Virtual Student Summer Talks

12:20 pm - 4:20 pm EDT

Jonathan Goesswein

Semi-supervised self-adjustment fine-tuning procedure for hearing aids for asymmetrical hearing loss

The individual fitting of hearing aids is still a challenge and usually requires several sessions. The audiologist typically fine-tunes the hearing aids based on the patient's reported perception. Recent research investigated the alternative of empowering the patient by means of self-adjustment. However, all known studies on self-adjustment procedures have so far focused on symmetric hearing loss and a symmetrical signal modification adjustable by the user. It is therefore still unknown how to deal with severe asymmetric hearing losses. In this study we examined a previously evaluated self-adjustment procedure for symmetric hearing losses with respect to its applicability for asymmetric hearing losses. For this purpose, experienced hearing-aid users with asymmetric hearing loss were fitted with real hearing aids and equipped with a self-adjustment user interface. Each fitting was performed in several realistic sound scenes in two conditions: first, the two hearing aids were fitted separately; second, both hearing aids were fitted in a coupled way and then fine-tuned separately. In addition to the comparison between the gain settings resulting from the self-adjustments the study examined also subjective sound impressions such as the balance of the sound in both ears.

Alaa Algargoosh

The Emotional Impact of Acoustic Environments

This presentation will discuss the emotional impact of acoustic environments shaped by architectural design. Considering that worship spaces are places in which acoustics play a significant role in the listeners' spiritual experience, this research adopts worship spaces as case studies to explore the acoustic characteristics of the space that can enhance the listeners' experience.

Marselle Barbo

Museum Acoustics - Objective Parameters

The architectural project involves the knowledge of variables of fundamental importance for a satisfactory and quality result. It is the activities developed in a building that will guide the program of acoustic comfort needs. To the designer, it is necessary to have technical knowledge of these variables in line with the design manipulation so that the desired acoustic comfort is achieved. This work addresses the acoustic characterization of museums, buildings that have historical and architectural value within the community in which they are inserted. Acoustically characterizing museums makes it possible to guide guidelines for the design of these buildings, meeting the precepts of environmental and acoustic comfort for users. Museums have an important educational role, consequently valuing these buildings goes beyond the principles of museology, such as, for example, the rules for conservation of works and lighting of the building. It is proposed to study objective parameters such as reverberation time, early decay time, definition. In addition, the criteria related to speech privacy will be proposed, such as distraction distance, spatial decay rate of speech and A-weighted sound pressure level of speech at 4 m. These parameters present interesting factors for analysis in museums that do not use audiovisual resources.

Ryan McCarthy

Braid representation of the shallow water acoustic channel to interpret temporal evolution of multipath arrivals

We present a morphological representation of the underwater acoustic channel that employs geometric braids to gain real-time knowledge of the dominant multipath activity. The key idea is to incorporate braid manifolds in tandem with non-convex mixed norm optimization techniques to track the temporal evolution of multipath arrivals. Specifically, we focus on channel braids that manifest as rapidly fluctuating high-energy taps in the delay spread as well as the delay-Doppler scattering function. Multiple braids can also be topologically combined using braid operations to interpret oceanic phenomena such as caustics and surface wave focusing, among others. We present techniques for adaptively updating the channel braids and their overlap patterns to reflect the temporal evolution of the shallow water acoustic channel. We evaluate the performance of proposed morphological channel estimation technique in terms of the normalized prediction error and computational time. Results based on numerical channel simulations based on diverse oceanic and experimental conditions as well as experimental field data from the SPACE08 experiment will be presented.

Bernice Kubicek

Sonar target classification using two-dimensional Gabor stripe features

Sonar target recognition suffers from the nonlinear combination of unpredictable environmental parameters (such as changing sound speed profiles, salinity, and signal deformations) and target geometry (such as shape and orientation). These factors introduce exceptional challenges in sonar target classification. A case-study of our target classification solution is introduced by describing a brief overview of a feature extraction algorithm used to create specific target class representations, or unique Gabor stripes. The algorithm uses an optimally selected two-dimensional Gabor wavelet as a kernel filter prior to a binarization and thresholding of the filtered acoustic color magnitude, Gabor stripes are then uniquely encoded. Classification results were generated from public domain experimental field data and are presented for the optimally selected two-dimensional Gabor filtered acoustic color magnitudes and unfiltered acoustic color magnitudes. This research was funded by the Office of Naval Research under grant number N000014-19-1-2436.

Kendal Leftwich

Estimating the range of Sperm whale clicks in the northern Gulf of Mexico from a single hydrophone

Using Kalman filters, attenuation of sound, measured sea water temperatures, salinities, and depths, combined with recorded sperm whale clicks from 2010 through 2017, we can estimate the range of sperm whales and apply the geographical soundings to produce a three-dimensional plot showing possible locations of the acoustical source. We are also investigating methods determining if the source is moving toward or away from the hydrophone and possibly identifying individual whales.

Lee Drown

Effects of phonetic and indexical variability on talker normalization

Our current work builds on past research demonstrating that listeners experience a processing cost when hearing speech from multiple talkers compared to a single talker. This processing cost is thought to reflect a normalization process during which listeners adjust the mapping to speech sounds to accommodate talker differences in speech production. In the current studies, we use a speeded word identification paradigm to measure processing time for word recognition in single- vs. mixed-talker blocks, and manipulate within-talker and between-talker variability along both phonetic (e.g., vowel formants) and indexical (e.g., fundamental frequency) dimensions. The results to date suggest that listeners incur processing costs given variability in either dimension, even in single-talker blocks, which raises critical methodological considerations for examining talker normalization in addition to informing theories of talker normalization.

Jerry Huang

Immersive Soundscape Reconstruction using Contextualized Visual Recognition and Deep Neural Network

The use of visual environments to generate corresponding acoustic environments has been of interest in audiovisual fusion research. The scope of works involved are currently limited by user-centered virtual reality devices with high computational demands. In this work, an immersive soundscape rendering system is developed using machine-learning-based visual recognition techniques. This system utilizes a hand-crafted panoramic image dataset, with their contents identified using pre-trained neural network models for semantic segmentation and object detection. The recognition process extracts spatial information of sound-generating elements in visual environments that are used to position and orient virtual sound sources and locate corresponding contents in pre-assembled audio datasets that consist of both synthetic sounds and pre-recorded audio. This process facilitates a plausible audiovisual rendering schema that could be presented both in binaural format and at the Collaborative-Research Augmented Immersive Virtual Environment Laboratory (CRAIVE-Lab) at Rensselaer Polytechnic Institute. This work intends to situate and enhance in human-scale and immersive context. This presentation covers in detail the designed and developed audiovisual rendering pipeline for this work and the computational performance of this approach.

Aleese Block

Phoneme-specific patterning in the production and perception of vowel quantity

This talk looks at the extent to which phonetic variation is used in phonological contrast through the lens of vowel quantity in Norwegian, testing the traditional view that vowel duration is the only cue used in quantity perception. I examine the perception of quantity in the vowels /i, u, a/ in 38 native Norwegian speakers at the University of Oslo who completed a 4AIX paired discrimination task with stimuli that were manipulated for both duration and quality. I found that not only in quality used, the extent to which it's used is phoneme specific; high vowels used quality while low vowels did not, and high front /i/ weighted quality heavier than duration. The results build upon previous literature and bring to light questions about our understanding of this contrast as well as the roles of enhancement and phoneme-specific patterns in speech perception.

Jayden Lee

Foreign language talker identification training does not generalize to new talkers

Listeners identify talkers less accurately in a foreign language than their native language, but it remains unclear whether this is due to lack of experience identifying foreign-language talkers or whether linguistic processing reveals additional talker-specific information in speech. Here, we investigated whether two types of training improved the ability to learn to identify foreign-language talkers. Participants completed four days of talker identification training in Mandarin, an unfamiliar foreign language. Participants were assigned to either the ‘same-voices’ condition, in which they trained on the same five voices during Days 1-3, or the ‘different-voices’ condition, in which they learned a new set of five voices on each day. Both groups learned five new voices on Day 4. Talker identification accuracy improved across Days 1-3 for the same-voices condition, but not for those in the different-voices condition. However, talker identification accuracy on Day 4 did not differ from the Day 1 baseline for either group. These results suggest that knowledge about foreign-language talkers is limited to the training set, and that, without specific linguistic knowledge, training on foreign-language talkers does not generalize to improved ability to learn to identify new foreign-language talkers.

Reilly Johnson & Nadia Kim

Articulatory strategies and their acoustic consequences: investigating tongue retraction and lip protrusion tradeoffs in talkers with amyotrophic lateral sclerosis

Observations of trading relations between lip protrusion and tongue retraction during productions of /u/ suggest that typical talkers can vary their articulatory strategies without affecting their acoustic output. In this study we tested the hypothesis that talkers with dysarthria due to amyotrophic lateral sclerosis (ALS) may take advantage of such trading relations and exaggerate lip protrusion to preserve speech acoustics in the presence of impaired tongue retraction. 14 talkers with mild to moderate dysarthria due to ALS (8 females, 6 males) and 14 age- and sex-matched controls produced “Tomorrow Mia may buy you toys again” five times at their habitual rate. Speech kinematics were recorded using electromagnetic articulography. Tongue retraction and lip protrusion were measured by calculating the displacement of the posterior tongue and upper lip sensors in the anterior-posterior dimension of the midsagittal plane during the production of “buy you”. F2 transition extent was measured by calculating the change in F2 (F2 maximum-minimum) during “you.” Group means of kinematic and acoustic variables were compared to determine between-group differences in articulatory and acoustic performance. Linear regressions were used to determine the contribution of each articulator as well as both articulators to F2 transition extent within each group.

Daniel Kim

Dysarthria Subgroups in Talkers with Huntington’s Disease: Free Classification versus Feature-Constrained Classification

In our recent study (Diehl et al., 2019), we examined the speech characteristics of 48 talkers with Huntington’s disease (HD) using the classic feature-based dysarthria rating scale (Darley, Aronson, & Brown, 1969). A cluster analysis based on speech feature ratings revealed four dysarthria subgroups within our cohort of talkers with HD. Talkers within each subgroup shared deviant speech features that set them apart from other talkers with HD. Presumably talkers with similar patterns of deviant speech features should sound alike. In the current study, we will test this notion by using a free classification task. Specifically, we will recruit 20 naïve listeners and ask them to sort the 48 speech samples that were used in the previous study into similar-sounding groups. Each listener’s grouping decision will be submitted to an additive similarity tree cluster analysis to determine dysarthria subgroups based on the free classification task. Moreover, speech features of each dysarthria subgroup in the current study will be determined using the speech feature ratings of the classic dysarthria rating scale. Subgroup findings of the current study will be compared to those of the previous study. Outcomes will provide insights into the saliency of specific speech features in talkers with HD.

Marissa Garcia

Eavesdropping on North Atlantic right whales in Cape Cod Bay: A comparison of acoustic, aerial, and eDNA survey techniques

With fewer than 411 remaining, the North Atlantic right whale (Eubalaena glacialis) is at risk of anthropogenically-induced extinction. They are susceptible to entanglement and vessel strike, causing 15 premature deaths from 2017-2019. In collaboration with Cornell University’s Center for Conservation Bioacoustics, we are documenting the dynamics of right whale presence during peak season in Cape Cod Bay (CCB) between winter and spring, where over 50% of the population is found. Through investigating spatiotemporal distribution patterns, we will compare the efficacy of various survey techniques. The Center for Coastal Studies conducted aerial surveys in CCB. This methodology, although foundational, underestimates right whales during winter, when diving for the copepods Pseudocalanus spp. They are thus less visible at the surface compared to spring, when they forage for surface-dwelling copepods Calanus finmarchicus. Passive acoustic monitoring may eliminate this bias. Marine Autonomous Recording Units

(MARUs), hydrophones, were deployed at five locations, positioned to optimize localization via triangulation of upcalls on at least three hydrophone channels. The upcall, a contact call universally used by males and females of all ages, is a proxy for population. eDNA, a DNA derivative left behind by the whales' blow, feces, or shed skin, is currently being processed to additionally detect presence. Determining the best survey technique will help policymakers redirect boating away from areas with high right whale density. Further directions include analyzing upcall diel patterns and comparing the upcall vocalization rate to other calls.

Emmanuelle Cook

Arctic Ambient Noise Measurements from a Real-time Observatory

The Barrow Straight Real Time Observatory (BSRTO) is a cabled underwater monitoring station operated by Fisheries and Oceans Canada, located in the Tallurutiup Imanga National Marine Conservation Area within the Canadian Arctic Archipelago. In addition to measuring a range of ocean properties, the observatory's hydrophone records a 1-minute pressure time series every 2 hours. Noise spectra with an acoustic bandwidth of 10 – 6390 Hz are then transmitted in near real-time to the Bedford Institute of Oceanography. Ice draft histograms are also transmitted every 6 hours. A nearby weather station in Resolute records hourly wind speed and direction, and atmospheric temperature. Hourly surface currents are provided from the Webtide tidal prediction model. Between August 2018 and September 2019, noise power over the entire measured frequency band ranged between 19 to 115 dB re 1 μPa^2/Hz. Monthly mean noise levels from 10-1000 Hz ranged between 45-75 dB re 1 μPa^2/Hz. Noise power follows seasonal ice variations. The quietest period is in the winter while there is shore fast ice. As mean ice draft starts to decrease in July, noise power increases. Noise power is the highest between August and September, which corresponds to a period of no ice cover. As mean ice draft starts to increase during ice freeze-up, noise power decreases. During ice freeze-up, when ice draft distributions are the most variable, and during complete ice cover, there is a near-24 hour periodic trend in noise at high frequencies linked to tidally forced surface currents.

Calder Robinson

Portability and model data comparison of semi-empirical ambient sound source level models through local transmission loss modelling

This project improves the portability of empirically parameterized wind-noise models in coastal regions by including the local transmission loss and regional environmental forcing biases. Audio recordings (Jan 2018 – present) from 3 regions of the Salish Sea, BC, a complex, shallow water region with heavy surface traffic, were collected and used to parameterize an empirical ambient noise model. Wind speed appears as the dominant term forcing the soundscape above 5 kHz, while the model is improved with the linear inclusion of terms capturing rainfall rate, tidally driven current speed. In-situ model-data comparisons achieve a root mean squared error of less than 3.5 dB re 1 μPa2/Hz. Portability of the wind-noise relationship is evaluated at neighboring sites through model-data comparison and improved by incorporating the effective listening area of the receiver using the locally computed transmission loss. The regional transmission loss models are validated by comparing the computed and measured vertical coherence of the noise field, ensuring the environments are adequately characterized. The increasing contributions of anthropogenic sources in all coastal regions, some without acoustic monitoring, require a modeled baseline for effective monitoring and comparison, necessitating acoustic model portability.

Najeem Shajahan

Quantifying ship-generated sound using the time series of ambient noise vertical coherence

This work presents the advantage of ambient noise analysis based on coherence in quantifying the influence of anthropogenic noise. Wind-generated ambient noise coherence is a stationary noise property of directionality. The variation due to ship noise induced coherence can be used to determine the impact of shipping on the ocean environment. A one-month continuous ambient noise recording off Long Island near Alvin canyon was used in this analysis. The time series of noise vertical coherence mainly consists of wind, distant ship, close-range shipping and flow noise components. Analytical models of ambient noise and sound propagation were used for the simulation of these noise components based on environmental inputs such as sound speed and sediment geo-acoustic properties. The ambient noise model is used for partitioning coherence into wind and shipping components and later used in conjunction with the sound propagation model to invert the relative contribution of shipping in the overall noise field.

Pravinkumar Ghodake

Preliminary Results of Topology Optimization for Design of Mechanical Filter in Nonlinear Ultrasonic Testing

The most common nonlinear ultrasonic technique for damage detection is based on the measurement of higher harmonics generated due to interaction of a monochromatic wave with the damaged material. One of the major challenges during nonlinear ultrasonic experiments is that the measured signals are corrupted by the generation of higher harmonics at the input stage due to the instrument itself. To address this problem, we proposed to design a mechanical filter based on the fact that a periodic arrangement of layers of materials with different acoustic impedance behaves like a filter. The filter is designed to inject a harmonic wave of desired frequency in the specimen. Filter consists of alternate arrangements of bonded disks of two different material and thicknesses but having the same cross-sectional area. The filter is first modeled using one dimensional wave analysis and then optimized to achieve the desired objective by formulating an optimization problem where the design variables are the thickness of the disks.