33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

A Low Cost, Adaptive Mixed Reality System for Home-based Stroke Rehabilitation

Yinpeng Chen, Michael Baran, Hari Sundaram and Thanassis Rikakis

Abstract— **This paper presents a novel, low-cost, real-time adaptive multimedia environment for home-based upper extremity rehabilitation of stroke survivors. The primary goal of this system is to provide an interactive tool with which the stroke survivor can sustain gains achieved within the clinical phase of therapy and increase the opportunity for functional recovery. This home-based mediated system has low cost sensing, off the shelf components for the auditory and visual feedback, and remote monitoring capability. The system is designed to continue active learning by reducing dependency on real-time feedback and focusing on summary feedback after a single task and sequences of tasks. To increase system effectiveness through customization, we use data from the training strategy developed by the therapist at the clinic for each stroke survivor to drive automated system adaptation at the home. The adaptation includes changing training focus, selecting proper feedback coupling both in real-time and in summary, and constructing appropriate dialogues with the stroke survivor to promote more efficient use of the system. This system also allows the therapist to review participant's progress and adjust the training strategy weekly.**

I. Introduction

Stroke is a leading cause of disability in the United States. Every 40 seconds, someone in the United States suffers a stroke [11], often leading to physiological impairment. Up to 85% of stroke survivors have a sensorimotor deficit in the arm, such as muscle weakness, abnormal muscle tone, and lack of coordination during voluntary movement [5]. It has been shown that rehabilitation involving motor learning can lead to recovery of lost functionality [10].

Research groups that have applied interactive therapy to stroke rehabilitation have demonstrated improvements in kinematic and functional performance of the upper extremity [9], [12]. However, the feedback provided by most existing systems does not communicate multiple aspects of the movement simultaneously and in an integrated manner. Our lab developed an Adaptive Mixed Reality Rehabilitation (AMRR) system [6] to address this limitation. It utilizes an integrated environment of both physical elements and interactive audio and visual feedback to train reach and grasp activities. The AMRR system was demonstrated improvement in movement performance and clinical scores [2].

Repeated visits for clinical-based therapy are costly to a patient, both financially and logistically. Recent research and development has led to telerehabilitation systems for home-based therapy. These systems range in approaches to utilize virtual environments [8] and hand-adorned sensors [7]. These systems face challenges in terms of engagement of feedback, the lack of system adaptation to individual performance and the lack of a plan for evolving tasks and feedback paradigms over an extended period of time.

Y. Chen, M. Baran, H. Sundaram and T. Rikakis are with the School of Arts, Media and Engineering, Arizona State University, Tempe, AZ 85287. Email: {yinpeng.chen, michael.baran, hari.sundaram, thanassis.rikakis}@asu.edu

Therefore, we develop a Home-based Adaptive Mixed Reality Rehabilitation system (HAMRR) to integrate engaging multimedia feedback, semi-automatic adaptation and evolving task and feedback plan in a low-cost multimedia environment. This HAMRR system is designed to support unsupervised reach and grasp training at home over 12-24 months (after clinical therapy using our AMRR system). There are three key contributions in this system - (a) a low-cost physical design that supports multimodal sensing and allows stroke survivors to setup easily at home, (b) an engaging and hierarchical multimedia feedback design, from real-time to summary feedback, which allows stroke survivors to experience evolving feedback and facilitates active learning by gradually reducing dependency on media feedback, and (c) a semi-supervised adaptation framework that customizes the system based on the participant's progress automatically and allows a therapist to review the progress and adjust training strategy weekly. In this paper, we refer to a stroke survivor as a *participant*.

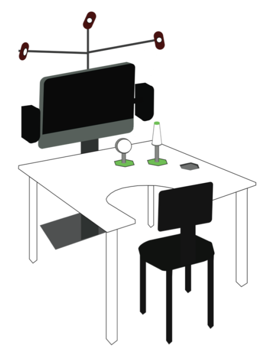


Fig. 1. The physical setup of the HAMRR system.

This paper is organized as follows. We summarize key features in our system design in Section II. In Section III, we present the adaptation framework. We show the functionality test in Section IV and conclude the paper in Section V.

II. Home-Based System Design

In this section, we introduce the physical setup, system architecture, and feedback design of our HAMRR system for training *reach and grasp* movement for stroke survivors.

A. Physical Setup

The physical setup of our HAMRR system includes *a media center*, *a table* and *a chair* (see Fig. 1). The *media center* is a slender aluminum tube frame with a cantilevered base that supports a 27-inch iMac computer, two speakers and three infrared Opti-Track cameras. The *table* is constructed from an aluminum tube frame. It is lightweight, such that it can be removed from the media center when not in use. The table contains three forms of embedded sensors. The rest area of the participant's impaired arm is covered by pressure sensors to check for starting location. Two buttons with capacitive sensing are placed close to the participant's unimpaired hand that allows the participant to interact with the system dialog. There are three tangible objects (base, button and cone) embedded on the table for the participant to reach. Each object has embedded color LEDs to provide feedback. Both button and cone objects have sensors to track object manipulation. The *chair* is embedded with force sensing resistors (mounted on the back and seat) to track participant's torso movement. The sensors are connected to the system through an XBee connection, with a sampling rate of 70 Hz. Compared to the hospital-used AMRR system, the HAMRR system significantly reduces sensing cost, and simplifies participant's setup. The participant only need to wear a wristband that is attached a reflective marker for 3D hand tracking.

B. System Architecture

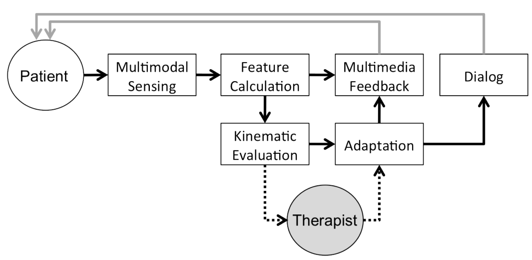


Fig. 2. The home-based system architecture.

The HAMRR system architecture is shown in Fig. 2. The participant's movement is captured by multimodal sensing that includes tracking 3D hand position by optical cameras, tracking object manipulation by force sensors on objects as well as tracking torso movement by force sensors on the chair. Based on the sensing data, we can derive kinematic features related to the reach and grasp movement. These features are used to generate tangible/visual/auditory feedback that helps the participant self-assess the movement performance. The kinematic features are used for quantitative performance evaluation, which informs the adaptation framework to customize training tasks and feedback environments. Our system provides a system dialog to guide the participant through training. Our system also allows a therapist to review the participant's performance data weekly and make appropriate adjustments in adaptation strategy.

C. Multimedia Feedback

Our HAMRR system intuitively communicates to the stroke survivor levels of his/her performance and direction for improvement by using three types of feedback - *tangible feedback* from color LEDs embedded within the object, *auditory feedback* from two speakers and *visual feedback* rendered on the iMac screen. The feedback is designed to provide dynamic feedback experience that evolves over time, incentivizes training, facilitates active learning, and accommodates more complex tasks. Therefore, the feedback is structured to provide information about the movement in three different levels - (a) *real time*, (b) *post-trial summary*, and (c) *post-set summary*. For any training set (including several reaches), only one of these three levels is selected.

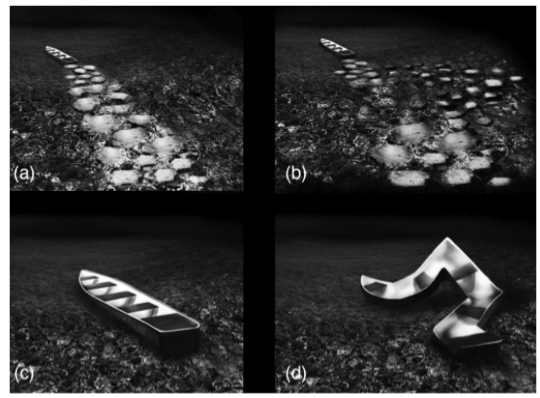


Fig. 3. Visual feedback summaries after a trial ((a) and (b)) and after a set ((c) and (d)). (a) represents an efficient reach and (b) represents trajectory error to the right. (c) is a representation of efficient and consistent task completions and (d) represents disjoint movement.

Real-time feedback is coupled to the participant's hand trajectory and speed, which are key features for completing a reach and grasp activity. The LEDs embedded within the object base change color based on the participant's trajectory error. The green, orange and red colors indicate no error, small error and big error respectively. Real-time auditory feedback communicates reaching speed to the participant by note density. The sound is designed to encourage the participant to reach in a natural speed. In *post-trial summary*, a visual summary of trajectory performance for a reach is communicated by the color and spatial distribution of stones in the water, as seen in Fig. 3a and 3b. This display summary helps the participant begin to think about the strategy of execution. In *post-set summary*, either a visual summary or an auditory summary communicates to the participant an *affective tag* with his/her performance (or one movement aspect) after a set of reaches, such as disjoint movement or slow movement. For example, in Fig. 3d, abrupt changes in the boats contour communicate a pattern of reaching without elbow-shoulder joint coordination. Visual tags include trajectory inaccuracy, jerkiness, disjoint movement, and task incompletion. Auditory tags includes ballistic movement, slow movement, inconsistent completion time, and hesitant movement. The details of feedback design can be found in [1].

D. Interactive Dialog

Our system provides an interactive dialog when a new task or feedback stream is introduced. During the dialog, the participant is firstly shown graphic instructions. Then the participant can practice with interactive feedback several times until he/she understands the instructions. As part of the interactive dialog, the participant is asked several questions with yes/no answer to check if he/she understands the feedback mappings. The participant can use the unimpaired hand to select the answer using the embedded table buttons.

III. Semi-Supervised Adaptation Framework

The HAMRR system is adaptable to address each participant's impairments based on his/her progress. Both tangible objects and interactive feedback can be adjusted to maintain a level of challenge and engagement appropriate for each stroke survivor. The adaptation is crucial to promote more efficient use of the system, resulting in better active learning and movement performance. The adaptation in home-based system is very challenging due to the absence of a therapist. Therefore, we need an automatic adaptation framework which customizes the system based on the participant's progress. Our basic idea is to integrate the real-time quantitative kinematic evaluation, correlation analysis between feedback and movement performance, and therapist's weekly recommendations into a semi-supervised adaptation framework. The adaptation framework customizes the system in the participant's daily training automatically and allows a therapist to review the rehabilitation progress weekly, and fine-tune the training strategy.

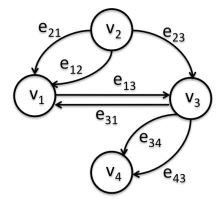
A. Adaptation Components

In our HAMRR system, both tangible objects and feedback parameters can be changed algorithmically. The adaptation for tangible objects includes changing the object position and the object type (e.g. base, button or cone).

The feedback adaptation includes three parts: (a) feedback level, (b) media selection, and (c) feedback sensitivity. First, the feedback can be generated either in real-time (e.g. feedback from color LEDs embedded within the object), or after a trial (e.g. visual summary of stone distribution in Fig. 3b), or after a set of reaches (e.g. boat shape in Fig. 3d). At each feedback level, we can select proper media streams for different training purposes. For example, if we focus on smooth reaching acceleration and deceleration at real-time feedback level, we might turn off the tangible feedback from embedded LEDs and turn on the auditory feedback. Finally, each feedback is associated to a sensitivity filter to map the raw data (finite or infinite) into a normalized movement feature that is used to generate feedback. Thus, the feedback sensitivity can be controlled by filter parameters.

B. Graphical Adaptation Representation

We now discuss how to represent system adaptation using a graph. First, we consider our system as a network G that has vertex set V and edge set E (i.e. $G = (V, E)$). Each vertex v_i indicates a training scenario (or system state) that sets up every component in our system, including object and feedback information. A system state can be represented by a state vector. Each component of the vector is corresponding to an object property (e.g. object ID or object type) or feedback information such as feedback level, on/off of media streams or feedback sensitivity parameters. The vertex set V includes all possible pre-defined system states $\{v_1, v_2, \ldots\}$. Each directional edge e_{ij} indicates a possible system change from state v_i to state v_j. Hence, an adaptation action is represented by an edge in the system network, indicating a change between two training scenarios (or system states). Fig. 4 shows an example of four training scenarios ($v_1 - v_4$) and possible adaptation options (jump between scenarios, e.g. e_{12}, e_{34}). The table on the right shows some system state parameters for these four scenarios.



System State Parameters	v_1	v_2	v_3	v_4
Feedback level	Real-time	Real-time	Post-trial	Post-set
Training focus	Trajectory	Speed	Trajectory	Jerkiness
Media stream	Tangible	Audio	Visual	Visual
… …	… …	… …	… …	… …

Fig. 4. Graphic representation of four training scenarios (or system states).

The system network is constructed by experts, including therapists and media experts. The experts first define all training scenarios including training foci and the corresponding object and feedback selection. Each training scenario is represented by a system state (or a vertex in the network). The experts determine the value for each component (i.e. object and feedback parameters) for each predefined system state. Secondly, the experts determine possible adaptation options (or move from a state to another) by drawing an arrow between two training scenarios (or state vertices). The experts determine adaptations carefully to balance the continuity and variation.

C. Decision-Making Process

The training starts from an initial training scenario which is determined by experts based on the participant's rehabilitation at hospital. After every set (including several reaches), an adaptation decision is made to either stay in the current training scenario (or current state) or move to another training scenario. Thus, an adaptation decision includes two steps: (a) determining *stay* or *move*, and (b) determining a *new scenario* if *move* is chosen.

The decision for *stay* or *move* is made based on the length of duration staying in the current scenario and the participant's performance in movement aspects that are focused. The *move* decision is made if either the participant stays in the current scenario for too long (e.g. over 5 sets) or the participant's performance in focused movement aspects satisfies the expectation (e.g. trajectory error is less than 2cm).

If we decide to *move*, the new scenario should focus on the movement aspects in which the participant has major deficit and at the same time does not introduce a big variation compared to the current scenario. We address this by searching over all neighbor scenarios of the current scenario in the system network for the best neighbor with maximum utility. Since the network topology is determined by experts to balance the variation and continuity, moving from the current scenario to any neighbor should not introduce a big variation. We developed a utility function to evaluate all neighbor scenarios in terms of effect on the participant's improvement. The neighbor with maximum utility provides the best assistance. Next, we discuss the utility function.

D. Utility Function

The utility function integrates three main components - *prior*, *history*, and *performance expectation*. The *prior utility* for each training scenario (or state vertex) represents the priority for that scenario to be selected. The prior utility is set at the beginning based on the participant's rehabilitation in hospital and fine-tuned weekly based on the participant's progress by a therapist. For example, if the therapist wants to only use scenario v_i, v_j, v_k for a week training, this can be achieved by setting the prior utility for these three states like $(u_p(v_i) = 0.5, u_p(v_j) = 0.3, u_p(v_k) = 0.2)$ and zero for all other system states. The *history utility* u_h signifies the usage distribution over all training scenarios in the past several days. The *performance expectation utility* u_e for a training scenario measures distance between the expected performance and the predicted performance for the participant after using the scenario. The participant's movement performance can be measured using kinematic impairment measurement (KIM) [3]. The expected performance can be determined by the therapist and the performance prediction after using a training scenario can be done using our previous research on media adaptation model [4]. The overall utility function is a linear combination of three utilities as follows:

$$U = \omega_p \cdot u_p + \omega_h \cdot (-u_h) + \omega_e \cdot (-u_e) \qquad (1)$$

where ω_p, ω_h, and ω_e are three weights. The history utility and performance expectation utility have negative sign because the higher usage and expectation-prediction distance for a scenario will reduce its chance to be selected.

E. Therapist Intervention

Our semi-supervised adaptation framework allows a therapist to fine-tune the adaptation strategy. The therapist may review the participant's training progress weekly and adjust the adaptation strategy using a GUI connected to the home system through internet. The therapist can change the system network topology by adding/removing edges to enable/disable adaptations between scenarios (or system states), update the prior utility over different scenarios, and adjust the weights for three utilities.

IV. Functionality Test

We tested the functionality of our HAMRR system by three unimpaired subjects. The average processing time for kinematic analysis per frame is 5ms which is less than the time interval between two frames (i.e. 10ms for frame rate 100Hz). Thus, the kinematic analysis does not introduce delay. The average processing time for post-trial kinematic evaluation is less than 0.5s, which is well below the 2s-3s rest time between trials. Each subject used the system for about 1.5 hours, without software/hardware problems.

V. Conclusion and Future Work

In this paper, we presents a Home-based Adaptive Mixed Reality Rehabilitation (HAMRR) system for upper extremity rehabilitation of stroke survivors. This HAMRR system has low cost sensing, engaging and hierarchical multimedia feedback and semi-supervised adaptation framework to customize the training. The system is designed to sustain gains achieved within the clinical phase of therapy and increase the opportunity for functional recovery. This system also allows the therapist to review participant's progress and adjust the training strategy weekly. Future work includes studying the efficacy of the HAMRR system. We plan to do a user study of 10 unimpaired subjects to determine if the feedback summary efficiently communicates the movement tags (e.g. slow movement, jerkiness). Each subject is asked to match the visual and auditory summary with movement tags. We also plan another user study with two stroke survivors to assess the integration and user-friendliness of use of HAMRR within the home environment.

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contribution of colleagues - Margaret Duff, Nicole Lehrer, and Diana Siwiak.

References

[1] M. Baran, N. Lehrer, D. Siwiak, and et al. Design of a home-based adaptive mixed reality stroke rehabilitation system. submitted to IEEE EMBC 2011, 2011.

[2] Y. Chen, M. Duff, N. Lehrer, and et al. A novel adaptive mixed reality system for stroke rehabilitation: principles, proof of concept and preliminary application in two patients. *Topics in Stroke Rehabilitation*, Accepted for publication, 2010.

[3] Y. Chen, M. Duff, N. Lehrer, H. Sundaram, J. He, S. L. Wolf, and T. Rikakis. A computational framework for quantitative evaluation of movement during rehabilitation. International Symposium on Computational Models for Life Sciences, 2011.

[4] Y. Chen, W. Xu, H. Sundaram, T. Rikakis, and S. M. Liu. A dynamic decision network framework for online media adaptation in stroke rehabilitation. *ACM Trans. On Multimedia Computing, Communications and Applications*, 5:1–38, 2008.

[5] M. C. Cirstea, A. B. Mitnitski, and et al. Interjoint coordination dynamics during reaching in stroke. *Exp Brain Res*, 151(3):289–300, 2003.

[6] M. Duff, Y. Chen, N. Lehrer, and et al. An adaptive mixed reality training system for stroke rehabilitation. *IEEE Trans. Neural Systems and Rehabilitation Engineering*, 18:531–541, 2010.

[7] W. Durfee, J. Carey, D. Nuckley, and J. Deng. Design and implementation of a home stroke telerehabilitation system. IEEE EMBC 2009, Sept. 2009.

[8] M. K. Holden, T. A. Dyar, and L. Dayan-Cimadoro. Telerehabilitation using a virtual environment improves upper extremity function in patients with stroke. *IEEE Trans. Nueral Systems and Rehabilitation Engineering*, 15:36–42, 2007.

[9] D. Jack, R. Boian, and et al. Virtual reality-enhanced stroke rehabilitation. *IEEE Trans. Neural Syst Rehabil Eng*, 9:308–318, 2001.

[10] J. W. Krakauer. Motor learning: its relevance to stroke recovery and neurorehabilitation. *Curr Opin Neurol*, 19:84–90, 2006.

[11] Donald Lloyd-Jones, Robert J. Adams, and et al. Heart disease and stroke statistics 2010 update: a report from the american heart association. *Circulation*, 121:146–215, Feb 2010.

[12] L. Piron, P. Tonin, F. Piccione, and et al. Virtual environment training therapy for arm motor rehabilitation. *Presence: Teleoperators and Virtual Environments*, 14:732–740, 2005.