An Inertial Sensor-based System to Develop Motor Capacity in Children with Cerebral Palsy

Shuo Qiao, Anil Prabhakar, Nitin Chandrachoodan, Namita Jacob, Harshvardhan Vathsangam

***Abstract*—Learning to communicate with alternative augmentative communication devices can be difficult because of the difficulty of achieving controlled interaction while simultaneously learning to communicate. What is needed is a device that harnesses a child's natural motor capabilities and provides the means to reinforce them. We present a kinematic sensor-based system that learns a child's natural gestural capability and allows him/her to practice those capabilities in the context of a game. Movement is captured with a single kinematic sensor that can be worn anywhere on the body. A gesture recognition algorithm interactively learns gesture models using kinematic data with the help of a nearby teacher. Learned gesture models are applied in the context of a game to help the child practice gestures to gain better consistency. The system was successfully tested with a child over two sessions. The system learned four candidate gestures: lift hand, sweep right, twist right and punch forward. These were then used in a game. The child showed better consistency in performing the gestures as each session progressed. We aim to expand on this work by developing qualitative scores of movement quality and quantifying algorithm accuracy on a larger population over long periods of time.**

***Index Terms*—Cerebral Palsy, Accelerometer, Gyroscope, Game, Motor Skills**

I. INTRODUCTION

Alternative Augmentative Communication (AAC) systems are commonly used for communication [1] by children with cerebral palsy (CP). Typical operation of these devices requires controlled motor interaction such as touching a switch or pointing. One issue with their operation is that accurate reach and press require a lot of training effort from the child, detracting from the process of communication itself, since the physical demands of the task are so high. The problem is compounded when the child is visually impaired, a condition found in 60% of children with CP. Visual impairment does not allow precise motor control due to lack of feedback. Thus, from a design perspective, what is needed is a communication device that harnesses a child's natural motor capabilities and provides the means to reinforce them.

In this paper, we describe a kinematic sensor-based system that is designed to learn a child's existing motor capacity. Movement is captured with a kinematic sensor (consisting of accelerometers and gyroscopes) that can be worn anywhere on the body. Attaching a kinematic sensor to the body allows direct capture of movement free of occlusions. Kinematic sensor-based activity monitors have been successfully used in monitoring of gross movement [2, 3]. Accelerometer-based sensors that recognize gestures have been developed to detect arm movements [4, 5] and in characterizing head motion disorders for gesture recognition [6]. We extend this work by employing a gesture recognition algorithm that uses movement captured with kinematic sensors and interactively learns gesture models that are natural to the child. Here, a teacher or caregiver uses the child's natural movements to teach the system an understanding of what gestures the child is capable of. These are then associated with a dictionary of sentences or actions.

Another goal in this work is to reinforce a child's natural gestures to increase repeatability. One way to reinforce motor skills and improve gesture consistency is through practice in a game. Packaging motor skill training in a video game format has the potential to improve sessions by reducing the stress of learning. Jannink et al. [7] successfully used the EyeToy game for training of upper extremity function in children. Sandlund et al. [8] used off-the-shelf consoles and reported high compliance. We adopt this approach by using learned gesture models in the context of a game to help a child practice gestures. An issue with previous systems is that they require the infrastructure of a living room for full usability. Better compliance could be achieved if the game were made even more portable. In our system, kinematic data captured by the sensor are transmitted wirelessly to a phone, thus allowing easy portability according to the convenience of the child.

The unique contribution of this system is that it harnesses existing motor capacity and movement in children to enable communication. The device has the potential to enable clear communication for students whose ability to control breathing and sound production reduce drastically as their emotions or discomfort rise. The sensor also makes it intuitive for a child who is visually impaired, as there is no need to interact with an object in space. It also has great potential for students who have severe cognitive deficits, are non-verbal and whose communication of needs are understood by subtle movements. Using our system, one can verbalize these movements, making their needs understood by a wider range of people who may otherwise not understand their signals.

II. GESTURE CAPTURE AND LEARNING

Learning a gesture model involves supplying appropriate (gesture, label) pairs in batch mode to a classifier that then optimizes the model's parameters. This approach was not feasible in our case for a number of reasons. For children with cerebral palsy without adequate training in gestures, it was not always possible to perform a gesture consistently. In addition, a child may be prone to involuntary movements which would need to be segmented out. Thus the input gestures are noisy and have many false positives. Additionally, in the absence of prior information about gesture structure, the quality of movement for the gesture needs to be assessed before passing it as an input to the learning algorithm.

To tackle these constraints, we adopted an interactive machine learning [9, 10] paradigm. This paradigm relies on

Anil Prabhakar and Nitin Chandrachoodan are with the Department of Electrical Engineering, Indian Institute of Technology - Madras, Chennai 600036, India.

Shuo Qiao is with Beijing Jiaotong University, Haidian District, Beijing, China

Namita Jacob is with Chetana Charitable Trust, Chennai 600036, India.

H. Vathsangam is with the Dept. of Computer Science, Univ. of Southern California, Los Angeles, CA 90089, USA. Support for H. Vathsangam was provided by the USC Annenberg Doctoral fellowship.

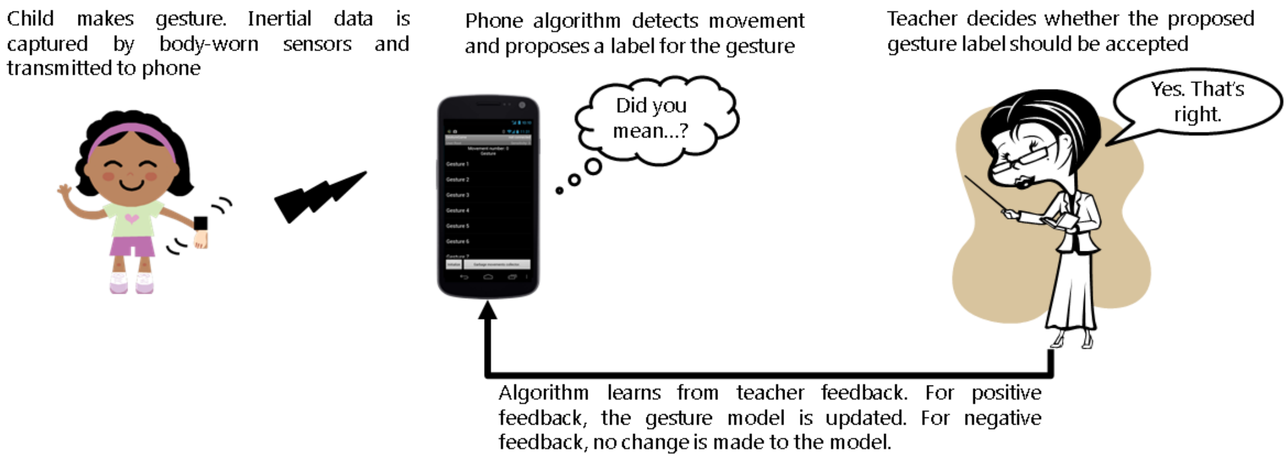

Figure 1: Models can be learned to take advantage of a child's individual motor capability. Here, rather than focusing on teaching the child to adapt to the device, the teacher teaches the device to adapt to the child.

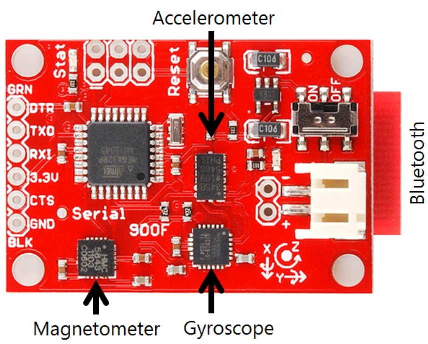

Figure 2: Sensor hardware used to collect data. Data was streamed wirelessly to a nearby phone. Image source: www.sparkfun.com

learning a gesture model in an online setting by providing continuous feedback to the algorithm with a human in the loop. As shown in Figure 1, the human in the loop was a teacher who supervised the child and the algorithm during the training procedure. The teacher decided whether a (gesture, label) pair was appropriate for the gesture model. If approved, the classifier would update its parameters to be consistent with the teacher's input. In this way, rather than simultaneously teaching the child how to communicate and make gestures appropriate to AAC hardware, the teacher could leave the burden of adaptation to the gesture learning system. The teacher could then focus on teaching the child the associations between the learned natural movements and a communicative output.

A. Hardware and Pre-processing

Kinematic data were captured using the Razor Inertial Measurement Unit [11] from Sparkfun. The sensor captures triaxial accelerations (using the ADXL345 accelerometer) and rotational rates (using the ITG-3200 gyroscope). These data are sampled at 50 Hz and transmitted via the RN-41 Bluetooth module. Data were received by the Samsung 5360 phone running Android v2.3.5. The phone has an 830 MHz ARMv6 processor. The sensor is typically worn on the particular limb segment for which movement is to be recorded. Before using the system, gyroscope readings needed to be calibrated by subtracting a DC bias. This was done by keeping the sensor still for 3 seconds, recording the mean bias values for each axis, represented as $\left[\begin{array}{ccc} \hat{g}_x & \hat{g}_y & \hat{g}_z \end{array}\right]$ and then continuously subtracting the same from each gyroscope sensor stream.

B. Gesture Segmentation and Parametrization

Segmentation was achieved using gyroscopes. Gestures were segmented using the rule that any movement outside of rest should be recorded as a gesture. At an instant of time, t, let each six dimensional data point be:

$$\begin{aligned} \mathbf{d}(t) &= \left[\begin{array}{cc} \mathbf{d}_{acc}(t) & \mathbf{d}_{gyr}(t) \end{array}\right] \\ &= \left[\begin{array}{cccccc} a_x(t) & a_y(t) & a_z(t) & g_x(t) & g_y(t) & g_z(t) \end{array}\right]. \end{aligned}$$

Given the gyroscopic component of the movement vector $\mathbf{d}_{gyr}$, we calculated a movement magnitude time series where each instantaneous term is given by:

$$m_{gyr}(t) = \sqrt{(g_x(t)-\hat{g}_x)^2 + (g_z(t)-\hat{g}_y)^2 + (g_z(t)-\hat{g}_z)^2}.$$

This time series was then filtered using a low pass filter with 3 dB cutoff that could be adjusted to anywhere between 1 Hz and 2 Hz, depending on the limb to which the sensor was attached and the child's movement capability. This was chosen to preserve smoothness of movement while allowing a minimum time lapse of 1 second (or 2 seconds) between gestures. With the filtered time series, if the movement magnitude was greater than a set threshold and remained so for a minimum amount of time, the original time series $\mathbf{d}(t)$ corresponding to this time period was segmented out as a gesture candidate. Each gesture could be represented by a $6 \times N_n$ matrix where N_n is the length of the n^{th} gesture.

The segmented gesture was split into 4 sub-segments each of dimension $6 \times \lfloor \frac{N_n}{4} \rfloor$. This resulted in a total 24 sub-gestures corresponding to 6 data streams and 4 sub-segments for each dimension. These numbers were chosen based on experience with the data. Each sub-gesture was parametrized into a feature vector to represent the information contained in it. The features calculated were mean, variance, and root-mean-squared value. This resulted in a 72 dimensional feature vector, x_n to describe the n^{th} gesture.

Algorithm 1 Algorithm to learn gesture models from data

Initialize:
Set $\mathbf{W} = 0 \left(W \in \mathbb{R}^{k \times d}\right)$ for k classes and d features.
Let $\mathbf{W}_r$ be the r^{th}row of this matrix corresponding to class r.
Loop:
Foreach training data input x_n

- Predict the label $\hat{y}_n \leftarrow \underset{r}{\text{argmax}} \mathbf{W}_r^T x_n$
- Provide the label $y_n$by pressing button.
- If $\hat{y}_n = y_n$, add $\{x_n, y_n\}$ to reference set $\mathcal{R}$ return
- else run function_remodel

endFor

function_remodel -
er $\leftarrow 1$
while(er $= 1$)
Foreach training pair $\{x_n, y_n\} \in \mathcal{R}$
errors $\leftarrow 0$

- Predict the label $\hat{y}_n \leftarrow \underset{r}{\text{argmax}} \mathbf{W}_r^T x_n$
- If $\hat{y}_n \neq y_n$
  - $W_{\hat{y}_n} \leftarrow W_{\hat{y}_n} - \tau_n x_n$
  - $W_{y_n} \leftarrow W_{y_n} + \tau_n x_n$
  - errors $\leftarrow$ errors $+ 1$

if errors $= 0$
er $\leftarrow 0$

C. Algorithm Description

The segmented gesture candidates were used to train a gesture model that maps gesture features to dictionary labels. We implemented an interactive algorithm using a modified version of the perceptron algorithm [12] with a winner-takes-all extension to handle multiple classes. Our aim was to pick a set of natural gestures that the child can perform and have the algorithm learn them for future use. The perceptron algorithm, described in algorithm 1, uses a modifiable linear model to predict labels for gestures. The child supplies a candidate gesture, represented by its feature vector x_n. The algorithm supplies a most likely label $\hat{y}_n$ predicted using a linear model as:

$$\hat{y}_n = \underset{r}{\text{argmax}} \mathbf{W}_r^T x_n$$

The teacher can decide whether to accept the label as correct, ignore the gesture or supply a correction label. If the algorithm's label is marked as correct, the feature-label pair is added to a historical reference set. If a correction label is supplied, then the algorithm adjusts its model parameters so as to be simultaneously consistent with the new label and historic reference set. Once the model is learned, the algorithm can guess future gestures using the rule $\hat{y}_n = \underset{r}{\text{argmax}} \mathbf{W}_r^T x_n$.

D. Design Issues

In addition to the feedback that the teacher provides, additional heuristics were necessary to ensure that the algorithm converged to the right model. We found that when providing multiple examples of an input gesture, it was better to train on a smaller set of "correct" gestures than a larger number of noisy gestures. The learning algorithm was thus dependent on the teacher making sure that input data quality was maintained. Additionally, it was required to set a dynamic learning rate η to account for the confidence of the learning algorithm. These included cases where the algorithm had high confidence and was wrong (fast learning rate to ensure quicker correction), low confidence and wrong (slow learning rate for slower correction) and corner cases such as encountering a gesture for the first time.

Figure 4: Mobile phone game that requires a child to minimize off-diagonal scores. With more practice while playing the game, the child showed more consistent performance as the session progressed due to increased familiarity.

III. System Evaluation

A. Game Design

A mobile game was developed to let the child practice the gestures learned by the algorithm. The design considerations for the game were that it should be easy to play, able to keep the child engaged for sufficient practice and provide meaningful outputs for a therapist to check the child's performance. Additionally, the game must also be visible for visually impaired children. We formulated the game in the form of a square matrix with each row corresponding to the performance for a particular gesture. Each square also had a number corresponding to the number of times the gesture was correct. Each main diagonal element of the square corresponded to the child performing the gesture correctly. These were colored yellow (for easy visibility), the remaining colored gray. The child was encouraged to perform the gesture corresponding to each row. Depending on what the algorithm predicted, the number of the corresponding box would increase. If performed correctly, the number in the yellow box corresponding to that gesture would increase. If performed wrong, the number in one of the gray boxes in the same row would increase. Controls for toggling between gestures were provided. Additionally, an animation such as a big smile face replaced the yellow box for half a second when the gesture was correct to attract the child's attention. The goal of the game was to get as high a number on the main diagonal entries as possible. Raw data were also recorded simultaneously for offline inspection.

B. Gesture Game Results

We evaluated the system on a single child. The child is a quadriplegic with athetoid cerebral palsy, uses a wheelchair for mobility, has limited speech and uses a communication chart for better communication. Prior permission was granted and the study was approved for safety. The sensor was worn on the right hand in the form of a wristband. The teacher

(a) An example of the sensor being worn on upper arm using a wrist-band. (b) A typical session with the child performing gestures and teacher training the algorithm. (c) Game setup once gesture models were learned.

Figure 3: Illustration of capture of triaxial accelerations and rotational rates from the wrist when performing gestures.

ascertained four candidate gestures: lift hand, sweep hand right, twist right and punch forward. These gestures were learned using the learning algorithm described in section II-C and the model was applied to the game. We then evaluated user performance across two sessions. All sessions were performed in the presence of a teacher who closely works with the child. In each session, the child was asked to perform each gesture 20 to 25 times. Figure 4 illustrates the display used in the game. The first session corresponded to the child playing the game for the first time. Here, the child had to get used to the game mechanics and had to use the gestures for the first time. Typical gesture accuracies were in the range of 60-70%. As the game progressed, the child became increasingly familiar with the gestures. This was repeated in the second session with the additional advantage the child was already familiar with the gestures. This resulted in higher accuracy. These preliminary results suggest that familiarity with the movement results in increased consistency of motor skills. We aim to explore this further in future work.

IV. Summary

This paper described a kinematic sensor-based system designed to capture natural gestural capability in children with cerebral palsy and practice them in the context of a game. Triaxial accelerations and rotational rates were sensed and transmitted wirelessly to a phone. The game software first learned gesture models from the the child's natural movements by interactively learning gestures with the assistance of a teacher. These learned gestures were used as templates for the child to practice in the context of a game. The system learned four gestures and we presented an evaluation of our system with a child over two sessions. The child's gesture consistency as predicted by the number of correct entries in our game display improved between two sessions. This indicates initial feasibility in improving motor skills.

We aim to improve our work in a number of ways. We plan to use measures that are relevant from a clinical assessment perspective to assess the quality of movements of the child [13]. We also plan on conducting a study on a larger cohort of children to examine in detail the effects of our algorithm across individuals and over time. We are currently building next generation hardware that records information on to an SD card for further analysis. This will allow us to track long term trends in movement and detect characteristic gestures that a child is capable of based on several hours of data. We also plan on improving the game mechanics by improving sensor hardware, improving the game design and supply audio outputs for each game.

Acknowledgments

The authors wish to thank the staff and students at Vidya Sagar, Chennai, Atishay Jain and Nathan Koilpillai for their support and help.

References

[1] B. Fossett and P. Mirenda., "Augmentative and Alternative Communication," *Handbook of Developmental Disabilities*, pp. 330–366, 2009.

[2] A. E. Bialocerkowski, "Wearable wrist activity monitor as an indicator of functional hand use in children with cerebral palsy," *Developmental Medicine & Child Neurology*, vol. 53, no. 11, pp. 976–976, 2011.

[3] A. Gordon, J. Schneider, A. Chinnan, and J. Charles, "Efficacy of a hand-arm bimanual intensive therapy (HABIT) in children with hemiplegic cerebral palsy: a randomized control trial." *Developmental Medicine & Child Neurology.*, vol. 49, pp. 830–8, 2007.

[4] M. Harrington, R. Daniel, and P. Kyberd, "A measurement system for the recognition of arm gestures using accelerometers." in *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 209, no. 2, 1995, pp. 129–34.

[5] R. Muscillo, M. Schmid, S. Conforto, and T. D'Alessio, "Early recognition of upper limb motor tasks through accelerometers: real-time implementation of a DTW-based algorithm." *Computers in biology and medicine*, vol. 41, no. 3, pp. 164–72, 2011.

[6] R. Raya, E. Rocon, R. Ceres, J. Harlaar, and J. Geytenbeek, "Characterizing head motor disorders to create novel interfaces for people with cerebral palsy: Creating an alternative communication channel by head motion," in *IEEE International Conference on Rehabilitation Robotics*, 2011, pp. 1 –6.

[7] M. Jannink, G. van der Wilden, D. Navis, G. Visser, J. Gussinklo, and M. Ijzerman, "A low-cost video game applied for training of upper extremity function in children with cerebral palsy: a pilot study." *CyberPsychology & Behavior*, vol. 11, pp. 27–32, Feb 2008.

[8] M. Sandlund, E. Lindh Waterworth, and C. Häger, "Using motion interactive games to promote physical activity and enhance motor performance in children with cerebral palsy," *Developmental Neurorehabilitation*, vol. 14, no. 1, pp. 15–21, 2011.

[9] J. A. Fails and D. R. Olsen, Jr., "Interactive machine learning," in *Proceedings of the 8th international conference on Intelligent user interfaces*. New York, NY, USA: ACM, 2003, pp. 39–45.

[10] A. Kapoor, B. Lee, D. Tan, and E. Horvitz, "Learning to Learn: Algorithmic Inspirations from Human Problem Solving," in *AAAI 2012*, 2012.

[11] "Sparkfun Razor IMU," Tech. Rep. [Online]. Available: https://www.sparkfun.com/products/10736?

[12] F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Psychological Review*, vol. 65, pp. 386–408, 1958.

[13] T. Bach, D. Reddihough, G. Burgess, and T. Johnson, L.M. Byrt, "Comparison of subjective and objective measures of movement performance in children with cerebral palsy." *Developmental Medicine & Child Neurology*, vol. 36, pp. 974–9, 1994.