International Journal of Information Processing, 8(4), 45-55, 2014
ISSN : 0973-8215
IK International Publishing House Pvt. Ltd., New Delhi, India

Enhanced Statistical Descriptors of the Intrinsic Mode Functions for Query by Humming Systems

Trisiladevi C Nagavi[a], Nagappa U Bhajantri[b]

[a]Department of Computer Science and Engineering, S J College of Engineering, Mysore, 570006 Karnataka, India, Contact: tnagavi@yahoo.com

[b]Department of Computer Science and Engineering, Government Engineering College, Chamarajanagar, 571313, Karnataka, India.

The advancement in the field of music signal processing insists on development of specific and adept Music Information Retrieval (MIR) techniques. In this paper, we propose an effective Query by Humming (QBH) MIR system for retrieving the favorite song based on enhanced statistical descriptors of the Intrinsic Mode Function (IMF) and Humming Query (HQ). Initially we have adopted the Empirical Mode Decomposition (EMD) technique for music processing and imparted outcome, to be exact IMFs as the significant source for the extraction of statistical descriptors. In the proposed system, nine statistical descriptors corresponding to five statistical properties are extracted. The proposed system's effectiveness is appraised through series of experiments on a music database consisting of 1495 fragments extracted manually from 1200 songs. The experimental findings effectively retrieve the favorite song based on HQ and affirm the importance of EMD.

Keywords : Empirical Mode Decomposition (EMD), Humming Systems, Intrinsic Mode Functions (IMF), Query by Humming (QBH), Statistical Descriptors.

1. INTRODUCTION

The precincts of syntactic and text based music retrieval techniques are overcome by QBH music retrieval system, which allows users to discover favorite song through humming a piece of a song [1]. The versatility and rising availability of contemporary digital music necessitates QBH music retrieval systems for applications like Radio or Disco Jockey (DJ) play list generation, on line music access, personal and public music collections management [2].

Music melody representation through the set of discriminating features, query processing and melody matching are some of the key challenges in QBH system. Music signal often exhibits rich and complex dynamic structure. In order to capture this structural information various signal transformations such as Fourier Transforms (FT), Wavelet Transforms (WT) are applied [3][4]. FT is more suitable for linear and stationary signals and WT for linear and non-stationary signals respectively. However, music signal's non-stationarity and dynamisms are quite difficult to capture.

The Hilbert-Huang Transform (HHT) is one among the options available that facilitates to quantify the complex structures and dynamism in music signals [5]. Hence we have adopted HHT as a reliable means to analyze and transform music signal before feature extraction [6]. The main constituent of HHT is the Empirical Mode Decomposition (EMD), which decomposes the signal into a summation of zero-mean Amplitude Modulated (AM) or Frequency Modulated (FM) components, called Intrinsic Mode Functions (IMF) [5].

Then enhanced statistical descriptors such as mean or first moment, Standard Deviation (SD) or second moment, skewness or third moment, kurtosis or fourth moment, three percentiles, Shannon entropy and Zero Crossing Rate (ZCR) are extracted from significant IMFs to represent music melody.

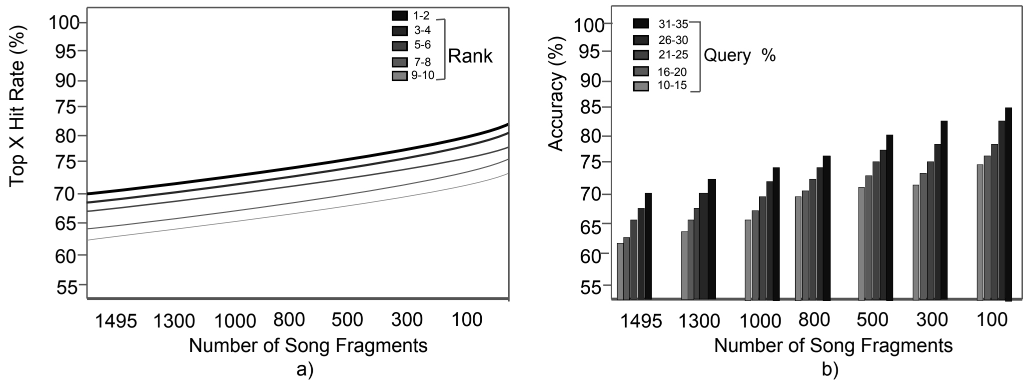

Figure 8. Performance Analysis (a) **Top X Hit Rate vs. Number of Song Fragments** (b) **Accuracy vs. Number of Song Fragments**

Table 2
Percentage of Songs Precisely Retrieved at Various Number of IMFs and Number of Song Fragments

IMFs	**Number of Song Fragments**						
	1495	**1300**	**1000**	**800**	**500**	**200**	**100**
1	58.90	59.20	61.99	64.53	65.94	67.81	**70.03**
2	57.19	58.30	61.20	63.58	64.00	66.34	69.89
3	56.10	57.14	60.10	62.34	64.39	65.48	68.22
4	54.89	56.70	59.10	60.30	62.80	64.00	66.90
5	53.77	55.12	58.55	59.11	61.44	62.12	65.32
6	52.70	54.40	57.30	58.10	60.40	61.70	64.10
7	51.11	53.00	55.11	57.35	59.91	60.21	62.14
8	50.90	52.40	53.90	56.60	58.80	59.30	61.10
9	49.87	51.14	52.99	54.77	56.14	58.40	59.33
10	48.10	50.20	51.70	52.30	54.20	56.20	58.40

gent Systems and Computing, 309:139–150, Bhubaneswar, India, 2014.

3. Peyman Heydarian and Joshua D Reiss. Extraction of Long Term Structures in Musical Signals using the Empirical Mode Decomposition, *In Proceedings of 8th International Conference on Digital Audio Effects (DAFX-05)* , Madrid, Spain, 2005.
4. Sangjin Cho and Yejin Seo. Extraction of Significant Features using Empirical Mode Decomposition and its Application, *In Proceedings of World Congress on Engineering and Computer Science (WCECS)*, 1, SanFrancisco, USA, 2013.
5. James M Hughes, Dong Mao, Daniel N Rockmore, Yang Wang and Qiang Wu. Empirical Mode Decomposition Analysis for Visual Stylometry, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(34):2147–2157, 2012.
6. Trisiladevi C Nagavi and Nagappa U Bhajantri. Effectiveness of Empirical Mode Decomposition for Query by Humming Systems, *In Proceedings of the Eighth International Conference on Data Mining and Warehousing (ICDMW) , Elsevier*, pages 259-267, Bangalore, India, 2014.
7. Tee-Ann Teo and Chi-Chung Lau. Pyramid based Image Empirical Mode Decomposition for the Fusion of Multispectral and Panchromatic Images, *In EURASIP Journal of Ad-*

vanced Signal Processing, 2012.
8. Md Khademul Islam Molla, Akimasa Sumi and M Sayedur Rahman. Analysis of Temperature Change under Global Warming Impact using Empirical Mode Decomposition, *In International Journal of Information Technology*, 3(2):131–143, 2007.
9. Tatiana Endrjukaite and Yasushi Kiyoki. Music Similarity Analysis through Repetitions and Instantaneous Frequency Spectrum, *In International Journal of Signal Processing Systems*, 1(2), 2013.
10. Michael Fulton and P J J Soraghan. Ensemble Empirical Mode Decomposition Applied to Musical Tempo Estimation, *In Proceedings of 15th Annual Conference on Systems, Signals and Image Processing* , Glasgow, pages 1–4, 2008.
11. Aggelos Pikrakis and Sergios Theodoridis. An Application of Empirical Mode Decomposition on Tempo Induction from Music Recordings, *In the ISMIR Proceedings*, pages 301–304, Austrian Computer Society, 2007.
12. Peyman Heydarian and Joshua D Reiss. Extraction of Long Term Rhythmic Structures using the Empirical Mode Decomposition, *In Proceedings of the* 122^{nd} *Audio Engineering Society Convention on Analysis and Synthesis of Sound*, Vienna, Austria, 2007.
13. Erwin Kreyszig. Advanced Engineering Mathematics, *John Wiley and Sons Inc.*, 2011.
14. Johannes Mayer. On Testing Image Processing Applications with Statistical Methods, *In Proceedings of the Software Engineering (SE 2005) , Lecture Notes in Informatics*, 64:69–78, 2005.
15. Vijay Kumar and Priyanka Gupta. Importance of Statistical Measures in Digital Image Processing, *In International Journal of Emerging Technology and Advanced Engineering*, 2(8):56–62, 2012.
16. A Arnaout, B Alsallakh, R Fruhwirth, G Thonhauser, B Esmael and M Prohaska. Diagnosing Drilling Problems Using Visual Analytics of Sensors Measurements, *In Proceedings of the IEEE International Conference on Instrumentation and Measurement Technology (I2MTC)* , pages 1750-1753, Graz, Austria, 2012.
17. Bilal Esmael, Arghad Arnaout, Rudolf K Fruhwirth and Gerhard Thonhauser. A Statistical Feature-Based Approach for Operations Recognition in Drilling Time Series, *In International Journal of Computer Information Systems and Industrial Management Applications*, 5:454–461, 2013.
18. Romain Fontugne, Jorge Ortiz, David Culler and Hiroshi Esaki. Empirical Mode Decomposition for Intrinsic Relationship Extraction in Large Sensor Deployments, *In the First International Workshop on Internet of Things Applications (CPS Week)*, Beijing, China, 2012.

Trisiladevi C Nagavi is an Assistant Professor at Department of Computer Science and Engineering, Sri Jayachamarajendra College of Engineering Mysore, Karnataka, India. She is pursuing her Ph.D in Computer Science and Engineering at the Visvesvaraya Technical University, Belgaum, Karnataka, India. She has received her M.Tech in Software Engineering at the Visvesvaraya Technical University, Belgaum, Karnataka, India in 2004. Her areas of interest are Audio, Speech and Image processing. She has published research papers at national and international journals, conference proceedings as well as chapters of books.

Nagappa U Bhajantri is a Professor and HOD of Computer Science and Engineering Department in Government Engineering College, Chamarajanagar, Karnataka, India. Received his Ph.D in Computer Science at the University of Mysore, Karnataka, India. He did his M.Tech in Computer Technology (Electrical Engineering Department) from Indian Institute of Delhi, India in 1999. His areas of interest are Image, Audio and Video processing. He is author of research papers published at national and international journals, conference proceedings as well as book chapters.