International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.4 (2016), pp. 59-66
http://dx.doi.org/10.14257/ijmue.2016.11.4.07

Implementation of an Immersive Hand Interface Using HNMA Gesture Learning Method in Real-time

Gi-Woo Kim[1], Hye-Youn Lim[1] and Dae-Seong Kang[1]

[1]*Dong-A University, Dept. of Electronics Engineering, 37 Nakdong-daero 550 beon-gil Saha-gu, Busan, Korea*
[1]*dskang@dau.ac.kr*

Abstract

We present a method of recognizing hand gestures using RGB values, depth value, hand center coordinates, and finger counting data from Microsoft's Kinect for implementing the immersive hand interface to overcome inconvenience using HMD. First, we set the RGB values and depth range to detect the hand. This processing can improve recognition rate. Then, through double labeling, outside labeling, and inside labeling, we detect the hand without noise. Then, we use the distance vector to obtain the hand center. It also removes everything except the hand area, including removal of the wrist. After detection of the hand, we use HNMA (Multi Information-Hippocampus Neuron Modeling Algorithm) to recognize the hand gesture. This algorithm helps to improve the recognize rate. It is difficult to use the interface when using an HMD (Head Mount Display) display machine. This algorithm can make an immersive environment.

Keywords: *Kinect, Hand recognition, Interface, Gesture, Hippocampus Neuron*

1. Introduction

Gesture recognition is being actively studied in computer vision. The gesture-based interface builds the computer environment and makes the interactive environment for the engagement. In particular, when using a display device, such as the HMD in Figure 1, immersion will be greatly increased.

Figure 1. Example of Development of Hand Gesture

After using the HMD display device, when the display device is used to remove the HMD, but you can easily use the mouse or keyboard, the immersive computer interface has the inconvenience of having to issue a command to use different interface devices. There has been a lot of research into ways of overcoming these inconveniences.

For example, using the method and the artificial neural network using the sensor leap motion of hand gesture recognition is possible and is research conducted through various methods, such as gesture recognition [1-2]. The gesture recognition method uses a neural network learning algorithm also when using multi-information for the gesture recognition due to the excessive complexity of the process in real time and has difficult disadvantages.

ISSN: 1975-0080 IJMUE

In particular, using a display device, such as the HMD in the Figure, will increase immersion greatly. Processing in real time in excess of the amount of calculation when using multiple types of information for the gesture recognition method using a neural network learning algorithm gesture recognition also has the difficult disadvantages.

In this paper, using one of several methods for hand gesture recognition, Kinect offers hand gesture recognition through the hippocampal neural network algorithm.

Chapter 2 describes the relevant theories, and Section 3 shows the flowchart of the proposed algorithm and experimental results. Chapter 4 provides conclusions and future research directions.

2. Background Theory

2.1. Hand Detection Based on Kinect

Kinect is used to detect the hand in this paper. Color image information through the Kinect as well as depth information via the infrared sensor can be obtained [3-4]. Because the hands and face in the image have the same color, it also detects when a hand is detected [5-6]. However, by setting the distance to be detected by using the depth information, it is possible to detect only the regions of the hand and around it. The hand is detected to remove noise present in the inner hand area through double labeling to detect the correct hand. By using convert color and the hand center point, we can achieve finger counting [7].

2.2. Hippocampus Neuron Modeling Theory

The most important aspect is the expansion of the neural networks in the hippocampus in the short-term memory into long-term memory [9-10]. The role of the hippocampus is to remember that you have just typed information and classified information for those who need it. The necessary information stored in the neurons will be recognized if the same pattern is input. The need to compare the information stored in the input pattern by measuring the good feeling with a highly favorable given pattern is to extend the long-term memory. Measurement is to give a highly favorable result when using a method to be used in statistical pattern recognition with respect to the values that appear more significant. Favorable measurement can measure both the learning and recognition system and use the weight of the value in the middle of receiving the input to scale the short-term memory to a long-term memory pattern. Figure 2 shows a model of the structure of the hippocampus, and the functional description of each block is as follows [11-12].

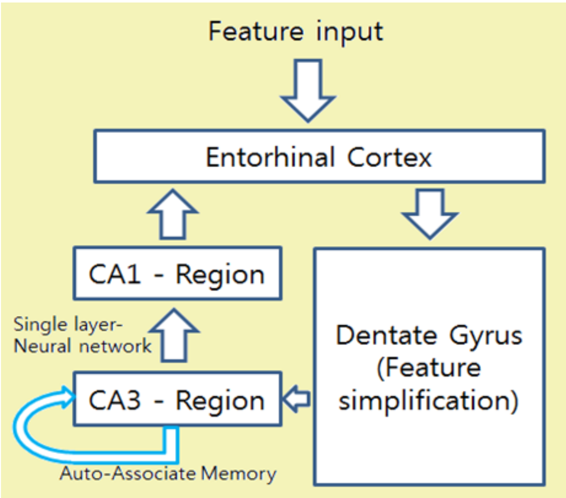


Figure 2. How the Hippocampus Works

The entorhinal cortex makes up the interface between the hippocampus and cortex, and the hippocampus shows the inputs and outputs of the neural network model [13]. The dentate gyrus teeth and directly connected structure within the olfactory cortex simplify the feature determined from the identity of the various components of the same object, the past input pattern feature in the model. In the pattern of the variation rate, if the average value exceeds the threshold value range of 1 to -1, then the identity of the binary feature. CA3 means when an event recalls that the association operation can be repeated to achieve better results, it is introduced as a concept of this circulation cycle of associative memory. The dentate gyrus and CA3 region are connected to receive the simplified information from the latter structure. In the CA3 region and the self-association, the noise feature information serves to make more sensible tidiness. The difference in the past learned characteristics to be a major feature in the input is classified as a new pattern. This study follows the Hopfield model for implementing an associative memory cycle. The CA1 region is connected to the CA3 region and indicates the end of data processing. Self-learning neural associative CA3 legislative information on the fault determines the long-term and short-term memory. In response to the output value, it does not match the learned weight and the ability to store and release a short-term memory pattern. CA1 is basically one of the Perceptron uses and has a modulated structure of Perceptron learning to be adaptive depending on the favorability.

3. Proposed Algorithm

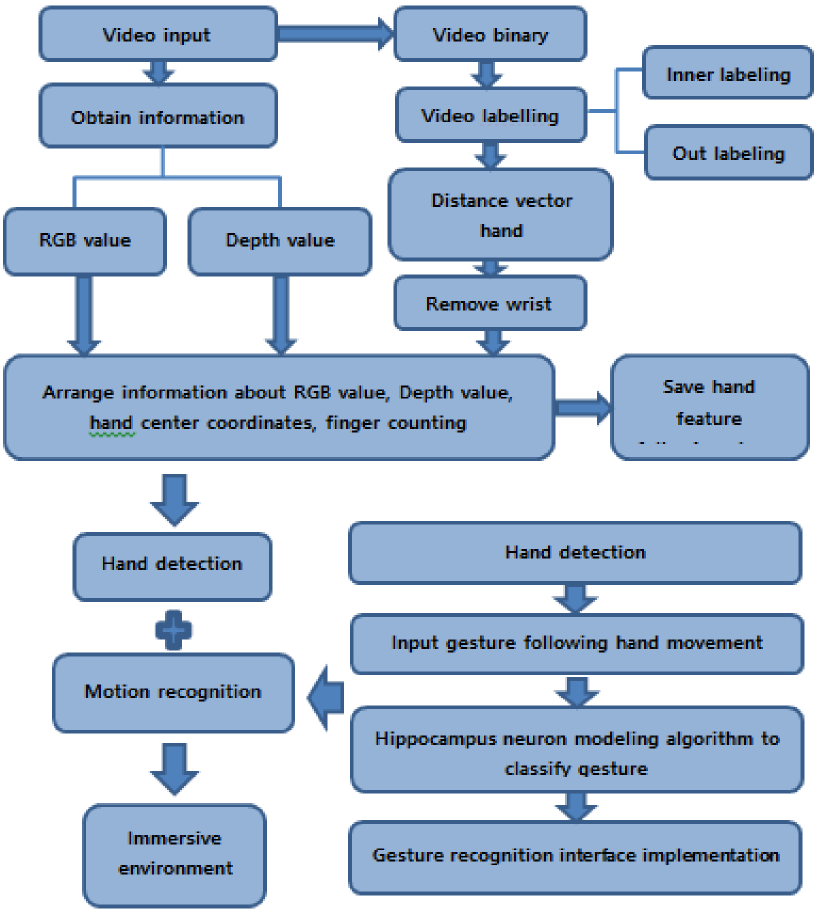


Figure 3. Proposed Method Flow Chart

The proposed algorithm flow chart is shown in Figure 3. Using the Kinect set to the value of the depth, the RGB stores the values in the gesture recognition to take advantage of the value. Save the binary video at the same time and use hand detection through the labeling. After the hand center coordinates are obtained through the streets transform vector. Hand-over-hand gives the center coordinates by removing part of the area and the bottom of the wrist. It also identified a number of fingers stretched through the colors reversed.

3.1. Hand Detection Processing

As a result of receiving video input, the depth information output of the RGB color model is shown in Figure 4.

Figure 4. Information about RGB Video and Depth

In this paper, a set of depth values in the range 50–80cm is used. It may vary Depending on the location of the Kinect depth setting, but when a hand gesture recognition is done in advance, it is always set to a value less than the face, which makes the depth values of the face position. The labeling hand result is shown in Figure 5.

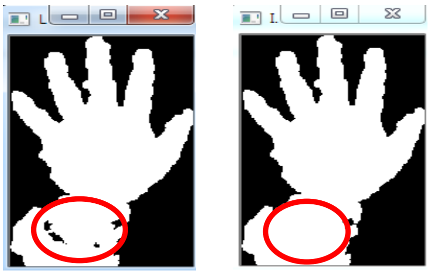


Figure 5. The Result of Labeling

In Figure 5, the hands are detected and noise is removed by discerning the external noise in the interior labeling by hand, and once more the detection area is detected through the streets of the center hands conversion filter mask.

If it then passes through a transformation vector, as shown in Figure 6, the distance can be seen with the lightest value in the center of the hand area.

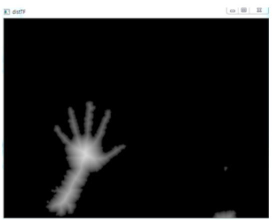


Figure 6. The Result of Distance Transform Vector

The lightest area represents the hand center coordinates. Using the center coordinates for the removal of the wrist portion is as in Figure 7.

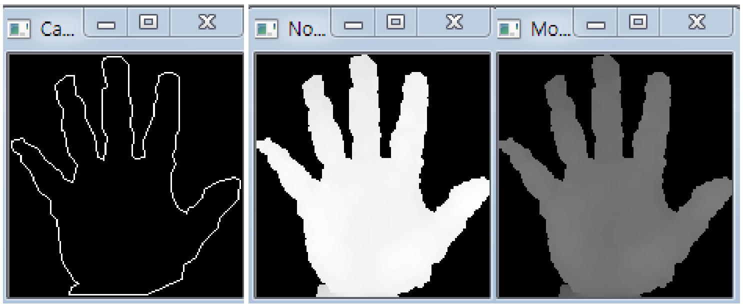


Figure 7. The Result of Detection Hand Area

Additionally, the hand detection result is used to achieve hand counting. We achieve hand counting though convert color counting. The result is shown in Figure 8.

Figure 8. The Result of Detection of the Hand Area

Figure 8 presents finger counting from 5 to 1. Until this processing, we obtain data for Table 1 and Table 2.

Table 1. Frame When Hand is Detected

Frame	RGB	Depth	Frame	RGB	Depth
Time 1	(251, 206, 177)	782	Time 11	(251, 206, 177)	791
Time 2	(251, 206, 177)	784	Time 12	(251, 206, 177)	789
Time 3	(251, 206, 177)	786	Time 13	(251, 206, 177)	789
Time 4	(251, 206, 177)	788	Time 14	(251, 206, 177)	788
Time 5	(251, 206, 177)	789	Time 15	(251, 206, 177)	788
Time 6	(251, 206, 177)	791	Time 16	(251, 206, 177)	786
Time 7	(251, 206, 177)	795	Time 17	(251, 206, 177)	786
Time 8	(251, 206, 177)	797	Time 18	(251, 206, 177)	784
Time 9	(251, 206, 177)	791	Time 19	(251, 206, 177)	784
Time 10	(251, 206, 177)	791	Time 20	(251, 206, 177)	782

Table 1 shows the RGB values and Depth values when the hand is detected. As you can see, the RGB value is same colors and the depth value ranges from 780–800.

Table 2. Hand Center Coordinates

Frame	x	y	Frame	x	y
Time 1	213	479	Time 11	229	305
Time 2	233	375	Time 12	227	303
Time 3	236	348	Time 13	225	304
Time 4	238	331	Time 14	222	308
Time 5	237	320	Time 15	232	309
Time 6	236	313	Time 16	241	310
Time 7	236	309	Time 17	250	313
Time 8	234	306	Time 18	255	314
Time 9	232	304	Time 19	261	314
Time 10	230	306	Time 20	270	316

3.2. Learning Algorithm Setting

To adjust the hippocampus neuron model theory, when we obtain information without hand detection, it erases the information on all things. We set the learning algorithm as in Table 3. We made eight gestures in Table 3.

Table 3. Hand Gesture Setting

Motion	RGB	Depth	Hand coordinate	Hand counting change
Back	(251,206,177)	780–800	Right -> left	1
Forward	(251,206,177)	780–800	Left -> right	1
Left mouse click	(251,206,177)	780–800	Mouse coordinate	2->1
Right mouse click	(251,206,177)	780–800	Mouse coordinate	2->3
Cut	(251,206,177)	780–800	Mouse coordinate	3
Paste	(251,206,177)	780–800	Mouse coordinate	3
Execute program size-up	(251,206,177)	780–800	Mouse coordinate	0->5
Execute program size-down	(251,206,177)	780–800	Mouse coordinate	5->0

3.3. Experimental Results Disseminate

The final experiment results are shown in Table 4.

Table 4. Detected Time and Recognition Rate

Motion	Detected time	Recognition rate
Back	0.89sec	97%
Forward	0.85sec	96.2%
Left mouse click	0.86sec	92%
Right mouse click	0.89sec	92%
Cut	0.92sec	95%
Paste	0.94sec	95%
Execute program size-up	0.91sec	96%
Execute program size-down	0.92sec	97%

We can see the detection time and recognition rate in Table 4. The recognition average is 95%. The detection time is less than 1sec. This means we can use eight gestures in real time.

4. Conclusion and Future Research

The Kinect is used to obtain RGB color information, the degree of depth detection, hand over the double center distance labeling, and converting vector of the image based on this. It could also be implemented in real time with a high recognition rate by introducing the hippocampal neural network interface concept based on this information. Additionally, this method can overcome HMD inconvenience. We will try to find additional ideas for a new event handler to implement additional interfaces in the event handler of eight degrees regarding the direction of future research.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0011735).

References

[1] N. H. Kang, J. Seo, S. Chae, D. C. Kim and T. D. Han, "Unistroke Korean Character Input Method using Leap Motion", The HCI Society of Korea, **(2014)**.

[2] L. K. Soo, P. S. Chun, L. H. Min and J. Y. Hoon, "The Study on the Forecasting of Runoff Applied the B.P. Algorithm of the Artificial Neural Network in the Young-San River", Korean Society of Civil Engineers., vol. 29, no.5, **(2007)**, pp. 679-688.

[3] H. Park, J. Choi, J. I. Park and K. S. Moon, "A Study on Hand Region Detection for Kinect-Based Hand Shape Recognition", the Korean Society of Broad Engineers, **(2013)**, pp. 393-400.

[4] M. Alsheakhali, A. Skaik, M. Aldahdouh and M. Alhelou, "Hand Gesture Recognition System", Computer Engineering Department, The Islamic University of Gaza Gaza Strip, Palestine, vol. 60, no. 11, **(2011)**, pp. 3592-3607.

[5] N. H. Dardas and N. D. Georganas, "Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques", IEEE Transactions on Instrumentation and Measurement.

[6] G. Bailly, R. Walter, J. Müller, T. Ning and E. Lecolinet, "Comparing free hand menu techniques for distant displays using linear, marking and finger-count menus", Human-Computer Interaction–INTERACT 2011. Springer Berlin Heidelberg, **(2011)**, pp. 248-262.

[7] J. L. Raheja, A. Chaudhary and K. Singal, "Tracking of Fingertips and Centres of Palm using KINECT", Proceedings of the 3rd IEEE International Conference on Computational Intelligence, Modelling and Simulation, Malaysia, **(2011)**, pp. 248-252.

[8] I. G. Oh and J. C. Namgung. "Facial Expressions Recognition using SIFT and Hippocampal Neuron Networks", Korean Institute of information Technology, **(2012)**, pp. 44-52.

[9] J. H. Kim, L. J. Joung and D. S. Kang, "An Object-Based Image Retrieval Technique using Brain Modeling of Hippocampus and Synapse", ICICS, **(2005)**.

[10] S. H. Jeong and D. S. Kang, "The Development of Efficient Multimedia Retrieval System of the Object-Based using the Hippocampal Neural Network", The Institute of Electronics Engineers of Korea, vol. 43, no. 2, **(2006)**, pp. 57-64.

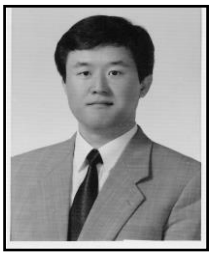
[11] I. G. Oh, H. J. Kim and J. C. Namgung, "Real Time Facial Recognition using SURF and Hippocampal Neuron Networks", Journal of KIIT, vol. 10, no. 1, **(2012)**, pp. 177-185.

[12] I. G. Oh and J. C. Namgung, "Facial Expressions Recognition using SIFT and Hippocampal Neuron Networks", Journal of KIIT, vol. 10, no. 2, **(2012)**, pp. 44-52.

Authors

Gi-Woo Kim, he received a B.S degree from Dong-A University, Busan, Korea, in 2014. He is in department of electronic engineering, Dong-A universty now. His research interests image processing.

Hye-Youn Lim, she received the BS degree in computation engineering from Daegu University, Daegu, Korea, in 2005 and her M.S. and Ph.D. degrees in Electronics Engineering from Dong-A University, Korea, in 2010 and 2013, respectively. Currently, she is anassistant professor in Department of Electronics Engineering, Dong-A University. Her research interests include image processing and object tracking



Dae-Seong Kang, he received a B.S. degree from Kyungpook National University, Daegu, Korea, in 1984, M.S. degree and D.Sc. degree in electrical engineering from Texas A&M University, in 1991 and 1994, respectively. He is currently full professor of the Department of Electronic Engineering, Dona-A University, Busan, Korea. His research interests are image processing and compression.