Immersive Audio Synthesis Algorithms

1. Research Team

Project Leader: Prof. Chris Kyriakakis, *Electrical Engineering*

Other Faculty: Prof. Tomlinson Holman, *CNTV*

Graduate Students: H. Huang, C. –S. Lin, M. Peterson

2. Statement of Project Goals

The work in this project is focused on the development of a neuro-fuzzy approach to remove spectral noise from audio signals. There are three major components in our method. The first is based on the complex semi-cepstrum technique that is employed to extract features from audio signals. The second element is a fuzzy cerebellar model articulation controller (FCMAC) that is used to quantize the signals. The final element was developed based on the theory of stochastic real values (SRV) that is used to search the optimal frequencies in the trained system. We developed a signal processing implementation of the SRV algorithm that performs well in removing spectral noise that is buried in audio signals. Real-time implementation of this method will be critical in all IMSC projects that rely on high-fidelity audio communication.

3. Project Role in Support of IMSC Strategic Plan

Immersive audio is one of the critical elements in IMSC's vision of Immersipresence and is intricately tied in to several projects at many levels. These include the 2020Classroom project and its Media Immersion Distance Learning Classroom (MIDLeC), the Distributed Immersive Performance project, as well as the Remote Media Immersion implementation within the MIE architecture. The noise removal methods presented here are critical because of the inherent background noise that is present in all real-world recording applications.

4. Discussion of Methodology Used

A primary difficulty for any reinforcement learning design is to identify an appropriate set of states, actions, and critics. The proposed signal-processing scheme covers all these issues and has a learning mechanism realized by a self-organizing fuzzy cerebellar model articulation controller. The motivation for integrating these different strategies is to develop a more effective approach for systems characterized by non-linearity and uncertainties. In the experiments we performed, a neural network was designed to remove the audio signals that are colored by the noise with multiple frequencies. The learning mechanism plays a role of a compensator for the unwanted sound from the loudspeakers to the listeners. The uncertainty and interaction between the subsystems can also be handled in the learning process of the neuro-fuzzy system. We demonstrated that this FCMAC-SRV neural network has the capability of approximating complex, nonlinear functions with multidimensional inputs. The proposed system can be applied to any nonlinear system by giving appropriate criteria as the performance measurement. The simulation results show that a noise-removing task can be learned even with little a priori knowledge. We are currently investigating how to generalize this approach to cover broad applications such as signal separation for virtual microphones and cross-talk cancellation for immersive audio rendering.

5. Short Description of Achievements in Previous Years

Research in previous years focused on innovations in the area of Virtual Microphones for synthesizing multichannel sound:

- First group to propose the use of virtual microphone algorithms
- Filters designed for various representative spaces that could be used to convert mono and stereo recordings to multichannel versions
- Real time synthesis of distant microphones demonstrated
- Off-line synthesis of spot microphones for percussive sounds demonstrated using Choi-Williams distributions

5a. Detail of Accomplishments During the Past Year

- Developed basic framework for FCMAC-SRV neural network
- Tested on representative cases
- Publication submitted to IEEE Transactions

6. Other Relevant Work Being Conducted and How this Project is Different

Conventional approaches to this problem have used adaptive filters or stochastic processes to reduce noise with given frequencies [1.2]. The major drawback of these approaches is that they require a priori knowledge of the noise signal characteristics. These reference models are typically unavailable and thus these methods are not suitable for practical applications. Moreover, analytical models tend to become complicated and nonlinear so that they cannot be solved by traditional methods. An alternative way to deal with the problem is to use neural networks to remove the unwanted components; however, in most cases the exact desired signals are unavailable for training. A major characteristic of our proposed system is the capability of using incomplete knowledge about the recorded signals to predict their inputs. Since most of the applications of reinforcement learning are designed for control issues, a preprocessor based on the complex semi-cepstrum and a quantizer based on the fuzzy cerebellar model articulation controller (FCMAC) are adopted for this purpose [3]. However, because the original FCMAC is a supervised learning algorithm and is not applicable for an ill-defined objective, the adaptive heuristic critic (AHC) and the stochastic real value (SRV) method are also applied to meet this requirement [4.5]. In other words, the synaptic weights of the FCMAC are updated by a stochastic reinforcement signal that represents the difference between the actual and the predictive failures. From the standpoint of the modified SRV, the FCMAC is employed as a signal quantizer. On the other hand, the FCMAC uses the modified SRV as its weight-update rule. Therefore this self-organizing FCMAC can cancel noise resulting from a given electroacoustic system including both the electrical and the acoustical domain.

7. Plan for the Next Year

We plan to demonstrate the effectiveness of this noise removal method on real two-way communication audio data collected from 2020Classroom and Distributed Immersive Performance Sessions. Initially, this will be done off-line. As we gain a better understanding of the performance trade-offs we will move towards a real-time implementation.

8. Expected Milestones and Deliverables

Implement real-time version that runs on DSP platform for two input channels.

9. Member Company Benefits

N/A

10. References

[1] Simon Haykin, “Adaptive Filter Theory,” Prentice Hall, Inc., 2002.
[2] Ben Gold and Nelson Morgan, “Speech and Audio Signal Processing: Processing and Perception of Speech and Music,” John Wiley and Sons, Inc., 2000.
[3] K. S. Hwang, C. S. Lin and C. L. Chang, “Fuzzy Cerebellar Model Articulation Controller”, Proceedings of The Second Chinese World Congress on Intelligent Control and Intelligent Automation, pp. 1536-1541, 1997.
[4] A. G. Barto, R. S. Sutton and C. W. Anderson, “Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problems,” IEEE Trans. Syst., Man, Cybern., vol. SMC-13, pp. 834-846, Sept./Oct. 1983..
[5] V. Gullapalli, J. A. Franklin, and H. Benbrahim, “Acquiring Robot Skills via Reinforcement Learning,” IEEE Control Systems Magazine, pp. 13-24, Feb. 1994.
[6] Alan. V. Oppenheim and Ronald. W. Schafer, Discrete-Time Signal Processing, Prentice-Hall, Englewood Cliffs, 1989.
[7] M. Brown and C. J. Harris, “A Perspective and Critique of Adaptive Neurofuzzy Systems Used for Modelling and Control Applications,” Int. J. of Neural Systems, vol. 6, no. 2, pp. 197-220, 1995.
[8] L. A. Zadeh, “Fuzzy Logic,” IEEE Computer, pp. 83-93, Apr. 1988.
[9] J. A. Franklin, “Refinement of Robot Motor Skills through Reinforcement Learning,” Proc. Of the 27th IEEE Conf. on Decision and Control, pp. 1096-1101, Dec. 1998.
[10] C. W. Anderson, “Learning to Control an Inverted Pendulum Using Neural Networks,” IEEE Control Systems Magazine, vol. 9, pp. 31-37, Apr. 1989.
[11] R. J. Williams, “Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning,” Machine Learning, vol. 8, pp.229-256, May 1992.
[12] J. S. Albus, “A New Approach to Manipulator Control: The Cerebellar Model Articulation Controller (CMAC),” Trans. ASME, J. Dynamic Syst., Meas., Contr., vol. 97, pp. 220-227, Sep. 1975.