

Motion Classification using IMU for Human-Robot Interaction

Kawroong Saktaweeikulkit and Thavida Maneewarn

Institute of Field Robotics, King Mongkut's University of Technology Thonburi, Bangkok, Thailand
(Tel : +66-2-470-9339; E-mail: kawroong_s@hotmail.com, praew@fibo.kmutt.ac.th)

Abstract: This paper proposed the method for identifying motion of the robot using signals from an Inertial Measurement Unit (IMU). When human interacted with the robot, the motion of the robot needed to be classified in order to create the suitable emotions of the robot during the human-robot interaction process. The output data from the IMU were classified by applying heuristic conditions to the high and low frequency data. Wavelet transform was added in the pre-processing step for finding the starting point of the motion state. We found that most of our tested actions could be correctly classified using our proposed method with 100% accuracy.

Keywords: Motion Classification, Inertial Measurement Unit (IMU)

1. INTRODUCTION

This research focused on the motion classification of a humanoid robot by using an Inertial Measurement Unit (IMU). In our laboratory, IMU was used in terrain classification for an autonomous vehicle [4]. For this reason, the IMU was chosen to be used on the humanoid robot in this research. At the present, emotional interaction between human and robot has increasingly become subject of interest. For example, Seong-Yong Koo et al [1] developed the bear robot, Gomy that could express its emotion when interacted with people. Gomy had an accelerometer, force sensitive resistors and a microphone inside its body for classifying its emotional state. Additionally Jamy Li et al [2] studied gesture generation and gesture interpretation in order to allow human to understand the messages from the robot. Li also studied about the contexts that affected the communication between human and robot. Moreover, there is a study about the robot design rubrics for interacting with children in the family context [3].

2. CONCEPT

Motion classification is important for identifying the states of the robot and the appropriate emotion corresponding to these states for the robot. For understanding the robot's state, we used the heuristic method for classifying the robot's motion from an IMU data. The data from IMU in this system consisted of roll (angular velocity around x-axis), pitch (angular velocity around y-axis), yaw (angular velocity around z-axis), accX (x-axis acceleration), accY (y-axis acceleration) and accZ (z-axis acceleration). These data were transformed into angles, thetaZY (angle around x-axis) and thetaXY (angle around z-axis).

The concept of the state definition system was shown in Fig. 1. The system was divided into 3 main parts. There were data collection, pre-processing and classification.

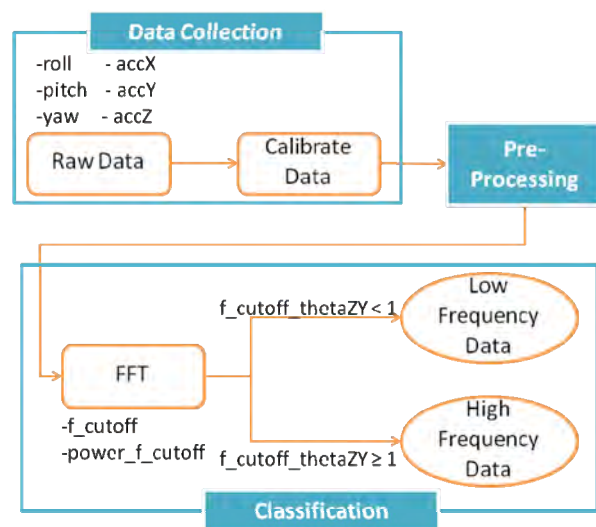


Fig. 1 The concept of the state definition system

2.1 Data Collection

The data were collected from an IMU at 171 Hz when the humanoid robot was “lifted up down”, “picked up and swayed”, “turned around”, “shaken”, “pushed” and “lifted up”. The data that consisted of 3-axis acceleration and 3-axis angular velocity were collected 2 times for each kind of actions.

2.2 Pre-Processing

The data were pre-processed using two methods for comparison: without wavelet transform and with wavelet transform, as shown in Fig. 2 and Fig. 3, respectively.

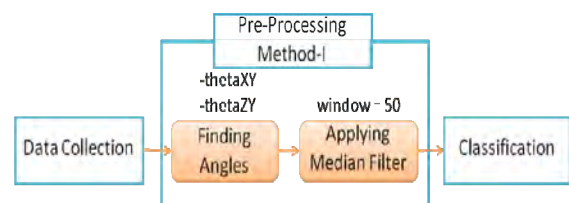


Fig. 2 Pre-Processing Method-I

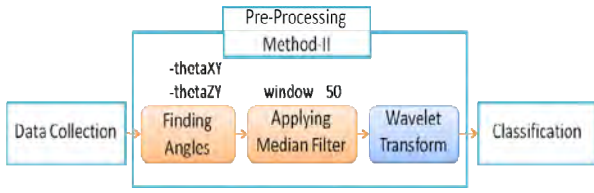


Fig. 3 Pre-Processing Method-II

In both methods, we divided the pre-processing procedure into 2 steps which were: finding angles and applying median filter.

1. Finding angles

This step was to calculate angles from acceleration data. There were thetaZY (the angle around x-axis) and thetaXY (the angle around z-axis). These angles showed the relation between the robot and the vector that perpendicular with the earth (refer to y-axis acceleration in Fig. 7). These angles could be calculated from the following relations.

$$\theta_{ZY} = \tan^{-1} \frac{accZ}{-accY}, \quad (1)$$

$$\theta_{XY} = \tan^{-1} \frac{accX}{-accY}. \quad (2)$$

2. Applying median filter

The median filter is a non-linear digital filter that can remove high frequency noise from data, so that the characteristic of the data can be shown more clearly. In this research, the median filter with a window size of 50 was used.

In the second method, we added the wavelet transform after applying the median filter in order to find the location that the motion starts. ThetaZY was transformed into low frequency component (approximations) and high frequency component (details) at Daubechies1 level 2. The details at level 1 and the details at level 2 were multiplied in order to reduce noises in signals.

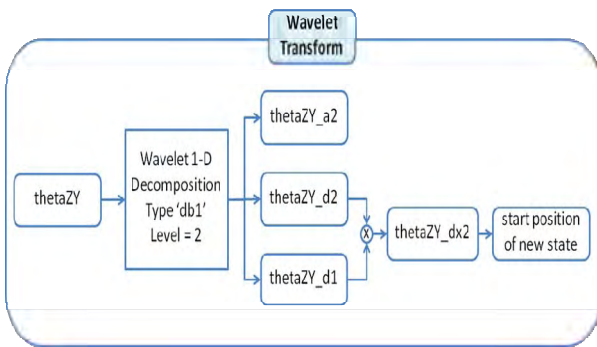


Fig. 4 The process of wavelet transform

The location that the new motion starts can be detected from the change in the signal after the wavelet

transform is applied. When the change of the signal's slope is greater than a given threshold value, it is considered the start position of the new state.

2.3 Classification

In the classification part, first we determined the appropriate cutoff frequency that can separate the signal into low and high frequency signals. Then the rule-based decisioning logic was applied according to the cutoff frequency and the characteristic of low/high frequency data. In order to explain this procedure, we divided the classification into 2 steps which were:

1. Finding the cutoff frequency by Fourier Transform

Fourier transform transformed data in time domain into frequency domain. The cutoff frequency was the frequency that had maximum power except at zero frequency. The power of each frequency could be calculated by following equation,

$$Signal_power@f = \frac{|signal_f^2|}{n}, \quad (3)$$

where n is the length of signal.

2. Classifying state of the robot

The dominant features were used in classifying state of the robot. They were applied for creating the conditions to verify each state.

First of all, the cutoff frequency of thetaZY was executed to separate the low frequency data and high frequency data. The data that contains the cutoff frequency of thetaZY less than 1 Hz would be the low frequency data. The data that contains the cutoff frequency of thetaZY larger than 1 Hz would be the high frequency data. (As shown in Fig. 1)

Condition1 Low frequency data

1. After the low frequency data had separated, the cutoff frequencies (f_c) and the power at cutoff frequencies (P_{f_c}) of all IMU's data were calculated.

2. If $P_{f_{c_roll}} == \max(P_{f_{c_roll}}, P_{f_{c_pitch}}, P_{f_{c_yaw}})$ and

$$P_{f_{c_accY}} == \max(P_{f_{c_accX}}, P_{f_{c_accY}}, P_{f_{c_accZ}}),$$

$$\text{If } P_{f_{c_roll}} > 1000,$$

The state is "lifted up down".

Else,

The state is "other poses".

Else,

$$angle_{ZY} = \text{average}(\theta_{ZY}).$$

If $angle_{ZY} \geq 100$ degrees,

The state is "turned around".

Elif $15 \text{ degrees} < angle_{ZY} \leq 40 \text{ degrees}$,

The state is "lifted up".

Elif $-60 \text{ degrees} < angle_{ZY} \leq -40 \text{ degrees}$,

The state is "picked up and swayed".

Else,

The state is "other poses".

3. EXPERIMENT

3.1 Experiment Methodology

The humanoid robot - KM4 that is one of the humanoid robots in KM-series of Team KMUTT designed for playing soccer in Robocup (kid size) soccer competition was used in this experiment. The 6 degrees of freedom IMU was mounted on KM4's back as shown in Fig. 7. The data were collected when he was “lifted up down”, “picked up and swayed”, “turned around”, “shaken”, “pushed” and “lifted up” as shown in Fig. 8.

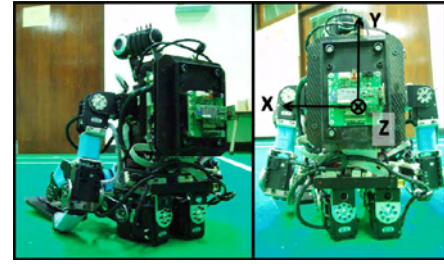


Fig. 7 The experimental humanoid robot – KM4 and 6-DOF IMU on his back

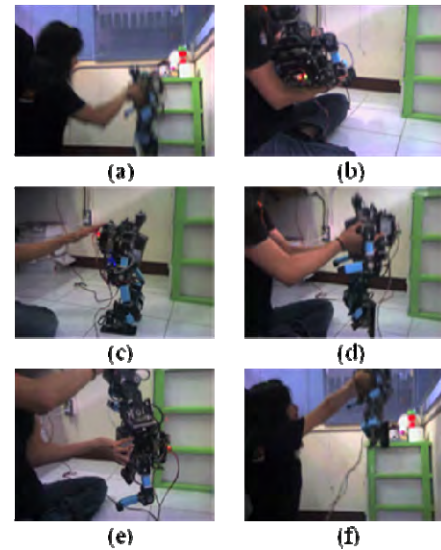


Fig. 8 The six kinds of actions in the experiment.
(a) lifted up down, (b) picked up and swayed, (c) pushed, (d) shaken, (e) turned around, (f) lifted up

3.2 Experiment Result

In this experiment, signals of all actions: “lifted up down”, “picked up and swayed”, “turned around”, “shaken”, “pushed” and “lifted up” (2 sets per each action) were used to test the classification system. Input of the system consisted of the data shown testing state and the data shown other state as shown in Figs. 9~10. The signals were divided into set of data that had 500 data per set in order to feed them into state defining system (as shown in Fig. 1) with pre-processing method-I.

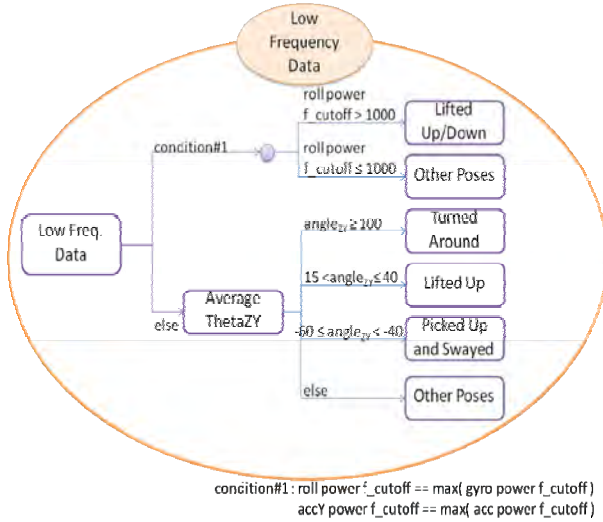


Fig. 5 Motion classifying conditions for the low frequency data

Condition2 High frequency data

1. For the high frequency data, the cutoff frequencies (f_c) and the power at cutoff frequencies (P_{f_c}) of all IMU's data were calculated.

2. If $P_{f_{c_roll}} == \max(P_{f_{c_roll}}, P_{f_{c_pitch}}, P_{f_{c_yaw}})$ and

$$P_{f_{c_accZ}} == \max(P_{f_{c_accX}}, P_{f_{c_accY}}, P_{f_{c_accZ}}),$$

If $2\text{Hz} < f_{c_roll} < 4\text{Hz}$ and $2\text{Hz} < f_{c_accZ} < 4\text{Hz}$,

The state is “shaken”.

Elif $1\text{Hz} < f_{c_roll} \leq 2\text{Hz}$ and $1\text{Hz} < f_{c_accZ} \leq 2\text{Hz}$,

The state is “pushed”.

Else,

The state is “other poses”.

Else,

The state is “other poses”.

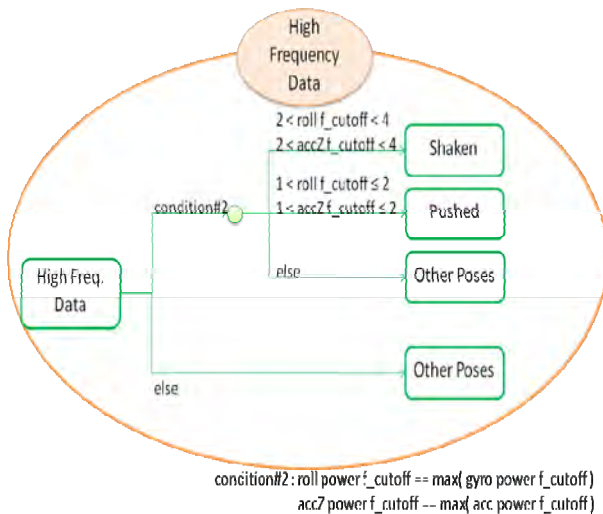


Fig. 6 Motion classifying conditions for the high frequency data

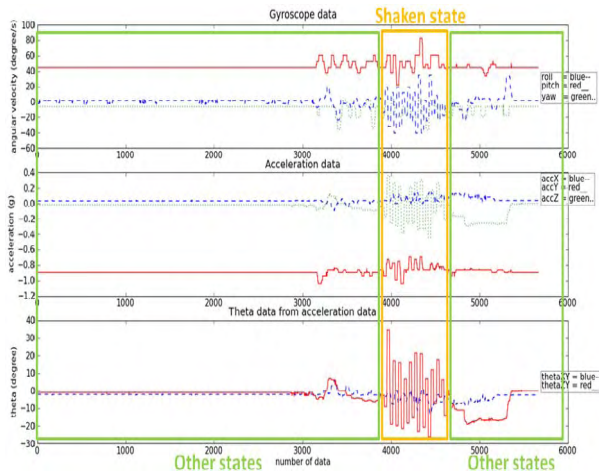


Fig. 9 Example signal using in testing the system from shaken action

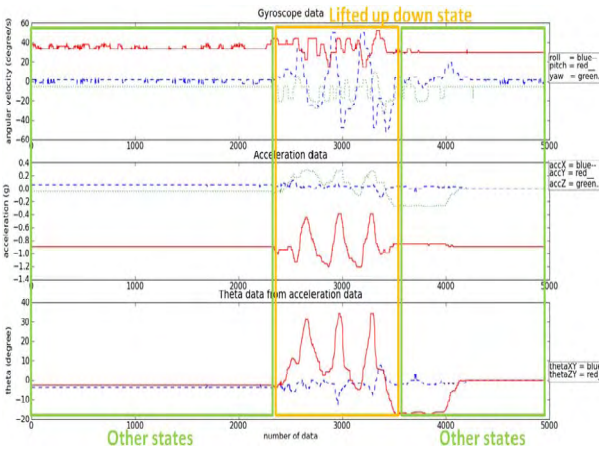


Fig. 10 Example signal using in testing the system from lifted up down action

Table 1 The resulting table of state defining system

No.	Test Action	Number of Data Sets	Result		Other Action	Number of Data Sets	Result	
			Correct (times)	Correct (%)			Correct (times)	Correct (%)
1	Lifted Up Down	2	2	100.00	Other	8	8	100.00
2	Lifted Up Down	2	2	100.00	Other	7	7	100.00
3	Picked Up and Swayed	3	3	100.00	Other	9	9	100.00
4	Picked Up and Swayed	4	4	100.00	Other	11	11	100.00
5	Pushed	4	4	100.00	Other	5	5	100.00
6	Pushed	3	3	100.00	Other	7	7	100.00
7	Shaken	1	1	100.00	Other	11	11	100.00
8	Shaken	2	2	100.00	Other	9	9	100.00
9	Turned Around	2	2	100.00	Other	10	10	100.00
10	Turned Around	2	2	100.00	Other	10	10	100.00
11	Lifted Up	6	5	83.33	Other	8	8	100.00
12	Lifted Up	6	4	66.67	Other	8	8	100.00

From the experiment, after each 500 data had fed into the system, the results were shown in Table 1. The Table 1 showed that the most actions (“lifted up down”, “picked up and swayed”, “turned around”, “shaken” and “pushed”) could be correctly classified by the system

with 100% accuracy. Only the “lifted up” action was identified with 75% accuracy which was slightly less than other actions.

From the above experiment, we found that the system with pre-processing method-I which was fed with 500 pre-recorded data (500 data per set) as input performed quiet well, however it could make a wrong decision when applied in real-time because of a mis-determination of the state transition. Therefore, the wavelet transform was applied as described in the pre-processing method-II (as shown in Fig. 3) in order to locate the start position of the observed state.

According to Fig. 4, the multiplication of the details at level 1 and the details at level 2 of thetaZY could reduce noises in the signals as Figs. 11~12.

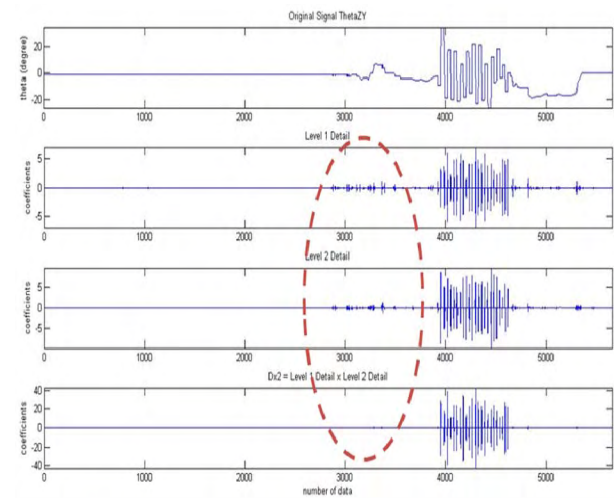


Fig. 11 Example signal after from wavelet transform in shaken action

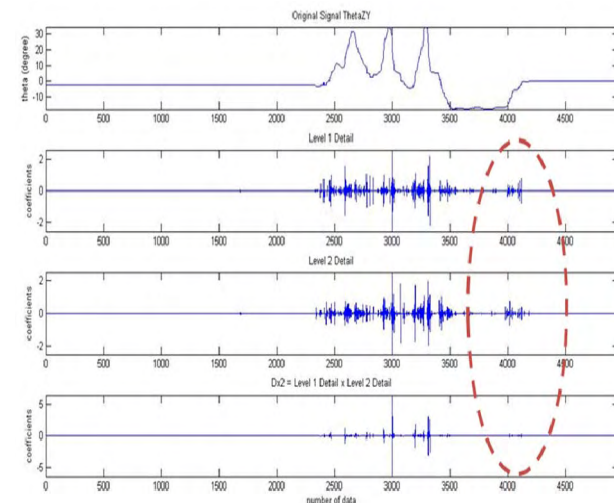


Fig. 12 Example signal after from wavelet transform in lifted up down action

From the experiment, we found that wavelet transform was able to locate the position of state transition correctly as shown in Figs. 13~14.

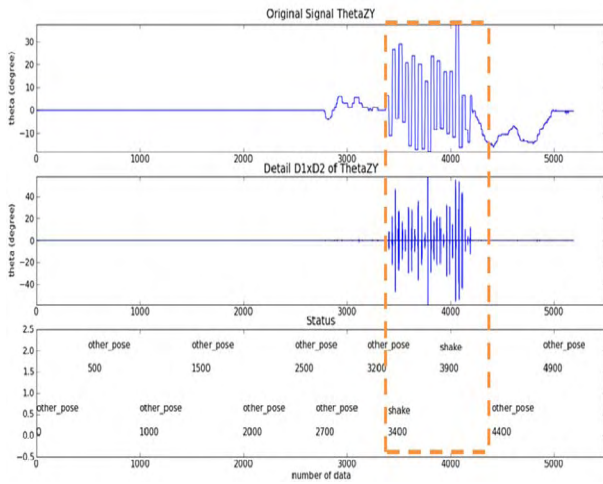


Fig. 13 Example result of locating state transition in the shaken action

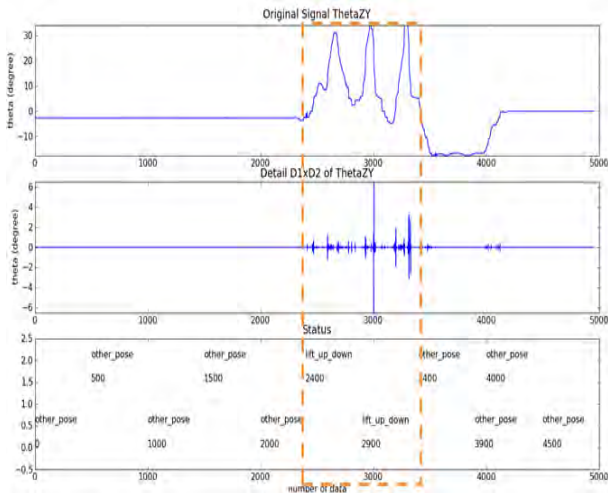


Fig. 14 Example result of locating state transition in the lifted up down action

Table 2 The result in defining state of the robot classify by sensitivity and specificity

		State	
		Testing State	Other State
Result from the System	Defined as Testing State	TP = 36	FP = 0
	Defined as other State	FN = 4	TN = 109

$$Sensitivity = \frac{TruePositives}{TruePositives + FalseNegatives}, \quad (4)$$

$$Specificity = \frac{TrueNegatives}{TrueNegatives + FalsePositives}. \quad (5)$$

4. CONCLUSION

Motion classification is used for identifying states of the robot in the human-robot interaction system so that suitable emotions can be shown by the robot. This paper focused on defining motion state of the robot using signals from an IMU. The data were collected when human interacted with the robot. After that, the output data from IMU were processed and classified by heuristic conditions. By adding the wavelet transform at the preprocessing step, we could correctly locate the starting position of the motion. We found that most actions could be classified with 100% accuracy. The sensitivity and specificity of the proposed classification algorithm are at 90% and 100% respectively (as shown in table 2).

REFERENCES

- [1] S. Koo and D. Kwon, "Gomy: The Baby Bear-like Robot for Emotional Human-Robot Interaction," *The 4th International Conference on Ubiquitous Robot and Ambient Intelligence (URAI 2007)*
- [2] J. Li, M. Chignell, S. Mizobuchi and M. Yasumura, "Emotions and Messages in Simple Robot Gesture," *Human-Computer Interaction, Part II, LNCS 5611*, pp. 331-340, 2009.
- [3] H. Knight, A. Chang, W. D. Stiehl, R. Toscano, Y. Wang and C. Breazeal, "Robot Design Rubrics for Social Gesture Categorization and User Studies with Children," *HRI Workshop on Acceptance of Robots in Society*, 2009.
- [4] R. Jitpakdee and T. Maneewarn, "Neural Networks Terrain Classification using Inertial Measurement Unit for an Autonomous Vehicle," *SICE Annual Conference*, 2008.