

Article

A Comparison of the Music Key Detection Approaches Utilizing Key-Profiles with a New Method Based on the Signature of Fifths

Michalina Kania ¹, Tomasz Łukaszewicz ², Dariusz Kania ^{2,*} , Katarzyna Mościńska ² and Józef Kulisz ²¹ Faculty of Electrical Engineering, Silesian University of Technology, 44-100 Gliwice, Poland² Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, 44-100 Gliwice, Poland* Correspondence: dariusz.kania@polsl.pl

Abstract: This paper compares approaches to music key detection based on popular key-profiles with a new key detection method that utilizes the concept of the signature of fifths. The signature of fifths is a geometrical music harmonic-content descriptor. Depending on the scenario, it may reflect either the multiplicities of occurrences or the aggregated durations of individual pitch-classes of the chromatic scale in a given fragment of music. In this study, we compare the efficacies of a few strictly correlational key recognition approaches based on music key-profiles (i.e., Krumhansl–Kessler, Temperley, and Albrecht–Shanahan) with a new method that implements the concept of the signature of fifths. All the experiments were performed in a collection of music pieces comprised of preludes by J. S. Bach (The Well-Tempered Clavier-Book I), preludes by F. Chopin (Op. 28), and songs by The Beatles (from the album A Hard Day’s Night). In the scenario implementing the aggregate durations of individual pitch-classes, based on the analysis of the shortest initial fragments of music for which the key was indicated in all the considered approaches, the key detection efficacy obtained with the method using the signature of fifths was greater than the efficacies obtained with the strictly correlational approaches utilizing key-profiles (on average by 9.27 pp). In the case of the analogous analysis carried out for the scenario implementing the multiplicities of occurrences of individual pitch-classes, on average, greater efficacy was observed for the strictly correlational approaches based on key-profiles (by 2.7 pp). The conducted experiments confirmed the new key-detecting method offers advantages in computational simplicity, stability of decision making, and the ability to successfully determine the key based on a very short fragment of music.

Keywords: music information retrieval; computer music analysis; music key detection; signature of fifths; key-profiles



Citation: Kania, M.; Łukaszewicz, T.; Kania, D.; Mościńska, K.; Kulisz, J. A Comparison of the Music Key Detection Approaches Utilizing Key-Profiles with a New Method Based on the Signature of Fifths. *Appl. Sci.* **2022**, *12*, 11261. <https://doi.org/10.3390/app122111261>

Academic Editors: Luigi Fortuna and Giancarlo Mauri

Received: 4 October 2022

Accepted: 31 October 2022

Published: 7 November 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Major–minor tonality has long been present in Western Classical Music, providing means for shaping the mood of musical compositions. Nowadays, it plays an important role in many music-related algorithms or systems, for example music genre recognition [1–3], assessment of musical tension [4–7], music visualization systems [8–11], music data mining [6], computer-aided composition software [12–14], and determining the harmonic structure of created pieces [15,16]. It is worth mentioning that in recent years, methods implementing neural networks [17–19] and other machine learning approaches [20–22] have gained popularity in many areas of music analysis.

Well-known key-detection methods use tonal models as reference points. The history of tonal models dates to Pythagoras, who defined the principles of the mathematical description of consonances. Consonant-sounding intervals, which form major and minor chords, were reflected in the Tonnetz proposed by Euler, which showed the most important harmonic relationships in major and minor scales. Extensions of Euler’s harmonic networks

are various types of spiral array models [23–25], which can constitute the basis for chord-detection algorithms [25–28]. They can also be used for the key recognition purposes [29,30].

There are many different models known to represent the relationships between tones and their associated keys. Interesting concepts of tonal description include, for example, Longuet–Higgins tonal maps [31,32], the geometrically regular helical models presented in [23], the spiral array models that show the interrelations among musical pitches [24], and the orbifold that represents the musical chord [11]. Unfortunately, they involve very complex calculations, which make hardware implementations impractical.

Key recognition methods based on key-profiles are much simpler [18,19,33–39]. Key-profiles generally assign larger weights to diatonic tones than to non-diatonic ones. The largest weights are assigned to the tones on which the triad chords are built: Tonic; Dominant; and Subdominant. Various methods have been used to create the key-profiles. For example, some have been based on extensive experimental studies [36,37], statistical analyses [10,40,41] or creation of probabilistic models [38]. In our opinion, the key-profiles created by Albrecht and Shanahan, with the help of artificial neural networks [33] deserve special attention because they recognize keys very effectively.

Some of the key-profiles resulted from an in-depth analysis of already existing profiles (e.g., the analysis of Krumhansl–Kessler key-profiles) supported by advanced models based on probabilistic reasoning. This led to the creation of the Temperley profiles [38], [39]. It is also worth noting that in some cases the process of creating key-profiles was based on the analysis of audio files [34,42].

In general, using key-profiles to detect music keys is rendered down to the calculation of the correlation coefficients of the input vector that represents the analyzed piece of music with 12 major and 12 minor key-profiles. The largest correlation coefficient indicates the key of the analyzed piece. It is an open question of how to select the analyzed music sample and which key-profiles to use to obtain the best results. Although the algorithms for key detection based on key-profiles are among the simplest, their implementations can be cumbersome due to the necessity of a relatively large number of multiplication operations, especially in the case of electronic instruments (e.g., keyboards). A much simpler approach to the key-detection problem can be established utilizing the concept of the music signature, introduced in [35], which in this paper is referred to as the signature of fifths. We decided to change the name of this construct to emphasize its strong association with the circle of fifths. As shown in [35], a structural analysis of the signature of fifths enables the determination of the key of a piece of music. It is worth noting that one can also use this method to determine the key signature without the need for computation of any correlation coefficients [43]. Observing the variations of the signature of fifths over time enables a “rough” evaluation of the harmonic structure of a given piece of music [44]. The signature of fifths can also be used to evaluate the style of an arrangement of carols [45]. In this paper, inspired by the observed properties of the signature of fifths [35,43], we present a broader, multi-criteria, comparison of key-recognition approaches based on the signature of fifths as well as the well-known key-profiles of Krumhansl and Kessler, Temperley, and Albrecht and Shanahan.

The main goal of this article is to compare the key detection method based on the signature of fifths with the simplest key detection approaches known from the literature that utilize key-profiles. The greatest value of the method implementing the signature of fifths is its computational simplicity. As it turns out, the orientation of the main directed axis of the signature of fifths is informative enough to enable narrowing of the set of 24 potential keys to just 2 relative keys, hence reducing the number of multiplication operations associated with the performed correlations. A great advantage of the key detection based on the signature of fifths is the stability of made decisions, i.e., the indicated key is most likely correct and generally not prone to changing when the analyzed fragment of music is extended. This property results from the way in which the main directed axis of the signature of fifths is determined. The foundational assumptions behind the key detection method using the signature of fifths are that the diatonic tones are always located

on one side of the main directed axis, and the orientation of the main directed axis is not very sensitive to the presence of a few off-scale tones.

Following this introduction, in Section 2 we provide the basic theory behind the concept of the signature of fifths and discuss the key recognition schemes considered in the study. Section 3 presents the results of our experiments along with a discussion of the strengths and weaknesses of the key detection approaches we compared. In Section 4, we summarize the study and indicate our further research activities.

2. Materials and Methods

Musical compositions can be represented with tones corresponding to the twelve pitch-classes of the chromatic scale: C; C \sharp /D \flat ; D; D \sharp /E \flat ; E; F; F \sharp /G \flat ; G; G \sharp /A \flat ; A; A \sharp /B \flat ; B. Of course, the tones comprising a particular piece may belong to different octaves. Knowing the number of occurrences of notes associated with individual pitch-classes in a given fragment of music, or their durations, it is possible to define a set of values X (1):

$$X = \{x_C, x_{C\sharp/D\flat}, x_D, x_{D\sharp/E\flat}, x_E, x_F, x_{F\sharp/G\flat}, x_G, x_{G\sharp/A\flat}, x_A, x_{A\sharp/B\flat}, x_B\} \quad (1)$$

where x_i is the multiplicity of occurrences of tones associated with a particular (i -th) pitch-class or the aggregate duration of tones corresponding to that pitch-class. Using (1) we define the vector K (2):

$$K = [k_A, k_D, k_G, k_C, k_F, k_{B\flat}, k_{E\flat}, k_{A\flat}, k_{D\flat}, k_{F\sharp}, k_B, k_E] \quad (2)$$

where:

$$k_i = \frac{x_i}{\max(X)} \quad (3)$$

The elements of the vector K are sorted in accordance with the succession of the pitch-classes defining the circle of fifths, beginning from the A and proceeding counterclockwise.

Definition 1. *The signature of fifths.*

The signature of fifths (also referred to as the music signature [35]), corresponding to a given fragment of music, is a set of twelve polar vectors $\{\vec{k}_i : i = A, D, G, \dots, E\}$, whose coordinates (r_i, ϕ_i) are determined with the following assumptions:

- the length of the i -th polar vector is equal to the i -th value defining the vector K, which can be obtained either using the multiplicities of occurrences of tones associated with individual pitch-classes or the aggregate duration of tones corresponding to those pitch-classes:

$$r_i = \left| \vec{k}_i \right| = k_i \quad (4)$$

- the angle of the i -th vector is determined with the following relationship:

$$\phi_i = j \cdot 30^\circ \quad (5)$$

where $j = 0|_{i=A}$, $j = 1|_{i=D}$, and so on.

Example 1. *The two ways of creating the signature of fifths.*

Let us consider the two ways of creating the signature of fifths corresponding to the fragment of music shown in Figure 1. In the first case, we assume that the length of each constituent vector of the signature of fifths reflects the multiplicity of occurrences of notes associated with a particular pitch-class. In the second case, the lengths of the vectors will reflect the aggregate durations of individual pitch-classes. In either case, a certain time resolution of the analysis needs to be assumed, e.g., eight-note or quarter-note. In this

example, we will create the quarter-note signatures of fifths corresponding to the fragment of music shown in Figure 1. The multiplicities of occurrences and aggregate durations of individual pitch-classes comprising the chromatic scale, obtained for this fragment, are presented in Table 1.



A directed axis of the circle of fifths $Y \rightarrow Z$, for which $[Y \rightarrow Z]$ assumes the maximum value, is called the main directed axis of the signature of fifths.

The main directed axes obtained for the signatures of fifths shown in Figure 2 are illustrated in Figure 3. The values on the outer edges of the drawn plots correspond to the characteristic values of individual directed axes. The dashed arrow present in each plot represents the main directed axis of a given signature of fifths. For the considered fragment of music, the directions of these axes happen to be the same.

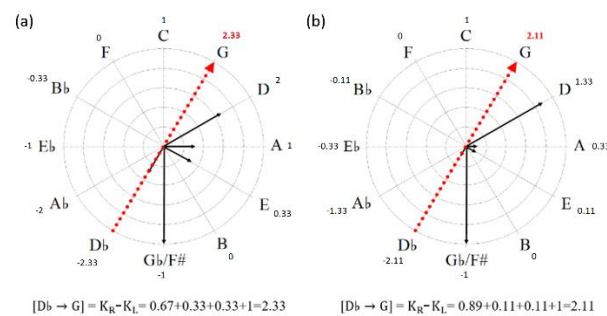


Figure 3. The directed axes of the signatures of fifths corresponding to the fragment of the piece of music presented in Figure 1: (a) obtained for the multiplicities of occurrences of individual pitch-classes; (b) obtained for the aggregate durations of individual pitch-classes.

There are multiple ways to detect the tonality of a piece of music [18,24,36,37]. One simple and effective method is based on key-profiles, which depict a particular key by assigning weights to individual pitch-classes. It is worth mentioning, though, that the key-profiles currently available tend to assign higher weights to diatonic tones than to non-diatonic ones.

The first key-profiles were developed by Krumhansl and Kessler [36,37] as a result of an experiment in which listeners were asked to rate the degree to which each tone of the chromatic scale matched a seven-tone major and a seven-tone minor scale. The degree to which the tones were related to each other was also examined and the obtained relationships were expressed as the stability of a given tone in a particular key. Findings indicated that the perceived stability of a tone was dependent on the tones that preceded it. This aspect is important because the same tones are characterized by different harmonic relationships in different keys. For example, the “C” tone in the key of C-major is a stable tonic, whereas in the D-major key it is an unstable tone (i.e., extraneous in the harmonic context). The obtained matching degrees comprise the normalized 12-element vectors featuring a given key and correspond to the weights associated with particular pitch-classes. Key-profiles representing any tonality can be obtained by relating the largest weight of such a vector with a tonic. An algorithm for key detection using key-profiles was presented by Krumhansl in [37]. It involved the calculation of the Pearson’s linear correlation coefficients between the weights of the individual pitch-classes found in the analyzed fragment of music and the weights associated with 12 major and 12 minor key-profiles. The correlation coefficient with the largest value indicates the key of the analyzed piece. This algorithm is versatile and can be used to determine the key using any set of key-profiles.

David Temperley proposed his key-profiles [38,39] after observing the results obtained using Krumhansl’s algorithm [37]. Krumhansl–Kessler profiles led to good results primarily for very short fragments of compositions (e.g., the first four notes of a given piece). These key-profiles have some disadvantages, though; for example, the detected key frequently changes when the number of notes of the analyzed fragment of music increases. In addition, according to Temperley, the Krumhansl–Kessler profiles exhibit a bias towards minor scales because of the different cumulative values obtained for the component tones of the harmonic triads in major and minor keys. According to Temperley, unifying the cumulative value of the weights for the component tones of the harmonic triad for both tonalities (major/minor), as applied in Temperley’s profiles, eliminates this bias.

There are other key-profiles in the literature [9,33,40,41]; however, in our opinion, those created by Albrecht and Shanahan deserve particular attention [33]. They were developed utilizing artificial neural networks trained on pieces created by well-known composers such as Johann Sebastian Bach, Frédéric Chopin, Ludwig van Beethoven, and Joseph Haydn. Only the first and last eight bars of each piece were considered. The key detection was based on the analysis of the Euclidean distance between the points representing the analyzed fragments of compositions and the points associated with 12 major and 12 minor keys. The smallest distance indicated the key. A comparison of the major and minor key-profiles proposed by Krumhansl and Kessler (K-K), Temperley (T), and Albrecht and Shanahan (A-S) is shown in Figure 4. The values presented in the graphs are normalized by the maximum value corresponding to the tonic (0).

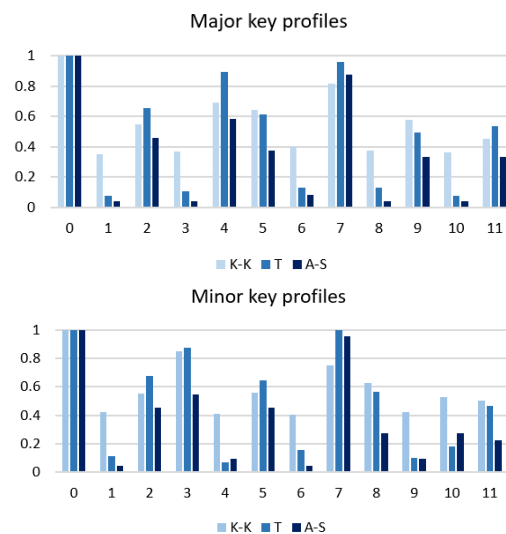


Figure 4. A comparison of the major and minor key-profiles proposed by Krumhansl and Kessler (K-K), Temperley (T), and Albrecht and Shanahan (A-S).

We have found that the key of a piece of music can be detected in a simpler manner. By calculating the signature of fifths (Definition 1), determining its main directed axis, and then rotating this axis by 30 degrees in the clockwise direction, it is possible to identify a pair of relative keys [35]. In the next step, the actual key is determined via correlation of the vector of weights assigned to individual pitch-classes found in the analyzed fragment of music (the vector K) with the major and minor key-profiles associated with the position of the rotated main directed axis (only two correlation coefficients are calculated). The recognized key will be the one for which the value of the correlation is higher.

Example 2. Finding the key of a music piece based on the signature of fifths.

Let us consider the signature of fifths shown in the left part of Figure 3. Its main directed axis indicates the G tone. In this situation, the key of the piece of music is either D-major or b-minor, which are associated with the obtained main directed axis rotated by 30 degrees in the clockwise direction. Calculating the two correlation coefficients r_{xy} , where x denotes the weighted pitch-class representation of the analyzed fragment of music and y corresponds, individually, to the D-major or b-minor Krumhansl–Kessler key-profiles, the following results are obtained: $r_{x \text{ D-major}} = 0.66$, $r_{x \text{ b-minor}} = 0.44$. The correlation coefficient calculated for the D-major key-profile is larger, hence the detected key is D-major (Figure 5).

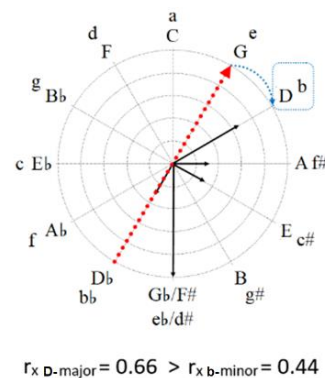


Figure 5. Finding the key of a piece of music based on the signature of fifths.

Proper selection of the analyzed fragment of music is an important element of the key recognition process. In general, the longer the sample, the better the chance that the correct key will be recognized. Obviously, we would like to be able to correctly detect keys based on the analysis of very short fragments of music. In this context, there is an important difference between key recognition approaches based strictly on the key-profiles and the method utilizing the signature of fifths. In the case of the former, there is a tendency for changing the decision (the detected key) as the length of the analyzed sample is increased, especially for very short fragments of music (the key is indicated even for a single note, treating it as a tonic). In the case of the method based on the signature of fifths, the decision is made only after the main directed axis of the signature of fifths is determined. More notes are usually needed to recognize the key, but once the decision is made, it is much less likely to undergo further changes.

Example 3. *The comparison of different key-finding approaches.*

Let us examine the results of the key recognition performed on the fragment of the composition shown in Figure 1. We will use the two ways of creating the signature of fifths, first by calculating the multiplicities of occurrences of individual pitch-classes and then by calculating the aggregate durations of the pitch-classes. The key will be evaluated for a successively increased number of initial notes. The results obtained for the two representations of the analyzed fragment are presented in Figures 6 and 7. These figures also illustrate the results obtained as the outcome of the strictly correlational analysis using the key-profiles of Krumhansl–Kessler (K-K), Temperley (T), and Albrecht–Shanahan (A-S). Incorrectly recognized keys are marked in red. It is clear that, in the case of the approach based strictly on the key-profiles, decisions were made at any sample size of the musical piece; thus, for very short numbers of notes they were not correct. However, in the case of the method implementing the signature of fifths the decision was made when the sample length reached four notes and it was correct. Expanding the sample size to five notes resulted in the detection of the correct key (D-major) for the approaches based on Temperley and Albrecht–Shanahan key-profiles. In the case of the sample comprised of six notes, the direction of the main directed axis of the signature of fifths could not be determined, making it impossible to indicate the key. Further extension of the sample size, however, resulted in confirmation of the previous decision, indicating the key of D-major. Interestingly, the strictly correlational method using Krumhansl–Kessler (K-K) profiles identified incorrect keys regardless of the number of notes that were analyzed. Besides the algorithm that utilized the signature of fifths, the best results were obtained with the strictly correlational method based on Temperley (T) key-profiles. In the case of the approach implementing the Albrecht–Shanahan (A-S) key-profiles, the correct key was indicated for six notes; however, the decision was changed after extending the sample by another two notes. The results obtained for all the considered key detection approaches, and for both ways of representing the sample of music (i.e., the weighting of the tones comprising the

analyzed sample with the multiplicities of occurrences or aggregate durations of individual pitch-classes) are shown in Figures 6 and 7.

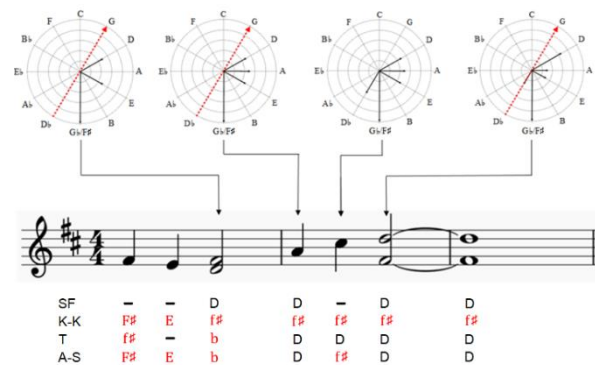


Figure 6. The results of the music key detection obtained for the method implementing the signature of fifths (SF) as well as for the strictly correlational approaches based on the music key-profiles of Krumhansl–Kessler (K-K), Temperley (T), and Albrecht–Shanahan (A-S), observed for the successively increased lengths of an analyzed music fragment—the weighting of individual pitch-classes was realized with the multiplicities of their occurrences (incorrectly detected keys are marked in red).

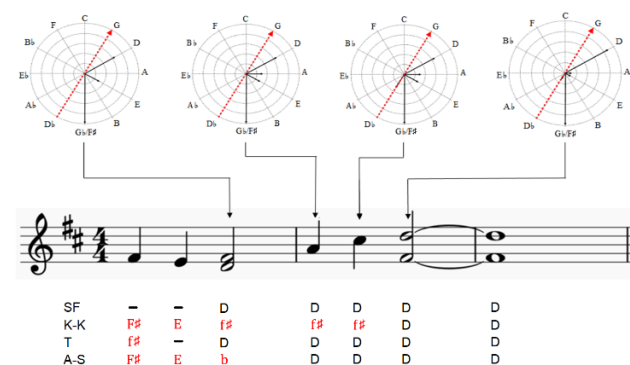


Figure 7. The results of the music key-detection obtained for the method implementing the signature of fifths (SF) as well as for the strictly correlational approaches based on the music key-profiles of Krumhansl–Kessler (K-K), Temperley (T), and Albrecht–Shanahan (A-S), observed for the successively increased length of an analyzed music fragment—the weighting of individual pitch-classes was realized with their aggregate durations (incorrectly detected keys are marked in red).

3. Experiments and Discussion

The main goal of the experiments we conducted was a multi-criteria comparison of the method to find keys based on the signature of fifths with the strictly correlational key-finding approaches that implement well-known key-profiles (i.e., Krumhansl–Kessler [36,37], Temperley [38], and Albrecht–Shanahan [33]). Of course, many more key-profiles are known and could be used in the study. The selected ones, however, seemed to provide the best reference.

The experiments were performed with the following three groups of pieces of music: the preludes by J. S. Bach from the collection *The Well-Tempered Clavier* (Book I); the preludes of F. Chopin from Op. 28; and songs by The Beatles from the album *A Hard Day's Night*. It is worth mentioning that the analyzed collections of Bach and Chopin, individually, covered all 24 keys.

It needs to be emphasized that the key detection method that implemented the signature of fifths required the computation of only two correlation coefficients for any given music sample—the correlation of the obtained weighted pitch-class representation with the Albrecht–Shanahan [28] key-profiles corresponding to the two relative keys associated with the main directed axis of the signature of fifths. In other approaches, the key was determined based on the values of the correlation coefficients, which were individually

calculated for all 24 key-profiles of Krumhansl–Kessler, Temperley, or Albrecht–Shanahan. In all cases, the largest correlation coefficient indicated the key of a piece of music.

The performed experiments involved the analysis of fragments of music extracted from the beginning (Beginning) and the end (End) of a given composition. Evaluation of the key was based on the first and last bars of a piece. The experiments were also conducted on whole (Whole) music compositions as well as on the concatenations of their first and last bars (Beginning_End). For each of the above-mentioned sample selection criteria (Beginning, End, Beginning_End, and Whole), the following weighted pitch-class representations of an analyzed fragment were used:

- the multiplicities of occurrences of individual pitch-classes (R_M);
- the aggregate durations of individual pitch-classes (R_D).

In all the experiments, the quarter-note resolution was applied.

Let us first examine the results of the conducted experiments in terms of the applied pitch-class representations. Tables 2–5 present the overall efficacies of all the considered key detection approaches, obtained for different weighted pitch-class representations in all three collections of compositions, regardless of the analyzed sample extraction method. The tables contain the results associated with the method implementing the signature of fifths as well as the approaches based on the strictly correlational analysis with popular key-profiles, i.e., Krumhansl–Kessler, Temperley, and Albrecht–Shanahan. The rows labeled with “All pieces” correspond to the results associated with the analysis of all the music pieces considered in the study (from all three collections).

Table 2. Overall * efficacies of the key detection obtained using the signature of fifths for different weighted pitch-class representations of analyzed fragments of music, where R_M —the multiplicities of occurrences of individual pitch-classes; R_D —the aggregate durations of individual pitch-classes.

Collection/Album	Efficacy for R_M	Efficacy for R_D
J. S. Bach—The Well-Tempered Clavier (Book I)	86.5%	78.1%
F. Chopin—Preludes Op. 28	87.5%	88.5%
The Beatles—A Hard Day’s Night	78.8%	82.7%
All pieces	85.2%	83.2%

* Based on the results obtained for all of the sample selection criteria (i.e., Beginning, End, Beginning_End, and Whole).

Table 3. Overall * efficacies of the key detection obtained for different weighted pitch-class representations of analyzed fragments of music using a strictly correlational approach based on Krumhansl–Kessler’s profiles (K-K), where R_M —the multiplicities of occurrences of individual pitch-classes; R_D —the aggregate durations of individual pitch-classes.

Collection/Album	Efficacy for R_M	Efficacy for R_D
J. S. Bach—The Well-Tempered Clavier (Book I)	81.3%	74%
F. Chopin—Preludes Op. 28	80.2%	79.2%
The Beatles—A Hard Day’s Night	73.1%	75%
All pieces	79.1%	76.2%

* Based on the results obtained for all of the sample selection criteria (i.e., Beginning, End, Beginning_End, and Whole).

Based on the results of the experiments, it is not possible to unequivocally indicate which type of pitch-class representation gives better results. For example, for the set of J. S. Bach’s preludes, better key-detection results were obtained using the method based on the multiplicities of occurrences of individual pitch-classes. In some cases, the differences were relatively large, for example, for the method using the signature of fifths (8.4 pp) or the strictly correlational approach with the Krumhansl–Kessler key-profiles (7.3 pp). However, for Chopin’s preludes, the opposite was observed in most cases. The exception was the strictly correlational approach based on Krumhansl–Kessler profiles, which gave slightly better results when applying the multiplicities of occurrences of individual pitch-classes. The differences observed with regard to the pitch-class representation using the aggregate

durations did not exceed 2.1 pp. Slightly larger differences (2–4 pp) were noticed for the songs by The Beatles.

Table 4. Overall * efficacies of the key detection obtained for different weighted pitch-class representations of analyzed fragments of music using a strictly correlational approach based on Temperley’s profiles (T), where R_M —the multiplicities of occurrences of individual pitch-classes; R_D —the aggregate durations of individual pitch-classes.

Collection/Album	Efficacy for R_M	Efficacy for R_D
J. S. Bach—The Well-Tempered Clavier (Book I)	82.3%	80.2%
F. Chopin—Preludes Op. 28	87.5%	89.6%
The Beatles—A Hard Day’s Night	76.9%	73.1%
All pieces	83.2%	82.4%

* Based on the results obtained for all of the sample selection criteria (i.e., Beginning, End, Beginning_End, and Whole).

Table 5. Overall * efficacies of the key detection obtained for different weighted pitch-class representations of analyzed fragments of music using strictly correlational approach based on Albrecht—Shanahan’s profiles (A-S) where R_M —the multiplicities of occurrences of individual pitch-classes; R_D —the aggregate durations of individual pitch-classes.

Collection/Album	Efficacy for R_M	Efficacy for R_D
J. S. Bach—The Well-Tempered Clavier (Book I)	88.5%	84.4%
F. Chopin—Preludes Op. 28	87.5%	88.4%
The Beatles—A Hard Day’s Night	80.8%	76.9%
All pieces	86.5%	84.4%

* Based on the results obtained for all of the sample selection criteria (i.e., Beginning, End, Beginning_End, and Whole).

Taking into account all the experiments and comparing the outcomes observed for the two types of pitch-class representation, slightly better results were noticed when using the multiplicities of occurrences of individual pitch-classes (they were not better by more than 3 pp). Figures 8–13 illustrate the key detection efficacies obtained for different collections of music, different ways of creating the pitch-class representations of the analyzed fragments of compositions, and different approaches to the extraction of the analyzed fragments. The bar plots indicate percentages of the pieces for which the key was correctly recognized. Figures 8–10 present the results obtained for the scenarios implementing sample representations based on the multiplicities of occurrences of individual pitch-classes, whereas Figures 11–13 summarize the results corresponding to scenarios with the samples represented using the aggregate durations of the pitch-classes.

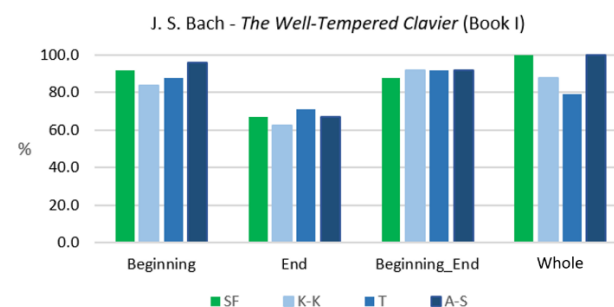


Figure 8. The key-finding efficacies obtained for the preludes of J. S. Bach, from The Well-Tempered Clavier collection (Book I); the analyzed samples represented using the multiplicities of occurrences of individual pitch-classes.

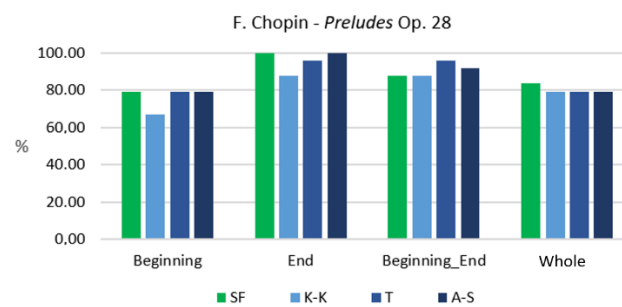


Figure 9. The key-finding efficacies obtained for Preludes Op. 28 by F. Chopin; the analyzed samples represented with the multiplicities of occurrences of individual pitch-classes.

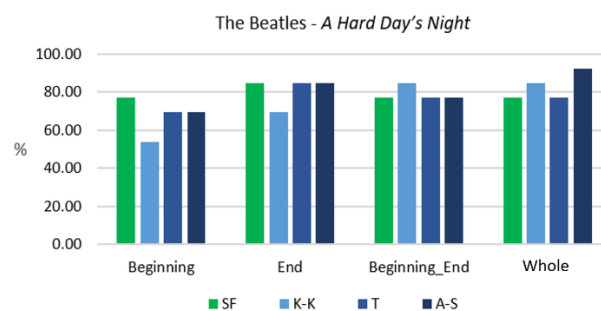


Figure 10. The key-finding efficacies obtained for the songs from the album A Hard Day's Night by The Beatles; the analyzed samples represented with the multiplicities of occurrences of individual pitch-classes.

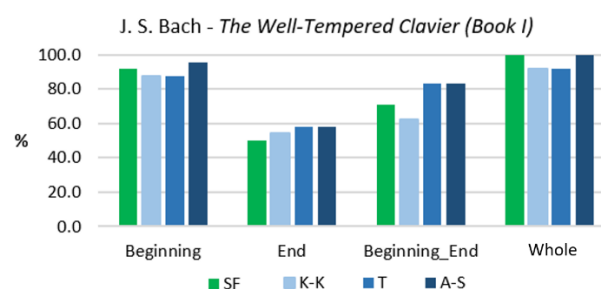


Figure 11. The key-finding efficacies obtained for the preludes of J. S. Bach, from The Well-Tempered Clavier collection (Book I); the samples represented using the aggregate durations of individual pitch-classes.

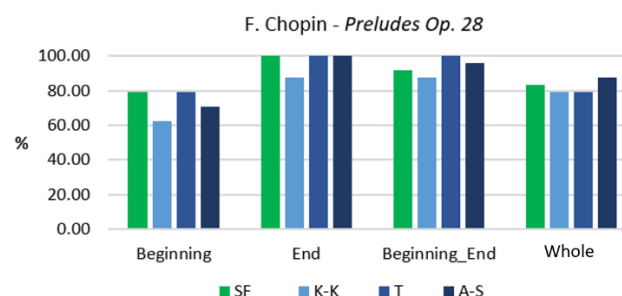


Figure 12. The key-finding efficacies obtained for the Preludes Op. 28 by F. Chopin; the samples represented with the aggregate durations of individual pitch-classes.

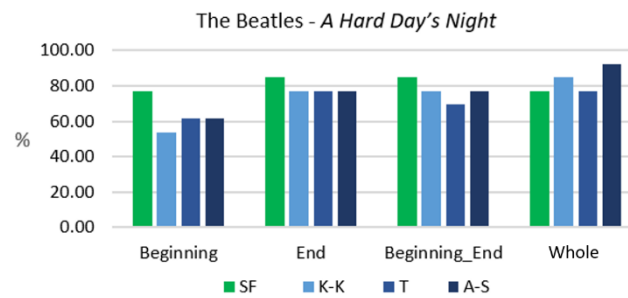


Figure 13. The key-finding efficacies obtained for the songs from the album *A Hard Day's Night* by The Beatles; the samples represented with the aggregate durations of individual pitch-classes.

Figures 8–10 show large differences in the key detection efficacies obtained for the given collections of music. The results depend on the selection of the analyzed fragments of pieces, for example, the beginning or the end of the piece. In the case of J. S. Bach's preludes, this can be explained by the fact that many of the compositions, which are written in minor keys, end in the relative major key. Such endings are found in preludes No. 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24.

When determining the key using samples obtained as the concatenation of the chords located at the beginning and the end of a composition, the major character of the ending chord is often masked by the minor sounding of the beginning and the ending chords just before the final ending in a major key. This is reflected by an efficacy of about 90%. When analyzing entire pieces of music (Whole), the efficacy obtained for the set of preludes by J. S. Bach reached 100% both using the method implementing the signature of fifths and using the strictly correlational approach based on 24 key-profiles of Albrecht–Shanahan. A much lower efficacy (79.2%) was obtained in the case of the strictly correlational method using Temperley key-profiles: incorrect keys were indicated for five preludes (Preludes No. 10, 14, 18, 20, and 24), all composed in minor keys.

In the collection of preludes by F. Chopin, the highest efficacies of key detection were observed while analyzing the ending chords of pieces. The key was detected with 100% efficacy using the method based on the signature of fifths as well as utilizing the strictly correlational approach with the Albrecht–Shanahan key-profiles. Significantly lower efficacies (<80%) were observed while analyzing the initial fragments of the preludes (Beginning). This can be explained by the uniqueness/originality of the beginning of many preludes, which mask the true key of the overall composition. Good examples of such pieces are preludes No. 2, 5, and 11. It is also worth noticing that the efficacy of the key detection performed on entire pieces was lower (reaching about 80%) than the efficacy corresponding to the analyses conducted on the beginning or the ending fragments (>87%). This was undoubtedly caused by the unambiguous setting of the ending of the preludes in the correct key but also due to the presence of pieces whose middle parts were composed in a different key than the beginning and ending ones (e.g., Prelude No. 15).

As far as the songs by The Beatles are concerned, the worst results were obtained when analyzing the initial fragments of pieces (Beginning). In other cases, the results were slightly better.

The results obtained for the sample representations using the aggregate durations of individual pitch-classes, illustrated in Figures 11–13, were in most cases very similar to the ones obtained for the multiplicities of pitch-classes, illustrated in Figures 8–10. In the case of J. S. Bach's preludes, there was a rather significant discrepancy between the key detection efficacies observed for the approaches with the multiplicities of occurrences and the aggregate durations of individual pitch-classes, when analyzing the ending chords (End) as well as the concatenation of the beginning and ending fragments of these pieces (Beginning_End). This observation can be explained by the major chord endings of many of the preludes written in minor keys (Preludes No. 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24). In these cases, the final chords were usually relatively long, hence the chances of detecting a major key were higher. Better results were obtained for the analysis of samples comprised

of the concatenated beginning and ending fragments of pieces (Beginning_End). This is the result of the temporal dominances of the minor parts of the beginning and the ending of the analyzed pieces (e.g., Preludes No. 6, 16, and 22).

One of the purposes of the experiments we conducted was to find out whether it is worth analyzing the entire piece of music in order to determine its key, or if it is enough to analyze much shorter fragments. In general, determining the key based on the entire piece should provide better results, however, the analysis of the entire piece may not always be desired or possible. We decided to compare the different key detection approaches in terms of various sample selection options (Beginning, End, Beginning_End, Whole), regardless of the type of the sample's pitch-class representation. The outcome of this comparison is shown in Figure 14, which indicates that the most accurate way of determining the key was achieved when the analysis was performed on the whole piece. The exception to this was the approach using Temperley key-profiles, in which case the best results were observed for the analysis of the concatenation of fragments located at the beginning and the end of a given composition.

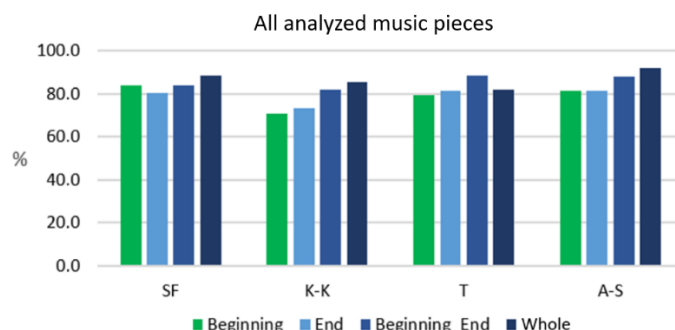


Figure 14. The key-finding efficacies obtained for various sample selection options, regardless of the type of the analyzed sample's pitch-class representation (results for all the considered music pieces).

There are scenarios when one may wish to evaluate the key in a real-time manner. In such cases, it is important to have an effective way of determining the key based on the shortest possible fragment of music. Figure 14 clearly shows that the key detection approach based on the signature of fifths gave the best results when the analysis was performed on the initial fragments of music. In order to confirm the high efficacy of the method utilizing the signature of fifths while dealing with short fragments of music, additional experiments were performed.

Figure 15 illustrates the overall key-detection efficacies corresponding to the shortest initial fragments of music for which the key was indicated in all considered approaches.

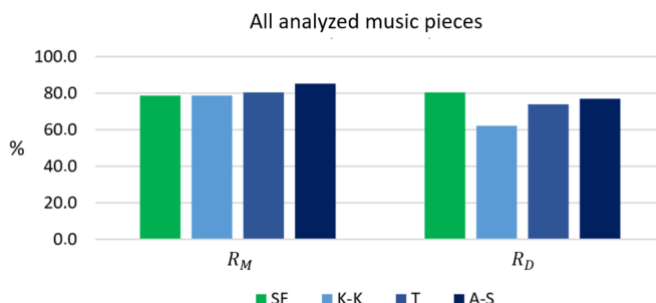


Figure 15. The overall key-finding efficacies obtained for all the considered music pieces, corresponding to the shortest initial fragments of music for which the key was indicated in all the examined approaches, where R_M —the multiplicities of occurrences of individual pitch-classes, R_D —the aggregate durations of individual pitch-classes.

Algorithms using key-profiles indicate the key after analyzing any non-empty sample fragment (with at least one occurrence of any tone). In the case of the method utilizing the

signature of fifths, the decision can be made only after determining the main directed axis. This requires at least two tones, but in most cases a larger number is needed to determine the key. Figure 15 indicates that the method employing the signature of fifths leads to very good results, especially when the analyzed samples are represented using the aggregate durations of the pitch-classes. In this particular case, the results were the best (Figure 15).

The results yield some overall conclusions regarding the specifics of particular key-finding approaches. We can confidently state the method based on the signature of fifths does not make hasty decisions; additionally, once the key is determined it is generally not prone to further changing. This is because of the way in which the main directed axis of the signature of fifths is determined. In contrast, methods based strictly on the key-profiles are more inclined towards changing their decisions. In order to make this clearer, let us inspect two preludes by F. Chopin. First, we will consider the beginning of Prelude No. 4, Op. 28, written in e-minor. The corresponding musical notation and the results of the key detection process obtained using different approaches (SF, K-K, T, A-S) are presented in Figure 16.



Figure 16. The music notation and key detection results obtained for the beginning of F. Chopin's prelude No. 4, Op. 28, written in e-minor, where: SF is the method using the Signature of Fifths and K-K, T, and A-S are the strictly correlational approaches based on the Krumhansl–Kessler, Temperley and Albrecht–Shanahan key-profiles, respectively (incorrectly detected keys are marked in red).

As we can see, the key-finding method using the signature of fifths (SF) indicated the correct key (e-minor) after analysis of six notes, as in the case of the strictly correlational approach implementing Albrecht–Shanahan key-profiles. Due to the dominance of the note B in the initial fragment, the use of the strictly correlational method based on the Krumhansl–Kessler and Temperley profiles (K-K, T) led to the detection of the b-minor key.

Detecting the key associated with a given fragment of music by relying only on the correlations of its pitch-class representation with the key-profiles is prone to less accurate decisions than detection using the signature of fifths. As an example, we will consider the beginning of F. Chopin's Prelude No. 7, Op. 28, in A-major. The musical notation corresponding to this fragment, along with the key detection results obtained when gradually increasing the number of analyzed notes, are presented in Figure 17. The key-finding method using the signature of fifths (SF) indicated the correct key of A-major after analysis of four notes and did not change the decision until the end of the fragment. In the case of the strictly correlational approaches based on the Temperley and Albrecht–Shanahan key-profiles the presence of E7 chords located in the second bar determined the incorrect key of E-major.

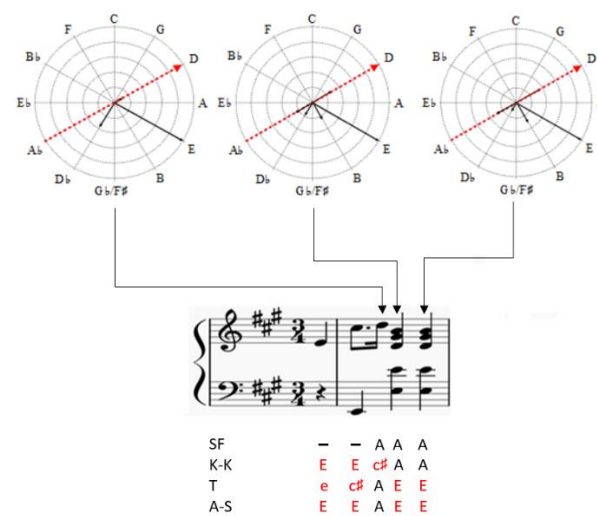


Figure 17. The beginning of F. Chopin’s prelude No. 7, Op. 28, in A-major, where SF is the key-finding method using the signature of fifths and K-K, T, and A-S are the strictly correlational key-finding approaches based on the Krumhansl–Kessler, Temperley, and Albrecht–Shanahan key-profiles, respectively (incorrectly detected keys are marked in red).

The results of the experiments discussed in this paper can be summarized in a few points:

- Generally, slightly better key detection efficacies are observed when representing the analyzed samples using the multiplicities of occurrences of individual pitch-classes than in the case of using the aggregate durations of the pitch-classes;
- In the scenario implementing the aggregate durations of individual pitch-classes, based on the analysis of the shortest initial fragments of music for which the key was indicated in all the considered approaches, the key detection efficacy obtained with the method using the signature of fifths was greater than the efficacies obtained with the strictly correlational approaches utilizing key-profiles (on average by 9.27 pp; Figure 15);
- In the case of the scenario implementing the multiplicities of occurrences of individual pitch-classes, based on the analysis of the shortest initial fragments of music for which the key was indicated in all the considered approaches, the average efficacy obtained with the strictly correlational approaches utilizing key-profiles was greater than the efficacy achieved with the method using the signature of fifths (by 2.7 pp; Figure 15);
- The key-finding method utilizing the signature of fifths is competitive with the strictly correlational approaches implementing the key-profiles of Krumhansl–Kessler, Temperley, and Albrecht–Shanahan, especially when one wants to detect the key by analyzing a very short fragment of music;
- The approach based on the signature of fifths usually needs a larger number of notes to determine the key, but once the decision is made, it is most often correct and less prone to further changing (as the number of analyzed notes increases) than the strictly correlational methods based on the key-profiles.

4. Conclusions

The multi-criteria experiments performed on different key-finding approaches allowed us to evaluate their strengths and weaknesses. It turned out that in the scenario implementing the aggregate durations of individual pitch-classes, based on the analysis of the shortest initial fragments of music for which the key was indicated in all the considered approaches, the key detection efficacy obtained with the method using the signature of fifths was greater than the efficacies obtained with the strictly correlational approaches utilizing key-profiles (on average by 9.27 pp). In the case of the analogous analysis carried out for the scenario implementing the multiplicities of occurrences of individual pitch-classes, the average efficacy obtained with the strictly correlational approaches utilizing key-profiles

was greater than the efficacy achieved with the method using the signature of fifths (by 2.7 pp; Figure 15).

In the light of the above, representing the analyzed compositions with the aggregate durations of individual pitch-classes seems more appropriate for the method utilizing the signature of fifths. On the other hand, the use of the multiplicities of individual pitch-class seems a better choice when one utilizes the other approaches.

It is evident that the method implementing the signature of fifths is a good alternative to the strictly correlational key detection approaches based on the key-profiles. There is no doubt that it is a computationally simpler technique that yields comparable results and provides greater stability of the detected keys. The conducted experiments showed that the method utilizing the signature of fifths was superior if the key detection process was performed on the initial fragments of the analyzed pieces of music. This results from the fact that the applied algorithm indicates the key only when the main directed axis of the signature of fifths is determined. In other approaches, the key is indicated based only on the values of correlation coefficients, which can be calculated for any number of notes/chords comprising a given fragment of a composition. The problem is that in the strictly correlational approaches, the keys detected based on the analysis of short fragments of music are often incorrect. Additionally, it is difficult to determine the optimal length of an analyzed fragment of music. In the case of the key detection method utilizing the signature of fifths, the number of chords/notes needed to detect the key is chosen adaptively. The key is pointed out only when there are enough notes to make a sensible decision.

The method based on the signature of fifths is well suited for real-time implementations where the analyzed fragments of music need to be very short. In particular, it is suitable for hardware applications in electronic musical instruments in which the musical notation representing the piece being played is displayed on an LCD screen.

It is worth emphasizing that, for the purposes of the discussed study, in order to visualize the signature of fifths, we used a normalization procedure so that the signature of fifths inscribed in the circle of fifths had a unitary radius. In hardware applications, though, this step is not needed because the main directed axis of the signature of fifths can be determined without it. This leads to further simplification, which significantly reduces the computational complexity of the method. Currently, we are focused on simplifying the presented technique further to eliminate the need for calculation of two correlation coefficients used for the final selection of one of two relative keys. This would make the key detection process dependent only on the analysis of the signature of fifths.

Author Contributions: Conceptualization, M.K., T.L., and D.K.; methodology, T.L. and D.K.; software, T.L. and D.K.; validation, M.K., T.L., D.K., K.M., and J.K.; formal analysis, M.K. and D.K.; investigation, M.K. and D.K.; resources, M.K. and D.K.; data curation, M.K. and D.K.; writing—original draft preparation, M.K., T.L., and D.K.; writing—review and editing, T.L., K.M., J. K., and D.K.; visualization, M.K., T.L., and D.K.; supervision, T.L. and D.K.; project administration, T.L., D.K., K.M., and J.K.; funding acquisition, D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Polish Ministry of Science and Higher Education.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anglade, A.; Benetos, E.; Mauch, M.; Dixon, S. Improving Music Genre Classification Using Automatically Induced Harmony Rules. *J. New Music Res.* **2010**, *39*, 349–361. [\[CrossRef\]](#)
2. Perez-Sanchio, C.; Rizo, D.; Inesta, J.M.; Ramirez, R. Genre classification of music by tonal harmony. *Intell. Data Anal.* **2010**, *14*, 533–545. [\[CrossRef\]](#)
3. Dai, J. Intelligent Music Style Classification System Based on K-Nearest Neighbor Algorithm and Artificial Neural Network. In Proceedings of the 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 25–27 February 2022; pp. 531–543.
4. Chapin, H.; Jantzen, K.; Kelso, J.S.; Steinberg, F.; Large, E. Dynamic Emotional and Neural Responses to Music Depend on Performance Expression and Listener Experience. *PLoS ONE* **2010**, *5*, e13812. [\[CrossRef\]](#)

5. Yang, S.; Reed, C.; Chew, E.; Barthet, M. Examining emotion perception agreement in live music performance. *IEEE Trans. Affect. Comput.* **2021**. [[CrossRef](#)]
6. Yanase, A.; Nakanishi, T. Musical impression extraction method by discovering relationships between acoustic features and impression terms. In Proceedings of the 10th International Congress on Advanced Applied Informatics (IIAI-AAI), Niigata, Japan, 11–16 July 2021; pp. 810–817.
7. Zhao, J.; Ru, G.; Yu, Y.; Wu, Y.; Li, D.; Li, W. Multimodal music emotion recognition with hierarchical cross-modal attention network. In Proceedings of the 2022 IEEE International Conference on Multimedia and Expo (ICME), Taipei, Taiwan, 18–22 July 2022.
8. Chacon, C.E.C.; Lattner, M.S.; Grachten, M. Developing tonal perception through unsupervised learning. In Proceedings of the 15th International Society for Music Information Retrieval Conference, Taipei, Taiwan, 27–31 October 2014.
9. Sapp, C. Harmonic Visualizations of Tonal Music. *Int. Comput. Music Assoc.* **2001**, *1*, 419–422.
10. Tymoczko, D. The geometry of musical chords. *Science* **2006**, *313*, 72–74. [[CrossRef](#)]
11. Milošević, I.; Bogavac, Ž.M.; Regodić, D.; Milošević, B. Visualization of Music-MIDI Content. In Proceedings of the 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), Ohrid, North Macedonia, 16–18 June 2022.
12. Huang, C.Z.A.; Duvenaud, D.; Gajos, K.Z. Chordripple: Recommending chords to help novice composers go beyond the ordinary. In Proceedings of the 21st International Conference on Intelligent User Interfaces, Sonoma, CA, USA, 7–10 March 2016; pp. 241–250.
13. Sabathé, R.; Coutinho, E.; Schuller, B. Deep recurrent music writer: Memory-enhanced variational autoencoder-based musical score composition and an objective measure. In Proceedings of the 2017 International Joint Conference on Neural Networks, Anchorage, AK, USA, 14–19 May 2017; pp. 3467–3474.
14. Tseng, B.; Shen, Y.; Chi, T. Extending music based on emotion and tonality via generative adversarial network. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021.
15. Weis, C. Global key extraction from classical music audio recording based on final chord. In Proceedings of the 10th Sound and Music Computing Conference, Stockholm, Sweden, 30 July–3 August 2013; pp. 1–6.
16. Roig, C.; Tardon, L.J.; Barbancho, I.; Barbancho, A.M. Automatic melody composition based on a probabilistic model of music style and harmonic rules. *Knowl. Based Syst.* **2014**, *71*, 419–434. [[CrossRef](#)]
17. Deng, J.; Kwok, Y.-K. Large vocabulary automatic chord estimation using deep neural nets: Design framework, system variations and limitations. *arXiv* **2017**, arXiv:1709.07153.
18. Dawson, M. *Connectionist Representations of Tonal Music: Discovering Musical Patterns by Interpreting Artificial Neural Networks*; AU Press, Athabasca University: Edmonton, AB, Canada, 2018.
19. Korzeniowski, F.; Widmer, G. End-to-end musical key estimation using a convolutional neural network. In Proceedings of the 25th European Signal Processing Conference (EUSIPCO), Kos, Greece, 28 August 2017–02 September 2017; pp. 966–970.
20. Masada, K.; Bunesco, R. Chord recognition in symbolic music using semi-markov conditional random fields. In Proceedings of the 18th International Society for Music Information Retrieval Conference, Suzhou, China, 23–27 October 2017; pp. 23–27.
21. McFee, B.; Bello, J.P. Structured training for large-vocabulary chord recognition. In Proceedings of the 18th International Conference on Music Information Retrieval (ISMIR), Suzhou, China, 23–27 October 2017; pp. 188–194.
22. Zhou, X.; Lerch, A. Chord Detection Using Deep Learning. In Proceedings of the 16th ISMIR Conference, Malaga, Spain, 26–30 October 2015.
23. Shepard, R. Geometrical approximations to the structure of musical pitch. *Psychol. Rev.* **1982**, *89*, 305–333. [[CrossRef](#)]
24. Chew, E. Towards a Mathematical Model of Tonality. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2000.
25. Chew, E. Out of the Grid and Into the Spiral: Geometric Interpretations of and Comparisons with the Spiral-Array Model. *Comput. Musicol.* **2008**, *15*, 51–72.
26. Mauch, M.; Dixon, S. Approximate note transcription for the improved identification of difficult chords. In Proceedings of the 11th International Society for Music Information Retrieval Conference, Utrecht, The Netherlands, 9–13 August 2010; pp. 135–140.
27. Osmalskyj, J.; Embrechts, J.; Piérard, S.; Van Droogenbroeck, M. Neural Networks for Musical Chords Recognition. In Proceedings of the Actes des Journées d’Informatique Musicale (JIM 2012), Mons, Belgique, 9–11 May 2012; pp. 39–46.
28. Sigia, S.; Boulanger-Lewandowski, N.; Dixon, S. Audio chord recognition with a hybrid recurrent neural network. In Proceedings of the 16th International Society for Music Information Retrieval Conference, Malaga, Spain, 26–30 October 2015; pp. 127–133.
29. Chuan, C.-H.; Chew, E. Polyphonic Audio Key Finding Using the Spiral Array CEG Algorithm. In Proceedings of the 2005 IEEE International Conference on Multimedia and Expo, Amsterdam, The Netherlands, 6 July 2005; pp. 21–24.
30. Chuan, C.-H.; Chew, E. Audio Key Finding: Considerations in System Design and Case Studies on Chopin’s 24 Preludes. *EURASIP J. Adv. Audio Signal Process.* **2007**, 1–15. [[CrossRef](#)]
31. Longuet-Higgins, H.C. Letter to a musical friend. *Music Rev.* **1962**, *23*, 244.
32. Longuet-Higgins, H.C. Second letter to a musical friend. *Music Rev.* **1962**, *23*, 280.
33. Albrecht, J.; Shanahan, D. The Use of Large Corpora to Train a New Type of Key-Finding Algorithm: An Improved Treatment of the Minor Mode. *Music Percept. Interdiscip. J.* **2013**, *31*, 59–67. [[CrossRef](#)]
34. Gomez, E.; Herrera, P. Estimating the tonality of polyphonic audio files: Cognitive versus machine learning modeling strategies. In Proceedings of the 5th International Conference on Music Information Retrieval, Barcelona, Spain, 10–14 October 2004; pp. 92–95.

35. Kania, D.; Kania, P. A key-finding algorithm based on music signature. *Arch. Acoust.* **2019**, *44*, 447–457.
36. Krumhansl, C.L.; Kessler, E.J. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol. Rev.* **1982**, *89*, 334–368. [[CrossRef](#)]
37. Krumhansl, C.L. *Cognitive Foundations of Musical Pitch*; Oxford University Press: New York, NY, USA, 1990; pp. 77–110.
38. Temperley, D. Bayesian models of musical structure and cognition. *Musicae Sci.* **2004**, *8*, 175–205. [[CrossRef](#)]
39. Temperley, D.; Marvin, E. Pitch-Class Distribution and Key Identification. *Music Percept.* **2008**, *25*, 193–212. [[CrossRef](#)]
40. Aarden, B. Dynamic Melodic Expectancy. Unpublished. Ph.D. Thesis, Ohio State University, Columbus, OH, USA, 2003.
41. Bellman, H. About the determination of key of a musical excerpt. In *Proceedings of Computer Music Modeling and Retrieval*; Springer: Pisa, Italy, 2005; pp. 187–203.
42. Chuan, C.-H.; Chew, E. The KUSC classical music dataset for audio key finding. *Int. J. Multimedia Appl.* **2014**, *6*, 1–18. [[CrossRef](#)]
43. Kania, P.; Kania, D.; Łukaszewicz, T. A hardware-oriented algorithm for real-time music key signature recognition. *Appl. Sci. Comput. Artif. Intell.* **2021**, *11*, 8753. [[CrossRef](#)]
44. Kania, D.; Kania, P.; Łukaszewicz, T. Trajectory of fifths in music data mining. *IEEE Access* **2021**, *9*, 8751–8761. [[CrossRef](#)]
45. Kania, P.; Kania, D. Sygnatura utworu w procesie reprezentacji i analizy treści utworu muzycznego. *Przegląd Elektrotechniczny* **2018**, *94*, 196–200. [[CrossRef](#)]