Hand Tracking Accuracy Enhancement by Data Fusion Using Leap Motion and MYO

Jingxiang Chen, Chao Liu, Rongxin Cui, Chenguang Yang

Hand Tracking Accuracy Enhancement by Data Fusion Using Leap Motion and MYO

Jingxiang Chen, Chao Liu, Rongxin Cui, Chenguang Yang[*]

Abstract—In this paper, two methods for hand tracking and and online hand guesture identification is proposed by using the combination of Leap Motion and MYO armband. With the proposed methods, We have improved the measurement accuracy of the palm direction and solved the problem of insufficient accuracy when the palm are at the limit of the measurment range. We use the Kalman filter algorithm and the neural network classification method to process and analyze the data measured by Leap Motion and MYO, so that the tracking of the operator's hand gesture is more accurate and robust even when the hand is at positions close to the measurement limit of one single sensor. The improved hand tracking method can be used for robotic control, teaching by demonstration or teleoperation. The effectiveness of the proposed methods have been demonstrated through comparative experiments.

Index Terms—Leap Motion, MYO, sensor fusion, hand tracking, hand guesture identification.

I. INTRODUCTION

Nowadays, robot technology is now widely used in more and more fields, and robotic hand control is one of the important areas. The real environment is much more complicated than the ideal environment. When the robot relies on a single vision sensor, it is sometimes difficult to get enough accurate information to perform related tasks [1]. As a result, in order to improve the accuracy of robot control, it is necessary to study the method of sensor fusion and apply it to the control of the robotic hand.

In some areas, robotic teleoperation has done a lot of work, and these efforts have made great progress on this topic. In the study of teleoperation, K. Ogawara et al. studied the problem of visual occlusion in robot remote teaching tasks [2], and J. Luo, C. Yang et al. conducted a detailed study on enhanced teleoperation performance using hybrid control and virtual fixture [3].

In addition to teleoperation and teaching research, the process of obtaining signals for robot control by sensors is also worthy of attention. Among the many sources of signals that can be selected, hand movements are an accurate and convenient source of signals [4], which can be used for hand tracking and gesture recognition. Currently, Kinect developed by Microsoft is widely used in the work of capturing human movements. For example, Z. Ping et al. developed a system that uses the operator's arms to control the robot arm using Kinect sensors [5]. However, Kinect is a sensor designed for full body movements that is not accurate enough to track hand movements.

Leap Motion, designed to track hand movements, is a better choice. Compared to Kinect, Leap Motion is a cost-effective visual sensor. Moreover, Leap Motion is small enough to be easily used with other devices or installed in a variety of locations. In addition, Leap Motion officially provides a mature software interface for developers to use, developers can easily carry out secondary development. Therefore, there has been great progress in the research and application of Leap Motion. For example, T. V. S. N. Venna et al. used Leap Motion to control the 4-DOF robot [6]. H. Jin et al. use Leap Motion to capture human hand information and recognize motion gestures for robotic [7]. D. Bassily et al. use Leap motion for intuitive and adaptive robotic arm operations to assist the elderly and the disabled in their daily lives [8]. J. C. Coelho et al. present an evaluation of 3D pointing tasks using Leap Motion sensor to support 3D object manipulation [9].

Although Leap Motion has high precision in tracking hand activity, when the target we are tracking is in a visual blind spot state, the tracking accuracy will drop a lot, such as when the palm or arm is blocked or the fingers overlap. This problem can cause system tracking performance to degrade, which may lead to some errors in subsequent gesture recognition and robot control.Therefore, in order to solve such problems, it is possible to solve the problem by adding another sensor. For example, in order to solve the problem of occlusion after hand rotation, H. Jin et al.'s research uses Multi-Leap Motion for tracking [10]. And for sensor fusion, G. Du et al. [11] proposed a matrix-weighted multi-sensor optimal information fusion criterion based on linear minimum variance in the study of Kalman filtering (KF) and particle filtering (PF).

There are a lot of research to make up for the shortcomings of single-sensor by multi-sensor. Among them, MYO armband is a good choice for sensors that work with Leap Motion. The MYO sensor can provide different types of signals than Leap Motion, reflecting the state of the operator from another aspect, and has the advantages of being easy to wear and cost-effective. A. Boyali et al. used MYO and spectral collaborative representation based classification for gesture recognition [12].M. E. Benalcazar et al. used a model based on K-nearest

J. Chen is with Key Laboratory of Autonomous Systems and Networked Control, College of Automation Science and Engineering, South China University of Technology, Guangzhou, China.

C. Liu is with LIRMM, CNRS-University of Montpellier, France. liu@lirmm.fr.

R. Cui is with School of Marine Engineering, Northwestern Polytechnical University, Xi'an, China. Email: r.cui@nwpu.edu.cn.

C. Yang is with Bristol Robotics Laboratory, University of the West of England, Bristol, BS16 1QY, UK. Email: cyang@ieee.org.

neighbor and dynamic time warping algorithm to identify five types of gestures for MYO EMG signals [13]. S Rajan et al. used MYO's gesture recognition function to conduct research on physiotherapy healthcare [14].

Using Leap Motion and MYO data to track the hand or arm at the same time is a study related to sensor fusion. At present, research on sensor fusion, there are many, for example, C Yang et al. based on the Kalman filter algorithm for data fusion of sensors, and had a research about the teleoperation control of Baxter robot [15]. ECP Silva et al. combined the data of Leap Motion and MYO to study the tracking problem of the arm [16].

Working with two Leap Motions simultaneously allows them to compensate for each other's measurement dead zones. However, the method of using two Leap Motion for data fusion requires two devices to be placed at 60° [10], which constrains the hand motion space and also causes inconvenience in hardware configuration of the measurement system, as shown in Fig. 1. At present, the research using the visual sensor and Leap Motion is mainly for tracking the position and motion of the arm or the entire palm. Therefore, work of this paper not only considers the movement of the palm, but also tackles the problem of motion tracking of a specific finger.
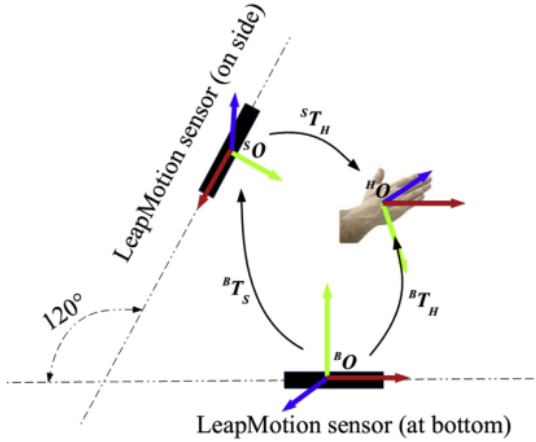

Fig. 1. The placement of Multi-Leap Motion

In this paper, in order to improve the accuracy of Leap Motion's tracking of the operator's hand, we propose two methods to combine Leap Motion and MYO data. We combine Leap Motion data with MYO's sEMG signal and rotation signal for better measurement in tracking.

The main contributions of this paper are listed below:

(1) A method of combining Leap Motion and MYO information using Kalman filtering algorithm is proposed to improve the accuracy of the direction of the operator's hand.

(2) Improve the motion recognition accuracy of Leap Motion under abnormal working conditions by using MYO's sEMG signal. In the experiment where the fingers overlap each other, the method we used accurately identified the active finger with the accuracy of 99%.

(3) The two methods were verified by experiments. Based on the results of the experiment, we found that the first method effectively reduces the deviation between the measured angle and the actual angle. The second method has 99% accuracy in identifying single finger activity. Finally, the validity of the above two methods can be verified.

II. SYSTEM SETUP

In this paper, we use Leap Motion, MYO armband, and a PC to construct a hand tracking system. The components of the system are shown in Fig.2. Among them, Leap Motion collects the finger joint angle of the operator's hand and the direction of the hand. In addition, MYO provides the operator's hand sEMG signal and rotation signal.
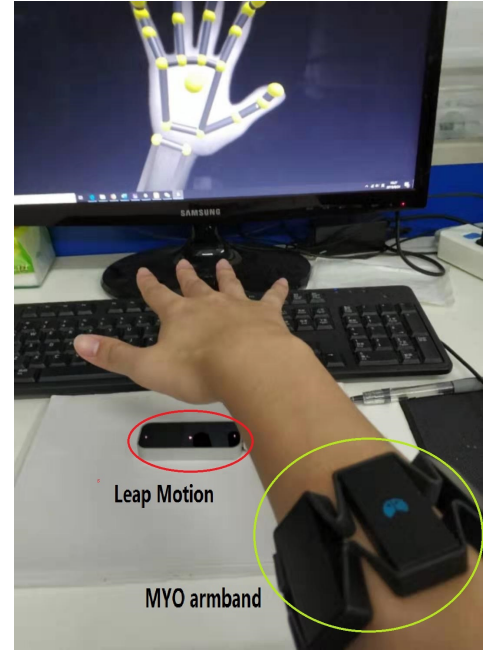

Fig. 2. System Components

A. Leap Motion sensor

As shown in Fig.3, Leap Motion is an important sensor in the simulation system to track the operator's finger joint angle and palm rotation information. Leap Motion's detection range is between 25 mm and 600 mm above the sensor, and the space tested is an inverted quadrangular pyramid. As can be seen from Frank Weichert's research [17], in the effective detection range of Leap Motion, the static accuracy of the detected position information can reach 0.2 mm, and the dynamic precision can reach 1.2 mm.

Leap Motion has two built-in cameras and three LED lights. It works according to the principle of binocular imaging and therefore performs poorly when occluded. The internal structure of the Leap Motion is shown as Fig.4.
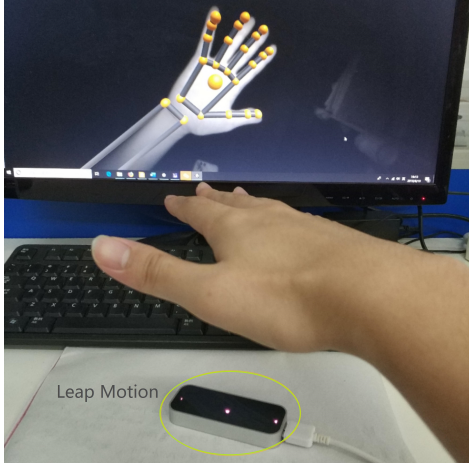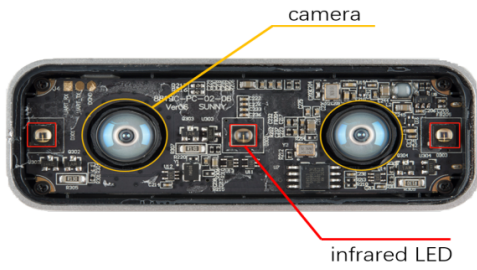
Fig. 3. Leap motion in practical applications

Fig. 4. Leap Motion inside

B. MYO

MYO is a device with built-in IMU modules and EMG sensors developed by Thalmic Labs. Through the IMU, we can get a rotation signal that represents the operator's hand [18]. In addition, using the 8 EMG sensors in MYO, we can get the sEMG signal of the arm.Fig.5 shows the structure of MYO armband:
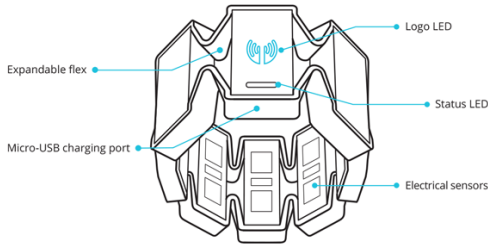

Fig. 5. MYO armband

In addition, some simple gesture recognition algorithms are built into MYO. These algorithms can use the sEMG signal to roughly recognize some gestures of the operator's hand. In this paper, the MYO armband measures the sEMG signal and the

rotation signal of the operator's hand, which helps to improve the accuracy of the data returned by Leap Motion.

III. HAND MOVEMENT TRACKING AND GESTURE IDENTIFICATION

This section mainly introduces the two methods on how to fuse the data from Leap Motion and MYO armband and consequently how to improve the tracking accuracy of the operator's hand.

As we can see from section 2, Leap Motion is a sensor consisting of a pair of cameras. Under normal circumstances, Leap Motion is a good way to track the operator's hand. However, when the operator's hand is rotated to a certain angle, mutual occlusion occurs between the fingers of the hand, which makes it impossible for Leap Motion to accurately track each finger. For example, when the operator's hand is rotated to near 90°, Leap Motion's built-in algorithm sometimes incorrectly recognizes the moving finger as another finger. As shown in Fig.6, Leap Motion mistakenly identifies the middle finger as a ring finger:
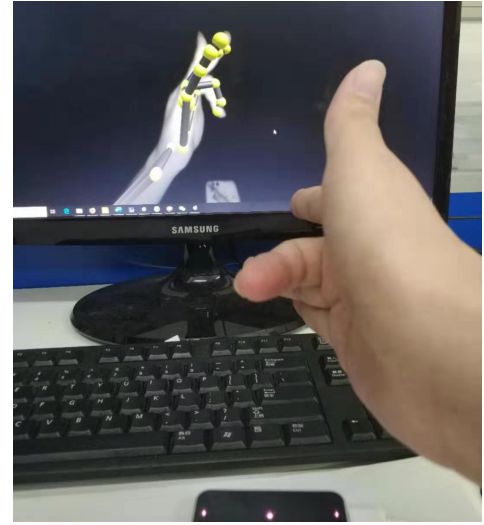
Fig. 6. An example of a mistake in identifying a finger

In order to solve the above-mentioned problems, we propose two methods of combining Leap Motion and MYO. These methods can improve the tracking accuracy of the operator's hand rotation angle and help Leap Motion correctly identify active fingers.

A. Direction information fusion between Leap Motion and MYO

Combining the direction information of Leap Motion with the rotation signal of MYO is one of the key points. Because determining an accurate palm rotation angle can be used not only to control the rotation of the robotic hand, but also to

determine whether the palm can be tracked normally by Leap Motion.

In this paper, two sensor data fusion is achieved by the Kalman filter [19]. The idea of Kalman filtering is to obtain an optimal estimate by synthesizing the results of "speculation" and "observation", and based on this best estimate, proceed to the next iteration. The following formulas represent the main steps of the Kalman filter algorithm and all the variables used are defined as in Table 1:

$$\overline{x}_t = A_t x_{t-1} \tag{1}$$

$$\overline{P}_t = A_t P_{t-1} A_t^T + Q_t \tag{2}$$

$$z_t = C_t x_t + R_t \tag{3}$$

$$K = \overline{P}_t C_t^T \left(C_t \overline{P}_t C_t^T + R_t \right)^{-1} \tag{4}$$

$$x_t = \overline{x}_t + K \left(z_t - C_t \overline{x}_t \right) \tag{5}$$

$$P_t = \left(I - K C_t \right) \overline{P}_t \tag{6}$$

In the method proposed in this paper, we use the rotation information provided by MYO to obtain the result of the "speculation" process through the state transition equation, and use the direction information detected by Leap Motion as the result of "observation". Finally, the Kalman gain coefficient K is obtained by using equation (4), and the results obtained by two different processes are combined to obtain a best estimated direction vector. This direction vector x_t represents the direction vector of the palm of the hand being tracked. A fixed space coordinate system is established with Leap Motion as the coordinate origin. In the process of only flipping the palm, the direction vector of the palm of the hand can indicate the posture of the entire palm rotation.

Among them, equations (1) and (2) represent the process of "speculation". Equation (1) is a state transition equation used to calculate the estimated value $\overline{x}_t$ at the current t time. In equation (1), At is the state transition matrix. In the data fusion process studied in this paper, we use the rotation matrix provided by MYO armband as A_t to infer the direction vector $\overline{x}_t$.

In addition, equation (3) represents the process of "observation", z_t is the palm direction vector directly observed by Leap Motion, and C_t is the matrix of the vector space where the vector x_t is mapped to the observed vector (here z_t). In the study of this paper, since the observed direction vector and the best estimation vector are the same type of information, we take C_t as the identity matrix.

Finally, equations (5) and (6) solve the optimal estimated value x_t at the current t time and its covariance matrix P_t after the above operation. After the result of the operation, the best estimate and the covariance matrix at the current time t can participate in the iterative process of the best estimate at the next moment.

Through the above process, we can get a more accurate palm posture, instead of using Leap Motion alone.

x_t	A state variable representing a direction vector with three components representing the values in the x, y, and z directions
$\overline{x}_t$	Predicted state variables obtained by state transition equation
A_t	State transition matrix, here is the rotation matrix provided by MYO
P_t	Best estimated covariance matrix at time t
$\overline{P}_t$	Prediction covariance matrix at time t
Q_t	Gaussian noise, as an approximation of the interference situation in the prediction process
z_t	Observations obtained by Leap Motion (direction vector)
C_t	A the gain matrix of the state variable mapped to the observation, here is the identity matrix
R_t	Gaussian noise, as the approximate uncertainty in the sensor measured values
K	Kalman gain coefficient

B. Improve finger recognition accuracy with sEMG signal

Although MYO has some built-in gesture recognition algorithms, these algorithms cannot accurately identify the activity of a single finger. When MYO is working, eight sEMG sensors are placed around the arm to measure the sEMG signal at eight different positions on the arm surface. When different finger is active (only one finger is active), the sEMG signals of these 8 channels are different. Thus, When the palm is at 90°, we can use the 8 channel sEMG signal provided by MYO armband to help Leap Motion identify the active finger. In this paper, we mainly study the sEMG signal when a single finger is active.

The eight channels of sEMG signals provided by MYO represent the signals obtained by MYO's eight sEMG signal sensors that surround the operator's arm. During the sampling process, the time series signals of these 8 channels are important original signals for us to identify the active fingers.

In this paper, we use Convolutional Neural Network(CNN) [20] to classify the collected sEMG signals. The schematic structure of CNN used in this paper includes 10 convolutional layers, 5 pooled layers, and 2 fully connected layers. The schematic structure of the CNN model we use is shown in Fig.7:

Input

Conv1_1: 64*8*16
Conv1_2: 64*8*32

Pool1: 32*8*32
Conv2_1: 32*8*32
Conv2_2: 32*8*32

Pool2: 16*8*32
Conv3_1: 16*8*32
Conv3_2: 16*8*32

Pool3: 8*8*32
Conv4_1: 8*8*64
Conv4_2: 8*8*64

Pool4: 4*4*64
Conv5_1: 4*4*128
Conv5_2: 4*4*128

Global average pooling: 1*128
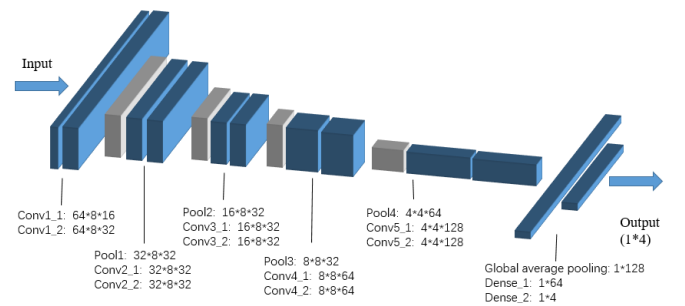Dense_1: 1*64
Dense_2: 1*4

Output
(1*4)

Fig. 7. The schematic structure of the CNN model

In order to improve the robustness of the model and avoid over-fitting as much as possible, we added the Maxpooling layer. And in the training process we use the dropout method, temporarily remove some of the neurons randomly for training during each training iteration. In this paper, we set the dropout value to 0.5. During the training process, the weight_size of the convolutional neural network is set to 48 for each update weight selection. After experimental verification, such a value can not only ensure the training speed, but also avoid the large oscillation of the learning curve.

In the work of classifying sEMG using neural networks, the sampling frequency of the sEMG signal is 200 Hz. And the sEMG signal of 8 channels is extracted as the input of the neural network with a sliding window of 320 ms duration. Finally, the neural network outputs four values ranging from 0-1 through the softmax output layer, representing the possibility of the current four fingers. Through the output of the neural network, we can determine which finger is most likely to have activity.

With the palm rotated to 90° to the Leap Motion plane, when Leap Motion detects finger activity, we can find the most reliable active finger through the trained neural network model. Finally, we compare the output of the neural network with the results of Leap Motion feedback and correct the results correctly.

IV. EXPERIMENT

In order to verify the above mentioned method, we conducted two experiments separately. The first experiment was to verify the validity of the hand direction information fusion method. The second experiment was to verify the validity of gesture recognition with the aide of EMG signal.

A. Experiment of direction information fusion

First, in the experiment of verifying the direction signal data fusion, we let the operator's hand rotate for a certain period of time, and then record the operator's hand rotation angle detected by Leap Motion and the rotation angle obtained after the Kalman filter algorithm. The results of the experiment are shown in the TABLE II.

The resulting data is averaged by an operator after 20 measurements at each particular angle. The ground truth angle data is measured using a protractor. During the experiment, we strictly ensure that the operator's palm is parallel to the viewing plane of Leap Motion in the initial state, and the arm has no obvious displacement and pitch during the rotation. We ensure the above two constraints as much as possible by observing the data from Leap Motion and MYO feedback.

From the experimental results in the TABLE II above, we can see that there is a certain deviation between the measured angle and the true angle when only Leap Motion is used. After the data fusion of MYO's rotation signal and Leap Motion, the best estimate obtained can not fully reach the true value, but the deviation from the true value is greatly reduced.

TABLE II
COMPARISON OF EXPERIMENTAL RESULTS

Actual angle	0	45°	60°	90°	120°	180°
Angle measured by Leap Motion	2.29°	37.88°	52.9°	112.7°	138.7°	177°
Angle after data fusion	2.1°	40°	63°	91°	124°	176°

However, as the angle of rotation of the palm increases, the angle of rotation of the MYO armband around the arm axis will deviate from the true angle of rotation of the palm. The farther the MYO is worn from the wrist, the greater the angular deviation produced. In order to minimize the deviation and allow MYO to acquire a valid sEMG signal, we can wear two MYO armbands at the same time, one on the wrist and the other on the arm. Despite the above method, when the angle of rotation of the arm is relatively large, the deviation generated cannot be ignored. As shown in the last column of the data in Table II, when the true angle is 180°, the data obtained after data fusion is more biased.

B. Experiment of sEMG signal for hand gesture recognition

In the experiment of using sEMG signal to improve the recognition accuracy, we used the Convolutional Neural Network(CNN) mentioned in Section 3 for classification. In order to construct a classifier that accurately identifies the activity of a single finger, we obtain the data used to train the model by collecting the operator's finger activity. Since the EMG signals of different people performing the same action are different, the data composing the data set comes from an operator's hand activity. In addition, different wearing positions will also cause differences in myoelectric signals, so each acquisition ensures that the MYO armband is always at the same position. In the course of the experiment, we asked the experimenter wearing MYO armband to only move one finger in each round of data acquisition, and each finger had to perform a signal acquisition process of up to 90s each time. In order to make the collected data enough and universal, the number of signals collected by each finger of MYO was as equal as possible, and each finger was subjected to 4 times of signal acquisition for 90 seconds.

After obtaining the long-term sequence signals of the four fingers, we intercept the original sequence signal with a sliding window of 320ms duration and 30ms steps, and finally get more short time series signals. Each time-series signal intercepted by the sliding window represents an 8-channel myoelectric signal over a short period of time. Finally, we obtained a total of 30,361 time-series signals with a length of 320 ms as a data set. Before training the model, we set the training Batch_size to 48 and the epoch to 60. The training set, the validation set and the test set are randomly selected from the data set, accounting for 60%, 20% and 20% of the data set, respectively. The loss in the learning curve is obtained by the categorical crossentropy cost function.The resulting Learning curve is in Fig.8:
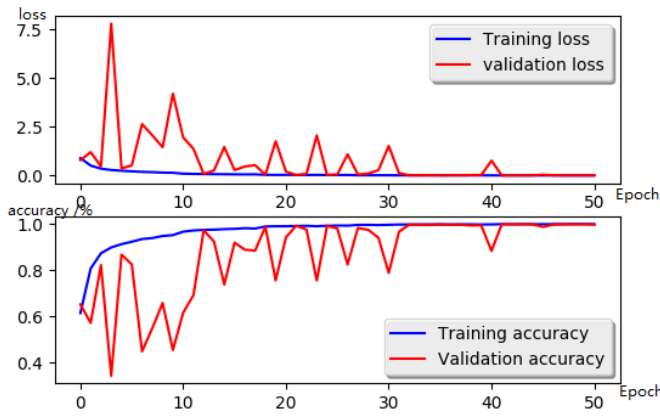
Fig. 8. The learning curve of model training

From the above figure, we can see that although there are more oscillations in the validation set curve at the beginning of the training iteration, when the number of iterations increases gradually, the validation set curve gradually converges to the training set curve, indicating that the model has certain generalization ability. Finally, the trained model is verified by the test set, and the accuracy rate is 99.58%. Therefore, we can conclude that the model trained by the above method has a better performance in solving the problem of identifying single finger activity.

V. Conclusion

In this paper, we propose two methods to fuse the data of Leap Motion and MYO sensors, which effectively improves the tracking accuracy of the operator's hand. In addition, we use the sEMG signal of MYO sensor and convolutional neural network to overcome Leap Motion's shortcomings in distinguishing the active finger quite accurately.

Our experiments show that by using the Kalman filter algorithm to combine the rotation information of MYO with the information measured by Leap Motion can improve the operator's hand tracking accuracy. In addition, from the experimental results of the classification of MYO's sEMG signal by the CNN, the 8-channel sEMG signal can distinguish the activity state of a single finger under certain conditions. Therefore, using the sEMG signal to distinguish the finger activity state, we overcome the shortcomings of Leap Motion under challenging conditions to distinguish active fingers, and improve the robustness of Leap Motion tracking.

References

[1] W. Yue, C. Jie, Y. Wang, Y. Hu, X. Rong, L. Yong, J. Zhang, and L. Qi, "Probabilistic graph based spatial assembly relation inference for programming of assembly task by demonstration," in *IEEE/RSJ International Conference on Intelligent Robots & Systems*, 2015.

[2] K. Ogawara, J. Takamatsu, H. Kimura, and K. Ikeuchi, "Generation of a task model by integrating multiple observations of human demonstrations," in *IEEE International Conference on Robotics & Automation*, 2002.

[3] J. Luo, C. Yang, N. Wang, and M. Wang, "Enhanced teleoperation performance using hybrid control and virtual fixture," *International Journal of Systems Science*, vol. 50, no. 3, pp. 451–462, 2019.

[4] A. Simorov, R. S. Otte, C. M. Kopietz, and D. Oleynikov, "Review of surgical robotics user interface: what is the best way to control robotic surgery?" *Surgical Endoscopy*, vol. 26, no. 8, pp. 2117–2125, 2012.

[5] G. Du and Z. Ping, "Markerless human-robot interface for dual robot manipulators using kinect sensor," *Robotics & Computer Integrated Manufacturing*, vol. 30, no. 2, pp. 150–159, 2014.

[6] T. V. S. N. Venna, "Real-time robot control using leap motion technology," 2015.

[7] H. Jin, L. Zhang, S. Rockel, J. Zhang, H. Ying, and J. Zhang, "A novel optical tracking based tele-control system for tabletop object manipulation tasks," in *IEEE/RSJ International Conference on Intelligent Robots & Systems*, 2015.

[8] D. Bassily, C. Georgoulas, J. Guettler, T. Linner, and T. Bock, "Intuitive and adaptive robotic arm manipulation using the leap motion controller," in *Isr/robotik ; International Symposium on Robotics*, 2014.

[9] J. C. Coelho and F. J. Verbeek, "Pointing task evaluation of leap motion controller in 3d virtual environment," *Creating the difference*, vol. 78, pp. 78–85, 2014.

[10] H. Jin, Q. Chen, Z. Chen, Y. Hu, and J. Zhang, "Multi-leapmotion sensor based demonstration for robotic refine tabletop object manipulation task," 2016.

[11] G. Du and Z. Ping, "A markerless human-robot interface using particle filter and kalman filter for dual robots," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2257–2264, 2015.

[12] A. Boyali, N. Hashimoto, and O. Matsumoto, "Hand posture and gesture recognition using myo armband and spectral collaborative representation based classification," in *2015 IEEE 4th Global Conference on Consumer Electronics (GCCE)*. IEEE, 2015, pp. 200–201.

[13] M. E. Benalcázar, A. G. Jaramillo, A. Zea, A. Páez, V. H. Andaluz *et al.*, "Hand gesture recognition using machine learning and the myo armband," in *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 1040–1044.

[14] M. Sathiyanarayanan and S. Rajan, "Myo armband for physiotherapy healthcare: A case study using gesture recognition application," in *2016 8th International Conference on Communication Systems and Networks (COMSNETS)*. IEEE, 2016, pp. 1–6.

[15] C. Li, C. Yang, J. Wan, A. S. Annamalai, and A. Cangelosi, "Teleoperation control of baxter robot using kalman filter-based sensor fusion," *Systems Science & Control Engineering*, vol. 5, no. 1, pp. 156–167, 2017.

[16] E. C. Silva, E. W. Clua, and A. A. Montenegro, "Sensor data fusion for full arm tracking using myo armband and leap motion," in *2015 14th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. IEEE, 2015, pp. 128–134.

[17] W. Frank, B. Daniel, R. Bartholom?Us, and F. Denis, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013.

[18] S. Rawat, S. Vats, and P. Kumar, "Evaluating and exploring the myo armband," in *2016 International Conference System Modeling & Advancement in Research Trends (SMART)*. IEEE, 2016, pp. 115–120.

[19] G. Welch and G. Bishop, "An introduction to the kalman filter," 1995.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.