International Journal of Advanced Engineering and Management Research
Vol. 3 Issue 5; 2018

www.ijaemr.com ISSN: 2456-3676

FEATURE EXTRACTION FOR MUSIC WITH WAVELETS

Chih-Hsu Hsu

Department of Applied Information and Multimedia, Ching Kuo Institute of Management and Health,
No.336, Fu Hsin Rd., Keelung 203, Taiwan, R.O.C.

Abstract

Music is widely used for various purposes, it contributes to the treatment of patients since people appreciate it and use it for music. Music signal processing researchers are expected to propose more efficient and effective solution for real world. This paper proposed a multi-resolution feature extraction technique to deal with music signals. I utilized a fuzzy system to search their clusters of features. The fuzzy rules with variable fuzzy regions were defined by activation rectangles, which show the existence region of data for a class and overlapping rectangles, which overlapping the existence of the data for the other classes. The effectiveness of the clusters is encouraging.

Key Words: feature extraction, music signal, wavelet transform, fuzzy system.

INTRODUCTION

Music has many features including their originality, structure, flexibility and improvisation to decrease the limitations of participants, times and spaces. At the same time, through the group interaction and respecting their willing, music could inspire participants' motivation and participation. In this paper, a music signal based on wavelet transform is proposed. Music is widely used for various purposes, it contributes to the treatment of patients since people appreciate it and use it for music therapy. Found in all cultures around the world, music has the power to touch the human spirit at a deep level, often without the use of words.

Music signal processing researchers are expected to propose more efficient and effective solution for real world. Using real world data is one of the good ways to think such solutions. In addition, we are required to grow the real fields by ourselves where the technologies potentially can be used. In music signal, we utilize the technology of signal processing to analyze the features of each music melody. Music signal could be extracted by a set of features of each frame from wavelet transform and Fouier transform.

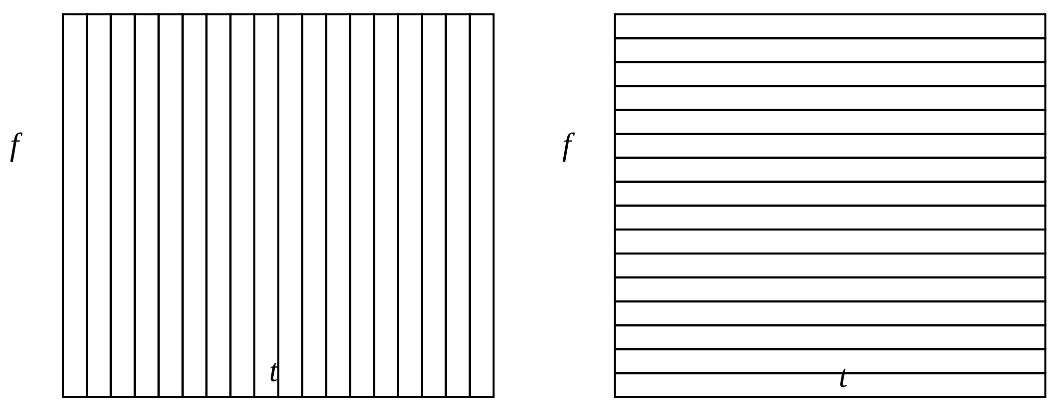
Feature extraction is commonly treated as a problem of classification. In music signal, we utilize the technology of signal processing to analyze the features of music. Music signal could be extracted by a set of features of each frame from wavelet transform. Pitch is a common parameter in many types of music signal processing. It is defined as the perceived fundamental frequency of a signal and is used in many applications of music signals. Pitch estimation algorithms can be classified in two separate categories, spectral-domain based and time-domain based detection. Spectral pitch detectors estimate the pitch period directly using windowed

segments of music signal. However, time based pitch detectors estimate the pitch period by determining the glottal closure instant and measuring the time period between each "event" [1].
This paper is organized as follows. In section 2 we discuss the feature extraction based on wavelet transform. Section 3 briefly describes the fuzzy system. The experimental results are given in Section 4. Finally, some concluding remarks are presented in Section 5.

FEATURE EXTRACTION BASED ON WAVELET TRANSFORM

Wavelet Transform

The multi-resolution formulation of wavelet transform is obviously designed to represent signals where a single event is decomposed into finer and finer detail, but it turns out also to be valuable in representing signals where a time-frequency or time-scale description is desired even if no concept of resolution is needed. In many applications, one studies the decomposition of a signal in terms of basis function. For example, stationary signals are decomposed into the Fourier basis using Fortier transform. For non-stationary signals (i.e. signals whose frequency characteristics are time-varying like music, speech, image, etc.) the Fourier basis is ill-suited because of the poor time-localization. The classical solution to this problem is to use the short-time (or windowed) Fourier transform. However, the short-time Fourier transform has several problems, the most severe being the fixed time-frequency resolution of the basis functions. Wavelet techniques give a new class of bases that have desired time-frequency resolution properties. The "optimal" decomposition depends on the signal studied.
Each function in a basis can be considered schematically as a tile in the time-frequency plane, where most of its energy is concentrated. Non-overlapping tiles can schematically capture or thonormality of the basic functions. With this assumption, the time-frequency tiles for the standard basis and the Fourier basis are shown in Fig. 1.



(a) (b)

Fig. 1 (a) Standard time domain basis, (b) Standard frequency domain basis.

The discrete wavelet transform is another signal independent tiling of the time-frequency plane suited for signals where high frequency signal components have shorter duration than low frequency signal components. The discrete wavelet transform coefficients are a measure of the energy of the signal components in the time-frequency plane, giving another tiling of time-frequency plane. Fig. 2 shows the corresponding tiling description, which illustrates time-frequency resolution properties of a discrete wavelet transform basis.

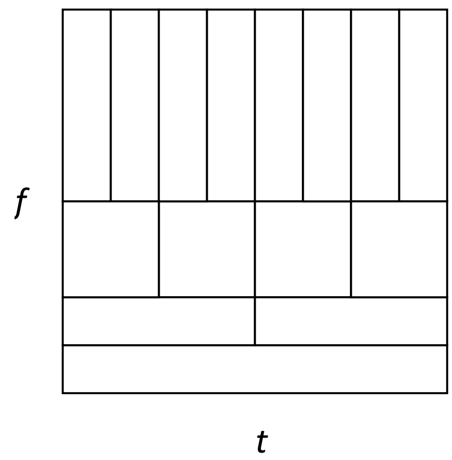


Fig. 2 Three-scale of wavelet basis.

The definition of the scaling function $\phi_{j,k}(t)$ and wavelet function $\psi_{j,k}(t)$ is given by [2].

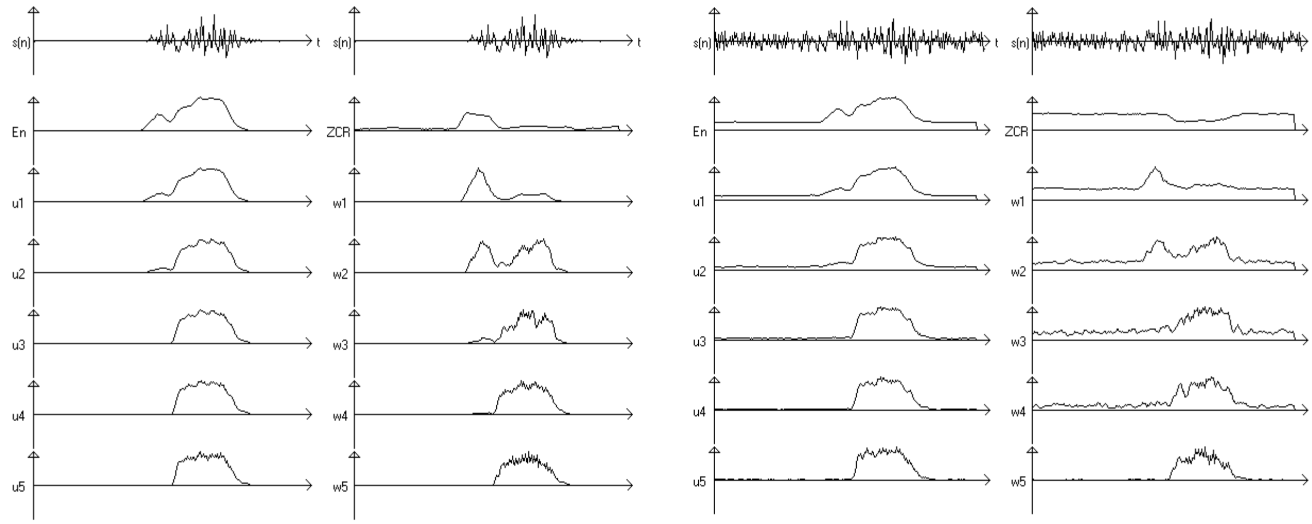
$$\phi_{j,k}(t) = 2^{j/2}\phi(2^{j}t - k) \qquad j,k \in Z \qquad (1)$$

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^{j}t - k) \qquad j,k \in Z \qquad (2)$$

This two-variable set of basis function is used in a way similar to the short time Fourier transforms. A signal space of multi-resolution approximation is decomposed by wavelet transform in an approximation (lower resolution) space and a detail (higher resolution) space. In order to generate a basis system that would allow higher resolution decomposition at higher frequencies, we will iterate the wavelet transform recursively to divide the approximation space, giving a left binary tree structure. The wavelet packet was allowed a finer and adjustable to particular signals or signal classes. The wavelet packet decomposes the detail spaces as well as approximation ones.

Feature Extraction with Wavelets

The signal produce each approximation signals u(t) and detail signals w(t). Fig. 3 shows the contours of energy (En), zero crossing rate (ZCR), variance of five levels of approximation (u1, … u5) and detail (w1, … w5) for the utterance /chii/, in the higher and lower SNR environment, respectively.



(a)

(b)

Fig. 3 (a) Contours of energy (En), zero crossing rate (ZCR), variance of five levels of approximation (u1, … u5) and detail (w1, … w5) for the utterance /chii/, in the higher SNR environment. (b) Contours of energy (En), zero crossing rate (ZCR), variance of five levels of approximation (u1, … u5) and detail (w1, … w5) for the utterance /chii/, in the lower SNR environment.

Several approaches have been proposed for classification problems. Some works focus on conventional probabilistic and deterministic classifiers [3]. Another approach uses neural networks to classify patterns [4]. In music signal, we utilize the technology of signal processing to analyze the features of music. Music signal could be extracted by a set of features of each frame from wavelet transform and Fourier transform. The windowed Fourier transform has uniform resolution over the time frequency plane. It is difficult to detect sudden burst in a slowly varying signal by Fourier transform. Wavelet transform overcomes the problem of fixed resolution, using adaptive window sizes, which allocate more time to the lower frequency and less time for the higher frequency [5]-[7].

FUZZY SYSTEM

Fuzzy System Architecture

The construction of a rule-based expert system involves the process of acquiring production rules [8]. Production rules are often represented as " IF condition THEN act. The class of fuzzy system provides a tool for machine learning. The classification knowledge is easily extracted from the weights in a rectangle. First, we divided the range of an output variable into many intervals and using the input data belonging to each interval. Each rule is composed of an activation rectangle, which defines the existence region of a class and, if necessary, an overlapping rectangle which overlapped the existence of data in that activation rectangle. We determine activation rectangle, which define the input region corresponding to the class, by calculating the maximum and minimum values of input data for each class. Fig. 4 illustrates the architecture of a fuzzy system.

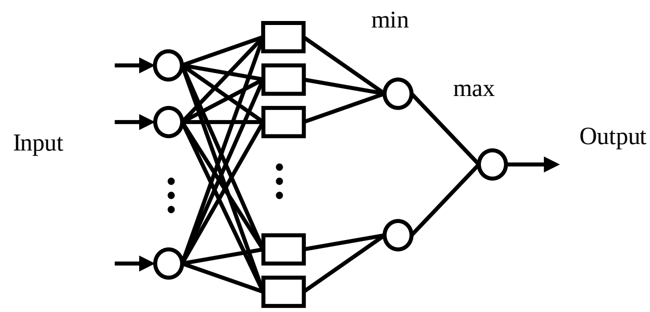


Fig. 4 A fuzzy system architecture.

Rule Extraction of a Fuzzy System

Let a set of input data for class i be X_i, where $i = 1,...n$. We define the activation rectangle A_{ij} as

$$A_{ij} = \left\{ x \,\middle|\, u_{ijk} \le x_k \le U_{ijk}, k = 1,...n \right\}$$

(3)

and define the fuzzy rule r_{ij} without overlapping as follows:

If x is A_{ij}, then x belongs to class i,

(4)

the overlapping rectangle I_{ij} as

$$I_{ij} = \left\{ x \,\middle|\, v_{ijk} \le x_k \le V_{ijk}, k = 1,...n \right\}$$

(5)

and define the fuzzy rule r_{ij} with overlapping rectangle as follows:

If x is A_{ij} and x is not I_{ij}, then x belongs to class i,

(6)

Fuzzy Rules Inference of a Fuzzy System

The degree of membership of the fuzzy rule for a given input x is determined by the membership function of the activation rectangle. While the degree of membership of the fuzzy rule for a given input x is determined by the difference between the membership function of the activation rectangle and that of the overlapping rectangle. The membership function for each input variable is a trapezoidal shape. Fig. 5 shows one-dimension membership function for the rectangle, where u_k and U_k denote the minimum and maximum values of the k-th dimension of the rectangle, respectively.

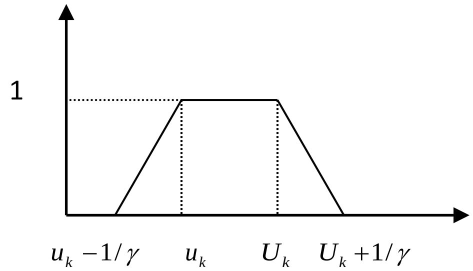


Fig. 5 The *k*-th dimension membership function for the rectangle.

$$m_X(x) = \min_{k=1,\ldots,n} m_X(x,k)$$

(7)

$$m_X(x,k) = \begin{cases} 1 & \text{for } u_k \le x_k \le U_k \\ 1 - \max(0, \min(1, \gamma(u_k - x_k)) & \text{for } x_k \le u_k \\ 1 - \max(0, \min(1, \gamma(x_k - U_k)) & \text{for } x_k \ge U_k \end{cases}$$

(8)

where γ is a sensitive parameter. The minimum value in (8) is taken so that the degree of membership within the rectangle and on the surface of the rectangle becomes 1.
The degree of membership of a fuzzy rule respected by (4) is:

$$d_{r_{ij}}(x) = m_{A_{ij}}(x)$$

(9)

The degree of membership of a fuzzy rule respected by (6) is:

$$d_{r_{ij}}(x) = \max(0, m_{A_{ij}}(x) - m_{I_{ij}}(x))$$

(10)

EXPERIMENTAL RESULTS

In feature extraction experiments, the database of music is adopting western classical music and the original songs. We utilize the technology of signal processing to analyze the features of music signals. Music signals could be extracted by a set of features of each frame from wavelet transform and Fourier transform. The windowed Fourier transform has uniform resolution over the time frequency plane. It is difficult to detect sudden burst in a slowly varying signal by Fourier transform. Wavelet transform overcomes the problem of fixed resolution, using adaptive window sizes, which allocate more time to the lower frequency and less time for the higher frequency.

The melody and rhythm play important roles in music care. Table 1 shows the results of clusters with Daubechies wavelets, the effectiveness of the clusters is encouraging. The results of this study can be applied as the reference of music intervention.

Table 1 The results of clusters with Daubechies wavelets.

Music / Features	1	2	3	4	5	Average

Training Set (%)	99.9	99.6	99.3	99.8	98.9	99.50
Testing Set (%)	96.2	95.9	95.6	96.0	93.9	95.52

CONCLUDING REMARKS

In this paper, a fuzzy model for music is presented. I utilize a multi-resolution feature extraction technique to deal with music signals. The fuzzy rules with variable fuzzy regions were defined by activation rectangles, which show the existence region of data for a class and overlapping rectangles, which overlapping the existence of the data for the other classes. The effectiveness of the clusters with Daubechies wavelets is encouraging. In the near future, we will try to apply the fuzzy system to adjust features to music system.

References

C. Wendt and A.P. Petropolis, "Pitch determination and speech segmentation using the discrete wavelet transform," In Proceedings of IEEE International Symposium on Circuits and Systems, pp.45-48, 1996.

C.S. Burrus, R.A. Goliath, and H. Guo, Introduction to Wavelets and Wavelet Transforms, Prentice-Hall, 1998.

X. Huang, A. Acero, and H.W. Hon, Spoken language processing: A guide to theory, algorithm, and system development, Prentice-Hall, 2001.

J. Jose de Jesus Guerrero-Turrubiates, S.E. Gonzalez-Reyna, S.E. Leadsman-Orozco, and J.G. Avina-Cervantes, "Pitch Estimation For Musical Note Recognition Using Artificial Neural Networks," In Proceedings of International Conference on Electronics, Communications and Computers, Cholula, Mexico, pp.53-58, Feb. 26-28, 2014.

W.C. Chen, C.T. Hsieh, and C.H. Hsu, "Two-stage vector quantization based multi-band models for speaker identification," In Proceedings of International Conference on Convergence Information Technology, Gyeongju, Korea, pp.2336-2341, Nov. 21-23, 2007.

T.H. Luo and C.H. Hsu, "Signal Analysis for "Kagaya Miyamoto Shiki" Music Therapy," In Proceedings of International Scientific Conference on Engineering and Applied Sciences, Singapore, pp. 283-284, Aug. 15-17, 2014.

T.H. Luo and C.H. Hsu, "Feature Extraction for the Classical Music of "Kagaya Miyamoto Shiki Music Care"," In Proceedings of Asian Conference on Engineering and Natural Sciences, Fukuoka, Japan, pp. 624-625, Feb. 1-3, 2016.

S. Abe and M.S. Lan, "Fuzzy Rules Extraction Directly from Numerical Data for Function Approximation," IEEE Trans. on System, Man, and Cybernetics, Vol. 25, No. 1, pp. 119-129, Jan. 1995.