

Key Estimation in Electronic Dance Music

Ángel Faraldo, Emilia Gómez, Sergi Jordà, and Perfecto Herrera

Music Technology Group, Universitat Pompeu Fabra,
Roc Boronat 138, 08018 Barcelona, Spain
`name.surname@upf.edu`

Abstract. In this paper we study key estimation in electronic dance music, an umbrella term referring to a variety of electronic music sub-genres intended for dancing at nightclubs and raves. We start by defining notions of tonality and key before outlining the basic architecture of a template-based key estimation method. Then, we report on the tonal characteristics of electronic dance music, in order to infer possible modifications of the method described. We create new key profiles combining these observations with corpus analysis, and add two pre-processing stages to the basic algorithm. We conclude by comparing our profiles to existing ones, and testing our modifications on independent datasets of pop and electronic dance music, observing interesting improvements in the performance of our algorithms, and suggesting paths for future research.

Keywords: Music Information Retrieval, Computational Key Estimation, Key Profiles, Electronic Dance Music, Tonality, Music Theory.

1 Introduction

The notion of tonality is one of the most prominent concepts in Western music. In its broadest sense, it defines the systematic arrangements of pitch phenomena and the relations between them, specially in reference to a main pitch class [9]. The idea of key conveys a similar meaning, but normally applied to a smaller temporal scope, being common to have several key changes along the same musical piece. Different periods and musical styles have developed different practices of tonality. For example, modulation (i.e. the process of digression from one local key to another according to tonality dynamics [20]) seems to be one of the main ingredients of musical language in euroclassical¹ music [25], whereas pop music tends to remain in a single key for a whole song or perform key changes by different means [3], [15].

Throughout this paper, we use the term *electronic dance music* (EDM) to refer to a number of subgenres originating in the 1980's and extending into the present, intended for dancing at nightclubs and raves, with a strong presence of percussion and a steady beat [4]. Some of such styles even seem to break up

¹ We take this term from Tagg [25] to refer to European Classical Music of the so-called common practice repertoire, on which most treatises on harmony are based.

with notions such as chord and harmonic progression (two basic building blocks of tonality in the previously mentioned repertoires) and result in an interplay between pitch classes of a given key, but without a sense of tonal direction.

These differences in the musical function of pitch and harmony suggest that computational key estimation, a popular area in the Music Information Retrieval (MIR) community, should take into account style-specific particularities and be tailored to specific genres rather than aiming at all-purpose solutions.

In the particular context of EDM, automatic key detection could be useful for a number of reasons, such as organising large music collections or facilitating *harmonic mixing*, a technique used by DJ’s and music producers to mix and layer sound files according to their tonal content.

1.1 Template-Based Key Estimation Methods

One of the most common approaches to key estimation is based on pitch-class profile extraction and template matching. Figure 1 shows the basic architecture of such key estimation system. Regular methodologies usually convert the audio signal to the frequency domain. The spectral representation is then folded into a so-called pitch class profile (PCP) or chromagram, a vector representing perceptually equal divisions of the musical octave, providing a measure of the intensity of each semitone of the chromatic scale per time frame. For improved results, a variety of pre-processing techniques such as tuning-frequency finding, transient removal or beat tracking can be applied. It is also common to smooth the results by weighting neighbouring vectors. Lastly, similarity measures serve to compare the averaged chromagram to a set of templates of tonality, and pick the best candidate as the key estimate. We refer the reader to [7], [17] for a detailed description of this method and its variations.

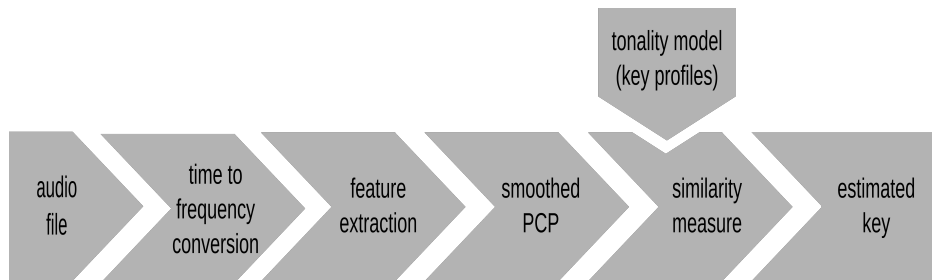


Fig. 1. Basic template-based key estimation system.

One of the most important aspects of such an approach is the model used in the similarity measure. Different key profiles have been proposed since the pioneering Krumhansl-Schmuckler algorithm, who proposed weighting coefficients derived from experiments involving human listeners ratings [12]. Most of them consist of deviations from the original coefficients to enhance performance on specific repertoires [6], [22], [26]. Among existing algorithms, *QM Key Detector*

[18] and *KeyFinder* [22] deserve special attention, particularly since both are publicly available. The former has provided the best results in previous editions of MIREX², whereas the latter appears to be the only open-source algorithm specifically tailored to key detection in EDM.

2 Timbral and Tonal Characteristics of EDM

Most subgenres falling under the umbrella term EDM give a central role to percussion and bass, over which other pitched materials and sound effects are normally layered. This idiosyncrasy results in a number of generalisable spectral features that we list hereunder:

- The ubiquity of percussive sounds tends to flatten the spectrum, possibly masking regions with meaningful tonal content.
- Tonal motion often concentrates on the lower register, where algorithms normally offer less resolution.
- Some types of EDM are characterised by tonal effects such as *glissandi* and extreme timbral saturation, that could make difficult to identify pitch as quantised and stable units.

With regard to tonal practices in EDM, pitch relations are often freed from the tonal dynamics based on the building up of tension and its relaxation. Some idiomatic characteristics from empirical observation follow, that could be taken into account when designing methods of tonality induction for this repertoire:

- Beginnings and endings of tracks tend to be the preferred place for sonic experimentation, and it is frequent to find sound effects such as field recordings, musical quotations, un-pitched excerpts or atonal interludes without a sense of continuity with the rest of the music.
- There are likely occasional detuning of songs or excerpts from the standard tuning due to manual alterations in the pitch/speed control present in industry-standard vinyl players, such as the *Technics SL-1200*, for the purpose of adjusting the tempo between different songs.
- Euroclassical tonal techniques such as modulation are essentially absent.
- The dialectics between consonance and dissonance are often replaced by a structural paradigm based on rhythmic and timbral intensification [23].
- The music normally unfolds as a cumulative form made with loops and repetitions. This sometimes causes polytonality or conflicting modality due to the overlap of two or more riffs [23].
- According to a general tendency observed in Western popular music since the 1960's, most EDM is in minor mode [21].

² The Music Information Retrieval Evaluation eXchange (MIREX) is an international committee born to evaluate advances in Music Information Retrieval among different research centres, by quantitatively comparing algorithm performance using test sets that are not available beforehand to participants.

- Tritone and semitone relationships seem to be characteristics of certain sub-genres [23], such as breakbeat or dubstep.
- In minor mode, there is hardly any appearance of the leading tone (\sharp VII) so characteristic of other tonal practices, favouring other minor scales, especially *aeolian* (\flat VII) and *phrygian* (\flat II) [24].
- It is also frequent to find pentatonic and hexatonic scales, instead of the major and minor heptatonic modes at the basis of most tonal theories [25].

3 Method

For this study, we gathered a collection of complete EDM tracks with a single key estimation per item. The main source was Sha’ath’s list of 1,000 annotations, determined by three human experts³. However, we filtered out some non-EDM songs and completed the training dataset with other manually annotated resources from the internet⁴, leaving us with a total of 925 tracks to extract new tonal profiles.

To avoid overfitting, evaluations were carried on an independent dataset of EDM, the so-called *GiantSteps key dataset* [11], consisting of 604 two-minute long excerpts from *Beatport*⁵, a well-known internet music store for DJs and other EDM consumers. Additionally, we used Harte’s dataset [14] of 179 songs by The Beatles reduced to a single estimation per song [19], to compare and test our method on other popular styles that do not follow the typical EDM conventions.

Despite the arguments presented in Section 2 about multi-modality in EDM, we decided to shape our system according to a major/minor binary model. In academic research, there has been little or no concern about tonality in electronic popular music, normally considered as a secondary domain compared to rhythm and timbre. In a way, the current paper stands as a first attempt at compensating this void. Therefore, we decided to use available methodologies and datasets (and all of these only deal with binary modality), to be able to compare our work with existing research, showing that current algorithms perform poorly on this repertoire. Furthermore, even in the field of EDM, commercial applications and specialised websites seem to ignore the modal characteristics referred and label their music within the classical paradigm.

Tonal Properties of the Datasets. The training dataset contains a representative sample of the main EDM subgenres, including but not limited to dubstep, drum’n’bass, electro, hip-hop, house, techno and trance. The most prominent aspect is its bias toward the minor mode, which as stated above, seems representative of this kind of music. Compared to *beatles* dataset, of which only a

³ <http://www.ibrahimshaath.co.uk/keyfinder/KeyFinderV2Dataset.pdf>

⁴ <http://blog.dubspot.com/dubspot-lab-report-mixed-in-key-vs-beatport>
<http://www.djtechttools.com/2014/01/14/key-detection-software-comparison-2014-edition>

⁵ <https://pro.beatport.com/>

10.6% is annotated in minor (considering one single key estimation per song), the training dataset shows exactly the inverse proportion, with 90.6% of it in minor. The *GiantSteps* dataset shows similar statistics (84.8% minor), confirming theoretical observation [21].

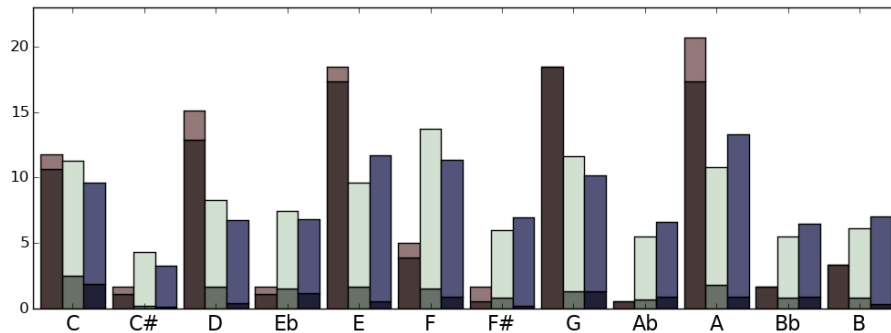


Fig. 2. Distribution of major (bottom) and minor (top) keys by tonal center in *beatles* (left), *GiantSteps* (center) and training (right) datasets.

Figure 2 illustrates the percentage distribution of keys in the three datasets according to the tonal centre of each song. We observe a tendency toward certain tonal centres in *beatles*, which correspond to guitar *open-string* keys (C, D, E, G, A), whereas the two EDM collections present a more even distribution among the 12 chromatic tones, probably as a result of music production with synthesisers and digital tools.

3.1 Algorithm

In the following, we propose modifications to a simple template-based key estimation method. We study the effect of different key profiles and create our own from a corpus of EDM. We modify the new templates manually, in the light of the considerations outlined in Section 2. Then, we incorporate a *spectral whitening* function as a pre-processing stage, in order to strengthen spectral peaks with presumable tonal content. Finally, taking into account the potential detuning of fragments of a given track due to hardware manipulations, we propose a simple *detuning correction* method.

The simple method we chose is implemented with *Essentia*^{6 7}, a C++ library for audio information retrieval [1], and it is based on prior work by Gómez [6, 7].

⁶ <http://essentia.upf.edu/>

⁷ After informal testing, we decided to use the following settings in all the experiments reported: mix-down to mono; sampling rate: 44,100 Hz.; window size: 4,096 hanning; hop size: 16,384; frequency range: 25-3,500 Hz.; PCP size: 36 bins; weighting size: 1 semitone; similarity: cosine distance.

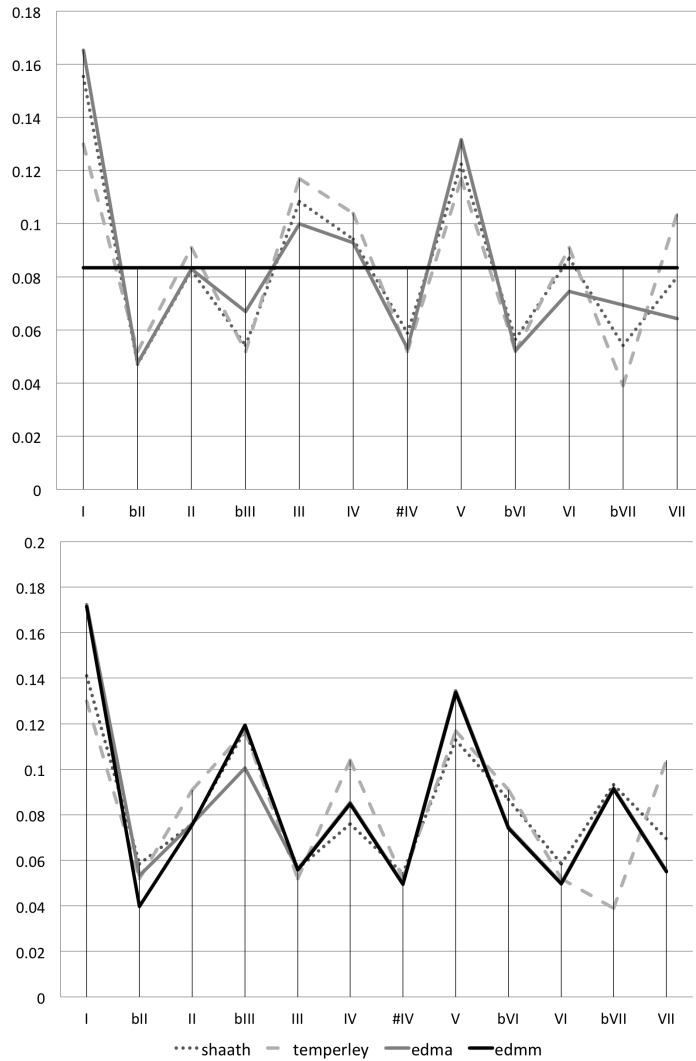


Fig. 3. The four major (above) and minor (below) key profiles. Note that the major profile of edmm is flat.

New Key Profiles. As explained above, one of the main ingredients in a template-based key estimator is the tonality model represented by the so-called key profile, a vector containing the relative weight of the different pitch classes for a given key. In this paper we compare four different profiles:

1. The ones proposed by Temperley [26], which are based on corpus analysis of euroclassical music repertoire.

2. Manual modifications of the original Krumhansl profiles [12] by Sha’ath [22], specifically oriented to EDM. The main differences in Sha’ath’s profiles are (a) a slight boost of the weight for the ♯VII degree in major; and (b) a significant increment of the subtonic (♭VII) in minor. Other than these, the two profiles remain essentially identical.
3. Major and minor profiles extracted as the median of the averaged chromagrams of the training set. This provided best results compared to other generalisation methods (such as grand average, max average, etc.). Throughout this paper we refer to these profiles as *edma*.
4. Manual adjustments on the extracted profiles (referenced as *edmm*) accounting for some of the tonal characteristics described in section 2, especially the prominence of the aeolian mode, and the much greater proportion of minor keys. In that regard, given the extremely low proportion of major tracks in the corpus, we decided to flatten the profile for major keys.

Figure 3 shows a comparison between these four profiles. They are all normalised so that the sum of each vector equals 1. It is visible how the profiles by Temperley favour the leading-tone in both modes, according to the euro-classical tonal tradition, whilst the other three profiles increase the weight for the subtonic. We can see that automatically generated profiles (*edma*) give less prominence to the diatonic third degree in both modes, reflecting the modal ambiguity present in much EDM. We compensated this manually, raising the ♭III in the minor profile, together with a decrement of the ♭II (*edmm*).

Spectral Whitening. We inserted a pre-processing stage that flattens the spectrum according to its spectral envelope, based on a method by Röbel and Rodet [27]. The aim was to increase the weight of the predominant peaks, so that notes across the selected pitch range contribute equally to the final PCP. This technique has been previously used by Gómez [7], and other authors have proposed similar solutions [10], [13], [16].

Detuning Correction. We noted that some of the estimations with the basic method produced tritone and semitone errors. Our hypothesis was that these could be due to possible de-tunings produced by record players with manual pitch/tempo corrections [22]. In order to tackle this, our algorithm uses a PCP resolution of 3 bins per semitone, as it is usual in key detection algorithms [8], [18]. This allowed us to insert a post-processing stage that shifts the averaged PCP ± 33 cents, depending on the position of the maximum peak in the vector.

Various tuning-frequency estimation methods have been proposed, mostly based on statistics [5], [28]. Our approach is a simplification of that described in [8]. The algorithm finds the maximum value in the averaged chromagram and shifts the spectrum ± 1 bin, depending on this unique position. This shift is done only once per track, after all the PCPs are averaged together.

Our approach is reference-frequency agnostic: it takes the reference pitch to be between 431.6 and 448.5 Hz (i.e. 1/3 of a semitone lower or higher than the

pitch standard). We regard this margin as comfortable enough for the repertoire under consideration, assuming that most music would fit within the range mentioned above.

3.2 Evaluation Criteria

In order to facilitate reproducibility, we compared the performance of our method with two publicly available algorithms, already mentioned in Section 1.1: Sha’ath’s *KeyFinder*⁸ and the *QM Key Detector* vamp-plugin⁹ by Noland and Landone [2], which we assume to be a close version of the best performing algorithm in MIREX 2014.

The MIREX evaluation has been so far carried on 30-second excerpts of MIDI renders of euroclassical music scores. This follows an extended practice of performing key estimation in fragments of short duration at the beginning or end of a piece of music. Contrary to this tendency, informal experiments suggest that computational key estimation in popular music provides better results when analysing full-length tracks. One of the motivations of observing the beginning of a piece of music is to skip modulational processes that can obstruct the global-key estimation task; however, modulation is not characteristic of EDM neither of pop music. Moreover, given the timbral complexity of most EDM, averaging the chromagrams over the full track likely provides a cleaner tonal profile, minimising the effect of transients and other unwanted spectral components. Based on these arguments, we performed all of our evaluations on complete tracks.

The ranking of the algorithms was carried out following the MIREX evaluation procedure, by which neighbouring keys are weighted by various factors and averaged into a final score.

4 Results

Table 1 presents the weighted scores of our basic algorithm with the variations we have described, now tested with 2 independent set collections, different than those used for the algorithm development. The top four rows show the effect of the key profiles discussed without further modifications. As expected, different profiles provide quite different responses, depending on the repertoire. Temperley’s profiles perform well on the *beatles* set, whereas they offer poor performance for *GiantSteps*; Shaa’th’s provide a moderate response in both datasets, whilst the two edm-variants raise the score on the EDM dataset, at the expense of a worse performance on *beatles*. This is especially true for the edmm profiles, provided the major profile is a flattened vector, and major keys are majority in the *beatles* collection (89.4%).

We observe that spectral whitening offers improvement in all cases, from a slight increment of 1.5 points in the more extreme profiles (edmm and temperley)

⁸ <http://www.ibrahimshaath.co.uk/keyfinder/>

⁹ <http://isophonics.net/QMVampPlugins>

<i>beatles GiantSteps training</i>			
temperley	62.0	47.6	<i>46.5</i>
sha'ath	55.8	55.8	<i>57.1</i>
edm-auto (edma)	50.1	60.1	<i>61.8</i>
edm-manual (edmm)	21.3	64.0	<i>66.7</i>
temperley + sw	66.4	48.9	<i>48.0</i>
sha'ath + sw	64.5	63.2	<i>64.5</i>
edma + sw	60.5	66.8	<i>67.3</i>
edmm + sw	22.7	71.5	<i>74.4</i>
temperley + dc	75.6	47.2	<i>46.7</i>
sha'ath + dc	71.3	56.1	<i>58.9</i>
edma + dc	68.7	60.8	<i>64.0</i>
edmm + dc	27.3	64.6	<i>69.0</i>
temperley + dc + sw	81.2	48.6	<i>48.4</i>
sha'ath + dc + sw	79.9	63.9	<i>66.1</i>
edma + dc + sw	76.0	67.3	<i>69.0</i>
edmm + dc + sw	28.8	72.0	<i>76.6</i>

Table 1. MIREX weighted scores for the two datasets with four different profiles and the proposed modifications: spectral whitening (sw) and detuning correction (dc). We additionally report the weighted scores on the training set on the third column.

to a raise of 10 p.p. in the performance of edma in the *beatles* dataset. Profiles by Sha'ath get a fair boost in both collections.

Similarly, the detuning correction method alone pushes up all the scores except temperley's profiles on the EDM dataset. Significant improvement is only observed in *beatles*, with increments between 6 and 18.6 p. p. It is known that some of The Beatles' albums were recorded with deviations from the pitch standard (this is specially the case in *Please Please Me* and *Help!*) [8], and our method seems to detect and correct them. On the other hand, the neutrality of this parameter on *GiantSteps* suggests further experimentation with particular subgenres such as hip-hop, where tempo adjustments and vinyl-scratching techniques are commonplace, which is sparsely represented in the *GiantSteps key dataset* [11].

In any case, the combination of both processing stages gives the best results. It is noteworthy that these modifications address different problems in the key-estimation process, and consequently, the combined score results in the addition of the two previous improvements. With these settings, the edma profiles yield 25.9 p.p. over the default settings in *beatles*, on which all key profiles obtain significant improvement. On *GiantSteps*, however, we observe more modest improvement.

4.1 Evaluation

Table 2 shows the results of our evaluation following the MIREX convention. Results are organised separately for each dataset. Along with *QM Key Detector* (qm) and *KeyFinder* (kf), we present our algorithm (with spectral whitening and detuning correction) with three different profiles, namely Temperley’s (edmt), automatically extracted profiles from our training dataset (edma), and the manually adjusted ones (edmm).

Both benchmarking algorithms were tested using their default settings. *KeyFinder* uses Sha’ath’s own profiles presented above, and provides a single estimate per track. *QM Key Detector*, on the other hand, uses key profiles derived from analysis of J. S. Bach’s *Well Tempered Klavier I* (1722), with window and hop sizes of 32,768 points, providing a key estimation per frame. We have reduced these by taking the most frequent estimation per track.

Edmt yields a weighted score of 81.2 in the *beatles* dataset, followed by edma (76), above both benchmarking algorithms. Most errors concentrate on the fifth, however, other common errors are minimised. Edmm produces 48% parallel errors, identifying all major keys as minor due to its flat major profile. For *GiantSteps*, results are slightly lower. The highest rank is for edmm, with a weighted score of 72.0, followed by edma and *KeyFinder*.

	<i>beatles</i>					<i>GiantSteps</i>				
	qm	kf	edmt	edma	edmm	qm	kf	edmt	edma	edmm
correct	61.4	46.9	73.7	67.0	10.1	39.4	45.3	33.4	58.1	64.2
fifth	17.3	17.3	12.3	12.3	6.7	16.9	20.7	16.7	10.1	10.8
relative	3.3	9.5	2.2	5.0	19.5	13.4	6.8	15.7	06.6	03.3
parallel	6.7	4.5	3.4	6.7	48.0	5.1	7.8	10.6	10.6	06.8
other	11.2	21.8	8.4	8.9	15.6	25.2	19.4	23.5	14.6	14.9
weighted	72.4	59.3	81.2	76.0	28.8	52.9	59.3	48.6	67.3	72.0

Table 2. Performance of the algorithms on the two evaluation datasets. Our method is reported with spectral whitening and detuning correction, on three different profiles: temperley (edmt), edm-auto (edma) and edm-manual (edmm). Under the correct estimations, we show results for different types of common errors.

4.2 Discussion

Among all algorithms under comparison, edma provides the best compromise among different styles, scoring 76.0 points on *beatles* and 67.3 on *GiantSteps*. This suggests that the modifications described are style-agnostic, since they offer improvement over the compared methods in both styles. Spectral whitening and detuning correction address different aspects of the key estimation process,

and their implementation works best in combination, independently of the key profiles used. However, results vary drastically depending on this last factor, evidencing that a method based on tonality profiles should be tailored to specific uses and hence not suitable for a general-purpose key identification algorithm. This is especially the case with our manually adjusted profiles, which are highly biased toward the minor modes.

5 Conclusion

In this paper, we adapted a template-based key estimation method to electronic dance music. We discussed some timbral and tonal properties of this metagenre, in order to inform the design of our method. However, although we obtained improved results over other publicly available algorithms, they leave room for improvement.

In future work, we plan to incorporate some of the tonal characteristics described more thoughtfully. In particular, we envision a model that expands the major/minor paradigm to incorporate a variety of modes (i.e. dorian, phrygian, etc.) that seem characteristic of EDM. Additionally, a model more robust to timbre variations could help us identifying major modes, in turn minimising the main flaw of our manually adjusted profiles. This would not only improve the performance on this specific repertoire, but also make it more generalisable to other musical genres.

Summarising, we hope to have provided first evidence that EDM calls for specific analysis of its particular tonal practices, and of computational methods informed by these. It could be the case that its subgenres also benefit from such kind of adaptations. In this regard, the bigger challenge would be to devise a method for adaptive key-template computation, able to work in agnostic genre-specific classification systems.

Acknowledgement. This research has been partially supported by the EU-funded *GiantSteps* project (FP7-ICT-2013-10. Grant agreement number 610591).

References

1. Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O.: ESSENTIA: an open-source library for sound and music analysis. Proc. 21st ACM-ICM, 855–858 (2013)
2. Cannam, C., Mauch, M., Davies, M.: MIREX 2013 Entry: Vamp plugins from the Centre For Digital Music. www.music-ir.org (2013)
3. Everett, W.: Making sense of rock’s tonal systems. Music Theory Online 10(4), 2004.
4. Dayal, G., Ferrigno, E.: Electronic Dance Music. Grove Music Online, Oxford University Press (2012)
5. Dressler, K., Streich, S.: Tuning frequency estimation using circular statistics. Proc. 8th ISMIR, 2–5 (2007)

6. Gómez, E.: Tonal description of polyphonic audio for music content processing. *INFORMS Journal on Computing* 18(3), 294–304 (2006)
7. Gómez, E.: Tonal description of music audio signals. PhD Thesis, Universitat Pompeu Fabra, Barcelona (2006)
8. Harte, C.: Towards automatic extraction of harmony information from music signals. PhD Thesis, Queen Mary University of London (2010)
9. Hyer, B.: *Tonality*. Grove Music Online, Oxford University Press (2012)
10. Klapuri, A.: Multipitch analysis of polyphonic music and speech signals using an auditory model. *IEEE Transactions on Audio, Speech, and Language Processing* 16(2), 255–266 (2008)
11. Knees, P., Faraldo, Á., Herrera, P., Vogl, R., Böck, S., Hörschläger, F., Le Goff, M.: Two Data Sets for Tempo Estimation and Key Detection in Electronic Dance Music Annotated from User Corrections. *Proc. 16th ISMIR* (2015)
12. Krumhansl, C. L.: *Cognitive foundations of musical pitch*. Oxford University Press, New York (1990)
13. Mauch, M., Dixon, S.: Approximate note transcription for the improved identification of difficult chords. *Proc. 11th ISMIR*, 135–140 (2010)
14. Mauch, M., Cannam, C., Davies, M., Dixon, S., Harte, C., Kolozali, S., Tidjar, D.: OMRAS2 metadata project 2009. *Proc. 10th ISMIR, Late-Breaking Session* (2009)
15. Moore, A.: The so-called “flattened seventh” in rock. *Popular Music* 14(2), 185–201 (1995)
16. Müller, M., Ewert, S.: Towards timbre-invariant audio features for harmony-based music. *IEEE Transactions of Audio, Speech & Language Processing* 18(3), 649–662 (2010)
17. Noland, K.: *Computational Tonality Estimation: Signal Processing and Hidden Markov Models*. PhD Thesis, Queen Mary University of London (2009)
18. Noland, K., Sandler, M.: *Signal Processing Parameters for Tonality Estimation*. *Proc. 122nd Convention Audio Engineering Society* (2007)
19. Pollack, A. W.: Notes on . . . series. www.icce.rug.nl/~soundscapes/DATABASES/AWP/awp-notes_on.shtml. Accessed: February 1st, 2015
20. Saslaw, J.: *Modulation (i)*. Grove Music Online, Oxford University Press (2012)
21. Schellenberg, E. G., von Scheve, C.: Emotional cues in American popular music: Five decades of the Top 40. *Psychology of Aesthetics, Creativity and the Arts* 6(3), 196–203 (2012)
22. Sha’ath, I.: *Estimation of key in digital music recordings*. Dep. Computer Science & Information Systems, Birkbeck College, University of London (2011)
23. Spicer, M.: (Ac)cumulative form in pop-rock music. *Twentieth Century Music* 1(1), 29–64 (2004)
24. Tagg, P.: From refrain to rave: the decline of figure and raise of ground *Popular Music* 13(2), 209–222 (1994)
25. Tagg, P.: *Everyday tonality II (Towards a tonal theory of what most people hear)*. The Mass Media Music Scholars’ Press. New York and Huddersfield (2014)
26. Temperley, D.: What’s key for key? The Krumhansl-Schmuckler key-finding algorithm reconsidered. *Music Perception: An Interdisciplinary Journal* 17(1), 65–100 (1999)
27. Röbel, A., Rodet, X.: Efficient spectral envelope estimation and its application to pitch shifting and envelope preservation. *Proc. 8th DAFX* (2005)
28. Zhu, Y., Kankanhalli, M. S., Gao, S.: Music key detection for musical audio. *Proc. 11th IMMC*, 30–37 (2005)