

# The Use of Visual and Auditory Feedback for Assembly Task Performance in a Virtual Environment

Ying Zhang, Terrence Fernando, Reza Sotudeh, Hannan Xiao

University of Hertfordshire, University of Salford, University of Hertfordshire, University of Hertfordshire  
{y.10.zhang@herts.ac.uk, t.fernando@salford.ac.uk, r.sotudeh@herts.ac.uk, h.xiao@herts.ac.uk}

## Abstract

*This paper presents our creation and evaluation of multi-modal interface for a virtual assembly environment. It involves implementing an assembly simulation environment with multi-sensory feedback (visual and auditory), and evaluating the effects of multi-modal feedback on assembly task performance. This virtual environment experimental platform brought together complex technologies such as constraint-based assembly simulation, optical motion tracking technology, and real-time 3D sound generation technology around a virtual reality workbench and a common software platform. A peg-in-a-hole and a Sener electronic box assembly tasks have been used as the task cases to perform human factor experiments, using sixteen subjects. Both objective performance data (task completion time, and human performance error rates) and subjective opinions (questionnaires) have been gathered from this experiment.*

**Keywords** --- **Virtual Environment, Assembly Simulation, Multi-sensory Feedback, Usability, Task Performance.**

## 1. Introduction

In the manufacturing industry arena, Virtual Environment (VE) technology has the potential for offering a useful method to interactively evaluate assembly-related engineering decisions through analysis, predictive models, visualisation, data presentation, and to factor the human elements and considerations into finished products very early in the development cycle, without physical realisation of the products [1]. This could potentially lead to lower cost, higher product quality, shorter time-to-market, thus improving competitiveness of the innovative products. Assembly is an interactive process involving the operator (user) and the handled objects, and hence simulation environments must be able to react according to the user's actions in real time. Furthermore, the action of the user and the reaction of the environments must be presented in an intuitively comprehensible way. Therefore, it is of great importance to investigate the factors related to information presentation modes and integration mechanisms, which affect the human performance in

performing assembly task in VEs. The multi-modal information presentation, integrated into the VE, has potential for stimulating different senses, increasing the user's impression of immersion and the amount of information that is accepted and processed by the user's perception system. Consequently, the increase of useful feedback information may enhance the user's efficiency and performance while interacting with VEs. However, despite of recent efforts in assembly simulation [2,3,4,5,6] and 3D sound performance modelling in VEs [7,8,9,10,11], very limited research has been conducted to investigate and evaluate the effects of multi-modal feedback mechanisms, especially 3D auditory and visual feedback, on assembly task performance within VEs [12].

This paper presents the overall system architecture implemented for creating a multi-modal virtual assembly environment (VAE), the approaches adopted to evaluate the factors affecting the user's performance in performing the assembly tasks. In particular, it addresses whether the introduction of auditory and/or visual feedback into VAE improves the assembly task performance and user's satisfaction; which type of the feedback is the best among neutral, visual, auditory and integrated feedback (visual *plus* auditory); and whether the factors of gender, age and task complexity have impacts on the assembly task performance with the introduction of visual and/or auditory feedback into VEs.

## 2. Experimental Platform of the Assembly Task Performance

The hardware configuration and software architecture of the experimental system platform for multi-modal virtual assembly task performance evaluation are addressed in this section.

### 2.1. Hardware Configuration of the Platform

The hardware configuration of the experimental system platform for virtual assembly task performance is comprised of three major parts: visualisation subsystem, auralisation subsystem, and the real-time optical motion tracking system (see Figure 1). The core of the visualisation subsystem is the Trimension's V-Desk 6, a fully integrated immersive L-shaped responsive workbench driven by Silicon Graphics Incorporated (SGI) desk-side Onyx2 supercomputer with four

250MHz IP 27 processors and an InfiniteReality-2E Graphics board. The Trimension's V-Desk 6 is integrated with StereoGraphics' Crystal Eyes3 liquid crystal shutter glasses and the infrared emitter that is connected to the Onyx2 workstation. These are used to generate stereoscopic images of the virtual world; one from the viewer's left eye perspective, and the other one from the right eye. When the user uses a pair of Crystal Eyes liquid crystal shutter glasses to view the virtual world, these images are presented to the corresponding eye, providing the user depth cues that make the immersive experience realistic.

The auralisation subsystem is based on a sound server (Huron PCI audio workstation), which is a specialised Digital Signal Processing (DSP) system. It employs a set of TCP/IP protocol-based procedures in terms of Spatial Network Audio Protocol (SNAP) to allow the VE host (i.e. visualisation subsystem) to transmit the attributes of the assembly scene, positional

information of the user and the sound-triggering event to the sound server through a local area network. The VE host sends packets specifying the auditory-related attributes of the scene and the events, such as collisions and motions between the manipulated objects, the position of the event, the position of the user, and the environmental attributes, which are derived from the geometry of the assembly environment. From these packets, the auralisation subsystem generates a set of auralisation filters and sends them to the DSP boards. Based on an event-driven scheme for the presentation of objects' interaction, the DSP board samples and processes sound materials (data streams) with specified filters. Processed sound materials are then sent back to a set of headphones or an array of loudspeakers within the VE area in analogue form through coaxial cables. The auditory feedback in this experiment was presented to the user using a pair of the Sennheiser HD600 headphone.

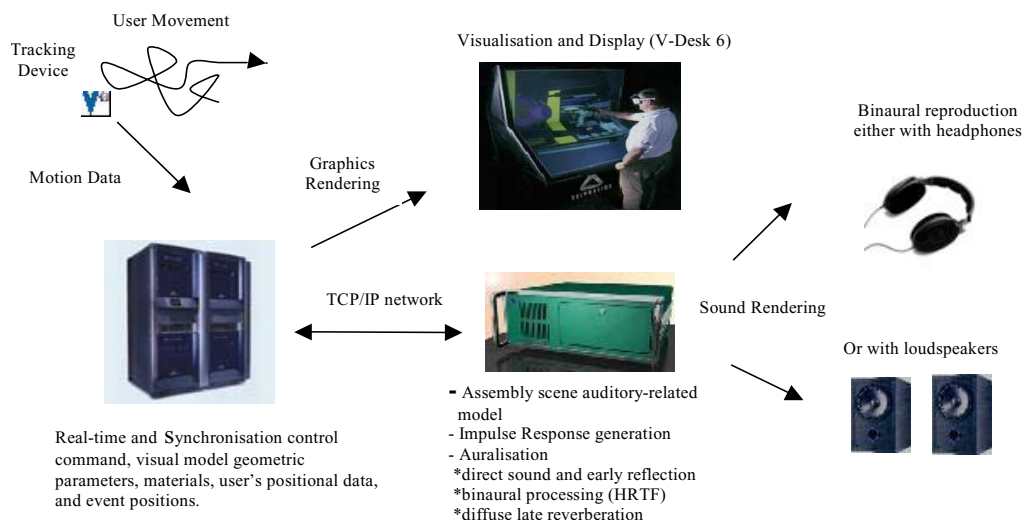


Figure 1 Infrastructure of the System Platform

The optical motion tracking system (Vicon's 612 workstation) provides dynamic, real-time measurement of the position (X, Y and Z) and the orientation (Azimuth, Elevation, and Roll) of the tracked targets such as the user's head and hands, and manipulation tools, using passive-reflective markers and high speed, high resolution cameras. It is connected to the VE host using the TCP/IP protocol over a local area gigabit Ethernet. A Wand is used to support interactive object selection and virtual assembly operations. A virtual 3D pointer with ray-casting and a virtual hand are utilised as the interaction metaphor for the assembly operation.

## 2.2. Software Architecture of the Platform

The software environment is a multi-threaded system that runs on SGI IRIX platforms. It consists of the *User-Interface/Configuration Manager*, the *World-*

*Manager*, the *Input-Manager*, the *Viewer-Manager*, the *Sound-Manager*, the *Assembly-Simulator*, the *CAD Translator* and the *CAD Database* (see Figure 2). The *User-Interface/ Configuration Manager* tracks all master processes to allow run time configuration of different modules.

The *World-Manager* is responsible for the administration of the overall system. It coordinates the visualisation, user's inputs, databases, assembly simulation, and visual and auditory feedback generation. The *World-Manager* fetches the user's inputs for manipulation, produces constrained motion using the *Assembly-Simulator*, and passes the corresponding data (e.g. the position and orientation information of the objects and the user) to the *Viewer-Manager* and the *Sound-Manager* for auditory and visual feedback generation. The new data is used to update the scene graph and control the sound server via the *Sound-*

*Manager*. The *World-Manager* also has the responsibility to synchronise various threads such as rendering and collision detection. Extensions to the OpenGL Optimiser have been made to view the scene using different display technologies (e.g. L-shaped Workbench, CAVE and Reality Room). The *Viewer-Manager* renders the scene to the selected display facility in the appropriate mode. Rendering is performed using parallel threads to provide real time response.

The *Input-Manager* manages user-object interactions, establishing the data flow between the user's inputs and the objects that are held by the *World-Manager*. It supports devices such as pinch gloves, Wands and Vicon's optical motion tracking system. These inputs describe the user's actions/commands in the VE. Each device has a thread to process its own data. These threads run in parallel with the rendering threads to achieve low latency. Once the assembly components are loaded into the scene graph via the *CAD-Translator*, the *Input-Manager* allows the user to select and manipulate objects in the environment. The *Sound-Manager* gets the location data of the user (listener/viewer), the positions of the collisions and motions (sound sources), and the parameters relating to sound signal modulation from the *World-Manager* and

the *Assembly-Simulator*, and then uses the Application Programming Interface (API) of the Huron audio workstation to manage the audio workstation via local network using the TCP/IP protocol.

The *Assembly-Simulator* carries out the detection of collisions between the manipulated object and the surrounding objects, supporting interactive constraint-based assembly operations. During object manipulation, the *Assembly-Simulator* samples the position of the moving object to identify new constraints between the manipulated object and the surrounding objects. Once new constraints are recognised, new allowable motions are derived by the *Assembly-Simulator* to simulate realistic motion of assembly objects. Parameters such as the accurate positions of the assembly objects are sent back to the *World-Manager*, which defines their precise positions in the scene. When a constraint is recognised, the matching surfaces are highlighted to provide visual feedback, and/or 3D auditory feedback is generated through the *Sound-Manager* and the sound server.

The details of the virtual assembly scene, and auditory feedback rendering, and the unifying mechanism of visual and auditory feedback generation can be found in [13, 14, 15].

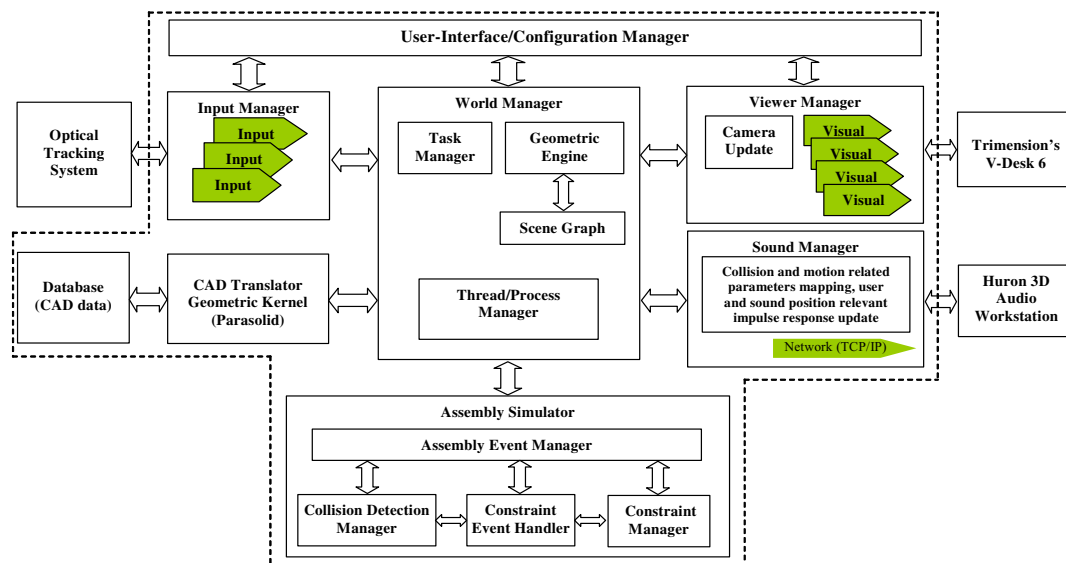


Figure 2 Software Architecture

### 3. Task Performance Evaluation

This section presents the experiment of assembly task performance evaluation including experiment hypotheses, objective evaluation, and subjective evaluation.

This research evaluated the effects of auditory and visual feedback on the assembly task performance, respectively, with the hypothesis that the performance could differ significantly between different feedback conditions. The performance is measured on the basis of objective and subjective means, where objective means

is the time to taken to complete the assembly task and the number of performance failure, and subjective means is the questionnaires for subjective ratings and preferences. There are two independent variables in the experiment: auditory feedback and visual feedback, which can be present and absent. The variations of the independent variables form the different feedback conditions of the multi-modal VAE system as described in Table 1, namely, neutral condition, visual condition, auditory condition and integrated feedback condition. The dependent variables are the Task Completion Time (TCT) and the Human Performance Error Rate (HPER)

under each experiment condition, and subjective ratings and preferences.

Conditions	Colour	Sound
Neutral	× (Absent)	× (Absent)
Visual	√ (Present)	× (Absent)
Auditory	× (Absent)	√ (Present)
Integrated	√ (Present)	√ (Present)

Table 1: Four Experimental Conditions

### 3.1. Experiments Hypotheses

The following hypotheses were assumed in the experiment:

- **The use of visual feedback can lead to better task performance than neutral condition.** Task performance is measured by TCT, HPER and subjective satisfaction. TCT is expected to decrease by providing essential collision, interaction and constraint cues by visual feedback for the assembly task. HPER is expected to decrease by introducing visual feedback into the VAE, especially for the complex task case. The subjective preference to and satisfaction with the interface with visual feedback is expected to be higher than without any feedback. It is expected that this could be indicated by the visual feedback condition having statistically significant higher scores on the rating scales by the questionnaires as compared to the neutral condition.
- **The use of 3D auditory feedback can lead to better task performance than neutral condition.** Better task performance is expected to be shown by shorter TCT, lower HPER and better subjective satisfaction for the auditory feedback condition than the neutral condition. Auditory feedback provides more information for producing a realistic and productive application than no any sensory cues, and the user could be better immersed with this information. Subjective preference to and satisfaction with the interface with auditory feedback is expected to be higher than without any feedback. This could be demonstrated by the auditory feedback condition having statistically significant higher scores by the questionnaires as compared to the neutral condition.
- **The use of integrated feedback (visual plus auditory) can lead to better task performance than either feedback used in isolation.** It is anticipated that this could be shown by shorter TCT, lower HPER, and statistically significant differences between the related rating scale results for the integrated feedback as compared to the conditions with just auditory or visual cues.
- **The factors of gender, age and task complexity have impacts on assembly task performance with the introduction of visual and/or auditory feedback**

*into virtual assembly environment.* It is expected that females exhibit better task performance improvement than males, and seniors exhibit better task performance improvement than youngsters, when introducing visual and/or auditory feedback into VAE.

### 3.2. Objective Evaluation

For the objective evaluation, a peg-in-a-hole assembly task (see Figure 3), which is relatively simple but geometrically well defined and accurate for TCT measurement, was used to explore and evaluate the effectiveness of neutral, visual, auditory and integrated feedback mechanisms on the assembly task performance. The peg-in-a-hole assembly task has several phases: (a) Placement of the peg to the upper surface of the plate (see Figure 3a); (b) Collision between the bottom surface of the peg and the upper surface of the plate (see Figure 3b); (c) Constraint recognition (see Figure 3b); (d) Constrained motion on the plate (see Figure 3c); (e) Alignment constraint between the peg cylinder and the hole cylinder (see Figure 3d); (f) Constrained motion between two cylinders (see Figure 3e); (g) Collision between the bottom surface of the peg ear and the upper surface of the plate (see Figure 3f); and (h) Constraint recognition (see Figure 3f). Different realistic 3D localised sounds and/or colour intensity/modification of the colliding polygons are presented as the action cues for each of the aforementioned phases.

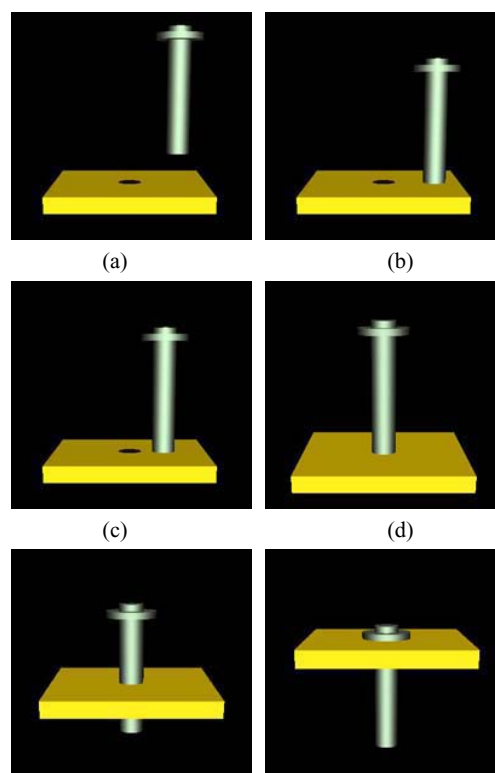


Figure 3: Virtual Assembly Scenario of Peg-in-a-hole Task



The objective evaluation is based on the TCT and HPER. The TCT, which represents the time span between the start and the end of the peg-in-a-hole task, were recorded by the experimental platform. The software timer was set to start when the subjects grabbed the peg to begin the assembly task process, and to stop when the subjects completed the assembly process and released the peg. The system clock drove the timer. The number of failures under different feedback conditions was counted by the experimental platform. A trial was considered to be a failure, when the subject made some errors and thus did not complete the task successfully, or he/she completed the trial beyond a fixed time period. The HPER was calculated by using the number of the failures and the total number of trials.

### 3.3. Subjective Evaluation

For the subjective evaluation of neutral, visual, auditory and integrated feedback mechanisms on the assembly task performance, the Sener electronic box assembly case from an aerospace company called Sener in Spain was used (see Figure 4).

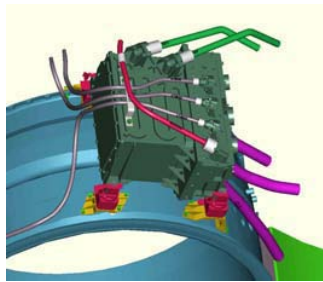


Figure 4: Sener Electronic Box Assembly Task

The Sener electronic box and its brackets assembly task scenarios have been implemented (see Figure 5). The assembly task involves several phases:

- (a) Inspect the environment and identify the parts to be assembled, this allows the subjects to be familiar with the assembly parts and its final assembly status (Figure 5a).
- (b) Mount the supporting brackets and bolt them to the frame. This needs subjects to undertake some exploring and reasoning to perform the assembly operations (Figure 5b). It involves: (i) pick up a bracket and identify its position; (ii) place the bracket into its position; (iii) identify and pick up the bolts; and (iv) bolt the bracket to the frame.
- (c) Slide the electronic box into the brackets (Figure 5c). This is expected to measure the performance when assembling large objects. It

involves: (i) pick up the box and determine its correct orientation; and (ii) slide the box into the brackets.

- (d) Plug the pipes into the electronic box (Figure 5d). It involves: (i) pick up the pipes and identify their correct locations; and (ii) attach the pipes to the box.

The subjective evaluation used the questionnaires to perform the subjective measurements including 10-point rating scales of the overall satisfaction, the realism, perceived task difficulty and performance, ease learning, perceived system speed and overall reaction to the received feedback. Additionally, after the subjects completed the tasks under all conditions they were required to rank the four feedback conditions in the order according to their preference from liking the best to the worst, and completed a set of 7-point rating scales and open-ended questions comparing the different feedback cues. The 7-point rating scales asked the subjects to compare how well the different feedback cues helped them to complete the task, how they foresaw these cues helpful in a real design application, and which kind of feedback cues they preferred. Finally, subjects were asked to provide general opinions and comments about their experiences. The answers of the subjects were recorded and analysed.

The experimental results are being analysed in various statistical methods such as *pair-wise t-test*, *repeated measures ANOVA* and *Friedman ANOVA* etc.

### 4. Conclusions

A VAE system platform, integrated with visual and 3D auditory feedback, has been developed in order to explore and evaluate the effects of neutral, visual, auditory and integrated feedback mechanisms on the task performance in the context of assembly simulation. A peg-in-a-hole and a Sener electronic box assembly tasks have been used as task cases to perform evaluation experiments, using sixteen subjects. At present, the task performance evaluation experiments have completed. The data are being analysed in order to testify the hypotheses. These mainly relate to the best type of feedback among neutral, visual, auditory and integrated feedback mechanisms, whether the integration feedback mechanisms of visual and auditory improves the assembly task performance more than the individual one within the VAE, and whether the method to integrate them together affects the task performance. For the future research, it requires determining how auditory feedback affects performance in specific design and tasks, and determines the substitution of 3D auditory feedback for force feedback in the assembly and manipulation tasks in VEs and how the 3D auditory feedback should be presented to maximize its utility.

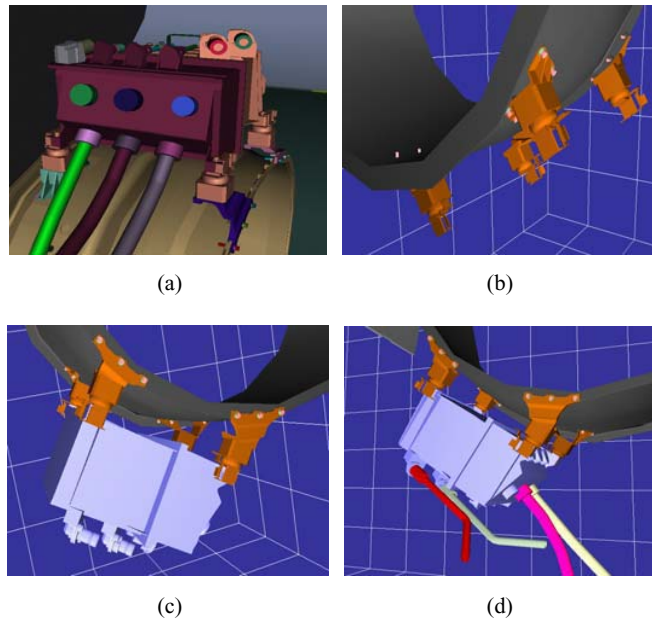


Figure 5: Virtual Assembly Scenario of Sener Electronic Box Task

## References

- [1] F. Dai (ed.) (1998). *Virtual Reality for Industrial Application*, Springer Verlag, Berlin, Heidelberg, Germany.
- [2] J. M. Maxfield, T. Fernando and P. Dew (1998). A Distributed Virtual Environment for Collaborative Engineering. *Presence*, Vol.7, No.3, 241-261, June.
- [3] M. Lin and S. Gottschalk (1998). Collision Detection between Geometric Models: A Survey, Proceedings of IMA Conference on Mathematics of Surfaces.
- [4] S. Jayaram, U. Jayaram, Y. Wang, H. Tirumali, K. Lyons and P. Hart (1999). VADE: A Virtual Assembly Design Environment, IEEE Computer Graphics & Application, November.
- [5] R. Steffan and T. Kuhlen (2001). MAESTRO – A Tool for Interactive Assembly Simulation in Virtual Environments, Proceedings of the joint IAO and EG workshop, 16-18 May, Stuttgart, Germany.
- [6] L. Marcelino, N. Murray and T. Fernando (2003). A Constraint Manager to Support Virtual Maintainability, *Computers & Graphics*, Vol. 27, No.1, 19-26, February.
- [7] E. M. Wenzel (1992). Localisation in Virtual Acoustic Displays, *Presence*, Vol.1, No.1, 80-107, Winter.
- [8] D. R. Begault (1994). *3D Sound for Virtual Reality and Multimedia*, Academic Press, Cambridge, Massachusetts, USA.
- [9] J. K. Hahn, H. Fouad, L. Gritz and L. W. Lee (1998). Integration Sounds and Motions in Virtual Environments, *Presence*, Vol.7, No.1, 67-77, February.
- [10] K. Doel, P. G. Kry and D. K. Pai (2001). Physically-based Sound Effects for Interactive Simulation and Animation, Proceedings of ACM SIGGRAPH'2001, 12-17 August, Los Angeles, CA, USA.
- [11] J. F. O'Brien, P. K. Cook and G. Essl (2001). Synthesising Sounds from Physically Based Motion, Proceedings of ACM SIGGRAPH'2001, Los Angeles, CA, USA.
- [12] Y. Kitamura, A. Yee and F. Kishino (1998). A Sophisticated Manipulation Aid in a Virtual Environment using Dynamic Constraints among Object Faces, *Presence*, Vol.7, No.5, 460-477, October.
- [13] Y. Zhang, N. Murray and T. Fernando (2003). Integration of 3D Sound Feedback into a Virtual Assembly Environment, Vol. 1 of the Proceedings of the 10<sup>th</sup> International Conference on Human-Computer Interaction (HCI International 2003), Crete, Greece, July.
- [14] Y. Zhang and T. Fernando (2003). 3D Auditory Feedback Act as Task Aid in a Virtual Assembly Environment, Proceedings of the 21<sup>st</sup> Eurographics UK Chapter Conference (EGUK 2003), Birmingham, England, IEEE Computer Society Press, June.
- [15] Y. Zhang and R. Sotudeh (2004). Evaluation of Auditory Feedback on Task Performance in a Virtual Environment, Proceedings of the 4<sup>th</sup> International Conference on Computer and Information Technology (CIT2004), Wuhan, China, IEEE Computer Society Press, September.