

Chord Recognition in Music Using a Robust Pitch Class Profile (PCP) Feature and Support Vector Machines (SVM)

Suwatchai Kamonsantiroj^{1,*}, Lita Wannatrongs², Luepol Pipanmaekaporn³

^{1,2,3}*Department of Computer and Information Science, King Mongkut's University of Technology, Thailand*

(Received: September 14, 2023; Revised: October 12, 2023; Accepted: November 18, 2023; Available online: January 7, 2024)

Abstract

Music serves as a powerful and immediate avenue for the expression of emotions, and a nuanced understanding of musical compositions is crucial for accurately interpreting and appreciating them. This research centers on the examination of robust Pitch Class Profile (PCP) features and Support Vector Machine (SVM) in the realm of music analysis. The initial phase of the study delves into the exploration of pertinent concepts and a myriad of resilient Constant-Q Transform (CQT) methods used in describing chord spectra for audio analysis. Subsequently, the paper elucidates the intrinsic correlation between SVM and speech tonality, outlining the design of a comprehensive system for music chord recognition. Rigorous testing of the system's performance follows, with a particular emphasis on evaluating the recognition rate. The results of these tests underscore the significant enhancement in music chord recognition achieved by the system, highlighting the pivotal role played by robust feature optimization and SVM pattern in bolstering its efficacy. This research not only contributes to the theoretical understanding of music analysis but also provides practical insights into improving the accuracy of music chord recognition systems through innovative feature selection and machine learning techniques.

Keywords: Robustness, PCP Feature, SVM, Music Chord

1. Introduction

In many music information retrieval applications, it is necessary to analyze the harmonic structure of music. In western music, harmonic structure is usually described by information about chord structure and chord progression [1][2]. The human brain can not only reorganize more objective things, but also deal with and understand more complex subjective things, such as music. Even so, extracting important information from music is still a hot research field for computer lovers. The development of music is inseparable from chords. In life, people usually think that music is "sound". However, with the continuous improvement of social and economic level and material conditions, as well as the gradual increase of spiritual and cultural demand, a large number of modern music creation have been produced [3][4].

Music chord recognition based on robust PCP feature and SVM has been explored and studied by many scholars. In the middle of the 20th century, the search of vocal music content mainly focused on the basic theoretical scheme to improve the signal. Note recognition was the focus of research at that time. Some scholars used chromaticity features and HMMs with EM algorithm to realize the recognition of chord recognition system. The innovation of this method is to combine music knowledge into the model by defining a state transition matrix based on the pitch distance in the pentatonic ring [5][6]. At the same time, it also avoids the random initialization of the mean vector and the covariance matrix of the observed symbol distribution. In addition, when training the model parameters, they assume that the chord distribution does not consider the type of music, that is, they do not model separately for the type of music, and then selectively update the weight of the parameters.

*Corresponding author: Suwatchai Kamonsantiroj (suwatchai.kamon@sci.kmutnb.ac.th)

DOI: <https://doi.org/10.47738/ijiis.v7i1.191>

This is an open access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>).

© Authors retain all copyrights

However, in this way, the distribution of parameters cannot be corrected [7][8]. Hidden Markov model and artificial neural network are two technologies with important theoretical and application value [9][10]. The gradual maturity of these technologies has laid a solid theoretical foundation for music chord recognition based on robust PCP feature and SVM.

This paper first introduces the research background, then briefly describes the concept of music, and classifies and distinguishes it. Then the characteristics of SVM are described in detail. Finally, the trained robust PCP feature is used to test whether the composition of music chords and the change of strength, and specific methods are given to improve its recognition rate and accuracy. At the same time, it is also verified by experiments that the conditioning numbers we have learned can be detected under different intensities and different scenarios, and these algorithms can be applied to real life for processing and application.

2. Literature Review

2.1. Characteristics and Definitions of Musical Chords

Since ancient times, music art has always been an important part of human spiritual life. With the development of society and times, people have a new understanding of music. With the continuous improvement of the level of science and technology, the acceleration of the process of economic globalization and the enhancement of the trend of cultural diversity, all these factors provide more possibilities and creative ideas for music creation. In addition, due to the scientific and technological progress and innovation, more talents emerge to the public, which makes composers need to consider a variety of complex emotional factors and personal feelings under the cultural background [11][12]. With the development of acoustic research, people began to study the methods of producing and transmitting sound, such as the research of various speakers, including pipes, musical instruments and strings. In the mid-19th century, people began to study the properties of the most advanced sounds, such as harmony. This marks a new stage in sound research. The study of the basic principles of acoustics and music has laid the foundation for today's use of computers to process audio and music signals.

As music has evolved over time, the concept of chords has undergone a transformative journey, progressing from the simplicity of single tones to the complexity of polyphony. This paper delves into a nuanced exploration of chords, focusing on the intricate and abstract relationships forged between the musician and the instrument. Two fundamental types of chords are identified: those formed through finger placement on the instrument and those created through the production of sound (hereinafter referred to as sight play). The latter category encompasses three distinct types of sight play, each contributing to the multifaceted nature of chords [13][14]. These include the harmonious interplay within the bass area, the melodic richness of the treble area, and the intricate resonance found in high octave chords. By dissecting the components of these chords, this research aims to unravel the intricate tapestry that defines the musical experience, shedding light on the dynamic interplay between the player and the music they create.

2.2. Introduction to Proper Terms of Music Specialty

Pitch: some music sounds high and some sounds low. This is called pitch. Pitch is expressed in vibrations per second. The higher the frequency, the higher the tone, the lower the tone, the lower the frequency, the lower the tone [15][16]. This sound that vibrates 440 times per second is called "a", which is the current international standard sound.

Interval: Specifies the distance between two notes. The unit used to calculate the interval is called "degree", and the number of syllables between two notes is called several degrees.

Octave: if the frequencies of a group of sounds are arranged strictly according to $x_1, X_2, x_4, \dots$, that is, according to the rule of 2^n , they sound like "the same pitch sequence". Because the human ear is sensitive to the frequency index, the above-mentioned " x_2 represents isometric" relationship is the most basic relationship in music. In music, X_2 is an octave.

Consistency: it is composed of three or more scales according to three degree interval relationship or non three degree interval relationship.

Harmony: harmony is a chord progression in which each chord is connected according to certain rules. Harmony is based on the mode that makes the melody more richly express the music content.

Semitone and whole tone: divide an octave into twelve equal parts, one semitone for each, and two semitones correspond to a whole tone. Half tones correspond to small seconds and whole tones correspond to large seconds.

2.3. Principle of Music Recognition

In modern society, with the continuous development of science and technology and economic level, people's demand for material life is also higher and higher. This phenomenon makes people have higher requirements and pay more attention to music art to a new point, that is, rhythm and beat, so that it can meet the needs of popularization and personalization. Music feature detection for audio files, analyze and identify many characteristic elements describing music: pitch, duration, intensity, beat, color, melody, melody, rhythm, music style characteristics, etc. Music is one of the most important parts of human spiritual world. Frequency file refers to the direct recording of the data obtained by binary sampling of the waveform of simulated real sound, that is, the reflection of real sound. In this way, the storage space of the sound file generated by storing the sound information is relatively large [17][18]. Module files (mod, S3M, XM, MTM, far, Kar, it, etc.) have the same attributes as MIDI files and sound files, which means that module files contain not only instructions on how to play musical instruments, but also sample data of beep. However, the difference is that module files have many different formats, depending on how they are encoded. In order to extract the characteristic parameters of melody, different coding formats must be processed accordingly.

2.4. Role of Identifying Chord Features

Due to the growing demand for content-based search with intermediate functions, the search engine based on text keywords can no longer meet the needs of users. If computers can more accurately identify and transcribe chords in music, tags can be used to match the chord process of text blur, downloading and playing songs.

Understanding chords can make people analyze the structure of music more meaningful and appropriate. Chords often have a certain length, so music can be segmented and spliced according to the results of chord recognition [19][20]. On the one hand, it can deepen the understanding and analysis of complex music structure, on the other hand, it can create more complex music emotional expression forms and enrich music expression forms.

2.5. SVM multi-classification algorithm

SVM is a binary model, which is a linear classifier with the largest interval defined in feature space. The goal is to find the maximum interval. The kernel method of support vector machine is the main embodiment of its superiority. It maps linear and indivisible data into high-dimensional feature space through nuclear technology, so as to classify them in high-dimensional space. The following is a brief introduction to its principle [21].

For a dichotomic problem, the training data set on the feature space is written in the form: where, $x \in x=R^n$, $y_i \in Y=[-1,+1]$, $i=1,2,\dots,N$, that is, the total number of samples is N . The characteristic parameters of each sample data are column vectors, and the vector is n -dimensional. The goal of learning is to find a separate hyperplane in the feature space, which separates instances of different types on both sides. The formula of the hyperplane is formula (1) :

$$R \times x + a = 0 \quad (1)$$

In general, if the data set can be classified, the number of corresponding hyperplanes is countless. The purpose of support vector machine is to find the hyperplane with the largest interval from the nearest instance, which can be expressed as a function. Formula (2) represents the distance between the i th sample and the separation hyperplane:

$$y = \min (y_i) \quad (2)$$

In the realm of machine learning and support vector machines (SVMs), the significance of identifying support vectors cannot be overstated. The preceding formulas ascertain the equality criteria that designate a particular sample vector as X , thereby classifying it as a support vector. These support vectors play a pivotal role in the SVM algorithm, serving as the foundation for constructing the decision boundary that optimally separates different classes within the data.

It is imperative to acknowledge, however, that real-world datasets often exhibit nuances that challenge the strict adherence to linear separability. In practice, data points may not conform perfectly to a linear boundary, and as a consequence, the previously outlined formula lacks a direct solution. This departure from strict linear separability underscores the need for advanced techniques and modifications to SVMs, such as kernel methods, which enable the algorithm to handle more complex patterns and achieve improved classification performance in scenarios where linearity is not strictly maintained. As researchers continue to delve into the intricacies of SVMs, their adaptability to diverse datasets becomes a focal point, with the overarching goal of enhancing the algorithm's robustness and applicability to real-world problems.

3. Research Methodology

3.1 System design

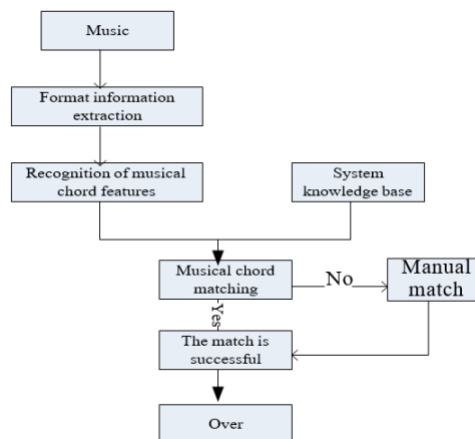


Figure 1. Music and Spin recognition system processes

Figure 1 is the flow chart of music and rotation recognition system designed according to the robust PCP feature and the application of SVM. As can be seen from the figure, when the music sounds, the system will record the music audio, then extract the music format information, and identify the music features after the extraction. At the same time, the system will match and compare the extracted music format information from the system music database. If the matching is successful, the matching result will be directly displayed and the system operation will be ended. However, if the automatic matching fails, the matching can be performed manually, and then the matching will be performed again until the matching is successful and the operation will be ended.

3.2. Music and Spin Recognition Needs

Chords can be divided into sound level and silent level. The classification standard of sound level is to distinguish according to the pitch. Therefore, the recognition of music features can also be divided into the following levels: on the one hand, extract the basic features of chords, on the other hand, analyze the complex features on this basis, and finally analyze the music. Based on the basic and complex characteristics, style and emotional connotation of music, note is the smallest basic unit of music. Each piece of music is a sequence of many different notes on a time axis. The basic properties of music are related to the information of notes, which can be obtained directly from music.

With the development of music, the sequence of notes plays a more and more important role in people's life. As the most active, intuitive and representative of rhythm and melody, beat stress. THM and melody represent the change law of tonal order on the time axis, and the description of music rhythm and melody is clearer. The structure, subdivision and emotional connotation of music are the general characteristics of music. They represent the most complete description of music.

Among all recognition categories, basic function recognition is the simplest and can be extracted directly from music format files. Complex characters must be analyzed through the recognized basic features to analyze the order of notes. They are recognizable but relatively complex. In this paper, the basic features, complex features and overall features of music structure are preliminarily identified by combining the robustness features and SVM model. This paper will

directly extract the basic attributes of music from the music format, then analyze the extracted note sequence to obtain the attributes of melody and rhythm, and then compare and analyze the music structure according to the attributes of beat and phrase.

4. Result and Discussion

4.1. System Recognition Rate Analysis

In this study, Table 1 serves as a comprehensive repository of data, presenting a meticulous categorization of various music genres alongside their corresponding spin recognition outcomes. The table encapsulates the essence of our investigation, acting as a visual representation of the intricate relationship between different types of music and their influence on spin recognition within the context of our experimental setup. Each entry in the table is a distinct point of reference, highlighting the nuances and distinctions among the diverse musical genres employed in the test. This tabulated data becomes a crucial tool for analysis, allowing for a systematic exploration of patterns, correlations, and trends in spin recognition across the musical spectrum. As we delve into the intricacies of the information encapsulated in Table 1, our research aims to unravel the underlying factors that contribute to the varying degrees of spin recognition associated with different genres of music, ultimately shedding light on the intricate interplay between auditory stimuli and the cognitive processes involved in spin perception.

Table 1. Identification data

Test music	Characteristic	Recognition mode	Discrimination (%)
Trouble I'm in	PCP	SVM	98
The royal wedding song	PCP	SVM	95
Ghost	PCP	SVM	91
Good to be live	PCP	SVM	93

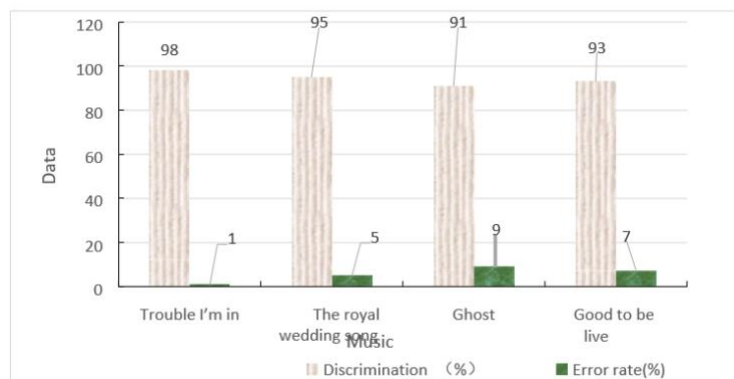


Figure 2. Comparison diagram of the music and spin recognition ratio

In conjunction with the insights gleaned from Table 1 and Figure 2, the comprehensive analysis reveals that the accuracy levels of all songs post-system recognition surpass an impressive 91%. This noteworthy performance underscores the efficacy of the recognition system in accurately identifying and classifying diverse musical compositions. Notably, the evaluation metrics presented in Table 1, coupled with the visual representation in Figure 2, provide a comprehensive overview of the system's robust performance across various songs.

It is pertinent to note that the conventional support vector machine (SVM) kernel function, employed for distance measurement, typically relies on the Euclidean distance metric. While this approach is recognized for its simplicity and ease of implementation, its applicability to specific datasets may encounter limitations. The intricacies of certain datasets may render the traditional SVM kernel function less effective in achieving the desired outcomes. This observation emphasizes the need for further investigation into alternative methods or enhancements to the existing system to address potential challenges posed by specific datasets. As we delve deeper into the intricacies of music recognition technologies, it becomes imperative to explore innovative approaches that can elevate the system's performance and adaptability to diverse and complex datasets.

5. Conclusion

In the realm of music, the identification and classification of entire musical compositions stand as pivotal endeavors. The expeditious and accurate recognition of a song is crucial for employing effective training methods that enable us to interpret and convey the emotional nuances encapsulated in its lyrics with precision. Consequently, harnessing the power of the robust Pitch Class Profile (PCP) feature becomes imperative. Through a meticulous exploration of whether the chord structure is meticulously composed and an adept distinction of the variations in strength, the PCP feature serves as a potent classifier. Its application allows for a comprehensive study and comprehension of the emotional characteristics inherent in a composer's work.

The utilization of PCP as a classifier not only facilitates the swift and accurate interpretation of a song's emotional underpinnings but also holds broader implications. It contributes valuable insights to the theoretical frameworks within the musical domain, thereby enhancing our understanding of relevant concepts. Moreover, this approach has the potential to elevate our comprehension of the integrity and precision of musical compositions. By shedding light on the intricate details of chord composition and strength variations, the study aids in refining and expanding our knowledge of music, offering a significant contribution to the broader discourse in the field.

6. Declarations

6.1. Author Contributions

Conceptualization: S.K., L.W., and L.P.; Methodology: L.P.; Software: S.K.; Validation: S.K., L.W., and L.P.; Formal Analysis: S.K., L.W., and L.P.; Investigation: S.K.; Resources: L.P.; Data Curation: L.P.; Writing Original Draft Preparation: S.K., L.W., and S.K.; Writing Review and Editing: L.P. and S.K.; Visualization: S.K.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

References

- [1] X. Liu and Y. Zhang, "Gesture recognition based on multi feature and SVM classification," *Computer Engineering and Design*, vol. 038, no. 004, pp. 953-958, 2017.
- [2] Z. Long, X. Zhou, X. Zhang, et al., "Recognition and Classification of Wire Bonding Joint via Image Feature and SVM Model," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 2019, no. 1, pp. 1-6, 2019.
- [3] S. Aziz, E. A. Mohamed, and F. Youssef, "Traffic Sign Recognition Based On Multi-feature Fusion and ELM Classifier," *Procedia Computer Science*, vol. 127, no. 1, pp. 146-153, 2018.
- [4] X. Sui, K. Wan, and Z. Yang, "Pattern recognition of SEMG based on wavelet packet transform and improved SVM," *Optik*, vol. 176, no. 1, pp. 228-235, 2019.

-
- [5] W. Zhang, D. Zhao, Z. Chai, et al., "Deep learning and SVM-based emotion recognition from Chinese speech for smart affective services," *Software Practice and Experience*, vol. 47, no. 8, pp. 1127-1138, 2017.
 - [6] Q. Song, X. Chen, P. Han, et al., "Traffic signboard detection and recognition based on feature matching algorithm," *Computer Knowledge and Technology*, vol. 014, no. 011, pp. 192-195, 2018.
 - [7] A. U. Haq, J. Li, M. H. Memon, et al., "Feature Selection Based on L1-Norm Support Vector Machine and Effective Recognition System for Parkinson's Disease Using Voice Recordings," *IEEE Access*, vol. 2019, no. 1, pp. 37718-37734, 2019.
 - [8] O. S. Faragallah, "Robust noise MKMFCC-SVM automatic speaker identification," *International Journal of Speech Technology*, vol. 21, no. 2, pp. 1-8, 2018.
 - [9] L. Zhang, H. Ma, N. Jiang, et al., "Recognition and location of transformer winding loss based on singular value spectrum entropy and SVM," *Power System Protection and Control*, vol. 045, no. 018, pp. 69-75, 2017.
 - [10] M. A. Khan, M. Sharif, M. Y. Javed, et al., "License Number Plate Recognition System using Entropy based Features Selection Approach with SVM," *IET Image Processing*, vol. 12, no. 2, pp. 200-209, 2018.
 - [11] Y. Chen, W. Xu, J. Zuo, et al., "The fire recognition algorithm using dynamic feature fusion and IV-SVM classifier," *Cluster Computing*, vol. 2018, no. 10, pp. 1-11, 2018.
 - [12] S. Ma, B. Xia, W. Yang, et al., "Leap motion gesture recognition based on SVM," *Modern Computer (Professional Edition)*, vol. 000, no. 023, pp. 55-58, 2017.
 - [13] D.-T. Nguyen and K.-H. Le, "The robust scheme for intrusion detection system in internet of things," *Internet of Things*, vol. 24, no. 1, pp. 100999-101009, 2023. doi:10.1016/j.iot.2023.100999
 - [14] L. Ran and Y. J. Wu, "Development of Computer Intelligent Control System Based on Modbus and WEB Technology," *J. Appl. Data Sci.*, vol. 4, no. 1, pp. 15-21, Jan. 2023,
 - [15] M. LI, X. YU, D. MENG, G. AN, and B. LIU, "Development of a deviation package method for low-cost robust optimization in compressor blade design," *Chinese Journal of Aeronautics*, vol. 1, no. 1, pp. 1-7, 2023. doi:10.1016/j.cja.2023.12.021
 - [16] Y. J. Lin, "An Exploration Into Trust and Privacy Management in a Digital Age", *Int. J. Appl. Inf. Manag.*, vol. 2, no. 1, pp. 34-43, Sep. 2021.
 - [17] L. Stanca, "Robust bayesian choice," *Mathematical Social Sciences*, vol. 126, no. 1, pp. 94-106, 2023. doi:10.1016/j.mathsocsci.2023.10.002
 - [18] R. Widayanti, M. Heru, R. Chakim, C. Lukita, U. Rahardja, and N. Lutfiani, "Improving Recommender Systems using Hybrid Techniques of Collaborative Filtering and Content-Based Filtering," *J. Appl. Data Sci.*, vol. 4, no. 3, pp. 289-302, Sep. 2023
 - [19] D. Bertsimas and T. Koukouvinos, "Robust linear algebra," *European Journal of Operational Research*, vol. 1, no. 1, pp. 13-20, 2023. doi:10.1016/j.ejor.2023.11.036
 - [20] K. Fujishima, "Knowledge-Driven Automated Service Composition as a Method for Developing Decision Support Systems", *Int. J. Appl. Inf. Manag.*, vol. 2, no. 1, pp. 44-49, Sep. 2021.
 - [21] X. Fang et al., "On robust estimation of the gauss-markov model with a singular covariance matrix," *Measurement*, vol. 223, no. 1, pp. 113834-113840, 2023. doi:10.1016/j.measurement.2023.113834