

RESEARCH ARTICLE

Quantifying Upper-Arm Rehabilitation Metrics for Children through Interaction with a Humanoid Robot

Douglas A. Brooks* & Ayanna M. Howard

Email: douglas.brooks@gatech.edu ayanna.howard@gatech.edu

Phone: 404.385.4066 404.385.4824

Georgia Institute of Technology

85 Fifth Street NW, Atlanta, GA 30308

(v2.0 released August 2011)

The objective of this research effort is to integrate therapy instruction with child-robot play interaction in order to better assess upper-arm rehabilitation. Using computer vision techniques such as Motion History Imaging (MHI), edge detection, and Random Sample Consensus (RANSAC), movements can be quantified through robot observation. In addition, incorporating prior knowledge regarding exercise data, physical therapeutic metrics, and novel approaches, a mapping to therapist instructions can be created allowing robotic feedback and intelligent interaction. The results are compared with ground truth data retrieved via the Trimble 5606 Robotic Total Station and visual experts for the purpose of assessing the efficiency of this approach. We performed a series of upper-arm exercises with two male subjects, which were captured via a simple webcam. The specific exercises involved adduction and abduction and lateral and medial movements. The analysis shows that our algorithmic results compare closely to the results obtain from the ground truth data, with an average algorithmic error is less than 9% for the range of motion and less than 8% for the peak angular velocity of each subject.

Keywords: Rehabilitation; Humanoid; Motion History Imaging; Dynamic Time Warping; Random Sample Consensus

1. Introduction

In the United States, the Individuals with Disabilities Education Act (IDEA) states that children with a physical disability are entitled to a free public education that emphasizes special education and related services designed to meet their unique needs and prepare them for further education, employment, and independent living. 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. 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 (Kwakkel et al 2008; Takahashi et al 2007; Colombo et al 2005). Sensory-motor rehabilitation techniques based on the use of robotic and

*Corresponding author. Email: douglas.brooks@gatech.edu

mechatronic devices have been applied with stroke patients (Krebs et al 1998; Reinkensmeyer et al 1999; Burgar et al 2000; Volpe et al 2000; Lum et al 2002; Hesse et al 2003; Loureiro et al 2003; Patton and Mussa-Ivaldi 2004; Mehrholz et al 2009). Furthermore, 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 (Volpe et al 2000; Hesse et al 2003; Hurley 2009; Casadio et al 2009; Federico Posteraro et al 2009; Brewer et al 2007). The major barrier is that, to date, most assistive robotic devices are not designed for children.

Studies in child cognitive behavior indicate that with repetitive and monotonous conditions over time, performance decrement occurs due to a reduction in arousal (Cooley and Morris 1990). This, in-turn, leads to a decrease in sustained attention for the current activity and overall motivation for continued repetitions of the exercises. Given that physical rehabilitation sessions consists of many repetitive and monotonous actions and sustained attention improves with age (Cooley and Morris 1990), it is important that researchers incorporate various methods for motivating children during this process.

There are several reasons that may be fueling the lack of necessary technological devices geared towards aiding children, both with and without disabilities. In general, researchers must first and foremost obtain the approval of their Institutional Review Board (IRB) before conducting any studies involving humans. For those who wish to conduct studies involving children, researchers must additionally obtain parental consent. This alone may cause a hindrance for collecting data from children that could ultimately lead to groundbreaking technology for this particular group. When considering children with disabilities, parents are typically even more protective and more reluctant to allow their child to participate in such studies. Furthermore, because disabilities vary from one individual to the next, it becomes quite challenging to design an assistive aid that is capable of providing assistance that is unique to each person. For instance, even within the realm of cerebral palsy there are various types and levels. One may have an issue of spasticity, where the muscles are very stiff and causes abnormal or jerky movement. However, a separate individual may have full motor functionality, but has an issue with speech. These variations cause a unique challenge for deploying 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. The robotic playmate, consistent with prior theory from physical therapy metrics, will incorporate visual queues that will determine the amount of movement that a child makes and his or her consistency with respect to the movement of the robot. Of particular importance within this proposed work are approaches that allow therapists to assess the information obtained by the robot during play scenarios, thus requiring that the data is conveniently decipherable.

In this article, several successful trials of employing this novel approach to physical therapy scenarios will be illustrated. Section 2 gives a brief overview of the current state of the field of therapeutic robots. Section 3 provides an overview of the current methods that are used by physical therapists to administer upper-arm rehabilitation sessions. Section 4 gives a detailed account of the methodology used in this research to achieve the aforementioned goals of attaining appropriate physical therapy metrics via image processing. Sections 5 and 6 are a presentation of

the results and the advantages and disadvantages of this approach, respectively. Finally, Section 7 concludes the article with a discussion of the future direction of this research.

2. Therapeutic Robotics

The use of robotics to aid in therapy is an on-going area of research. Scientists continue to design new methods for administering aid to those in need of cognitive and physical development. Even though no robot can replicate the knowledge and experience of a therapist in assessing the needs of the patient and the outcomes of the therapeutic program, robotic technology can be used to carry out the repetitive practice that is needed to facilitate functional gains (Lum et al 2002). Whether using these new methods to assist with developing communication skills or aiding in patient motor development, researchers have shown that the use of robots can greatly improve patient therapy sessions.

2.1. *Keepon*

One robotic platform that has been developed for therapy and play is *Keepon*. Introduced by researchers and physicians in Japan, *Keepon* was designed to get autistic and non-autistic children (ages 2-4) involved in playful interaction that was generally initiated and directed by the child (i.e. without any experimental setting or instruction) (Kozima et al 2005). A point of emphasis for the *Keepon* design was to keep it as simple as possible in order to abstain from overwhelming and/or frightening the children, a notion that all researchers working with therapeutic robots should keep in mind. Therefore, the simple structure and function only contained two color CCD cameras as eyes, a microphone as a nose, a small gimbal and four wires, by which the body is manipulated like a marionette, four motors and two circuit boards (an SH2-based PID controller and a motor driver). Since the body is made of silicone rubber and its inside is relatively hollow, *Keepons* head and belly deform whenever it changes posture and when people touch it (Kozima et al 2005).

The resulting information, obtained from a year and a half of data collection, indicated that *Keepons* simple appearance and predictable responses gave the autistic children a playful and relaxed mood, in which they spontaneously engaged in dyadic play with *Keepon*. This in turn expanded into interpersonal communication where *Keepon* worked as the pivot of triadic play with adults or other children (Kozima et al 2005).

2.2. *CosmoBot*

Another robotic system that has been designed for therapy, education, and play is the *CosmoBotTM*. Children interact with the *CosmoBotTM*, controlling the robots movements and audio output, using a variety of gestural sensors and speech recognition, while actively targeting their therapy goals (Brisben et al 2005). The system consists of the physical robot, an interface board, interface software, and plug-in gestural interfaces (joystick, head sensor, and arm sensor).

Researchers recruited six children who were already receiving outpatient physical and occupational therapy services with a physical therapist. More specifically, the subjects ranged in age from 4 – 10 years old, and each subject had a diagnosis of cerebral palsy and received physical and/or occupational therapy for treatment of

upper extremity motor deficits. The physical therapist provided feedback regarding the system, noting that it was easy to use and that it provided a motivating factor for client participation. Furthermore, researchers reported improvements in the physical abilities of patients, reporting patient novelties such as multi-tasking, improved attention span, improved upper extremity strength and coordination, and an improvement in daily activities of living (Brisben et al 2005).

2.3. Moving Forward

These humanoids provide a brief overview of state-of-the-art robotic platforms used in therapy. While these robotic systems are among the first to make significant progress toward aiding individuals with severe disabilities, there is still much that needs to be solved. First, neither of these designs give a physical assessment of the individuals' motor skills. Even though children with disabilities may never have full usage of certain body parts, significant improvements can be made in hopes that these individuals will be able to lead healthy, self-contained lifestyles. Also, of the systems that do incorporate a physical assessment of the individual, only few, if any, are fully autonomous with a completely on-board system; each has to connect to a computer which incorporates its intelligence system or is teleoperated.

As an example, MIME or Mirror-Image Motion Enabler (Burgar et al 2000), was designed to aid in the physical rehabilitation of stroke victims. This platform functions by strapping the subject into a wheel-chair that has mobile arm supports, which limit horizontal movement, and a robotic arm that applies forces and torques to the arm through those supports. The desire is to monitor patient recovery through parameters obtained via force-feedback. Another robotic platform with the capability of providing an assessment of a patient's physical attributes was introduced by Gockley and Matarić in 2006. The researchers in this project attached a laser range finder to a Pioneer-2DX robotic platform that was used to track and locate a patient wearing a reflector on his or her leg. The researchers actually used the system as a way of encouraging the patient to continue the rehabilitation session through positive reinforcement, which was provided via a series of beeps and movements based on the patient's movements. As a control variable, they varied the robot's level of engagement (or amount of movement) and they reported the patients' level of engagement and comfort with the robot (Gockley and Matarić 2006).

With the introduction of these robotic systems, we have research showing approved attention span and, separately, contact assistive robotics. By bridging the gap between these two research areas, we can design a system that allows a patient to move freely while keeping them engaged and motivated during a continuous rehabilitation session. Furthermore, since the progression of robotics has been to achieve full autonomy, it is ideal to produce systems with the capabilities of making intelligent decisions and performing actions without the aid of a separate system, be it human or computer. This work is an attempt to alleviate these issues.

3. Current Physical Therapy Metrics

Until recently, many of the current physical therapy metrics were not designed with the intent to administer such practices to children with disabilities. For instance, the Fugl-Meyer Test (Fugl-Meyer et al 1975), an assessment that consists of 155 items and is used to determine the impairment in the upper and lower extremities, was designed specifically for stroke victims. Utilizing a scoring scale of 226, phys-

ical therapists administer a series of tests, which include reflex activity, balance, sensation, position sense, and range of motion (ROM) for evaluating a patient's degree of impairment (Duncan et al 1983). However, this scoring scale is based upon the movement expectations of a healthy adult subject.

The lack of a legitimate metric of child motor development lead therapists to design scales specifically for assessing the aforementioned group. The Gross Motor Function Measure (GMFM) (Russell et al 1989) and the Peabody Developmental Motor Scales (PDMS) (Folio and Fewell 1983) are the two well-known motor instruments for children, specifically those diagnosed with cerebral palsy (CP). However, the GMFM measures the gross motor (GM) domain only (Russell et al 1989). For measurement of the fine motor (FM) domain, the GMFM is inadequate as an evaluative tool (Wang et al 2006). The opposite is true of the original PDMS.

During the late 1990's, with the culmination in 2000, the PDMS was revised to the Peabody Developmental Motor Scales–Second Edition (PDMS-2), with new norms, revised testing materials, more precise scoring criteria, and more information on norm samples (Folio and Fewell 2000). Scores include 1) a Gross Motor Quotient which is a composite of the Reflexes, Stationary, Locomotion, and Object Manipulation subtests, 2) a Fine Motor Quotient, a composite of the Grasping and Visual-Motor Integration subtests, and 3) a Total Quotient, a combination of the gross and motor subtests (Folio and Fewell 2000). The new normative data on the PDMS-2 was collected on 2,003 children both with and without disabilities. There are also more reliability and validity data for the PDMS-2 than for the PDMS (Folio and Fewell 2000). Therefore, the PDMS-2 is potentially appropriate for investigating the progress of the gross and fine motor domains for children with and without disabilities because it assesses both GM and FM composites and incorporates both quantitative and qualitative rating criteria (Wang et al 2006).

Since the research presented in this work focuses on non-contact, upper-arm rehabilitation, specifically for the shoulder joint, we present the successful analysis of a typical metric used by physical therapist, ROM. As an extension of ROM, we also quantify the patients' peak angular velocity (PAV) for the purpose of providing a more accurate quantitative analysis. These types of measurements can coincide with a child's ability to catch and throw objects as well as his or her reaction to environmental events; each of those items are within the object manipulation and reflexes subtests of the PDMS-2. Simultaneously, we desire to provide a new method for obtaining this data, namely a way to keep the child subjects motivated during the process.

4. Approach

In order to achieve the goal of obtaining physical therapy metrics by using computer vision as the only means of collecting data, several image processing techniques are employed. First, the video sequence is partitioned into separate gray-scale images containing information regarding human movement via a process termed Motion History Imaging. Then, the incoming data is aligned with the robot's notion of its own motion (i.e. a set of images representing the robot's movements). Next, a contour extraction process is applied to each of the remaining images, which creates an ideal representation of the two movements. Finally, 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; adding the known factor of the number of frames per second (fps), the PAV is also calculated.

4.1. Motion History Imaging

4.1.1. Background

The initial step in determining the consistency between robotic movements and child movements is to segment a video sequence into individual images that contain pertinent information as they relate to the overall representation of recent movement (Brooks and Howard 2009). 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 (Bobick and Davis 2001). This generally requires a strong segmentation of foreground/background and also of individual body parts to aid the model alignment process. Furthermore, background uniformity is typically a necessity when processing images for motion estimation (Goncalves et al 1995; Rehg and Kanade 1995). However, in this work, it is desirable to enable child-robot interaction and data collection immediately rather than require the child to wait during an initiation process.

Since the purpose of this work is to analyze the movement of specific body parts, similar to Campbell and Bobick (Campbell and Bobick 1995), the algorithmic approach is to use temporal templates. While some algorithms utilize sequences of static configurations, which require recognition and segmentation of the person (Polana and Nelson 1994), 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.

4.1.2. Methodology

In a MHI, pixel intensity, H_τ , is a function of the temporal history of motion at that point in physical space (Bobick and Davis 2001). Similar to Bobick and Davis, 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 τ is the current intensity value (Bobick and Davis 2001). The result, as illustrated in Figure 1c, is a scalar-valued image where more recently moving pixels are brighter in intensity.

Once the MHI has been determined for a specific frame, a feature vector is created for that frame by dividing the MHI into an 8x8 grid, calculating the mean, μ , and standard deviation, σ , from each grid, and then constructing a 64 length feature vector. Various size grids could be used, but it is ideal to adhere to the law of diminishing returns by utilizing a feature vector that will allow the most features for image recognition without compromising a significant amount of processing time (i.e. minimize the classification error using the appropriate feature vector while maximizing the classification rate of the classifier). Figure 2 and Table 1 show the results of testing different size grids, which indicate that a 64 length feature vector is ideal.

Utilizing the feature vectors of the reference image and the input image, a normalized Manhattan distance is calculated, for ease of use, between the two as shown

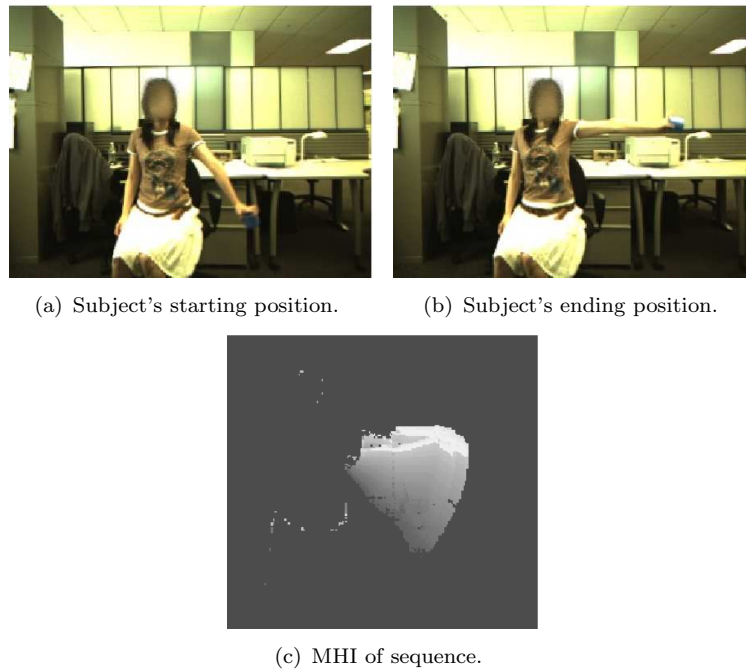


Figure 1. Motion History Image of the left arm movement.

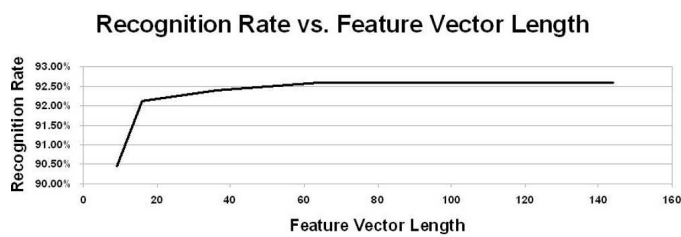


Figure 2. Graphical representation of feature vector selection.

Table 1. Resulting Processing Times and Recognition Rates for Various Feature Vector Lengths

Feature Vector Length	Processing Time (s)	Recognition Rate(%)
9	76.59	90.45
16	80.96	92.11
36	83.35	92.40
64	86.06	92.59
81	89.58	92.59
144	93.17	92.59

in Equation (2):

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

where x and y are the the feature vectors of the reference and input frames and d is the Manhattan distance. However, the normalized Manhattan distance only gives information on a per frame basis. Once the child's movements have been effectively represented using MHI, the next step invokes aligning that movement with the robot's perception of its own movement.

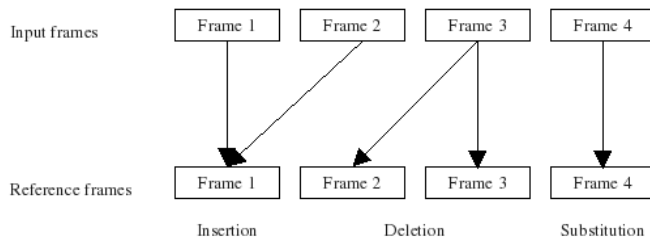


Figure 3. Pictorial representation of the basic operations of Dynamic Time Warping.

4.2. Dynamic Time Warping

Dynamic Time Warping (DTW) is a process typically used in communication protocols for speech recognition (Parsons 1987). The overall idea is to align the input speech data with data from a reference speech signal by repeating certain portions of a set of data that do not match the current position of the opposing set. For example, when someone wishes to place a phone call via voice dialing, he or she may simply say “Call, John Doe”. The next time that the same individual wishes to call John Doe via voice dialing, his or her voice will undoubtedly have variations from the first time that voice dialing was initiated. For instance, the second time, the individual may say “Call, Joohn Doe”. Obviously, these two statements are not identical; however, the desire for the user to call John Doe has not changed. Therefore, DTW would be used to correctly match the vocal inflections made during the second speech (i.e. the input signal) with those of the previously stored first speech (i.e. the reference signal). Here, the three “o’s” in the input signal would simply match with the single “o” in the reference signal, and John Doe would receive a phone call.

In this work, DTW is used to align image representations of robotic movements with those of the patient’s movements. Since there will be variations in segment lengths and duration due to varying velocities between the two movements, this process becomes necessary for deciding whether or not two movements are similar. Figure 3 is a pictorial representation of the DTW process and is described as follows: 1.) **Insertion**: The current reference frame does not match the current input frame, based on some threshold value γ , thus the current reference frame is repeated until a match is found during the linear search. 2.) **Deletion**: The current input frame does not match the current reference frame, thus the current input frame is repeated until a match is found. 3.) **Substitution**: The current reference frame matches the current input frame, and the sequence progresses forward.

The Manhattan distances calculated from the feature vectors obtained during the MHI process are stored in matrix form and used 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:

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

where W_I , W_D , and W_S are weights for performing an “insertion”, “deletion”, or “substitution”, d is the Manhattan distance between each frame, and m and n are the row and column positions in the matrices, respectively. 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

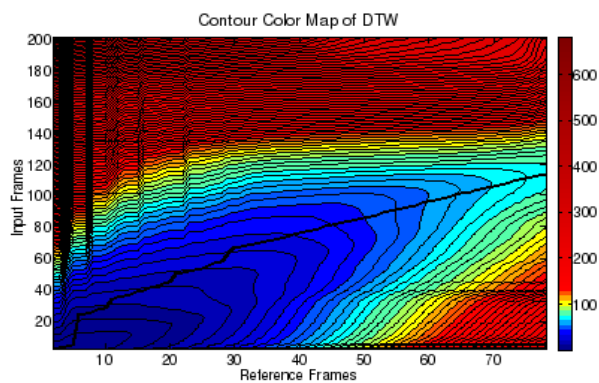


Figure 4. Contour color map of DTW illustrating the optimal path for mapping an input sequence to a reference sequence for one participant with a separate input sequence provided by the same participant.

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, here, it is not assumed that the starting and ending points of the exercise are known. Thus, DTW is employed in order to determine a minimal exit point.

Utilizing this method, the least costly path (i.e. the optimal matching sequence) is calculated and a mapping between the input sequence to the reference sequence is obtained thus minimizing the effects of varying velocities. In order to quantify this method for the purpose of this work, contour color maps were generated from patient exercise data. 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 darker contours represent points of lower cost, an optimal path would contain the maximum number of dark contours reaching the end of each sequence; the bold black line shows the chosen path for mapping the two sequences in this specific scenario.

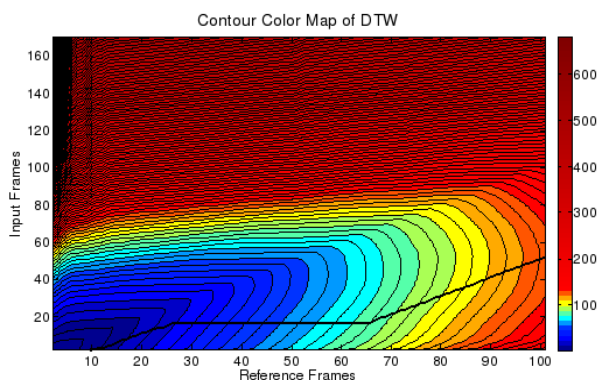
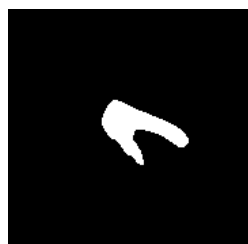


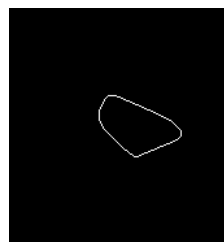
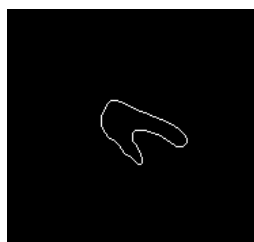
Figure 5. Contour color map of DTW illustrating the optimal path for mapping a reference sequence provided by the first participant with a separate input sequence provided by the same participant performing two different exercises.

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. 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 numerous light colored contours which are of high values. Second, only 50 frames of the 160



(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.

Figure 6. Example image processing sequence used to extract an ideal contour from the human's movements.

possible frames for the input image provided comparable feature vectors. In other words, this would not be considered a match for movements.

4.3. Contour Extraction

Once the humanoid and child movements have been aligned using DTW, a contour representing the shape of the respective 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 (e.g a round disc, square, rectangle, or cross). Here, a 16x16 square window was used. Figure 6b illustrates the use of a median filter on image data.

When the smaller contours have been removed, a canny edge detection algorithm (Canny 1986) is utilized in order to extract the edges of the contour representing the upper-arm movement, see Figure 6c. 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 6a illustrates an inaccuracy caused by camera lag time; even though the subject moved his arm in one complete motion, the video sequence shows a gap between arm positions. Therefore, the convex hull of the edge detected image is determined.

The convex hull can be thought of as the boundary of a minimal convex set of

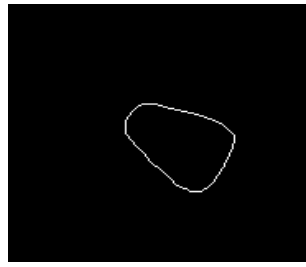


Figure 7. Illustration of the original convex hull image.

points containing a given non-empty finite set of points in a plane. The convex hull that is utilized is that of a simple polygon. 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 the each of the points in the triple is shifted by one vertex along the polygon (Cormen et al 1998). Melkman's Algorithm was employed to ensure correct outcomes (Melkman 1987). This implementation gives a somewhat ideal outline as seen in Figure 6d.

4.4. 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 (Duda and Hart 1972) on either the upper or lower region of the contour could enable a determination of upper or lower line for the purpose of finding the angle between 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 **RAN**dom **SAM**ple **CON**sensus (RANSAC) algorithm.

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 a model with unknown parameters from parameter space, Θ . These data points are inliers, and all others are outliers. The goal is to find model parameters, θ^* , from a parameter space, Θ , that maximizes a cost function, $J_S(\theta, U, \Delta)$. In the standard formulation, the cost function, J_S , is the size of the support of the model with parameters, θ , i.e. how many data points from U are consistent with it. Data points with error smaller than a threshold, Δ , are considered to be consistent or to support the model; Δ is an input parameter to RANSAC. An error function, $\rho(\theta, x)$, representing a distance of a data point to a model is typically given (Chum 2005; Fischler and Bolles 1981).

The RANSAC algorithm carries out the maximization of J_S by repeatedly ex-

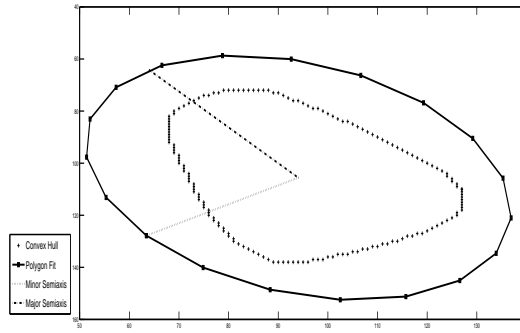


Figure 8. Illustration of the Major (dashed line) and Minor (dotted line) semi-axes located on a contour obtained from human upper-arm movement. The polygon boundary used to calculate the two semi-axes is shown with square markers.

ecuting two steps: (i) a **hypothesis generation** step, where a hypothesis, θ_k , of the model parameters is computed from a subset of points, S_k , selected from the input data points, U , at random and (ii) the **verification** step, where the quality of the hypothesized model parameters is calculated by utilizing a user predefined probability (confidence = $1 - \eta_o$, typically set to 95%) to recover maximum of the cost function J_S (Chum 2005; Fischler and Bolles 1981).

The algorithm can be described mathematically as follows: Let P be the probability that a sample of size m is randomly selected from a set U of N data points

$$P(I) = \frac{\binom{I}{m}}{\binom{N}{m}} = \prod_{j=0}^{m-1} \frac{I-j}{N-j} \leq \epsilon^m, \tag{4}$$

where ϵ is the fraction of inliers $\epsilon = I/N$. The number of inliers, I , is not known beforehand. Let χ_k^* be the largest support of a hypothesized model found up to the k -th sample inclusively, $I_k^* = |\chi_k^*|$. The sampling process is terminated when the likelihood of finding a better model (with larger support than I_k^*) falls under a threshold, i.e. when the probability, η , of missing a set of inliers, χ^+ , of size $|\chi^+| \leq I_k^*$ within k samples falls under a predefined threshold, η_o ,

$$\eta = (1 - P(I_k^*))^k \tag{5}$$

The number of samples that has to be drawn to satisfy $\eta \leq \eta_o$ is

$$k_{\eta_o}(I_k^*) = \frac{\ln(\eta_o)}{\ln(1 - P(I_k^*))}. \tag{6}$$

This method is a more accurate approach 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 7, 8, and 9 illustrate this process.

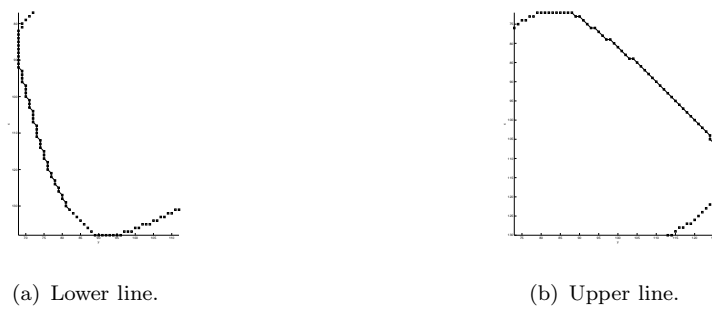


Figure 9. Lower line, determined by RANSAC, used to find the ROM.

4.5. Range Of Motion and Angular Velocity

4.5.1. Range Of Motion

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 7:

$$\begin{aligned}
 m_1 &= \frac{y_2 - y_1}{x_2 - x_1} \\
 m_2 &= \frac{y_4 - y_3}{x_4 - x_3} \\
 \Theta &= \arctan\left(\frac{m_2 - m_1}{1 + m_2 m_1}\right) \\
 ROM &= \left| \Theta \frac{180^\circ}{\pi} \right|
 \end{aligned} \tag{7}$$

where x_i and y_i are the coordinates of points on each line segment, m_1 and m_2 are the slopes of each line, and Θ and ROM are the current angle in radians and degrees, respectively. The maximum angle found over the length of the video sequence gives the total ROM of the child's movements.

4.5.2. Peak 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 8, 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.

$$\omega = \frac{d\theta}{dt} \tag{8}$$

The maximum angular velocity would be considered the PAV. Thus, this algorithm gives two physical therapeutic metrics that can be easily read by physical therapists for the purpose of analyzing a patient's current status and overall progress.

5. Results and Analysis

For the initial testing of the methodology of this research, two subjects were utilized, both male. The subjects were asked to perform a series of upper-arm exercises,



(a) Subject's starting position.

(b) Subject's ending position.

Figure 10. Lateral and medial movement demonstrated by participant.

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 10. The images were then processed by our algorithm in order to obtain the ROM and PAV, which then was compared to the data obtained from separate subjects who performed visual observation techniques (deemed expert users), and data captured via the Trimble 5606 Robotic Total Station. The specific task of the expert users was to locate the top portion of each patient's arm in several sets of images using mouse clicks. Using that information, in congruence with a constant location for the patient's beginning position, it was trivial to determine the ROM and, by extension, the PAV as viewed by the expert users. Expert users were adults between the ages of 30–35. They were required to have at least five years of experience in image-processing research. Recruitment of the expert users was done by word-of-mouth. Since physical therapists typically use visual feedback and estimation as the sole method for determining a patient's ROM, the author's deemed that this method is at least equivalent to such an analysis.

Occasionally, physical therapists use a measurement device known as a goniometer, which measures the initial and final position of a joint during an exercise to determine a patient's ROM. However, the author's decided to utilize a more accurate approach for retrieving this information. 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. The long-range 5600 series DR200+ total station allows measurements up to 600 m (1,968 ft) to a 90% reflective Kodak Gray Card and 200m (656 ft) to an 18% reflective Kodak Gray Card. The range using a single prism is 5,500 m (18,040 ft) with an accuracy of (3 mm +2 pixels per meter (ppm)). Coupling that with the robotic tracker gives a range up to 1,200 m (3,937 ft) with 2" (0.5 milligon (mgon)) (Trimble 2005). (A gon is a unit of angle measurement equal to 1/400 circle, 0.01 right angle, 0.9°, or 54'). As such, the data obtained from the robotic station represents ground truth. Tables 2 and 3 show the ROM comparison between the expert users, algorithmic data, and Trimble 5606 for each subject per exercise. Tables 4 and 5 show the PAV comparison between the aforementioned data sets.

As shown in the tables, the ROM values calculated via our algorithm are closely related to the ground truth data (overall average expert user error < 11% and overall average algorithmic error < 9%). Figures 11a and 11b show a comparison between the expert users and algorithm, with respect to the ground truth data, of the averages across all subjects for the ROM data. 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 PAV calculated via our algorithm is related to the trend of those calculated

Table 2. Measurements of ROM for Adduction/Abduction Exercise.

Repetition	Subject A			Subject B		
	1	2	3	1	2	3
Expert User ($^{\circ}$)	90.1	65.5	39.2	86.7	57.0	45.8
Algorithm ($^{\circ}$)	86.4	56.5	40.1	88.1	54.9	45.6
Ground Truth ($^{\circ}$)	90.3	56.5	31.1	90.0	65.0	42.1
Expert User Difference ($^{\circ}$)	0.2	9.0	8.2	3.3	7.9	3.7
Algorithm Difference ($^{\circ}$)	3.9	0.0	9.0	1.9	10.1	3.5
Expert User Error (%)	0.18	15.96	26.33	3.65	12.20	8.81
Algorithm Error (%)	4.32	0.00	29.07	2.14	15.53	8.43

Table 3. Measurements of ROM for Lateral/Medial Exercise.

Repetition	Subject A			Subject B		
	1	2	3	1	2	3
Expert User ($^{\circ}$)	86.2	57.3	29.8	84.9	56.0	44.7
Algorithm ($^{\circ}$)	80.5	55.5	28.0	84.5	62.8	38.4
Ground Truth ($^{\circ}$)	90.0	67.2	26.4	88.4	63.2	40.5
Expert User Difference ($^{\circ}$)	3.8	9.9	3.4	3.5	7.3	4.2
Algorithm Difference ($^{\circ}$)	9.5	11.7	1.6	3.9	0.4	2.1
Expert User Error (%)	4.22	14.70	13.07	3.95	11.46	10.40
Algorithm Error (%)	10.57	17.43	6.25	4.42	0.67	5.11

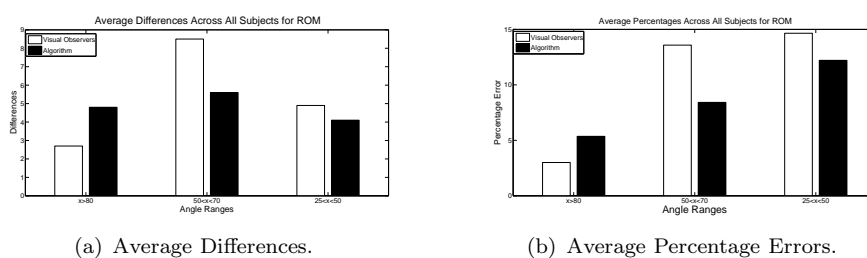


Figure 11. Comparison between the expert users and algorithm, with respect to the ground truth data, of the averages across all subjects for the ROM data.

Table 4. Measurements of PAV for Adduction/Abduction Exercise.

Repetition	Subject A			Subject B		
	1	2	3	1	2	3
Expert User (ω)	82.1	62.2	57.0	81.9	29.1	26.0
Algorithm (ω)	79.6	57.7	67.4	71.3	22.8	52.3
Trimble 5606 (ω)	87.8	59.9	67.5	62.4	23.9	54.6
Expert User Difference (ω)	5.7	2.3	10.5	19.5	5.2	28.6
Algorithm Difference (ω)	8.2	2.2	0.1	8.9	1.1	2.3
Expert User Error (%)	6.51	3.77	15.53	31.24	21.72	52.36
Algorithm Error (%)	9.33	3.66	0.09	14.32	4.41	4.13

via ground truth data (average expert user error $< 16\%$ and average algorithmic error $< 8\%$). Figures 12a and 12b show a comparison between the expert users and algorithm, with respect to the ground truth data, of the averages across all subjects for the ROM data. It should 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.

Of particular importance are the comments made by the expert users when attempting to determine the range of motion of the patients. Both users noted

Table 5. Measurements of PAV for Lateral/Medial Exercise.

Repetition	Subject A			Subject B		
	1	2	3	1	2	3
Expert User (ω)	50.8	40.9	47.1	39.7	37.3	30.8
Algorithm (ω)	48.0	39.3	65.4	34.1	40.2	27.4
Trimble 5606 (ω)	51.3	43.8	65.9	36.5	34.5	31.1
Expert User Difference (ω)	0.5	2.9	18.8	3.2	2.8	0.3
Algorithm Difference (ω)	3.3	4.6	0.5	2.4	5.7	3.7
Expert User Error (%)	1.00	6.60	28.53	8.83	8.05	1.09
Algorithm Error (%)	6.48	10.42	0.74	6.51	16.62	11.97

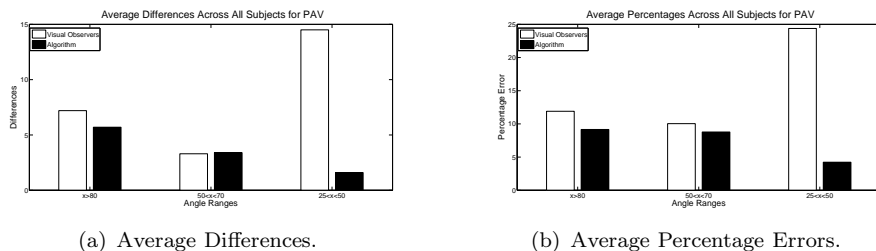


Figure 12. Comparison between the expert users and algorithm, with respect to the ground truth data, of the averages across all subjects for the PAV data.

that it was quite difficult to determine the exact location of Subject B's arm due to the fact that he was wearing a large blue sweater. One advantage that our algorithm has is that regardless of the clothing that the patient is wearing, we can still track the overall movement and obtain reasonable data.

6. Advantages and Disadvantages

When considering this approach for gathering physical therapeutic data, one has to be knowledgeable of the limitations or potential concerns. Many physical therapists use sensation or force-feedback as a method of assessing the patient's current state and determining the effects of the physical therapy sessions. Since our algorithm focuses on non-contact assessments, it will not be possible to gather the important information obtained from sensation analysis. This limitation in-turn creates a concern as to whether or not this method will produce comparable effects to robotic rehabilitation devices that incorporate force-feedback with subtle, unrecognizable to the patient, changes that induce motor movements that the patient believes he or she is incapable of achieving due to injury or disability. However, noting the earlier references to sustained attention, it is believed that this approach will have positive results in that it will provide a motivating factor for individuals during the sessions. Thus, with an increase in motivation and attention, we hypothesize that this will indeed produce comparable results to robotic rehabilitation devices that incorporate force-feedback.

The use of a robotic platform for this specific application also reintroduces the debate of robotics versus unembodied (i.e. simulated) agents. Specifically, what are the benefits of using a robotic platform rather than a simulated agent such as those that currently exist within the gaming realm (e.g. Nintendo Wii and Microsoft Xbox)? For years, the goal of artificial life has been to move from simulation to embodied systems. Two main reasons are that 1.) without regular validation on real robots there is a great danger that much effort will go into solving problems that simply do not come up in the real world and 2.) there is a real danger that

programs which work well on simulated robots will completely fail on real robots because of the differences in real world sensing and actuation (Brooks 1992). These statements were true in the early 1990's and are still prevalent today. If we stay within the realm of simulated agents we may never reach fully autonomous robotic placement within our society. The major role of simulated agents should be the initial modeling of real robots. Human robot interaction, even at the simplest level, enables researchers to gain valuable insight regarding robotic acceptance. For this specific research, as it was stated earlier, robots have been shown to be engaging to children both with and without disabilities. We hypothesize that having an agent that is embodied in the real world increases the motivation and enthusiasm for patients.

7. Conclusions and Future Work



Figure 13. Illustration of the robotic platform.

In this paper, an approach to matching child movements with robotic movements for the purpose of evaluating child upper-arm rehabilitation exercises has been discussed. Specifically, the physical therapeutic metrics range of motion and peak angular velocity have been calculated via computer vision techniques and can be utilized in a robotic system. When analyzing the data, one should note that the Trimble 5606 captured 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 than there are for the ground truth data without any interpolation. With this in mind, there is a potential issue in cases where the patient is moving at very high velocities (i.e. there may be issues in tracking efficiency when using devices with low frame rate capabilities). Ideally, most patients will move at a rate that is capable of being tracked due to injury (i.e. relatively slowly). However, in this pilot study, patients were not injured. This may give some insight into the high variances in percent errors between the algorithmic and Trimble data such as those found in Tables 2 and 3. As such, the immediate future work consists of obtaining ground truth data from a tracking device with a faster frame rate, giving us even greater accuracy.

Concurrently, the next step is to incorporate another physical therapy metric, namely movement smoothness and apply each of these principles to other activities such as reaching. One possible method of determining movement smoothness is to analyze the acceleration of the patient's motion over time, where drastic increases or

decreases may signify a jittery motion. Ultimately, this will enable the final robotic system to better assess the physical disability of the patient. Another seemingly simplistic task is to decrease the processing time that is required to achieve accurate results when using RANSAC. Because the algorithm requires a variable number of iterations to achieve high accuracy (99% is used in this work), some compromise between the number of iterations allowed, accuracy, and number of images available may have to be forced in order to speed things up. Another component that will decrease the processing time is use of categorical information. By allowing the humanoid to categorize movements and store the data representing its own motion into specific locations *a priori*, only one set of images (the input from the patient's exercise) will need to be captured and processed. The overall desire to decrease time comes from the logical fact that children will become restless when required to wait for extended periods when engaged in a turn-taking scenario.

The final step of this research is to equip a robotic platform, shown in Figure 13, with a small camera and Gumstix OveroTM Earth that will enable the robot to perform its movements, video capture, and image processing completely on-board (Brooks and Howard 2009). Currently, our vision processing is done on a separate host computer. Once this phase is complete, the researchers intend to conduct testing with physically injured children in realistic physical therapy scenarios.

8. Acknowledgments

The authors gratefully acknowledge the contributions of the individuals in the HumAnS Lab, Dr. Patricio Vela, Dr. Jochen Teizer, Miguel Serrano, NASA JFPF, UNCFSP, and those who participated in the therapy sessions.

References

- Bobick AF, Davis JW. 2001. The recognition of human movement using temporal templates. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. March (. ; vol. 23). pp. 257–267.
- Brewer B, McDowell S, Worthen-Chaudhari L. 2007. Thomas Land; Poststroke upper extremity rehabilitation: a review of robotic systems and clinical results. *Topics in Stroke Rehabilitation* 14(6):22–44.
- Brisben A, Safos C, Lockerd A, Vice J, Lathan C. 2005. The CosmoBotTM System: Evaluating its Usability in Therapy Sessions with Children Diagnosed with Cerebral Palsy .
- Brooks D, Howard AM. 2009. Upper-limb Rehabilitation and Evaluation of Children using a Humanoid Robot. In: *ICMI-MLMI09 Workshop on Child, Computer and Interaction*. Cambridge, MA. November.
- Brooks R. 1992. Artificial life and real robots. In: *Proceedings of the first European conference on artificial life*. Citeseer. pp. 3–10.
- Burgar C, Lum P, Shor P, derLoos MV. 2000. Development of robots for rehabilitation therapy: the Palo Alto VA/Stanford experience. *J Rehabil Res Dev*. 37(6):663–673.
- Campbell L, Bobick AF. 1995. Recognition of Human Body Motion Using Phase Space Constraints. In: *Proc. Int'l Conference Computer Vision*. pp. 624–630.
- Canny J. 1986. A computational approach to edge detection. *Readings in Computer Vision: Issues, Problems, Principles and Paradigms* 8(6):184–203.
- Casadio M, Giannoni P, Morasso P, Sanguineti V. 2009. A proof of concept study for the integration of robot therapy with physiotherapy in the treatment of stroke patients. *Clinical Rehabilitation* 23(3):217.
- Chum O. 2005. Two-view geometry estimation by random sample and consensus. Czech Technical University.
- Colombo R, Pisano F, Micera S, Mazzone A, Delconte C, Carrozza M, Dario P, Minuco G. 2005. Upper Limb Rehabilitation and Evaluation of Stroke Patients Using Robot-Aided Techniques. In: *ICORR 2005 Conference Proceedings*. June 28 - July 1. pp. 515–518.
- Cooley E, Morris R. 1990. Psychology Press; Attention in children: A neuropsychologically based model for assessment. *Developmental Neuropsychology* 6(3):239–274.
- Cormen T, Leiserson C, Rivest R. 1998. 0-262-03141-8.
- Duda R, Hart P. 1972. ACM New York, NY, USA; Use of the Hough transformation to detect lines and curves in pictures .
- Duncan P, Propst M, Nelson S. 1983. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Physical Therapy* 63(10):1606.
- Federico Posteraro M, Mazzoleni S, PhD S, MD B. 2009. Robot-mediated therapy for paretic upper limb of chronic patients following neurological injury. *REHABILITATION MEDICINE* 41:976–980.

- Fischler M, Bolles R. 1981. ACM New York, NY, USA; Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography .
- Folio M, Fewell R. 1983. Peabody Developmental Motor Scales and Activity Cards (Manual): Hingham. MA: Teaching Resources .
- Folio M, Fewell R. 2000. Peabody Developmental Motor Scales: Examiner's Manual. Pro-ed.
- Fugl-Meyer A, Jääskö L, Leyman I, Olsson S, Steglind S. 1975. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance.. *Scandinavian Journal of Rehabilitation Medicine* 7(1):13.
- Gockley R, Matarić M. 2006. Encouraging physical therapy compliance with a hands-off mobile robotIn: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM. pp. 150–155.
- Goncalves L, DiBernardo E, Ursella E, Perona P. 1995. Monocular Tracking of the Human Arm in 3DIn: *Proc. Int'l Conference Computer Vision*. August. pp. 764–770.
- Hesse S, Schulte-Tiggens G, Konrad M, Bardeleben A, Werner C. 2003. 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.
- Hurley M. 2009. Effectiveness of robot-assisted therapy on stroke patients with upper extremity impairment. *Physical Function* p. 14.
- Kozima H, Nakagawa C, Yasuda Y. 2005. Interactive Robots for Communication-Care: A Case Study in Autism TherapyIn: *ROMAN 2005*. August. pp. 341–346.
- Krebs H, Hogan N, Aisen M, Volpe B. 1998. Robot-aided Neurorehabilitation. *IEEE Trans. Rehab Eng.* 6(1):75–87.
- Kwakkel G, Kollen B, Krebs H. 2008. Am Soc Neurorehabil; Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review. *Neurorehabilitation and Neural Repair* 22(2):111.
- Loureiro R, Amirabdollahian F, Topping M, Driessen B, Harwin W. 2003. Upper Limb Robot Mediated Stroke Therapy Gentle/s Approach. *Autonomous Robots* 15(1):35–51.
- Lum P, Burgar C, Shor P, Majmundar M, derLoos MV. 2002. 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.
- Lum P, Reinkensmeyer D, Mahoney R, Rymer W, Burgar C. 2002. Thomas Land; Robotic devices for movement therapy after stroke: current status and challenges to clinical acceptance. *Topics in stroke rehabilitation* 8(4):40–53.
- Mehrholz J, Platz T, Kugler J, Pohl M. 2009. Am Heart Assoc; Electromechanical and robot-assisted arm training for improving arm function and activities of daily living after stroke. *Stroke* 40(5):e392.
- Melkman A. 1987. On-line construction of the convex hull of a simple polyline. *Information Processing Letters* 25(1):11–12.
- Parsons T. 1987. *Voice and Speech Processing*. New York: McGraw-Hill.
- Patton JL, Mussa-Ivaldi FA. 2004. Robot-assisted adaptive training: custom force fields for teaching movement patterns. *IEEE Trans Biomed Eng.* 21(4):636–646.
- Polana R, Nelson R. 1994. Low Level Recognition of Human MotionIn: *Proc. IEEE Workshop Non-Rigid and Articulated Motion*. pp. 77–82.
- Rehg J, Kanade T. 1995. Model-Based Tracking of Self-Occluding Articulated ObjectsIn: *Proc. Int'l Conference Computer Vision*. pp. 612–617.
- Reinkensmeyer DJ, Schmit BD, Rymer WZ. 1999. Assessment of active and passive restraint during guided reaching after chronic brain injury. *Ann Biomed Eng.* 27(6):805–814.
- Russell D, Rosenbaum P, Cadman D, Gowland C, Hardy S, Jarvis S. 1989. Wiley Online Library; The gross motor function measure: a means to evaluate the effects of physical therapy. *Developmental Medicine & Child Neurology* 31(3):341–352.
- Takahashi C, Der-Yeghiaian L, Le V, Motiwala R, Cramer S. 2007. Oxford Univ Press; Robot-based hand motor therapy after stroke. *Brain* .
- Trimble. 2005. Trimble - 5605 Total Station and LM80 Layout Manage. Datasheet .
- Volpe BT, Krebs HI, Hogan N, Edelstein L, Diels C, Aisen M. 2000. A novel approach to stroke rehabilitation: robot-aided sensorimotor stimulation. *Neurology* 54(10):1938–1944.
- Wang H, Liao H, Hsieh C. 2006. American Physical Therapy Association; Reliability, sensitivity to change, and responsiveness of the Peabody Developmental Motor Scales—second edition for children with cerebral palsy. *Physical Therapy* 86(10):1351.