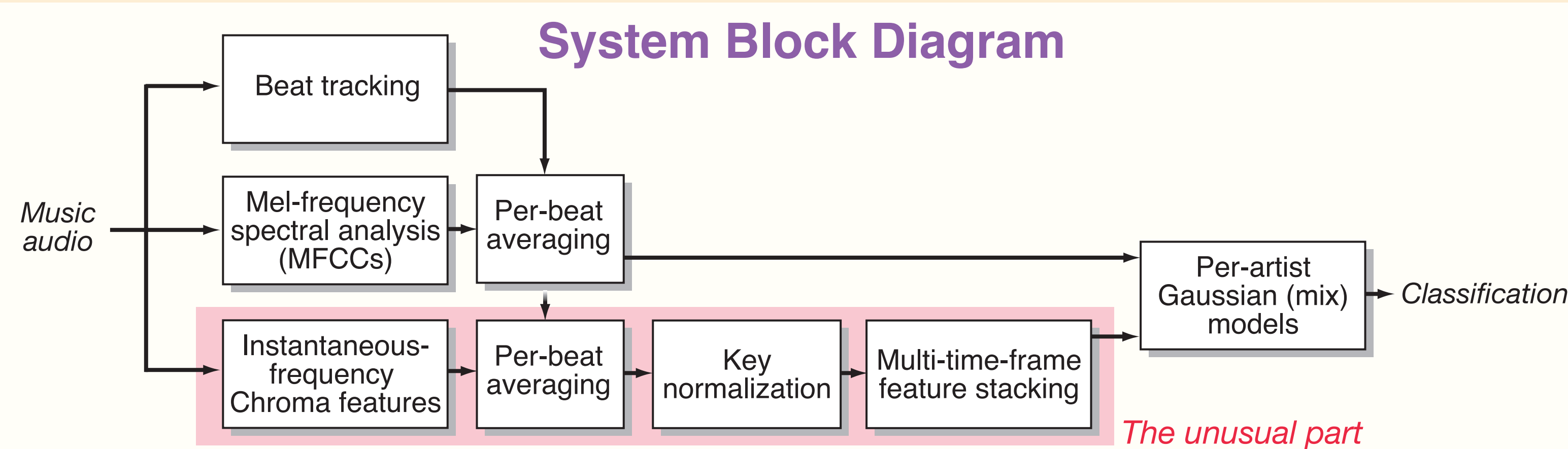


Summary: In addition to conventional cepstra, we model the **covariance of chroma** (melodic/harmonic) features and gain a small improvement in a 20-way pop music artist identification.

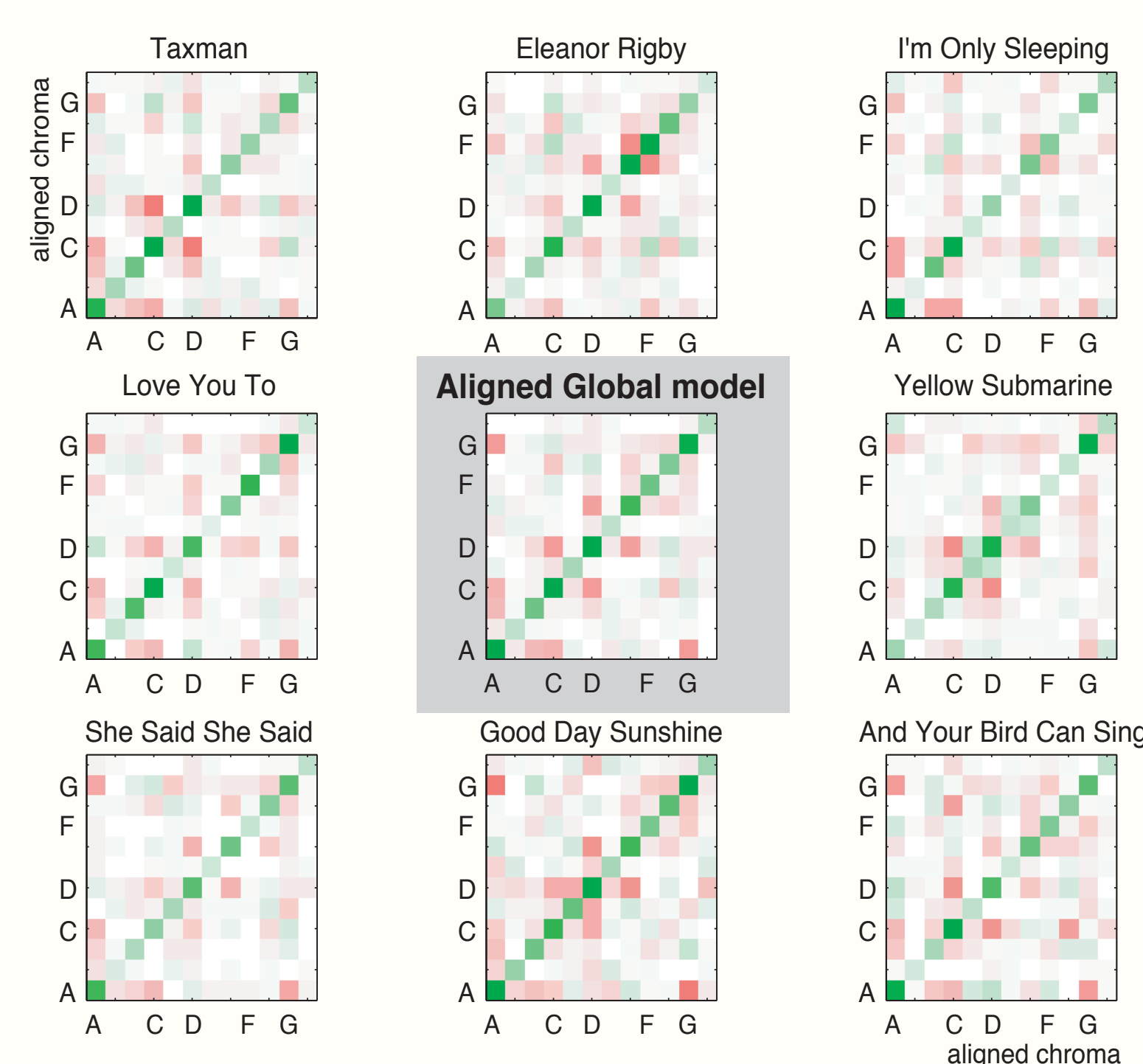
Introduction

- “Classic” approaches to music audio classification model the covariance of **spectral features** (e.g. MFCCs) as distributions [1] or discriminatively [2]. These appear to work by reflecting the instrumentation, which correlates well with genre or band [3].
- **Chroma features** [4] attempt to represent the pitch content (melody and harmony) while minimizing the influence of instrumentation.
- We investigate using the **covariance of chroma features** as a basis for music classification by artist. Although weak on their own, chroma features can improve classification when combined with spectral features
- This suggests that artists have **particular harmonic combinations** or motifs that can be automatically recognized.



Key Normalization

- Similar melodic/harmonic gestures occur ‘relative’ to the **different keys** of individual songs.
- **Key normalization** attempts to transpose (rotate) the chroma features to a canonical key prior to modeling.
- We do this by:
 - build a **chroma covariance matrix** from **all songs**
 - **transpose** each song to maximize likelihood under global model
 - **re-estimate** global model from transposed songs and repeat.



Data - “artist20” dataset

- 1413 tracks, from 120 albums (20 artists x 6 albums) [5]
- Contemporary pop music drawn from uspop2002 [6] and others
- Studio albums, chosen for chronological & stylistic consistency
- MFCCs, Chroma features, etc. available for **download** at <http://labrosa.ee.columbia.edu/projects/artistid/>

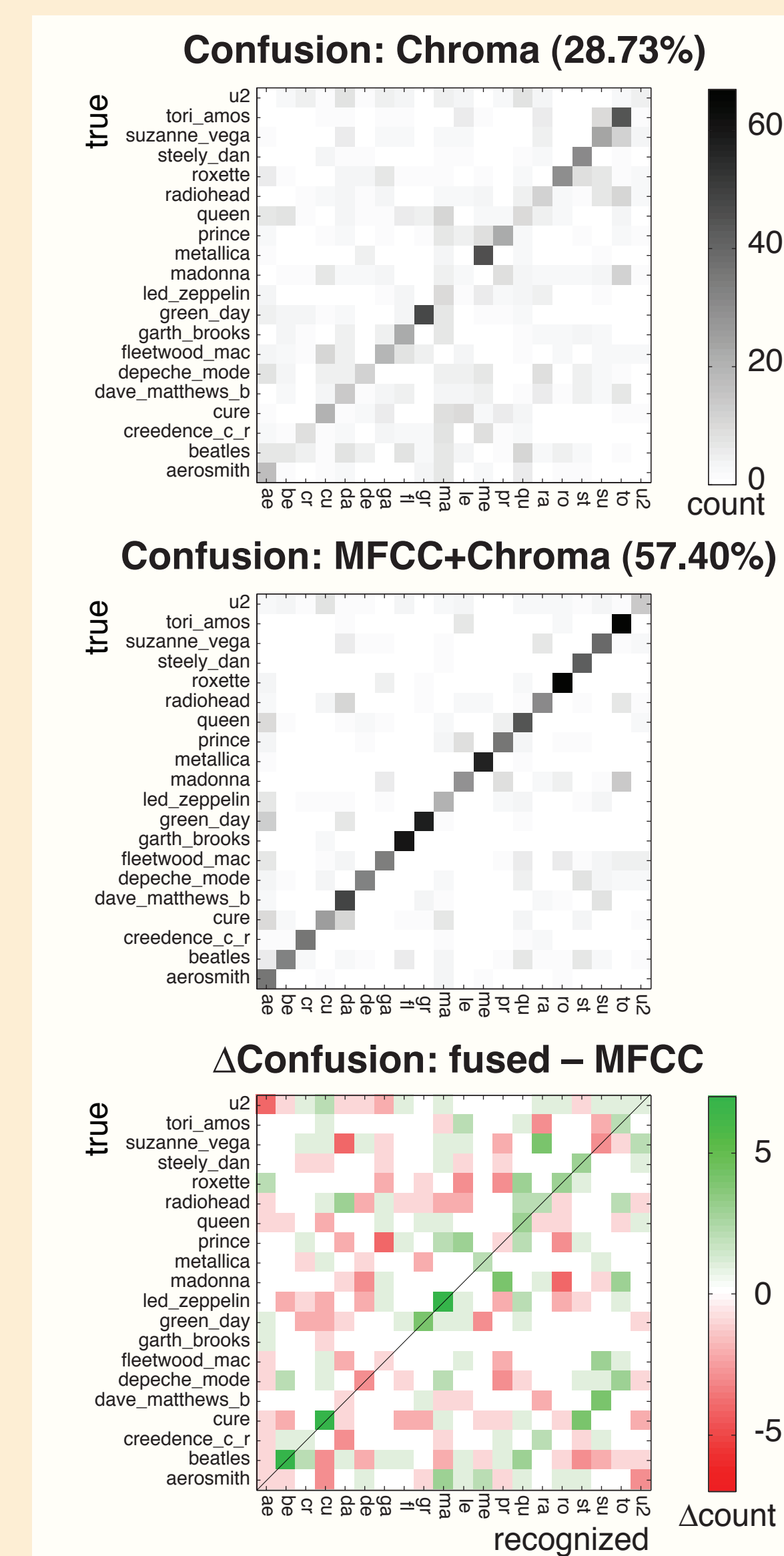
Results

- MFCC-based classification is much more accurate (55%) than using Chroma features (29% at best), but **combining the two** does give further gains (to 57% correct; McNemar $p < .01$).
- A single, full-covariance Gaussian is adequate to model the MFCC data, but Chroma data need a **64-mix GMM**.
- **Key normalization** (ChromaKN) is important for Chroma features.
- Stacking Chroma features from up to 4 adjacent beat-times (T win = 4) actually **hurts performance** (larger model space?).

Feature	Model	T win	Acc	Exec time
MFCC20	FullCov	1	55%	181 s
MFCC20	64 GMM	1	54%	1282 s
Chroma	FullCov	1	15%	31 s
Chroma	FullCov	4	15%	50 s
Chroma	64GMM	1	25%	581 s
Chroma	64GMM	4	16%	1501 s
ChromaKN	FullCov	1	23%	113 s
ChromaKN	FullCov	4	23%	255 s
ChromaKN	64GMM	1	29%	957 s
ChromaKN	64GMM	4	23%	1717 s
MFCC + Chroma fusion			57%	

Discussion and Conclusions

- Covariance of single beat-chroma vectors **carries some information** about artist, beyond that captured by MFCCs. This could relate to e.g. an artist’s **preferred chords**.
- Used alone, chroma features do quite well on a **few artists** (tori_amos, metallica) but learn almost nothing about others (madonna, beatles).
- Our best performance came from a simple **weighted sum of likelihoods** from separate MFCC and Chroma models. Differences were small, but gains were on different artists than those classified best by Chroma alone (beatles, led_zeppelin).



References

- [1] J.-J. Aucouturier & F. Pachet, “Music similarity measures: What’s the use?” ISMIR, Paris, 2002.
- [2] M. I. Mandel & D. P. W. Ellis, “Song-level features and support vector machines for music classification,” ISMIR, London, 2005.
- [3] J. H. Jensen, M. G. Christensen, S. H. Jensen, “A framework for analysis of music similarity measures,” EUSIPCO, 2007.
- [4] M. A. Bartsch & G. H. Wakefield, “To catch a chorus: Using chroma-based representations for audio thumbnailing,” WASPAA, Mohonk, 2001.
- [5] D. P. W. Ellis, “Music artist identification: artist20 baseline system in Matlab,” web resource, 2007. <http://labrosa.ee.columbia.edu/projects/artistid/>
- [6] D. P. W. Ellis, A. Berenzweig, B. Whitman, “The uspop2002 Pop Music data set,” web resource, 2003. <http://labrosa.ee.columbia.edu/projects/musicsim/uspop2002.html>

MATLAB code to run this system is available at:

<http://labrosa.ee.columbia.edu/projects/timbrechroma/>