

MAT 240F - Digital Audio Programming: Audio Analysis and Music Information Retrieval (Spring, 2008)

The MAT 240 course sequence is a six-part (two-year) practical programming course; it consists of hands-on software development devoted to digital audio and multimedia applications. Students read a selection of papers from the literature, with the emphasis on learning to use and extend the current state-of-the-art programming methods, tools, and programming interfaces. Class assignments involve C/C++/Java programming on Linux, Macintosh, MS-Windows, various plug-in APIs, and other platforms.

The focus of the MAT 240F course is on audio analysis and signal processing techniques applied to sound/music databases and music information retrieval systems. We will work with libraries for signal analysis and feature extraction to develop skills in time-domain processes such as beat following, tempo analysis, and song segmentation, and in spectral-domain analysis techniques such as pitch estimation, spectral peak analysis and tracking, and instrument signature identification. The topic of feature vector design will play an important role in the development tasks. Applications will include music segmentation, finger-printing, thumb-nailing, clustering and genre classification, and user preference matching.

Students are expected to know the basics of digital audio signal representation and processing, and to be proficient in C, C++, or Java (Smalltalk, SuperCollider or LISP are a plus). Grading will be on the basis of in-class participation and programming projects.

Course Outline

- Multimedia Database Applications
- Signal Processing for Feature Extraction
- Time-domain, Frequency-domain Analysis
- Other Kinds, Cross-domain analysis
- Audio Segmentation and Musical Form
- Clustering and Classification
- Handling of Large or Dynamic Feature Vectors
- Application Requirements and Design

Instructor

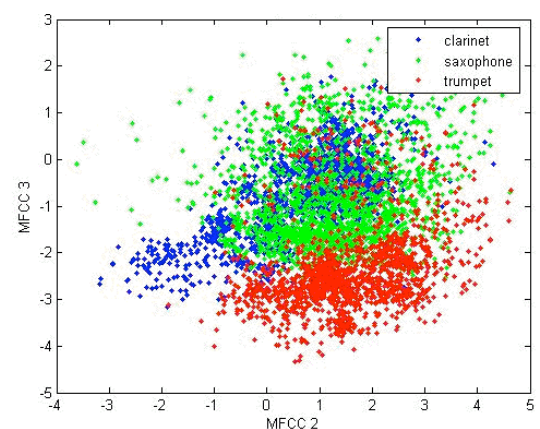
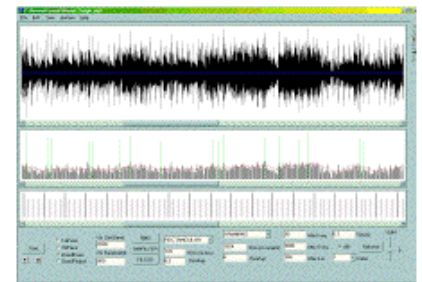
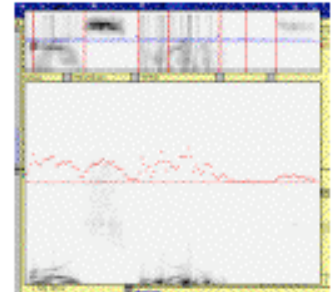
- Stephen T. Pope (stephen@mat.ucsb.edu)

Meeting time and place

- Tues/Thurs 2:00 - 3:50 PM, Music 2215

Electronic Resources

- Course Web Site: <http://www.create.ucsb.edu/240>
- Email Mailing List
See <http://www.mat.ucsb.edu/mailman/listinfo/240> to join



MAT 240F Reader Contents

Reader Sections

- Introductions, Overviews, Tutorials
- Applied DASP Techniques
- Numerical and Mid-level Processing: Segmentation, Pattern Recognition, Clustering, Classification
- Databases and Applications
- Tools and APIs

Introductions, Overviews, Tutorials

- Foote, J. 1999. "An Overview of Audio Information Retrieval." *Multimedia Sys*, 7(1): 2-11.
- Tzanetakis, G. 2002. "Music Information Retrieval Tutorial." *Proc. ISMIR 2002*. slides + extended annotated bibliography
- Pope, S. T., P. Roy and N. Orio. 1999. "Content Analysis and Queries in a Sound and Music Database" *Proc. ICMC 1999*.
- Tzanetakis, G., and P. Cook. 2000. "Audio Information Retrieval Tools." *Proc. ISMIR 2000*.
- Typke, R., F. Wiering, and R. Veltkamp. 2005. "A Survey of Music Information Retrieval Systems." *Proc. 2005 ISMIR*.
- Wright, M, et al. 2000. "Analysis/synthesis Comparison." *Organised Sound* 5(3): 173–189

Applied DASP Techniques

- Pope, S., Holm, F., and Kouznetsov, A. 2004. "Feature extraction and database design for music software." *Proc. ICMC 2004*. paper + slides
- Klapuri, A. 2003. Pitch and Multipitch Estimation. slides
www.cs.tut.fi/sgn/arg/summer2000/klap/mpe.pdf
- Meng, L., 2004. Pitch Detection slides.
www2.cs.uregina.ca/~gerhard/courses/Audio/PitchDetection.pdf
- de la Cuadra, P. 2005. *Pitch Detection Methods Review*.
<http://www-ccrma.stanford.edu/~pdelac/154/m154paper.htm>
- Serra, X., *Spectral Analysis, Peak Detection, Peak Continuation*.
<http://www.iaa.upf.es/~xserra/articles/msm>, <http://www.iaa.upf.es/~xserra/articles/spectral-models>
- Schwarz, D., and X. Rodet. 1999. "Spectral Envelope Estimation and Representation for Sound Analysis-Synthesis." *Proc. 1999 ICMC*. slides
- Hosom, J-P., R. Cole, and M. Fanty. 1999. *Speech Recognition Using Neural Networks*.
http://cslu.cse.ogi.edu/tutordemos/nnet_recog/recog.html
- Kasprzak, W., A. Okazaki, and A. Kowalski. 2006. "ICA-Based Speech Features in the Frequency Domain." *Proc ICA 2006*.
- Klapuri, A. 2001. "Means of Integrating Audio Content Analysis Algorithms." *Proc. 110th AES Conv.*

Numerical, Mid-level Processing: Segmentation, Clustering, Classification

- Ellis, D., 2003. Pattern Recognition Applied to Music Signals. CLSP Summer School slides.
www.ee.columbia.edu/~dpwe/muscontent/JHU-2003-07.pdf
- Dannenberg, R., and N. Hu. 2002. "Pattern Discovery Techniques for Music Audio." *Proc. ISMIR*.

- Tzanetakis, G., and P. Cook. 1999. "Multifeature audio segmentation for browsing and annotation." *Proc. IEEE WASPAA 99*.
- Foote, J., and M. Cooper. *Media Segmentation using Self-Similarity Decomposition*. www.fxpal.com/people/cooper/Papers/SPIE02.pdf
- Lu, L., S. Li, and H-J Zhang. "Content-Based Audio Segmentation using Support Vector Machines." *Proc. ICME'01*
- Lefèvre, S., B. Maillard, and N. Vincent. 2002. "A Two-level Classifier Process for Audio Segmentation." *Proc. IAPR Int'l. Conf. on Pattern Recognition*.
- Frühwirth, M., and A. Rauber. 2001. "Self-Organizing Maps for Content-Based Music Clustering." *Proc. 12th Italian Workshop on Neural Nets*. paper + slides
- Rudi Cilibrasi, R., P. Vitanyi and R. de Wolf. 2004. "Algorithmic Clustering of Music Based on String Compression." *Computer Music Journal* 28(4): 49-67
- Pope, S. T. 2004. "FMAK 2.0 Analysis and Clustering Utilities." FASTLab, Inc. <http://fastlabinc.com>
- Langner, J., and W. Goebel. 2003. "Visualizing Expressive Performance in Tempo-Loudness Space." *CMJ* 27(4)

Databases and Applications

- Bloch, J., and R. Dannenberg. 1985. "Real-Time Computer Accompaniment of Keyboard Performances." *Proc. 1985 ICMC*
- Bello, J-P., C. Monti, and M. Sandler. 2000. "Techniques for Automatic Music Transcription." *Proc ISMIR 2000*.
- Klapuri, A., 2003. "Automatic of Music Transcription as We Know it Today." *JNMR* 33(3)
- Peltonen, V., et al. 2004. *Computational Auditory Scene Recognition*.
- Tzanetakis, G., and P. Cook. 2001. "Musical Genre Classification of Audio Signals," *Proc. ISMIR*.
- S. Dixon., E. Pampalk, G. Widmer 2003. "Classification of dance music by periodicity pattern." *Proc. ISMIR*.
- Scheirer, E., (1998) "Tempo and Beat Analysis of Acoustic Musical Signals." *JASA*. 103(1)
- Cemgil, A., et al. 2001. "On Tempo Tracking: Tempogram Representation and Kalman Filtering." *JNMR* 28(4).
- Martin, K., and Y. E. Kim. 1998. "Musical instrument identification: A pattern-recognition approach." *Proc. 136th ASA*.
- Pope, S. T. 2004. EMA: The Expert Mastering Assistant. FASTLab, Inc. slides + description. <http://fastlabinc.com>
- Haitsma, J., and T. Kalker. 2002. "A Robust Audio Fingerprinting System" *Proc. ISMIR 2002*.
- Lamere, P. 2005. *Search Inside the Music*. Slides, Sun Labs.
- MusicIP Inc. 2006. *Open FingerPrint Architecture White Papers*. <http://www.musicdns.org>

Tools and APIs

- Pope, S. T. 2004. *The FASTLab Music Analysis Kernel (FMAK) System Description and Application Developer's Notes and FMAK Driver Source Code Example*.. FASTLab, Inc. <http://fastlabinc.com>
- Beauchamp, J., 2000. *SNDAN User Guide and MQ Analysis Source Code Example*.
- Tzanetakis, G. 2006. *MARSYAS 0.2 Reference Manual*.