Alan Liu
1/23/18

CPSC 490 Proposal

Title

Machine Learning for Sound Source Localization with Bone Conduction Transfer Functions

Advisor

Professor Steven Zucker, Yale Computer Science Department

Abstract

In everyday settings, ordinary conversations occur in a multitude of environments. For these conversations to be correctly understood, it's important to distinguish voices and sources from each other just as the human brain can with the cocktail party effect. Localization of the sound takes this a step further. Source localization using interaural time differences (ITD) and interaural level differences (ILD) in binaural processing is a fairly developed field[1], but recent advances in using a Convolutional Neural Network (CNN) for image identification have raised the question if similar advances can be used for CNNs in sound analysis in various noisy environments[2]. After the CNN has filtered the sound, tools like independent component analysis (ICA) will be useful for improving localization of sound sources.

Using the results from source localization, there are also ways to directly transmit the sound to listeners in applications such as VR binaural audio using head-related transfer

[1] Kohlrausch A., Braasch J., Kolossa D., Blauert J. (2013) An Introduction to Binaural Processing. In: Blauert J. (eds) The Technology of Binaural Listening. Modern Acoustics and Signal Processing. Springer, Berlin, Heidelberg

[2] M. Bi, Y. Qian, and K. Yu, "Very deep convolutional neural networks for LVCSR", in Proc. Interspeech, pp. 3259–3263, 2015.

functions (HRTF). These functions take a position in space and map them to sounds to be transmitted through air to the left and right ears using different ITD and ILD to provide the same localization effect. However, there has not been extensive research into similar functions in bone conduction, which instead, transmit sound using vibration through the bone needing to be optimized for different levels of ITD and ILD due to a faster speed of transmission.

Proposal

There are two main parts to the project: separating and identifying sound sources, and transmitting binaural audio using bone conduction.

I will focus first on using CNNs with ICA to filter separate audio streams and determine their location. Audio filtering has applications in multiple fields including smart home devices like the Amazon Echo and location is similarly important in hands-free voice control while driving to determine if the driver is attempting to control the car. To test this application, I will be using a publically available binaural audio dataset. I plan to publish the results along with the code.

Afterwards, I will focus on transmitting binaural audio using bone conduction transceivers. Bone conduction is a relatively newer technology, but has already been used in some commercial applications like the Google Glass. As a more powerful, but costlier method to transmit sound, there has yet to be extensive research done on binaural bone conduction devices especially with regard to localization. Similar to the last part of the project, I will create a wearable device that can be used to test localization with two bone conduction speakers and use it to test various methods of localization including different ITD and ILD.

Alan Liu
1/23/18

Plan

Date	**Progress**
January 15th – February 4th (3 weeks)	**Literature Review** – Thoroughly review existing research on source localization, ICA, and bone conduction transfer functions
February 5th – February 18th (2 weeks)	**CNN Design and Implementation** – Plan and create a Tensorflow model using a CNN that uses the available binaural audio dataset to perform ICA. Perform localization and source separation. Suggest improvements on existing work and testing methods.
February 19th – March 4th (2 weeks)	**Experiment Setup/Construction** – Build a wearable device with bone conduction transceivers that can be fed an audio source in order to test sound localization.
March 5th – March 27th (1 week, Spring Break)	**Preparation for Mellon Forum Presentation –** Create a 30 minute presentation on existing progress and demonstrate properties of localization and bone conduction.
March 28th – April 18th (3 weeks)	**Binaural Bone Conduction Location Sensing** – Develop an experimental setup using the wearable device to test various audio delay and intensity settings for location transmission when wearing a bone conduction device
April 18th – May 3rd (2 weeks)	**Final Written Report –** Collect documents and research and present them in a written report.

Deliverables

- Github repository with machine learning code
- Graphs & tables of experimental results
- Prototype of experimental devices
- Written report & poster
- Mellon Forum oral presentation