Semi-automated home-based therapy for the upper extremity of stroke survivors

Thanassis Rikakis
Virginia Tech
Blacksburg, Virginia, USA
rikakis@vt.edu

Aisling Kelliher
Virginia Tech
Blacksburg, Virginia, USA
aislingk@vt.edu

Jinwoo Choi
Virginia Tech
Blacksburg, Virginia, USA
jinchoi@vt.edu

Jia-Bin Huang
Virginia Tech
Blacksburg, Virginia, USA
jbhuang@vt.edu

Kris Kitani
Carnegie Mellon University
Pittsburgh, PA, USA
kkitani@cs.cmu.edu

Setor Zilevu
Virginia Tech
Blacksburg, Virginia, USA
zilevus@vt.edu

Steven L Wolf
Emory University
Atlanta, Georgia, USA
swolf@emory.edu

ABSTRACT

Technology assisted home based rehabilitation therapy offers a potentially cost-effective and convenient solution for those affected by neuro and musculoskeletal impairments. Home based solutions, however, face many challenges, the most significant of which is trying to reproduce a complex adaptive therapy experience in the home without the continuous presence of the therapist. Building on our prior work creating interactive systems for the clinic, we present our home-based system that integrates customized therapy objects, camera based movement capture and assessment techniques, and a flexible exercise protocol aimed at generalizing to variable daily life activities. We present findings from two pilot studies with unimpaired and impaired users and describe how insights from these studies will guide future work.

CCS CONCEPTS

• **Human Centered Computing** → *Usability Testing;* Interaction design process and methods

KEYWORDS

Home Based Therapy Systems, Interactive Neurorehabilitation, Stroke Rehabilitation

1 INTRODUCTION

According to the American Heart Association [6] every year over 795,000 people in the United States experience stroke, with approximately 80% of stroke survivors experiencing hemiparesis, meaning weakness of the left or right side of the body. As a result, those impacted may have trouble performing everyday activities such as eating, dressing, using the bathroom, and manipulating daily life objects like door handles, driving wheels, etc. In additional to its effect on everyday life, stroke costs an estimated $36 billion each year within the United States alone. This total includes the cost of health care services, medications to treat stroke, and missed days of work, making stroke the leading cause of serious long-term disability.

Stroke survivors can benefit from long-term rehabilitation therapy, as determined by findings from several large-scale studies [14, 31]. However, there are multiple challenges with this approach, notably the financial cost, access to and availability of therapists, and difficulties arranging regular transportation to hospital and clinic facilities over a prolonged period of time. A potential alternative has emerged in recent years in the form of technology assisted home-based unsupervised or lightly supervised therapy. Home based interactive rehabilitation can be defined as computer-assisted therapy with limited engagement of the therapist, either through remote supervision [25] or a limited number of visits to the home [4]. Recent developments using relatively cheap technologies such as tablet computers and depth cameras such as the Kinect increase the possibility of long-term, affordable home monitored care [5, 18, 25, 30]. This form of therapy can be effective as the primary mode of treatment [2], or as a support mechanism for traditional therapy in the clinic [1]. However, the scaling of interactive home-based therapy faces significant challenges [5], beginning with the replication of the therapist functions, the cost and potentially intrusive nature of domestic systems [3], the lack of consensus regarding quantitative movement measurements [16], and challenges in motivating

PETRA '18, June 26–29, 2018, Corfu, Greece

ACM ISBN 978-1-4503-6390-7/18/06…$15.00
https://doi.org/10.1145/3197768.3197777

adherence to home based training [22]. A viable system for interactive home-based rehabilitation must coherently address these nested problems from a multidimensional perspective, integrating technical, social, medical, design, and HCI knowledge.

In this paper, we describe the development of our approach to home based stroke rehabilitation, building on findings from the literature and our own work carried out in hospitals and clinics over the last decade. We introduce the current implementation of our interactive system and describe findings from two pilot studies with unimpaired and impaired users. We conclude with discussion of the opportunity space moving forward and proposed future work.

2 RELATED WORK

2.1 Fundamentals of Contemporary Adaptive Therapy

Traditional therapy of the upper extremity in the clinic is usually composed of repetitive movement tasks such as reaching and grasping an object. A participant performs these movement tasks in the clinic under the supervision of a physical therapist, who visually monitors the improvement in functionality and quality of movement over time, to provide personalized rehabilitation therapy. To make all these decisions, the therapist needs to track and analyze the activity of the stroke survivor (including interactions with different objects, environments, and people) over multiple time brackets (from real-time performance to progress over months). The observations include an overall assessment of functionality, as well as detailed aspects of movement quality that may affect functionality.

Approaches to stroke rehabilitation have evolved considerably in the last two decades, with an emerging focus on patient use of goal-driven, active problem solving strategies [24] and experimentation with training tasks that can generalize to multiple activities of everyday relevance [22, 24, 31]. Adherence to long term training and functional use of the affected extremity improves with enhanced self-efficacy [22]. The therapist therefore needs to carefully adapt the protocol to be challenging, but not frustrating. The therapist must support active participation by the patient, reinforce success and help the patient construct a narrative of improvement that can motivate and overcome potential failures, fatigue, or boredom. The therapist also needs to help the patient connect improvement in training to empowerment in everyday life, especially in daily life functions important to each patient. Finally, the therapist needs to also monitor and address the presence of physical or emotional discomfort. This very complex work by the therapist results in a continuous Adaptive Training Plan (which we term ATP). The ability of a therapist to produce successful ATPs results from years of collective and individual experience that is not readily observable or easily quantifiable. Therefore, the first and most significant challenge for automated, home-based interactive rehabilitation is to *reproduce a complex adaptive therapy experience in the home without the continuous presence of the therapist, while advancing patient self-efficacy, improving adherence and increasing patient quality of life.*

2.2 Movement Capture and Assessment

Automated movement capture and analysis is a fundamental component of the above challenge. Automated approaches for tracking and assessing upper extremity activities require capturing both the movement of the person and the relation of that movement to artifacts of various shapes and functions [13]. Detailed information on the hand and object manipulation needs to be captured in conjunction with the movement of the whole upper limb and the torso [16, 29, 41]. The arms and hands of stroke survivors are often deformed and contorted due to a loss of neuromuscular control. Because of such physiological limitations, stroke survivors often use compensatory movements to perform functional tasks, which further increases variance in movement performance. Results from prior clinical upper extremity studies emphasize the most promising features for functional movement analysis, including end effector activity, torso compensation, hand shape, and object behavior [5, 11, 15]. The technical infrastructure for home based systems must be appropriately reduced, which necessitates a concurrent simplification of the sensing approach. Insights from the work above, particularly with regard to a reduced set of movement features for analysis, are extremely useful in helping to determine the selection and implementation of ideal sensing technologies and methods.

A key feature used by therapists to evaluate the quality of upper extremity functionality is the shape of the hand during a grasping task. Understanding hand configurations in terms of functional grasps categories is a critical technology needed for understanding hand use overall [13] and especially in the context of rehabilitation [19, 28, 31]. From a hardware perspective, the creation of smart objects presents a potential support solution in this context, with several clinical systems developing artifacts embedded with sensors to assist in the measurement of movement quality, and in particular, with tracking the hand [5, 11]. This technical approach has proven effective in supervised settings (including our own work), where issues with charging, object setup, and object maintenance can be handled by a dedicated development team. However, translating this approach to an unsupervised home setting presents difficult challenges as each electronic or technical component adds another layer of complexity for the patient and/or the caregiver to deal with.

Once the movement capture and assessment challenges are addressed, the issue of automated feedback needs to be tackled. Since internal feedback and assessment mechanisms are affected by the stroke, during therapy, patients rely on assistive feedback from the therapist to effectively and consistently evaluate their performance. These feedback functions need to be transferred to computational agents during interactive therapy [15]. Finally, the adaptation decisions made by the therapists also need to be gradually transferred to a computational agent [7]. All of these issues need to addressed within systems comprised of inexpensive infrastructure, which can fit unobtrusively in small homes and not be too complicated or confusing to operate.

2. 3 Our Approach

In working together over many years, our diverse team of physiotherapists, computer scientists, engineers, designers, and HCI experts have tackled the problem of interactive neurorehabilitation for the upper extremity of stroke survivors over three key stages.

In the first stage, we developed systems for supervised interactive rehabilitation in the clinic [8, 9, 11]. Supervised use in the clinic allowed for expert observers to monitor the therapy, document challenges and areas of improvement, and intervene when necessary when, and if, problems appeared. This system was marker based and used an extensive array of motion capture cameras [11]. Through a series of studies with stroke survivors, we were able to identify key kinematic features for evaluating movement quality and model their interrelation and their correlation to standardized clinical measures of functionality [9]. We used arts and interactive computing principles to map each feature to an appropriate stream of multimodal feedback (i.e. time series data mapped to sound). Through appropriate multimodal compositions techniques, we were able to present as many as 6 streams of feedback together that could be parsed and used by the patient for performance evaluation [11]. We were furthermore able to analyze the adaptation processes used by the therapists and successfully train a computational agent to imitate and predict a limited set of adaptation decisions [7]. However, to achieve all these computational advances, we needed to limit the set of movement tasks that could be performed with the system so that we could get a significant amount of coherent data for the actions being modeled. A clinical study with 24 stroke survivors showed that our system could promote active learning and meaningful improvements to movement quality [11]. The study also showed that the translation of the training gains to activities of daily living required a much more varied set of training activities and increased possibilities of therapy adaptation so as to address the unique characteristics of each patient [11].

During the second stage, we developed a self-contained system (a custom table and screen) aimed towards home use. The movement capture and analysis approach combined a significantly reduced marker set, with three motion capture cameras mounted on the screen and smart objects embedded in the table. Therefore, only a reduced set of movement features could be captured and assessed and used to drive automated feedback. The second version increased the variability of tasks that could be performed. We tested the system in the clinic with 15 patients using the system under light supervision by the therapist [5]. Training with the system promoted active learning by the patients, but even this reduced infrastructure proved too technologically complex to embed in the home. The multimodal feedback provided by the system was effective, but could be overwhelming to interpret without the assistance from a therapist. The variability of tasks and the adaptation options were also still too limited for producing generalizable learning. This cycle of testing included offline rating of videos of patient task performance by expert therapists in order to further inform the computational assessment of tasks by our system. Although the therapist ratings were limited to a sub-group of the tasks performed, this work showed that an expert constructed movement rating system, used in a consistent manner by trained therapists, could provide good data for training computational agents.

We concluded that successful training at the home needs an even simpler computational infrastructure. However, the variability of training tasks needs to be increased, along with their ability to promote generalizable learning that maps to meaningful activities of daily living. Effective adaptive training at the home needs to be lightly supervised by the therapist and the most effective mode of light supervision still need to be discovered.

The results of this extensive and varied evaluation helped us identify key improvements for the third stage in our process, namely the redesign and evaluation of the system as proposed and presented in this paper. For the third version of the system, we adopted a cyber human system approach, with several complex functions offloaded to human actors (patient, therapist, system designer). This enables us to reduce the cost and scale of the technology infrastructure for a rehabilitation system, while also providing a solution that is unobtrusive, adaptable, and suitable for varying home environments. We tested the functionality of this version with a group of 15 unimpaired subjects and the feasibility of the system with a small group of stroke survivors. In the next section we review our system design for the third phase of our work and the testing and outcomes from the pilot testing.

Figure 1. a) The interactive stroke rehabilitation system including mat, objects, tablet and mounted Kinect; b) set of 6 objects; c) combining two objects by screwing together

3 HOME BASED INTERACTIVE REHAB SYSTEM

Our current home based interactive rehabilitation system consists of a laser etched mat (see Fig 1a), six customizable therapy artifacts (see Fig 1b), a table mounted Kinect camera and mini-computer module clamped underneath, and a tablet device (see Fig 1a). The mat is placed on a table and acts as a stage on which the patient performs each rehabilitation activity using the artifacts individually or in combination. Visual markers etched on the mat are used to guide the user and can also support system calibration. The Kinect camera is integrated into an adjustable stand which allows for maximum visibility of the upper body of the user. The tablet device hosts the dynamic web application presenting the training protocol, including the activity instructions.

3.1 System design and implementation

In the following section, we describe how the design and implementation of our system aims to address the key challenges identified earlier in creating interactive home based rehabilitation systems and developing suitable training environments and protocols.

Embedding systems in the home

Our system is designed to fit on typical tabletop surfaces found in the home, such as a kitchen, dining room, or computer/office table. The components are compact and lightweight (none individually weighing more than a pound) and can be lifted, installed, adjusted, and disassembled using one hand. If desired, the system (entirely or individual components) can be removed and stored relatively quickly after use, meaning it does not require a dedicated permanent installation space, and thus will not intrude unnecessarily in the existing setup of the home. Physical system components also contain assistive features to both support the participant in the straightforward system (re)setup and to assist aspects of the computer vision mechanism. For example, metal hinges on the outer plane of the mat help setup and maintain alignment with the table edge, while embedded magnets in the mat, tablet stand, and objects container snap gently together to ensure consistent placement.

The mini computer, camera, and tablet (when charging) are plugged into one power strip, thus minimizing effort in ensuring all components are powered and operational. The mat size and markings can be custom cut and inscribed for each user, depending on their height, handedness, and limb span and reach. Removable elements such as a customized ergonomic seat cushion (for comfort and stability), Velcro-mounted cushion pads on the seat back (to encourage correct posture), and floor tape markings (to ensure correct seat location) are low-cost and temporary interventions that support participant and system goals, without requiring permanent alterations to the home. Finally, the aesthetic look and feel of the mat, objects, and object container can also be customized for the participant to amplify the potential for seamless integration in their preferred home environment.

The technical infrastructure of our system is greatly reduced compared to interactive systems found in the clinic. A Kinect 2 camera, an ultra compact GIGABYTE mini-pc, and an iPad comprise the computational components of our system, which keeps the current overall technology cost at around $1300. The Kinect provides a relatively inexpensive (<$100) and unencumbered solution for tackling hand detection and torso movement, while its lightweight form factor means it can be easily and temporarily mounted to any available table. The iPad touchscreen tablet provides an accessible platform for the coordination of communication between the participant and the system. We choose to use a tablet computer as studies indicate that older adults (who make up the majority of stroke survivors) find touch computers 'less intimidating, less frustrating, and less overwhelming' than traditional computers [23]. The tablet is supported on the mat by a custom designed stand with adjustable angle settings, allowing the participant to orient the tilt of the device to their user preference.

Movement Capture and Assessment

Our system adopts a computer vision approach to movement capture as this obviates the need for placing specialized markers or other devices on the participant themselves. In addition, our prior work investigating hand detection using wearable cameras under changing light conditions and contexts indicates the potential of a vision based approach [17]. In our current implementation, we use the Kinect to passively measure hand movements and the Kinect SDK body tracking algorithm to track the positions of both shoulders as a proxy for torso movement. This approach provides coverage for the key movement features as defined by prior clinical work [5, 10] and as verified by the physiotherapists on our team, namely: 1) end effector activity over time and space (reach time, trajectory, velocity profile); 2) hand shape (grasp analysis); 3) torso compensation; and 4) object behavior.

For hand detection, we use a state-of-the-art regressor as proposed in [17] on the RGB images only. The regressor learns a sparse combination of color, texture, and gradient histogram features to detect hand regions over a variety of illumination condition and hand poses. We also apply this regressor for object detection. We designed the objects so that their colors are significantly different from skin color. As a result, the hand and object detectors can robustly detect hand and objects simultaneously within the environment. The detection of an object is important because it allows us to determine when the subject's hand is in contact with the object. Such information is useful to determine the correctness of the task. We train grasp classifiers for the three primary grasps (medium wrap, power sphere and precision pinch) using an approach used by Minnen and Zafrulla [20]. We use hand masks obtained from the hand detector along with depth data to learn random forest classifiers for each individual grasp category related to a task. Finally, we use the pose detection algorithm in [26] to compute 3D joint positions from depth images. In our analysis, we use the left and the right shoulder joints to keep track of movement of the torso.

Using the above movement capture system (hand/object detection, grasp analysis, and torso tracking), we implement the movement assessment module to judge the performance of the activity. Specifically, we aim to track and evaluate, through the computational system, key performance errors that are tracked by therapists in evaluating performance. The six identified primary performance errors are: 1) *Low speed*: the time for completing the activity exceeds a pre-defined threshold; 2) *Indirect path*: the length of the trajectory of the hand movement for object manipulation exceeds a pre-defined threshold; 3) *Dropped object*: the object is moving without being held in the hand; 4)

Misplacement: the object is placed in an incorrect region for one of the activities; 5) *Incomplete task*: at least one element of the activity is not completed; and 6) *Torso compensation*: large torso displacement. The user starts and stops the recording at the beginning and end of the performance of each task using the tablet interface. This provides a user defined video segment for each task performance. We implement a heuristic state transition estimator to separate the task into sub-activities by identifying when a hand returns/leaves a predefined hand rest zone in the mat. The movement capture and assessment algorithms are implemented in C++ and the captured videos can be processed both in real-time or offline. On average, our implementation achieves 5-6 frames-per-second on high-resolution videos (1920x1080 pixels for RGB images and 512x424 for depth images) on a mini PC with a 2.5 GHz Intel i7 CPU and 16 GB memory.

Training environment and protocol

The set of objects in our system are designed to support cross-mapping, problem solving, and generalizable activity strategies through their open-ended affordances, combinatorial possibilities, and perceived correlation with diverse artifacts of daily living. The design of the objects is informed by 1) comparative analysis of three complementary standards related to stroke rehabilitation, human hand grasps, and daily living activity assessment; 2) consideration of current and historical therapy artifacts; and 3) consultation with leading stroke rehabilitation experts.

We began by performing an analysis of three complementary standards of direct relevance – the Wolf Motor Function Test (WMFT); the Taxonomy of Human Grasp Types (GRASP); and the Motor Activity Log (MAL). The WMFT [31] provides a well known method for evaluating upper extremity movement using a clearly defined set of common household objects (e.g. soda can, pencil, towel, etc.). The GRASP Taxonomy [12] systematically classifies 33 different grasp types that have proven useful for computational recognition and assessment of human hand grasps (e.g., precision grip, power sphere etc.). The MAL [27] is a structured set of interview questions used to measure the effects of stroke therapy on the use of the impacted limb during everyday activities outside the clinic environment. In studying and comparing these standards, we sought to identify commonalities (e.g. encountered everyday objects, dominant grasp types, connections between activities etc.) that would point towards key design functions, attributes, and conditions. Based on our studies and working in consultation with the doctors and physiotherapists on our team, we created a set of three base objects (tapered can, hourglass, covered tripod), and three tops (teardrop, checker, round ball) that can be used individually or variously combined by stacking, or in the case of the can and teardrop top, screwing together (see Fig. 1b and 1c). The objects can be grasped and manipulated in a wide variety of ways corresponding to the primary identified grasps, thus lending themselves to rich cross-mapping opportunities.

Renowned stroke expert Dr. Steve Wolf led the development of the protocol for the series of 12 activities involving reach (e.g. reach and lightly touch two objects); reach and grasp (e.g. reach and grip one of the tripod object legs), reach, grasp and transport (e.g. reach and lift the can object up towards the participant's face); and reach, grasp, transport, and manipulate components (e.g. reach and hold the can object with the left hand and reach and pick up the tear drop object with the right hand, then screw the tear drop object into the can object).The activity tasks scaffold in complexity (from simple reaching exercises with single objects, to two-handed multi-stage manipulations with two objects) and are crafted to map to various activities of daily living, including those featured in the MAL set.

4 PILOT STUDIES

We evaluated our system in two pilot studies with unimpaired and impaired users. We now describe these studies in detail.

4.1 Pilot Study 1

Set up: We evaluated the functionality of the system through a pilot user study with 15 unimpaired subjects (using a similar approach to [25]). The first study aimed to assess and understand the following: 1) the overall functionality of the entire system; 2) the ability of the participant to understand and use the system; 3) the accuracy of the computer vision approach in assessing movement quality. The study took place in our research lab at Virginia Tech, with eight women and seven men participating, ranging in age from 20 to 63, with an average age of 34.

During the session, participants were tasked with performing 12 distinct activities (4 repetitions each) using the set of 6 objects. For the third and fourth repetitions, the participants were instructed to complete the activity, but each time with one of the identified 6 errors (assigned using a standard randomizing function), designed to simulate typical performance errors observed in stroke survivors during therapy: do the activity slowly; use an indirect path; drop an object; put the object in the incorrect place; do not fully complete the activity; and lean forward in the chair while completing the activity. It is important to note that while the errors as performed by unimpaired subjects might not 100% approximate stroke survivors, they are considered close enough by the therapists on our team to test the robustness and accuracy of our system.

Results and Discussion

System functionality: The system functioned smoothly throughout all 15 study sessions, with no technical hardware failures and correct communication maintained between the various system devices (Kinect, tablet, and mini-computer) throughout. The Kinect camera was able to capture skeletal data from all participants, regardless of height, without having to be adjusted. Participants were observed wearing accessories on their wrists and hands including bracelets, watches, large rings, and fitness devices, but the computer vision system handled the reflections and potential interference from them correctly. However, some errors were encountered with P2, who was wearing a blue shirt that was very similar in color to the blue objects.

User comprehension: All 15 participants completed the study within the assumed 1-hour timeframe and all demonstrated observable growth in fluency and confidence with the system as they progressed through the activities. Several of the participants moved quite briskly through the activities, particularly during the repetitive sections, clicking quickly through the repeated setup and instruction screens. All participants watched the majority of the instruction videos at least once per activity, while the videos for the more challenging and involved activities were watched by many participants two or three times. Several of the participants practiced doing the activities before electing to record their performance.

Movement assessment: Each user defined recorded task performance comprised one video of our data base. For the performance evaluation of our movement assessment algorithms, we manually excluded 70 problematic videos from our test dataset, for reasons such as the subject did not follow the instructions correctly, the video contains more than one simulated error, or the video was not recorded properly. As a result, we constructed a test set consisting of 410 videos in total.

We quantify the performance of the movement assessment algorithms using precision and recall for each type of performance error. The recall values indicate the fraction of all videos with a particular simulated performance error (e.g. low speed) to the number of videos whose performance errors are correctly identified by the movement assessment system. The precision values, on the other hand, indicate the fraction of videos identified by the assessment system that actually contains the particular error. In Figure 2, we show the precision and recall plot for the six performance errors. We also report the precision and recall in identifying videos with 'no error' (i.e. the videos captured in the first two repetitions for each activity) by treating them as an independent class.

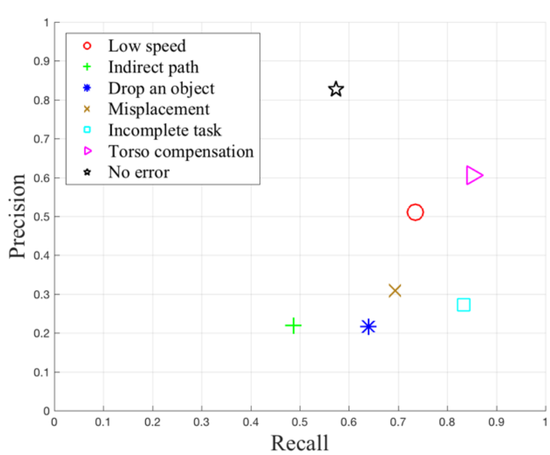

Figure 2: Precision-Recall plot for the six primary error types and the 'no error' case.

Among the six performance errors, five have recall values higher than 0.7. This suggests that our current movement assessment implementation is able to correctly identify a particular simulated performance error with decent accuracy. We note however that four of the performance errors have relatively low precision values, which can be attributed in the case of "Misplacement" and "Incomplete task" to the definition of these error types. For example, placing an object in the wrong location could result in both 'Misplacement' and 'Incomplete task", which therefore decreases the precision of the 'Incomplete task' error detection. 'Drop an object' and 'Indirect path' also show somewhat lower precision values. We use hand and object trajectory information to detect whether the path is/is not indirect and if the object is/is not dropped. It is possible that unstable trajectory information is caused by object occlusion by the hand or failure to detect the hand or object.

4.2 Pilot Study 2

Set up: We evaluated the feasibility of our system with 3 stroke survivors in a lightly supervised study conducted at Emory University Rehabilitation Hospital. The 3 participants were 1 woman (left-handed, mild moderate impairment); 2 men (1 left-handed, moderate impairment; 1 right-handed, mild moderate impairment). The patients were asked to complete a similar movement protocol as the unimpaired subjects in pilot study 1, except they were not instructed to simulate errors in task repetitions 3 and 4, and the entire study also took place under the supervision of a physiotherapist.

For this second study, we introduced an automated patient feedback mechanism to the system, which is critical for interactive therapy with patients. We know from previous work that feedback during unsupervised home-based therapy has to be simple but effective in promoting adherence and active learning. Both performance of the activity and movement quality, in that order of significance, has to be addressed by the feedback. The limitations of the low cost automated movement tracking and assessment system by definition, also necessitate simplified feedback structures. Feedback in this case was provided to patients through a simple "rating of performance" scheme shown on the tablet. All task performances with minimal detected errors produced an "excellent" rating. All task performances with significant errors in object placement or task completion but with minimal errors in movement quality were given a "very good" rating. Task performances with a significant error in both activity and movement quality were given a "nice try" rating.

Results and Discussion

The system successfully captured the performance of all three patients. The more impaired participants could not complete all four repetitions of the 12 tasks within the allotted one-hour time frame but all participants completed at least two repetitions of each task. The therapist intervened to advance the system to the next task for activities that participants found challenging after two repetitions. The tasks that each participant found challenging depended on the specific movement challenges of each patient.

The adaptation of therapy to each patient's profile (dosage of each task, sequence of tasks, type of feedback) will need to be addressed in our future adaptation algorithms that are currently under development.

All three participants understood the functioning of the system and demonstrated observable growth in fluency and confidence as their session progressed. All three patients also asked that written instructions be presented in a larger font (which is a typical request for older populations). The therapist had to intervene a number of times at the beginning of each session to explain some of the system functions, especially for patients not very familiar with interactive technologies. During the debrief sessions with the patients and the therapists after the study, they proposed that this issue could be addressed through the development of an introductory video explaining system functionality. During the study, all three patients watched the instruction videos describing the activity tasks at least two times. One of the patients suggested that task comprehension could be enhanced if the videos were accompanied by audio instructions that described with accuracy the goal of each action, while allowing for adaptation to each patient's ability. For example, "reach and touch the object with any digit"; or "reach and grasp the object and move it from the second to the third quadrant on the mat using the type of grasp most effective for you". The therapists further suggested that numbering the mat quadrants and the objects could assist with task comprehension and help minimize error.

The participants expressed mild annoyance at having to repeatedly click through extensive instructions for each task repetition (an issue broadly observed also with the unimpaired participants). We will address this issue by allowing the user to switch the system to expert mode where only the videos (with audio narration) are repeated ahead of each task repetition and the basic set up instructions for each task are skipped. In the debrief session, the participants and the therapists discussed their overall impressions of the cross-mapping potential between the objects. They described how they liked how the objects didn't obviously "give away" what they could be, or what they could be used for, indicating that the setup succeeded in prompting patients to think actively and creatively during the exercise.

The movement capture and analysis challenges observed with unimpaired subjects persisted and were exacerbated in the pilot with patients. Patients dropped and misplaced objects and failed to complete tasks much more often than unimpaired subjects. Patients needed to occasionally use the unimpaired limb to recover from task failure or task associated challenges, thus creating even more occlusions. The lighting conditions in the hospital were not as controlled as in the lab thus creating some instances where object tracking was dropped, which influences the rating of task performance.

Patients paid significant attention to the ratings of task performance since the ratings provided motivation for adherence, and active self-evaluation of movement and improvement. In instances where uncertainties in movement tracking produced inconsistent ratings (i.e. an excellent performance receiving a very good rating because of occlusion) the therapist intervened to mitigate the potential confusion. Given that the therapist will ultimately not be present at the home, ratings need to be robust and more reliable in our system or the patient will quickly get frustrated and abandon it. It is important to note that small errors in movement tracking and analysis can also happen when an unimpaired subject plays Kinect based games, but these are not catastrophic. The game is counting on the unimpaired user to analyze their error in performance and to filter out small system errors. This kind of self-evaluation and error filtering is an impossible task for impaired patients who do not have access to good internal feedback on their performance. Finally, the study also emphasized that our vision system cannot track in detail three elements that are critical to performance evaluation by the therapists: elbow extension, wrist flexion and hand digit opening and closing in relation to object engagement and manipulation.

In response to these findings, we are now investigating different movement capture approaches through deep learning methods to further improve the generalizability of the system without increasing the cost or complexity of the infrastructure. Our learning-based approaches will rely on expert ratings of a large number of video recordings of task performances by patients with different movement challenges. Over the past three years, we have worked with expert therapists to develop a set of limited movement quality and activity performance features that can be rated by therapists in a standardized manner using a limited scale (0 to 3). To train our system we will combine supervised and unsupervised learning approaches so we can also capture potential latent features that therapists may use for ratings without being fully aware of their use [31]. This approach, however, requires that we switch our movement capture infrastructure. We require high resolution videos of the hand shape as well as the torso and limb that are captured from an angle that most therapists use to evaluate movement (profile view from the side of the impaired limb). We are thus replacing the Kinect with two small high resolution video cameras with one focused on the torso and limb and the other on the hand shape.

5 CONCLUSION

Developing interactive systems for home-based rehabilitation presents numerous challenges in terms of sensing approach, movement assessment, protocol adaptation, and patient adherence. Building on our prior work in the clinic, we introduce a simplified computer vision based system, driven by the patient using an interactive tablet interface. Based on findings from two pilot studies with unimpaired and impaired users, the system was deemed understandable, relatively straightforward to use, and functioned well in more controlled conditions. Important areas for improvement include increasing the robustness of the movement capture and assessment functions, refining the patient-system interactions, and developing a more generalized approach through learning based methods.

REFERENCES

[1] Craig Anderson, Sally Rubenach, Cliona Ni Mhurchu, Michael Clark, Carol Spencer, and Adrian Winsor. Home or hospital for stroke rehabilitation? Results of a randomized controlled trial: I: Health outcomes at 6 months. Stroke, 31(5): 1024–1031, 2000.

[2] Craig Anderson, Cliona Ni Mhurchu, Paul M Brown, and Kristie Carter. Stroke rehabilitation services to accelerate hospital discharge and provide home-based care. Pharmacoeconomics, 20(8): 537–552, 2002.

[3] Lesley Axelrod, Geraldine Fitzpatrick, Jane Burridge, Sue Mawson, Penny Smith, Tom Rodden, and Ian Ricketts. 2009. The reality of homes fit for heroes: design challenges for rehabilitation technology at home. Jl of Assist. Technologies, 3(2):35–43, 2009.

[4] Madeline Balaam, Stefan Rennick Egglestone, Geraldine Fitzpatrick, Tom Rodden, Ann-Marie Hughes, Anna Wilkinson, Thomas Nind, Lesley Axelrod, Eric Harris, Ian Ricketts, Susan Mawson, and Jane Burridge. 2011. Motivating mobility: designing for lived motivation in stroke rehabilitation. Proc CHI '11, 3073-3082.

[5] Michael Baran, Nicole Lehrer, Margaret Duff, Vinay Venkataraman, Pavan Turaga, Todd Ingalls, W Zev Rymer, Steven L Wolf, and Thanassis Rikakis. 2014. Interdisciplinary concepts for de- sign and implementation of mixed reality interactive neurorehabilitation systems for stroke. Physical therapy, 2014.

[6] Emelia Benjamin et. al. 2017. Heart Disease and Stroke Statistics – 2017 Update: A report from the American Heart Association, Circulation, Volume 135, Issue 10, pp 146 – 603

[7] Yinpeng Chen, Weiwei Xu, Hari Sundaram, Thanassis Rikakis, and Sheng-Min Liu. 2007. Media adaptation framework in biofeedback system for stroke patient rehabilitation. In Proceedings of the 15th ACM international conference on Multimedia (MM '07). ACM, New York, NY, USA, 47-57.

[8] Yinpeng Chen, Margaret Duff, Nicole Lehrer, Sheng-Min t Liu, Paul Blake, Steve Wolf, Hari Sundaram, and Thanassis Rikakis. 2011. A Novel Adaptive Mixed Reality System for Stroke Rehabilitation: Principles, Proof of Concept and Preliminary Application in Two Patients". Topics in Stroke Rehabilitation 2011;18 (3), 212–230

[9] Yinpeng Chen, Margaret Duff, Nicole Lehrer, Hari Sundaram, Jipeng He, Steven L Wolf & Thanassis Rikakis. 2011. A computational framework for quantitative evaluation of movement during rehabilitation. Proc AIP '11, 1371:317.

[10] Mark Cutkosky. On grasp choice, grasp models, and the design of hands for manufacturing tasks. Robotics and Automation, IEEE Transactions on, 5(3):269–279, 1989.

[11] Margaret Duff, Yinpeng Chen, Long Cheng, Sheng-Min Liu, Paul Blake, Steven L Wolf, and Thanassis Rikakis. Adaptive mixed reality rehabilitation improves quality of reaching movements more than traditional reaching therapy following stroke. Neurorehabilitation and neural repair, page2012.

[12] Thomas Feix, Roland Pawlik, Heinz-Bodo Schmiedmayer, Javier Romero, and Danica Kragic. A comprehensive grasp taxonomy. In Robotics, Science and Systems: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation, pp 2–3, 2009.

[13] Qiushi Fu and Marco Santello. Towards a complete description of grasping kinematics: a framework for quantifying human grasping and manipulation. In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, pages 8247–8250. IEEE, 2011.

[14] Sarah J Housman, Kelly M Scott, and David J Reinkensmeyer. A randomized controlled trial of gravity- supported, computer-enhanced arm exercise for individuals with severe hemiparesis. Neurorehabilitation and neural repair, 2009.

[15] Nicole Lehrer, Yinpeng Chen, Margaret Duff, Steven L Wolf, and Thanassis Rikakis. Exploring the bases for a mixed reality stroke rehabilitation system, part ii: Design of interactive feedback for upper limb rehabilitation. Journal of neuroengineering and rehabilitation, 8(1):54, 201

[16] Mindy Levin, Jeffrey Kleim, and Steven L Wolf. 2008. What do motor recovery and compensation mean in patients following stroke? Neurorehabilitation and neural repair, 2008.

[17] Cheng Li and Kris Kitani, (2013) "Model Recommendation with Virtual Probes for Ego-Centric Hand Detection". International Conference on Computer Vision (ICCV 2013). Dec 2013.

[18] Peter Kirk, P. et al. Motivating Stroke Rehabilitation Through Music: A Feasibility Study Using Digital Musical Instruments in the Home. In Proc CHI '16, 1781-1785.

[19] John W Krakauer. Arm function after stroke: from physiology to recovery. In Seminars in neurology, volume 25, pages 384–395. [New York]: Thieme-Stratton Inc. [c1981-, 2005.

[20] David Minnen, and Zahoor Zafrulla. "Towards robust cross-user hand tracking and shape recognition." IEEE International Conference on Computer Vision Workshops, 2011.

[21] David Nelson (1988). Occupation: Form and performance. American journal of Occupational Therapy, 42, 633-641.

[22] Kelsey Picha and Dana Howell. 2017. A model to increase rehabilitation adherence to home exercise programmes in patients with varying levels of self-efficacy. Musculoskeletal Care, 2017, April 12, 10.1002/msc.1194.

[23] Anne Marie Piper, Ross Campbell, and James D. Hollan. 2010. Exploring the accessibility and appeal of surface computing for older adult health care support. Proc CHI '10, 907-916.

[24] David Reinkensmeyer et al. 2016. Computational neurorehabilitation: modeling plasticity and learning to predict recovery. Jrnl of Neuroeng. and Rehab., vol. 13, (1), pp. 42, 2016.

[25] Elham Saraee, Saurabh Singh, Kathryn Hendron, Mingxin Zheng, Ajjen Joshi, Terry Ellis, and Margrit Betke. 2017. ExerciseCheck: Remote Monitoring and Evaluation Platform for Home Based Physical Therapy. Proc PETRA '17, 87-90.

[26] Jamie Shotton, Ross Girshick, Andrew Fitzgibbon, Toby Sharp, Mat Cook, Mark Finocchio, Richard Moore, Pushmeet Kohli, Antonio Criminisi, Alex Kipman, and Andrew Blake. "Efficient Human Pose Estimation from Single Depth Images", Trans. PAMI, IEEE, 2012.

[27] Edward Taub, Miller NE, Novack TA, Cook EW III, Fleming WC, Nepomuceno CS, Connell JS, Crago JE. Technique to improve chronic motor deficit after stroke. Arch Phys Med Rehabil. 1993;74:347–354

[28] Catherine Trombly. Occupational therapy for physical dysfunction, baltimore: Wlliams & wfilkins. Augmentative and Alternative Communication, 283, 1989.

[29] Roland Van Peppen, Gert Kwakkel, Sharon Wood-Dauphinee, H JM Hendriks, Ph J Van der Wees, and Joost Dekker. The impact of physical therapy on functional outcomes after stroke: what's the evidence? Clinical rehabilitation, 18(8):833–862, 2004.

[30] David Webster, Ozkan Celik. Systematic review of Kinect applications in elderly care and stroke rehabilitation, J. of Neuroengineering and Rehab., vol. 11, pp. 1–24, 2014.

[31] Steven L Wolf, Pamela A Catlin, Michael Ellis, Audrey Link Archer, Bryn Morgan, and Aimee Pia- centino. Assessing wolf motor function test as outcome measure for research in patients after stroke. Stroke, 32(7):1635–1639, 2001.