Evaluation of an Enhanced Human-Robot Interface*

Carlotta A. Johnson
Center for Intelligent Systems
Vanderbilt University
Nashville, TN 37235
carlotta.a.johnson@vanderbilt.edu

Julie A. Adams
Center for Intelligent Systems
Vanderbilt University
Nashville, TN 37235
julie.a.adams@vanderbilt.edu

Kazuhiko Kawamura
Center for Intelligent Systems
Vanderbilt University
Nashville, TN 37235
kaz.kawamura@vanderbilt.edu

Abstract - *A human-robot interface for a mobile robot was extended to include a discrete geodesic dome called a Sensory EgoSphere (SES) The SES is a two-dimensional data structure, centered on the robot's coordinate frame. The SES provides the robot's perspective of a remote environment via images, sonar, and laser range finder representations. It was proposed that the SES would enhance the general interface usability by decreasing perceived workload and increasing situational awareness. A human factors evaluation was performed to evaluate the established hypothesis. Novice users participated in the evaluation. The purpose of this paper is to review some of the evaluation results.*

Keywords: Sensory EgoSphere, Human-Robot Interface, Graphical User Interface, Workload, and Situational Awareness.

1. Introduction

Determining a mobile robot's present status when the supervisor is located at a remote location can be difficult. A remote supervisor is necessary when environmental hazards or harsh working conditions exist. This paper focuses on a user evaluation of a human-robot interface (HRI) that incorporates a discrete geodesic dome, called the Sensory EgoSphere (SES).

The extraction of environmental information from landmarks and sensor readings is a catalyst for the SES research. The SES links the HRI to the mobile-robot's short-term memory database. The memory database is indexed by azimuth and elevation. The geodesic dome and associated database are called the Sensory EgoSphere.

The SES is a proposed solution to coordinating distributed sensors during mobile robot navigation [7]. The SES may enhance an HRI by providing a robot-centric display of the robot's sensory data to the human [6].

It was believed that a graphical based HRI incorporating the SES would provide a more intuitive sensory data display. The SES display permits the user to mentally fuse notable events that occur in close proximity. The display provides the robot's egocentric view of the environment as the dome is centered on the robot's frame.

The overall research objective was to determine if the Sensory EgoSphere enhanced a human-robot interface. Two research hypotheses were tested. 1. *The SES decreases participant mental workload with the addition of a more intuitive display of sensory data.* 2. *The SES increases participant situational awareness of the robot status and the task/mission status.*

This paper discusses the user evaluation designed to test the stated hypotheses. Section 2 provides an SES overview while Section 3 describes the SES display design and enhanced HRI. Section 4 describes the experimental design. The evaluation results are presented in Section 5. Section 6 provides a discussion of the relevant results and Section 7 provides the conclusions and future work.

2. Sensory EgoSphere

Albus proposed an egosphere. He defined it as "a dense spherical coordinate system with the self (ego) at the origin" [1]. Visible points on regions or objects in the world are projected on to the egosphere. Each of us resides at the origin of our own egosphere. Everything humans observe can be represented by a location with an azimuth elevation and range measured from the center of our ego. To the observer, the world is seen through a transparent sphere. Each observed world point appears on the egosphere at a location defined by the azimuth and elevation.

The SES definition for this work provides a two-dimensional spherical data structure, centered on a robot's coordinate frame. The primary difference between this and Albus' work is that our SES is also a short-term memory used for robot navigation. The SES provides a sparse environmental map containing pointers to object or event descriptors detected by the robot. As the robot operates, both external and internal events stimulate the robot's sensors. Sensory processing modules write to the SES at the node closest to the direction from which the stimulus arrived. Sensory data of different modalities, from similar

* 0-7803-7952-7/03/$17.00 © 2003 IEEE.

directions at similar times, register close to one another on the SES [10].

The work described in this paper employed an ATRV-JR robot as the ego center. The robot was equipped with two cameras, seventeen ultrasonic sonar, and a laser range finder. The camera head was the center of the SES. Since most robots do not have 360-degree sensory data, the SES is an incomplete geodesic dome and is restricted to the vertices within the robot's sensory field. One camera is mounted on a pan-tilt head; therefore image features are stored at the vertex closest to the camera direction. The sonar and laser range finder provide data around the robot's equator, thus the SES equator [10].

The SES was implemented using OpenGL® and Visual Basic™. This implementation is compatible with the agent-based software architecture-programming environment, the Intelligent Machine Architecture (IMA) [9]. An octahedron-based tessellated dome was used. Figure 1 provides several SES views including images on nodes as well as sonar and laser data along the equator.

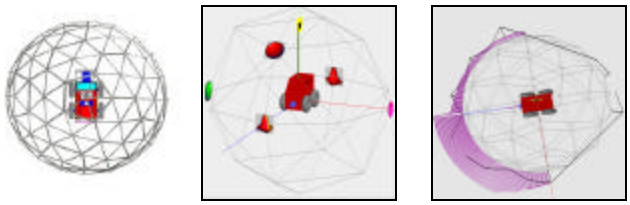

Figure 1. Sensory EgoSphere instances.

3. Human-Robot Interface

The human-robot interface was developed using the Intelligent Machine Architecture (IMA). IMA allows distributed software agents to concurrently execute while facilitating inter-agent communication. The HRI includes the following agents: SES, map, sonar, laser, and camera. This interface was based on the work of Nilas et. al. [8].

The ATRV-JR mobile robot has a sensor suite providing odometry and heading (x, y, θ), compass data, GPS, DGPS, ultrasonic sonar data, laser range data, a forward facing camera, and a backward facing camera. The laser range finder is mounted on the front of the robot. The forward facing pan-tilt-zoom camera system provides a high-speed range of –100 to 110 degrees. This work employed the odometry and heading, compass, ultrasound, laser, and forward facing camera information.

The original HRI, shown in Figure 2, provided the user with an a priori environmental map, the forward facing camera image, and displays of the sonar, laser, and compass readings. The map indicates the robot position and various detected landmarks.

As Figure 3 shows, the SES display was added to the original HRI (Figure 2). The SES display agent communicates with the other sensory and HRI agents. The interface agents provide sensory data displays including the camera, sonar, laser, and compass displays.

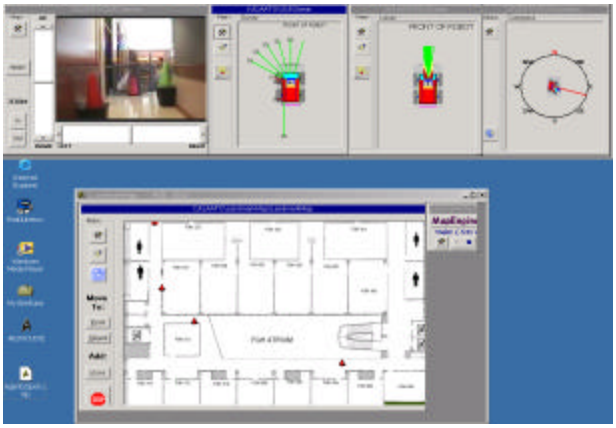

Figure 2. The original human-robot interface.

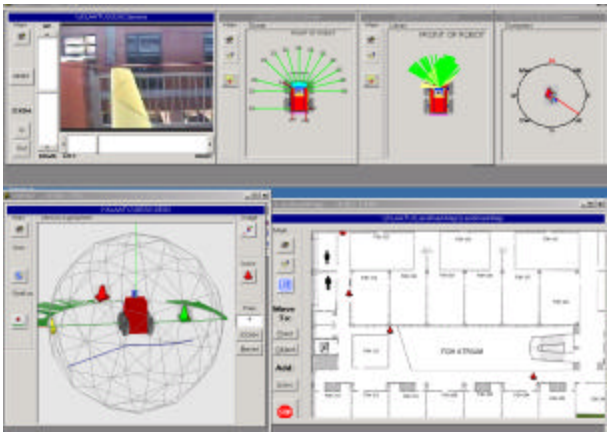

Figure 3. The enhanced human-robot interface.

4. Experimental Design

Twenty-seven novice robotics users from the Nashville community participated in the evaluation. The ability to visualize 3D relationships was deemed important given the remote mobile robot operation. Participants' spatial reasoning abilities were measured via a spatial rotation test. A pre-experimental questionnaire determined familiarity with computers, video games, mobile robots, and graphical user interfaces.

Each user completed tasks with both interfaces (Figures 2 and 3). The experiment consisted of a set of training tasks and a set of evaluation tasks. Each task set required activities with both interfaces. The task and interface presentations were randomized over all participants. There were two evaluation sessions.

The first session included an orientation followed by a training session. Each participant received a fifteen-minute training session. The participants then completed the training task, and then repeated the task with the remaining interface. During the second session, the participants completed the evaluation tasks followed by the same task with the second interface. The participants completed all four tasks. During the task execution, quantitative data was collected via automatic data recording. After each task completion a post-task questionnaire was completed. At the end of the second session a post-experimental questionnaire was completed.

The training task required participants to search for environmental landmarks via the interface displays. The evaluation task required participants to teleoperate the robot along a seventy-meter path and locate pre-defined landmarks. Natural landmarks; such as people, doors, and water fountains were the only environmental obstacles. The landmarks were located along the path and in corners, doorways, and side hallways. The participants provided a navigation command and the robot moved to the defined point and then signaled the participant. During all tasks, the participants were able to change the data display views.

The participants provided navigation commands using point and click interaction on the environment map. The move to point navigation required the definition of waypoints. The participants selected object icons in order to command the robot to go to a particular object.

5. Data Analysis and Results

The training task involved determining the robot's position via the interface displays. The teleoperation tasks entailed driving the robot through an obstacle course while documenting all significant objects. This section discusses some of the results, full details can be found in [5].

Twenty-seven individuals participated but the statistical analysis was performed on the ten data sets from participants who completed both teleoperation tasks with no major system or hardware failures. This group included five males and five females. Two participants had low spatial reasoning capabilities, four participants had average ratings, and four had high ratings. The participant ages ranged between 18 and 70 years, with an average of 30. Due to the small sample size, non-parametric evaluations were performed, such as the Kruskal-Wallis Rank and Spearman Rank Correlation tests.

A training task score was calculated based upon the robot's placement, orientation, and location as well as the color of landmarks. The teleoperation task score was calculated based upon the placement and color of landmarks. A task score comparison across the interfaces was used to evaluate the participants' situational awareness. The raw data showed that the training task scores were higher with the original interface, as shown in Table 1. Conversely the results from the enhanced interface teleoperation task found a higher overall score, as shown in Table 2. These results may imply that the enhanced interface is more useful when the robot is in motion but these results were not statistically significant.

Table 1. Training task scores.

Training Task		
Sub-tasks	Original Interface	Enhanced Interface
Robot Placement	100	87.5
Robot Orientation	100	75
Cone Placement	89.58	67.71
Cone Color Scores	95.83	86.46
Driving Direction	85	97.14
Overall Score	93.98	81.4

As Table 1 demonstrates, the original interface training task sub-scores rated higher for robot placement, robot orientation, cone placement, and cone color scores. The driving direction score was the only higher enhanced interface training task score. The teleoperation sub-task score results differed, as shown in Table 2. The original interface cone color score was higher while the enhanced interface cone placement score was higher.

Table 2. Teleoperation task scores.

Teleoperation Task		
Sub-tasks	Original Interface	Enhanced Interface
Cone Placement	67.14	78.57
Cone Color Scores	92.85	84.29
Overall Score	80	81.43

An analysis of the task score versus the number of camera clicks found a majority of negative correlations, indicating that increased camera usage decreased the task scores. The analysis of the original interface training task found a negative correlation between the driving direction score and the total pan clicks ($r = -0.859$, $p = 0.029$). There were negative correlations between the driving direction score and the total reset clicks ($r = -0.959$, $p = 0.002$) as well as the driving direction score and total camera clicks ($r = -0.826$, $p = 0.043$). The original interface

training task evaluation found three negative correlations: total tilt clicks and driving direction score ($r = -0.987$, $p = 0.0$), total zoom-out clicks with the robot placement score ($r = -0.764$, $p = 0.046$), and the total reset clicks with the driving direction score ($r = -0.956$, $p = 0.003$). A negative correlation between the total zoom-out clicks and the overall score ($r = -0.748$, $p = 0.013$) was found for the enhanced interface teleoperation task. Finally, there was a positive correlation for the enhanced interface teleoperation task between the cone placement score and the total reset clicks ($r = 0.717$, $p = 0.02$).

The task score decreased with increased SES usage for the enhanced interface training and teleoperation tasks. Several negative correlations were found for the training task: total pan left clicks vs. cone color score ($r = -0.679$, $p = 0.064$), total pan right clicks vs. robot orientation score ($r = -0.917$, $p = 0.001$), total clicks with robot orientation score ($r = -0.810$, $p = 0.015$), and cone color score vs. total clicks ($r = -0.844$, $p = 0.008$). One negative correlation existed for the teleoperation task between the cone color score and the total tilt up clicks ($r = -0.675$, $p = 0.032$).

The Multiple Resources Questionnaire (MRQ) [2] was used to detect differences in operator resource usage between the interfaces. The MRQ measures the resources participants expend in order to accomplish a given task.

The training task results found overall resource scores slightly higher for the original interface (2.57) than the enhanced interface score (2.28). Although these results were not statistically significant, they imply that more resources may have been used with the original interface. The teleoperation task overall resources with the original interface rated 2.30 and 2.44 for the enhanced interface ($z = -0.89$, $p = 0.37$). Therefore, the results suggest that the SES did not reduce the multiple resource ratings during teleoperation.

Differences in some MRQ sub-processes were found. The original interface training task resulted in higher manual, short-term memory, spatial attentive, spatial quantitative, visual lexical, visual temporal and overall resources ratings. The original interface teleoperation task resulted in higher spatial quantitative and visual lexical processing ratings. These results suggest that the enhanced interface may reduce the resource demand during the training task, but these results are not significant.

The enhanced interface teleoperation task MRQ scores increased with decreased camera usage ($r = -0.667$, $p = 0.035$). Positive correlations between the enhanced interface MRQ ratings and the SES clicks existed. For example, total zoom-out clicks vs. spatial quantitative resource ($r = 0.861$, $p = 0.006$), total pan left clicks vs. spatial positional resource ($r = 0.772$, $p = 0.025$), and total pan right clicks vs. overall rating ($r = 0.764$, $p = 0.027$). Correlations for the teleoperation task were found between total zoom-out clicks and spatial quantitative resource ($r = 0.69$, $p = 0.027$), total pan left clicks and spatial quantitative resource ($r = 0.717$, $p = 0.02$), and total pan right clicks and spatial quantitative resource ($r = 0.878$, $p = 0.0001$).

The enhanced interface MRQ score increased with increased SES usage ($r = 0.807$, $p = 0.005$). Correlations between the spatial reasoning score and total SES clicks demonstrated that participants with spatial reasoning ability tended to use the SES more. The two primary types of SES clicks were the scan ($r = 0.683$, $p = 0.037$) and reset ($r = 0.894$, $p = 0.026$).

A negative correlation existed between the overall MRQ rating and the task score during the original interface teleoperation task ($r = -0.77$, $p = 0.009$). Additionally, a negative correlation existed between the overall MRQ rating and the overall training task score when using the enhanced interface ($r = -0.72$, $p = 0.04$). There was a positive correlation between the driving direction score and the spatial quantitative sub-process ($r = 0.88$, $p = 0.009$) for the enhanced interface training task. A positive correlation existed between the driving direction score and the visual temporal process for the enhanced interface training task ($r = 0.76$, $p = 0.046$). There were no correlations for the original interface teleoperation task.

The NASA-TLX [4] tool measures perceived workload. The participants completed the first portion of the tool, ranking each tool component on the scale between 0 and 100. The participants did not complete the pair wise comparison selection. The overall workload rating was determined by averaging all sub-scale responses.

The following workload sub-ratings were rated higher during the training task for the enhanced interface over the original interface: necessary thinking (original: 51.2, enhanced: 57.6), task difficulty (original: 26.8, enhanced: 29.1), physical effort (original: 1.25, enhanced: 1.5), and stress level (original: 2.38, enhanced: 6.88). The enhanced interface exhibited a lower overall perceived workload for the teleoperation task but not for all individual measurements. The enhanced interface rated higher for the frustration (original: 17.3, enhanced: 33.6) and stress (original: 13.0, enhanced: 17.5) levels.

There were negative correlations between the number of SES clicks and the NASA-TLX ratings. For example the frustration level was reduced with increased SES use ($r = -0.83$, $p = 0.002$) for the training task. The NASA-TLX ratings with the training task also demonstrated positive correlations for the SES. For example, the task difficulty with total zoom-in clicks ($r = 0.71$, $p = 0.04$), and the

mental effort and scan clicks (r = 0.719, p = 0.04). A negative correlation existed between total pan right clicks and necessary thinking (r = -0.636, p = 0.04) for the teleoperation task. Additionally, there was a positive correlation between total scan clicks and the mental effort (r = 0.66, p = 0.04) for the teleoperation task.

Several positive correlations existed between total camera clicks and the NASA-TLX sub-ratings. The original interface training task results found a correlation between the necessary thinking and total zoom-out clicks (r = 0.88, p = 0.02), and time required with total reset clicks (r = 0.893, p = 0.02). The enhanced interface results for the same task found positive correlations between time required and total zoom-in clicks (r = 0.861, p = 0.013), mental effort and total zoom-in clicks (r = 0.975, p = 0.0), physical effort with total pan clicks (r = 0.77, p = 0.04), and frustration level with total tilt clicks (r = 0.788, p = 0.035). Correlations existed for the original interface teleoperation task between total zoom-in clicks and time required (r = 0.664, p = 0.036) and total zoom-out clicks and time pressure (r = 0.693, p = 0.026). No positive correlations existed for the enhanced interface teleoperation task.

It was found that increased map clicks decreased the NASA-TLX workload rating for the original interface teleoperation task. There were several negative correlations between the NASA-TLX and the number of map clicks for this interface. Negative correlations existed between necessary thinking and total add icon clicks (r = -0.74, p = 0.021), frustration level and total add icon clicks (r = -0.67, p = 0.05) as well as between overall workload rating and total add icon clicks (r = -0.68, p = 0.04). Positive correlations between the overall rating with total map (r = 0.67, p = 0.05) and add icon (r = 0.691, p = 0.039) clicks existed for the enhanced interface. Correlations between the number of SES clicks and the NASA-TLX workload rating found no consistent positive or negative correlation for either task set.

6. Discussion

The Multiple Resource Questionnaire (MRQ) [2] and NASA-TLX [4] methodologies were used to evaluate the first hypothesis. The MRQ evaluation was to determine if the enhanced interface reduced the resources participants used to complete a task. The assumption was that reduced resources would imply reduced perceived mental workload. It was shown via the correlation analysis that a relationship between the resources and workload existed.

A comparison of participants' responses found a higher numerical resource value implied higher usage of that resource to complete a task independent of task order. In a comparison of the enhanced and original interfaces, it was shown that the enhanced interface required fewer multiple resources. This was true for all categories except the spatial emergent. Since the enhanced interface included the SES, this may account for the increased resource usage.

The teleoperation task results were contradictory. The manual resources were the same for both interfaces. The original interface had higher spatial quantitative and visual lexical resources. The remaining resource ratings were higher for the enhanced interface, including the overall rating. The MRQ results disprove the concept that the enhanced interface reduced multiple resources usage. The enhanced interface actually increased the multiple resource demands by approximately 5%. However, the training task showed a resource demand reduction by approximately 11%. This increase may exist because the SES not as useful when the robot was moving, therefore causing an increase resource usage. The enhanced interface teleoperation tasks may have increased mental workload based upon the increased resources.

The hypothesis was that the SES display would reduce perceived mental workload. The training task enhanced interface demonstrated higher demands for necessary thinking, task difficulty, physical effort, and stress level. These findings may be due to the addition of the SES display.

In a comparison of the interfaces for the teleoperation task, it was found that the enhanced interface received higher ratings for the frustration and stress levels. These findings could be attributed to the odometry error as well as the SES display. The overall comparison showed that the perceived mental workload fell with the enhanced interface by approximately 13%. This result does imply that the enhanced interface would reduce the perceived mental workload.

Several positive as well as negative correlations were found related to the MRQ and NASA-TLX results. The positive correlations suggest that perceived workload might be related to the corresponding MRQ processes. Several negative correlations also existed suggesting that it may not be possible to predict perceived workload based on some MRQ resources. Therefore, the relationship between the two tools is inconclusive.

In conclusion, the raw data implies confirmation of the hypothesis that the SES display would decrease perceived mental workload but the analysis did not statistically support this hypothesis.

The three levels of situation awareness are perception, comprehension, and prediction [3]. This work proposed that the SES display would move the

participants' situation awareness level from the perception to the comprehension level; therefore increasing the participants' situation awareness. Situation awareness was evaluated by examining the task scores. For the two training tasks, the theory was that the cone color score might not differentiate between the two tasks, as this would be considered the perception level. However, it was believed that the remaining scores would improve with the enhanced interface. These scores correspond to the comprehension level. The results found that the enhanced interface only improved the driving direction score. This improvement suggests that the second hypothesis may be partially validated.

With respect to the teleoperation tasks, the cone color score represented the situational awareness perception level. The cone placement score should have improved with the addition of the SES display. The results found that there was on average a twenty-one-point improvement for the cone placement score with the enhanced interface. Therefore, it was suggested that the raw score for the enhanced interface improved situation awareness in the cone placement task. In summary, the raw data weakly suggests that the hypothesis is correct for both tasks.

7. Conclusions and Future Work

The user study did not statistically support the research hypotheses, but the raw data did weakly imply confirmation of the hypotheses. There is need for additional well-controlled evaluations. The results related to perceived workload and situation awareness along with the usability data suggest modifications to the interface and SES display. New evaluations with tasks that are more stringent on a larger sample size should be completed. Additionally, some influences on the workload, task time, and task score should be minimized.

This work has presented an enhanced HRI that included the addition of the SES. A statistical analysis was performed using the data from ten participants who completed both teleoperation tasks with no major failures. The non-parametric analysis included the Spearman rank correlation and the Kruskal-Wallis rank test. These results were analyzed in order to determine the validity of the research hypotheses.

Acknowledgements

The authors wish to thank Phongchai Nilas for his assistance when developing the interfaces evaluated in this work.

References

[1] J. A. Albus, "Outline for a theory of Intelligence", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 21, No. 3, pp. 473-509, May/June 1991.

[2] D. P. Boles and L. P. Adair, "The Validity of the Multiple Ratings Questionnaire (MRQ)", Proc. of the *Human Factors and Ergonomics Society 45th Annual Meeting*, Minneapolis, pp. 1795-1799, Oct. 2001.

[3] M. Endsley, "Theoretical Underpinnings of Situation Awareness: A Critical Review", *Situation Awareness Analysis and Measurement*, M. Endsley and D. Garland (Eds.). Lawrence Erlbaum Associates, London, pp. 1-32, 2000.

[4] S. G. Hart and L.E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, in Human Mental Workload", *Human Mental Workload*, P. A. Hancock and N. Meshkati (Eds.), Elsevier Science Publishing Company, New York, pp. 139-183, 1988.

[5] C.A. Johnson, "Enhancing a Human-Robot Interface Using a Sensory EgoSphere", Ph.D. Thesis, Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, March 2003.

[6] C. A. Johnson, A. B. Koku, K. Kawamura, and R. A. Peters II, "Enhancing a human-robot interface using Sensory EgoSphere", Proc. of the *2002 IEEE International Conference on Robotics and Automation*, Washington DC, pp. 4132-4137, May 2002.

[7] K. Kawamura, A.B. Koku, D.M. Wilkes, R. A. Peters II, and A. Sekmen, "Toward Egocentric Navigation", *International Journal of Robotics and Automation*, Vol. 17, No. 4, pp.135-145, Nov. 2002.

[8] K. Kawamura, P. Nilas, K. Muguruma, J.A. Adams, and C. Zhou, "An Agent-Based Architecture for an Adaptive Human-Robot Interface", Proc. of the *36th Hawaii International Conference on System Sciences*, Hawaii, 2003.

[9] R. T. Pack, "IMA: The Intelligent Machine Architecture", Ph.D. Thesis, Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, 1998.

[10] R. A. Peters II, K. A. Hambuchen, K. Kawamura, and D. M. Wilkes, "The Sensory EgoSphere as a Short-Term Memory for Humanoids", Proc. of *IEEE Robotics and Automation Society International Conference on Humanoid Robots*, Tokyo, pp. 451-459, Nov. 2001.