UTD ATEC ARTSCI LAB
Grey Paper #6
Machine Listening
Extracting Information from Sounds and Signals

April 2 , 2019

Primary Author (Owner) : Sharath Chandra Ram https://orcid.org/0000-0001-7984-9442

Co-authors : a) Scot Gresham Lancaster https://orcid.org/0000-0002-8400-6041
b) Roger Malina - https://orcid.org/0000-0003-3399-3865

Abstract:
Machine Listening is a mode of the data extraction used to extract information and data from sound. While ‘Data Listening’ techniques operate on raw numerical data sets to convert them into an acoustic image, ‘Machine Listening’ does the reverse. Machine Listening techniques operate on signals and audio in the acoustic domain to extract information, features and insights. These extracted data points and features may be further converted to sonic or visual representations. Machine Listening models of audition may also be used to simulate human hearing and allow for the development of novel hearing tests.

Framework of Machine Listening:
Computer algorithms can be modeled to listen to and analyze audio effectively, especially when there are competing sounds and signal sources embedded within noise. The methods of feature extraction and audio analysis that accompany machine listening, are modelled after what we know about the sense of hearing and perception from fields like psychoacoustics (sound perception); cognitive sciences (neuroscience and artificial intelligence); acoustics (physics of sound production); and music cognition (harmony, rhythm & timbre). In this field of machine listening, machines provide a means of isolating , segmenting and deciphering understandable parts of any audio signal. This allows for the capability to use techniques of information extraction upon audio and acoustic representation of signals.

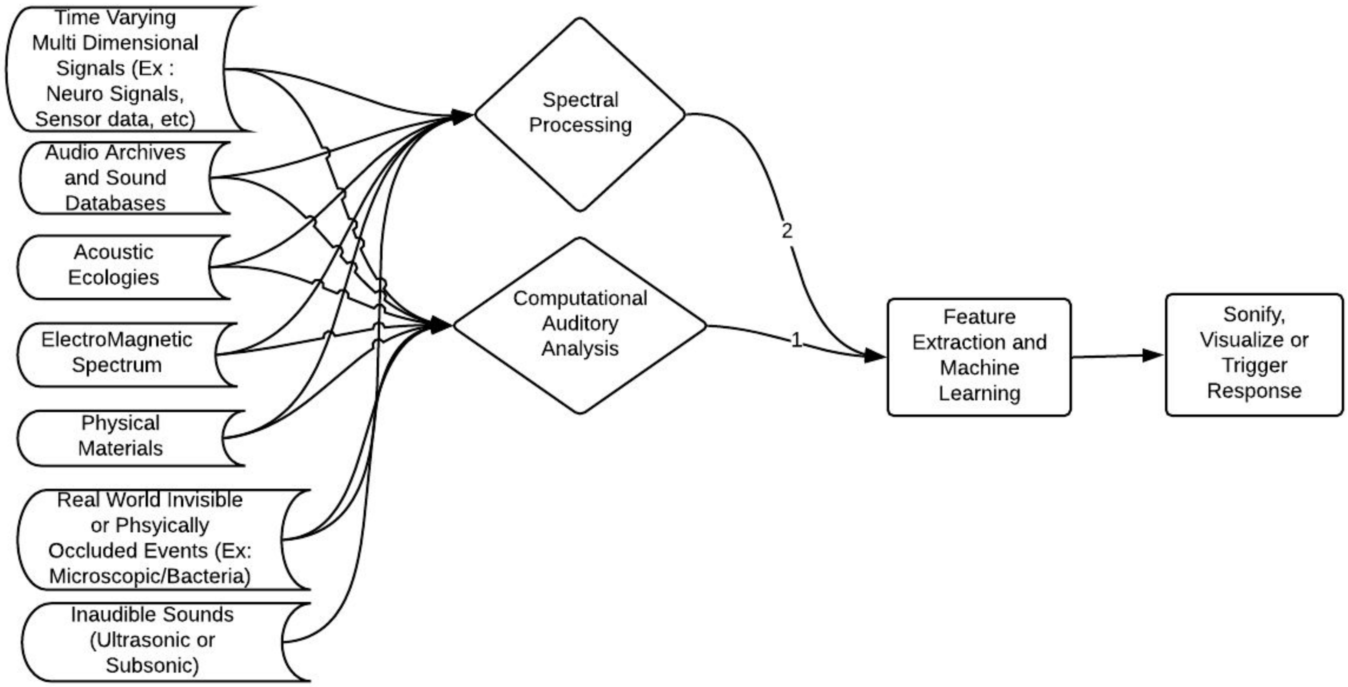

Figure 1 : Framework of Machine Listening

Machine listening presents an opportunity to 'listen' into large audio and signal archives like acoustic ecologies, realtime signal streams such as multi channel EEG data, as well as audio-visual media and podcast archives to effectively sift, sort, search and extract features and insights from them. Also special kinds of signals may be represented and encoded into the audio domain, such as information in the electromagnetic spectrum, material properties of physical objects, electronic signals from sensors and networks, ultrasonic & infrasonic sonic events that are inaudible to humans, and real world events that are not visible to the eye or may be occluded. Further, Machine Listening presents itself in a variety of mission-critical contexts where visualization of data may not be sufficient or could be overwhelming -- for instance, in the case of time varying multi-channel signals where co-variation in signal characteristics may be difficult to decipher visually. Another aspect where Machine Listening comes into play is when listening is performed amongst machines. An instance of this are ultrasonic beacons or signals in sensor networks and internet-of-things (IoT) for data transmission and secure authentication, as well as image sonification used by weather satellite systems to transmit imagery to ground stations.

Research & development in this area involves the use of skills and tools that span the fields of Audio Cognition, Signal Analysis , Artificial Intelligence and Virtual Environment Design. Machine listening interfaces make for new ways and augmented means of fully examining all the aspects of a signal or acoustic data to become achievable.

Conclusion:

Machine Listening promises to add a new and powerful set of infrastructure and tools to the practice of Data listening, where the 'data' in this case is already in the sound domain or are signals that may be processed using the same techniques used to decipher audio signals. As techniques in audio signal processing, AI & machine learning improve and begin to be more and more integrated into our human machine interaction ecology, the use of machine listening will augment the reach of information gathering beyond just the visual and include the multidimensional qualities available via auditory, speech and music perception. Additionally the audio spectrum can be advantaged to allow intermediary audio listening machines to monitor each other and guide information flows and parallel streams of real time data analysis.

References:

1. Auditory Scene Analysis: Perceptual Organization of Sound, Albert S. Bregman.
2. Human and Machine Hearing - Richard F Lyon
3. Computational Auditory Scene Analysis: Principles, Algorithms, and Applications , DeLiang Wang ; Guy J. Brown
4. A. Moorer (1975) On the segmentation and analysis of continuous musical sound by computer, Ph.D. thesis, CS dept., Stanford Univ
5. G. Tzanetakis & P. Cook (2002) “Musical genre classification of audio signals,” IEEE Tr. Speech & Audio Proc. 10(5)