University of Miami Frost School of Music
Spring 2008

MMI 505/606
Current Topics in Audio Analysis and Signal Processing

Colby Leider, Ph.D.
3 Credit Hours
Office Hours: F 10-12

Course Description

MMI 505/606 surveys recent topics related to audio analysis, synthesis, and signal processing with an emphasis in software programming and practical applications.

Prerequisites

MMI 504 (undergraduate) or MMI 601 (graduate).

Course Objectives

Students in MMI 505/606 will accomplish three primary goals: (1) gaining familiarity and facility with current tools (both theoretical and practical) used in audio analysis, synthesis, and signal processing; (2) development of critical thinking skills and competence in current audio research techniques; and (3) gaining experience in audio software design, programming, and research methodologies.

Instructional Methodology

MMI 505/606 will be taught as an investigational seminar for advanced undergraduates and graduate music engineering students. The course will examine nine primary areas: current audio APIs and plug-in architectures, modern theories of musical timbre, DSP hardware programming, machine listening, spatial audio, digital audio effects, new digital audio synthesis techniques, machine-musician interaction modalities, and psychoacoustic testing and data mining of multidimensional data sets. The course will require a good deal of reading and a good deal of C++ and MATLAB programming. (We will have a programming assignment approximately every week or two).

Class time will be used for traditional lectures as well as student presentations of the articles, discussions, and laboratory exercises. Occasional pop quizzes may be used to emphasize the required reading. Each week, we will discuss the required reading in class. Please come prepared for each class meeting.

Required Texts

In lieu of a textbook, a course packet of articles is available for you to download and/or photocopy. I will announce details in class.

Additional reference materials in the library and/or my office:

1. Gold, B., and N. Morgan. 2000. *Speech and Audio Signal Processing*. New York: Wiley.

2. Kim, H.-G., N. Moreau, and T. Sikora. 2005. *MPEG-7 Audio and Beyond*. New York: Wiley.
3. Zölzer, U. 1997. *Digital Audio Signal Processing*. New York: Wiley.
4. Roads, C. 1996. *The Computer Music Tutorial*. Cambridge, Massachusetts: MIT Press.
5. Lane, J., et al. 2001. *DSP Filters*. Indianapolis: Prompt/Sams.
6. Rumsey, F. 2001. *Spatial Audio*. Oxford: Focal Press.
7. Deutsch, D., ed. 1998. *The Psychology of Music*, 2nd ed. San Diego: Academic Press.
8. Steiglitz, K. *A Digital Signal Processing Primer*. Menlo Park, California: Addison-Wesley.
9. Cook, P. R. 2002. *Real Audio Synthesis for Interactive Applications*. Wellesley, Massachusetts: AK Peters.

Examinations

One take-home examination will be given during the semester. A final project is due at the conclusion of the course in lieu of a final exam.

Grading Policy

Attendance, Participation, and Quizzes		10%
Programming Assignments and Homework		30%
Take-Home Examination		20%
Presentations	10%	
Final Project		30%

Students will be required to present additional material throughout the semester and must write a final article as part of their final project.

Class Topics

Introduction

DSP review
Quick evaluation quiz (not graded)

Part One: Audio Programming

Modern Audio APIs
Audio File Formats
Plug-in architectures (VST and AU)

Assignment One:

Selectable low-pass/high-pass filter plug-in

Assignment Two:

Standalone executable in C/C++ that synthesizes and mixes sinusoids and can read/write sound files

Reading:

Bencina, R., and P. Burk. 2001. "PortAudio: An Open Source Cross Platform Audio API." *Proceedings of ICMC 2001*.
Scavone, G. P. 2002. "RtAudio: A Cross-Platform C++ Class for Real-Time Audio Input/Output." *Proceedings of ICMC 2002*.
OpenAL Programmer's Guide. http://www.openal.org

Music and Sound Effects System. http://www.fmod.org/
Apple CoreAudio and AU documentation at
http://developer.apple.com/audio/
Steinberg VST documentation at http://www.steinberg.de

Part Two: Timbre

Timbre classification and taxonomy
Feature extraction
Fourier analysis
Wavelet analysis
Cepstral analysis
Multidimensional Scaling
Timbral Similarity
Assignment Three:
Standalone C++ app that can extract a standard audio feature from an input sound file
Reading:
Grey, J. 1977. "Multidimensional Perceptual Scaling of Musical Timbres." *J. Ac. Soc. Amer.* 61:1270–1277.
Iverson, P., and K. Krumhansl. 1992. "Isolating the Dynamic Attributes of Musical Timbre." *J. Ac. Soc. Amer*. 94(5):2595–2603.
McAdams, S., et al. 1995. "Perceptual Scaling of Synthesized Musical Timbres: Common Dimensions, Specificities, and Latent Subject Classes." *Psychological Results* 58:177–192.
Wessel, D. 1979. "Timbre Space as a Musical Control Structure." *Computer Music Journal* 3(2):45–52.
Tzanetakis, G., G. Essl, and P. R. Cook. 2001. "Audio Analysis Using the Discrete Wavelet Transform." *Proceedings of the Conference on Acoustics and Music Theory and Applications.*

Part Three: Programming a Digital Signal Processor

SHARC ADSP-21364
In-class handouts
Assignment Four:
Real-time hardware-based vocal sound processor

Part Four: Machine Listening

Modeling the cochlea
Tempo and beat detection
Pitch detection
Automated music transcription; OCR
Models of perceptual similarity (melodic, timbral, and rhythmic)
Automated semiotic analysis of music
Dissonance analysis
Content retrieval
Source segregation
Assignment Five:
Code and evaluate a pitch or tempo tracker in MATLAB
Reading:
Lane, J. 1990. "Pitch Detection Using a Tunable IIR Filter." *Computer Music Journal* 14(3):46–59.

Andrew Choi. 1997. “Real-Time Fundamental Frequency Estimation by Least-Square Fitting.” *IEEE Transactions on Speech and Audio Processing* 5(2).
Slaney, M., and R. Lyon. 1990. “A Perceptual Pitch Detector.” *Proceedings of ICASSP 1990* pp. 357–360, vol. 1
Scheirer, E. 1998. “Tempo and Beat Analysis of Acoustic Musical Signals.” *J. Ac. Soc. Amer.* 103(1):588–601.
Foote, J. 1999. “An Overview of Audio Information Retrieval.” *Multimedia Systems* 7(1):2–10.
Tzanetakis, G., and P. R. Cook. 2000. “Audio Information Retrieval (AIR) Tools.” *Proceedings of ISMIR 2000*.
Casey, M. A., and A. Westner. 2000. “Separation of Mixed Audio Sources by Independent Subspace Analysis.” *Proceedings of ICMC 2000*.

Part Five: Spatial Audio

Psychoacoustics of spatial perception
Multichannel recording
Multichannel playback
Multichannel reverberation
Real-time sound spatialization
Head-related transfer functions
Ambisonics
Wave Field Synthesis

Assignment Six:

Create an ambisonic decoder or HRTF-based spatializer in MATLAB

Reading:

http://www.york.ac.uk/inst/mustech/3d_audio/
Malham, D. 2000. “Ambisonics.” Paper presented at the Sound in Space Symposium, UCSB, 18–19 March.
Gerzon, M. A. 1985. “Ambisonics in Multichannel Broadcasting and Video.” *J. Audio Eng. Soc.* 33(11):859–871.
Malham, D.G., and A. Myatt. “3-D Sound Spatialization using Ambisonic Techniques.” *Computer Music Journal* 19(4):58–70.
Kendall, G. 1995. “A 3-D Sound Primer: Directional Hearing and Stereo Reproduction.” *Computer Music Journal*
Cheng, C. I., and G. H. Wakefield. 1999. “Introduction to Head-Related Transfer Functions (HRTF’s): Representations of HRTFs in Time, Frequency, and Space.” *Proceedings of the 107th Audio Engineering Society (AES) 107th Convention*, New York.
Cheng, C. I,, and G. H. Wakefield. 2001. “Moving Sound Source Synthesis for Binaural Electroacoustic Music Using Interpoplated Head-Related Transfer Functions (HRTFs).” *Computer Music Journal* 25(4):57–80.
Clozer, C. “The Gmebaphone Concept and the Cybernophone Instrument.” *Computer Music Journal* 25(4).
A. J. Berkhout, D. de Vries, and P. Vogel. 1993. “Acoustic Control by Wave Field Synthesis.” *J. Acoust. Soc. Am.* 93:2764–2778.

Part Six: Digital Audio Effects

The phase vocoder
Timbre morphing
Reverberation methods

Assignment Seven:

Implement an FFT-based analysis/resynthesis sound processor

Reading:

Flanagan, J. L., and R. M. Golden. 1966. "Phase vocoder." *Bell Syst. Tech. J.* 45:1493–1509.

Dolson, M. 1986. "The Phase Vocoder: A Tutorial." *Computer Music Journal* 10(4):14–27.

Laroche, J., and M. Dolson. "New Phase-Vocoder Techniques for Pitch-Shifting, Harmonizing, and Other Exotic Effects." *Proc. WASPAA 1999*. New York: IEEE, pp. 91–94.

Schroeder, M. R., and B. F. Logan. 1961. "'Colorless' Artificial Reverberation." *IRE Trans. Audio* 9(6):209–214.

Tellman, E.; L. Haken; B. Holloway. 1995. "Timbre Morphing of Sounds with Unequal Number of Features" *J. Audio Eng. Soc.* 43(9).

Fitz, K., et al. 2003. "Cell-utes and Flutter-Tongued Cats: Sound Morphing Using Loris and the Reassigned Bandwidth-Enhanced Model" *Computer Music Journal* 27(3):44–65.

Cook, P. R. 1998. "Toward the Perfect Audio Morph? Singing Voice Synthesis and Processing." *DAFX 1998 Proceeding*. Available online at http://ccrma-www.stanford.edu/~serafin/UVA/MUSI445/Cook98.pdf.

Essl, G., Serafin, S., Cook, P. R., Smith, J. O., "Theory of Banded Waveguides," Computer Music Journal, 28:1, 37-50, 2004.

Essl, G., Serafin, S., Cook, P. R., Smith, J. O., "Musical Applications of Banded Waveguides," Computer Music Journal, 28:1, 51-63, 2004.

Part Seven: Digital Audio Synthesis

A brief history
Spectral modeling techniques
Physical modeling techniques

Assignment Eight:

Create a tunable plucked string or other physical model in C++

Reading:

De Poli, G. 1983. "A Tutorial on Digital Sound Synthesis Techniques." *Computer Music Journal* 7(4):429–447.

Smith, J. O. 1992. "Physical Modeling Using Digital Waveguides." *Computer Music Journal* 16(4):74–91.

Jaffe, D. A., and J. O. Smith. 1983. "Extensions of the Karplus-Strong Plucked-string Algorithm." *Computer Music Journal* 7(2):56–67.

Cook, P. R., and G. P. Scavone. "The Synthesis Toolkit (STK)." *Proceedings of ICMC 1999.*

Smith, J. R. 2005. "Virtual Acoustic Musical Instruments: Review of Models and Selected Research." Keynote presentation, WASPAA 2005, 19 October.

Scavone, G., and P. R. Cook. 1998. "Real-Time Computer Modeling of Woodwind Instruments." *Proceedings of the 1998 International Symposium on Musical Acoustics ISMA-98*, pp. 197–202.

Part Eight: HCI and Auditory Display

The "musical" interface
New musical instruments
Haptics
Sensors

Sonification
Earcons and auditory icons
Assignment Nine:
Experiments with Fitts' Law
Making a simple audio-synthesis controller
Reading:
Skim through *NIME 2007 Proceedings*. Available online at http://itp.nyu.edu/nime/2007.
Cook, P. 2001. "Principles for Designing Computer Music Controllers." *Proceedings of the 2001 CHI Conference.*
Cook, P., and C. Leider, "SqueezeVox: A New Controller for Vocal Synthesis Models," *Proceedings of the 2000 International Computer Music Conference*, Berlin, Aug. 2000.

Part Nine: Data Mining and Multidimensional Data Sets
Psychoacoustic testing
Clustering and unsupervised machine learning
Visualization of multidimensional data
Assignment Nine:
Visualization and analysis of listening test data
Reading TBA

Final Project Presentations

Attendance Policy

You must attend class.

Honor Code

Students will be bound by the University of Miami Honor Code. All reports, papers, written assignments, test papers, and examination papers must include a signed honor pledge that states: "On my honor, I have neither given nor received any aid on this assignment." Academic dishonesty may result in a lower grade or a failing grade for the entire course.

Disabilities

Any student with a documented disability (e.g., physical, learning, psychiatric, visual, aural, etc.) who needs to arrange reasonable accommodations must contact the instructor and Disability Services at the beginning of the semester.