HMD 환경에서 사용자 손의 자세 추정을 위한 MLP 기반 마커 분류

룩콩부*,**, 최은석*,**, 유범재*,**
*HCI 및로봇공학과, 과학기술연합대학원대학교
**(재)실감교류인체감응솔루션연구단
e-mail : vucongluc412@gmail.com, riosnai@chic.re.kr, ybj@kist.re.kr

Marker Classification by Sensor Fusion for Hand Pose Tracking in HMD Environments using MLP

Luc Cong Vu*,**, Eun-Seok Choi*,**, Bum-Jae You*,**
*Dept. of HCI & Robotics Engineering, Univ. of Science and Technology(UST)
**Center of Human-centered Interaction for Coexistence

Abstract: This paper describes a method to classify simple circular artificial markers on surfaces of a box on the back of hand to detect the pose of user's hand for VR/AR applications by using a Leap Motion camera and two IMU sensors. One IMU sensor is located in the box and the other IMU sensor is fixed with the camera. Multi-layer Perceptron (MLP) algorithm is adopted to classify artificial markers on each surface tracked by the camera using IMU sensor data. It is experimented successfully in real-time, 70Hz, under PC environments.

1. Introduction

Hand pose detection is critical in HMD-based virtual reality and/or augmented reality environments for natural interaction with virtual information using user's hand. There have been proposed a few sensors for hand pose tracking such as lighthouse trackers from HTC, Oculus Lift and Samsung, and a stereo IR camera from Leap Motion, Inc. Leap Motion Camera has much interests since it has a potential to provide an interaction by bare hands.

Leap Motion Camera (LMC) has been applied for hand gesture recognition and sign language recognition of Arabic Sign Language, Indonesian Sign Language, and Arabic numbers [1-5]. Leap Motion Inc. showed a demonstration on hand interaction with virtual information in AR environments. It is not shown, however, how accurately the camera can detect the hand pose.

There have been proposed several approaches to decide hand pose accurately in [6-8]. IR LED markers are used together with LMC to track different kinds of objects [6]. A hardware board is designed and attached on a target for tracking to control and provide power for IR LED markers. In [7], a sensor fusion approach by the combination of an IMU sensor with a fixed external camera is proposed for precise position estimation of fiducial markers attached with an IMU on the back of hand. Its 3D pose is decided from the IMU sensor data and compensated by position data from the video tracking using extended Kalman filter model. [8] introduces a novel pose tracking approach for mobile devices by attaching a tricolored strip on AR books and using an embedded inertial sensor and camera in a mobile device.

In previous approaches, however, the pose of a target cannot be detected when one or more markers are lost during changes of line-of-sights by motions of user's hand and HMD. To overcome the difficulty, there is under development a new approach to detect hand pose by using simple circular artificial markers on a box with an IMU sensor attached at the back of hand tracked by Leap Motion camera with an IMU sensor.

In this paper, there is proposed a method to classify circular markers on a surface of a box by using IMU sensor data. By classifying a surface that includes circular artificial markers tracked by Leap Motion camera, the position of a box is decided by using relative relationship between the positions of makers and center of the box. The algorithm is experimented successfully in real-time, 70Hz, under PC environments.

2. Proposed Sensor System & Algorithms

2.1. Sensor System

In order to detect the pose of user's hand in HMD environments, a Leap Motion camera is attached in the front of a Head-mount display together with an IMU sensor and a cube (box) with circular markers in each surface is located at the back of user's hand. One IMU sensor is included in the cube as shown in Fig. 1.

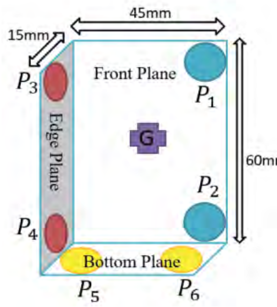


Figure 1. Cube with circular markers and an IMU sensor

The circular markers are made by using a retro-reflective sheet with high reflection characteristics with respect to infrared lights emitted from Leap Motion camera. Its radius is around 5 mm. And two 9-axis motion tracking IMU sensors, MPU-9250, are adopted since it combines a 3-axis gyroscope, 3-axis accelerometer, and 3-axis magnetometer. Sensor data from the camera and two IMU sensors is transferred to a PC and processed in the PC.

2.2. Marker Tracking

Raw images from Leap Motion camera are converted into binary images by thresholding at 10 levels and effective blobs are found by using OpenCV library [9]. Parameters for filtering such as radius, area, height of a blob are set experimentally. To remove noisy blobs, a reference distance between two markers is adopted. The reference distance between two i^{th} and j^{th} blobs is calculated as function (1).

$$D_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \qquad (1)$$

The i^{th} and j^{th} blobs are detected as markers if $D_{i,j} \in [D_{min}, D_{max}]$ in which D_{min}, D_{max} are a minimum and maximum distance that are determined from experiments for two markers. Then, the set of markers from the left and right camera images are correlated into pairs for 3D geometry information extraction. The correlation method is used in this paper based on the constrain rectangle conditions as the function (2).

$$\|rightmarker.x - leftmarker.x\| < a$$
$$\|rightmarker.y - leftmarker.y\| < b \qquad (2)$$

Where a and b are the constrain constants of the rectangle. 3D position of each marker, $P_i = (x_i, y_i, z_i)$ in space, is decided by stereoscopic geometry of Leap Motion camera using disparities between corresponding markers. Then, the position of the cube (user's hand) is determined finally by the known relationship between markers and the center point of the cube.

2.3. Marker Classification

Each surface of the cube has same circular makers and the markers in one surface may be confused with markers in the other surface when users rotate his/her hand if visual information is used only. That is, even though the camera tracks markers correctly, the surface may be changed following the user's hand motion. To solve this problem, a method using two IMU sensors, one sensor in the box and the other sensor in Leap Motion camera, is proposed to detect the front plane, edge plane or bottom plane of the box as Fig. 1 by using MLP algorithm (Artificial Neural Network) [10].

Since one IMU sensor is fixed in the box, the relative Euler angles between two IMU sensors are different for each surface of the box when a surface becomes a front plane in view of the camera by user's hand motion. So we can get the relative Euler angles including Roll, Pitch, Yaw for different surfaces when users move and rotate hand in the field of view of the camera before real-time operation. The obtained data is shown in Fig. 2 while three data sets are shown in different colors, blue, green and red. Each axis shows roll, pitch, and yaw data respectively.

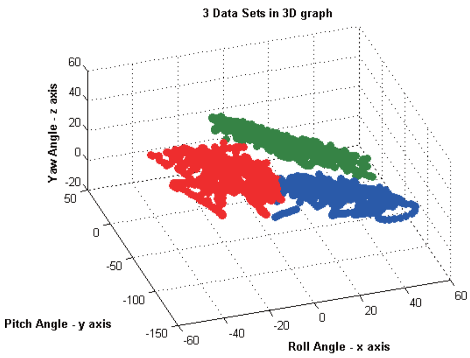


Figure 2. Three data sets captured from IMU sensors

In addition, we also get the distance of x-axis and y- axis between two markers, given by Δx and Δy in order to avoid some case that the relative Euler angles is overlapped. Thus, we have the data set with five elements (Roll, Pitch, Yaw, Δx, Δy) for each surface.

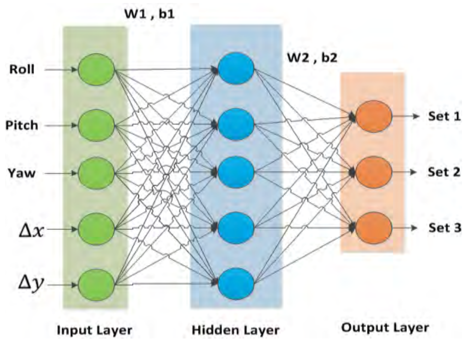


Figure 3. Neural Network model

The Neural Network model as Fig. 3 to classify (or separate) three data sets can be found by applying Multi-layer Perceptron algorithm. Finally, the surface including markers tracked by the camera is determined by the Neural Network model using two IMU sensor data in real-time and the position of the box and user's hand is decided.

3. Experimental Results

The implemented sensors are shown in Fig. 4. The first one is the box with retro-reflective markers and an IMU sensor with an Arduino board to read IMU data and send to PC. The second is a Leap Motion camera with a fixed IMU sensor.

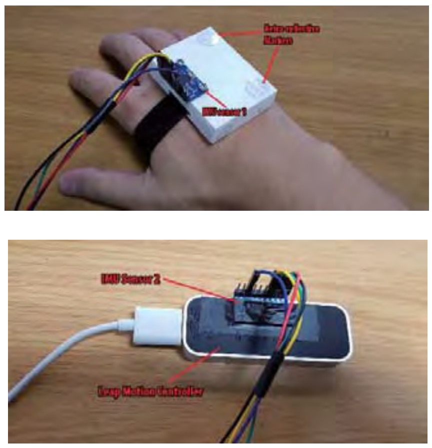


Figure 4. Implemented Sensors

3.1. Neural Network training for classification

Three data sets for three surfaces of the box are obtained from two IMU sensors and Leap Motion controller during the tracking of markers on three different planes by the camera. The number of data in each data set is around 1,000. We used these data as input data for training Neural Network model with the increase of the number of units in hidden layer. The Table 1 shows the training accuracy for model.

From the training accuracy on the table, we use the Neural Network with 5 units in hidden layer that is represented by weight matrices (W_1 , W_2) and bias vectors (b_1, b_2).

Table 1. The training accuracy

Number of units in hidden layer	Training accuracy
3 units	67.28%
4 units	99.69%
5 units	99.97%
6 units	99.97%

3.2. Marker Tracking and Surface Detection

After training, the Neural Network model is applied to classify the data and determine the front plane, edge plane or bottom plane in real-time. The tracking results are shown in Fig. 5, Fig. 6 and Fig. 7 for three positions of user's hand in each figure after real-time classification. Fig. 5 shows blue circles are drawn for marker pair (P_1, P_2) in the front plane. Fig. 6 represents red circles for marker pair (P_3, P_4) in the edge plane. The marker pair (P_5, P_6) in the bottom plane is drawn in Fig. 7 by yellow circles.

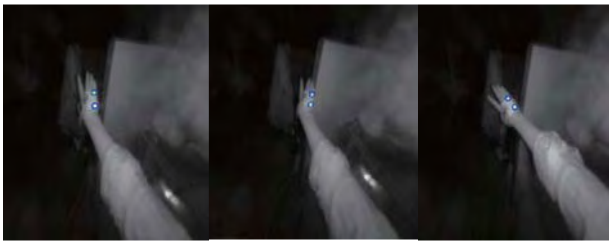


Figure 5. Blue circles are drawn for marker pair (P_1, P_2)

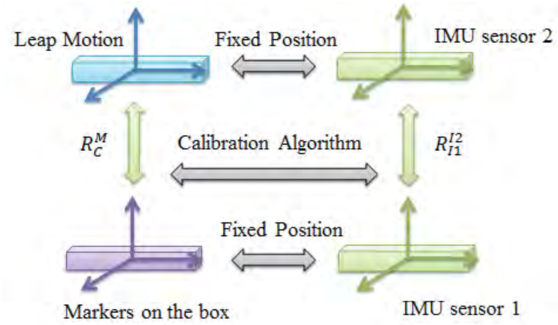
Figure 6. Red circles are drawn for marker pair (P_3, P_4)

Figure 7. Red circles are drawn for marker pair (P_5, P_6)

The marker classification and tracking are done successfully in real-time with 70Hz when the hand is moving in space freely.

4. Conclusions and Future Works

In this paper, we present a method to classify and track circular markers on a surface of a box by using a Leap Motion camera to track retro-reflective markers and two IMU sensors to detect the pose of user's hand. The model to separate three surfaces is extracted by the MLP algorithm and applied for tracking and position detection of a box successfully. In the future works, a real-time hand pose detection approach in the following figure will be investigated by including the orientation detection, and we apply it to spatial hand interaction in VR/AR environments.



References

[1] C. Naidu1, A. Ghotkar, "Hand Gesture Recognition Using Leap Motion Controller". In International Journal of Science and Research (IJSR), Volume 5 Issue 10, October 2016.

[2] S. Ameur, A. B. Khalifa, M. S. Bouhlel "A Comprehensive Leap Motion Database for Hand Gesture Recognition". In the 7th International Conference on SETIT, Dec. 2016.

[3] B. Khelil, H. Amiri, "Hand Gesture Recognition Using Leap Motion Controller for Recognition of Arabic Sign Language". In 3rd International Conference on ACECS, PET, pp. 233-238, 2016.

[4] M. D. Wibowo, I. Nurtanio, A. A. Ilham, "Indonesian sign language recognition using leap motion controller". In 2017 ICTS, Oct. 2017

[5] Q. Wang, Y. Wang, F. Liu, W. Zeng "Hand Gesture Recognition of Arabic Numbers Using Leap Motion via Deterministic Learning". In CCC, 2017 36th Chinese, July 2017.

[6] E. Urtans, A. Nikitenko "Active infrared markers for augmented and virtual reality". In proceedings of the international scientific conference 2016.

[7] H. Bastian, G. F. Trommer "Indoor 3D Position Estimation Using Low-Cost Inertial Sensors and Marker-Based Video-Tracking" IEEE/ION Position, Location and Navigation Symposium, 2010.

[8] J. Li, H. Aghajan, J. R. Casar. W. Philips "Camera Pose Estimation by Vision-inertial Sensor Fusion: An Application to Augmented Reality Books" Society for Imaging Science and Technology, 2016

[9] "The OpenCV Tutorials" Release 2.4.13.7, September 25, 2018

[10] Ian Larsen "Introduction to Artificial Neural Networks" Department of Mathematical Modelling Technical University of Denmark, November 1999