Automatic Music Composition based on HMM and Identified Wavelets in Musical Instruments

M.S. Sinith, *K.V.V. Murthy
Dept. of Electronics and Communication Engineering
College of Engineering, Trivandrum, India
sinith@ieee.org
Dept. of Electrical Engineering
Indian Institute of Technology Gandhinagar, India.
kvvm@iitgn.ac.in

***Abstract*—Automatic Music Composition plays a crucial role in the musical research and can become a tool for the incorporation of artificial intelligence in computer musicology. This paper finds an efficient method for identifying the wavelets and filter bank coefficients in musical instruments using NLMS algortithm and the usage of these wavelets for Automatic music Composition using Hidden Markov Model. In this paper, a technique to identify the scaling function and the wavelet functions of the wavelets present in musical instruments, violin and flute, is presented. NLMS algorithm is used to identify the filter bank coefficients of wavelet-like elements, found repeating in musical notes of the instruments. Pre-trained hidden markov models for each raga of South Indian Music is used for the composition. The HMM selected has twelve states which represent the twelve notes in South Indian music. Fundamental frequency tracking algorithm, followed by quantization is done. The resulting sequence of frequency jumps of different musical clips of same musical pattern (Raga) is presented to Hidden Markov Model of a particular Raga for training. The HMM model of that Raga along with the filter coefficient is used to regenerate a piece of music in that particular raga. The methodology is tested in the context of South Indian Classical Music, using the wavelet of classical music intruments, Flute and Violin.**

I. Introduction

The earliest and most well known survey of digital signal processing techniques for the production and processing of musical sounds was authored by [1] in 1976. Wavelets form an integral part of digital signal processing today. Analysis of musical notes of different musical instruments shows that each instrument is associated with a unique wavelet [2]. In widely adopted Fourier representation in signal processing, one can get the features of the signal either in Time Domain or in Frequency Domain, one at a time. Both are extensively used in analysis, design and various other applications. However there are many instances in which the localization in time as well as localization in frequency, both are required simultaneously. Short duration signals need to be localized in time and small bandwidth signals localized in frequency. In musical signals, small duration signals or small bandwidth musical pieces are placed at an effective temporal position to give special effects. They need to be captured in time as well as in frequency simultaneously. Fourier representation is not suited for such requirements. Wavelet transform of finite energy can effectively capture the above requirement. Wavelet transform replaces Fourier Transforms sinusoidal waves by a family generated by translations and dilations of finite energy signals called wavelet. Wavelets are manipulated in two ways. The first one is translation where the position of the wavelet is changed along the time axis. The second one is scaling. Scaling means changing the frequency of the signal. In wavelet analysis, two special functions: the (1) wavelet function and (2) the scaling function are used. They have unique expressions:

$$\psi(t) = \Sigma g(n)\phi(2t-n) \quad (1)$$

$$\phi(t) = \Sigma h(n)\phi(2t-n) \quad (2)$$

$$where, g(n) = (-1)^{1-n}h(1-n) \quad (3)$$

II. Filter bank Theory

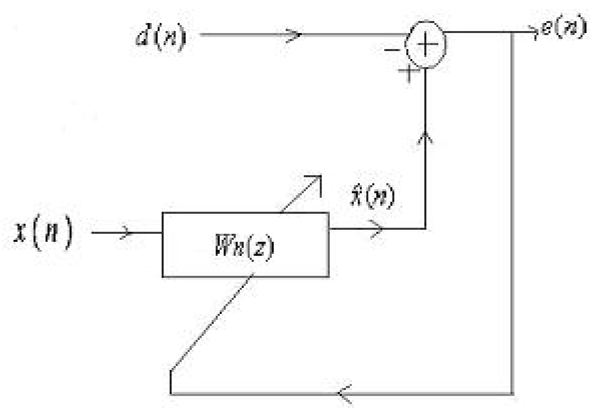
As studies reveal wavelets can be generated from a set of transfer functions H(z) and G(z). A wavelet can be reconstructed from its approximation and detail coefficients [3]. The wavelet acts as a bank of band pass filters [4]. Most of the wavelet applications are dealt with the coefficients h(n) and g(n) from (1) and (2). These filters form Quadrature Mirror Filters(QMF) whose spectra are mirror images of each other. The filter formed by the mother wavelet acts as constant-Q filters. Qfactor is given by:

$$Qfactor = centrefrequency/bandwidth \quad (4)$$

QMF filters can be used for the filter bank implementation of Discrete Wavelet Transform(DWT).The filters form the high pass and low pass filter segment of the QMF filter that gives the approximation and detailed coefficients of the signal. The approximation coefficients obtained at the output of low pass filter is down sampled and passed through another pair of QMF filters and the process of decomposition continues. We can reconstruct the wavelet by using the same QMF filters by giving the approximation and detail coefficients as filter inputs. In this paper, a new algorithm is proposed to find out the h(n) and g(n) in an iterative manner. Once h(n) is obtained, g(n) can be obtained using (3). The Scaling function of the standard wavelets are fed into the Normalized Least Mean Square(NLMS) algorithm. The role of an adaptive filter is to find the filter coefficients that give minimum error. The block diagram of the adaptive filter is given in fig1. Our algorithm

978-1-61284-653-8/11/$26.00 ©2011 IEEE

Fig. 1. NLMS filter



suggests that by up-sampling the estimated signal obtained at output end of the adaptive filter and giving this new signal as input and the original signal as desired signal, a new set of filter weights can be obtained that would give minimum error. This result however is accomplished by repeating the above steps iteratively for many times. This set of filter weights help to reconstruct the $h'(n)$ with minimum error. This process was carried out with all the standard wavelets. For all these wavelets, with this set of filter weights, their respective $h'(n)$ and $g'(n)$ were successfully reconstructed. Thus the wavelet and scaling functions of each wavelet for flute and violin is obtained using this algorithm.

III. Normalized Least mean square algorithm

In designing an FIR adaptive filter [5], [6] the goal is to find the weights of filter, w_n, at time n that minimizes the quadratic equation, i.e. the prediction error:

$$\zeta(n) = E[e^2(n)] \tag{5}$$

The vector that minimizes $\zeta(n)$ may be found by setting the derivates of $\zeta(n)$ with respect to w*(k) equal to zero. The weight-vector update equation used here is

$$w_{n+1} = w_n + \mu E[e(n)x^*(n)] \tag{6}$$

An LMS adaptive filter having p+1 coefficients require p+1 multiplication and p+1 addition to update the filter coefficients. In addition, it is necessary to compute the error Since w_n is a vector of random variables, the convergence of the LMS algorithm [7] must be considered within a statistical framework. The main disadvantage of the LMS algorithm is that it is very sensitive to the changes in the input signal . As a result, it is very difficult to find an optimum step size or convergence parameter , which guarantees convergence of the algorithm, as well as minimizes the time of computation. A suitable alternative is to use the Normalized LMS algorithm [8], which normalizes the LMS step size with the power of the input. When the input becomes too small, the NLMS algorithm can be modified by adding a small positive value ϵ to the power of the input signal.

$$w_{n+1} = w_n + \beta \frac{e(n)x^*(n)}{\epsilon + x(n^2)} \tag{7}$$

In the new algorithm developed [2], $h'(n)$ was tried to be reconstructed using LMS algorithm. Here the desired signal was down sampled and then up sampled before being given to the filter. But it encountered the problem of sensitivity of step size as mentioned above. The obtained $h'(n)$ varied much from the actual value as the step size is varied . This led us to modify our algorithm by incorporating NLMS algorithm.

IV. Hidden Markov Model

Hidden Markov models (HMMs) are mathematical models of stochastic processes, i.e. processes which generate random sequences of outcomes according to certain probabilities. A simple example of such a process is a sequence of coin tosses. More concretely, an HMM is a finite set of states, each of which is associated with a (generally multidimensional) probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state, an outcome or observation can be generated, according to the associated probability distribution. It is only the outcome not the state that is visible to an external observer. So states are hidden and hence the name hidden Markov model.

In order to define an HMM completely, the following elements are needed
The number of states of the model, N
The number of observation symbols in the alphabet, M.
A set of state transition probabilities

$$A = \{a_{ij}\}$$

$$a_{ij} = P\{q_{t+1} = j/q_t = i\}, 1 \le i, j \le N$$

where q_t denotes the current state.
A probability distribution in each of the states,

$$B = \{b_{jk}\}$$

$$b_{jk} = P\{\alpha_t = v_k/q_t = j\}, 1 \le j \le N, j \le N, 1 \le k \le M$$

where v_k denotes the k_{th} observation symbol in the alphabet and t the current parameter vector.
The initial state distribution,

$$\pi = \{\pi_i\}$$

where $\pi_i = p(q_1 = i), 1 \le i \le N$
Thus, an HMM can be compactly represented as

$$\lambda = \{A, B, \pi\}$$

Hidden Markov models and their derivatives have been widely applied to speech recognition and other pattern recognition problems [9]. Most of these applications have been inspired by the strength of HMMs, ie the possibility to derive understandable rules, with highly accurate predictive power for detecting instances of the system studied, from the generated models. This also makes HMMs the ideal method for solving Raga identification problems.

A. Selection of Hidden Markov Model

The HMM used in our solution is significantly different from that used in, say word recognition. In the HMM, used in our context [9], each note in each octave represents one state in λ . Thus, the number of states actually required is N=12x3=36 (Here, we are considering the three octaves of Indian classical music, namely the Mandra, Madhya and Tar Saptak, each of which consist of 12 notes) as used in [6]. Instead we have first mapped the frequencies in the Mandra and Tar Saptak into the corresponding frequencies in the Madhya. Hence the total number of states for the HMM is N=12. HMM in [10] is based on the frequency difference, where the total number of states needed for the HMM is very high compared to the HMM whose states are taken on the basis of absolute frequencies.So we have adopted absolute frequency as the basis for HMM.

The transition probability $A = \{a_{ij}\}$ represents the probability of note j appearing after note i in a note sequence of the Raga represented by λ.

The initial state probability $\pi = \{\pi_i\}$ represents the probability of note i being the first note in a note sequence of the Raga represented by λ .

The outcome probability $B = \{b_{ij}\}$ is set according to the following formula
$B_{ij} = 0, \forall i \neq j,$
$B_{ij} = 1, \forall i = j.$

The last condition takes the hidden character away from, but it can be argued that this setup suffices for the representation of Ragas, as our solution distinguishes between performances of distinct raga on the basis of the exact order of notes sung in them and not on the basis of the embellishments used.Thus, at each state α in λ, the only possible outcome is note α.

V. RESULTS

The filter bank coefficient of the standard wavelets db2 was found iteratively using the NLMS algorithm as in Table I. With this algorithm the wavelet and scaling functions were reconstructed successfully. New scaling functions and wavelet functions for two music instruments-violin and flute are found. The results are shown in Table II. The repeating elements were extracted from the classical music signals played. Average of these signals were found out with respect to each instrument. With this average being fed to the algorithm defined, scaling and wavelet functions of each instrument were obtained. The scaling and wavelet functions are shown in figures [2-5]. The Composition of music is automatically done using the probability values of the HMM

Fig. 2. Scaling function for violin

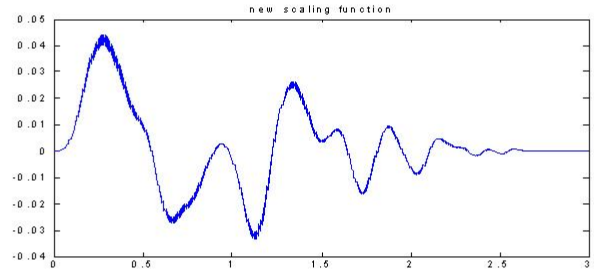
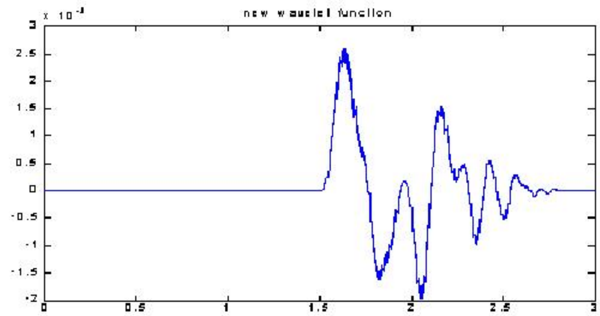
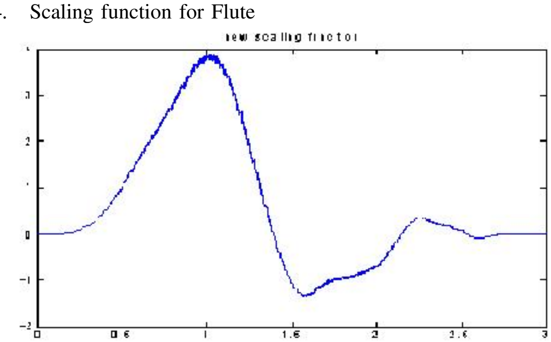


Fig. 3. Wavelet function for violin



obtained after training the model with several musical clips of same $Raga$. Each required note of the 12 notes in South Indian Classical music is obtained by Upsampling and then downsampling to get the desired pitch and these waves are concatenated to get the Composition. The HMM of a typical Raga $Kalyani$ obtained is shown in table III. If the wavelet of flute is replaced by Violin, beautiful Composition of Violin is also obtained.

Fig. 4. Scaling function for Flute



.

Acknowledgments

First Author Sinith thanks all his friends who have given valuable information relating to the work. we also thank

TABLE I
OBTAINED FILTER COEFFICIENT USING NLMS

Actual Filter Coefficients h(n)	Step-Size	Obtained filter coefficients
$[0.683, 1.183, 0.317, -0.183]$	0.3	$[0.682, 1.666, 0.3172, -0.1801]$
$[0.683, 1.183, 0.317, -0.183]$	0.7	$[0.683, 1.183, 0.317, -0.183]$
$[0.683, 1.183, 0.317, -0.183]$	1.2	$[0.68291.1830.317 - 0.183]$
$[0.683, 1.183, 0.317, -0.183]$	1.9	$[0.726, 1.994, 0.329, -0.1802]$

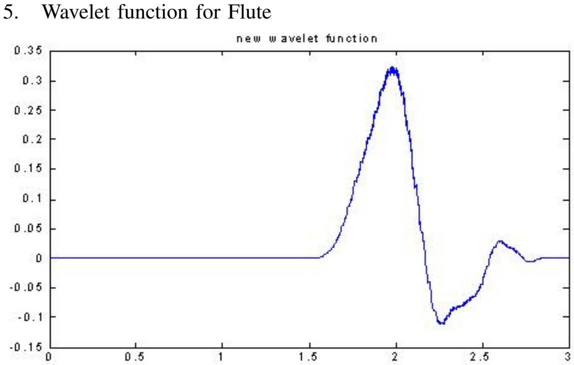
TABLE II
FILTER COEFFICIENTS FOR FLUTE AND VIOLIN

$Instrument$	$FilterCoefficient$
Violin	$[0.0388, 0.0640, 0.0777, 0.0839, 0.0865, 0.0871, 0.0878, 0.0902, 0.0955, 0.1017, 0.1088, 0.1139, 0.1175, 0.1191, 0.1191, 0.1201, 0.1233, 0.1283, 0.$
Flute	$[0.0833, 0.1249, 0.1458, 0.1562, 0.1614, 0.1640, 0.1654, 0.1660, 0.1663, 0.1665, 0.1666, 0.1666, 0.1666, 0.1667, 0.1667, 0.1667, 0.1667, 0.1667, 0.$

TABLE III
HMM FOR A TYPICAL $Raga$

$states$	1	2	3	4	5	6	7	8	9	10	11	12
1	0.3548	0	0	0.2258	0.0323	0.0645	0.0323	0.129	0.1613	0	0	0
2	0.6667	0.1667	0	0	0	0.1667	0	0	0	0	0	0
3	0	0	0	0	0	0	1	0	0	0	0	0
4	0.0833	0.3333	0	0.1667	0.0833	0.1667	0.0833	0	0.0833	0	0	0
5	0.3333	0.3333	0	0	0.3333	0	0	0	0	0	0	0
6	0	0	0.1667	0.5	0.1667	0.1667	0	0	0	0	0	0
7	0	0	0	0	0	0.5	0.5	0	0	0	0	0
8	0.1	0	0	0	0	0	0	0.6	0.3	0	0	0
9	0.4444	0	0.1111	0	0	0	0.1111	0.2222	0.1111	0	0	0
10	0	0	0	0	0	0	0	0	0	1	1	1
11	0	0	0	0	0	0	0	0	0	1	1	1
12	0	0	0	0	0	0	0	0	0	1	1	1

Fig. 5. Wavelet function for Flute



Mr.Thor, University of Sussex, England for the valuable insight given by him during the work.

REFERENCES

[1] J. A. Moorer, “Signal processing aspects of computer music A survey”, *Proceedings of the IEEE* ,vol. 65, no. 8.

[2] M.S.Sinith, Madhavi N. Nair, Niveditha P. Nair, S.Parvathy”Identification of Wavelets and Filter Bank Coefficients in Musical Instruments.” *IEEE International Conference on Audio Language and Image Processing (ICALIP 2010)*, Shanghai, China.

[3] Joseph .O. Chapa and Raghuveer M.Rao. Algorithms for designing wavelets to match a specified signal *IEEE Transactions on Signal Processing, vol.48, No.12, December 2000*

[4] Martin Vetterli and Cormac Harley. Wavelets and Filter Banks: Theory and Design. *IEEE Transactions on Signal Processing. Vol.40. No.9, September 1992.*

[5] Monson .H. Hayes, “Statistical Digital Signal Processing and Modelling”, *John Wiley&Sons, Inc.*

[6]] Markus Rupp, ”The Behaviour of LMS and NLMS Algorithms in the Presence of Spherically Invariant Process.”, *IEEE Transactions on Signal Processing.* Vol.41. No.3. March 1993.

[7]]Dirk T.M. Slock, “On the convergence Behaviour of LMS and Normalized LMS Algorithms”, *IEEE Transactions on Signal Processing* .Vol.41. No.9. September 1993.

[8] Moshe Tarrab and Arie Feuer. ” Convergence and Performance Analysis of the Normalized LMS Algorithm with Uncorrelated Gaussian Data.” *IEEE Transactions on Information Theory* Vol.34.No.4. July 1988.

[9] M.S.Sinith, K.Rajeev”Pattern Recognition in South Indian Classical Music using A Hybrid of HMM and DTW.” *IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2008)*, India.

[10] Gaurav Pandey, “Tansen:A system for automatic raga identification,” *Indian International conference on Artificial Intelligence 2003*, Hyderabad, India, 2003