

January 2013

A Study : Analysis of Music Features for Musical Instrument Recognition and Music Similarity Search

J. Sirisha Devi

Dept of CSE, CMRCET, JNTU (H), Hyderabad, 501401, Andhra Pradesh, India, jsirisha.devi@gmail.com

Follow this and additional works at: <https://www.interscience.in/ijcsi>



Part of the [Computer Engineering Commons](#), [Information Security Commons](#), and the [Systems and Communications Commons](#)

Recommended Citation

Devi, J. Sirisha (2013) "A Study : Analysis of Music Features for Musical Instrument Recognition and Music Similarity Search," *International Journal of Computer Science and Informatics*: Vol. 2 : Iss. 3 , Article 6.

DOI: 10.47893/IJCSI.2013.1088

Available at: <https://www.interscience.in/ijcsi/vol2/iss3/6>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in International Journal of Computer Science and Informatics by an authorized editor of Interscience Research Network. For more information, please contact sritampatnaik@gmail.com.

A Study : Analysis of Music Features for Musical Instrument Recognition and Music Similarity Search

J. Sirisha Devi¹, Y. Srinivas² & N. Murali Krishna³

¹Dept of CSE, CMRCET, JNTU (H), Hyderabad, 501401, Andhra Pradesh, India

²Dept of IT, GITAM University, Visakhapatnam, 530045, Andhra Pradesh, India

³Dept of CSE, GITAM University, Visakhapatnam, 530045, Andhra Pradesh, India

Abstract - Lots of work has been done on speech and speaker recognition. Many technologies were developed for the analysis of speech waveforms. Musical instrument recognition is an important aspect of music information retrieval system. In this paper we analyzed features for musical instruments recognition and a brief study on music similarity. Music similarity search is done by using mel-frequency cepstral coefficients (MFCC), linear predictive coefficients (LPC) and the classifier used is K- Nearest Neighbor with Dynamic Time Warping.

Keywords - *mel-frequency cepstral coefficients (MFCC), Linear predictive coefficients (LPC), timbre analysis, timbre features.*

I. INTRODUCTION

Musical signal analysis has not been able to attain as much as commercial interest as, for instance, speaker and speech recognition. This is because the topic around speech processing is more readily commercially applicable although both areas are considered as being highly complicated. Through constructing computer systems that “listen”, we may also gain some new insights into human perception.

A central topic in our paper is the analysis of musical instrument features for their recognition and a study on music similarity. In section II we describe the basics of speech processing and the main features that distinguish instruments. In section III we discuss regarding timbre, which is one of the main features for musical instrument recognition or recognition. In section IV we present different features of instruments in a tabular form.

II. MUSICAL INSTRUMENT SIGNAL PROCESSING BASICS

Sounds can be generally described [1] via four basic but important parameters – duration, pitch, amplitude, and timbre. All of these basic parameters are crucial in the domain of hearing sounds, whether be it sounds that one experiences in nature, music that blares through loud speakers in concert halls. Main features that distinguish musical instrument are pitch, harmonics,

timbre (amplitude envelope, spectrum envelope, attack time, decay time, sustain time) and loudness.

Music processing can be divided into two main branches: music signal modeling and music content creation. The music signal modeling approach is based on a well- developed body of signal processing techniques using signal analysis and synthesis tools such as filter banks, Fourier transforms, cosine transform, harmonic plus noise model, wavelet transform, linear prediction models, decision- tree clustering, Bayesian inference and perceptual models of hearing.

Music creation has a different set of objectives concerned with the methods of composition of music content.

III. TIMBRE

Sound quality or timbre or tone color, is oddly defined in terms of what it is not. When two sounds are heard that match for pitch, loudness, and duration, and a difference can still be heard between the two sounds, that difference is called timbre. There are two physical correlates of timbre spectrum envelope and amplitude envelope. It is mainly determined by the harmonic content of a sound and the dynamic characteristics of the sound as vibrato and the attack-delay envelope of the sound [9].

A. *Timbre and Spectrum Envelope*

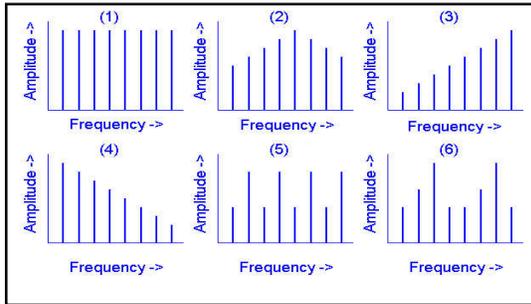


Fig 1: Six Synthesized Sounds Differing in Spectrum Envelope.

Timbre differences between one musical instrument and another are partly related to differences in spectrum envelope (figure-1) -- differences in the relative amplitudes of the individual harmonics. In the examples above, figure 2, we would expect all of these sounds to have the same pitch because the harmonic spacing is the same in all cases. The timbre differences that you would hear are controlled in part by the differences in the shape of the spectrum envelope.

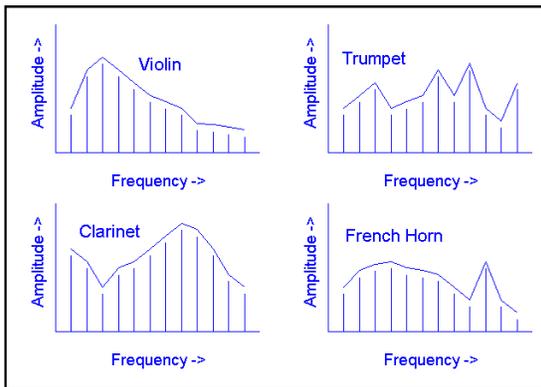


Fig 2: Timbre in different instruments.

A. *Amplitude Envelope*

Of particular importance to the timbre of a sound are the attack and the decay. Experiments where the attack and decay of a recorded sound have been removed have shown listeners have difficulty identifying the instrument.

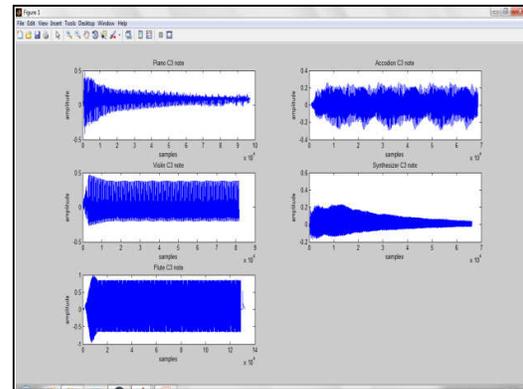


Fig 3: Attack, decay, sustain and release

The first part of the envelope is called the attack. How long it takes to go from silence to the maximum volume. A drum and a piano usually have short attack times. The next section is called decay. During the decay section the amplitude would decrease from the maximum level to some constant level. Drums have short decay times; a piano might have a slightly longer decay time and a horn longer still. The sustain level is a constant level that the sound maintains during the middle part of the sound. The final part of the envelope is called the release. The release time is the time it takes the sound to fade from the sustain level to silence. Figures 3 and 4 explain attack, decay, sustain and decay.

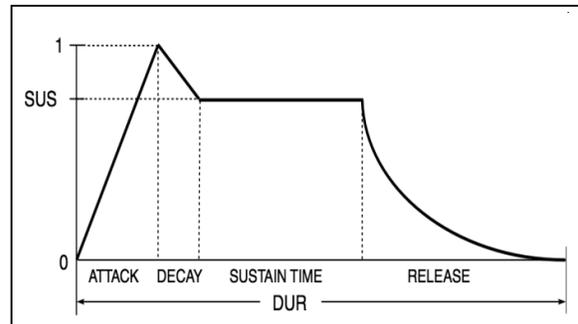


Fig 4: Attack, decay, sustain and release

A. *Slope of the Attack*

Another feature of timbre is the slope of attack of the waveform. The six diagrams showed in figure 5, the attack envelopes of the waveforms of six instruments, including acoustic piano, flute, violin, and acoustic guitar, trumpet, and string ensembles.

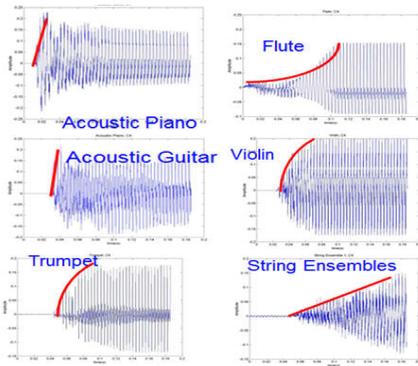


Fig 5: Attack envelopes waveforms of six musical instruments.

IV. DIFFERENT FEATURES OF INSTRUMENT

The term time domain generally refers to topics (analysis, synthesis, signal modulation, etc.) that have to do with two-dimensional data types, amplitude and time being the two dimensions.

Time Domain Features
Zero crossing rate (ZCR) (1)
Energy (2)
Attack time (3)
Decay time (4)
Sustain time (5)
Release time (6)
rms energy (7)
Autocorrelation coefficients (8)

Table 1: Time domain features

The basic concept behind frequency domain revolves around Fourier transform.

Frequency Domain Features
Spectral Centroid (9)
Relative power of fundamental component (10)
Relative power of Fundamental to ith component (10-17)
Relative power in odd and even component(18)
Fundamental frequency (19)
No. of harmonics (20)
Constant Q transform (21)
LPC coefficient (22-35)
MFCC coefficients (36-75)
Wavelet coefficients (76-86)
No. of odd harmonics (87)
No. of even harmonics (88)

Table 2: Frequency domain features

V. SIMILARITY SEARCH

Basic Challenge for instrument recognition is optimal and compact feature set to be found for allowing classifier to be built quickly and accurately. To find optimal and compact feature set for Musical Instrument recognition [3],[8]. Music search using DTW-MFCC, DTW-LPC KNN classifier. Music similarity search, we propose a stepwise method

- Step 1: Selection of Musical Instruments
- Step 2: Collection of music data base
- Step 3: Study and extraction of Musical Instrument features [2].
- Step 4: Reduction of feature dimension using dimension reduction technique [3].
- Step 5: Recognition of Musical instrument and instrument family using Machine learning techniques like KNN , ANN , SVM [4].
- Step 6: Finding of best features and recognition technique [5].
- Step7: Music Similarity search

VI. RESULTS

The instruments used for this experiment are Piano, Flute, Guitar, Accordion, Synthesizer, and Harmonica. Total number of notes considered is 840. In which 420 notes are used for training and 420 notes for testing. Features that are considered are LPC (13 features), MFCC (39 coefficients). Classifier used is KNN along with DTW. The results are shown in table-3 in terms of recognition rates.

	LPC without DTW	LPC with DTW	MFCC without DTW	MFCC with DTW
Recognition Rate	66%	72%	75%	82%

Table 3: Recognition rates

REFERENCES

- [1] Da Deng , Christian Simmermacher, Stephen Cranefield ,“A Study on Feature Analysis for Musical Instrument Recognition “ , IEEE transaction on Systems , Man , and Cybernetics part B: cybernetics , Vol. 38, No. 2, April 2008
- [2] Tao Li, Rochester “Music feature extraction using wavelet coefficient extraction “,U.S.Patent No. US7,091,409 B2 , Date 15 Aug 2006.

- [3] Jayme Garcia , Arnal Barbedo and George Tzanetakis, “Musical Instrument Recognition Using Individual Partial” , IEEE transaction on audio , speech and Language processing , Vol 19, No. 1, Jan 2011.
- [4] Cyril Joder, Slim Essid, and Gaël Richard,“ Temporal Integration for Audio Recognition With Application to Musical Instrument Recognition “IEEE transaction on audio , speech and Language processing , Vol 19, No. 1, May 2009.
- [5] Aggelos Pikrakis ,“ Recognition of Isolated Musical Patterns Using Context Dependent Dynamic Time Warping “IEEE transaction on audio , speech and Language processing , Vol 11, No. 3, Jan 2003
- [6] WEI-MIN LI1, FENG LI1, JIAN-WEI LIU2 , JIA-JIN LE1, “Similarity search over data stream using LPC-DTW “ , : Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007
- [7] Mahmoud I. Abdalla and Hanaa S. Ali , ”Wavelet-Based Mel-Frequency Cepstral Coefficients for Speaker Identification using Hidden Markov Models “ ,JOURNAL OF TELECOMMUNICATIONS, VOLUME 1, ISSUE 2, MARCH 2010.
- [8] Toni Heittola, Anssi Klapuri, Tuomas Virtanen ,” MUSICAL INSTRUMENT RECOGNITION IN POLYPHONIC AUDIO USING SOURCE-FILTER MODEL FOR SOUND SEPARATION” in information retrieval, 2009
- [9] Jiang, W. et al. “Multiple Classifiers for Different Features in Timbre Estimation.” In Z. W. Ras & L.-S. Tsay, eds. Advances in Intelligent Information Systems. Springer.

