「様式３」

課題番号 ：P-15

Multimedia Interaction Interfaces in Collaborative E-learning Environments

[1] Organization

Leader: *Noriyuki Matsuda* (Faculty of Systems Engineering, Wakayama University, Japan)
Representatives at RIE: *Kamen Kanev* & *Hiroshi Inokawa* (Research Institute of Electronics, Shizuoka University)
Participants:
Hirokazu Taki (Faculty of Systems Engineering, Wakayama University, Japan)
Hidetoshi Nonaka (Graduate School of Information Science and Technology, Hokkaido University, Japan)
Julita Vassileva (University of Saskatchewan, Canada)
Michael Jenkin (Department of Computer Science and Engineering, Faculty of Engineering, York University, Canada)
Magdalina Todorova (Department of Computer Science, Faculty of Mathematics and Informatics, Sofia University, Bulgaria)
Volodymyr Gnatyuk (V.E. Lashkaryov Institute of Semiconductor Physics, National Academy of Sciences of Ukraine)
Karen Collins (Canadian Centre of Arts and Technology at the University of Waterloo, Canada)
Bill Kapralos (Faculty of Business and Information Technology, University of Ontario Institute of Technology, Canada)
Andrew Hogue (Faculty of Business and Information Technology, University of Ontario Institute of Technology, Canada)
Boris Brimkov (Department of Mathematics, University at Buffalo)
Sanshiro Sakai (Department of Computer Science, Faculty of Informatics, Shizuoka University)

[2] Progress of the research

This cooperative research project explores possibilities for combining expertise from different research and fostering the design and development of innovative interfaces and environments for e-learning based on natural interaction paradigms that are adjustable to the needs of specific user groups. Smart table computer interactions have been identified as a main vehicle of this project and sound interface design for smart tables is being addressed where users position themselves around a horizontal computer screen and a combined audio/vision-based interface is employed for sound source and user tracking in real-time.

Tabletop displays are a relatively new display and interaction technology in which a number of users interact with a large display surface that is placed horizontally. Such devices have already found limited use as general purpose displays, gaming devices, and have been proposed for a range of other applications including their use as control panels for industrial processes. Tabletop displays present a large number of interesting technical and operational questions that must be addressed if the underlying technology is to make inroads into serious applications such as their use in industrial operations and emergency planning and response.

One critical question that must be addressed in order to fully operationalize tabletop displays is the user's ability to integrate multimodal display cues within the device. Of particular interest in this project has been the ability of the user to integrate spatialized visual and audio cues into a coherent perception of a spatialized event. To take a specific example of this, suppose that a warning display is to be presented on a tabletop display (indicating some off-nominal condition in an industrial plant, for example). In a physical embodiment of such a display a warning buzzer would be located with the display and when the off nominal condition occurs the display would indicate the condition and the buzzer would sound to alert the operator of the event. When this physical display is rendered on a tabletop console it is necessary to provide both the visual display and also a localized audio cue that directs the operator's attention to the appropriate visual console. The fundamental question then becomes, how does one generate a spatialized audio cue within the constraints of a tabletop display?

In answering this question it is instructive to look at how this spatialization is accomplished in a movie theatre with a pair of speakers. The audience sits in front of a vertical display with speakers presented to the left and right of the screen. When an event occurs on screen it is not possible to generate a sound at the exact visual location of the event (there is no speaker there), rather sounds are mixed through the two speakers to as to generate a sound pattern that is perceived as though it is at the location of the visual event. This is a well studied problem and a large number

of solutions have been proposed and the basic approach is easily generalized to a collection of loudspeakers that surrounds the listeners. The tabletop display environment presents a number of interesting challenges to the tradition theatre-like audio display environment.

First, the users surround the display rather than all being located on one side of it. This provides a radically different set of constraints in terms of sound field generation and also opens up some interesting questions about the location of the speaker array. Should the speakers be located around the table "in front" of the various participants, or should they be positioned outside of the circle of potential operator locations? Neither solution seems particularly attractive as in the first listeners are very close to speakers while in the second the sound sources are behind some of the listeners. Second, the display surface is horizontal, rather than vertical. This changes the geometry substantively in terms of the direction of viewing and listening. In the movie theatre you hold your head erect and sounds should appear to emanate from a vertical surface that is orthogonal to the viewing direction. In the tabletop environment the surface is oblique to the viewing direction. Third, there is the question of the effect of speaker distance on any mixing strategy. In the theatre model all speakers are the same distance from the listener. In the tabletop version this is likely to no longer be true, especially for two-dimensional spatialization over the tabletop surface. How does this change in speaker distances influence the various mixing models. Finally, and perhaps most challenging, is how to deal with the fact that all of the listeners should perceive the same synthetic sound source location. In the theatre application this task is simplified by having all of the speakers on one side of the display. For the tabletop display however, this is not the case leading to a number of issues in terms of synthesizing spatialized audio.

A number of research exchanges took place during the year. Prof. Kanev undertook two trips to Ontario to visit UOIT, University of Waterloo and York University in 2011. The first trip in the summer was to set up the basic parameters for a more extended visit throughout the fall and winter. During this second visit an experimental apparatus (see results, below) was constructed at York University and installed at UOIT. Prof. Jenkin, funded by a JSPS Short Term Fellowship spent much of February and March 2012 at Shizuoka University. Andrew Speers, a student of Prof. Jenkin also visited Shizuoka in March in order to present some of his research results and to observe the operation of the experimental hardware so that he can utilize a similar system in Canada upon his return.

[3] Results

(3.1) Research results

During this last year work within the project has concentrated on understanding how humans perceive spatialized audio cues in a tabletop environment and how traditional mixing algorithms for vertical displays must be adapted for the generation of effective spatialized audio cues. Experiments at Shizuoka University and UOIT have begun to explore the ability of humans to localize audio cues in a tabletop environment versus their ability to localize such cues in a theatre-like environment. At Shizuoka University this work has lead to the construction of a novel 5x5 speaker array that can be used to generate true sounds at specific locations while at UOIT a 2D mockup surface has been constructed that allows a single sound source to be positioned out of the sight of the user over the entire table surface.

Experiments have demonstrated that humans make consistent errors in sound source localization for table top displays (2). Left-right localization is quite good although forward-backward localization shows systematic errors with sounds being pulled 'towards' the listener. Although the nature of this pulling is unknown it is an open question if this effect is found in the vertical plane. The experimental apparatus constructed at UOIT is being used to answer this question now.

The question of how best to synthesize different virtual sound source locations is considered in (6). Two traditional stereo panning approaches were evaluated in the tabletop environment setting. Systematic errors were identified using both techniques and ongoing work is exploring how to improve on these previous mixing approaches.

Other work within the project has investigated different educational frameworks (3) and technologies for interaction localization on a tabletop display using marked surfaces. (1) and (4) describe various approaches to invisibly mark an interaction surface or object so that the marks provide a map for decoding later interactions with the surface. (1) demonstrates the approach generalized to non-planar surfaces which is of particular interest in the application of tabletop display technology in industrial plant displays where nonlinear (curved) surfaces may be an appropriate surface upon which to mount any display. (5) describes recent work in the

acquisition of 3D models for display in tabletop settings. During the trip of Andrew Speers to Japan a number of 3D external structures were captured suitable for later display on tabletop surfaces

(3.2) Future work

During Prof. Jenkin's stay at Shizuoka University an experimental protocol was established for measuring systematic errors in sound source localization using a pair of speakers. The experimental apparatus has been assembled and software written to allow the experiments to be conducted. A student is currently collecting data and we expect to be able to report the results of this experiment shortly. The experimental setup is being replicated in Canada although with some slight differences. The setup at Shizuoka University is designed to work with the existing speaker array system built there. The test will be to compare a real sound from a number of different locations with a spatialized signal within a psychophysical testing regime. The system at York University will utilize the active display surface which will display a localized visual cue on the tabletop with a similar acoustic panning system and the same experimental protocol.

Further exchanges are planned in order to support the research. Prof. Kanev plans to visit York University and UOIT in the fall of 2012, and Prof. Kapralos is planning to visit Shizuoka in the spring of 2013. Initial planning is underway for a joint graduate course in tabletop display technology to take place in Japan in 2013. In 2011, an application for a Social Sciences and Humanities Research Council Partnership Grant was submitted and we are awaiting results. It is planned that in fall of 2012 we will apply for a Natural Sciences and Engineering Research Council grant with Dr. Collins as Principle Investigator.

[4] Publications

(1) Kanev, K., Mizeikis, V., Gnatyuk, V., Localization Encoding in the Bulk of Physical Objects by Laser-Induced Damage, In *Proceedings of The Joint Int. Conf. on Human-Centered Computer Environments HCCE 2012*, Aizu-Wakamatsu, Japan, March 8-13, 2012, pp.93-98.

(2) Nakano, D., Lam, J., Kapralos, B., Kanev, K., Collins, K., Hogue, A., Jenkin, M., A Framework for Sound Localization Experiments and Automation, In *Proceedings of The Joint Int. Conf. on Human-Centered Computer Environments HCCE 2012*, Aizu-Wakamatsu, Japan, March 8-13, 2012, pp.137-144.

(3) Todorova, M., Kanev, K., Educational Framework for Verification of Object-Oriented Programs, In *Proceedings of The Joint Int. Conf. on Human-Centered Computer Environments HCCE 2012*, Aizu-Wakamatsu, Japan, March 8-13, 2012, pp. 23-27.

(4) Gnatyuk, V., Kanev, K., Gagarsky, S., Features of Transparent Material Marking with Nano- and Subnanosecond Laser Pulses, *The 9th Int. Conf. on Global Research and Education InterAcademia2011*, Sicevita, Romania, September 26-29, 2011, pp.26.

(5) Speers, A., Jenkin, M., Tuning Stereo Image Matching with Stereo Video Sequence Processing, In *Proceedings of The Joint Int. Conf. on Human-Centered Computer Environments HCCE 2012*, Aizu-Wakamatsu, Japan, March 8-13, 2012, pp.208-214.

(6) Lam, J., Kapralos, B., Collins, K, Hogue, A., Kanev, K., Jenkin, M., Sound Localization on Table-top Computers: A Comparison of Two Amplitude Panning Methods. (accepted for publication by Computers in Entertainment)

「様式３」

Traveling Report

Name:	Noriyuki Matsuda
Affiliation:	Faculty of Systems Engineering, Wakayama University, Japan
Period of time:	March 10, 2012 – March 11, 2012
Destination:	Shizuoka University, Japan
Purpose:	To discuss the research and project organization, to schedule future work, and to participate in The Joint International Conference on Human-Centered Computer Environments HCCE 2012.
Name of receiver:	Prof. Kamen Kanev

Name:	Volodymyr Gnatyuk
Affiliation:	V.E. Lashkaryov Institute of Semiconductor Physics, National Academy of Sciences, Ukraine
Period of time:	March 10, 2012 – March 16, 2012
Destination:	Shizuoka University, Japan
Purpose:	To discuss continuing cooperative research, to plan publications, and to participate and report obtained results at The Joint International Conference on Human-Centered Computer Environments HCCE 2012.
Name of receiver:	Prof. Kamen Kanev