IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CONVERSION OF MUSIC INTO CHORDS USING MACHINE LEARNING

Santosh Chobe[1*], Neeraj Saini[1], Prasanna Jagtap[1], Ashutosh Kakde[1], Yuvraj Kokare[1]

[1] Department of Computer Engineering, Pimpri Chinchwad College of Engineering and Research, Ravet, Pune

Abstract- Music chord recognition is an essential task in music analysis and understanding, and machine learning has been shown to be an effective approach for accomplishing this task. The survey reviews different approaches to chord recognition, including supervised and unsupervised learning, deep learning, and feature-based methods. It explores the strengths and weaknesses of each approach and compares their performances in various datasets. Additionally, the survey discusses the challenges and future directions for research in this area, including improving accuracy and robustness of the models, developing new evaluation metrics, and exploring new applications of chord recognition in music processing and analysis. Overall, the survey provides a comprehensive overview of the state-of-the-art in chord recognition using machine learning, offering a valuable resource for researchers and practitioners working in this field.

Keywords: Chord, Conversion, Machine Learning, Music.

1. INTRODUCTION

Music is a form of expression that has been an integral part of human culture for thousands of years. One of the essential elements of music is chords, which are groups of notes played together to create harmony. Recognizing chords in music is a critical task in music analysis, understanding, and creation. However, manually transcribing chords from audio recordings can be a time-consuming and challenging task, especially for long and complex pieces of music. In recent years, machine learning has emerged as a powerful approach to automate the process of chord recognition. Machine learning algorithms can learn from large datasets of annotated musical recordings to recognize chords accurately and efficiently.

The goal of this literature survey is to provide a comprehensive overview of various techniques for converting music into chords using machine learning. The survey will review different approaches to chord recognition, including supervised and unsupervised learning, deep learning, and feature-based methods. It will explore the strengths and weaknesses of each approach and compare their performances in various datasets. Additionally, the survey will discuss the challenges and future directions for research in this area, including improving accuracy and robustness of the models, developing new evaluation metrics, and exploring new applications of chord recognition in music processing and analysis.

Overall, this literature survey aims to provide a valuable resource for researchers and practitioners working in the field of music analysis and machine learning, helping them to understand the state-of the-art in chord recognition and identify opportunities for future research.

2. LITERATURE REVIEW

The literature survey shows the study of various machine learning algorithms

[1]Taixiang Li presents a novel approach for recognizing musical chords from audio signals using a combination of convolutional neural networks (CNNs) and hidden Markov models (HMMs). The proposed method first extracts spectrograms from audio signals, which are then fed into a CNN to learn high-level features that capture chord characteristics. The CNN output is then used to train an HMM, which models the temporal relationship between chord labels and provides a more accurate chord recognition result.

To evaluate the proposed method, the authors conducted experiments on a dataset of 2,216 audio clips of various musical genres. The results show that the CNN-HMM approach outperforms several state-of-the-art methods in terms of chord recognition accuracy, achieving a mean accuracy of 72.56%. The authors also conducted a sensitivity analysis to investigate the impact of different hyper parameters on the performance of the proposed method.

[2]J. Osmalskyj, et al. investigates the use of artificial neural networks (ANNs) for recognizing musical chords from audio signals. The authors propose a novel method that involves using a combination of feed forward and recurrent neural networks to model the temporal relationships between chord labels in music. The feed forward neural network extracts high-level features from the audio signals, which are then fed into the recurrent neural network to capture the temporal dependencies between chord labels.

To evaluate the proposed method, the authors conducted experiments on a dataset of 350 audio clips of various musical genres. The results show that the proposed method outperforms several state-of-the-art methods in terms of chord recognition accuracy, achieving a mean accuracy of 86.3%. The authors also conducted a sensitivity analysis to investigate the impact of different hyper parameters on the performance of the proposed method.

[3]Heng-Tze Cheng, et al. proposed a method for automatic chord recognition from audio signals, which can be used for music classification and retrieval. The proposed method involves extracting chroma features from the audio signals, which represent the distribution of energy across the 12 semitones in an octave. The chroma features are then used to train a hidden Markov model (HMM) for chord recognition.

To evaluate the proposed method, the authors conducted experiments on a dataset of 700 audio clips of various musical genres. The results show that the proposed method outperforms several state-of-the-art methods in terms of chord recognition accuracy, achieving a mean accuracy of 75.8%. The authors also conducted experiments to investigate the impact of different HMM topologies on the performance of the proposed method.

Furthermore, the authors demonstrated the effectiveness of the proposed method for music classification and retrieval. They used the chord recognition results to construct chord histograms for each audio clip, which can be used as a feature vector for music classification and retrieval. The authors conducted experiments on two music datasets and showed that the proposed method achieved better classification and retrieval performance compared to several baseline methods.

[4]Fernandes Saputra, et al. proposed a method for automatic piano sheet music transcription from audio signals, which involves using machine learning techniques. The proposed method involves preprocessing the audio signals to remove noise and other unwanted sounds, and then extracting relevant features such as pitch, duration, and intensity. These features are then used to train a machine learning model for transcription.

To evaluate the proposed method, the authors conducted experiments on a dataset of 20 audio recordings of piano performances. The results show that the proposed method achieved an average accuracy of 90.4% in transcribing the piano sheet music, which outperformed several baseline methods.

Furthermore, the authors demonstrated the effectiveness of the proposed method by applying it to transcribe a popular song, "Fur Elise" by Beethoven, and comparing the transcribed sheet music with the original sheet music. The results show that the proposed method achieved a high level of accuracy in transcribing the piano sheet music.

[5]Jonathan Sleep proposed a method for automatic music transcription using convolutional neural networks (CNNs) with intuitive filter shapes. The proposed method involves converting audio signals into spectrograms, which are then fed into a CNN for transcription. The CNN uses intuitive filter shapes, which are designed to capture the relevant features of the spectrograms, such as the pitch, duration, and intensity of the notes.

To evaluate the proposed method, the author conducted experiments on a dataset of 50 audio recordings of piano performances. The results show that the proposed method achieved a high level of accuracy in transcribing the piano sheet music, outperforming several state-of-the-art methods in terms of both note-wise and frame-wise F1 score.

Furthermore, the author demonstrated the effectiveness of the proposed method by applying it to transcribe a popular song, "Für Elise" by Beethoven, and comparing the transcribed sheet music with the original sheet music. The results show that the proposed method achieved a high level of accuracy in transcribing the piano sheet music.

[6]Leon Tran proposed a method for visual guitar chord classification using transfer learning with convolutional neural networks (CNNs). The proposed method involves using pre-trained CNN models, such as VGG16 and ResNet50, and fine-tuning them on a dataset of guitar chord images. The fine-tuned models are then used for chord classification.

To evaluate the proposed method, the authors collected a dataset of 15,600 guitar chord images, consisting of 33 chord types. The dataset was split into training, validation, and test sets. The results show that the proposed method achieved a high level of accuracy in chord classification, with an average accuracy of 98.3% using VGG16 and 98.9% using ResNet50 on the test set.

Furthermore, the authors demonstrated the effectiveness of the proposed method by applying it to classify guitar chords in a real-time video stream. The results show that the proposed method achieved a high level of accuracy in chord classification, even with variations in lighting and camera angles.

[7]Darrell A. , et al. Talavera presents a method for automatic transcription of guitar chords from acoustic audio signals. This method involves extracting audio features from the input signal using a short-time Fourier transform (STFT) and a modified Mel-frequency cepstral coefficients (MFCC) approach. The extracted features are then fed into a classifier,

which is trained on a dataset of guitar chord audio recordings.

To evaluate the proposed method, the authors collected a dataset of 360 audio recordings of guitar chords played by different guitarists. The results show that the proposed method achieved a high level of accuracy in chord transcription, with an average accuracy of 93.06% across all chords.

Furthermore, the authors evaluated the proposed method on real-world audio recordings and found that it was able to accurately transcribe chords even in the presence of background noise and other audio artifacts.

[8]Tristan Carsault, et al. proposed a method for real-time extraction and prediction of musical chord progressions using machine learning techniques.

The proposed method involves extracting features from an input audio signal in real-time using a sliding window approach and a chroma-based representation of the audio signal. The extracted features are then used to train a Hidden Markov Model (HMM) to predict chord progressions.

To evaluate the proposed method, the authors conducted experiments on a dataset of 100 MIDI files, consisting of various musical styles and genres. The results show that the proposed method achieved a high level of accuracy in predicting chord progressions, with an average accuracy of 86.8% on the test set.

Furthermore, the authors demonstrated the potential of the proposed method for creative applications by using it to generate chord progressions in real-time and incorporating them into a music composition system. The results show that the proposed method can produce interesting and creative chord progressions that can be used as the basis for musical compositions.

[9]Mu-Heng Yang, et al. proposed a method for automatic melody-to-chord transcription using machine learning techniques.

This method involved using a paired model and multi-task learning language modeling approach. The paired model consists of a melody model and a chord model, which are trained jointly to predict chords from melodies. The melody model is trained to predict the next note in a melody sequence, while the chord model is trained to predict the next chord in a chord sequence. The multi-task learning language modeling approach involved sharing the same neural network architecture and embedding layers for both models.

To evaluate the proposed method, the authors used a dataset of 200 songs with melody and chord annotations. The results show that the proposed method achieved a high level of accuracy in chord transcription, with an average accuracy of 79.69% on the test set.

Furthermore, the authors compared the proposed method with several existing methods and found that it outperformed them in terms of accuracy and robustness

[10] Yoonchang Han, et al. proposed a method for predominant instrument recognition in polyphonic music using deep convolutional neural networks (CNNs).

The proposed method involved using a CNN architecture that takes in a spectrogram of a polyphonic audio signal as input and outputs a probability distribution over a set of instrument classes. The CNN architecture consists of multiple layers of convolutional, pooling, and fully connected layers, and is trained using a large dataset of labeled audio recordings.

To evaluate the proposed method, the authors conducted experiments on two datasets: the RWC pop music dataset and the MedleyDB dataset. The results show that the proposed method achieved a high level of accuracy in predominant instrument recognition, with an average accuracy of 87.2% on the RWC dataset and 85.9% on the MedleyDB dataset.

Furthermore, the authors conducted a comparative study with several existing methods and found that the proposed method outperformed them in terms of accuracy and robustness.

3. DATASET

The database contains 2000 chords split up in 10 classes, giving up to 200 chords per chord type. The files are stored in raw WAV 16 bits mono 44100Hz format.

4. RESEARCH GAP

There is currently a lack of standardization in the field of music analysis, with different researchers using different approached to analyse and annotate music. This can make it difficult to compare results across studies and may limit the effectiveness of machine learning algorithms for chord recognition.

While many studies have developed machine learning models for chord recognition, there is often limited evaluation of these models on real-world music datasets. This can make it difficult to determine how well these models perform in practice and may limit their usefulness for real-world applications.

Many machine learning models used for chord recognition are highly complex and may be difficult to interpret or explain. This can make it difficult for researchers and musicians to understand how these models are making chord predictions, which may limit their adoption in practice.

5. CONCLUSION

This literature survey provides a comprehensive overview of the state-of-the-art in chord recognition using machine learning, offering a valuable resource for researchers and practitioners working in this field. The survey provided insights into the strengths and weaknesses of different approaches, identifies research challenges and opportunities, and suggests potential directions for future work. The findings of this survey can be used to guide further research in the area of chord recognition and contribute to the development of more accurate and robust models for music processing and analysis.

6. REFERENCES

[1] Study on a CNN-HMM Approach for Audio-Based Musical Chord Recognition by Taixiang Li

[2] Neural Networks for Musical Chords Recognition by J. Osmalskyj, J-J. Embrechts, S. Piérard, M. Van Droogenbroeck

[3] Automatic Chord Recognition for Music Classification And retrieval by Heng-Tze Cheng, Yi-Hsuan Yang, Yu-Ching Lin, I-Bin Liao* , and Homer H. Chen

[4] Automatic Piano Sheet Music Transcription with Machine Learning by Fernandes Saputra, Un Greffin Namyu, Vincent, Derwin Suhartono and Aryo Pradipta Gema

[5] Automatic Music Transcription with Convolutional Neural Networks using Intuitive Filter Shapes by Jonathan Sleep

[6] CNN Transfer Learning for Visual Guitar Chord Classification by Leon Tran Shawn Zhang Eric Zhou

[7] Transcription of Guitar Chords from Acoustic Audio by Darrell A. Talavera, Erika Svetlana C. Nase, Leonard D. Pancho, and Adomar L. Ilao

[8] Combining Real-Time Extraction and Prediction of Musical Chord Progressions for Creative Applications by Tristan Carsault *, Jérôme Nika *, Philippe Esling * and Gérard Assayag

[9] Melody-to-Chord using Paired Model and Multi-task Learning Language Modelling by Mu-Heng Yang , Wei-Ting Hsu , Nicholas Huang

[10] Deep Convolutional Neural Networks for Predominant Instrument Recognition in Polyphonic Music by Yoonchang Han, Jaehun Kim, and Kyogu Lee