Visualization of a Clustering Algorithm of Sound Sources based on Localization Errors

Jens Herder
Spatial Media Group
University of Aizu
Fukushima-ken 965-8580
Japan
voice: [+81](242)37-2537; fax: [+81](242)37-2549
e·mail: herder@u-aizu.ac.jp
www: http://www.u-aizu.ac.jp/~herder

Abstract

A module for soundscape monitoring and visualizing resource management processes was extended for presenting clusters, generated by a novel sound source clustering algorithm. This algorithm groups multiple sound sources together into a single representative source, considering localization errors depending on listener orientation. Localization errors are visualized for each cluster using resolution cones. Visualization is done in runtime and allows understanding and evaluation of the clustering algorithm.

Keywords: Visualization, sound spatialization, resource management, audio rendering, clustering, and human perception

1   Introduction

The spatialization resource visualizer is an inspector for sound objects in the soundscape, a visual debugger for sound objects in virtual reality environments. These are sound sources and sinks, generalization of listener and microphone. Figure 1 shows on the left side a test scene with the corresponding sound objects in the visualizer on the right side. A special case involves virtual sound sources which do not exist in the virtual reality environment scene and are generated during the clustering process of the spatialization resource manager. The set of sound sources can be selected (using the preference menu) to focus on all, active, virtual, or ambient sound sources. A sound source can be displayed using its core range, which is an ellipsoid representing a zone with maximum intensity [BCM97], and its audible range, shown as a translucent ellipsoid representing the space in which the source is audible. Between the two ranges the intensity drops off according to the square of the distance. Priority and intensity of the sound node may be included as text values facing the user. Different states of a sound node can be conveyed using color codes for the core range.

The novel clustering algorithm, which groups sound sources together and represents them using one representative sound source in case the localization error is smaller than predefined limits depending on listener configuration relative to all sound sources is described in [Her99a][Her99b]. The limits are based on average data gathered in listening experiments [MM90].

2   Visualization

The spatialization resource visualizer [Her98b] is used to visualize the clustering process. The application is the Helical Keyboard [HC96], but this time the sources are activated interactively. The number of spatialization channels is restricted to two. Figure 2 shows 88 sound sources and one sound sink.

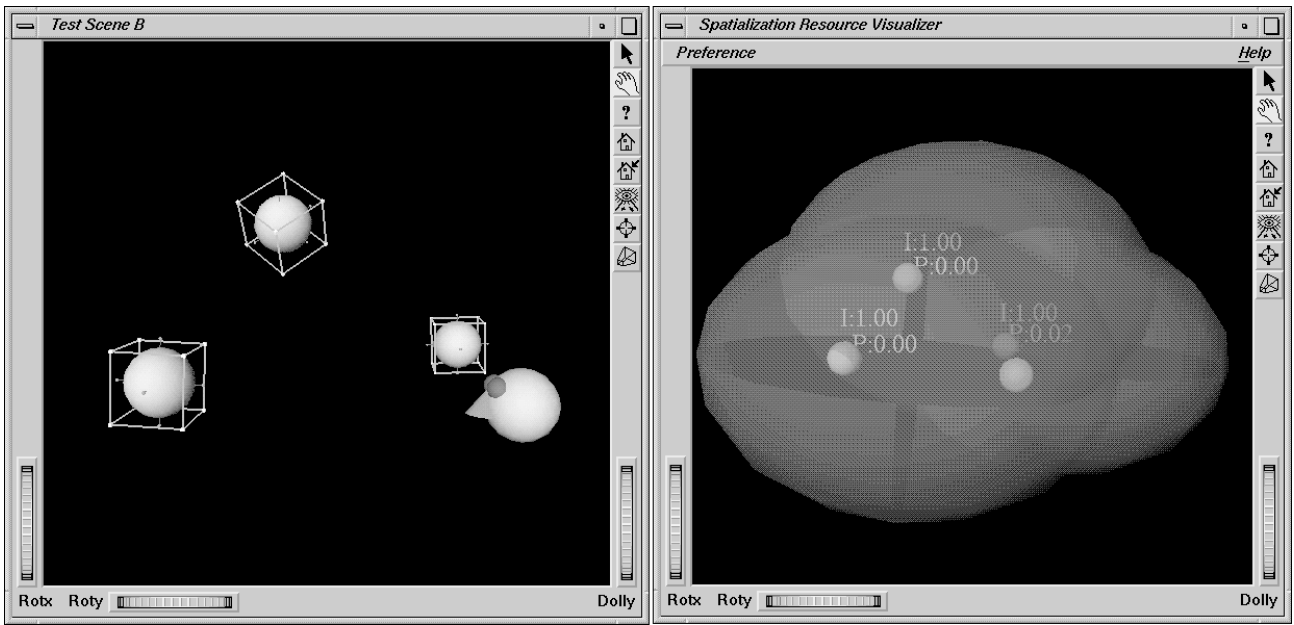Active sound sources are highlighted. The

Figure 1: Test scene with sound source visualization: left side shows a virtual environment with graphical objects; right side shows the visualization of the soundscape with three sound sources and one sound sink for that environment, the translucent spheres represent the audible ranges of the sound sources, priority and intensity of the sound sources are given as text values

sound sink is presented with an up vector, a front vector, and a lateral axis, orthogonal to both vectors. This shows the orientation clearly. Additional labels L and R mark the left and right sides from the sink perspective. Two sound sources are active and highlighted in Figure 2. The clustering process starts only when the limit of available spatialization channels is hit. As long enough resources are available, no clustering is required.

In Figure 3 three sound sources are requested. The clustering process starts, finding two clusters, one representing two sound sources. The other contains only one sound source. The representative sound sources of each cluster are highlighted. A small tether associates the sound sources in each cluster with their representative sound source. The sound sources are within the resolution cone of the representative sound source. The resolution cone shape varies depending on azimuth and elevation. The ellipses axes of the cones denote error in azimuth and elevation.

Four sound sources are requested in Figure 4. Two clusters, each with two sound sources, are formed.

Rotation of the sound sink changes the cluster allocation, as shown in Figure 5, redistributing the
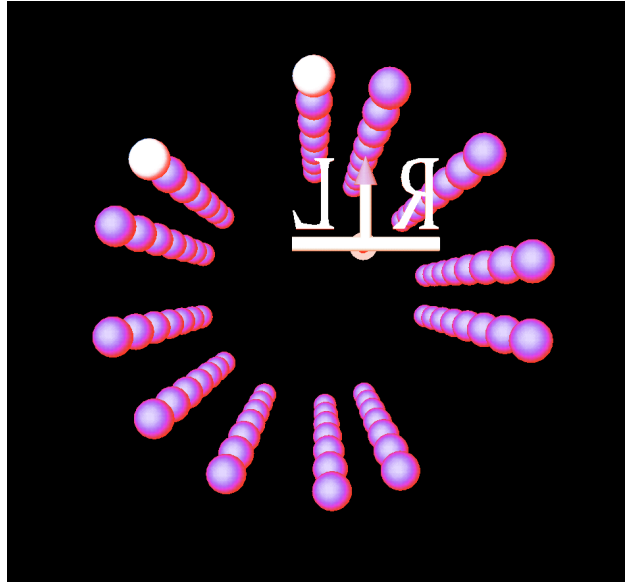

Figure 2: Two sound source are requested

four sources into a single cluster. The resolution cones are listener orientation-dependent. To the side of a listener, localization errors in elevation are lower and localization errors in azimuth higher (see also [Her99a]).
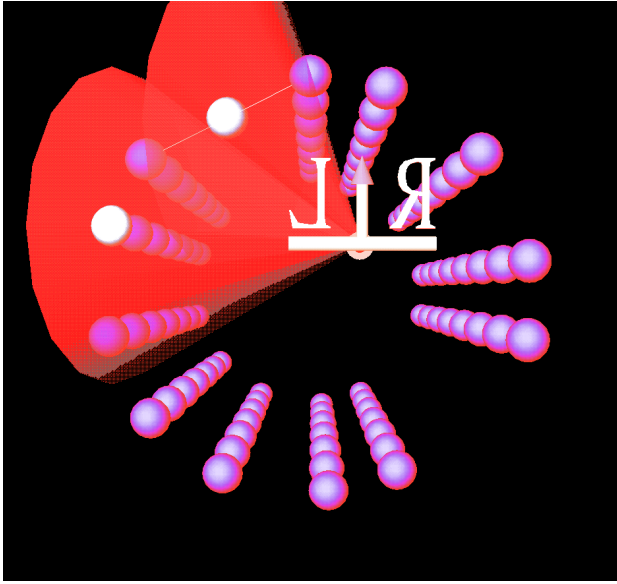

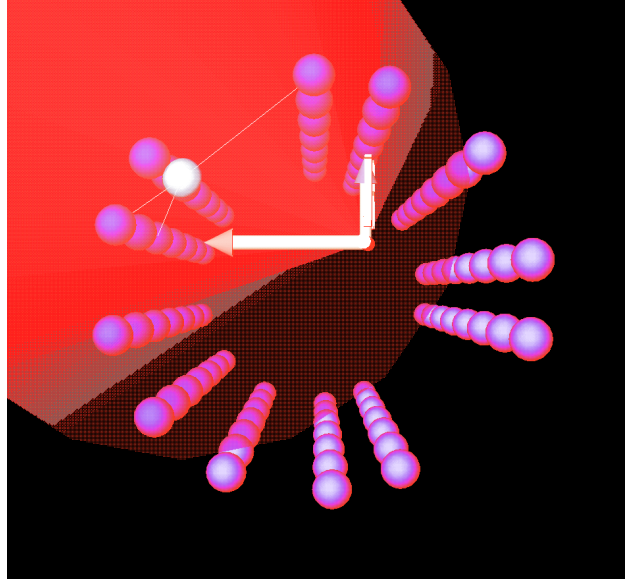Figure 3: Three sound sources are requested; clustering algorithm becomes active

Figure 5: Rotation of the sound sink changes the cluster allocation
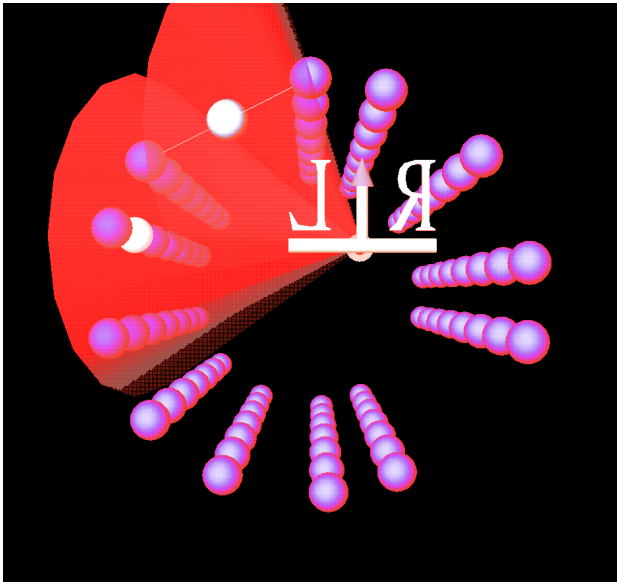

Figure 4: Four sound sources are requested

Moving the sound sink closer to the sources in Figure 6 again changes the cluster distribution. Three sound sources on the left are in one cluster. Localization errors in azimuth to the front are small, so that the sound source in the upper front direction cannot be grouped with the other requested sound sources.

Moving the sink closer to the sources changes the number of sound sources which can be clustered. In Figure 7, the clusters get split up after further movement of the sink close to the sound sources. This shows that in the near field of the head, clustering does not reduce the number of required spatialization channels.

In Figure 8, three requested sound sources are passed. One active sound source is still in front of the sink. Two sound sources which are now in the back of the sink get clustered. One cluster must be processed ambiently.

The side view in Figure 9 shows that the resolution cones to the back of the listener are larger.
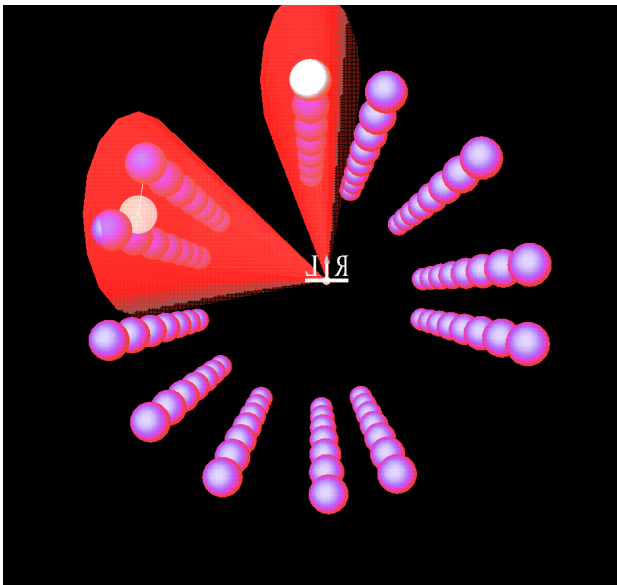
Figure 6: Moving closer with the sound sink to all sound sources
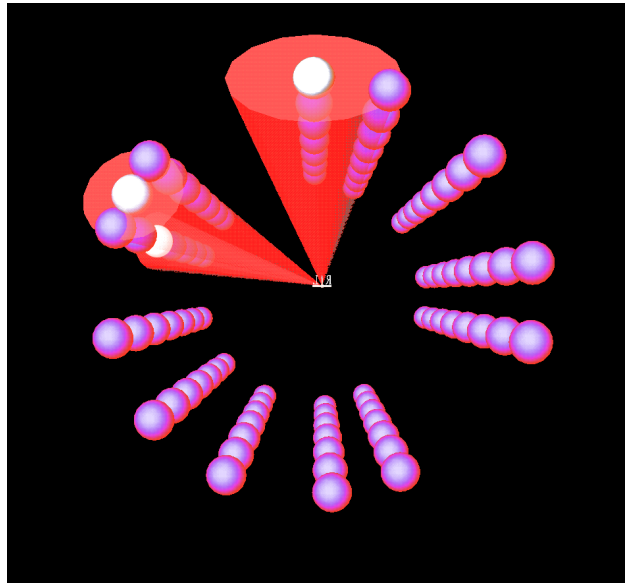

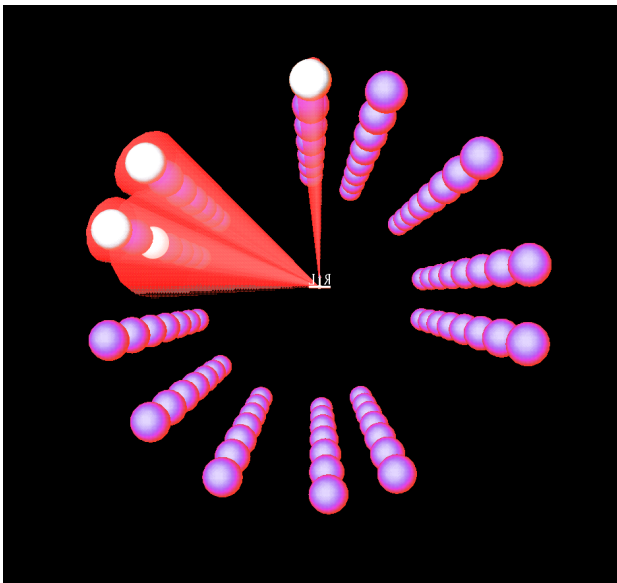Figure 8: Besides on sound source, active sound sources are passed

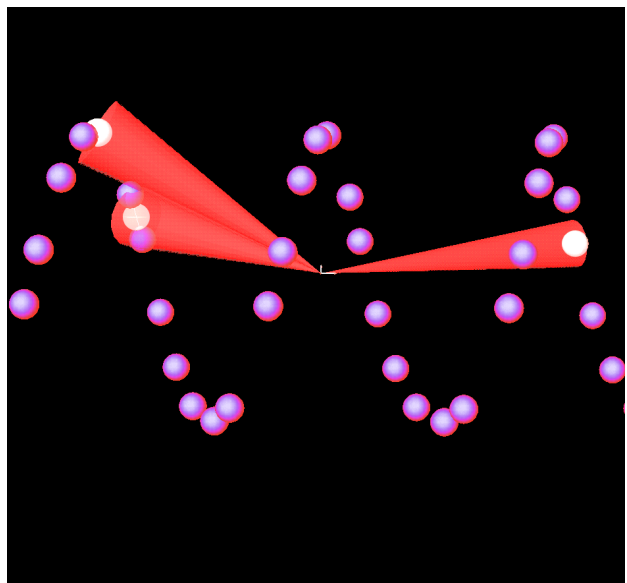Figure 7: Moving closer again; cones become smaller; clusters get split up

Figure 9: Side view shows that resolution cones to the back are larger

3 Implementation

All implementation were done in C++ using Open Inventor and the Sound Spatialization Framework[Her98a] which is freely available as binary distribution. The visualization process takes computation resources, so that the viewer for the soundscape were extended to control its own framerate independently from the framerate of the audio rendering process. This framerate can be set interactively. This allows monitoring without interrupting the main task like application graphics and audio rendering.

4 Conclusion

Realtime visualization of algorithms and related data can give an understanding of those and its performance [FCE$^+$98]. The visualization showed that the clustering algorithm does not reduce resource use (i.e., spatialization channels) in case the sound sources are relative close to the listener.

5 Future research

Not all parameters important for the clustering algorithm were visualized. Distance errors and Doppler effects can be visualized in a future implementation.

Acknowledgments

References

[BCM97] Gavin Bell, Rikk Carey, and Chris Marrin, ISO/IEC *14772-1:1997: The Virtual Reality Modeling Language (*VRML97*)*, 1997, http://www.vrml.org /Specifications/VRML97/.

[FCE$^+$98] Thomas Funkhouser, Ingrid Carlbom, Gary Elko, Gopal Pingali, Mohan Sondhi, and Jim West, *A beam tracing approach to acoustic modeling for interactive virtual environments*, SIGGRAPH 98 conference (held in Orlando, Florida), July 1998.

[HC96] Jens Herder and Michael Cohen, *Design of a Helical Keyboard*, ICAD'96 — Int. Conf. on Auditory Display (Palo Alto, CA; USA) (Steven P. Frysinger and Gregory Kramer, eds.), November 1996.

[Her98a] Jens Herder, *Sound Spatialization Framework*, Web site, University of Aizu, Japan, 1998, http://www-ci.u-aizu.ac.jp/SF/.

[Her98b] Jens Herder, *Sound spatialization framework: An audio toolkit for virtual environments*, Journal of the 3D-Forum Society, Japan **12** (1998), no. 9, 17–22.

[Her99a] Jens Herder, *Optimization of sound spatialization resource management through clustering*, Second Int. Conf. on Human and Computer (Aizu-Wakamatsu, Japan), University of Aizu, September 1999.

[Her99b] Jens Herder, *A sound spatialization resource management framework*, Dissertation, University of Tsukuba, July 1999.

[MM90] James C. Makous and John C. Middlebrooks, *Two-dimensional sound localization by human listeners*, JASA **87** (1990), no. 5, 2188–2200.