The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Full-body Gesture Recognition Using Inertial Sensors for Playful Interaction with Small Humanoid Robot

Martin D. Cooney, Christian Becker-Asano, Takayuki Kanda, Aris Alissandrakis, Hiroshi Ishiguro

Abstract—**People like to play, and robotic technology offers the opportunity to interact with artifacts in new ways. Robots co-existing with humans in domestic and public environments are expected to behave as companions, also engaging in playful interaction. If a robot is small, we foresee that people will want to be able to pick it up and express their intentions playfully by hugging, shaking and moving it around in various ways. Such robots will need to recognize these gestures--which we call "full-body gestures" because they affect the robot's full body. Inertial sensors inside the robot could be used to detect these gestures, in order to avoid having to rely on external sensors in the environment. However, it is not obvious which gestures typically occur during play, and which of these can be reliably detected. We therefore investigate full-body gesture recognition using Sponge Robot, a small humanoid robot equipped with inertial sensors and designed for playful human-robot interaction.**

Keywords: full-body gestures, small robots, inertial sensors, playful human-robot interaction

I. Introduction

Small robots equipped with gesture recognition capability offer great promise for new and fun interactions: people will be able to communicate in a playful fashion with the robots by picking them up and hugging them, shuffling them about, shaking them, dancing with them, and performing other full-body gestures. These gestures are new because they were not possible with previous larger, heavier robots, and are expected to be fun as physically holding the robot allows for up-close, hands-on interaction. We call these gestures "full-body gestures" because they affect the entire body of the robot (position and orientation).

The problem is that the recognition of such gestures by a robot is difficult. In a small-sized robot, limited space restricts the available internal sensor modalities. Secondly, it's not obvious which full-body gestures are typical during play and therefore should be recognized. Thirdly, people perform the same gesture (trying to communicate the same intention) in different ways, affecting the reliability of the detection. Fourthly, there are many kinds of features which could be used by a full-body gesture recognition system and which should be investigated. Finally, while interacting, robots move, and these movements, which are not necessarily related to the current gesture being enacted, can affect the data from the sensors.

Fig. 1. Playful full-body interaction with Sponge Robot

The rest of the paper is structured as follows. Section II describes some related work. Section III introduces Sponge Robot, a novel robotic system with online interactive full-body gesture recognition capability using inertial sensors and Support Vector Machines (SVMs). Sponge Robot is used to investigate the unique problems described above, which arise when small robots are given full-body gesture recognition capability. Section IV describes the developed gesture recognition system, which is then evaluated in section V. Finally section VI summarizes the paper's contributions.

II. Related Work

Playful human-robot interaction has been conducted with a small animal robot [11] and with a small humanoid robot in a classroom setting [14, 15]. However, these did not involve detection of full-body gestures with inertial sensors.

Inertial sensors have been used for various purposes including activity recognition using wearable sensors [13] and teleoperation of a large humanoid robot [5]; in particular, gesture recognition using only inertial sensors has been performed with Wii controllers [10], and a spinning robotic

Manuscript received February 28, 2009. This research was supported by the Ministry of Internal Affairs and Communications of Japan. The first author is partially supported by a MEXT education scholarship. The second author is supported by a JSPS post-doctoral fellowship.

Martin Cooney, Christian Becker-Asano, Aris Alissandrakis, Takayuki Kanda, and Hiroshi Ishiguro are with ATR Intelligent Robotics and Communication Laboratories, 2-2-2 Hikaridai, Keihanna Science City, Kyoto 619-0288, Japan. Martin Cooney and Hiroshi Ishiguro are also with the Department of Systems Innovation, Osaka University, 1-3 Machikaneyama Toyonaka Osaka 560-8531 Japan. (E-mail: martin.cooney@irl.sys.es.osaka-u.ac.jp, christian@ becker-asano.de, [ishiguro, kanda, alissandrakis]@atr.jp)

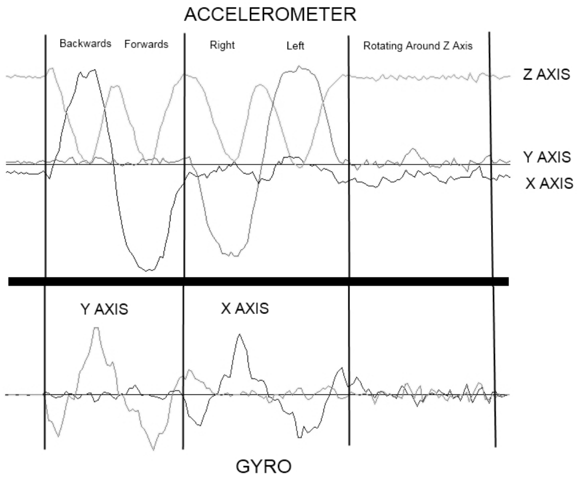

Fig. 2. Effect of Rotating Robot on Inertial Data

ball [8, 9]. In the latter, Salter *et al.* sought to recognize four categories of full-body interaction—alone, interaction, carrying, spinning—over five minute trials, using estimated thresholds on average sensor values from accelerometer and tilt sensors.

However, for our purposes these interaction categories are few; we expect people to perform a variety of gestures—some complex—when interacting with small humanoid robots. Such small humanoid robots with various sensors and gesture recognition capability are being developed for use in up-close [3] and playful interactions [4, 12], [6, 7]. In particular, the Huggable, a small Teddy Bear robot developed for remote operation by the Personal Robots group at MIT, can recognize three full-body gestures directed toward a humanoid robot—pick up, bounce, and rock—using inertial sensors and features based on frequency and "jerk" [4, 12].

However, none of the previous studies have identified what the typical full-body gestures are, proposed a method for identifying these gestures, or reported on which of these gestures can be detected reliably.

III. SPONGE ROBOT

Sponge Robot, the robot developed for playful interaction (see Fig. 1), is a small humanoid robot based on the Robovie-X platform developed at ATR Robotics and Vstone Co., Ltd., Japan. Information on the Robovie-X platform can be found on the Vstone website[1], and a short video showing interaction with Sponge Robot has been submitted with this paper.

Sponge is covered in soft yellow urethane foam, measures roughly 37cm in height and weighs 1.4kg, making it easy to hold and play with.[2] It features a total of 13 degrees of freedom, comprising 2 degrees of freedom in each arm, 4 in each leg, and 1 in its head.

[1] http://www.vstone.co.jp/ [Japanese]

[2] Four motors were removed from the original Robovie-X base to make Sponge lighter for easier interaction, and to make it easier to cover in foam

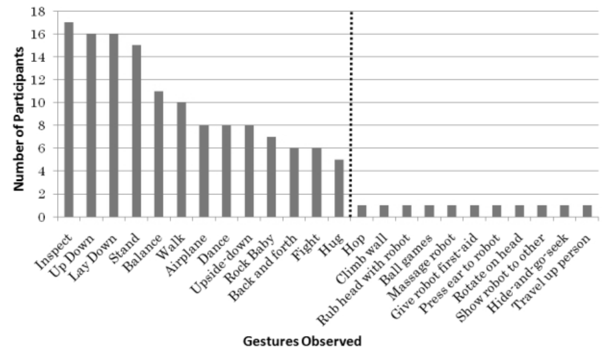

Fig. 3. Gestures observed, sorted by number of participants who performed each gesture. The dotted vertical line indicates the classification targets (left) and those gestures that were not considered (right)

Inertial data are obtained from a 3-axis accelerometer and a 2-axis gyro sensor on Sponge's VS-IX001 inertial sensor board (located in the robot's abdomen). The data are harvested by Sponge's VS-RC003 CPU board over an IXBUS connection, and sent using an AG-BT20E serial Bluetooth Wireless Module (located in the robot's chest) to a laptop computer for processing. In total, it takes an average of 80ms to acquire each new data point consisting of 3 accelerometer and 2 gyro sensor values (a rate of 12.5Hz).

The wireless module is also used to trigger motions. The motions are pre-defined and uploaded to the robot's firmware. This allows motions to be called quickly (~60ms) via wireless commands in a fashion emulating gamepad control.

The accelerometer measures the acceleration due to gravity, and hence its output changes as the robot's posture changes. The gyro sensor measures angular velocity about the X and Y axes. When there is no rotation the readings return to zero. Fig. 2 illustrates how changes in the robot's orientation affect the data obtained from the sensors.

Due to similarity of the robot's shape to that of a human baby, we expect people to interact with the robot in various complex ways, the possible space of which needs to be investigated.

IV. GESTURE RECOGNITION

A. Classification Target

People playfully interacting with a robot would probably lose interest or get a negative impression if the robot does not respond to any of their actions directed towards the robot. Therefore, Sponge Robot should appear to be responsive to such actions. Towards this, a set of typical gestures that occur during play had to be identified; in a free-interaction scenario, 17 participants were asked to play with Sponge and each full-body gesture that occurred was ranked according to the number of participants who performed it. Gestures which required modalities such as vision, sound, or touch, were not noted. The results, with each observed gesture labeled by the experimenter, are shown in Fig. 3. During these sessions, the

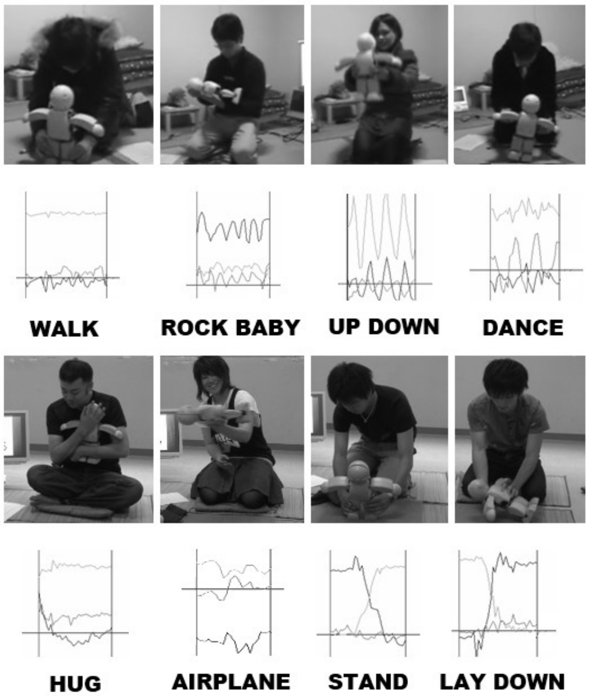

Fig. 4. Participants performing some of the full-body gestures; inertial data is shown below each gesture

robot's power was on, and its arms outstretched in a neutral pose.

The Inspect gesture was the most common; the participants turned Sponge in various directions, examining it from different angles. Also common were Up Down—the robot was raised and/or lowered—Lay Down, and Stand. In contrast, some gestures were performed only by a single participant, such as Ball Games or Rub Head With Robot.

We decided to select as our classification target all gestures which were performed by at least two participants:

1) Inspect – look at different parts of the robot from various angles
2) Up Down – move the robot up and down
3) Lay Down – lay the robot down
4) Stand – raise the robot to a standing position
5) Balance – balance the robot and try to make sure it does not fall
6) Walk – make the robot look like it is walking
7) Airplane Game – make the robot look like it is flying
8) Dance – make the robot do a little dance
9) Upside-down—put the robot upside-down
10) Rock Baby – hold the robot like a baby and rock it
11) Back and forth—shake the robot back and forth
12) Fight – make the robot fight
13) Hug – hug the robot

It is worth noting that the gestures here are defined semantically and not physically. There should not be a need to tell people how they are supposed to play with the robot; instead they should be free to play in their own way. For example, Lay Down and Stand can be performed differently

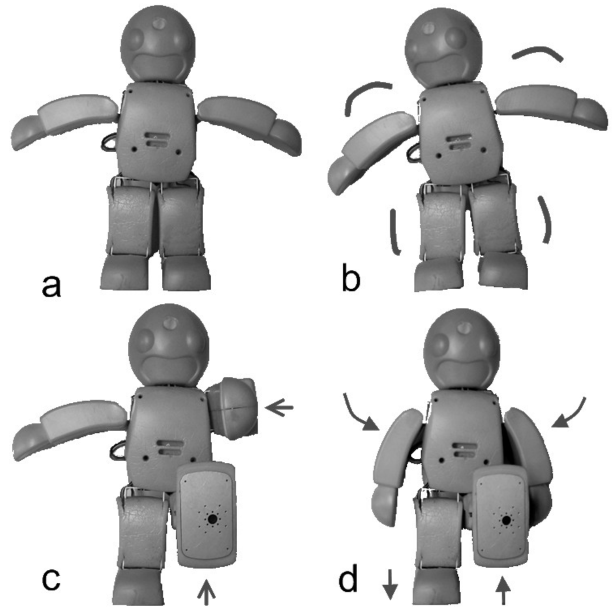

Fig. 5. The four motion conditions for Sponge Robot

depending on whether the robot is facing up or down.

We expected the degree of variation in interpretation to be closely related to the difficulty of recognition. In order to verify this, data were acquired for each of the target gestures.

B. Data Collection

Inertial data was collected from 21 participants in their 20s at Advanced Telecommunications Research Institute International (ATR) and Osaka University, both in Japan.

At both locations, participants sat on pillows over "tatami" floor mats (ATR) or a similar material (Osaka U.), but were allowed to stand and act freely (see Fig. 4). A separate monitor to one side ran a simple clock program to allow identification of when gestures started and ended. Sessions lasted approximately 15 minutes. First, the participants were handed a sheet with a list of gestures and given simple instructions. Next, the robot was turned on in a neutral pose with its arms outstretched to each side, and the participants were instructed to perform the 13 different candidate gestures.

In order to explore the effect of the robot's motion on recognition, the participants were asked to repeat the gestures over four different robot motion conditions (one where the robot was not moving, and three where the robot was moving). These motion conditions are shown in Fig. 5:

a) No motion; the robot's joints were stiff and the robot was in its initial neutral pose, with arms outstretched and legs together
b) Idling – slight, but continuous motion; Gaussian noise was applied to the robot's servo positions
c) Try to Turn – a sudden motion; the robot quickly tucks in one arm and raises its leg to create an unbalanced state
d) Flap arms and legs – a large motion; the robot makes a motion which could interfere with the participant's

ability to grasp the robot.

The idling motion (b) was triggered at the beginning and lasted throughout the condition. The latter two motions (c) and (d) were triggered to occur during each gesture. It was guessed that the robot's motion would significantly disrupt the gestures, thereby reducing recognition accuracy.

Afterwards, a total of 1748 gesture instances were manually labeled using video recordings of the sessions. This involved making subjective decisions about when gestures started and ended. In a few rare cases where the connection between robot and the computer that was used for collecting the data was temporarily interrupted or slower than expected, any missing value was replaced with the previous data value.

After labeling, a learning system was required in order to learn from the data and provide gesture recognition capability.

C. Learning System

A fixed size window was used to classify gestures. The alternative involved finding "breakpoints" where gestures start and end. But, it was assumed that people interacting with the robot would find it disruptive to have to pause between gestures or return the robot to some neutral position. Also, finding breakpoints would result in long delays (not desirable for playful interaction) when waiting for long gestures to end, even if the information needed to recognize the gesture could be found by a short window. Furthermore, we didn't want our results to depend on the efficacy of the breakpoint-finding algorithm, as this was not our main focus. For these reasons, a fixed sized window was selected.

We found a window of about 3 seconds to be sufficient for capturing information from the gestures. This means we expected gestures to last a few seconds, but not that the system must necessarily wait for 3 seconds. Gesture recognition can take place each time a new data point has been added to the window (with a delay of around 80ms). Thus, for short gestures the probability output for that gesture is likely to go high before the full 3 seconds has passed, and the system does not need to wait the entire time. This timing depends on the training samples and how the gestures are temporally defined; e.g. when does "Hug" start? Does the gesture start when the robot is picked up? When the robot is raised and (usually) tilted slightly backward? When the robot is tilted forward and first comes into contact with the person's chest? Or just before the robot is tilted backward and released from physical contact? These decisions affect when the probability output goes high, and when the system can recognize a gesture.

In order to classify the windows, we decided to use standard one-vs.-one RBF kernel SVMs with probabilistic output using LIBSVM [1, 2].[3]

For the SVM classifiers, a one-vs.-one system was chosen for accuracy at the cost of using more binary classifiers than a one-vs.-all system. The RBF kernel was chosen for its applicability to nonlinear problems and the other reasons listed in [1], including avoidance of numerical problems by constraining kernel coefficients to be between 0 and 1, and the small number (2) of hyper-parameters which must be found.

[3] We also experimented with several other approaches including k-NN, a k-means algorithm which classified samples using learned centroids, a Mahalanobis distance-based classifier, a one-vs.-all form of Adaboost, and nu-SVMs using a different SVM library, but we obtained the greatest accuracy and speed with LIBSVM

TABLE 1
COMPARING FEATURE TYPES

Feature Type	Cross-validation accuracy
Various statistics	74.3
DFT coefficients	62.1
Haar coefficients	51.4

Regarding these hyper-parameters, C and gamma, the algorithm was set to automatically find values for C and gamma for each fold using LIBSVM when doing cross-validation. For the entire dataset, we found values of C = 8 and gamma = 0.5.

After defining the overall system, we needed to determine what useful information (i.e. features) could be extracted from the data and used by the system to recognize the target gestures, but it was not evident which features would be best suited for our problem.

D. Features

We investigated several types of candidate features. The use of frequency-based features in [4] suggested the applicability of Haar and Discrete Fourier Transform (DFT) magnitude coefficients. Haar coefficients capture both time and frequency information, and are simple and fast to calculate; the cyclic nature of many of the gestures also suggested purely frequency-based features such as DFT magnitude coefficients could capture valuable information. In addition, we considered a group of various statistics, which included mean axis values (also used by Salter *et al.* [8, 9]) as well as features we thought might work well for our problem such as the overall "trends" (the change between first input value and last input value) for each axis.

We ran a wrapper-based feature selection algorithm based on the system described in the preceding section in order to decide which type of feature to use. This yielded a cross-validation accuracy score for each full group of features, which was used to rank feature groups. The results can be seen in Table 1. The "various statistics" group (composed of 40 different features) performed the best.

We explored both increasing the size and decreasing the size of this group. We found a slight decrease in accuracy when cross-axis variants of the statistics were added. Next we tried reducing the features in order to increase accuracy, prevent over-fitting, and better understand what qualities of the data change for different gestures; eliminating related features from an initially full set gave a slight improvement in cross-validation accuracy. This resulted in the following list of 19 features used:

1) Mean values for accelerometer (3)
2) Standard deviations for accelerometer, gyro (5)
3) Overall "trends" for accelerometer (3)

Fig. 6. Variations for (a-c) Fight (d-f) Hug (g-i) Inspect

Gestures		I	UD	LD	S	B	W	A	D	U	RB	BF	F	H
Inspect	I	**49**	0	3	3	18	9	6	0	0	9	0	3	0
Up Down	UD	0	**89**	2	0	0	0	3	0	0	3	0	0	3
Lay Down	LD	3	0	**86**	0	0	0	3	0	0	5	0	3	0
Stand	S	0	0	0	**98**	0	0	2	0	0	0	0	0	0
Balance	B	0	0	2	0	**81**	10	0	0	0	0	0	5	2
Walk	W	0	0	0	0	35	**41**	0	13	0	0	0	11	0
Airplane	A	8	0	3	3	0	0	**83**	0	0	3	0	0	0
Dance	D	3	0	0	0	0	3	0	**80**	0	0	0	9	5
Upside-down	U	2	0	0	0	0	0	0	0	**98**	0	0	0	0
Rock Baby	RB	12	3	0	0	0	3	0	3	0	**64**	3	3	9
Back and Forth	BF	0	0	2	0	0	0	0	0	0	0	**98**	0	0
Fight	F	3	3	0	0	8	3	0	16	0	3	3	**58**	3
Hug	H	0	0	0	0	9	0	0	0	0	15	0	12	**64**

Fig. 7. Confusion Matrix for the 13 target gestures

4) Medians for accelerometer (3)
5) Minimums for accelerometer (3)
6) Maximums for gyro (2)

Having made the necessary decisions about the candidate gestures, the nature of our gesture recognition system, and the features to use, the next step was to use the collected data to evaluate the proposed approach.

V. EVALUATION

A. Results

1) Gesture detection

During data collection, we observed overlap between some gestures. Some gestures such as Upside-down had a stronger effect on the inertial data than others. Some gestures were also interpreted in many different ways. This variance was not just due to a difference in the way different participants chose to interpret the gestures, but was even observed within single participants' data as they varied the gestures each time they were asked to perform them.

Fig. 6 shows examples of the variations observed for several of the gestures. The top row, Fight, shows participants making Sponge punch, kick, and body slam. For Hug, we see participants facing the robot, or hugging Sponge from behind, or only half-hugging the robot. For the last row, Inspect, participants can be seen rotating Sponge, examining the robot without touching it, and lifting the robot while craning their heads to see it from various angles.

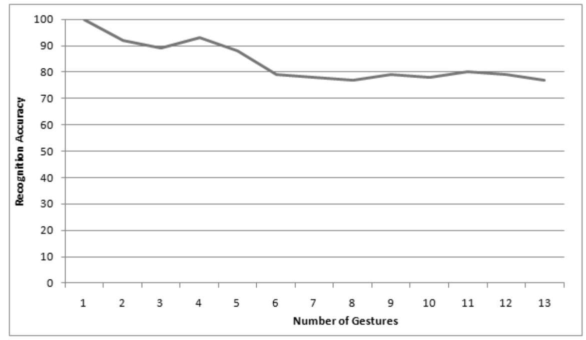

Fig. 8. Accuracies obtained for different size sets

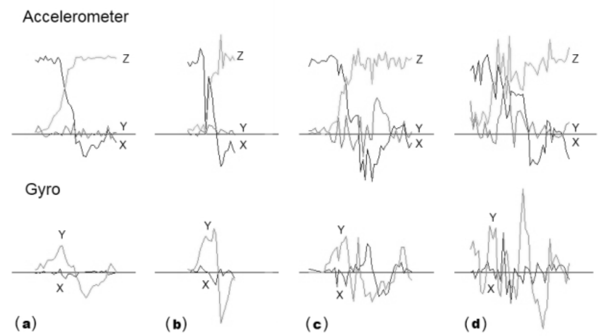

Fig. 9. Effect of robot's motion on gesture (Stand) a None b Idling c Try to Turn d Flap Arms and Legs

Fig. 7 shows a confusion matrix obtained for the gestures using leave-one-out cross-validation. We can see that Walk (41%), Inspect (49%), Fight (58%), Hug (64%), and Rock Baby (64%) were the most difficult to distinguish from other gestures. We think overlap, variance, and impact on inertial data were the cause for the low recognition accuracies for these gestures. For example, participants sometimes did a floating motion for Walk which resembled the start for Balance and Fight when the robot was being transported somewhere to be balanced or being brought close to its adversary. In addition, a great deal of variation was observed for Inspect and Fight. Also, Rock Baby and Hug in particular were often performed gently, and did not change the inertial data input as strongly as gestures such as Back And Forth or Upside-Down.

An accuracy of 77% was obtained for the system. But, for certain contexts perhaps not all gestures are required. In those cases, higher accuracies could be obtained. Fig. 8 shows, for example, that an accuracy of 93% can be realized for the 4 most common gestures.

2) Effect of robot's motion on recognition

Inspection of the inertial data showed that the robot's motion had a visible impact on the data, as can be seen in the example shown in Fig. 9.

In order to investigate the degree to which accuracy was affected, we compared the accuracy of a standard system

Gestures		I	UD	LD	S	B	W	A	D	U	RB	BF	F	H
Inspect	I	**39**	9	14	0	0	0	3	8	2	14	0	9	2
Up Down	UD	2	**70**	2	5	0	0	1	5	0	10	2	2	1
Lay Down	LD	9	3	**74**	0	2	0	11	0	0	1	0	0	0
Stand	S	2	5	0	**82**	0	0	3	0	0	6	0	1	1
Balance	B	11	5	5	0	**7**	16	1	5	0	5	4	35	6
Walk	W	3	17	0	0	1	**9**	12	10	0	3	5	36	4
Airplane	A	9	0	0	2	0	0	**85**	1	0	3	0	0	0
Dance	D	3	8	0	0	0	1	0	**58**	0	3	0	26	1
Upside-down	U	4	0	0	0	0	0	1	0	**95**	0	0	0	0
Rock Baby	RB	3	11	1	2	2	0	0	7	1	**61**	1	6	5
Back and Forth	BF	0	0	1	0	1	0	0	2	0	5	**78**	13	0
Fight	F	3	20	1	1	0	3	3	5	0	4	3	**41**	16
Hug	H	1	16	2	3	0	4	0	3	0	14	2	31	**24**

Fig. 10. Confusion matrix for motion set
Changes to accuracy: I (-10), UD (-19), LD (-13), S (-17), B (-74), W (-33), A (+1), D (-23), U (-3), RB (-4), BF (-21), F (-18), H (-40)

trained using samples from the non-motion case on two different sets: non-motion samples, versus motion samples. The motion set was made to be the same size as the non-motion set by random sampling without replacement, and the process was repeated 10 times with the resulting accuracies averaged in order to avoid lucky or unlucky draws.

Cross-validation accuracy for the non-motion set was 77%, compared with an accuracy of 56% for the motion set. This result clearly shows an adverse effect from the robot's motion on gesture recognition accuracy.

We attempted to gain insight into this issue. Simple approaches such as smoothing or training with motion data did not fix the problem—probably because people's reactions were not easily predictable and their effect not simple—but the confusion matrix for the motion set (Fig. 10) revealed that Balance, Walk, and Hug in particular were highly sensitive to the robot's motion. These gestures become very difficult to detect (7, 8, 24) with sharp decreases in accuracy (-74, -33, -40), and become increasingly confused with gestures with a relatively stronger inertial effect (e.g., Fight). We think that knowledge of which gestures are sensitive could be useful when deciding a target application; also, when desired, this knowledge could be combined with an uncertain response from the robot to reduce errors and provide a more consistent system for playful interaction.

In summary, we found an effect of the robot's motion on recognition accuracy, but the obtained system accuracy for 13 gestures was still far in excess of random chance (1/13 = 8%). We think this is because participants were observed trying hard to compensate for the robot's motions when carrying out gestures, and we expect to see similar results when Sponge Robot is used in playful interactions with real users in the future.

B. Discussion

We observed interesting phenomena related to gesture detection. First, during the free-interaction trials many participants rubbed the top of Sponge's head and squeezed its hands; as well, two participants were also seen trying to dress Sponge (using their own glasses and handbag) and another greeted the robot. Unfortunately, this could not be detected by Sponge's inertial sensors. Second, for data collection, it was noted that the participants varied their interpretations of gestures, although they were not asked to do so. Based on the participants' comments we assume that the robot's motion added to the smoothness and playfulness of the interaction.

Noteworthy with regard to the robot's movements was that we found they had a complex effect on free interaction; it seemed possible to suggest or deter gestures, but criteria for choosing motions and timing were not obvious. Second, during data collection participants noted how the robot's motion caused them to change their grasps on the robot; this suggests it could be possible to estimate how people are holding the robot, in order to avoid disrupting (or deliberately disrupt) people's grasps. Third, when we outfitted the system with responses, we found another complication due to the robot's motion in which Sponge would trigger its own responses; e.g., Sponge would walk a few steps forward when the Walk gesture was detected, but the walking motion would often cause Walk to be recognized again. Although this problem can be solved by, e.g., waiting before recognizing subsequent Walks, by increasing the probability output threshold for Walk, or by checking that the sum of the gyro activity motion is greater than some threshold, we think this recursive behavior could be a cause for playfulness and fun during the interaction.

For the developed system, we noted regarding the Rock Baby gesture that the robot's right shoulder tended to be lower when it carried on the left, and vice versa. Out of 22 recorded gestures—11 carrying the robot on the left side, and 11 carrying the robot on the right—a simple threshold on the average accelerometer Y axis yielded 91% accuracy (20/22 cases labeled correctly). This could be used for Sponge to look toward (or away from) the person holding it when Rock Baby (or Hug) is detected.

VI. Conclusions

In summary, this paper reported on the unique problem of full-body gesture recognition for a small humanoid robot designed for playful interactions. First, 13 typical full-body gestures were identified from observing free interaction with the robot. Next, we found that statistics such as the mean, standard deviation, and change across a window of data for each axis performed better for gesture recognition than frequency-based features such as Discrete Fourier Transform Coefficients or Haar Transform Coefficients. We reported on a SVM-based system which recognizes these typical gestures with an average accuracy of 77%, and identified gestures which were not easily detectable, proposing that variation, overlap, and inertial effect could be related to ease of gesture recognition. In addition, we explored the extent of the effect of the robot's movement on classification accuracy, identifying three gestures particularly sensitive to the robot's motion, and found the system still performed quite well despite the difficulty of the task. Lastly, this paper introduced Sponge Robot, a new small humanoid robotic system developed for playful interactions which can recognize full-body gestures using inertial sensors and respond in an equally complex fashion.

Future work will involve extending the present recognition system to use multiple sensors (e.g., inertial and touch), extracting gestures without using a fixed sized window, and increasing the robustness of the system to the effects of the robot's motions (possibly by implementing a form of self-motion perception such as may be observed in humans). Knowledge of context, in conjunction with the recognized gestures, can be employed toward inferring the users' intentions. At the interaction level, identifying users' patterns of interaction during play and the effects of the robot's motion responses to recognized gestures on these patterns of interaction remain topics to be further explored.

ACKNOWLEDGMENT

We'd like to thank Reo Matsumura at ATR and Takuro Imagawa at Vstone for much help with Sponge Robot, as well as Tomoko Yonezawa; we are grateful for all the assistance we received.

REFERENCES

[1] Chang C., Hsu C., and Lin C. "A Practical Guide to Support Vector Classification". Taipei 106, Taiwan, 2003

[2] Chang C. and Lin C. "LIBSVM: a library for support vector machines", http://www.csie.ntu.edu.tw/ cjlin/libsvm, 2001

[3] Hayashi M., Sagisaka T., Ishizaka Y., Yoshikai T., and Inaba M. "Development of Functional Whole-Body Flesh with Distributed Three-axis Force Sensors to Enable Close Interaction by Humanoids", in Proceedings of The 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.3610-3615, 2007.

[4] Lee, J. K., Toscano, R. L., Stiehl W. D., and Breazeal, C. "The Design of a Semi-Autonomous Robot Avatar for Family Communication and Education", Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2008.

[5] Miller, N., Jenkins, O. C., Kallmann, M., and Matarić, M. J. "Motion Capture from Inertial Sensing for Untethered Humanoid Teleoperation", IEEE/RAS Int. Conf. on Humanoid Robots, pp. 547-565, 2004.

[6] Saldien, J., Goris, K., Vanderborght B., Verrelst, B., Van Ham, R., and Lefeber, D. "ANTY: The development of an intelligent huggable robot for hospitalized children", CLAWAR, 2006.

[7] Saldien, J., Goris, K., Yilmazyildiz S., Verhelst, W., and Lefeber, D. "On the design of the huggable robot Probo". Journal of Physical Agents, Special Issue on Human Interaction with Domestic Robots, Vol. 2, No. 2 2008.

[8] Salter, T., Michaud, F., Dautenhahn, K., Létourneau, D., and Caron, S. "Recognizing interaction from a robot's perspective," Proceedings IEEE International Workshop on Robot and Human Interactive Communication, Nashville USA, 178-183, 2005

[9] Salter,T., Michaud F., Letourneau D., Lee D.C., and Werry I.P. "Using Proprioceptive Sensors for Categorizing. Human-Robot Interactions," 2nd ACM/IEEE International. Conference on Human-Robot Interaction, 2007.

[10] Schloemer, T., Poppinga, B., Henze, N., and Boll, S. "Gesture Recognition with a Wii Controller", Proceedings of the 2nd international Conference on Tangible and Embedded interaction, 2008

[11] Shibata, T. and Tanie, K. "Physical and Affective Interaction between Human and Mental Commit Robot", IEEE Int. Conf. on Robotics and Automation (ICRA2001), pp. 2572-2577, 2001.

[12] Stiehl, W. D., Lieberman, J., Breazeal, C. , Basel, L., Lalla, L., and Wolf, M. "Design of a Therapeutic Robotic Companion for Relational, Affective Touch," in IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN), Nashville, TN, 2005.

[13] Subramanya, A., Raj, A., Bilmes, J. and Fox, D. "Recognizing Activities and Spatial Context Using Wearable Sensors". Conf. on Uncertainty in Artificial Intelligence, 2006.

[14] Tanaka, F., Cicourel, A., and Movellan, J. R., "Socialization between Toddlers and Robots at an Early Childhood Education Center". Proceedings of the National Academy of Sciences of the USA (PNAS), pp. 17954-17958, 2007.

[15] Tanaka, F., Movellan, J. R., Fortenberry, B., and Aisaka, K. "Daily Hri Evaluation at a Classroom Environment:Reports from Dance Interaction Experiments", ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI2006), pp. 3-9, 2006.