yang, J. Jo, P. Sincak, and H. Myung, Eds. Cham: Springer International Publishing, 2015, pp. 719-729. http://dx.doi.org/10.1007/978-3-319-16841-8_65
- [5] W. R. Schwartz, A. Kembhavi, D. Harwood, and L. S. Davis, "Human detection using partial least squares analysis," in *Proceedings of 2009 IEEE 12th International Conference on Computer Vision*, Kyoto, Japan, 2009, pp. 24-31. <http://dx.doi.org/10.1109/ICCV.2009.5459205>
- [6] Z. Lin, L. S. Davis, D. Doermann, and D. DeMenthon, "Hierarchical part-template matching for human detection and segmentation," in *Proceedings of 2007 IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, Brazil, 2007, pp. 1-8. <http://dx.doi.org/10.1109/ICCV.2007.4408975>
- [7] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight sensors in computer graphics," in *Proceedings of the 30th Annual Conference of the European Association for Computer Graphics*, Munich, Germany, 2009.
- [8] Y. Cui, S. Schuon, D. Chan, S. Thrun, and C. Theobalt, "3D shape scanning with a time-of-flight camera," in *Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, 2010, pp. 1173-1180. <http://dx.doi.org/10.1109/CVPR.2010.5540082>
- [9] Microsoft, "Kinect," Available <http://www.xbox.com/en-us/kinect/>
- [10] L. Xia, C.C. Chen, and J. K. Aggarwal, "Human detection using depth information by Kinect," in *Proceedings of 2011 Computer Vision and Pattern Recognition Workshops*, Colorado Springs, CO, 2011, pp. 15-22. <http://dx.doi.org/10.1109/CVPRW.2011.5981811>
- [11] Z. Ren, J. Meng, and J. Yuan, "Depth camera based hand gesture recognition and its applications in Human-Computer-Interaction," in *Proceedings of 2011 8th International Conference on Information, Communications and Signal Processing*, Singapore, 2011, pp. 1-5. <http://dx.doi.org/10.1109/ICICS.2011.6173545>
- [12] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037-2041, 2006. <http://dx.doi.org/10.1109/TPAMI.2006.244>
- [13] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002. <http://dx.doi.org/10.1109/TPAMI.2002.1017623>
- [14] M. Heikkila, M. Pietikainen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425-436, 2009. <http://dx.doi.org/10.1016/j.patcog.2008.08.014>
- [15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, 2005, pp. 886-893. <http://dx.doi.org/10.1109/CVPR.2005.177>
- [16] Q. Zhu, M. C. Yeh, K. T. Cheng, and S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients," in *Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, NY, 2006, pp. 1491-1498. <http://dx.doi.org/10.1109/CVPR.2006.119>
- [17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004. <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>
- [18] O. Biglari, R. Ahsan, and M. Rahi, "Human detection using SURF and SIFT feature extraction methods in different color spaces," *Journal of Mathematics and Computer Science*, vol. 11, pp. 111-122, 2014.
- [19] <http://openni.org/>



Joongrock Kim received his M.S. degree in Graduate Program in Biometrics from Yonsei University, Seoul, Korea. Currently, he is a candidate of Ph.D. degree in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea. Since 2014, he has been a senior research

engineer in LG Electronics. His research interests include human computer interaction, biometrics and computer vision.

E-mail: jurock.kim@lge.com



Changyong Yoon received B.S., M.S., and Ph.D. degrees in Electrical and Electronic Engineering from Yonsei University, Seoul, Korea, in 1997, 1999, and 2010, respectively. He was a senior research engineer in LG Electronics, Inc., and LG-Nortel, and he developed system

software for the DVR and WCDMA from 1999 to 2006. From 2010 to February 2012, he was a chief research engineer in LG Display and developed the circuit and algorithms in touch systems. Since 2012, he has been a Professor in the Department of Electrical Engineering, Suwon Science College, Gyeonggi-do, Korea. His main research interests include intelligent transportation systems, pattern recognition, robot vision, and fuzzy application systems.

E-mail: cyyoon@ssc.ac.kr