



CLUSTERING BEAT-CHROMA PATTERNS IN A LARGE MUSIC DATABASE

Thierry Bertin-Mahieux [†], Ron J. Weiss [‡] and Daniel P. W. Ellis [†]

[†] LabROSA, Dept. of Electrical Engineering, Columbia University

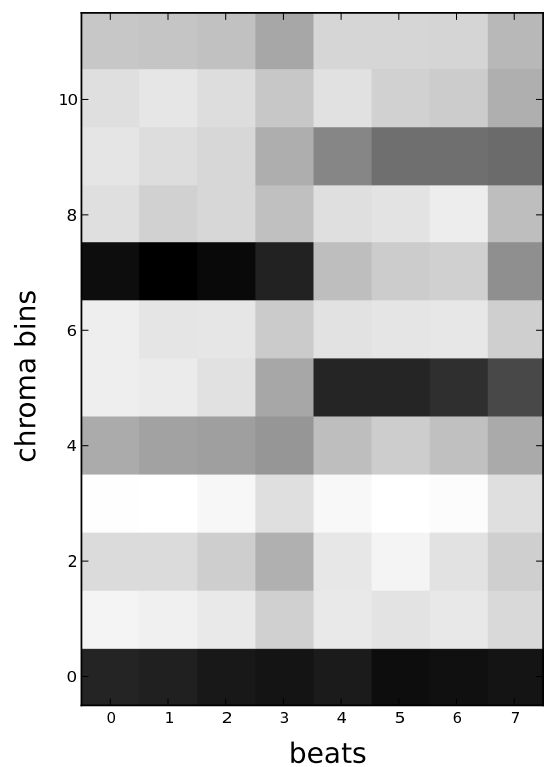
[‡] MARL, New York University

{[@ee.columbia.edu](mailto:thierry,ronw,dpwe)}



Introduction

- Availability of very large collections of music audio: can we infer anything about the underlying structure and common features of e.g. commercial pop music?
- Our interest: tonal content of the music – i.e. the harmony and melody.
- Beat-synchronous chromagrams: rich enough to generate musically-relevant results, simplified enough to abstract away instrumentation and other stylistic details.
- This paper identifies common patterns in beat-synchronous chromagrams by learning codebooks from a large set of examples. The individual codewords consist of short beat-chroma patches of between 1 and 8 beats, optionally aligned to bar boundaries.



Our goal is to identify meaningful information about the musical structure represented in the entire database by examining individual entries in this codebook. Since the common patterns represent a higher-level description of the musical content than the raw chroma, we also expect them to be useful in other applications, such as music classification and retrieving tonally-related items.

Prior work using small patches of chroma features includes the “shingles” of [?], which were used to identify “remixes”, i.e., music based on some of the same underlying instrument tracks, and also for matching performances of Mazurkas [?]. That work, however, was not concerned with extracting the deeper common patterns underlying different pieces (and did not use either beat- or bar-synchronous features). Earlier work in beat-synchronous analysis includes [?], which looked for repeated patterns within single songs to identify the chorus, and [?], which cross-correlated beat-chroma matrices to match cover versions of pop music tracks. None of these works examined or interpreted the content of the chroma matrices in any detail. In contrast, here we hope to develop a codebook whose entries are of interest in their own right.

- A bullet point
- A bullet point
- A bullet point
- A bullet point

Audio Features - Echo Nest

- A bullet point
- A bullet point
- A bullet point

Vector Quantization

- A bullet point
- A bullet point
- A bullet point
- A bullet point

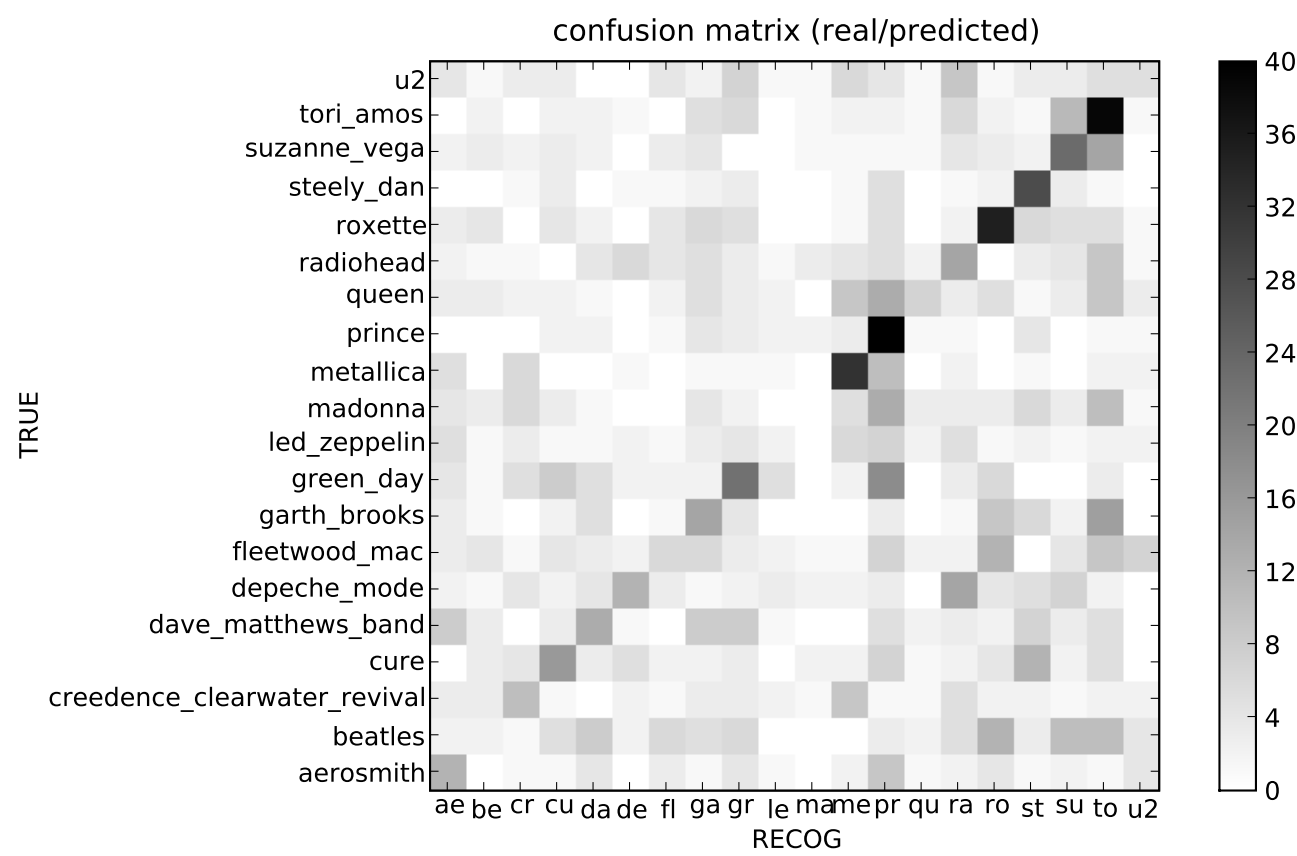
Pattern Analysis

Experiments

We present two applications of the beat-chroma codebooks to illustrate how the “natural” structure identified via unsupervised clustering can provide useful features for subsequent supervised tasks.

Artist recognition task. We use the *artist20* data set: 1402 songs from 20 artists, mostly rock and pop of different subgenres. Previously published results using GMMs on MFCC features achieve an accuracy of 50%, whereas using only chroma as a representation yields an accuracy of 33% [?].

We get an accuracy of **23.4%**, random baseline is around 5%. The confusion matrix is shown here, note that certain artists are recognized at an accuracy far above the average.



Bar alignment task. Since the clustering described is based on the segmentation of the signal in to bars, the codewords should contain information related to bar alignment, such as the presence of a strong beat on the first beat.

Offset	% of times chosen
0	62.6
1	16.5
2	9.4
3	11.5

- A bullet point

- A bullet point

- A bullet point

- A bullet point

Conclusion

- A bullet point

- A bullet point