Constant Q Profiles for Tracking Modulations in Audio Data

Cq-profiles are 12-dimensional vectors, each component referring to a pitch class. They can be employed to represent keys. Cq-profiles are calculated with the constant Q filter bank [2]. They have the following advantages: (i) They correspond to probe tone ratings. (ii) Calculation is possible in real-time. (iii) Stability is obtained with respect to sound quality. (iv) They are transposable. Cq-profiles are reliably applied to modulation tracking by introducing a special distance measure.

Constant Q transform The calculation of the cq-profiles is based on the constant Q transform [1]. The letter 'Q' refers to the constant quotient of center frequency and bandwidth for each filter. The constant Q transform is useful in establishing a direct correspondence between filters and musical notes by identifying appropriate center frequencies. To minimize spectral leakage (cf. [3]), we use 36 filters per octave rather than 12.

Figure 1 shows how cq-profiles are calculated from the output of the transform. Hence, only every third filter output maps to a tone of the chromatic scale. Cq-profiles can be used to study pitch use in different composers and for modulation tracking. A cq-reference set is a sequence of 24 cq-profiles, one for each key. Every profile should reflect the tonal hierarchy that is characteristic for its key. Typically cq-reference sets are calculated from sampled cadential chord progressions or from small pieces of music.

Calculation of cq-transform Like the Fourier transform, a constant Q transform [1] is a bank of filters, but in contrast to the former it has geometrically spaced center frequencies $f_k = f_0 \cdot 2^{\frac{k}{b}}$ and a constant ratio of frequency to bandwidth $Q = \frac{f_k}{\Delta_k^{cq}} = (2^{\frac{1}{b}} - 1)^{-1}$ ($k = 0, \ldots$), where b dictates the number of filters per octave. This is achieved by choosing an appropriate window length N_k individually for each component of the constant Q transform (cq-bin). For integer values Q the k-th cq-bin is the Q-th DFT-bin with window length $Q\frac{f_s}{f_k}$. Calculation: First choose minimal frequency f_0 and the number of bins per octave b according to the requirements of the application and let[1]: $K := \lceil b \cdot \log_2(\frac{f_{\max}}{f_0}) \rceil$, $Q := (2^{\frac{1}{b}} - 1)^{-1}$, and $N_k := \lceil Q\frac{f_s}{f_k} \rceil$ (for $k < K$). Then the k–th cq-bin is equal to $N_k^{-1} \sum_{n<N_k} x[n]\ w_{N_k}[n]\ e^{-2\pi i n Q/N_k}$. Following [2] we use Hamming windows $\langle w_N[n] : n < N \rangle$. Using Parseval's rule a filter matrix is calculated in advance. Exploiting sparsity accelerates the calculation of the constant Q transform very much [2].

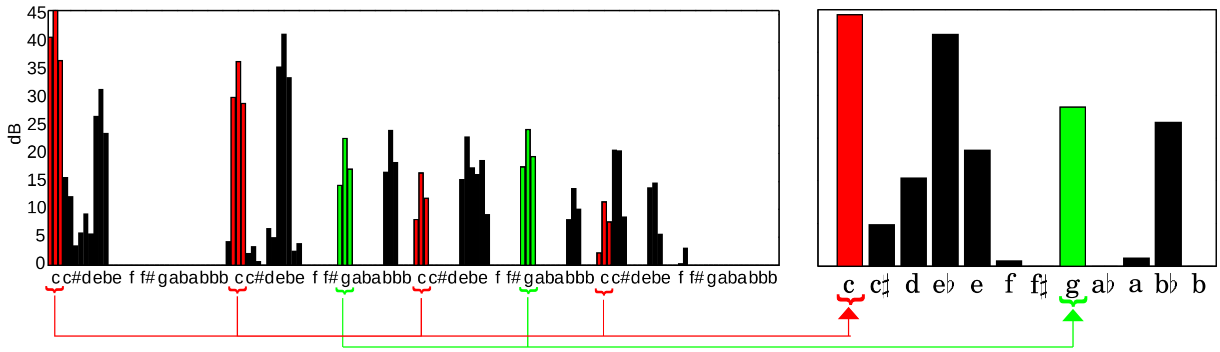
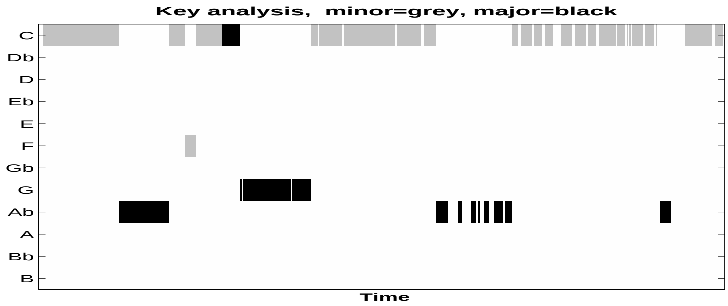


FIGURE 1: The constant Q transform is calculated from a minor third c – e♭ (played on piano) with three bins per half-tone (left figure). We yield the constant Q profile (right figure) by summing up bins for each tone over all octaves.

Application in tone center tracking How can a piece be classified according to a cq-reference set? Generally we have the problem of matching a given cq-profile with a profile of the cq-reference set. A typical matching criteria is the closest fuzzy distance: Let y be a value subject to an uncertainty quantisized by a value σ (typically y is the mean and σ the standard deviation of some statistical data). The fuzzy distance of some value x to y regarding σ is defined by $d_\sigma(x, y) := |x - y| \cdot \left(1 - \frac{\sigma}{|x-y|+\sigma} e^{-\frac{|x-y|^2}{2\sigma^2}}\right)$. The fuzzy distance is similar to the Euclidean metric, but the greater the uncertainty the more relaxed is the metric. As an example, we present an analysis of Chopin's c-minor Prélude op. 28, No. 20. The reference vectors were calculated from all 24 Chopin Préludes in audio format.

[1] $\lceil x \rceil$ denotes the least integer greater than or equal to x.



(a) Result of automatic tone center analysis

(b) Score

FIGURE 2: Chopin's c-minor Prélude, op. 28, No. 20. In (a) grey indicates minor, black indicates major. If there is neither black nor grey at a certain time, the significance of a particular key is below a given threshold. There is no distinction between enharmonic equivalent keys.

In the score (Figure 2 (b)) tone centers are marked. They were determined by a musical expert. Tone centers in parentheses indicate tonicizations on a very short time scale. Since the automatic tone center recognition (Figure 2 (a)) does not look ahead, there is a delay in recognizing tone centers. The program captures the prevailing key c–minor and the modulations.

This result is astonishing. The only explicit musical knowledge utilized is the display of the signal in terms of pitch classes. The system receives musical knowledge only by choice of the music pieces, which lead to the reference vectors.

The simple constant Q profile method incorporates context processing by averaging over the entire piece. Only very basic music theoretical assumptions like octave equivalence and the chromatic scale are used explicitly. However it can capture a large amount of harmonic structure including modulation and keys. Other musical knowledge is not explicitly used, like voice leading, harmony, metric, and rhythm. The constant Q profile method is a powerful tool that can be extended to different tunings, and to real time analysis. It could be improved by modeling masking phenomena.

We intend to present modulation tracking applied to a larger body of music samples. Other applications include analysis of pitch use in different composers and epochs.

References

[1] J. Brown. Calculation of a constant Q spectral transform. *J. Acoust. Soc. Am.*, 89(1):425–434, 1991.

[2] Judith C. Brown and Miller S. Puckette. An efficient algorithm for the calculation of a constant Q transform. *J. Acoust. Soc. Am.*, 92(5):2698–2701, 1992.

[3] F. J. Harris. On the use of windows for harmonic analysis with discrete fourier transform. In *Proc. IEEE*, volume 66, pages 51–83, 1978.

[4] E. Zimmermann. *Chopin Préludes op. 28: Kritischer Bericht*. Henle, 1969.