CQ-PROFILES FOR KEY FINDING IN AUDIO

Hendrik Purwins
Universitat Pompeu Fabra IUA (MTG)
Ocata, 1
08003 Barcelona
hpurwins@iua.upf.es

Benjamin Blankertz
Fraunhofer FIRST (IDA)
Kekuléstraße 7
12489 Berlin
benjamin.blankertz@first.fhg.de

ABSTRACT

Key finding in audio is based on the constant Q transform. A heuristics is suggested how to compress the constant Q transform into a 12-dimensional short-term pitch class profile. Short-term profiles are weighted by a cosine window and summed up yielding long-term profiles. The latter are matched against averaged major and minor prototype profiles.

Keywords: key, pitch class, chroma, profile, probe tone rating

1 METHOD

For each pitch class calculating the strength in a musical recording, our constant quotient (CQ-) profile method [3,4] is a) consistent with psychological probe tone ratings, b) highly efficient, c) computable in real-time, d) stable with respect to sound quality, e) applicable to transposition, f) free of musical presupposition, except approximately stable equal temperament, and g) sensitive to substantial musical features, especially key, in a highly compact reduction. The constant Q(uotient) transform [1,2] gives equal resolution in the logarithmic frequency domain. Hence the quotient of center frequency and bandwidth is constant for each filter. To cope with spectral leakage effects, we use a resolution of b=36 bins per octave. Every third CQ-bin corresponds to a semitone. In the outcome of the CQ transform, only CQ middle bins of a triple are considered that tower above their neighbors and a given threshold. Summing up values of middle bins of multiple octave distance yields a 12-dimensional short-term CQ profile. The latter are calculated every 50 ms and summed up to long-term profiles. In this summation process only the short-term profiles of the first 15 s are considered. They are weighted by a cosine function in the range of 0-$\frac{\pi}{2}$. For the beginning of each piece, a long-term profile is calculated. To determine the tuning of C we employ the above described middle-bin-maximum heuristics again. Varying the frequency of C within a range of $\pm$ 16 Cent, the resulting profile with maximum length indicates the right frequency. Long-term profiles are transposed to a common keynote. Then a mean major and minor profile is calculated. These mean profiles are transposed back to all keys to constitute a set of 24 reference key profiles. The key of the incoming piece is determined by maximal correlation of its long-term CQ-profile with one of the reference key profiles.

This algorithm was placed second in the MIREX audio key finding competition.

2 REFERENCES

[1] J. Brown. Calculation of a constant Q spectral transform. *JASA*, 89(1):425–434, 1991.
[2] J. Brown and M. Puckette. An efficient algorithm for the calculation of a constant Q transform. *JASA*, 92(5):2698–2701, 1992.
[3] H. Purwins. *Profiles of Pitch Classes*. PhD thesis, Berlin University of Technology, 2005.
[4] H. Purwins, B. Blankertz, and K. Obermayer. A new method for tracking modulations in tonal music in audio data format. In IJCNN-00(6):270–275, 2000.