

The impact of system latency on dynamic performance in virtual acoustic environments

Elizabeth M. Wenzel

*Human Information Processing Research Branch, MS 262-2, NASA Ames Research Center,
Moffett Field, CA 94035-1000.*

Abstract. Engineering constraints that may be encountered when implementing interactive virtual acoustic environments are examined. In particular, system parameters such as the update rate and total system latency are defined and the impact they may have on perception is discussed. A testbed for measuring total system latency is described and some preliminary measurements for an anechoic simulation are summarized.

INTRODUCTION

In a virtual acoustic environment (VAE), the total system latency (TSL) and the effective update rate are distinct parameters although they may be related in practice. The total system latency, or end-to-end latency, refers to the time elapsed from the transduction of an event or action, such as movement of the head, until the consequences of that action cause the equivalent change in the virtual sound source location. Latencies are contributed by individual components of a VAE system, including tracking devices, signal processors, software to control these devices, and communications lines. TSL differs from the "internal latency" (1, 2) of each system component; e.g., in a spatialization device, the internal latency is the delay between acquisition of location data and the rendered audio output. Update periods (period = 1/rate) in a VAE system refer to various sampling or rendering intervals which may be present in the system; for example, the time elapsed between successive samples of the listener's head motion and the time elapsed between calculation of one spatial location and a new spatial location by a spatialization engine (i.e., 1/frame rate). Due to differences in sampling rates, the effective update rate usually corresponds to the update rate of the slowest component in a VAE system. As with latency, there is no reason to expect that a system's update rate remains constant over time. Thus, measurements of the mean, standard deviation, and range of the TSL and update rate provide a better characterization of these parameters.

A previous examination (3) of performance parameters reported for VAEs suggested that commonly-specified parameters such as the audio update rate determine only the "best-case" latency possible in a VAE, and that TSL and internal latencies are frequently not measured by VAE developers. Since such parameters are critical for characterizing the dynamic performance of VAEs, it is important that they be carefully defined and measured. Psychoacoustic parameters such as the minimum audible movement angle (MAMA) can then be used as target guidelines to assess whether a given system meets perceptual requirements. In addition, such measurements enable systematic perceptual studies of the impact of degrading system latency or update rate on localization behavior.

LATENCY MEASUREMENTS

This paper reports on measurements of TSL for the virtual audio system used in previous studies of localization with and without head motion (3, 4). The system consisted of a Convolvotron spatialization system using an anechoic model with 256-pt., minimum-phase head-related transfer functions (HRTFs) interpolated between the 4 nearest map locations. The Convolvotron's specifications state an update rate of 33 Hz and latency of 32 msec. It received head-position data from a Polhemus Fastrak (set up for 3 sources, 40 Hz update rate) over a 19.2 kBaud serial line. The host computer was a 90 MHz Pentium running DOS.

In order to measure latency, a special HRTF map was constructed which contained a single impulse at one map location and zeroes at all other locations; otherwise, the computational load was the same as in (3, 4). Latency measurements were conducted using a testbed (Figure 1) adapted from (1, 2). The Fastrak receiver was mounted on the end of a mechanical swing arm with an optical switch which detected when the arm passed through a pre-set threshold position. This event threshold was considered to be analogous to the initiation of head (and/or source) motion in a VAE and began the universal counter's TSL timing cycle. At the same time, the tracker sent data to the Convolvotron in polled mode via a serial line, and a signal generator fed a 6000 Hz square wave to one of the input A/D channels. Before the threshold was crossed, the Convolvotron was set to a zeroed map location so that no signal passed through the D/A. The experimenter then pushed the swing arm through the threshold position. The next tracker data sample received after threshold-crossing caused the Convolvotron to switch to the non-zero map location. The square wave then propagated through the system and terminated the timing cycle of the universal

counter whose reading was considered to be the TSL. 117 such measurements were taken and are summarized in Table 1. Measurements of tracker latency and the Convolvotron's internal latency were inferred via software which registered when the first tracker sample past threshold occurred.

TABLE 1. Component & total system latencies (msec).

	Tracker	Serial Line	Convolvotron	TSL
Mean	16.2	15.1	23.0	54.3
St. Dev.	7.0	0	5.2	8.8
Min	4.03	15.1	14.2	35.4
Max	28.7	15.1	32.0	74.6

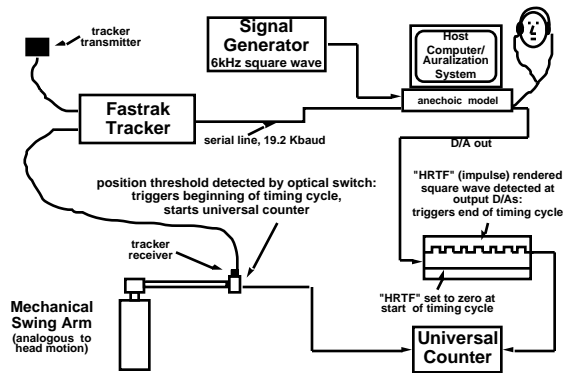


FIGURE 1. Testbed for measuring total system latency in a virtual audio system.

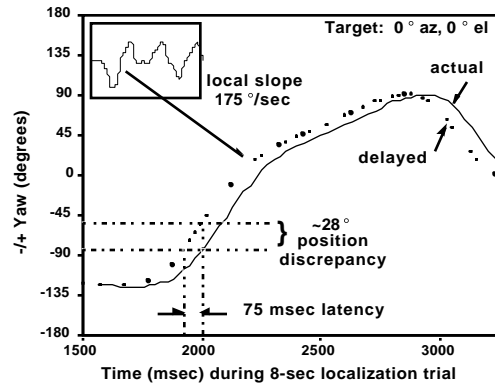


FIGURE 2. Illustration of position displacement due to latency.

Examination of the head motions that listeners used to aid localization in (3, 4) suggests that the angular velocity of some head motions (in particular, left-right yaw) may be as fast as about $175^\circ/\text{sec}$ for short time periods (e.g., about 1200 msec). A maximum TSL of 75 msec could potentially result in short-term under-sampling of relative listener-source motion as well as positional instability of the simulated source. For example, in Figure 2, head motion yaw for an individual subject is plotted as a function of time (undelayed and delayed by 75 msec) during localization of a virtual source fixed at $0^\circ, 0^\circ$. The inset shows the entire 8-sec trial. In regions of the head-motion trace where angular head motion is large (local slope = $175^\circ/\text{sec}$), a TSL of 75 msec could result in a relative position discrepancy of about 28° between actual head orientation and the rendered direction of the source. Psychophysical studies of the MAMA (6) for real sound sources (listener position fixed) indicate that angular source velocities of $90^\circ, 180^\circ,$ and $360^\circ/\text{sec}$ result in threshold MAMAs of $8.3^\circ, 12.5^\circ,$ and 21.2° , respectively. From these data, one can infer that the minimum perceptible TSL for a virtual audio system should be no more than about 92, 69, and 59 msec for the slow, moderate, and fast source velocities, respectively. If one assumes that these thresholds apply to relative source-listener motion in general (e.g., when the source is fixed and the listener is moving), then the positional displacement of the simulated source due to TSL in (3, 4) may have occasionally exceeded the perceptible threshold. In fact, listeners did not report any obvious instability in source position.

ACKNOWLEDGEMENTS

Supported by NASA and by the Naval Command, Control and Ocean Surveillance Center, San Diego. Thanks to Mark Anderson and Joel Miller for implementing the measurement testbed and software.

REFERENCES

- Adelstein, B. D., Johnston, E. R., and Ellis, S. R. *Presence: Teleoperators & Virtual Environments*, **5**, 302-318 (1996).
- Jacoby, R. H., Adelstein, B. D., and Ellis, S. R. "Improved temporal response in virtual environments through system hardware and software." *Proceedings of SPIE: Stereoscopic Displays & Virtual Reality Systems*, San Jose, CA, 1996.
- Wenzel, E. M. "Analysis of the role of update rate and system latency in interactive virtual acoustic environments" *103rd Convention of the Audio Engineering Society*, New York, NY, 1997, preprint 4633.
- Wenzel, E. M. "Effectiveness of interaural delays alone as cues during dynamic sound localization." *Journal of the Acoustical Society of America*, **100**, 2608 (1996).
- Wenzel, E. M. "The relative contribution of interaural time and magnitude cues to dynamic sound localization." *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio & Acoustics*, New Paltz, NY, 1995.
- Perrott, D. and Musicant, A. *Journal of the Acoustical Society of America*, **62**, 1463-1466 (1977).