Generative Soundscapes for Experiential Communication

David Birchfield Nahla Mattar Hari Sundaram
Preetha Appan Ankur Mani Bageshree Shevade
Arts Media and Engineering Program
Arizona State University
e-mail:{david.birchfield, nahla.matar, hari.sundaram, preetha.appan, ankur.mani, bageshree.shevade }@asu.edu

1. Abstract

This paper describes the design and implementation of a generative model for the creation of soundscapes in real time. The model extends the work of the Acoustic Ecology community, and offers several extensions. In particular, the model is adaptive to individual user contexts and integrates audio techniques for creating immersive sonic environments. The authors outline tools and methods for the creation of annotated media databases that document daily experiences, and describe relevant applications that utilize this work.

2. Introduction

Soundscapes and sound collages are capable of creating a strong sense of time, place, and context. This has been documented in analytical work from musicians [36], theorists, [3,7] and ethnomusicologists [17, 19]. Soundscape collages from the real world have been captured and studied [29, 30], and have been synthesized for artistic purposes [35,36]. However, there has been little focus on computational, generative mechanisms for the automated creation of convincing sonic environments.

In our recent work, we have been developing frameworks for interaction with multimodal digital media that are intended to communicate experiences among participants in a social network [2]. We are concentrating on the development of applications that allow users to both reflect upon and communicate their experiences using collected media samples from their daily lives. Through the remediation of their experiences, we seek to engage users with the experience of another's activities, and to provide paths by which users will better understand one another and their experiences.

Our multimodal interaction frameworks include presentation schemes for digital still images, videos, and sound samples. This paper describes our work in the creation of dynamic soundscapes that can be used in multiple applications to generate compelling, communicative, immersive sound environments. We begin by describing the system architecture from acquisition to presentation. Next we discuss our work with context models that inform and influence the generative model. Then we describe the generative model itself and present several applications that utilize the model. Finally, we propose future work and conclusions.

3. System Architecture

Each application that is describe in Section 6 shares the same basic architecture and data flow. First, users collect digital media from their environments. Second, each user uploads this data to a central, online database. Next, through an active, incentive-based annotation system, participants can annotated their media. Finally, the collected data and annotations are search and retrieved in real time to deliver soundfiles and other digital media to our generative sound applications.

3.1 Acquisition Rate

There is extensive prior and current work in the documentation of daily experiences [18]. Many of these documentation systems are geared toward pervasive, 'always-on' multimedia recording that use idiosyncratic tools and methods.

In contrast, we view the recording and annotation of daily experience as a question of finding an appropriate sample rate that mitigates concerns of practicality and documentation. We are exploring a range of sample rates and media to determine what quantity and type of data is required in order to remediate and communicate everyday experiences. For example, can only three or four still images effectively communicate the experience of a day in one's life? Does this limited data rate allow one's subjective memory to inform the documented moments? Can one convey the experience of a summer evening in a rural setting with only one recorded ambient soundscape instead of with a narrated video? How might this change the way in which the user reflects on this experience? Though the scope of this paper is focused on the construction and presentation of soundscapes, our related work [2, 23] provides a more detailed description of tools and techniques for multimodal acquisition and presentation. Our initial work has demonstrated that this low sample rate provides advantages over 'always-on' recording techniques in terms of practicality, convenience, data management, and communication.

3.2 Tools

Our application is geared toward a large group of motivated users who will carry tools for recording their everyday experiences. There are three primary criteria that must be met by our recording tools: (a) availability, (b) fidelity, and (c) convenience. Though this paper describes a number of custom software tools for annotation and display of collected media, we are limiting our recording tools to devices that are commercially available, and are not prohibitively expensive.

We are currently using MiniDisc recorders (Sony MZ-R700 or Sony MZ-NH1, with an external microphone such as Sony ECM-MS709) to record audio events and ambient sounds from our daily experiences. Though, unlike DAT and CD recorders, they apply some compression, we have been pleased with the sound quality for our purposes. Dictaphones and even the iPod offer convenient and portable solutions for audio recording, but most models are limited by a 8 or 11 kHz sampling rate that provides insufficient fidelity.

Our current pilot study demonstrates that recording tools must be lightweight and small if users are realistically expected to carry these tools to document everyday experience and not just special events. Though they would provide optimal quality and download convenience, most standard high quality audio recording devices, such as DAT, CD, and solid-state recorders, are bulky and often weigh in excess of ten pounds. Users cannot be expected to carry such a large device, and this has informed our decision to use MiniDisc recorders with small external microphones for audio recording.

3.3 Uploading Media

A second important concern with convenience is the transfer of data from the acquisition tool to the media database. Manufacturers of digital camera technologies have recognized that convenience is paramount for users to regularly capture samples from their daily environments. In our related work we have used standard commercial software to download pictures from a digital camera, and then batch uploaded these images to the database [2]. Although this procedure has not proven prohibitive, we have found that any step toward convenience will greatly increase the success of true recording of daily events. Thus, we are in the process of developing our own tool that will download pictures, extract meta data, and populate the media database with this information in one seamless step.

Unfortunately, data transfer of recorded audio poses more of a challenge. To meet our goal of a low cost (i.e. - time, money, learning curve) for participation, we are in the process of developing tools that will allow the upload of sounds to be as convenient as for digital images. We have developed software to normalize the audio data, extract fundamental features of the file (e.g. duration, format, spectral characteristics), and populate the media database with this information.

3.4 Media Annotation

As described in Section 3.3, some basic signal-level audio features are extracted from sound samples at the time of upload. Though these details are useful, our generative framework requires semantic annotations of the soundfiles. Furthermore, the system can be most adaptable to multiple users, if it has semantic knowledge that is provided from numerous viewpoints. Thus, the uploaded soundfiles must be annotated by participants in the social network.

We have been working on the problem of creating media annotation systems, that provide incentives to annotate [33]. The goal of our work is to create novel semi-automated, intelligent annotation algorithms that bridge manual methods for annotation and fully automatic techniques. There has been prior work in semi-automatic image annotation using relevance feedback [14]. While there are rich mathematical models used, we believe that there are two shortcomings of the current work:

- **Experience**: None of the systems focus on the end-user experience. There is very little return on the enormous time invested by the users to annotate media. Currently annotations only enhance the search capability and *not* the presentations. *An annotation system that provides insight will create incentive in the user to enter richer annotations.*
- **Semantics**: Current approaches to annotate media [14] essentially treat words as symbols regardless of their semantic relationships with other words, which is no different than any normal image feature. The lexical meaning of the keywords/annotations is not exploited.

In our work we address issues relating to both semantics and the end user experience. We establish the semantic inter-relationships amongst the annotations by using WordNet [27]. We attempt to map the annotation problem as one of an experiential system [20, 34] – the key idea being that the user will gain insight about the media in relation to her context, thus providing a strong incentive for the user to annotate the media.

The annotation procedure is as follows. Our annotation system uses a combination of low-level as well as WordNet distances to propagate semantics. As the user begins to annotate the media objects, the system creates positive example (or negative examples) media sets for the associated WordNet meanings. These are then propagated to the entire database, using low-level features as well as WordNet distances. The system then determines the media object that is least likely to have been annotated correctly and presents it to the user for relevance feedback.

The system also attempts to provide insight by presenting knowledge sources to the user. This is done using context-aware hyper-mediation. In our approach, we use Google to automatically generate hyperlinks. This is done by taking into account the user profile, the semantics of the media and the semantic relationship between the media item and the user profile.

3.5 Database access

All media objects and annotations are stored on a centrally located server. We employ a MySQL database, and have developed database clients for uploading, annotation, and retrieval in PHP, Java, and Max/MSP. Practical experience has demonstrated that if the sound samples are propagated to a local computer, the database can be searched by the generative algorithm in real time as the soundscapes are being dynamically generated.

4. Modeling Context

A critical aspect of the generative model is adaptation to the user's needs and the environmental setting - in short, the context. The Merriam-Webster [1] defines context as "*the interrelated conditions in which something exists or occurs*." Context is very important in a multimedia system since perception and interpretation of any media is relative to the context in which it is consumed. Secondly, actions are also user context dependent.

Prior work on context includes work in the Natural Language Understanding [13], and ubiquitous computing [15]. In our multimodal feedback systems, we require a broader definition of *multimodal context*. The three parallel phases of the context model depend upon the set of questions are *context acquisition, context representation* and *context evolution*. Research on the architecture of context aware systems have focused on a modular approach to context modeling that attempts to maintain a demarcation in these phases [16]. Based upon the set of questions and the phases the context can be classified into three types as follows.

Environmental context: The environmental context answers a set of questions related to the user environment such as, *"where", "when", "what"* and similar questions [16]. *User context:* The user context is related to the meaning and user understanding. It answers the questions such as, *"who", "what are the skills", "what are the interests", "what are the goals", "what is the* *understanding"* among others. This pertains both to the users who acquires media samples, and to the user who interacts in an application environment. Hence, it is concerned with semantic inter-relationships between concepts, which can be arbitrary.

Application context: The application context answers the questions such as, "what situation and state of the application".

We address the issues related to context in our work on context models [12]. The relationships between the concepts are arbitrary. Our previous work on context models has successfully been applied to different problems [12, 32]. The formal model is defined using a semantic-net – a graph $G = <V,E,W>$ where the nodes $v_i \in V$ represent the concepts, the edges $e_{ij} \in E$ represent the type of relationship (semantic, spatio-temporal, feature-level) between the nodes i and j and $w_{ij} \in$ W, specifies the strength of the relationship between the two nodes. The notion of a concept is multimodal. We define the context be the union of semantic-nets:

<1>

where C is the context, k is the total number of semantic-nets and G_i is the i^{th} semantic-net. We also discussed (a) the composition and construction of user context (b) the relationship between media and the user context and (c) the context evolution during the user's interaction with the environment.

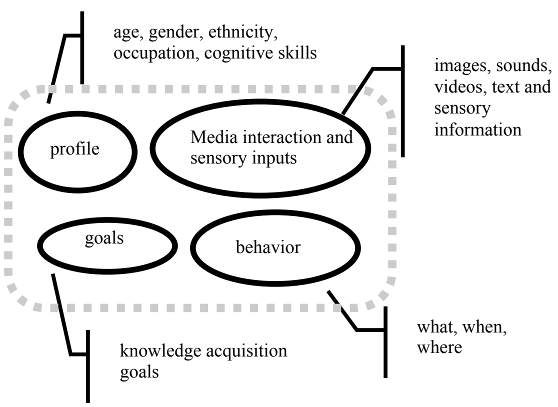

Figure 1: The user context model has four components – (a) profile (b) media history (c) behavior and (d) goals

As depicted in Figure 1, the user context is comprised of concept nets composed from (a) the initial user profile (stating the user's interests, background etc.), (b) user interaction with the media and the system and the sensory information about the user (c) user behavior (d) user's goals. This information is used to make concept nets that together form the user context. The relationships between the concepts like the concept cover and the concept distances were also introduced in the paper. Finally a memory model was introduced to model the temporal evolution of context. The work gives a novel method of modeling multimodal user context with information coming from different sensors. The context evolves over time, with user interaction.

The context model introduced is multimodal. However, the cross-modality relationships and the relationships between the concepts represented by media such as images and audio are only statistical and very elementary. More sophisticated and arbitrary relationships are required for multimodal sensory fusion and context modeling.

5. Model for Generative Soundscape Construction

In [28] the authors propose an approach to the automated generation of soundscapes in their work on Audio Aura. There they dynamically generate sonic seascapes to convey information in specific applications such as email alerts and ambient displays. However, their generative mechanism is hindered by a small palette of sound sources and a narrow focus on information display. In our recent work [4-6], we have developed generative models for the automated synthesis of sonic environments and music. Although these models have yielded successful musical results, they do not focus on the creation of immersive environments or explicit narratives. Our framework for the generation of immersive soundscapes extends this work through the use of user-adaptive models derived from research in acoustic ecology, and combining this with the tools for acquisition and annotation described in Section 3.

5.1 Acoustic Ecology

The Acoustic Ecology community has undertaken important and influential work in the empirical study of sound and soundscapes in our communities [30, 36]. Through their work in recording and analyzing real world soundscapes, Schafer and Truax propose a powerful system of semantic sound classification [29, 35].

Four classes of sounds emerged from the work of Shafer: *keynotes, signals, soundmarks, and sound romances*. A keynote sound is the tonal center of a soundscape. For example, the keynote of an oceanside town is the sound of the beach. The hum of the computer hard drive is the keynote for many of our work environments. An infrequent, and sometimes alarming sound, is classified as a signal. Due to their infrequency, the appearance of a signal sound will convey a kind of message to the listener. A police siren or email alert are typical examples. A soundmark is the acoustic equivalent of a land mark. It is a sound that distinguish and identifies a particular sonic environment. The carillon of Big Ben or a church's bell are soundmarks. Finally, sound romances are those that inspire a feeling of nostalgia or longing in a listener. Often they are anachronistic sounds that have since disappeared. For example, the sound of a wooden wagon wheel on a cobblestone street.

Our model for the generation of immersive soundscapes is fundamentally based on principles of Acoustic Design outlined by Shafer and Truax. Using the annotations provided by participants in the social network, the model generates appropriate sonic environments with keynotes, signals, soundmarks, and sound romances that are relevant to the application. In addition, we have introduced several extensions that allow the sonic environments to dynamically adapt to and embody individual listeners.

5.2 Integration of User Context

As described in Section 4, our model of user context includes dynamically updated knowledge about a the user, her environments, and her interactions with an application. This contextual knowledge is used to generate soundscapes that adapt to the particular experiences and expectations of individual users.

Shafer proposes four classes of sound events, but the distinctions are not always clear cut and sound classifications can evolve and transform over time. For example, the sound of an ambulance siren is often a clear *signal* sound. However, through repeated exposure to that signal – either in the physical world if one lives near a hospital, or through repetition in a synthesized soundscape – the sound of a siren can lose its immediacy and impact as a signal.

In addition, the classification of sounds cannot be universal for all users. For example, in Islamic countries the call to prayer sounds five times a day. For native Muslims it would be a signal to gather for prayer. For native, but non-Muslims the event will be a keynote, and for non-natives the event will be a soundmark. As a second example, if the goal of a given soundscape is to convey a sense of nostalgia and familiarity for home, the model might introduce keynotes of the ocean, with soundmarks such as channel buoy bells. While this might generate a compelling and immersive sound environment, for a listener raised in the plains states of the United States, the intended result would be lost. Rather, keynotes taken from rural environments, with the sound of the wind in trees, and signals such as cow bells would be more effective.

Our user context model plays an integral role in ensuring that the generated soundscapes are well matched to the experiences and responses of individual users.

5.3 Modeling Acoustic Spaces

One of the goals of our generative model is to create compelling, immersive sonic environments. To enhance the sense of immersion for listeners, we are using reverberation and frequency filtering techniques to simulate 3D acoustic spaces. In some cases, these virtual rooms or environments are intended to mimic real world acoustic spaces such as a home or office space for a particular participant. In other applications, these virtual rooms are imaginary and are realized for affect. For example, the model might communicate a sense of expansiveness or confinement using the same set of keynotes and signals, but altering the virtual room acoustic in which these samples are presented.

5.4 Embodiment in Soundscapes

The focus of analytical work in the World Soundscape Project, and in acoustic ecology more broadly, is to better understand the sonic environments in which we live and how they have evolved. While this research provides an excellent basis for the creation of soundscapes that evoke a sense of place, it does not suggest how we might represent peoples in those places. In recent work, we have examined effective techniques of representation in western common practice music [24, 25], and we borrow concepts from program music, opera, and suggestive music to create a sense of embodiment in our soundscapes.

In addition to sampling and uploading their sound environments, participants are encouraged to upload and annotate music samples that represent their musical preferences, heritage, and cultural identities. Employing the concept of the *leitmotif* (leading motif) [31], when appropriate, the generative model will represent individual participants in the soundscape, using a selection of their music samples. This representation can be used to convey information about a particular social event, to associate a person with a place, or to more broadly introduce the rich musical communities of the participants in conjunction with their shared sonic environments.

6. Applications

Thus far, we have presented our framework for the acquisition and annotation of a database of sound samples and media objects. We have described our model for the dynamic creation of user-adaptive immersive soundscapes. In this section we describe three applications in which we are integrating this work.

6.1 Social Network Navigation

We are using generative models for soundscape creation to enhance our work in interactive frameworks for viewpoint based exploration of events. We are creating immersive, engaging, multimodal environments that allow users to reflect upon and communicate their experiences. Participants in social networks are encouraged to document their daily experiences using recording tools described in Section 3.2. They annotate these media objects and those of others in their spheres. Using the sound samples captured by the group, in the context of a visual interactive interface, we use the generative model described above to remediate real and imagined sonic environments.

The idea behind this application is that users desire *agency* [8] – users proactively interact with the system to control the flow of story, is incorporated in this visualization. Prior work in [21, 26], discussed the idea of multiple authorship as well as tangible interfaces with multiple viewpoints as a storytelling mechanism. Our work combines creation and visualization of everyday stories into a unified web – based framework. Our interaction is event based and uses time and space as visual cues to make the viewer understand how the events took place.

We provide an exploratory environment that provides users with different ways to navigate through the same event space. Users are presented with a dynamic visualization and sonic presentation that lets them explore events based on the viewpoints associated with them. The interface allows them to navigate across time.

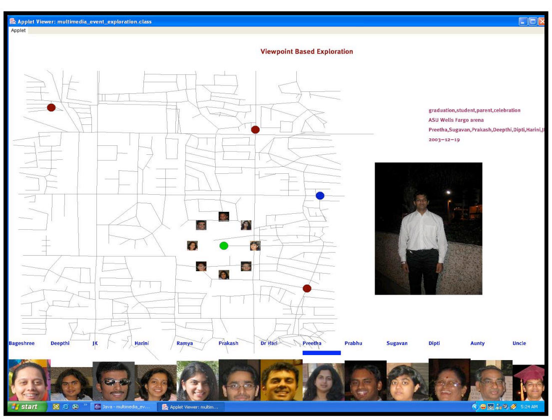

Figure 2: Snapshot from the application

6.2 Situated Media Systems

Situated multimedia systems are physically grounded systems, often located in everyday environments such as the home, office, or automobile (e.g. an intelligent home). They focus on how meaning can be communicated through novel approaches to sensing, modeling, interaction, and feedback (non-computer screen based feedback – e.g. changes in ambient sound). The situated media environment is an ideal application for the integration of soundscapes.

Our situated media systems are inspired Rodney Brooks' [9-11] and Pattie Maes' earlier work on autonomous agents [22]. Though we are building on their work, there are important differences. In a situated multimedia system, the human being (unlike a robot in Brooks' case) is the embodied agent, that is also not controlled. However, the world that the human inhabits is partially controlled by the environment (unlike Brooks' chaotic world), using sonic and visual displays. In this situated media system application, our model for soundscape creation is oriented in two directions: ambient display of information, and the creation of immersive sound environments that are matched to the environmental context.

6.3 Soundscapes in Performance

In addition the interactive applications described above, we are utilizing these methods for soundscape creation in recent media performance works. *Scars*, is a large-scale performance work that integrates soundscapes with live music, video, animation, and theater.

The work exists in two imaginary soundscapes. The *mindscape* of the piece is derived from the Egyptian culture of its composer, and the *inscape* conveys the worlds of other characters and cultures. Generated soundscapes in *Scars* connect these imagined environments, the composer's experiences (sampled from her daily experiences), and the real time sounds of onstage performers. The resulting immersive environments communicate the real and imagined experiences of the work's characters, and provide performance environments that provide feedback to the performers themselves. Interactive mechanisms in the piece utilize a modified context model that affords flexibility and control in the performance at both the local and structural levels.

7. Conclusions and Future work

We have described a model for the dynamic creation of soundscapes in real time. The model is based on the work of the Acoustic Ecology community, but offers several extensions. In particular, the model is adaptive to individual user contexts and introduces additional techniques for creating immersive sonic environments. We have outlined tools and methods for the creation of media databases the document daily experiences, and have outlined several of our applications that utilize this work. Though our initial work has been promising, we anticipate several directions in our future work that will refine both the usability and effectiveness of the model.

We must further improve the automated routines for audio feature extraction that employ more sophisticated machine listening techniques. We briefly discussed software tools that are under development to streamline the transfer of data from acquisition tools to the media database. Any improvements in this stage of acquisition and analysis will reduce the annotation burden, and will lead to greater accuracy in database search and retrieval.

In Section 4 we discussed that user interaction comprise part of the user context model. Though we have made some small steps toward integrating this knowledge into the generative model, further work is required to improve this important feature.

We have formalized the use of Acoustic Design principles in our model. However, we have only implemented rudimentary mechanisms for managing the layering and organization of sound samples. This is a crucial aspect of the model that will require more extensive work. We intend to draw on models from the music theoretical literature and design user feedback studies to address this need.

8. References

[1] *Merriam Webster Dictionary* http://www.m-w.com.

[2] P. Appan, H. Sundaram, D. Birchfield (2004). *Communicating Everyday Experiences*. Proceedings ACM Workshop on Story Representation, Mechanism and Context, New York, NY.

[3] J. BALLAS and J. HOWARD (1987). *Interpreting the Language of Environmental Sounds*. Environment and Behavior.

[4] D. BIRCHFIELD (2002). for Susan, http://ame2.asu.edu/faculty/dab/music.php.

[5] D. BIRCHFIELD (2003). *Genetic Algorithm for the Evolution of Feature Trajectories in Time-Dependent Arts*. 6th International conferene on Generative Art. Milan, Italy.

[6] D. BIRCHFIELD (2003). *Generative Model for the Creation of Musical Emotion, Meaning, and Form*, ACM SIGMM 2003 Workshop on Experiential Telepresence, Berkeley, CA.

[7] A. BREGMAN (1990). Auditory Scene Analysis, MIT Press.

[8] K. M. BROOKS (1997). *Do story agents use rocking chairs? The theory and implementation of one model for computational narrative*, Proceedings of the fourth ACM international conference on Multimedia, 317 – 328.

[9] R. BROOKS (1986). *A robust layered control system for a mobile robot*. IEEE Journal of Robotics and Automation **2**(1): 14-23.

[10] R. A. BROOKS (1991). *Intelligence Without Reason*, International Joint Conference on Articial Intelligence, Sydney, Austrailia, pp. 569-595, Aug. 1991.

[11] R. A. BROOKS, C. BREAZEAL, R. IRIE, et al. (1998). *Alternate Essences of Intelligence*, Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98), Madison, Wisconsin, pp. 961-976, July 1998.

[12] J. BRUNGART, H. SRIDHARAN, A. MANI, et al. (2004). *Adapting Multimedia Design To Context: A design framework for interactive, user context-adaptive multimodal learning environments,* Arts Media and Engineering Program, Arizona State University, AME-TR-2004-08, Jun. 2004

[13] S. W. K. CHAN and J. FRANKLIN (2003). *Dynamic Context Generation for Natural Language Understanding: A Multifaceted Knowledge Approach*. IEEE Transactions on Systems, Man, and Cybernetics **33**(1): 23-41.

[14] E. CHANG, K. GOH, G. SYCHAY, et al. (2003). *CBSA: content-based soft annotation for multimodal image retrieval using Bayes point machines*. IEEE Transactions on Circuits and Systems for Video Technology **13**(1): 26-38.

[15] A. K. DEY and G. D. ABOWD (1999). *Towards a Better Understanding of Context and Context-Awareness*, Proceedings of the 3rd International Symposium on Wearable Computers, San Francisco, CA, pp. 21-28, October 20-21, 1999.

[16] A. K. DEY (2000). *Providing Architectural Support for Building Context-Aware Applications*. College of Computing. Atlanta, Georgia Institute of Technology, PhD.

[17] S. FELD (1982). Sound and Sentiment: Birds, Weeping, Poetics, and Sound in Kaluli Expression. Philadelphia, PA, University of Pennsylvania Press.

[18] J. GEMMELL, G. BELL, R. LUEDER, et al. (2002). *MyLifeBits: fulfilling the Memex vision*, Proceedings of the 10th ACM international conference on Multimedia, Juan Les-Pins, France, pp. 235-238, Dec. 2002.

[19] B. GOODEY (1974). Images of Place: Essays on Environmental Perception, Communications and Education. Birmingham, England, University of Birminghan, Centre for Urban and Regional Studies.

[20] R. JAIN (2003). *Experiential Computing*. Communications of the ACM **46**(7): 48-55.

[21] A. KELLIHER, A. MAZALEK and G. DAVENPORT (2003). *Documenting Digital Dialogues: Engaging Audience in the Construction of a Collective Documentary Across Time and Space*, Proceedings of TIDSE 2003, Darmstadt, Germany, Springer-Verlag, pp. 248-259, 2003.

[22] P. MAES (1994). *Modeling adaptive autonomous agents*. Artif. Life **1**(1-2): 135--162.

[23] A. Mani, H. Sundaram, D. Birchfield, G. Qian, *The Networked Home as a User-Centric Multimedia System* , to appear in Proceedings ACM Workshop on Next-generation Residential Broadband Challenges, New York, NY, 2004.

[24] N. MATTAR, (1998) *Analytical Study to Reffat Granna's program Music Composition*. Masters Thesis. Cairo, Egypt.

[25] N. MATTAR, (2000). *Art's Incarnation to Life: Beethoven and Tolstoy's*. Afaq. Cairo: The Supreme Culture Council.

[26] A. MAZALEK, G. DAVENPORT and H. ISHII (2002). *Tangible viewpoints: a physical approach to multimedia stories*. Proceedings of the tenth ACM international conference on Multimedia(eds), ACM Press**:** 153--160.

[27] G. A. MILLER, R. BECKWITH, C. FELLBAUM, et al. (1993). *Introduction to WordNet: An On-line Lexical Database*. International Journal of Lexicography **3**(4): 235-244.

[28] E. MYNATT, BACK, M., WANT, R., BAER, M., ELLIS, J.B. (1998). *Designing Audio Aura*, SIGCHI conference on Human factors in computing systems.

[29] R. M. SCHAFER. (1977). The Tuning of the World. New York, Random House.

[30] R. M. SCHAFER (1973-78). The Music of the Environment Series. Vancouver, A.R.C. Publications.

[31] R. Scruton. *Programme Music*, Grove Music Online ed. L. Macy (Accessed [27 10 2004]), http://www.grovemusic.com.

[32] B. SHEVADE and H. SUNDARAM (2003). *Vidya: An Experiential Annotation System*, 1st ACM Workshop on Experiential Telepresence, in conjunction with ACM Multimedia 2003, Berkeley, CA, Nov. 2003.

[33] B. SHEVADE and H. SUNDARAM (2004). *Incentive Based Image Annotation*. Arts Media and Engineering Program, Arizona State University, AME-TR-2004-02, Jan. 2004.

[34] H. SRIDHARAN, H. SUNDARAM and T. RIKAKIS (2003). *Computational models for experiences in the arts and multimedia*, 1st ACM Workshop on Experiential Telepresence, in conjunction with ACM Multimedia 2003, Berkeley CA.

[35] B. TRUAX (1996). *Soundscape, Acoustic Communication and Environmental Sound Composition*. Contemporary Music Review **7(1)**: 49-65.
[36] B. TRUAX (1985). *Acoustic Communication, 2nd Ed*. Vancouver, Ablex Publishing.
[37] K. WRIGHTSON (2000). *An Introduction to Acoustic Ecology*. Soundscape, The Journal of Acoustic Ecology(1(1)): 10-13.