KEY-FINDING WITH INTERVAL PROFILES

Søren Tjagvad Madsen

Gerhard Widmer

Austrian Research Institute
for Artificial Intelligence
(OFAI), Vienna

Department of
Computational Perception,
Johannes Kepler
University, Linz

ABSTRACT

Comparing pitch class distributions with predefined key profiles has become the preferred method for key-finding in tonal music, since it was first proposed by Krumhansl and Schmuckler in 1990 [6].

When determining keys using this strategy, information about the temporal order of the notes is not taken into account, although this might contribute additional information relevant for key-finding.

An obvious extension of the pitch class profiles is to look at distributions of intervals – calculate scale degree transition profiles. This idea has not been given much attention in previous research. We conduct a data driven experiment where pitch class profiles and interval profiles are learned from key-annotated music and evaluated on a key-finding task.

1. INTRODUCTION

Pitch class profiles have in the last decades proven to be useful for key-finding. In this paper we propose and evaluate a natural extension of the pitch class profiles: *interval profiles*. A pitch class profile weights the 12 chromatic tones within an octave according to prevalence within a mode – i.e. pitch class profiles have been suggested for major and minor keys. Similarly an interval profile weights for two successive notes the 12×12 possible tone transitions within the octave.

2. KEY PROFILES

Krumhansl and Kessler derived key profiles for the major and minor modes – representing the relative importance of the tones in the chromatic scale [7]. These pitch class profiles were determined by asking listeners to rate how well 'probe tones' fitted into various musical contexts (cadences in major and minor). Figure 1 present major and minor profiles resulting from the experiments. The profiles shown are rooted at C – the distribution for a transposed key is equivalent, which gives us a total of 24 pitch class profiles.

A key-finding algorithm known as the Krumhansl-Schmuckler algorithm was proposed, based on the idea that the pitch class distribution of the notes in a piece of music could reveal its tonality simply by calculating the correlation of the distribution with each of the 12 major and 12 minor profiles, and predicting the highest correlated key [6].

Pitch class distributions can also be estimated from audio (e.g. [4]), and calculating the correlation with key profiles seems to be the preferred way to do key-finding in audio as well, although other approaches exist (e.g. [1]).

The correlations with key profiles are often used as a basic measure of key in a (short) passage of music, and more elaborate key-finding algorithms can be built from these basic measures (see for example [13, 12, 5, 11].

In this paper we will not propose a full-blown key-finding algorithm, but the aim is (at least for the moment) to examine the capability of interval profiles and make a comparison with the well known pitch class profiles.
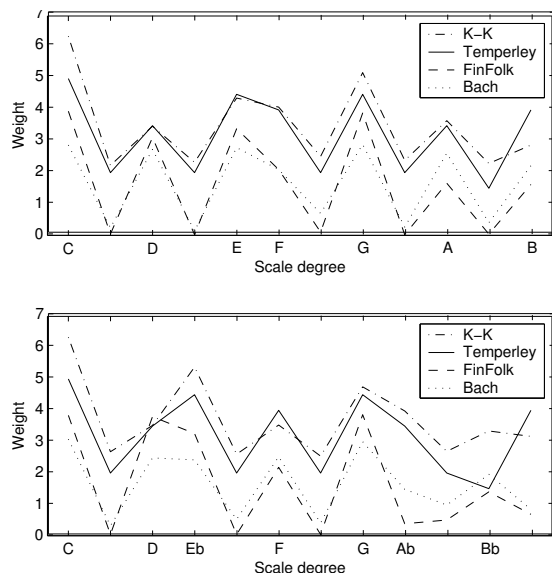
2.1. Pitch Class Profiles

From the Krumhansl-Kessler (K-K) profiles in Fig. 1 it seems that the tonic is the most stable scale degree, followed by the other two members of the triad. In major, the fifth is more emphasised than the third, whereas in minor it is the other way around. Then the fourth, sixth, second and seventh degree of the diatonic scale follows. Non-diatonic scale degrees are considered the least stable.

Parncutt notes that the actual distribution of pitch classes of a passage of tonal music corresponds closely to the key profile [10], suggesting that the profiles can be learned from data. In Fig. 1 we have also depicted the profiles learned from a collection of Finnish folk songs (see below) along with the profiles learned from inventions and fugues by J. S. Bach. Indeed the similarity with the K-K profiles is noticeable. The most striking difference is that many non-diatonic tones are given a much lower weight.

Temperley proposed some changes to the profiles, giving the major and minor profile the same mean (in order to remove the inherent preference for the minor profile) [14]. Also the 7th degree (11th pitch class) in the minor profile was given a higher value, resulting in a more correct classification.

It seems plausible that it takes different key profiles to perform well in key-finding on different data sets. Since we are examining how pitch class profiles and interval profiles can be used for key-finding, we will test profiles

Figure 1. Major and minor pitch class profiles.

Figure 2. Major and minor interval profiles. Dark squares correspond to high values. Rows correspond to the first note of the intervals, columns to the second.

Figure 3. Interval profiles for ascending and descending intervals in minor.

learned from real data, but we also include the K-K profiles and the profiles by Temperley as a reference.
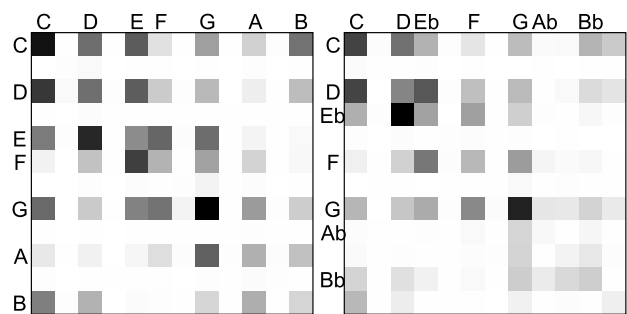
2.2. Interval Profiles

An obvious extension of the pitch class profiles is to look at distributions of intervals, or more correctly, calculate scale degree transition profiles (henceforth 'interval profiles'). The rationale is that the order of the notes might convey knowledge about the key. Two equal pitch class distributions might have different note transitions which might imply different tonalities. [2, 9]

Krumhansl did an additional probe tone experiment, empirically collecting relatedness ratings for all possible ordered pairs of tones presented after C major and C minor key-defining contexts (as cited by [15]). In [15] a key-prediction model based on the K-K pitch class profiles was compared to a key-prediction model based on scale degree transitions. Both models were tested on a dynamic key-prediction task, and compared to human key labellings of the same piece. The predictions from both models were found to correlate equally well with the human evaluation.

The interval profiles thus did not seem to be more powerful than the simpler pitch class profiles, and maybe this is why the method has not been developed further for key-finding. However, Li and Huron discovered that a scale degree transition model turned out to be successful in melodic modeling [8]. The model was shown to be more capable of learning note transitions (in both major and minor modes) than a model trained on intervals alone.

Our scale degree transition profiles will be learned from key-annotated data. A profile is now a 12×12 matrix with a transition probability for each pitch class transition. As with the pitch class profiles, separate profiles will be learned for major and minor keys. We will experiment with undirected interval profiles as well as directed profiles (distinguishing ascending and descending intervals).
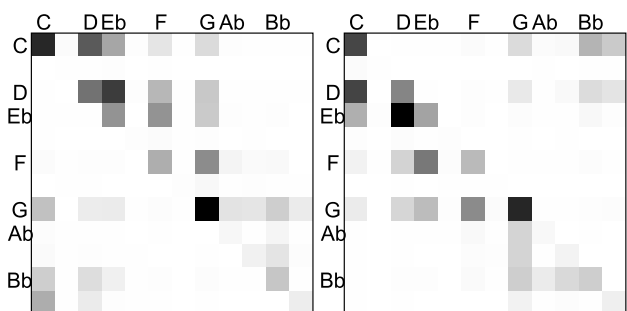
3. APPROACH

3.1. Learning Profiles

Given a sequence of notes, along with a key descriptor stating root note as an integer $0 \leq r \leq 11$ and mode $m \in \{maj, min\}$, we can learn pitch class profiles and interval profiles. Count tables are maintained for this purpose. Music pieces in major and minor will update different tables. The tables are kept with respect to C as root, so a note with (MIDI) pitch p from a file labeled as q major will update entry $(p - q \mod 12)$ in the major pitch class count table (with the value 1).

The interval count tables are updated for every pair of consecutive notes in the 'training' data. Entry $((p_1 - q \mod 12),(p_2 - q \mod 12))$ is updated in the interval count table for notes with pitches p_1 and p_2 occurring in a key rooted on q (also according to mode). Thus intervals greater than 12 semitones are reduced by the octave(s). We also keep directed interval tables for both major and minor. The ascending tables are updated when $p_1 \leq p_2$ and the descending tables when $p_1 \geq p_2$.

After updating count tables from a number of melodies, a profile is calculated stating the probability for every entry. Fig. 1 shows the pitch class profiles (scaled) from the Finnish folk song database, and Fig. 2 shows interval profiles for the same data set. Note the interval profiles are not symmetric; i.e. in minor E♭-D seems more frequent than D-E♭. Fig. 3 depicts the directed interval profiles for minor keys.

3.2. Predicting a Key

Given a melody of which we want to determine the key, a count table is computed from its pitches (using the same methods as described above), and we calculate the correlation of this 'input profile' with each of the major and minor profiles when shifting the root one semi-tone at a time through the 12 possible positions (interval profiles are shifted along the diagonal). The key giving the highest correlation is predicted. The correlation of a pitch class profile and a pitch class distribution vector (input vector) is calculated as the inner product; the correlation between interval profile P_i and interval distribution D_i matrices is found by summing the product of corresponding entries: $cor(P_i, D_i) = \sum_j \sum_k P_i(j,k) D_i(j,k)$. [1]

4. EVALUATION

4.1. Key-Annotated Data

The Finnish Folk Song Database [3] contains more than 8000 key-annotated melodies from different areas of Finland. This collection of MIDI files is very suitable for our experiments. A total of 8325 files have been examined in the experiments. 4956 melodies were labeled with a major key, 3369 files were annotated as being in minor. A small number of files (288) having ambiguous or no key information were discarded.

The files were split into three sets of each 2775 files. In turn two sets were used for building profiles and the third was left out for evaluation, so each set served as evaluation once (three fold cross-validation).

A second corpus of data has been compiled of 384 chorales and 30 inventions by J. S. Bach. In these polyphonic files, note transitions were determined from voice information. We will test the different profiles' ability to determine the key of each of the 48 fugue subjects from the two books of 'Das Wohltemperierte Klavier'.

4.2. Experiments and Results

We are going to test four (pairs of) pitch class profiles: The K-K profiles, the profiles modified by Temperley, flat triad profiles (having all entries of the triad 1.0 and all other 0.0), and the learned pitch class profiles (relative to the data). In addition we will evaluate the learned interval profiles and directed interval profiles.

Temperley argues that flattening the input vector (setting all nonzero entries to 1.0) can be an advantage in some cases [14]. We will run every experiment twice, with the weighted input and with the flattened input vector/matrix.

The left half of Table 1 shows key-prediction correctness scores for our algorithms. The interval profile (with flat input) performs here best overall. Krumhansl proposed to weight the input stimulus relative to the duration of the notes. When determining input values (and when

Dur. weighting	No		Yes	
Flat input	No	Yes	No	Yes
K-K	58.3%	64.8%	61.4%	64.8%
Temperley	69.8%	62.2%	71.0%	62.2%
Triad	63.6%	29.3%	67.6%	29.3%
Learned	**76.5%**	62.3%	**80.2%**	63.2%
Interval	74.8%	**77.5%**	78.7%	**78.6%**
Directed Interval	72.9%	74.7%	76.1%	75.4%

Table 1. Key-prediction correctness for the Finnish Folk Song Database (three fold cross-validation).

Flat input	No	Yes
K-K	67.7%	64.8%
Temperley	70.9%	62.2%
Triad	68.8%	29.3%
Learned	79.7%	65.8%
Interval	**80.7%**	**79.4%**
Directed Interval	79.6%	77.4%

Table 2. Duration weighted and per-file equalised.

learning profiles) we can update the count tables with values proportional to the durations of the notes. The right half of Table 1 shows the prediction scores, when applying note duration weighting. Weighting certainly has a positive effect – all methods increase in correctness. The learned pitch class profile is now the most successful. [2]

When learning note transitions from melodies, not only information about note transitions are learned – also melody specific information is learned. Each file used for training will to some extent bias the model toward a preference for similar melodies. This side effect is thought to cancel itself out when using a large corpus. However, we tried a more active approach: when keeping track of the intervals occurring in a training file, only the square root of the count value for each interval was entered into the table that in the end the profile is constructed from. In this way frequently occurring intervals in one file were given less importance. Table 2 shows that this file equalisation has a positive effect on the interval profiles. [3] In fact the interval profiles perform overall slightly better than most successful pitch class profile.

A benchmark problem in key finding seems to be the 48 fugue subjects from 'Das Wohltemperierte Klavier'. For this problem, profiles were learned from the Bach corpus described in section 4.1. All profiles were given duration weighted input, and in addition the interval profiles were subjected to the aforementioned file equalisation. Results

---

[1] When determining the key based on directed intervals, the correlation is found by averaging the correlation of the input with the ascending and descending profiles respectively.

[2] The weighting according to duration seems to make less sense when speaking about intervals, but nevertheless when weighting the interval count tables proportional to the average duration of the two notes, a small improvement can be noticed.

[3] The effects on the other profiles have also been reported, although this is an interval profile specific feature. Since the same methods are used for calculating a profile from a set of files and an input vector/matrix for a single file, the prediction rates of the fixed (not learned) profiles change (compared to the last experiment) because the input vectors are different.

Flat input	No	Yes
K-K	32	41
Temperley	42	38
Triad	32	10
Learned	33	26
Interval	34	36
Directed Interval	32	31

Table 3. Determining the keys of 48 fugue subjects.

are shown in integers in Table 3. The profile proposed by Temperley is clearly capturing most of the concept here – this time the learned profiles were beaten by expert knowledge. Again we notice a small advantage of the interval profiles over the learned pitch class profiles. Also in this experiment, the directed pitch class model was found to be inferior to the joint profile.

When looking more closely at the prediction results, we notice that 7 times it occurs that the interval profile (flat input) is correct when the learned pitch class profile is wrong. Conversely, in 4 cases the learned pitch class profile is correct while the interval profile (flat input) is wrong.

A similar observation can be made on the Finnish folk song data set. When comparing the results from the learned pitch class profile and the interval profile (weighted input) it shows that at least one of the two profiles is correct on 84.7 % of the files.

The fact that the profiles are not making the same mistakes indicates that they are capturing different concepts, and that it is likely that they can be combined into a better model.

5. CONCLUSION

We introduced the idea of learning scale degree transition profiles for key-finding as an alternative to, and a natural extension of the pitch class profiles. The Krumhansl-Schmuckler key-finding algorithm was extended to handle interval profiles, and an evaluation was performed.

Prediction rates using interval profiles were found to be fully comparable with the methods using pitch class profiles. We believe however, that a real advantage can be achieved by combining the methods. Future research will determine how correlated the prediction results from the different approaches are, and based on that we will look into the possibility of combining the approaches.

We are also considering ways of transferring the concept of interval profile into key-finding in audio.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Ching-Hua Chuan and Elaine Chew. Audio Key Finding: Considerations in System Design and Case Studies on Chopin's 24 Preludes. *EURASIP Journal on Advances in Signal Processing*, 2007.

[2] Diana Deutsch. *The Psychology of Music*. Academic Press, 2nd Edition, 1999.

[3] Tuomas Eerola and Petri Toiviainen. Suomen Kansan eSävelmät. Finnish Folk Song Database. Available: http://www.jyu.fi/musica/sks/, 2004.

[4] Emilia Gómez. Tonal description of polyphonic audio for music content processing. *INFORMS Journal of Computing*, 18(3):294–304, 2006.

[5] Özgür İzmirli. Template Based Key Finding From Audio. In *Proceedings of the International Computer Music Conference*, Barcelona, 2005.

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

[7] Carol L. Krumhansl and E. J. Kessler. Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, 89:334–368, 1982.

[8] Yipeng Li and David Huron. Melodic modeling: A comparison of scale degree and interval. In *Proceedings of the International Computer Music Conference*, New Orleans, 2006.

[9] Rie Matsunaga and Jun ichi Abe. Cues for key perception of a melody: Pitch set alone? *Music Perception*, 23(2):153–164, 2005.

[10] Richard Parncutt. Tonality as Implication-Realisation. In P. Vos and M. Leman, editors, *Proceedings of Expert Meeting on Tonality Induction*, pages 121–141, Holland and Belgium, 1999.

[11] Geoffroy Peeters. Musical key estimation of audio signal based on hidden markov modelling of chroma vectors. In *Proceedings of the 9th International Conference on Digital Audio Effects*, Montreal, 2006.

[12] Ilya Shmulevich and Olli Yli-Harja. Localized key-finding: Algorithms and applications. *Music Perception*, 17(4):531–544, 2000.

[13] David Temperley. What's Key for Key? The Krumhansl-Schmuckler Key-Finding Algorithm Reconsidered. *Music Perception*, 17(1):65–100, 1999.

[14] David Temperley. *The Cognition of Basic Musical Structures*. MIT Press, Cambridge, MA, 2001.

[15] Petri Toiviainen and Carol L. Krumhansl. Measuring and modelling real-time responses to music: The dynamics of tonality induction. *Perception*, 32, 2003.