Introduction to the Proceedings for ICAD 94

Gregory Kramer
Clarity/Santa Fe Institute
2335 SW 19th Avenue
Portland, OR 97201
Email: kramer@listen.com

The Conference

The Second International Conference on Auditory Display, ICAD 94, was held November 7-9, 1994 in Santa Fe, New Mexico under the sponsorship of the Santa Fe Institute. The Steering Committee consisted of Steve Frysinger, Gregory Kramer (Chair), Stuart Smith, and Elizabeth Wenzel.

The announcement for the conference defined ICAD as follows:

ICAD is a forum for presenting research on the use of sound to display data, monitor systems, and provide enhanced user interfaces for computers and virtual reality systems. It is unique in its singular focus on auditory displays and the array of perception, technology, and application areas that this encompasses. Research areas covered by ICAD include:

- Auditory exploration of data via sonification (data controlled sound) and audification (audible playback of data samples)
- Real Time monitoring of multivariate data
- Sound in Immersive Interfaces (Virtual Reality) and Teleoperation
- Perceptual issues in Auditory Display
- Sound in generalized computer interfaces
- Technologies supporting Auditory Display creation
- Data handling for Auditory Display systems
- Applications of Auditory Display

Included within each of these areas of inquiry are many issues concerning application, theory, hardware/software, and human factors. Integration with speech-audio implementations and with graphical display techniques and their concomitant perception issues also pose significant challenges in each area.

By way of reference, the announcement described some of the topics at ICAD '92. Among the areas of application covered at ICAD '92 were:

- Comprehending complex, high dimensional systems
- Telepresence and Virtual RealityInterfaces for the vision impaired

- Geophysical
- Financial
- Census
- Chemistry data display for blind users
- Software
- Sonar
- Parallel computation analysis
- General user interfaces for vision impaired persons
- Monitoring background processes
- Medical

The Papers

Selection Procedure

Papers submitted to ICAD 94 were sent to the Santa Fe Institute and from there forwarded, without cover pages, to the Review Committee for a blind review process. The Review Committee consisted of:

James Ballas (Naval Research Lab), Nat Durlach (MIT-VETREC), Steve Frysinger (ATT), William Gaver (RANK Xerox), Gregory Kramer, Chair (Clarity, SFI), Elizabeth Mynatt (Georgia Tech), Carla Scaletti (Symbolic Sound), Julius Smith (Stanford), Stuart Smith (U. Mass. Lowell), Elizabeth Wenzel (NASA-Ames), Sheila Williams (U. of Sheffield).

The reviews were then returned to the Steering Committee and the final recommendations were sent to the authors. Papers were accepted as long, short (including student papers), or posters. Long papers were given 30 minutes, short papers 20 minutes, and posters were presented informally at two poster sessions. The poster sessions also included several demonstrations of sound spatialization hardware.

Topics

The title of the conference asserts the thread that connects this collection of papers. All of the papers here relate in some way to the field of auditory display (AD), or using sound (primarily non-speech) to convey information to humans. AD can be understood as a synthesis of research areas, with papers submitted in the following categories:

- Applications
- Sound Synthesis and Design
- Perception
- Software Systems

Of course any given paper is likely to cross boundaries. Indeed, these categories themselves are not as distinct as the disciplines from which they emerged may suggest. The fields of research represented at ICAD have been subdivided in a number of ways, as indicated by the journals and conferences in which related work is presented. The major fields of research represented at ICAD 94 and their associated conferences include:

Human-Machine Systems:

Human Factors and ACM SIG-CHI (Computer Human Interface), which are forums for discussing cognitive science and applications in the context of human-machine systems.

Computer Science:

A variety of computer science conferences such as Visualization, Supercomputing, and MULTIMEDIA are among the usual venues for presenting work on software systems.

Sound Synthesis:

Sound synthesis or audio systems are described in papers presented at the International Computer Music Conference or Audio Engineering Society conference.

Process Of Sound Perception:

The Acoustical Society meetings or gatherings on music perception and cognition serve as outlets for how we hear or process sound.

Auditory Display Applications:

A variety of fields served by auditory displays might attract papers that were presented at ICAD, such as VRAIS for virtual environments, the Statistical Society for data analysis, or the SFI Artificial Life conferences for certain types of large scale simulations.

Auditory Prosthesis:

Blind-related research presented at ICAD would also be appropriate to many of other conferences concerning persons with visual and other disabilities.

Speech Audio:

There is the plethora of speech related conferences and workshops, including ICASSP–International Conference on Acoustics, Speech, Signal Processing and EUROSPEECH.

An Overview

The papers are described here as falling under four categories, viz. *Applications, Sound Synthesis and Design, Perception*, and *Software Systems.* As mentioned above, these categories are simply an organizing tool. Most of the papers presented might have been placed in other categories or required the creation of a new category. For example, Smith et al. could have as readily been placed under Perception as under Software Systems, while Shinn-Cunningham and Durlach could have been placed in Perception rather than Sound Synthesis and Design, where it ended up because of its focus on design guidelines.

Applications

The high information, high stress environments of air traffic control and flight cockpits have seen more AD work than most applications. Following in this tradition, Albers' paper describes a proposed satellite ground control environment that combines sonification techniques with auditory icons and earcons. By applying each technique to a problem best suited to its strengths, the author attempts to tame a difficult human factors problem. Cohen's OutToLunch software, on the other hand, addresses the more subtle problems of group awareness. Using background sounds to provide a sense of presence and information as to how busy other workers are, the author also points out the importance of sound design in the acceptance of an auditory display.

Since the development of tone-based readers, speech interfaces, and digital oscilloscopes and thermometers, interfaces for the blind has been an active AD research area. Stevens et al. describe their work in designing an auditory interface to assist blind readers in understanding the structure of algebraic equations. By associating different timbres with different operations and using pitch and rhythm as they might be employed when an equation is spoken, the authors attempt to indicate the syntax of an equation and thereby make it more comprehensible when spoken. Axen and Choi present a technique for using sound to orient oneself when traversing high-dimensional objects in a software environment. Using a traversal algorithm to extract temporal information from static geometric structures, the authors use the derived data to control frequency, waveform, duration, and amplitude of an additive synthesis system. The complexity of the shapes is thus made perceivable auditorily by utilizing sounds of varying complexity.

Portigal and Carey describe an experiment in which they tested a system for representing the structure of a text document using a sequence of tones. The number of tones indicated the number of chapters and a pause was inserted after the subject's current chapter. Inferior performance with the audio-only cue, and auditory plus visual cues, as compared to that of visual cues alone indicate less than favorable potential for this approach. Nevertheless, the search strategies employed by some subjects did benefit from the audio cues, particularly those involving recall of cues after presentation. The only work presented at ICAD on purely speech audio was Steeples' paper on voice annotation in computer-mediated collaborative work. After providing a background on collaborative writing, Steeples reports on an experiment in which voice annotations were found to be useful in articulating complex ideas and as a complement to visual information.

Sound Synthesis and Design

Shinn-Cunningham and Durlach's paper provides a bridge between research on spatial hearing and the application of this research to AD design. In addition to providing background on the mechanisms, thresholds, and limitations of spatial hearing, this paper provides a useful model for applying known perceptual capacities to display and interface design. Choi begins her paper with a discussion of the role of the observer in scientific and artistic explorations. Recognizing the unavoidable tendency of observers to infer structure in perceived phenomena, Choi goes on to describe her explorations of chaotic systems, most notably the Chua's circuit,an electronic circuit designed to produce chaotic behavior from an analog system. The author focuses on time scaling as a means for ordering observations of complex auditory phenomena.

As the AD field matures, there is beginning to be sufficient published work from which to draw generalized design principles. Mynatt suggests using a two step procedure when designing with auditory icons: first evaluate the identifiability of the icons, and then proceed to evaluate the mappings between cues and the interface concepts they represent. She describes two recent experiments which apply this method.

Lucas' paper likewise emphasizes generalized design concepts by presenting the results of an experiment in which he compared the effectiveness of auditory icons (realistic cues), earcons (cues constructed from abstract [musical] sounds), and speech displays in a generalized computer interface display. Speech turned in the best performance. Lucas found that after the logic of the display was explained, association time and errors for non-speech displays decreased. Perhaps this points to a significant cognitive load in both techniques, a problem that has been well known in the case of earcons.

Perception

The development of a system capable of rendering relationships between data values in a natural way remains one of the key challenges facing sonification researchers. Nonlinearities in our perceptual systems combined with the nascent state of research into sound synthesis and timbre perception add up to a formidable difficulty for sonification investigators. Barrass addresses this challenge head-on by building on John Grey's timbre space research and formulating a framework for sonification capable of representing three data dimensions in a more or less orthogonal manner.

Ballas' paper, building on the work he presented at ICAD 92, describes the results of two experiments in which he sought to discover what aspects of a complex sound composed of several transients are crucial for accurate identification and how much of a sound must be heard in order to identify it. The detail with which Ballas investigates this recognition process provides a useful standard by which we may appraise both synthesized and sampled sound cues. Ballas concludes that acoustic displays employing everyday sounds must take into account not only the sound being used, but also the acoustic details of those sounds.

Using audio spatializers, a so-called "virtual speaker array," Hollander and Furness performed a series of experiments to determine the extent to which geometric shapes could be recognized by use of an audio display that traced the contours of the shapes with moving sound. Building on earlier work that used physical speakers, the authors found that the "spatial synthesizers" were capable of providing a viable alternative to physical speakers. These authors concluded that recognizing spatial patterns in a moving sound was a cognitively loaded task and therefore probably would not be suitable to applications in which the intended benefit of the display is the reduction of cognitive loading.

In his ICAD presentation, Arons reviews the research on time compression of recorded speech and then goes on to describe a technique which pursues efficient presentation of speech information by presenting different signals to each ear. Arons' technique sends alternate segments of a sound to each ear (a technique developed in the 1960's) and adds spatialization of the segments to help overcome intelligibility issues introduced by the unnatural presentation. The informal finding was that this technique produced an externalized audio image that, while still somewhat choppy, was comprehensible as speech.

Zahorik et al. examined localization performance in echoic and anechoic virtual audio environments. The group found that, given the technology they used, spatialized room reflections added little to localization performance. Recent developments emphasizing spatialized reflection information bear reexamination in light of this finding. Macpherson, working in Wightman and Kistler's lab alongside Zahorik, investigated whether the location of a notch in the high-frequency response could account for elevation judgments in virtual auditory displays. The author did not find such a causal relationship. The research into other spectral cues for elevation judgement continues.

Lueck's paper proposes a computationally efficient system for modeling the head related transfer function used in spatial displays. He suggests the use of autoregressive moving average systems

and presents results that indicate such systems are well suited to HRTF computation. Given the computational intensity of virtual acoustic displays, techniques for reducing the computational load are welcome. The author has yet to run experiments indicating the integrity of displays so designed.

Systems

The variety of perception issues and applications associated with ADs point towards a need for generalized display systems. Such systems would be flexible enough to implement a variety of design concepts or examine divergent issues in perception of complex auditory phenomena. Without such systems, researchers in this field are obligated to direct most of their efforts towards coding synthesis algorithms, schedulers, subject response recorders, user interfaces, and so on.

Das et al. describe a high-level system for control of sound designed to be used in conjunction with the CAVE, a virtual environment developed at UI Chicago and more recently at NCSA. The architecture and design philosophy discussed in their paper have broad implications. They discuss the need for interdisciplinary teams, multiple levels of control, and distributed working arrangements, thus encompassing a wide array of important issues in AD system design. This paper introduces the organizing principle of "auditory actors," which are active objects that can have an arbitrarily complex state and behaviors. By establishing hierarchies of actors, their system is capable of providing the appropriate degrees of complexity and control for a wide variety of AD situations. Such a system also enables independent development of different code modules, which, should all elements of the code become widely available, holds promise of enhancing the development process.

In contrast to all of the papers focusing on the implementation of auditory displays, Smith et al. describe a system for psychometric testing of data sonification techniques. Developing work presented at ICAD 92, the authors express their aim as the creation of an interactive computing environment in which sonifications can be created and evaluated, the development of several specific auditory data representations, and the evaluation of one such representation. The authors go on to describe an object oriented system in which the means for flexible sound generation are provided, along with a variety of testing methods, low level tools for manipulating the tests, higher level tools for coordinating a series of tests, and a module for recording the specific performance of each subject. Smith's project points towards a real need in AD research: the integration of sound generation, sonification control, and perceptual testing. The important contribution of Smith et al.'s work, in my opinion, lies not so much in its interesting but inconclusive results concerning the specific characteristics of the sonification technique being tested, but rather in the authors' methodical investigation of underlying dynamics, and the consequent implications for a scientific approach to sonification research in general.

Papp and Blattner have also extended the work they presented at ICAD92, including an exploration of generalized issues in system design for auditory interfaces. They emphasize the importance of the priority given by the observer to different audio notifications, particularly when several applications may be producing audio output. The authors go on to define several notification types and factors that will influence the scheduling of event presentation.

Jameson, another participant following up on work presented at ICAD92, describes the run-time components of Sonnet, a monitoring and debugging program written to investigate how sound can be used in software development. Jameson's contribution to ICAD 94 includes a discussion of the software modules he has designed along with his rationale for design decisions. Bock presents another system for the diagnosis of software bugs. He describes his "auditory domain specification language" with which users can create their own sound domains to employ in the

analysis of executing sequential programs. Boardman et al. address the same issue with their "listen specification language" (LSL). LSL provides a notation for specifying sonifications of program execution and includes a specific tool for sonifying programs written in the C programming language. More papers were presented at ICAD 94 in the use of sound for software diagnosis than any other single application. While all three of the papers presented above describe work in progress, judged by the popularity of this application, software development may very well turn out to be the first source of generally applicable sonification systems.

Posters and other presentations

The posters presented at ICAD equaled the papers for breadth of scope and were connected by the same AD thread. The papers ranged from a discussion of head movements in localization of virtual auditory display objects [Wightman and Kistler] to Mayer-Kress' *Dynamical Resonance and Synchronization of Auditory Stimuli.*

The papers in this volume were augmented at the conference by two invited presentations on complex data representation issues. Chris Langton, of the Santa Fe Institute, gave a talk entitled "Data Presentation Issues in Swarm and other Large Scale Simulations," and Steve Bryson, of NASA-Ames spoke on data management and interactive visualization for computational fluid dynamics. In a presentation which was not written up for publication, Watson and Kidd discussed the application of traditional auditory perception research to AD in *Factors in the Design of Effective Auditory Displays.*

Conclusions

The multi-disciplinary but topically unified quality of the presentations at ICAD 94 were congruent with and a logical extension of the material introduced at the First International Conference on Auditory Display, held in Santa Fe two years prior. ICAD 92 was a small workshop, by invitation only, while ICAD 94 was three times the size and open to all who wanted to attend.

When assembling the first gathering, some doubt was expressed by members of the Steering Committee as to whether even 20 researchers could be found who were working within the field (a number we felt was necessary to make the gathering worthwhile). After working very hard to assemble a small group of researchers, we ended up, to our surprise, with 36 attendees and had to turn a few away. By contrast, with no effort beyond a limited mailing and the posting of an announcement on a couple of Internet lists, about 110 people flew into New Mexico in November to attend ICAD 94.

The strongly international flavor of the gathering is also worth mentioning. Nearly a third of the attendees were from outside of theUnited States. Regions represented included Europe, Australia, Japan,and Scandinavia. Likewise, the financial base of ICAD grew considerably. This year the conference received major sponsorship from NASA. ICAD 94 was also funded by the Department of Energy, the Advanced Computation Lab and Center for Nonlinear Studies at Los Alamos National Laboratory, and Microsoft. Post conference activities received additional funding from the Office for Naval Research.

This increase in interest and support appears to demonstrate a compelling need for such a manifestly multi-disciplinary gathering. Given the strong tradition of narrow focus at most scientific meetings, the import of this point should not be minimized. After all, most of the disciplines that we have discussed above do, indeed, have forums of their own.

For example, ACM SIGGRAPH falls within the genre of multi-media/arts/technology gatherings, as do ISEA and many other meetings. These conferences provide very broad venues for

presentation of new research. Like ICAD, all of these meetings have emerged with the mandate to provide communication channels across disciplines and interests. They are all born of the understanding that we cannot work productively exclusively in isolation. It is becoming increasingly recognized that to address many complex problems we need an interchange of information with people in divergent and sometimes alien fields. As has often been observed, it is at the intersection, the margins, that new ideas are most likely to emerge.

Perhaps the highest contribution that the new, multi-disciplinary gatherings can make is to steer a course broad enough to serve the intellectual and practical need for cross-pollination between disciplines and yet specific enough to provide clear focus, strong relevancy, and thus relatively broad potential application. At future ICADs, one can assume that the papers will have something to do with auditory display applications, the systems that make them, or the perceptual issues associated with such displays. One-track designs will no doubt vary in interest level, depending on the attendee; however, given sufficient quality of presentations, it is to be hoped that interest in the conference as a whole will remain reasonably high for most participants while at the same time regions of peak relevance will continue to appear with good frequency.

It may be that ICAD has the potential to deliver even more. A frequently heard comment at ICAD 94 was, "I wouldn't have attended a paper session that included (the work in question), but I'm glad I heard these papers. They gave me new ideas." The comments of geophysicist Chris Hayward were illustrative: "I came because the previous conference (ICAD 92) generated a rich set of ideas and concepts, some of which may, in the future, have an application in geophysics and because the previous ICAD had a rich collection of creative, imaginative people that, for a brief week, collaborated to build a field which none of us has yet been able to adequately define."

Attendees were also not hesitant to express their observations as to deficiencies they saw with ICAD 94. These included a shortage of empirical findings, far too few sounds played with the paper presentations, and too few data representation papers. The quality of the papers was felt by many to be uneven. Perhaps this is to be expected, considering the youth of the field and the brief history of the conference itself. The overwhelmingly positive response to ICAD was all the more remarkable in light of these weaknesses.

Perhaps it was the emergence of a research community capable of supporting this interdisciplinary field that generated the cross-disciplinary appreciation discussed above. Community formation was enhanced, I think, by the bus trip to the Anasazi cliff dwellings, out past the mesas near the Los Alamos National Laboratories. Against the shocking blue skies of the northern New Mexico desert, a varied and colorful group of participants climbed ladders and cliffs to wander amongst ancient dwellings and religious sites. The good food and generous break time found computer scientists speaking with perception researchers and people from industry or government labs speaking with composers and professors. The formation of interdisciplinary research communities is one of the missions and hallmarks of the Santa Fe Institute. Perhaps the character of the ICAD gathering may be credited with encouraging the formation of such a heterogeneous community.

Next Steps

As a result of the interest generated at ICAD 94 and elsewhere, there now exists demand for an auditory display organization to serve the needs of this research community. The general sense at the conference seemed to be that such an organization could support the ICAD email list. There was also very strong support for an ICAD World Wide Web page. This page, located at http://www.santafe.edu/icad/ has been under developmentand shows promise as a tool for

research and community formation. While there has also been talk of a journal, the general sense seemed to be that the field needs to mature further before a regular publication schedule could be supported.

One unambiguous result of ICAD 94 was support for another conference. ICAD 96 is now being planned, again with the sponsorship of the Santa Fe Institute. With an increased focus on empirical results and sound examples, along with tutorials to help bridge the gap between disciplines, this meeting will carry on the tradition of providing our auditory research community with a productive and broadly relevant forum. Future ICADs, in turn, promise to sustain the emergence of auditory display as an essential component in human's orchestration of their own perceptual faculties for purposes of comprehending their world.