Upper Limb Rehabilitation and Evaluation of Children Using a Humanoid Robot

Douglas Brooks
Georgia Institute of Technology
School of Electrical and Computer Engineering
777 Atlantic Drive
Atlanta, Georgia
douglas.brooks@gatech.edu

Ayanna M. Howard
Georgia Institute of Technology
School of Electrical and Computer Engineering
777 Atlantic Drive
Atlanta, Georgia
ayanna.howard@ece.gatech.edu

ABSTRACT

This paper discusses a preliminary approach to matching child movements with robotic movements for the purpose of evaluating child upper limb rehabilitation exercises. Utilizing existing algorithms termed Motion History Imaging and Dynamic Time Warping for determining areas of movement and video frame mapping respectively, we are able to determine whether or not a patient is consistently performing accurate rehabilitation exercises. The overall goal of this research is to fuse play and rehabilitation techniques using a robotic design to induce child-robot interaction that will be entertaining as well as effective for the child.

Categories and Subject Descriptors

J.2 [**Physical Sciences and Engineering**]: Engineering

General Terms

Algorithm

Keywords

Motion History Image, Dynamic Time Warping, Manhattan Distance, Fuzzy Logic

1. INTRODUCTION

Mechatronic and robotic systems for neurorehabilitation can be generally used to record information about the motor performance (position, trajectory, interaction force/impedance) during active movements [5]. A relatively new sensory-motor rehabilitation technique based on the use of robotic and mechatronic devices has been applied in stroke patients [3, 9, 11–13, 17, 20, 21]. Being able to objectively assess the performance of a patient through repeatable and quantifiable metrics has shown to be an effective means for rehabilitation therapy [9, 21]. However, to date, we are unaware of any research regarding child upper limb rehabilitation techniques using robotic systems; rather, the majority of these systems are only being applied to stroke patients.

Logically, children are naturally engaged by toys, especially those that are animate. However, while there are a number of robotic toys that have been shown to be engaging to children [1, 6, 10, 14, 15], many of the studies focus solely on children with autism [1, 10]. The goal of this project is to fuse play and rehabilitation techniques using a robotic design to induce child-robot interaction that will be entertaining as well as effective for the child. Here, we discuss our approach to matching child movements with robotic movements for the purpose of evaluating rehabilitation exercises. In Section 3, we provide preliminary results of our initial implementation, and in Section 4 we discuss our thought process for future work.

2. APPROACH

2.1 Motion History Imaging

2.1.1 Background

The initial step in determining a match between the robotic upper limb movement and the human upper limb movement is to produce images from the video sequence that will give an overall representation of the recent movement. The most common technique for attaining the three-dimensional information of movement is to recover the pose of the person at each time instant using a three-dimensional model [2]. This generally requires a strong segmentation of foreground/background and also of individual body parts to aid the model alignment process [2]. Many algorithms also utilize a uniform background when processing images [8, 19]. However, it is our hope to enable the child to immediately begin interaction with the robot, regardless of the background setting, thus decreasing the required waiting time for the child.

Since we only wish to analyze the movement of specific body parts, much like Campbell and Bobick's method of analyzing human body limb positions [4], our algorithmic approach is to use temporal templates. While some algorithms utilize sequences of static configurations, which require recognition and segmentation of the person [18], we specifically form a motion-history image (MHI) to represent how motion in the image is moving. This essentially allows real-time processing of the input data.

2.1.2 Methodology

ICMI-MLMI'09 Workshop on Child, Computer and Interaction November 5, 2009 Cambridge, MA, USA

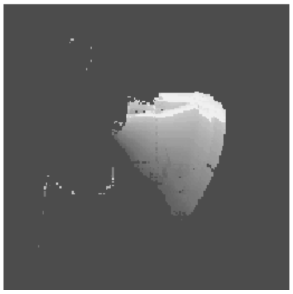

Figure 1: Motion History Image of the left arm movement.

In a MHI, H_τ, pixel intensity is a function of the temporal history of motion at that point [2]. Similar to Bobick and Davis [2], we use a replacement and decay operator, as shown in Equation (1), to obtain our MHIs:

$$H_\tau(x,y,t) = \begin{cases} \tau & if\ D(x,y,t)=1 \\ max(0, H_\tau(x,y,t-1)-1) & otherwise \end{cases} \tag{1}$$

where D is a binary image sequence indicating regions of motion. The result, as illustrated in Figure 1, is a scalar-valued image where more recently moving pixels are brighter.

Once the MHI has been determined for a specific frame, we create a feature vector for said frame by dividing the MHI into a 9x9 grid, calculating μ and σ from each grid, and then constructing an 81 length feature vector; various size grids could be used, but, after testing different grids, a 9x9 grid creates a relatively small feature vector which provides sufficient data with a low processing time. Utilizing the feature vectors of the reference image and the input image, we calculate a normalized Manhattan distance, for ease of use, between the two as shown in Equation (2):

$$d(\overrightarrow{x}, \overrightarrow{y}) = \sum_{i=1}^{n} |x_i - y_i| \tag{2}$$

where ***x*** and ***y*** are the the feature vectors of the reference and input frames. However, the normalized Manhattan distance only gives us information on a per frame basis.

2.2 Dynamic Time Warping

In order to determine that an individual is correctly performing a specific exercise, we need to utilize a sequence of movements. Furthermore, there will, undoubtedly, be variations in duration and segment lengths due to varying velocities of movements (between the reference and input sequences) that will make it difficult to determine a match. Our approach to combat this issue is to implement Dynamic Time Warping (DTW), which can be loosely defined as a nonlinear stretching or squeezing of the input sequence to line up optimally with the reference sequence.

Typically used in speech recognition, DTW has three basic operations if a left-to-right time constraint is imposed [16]:

1. Repetition of the Reference Frame: The input frame advances while the reference frame does not ("insertion").

2. Repetition of the Input Frame: The reference frame advances while the input frame does not ("deletion").

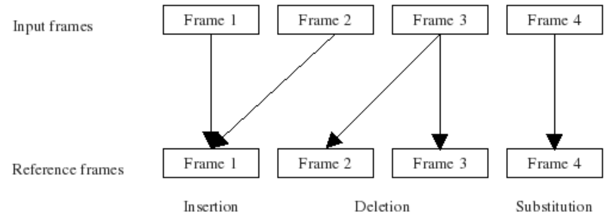

Figure 2: Pictorial representation of the basic operations of Dynamic Time Warping.

3. Both frames advance ("substitution").

Figure 2 is a pictorial representation of the operations previously described. We specifically use the Manhattan distances calculated from our feature vectors, and stored in matrix form, to determine the best possible choice between the operations for a specific frame comparison. In other words, the dtw matrix is populated by performing the following comparison:

$$dtw(m,n) = min \begin{cases} W_I d(m,n) + dtw(m-1,n) \\ W_D d(m,n) + dtw(m,n-1) \\ W_S d(m,n) + dtw(m-1,n-1) \end{cases} \tag{3}$$

where $\boldsymbol{W}_I$, $\boldsymbol{W}_D$, and $\boldsymbol{W}_S$ are weights for performing an "insertion", "deletion", or "substitution", and ***d(m,n)*** is the Manhattan distance, calculated earlier, between each frame. The weights are chosen such that choosing to perform an "insertion" or "deletion" has more of a penalty than choosing to perform a "substitution" because the data has been somewhat skewed in choosing the former. It should be noted that the standard form of a DTW implementation assumes a known starting and ending point in the sequence; however, we do not assume that we aware of the starting and ending points of the exercise. Thus, we employ DTW in order to determine a minimal exit point.

Utilizing this method, we are able to calculate the least costly path (i.e. the optimal matching sequence) and map the input sequence to the reference sequence thus minimizing the effects of varying velocities. The reader is encouraged to review [16] for a more detailed description of how to implement DTW. The results of our DTW implementation are presented in Section 3.

2.3 Fuzzy Logic

Finally, there must be some method of determining whether or not the arm trajectory in the input sequence matches, within a reasonable measure, with the arm trajectory of the reference sequence. Our approach is to utilize a ground truth from previously known images and costs and employ a fuzzy logic system for determining the degree of "correctness" of the exercise in question. The total cost is calculated by summing each of the distances of the optimal path obtained from the DTW output; logically, a high cost indicates that the exercises are not similar. In order for the total cost to be a consistent metric, we assume that the total number of frames per sequence is the same; that is to say, the initial number of reference frames is equal to the initial number of input frames.

Utilizing a ground truth sequence, testing it against multiple sequences, and establishing the means and standard

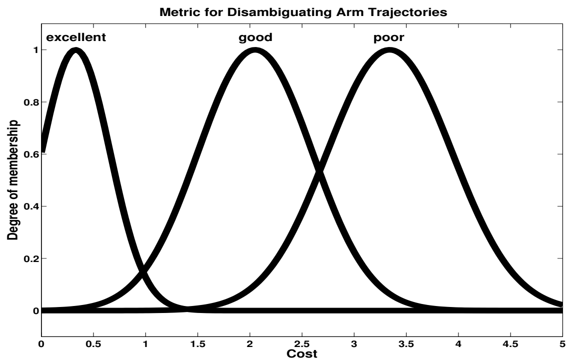

Figure 3: Metric for determining the degree of similarity between sequences

deviations of the total costs per category enables us to establish the fuzzy logic graph shown in Figure 3 as a metric for determining the degree of similarity between the input sequence and reference sequence. Once the degree of "correctness" has been determined, the humanoid will be able to inform the patient and/or physical therapist whether or not an exercise should be repeated at an attempt to achieve proper motion.

3. PRELIMINARY RESULTS

Currently, we have collected exercise data from three different participants; seven exercises for the first participant spanning a two day period, nine exercises for the second participant also spanning a two day period, and one exercise for the third participant conducted in a single session. Table 1 shows the averages and standard deviations of the costs based on the criteria set forth for generating the fuzzy logic memberships.

Table 1: μ and σ used to Create Fuzzy Logic Memberships

Classification	μ	σ
Excellent	0.332	0.336
Good	2.05	0.553
Poor	3.24	0.654

Figure 4 illustrates the results of mapping a reference sequence provided by the first participant with a separate input sequence provided by the same participant. As shown in the graph, the input sequence was significantly longer than the reference sequence; in essence, the participant performed the exercise at a faster velocity in the input sequence as opposed to the reference sequence. Given that dark blue values represent points of lower cost, an optimal path would contain the maximum number of dark blue contours reaching the end of each sequence; the black line shows the chosen path for mapping the two sequences in this specific scenario.

As a comparison, Figure 5 illustrates the results of our DTW algorithm mapping a reference sequence provided by the first participant with a separate input sequence provided by the same participant performing two different exercises.

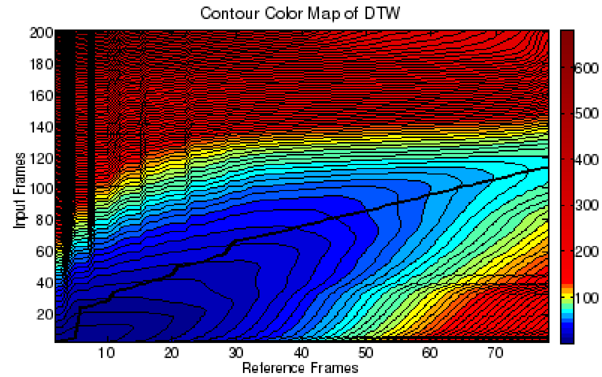

Figure 4: Contour color map of DTW illustrating the optimal path for mapping an input sequence to a reference sequence for one participant with a separate input sequence provided by the same participant.

There are two important points to note from the graph. First, although the DTW algorithm did find an "optimal" path for the two sequences, the path contains orange contours which are high values. Second, only 50 frames of the 160 possible frames for the input image provided comparable feature vectors. In other words, this would not be considered a match for movements.

4. FUTURE WORK

Although DTW enables us to map an input sequence with a reference sequence, there are still uncertainties that must be addressed. One potential issue is the variable heights and arm spans of the individuals involved in rehabilitation. The processed image of an individual with an arm span of 190cm may be quite different from someone with an arm span of 140cm; the main difference is the overall size of the contours. This variation may in turn create differences in our feature vectors and, on the whole, our DTW outputs. Currently, we foresee two directions that may combat this problem. One approach could be to incorporate the exact wingspan information gathered from the child and scale the contours of our reference frames to match that of the child's contours based upon a predefined ratio. Another approach could be to simply utilize a typical, minute range of wingspans for a target age group. This would, of course, make our algorithm less specific to the individual, but may increase robustness.

Our immediate future work is to incorporate a more intelligent scheme. First, rather than utilizing the feature vectors from the MHI output, which incorporates all areas of the image, we intend to minimize this process. A possible way to achieve this step is by using an edge detection algorithm and assuming only the largest edge detected contour is of importance (i.e. the actual area of the arm trajectory). Following this step, a shape matching algorithm such as matching by correlation or matching through Maximally Stable Extremal Regions [7] could be used for a more accurate method of quantizing the two shapes. After quantizing the two shapes (reference and input) based on this algorithm, there will still be a need to employ a more scientific metric for categorizing correct versus incorrect motions. A possible solution could be to gather information from the field of Physical Therapy.

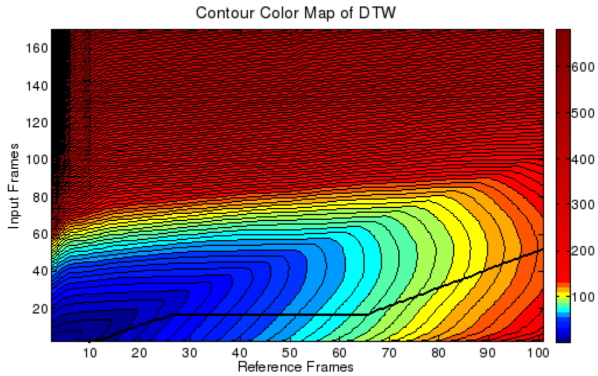

Figure 5: Contour color map of DTW illustrating the optimal path for mapping a reference sequence provided by the first participant with a separate input sequence provided by the same participant performing two different exercises.

The final step in our project is to equip the Manoi AT01, shown in Figure 6, with a small camera and a Gumstix Overo^{TM} Earth that will enable the robot to perform its movements, video capture, and image processing completely on-board. The Manoi AT01 was chosen for its robust movement capabilities and the entertaining affect that we expect it to have when interacting with children.

5. CONCLUSIONS

In this paper we have discussed an approach to matching child movements with robotic movements for the purpose of evaluating child upper limb rehabilitation exercises. Utilizing Motion History Imaging and Dynamic Time Warping for determining areas of movement and video frame mapping respectively, we were able to illustrate how a cost-based approach can be used to determine consistency in patient rehabilitation exercises.

Figure 6: Illustration of the Manoi AT01. **(www.trossenrobotics.com/Manoi-AT01-Humanoid-Robot-Kit.aspx)**

6. REFERENCES

[1] A. Billard. Play, dreams, and imitation in robotat. *Multiagent Systems, Artificial Societies, and Simulated Organizations*, 3, 2002.

[2] A. F. Bobick and J. W. Davis. The recognition of human movement using temporal templates. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 23, pages 257–267, March 2001.

[3] C. Burgar, P. Lum, P. Shor, and M. V. der Loos. Development of robots for rehabilitation therapy: the palo alto va/stanford experience. *J Rehabil Res Dev.*, 37(6):663–673, Nov-Dec 2000.

[4] L. Campbell and A. F. Bobick. Recognition of human body motion using phase space constraints. In *Proc. Int'l Conference Computer Vision*, pages 624–630, 1995.

[5] R. Colombo, F. Pisano, S. Micera, A. Mazzone, C. Delconte, M. Carrozza, P. Dario, and G. Minuco. Upper limb rehabilitation and evaluation of stroke patients using robot-aided techniques. In *ICORR 2005 Conference Proceedings*, pages 515–518, June 28 - July 1 2005.

[6] K. Dautenhahn. Robots as social actors: Aurora and the case of autism. In *Proceedings of The Third International Cognitive Technology Conference*, 1999.

[7] P. Forssen and D. Lowe. Shape descriptors for maximally stable extremal regions. In *Proc. Int'l Conference Computer Vision*, pages 1–8, October 2007.

[8] L. Goncalves, E. DiBernardo, E. Ursella, and P. Perona. Monocular tracking of the human arm in 3d. In *Proc. Int'l Conference Computer Vision*, pages 764–770, August 1995.

[9] S. Hesse, G. Schulte-Tigges, M. Konrad, A. Bardeleben, and C. Werner. Robot-assisted arm trainer for the passive and active practice of bilateral forearm and wrist movements in hemiparetic subjects. *Arch Phys Med Rehabil*, 84(6):915–920, June 2003.

[10] H. Kozima, C. Nakagawa, and Y. Yasuda. Interactive robots for communication-care: A case study in autism therapy. In *ROMAN 2005*, pages 341–346, August 2005.

[11] H. Krebs, N. Hogan, M. Aisen, and B. Volpe. Robot-aided neurorehabilitation. *IEEE Trans. Rehab Eng.*, 6(1):75–87, March 1998.

[12] R. Loureiro, F. Amirabdollahian, M. Topping, B. Driessen, and W. Harwin. Upper limb robot mediated stroke therapy gentle/s approach. *Autonomous Robots*, 15(1):35–51, July 2003.

[13] P. Lum, C. Burgar, P. Shor, M. Majmundar, and M. V. der Loos. Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Arch Phys Med Rehabil*, 83(7):952–959, July 2002.

[14] F. Michaud. Assistive technologies and child-robot interaction. In *Proceedings of AAAI Spring Symposium on Multi-disciplinary Collaboration for Socially Assistive Robotics*, pages 45–49, July 2007.

[15] F. Michaud and S. Caron. Roball , the rolling robot. *Autonomous Robots*, 12(2):211–222, March 2002.

[16] T. Parsons. *Voice and Speech Processing.* McGraw-Hill, New York, 1987.

[17] J. L. Patton and F. A. Mussa-Ivaldi. Robot-assisted adaptive training: custom force fields for teaching movement patterns. *IEEE Trans Biomed Eng.*, 21(4):636–646, April 2004.

[18] R. Polana and R. Nelson. Low level recognition of human motion. In *Proc. IEEE Workshop Non-Rigid and Articulated Motion*, pages 77–82, 1994.

[19] J. Rehg and T. Kanade. Model-based tracking of self-occluding articulated objects. In *Proc. Int'l Conference Computer Vision*, pages 612–617, 1995.

[20] D. J. Reinkensmeyer, B. D. Schmit, and W. Z. Rymer. Assessment of active and passive restraint during guided reaching after chronic brain injury. *Ann Biomed Eng.*, 27(6):805–814, Nov-Dec 1999.

[21] B. T. Volpe, H. I. Krebs, N. Hogan, L. Edelstein, C. Diels, and M. Aisen. A novel approach to stroke rehabilitation: robot-aided sensorimotor stimulation. *Neurology*, 54(10):1938–1944, May 2000.