Stephen Sinclair

***Music Technology Area*, McGill University. Montreal, Canada**

Annotated bibliographies for presentations in MUMT 611, Winter 2006

- *Presentation 4:* **Musical Genre Similarity**

Aucouturier, J.-J. and F. Pachet. 2002. Music similarity measures: What's the use? *Proceedings of the 3rd International Symposium on Music Information Retrieval*. [online]

> This paper uses a Gaussian mixture model for estimating the MFCC-derived "global" timbre of a song. It also includes a discussion of evaluation methods for similarity measurements. Similarity between same song, same artist, and same genre are mentioned. The paper addresses the question of what kind of automatic similarity measures can actually be useful when interacting with the subjective opinions of an average user.

Aucouturier, J.-J. and F. Pachet. 2003. Representing musical genre: A state of the art. *Journal of New Music Research. 32*(1), 83-93. [online]

> This paper is a literature review of the field of genre classification as it was in 2003, which is still very useful for an overview of the various methods. It is a good starting point for any researcher in the field. Discussed are the two main approaches to handling taxonomy, three typical ways of addressing feature extraction, including timbre-, rhythm-, and pitch-based features, and finally a discussion of collaborative filtering and textual data mining techniques.

Bağci, U. and E. Erzin. 2005. Boosting classifiers for music genre classification. *International Symposium on Computer and Information Sciences*. [online]

> This paper discusses the technique of "boosting" classifiers for use in genre classification. "Boosting" is a method of training multiple classifiers and combining them to improve over-all classification performance. The AdaBoost algorithm is employed on various Gaussian Mixture Model classifiers to improve accuracy by about 2 to 10%.

Bergstra, J., N. Casagrande, and D. Eck. 2005. Genre classification: Timbre- and rhythm-based multiresolution audio classification. MIREX genre classification contest. [online]

> This paper is discusses the methods use for the winning entry in the 2005 MIREX competition in the area of genre classification. They used a relatively large number of timbral features, and employed the AdaBoost algorithm on a small decision tree classifier. The best result from this algorithm was an accuracy of 86.92% on the competition data set.

Dannenberg, R. B., B. Thom, and D. Watson. 1997. A machine learning approach to musical style recognition. *Proceedings of the 1997 International Computer Music Conference*, pp. 344-7. International Computer Music Association. [online]

> This paper looked at the use of machine learning techniques to analyse and separate instrumental playing style in symbolic MIDI data. They employed the use of Bayesian filters, a linear classifier, and neural networks. The NN approach was the most successful.

Foote, J. T. 1997. Content-based retrieval of music and audio. In C.-C. J. Kuo et al. (Eds.), *Multimedia Storage and Archiving Systems II, Proceedings of SPIE*, Volume 3229, pp. 138-47. [online]

> This was the earliest paper found which addressed the idea of analysing frames of an audio recording into feature vectors and subsequently use a distance classifier to determine similarity. They used MFCC values as features, and used Euclidean and Cosine distance metrics to build a tree structure intended for fast searching.

Jiang, D.-N., L. Lu, H.-J. Zhang, J.-H. Tao, and L.-H. Cai. 2002. Music type classification by spectral contrast feature. *Proceedings of IEEE International Conference on Multimedia and Expo*. [online]

> This paper discusses the advantages of using Octave-based Spectral Contrast feature vectors as a replacement for the popular MFCC feature vector. They suggest that spectral contrast, which considers peaks and valleys of spectral subbands, can be used to estimate the distribution of harmonic and inharmonic components. This is meant to improve over the averaging that takes place in MFCC representation.

Kranenburg, P. V. and E. Backer. 2004. Musical style recognition - a quantitative approach. *Proceedings of the Conference on Interdisciplinary Musicology*. [online]

> This paper examines the use of classification techniques for determine musical style within the Classical genre. Some interesting features used were Pitch Entropy, Voice Density, and Intervals. Several standard classifiers were tried. It is written from a more musicological point of view than the audio-based papers listed here and may be of interest to researchers in that area.

Logan, B. and A. Salomon. 2001. A music similarity function based on signal analysis. *Proceedings of the IEEE International Conference on Multimedia and Expo*. [online]

> This paper preceded many others in using spectral signature to evaluate musical similarity. It builds on the work of Foote to better make use of MFCC features by creating a "personalized" signature for each song instead of allocating histogram bins across a whole data set. It also includes a discussion of automatic playlist generation.

Mandel, M. I. and D. P. Ellis. 2005. Song-level features and support vector machines for music classification. MIREX genre classification contest. [online]

> This 2005 MIREX submission uses a set of MFCC Gaussians as features, comparing them using Kullback-Leibler divergence. Support Vector Machines (SVM) are used as the final classifier. It came in third place after Bergstra et al.

McKay, C. and I. Fujinaga. 2004. Automatic genre classification using large high-level musical feature sets. *Proceedings of the International Conference on Music Information Retrieval*. [online]

> This paper discusses the use of classification for determining musical genre based on symbolic data. Encouraging results showed that symbolic data can generally provide much information about genre that is not available from low-level timbral characteristics.

McKay, C., R. Fiebrink, D. McEnnis, B. Li, and I. Fujinaga. 2005. ACE: A framework for optimizing music classification. *Proceedings of the International Conference on Music Information Retrieval*. 42-39. [online]

> This paper presents a system called ACE, which is a Java-based engine for automatic classification of musical data. It can experiment on a given data set using multiple combinations of feature vectors and classifiers to determine maximal performance. It has several advantages over similar systems which are specific to music research.

McKay, C., and I. Fujinaga. 2005. Automatic music classification and the importance of instrument identification. *Proceedings of the Conference on Interdisciplinary Musicology*. CD-ROM. [online]

> This paper discusses the need for high-level instrument identification in addition to simple low-level timbral features. It discusses the use of genetic algorithms for determining feature sets of higher importance, discovering that instrumentation-related features often rise to the top.

McKay, C. 2004. Issues in automatic musical genre classification. Presented at the *McGill Graduate Students Society Symposium*. [online]

> This paper presents an in-depth overview of many problems often encountered when dealing with automatic genre classification. Problems related to taxonomy and feature identification are discussed. This is certainly recommended reading for new researchers in the field who wish to avoid common pitfalls in classification tasks.

Pampalk, E. 2001. Islands of music: analysis, organization, and visualization of music archives. Master's thesis, Vienna University of Technology. [online]

> This master's thesis discusses an implementation of a system for automatic genre classification and the metaphor of "islands" for organization and visualization of this data.

Pampalk, E. 2004. A matlab toolbox to compute music similarity from audio. *Proceedings of the International Conference on Music Information Retrieval*. [online]

> This paper presents a toolbox of Matlab functions for computing feature vectors and distance measures for musical similarity. Among provided functions are computation of spectrum histograms, periodicity histograms, and fluctuation patterns.

Pampalk, E. 2005. Speeding up music similarity. MIREX genre classification contest. [online]

> This 2005 MIREX submission uses two features derived from fluctuation patterns in addition to MFCC-derived features. It placed 6th overall in the genre classification category.

Pampalk, E., S. Dixon, and G. Widmer. 2003. On the evaluation of perceptual similarity measures for music. *Proceedings of the International Conference on Digital Audio Effects*. [online]

> This paper discusses some methods for evaluation of perceptual similarity measures. They performed a large-scale and small-scale comparison between 5 recently-published similarity measures (as of 2003), using expert groupings from the All Music Guide as baseline reference. They found that spectrum histograms outperformed all other measures.

Pampalk, E., A. Flexer, and G. Widmer. 2005. Improvements of audio-based music similarity and genre classification. *Proceedings of the International Conference on Music Information Retrieval*. [online]

> This paper discusses in further depth the fluctuation patterns features used in their MIREX entry. Two derived similarity features are introduced, called Focus and Gravity. An analysis of performance on the Magnatunes database is included. Additionally, the necessity of ensuring that the same artist is not present in both training and testing data sets is discussed.

Ponce de León, P. J., C. Pérez-Sancho, and J. Iñesta. 2004. A shallow description framework for music style recognition. In A. Fred et al. (Eds.), *Lecture Notes in Computer Science*, Volume 3138, pp. 876-884. Springer-Verlag. [online]

> This paper examines the use of classification techniques on symbolic data to model musical structure and separate musical style. The categories of Jazz and Classical music are discussed.

Soltau, H. 1998. Recognition of musical types. *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [online]

> This paper examines the hidden units in feed-forward neural networks used to classify audio signals. The hypothesis is that abstract acoustic events can be determined from the hidden units, and subsequent temporal analysis of event activation can lead to genre identification. Included is a mention of a perceptual study, in which 37 subjects overall made mistakes similar to those of the machine learning algorithm.

Tzanetakis, G., G. Essl, and P. Cook. 2001. Automatic musical genre classification of audio signals. *Proceedings of the International Conference on Music Information Retrieval*. [online]

> This important paper introduced the use of rhythm histograms as an additional feature vector for analysis of musical genre. It also used a tree-structure taxonomy for classification, and presented two user interfaces for visualization of classification data.

West, K. 2005. Mirex audio genre classification. MIREX genre classification contest. [online]

> This 2005 MIREX submission used spectral and rhythmic features as input to a modification of Brieman's "trees" algorithm for classification. It placed 4th overall in the genre classification category, after the two entries by Bergstra et al., and that of Mandel et al.

Zhu, J., X. Xue, and H. Lu. 2004. Musical genre classification by instrumental features. *Proceedings of the International Computer Music Conference*. [online]

> This paper discusses a method of spectral analysis for explicitly identifying instruments found in an audio recording. This is intended to be used as an important high-level feature for genre classification. The algorithm improved over MFCC-only techniques on a simple 4-bin genre taxonomy.