

Head Motion and Latency Compensation on Localization of 3D Sound in Virtual Reality

Giann-Rong Wu, Cha-Dong Duh, Ming Ouhyoung

Communication and Multimedia Lab.

Department of Computer Science and Information Engineering

National Taiwan University, Taiwan, R.O.C.

<http://www.cmlab.csie.ntu.edu.tw/~ming>

Jei-Tun Wu

Department of Psychology

National Taiwan University, Taiwan, R.O.C.

ABSTRACT

As part of designing a multi-sensory VR environment, in addition to evaluating the visual subsystem, acoustic presence must be evaluated. This paper proposes two experiments to examine the effects of human head movement and latency compensation in 3D sound localization. There are two hypotheses, the first hypothesis is that through the computer simulation of 3D sound, dynamic head movement can help in the localization of sound in space, as compared to fixed head position. The second hypothesis is when there is latency introduced in the computer generation of 3D sound, a human subject can perform better in a sound locating task with latency compensation than the one without compensation. The results of two proposed experiments corroborate with the above two hypotheses. Moreover, we are able to identify that with dynamic head movement, the human capability can be enhanced by more than 90% in the localization of sound in space and at the same time reduce the front-back ambiguity. In the second experiment involving a typical system latency at 300ms, it is shown that with compensation for latency the average time to perform sound locating task can be reduced by more than 50% than that without compensation. A priori study of the latency effects indicates that if the latency is larger than 150ms, human performance of locating a 3D sound source is noticeably decreased.

KEYWORDS: Virtual reality, 3D sound, localization of sound, latency compensation.

1. INTRODUCTION

Overall system latency, the elapsed time from input human motion to the immediate response of that input in the display, is one of the most frequently cited shortcomings of current virtual environment (VE) technology when the latency is relatively large. In a head-mounted display (HMD) based VE system with head tracker, if overall latency is longer than 200 ms, for instance, it will cause motion sickness for long time wearing.

In a multi-sensory system, the situation is more complicated if

each sensory display has different latency. For example, in our interactive building walkthrough system, there are three kinds of sensory perception (vision, 3D sound, and mechanical sensory by treadmill) combined in the system with different lags. A preliminary research on the visual part of lag problem has been examined by a 3D tracking experiment on latency and its compensation method in virtual environments in 1995. The results showed that the Grey system predictor we proposed can reduce the latency significantly. Similarly, the auditory system may be sensitive to time lags in a VE. However, there is yet no research data available that describes the relationship between localization performance of sound source and head tracker latency [1].

Although the power of a modern personal computer is enough to compute and playback 3D sound in real time purely by software, the requirement of temporal consistency between audio and visual data makes the auditory subsystem suffer from the same latency problem of visual subsystem in an immersive VE.

The new technology of 3D sound enables us to conduct some experiments on the topics we always want to do for studying the effect of combining 3D acoustic fidelity in VE [1,2,3]. For example, in the experiments on the latency effect of a walkthrough project, a user can walk around an environment wearing an HMD. However, the graphics subsystem plus the space tracker hooked onto an HMD and the LCD panel display on the HMD altogether introduce a latency of 300 ms or more [4]. Therefore, when a user was using the system for more than 5 minutes, he or she sometimes felt dizzy and thus got motion sickness because of the relative high latency. This makes us think about another problem also related to virtual reality: that in order to make the artificial environment look real and sound real, we should also introduce the 3D sound effect. What would be the effect of the latency in a 3D sound environment? Is that similar to the visual subsystem when the human perception can tell the difference and behave differently?

Besides, there is another problem we are interested in, that is, in some reports we know that motion parallax can help human beings in the perception of depth better than just wearing stereo glasses [5]. Similarly, there are also reports on 3D sound experiments in free-field conditions [6,7], which says that if a system let a human user move his or her head around in locating a sound source in space, the precision is better than that of keeping the human head fixed in space. In Pollack and Rose's study [8], low bandwidth and high bandwidth thermal noise sound with short duration were used in these experiment, and an improvement of 10-15% was observed. Therefore, from the results,

Permission to make digital/hard copies of all or part of this material for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copyright is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires specific permission and/or fee

ACM VRST '97 Lausanne Switzerland
Copyright 1997 ACM 0-89791-953-x/97/9..\$3.50

we know that dynamic head movement is better for locating sound, but how much better in a computer generated 3D sound system?

In this paper, our first experiment examines the benefits of head movements in sound localization, and if so, by what range in degrees. The hypothesis of the first experiment is that dynamic head movement can improve the localization of sound in space as compared to that of static head position. Our second experiment focuses on the benefits of reducing the latency in the audio signal that is typically introduced in virtual environments. The hypothesis of the second experiment is that when a human subject is performing a real task and there is latency involved in the system, one's capability would be lowered. Since we can use a computer to compensate for latency using prediction algorithms, in a system where there is significant latency, the hypothesis becomes that with latency compensation, human subjects can perform better than that without.

In the following, we will introduce the implementation of our software based 3D sound system used in the above two experiments. Experienced readers in 3D sound can skip most of Section 2.

2. 3D SOUND SYSTEM IMPLEMENTATION

3D sound generally means that a listener hears sounds in all directions, where the sound is simulated by a computer. For a headphone based 3D sound system, the system should be able to place sounds outside one's head, as well as to the listener's front and rear. In the following, we will introduce the implementation of our headphone based 3D sound system used in our two experiments. This is a real-time software based 3D sound generation system.

Consider the spatial environment as a digital system, one can generate perceptible audio by means of digital signal processing [9,10,11]. If one treat the free-field spatial environment as a linear, time-invariant digital system, one can understand the behaviour of sound in space by its impulse response. Basically, the spatial hearing environment has several factors that contribute to the result of its impulse response: the azimuth, the elevation, and the position of the sound source, including the distance between the sound source and the listener, all can affect the value of the impulse response. Furthermore, some factors are human related, such as the shape of listener's pinnae and canals, the size of the listener's head, the height of the listener's nose, etc. From measurements of the impulse response, one can get the head-related transform functions (HRTF). If the HRTFs are available, the 3D sound effects can be produced from a linear convolution between sound signals and its corresponding spatial related impulse response [12,13]. However, the linear convolution is a time consuming computation. In some virtual environment systems, like NASA's VIEW system, real-time convolution is handled by a DSP chip. Because of that PCs' computation capability becomes more and more powerful, it is possible to generate CD quality 3D audio purely in software on a PC. Our system is consisted of a PC with a Intel Pentium-133MHz processor, a Creative Lab Sound Blaster audio card, and a Pro.2 headphone. By optimization of direct computation of convolution, we can achieve real time performance. In our design of simulating 3D sound in VEs, a space tracker is hooked to a headphone for reporting the head's position and orientation. Depending on the movement of the space tracker hooked on one's head, the 3D audio player can select the nearest impulse response and generate 3D sound accordingly.

The HRTFs used in our system are impulse responses obtained from the MIT Media Lab. The compacted data are equalized with speaker's impulse response and packed in stereo. The MIT Media Lab proposed a set of HRTFs measured by Bill Gardner and Keith Martin in 1994 [14]. They used a dummy head model, KEMAR mannequin head, as the listener's head. A probe microphone is attached to the position of eardrum, and can record the sound from outside. The impulse response is computed with maximum length sequence (MLS) technique. A speaker is mounted 1.4 meters from the KEMAR model. They measured in total 710 different positions at a sampling rate of 44.1KHz with two types of pinnae, ranging from the elevation from -40 degrees to 90 degrees and the azimuth from 0 degree to 360 degrees. Each impulse response is 128-point length.

Then we have to equalize the compact data by an inverse filter of the headphone. Because the original compact data used ear canal resonance, we have to remove the ear canal resonance to avoid "double resonance" [15]. An inverse filter of the headphone can be used to eliminate the effect of the superfluous canal resonance. Gardner and Martin provided a set of measured impulse response for several combinations of the headphone and pinnae. We choose the AKG-K240 headphone with normal pinnae as our target pairs and compute its inverse filter to equalize the impulse response.

The playback system reproduces the 3D audio output from a mono-audio source by using the 128 point linear convolution with the specified impulse responses (one for left ear and the other for right ear) continuously. The mono-audio source is based on pulse code modulation (PCM). During playback of 3D sound, changing the location of sound source is necessary for our experiments. Moving head in 3D sound environment should change the pair of impulse responses to simulate the correspondence of the newly sound source position related to head orientation. When changing impulse response, there will be some power gap between two impulse response which cause some clicks in the playback. Applying interpolation can eliminate most of the effect of click and make playback more smoothly.

3. EXPERIMENT 1: HEAD FIXED AND HEAD MOVEMENT OF 3D SOUND

The goal of the experiment is to verify whether dynamic movement of human head can really improve the localization of sound in space. The difference between this experiment and the previous experiments in free-field conditions is that our sound source is simulated by a computer. The hypothesis of this experiment is that dynamic head movement can improve the localization of sound in space as compared to that of static head position, here called *Head Movement Hypothesis*.

3.1 EXPERIMENTAL DESIGN

Assume that the sound source is fixed and continuous in space, the subjects must point out the direction of sound by a pointer either in the case of fixing one's head or allowed to rotate one's head. For tracking head movements, a magnetic space tracker ("Flock of Bird" from Ascension technology Co.) was introduced to report the head's position and orientation. Figure 1 shows a subject points out the location of sound in space. The degree of azimuth pointed is measured by referring to a large compass with the projected line of the pointer on the ground.



Figure 1. A photo of our experimental set-up for the Experiment 1.

In the set of impulse responses, the higher elevation has lower precision. From [16] point of view, we know that minimum audible movement angle (MAMA) at elevation 0 degree is smaller than that of another elevation. Therefore, to be precise, we choose elevation at zero to conduct the following two experiments to preserve higher precision [17].

For the motion based experiment, the static precision of sound source play an important role in conducting the experiments. That is, the position offset between the fixed sound source and the subject's head must be preserved as the same for getting the stable pair of (evaluation, azimuth). However, there is no guarantee that head position is fixed during rotation. The solution is to modify the location of sound source with respect to an offset of the difference of current head tracker's position and the initial tracker's position.

With different pinnae among subjects, in the ideal case, it is better to have different HRTFs for each subject. However, the impulse responses we got were reported from the standard KAMAR mannequin head, that's the reason why we chose this for later experiments.

3.2 PROCEDURE

In the experiment, we invited 10 volunteer subjects to participate in our experiments. All subjects were able to hear 3D sound. Each subject was given two sessions of experiments, the first one required that the subject's head be fixed in space and facing the front. The second allowed the human subject move around the space as he/she wished, of course when one was moving one's head around, a headphone set with space tracker is fixed on one's head. The sound source was always fixed in space even though the subject can turn his head around. Within each session, 12 directions were selected in random order. All directions were generated in advance. For one-half of subjects, the experimental order was keeping the head fixed first and then allowed dynamic head movement, and for the others, the order was reversed. Each subject was trained to be familiar with the 3D sound perception for about 5 minutes before the experiment.

3.3 RESULTS

Table 1 shows the localization errors, which are the average error in degrees within twelve trials, of ten subjects. Note that the front-

back ambiguity of the fixed head part of localization errors has been eliminated in advance. The results corroborate with our *Head Movement Hypothesis*, $t(9)=4.28, p<.01$.

Table 1. Localization errors (degrees) of ten subjects of the sound locating experiment.

Subject	fixed head location	dynamic head movement
S1	18.333	6.917
S2	16.417	13.083
S3	22.833	12.667
S4	16.667	10.833
S5	8.333	2.500
S6	22.500	7.500
S7	17.500	9.583
S8	13.333	4.583
S9	19.583	15.833
S10	24.416	11.083
Mean	17.992	9.458
* $t(9)=4.28, p<.01$		

By examining the data for dynamic head movement, the localization of sound in space is more precise, that is, the average error in dynamic head movement is 9.458 degrees while if the head is fixed in space, the error is about 18 degrees. As a result, we are able to identify that with dynamic head movement, the human capability is enhanced by more than 90% in locating sound in space.

Finally, there are more so called front-back ambiguity when the human head is fixed in space, since that's the case when one can not identify whether the sound locates in one's front or in one's back. When the head is fixed in space, on average, there are three front-back errors out of twelve tests caused by confusion, however, in the case of dynamic head movement, there is no such error. As a brief comment, the result shows that dynamic head movement can not only help in the precision of locating sound in space, but also help to reduce the front-back ambiguity.

4. EXPERIMENT 2: LATENCY AND ITS COMPENSATION OF 3D SOUND

Considering a typical architecture walkthrough system that can generate 3D sound as well as graphical objects, since the simulated sound has to be synchronized with the graphics subsystem, whenever there is latency in the graphics subsystem, there is equal amount of latency introduced to the 3D sound system. Therefore, since there is usually a latency of 300ms in a typical walkthrough system, we will use the same latency (300 ms) in the later 3D sound experiment.

The hypothesis of the experiment is that when there introduces relatively large latency in a VE, one's capability of locating sound source would be significantly reduced. Since we can use a computer to compensate for latency based on prediction algorithms, the hypothesis becomes that with latency compensation human beings can perform better than those without, and we called it *Latency Compensation Hypothesis*.

Before conducting the experiment, there is a question to be answered which is "What's the largest overall system latency that

Table 2. A table of mean difference between any pair of two cases. The symbol “*” indicates that the difference of two means is greater than the critical value set by Tukey’s HSD.

	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5
Condition 1 (Mean = 1.55)	-	0.4	0.41	1.01	2.12*
Condition 2 (Mean = 1.95)		-	0.01	0.61	1.71*
Condition 3 (Mean = 1.96)			-	0.60	1.70*
Condition 4 (Mean = 2.56)				-	1.11*
Condition 5 (Mean = 3.66)					-
Tukey’s HSD. $p < .05$; Critical Value=1.017					

will not affect the perception of 3D sound localization?” The answer is very important since it can help a VE’s designer to decide whether his/her system needs to deal with the latency problem when using 3D sound.

4.1 A PRIORI STUDY ON LATENCY AND RECOGNITION OF 3D SOUND

In the priori experiment, we would like to find the mythical threshold in the latency just mentioned above. There were five conditions of VEs indicating different latency lengths as follows.

- Condition 1: 0 ms, no latency included for reference.
- Condition 2: 50 ms latency included.
- Condition 3: 100 ms latency included.
- Condition 4: 150 ms latency included.
- Condition 5: 200 ms latency included.

A sound source tracing task is designed for the priori experiment, and its experimental design is described in Section 4.3 in detail. Five subjects took part in the priori experiment. Each subject accepted all of 5 conditions with sequential order (counter-balanced between subjects) For each condition, randomly selected sound source was put in space selected in random order. When the testing music (used as sound source) started, the subject was required to find the location of sound source by turning his/her head as fast as possible so that the sound source appears exactly in front of him/her. There were in total 15 trials examined and the corresponding reaction times of location used were reported. Every subject had a training for about 5 minutes.

The reaction time of 3D localization were tested by a one-way within-subject analysis of variance (ANOVA). The result indicated that there exists difference of reaction time of 3D localization among those five conditions, $F(4,16)=10.8$, $MSe=0.315$, $p<.01$. That is, among ten pairs of these conditions, there is at least one pair that is significantly different.

A multiple comparison test, Tukey’s HSD, was applied to find whether there is significant difference among five conditions. Table 2 shows the results of mean difference between any pair of conditions. When latency is less than 150 ms, there is no significant difference. However, when the latency reaches 200 ms, it showed that there exists difference between condition of 200 ms and any one of conditions with latency less than 150 ms and thus indicates that the performance in a VE with 200 ms latency is significantly worse than the one without latency in locating a sound in space. From our observation, if the latency is larger than 150 ms, the performance of localization of 3D sound is noticeably decreased. Since a typical application such as our architecture walkthrough system suffers from 300 ms latency, finding a way to compensate the latency is obviously necessary for improving the precision of locating a sound in space.

4.2 PREDICTION ALGORITHMS FOR LATENCY COMPENSATION

3D sound latency is the time delay between head movement and its corresponding motion of virtual sound source played on the headphone. To compensate for latency, many proposed methods used prediction in tracking. Several HMD systems have been implemented with head tracker prediction [18,19,20,21], where a “look ahead” algorithm is implemented which uses the 3D position and orientation as the input data.

Figure 2 shows the system diagram of a general prediction system. At each time t , a system behaviour formula can be generated from applying the historical data sequence X (the nearest data observed in time domain) to the prediction algorithm. In Figure 2, the number of historical data is set to i , and therefore the system will use i observed data as the inputs. Applying a specified prediction length to the system behaviour formula, a new predicted data at time t , P_t , will appear.

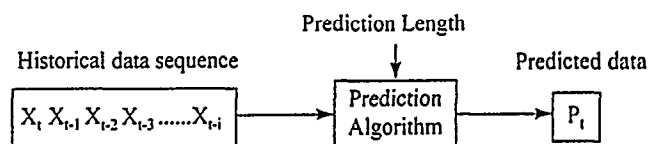


Figure 2. A system diagram of a general prediction system.

There are two useful prediction algorithms, Kalman filtering and Grey system theory, both of which have been evaluated to be almost the same in real task experiments on visual perception [22]. The Grey system based prediction algorithm is used in our experiment because of its lower computation complexity than that of the Kalman filtering based prediction. For detail information of the above two prediction algorithms, please refer to the original papers [4,18,19,20,21].

4.3 EXPERIMENTAL DESIGN

A tracing task was designed to evaluate the effects of latency on locating sound in space, where the sound source was randomly generated in space during the task. When a subject hears 3D sound in space, he/she is required to trace the 3D sound source by facing to the sound source, and when the subject was certain that he/she is exactly facing the sound source, he/she can press the left button of a mouse to signal his/her decision. If the azimuth degrees between head orientation and the sound source was close enough (under a given threshold of ± 5 degrees), the playback of 3D sound stopped. This means that the subject has finished the

tracing task. However, if the azimuth difference in degree exceeds a given threshold, the 3D sound played continuously. The localization process continued until the 3D sound stopped. For example, in Figure 3, assuming that the 3D sound source is fixed in azimuth angle of 300 degrees, the task will be finished if the subject faces between 295 to 305 degree and presses button to confirm his/her decision, otherwise, the 3D sound plays continuously.

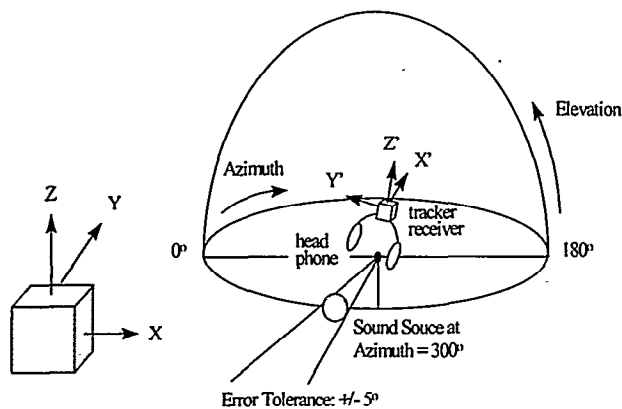
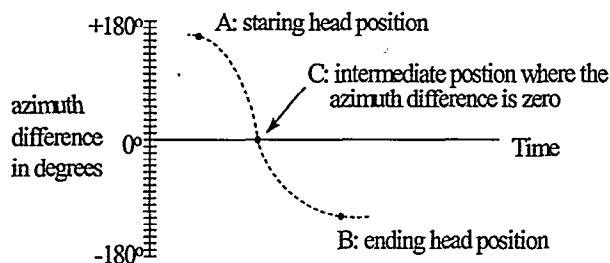


Figure 3. The tracing task for localization of sound in azimuth angle.

Task completion time spent in the tracing task was considered as the key parameter to evaluate the effects of different values of latency. From our previous experience on the visual subsystem, the larger latency introduced in a VE, the more time it took to finish a task [22].

On the other hand, average error distance may be considered to be a good indicator to evaluate the latency effect. However, there is a problem caused from the case that one can move his head slower and spend more time to locate a sound source more precisely even if large latency is included. Therefore, we still analyze task completion time in our experiments.

The reason that a subject has to explicitly press a button to signal his/her possible completion can be explained in the following figure.



Since head swing from one direction (point A) to the opposite direction (point B) is continuous, there exists a point C in the path that will have a zero value in the azimuth difference. However, this does not mean that a subject has exactly located the sound source, since he/she is just swinging from one direction to the other trying to locate the sound source, and so this intermediate zero value is meaningless.

According to our observation from Experiment 1, a useful heuristic can be used: when one tries to localize a sound source, one tends to move one's head to minimize the intramural difference. That is, when one hears sound in perfect balance, one

can identify the sound source in front of him more precisely. That's the reason why we used the technique of facing the sound source in the localization task of our experiments.

4.4 PROCEDURE

Ten subjects were invited to participate in the experiment. During the experiment, subjects were instructed to face the sound source by turning their head as fast as possible. Each subject was given two separate levels in random order, one is with latency compensation and the other is without latency compensation, where the latency is 300 ms. We designed six sets of cases with different initial sound source location, and the average task completion time is used for each subject. Since each test also involves two conditions, i.e., with and without latency compensation, each subject actually took 12 trials in total. The training phase was the same as in Experiment 1.

4.5 RESULTS

Table 3 shows the task completion time used, which is the average time spent in second within twelve trials, by ten subjects. The result corroborates with our *Latency Compensation Hypothesis*, $t(9)=3.439, p<.01$.

Table 3. Task completion time spent (seconds) of ten subjects of the experiment.

Subject	without latency compensation	with latency compensation
S1	6.81	3.24
S2	5.49	2.58
S3	10.49	5.19
S4	11.15	5.60
S5	8.79	2.60
S6	7.03	2.58
S7	3.03	1.65
S8	4.01	2.86
S9	4.10	1.54
S10	4.17	2.90
Mean	6.507	3.074

* $t(9)=3.439, p<.01$

On the average, the time to localize a target sound location in space with latency compensation is 3.074 seconds, and the time without latency compensation is about 6.507 seconds. Therefore, the average task completion time in sound localization is about 50% shorter when latency compensation technique is used.

Similarly, with latency compensation, on average, the number of times of pressing a button to signal the completion of the task is two, while without latency compensation, it is 4.5. That is a significant improvement in human computer interaction. For the button pressing part, please refer to Section 4.3 for detailed explanation. This indicates that with the same prediction algorithm as used in a visual subsystem, compensation for latency in the 3D sound system can also significantly improve human localization of 3D sound. That is, the prediction algorithms such as those based on Grey system or Kalman filtering not only can reduce the latency in HMD but also can improve the localization of 3D sound.

5. CONCLUSION AND FUTURE WORK

We have conducted two experiments to examine two hypotheses. The first hypothesis is that localization of sound in space with head movement is easier than that of keeping head fixed. The second hypothesis is that in a virtual reality application with 3D sound capability, with latency compensation, a subject in the task of locating 3D sound performs better than that without latency compensation. Two conducted experiments corroborate with our hypotheses. Note that the results of the second experiment are original. Moreover, a priori study of the latency effects indicates that if the latency is larger than 150ms, 3D sound localization performance is noticeably decreased. We have therefore established the similarity between computer graphics and 3D sound, namely, the effects of motion parallax and latency compensation.

ACKNOWLEDGEMENT

We would like to thank the MIT Media Lab's HRTF measurements, which they have put in WWW for public use. We thank detailed comments from anonymous reviewers when we first submitted this paper to SIGCHI'97. This project is partially supported by the grant NSC 86-2213-E-002-044.

REFERENCE

1. Kendall, G. S., "A 3-D Sound Primer: Directional Hearing and Stereo Reproduction," *Computer Music Journal*, Vol. 19, No. 4, pp.23-46, Winter 1995.
2. Hartmann, W. M., "Localization of Sound in Rooms," *Journal of the Acoustical Society of America*, Vol. 74, No. 5, pp.1380-1391, November 1983.
3. Middlebrooks, J. C., Green, D. M., "Sound Localization by Human Listeners," *Annual Review of Psychology*, Vol. 42, pp.135-159, 1991.
4. Wu, J.-R., Lei, Y.-W., Chen, B.-Y., Ouhyoung, M., "User Interface Issues for a Building Walkthrough System with Motion Prediction", *Proc. of IEEE 1996 International Conference on Consumer Electronics*, pp. 378-379, Chicago, 1996.
5. Kalawsky, R. S., *The Science of Virtual Reality and Virtual Environments*, Addison-Wesley, 1993.
6. Wallach H., "The Role of Head Movements and Vestibular and Visual Cues in Sound Localization," *Journal of Experimental Psychology*, Vol.27, pp.339-368, 1940.
7. Thurlow, W. R., Mangels, J. W., and Runge, P. S., "Head Movements During Sound Localization," *The Journal of the Acoustical society of America*, Vol.42, pp. 489-493, 1967.
8. Pollack, I., Rose, M., "Effect of Head Movement on the Localization of Sounds in the Equatorial Plane," *Precept. Psychophys*, Vol. 2, pp.591-596, 1967.
9. Rodgers, C. A. P., "Pinna Transformation and Sound Reproduction," *Journal of the Audio Engineering Society*, Vol. 29, pp. 226-234, 1981.
10. Morimoto, M., Ando, Y., "On the Simulation of Sound Localization," In R. W. Gatehouse (Ed.), *Localization of Sound: Theory and Applications*. Groton, CT: Amphora Press, 1982.
11. Wightman, F. L., Kistler, D. J., "Headphone Simulation of Free-field Listening I: Stimulus Synthesis," *Journal of the Acoustical Society of America*, Vol. 85, pp. 858-867, February 1989.
12. Fisher, H., Freedman, S. J., "The Role of the Pinnae in Auditory Localization," *Journal of Auditory Research*, Vol. 8, pp15-26, 1968.
13. Wenzel, E. M., Wightman, F. L., Foster, S. H., "A Virtual Display System for Conveying Three-dimensional Acoustic Information," In *32nd Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica: Human Factors and Ergonomics Society, 1988.
14. Gardner, B., Martin, K., "HRTF Measurements of a KEMAR Dummy-Head Microphone," *MIT Media Lab Perceptual Computing Technical Report #280*, May 1994.
15. Begault, D. R., "3-D Sound for Virtual Reality and Multimedia," *Academic Press*, 1994.
16. Strybel, T. Z., Manligas, C. L., Perrott, D. R., "Minimum Audible Movement Angle as a Function of the Azimuth and Elevation of the Source," *Human Factors*, Vol. 34, pp. 267-275, 1992.
17. Grantham, D. W., "Detection and Discrimination of Simulated Motion of Auditory Targets in the Horizontal Plane," *Journal of the Acoustical Society of America*, Vol. 79, No. 6, pp.1939-1949, June 1986.
18. Liang, J., Shaw, C., Green, M., "On Temporal-spatial Realism in the Virtual Reality Environment," *Proc. 4th Annual Symposium on User Interface Software and Technology*, Hilton Head SC, pp. 19-25, 1991.
19. Azuma, R. and Bishop, G., "Improving Static and Dynamic Registration in an Optical See-through HMD," *SIGGRAPH'94 Conference Proceedings*, pp. 197-204, 1994.
20. Mazuryk, T. And Gervautz M., "Two-Step Prediction and Image Deflection for Exact Head Tracking in Virtual Environments," *Computer Graphics Forum (Eurographics '95)*, Vol. 14, No. 3, pp. c30-c41, 1995.
21. Wu, J.-R., Ouhyoung, M., "Reducing The Latency In Head-Mounted Displays By a Novel Prediction Method Using Grey System Theory," *Computer Graphics Forum (EuroGraphics'94)*, Vol. 13, No. 3, pp. c503-c512, 1994.
22. Wu, J.-R., Ouhyoung, M., "A 3D Tracking Experiment on Latency and Its Compensation Methods in Virtual Environments", *Proc. of UIST'95 (User Interface and Software Technology 1995)*, pp. 41-49, ACM Press, Pittsburgh, 1995.