Aurelien Antoine and Eduardo Miranda (Interdisciplinary Ctr. for Comput. Music Res. (ICCMR), Univ. of Plymouth, The House Bldg. - Rm. 304, Plymouth PL4 8AA, United Kingdom, aurelien.antoine@postgrad.plymouth.ac.uk)

In this paper, we report on the development of a perceptually orientated and automatic classification system of timbre content within orchestral audio samples. Here, we have decided to investigate polyphonic timbre, a phenomenon emerging from the mixture of instruments playing simultaneously. Moreover, we are focusing on the perception of the entire orchestral sound, and not individual instrumental sound. For accessibility to non-Acoustics experts, we chose to use verbal descriptors of timbre quality, such as brightness and roughness, to represent the timbral content of the samples. We based our acoustic analysis on the existing research into the perception and description of timbre. However, with a lack of agreed metrics, we had to establish a comparative scale for each timbre attribute implemented in our system, which is based on an analysis of audio recordings, in order to identify the dominant timbral attribute. To improve the classification accuracy, the system continually calibrates this scale as new audio files are analyzed. Preliminary analysis of our results shows a correlation between the system’s classification and human perception, which is promising for further developments, such as standardizing metrics for perceived responses of timbral attributes or implementing systems for music production tasks.

3pMU2. A model for the observed decorrelation of partials in the overtone spectra of bowed stringed instruments. Sarah R. Smith and Mark F. Bocko (Univ. of Rochester, 405 Comput. Studies Bldg., P.O. Box 270231, Rochester, NY 14627, sarahsmith@rochester.edu)

It has been shown that the overtone frequencies in the spectra of bowed stringed instruments played with vibrato exhibit less than perfect pairwise correlations. However, these results are inconsistent with the mechanism of performing vibrato by changing the length of the string. Since modulating the string length affects the frequencies of all string modes proportionately, it is curious that the overtones exhibit less than perfect correlations. The observed decorrelations, therefore, may be attributed to the filtering of the string’s vibrations by the mechanical-acoustic resonant modes of the instrument body. The exact frequency deviations depend upon the frequencies of instrument’s resonant modes in relation to the string’s overtone frequencies and the width and rate of the vibrato. By modelling the instrument body as a sum of resonant modes driven by a frequency modulated saw tooth wave, we develop an analytical model relating the observed frequency deviations to the modal properties of the instrument. The effect of a single resonant mode on the instantaneous frequency trajectories is found analytically and informs numerical simulations of instruments with multiple modes. The simulated results compare well with data from recorded violin tones.

3pMU3. Characterization of free field radiation in brass instrument horns under swept sine excitation using a linear microphone array. Amaya López-Carromero (School of Phys. and Astronomy, Univ. of Edinburgh, 1608, James Clerk Maxwell Bldg., Peter Guthrie Tait Rd., Edinburgh, Scotland/Midlothian EH9 3FD, United Kingdom, s1374028@sms.ed.ac.uk), Jonathan A. Kemp (Univ. of St Andrews, St Andrews, Fife, United Kingdom), and D. Murray Campbell (School of Phys. and Astronomy, Univ. of Edinburgh, Edinburgh, United Kingdom)

A linear array with 23 microphones is used to scan a planar section of the sound field radiated into an anechoic environment by a range of brass instruments excited by a sinusoidal sine sweep. The planar section contains the symmetry axis of the bell and covers a rectangular area of 0.9 by 0.6 metres, starting at the plane of the bell and extending away from it along the longest side. The linear microphone array is perpendicular to the symmetry axis and is stepped along the axis. The resulting matrix of signals is processed to separate the linear and non-linear parts of the response; the three-dimensional pressure distribution in the sound field is deduced on the assumption that the field is cylindrically symmetric. This data then allows the visualization and further analysis of the frequency dependence of radiated wave fronts in brass instrument bells. Comparison is drawn between the observations and the predictions of several popular radiation models.

3pMU4. Non-invasive measurement of acoustic coupling between the clarinet bore and its player’s vocal tract. Steven M. Lulich (Speech and Hearing Sci., Indiana Univ., 4789 N White River Dr., Bloomington, IN 47404, slulich@indiana.edu)

As acoustically coupled resonators, the bore of a clarinet or other woodwind instrument and the vocal tract of the player interact in ways that affect timbre and pitch. Pitch in particular is strongly dependent on vocal tract acoustics when the bore-tract coupling is strong, such as when a tract impedance maximum is close in frequency and amplitude to a bore impedance maximum. Direct investigation of bore-tract coupling requires invasive measurement of bore and tract input acoustic pressures (or impedances), and one particular technique makes use of the ratio of these pressures (in the frequency domain, Pt/Pb) at harmonics of the reed vibration fundamental frequency. A non-invasive, model-based approach to investigating bore-tract coupling has been developed, which depends on a free-field microphone recording the sound produced by the instrument (P1), and an accelerometer placed against the skin of the neck recording skin vibrations (P2) related to intra-tract acoustic pressures. An additional, model-based calibration step is required. The ratio of these two signals (in the frequency domain, P2/P1) following calibration is qualitatively similar to the ratio Pt/Pb, and approaches quantitative identity as the model-based calibration step improves.
3pMU5. Construction of a finite element model of the Japanese koto. Angela K. Coaldrake (Music, Univ. of Adelaide, Elder Conservatorium of Music, University of Adelaide, Adelaide, SA 5005, Australia, kimi.coaldrake@adelaide.edu.au)

This paper presents the steps in developing a finite element model of the Japanese koto (13-stringed zither) in Combosol Multiphysics® v5.2a and some of the issues encountered. As the instrument is 1.8 meters in length and hand crafted there are many internal irregular shapes. Early attempts at creating a geometry were unsatisfactory. To address this issue a CT scan with 2400 cross-sections was used to measure the internal details. A mesh was created from the scan using Simpleware® software. The result was a mesh with 430,000 elements for the instrument alone, placed in a sphere of air resulting in over 7 million degrees of freedom. This new model therefore has required the use of high performance computing to produce a second of acoustic output. The issue of the physical properties of paulownia, the less well-characterized highly anisotropic wood used to construct the koto, has proven more intractable. Scanning electron microscopy, frequency response and acoustic camera studies of the original instrument provided important insights into paulownia in particular and developing the model in general. A number of studies have been undertaken to validate the model including comparing it with the original instrument. Further studies of the acoustics of the koto are in progress.


The exponential horn is known as the shape realizing the best matching between a source and the external field for frequencies higher than its cut-off frequency. In practice, the horn being of finite length the effective cut-off is significantly higher and resonances appear as waves are reflected at the end of the horn. So, the response of the horn is far from being flat. Then, a question arises: in what extent the shape of an exponential horn has to be strictly respected in order to keep its main acoustical properties. The present paper intend to answer this question. First, a review of the different ways to calculate the input impedance of a horn and their accuracy is made. Comparison with input impedance measurements show that plane wave approximation is often sufficient even when the horn is strongly bended. Second, some criteria are proposed to characterize horns ability to match with the external field. These criteria are finally used to compare the performances of some simplified geometries of horns to that of strictly exponential horns.

3pMU7. Numerical study of nonlinear distortion of resonant states in acoustic tubes. Roberto Velasco-Segura and Pablo L. Rendon (Laboratorio de Acústica y Vibrationes, Centro de Ciencias Aplicadas y Desarrollo Tecnológico, Universidad Nacional Autónoma de México, Circuito Exterior S/N, C. U., Delegación Coyoacán, Mexico City 04510, Mexico, roberto.velasco@ccadet.unam.mx)

A numerical study of nonlinear acoustic propagation inside tubes is presented. Thermoviscous attenuation is included, giving rise to wall losses associated with the boundary layer. The full-wave simulation is performed in the time domain, over a 2D spatial domain assuming axial symmetry, and it is based on a previously validated open source code, using Finite Volume Method implemented in GPU (FiVoNAGI) [Velasco & Rendón, A finite volume approach for the simulation of nonlinear dissipative acoustic wave propagation, 2015]. One intended application is the identification of resonance frequency shifts in the nonlinear regime in brass musical instruments as a function of bore profile and amplitude of the driving stimulus. To gain insight on the nonlinear processes taking place inside the tube, visualizations are presented, differentiating spectral components and traveling waves in both directions.

3pMU8. Use of H1-H2 to quantify formant tuning for notes and portions of the vibrato cycle in the second passagio of a professional female singer. Richard C. Lissencome (Speech-Language-Hearing Sci., Graduate Center/The City Univ. of New York, 499 Fort Washington Ave., Apt. #4F, New York, NY 10033, rlissencome@gradcenter.cuny.edu), Christine H. Shadle (Haskins Labs., New Haven, CT), Kevin Roon, and D. H. Whalen (Speech-Language-Hearing Sci., Graduate Center/The City Univ. of New York, New York, NY)

Singing voice pedagogy emphasizes that an acoustic change occurs in standard, classical, and symphonic sopranos voice between the musical notes D5 (587 Hz) and F5 (698 Hz). For low vowels, this involves a transition from the second resonance to the second harmonic (F2/H2) to tuning of the first resonance to the fundamental (F1/F0). In this single-subject study, we quantified the acoustics of this transition as the amplitude difference between the first and second harmonics (H1-H2). Results showed a clear and substantial change from negative to positive H1-H2 values at a pivot point between E5 and E5, implying the resonance tuning. Non-technical singing, with the same singer, showed no such change. F0 fluctuation (vibrato) of ±90 cent at the pivot point resulted in positive H1-H2 values at vibrato maxima and negative ones at vibrato minima. Additionally, H1-H2 values were consistently higher at vibrato maxima than minima throughout the transition area. Potential explanations for the latter result are: (i) vocal tract resonances are located just above the sung F0, or (ii) the vibrato cycle is accompanied by an articularatory change, possibly laryngeal movement. This illustrates the intricacies of formant tuning and suggests future possibilities for numerical assessment of vocal technique.

3pMU9. Velocity analysis of the vacuum-driven clarinet reed. Carter K. Richard and Whitney L. Coyle (Phys., Rollins College, 1000 Holt Ave., Winter Park, FL 32789, crichard@rollins.edu)

A vacuumed artificial mouth has been assembled and tested to measure reed velocity for a Bb clarinet along the width of the reed. Reed velocity measurements may be useful for better estimation of parameters in vacuum models, such as the relevant surface area of the vibrating reed. Use of a vacuum system instead of a pressurized mouth chamber allows straightforward observation and manipulation of the mouthpiece apparatus. Point measurements of reed velocity were obtained via a laser-Doppler vibrometer directed at the reed surface when artificially blown. Simultaneous high-speed exposures were recorded to visualize reed motion. Preliminary results indicate that the velocity amplitude of any torsional motion in the reed is negligible compared to an asymmetric reed velocity, likely caused by natural limitations of the clarinet ligature. Velocity measurements also indicate that the reed may sometimes rebound against the mouthpiece in its oscillatory period. High-speed exposures support this conclusion by visualizing the reed deformation as it collides with the mouthpiece. This "reboustic" deformation may contribute flow into the clarinet system. Further work will expand this measurement technique for a full grid along the surface of the reed, with various ligature mounts, and will seek to verify experimental measurement with analytic models.

3pMU10. Stepwise regimes in elephant trumpet calls: Similarities with brass instrument behavior. Joël Gilbert (Laboratoire d’Acoustique de l’Université du Maine - CNRS, Ave. Olivier Messiaen, LE MANS 72085, France, joel.gilbert@univ-lemans.fr), Angela Stoeger (Dept. of Cognit. Biology, Univ. of Vienna, Vienna, Austria), Benjamin Charlton (School of Biology & Environment Sci., Univ. College Dublin, Dublin, Ireland), and David Reby (School of Psych., Univ. of Sussex, Brighton, United Kingdom)

Trumpet calls are very loud voiced signals given by highly aroused elephants, and appear to be produced by a forceful expulsion of air through the trunk. Beyond their characteristic “brassy quality” previously attributed to shockwave formation, some trumpet calls are also characterized by stepwise fundamental frequency increase and decrease. Here we used spectral analysis to investigate the frequency composition of trumpet calls from one Asian and one African elephant. We found that the frequency interval between the steps were consistent with resonances expected in the exceptionally long elephant vocal tract. Such stepwise regimes are commonly observed in brass instruments as self-sustained oscillations transiently align on the bore’s resonance frequencies during arpeggios. We suggest that this production
Professional singers are trained to maintain vocal configuration by suppressing laryngeal elevation when singing high notes. The present study investigates the organization of vocal filter in sudden pitch raise condition by examining the corresponding acoustic correlates. More specifically, we are interested in whether this phenomenon may have differential effect on vowels with different height. In the experiment, high and low vowels were embedded in a nonsense Mandarin carrier sentence “kuai-lai-kuai-lai _yi-po” with C4E4G4E4_D4D4 pitch contour. Thirty amateur singers were recorded singing the sentence in four melodic conditions with the pitch intervals between the preceding syllable /lai/ to target word manipulated. The conditions are Micro (E4 to D4), Macro (E4 to D5), all High (every word sung in D5) and Null (only the last three words sung in D5-D5D5 contour). The results show that high pitch singing (Macro, High, and Null) indeed induces formant reconstruction when compared to Micro and the effect is markedly stronger in Macro. Furthermore, high vowel are more susceptible than low vowel and undergoes greater degree of formant reconstruction. The results provide acoustical grounding for the possible interplay between diction and pitch contour.

The voice of soprano singers reaches at the highest notes among female vocalists. Soprano singers usually have low clarity with respect to pronunciation because their jaw joints and mouth shapes tend to stay rigid in order to maintain high notes for a long time. Five formants found in human voice are differently shaped depending on his or her different physical structures. Especially, high clarity in pronunciation has a distinct formant shape from F1 through F5, because the jaw joint and the mouth shape have a large influence on F4 and F5 appearing in the high frequency range. This paper comparatively analyzes the vocal voices of four different Korean soprano singers concerning clarity in pronunciation. The results of acoustic analysis of these four singers shows that the singer A and the singer D show from F1 ~ up to F3 above 2 kHz but F4 and F5 do not appear in this range, and the singer C had somewhat inconsistency with her formant characteristics as a whole. In the singer B, formants from F1 through F5 are distinctly shown even above 2 kHz. The study concludes that the singer B has the best clarity in pronunciation.

The word “’trot” means “walk quickly” and the fox-trot, which refers to a terrestrial mammal. These preliminary observations also emphasize how the generalization of musical acoustic models can provide useful insight into the production of animal vocal signals.

In Korea, the trot began shaping its own style in the 1960s. Then, later in the 1970s, it became more specialized having four-fourths beat rhythm of fox-trot, and the Korean style has been finally established with a strong beat and a unique chopping technique. Although there are many trot singers in Korea, the singer Bae Ho was very special in using the heavy bass accompaniment which was later frequently used by other pop singers. The singing voice of Bae Ho gives us the feeling of special softness and appealingness. In addition, his vocalization features the addition of a deep vibe to the song, so that the listeners feel sympathetic and comfortable. This study compares the singing voice of Bae Ho to that of his mimic singers by examining acoustic characteristics of their voices including amplitude, frequency and duration. The acoustic analysis proved that the singing voice of Bae Ho is clearer and has longer duration of vibration in the bass section than that of mimic singers. Bae Ho has a natural talent in expressing the bass part as well as in the midrange part without changing the tone of his voice, while the mimic singers reveal many unnatural connections between measures.

The aim of the study is to conduct subjective tests on audio excerpt assignment to music genre and to carry out automatic classification of musical and vocal signals with the use of decision algorithms. First, the musicology background of classifying music into styles and genres is discussed. Then, an online survey is created to perform subjective tests with a group of listeners, whose task is assigning audio samples to selected music genres. Next, a set of music descriptors is proposed and all music excerpts are parametrized. For checking parameter redundancy the Principal Component Analysis (PCA) is performed. The created database containing feature vectors is then utilized for automatic music genre classification. Two classifiers, namely: Belief Networks and SMO (Sequential Minimal Optimization Algorithm) are employed for the purpose of music genre classification. The last step of this study is to compare the efficiency of the listeners classification with the automatic music genre classification system designed by the authors. Conducted tests show to what extent listeners’ assignment and the automatic classification results agree. It is also observed that very known performers are often rated without problems. Contrarily, songs of less known artists are more difficult to assign to the given genre.

Hidden Markov Models (HMMs) have been studied and used extensively in speech and birdsong recognition but they are not robust to limited training data and noise. This work present a novel method to training GMM-HMMs with extremely limited data—and possibly noisy—by sharing HMM components and generating more training samples that cover the variation of the models. We propose an efficient state-tying algorithm that takes advantage of unique characteristics of birdsongs. Specifically, the algorithm groups HMM states based on the spectral envelopes and fundamental frequencies, and the state parameters are estimated according to the group assignments. For noise-robustness, prominent time-frequency regions (time-frequency ranges expected
to contain high energy for a particular HMM state) are used to compute the state emitting probability. In Cassin’s Vireo phrase classification using 75 phrase types, the results show that the proposed state-tying algorithm significantly outperforms both traditional state-tying algorithms and baseline HMMs in most training conditions (using 1, 2, 4, 8, and 16 samples). Factors such as number of training data, number of shared components, and level of background noise are also studied in this work.


This research describes a method for dynamic beamforming with a particle filter to localize and track musical instruments in a real-time context. Using a spherical harmonic framework, spherical microphone arrays are able to decompose three-dimensional sound fields into their basic components, which enables detailed analysis and efficient spatial filtering algorithms. In recent years, methods for determining relative source positions around an array using steerable beams have been studied. By creating multiple weighting functions based on spherical harmonic components, many beams can be generated simultaneously and can be used to dynamically track instruments via an iterative process.

3pMU18. Acoustical characteristics of Chinese musical instrument bamboo flute. Linhui Peng (Ocean Technol., Ocean Univ. of China, 238 Songling Rd., Information College, Qingdao, Shandong 266100, China, penglh@ouc.edu.cn) and Tao Geng (Music performance, School of Music of South China Normal Univ., Guangzhou, Guangdong, China)

There are few reports on acoustical study about Chinese music instruments, which is an area worthy of researching. Nowadays there are more and more new and comprehensive ways that can be used to analyze acoustical characteristics and quality of musical instruments, such as experimental modal analysis, finite element software. It is said that music is a kind of a world language. However, the music is an expression of a culture, which is expressed by the musical instrument with its specific acoustical property and character. Therefore, it is necessary to research the acoustical characteristics that can express Chinese culture in the acoustical music instrument study. Bamboo flute is one of the most important musical instruments in any Chinese music ensemble or Chinese orchestra. Acoustical structure and characteristics of bamboo flute tone are researched. Meanwhile, acoustical characteristic features for some main playing technique of bamboo flute are also researched. Then, the identified characteristics of bamboo flute tone related with Chinese music and culture are analyzed.

3pMU19. Intonation detection in a melodic context. Gabriella Marrone (Commun. Disord., Stockton Univ., 101 Vera King Farris Dr., Galloway, NJ 08205, marroneg@go.stockton.edu) and Neil L. Aaronson (Natural Sci. and Mathematics, Stockton Univ., Galloway, NJ)

Listeners with a wide range of formal and informal musical experience were asked to listen to an eight-tone diatonic C Major scale, generated using a piano sample library, in which one of four notes (D4, F4, A4, or C5) would be mistuned in 13 different mounts between -32¢ and +32¢. Listeners were told which note might be mistuned and were simply asked to indicate whether the scale was in-tune or not. Each listener was exposed to each degree of mistuning ten times. The frequency with which they said a scale was in-tune as a function of the degree of mistuning was plotted for each note and listener, to which a three-parameter pseudo-normal distribution (mean, standard deviation, height) was fitted. The standard deviation indicated the sensitivity of the listener to intonation in each case (large deviation implied low sensitivity to intonation). Listeners were then ranked based on their musical background, training, and experience. The effect of musical training on intonation sensitivity was a significant factor ($p<0.001$). There was also a significant effect of the particular note on the sensitivity of listeners, with the intonation of A4 being most difficult to detect across listeners ($p<0.001$).

3pMU20. Exploring some questions on occlusion effect in the human auditory system when a musician or singer’s external ear canal is blocked. Amitava Biswas (Speech and Hearing Sci., Univ. of Southern MS, 118 College Dr. #5092, USM-CHS-SHS, Hattiesburg, MS 39406-0001, Amitava.Biswas@usm.edu)

Sometimes some musicians and singers prefer to use their palm or other objects to cover or occlude at least one ear during performance. This practice may be helpful to enhance their self monitoring of the sound production using the occlusion effect. The basic occlusion effect in the human auditory system has been explored and reported in the literature by several individuals. According to those reports, the musician or singer can hear his or her own voice or musical instrument significantly louder when the ear canal is blocked at the outer end. Many clinicians routinely utilize detectability of vibrations from a tuning fork when placed on the mastoid process and the ear canal is occluded. This is popularly known as the Bing effect. These empirical data suggest existence of an efficient acoustic connectivity from the vocal tract to the ear canal for healthy individuals. Therefore, this study will explore quantified effects of the classic Bing effect for normal healthy musicians and singers across the audio spectrum.