Real-Time 3D Avatars for Tele-rehabilitation in Virtual Reality

Gregorij KURILLO [a,1], Tomaz KORITNIK [b] Tadej BAJD [b] and Ruzena BAJCSY [a]
[a] *University of California, Berkeley, USA*
[b] *University of Ljubljana, Slovenia*

Abstract. We present work in progress on a tele-immersion system for tele-rehabilitation using real-time stereo vision and virtual environments. Stereo reconstruction is used to capture user's 3D avatar in real time and project it into a shared virtual environment, enabling a patient and therapist to interact remotely. Captured data can also be used to analyze the movement and provide feedback to the patient as we present in a preliminary study of stepping-in-place task. Such tele-presence system could in the future allow patients to interact remotely with remote physical therapist and virtual environment while objectively tracking their performance.

Keywords. Teleimmersion, Rehabilitation, Telerehabilitation, Lower extremities

Introduction

One of the major goals in rehabilitation is to make quantitative and qualitative improvements in daily activities in order to improve quality of independent living [15]. Rehabilitation process often includes task-oriented training and repetition of different motor activities involving impaired neuromuscular or musculoskeletal system [10]. In traditional rehabilitation approach, patient is guided by a trained physical therapist who observes and assists the patient to perform the tasks correctly. This process, however, is labor intensive, time consuming and often very subjective. Patient, on the other hand, often perceives the repetitive tasks as dull and non-engaging, which is consequently reducing patient's level of involvement in the rehabilitation. Several studies have shown importance of patient's psychological response, which greatly depends on the cognitive feedback associated with the performance of the tasks and affects success of the rehabilitation [6][7].

1. Related Work

Virtual reality (VR) as such can enrich the visual feedback associated with the performance of rehabilitation tasks [14]. In the past, many of the VR-based rehabilitation systems relied on custom-built devices to be used for input or feedback in virtual environ-

[1] Corresponding Author: Gregorij Kurillo, University of California, Berkeley, #736 Saturdja Dai Hall, CA 94720, Berkeley; E-mail: gregorij@eecs.berkeley.edu.

ments. Such devices, however, were usually associated with high costs, low reliability, low accessibility, and poor ergonomic design, making them unsuitable for clinical use. More recently, some of the commercial systems for gaming have been adopted for use in rehabilitation applications, such as Wii Remote and Wii Fit by Nintendo [3][11]. Several studies have also included live video-feedback of the patient to be integrated with the virtual environment to enhance patient's feeling of presence in the interactive space [3][8]. In these applications, the captured video was used to provide the visual feedback (as a virtual mirror) of the patient immersed in the environment. The video data was also used to interact with the graphics environment (e.g. popping bubbles with your hand) [3]. The technology behind it, however, does not allow for a full three-dimensional (3D) interaction in the virtual space as the users are only captured by a regular (2D) video camera.

2. Tele-immersion

In our framework (Fig. 1) we address some of the issues associated with creating the immersive feedback and data acquisition when using such video systems for rehabilitation and training of motor skills. We employ one or more stereo cameras to capture 3D avatar of the user/patient in real-time [16]. In contrast to traditional video cameras, the data produced preserves geometry of the movement with respect to the body and the environment allowing for accurate mapping of the movement into the virtual environment. The 3D video stream can be sent remotely or displayed locally while being seamlessly integrated with the virtual scene. Generated 3D mesh can be enhanced with dynamic texture to improve the visual quality of the video. In addition the 3D data captured by the cameras can be analyzed in real time to provide feedback on the screen while posing no restrictions on user's movements, such as in the case of motion capture systems with markers. This tele-immersive technology in connection with a virtual reality can provide a feeling of remote presence (i.e. tele-presence). A shared virtual environment can host several individuals from mutually distant locations and enable them to interact with each other in real time via a system of video cameras, audio communication, and fast network-based data transmission. Such approach can represent a basis for tele-rehabilitation practices that we are addressing in our current and future work.

In the past we have demonstrated benefits of immersive training in teaching of Tai-Chi movements [1], dance [14] and remote interaction. Similar tele-immersive technology has also been applied by one of our collaborators in the coaching of basketball players in wheelchairs [2]. Recently we have made significant advancements in the stereo algorithms to allow for real-time (25+ FPS) capture of 3D videos from one or more stereo cameras. Application of our technology in the area of telemedicine is in the early stage; however, we feel it is important to share the experience of this new emerging technology with the VR-based medical and rehabilitation community.

In this paper we present a pilot study on a group of healthy individuals using stepping-in-place (SIP) task which has a long history in evaluation of movement patterns in lower extremities [4]. In our work we apply 3D video in two ways, (a) to generate lifelike visual feedback of the remote therapist and patient as their reflection in a virtual mirror and (b) to measure the hip angles during task performance directly from the data without using markers. Finally, we outline our future work in the tele-rehabilitation and full body tracking using the tele-immersion technology and VR.

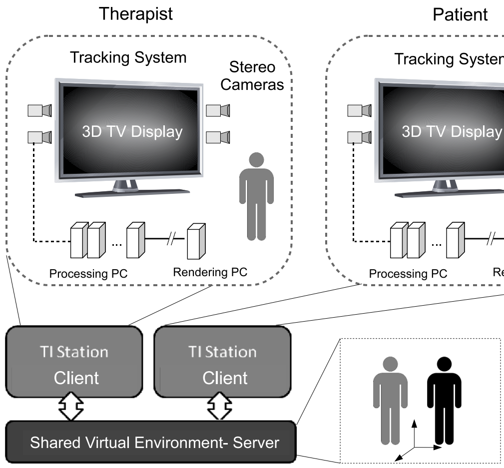

Figure 1. Diagram of the proposed setup for tele-immersive rehabilitation in system using a shared virtual environment and 3D avatars.

3. Methods & Materials

3.1. Real-time 3D Video

To generate patient's real-time 3D avatar for the VR rehabilitation task, a stereo camera is needed to capture two slightly displaced images. Our stereo algorithm with an adaptive meshing scheme [16] allows for fast and robust stereo matching on the triangle nodes. The stereo reconstruction can achieve the frame rate of more than 25 FPS on a pair of re-sampled images with 320x240 pixels or about 10 FPS on images with the resolution of 640x480 pixels. Result is a 3D mesh of the surface, which can be further mapped with a high resolution dynamic texture to achieve better visual fidelity. Further details on the algorithm can be found in [16].

The accuracy of the depth reconstruction depends on several factors related to the camera arrangement and typically ranges from 1 cm to 3 cm in our setup. Several stereo views can be combined by externally calibrating the cameras to a common coordinate system to increase user workspace or to generate 360-degree avatars. 3D video is then streamed through a gateway to the local renderer or remote location enabling telepresence by linking two or more remote sites. The captured 3D data is also used to perform body tracking and extract simplified kinematics which is used in the feedback loop for augmenting the video with virtual objects. Since user's body is captured by a calibrated stereo camera, the body movement is accurately mapped into the virtual environment and the geometry of the workspace is preserved. The generated data is also suitable for display on a 3D screen.

3.2. Stepping-in-Place (SIP) Task

The SIP task consists of guided performance of rhythmic movement of the lower extremities. It allows for the assessment of basic temporal parameters closely related to gait, such as stance and swing phase, double-stance phase, and step frequency. In the previous studies in connection with VR, the SIP task was considered as a modality of lower-extremities training for rehabilitation [10]. In this study, a virtual mirror was applied, displaying a generic avatar driven by a motion capture system to provide feedback during the training. The subjects were asked to track the steps performed by a virtual teacher. One of the concerns reported was the use of the same human figure model for all subjects as compared to using a personalized avatar as a form of feedback. Our tele-immersion framework enables exactly that by real-time 3D capture of a patient and a therapist. The therapist guiding the rehabilitation tasks can be located in the same room, at different geographical location or pre-recorded to replicate the same movements in every session. In our experiments we chose the latter to ensure a consistent reference motion pattern. In addition, the 3D data generated by the stereo system was also applied for the lower extremity kinematics extraction, thus allowing for markerless capture of hip angles.

As mentioned above, the 3D video of the 'therapist' was pre-recorded for our task to achieve consistency across subjects. We enabled visualization of two persons at the same time in the same virtual environment. The subjects observed themselves on the screen with the virtual 'therapist' rendered next to them (Fig. 2). They were instructed to track the 'therapist' avatar's movement as closely as possible. Two scenarios were tested: (1) 3D video only and (2) 3D video with overlaid virtual tracking targets to enhance visual feedback. In the first scenario, the left and right leg was shaded with two different color tones. The tracking targets in the second scenario marked the location of the knee joints for both persons. The targets were scaled and superimposed on the therapist's avatar.

3.3. Experimental Setup

The hardware setup consisted of one stereo camera Bumblebee2 (Point Grey, Inc.), with the resolution of 1024x768 pixels and the focal length of 3.8mm, was positioned above 65-inch LCD screen in front of the subject. The subject was standing upright at a marked position about 3m from the display and the camera. During the execution of the tasks, the subject was instructed to keep the arms close to his/her body to allow the algorithm to perform the segmentation based on the body symmetry along the sagittal plane. The stereo reconstruction was performed by a two dual core 2.33 GHz machine with the connected camera while the rendering and segmentation was performed on a dual quad core 2.00 GHz graphics server with GeForce GTX 8800.

Figure 2. Stereo reconstruction is used to capture user's 3D avatar in real time (left) and project it into a shared virtual environment (right), enabling a patient and therapist to interact remotely. Color shading of the legs assists with focus and orientation within the virtual mirror projection.

3.4. Trials

We have performed preliminary experiments on the group of 12 healthy individuals with the average age of 26.7 +/- 5.7 years (minimum age 20 years; maximum age 37 years). None of the subjects had a medical history of significant lower limb injuries or any other known medical condition that would impair movement. All subjects gave informed consent to participate in the experiment. Subjects performed the task in each of the two scenarios three times, starting with the 3D video only feedback. The reference recording required the subjects to exert hip angles of about 30° during the SIP task.

3.5. Data Analysis

To obtain the hip angles and knee joint positions from the stereo data, a simplified 3D video kinematics analysis was performed online on the renderer side (Fig. 4). The algorithm first segmented the data into left and right body half, assuming the symmetry of the body with respect to the camera coordinate system. Position of the lower part of the body was calculated from the ergonomics table while the segmented left and right leg was projected onto a plane aligned with the sagittal body plane. From the projection, hip and knee angles were calculated using a line fitting algorithm (using least squares method). Only the hip angle was used for the feedback and the analysis. The hip angle of 0° was defined in an upright standing position with the angle increasing to 90° as the leg was lifted from the floor. The accuracy of the method for the hip angle calculation was evaluated using a motion capture system and was within 10-degree error margin. The results of the measured hip angles were analyzed using correspondence algorithm presented in [5] and variance analysis of the spatial and temporal adaptation [9].

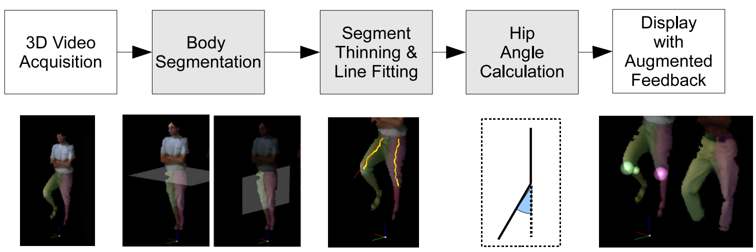

Figure 3. Block diagram of the segmentation and angle extraction process with intermediate results.

4. Results

The results of the experiments were analyzed for the two different scenarios by calculating the error of the aligned signals between the teacher and each subject. Fig. 4 (left) shows the output of the right and left hip angles as acquired in one of the subjects as compared to the reference recording. The shown output was captured for the 3D video only scenario. The subject closely followed the reference, with only small delays (200ms). The result in Fig. 4 (left) shows more precise tracking with the left leg as compared to the right one where larger deviations can be observed.

Fig. 4 (right) shows the distribution of the tracking results for the spatial and temporal adaptation in all subjects (n=12) with (1) and without (2) superimposed targets. The box diagram presents 5.0, 25.0, 50.0, 75.0, and 95.0 percentiles of the distribution. The spatial and the temporal adaptation was statistically different between the two scenarios ($p < 0.001$, ANOVA) suggesting that the augmented feedback (with superimposed tracking targets) helped subjects perform the task with greater accuracy.

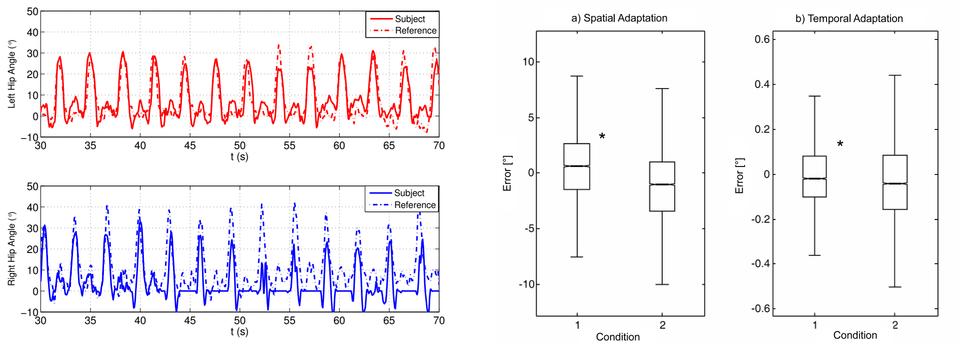

Figure 4. Sample output of the measured left (above) and right (below) hip angle trajectories during the trial. The angles were extracted in real time with a markerless method from the captured 3D video. Mean distribution (n=12) of the error for spatial and temporal adaptation during the two conditions: (1) 3D video with superimposed tracking targets and (2) 3D video only.

5. Discussion & Future Work

The focus of this exemplar study was to investigate the feasibility of the tele-presence system for use in tele-rehabilitation. The group of healthy individuals successfully performed the tracking task of stepping in place. We compared the tracking results of 3D video-only with 3D video enhanced by virtual targets, and concluded that better spatial and temporal adaptation was achieved when additional tracking targets were displayed.

Our ongoing work is directed towards development of tele-immersive framework for use in different application areas that would benefit from collaborative aspect of real-time 3D video, such as in tele-rehabilitation, medical evaluation, sports medicine, teaching of dancing and several areas of collaborative work. In many of the application, full-body segmentation and tracking is crucial for extraction of kinematic parameters which can be used to provide online feedback (such as in the presented example), perform gesture-based interaction or to drive computer-generated avatars. In this preliminary study, the extraction of the kinematics was simplified and closely related to the task. Our goal is to achieve more general human body segmentation and tracking from the 3D data in real-time (e.g. [12]) which could be applied to various tasks in VR-based (tele-) rehabilitation and applications of markerless motion capture. The performance of the current algorithms for full-body tracking is limited due to real-time constraints and sensitivity to outliers in the stereo data.

In this study and our past research work [1][14] we showed that current technology provides the user with feeling of tele-presence, suitable for remote teaching of body motion (such as in tele-rehabilitation), using relatively affordable equipment. The system also produces data which can be used online for feedback or offline for analysis which can quantify patient's performance during different motor activities. In this way, the patients could in the future participate in rehabilitation process from their homes or smaller medical office without the need to travel to large urban rehabilitation centers.

Acknowledgements

Research work on stepping in place task was supported by Slovenian Research Agency. Development of the Tele-immersion framework was partially supported by NSF (grants: 0703787, 0724681, 0937060), HP Labs, EADS and CITRIS at University of California, Berkeley.

References

[1] J.N. Bailenson, K. Patel, A. Nielsen, R. Bajcsy, S. Jung, G. Kurillo. The effect of interactivity on learning physical actions in virtual reality, *Media Psychology* **11** (2008), 354-376.

[2] P. Bajcsy, K. McHenry, H.J. Na, R. Malik, A. Spencer, S.K. Lee, R. Kooper, M. Frogley. Immersive environments for rehabilitation activities, *Proceedings of ACM International Conference on Multimedia*, Beijing, China, (2009).

[3] J.W. Burke, M.D.J. McNeill, D.K. Charles, P.J. Morrow, J.H. Crosbie, S.M. McDonough. Serious games for upper limb rehabilitation following stroke, *Proceedings of Games and Virtual Worlds for Serious Applications*, Coventry, (2009), 103-110.

[4] T. Fukuda. The stepping test: two phases of the labyrinthine reflex, *Acta Oto-Laryngol* **50** (1958), 95-108.

[5] M.A. Giese, T. Poggio. Synthesis and recognition of biological motion patterns based on linear superposition of prototypical motion sequences, *Proceedings of the 1999 IEEE Workshop on multi-view modeling and analysis of visual scene*, Fort Collins, CO, USA, (1999), 73-80.

[6] R.L. Hewer, Rehabilitation after stroke, *Neurological Rehabilitation*, Blackwell Scientific Publications, Inc., Oxford, UK, (1994), 157-166.

[7] M.K. Holden, T. Dyar. Virtual environment training: a new tool for neurorehabilitation, *Neurology Report* **26** (2002), 62-71.

[8] R. Kizony, N. Katzand, P.L. Weiss. Adapting an immersive virtual reality system for rehabilitation, *J. Visual. Com. Anim.* **14** (2003), 261-268.

[9] T. Koritnik, T. Bajd, M. Munih. Virtual environment for lower-extremities training, *Gait & Posture* **27** (2008), 323-330.

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

[11] R.S. Leder, G. Azcarate, R. Savage, S. Savage, L.E. Sucar, et al. Nintendo Wii Remote for computer simulated arm and wrist therapy in stroke survivors with upper extremity hemipariesis, *Proceedings of Virtual Rehabilitation*, Vancouver, BC, (2008), p. 74.

[12] J.M. Lien, G. Kurillo, R. Bajcsy. Multi-camera tele-immersion system with real-time model driven data compression, *The Visual Computer* **26**, (2010), 3-15.

[13] J. Moline. Virtual reality for health care: a survey, *Virtual Reality in Neuro-Psycho-Physiology*, IOS Press, Amsterdam, Netherlands, (1998).

[14] K. Nahrstedt, R. Bajcsy, L. Wymore, R. Sheppard, K. Mezur. Computation model of human creativity in dance choreography, *Proceedings of Association for the Advancement of Artificial Intelligence (AAAI) Spring Symposia*, (2008).

[15] H. Sveistrup. Motor Rehabilitation using virtual reality, *J. NeuroEngineering & Rehabilitation* **10**, (2004), 1-10.

[16] R. Vasudevan, Z. Zhou, G. Kurillo, E. Lobaton, R. Bajcsy, K. Nahrstedt. Real-time stereo-vision system for 3d tele-immersive collaboration, *Proceedings of IEEE International Conference on Multimedia & Expo*, Singapore, (2010).