

# A Computational Method for Physical Rehabilitation Assessment

Douglas Brooks & Ayanna M. Howard

**Abstract**—The objective of this research effort is to advance the process of quantifying physical rehabilitation techniques by developing and validating the core technologies needed to integrate therapy instruction with human-robot interaction in order to improve upper-arm rehabilitation. The method presented uses computer vision techniques such as Motion History Imaging (MHI), edge detection, and Random Sample Consensus (RANSAC) to quantify movements through robot observation. The results are compared with ground truth data retrieved via the Trimble 5606 Robotic Total Station for the purpose of assessing the efficiency of this approach.

## I. INTRODUCTION

It is essential that society as a whole continuously and persistently strives to provide the basic means toward the fulfillment of the lives of all its inhabitants, including those with disabilities [1]. Unfortunately, access to necessary assistive technology remains unequal and persons with severe or multiple physical disabilities are largely overlooked. However, recent successes in commercial robots appear to foreshadow an explosion of promising robotic applications for individuals with disabilities [2].

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 [3], [4]. Quantitative assessment of motor abilities in physically disabled individuals can provide valuable feedback to guide physical therapists during interventions. 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 [5]–[7]. The major barrier is that, to date, most assistive robotic devices are not designed for children, especially those with severe disabilities. This causes a unique challenge for deploying such robotics for this target demographic.

To overcome this barrier, state-of-the-art techniques must be created to facilitate the interaction necessary to be useful for therapeutic rehabilitation with respect to children. Utilizing the logical fact that animate toys naturally engage children, this research focuses on the design of a robotic therapeutic playmate that will aid children in physical rehabilitation by fusing play and rehabilitation techniques that are both entertaining for the child and effective for upper-arm rehabilitation [2]. In this paper, we present the initial phase of developing this robotic system to aid in physical rehabilitation by presenting a novel method for automating

and quantizing two physical therapeutic metrics, obtained from the Fugl-Meyer test, namely Range of Motion (ROM) and Peak (angular) Velocity.

In performing the Fugl-Meyer assessment, physical therapists administer a series of tests, which include reflex activity, balance, sensation, position sense, ROM and peak velocity, for evaluating a patient's degree of impairment [8]. Since the research presented here focuses on non-touch, upper-arm rehabilitation, the Fugl-Meyer assessment is limited to ROM and angular velocity, for now. We specifically present an approach which solely utilizes computer vision techniques for accurately determining these metrics for a patient through a non-touch exercise scenario. The use of a simple webcam based system would drastically reduce the cost of determining a patient's physical abilities as opposed to utilizing motion analysis systems such as magneto-inertial and stereo-photogrammetric systems. The method presented uses Motion History Imaging (MHI), edge detection, and Random Sample Consensus (RANSAC) to quantify movements through robot observation. Ground truth data, provided via the Trimble 5606 Robotic Total Station, was used as a baseline measurement for determining the efficiency of our approach.

Section II gives a detailed account of the methodology used in this research to achieve the aforementioned goals of attaining appropriate physical therapy metrics. Section III is a presentation of the results that were obtained when applying this algorithm, and Section IV provides a discussion of the overall efficiency. Finally, Section V concludes the article with a discussion of the advantages and disadvantages and future direction of this research.

## II. APPROACH

In order to achieve the goal of obtaining physical therapy metrics by using computer vision as the only means of collecting data, our approach utilizes several well-known image processing techniques. First, the video sequence is partitioned into separate grayscale images containing information regarding human movement via a process termed Motion History Imaging. Then, a contour extraction process is applied to each of the remaining images, which creates an ideal representation of the two movements. Following, a Random Sample Consensus is applied to the ideal contours, which enables the determination of straight line segments ultimately allowing the calculation of the ROM. Finally, the angular velocity is calculated by applying physics equations to already known angles and frame rate.

D. Brooks and A. Howard are with the HumAnS Lab at Georgia Tech in the department Electrical and Computer Engineering, 85 Fifth Street NW, Atlanta, GA 30308 (email: douglas.brooks@gatech.edu, ayanna.howard@ece.gatech.edu)

## A. Motion History Imaging

1) *Background:* The initial step for recognizing the patient's movements is to segment a video sequence into individual images that contain pertinent information as they relate to the overall representation of recent movement [9]. One common technique for attaining the three-dimensional information from a particular movement is to recover the pose of the person at each time instant using a three-dimensional model [10]. This generally requires a strong segmentation of foreground/background and also of individual body parts, in congruence with background uniformity [11], [12]. However, in this work, it is desirable to enable human-robot interaction and data collection immediately rather than require idle waiting time during an initiation process. As such, the use of Motion History Imaging

Since the purpose of this work is to analyze the movement of specific body parts, similar to [13], the algorithmic approach is to use temporal templates. While some algorithms utilize sequences of static configurations, which require recognition and segmentation of the person [14], here, a Motion History Image (MHI) to represent how motion in the image is moving is specifically formed. This essentially allows real-time processing of the input data.

2) *Methodology:* In a MHI,  $H_\tau$ , pixel intensity is a function of the temporal history of motion at that point in physical space [10]. Similar to [10], a replacement and decay operator is used, as shown in Equation (1), to obtain the MHIs:

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

where  $D$  is a binary image sequence indicating regions of motion,  $x$  and  $y$  are the horizontal and vertical directions in the image, respectively,  $t$  is the current time step, and  $\tau$  is the current intensity value. The result, as illustrated in Figure 1, is a scalar-valued image where more recently moving pixels are brighter in intensity.

Once the patient's movements have been effectively represented using MHI, the next step invokes extracting the overall contour of the movement for quickly and efficiently quantifying the physical therapy metrics previously mentioned.

## B. Contour Extraction

Once the patient's movements have been recognized, a contour representing the shape of the movements is then extracted. The first step in this process is to use a median filter that will remove smaller, unwanted contours in the image typically caused by camera jitter or human inaccuracies (i.e. movements of body parts other than the desired limb). The median filter is a sliding-window spatial filter that replaces the center pixel value in the window with the median of all the pixel values in the window, and it can be of any central symmetric shape. Here, a 16x16 square window was used. Figure 2b illustrates the use of a median filter on the image data.

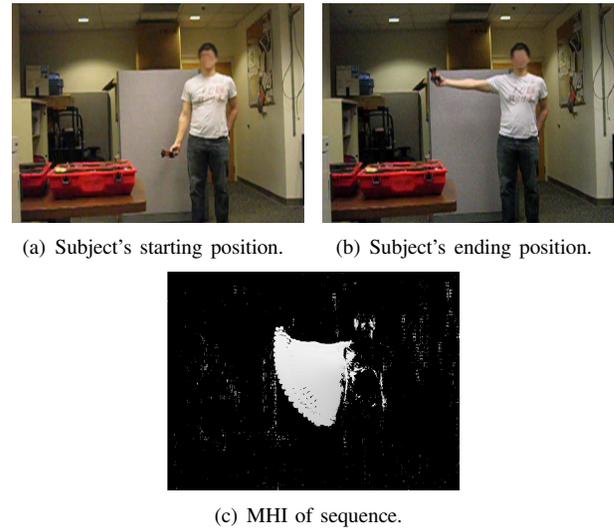


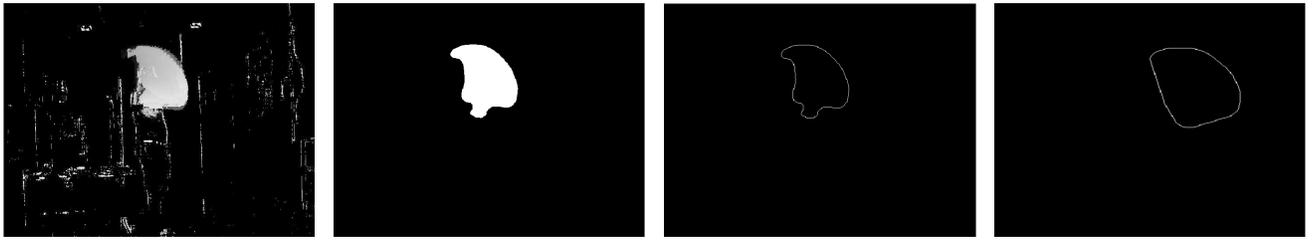
Fig. 1. Motion History Image of the left arm movement.

When the smaller contours have been removed, a canny edge detection algorithm [15] is utilized in order to extract the edges of the contour representing the upper-arm movement, see Figure 2c. Utilizing the edge detected shape, a proper representation of the sequence is then created. Again, due to camera inaccuracies there will undoubtedly be areas in the image where actual movement is not properly represented, even after the initial filtering and edge detection processes. Figure 2a illustrates an inaccuracy caused by camera lag time; even though the subject's arm moved in one complete motion, the video sequence shows a gap between arm positions. Therefore, the convex hull of the edge detected image is calculated and utilized.

The convex hull can be thought of as the boundary of a minimal convex set of points containing a given non-empty finite set of points in a plane. Here, we utilize a simple polygon model in order to construct our convex hulls. By looking at three consecutive vertices of the polygon, during a recursive progression around the polygon, this algorithm simplifies to determining whether the resulting angle between the three vertices is concave or convex. If the resulting angle is concave, then the middle point is removed and the next (along the polygon) vertex is added to the triple to be tested. If the angle is convex, then each of the points in the triple is shifted by one vertex along the polygon [16]. Melkman's Algorithm was employed to ensure correct outcomes [17]. This implementation gives a somewhat ideal outline as seen in Figure 2d.

## C. RANSAC

Now that a somewhat ideal outline has been obtained, determining the best method for finding the range of motion using only the image data is needed. Utilizing the major axis as a symmetrical dissection of the polygon and employing a Hough Transform [18] on either the upper or lower region of the contour could enable a determination of the upper or lower lines for the purpose of finding the angle between



(a) Original image obtained from the MHI process. (b) Image obtained from the Median Filter. (c) Image obtained from the Canny Edge Detection algorithm. (d) Image obtained from the Convex Hull.

Fig. 2. Example image processing sequence used to extract an ideal contour from the human's movements.

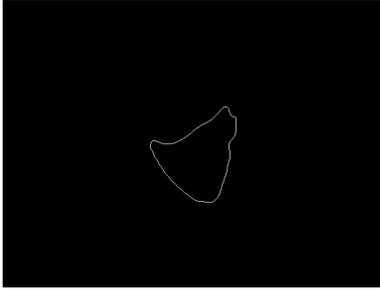


Fig. 3. Illustration of the original convex hull image.

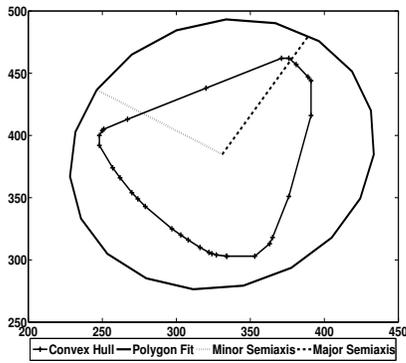


Fig. 4. Illustration of the Major (dashed line) and Minor (dotted line) semi-axes located on a contour obtained from human upper-arm movement. The polygon boundary used to calculate the two semi-axes is shown with square markers.

either line and the major axis. The Hough Transform is a method used in computer vision to detect simple shapes, such as straight lines, by using the parameters of a line,  $y = mx + b$  and representing the slope and intercept in parameter space  $(b, m)$ . However, after much deliberation and testing, it was decided that since the Hough Transform merely makes estimations of the best possible line to fit the upper or lower region, a more accurate approach would be beneficial. Therefore, it was decided to use the RANdom SAMple Consensus (RANSAC) algorithm [19].

RANSAC determines the best possible line fit by iteratively selecting a random subset of the original input data and returns points from the original input data that are inliers. Given a set of data points  $U$ , there is an unknown number of data points that are consistent with the model with unknown parameters from parameter space  $\Theta$ . These data points are inliers, and all others are outliers. The goal is to find model parameters  $\theta^*$  from a parameter space  $\Theta$  that maximizes a

cost function  $J_S(\theta, U, \Delta)$  [20]. Here, RANSAC is employed in the following manner:

- 1) The input data is provided via the upper and lower regions provided by the convex hull and the major or minor axis (dissection line), depending upon the orientation.
- 2) A subset of the data is randomly selected from  $U$ .
- 3) The model parameters are estimated in order to fit the sample created from the previous step.
- 4) The support value (cost) of the model is calculated.
- 5) If the current support is greater than the previously calculated support, then store the current model parameters.
- 6) The process is repeated until the probability of finding a model with support larger than  $I_k^* - 1$  in the  $k$ -th step falls under a predefined threshold  $\eta_0$ , where  $I_k^*$  is the largest support of a hypothesized model found up to the  $k$ -th sample inclusively.
- 7) The inlier values are returned and are used to create the upper and lower straight line segments of the convex hull.

The basic structure can be found in [19].

This method is a more accurate approach, for our needs, than the Hough Transform because it only returns points from the original input data (inliers) rather than creating its own values when predicting the line segment. Also, rather than using the major axis as one of the two lines used to find the range of motion, it was determined that a more accurate measure would be to perform RANSAC on the upper and lower regions (created by the major axis's dissection of the contour), thus creating a measure of the highest and lowest positions of the subject's arm. Figures 3, 4, and 5 illustrate this process.

#### D. ROM

Once the points that create the upper and lower lines are recognized, the slopes of each are used to calculate the angle between the two lines via simple geometry, shown in Equation 2:

$$\begin{aligned}
 m_1 &= \frac{y_2 - y_1}{x_2 - x_1} \\
 m_2 &= \frac{y_4 - y_3}{x_4 - x_3} \\
 \Theta &= \arctan\left(\frac{m_2 - m_1}{1 + m_2 * m_1}\right) \\
 ROM &= \left| \Theta \frac{180^\circ}{\pi} \right|
 \end{aligned} \tag{2}$$



Fig. 5. Lower and Upper lines, determined by RANSAC, used to find the ROM.

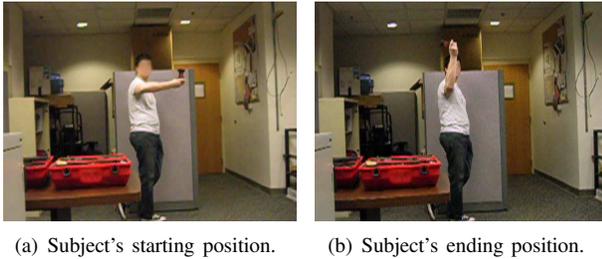


Fig. 6. Lateral and medial movement demonstrated by participant.

where  $x$  and  $y$  are the coordinates of points on each line segment and  $m_1$  and  $m_2$  are the slopes of each line, respectively. The maximum angle found over the length of the video sequence gives the range of motion of the patient's movements.

### E. Angular Velocity

Given that the frame rate of the camera used to capture the patient's movements was 15fps, calculating the angular velocity of the arm was trivial. We chose to use the initial lower line and subsequent upper lines that were recognized via RANSAC to determine the angular velocity as it relates to each frame. Meaning, the lower line found during the first RANSAC calculation over the convex hull in frame one of the video sequence was used as the initial position of the arm, while the current upper line changed as the subject moved his arm upward during the exercise. Using the standard equation for angular velocity, shown in Equation (3), the angular velocity of the patient can easily be determined with the angle obtained from each pass of the ROM calculation and known frame rate.

$$\omega = \frac{d\theta}{dt} \quad (3)$$

## III. RESULTS

For the initial testing of the methodology of this research, two subjects were utilized, both male adults. The subject was asked to perform a series of upper arm exercises, which were captured via a simple webcam. The specific exercises involved adduction and abduction, shown in Figure 1, and lateral and medial movements, shown in Figure 6. The images were then processed by our algorithm in

TABLE I  
GROUND TRUTH VERSUS ALGORITHMIC MEASUREMENTS OF ROM FOR ADDUCTION/ABDUCTION EXERCISE FOR SUBJECT A.

Repetition	1	2	3
Ground Truth (Degrees)	89.9979	41.3106	16.5313
Algorithm (Degrees)	89.5342	40.2887	16.6658
Difference (Degrees)	0.4637	1.0219	0.1345
Error (Percentage)	0.5152	2.4736	0.8136

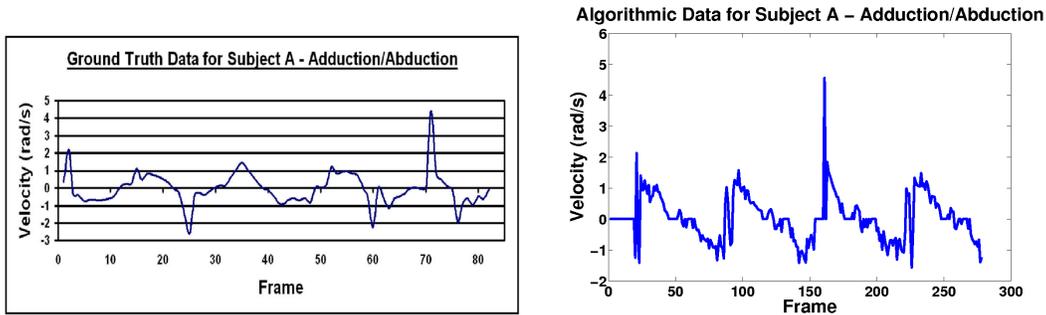
TABLE II  
GROUND TRUTH VERSUS ALGORITHMIC MEASUREMENTS OF ROM FOR ADDUCTION/ABDUCTION EXERCISE FOR SUBJECT B.

Repetition	1	2	3
Ground Truth (Degrees)	86.8948	61.7268	19.1962
Algorithm (Degrees)	90.0000	56.4174	16.6992
Difference (Degrees)	3.1052	5.3094	2.4970
Error (Percentage)	3.5735	8.6014	13.0077

order to obtain the ROM and angular velocity, which then was compared to the ground truth data captured via the Trimble 5606 Robotic Total Station. The Trimble 5606 uses a time-of-flight measurement technique based on the pulse measurement principle; it measures the time for a very short transmitted pulse to travel to a targeted prism, held by the subject, and back, thus calculating the position of the subject's end-effector. Tables I, II, III, and IV show the ROM comparison between ground truth and algorithmic data for each subject per exercise. Figures 7 and 8 illustrate the peak angular velocity for the ground truth versus our algorithm.

TABLE III  
GROUND TRUTH VERSUS ALGORITHMIC MEASUREMENTS OF ROM FOR LATERAL/MEDIAL EXERCISE SUBJECT A.

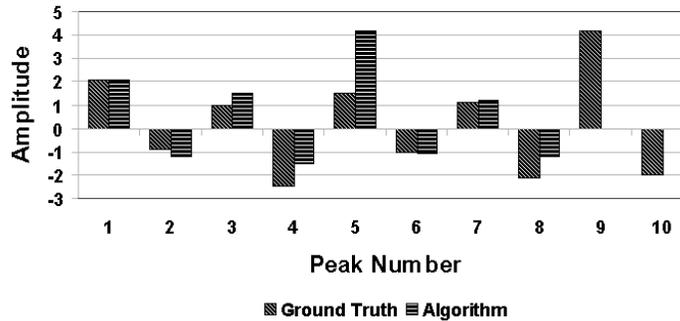
Repetition	1	2	3
Ground Truth (Degrees)	87.8455	61.0102	14.2693
Algorithm (Degrees)	81.2742	58.1582	12.3986
Difference (Degrees)	6.5713	2.8520	1.8707
Error (Percentage)	7.4805	4.6746	13.1099



(a) Ground Truth Data for Subject A - Adduction/Abduction. (b) Algorithmic Data for Subject A - Adduction/Abduction.

Fig. 7. Ground truth and algorithmic angular velocity for subject A performing the Adduction/Abduction exercise.

### Ground Truth vs. Algorithmic Peak Angular Velocity Data



(a) Peak Angular Velocity comparison for Subject A - Adduction/Abduction.

Fig. 8. Comparison of Ground truth and algorithmic peak angular velocity for subject A.

TABLE IV  
GROUND TRUTH VERSUS ALGORITHMIC MEASUREMENTS OF ROM FOR  
LATERAL/MEDIAL EXERCISE SUBJECT B.

Repetition	1	2	3
Ground Truth (Degrees)	88.9327	24.9530	14.1838
Algorithm (Degrees)	89.8538	23.3008	14.7707
Difference (Degrees)	0.9211	1.6522	0.5869
Error (Percentage)	1.0357	6.6212	4.1378

#### IV. ANALYSIS

As shown in the tables, the ROM values calculated via our algorithm are closely related to the ground truth data (average error < 6%). For a patient with a limited range of motion, our algorithm could be used to identify this condition in real-time, given a known standard ROM. This will allow the system to monitor patient progress between sessions.

The angular velocity calculated via our algorithm, shown in Figure 7b, is related to the trend of those calculated via ground truth data, shown in Figure 7a. When analyzing the data, one should note that the Trimble 5606 captures data at a rate that is equivalent to 4fps, which is a bit on the slow side, while our camera captured data at a rate of 15fps. Therefore, there is approximately four times the amount of data points for our graphs then there are for the ground truth graphs. This gives some explanation to the variation in frames for each

set of measurements. It should also be noted that because the Trimble 5606 is a real tracking system, and human motion is not ideal, instances where the patient is not moving may not be conveyed with zero velocity in the ground truth data; thus values that are approximately zero are categorized as non-movement. The other important aspect to note is that the magnitude of the velocity calculations in the ground truth data matches closely with that of our algorithm.

#### V. CONCLUSION AND FUTURE WORKS

In this paper, an approach to evaluate patient movements using robotic observation has been discussed. Specifically, the physical therapeutic metrics range of motion and angular velocity used in the Fugl-Meyer test have been calculated via computer vision techniques and can be utilized in a robotic system. While the efficiency of our method can be increased with the addition of extra sensors (e.g. retro-reflective markers in passive optical systems), our approach performs at a high degree of efficiency with a much lower cost. One disadvantage of our system is the fact that the exercises must be in planar view of the camera; thus the robot or patient will have to position himself or herself accordingly.

The immediate future work for this research is to incorporate another physical therapy metric, which is movement smoothness. This will enable the final robotic system to better assess the physical disability of the patient through the quantification of measurements received during a physical



Fig. 9. Illustration of the Manoi AT01.

therapy scenario, which can be reported to a physical therapist for further evaluation. The final step of this research is to equip a robotic platform, shown in Figure 9, with a small camera and Gumstix Overo<sup>TM</sup> Air, a small computer-on-module, that will enable the robot to perform its movements, video capture, and image processing completely on-board [9]. Again, the motivation for this work is to create a robotic playmate for children with severe disabilities, thus keeping them engaged during rehabilitation. Currently, our vision processing is done on a separate host computer. Once this phase is complete, all aspects of the observation and analysis will be accomplished via the robot, which will interact directly with physically injured/disabled individuals in realistic physical therapy scenarios.

## VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of the individuals in the HumAnS Lab, Dr. Patricio Vela, Dr. Jochen Teizer, Miguel Serrano, and NASA JFPF for supporting this work.

## REFERENCES

- [1] B. Wright, *Physical disability-A psychosocial approach*. Harper and Row New York, 1983.
- [2] D. Brooks and A. M. Howard, "Quantifying Upper-Arm Rehabilitation Metrics for Children through Interaction with a Humanoid Robot," *Journal of Applied Bionics and Biomechanics*, 2010 (submitted).
- [3] G. Kwakkel, B. Kollen, and H. Krebs, "Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review," *Neurorehabilitation and Neural Repair*, vol. 22, no. 2, p. 111, 2008.
- [4] C. Takahashi, L. Der-Yeghiaian, V. Le, R. Motiwala, and S. Cramer, "Robot-based hand motor therapy after stroke," *Brain*, 2007.
- [5] M. Hurley, "Effectiveness of robot-assisted therapy on stroke patients with upper extremity impairment," *Physical Function*, p. 14, 2009.
- [6] M. Casadio, P. Giannoni, P. Morasso, and V. Sanguineti, "A proof of concept study for the integration of robot therapy with physiotherapy in the treatment of stroke patients," *Clinical Rehabilitation*, vol. 23, no. 3, p. 217, 2009.
- [7] M. Federico Posteraro, S. Mazzoleni, S. PhD, and B. MD, "Robot-mediated therapy for paretic upper limb of chronic patients following neurological injury," *REHABILITATION MEDICINE*, vol. 41, pp. 976–980, 2009.
- [8] P. Duncan, M. Propst, and S. Nelson, "Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident," *Physical Therapy*, vol. 63, no. 10, p. 1606, 1983.
- [9] D. Brooks and A. M. Howard, "Upper-limb Rehabilitation and Evaluation of Children using a Humanoid Robot," in *ICMI-MLMI09 Workshop on Child, Computer and Interaction*, November 2009.
- [10] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, March 2001, pp. 257–267.
- [11] L. Goncalves, E. DiBernardo, E. Ursella, and P. Perona, "Monocular tracking of the human arm in 3d," in *Proc. Int'l Conference Computer Vision*, August 1995, pp. 764–770.
- [12] J. Rehg and T. Kanade, "Model-based tracking of self-occluding articulated objects," in *Proc. Int'l Conference Computer Vision*, 1995, pp. 612–617.
- [13] L. Campbell and A. F. Bobick, "Recognition of human body motion using phase space constraints," in *Proc. Int'l Conference Computer Vision*, 1995, pp. 624–630.
- [14] R. Polana and R. Nelson, "Low level recognition of human motion," in *Proc. IEEE Workshop Non-Rigid and Articulated Motion*, 1994, pp. 77–82.
- [15] J. Canny, "A computational approach to edge detection," *Readings in Computer Vision: Issues, Problems, Principles and Paradigms*, vol. 8, no. 6, pp. 184–203, 1986.
- [16] T. Cormen, C. Leiserson, and R. Rivest, "0-262-03141-8," 1998.
- [17] A. Melkman, "On-line construction of the convex hull of a simple polyline," *Information Processing Letters*, vol. 25, no. 1, pp. 11–12, 1987.
- [18] R. Duda and P. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [19] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [20] O. Chum, "Two-view geometry estimation by random sample and consensus," Ph.D. dissertation, Czech Technical University, 2005.