Body and Gesture Tracking, Representation and Analysis: Implications for HCI, Human Motor Function & Virtual Human Action

A NATIONAL SCIENCE FOUNDATION
ENGINEERING RESEARCH CENTER

User Centered Sciences

Yeh Shih-Ching, Jane Hwang, Tim Marsh, Baharak Zali, Issac Cohen, Shri Narayanan, Cyrus Shahabi, Gerry Kim & Skip Rizzo

Research Goals

This User Centered Sciences (UCS) area has multiple projects with unique goals. These involve:

- Gestural HCI – This project aims to design, develop and evaluate a hand gesture based methods for enhancing human computer interaction in collaboration with the IMSC 2020 Classroom Research Vision. The research in this area will also have impact on user interface components across a wide spectrum of IMSC research areas. Comparison of hand gestures for object selection and manipulation with mouse usage is being focused on within the context of development of a 3D User Interface benchmarking scenario.
- Vision-Based Tracking – This project aims to create a vision-based tracking and sensing system in collaboration with Issac Cohen and Shri Narayanan to be used for tracking, representing and quantifying human motor performance within the IMSC Communication Vision. Advances in vision-based tracking technology could form the basis for the capture and recognition of human action required for the creation of multimodal HCI options that utilize gestural behavior. This technology could also have significant impact on applications that target gait analysis and motor rehabilitation and we have developed three VR scenarios for this purpose.
- Body Gesture and the Expression of Emotional State in Virtual Human Representations – This project aims to investigate human capacity to decode the communication delivered via body action in virtual human representations (avatars/agents). Previously, IMSC UCS research has examined the factors that contribute to successful detection, by human users, of emotional facial expressions conveyed by avatars actuated using Performance Driven Facial Animation. The current work has expanded to investigate similar issues for full body non-verbal expression of emotion.

Research Approach

Multiple methodologies have been employed across the projects in this area.

• For the Gestural HCI project, we created a benchmarking scenario that would allow us to make comparisons between gesture interaction and traditional mouse/keyboard methods. Based on this need we have created a block configuration methodology that was derived from some of our previous work with Shepard and Metzler mental rotation stimuli (Figure 1).

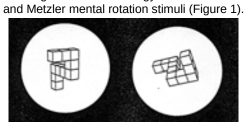

Figure 1 - Shepard & Metzler Block Configurations

These block stimuli have many features that make them attractive as 3D interaction benchmarking stimuli. There is a rich history of research examining human performance with these stimuli, particularly in the area of researching spatial abilities across age and gender. The stimuli are easily capable of being hierarchically presented from very simple renderings to much more complex configurations and in the types of rotations required to perform tasks with them. The influence of prior learning and experience on interaction performance measures with these stimuli is lessened when compared with using realistic objects. And the influence of a persons innate level of spatial ability can be statistically parceled out of the user's hands-on interaction performance with the blocks, by factoring out results from user performance on the highly reliable Vandenberg and Kuse Mental Rotations Test (Figure 2).

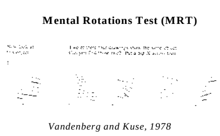

Figure 2. Vandenberg & Kuse Mental Rotations Test

We have now designed a series of progressively more complex block configurations that can be rendered in either mono or stereo mode on a common PC monitor (shutter glasses are required for stereo mode). The task presents pairs of identical block stimuli in orientations (that are readily controllable by the experimenter) and the user is instructed to select and manipulate one of the blocks and to superimpose it upon the target blocks using a designated input device. When the user completes the trial, data is automatically collected on the speed of successful superimposition and the efficiency of the actual movement as measured by the ratio of the shortest path of travel to the actual path. We are now able to configure the blocks to be actuated by any input device. With our performance metrics, we are able to compare speed and efficiency across whatever input devices that we chose to compare. Currently 20 subjects have been tested on this system comparing mouse interaction with a glove-based gesture approach.

•The UCS component of the Communication Vision project (in collaboration with Issac Cohen and Shri Narayanan) has involved the creation of a machine repair test system whereby users are tested on their ability to repair the device either using a paper manual (see Figure 3) or by way of a remote agent that senses the user's status and state and provides help and guidance. We have currently conducted the "manual" baseline test, completed a "Wizard of OZ" study with the same task and are gearing up to compare these data with a single camera/audio dialog sensing system to provide user feedback based on the sensing in an automatic fashion. We are also experimenting with the use of vision based tracking to capture user movement in a series of motor rehabilitation tasks. The goal here is to replace expensive marker based and magnetic tracking systems while the patient does physical rehabilitation in a game-based motor rehabilitation virtual environment (See Figure 4).

• The Body Gesture and the Expression of Emotional State in Virtual Human Representations project uses the Vicon motion tracking system at the USC Zemeckis Center to capture trained actors who are expressing dynamic body postures and gestures that implicitly communicate emotional states. The body actions captured in this system are rendered as "faceless" animated characters using 3D projection technology and are presented to research subjects who attempt to decode the nature of the emotional expressions. These test scenarios consist of a text-based description of the context in which pairs of these animated characters will be interacting. Selective emphasis cues will be provided to research subjects guiding them to observe the action of one of the figures for ratings of the animated characters state. No verbal or facial cues will be presented to subjects in an effort to isolate specific body gesture communication efficacy.

Figure 3. Manual for Machine Repair Baseline Test

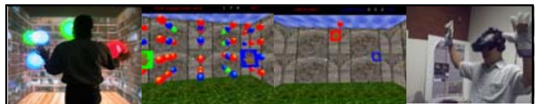

Figure 4. Motor rehabilitation test environments for future integration of vision based tracking

Accomplishments

•Gestural HCI –The block rotation scenario produced data indicating that users preferred to select and manipulate objects using glove-based gestural interaction over mouse interaction. As well speed of performance was also found to be superior on the 3D User interaction task with stereoscopic stimuli (see Figure 5). Analyses are ongoing to determine how age, gender and pre-existing visuospatial ability influences such performance. All data are being saved in order to construct a normative database by which other devices and interactional methods will be compared.

•Vision-Based Tracking – Baseline usability and performance data was collected on the machine repair task comparing manual-based instructions with a Wizard of OZ instructional dialog system (See Figure 6). Proof of concept magnetic tracked hand movement VR rehabilitation scenarios were created and initial user evaluation of these system is ongoing. Collaborations with leading international rehabilitation centers have been formalized and tasks/requirements analysis has begun with these groups to guide system design.

• Body Gesture and the Expression of Emotional State in Virtual Human Representations – The methodology for this project has been formalized after painstaking analysis of the non-verbal communication literature for best practices and via extensive viewing of taped "improv" interactions using professional actors. Actors from the USC School of Theatre have been trained on the key gestures for use in the study and Vicon motion capture of these actions is in progress.

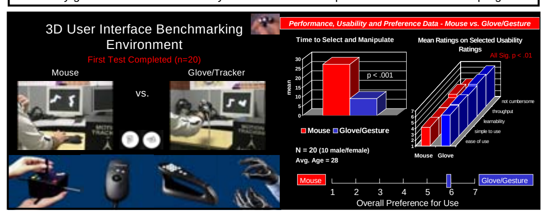

Figure 5. 3D User Interface Benchmarking Scenario and Key Findings

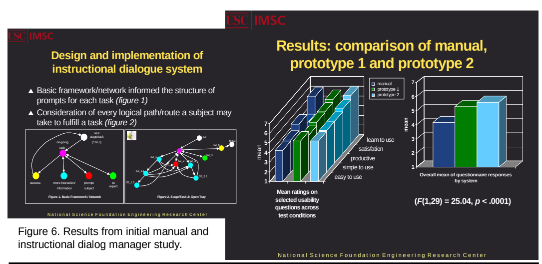

Figure 6. Results from initial manual and instructional dialog manager study.

5-Year Plan

- Gestural HCI – We will continue user testing across interface types and interactional methods with an aim to build "selection and manipulation" normative data and study the influence of preexisting spatial ability on user performance. Since object selection and manipulation are but two of the four Universal interaction task (the other are navigation and system control), we plan to follow a similar benchmarking approach for navigation testing via the design of standardized maze and obstacle environments. Integration of selection, manipulation and navigation tasks will likely follow contingent on the results of our current research. Exploration of vision-based tracking integration to replace current tracked glove solution will be done.
- Vision-Based Tracking – We will use our baseline data from the manual and instructional dialog manager test to compare future iterations of the vision/audio sensing applications as part of the Communication Vision and apply our efforts in collaboration with Gloria Mark's lab examining trust factors in crisis situations. We also expect to continue developing and evaluating the performances of the vision-based algorithms within specific 3D Virtual Environments. The evaluation of the accuracy of the vision-based body motion estimation will be performed by using magnetic sensors and optical markers for comparison studies. Our collaborators will provide input for the design of scenarios of interest for the applications.
- Body Gesture and the Expression of Emotional State in Virtual Human Representations – We intend to complete the first round of decoding studies with human raters. Knowledge gained in this round will be incorporated into the next study to expand our library of gestural behaviors that underlie non-verbal expression of emotion. This knowledge will also support our development of PUI technology.