Gesture Therapy: A Vision–Based System for Upper Extremity Stroke Rehabilitation

L. Enrique Sucar, Roger Luis, Ron Leder, Jorge Hernández and Israel Sánchez

Abstract—Stroke is the main cause of motor and cognitive disabilities requiring therapy in the world. Therefore it is important to develop rehabilitation technology that allows individuals who have suffered a stroke to practice intensive movement training without the expense of an always-present therapist. We have developed a low-cost vision-based system that allows stroke survivors to practice arm movement exercises at home or at the clinic, with periodic interactions with a therapist. The system integrates a virtual environment for facilitating repetitive movement training, with computer vision algorithms that track the hand of a patient, using an inexpensive camera and a personal computer. This system, called Gesture Therapy, includes a gripper with a pressure sensor to include hand and finger rehabilitation; and it tracks the head of the patient to detect and avoid trunk compensation. It has been evaluated in a controlled clinical trial at the National Institute for Neurology and Neurosurgery in Mexico City, comparing it with conventional occupational therapy. In this paper we describe the latest version of the Gesture Therapy System and summarize the results of the clinical trial.

I. INTRODUCTION

Stroke is the leading cause of long-term disabilities in the world. Only the U.S.A. it affects an estimated 6.4 million people [1]. Approximately 80% of acute stroke survivors lose arm and hand movement skills. Movement impairments after stroke are typically treated with intensive, hands-on physical and occupational therapy for several weeks after the initial injury. Unfortunately, due to economic pressures on health care providers, stroke patients are receiving less therapy and going home sooner. The ensuing home rehabilitation is often self directed with little professional or quantitative feedback. Even as formal therapy declines, a growing body of evidence suggests that both acute and chronic stroke survivors can improve movement ability with intensive, supervised training. Thus, an important goal for rehabilitation engineering is to develop technology that allows individuals with stroke to practice intensive movement training without the expense of an always-present therapist. Although their are some recent developments of robotic systems for rehabilitation [2], these are too expensive for their use at home or in small clinics. For example, the Armeo arm exoskeleton device has a cost of approx. $40,000 USD [3]. Thus, a low cost alternative is required for home therapy.

We have developed a vision–based system for rehabilitation after stroke, called “Gesture Therapy” (GT). The objective of the system is to allow individuals with stroke to practice arm movement exercises at home or at the clinic, without the need of an always present therapist. The system makes use of a virtual environment for facilitating repetitive movement training that provides simulation activities relevant to daily life. A web cam is used for tracking the hand of the patient using a color ball attached to a gripper. The vision algorithm locates and tracks the ball using color and texture information, and based on the apparent diameter of the ball, estimates its 3-D position in space. The coordinates of the hand are sent to the simulator so that the patient interacts with a virtual environment by moving his/her impaired arm, performing different tasks designed to mimic real life situations and thus oriented for effective rehabilitation. The specially designed gripper includes a pressure sensor to include hand exercises, which are also considered in the virtual environment. An additional software component detects movements of the patient’s head to avoid trunk compensation. The estimated hardware cost of the system is in the order of $1,000 USD, including a personal computer, a web cam, and the gripper.

The GT system has been installed at the rehabilitation unit at the National Institute of Neurology and Neurosurgery (INNN) in Mexico City. A first clinical trial has been conducted at INNN comparing Gesture Therapy with conventional therapy. The clinical evaluation was done using the Fugl-Meyer and Motricity Index scales, and also an Intrinsic Motivation scale was administered at the end of the treatment. The results show that both groups have a significant improvement according to the two standard clinical scales, but there is no significant difference between both groups. However, according to the Intrinsic Motivation scale, a stronger motivation and attachment to the treatment is observed for the patients that used Gesture Therapy. We consider that this is an important result since long term motivation and attachment are decisive for maximal recovery.

Preliminary parts of this work have been reported in [4], [5]. In this paper we describe a new and more complete version of the Gesture Therapy system, and present for the first time the complete results of the clinical trial.
II. Gesture Therapy

Gesture Therapy integrates a simulated environment for rehabilitation with a gesture tracking software in a low-cost system for rehabilitation after stroke. The movement of the patient’s affected hand is tracked based on an image sequence obtained by a low–cost camera. The tracker estimates the 3-D coordinates of the hand in each frame, and sends this information to the simulated environment, so that the patient can interact with the games and observe the results in the screen. The physical system has 3 main elements: (a) a personal computer, used to run the software for the simulated environment and the visual tracker; (b) a web cam, which follows the movements of the patient by tracking a colored ball attached to the gripper, tracking also the face of the patient to detect trunk compensation; and (c) a hand grip, used to follow the motion of the hand and to measure hand grasp strength.

A. Simulated environment

The ARMEO [3] simulated environment has three key elements: therapy activities that guide movement exercise and measure movement recovery, progress charts that inform users of their rehabilitation progress, and a therapist page that allows rehabilitation programs to be prescribed and monitored. The therapy activities are presented in the software simulation like games. These activities were designed to be intuitive even for patients with minimal cognitive or perceptual problems to understand. The simulated activities are for repetitive daily task-specific practice and were selected by their functional relevance and inherent motivation like stove cleaning, window mopping, fish cashing, fruit shopping, flower watering, driving, etc (see Fig. 1). The system configuration allows therapists to customize the software to enhance the therapeutic benefits for each patient, by selecting a specific therapy activity. It also provides facilities to define the range of motions of the hand of the patient, so it can be adapted according to each patient’s needs and progress in the therapy. Additionally, the system gives objective visual feedback of patient task performance as well as entertainment. The visual feedback has the effect of enhancing motivation and endurance along the rehabilitation process by a patient’s awareness of his/her progress, as we found in our clinical results.

B. Monocular tracker

Based on a single low–cost camera and a computer, the hand of the user (via a color ball attached to a gripper) is detected and tracked in a sequence of images to obtain its 3-D coordinates in each frame, which are sent to the simulated environment. First the object is tracked in 2-D and then the third coordinate (depth) is estimated. Tracking the object in 2-D is based on particle filters [6]. Initially, the object (color ball) to be tracked is captured, and color and texture histograms are obtained from the object region in the image. Color and texture information are combined with a simple motion model to track the position of the ball in the image using particle filters, and the 2-D position, \((x, y)\), is estimated as the mean of the particles. The color observation model is based on the HSV color representation, while the texture model uses an edge orientation histogram. Both estimates are combined using a simple Bayesian fusion, assuming conditional independence. The variance of the distribution of the particles is used to estimate the distance of the object to the camera, that is the depth, \(z\). Finally the tracker reports the existence of the object and its position in space, \((x, y, z)\), to the simulated environment.

C. Gripper

We have designed a gripper that has two main functions: (i) to aid in the 3-D tracking of the patient’s hand via a color ball attached to one extreme, and (ii) to measure the pressure of the patient’s hand, so that exercises that incorporate hand grasping (closing and opening) can be incorporated. The gripper being tested at our lab is depicted in figure 2.

The color ball is attached to the gripper using a screw, so this can be easily changed according to the patient and environment. Ideally, the color selected should be easily distinguishable from the patient’s clothes and the objects in the environment, to avoid potential confusions of the tracker. The ball is made of any matte (lambertian) material to minimize reflections; and its size should be so that it can be detected according to the distance of the patient to the camera. In the clinical trials, the patient is at a distance between one and two meters to the camera, so a ball of 2–3 inches in diameter is adequate.

The gripper has a “soft” part where the patient grasps it with his hand. The pressure is sensed by means of the compression of the air volume confined in the pressure sensitive part of the gripper, and converted to an electrical signal by means of a pressure transducer. The signal is digitized via an analog-digital converter, and sent to the base computer using a cable with a standard USB connection. The pressure sensor provides a digital quantity that is proportional...
to the force of the patient’s grip and is used in some of the games. For example, some games require the patient to hold an object (like an egg in the egg cooking game), and certain minimum (or maximum) force is required so that the object does not fall (or is damaged), and the task can be completed. These pressure thresholds (min., max.) can be configured according to the patient in the virtual environment.

D. Trunk compensation detection

Stroke patients frequently do trunk compensation when interacting with the virtual environment. This is not desirable, as they are not exercising the arm, therefor it could limit the benefits of the therapy. To limit trunk compensation we have incorporated an additional software module to detect and avoid trunk compensation. Observing several patients during the clinical trials, we notice that when they do trunk compensation, they usually tilt their trunk and head together.

As it is easier to detect and track the face of a person using computer vision, compensation detection is based on a face detector. We use the AdaBoost face detector [7], which is based on simple Haar features and a cascade of classifiers; it has shown a good performance and is also very efficient. This detector, however, is very sensitive to face orientation, so it generally does not detect a face when it is tilted more than 15 degrees. In this application we use this limitation to detect trunk compensation. We assume that the patient is in front of the camera so her face should be inside the field of view. The system detects the person’s face using the AdaBoost face detector; when it loses detection for \( n \) consecutive frames (\( n \) a configurable parameter), it assumes that the patient is doing trunk compensation.

The system provides two alternatives when it detects trunk compensation that can be set-up by the therapist. One option is to sound an alarm when trunk compensation is detected, so that the patient is aware and in principle avoids it. The other option is to block the communication with the virtual environment, so that virtual object does not move, and the patient is forced to stop trunk compensation to continue with the game.

### III. Clinical Study

A. Methodology

A clinical evaluation of the Gesture Therapy system has been conducted at the Rehabilitation Unit of the National Institute for Neurology and Neurosurgery (INNN) in Mexico City. It is a longitudinal and comparative study with 42 patients that have suffered a stroke. The patients were divided randomly into two groups; a control group with 22 patients, and a study group with 20 patients. The inclusion criteria were: (i) at least 4 weeks after stroke, (ii) able to lift their arm against gravity, (iii) free from additional orthopedic, neurological or rheumatological disease (iv) able to understand and follow instructions. The exclusion criteria were: (i) significant pain, (ii) instability of the affected shoulder, (iii) severe cognitive disfunction, (iv) hemispatial neglect. The main characteristics of both groups are summarized in Table I.

Both groups received treatment for 21 sessions, about 60 minutes each, during 7 weeks, 3 sessions per week. The control group received conventional occupational therapy, consisting of different exercises of the affected upper extremity guided by a therapist, using didactic material such as cones, balls, etc. The study group used the GT system guided by a therapist. Before each session a calibration procedure is done to define the range of motions of the hand (in \( x, y, z \)), and the therapist determines which games to use for each patient.

The impact of the therapy for both groups was evaluated using 3 different scales: (i) the Fugl-Meyer scale [8], (ii) the Motricity Index [9], and an Intrinsic Motivation Survey [10]. The Fugl-Meyer and Motricity Index scales were applied before and after the therapy to each patient in both groups; while the Intrinsic Motivation Scale was applied to each patient of both groups at the end of the clinical study. Next we summarize the results of this study.

#### B. Results

We first analyzed the evolution of both groups of patients in terms of the Fugl-Meyer scale and Motricity Index. Both groups present a significant improvement (according to the Wilcoxon statistical test with \( p < 0.5 \)) after the 21 therapy sessions in terms of motor and functional recovery of the impaired arm. The Motricity Index shows a significant improvement in both groups; increasing from 18 (39.34%) to 26.3 (46.8%) in the control group, and from 19.34 (32.75%) to 31.36 (52.75%) in the study group. There is also a notable

### Table I

<table>
<thead>
<tr>
<th></th>
<th>Study Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (average)</td>
<td>47.9</td>
<td>51.9</td>
</tr>
<tr>
<td>Gender</td>
<td>14 female, 6 male</td>
<td>12 female, 10 male</td>
</tr>
<tr>
<td>Months post-stroke</td>
<td>24.4</td>
<td>25.7</td>
</tr>
<tr>
<td>Hemiparetic side</td>
<td>4 right, 16 left</td>
<td>3 right, 19 left</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Hemorrhagic stroke</td>
<td>8</td>
<td>12</td>
</tr>
</tbody>
</table>
An additional advantage is that the patients can use the Gesture Therapy system at home without the need of having an always-present therapist. We are also starting a new clinical trial that includes wrist-worn actigraphy measurements of the patient's activities at home, to have a more comprehensive evaluation of the effects of rehabilitation in the long term. With this we will capture objective measures, patient compliance, and reduce reliance on anecdotal reports of behavior profiles between clinic visits. An fMRI analysis will also be performed during this new study, to try to understand the biological bases for rehabilitation.

### TABLE II

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Study Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>Competence</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Effort</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Pressure</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Utility</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>Pain</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

IV. Conclusions and Future Work

We have developed Gesture Therapy, a low-cost vision-based system that allows individuals with stroke to practice arm movement exercises at home or at the clinic. The system integrates a virtual environment for facilitating repetitive movement training, with computer vision algorithms that track the hand of a patient, using an inexpensive camera, a gripper, and a personal computer. The results of a first clinical study shows similar impact in rehabilitation in terms of standard clinical scales compared to conventional therapy. However, according to a motivation survey and feedback from the therapists, a stronger motivation and attachment to the treatment is observed for the patients that used GT, which is considered an important factor for maximal recovery.

We are currently working on the automatic adaptation of the system based on the progress of the patient, so it can be used at home without the need of having an always-present therapist. We are also starting a new clinical trial that includes wrist-worn actigraphy measurements of the patient’s activities at home, to have a more comprehensive evaluation of the effects of rehabilitation in the long term. With this we will capture objective measures, patient compliance, and reduce reliance on anecdotal reports of behavior profiles between clinic visits. An fMRI analysis will also be performed during this new study, to try to understand the biological bases for rehabilitation.

REFERENCES