International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.7 (2015), pp.229-236
http://dx.doi.org/10.14257/ijmue.2015.10.7.24

3D Wrist Gesture Recognition using Three-Dimensional Directional Codes and HMM

Young-ju Kim[1]

[1]*Department of Computer Engineering, School of Engineering, Silla University, Busan 617-736, Korea*
yjkim@silla.ac.kr

Abstract

This paper proposes a novel 3D wrist gesture recognition method as the human-machine interface for mobile robot control. A sequence of depth maps was extracted from Kinect sensor on wrist gestures of an operator and a wrist gesture was specified by a sequence of 24-directional codes, which represents 3D motions with feature values of a single type, including directional and depth information. HMM algorithm was used to recognize wrist gestures, in which a fully connected hidden Markov model was designed and trained by Viterbi method and Baum-Welch method and the forward-backward procedure was used for the likelihood evaluation. The performance evaluation considering various factors showed above 96% of recognition rate and higher recognition rate on wrist gestures with complex motions.

Keywords: *Pattern Recognition, 3D Gesture Recognition, HMM, 24-diectional codes, Kinect, Human-Machine Interface*

1. Introduction

Recently, various methods for robot control have been researched due to the development of computer technology [1-2]. To control mobile robots using only wrist gestures, this paper proposed and evaluated a novel 3D gesture recognition method among image processing-based control ones.

The proposed method processes depth-map sequences grabbed by Kinect sensor [3] being a 3D image input device and extracts 3-dimensional feature values, so-called 24-directional codes, to specify wrist gestures. In previous works [4-5], 3-dimensional feature values were extracted by processing images of the front and lateral sides or by adding depth data to 2-dimensional direction values. These feature values were of various types and a lot in number, causing recognition algorithms to be highly complex. The proposed method extracts 3D feature angles characterizing wrist motions using a sequence of depth maps and defines 8-directional codes in x-y plane, and by combining the codes with the variation of z-coordinates, generates 24-directional code patterns of a single type as 3-dimensional feature values. The proposed method applied HMM (Hidden Markov Model) algorithm [6-7] to recognize wrist gestures. A fully-connected hidden Markov model was designed having 24-directional codes as hidden states and the forward-backward procedure [7] and Viterbi method [8] was used to evaluate the HMM model and recognize wrist gestures. The Baum-Welch method [7] was used as the training algorithm for the normalization of likelihood values of the HMM model.

In the evaluation experiment with 420 training samples and 210 test samples for 7 wrist gestures of control, the proposed method showed the recognition rate of above 96% and higher recognition rate on wrist gestures with complex motions.

ISSN: 1975-0080 IJMUE

2. 3D Feature Values for Wrist Gestures

2.1 Definition of Wrist Gestures for Control

Considering the usage of Kinect sensor as an input device, this paper defined 7 wrist gesture patterns to control the movement of a mobile robot, as shown in Figure. 1. Each wrist gesture comprises of a series of continuous motions in left-right, top-bottom and forward-backward directions and needs 3-dimensional motion analysis to improve the probability of recognition success.

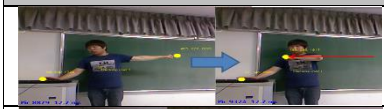
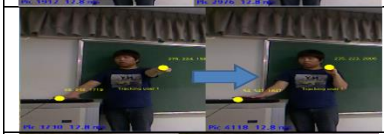
Wrist Gestures	Control Commands
	Left Rotation
	Right Rotation
	Forward
	Backward
	Stop
	Automatic
	Manual

Figure 1. Wrist Gesture Patterns for Mobile Robot Control

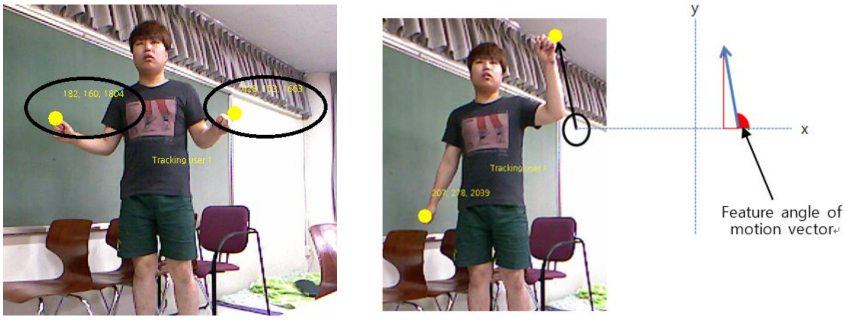
2.2 24-Diectional Codes for Wrist Gestures

At first, as shown in Figure. 2(a), using skeleton joint data read from Kinect sensor, and motion data of two wrists, left and right ones were extracted from a sequence of depth maps as a set of coordinates of 3-dimensional space.

Using only x and y coordinates from motion data extracted from successive depth maps, 2-dimensional motion vectors were generated, which represent a wrist motion in x-y plane, as shown Figure. 2(b). In this process, the nonlinearity of continuous x-y coordinates was filtered as noises, generating linear vectors.

Applying Eq. (1) to start and end coordinates of motion vectors, feature angles between motion vectors and x-axis in x-y plane were calculated. Using feature angles, 8-dictional codes in x-y plane were determined corresponding to motion vectors, as shown in Figure. 3

$$\theta = \arctan\left((y_2 - y_1) / |x_2 - x_1| \right) \tag{1}$$



(a) (b)

Figure 2. Extraction of Two Wrists and Wrist Motions: (a) Two Wrist Extracted and (b) a Motion Vector and a Feature Angle of Wrist Motion

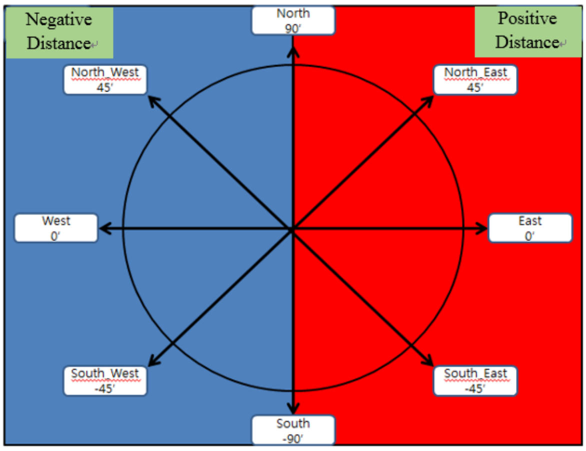


Figure 3. 8-Diectional Codes in X-Y Plane Corresponding to Wrist Motions

The proposed method calculated z-coordinate variation of points included in motion vectors and combined with 8-diectional codes, lastly defining 24-directional codes as 3-dimentional feature values of wrist motions, as shown in Table 1. Therefore, a wrist gesture comprising of consecutive wrist motions is able to be specified by a sequence of 24-directional codes, being used in recognition algorithms.

Table 1. 24-Diectional Codes as 3-D Feature Values for Wrist Motions

8-directional angles	Directional Codes		
	negative Z-variation	no Z-variation	positive Z-variation
0˚	Center _Backward	Center	Center _Forward
67.5˚ ~ 90˚	North _Backward	North	North _Forward
22.5˚ ~ 67.5˚	North_East _Backward	North_East	North_East _Forward
22.5˚ ~ -22.5˚	East _Backward	East	East _Forward
-22.5˚ ~ -67.5˚	South_East _Backward	South_East	South_East _Forward
-67.5˚ ~ -90˚	South _Backward	South	South _Forward
-22.5˚ ~ -67.5˚	South_West	South_West	South_West

	_Backward		_Forward
22.5° ~-22.5°	West _Backward	West	West _Forward
22.5° ~ 67.5°	North_West _Backward	North_West	North_West _Forward

3. Recognition of Wrist Gestures Using HMM

3.1 Design of Hidden Markov Model

As described in the previous chapter, a wrist gesture is represented as a sequence of 24-directional codes of a spatio-temporal pattern type, and this paper applies HMM algorithms to classify and recognize wrist gestures. A Hidden Markov Model [6] consists of N states, each of which is associated with a set of M possible observable states, and includes the model parameter set $\lambda = (A, B, \pi)$ defined in Eq. (2).

$$\lambda = (A, B, \pi) \tag{2}$$

where *A* donotes the *N*x*N* state-transition probability matrix, which specifies the probability that the state will transit from state *i* to state *j*. B denotes the observation probability matrix which specifies the observable state *i* will be generated at state *j* and at time *t*, and π is the initial state probability which state *i* may be the initial state.

This paper defined the fully-connected hidden Markov model for wrist gestures by specifying all 24-directional codes as hidden states and a subset of 24-direcional codes as observable states, as shown in Figure. 4.

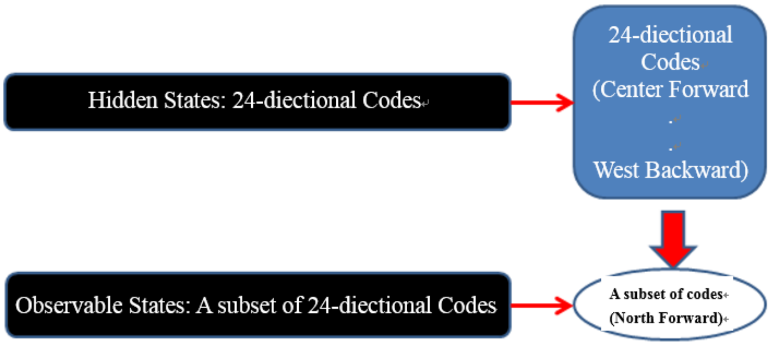


Figure 4. Hidden States and Observable States in HMM Model Design

3.2 Training of HMM

In training the hidden Markov model to compute the model parameters *A*, *B* and π, probability (or likelihood) evaluation was executed by the forward-backward procedure[7], which computes the output probability $P(O \mid \lambda)$ with which the HMM will generate on output observable state sequence *O* given an HMM parameter λ, as shown in Figure. 5. Viterbi algorithm [8] was used to find the optimal state sequences $q = (q_1, q_2, \cdots q_T)$ given the observable state sequence *O* and the HMM parameter λ in order to maximize $P(q \mid O, \lambda)$.

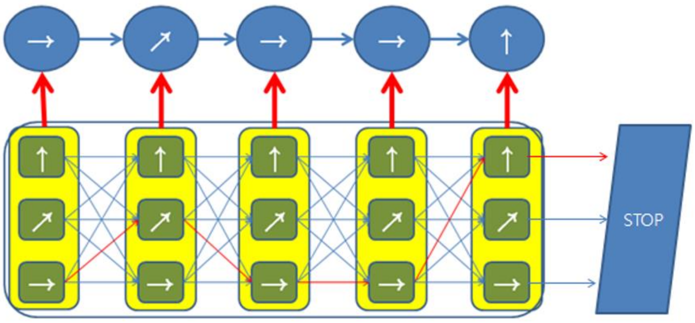


Figure. 5. Example of Path and Probability Evaluation in the Proposed HMM Model

This paper made training data by sampling a sequence of depth maps on 7 wrist gesture patterns by 3 persons in various distances, between 1.2 and 3.5 meters, from Kinect sensor being at the height of 1.1 meters, and the training data contains 60 samples for each wrist gesture and a total of 420 gesture samples. To estimate an optimal model parameter based on the training data samples, Baum-Welch method[7] was used to choose the maximum likelihood model parameter set $\lambda = (A, B, \pi)$ such that its likelihood function $P(O \mid \lambda)$ is locally maximized using an iterative procedure.

3.3 Evaluation and Recognition of HMM

Given a sequence of 24-directional codes specified by a wrist gesture from input of Kinect sensor, the code sequence is fed into the HMM model module to evaluate likelihood values between the given code sequence and trained HMM data and determine the HMM data with a maximum log likelihood value as recognized gesture.

4. Performance Evaluation

Recognition experiments were performed in the same condition of the sampling of training data for the performance evaluation of the proposed method, and the test data comprising of 30 samples for each wrist gesture and totally 210 samples was used.

Table 2 shows the rate of recognition success for each of 7 wrist gestures, and it can be known that the recognition rate of relatively dynamic gestures such as ‘Automatic’ and ‘Manual’ is higher than static gestures such as ‘Left Rotation’ and ‘Forward’.

Table 3 shows the gesture vote frequency generated in the HMM recognition experiment and informs that static gestures comprising of repeated similar motions are wrongly recognized by each other because static gestures generate similar directional code sequences with very little difference in order. Therefore, in the design of wrist gesture patterns for robot control using the proposed method, it must be considered that wrist gestures include some definitely different directional motions for the improvement of recognition performance.

Table 2. Result of Recognition Experiments Using 210 Samples

7’s Wrist Gestures	Recognition Rate
Left Rotation	96.7%
Right Rotation	93.3%
Forward	93.3%
Backward	93.3%
Stop	96.7%
Automatic	100%
Manual	100%

Table 3. Gesture Vote Frequency in Recognition Experiments Using 210 Samples

	Left Rotation	Right Rotation	Forward	Back-ward	Stop	Auto	Manual
Left Rotation	29	0	0	1	0	0	0
Right_ Rotation	0	28	0	2	0	0	0
Forward	0	0	26	4	0	0	0
Backward	0	0	0	28	0	0	0
Stop	1	0	0	0	29	0	0
Automatic	0	0	0	0	0	30	0
Manual	0	0	0	0	0	0	30

5. Conclusions

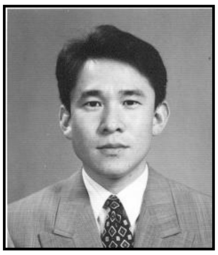
This paper proposed and evaluated a novel 3D wrist gesture recognition method as the human-machine interface for mobile robot control. Using Kinect sensor as an input device, a sequence of depth maps is extracted on wrist gestures of an operator and a wrist gesture is specified by a 24-directional code sequence of a single type being different with previous works, bringing down the complexity of recognition algorithms of 3D gestures. For the recognition of wrist gestures, a Hidden Markov model was designed and trained by using related algorithms such as the forward-backward procedure, Viterbi method and Baum-Welch method. The performance evaluation using 420 training samples and 210 test samples showed above 96% of recognition rate and higher recognition rate on wrist gestures with complex motions.

In future, for the improvement of recognition performance, various hidden Markov models will be designed and evaluated to find an efficient hidden Markov model for 24-directional code sequences and the work on training algorithms for a selected hidden Markov model will be performed continuously.

References

[1] M. W. Spong and M. Fujita, “Control in Robotic, T. Samad and A. M. Annaswamy (Eds.), the Impact of Control Technology,” IEEE Control Systems Society, **(2011)**, pp. 49-56.
[2] Wu Y. and Thomas S. H., “Vision-based gesture recognition: A review, Gesture-based communication in human-computer interaction,” Springer Berlin Heidelberg, **(1999)**, pp.103-115.
[3] “Microsoft Kinect,” http://en.wikipedia.org/wiki/Kinect
[4] Malima A., Erol O. and Müjdat Ç., “A fast algorithm for vision-based hand gesture recognition for robot control,” Signal Processing and Communications Applications, 2006 IEEE 14th, **(2006)**, pp.1-4.
[5] Jong W. K., Dong J. S. and Dong S. J., “A Study On the control Method of 3-Dimensional Space Application using KINECT System,” IJCSNS International Journal of Computer Science and Network Security, vol.11 no.9, **(2011)**, pp.55-59.
[6] J. Yamato, H. Ohya and K. Ishii, “Recognizing Human Action in Time-Sequential Images Using Hidden Markov Model,” Proc. 1992 IEEE Conf. on Computer Vision and Pattern Recognition, **(1992)**, pp.279-385.
[7] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” Proc. of The IEEE, vol. 77 no. 2, **(1989)**, pp.257-285.
[8] G. D. Forney, “The viterbi algorithm,” Proc. of The IEEE, vol. 48, **(1973)**, pp.268-278.
[9] Hyeon K. L. and Jin H. K., “An HMM-based threshold model approach for gesture recognition,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.21 no.10, **(1999)**, pp. 961-973.

Author



Young-ju, Kim, Department of Computer Engineering, School of Engineering, Silla University, 140 Baegyang-daero 700 beon-gil, Sasanggu, Busan, 617-736, Korea, E-mail: yjkim@silla.ac.kr.