Improvements of Key finding Methods

3 authors, including:

Pierre Hanna
Laboratoire Bordelais de Recherche en Informatique

IMPROVEMENTS OF SYMBOLIC KEY FINDING METHODS

Matthias Robine, Thomas Rocher and Pierre Hanna
SIMBALS Project
LaBRI - University of Bordeaux 1
F-33405 Talence cedex, France
firstname.name@labri.fr

ABSTRACT

Automatically estimating the key of a musical piece is an important part of a lot of musical applications such as music classification or music transcription. Existing methods rely on the comparison of pitch class profiles. Correlation computed between the input pitch profile and a key reference profile indicates the key of the musical piece tested. Other recent methods propose to consider the note intervals and introduce interval profiles. In this paper, we propose to investigate the possible combinations of these different methods. The study of the errors induced by these methods lead us to propose other improvements, since errors mainly lead to the detection of neighbour or relative keys. By considering particular tonal properties of the music piece and the information from the different profiles, new methods are proposed to correct such errors. Experiments performed with two different databases finally show the improvements induced.

1. INTRODUCTION

Retrieving the key of a musical piece is a major issue when it comes to music perception, music edition, or simply music experience. In [12], Schmuckler and Tomovski write that *perceiving the tonality of a musical passage is a fundamental aspect of the experience of hearing music*. Therefore, being able to automatically retrieve the key from a musical file allows the development of a lot of applications [5] such as music classification, music transcription [2], virtual instruments, etc.

Nowadays, music is generally classified by artist and title. Automatically retrieving a piece of music from its key would extend search engines. For example, retrieval results may be filtered according to the user's mood, since major and minor modes usually represent the general mood the music gives away. Homogeneous playlists, containing only songs in the same key, could also be automatically generated. Virtual instruments can also take advantage of such a research, especially those whose point is playing according to an harmonic context. Being able to retrieve the key of the harmonic background can, without a doubt, increase their efficiency.

Existing key finding methods are described in Section 2. Then, procedures of the experiments presented in this paper are proposed in Section 3. In Section 4, a new approach consisting of combining existing methods is detailed. Other improvements are described in Section 5 with a presentation of the different results obtained. Finally, perspectives are proposed in Section 6.

2. KEY FINDING METHODS

Most of the methods developed until now are based on a correlation between the pitch profile (or chromagram) of a piece of music and a key-reference profile. The input pitch profile is a vector representing the likelihood of the pitch occurrences in the piece. The key-reference profile defines the stability of each pitch class in a major or a minor mode [13].

The first key finding algorithm using pitch profile was used in 1990 by Krumhansl and Schmuckler [7] as a result of Krumhansl and Kessler psychoacoustic experiments [8]. Major and minor key reference profiles have been established from these experiments. The stability of each pitch class (tonic, supertonic, mediant..) according to the mode (major or minor) is represented by a single value. All the major key reference profiles can be obtained from the major profile by index rotation. For example, G would have the same value in C Major as D in G Major. Once the key reference profile has been established, the algorithm compares all these 24 profiles (12 Major and 12 Minor) with the input profile. A value is then computed according to the similarity between the input and the key reference profiles. For example, this value can be obtained by correlation between the two vectors. The closer to a reference profile the input profile is, the higher the correlation is. The highest correlation thus defines the preferred key.

Temperley suggested several modifications to this algorithm [14]. In the correlation calculation, he applied a simple scalar product instead of the standard formula used in the Krumhansl and Schmuckler algorithm. He also highlighted the distinction between diatonic and chromatic scale degrees by overweighting triad values (tonic, mediant, dominant). Furthermore, he proposed to raise the value of the leading tone. Several other improvements of Krumhansl's method were presented, using harmonic of pitches [6], specific mode profile to decide between major and minor [1] and Hidden Markov Model [11].

Based on Li and Huron observations [9], Madsen and Widmer recently adopted a different approach [10]. Their method consists of using the information induced by the temporal order of notes. Indeed, two equal pitch class distributions may have different note transitions which may imply different keys. Instead of simply taking the pitch occurrences into account, Madsen and Widmer thus consider interval occurrences and map them in a 12×12 matrix, giving a likelihood to each possible pitch class transition. As in the pitch profile method, an input interval profile is compared to 24 key reference interval profiles, and the highest correlation induces the preferred key.

The use of interval turned out to be very comparable (even slightly better in some cases) to pitch profiles. But since the two methods are successful on different pieces of music, it seems that they bring out different kinds of tonal information. It could be thus interesting to combine these two methods, in order to gather the more possible information.

3. PROCEDURE FOR EXPERIMENTS

In this section, we present the different databases used during our experiments and we discuss the different reference profiles applied.

3.1. Databases

The Music Information Retrieval Evaluation eXchange or MIREX is a contest whose goal is to compare state-of-the-art algorithms and systems relevant for Music Information Retrieval. During the first contest [3], an evaluation topic about audio and symbolic key finding has been organized. The training dataset contains nearly 100 classical music MIDI files selected from the Classical Music Archives [1] and labelled with the key stated in their title.

The procedure of the MIREX contest fixed an evaluation for the results of the different algorithms. It consisted on giving 1 point for a correct retrieved key, 0.5 for a neighbour of the key, 0.3 for a relative and 0.2 for a parallel. The neighbour of a key is a tonality which is distant from the key by a fifth (7 semitones). The relative is distant from a minor third (3 semitones) with a change of mode. The parallel key denotes the key with the same first note (tonic) using another mode. We use this evaluation measure (as MS, for MIREX Score) in the following.

Another collection is considered in this paper. The Finnish Folk Song Database [4] or FinFolk, contains more than 8000 key-annotated melodies collected by the Finnish Literary Society (SKS) from all over Finland. All the files composing this database have been manually labelled. A small number of files having ambiguous or no key information were discarded. This collection is somewhat different from the MIREX database, since the musical pieces composing it are not classical music pieces.

[1] http://www.classicalarchives.com

To obtain the results presented in this paper, we have principally trained our method on the MIREX database before testing it on the FinFolk one. The reference profiles can be learned from the databases. But we choose to apply reference profiles learned from another database or from the literature to be less dependent on the database.

3.2. Reference Profiles

We use two kinds of key reference profiles to test our method. The first, presented by Figure 1, is the key pitch profile proposed by Temperley [13]. This profile defines the stability of each pitch class in a major or a minor mode. The second is the key interval profile of Figure 2. It was proposed by Madsen and Widmer [10] and was learned from 384 Bach chorales (thanks to the author for the personal communication). It represents the frequency of interval occurrences for a given key and mode.

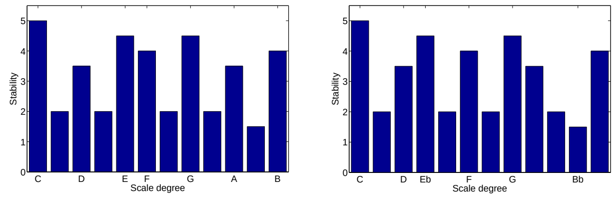

Figure 1. Key-reference pitch profiles from Temperley [13].

If some identical information could be retrieved from the two reference profiles, each can also give some special information. As the profile from Temperley is issued from psychoacoustic experiments, it does not depend on the database. We can observe thus that every pitch of the profile has a non-zero value of stability, even the note not belonging to the scale. For example, the interval profile shows that the interval tonic-supertonic (here C-D) is one of the most frequent intervals in major or minor. Such additional information reinforces our assumption that these reference profiles may be complementary, as Madsen and Widmer conclude in [10].

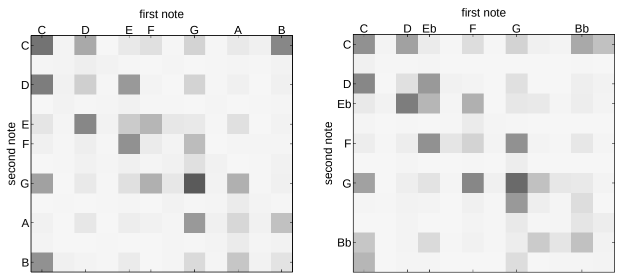

Figure 2. Key-reference interval profiles from Madsen and Widmer [10] learned from 384 Bach chorales (personal communication).

Method	Data	Correct Key	Errors rel.	Errors nei.	Errors oth.	MS
pitch	M	67.4	14.7	17.9	0.0	80.8
interval	M	83.2	4.2	11.6	1.1	90.2
hybrid	M	**83.2**	4.2	12.6	**0.0**	**90.8**
pitch	FF	71.0	7.0	16.5	5.5	81.4
interval	FF	72.4	7.0	15.6	5.0	82.3
hybrid	FF	**75.7**	4.7	15.7	**3.9**	**85.0**

Table 1. Results on MIREX (M) or FinFolk (FF) database. The results are obtained by pitch or interval correlation method, or by the sum of these correlations (hybrid method). They are given in % of the number of files or with the MIREX Score (MS, see Section 3). We can see that the hybrid method gives the best results in terms of correct key detection, of errors induced, and also for the MS value.

4. COMPLEMENTARITY OF METHODS

Since the pitch and the interval profiles give complementary information, we propose to combine them to achieve a new and more accurate key-finding method. We have chosen for now to sum the results of the correlations between the input and the reference profiles. The value of the correlation is obtained by applying a scalar product between the normalized reference profiles and the normalized input profile. The reference profiles are not learned from the tested database.

Results are presented in Table 1. We can see that the hybrid method gives the best results in terms of correct key detection, and in terms of errors induced. The MIREX Score is also improved using this method. One important improvement is that this combination minimizes the number of *real* errors (denoted oth. in tables), i.e. errors that are neither relatives (rel.) nor neighbours (nei.) of the correct key. Thus, there is a very high probability (100% for the MIREX database) that the right key can be found by correcting the first retrieved key.

5. CORRECTION OF KEY ERRORS

We expose here some simple elements of music theory which could help to differentiate potential errors in the key detection. We previously stated that the combination of pitch and interval correlation methods leads to a very accurate method in terms of *real* errors. The idea is then to test the retrieved key to decide whether a relative or neighbour error occurs.

5.1. Flat Input

A first idea is to use a more basic key-reference pitch profile to test the retrieved key. Thus, we have first chosen to use a triad profile $T1$:

$$T1_{major} = [1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0]$$

$$T1_{minor} = [1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0]$$

We have tried empirically other kinds of simple profiles, and we propose $T2$:

$$T2_{major} = [1.5, 0, 0.1, 0, 1, 0, 0, 1, 0, 0.1, 0, 0.1]$$

$$T2_{minor} = [1, 0, 0.5, 1, 0, 0, 0, 1, 0.5, 0, 0, 0.1]$$

The correlation between this profile and the supposed key, its relative or the two neighbours can decide the change of the supposed key. We have noted some great improvements with this technique. Results are presented with other improvements in Table 2.

5.2. Mode or Neighbour Decision

On the same way, we can also use the information given by both interval and pitch profiles of the musical piece to decide the change of the retrieved key, without computing any correlation. The difficulty is here to isolate each wrong case and to identify the quality of the error. Otherwise a correction could also be wrong (i.e. relative correction for a neighbour error can lead to a *real* error).

Considering a retrieved key, a mode decision may detect the correctness of its mode. Chai [1] proposed for example to differentiate the mode of a musical piece by correlating its pitch profile with the simple flat profile :

$$\theta_{major} = [0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0]$$

$$\theta_{minor} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0]$$

However we deal here with musical pieces which could not be analyzed by applying a correlation-based method. Therefore, such approach cannot detect mode errors.

The interval profile of the piece also gives information about the mode. It is obvious that in a major mode, there are many more major third intervals (4 semitones) than minor third (3) from the tonic. The interval matrices presented by Figure 2 illustrates this property. Moreover, if a relative error occurs with a major retrieved key, then the semitone interval from the minor relative tonic and its leading tone may appear. Therefore, we can correct a major key if the number of minor thirds is important regarding to the number of major thirds from the tonic, and if the number of semitone intervals between the 8th and the 9th semitone of the interval profile is high. Note that the main difficulty to correct a key is not to detect its mode but to identify first a mode error.

Results presented in Table 2 show that the improvements we propose may correct some errors. For MIREX database, the number of neighbour errors has been reduced from 12.6 to 6.3. This improvement implies the increase of correct key detection (89.5 instead of 83.2). Concerning FinFolk database, the improvements are limited (76.8 instead of 75.7) but no more error has been introduced.

Data	Corr.	Correct Key	Errors rel.	nei.	oth.	MS
M	no	83.2	4.2	12.6	0.0	90.8
M	yes	**89.5**	4.2	6.3	**0.0**	**93.9**
FF	no	75.7	4.7	15.7	3.9	85.0
FF	yes	**76.8**	4.7	14.5	**3.9**	**85.6**

Table 2. Corrections (Corr.) of the retrieved key from the MIREX (M) or the FinFolk (F) database. The results are given in % of the number of files. They show that applying some flat profiles and other improvements (see Section 5) on the first retrieved key can correct some errors.

6. CONCLUSION AND PERSPECTIVES

In this paper we show that combining some different approaches for key finding can be more accurate than using each method separately. It is due to the complementary of the different methods used. For now, the combination adds the value of the scalar product for the pitch correlation method to the value of the product for the interval correlation method. We have to investigate other possibilities.

Once one candidate key is retrieved, we proposed to test it in order to confirm the estimation. A correction is eventually done, considering the correlation with a flat profile and the occurrences of particular intervals in the musical piece. After first promising results, we need now to refine this decision phase. In fact, as the addition of the number of correct keys to the relative and neighbour errors is generally very high, a solution to this point will help to reach a very good accuracy. We also plan to test the algorithms proposed using other databases with different musical styles.

After the pitch correlation approach then the interval one, we can imagine counting the occurences of longer sequences in the musical piece, to retrieve some additional tonality properties. However, statistics could be less significant since there are less occurrences of each sequence.

7. ACKNOWLEDGEMENT

This work is part of the SIMBALS project (JC07-188930), funded by the French National Research Agency.

8. REFERENCES

[1] Wei Chai and Barry Vercoe. Detection of Key Change in Classical Piano Music. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*, pages 468–473, London, UK, September 2005.

[2] Elaine Chew and Yun-Ching Chen. Real-Time Pitch Spelling Using the Spiral Array. *Computer Music Journal*, 29(2):61–76, 2005.

[3] J. Stephen Downie, Kris West, Andreas F. Ehmann, and Emmanuel Vincent. The 2005 Music Information retrieval Evaluation Exchange (MIREX): Preliminary Overview. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 320–323, London, UK, 2005.

[4] T. Eerola and P. Toiviainen. *Suomen Kansan eSävelmät. Finnish Folk Song Database.* Available: http://www.jyu.fi/musica/sks/, 2004.

[5] Emilia Gómez. *Tonal Description of Music Audio Signals*. PhD thesis, University Pompeu Fabra, Barcelona, Spain, July 2006.

[6] Emilia Gómez and Perfecto Herrera. Estimating The Tonality Of Polyphonic Audio Files: Cognitive Versus Machine Learning Modelling Strategies. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR)*, Barcelona, Spain, October 2004.

[7] Carol L. Krumhansl. *Cognitive Foundations of Musical Pitch*. Oxford University Press, New York, 1990.

[8] Krumhansl, Carol L. and Kessler, E. J. Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys Key-Finding with Interval Profiles. *Psychological Review*, 89(2):334–368, 1982.

[9] Y. Li and D. Huron. Melodic Modeling: A Comparison of Scale Degree and Interval. In *Proceddings of the International Computer Music Conference (ICMC)*, New Orleans, USA, November 2006.

[10] S. T. Madsen and G. Widmer. Key-Finding with Interval Profiles. In *Proceedings of the International Computer Music Conference (ICMC)*, Copenhagen, Denmark, August 2007.

[11] Katy Noland and Mark Sandler. Key Estimation Using a Hidden Markov Model. In *Proceddings of the 7th International Conference on Music Information Retrieval (ISMIR)*, pages 121–126, Victoria, Canada, October 2006.

[12] Schmuckler, Mark A. and Tomovski, Robert. Perceptual Tests of an Algorithm for Musical Key-Finding. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5):1124–1149, 2005.

[13] David Temperley. *The Cognition of Basic Musical Structures*. The MIT Press, 1999.

[14] David Temperley. A bayesian approach to key-finding. In *Proceedings of the Second International Conference on Music and Artificial Intelligence (ICMAI)*, pages 195–206, London, UK, 2002. Springer-Verlag.