A chord distance metric based on the Tonal Pitch Space and a key-finding method for chord annotation sequences

Lucas Marques[1] **Luiz Eduardo da Silva**[1] **José Maurício Schneedorf Ferreira da Silva**[1]

[1]Federal University of Alfenas (UNIFAL-MG)
Av. Jovino Fernandes Sales, 2600 – 37133-840 Alfenas, MG

a14017@bcc.unifal-mg.edu.br

Abstract. *Music Information Retrieval (MIR) is a growing field of research concerned about recovering and generating useful information about music in general. One classic problem of MIR is key-finding, which could be described as the activity of finding the most stable tone and mode of a determined musical piece or a fragment of it. This problem, however, is usually modeled for audio as an input, sometimes MIDI, but little attention seems to be given to approaches considering musical notations and music-theory. This paper will present a method of key-finding that has chord annotations as its only input. A new metric is proposed for calculating distances between tonal pitch spaces and chords, which will be later used to create a key-finding method for chord annotations sequences. We achieve a success rate from* 77.85% *up to* 88.75% *for the whole database, depending on whether or not and how some parameters of approximation are configured. We argue that musical-theoretical approaches independent of audio could still bring progress to the MIR area and definitely could be used as complementary techniques.*

1 Introduction

In western music, tonality is a basic concept thoroughly studied by authors as Riemann [1] and Schoenberg [2], and could be defined in many ways. For example, a brilliant summarization of Schoenberg's thoughts on tonality and tonal function can be found on Carpenter [3]:

> Tonality for Schoenberg is not merely a certain collection of pitches of a scale, but more importantly, a kind of centricity. All pitches of a key-collection are related to a single tonal center, each in a specific way. The function of a single tone is signified by the degree of the scale it represents. The function of a chord depends upon its root, which is, in turn, the scalar degree upon which the chord is constructed. Tonality, then, is a set of functions of scalar degrees.

Since tonality is a fundamental concept of musical theory and many other information can be derived from the tonality of a piece, it is intuitive that within the Music Information Retrieval area, finding the tonality – also called key – of pieces would also be a fundamental problem. This paper presents a method of key-finding based on previous musical-theoretical work done by Lerdahl [4], a model of tonality named Tonal Pitch Space (TPS) which was corroborated by psychological experiments and matches musical intuition.

In Section 2 we discuss other methods of key-finding and theoretical models of tonality, and explain why we chose Lerdahl's Tonal Pitch Space. In Section 3, we focus on the TPS model, explain how it works and show its psychological and musical-theoretical background. Section 4 we introduces a new metric based on Lerdahl's model to calculate the distance between a tonal pitch space and any chord. Section 5 describes a method of key-finding utilizing that previous metric and Section 6 discuss its results on the test database created for this paper. More discussion will be brought about how audio-independent approaches could contribute to the MIR area in general in Section 7.

2 Related Work

One important problem in MIR is key-finding. Several models, algorithms and techniques were presented in the past considering the simple task of finding the global key of a determined piece of music or a local key within a subset of said piece – and, of course, methods of key-finding sensitive to context, with the objective of detecting changes of tonality (modulations). Chew [5] proposed, in 2002, a geometrical model of tonality that used a "Spiral Array" as a way to represent keys, chords, intervals and pitches. Pauws [6] describes a way of extracting the key of an audio source using chromagram computation. İzmirli [7] presented a model that uses a low-dimensional tonal representation. Hu [8] developed a probabilistic model to determine the key and also modulations on a MIDI database containing classical pieces from artists such as Bach, Mozart and Rachmaninoff. However, for this paper, we give a special attention to another model of tonality, proposed by Lerdahl [4]. Lerdahl's *Tonal Pitch Space* (TPS) is a model that correlates with empirical data provided by Krumhansl [9] and matches music-theoretical intuitions about tonality.

This paper is based upon the TPS model specially because we do not use audio from musical pieces as an input, but their chords annotations only. Thus, any model that uses audio as its main source of data would not serve our purpose. Since chord notation does not consider different octaves, solutions related to MIDI with notes from all octaves would also not bring any benefit to our objective. So, on top of all previous reasons to use Lerdahl's TPS, its simplicity is the main reason why we choose it. It completely matches the simplicity of chord annotations.

3 Tonal Pitch Space

The TPS is a model of tonality supported by empirical data from psychology [9] and matching human intuitions. It can be used to calculate the distance between all imaginable chords and it is based upon a notion of hierarchy between musical intervals in western music. That hierarchy is defined according to stability and what precisely fits with the experimental data provided by Krumhansl [9].

The best way to understand the TPS is looking at its structure. In Figure 1 we built the TPS of C major. The space is defined by five levels of stability, from the most to the less stable. The first one, *level a*, is the root level, which contains only the root of the TPS basic chord. In this case, since we are looking at the space of C major, it is C. This is in line with the fact that the most stable and consonant intervals are the octave and the unison.

Tonal Pitch Space

level a:	C											
level b:	C							G				
level c:	C				E			G				
level d:	C		D		E	F		G		A		B
level e:	C	Db	D	Eb	E	F	Gb	G	Ab	A	Bb	B

Figure 1: The TPS of C Major

Level b contains the root and the fifth interval which, in this case, is G. The second most stable interval reflects on the second level of the TPS. On the third level, *level c*, we have the triadic level, containing all the notes of the chord that generates the harmonic field represented by the TPS. The C major chord is composed of {C, E, G}. This set notation {} will be used ahead in this paper, and it is also very common to use the integers notation for the TPS (Figure 2). Next, *level d* is the diatonic level, containing the natural scale of the TPS. Here, we have the major scale of C. The last one, *level e*, is the most unstable of all, containing all 12 notes used in western music. We call it the *cromatic level*.

level a:	0											
level b:	0							7				
level c:	0				4			7				
level d:	0		2		4	5		7		9		11
level e:	0	1	2	3	4	5	6	7	8	9	10	11

Figure 2: The TPS of C Major with the numerical notation

4 Proposed Metric

In this section we discuss the metric proposed by this paper and its main differences between the metric proposed by Lerdahl himself when he presented the framework of the Tonal Pitch Space. Our metric differs specially within the two compared objects.

When introducing the concept of the TPS, Lerdahl [4] defined a method for calculating the distance between any two chords in the context of a key, and it is composed of two elements: circle-of-fifths distance and uncommon tones[1]. Lerdahl proposes a distance formula $d(x, y) = j + k$ where $d(x, y)$ is the distance between chords x and y, j is the minimal applications of the circle-of-fifths rule needed to transform x into y and k is the number of non-common pitch classes in the levels (a-d) within the basic space of y compared to the levels (a-d) in the basic space of x.

Our metric follows the same universal circle-of-fifths rule, but a different approach when considering common tones. Our formula is also conceived with the notion that there are 24 tonal pitch spaces (considering all 12 notes and major/minor qualities) and a much larger number of chords. Those chords are not necessarily part of a space, they don't necessarily fit into major/minor scales and they don't belong *a priori* to any harmonic field – which we imply that here are represented by the tonal pitch spaces. Thus, we do not compare two equal objects from a same metrical space. That said, our metric could be described as it follows:

$$d(S, x) = j + k$$

Where $d(S, x)$ is the distance between any tonal pitch space S and any chord x, j is the same previous circle-of-fifths rule[2] and k is the sum of the number of uncommon tones between the levels (a-c) of S and x and the difference between all notes of x and the *level d* of S. Figure 3 shows an example of the rule described previously. There we compare the distance between the TPS of C major and the chord G7. Check Figure 1 in need of remembering the levels of the TPS of C major.

Levels of *S*	**Levels of *x***
a: {C}	a: {G}
b: {C, G}	b: {G, D}
c: {C, E, G}	c: {G, B, D, F}
d: {C, D, E, F, G, A, B}	

Distance Calculation

uncommon(a) = 2, {C, G}
uncommon(b) = 2, {C, D}
uncommon(c) = 5, {C, E, B, D, F}

difference between *c* of *x* and *d* of S = 0, {}
circle-of-fifths(C → G) = 1

Total distance = 10.

Figure 3: Distance between the TPS of C major and the chord G7

[1] An uncommon tone is a tone present only in one of the two compared objects

[2] If a chord is non-diatonic, the circle-of-fifths rule returns a maximum value of 3. Major and minor qualities are also considered and, for example, the distance between C and Am on the circle-of-fifths is here considered as 1 instead of 0, since we have one step down to change from the major to the minor circle.

It is interesting to notice that this metric has a maximum value of 23, which can be obtained when comparing a C major tonal pitch space with a B chord containing all of the 12 notes of the cromatic scale. Such chord containing all notes is not musically practical, but this information could be useful for design, computation and application purposes of this metric. The minimum possible value is, as expected for a metric, a value of 0 when comparing a tonal pitch space to its basic chord, such as comparing that same previous C major tonal pitch space to its C major basic chord.

In the test database discussed in Section 5, we have found, comparing to the TPS of C major, chords on all distances from 0 to 21 - surprisingly close to the 23 maximum value. From 0 to 6 we have only chords with C as the root of the chord. The two chords that result in a distance of 21 are G#m7(9) and B6(9). The results of distances from chords to tonal pitch spaces match with musical-theoretical intuitions and reinforce the value of the tonal pitch space as a model for tonality. The results also match intuitions about chords within a given harmonic field.

Table 1 shows the distances between all seven degrees of C major. It is interesting to note how IV, V and VI have the same value and VII is closer than II is. This could be interpreted as the presence of VII being an indication of the tonality since it is present in only one harmonic field and it is also a chord of tension that suggests resolution on the tone center, matching musical intuitions.

5 A Key-Finding Method

Considering we have a metric to determine the distance between any chord and any tonal pitch space, the next step is to use it to develop a method to estimate tonality. We have done it following the premise that if a piece has a global key, the sum of the distances between the chords present in that piece should be minimum when comparing to the TPS of that global key. This was also assumed by de Haas [10] in his works based on Lerdahl's Tonal Pitch Space.

Thus, for each musical piece analyzed, we create 24 tonal pitch spaces (one for each major and minor of the twelve notes of the cromatic scale) and 24 variables of distance that are the sum of the distances between each chord present in the musical piece and each one of the tonal pitch spaces. We also have a multiplication factor MF for when the first or the last chord of the song is the same as the basic chord from the analyzed TPS – for example, if we are calculating the distances considering a TPS of Am and a song that starts or ends with Am.

Chord	Degree	Distance
C	I	0
Dm	II	14
Em	III	10
F	IV	9
G	V	9
Am	VI	9
Bdim	VII	13

Table 1: Distances from the chords of the C major harmonic field to a C major TPS

This multiplication factor improves the estimation results because musical pieces commonly start with the tonality to introduce it or end with it to resolve. From the set of 240 songs used for this paper, 180 songs start with the chord of the tonality, 145 songs end with the chord of the tonality and, within these two subsets, there is an intersection of 105 songs that start and also end with the chord of the tonality. Only 20 songs from the 240 do not start nor end with the chord that defines their tonality – this indicates how approximation features based on this information could benefit estimation of tonality. That said, the following formula describes the calculation of the distance between a musical piece and a tonal pitch space:

$$total\_dist = MF \sum_{i=1}^{n} TPS\_distance(S, c_i)$$

Where MF = OC, if only one chord – the first or the last – is the same as the basic chord of S; or MF = BC, if both chords are the same. The $TPS\_distance()$ function is the same we introduced at Section 4, applied to a tonal pitch space S and each chord c_i from the n chords of the piece. For the database used in this paper, we have tested values from 0.80 to 0.99 for OC and from 0.75 to 0.99 to BC – two digits precision. We have found that, from all permutations, the best results are obtained with values of 0.90 and 0.83 for OC and BC, respectively. For every configuration tested within these two ranges, there was an increase in the success rate of the prediction – indicating that perhaps this approximation step is almost never prejudicial to this method and should always be used.

6 Tests and Results

The objective of this paper was, *a priori*, the harmonic analysis of song chords available on websites such as Ultimate Guitar Tabs [11] and Cifra Club [12]. We chose the latter because of its vast repertoire, specially on brazilian music, such as Bossa Nova and Samba. We gathered[3], then, 240 songs of 104 artists from 25 different genres – Table 2 shows the distribution of songs per genre. Cifra Club pages provide the tonality of each piece. However, when analyzed thoroughly, that information was usually wrong.

Because of this lack of trust in the available information, we were forced to listen, play and search for information about each one of the 240 songs and their primary or global tonality. This was done in order to increase the reliability of the test database and guarantee ground-truth. This also helped deciding the first primitive values of

[3]For this part we simply downloaded the webpages of the songs and then parsed them to extract, from the HTML code, the chords and other relevant information as genre, title, etc.

Genres	Number of Songs
Alternative	17
Blues	1
Bossa Nova	11
Country	12
Dance	1
Disco	9
Folk	15
Funk	2
Grunge	2
Hard Rock	2
Heavy Metal	1
Indie	3
Jazz	6
Jovem Guarda	11
MPB	64
Pagode	4
Pop and Pop Rock	11
R&B	3
Progressive Rock	9
Rock and Rockabilly	31
Romantic	1
Samba	11
Soul	12

Table 2: Database's distribution of musical genres. Those genres were not checked for each song, but still indicate a good level of variety in the sample

the multiplication factor MF previously described – this process was basically noting how distant was the correct verified tonality from the one mistakenly estimated by our algorithm and whether the song had the first and/or last chord matching the tonality. This way, we could deduce how much should be the reduction applied by each one of the two parameters.

For the whole set of 240 songs, we show here two tests: one where the values of BC and OC were 1, i.e, without having the multiplication factor reducing the distance depending on the first and/or last chords. Without this feature, our method was able to achieve a success rate on estimating the correct tonality of 78.75%, with a 51 mistakes. This is already a promising success rate, and yet using different values – 0.82 and 0.90, tuned by hand – on a second test we are able to achieve 88.75% of success, with only 27 mistakes.

We also created categories of complexity[4] based on the number of different chords used on each song. This database is reasonably diverse on that matter, with songs going from having only three chords up to the most complex song – O Caderno, by Toquinho – having a total amount of 40 different parsed chords. With that in mind, we have created five arbitrary categories of complexity to test if the algorithm performs better or worse when estimating the tonality of songs with a lot or just a few chords.

[4]The categories of complexity were created to separate songs that most probably use chords not present in the harmonic field of the song.

Complexity Category	No. of Songs	Success Rate (%)
20+ chords	16	68.75
15–20 chords	44	70.45
9–14 chords	75	84.00
5–8 chords	67	79.10
less than 5 chords	54	77.77

Table 3: Categories of complexity, number of songs in each of them and success rates ($BC = OC = 1$).

Complexity Category	No. of Songs	Success Rate (%)
20+ chords	16	81.25
15–20 chords	44	88.63
9–14 chords	75	89.33
5–8 chords	67	88.06
less than 5 chords	54	88.89

Table 4: Success rates for each category when BC and OC are equal to 0.83 and 0.90, respectively.

The tests results are exposed in Table 3, without using the multiplication factors and also in Table 4 with the optimized configuration. Despite a considerable smaller success rate on songs with 20 or more chords, the method achieves promising results in all categories, specially when using the approximation feature of the multiplication factors.

7 Discussion and Conclusions

In this paper we have introduced a metric based on Lerdahl's Tonal Pitch Space to create a key-finding method. This method finds a single global tonality, but could possibly be used in local key-finding scenarios, given a subset of a piece. There are limitations from the simplicity of the input, given that they were only chord annotations without information of the duration of each one of them. It could also be discussed in further research a way of cross-validating the values of the parameters OC and BC in order to increase the generalization capacity of this small yet very relevant predictive part of the model.

However, even with simple chord annotations as input and every context based limitation, this method still managed to have great success rates when estimating the tonality of western popular music. This method could be used along with others, specially the ones with audio as input, to help in the task of finding the central key of a determined piece or.

It is also our argument that Music Information Retrieval could possibly benefit even more from theory-based and annotation-based methods. While audio and MIDI based computational and mathematical approaches seem to be dominant in this area, we might still have large space for progress for audio-independent techniques and approaches more rooted in musical-theoretical concepts.

References

[1] RIEMANN, H. *Harmony Simplified: or, The Theory of the Tonal Fuctions of Chords.* London: Augener Ltd., 1893.

[2] Schoenberg, A. (2010). *Theory of Harmony*. University of California.

[3] Carpenter, P. (1983). *"Grundgestalt" as Tonal Function.* Music Theory Spectrum, 5, 15-38. doi:10.2307/746093. p. 16-7.

[4] Lerdahl F. (1982) *Tonal Pitch Space*. In Music Perception: An Interdisciplinary Journal, Vol. 5 No. 3, Spring, 1988; (pp. 315-349) DOI: 10.2307/40285402.

[5] Chew E. (2002) *The Spiral Array: An Algorithm for Determining Key Boundaries.* In Anagnostopoulou C., Ferrand M., Smaill A. (eds) Music and Artificial Intelligence. ICMAI 2002. Lecture Notes in Computer Science, vol 2445. Springer, Berlin, Heidelberg.

[6] Pauws S. (2004) *Musical key extraction from audio*. In Proceedings of the 5th ISMIR, Barcelona, Spain., pages 96–99, 2004.

[7] Izmirli O. (2006) *Audio key finding using low-dimensional spaces.* In Proc. of the International Symposium on Music Information Retrieval (ISMIR), pages 127–132, Victoria, Canada, 2006.

[8] Hu, D., & Saul, L. K. (2009) *A Probabilistic Topic Model for Unsupervised Learning of Musical Key-Profiles.* In ISMIR (pp. 441-446).

[9] Krumhansl, C. L., & Kessler, E. J. (1982). *Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys.* In Psychological Review, 89(4), 334-368.

[10] de Haas W.B., Robine M., Hanna P., Veltkamp R.C., Wiering F. (2011). *Comparing Approaches to the Similarity of Musical Chord Sequences.* In: Ystad S., Aramaki M., Kronland-Martinet R., Jensen K. (eds) Exploring Music Contents. CMMR 2010. Lecture Notes in Computer Science, vol 6684. Springer, Berlin, Heidelberg.

[11] ULTIMATE GUITAR TABS - 1,100,000 songs catalog with free Chords, Guitar Tabs, Bass Tabs, Ukulele Chords and Guitar Pro Tabs!. [online] Available at: `https://ultimate-guitar.com` [Accessed 12 Jun. 2019].

[12] Cifra Club. (1996). Cifra Club - seu site de cifras e tablaturas. [online] Available at: `https://www.cifraclub.com.br` [Accessed 10 Jun. 2019].