Multimodal Archiving, Real-Time Annotation and Information Visualization in a Biofeedback System for Stroke Patient Rehabilitation

Weiwei Xu Yinpeng Chen Hari Sundaram Thanassis Rikakis

Arts Media and Engineering Program, Arizona State University

e-mail: {wwxu, yinpeng.chen , hari.sundaram, thanassis.rikakis}@asu.edu

ABSTRACT

In this paper we present our work on a system to support real-time multimodal archiving, collaborative annotation and offline information visualization for a biofeedback stroke-rehabilitation application. Our archiving / annotation / visualization system can play a critical role in the long-term biofeedback stroke therapy by supporting cooperative data analysis and media feedback as well as by providing the therapist with insight into computing-supported therapy. There are three contributions of this paper: (a) the design of a robust archiving system that archives in real time parametric model data (motion capture, motion analysis and audio / visual synthesis parameters) as well as audio / video from the biofeedback environment. (b) a web-based annotation tool designed with low cognitive load (c) a hierarchical information visualization tool that enables the therapist and other team members to examine quantitative motion analysis of subject performance with the context of media feedback, thus enabling collaborative insights. Our user studies indicate that the system performs well.

Categories and Subject Descriptors

J.3 [**Life and Medical Sciences**]; Medical information systems H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *auditory (non-speech) feedback, screen design, interaction styles*; H.1.2 **[User/Machine Systems]:** *Human factors*, *Human information processing*.

General Terms

Design, Experimentation, Human Factors

Keywords

Biofeedback, Analysis, Archiving, Annotation, Information Visualization

1. INTRODUCTION

This paper deals with the problem of real-time continuous archiving of multi-modal datasets and multi-domain collaborative annotations, and post therapeutic visualization of the archived data for a hybrid biofeedback stroke-rehabilitation system. We are developing a biofeedback system [3] that integrates task dependent physical therapy and cognitive stimuli within an interactive, multimodal environment. The system decreased rehabilitation duration and promotes more extensive recovery. We are currently focusing on assisting stroke-patients with reaching and grasping movements.

The problem is important for several reasons – (a) every 45 seconds, someone in the United States suffers a stroke [12], leading to serious functional deficits of neuropsychological and physical functions in stroke survivors. (b) The rehabilitation process can last as long as ten years. A system that archives the large amounts of multimodal data generated during therapy (over long time periods), analyzes annotations and creates insightful information visualizations can have a significant impact in stroke therapy. Some of the key challenges for us include – (a) archiving of large amounts of data in real-time (b) customizable annotation interfaces with low-cognitive load and (c) novel information visualization tools. We believe that continuous archiving, collaborative annotation and visualization of long-tem therapeutic data have not attracted as much research interest as consumer data (e.g. photos).

1.1 Related Work

Traditional stroke-rehabilitation therapy is monotonous and does not incorporate detailed quantitative assessment of recovery. Additionally, clinical history is kept on paper records, and written by different domain experts, who do not have access to each others' observations. Virtual reality (VR) is an emerging and promising technology for task-oriented biofeedback therapy [7]. It can offer complex, highly detailed and engaging multimodal feedback in response to physical action. This has significant potential in augmenting traditional task-oriented therapy training. There has been related work on virtual reality based techniques [8,15] to perform the biofeedback intervention for functional task retraining.

There has been prior work in creating collaborative annotation systems [4,14]. In [14], the authors explore a collaborative annotation system for mobile devices. They allow users to annotate digital photos at the time of capture. There they used appearance based recommendations as well as location context to suggest annotations to mobile users. In [4], the authors describe a collaborative annotation procedure for scientific visualization tasks, that can be done remotely. The users of their system annotate the same media element. Their system supports digital ink, voice and text annotations. However, there is no notion of user-context and the system does not exploit the existence of annotations made by other users.

CARPE'06, October 28, 2006, Santa Barbara, California, USA.

There has been significant work in the CSCW / CHI communities on distributed collaborations either in real-time [9,10], or over documents [2,13]. In [10], the authors examine the gaze patterns of helpers in a collaborative task, and in [9], the authors examine the effect of having a shared visual space on a collaborative task - the space grounds the communications between the participants with common awareness of the state of the task , promoting efficient communication. In [2] the authors examine the use of annotations in a shared documents and in [13] the authors design an activity oriented model to mimic the rich functionality found in physical annotations. However in the last two cases, the work has focused on document collaboration. In this paper, we focus on real-time annotation of a collaborative experiment, and the participants do not work on a shared document.

There has also been recent interest in visualizing / exploring the personalized history of each user [5,11,12]. In [11], the authors develop visualization for exploring the personal medical history of a single person. The *MyLifeBits* [5] system is designed to store a lifetime of data from a user. This includes data from digitizing paper records such as bills, letters, photographs, as well as digital capture of email, web pages visited, radio and television. The *Haystack* system [12] offers the users a framework for an integrated approach to the management of e-mails / instant messaging etc.

1.2 Our Approach

Our approach to the problem has three parts. First, we develop a synchronized continuous archiving subsystem to manage and store multiple data streams (motion capture, motion analysis, audio / visual feedback, speech from the patient and team members, video of the rehabilitation process as well as real-time annotations.). The archiving is distributed over two servers – one to capture motion analysis and audio / visual synthesis parameters, and annotations, while the other stores the actual audio / visual feedback (and audio / video from the environment) in real-time.

Second, we develop a distributed, low-cognitive load annotation tool to support tagging of events. This is customized per domain expert (therapist, engineer, musician, visual artist and archivist), and allows query / navigation to the events at multiple time-scales, in a non-linear manner. One's annotations are multicast to other annotation clients, thus providing each domain expert with others' insights in real-time.

Finally, we developed an offline visualization of our large multimodal dataset aligned with the annotations by presenting a prototype that supports hierarchical navigation through the patient's rehabilitation history with motion analysis evaluation metrics or parameter values, and synchronized playback of contextual information (audio / visual feedback, speech audio and video of the patient during experimentation). Real-time collected annotations are editable on the visualization interface. We have conducted preliminary clinical trials with stroke patients. Our user studies indicate that the system was well liked.

The rest of the paper is organized as follows. In the next section, we provide a brief overview of our biofeedback system. We shall discuss continuous archiving in section 3, present our real-time collaborative annotation tool in section 4 and our offline visualization of the archived dataset in section 5. In section 6 we discuss our early stage real-time annotation experiments. Following the experimental user studies, we propose future work in section 7 and then present our conclusions.

2. THE BIOFEEDBACK SYSTEM

We now briefly review our real-time biofeedback system [3], as it its critical to understand the proposed archiving and annotation and visualization system design. In [3], the focus was only on motion analysis, and audio / visual feedback creation, not archiving, annotation and visualization.

2.1 System Design

Traditional physical stroke-rehabilitation is monotonous and we propose an integrated virtual reality therapy which is more interesting and effective. We envision it to be developed into a stand-alone home rehabilitation system. The biofeedback system integrates five computational subsystems: (a) Motion capture; (b) Motion analysis; (c) Audio feedback; (d) Visual feedback; and (5) Database for archiving and annotation. The current overall goal of the system is to enable therapy with respect to a functional task of reaching.

All five subsystems are synchronized with respect to a universal time clock. Figure 1 shows the system diagram. The motion capture subsystem uses six calibrated and infrared cameras with up to 100 frames per second to track the three-dimensional position of reflective markers that are placed on the subject. The real-time motion analysis subsystem smoothes the raw sensing data, and derives an expanded set of task specific quantitative features. It multicasts the analyzed data to the audio, visual and archiving subsystems at the same frame rate. The audio and visual subsystems adapt their auditory and visual response dynamically to selected motion features under different feedback environments. The audio-visual feedback subsystems respond to the subject movement in real-time (1/100th sec.). The archiving subsystem continuously stores the motion capture, motion analysis, audio and video feedback parameters with a hierarchical indexing scheme and universal timestamps. It also records audio and video media streams. It provides a synchronized, personalized and distributed web annotation interface in real time and an offline presentation interface as well.

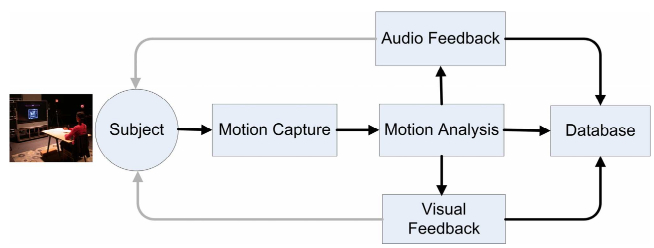

Figure 1: The biofeedback system showing different functional components of motion analysis, audio-visual feedback and archiving. Each component needs to operate in real-time (1/100th sec).

In the biofeedback system, we distribute four tasks (*motion analysis, audio feedback creation, video feedback creation, and archiving*) across four computers. These four nodes are connected by standard public ASU Ethernet. The motion analysis node is responsible for all the patient movement analysis. This analysis is then multicast to the other three nodes. We decided to use multicast, as it offers us the possibility of adding other analysis / feedback nodes, without affecting the motion analysis node.

2.2 Coupling Action to Feedback

We use 12 labeled 3D markers to represent arm and torso. The label specifies the location of each marker. Using the 3D positions of these 12 markers, we compute hand trajectory, shoulder trajectory, hand-target distance and eleven joint angles. For each feature, we also compute both the first and second time derivatives of features (velocity and acceleration). Based on domain knowledge of the reaching and grasping arm action, we group these features into six levels, with level 1 being the most important. With these motion features, we can derive three sub-goals of reaching and grasping task: (a) *reaching* the target, (b) *opening* of the arm joints and (c) *flow* of the movement.

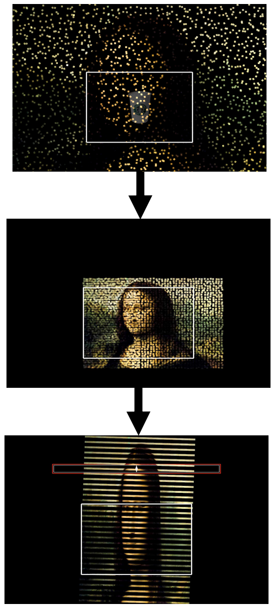

Figure 2: Visual feedback in the abstract environments. Top: particles begin to the form the image as the hand approaches the target. Middle: image pulled to the right when subject is off target. Bottom: vertical bands appear when the subject has wrong target height.

The structure of the feedback environment and its relationship to the achievement of the goals are based on well established principles regarding the role and function of art . At the highest level of its structure the environment must communicate to the patient the messages that can encourage the accomplishment of the movement goals. These messages are: *reach*, *open*, *flow*.

The feedback images used are all well known paintings and the music played is based on well established rules of western classical music. Thus the content has a high probability of attracting and engaging the subject and deepening their immersion in the experience.

The overall idea driving mappings from motion analysis to media feedback is that spatial and target information is better communicated through visuals, while temporal information is better communicated through audios. The calculated movement parameters that drive successful manipulation of the biofeedback environment are key parameters of an everyday reaching and grasping movement. Thus, the entire environment can be easily connected in a analysis-feedback loop, and in terms of action to its goal and does not require unintuitive movement learning, which is a natural artifact of the interaction.

The mapped media contents follow a similar structural hierarchy as the movement parameters and goals with sub-message levels supporting the communication of each larger message. As is the case of movement parameters, there are feedback parameters that the subject can quickly understand and control, parameters that require practice to control and subconscious parameters supporting the achievement of consciously controlled goals [14].

2.3 Validation

We use five validation metrics to evaluate the patient's performance with respect to three sub-goals (*reaching*, *opening* and *flow*). These five validation metrics are: (a) reaching spatial error, (b) reaching duration (c) arm openness, (d) flow error and (e) consistency. For *reaching* sub-goal evaluation, we use reaching spatial error and reaching duration. We evaluate *opening* sub-goal by arm openness and *flow* sub-goal by flow error. Consistency is used to evaluate how consistent of the patient's *flow* is in several consecutive trials. The computation detail is presented in [3] These validation metrics will be calculated offline and visualized in the offline presentation framework. Through validation visualization, we can see the biofeedback rehabilitation progress of stoke patient. We have reported validation results for the biofeedback system based on normal subjects' experiments in [2]. This paper is based on our subsequent preliminary clinical experiments with stroke patients.

2.4 The Rehabilitation Process

We now provide an overview of the rehabilitation process to show that the process is detailed, and requires significant attention from all members of the biofeedback team. The whole rehabilitation process includes two steps: calibration and real reaching process. In the calibration step, we need to calibrate six cameras and the patient's reaching and rest positions. This helps in determining the patients range of motion as well as in building the motion analysis model.

We start with the rehabilitation process after calibration. As noted earlier, the environment assists in the rehabilitation task of reaching. The process comprises five environments, which the patient (with the therapist's help) interacts with in sequence: (a) physical cup environment, (b) transitional environment, (c) abstract environment I, (d) abstract environment II, and (e) post physical cup environment. Before each environment, patient will be given clear instructions about the environment. In each environment, the patent is asked to do ten reaching trials. In (a) and (e), the patient reach a physical cup without any feedback. In (b), we only provide visual feedback through 3D virtual environment shown on the screen. The patient's movement controls the 3D virtual arm to reach a 3D virtual cup. In (c) and (d), both audio and visual feedbacks are provided. Although the virtual cup is not shown, feedback provides the patient with clear audio and visual cue to reach the virtual target. Each trial in the transitional environment and the two abstract environments includes six sequential states: *ready check*, *reaction*, *reaching*, *grasping*, *returning*, and *stop*. The motion analysis subsystem calculates the state automatically based on the patient's movement. This affects the audio / visual feedback and gives patient clear cues about the environment state during the trial.

While the patient is performing the functional task (i.e. during each trial), this requires careful attention from all members of the biofeedback team. There are several reasons – first, trial duration (while variable across patients) is usually short – just 10 sec. *Second* since different patients have different medical histories, arm control capability and audio/visual preference, the calculation of trial states and feedback creation need to be fine-tuned to each patient. Some of these adaptations have to be done manually – grasping condition parameters from the motion analysis subsystem, subtle alterations to the sound feedback, and image rotation sensitivity in visual feedback subsystem. Therefore, all team members need to be highly focused on the patient's movement and different parameter adaptations. These experimentation requirements add up to a high cognitive load to all team members. We additionally note that for each patient, the therapy session lasts about two hours.

In the rest of the paper, we shall employ the following notation. A visit by the patient will be called a *session.* Within a session, the rehabilitation process is grouped into *sets,* where a set represents rehabilitation by a patient with consistent physical and media feedback environments. A set comprises *trials,* where the patient performs the functional task (in our experiments the trial involves the functional task of reaching for a cup).

3. CONTINUOUS ARCHIVING

Continuous archiving of biofeedback data involves storing in real-time a diverse collection of data types. We plan to store motion capture and analysis parameters, and audio / visual synthesis model parameters. We also plan to store the actual audio / visual feedback media generated, audio data that capture the conversations between the patient and the therapist, as well as conversations amongst the biofeedback team members, and the patient video.

It is important to store the actual audio / visual feedback, in addition to the audio-visual synthesis models. This is because even though the audio / visual synthesis models are complete, storing the actual audio / visual data precludes the need to re-synthesize them in real-time, when it is necessary to review the audio / visual display shown to a patient at a particular trial. This decision leads to significant computational savings, at the expense of increased storage. The magnitude of information to be stored is a challenge, but this high data rate is important to capture the details of the interaction. Getting the details right is important to facilitate insight for the therapist (as well as the biofeedback team).

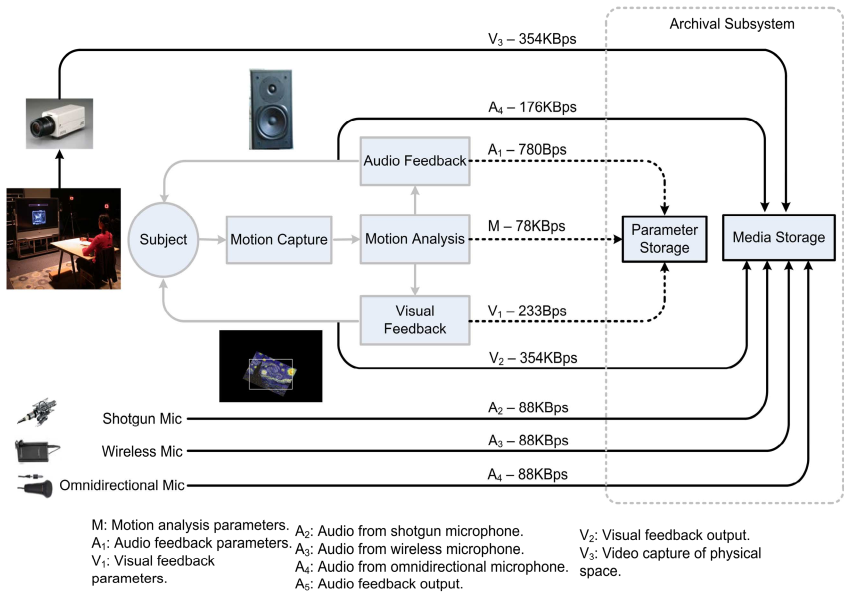

Figure 3: Multimodal dataflow and data transport rates within the biofeedback system

In this section we will cover issues related to archiving. We split archiving into two subsystems – one that archives parametric models and raw motion capture data, and the second which stores the audio-visual feedback data. As is shown by Figure 3 two groups of multimodal data converge onto separate archiving servers which distribute the computational overload of the entire subsystem. In section 3.1, we discuss continuous archiving of parameterized motion and feedback data. In section 3.2 we discuss our further efforts on capture and recording of biofeedback data of other modalities. Finally in section 3.3, key database design and implementation details are presented.

3.1 Archiving of Parametric System Models

In this section we focus on the transport of model parameters. We use one dedicated desktop machine to facilitate the archiving of the parametric data. Transportation of motion capture, motion analysis, audio feedback and video feedback parameters share the common multicast framework that we presented in section 2.1. The motion analysis subsystem packages and sends out data at 78KBps while the data rate of media feedback subsystems is at 1.01KBps. The reason for this difference is as follows. The motion analysis subsystem sends out *both* the raw motion capture data, as well as the derived motion features. This needs to be done per frame (the motion capture system sends out data at 100 frames per second). The audio / visual synthesis models do not change *within* a trial. Within each trial, the synthesis models are fixed and the production of the audio and visual feedback environments is deterministically coupled to the motion analysis features. Changes are only made between trials. Hence the synthesis models need to be updated (and stored) per trial. The models are multicast, thus taking advantage of the prior multicast code.

Motion analysis parameters are organized into 15 categories based upon bioengineering theories and our firsthand experimentation with normal subjects and patients. There are three criteria that guide the categorization of motion parameters. First, parameters dealing with geometrical relationships are grouped together (either spatial positions or joint angles). Corresponding velocity and acceleration values of all parameters are also stored. Second parameters are assigned to different levels according to their role of importance towards the biofeedback rehabilitation goals. This ordering is provided by a bioengineering domain expert. Third, some special purpose parameters are grouped independently. For example, a group of synergy table parameters are computed and stored at the initial patient calibration stage. The synergy table provides baseline spatial and angular relationships that are used for media feedback during the trials.

Audio and video feedback synthesis model parameters are manually or automatically initialized once at the beginning of each trial. The synthesis models do not change during the trial. As long as we store the algorithm and the initial parameter values, we will be able to simulate real-time media feedback offline with no loss of information.

3.2 Media Capture and Recording

In this section we discuss our approach to media capture and recording. We make use of a duel-core desktop server for fast and high quality archiving of multimodal media streams.

We record five audio channels. Three microphones are set up around our experimentation environment and connected to the server via a FireWire audio interface. The first one is a professional shotgun microphone directed towards our patient to record the patient's speech; the second microphone is a wireless microphone with the therapist who assists our patients of the experimentation; and the last one is a general-purpose omnidirectional microphone aimed to pick up conversations amongst the biofeedback researchers. The final two audio streams recorded are from the audio feedback module. Each stream is recoded at 44.1 kHz, 16 bits per sample.

We record two video channels. The first channel is the output of the video feedback module. The output is scan converted and sent to the archiving server. The other channel comes from a wide-angle video camera view of our experimentation area, directed at the patient's arm movement and additionally includes in its field of view, the video feedback shown on the monitor that the subject watches. Both video streams are then digitized using an mpeg hardware encoder.

We developed a media capture and recording tool on the Max/MSP platform [1] to archive all the streams. Max/MSP is a graphical programming environment that enables rapid access to multimedia devices, and processing and creation of multimedia content. We combine two separate audio streams from our soundboard into one single stream and two video streams into one using a Max/Jitter video grab/recording object at 30fps, 189Kbits per frame and image sizes of 720x480.

We have discussed the front-end component including parameterized model data streaming and organization, and media data capture and recording. We now introduce the back-end component of continuous real-time archiving.

3.3 Database Engine

We now present our design of the biofeedback for Rehabilitation SQL Server database (BfRDB), and associated technical challenges.

3.3.1 Design of BfRDB

We present our database scheme in this section. We store detailed patient info related to biofeedback rehabilitation into a unique table as a reference of patients' background and to evaluate their rehabilitation history with our biofeedback rehabilitation process.

We have a unique set of tables per patient. The idea is to access, store and retrieve the patient's personal rehabilitation history quickly, while maintaining privacy. We acknowledge that more sophisticated privacy schemes may be required as the scope of the project increases. Because each patient's dataset is very large, we speed up our database query and retrieval operations by separating them from the dataset associated with other patients. Within the set of tables belonging to one patient, a unique trial information table stores the trial indexes. A session id indicates an experimentation procedure from system calibration to the end of experimentation. A set id indicates a group of sequential trials with consistent physical and feedback environment setting. A trial id indicates trials within the set. A frame number indicates time instance within a trial and the motion data frame rate is at 100fps. Each trial has a task id that indicates the rehabilitation functional task type and an environment id corresponding to the feedback environment setting type. 15 categories of motion capture plus motion analysis parameters are stored into relevant tables.

We differentiate audio and visual feedback model parameters according to their semantics. For example, the audio parameters are organized as belonging to the general sound player object, the foreground player object and one of the eight background player objects. Therefore we have 10 tables for the audio feedback model. The set of visual parameters are changeable according to current feedback environment setting. Therefore we have a general environment setting table and other two tables for different visual feedback environment (transitional and abstract). The database also refers to multimodal media data capture and recording files on the other archiving server by a separate table.

3.3.2 Technical Issues

Continuous archiving of large datasets raises challenges in both temporal synchronization and high speed storage at 1.2MBps. We have distributed our overall archiving computational and storage overload into two desktop servers. While this distributes the computational load, this introduces inter-server synchronization issue. We address the time synchronization issue in three stages. In order to align captured media with model parameters on the time axis, we first synchronize the servers to make sure they have exactly the same system clock. Second the time-sensitive information of trials from separate models is stored with synchronized hierarchical indexing (session id/set id/trial id/frame number) as well as a universal system timestamp. Third we use the Max/MSP archiving tool to trace the timestamps of media capture and recording on the other server. With temporally aligned multimodal data, we are able to access, store, query, retrieve, simulate and present real-time biofeedback loops offline.

As is required by real-time and continuous archiving of large datasets, we developed reliable multicast network communication without any data loss. The archiving subsystem's front-end component streams in model parameters from the multicast framework, holds parameter values of several trials in a multi-buffered layer in memory in order to save the effect of database operation latency and response in real time to incoming data, and stores them into the back-end database with a bulk insert operation. With the bulk insert operation we are able to raise our database write rate from 20KBps to 1.5MBps. For the archiving of captured video feedback, we use a hardware encoder to compress data, and then store them using our programmed Max/MSP tool. Since the archiving of model parameters and the media occur on separate machines, there are no write conflicts.

4. REAL-TIME ANNOTATION

In this section, we shall discuss real-time annotation issues. Real-time annotation plays an important role in our biofeedback system. This is because collaborative annotations from interdisciplinary annotators (therapist, engineer, musician, visual artist and archivist) give us the firsthand and time-critical expert decisions in the biofeedback rehabilitation process. And we believe their annotations will explore the semantic space of biofeedback from different perspectives, and finally help build a real-time adaptive and automatic biofeedback system. Hence we developed a real-time web annotation tool for collecting collaborative annotations.

The cognitive load on the entire biofeedback team can be high (ref. Section 2.4). Since this is a hybrid rehabilitation system, it is important that every aspect of the system (physical set-up, visual, audio, therapist and patient, as well as the archiving) be working as needed. Each member usually makes changes to their module, per patient, as the specific medical history (specific muscles may be weak, patients cognitive capability may be impaired due to the stroke, and hence careful attention needs to be paid to the audio-visual feedback), as well as physical issues (the physical dimensions of the patient may be different) requires a different calibration process. These unpredictable variations across patients imply that things in the system may need to be adapted on the fly. It is imperative that these system level changes be documented and annotated. However, sometimes the process of making the changes can be complex, and since each trial is short (~10 sec.) there is a possibility that the important change will not be annotated. Hence an interface with low-cognitive load is important.

4.1 The Real-Time Web Annotation Tool

Our real-time web annotation tool is distributed, personalized and designed to facilitate trial annotation in real time. We decided to develop web based annotation tool, as this has the advantage of scaling to remote therapist / patient interaction. It is programmed as a Java applet and can be loaded from our Apache web server into any Java-enabled web browser. The web server is co-located with our continuous archiving server that hosts the SQL server database. Each biofeedback team member (therapist, engineer, musician, visual artist and archivist) is responsible for their own annotation applet customized to their discipline. They annotate on the applet while controlling their subsystem during experimentation. The annotation tool

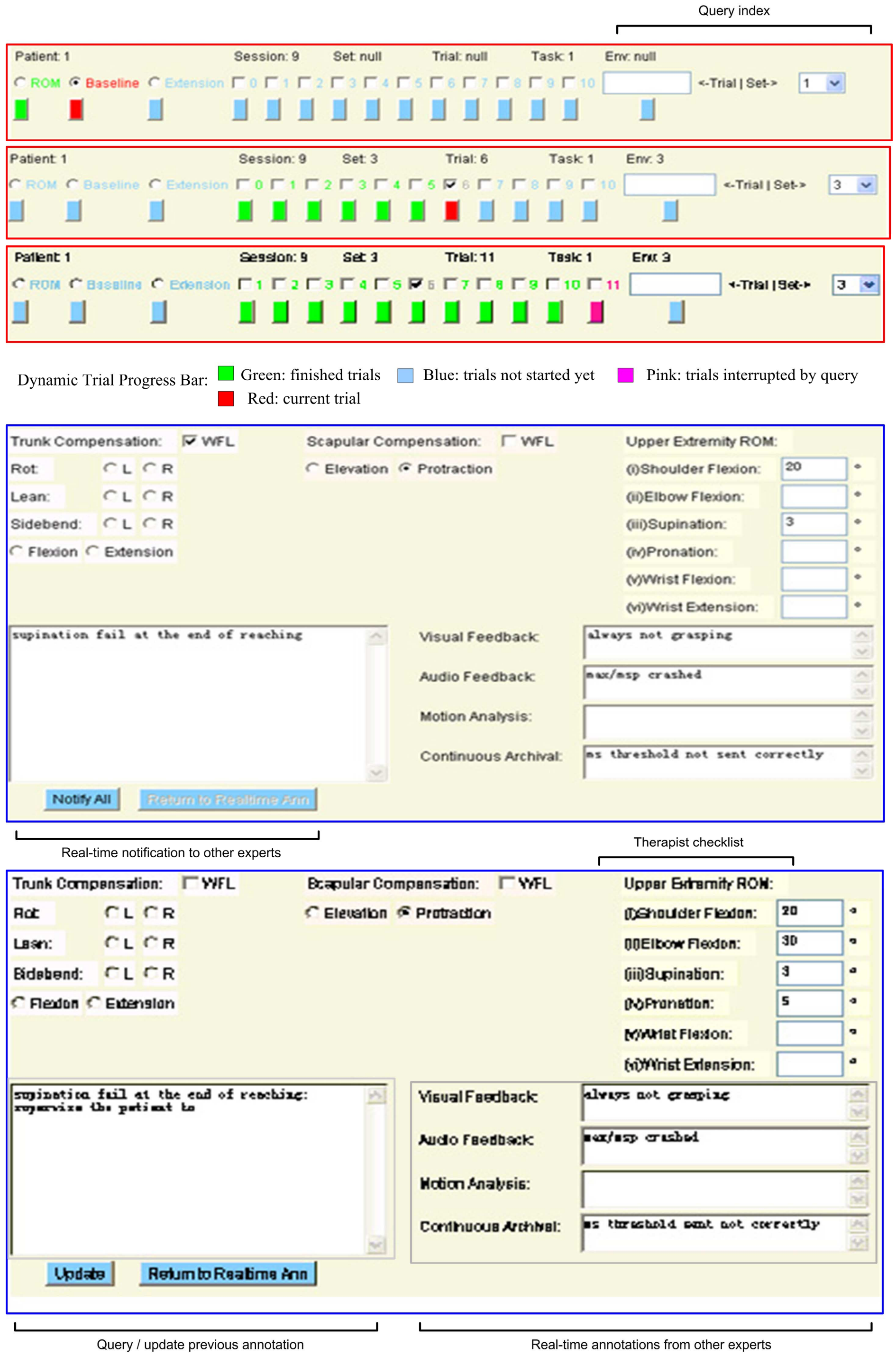

Figure 4: Top and Center/Bottom Panel Of Our Web Annotation Tool. During real-time annotation, annotators keep informed of current trial status with dynamic visual cues from the annotation sever; of other members' annotation from distributed annotation clients; and input their own annotations. When in query of any past trial, they are able to update previous annotation and review others' annotation as well. They can return to real-time annotation after intended update.

includes four main functions: (a) hierarchical indexing, (b) domain specific checklist, (c) collaborative annotation and (d) post modification.

4.1.1 Hierarchical annotation index

We propose a hierarchical annotation index that allows the annotator to modify at any time, any trial, in any previous set during the same session. Annotators can also go back to previous trials to check/modify annotation through selecting proper patient/session/set/environment/trial ID. Therefore the annotation interface changes according to different experimentation and annotation stages: (a) pre-experimentation calibration (b) real-time experimentation / annotation (c) post annotation query and modification.

The indexing panel changes dynamically to indicate the progression during the trial. Annotators can easily get patient/session/set/environment/trial ID and current trial state (progressing or finished). As shown in Figure 4, a colored progress bar with a fixed number of trials is presented. In our case the last 11 trials are always shown on the progress bar. Annotators can locate the current trial with red color, group the past trials with green colors and ignore the future trials with blue colors. If no red color is present, our system is in ready-check or stop status of trials (ref. section 2.4). If the real-time annotation process is interrupted by an annotator's intended query of past trial annotation, the interrupted trial display element will be colored pink.

We also have indicators on the bar to inform annotators of the calibration status at the calibration stage before real-time experimentation. This period is important as we get to understand better the specific constraints of the patients (e.g. physical dimensions, cognitive deficits), leading to specific decisions on how to adapt the system before the set begins. These are valuable observations and changes to the system at that stage, and therefore must be annotated and archived.

4.1.2 Domain Specific Checklist

The annotation space for the experiment is large – each domain expert participating in the project is looking out for very different set of events. Hence it is not reasonable for us to design a single annotation interface that contains all the possible checkboxes for the entire team. This would create a higher cognitive load for each team member, as they would need to spend unnecessary time to locate the item of interest. Hence our annotation tool is adapted to annotators from different domain perspectives (the engineer, the musician, the visual artist, the therapist and the archivist). In order to speedup annotation, we provide each team member a domain specific checklist (shown in the center panel of Figure 4). The checklist comprises their personalized frequently used annotation options. The checklists are made through consulting all team members based on their domain experiences. Note that each team member has a dedicated annotation workstation adjacent to where they sit, to input the annotation. For each item in checklist, we also provide keyboard shortcuts to facilitate annotation. Figure 4 also shows the detail checklist for therapist annotation.

4.1.3 Collaborative Annotation

It is very important for team members to share annotations with each other in real-time. This has two effects – (a) domain experts are alerted to events of interest to other domain experts. Since the underlying phenomena (rehabilitation of the subject) are common, it is possible that the events of interest to one domain may be of semantic relevance to other experts. (b) the nature of novel inter-domain relations can provide rich insight into the biofeedback problem, leading to further annotations.

Therefore, we incorporate the idea of collaborative annotation into our annotation tool (shown in the bottom panel of Figure 4). Each domain expert sees five text boxes – one corresponding to each domain. However, only the domain expert's own domain textbox is editable. The other four text boxes can contain annotations from other domain experts during the same trial (ref. Figure 4). We accomplish this using a Java multicast framework where each individual web annotation tool listens to not only a server multicast in order to update its trial indexing but also listens to four client multicasts from other Java applets in order to get others' annotation in real time. Therefore our annotators as a team are always aware of other members' input and collaboratively annotate on synchronized observations.

4.1.4 Post-Set Modification

The biofeedback rehabilitation process is time and effort consuming and all annotators put their major focus on their own subsystem or therapy process, their cognitive overload sometimes prevents them from timely annotation. Hence, we provide random access to past trials at any time. Annotators can easily go back to specific past trials for annotation modification during the experimentation break between sets by clicking on the button aligned with their intended trial or typing the trial index. They also can select an arbitrary combination of set and trial to perform any intended annotation modification. They can input any shorthand annotation in real time and refine them later. In the next section we present our key ideas in information visualization.

5. INFORMATION VISUALIZATION

The continuous archiving of all the different modalities in the biofeedback system creates significant amounts of data. This data (motion capture, extracted features, audio-visual feedback, and annotations) must be visualized in a manner that allows for insight for all members of the biofeedback team. We expected the visualization tool to be used by biofeedback team members, especially the therapist, to review the validated data in our system overlaid with annotations after each experimentation session. By examining these data, she/he can (a) draw insightful conclusions about the patient's stroke-rehabilitation progress; (b) annotate the contextual information and possibly make integral recommendations to changes of the cooperative data analysis and media feedback subsystems (e.g. new validation metrics, customized media feedback etc.). In the next section we present the design goals of our information visualization system and in Section 5.2, we present the approach to realizing our information visualization design goals.

5.1 Our Design Goals

We now discuss our design goals of information visualization: (a) hierarchical and selectable navigation through our information space; (b) synchronized contextual information playback; (c) built-in annotation facilities.

5.1.1 Hierarchical and Selectable Navigation

Any member of the biofeedback team (e.g. the therapist) should be able to navigate the archived motion analysis, the annotations and the audio visual feedback at different time-scales. This is important in understanding the long term progress of the patient. We would like to remind the reader that a key advantage of the hybrid rehabilitation system lies in our ability to analyze parameter variations over smaller time-scales. Subtle cross-parameter correlations, and small improvements (e.g. between two sessions) are very difficult for a therapist to determine. She can more easily look for larger variations over much larger (e.g. a few weeks) time-scales, but the *exact* nature of the improvement is still not available.

5.1.2 Synchronized Playback of Context and Content

In order to provide insight to the viewer, the information visualization system must reveal the *temporal dynamics* of the motion features / validation metrics in conjunction with media feedback. This is important as important changes to our validation metrics may be tied to specific aspects of the media feedback, or perhaps the physical setup of the experiment. Note we video record the actual experiment in real time (ref. Section 3.2). We believe that an integrated presentation will reveal relationships not readily apparent in the actual trial.

5.1.3 Incorporating Annotations

We need to overlay the real-time annotation results on the visualization interface. This has two effects – (a) boosts usage of real-time annotation from domain experts (b) encourages offline update on the annotations, based on new observations of the data. The observed data must be time-aligned with all domain annotations along the different visualization levels. This integrated view allows the context to be inferred for a specific event (e.g. a specific change in the feedback may always be triggered by a specific change to the motion analysis model parameter.).

5.2 The Information Presentation Interface

This section presents the information visualization interface that attempts to achieve our design goals. We developed our offline information presentation interface on the Visual C# platform with the Apple QuickTime API for multimedia manipulation

5.2.1 Navigation Through The Hierarchy

The interface supports hierarchical navigation as shown in the left part of Figure 5. At the top level of the hierarchy the entire biofeedback history is shown (once a patient id has been selected). The motion analysis validation metrics are shown and are grouped by sessions. Trials within the session are marked with orange squares. Annotations from different domain experts appear as differently colored dots that are vertically aligned with the corresponding trial marker.

The mouse-over user interaction at any trial marker displays a tool tip showing in sequence, the trial index, evaluation value and all annotations. Meanwhile the session curve that is being currently inspected with the mouse will be highlighted in red to provide visual cues to distinguish it from the other sessions. By clicking on the highlighted session curve the reviewer can navigate into the session level visualization where the validation curves are grouped by sets within the session. Note that only one validation metric is displayed at a time, and other metrics are selectable using a drop-down list.

The reviewer can navigate similarly into other levels. At the first three levels, five biofeedback validation metrics can be selected by the reviewer. At the trial and frame level, archived motion capture and analysis parameters can be visualized on selection. Because real-time annotations are collected on a trial basis, they are presented by tooltip at the entire history, session and set

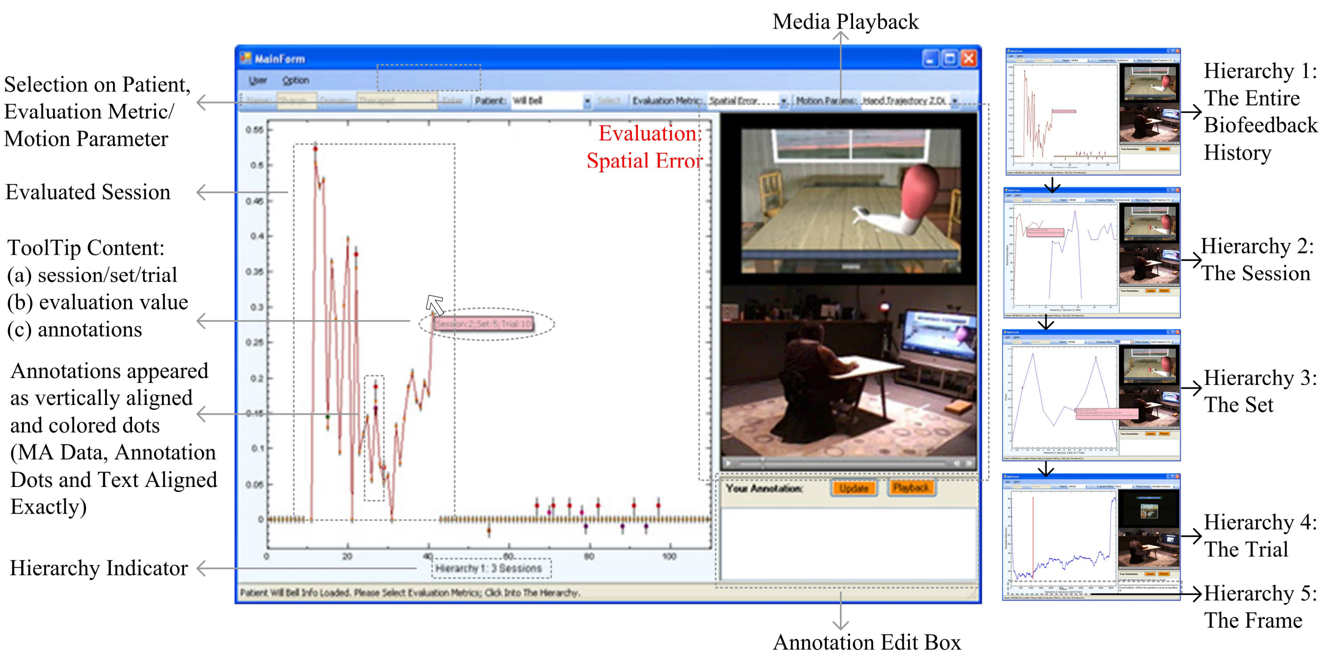

Figure 5 Information Presentation Interface. The left panel visualizes motion analysis model parameters and offline evaluation results in a hierarchical way. The top-right panel replays recorded video and audio in synchrony with parameters' presentation. The bottom-right panel shows annotations of the selected trial and allows offline annotation update. The presentation framework supports hierarchical navigation as well as direct access to intended trials.

levels, and presented on the right-bottom panel at the other two levels. Besides hierarchical navigation at multiple time-scales, the interface also supports direct access to any intended trial by right-clicking on the trial markers. A context menu with choices of direct playback (of media) and annotation update (to change the annotation) will be presented in that case.

5.2.2 Synchronized Model / Media Playback

After selection on their intended session, the corresponding session video and audio media will be loaded onto the presentation panel. Upon users' navigation into the trial level hierarchy and selection on any motion parameter, the left panel (showing the motion analysis validation data / motion analysis features depending upon the time-scale) and the right media panel will replay in synchrony the trial progression with the selected motion analysis model parameter, the audio and video feedback model in multimedia format. Synchronized contextual playback is achieved with the help of our archiving of the trial start and end time in the trial info table mentioned in section 3.3, and of the media capture and recording time in the media reference table. We are able to locate any trial in our video/audio file.

5.2.3 Annotation Facilities

Annotations are always aligned with the current selection on trial. They are presented in two formats. *First* they are visualized as dots colored by domain expert, and aligned with their corresponding trial marker on the curve. This is maintained when the user zooms in and out of the hierarchy. *Second* at the first three levels of our hierarchy (entire history, session and set), trials are represented as squares on the curve. Therefore annotations are shown in the tooltip when the reviewer positions his mouse over the square or dot marker on the curve. When navigation enters the trial/frame level, annotations are presented in the right-bottom panel. This is because our domain experts annotate at the level of each trial, not each frame. The user can update her annotation at any hierarchy. By right-clicking and select annotation, the trial pointed beneath the mouse will be re-annotated. And when annotations are presented in the right-bottom panel, the reviewer has access to all annotations up-to-date from different domains in that panel and can edit his annotation for the trial whose contextual information is played back in the other two panels.

6. ANNOTATION USER STUDIES

This section discusses our user study for the real-time annotation experimentation. We have recently begun clinical trials of our biofeedback system with stroke patients. We conduct two sessions each with two patients, and we now present initial observations of our annotation tool. As noted earlier, the actual rehabilitation process is stressful, as we are dealing with real patients, and correct treatment (we cannot keep the patients waiting, say during recalibration etc.) is of utmost ethical importance (the patients travel long distances to participate in our studies). Due to the high degree of variability in our patients, the entire system needs to be adjusted semi-automatically. Most of the adjustments happen prior to a set, and the team also makes periodic adjustments during a set. The annotators have to thus split their attention between their tasks (i.e. adjustments) and to annotate why they are making the adjustment. This is an issue with all the team members.

Three aspects of the annotation tool have received positive comments. First the colored visual cues of trial progression keep them updated on where they are in the process, in general, which trial and whether it is within trial or in between trials. Second random access to past trials help them go back to complete their partial annotations at break. They can annotate as concisely as possible in real-time and refine the annotations later for their annotation to be easily understood by others. Third they like the collaborative feature of our annotation tool, which keep them aware of other domain observations/opinions of the same trial, both in real-time and in post-set modification.

We have also collected valuable suggestions for our web annotation tool after the experiments. Some are interface design related. For example, our visual artist suggests that using different text font sizes on the webpage can help him locate indexing entries fast to save his annotation efforts. Our musician colleague asks for optional panel display rather than make the display fixed in our case. All interface suggestions touch the core issue of decreasing visual overload for the annotators, as they are required to watch on their subsystem workstation, annotation workstation as well as the patient during experimentation. Some suggestions are tool function related. The most discussed function is whether the tool should automatically archive the annotation and clear the webpage at the start of a new trial. We designed it to be automatic so that annotators have no need to worry about clicking any update button after each trial annotation. This can be an issue if they are actually in the middle of an annotation when their webpage gets cleared and ready for a new trial. While they can go back to continue the interrupted partial annotation, some annotators feel that it interrupts their thoughts on annotation as well. The workaround should be automatic detection of whether the textbox is still being edited when a new trial starts and if so prevention of webpage clearance. We will discuss another potential function-related improvement in the next section. Other suggestions are domain related. For example, the therapist requires annotation at a finer time-scale rather than on a trial basis. She annotates at different periods within the trial, specifically, the reaction, reaching, grasping and returning periods. Therefore rather than her inputting the exact period for each checklist entry, we can facilitate it by visualizing intra-trial statuses on the interface and providing a popup choice menu for each entry which allows specification of exact periods.

7. OPEN ISSUES / FUTURE WORK

Our initial experiments have raised interesting open issues that we plan to address in our future work.

- *Event Modeling:* An important issue that we plan to work on is the development of robust event model. We can then build event detectors to detect events automatically. Some events occur at the feature level (significant change in the variance of the hand trajectory), while others are more subjective (e.g. change in the timbre of the sound). We plan to develop a model in conjunction with domain experts (by helping us identify events and event inter-relationships), which allows us to detect certain low-level events, and allow users to annotate high level events. We plan to develop a dynamic Bayesian model to predict events.
- *Pre-emptive Annotation:* A key issue to reduce cognitive load is to be able to present to the domain expert interesting events that can be annotated. This could be presented as an event log to be analyzed after

each set. An advantage of this approach is that feature level events that may be missed by the human annotator (e.g. a significant change in the parameter value), can then be presented for her to annotate.

- *SenseCam Integration:* During our initial experiments with patients, we have had our therapist wear the SenseCam [6]. The goal was to be able to capture what she does / sees during the experiments. We believe that the integration of the SenseCam pictures into our information visualization panel will enrich the contextual information to be presented.

8. CONCLUSION

In this paper we have presented our design and implementation of (a) multimodal data continuous archiving, (b) real-time collaborative annotation, and (c) information visualization of the large dataset comprising (a) and (b) for a hybrid biofeedback rehabilitation system. The problem is important for long-term therapy (several years) of stroke patients.

Our approach to the problem had three parts. First we developed a synchronized continuous archiving subsystem to manage and store large multimodal dataset. Second we developed a distributed, low-cognitive load annotation tool to support tagging of events. This is customized per domain expert, and allows query / navigation/ modification on the events in a non-linear manner. Finally, we developed an offline visualization of our dataset by a prototype that supports hierarchical navigation through the rehabilitation history on motion analysis results, and synchronized playback of the contextual information.

We have conducted several preliminary trials with two stroke patients. User studies with members of the team on the efficacy of the annotation tool showed that the annotation / archiving system was well liked by all members. In particular, they liked multiple dynamic visual indicators of experiment progression, the personalization of their interface, the collaborative annotation facility and random access of their past annotations.

We are planning to extend this work along several directions – (a) development of a robust event model, (b) event detection and its use in pre-emptive annotation, and (c) integration of SenseCam in the information visualization framework.

9. ACKNOWLEDGEMENTS

We would like to acknowledge the help from the members of the biofeedback team – Todd Ingalls, Richard Wallis and Loren Olson and Sheng-Meg Liu for their help in realizing the archival and annotation subsystems.

10. REFERENCES

[1] *Max/MSP (media tool)*. http://www.cycling74.com.

[2] J. J. CADIZ, A. GUPTA and J. GRUDIN (2000). *Using Web annotations for asynchronous collaboration around documents*, Proceedings of the 2000 ACM conference on Computer supported cooperative work, 1-58113-222-0, ACM Press, 309-318, Philadelphia, Pennsylvania, United States.

[3] Y. CHEN, H. HUANG, W. XU, R. WALLIS, H. SUNDARAM, et al. (2006). *The Design Of A Real-Time, Multimodal Biofeedback System For Stroke Patient Rehabilitation*, Proc. SIG ACM Multimedia 2006, also AME-TR-2006-07, Oct. 2006, Santa Barbara, CA.

[4] S. E. ELLIS and D. P. GROTH (2004). *A Collaborative Annotation System for Data Visualization*, Proc. of the working conference on Advanced Visual Interfaces.,

[5] J. GEMMELL, G. BELL, R. LUEDER, S. DRUCKER and C. WONG (2002). *MyLifeBits: fulfilling the Memex vision*, Proceedings of the 10th ACM international conference on Multimedia, pp. 235-238, Dec. 2002, Juan Les-Pins, France.

[6] S. HODGES, L. WILLIAMS, E. BERRY, S. IZADI, J. SRINIVASAN, et al. *SenseCam: a Retrospective Memory Aid*, Proc. 8th International Conference on Ubicomp 2006, Sep. 2006, Orange County, CA.

[7] M. HOLDEN and T. DYAR (2002). *Virtual environment training: a new tool for neurorehabilitation.* Neurology Report **26**(2): 62-72.

[8] D. JACK, R. BOIAN, A. S. MERIANS, M. TREMAINE, G. C. BURDEA, et al. (2001). *Virtual reality-enhanced stroke rehabilitation.* IEEE Trans Neural Syst Rehabil Eng **9**(3): 308-318.

[9] R. E. KRAUT, D. GERGLE and S. R. FUSSELL (2002). *The use of visual information in shared visual spaces: informing the development of virtual co-presence*, Proceedings of the 2002 ACM conference on Computer supported cooperative work, 1-58113-560-2, ACM Press, 31-40, New Orleans, Louisiana, USA.

[10] J. OU, L. M. OH, J. YANG and S. R. FUSSELL (2005). *Effects of task properties, partner actions, and message content on eye gaze patterns in a collaborative task*, Proceedings of the SIGCHI conference on Human factors in computing systems, 1-58113-998-5, ACM Press, 231-240, Portland, Oregon, USA.

[11] C. PLAISANT, B. MILASH, A. ROSE, S. WIDOFF and B. SHNEIDERMAN (1996). *LifeLines: visualizing personal histories*. Proceedings of the SIGCHI conference on Human factors in computing systems(eds), ACM Press**:** 221--ff.

[12] D. QUAN, D. HUYNH and D. R. KARGER (2003). *Haystack: A Platform for Authoring End User Semantic Web Applications*, 2nd International Semantic Web Conference, Oct. 2003, Sanibel Island, Florida, USA.

[13] C. WENG and J. H. GENNARI (2004). *Asynchronous collaborative writing through annotations*, Proceedings of the 2004 ACM conference on Computer supported cooperative work, 1-58113-810-5, ACM Press, 578-581, Chicago, Illinois, USA.

[14] A. WILHELM, Y. TAKHTEYEV, R. SARVAS, N. V. HOUSE and M. DAVIS (2004). *Photo Annotation on a Camera Phone*, ACM Conference on Human Computer Interaction, Apr. 2004, Vienna, Austria.

[15] S. H. YOU, S. H. JANG, Y. H. KIM, M. HALLETT, S. H. AHN, et al. (2005). *Virtual reality-induced cortical reorganization and associated locomotor recovery in chronic stroke: an experimenter-blind randomized study.* Stroke **36**(6): 1166-1171.