International Journal of Research in Engineering, Science and Management
Volume 6, Issue 1, January 2023
https://www.ijresm.com | ISSN (Online): 2581-5792

Sound Analysis and Adaptive Synthesis in Time-Frequency Resolution

J. Hemalatha[1*], L. Papayee[2]

[1]*Associate Professor, Department of Electronics and Communication Engineering, Mookambigai College of Engineering, Kalamavur, India*
[2]*Assistant Professor, Department of Electronics and Communication Engineering, Mookambigai College of Engineering, Kalamavur, India*

***Abstract*: We propose a simple, efficient technique for sound analysis and synthesis with automatic adaptation of time-frequency resolution. The analysis and synthesis give a good approximation of the original signal from analysis with different time-varying resolutions with frequency bands using Mel- filter bank approach. Time-Frequency signal processing can be modeled in a formal mathematical framework. Here, mathematical frameworks are done for the analysis and synthesis of sound. The possibility of an adaptive analysis drastically limits the parameters to set and also without affecting the quality of sound. The input sound signal is analyzed and the same is synthesized with the best resolution and the reconstruction error is computed.**

***Keywords*: Time-Frequency analysis, adaptivity, entropy.**

1. Introduction

The time-frequency signal processing, sound processing can be modeled in mathematical framework. The basic components of complete analysis, transformation and, synthesis of a sound are,

a) The given sound signal can be decomposed by means of a given set of atoms, the resultant being a set of analysis coefficients.
b) These analysis coefficients are interpreted to deduce information about the original sound signal.
c) The analysis coefficients are modified to transform specific features of the representation.
d) A new sound signal is constructed as an expansion of the modified coefficients.

In general, sound visualization makes use of the first one alone. While feature extraction technique exploits the first two and the source separation or vocal transformations which are quite complicated exploit all of the above.

The main aim of this paper is to analyze and synthesize the sound efficiently. This paper is structured as follows: Section 2 presents the Time-Frequency Analysis. Section 3 contains the existing and proposed system analysis. Section 4 discussed about the main features of cepstral analysis. In section 5 is about the implementation of the MFCC algorithm. Further in section 6 the features of MFCC algorithm are discussed. Finally, section 7 presents our conclusions and future work.

2. Time-Frequency Analysis

In signal processing, time–frequency analysis is a body of techniques and methods used for characterizing and manipulating signals whose statistics vary in time, such as transient signals.

It is a generalization and refinement of Fourier analysis, for the case when the signal frequency characteristics are varying with time. Since many signals of interest – such as speech, music, images, and medical signals – have changing frequency characteristics, time–frequency analysis has broad scope of applications. Whereas the technique of the Fourier transform can be extended to obtain the frequency spectrum of any slowly growing locally integrable signal, this approach requires a complete description of the signal's behavior overall time [5]. Indeed, one can think of points in the (spectral) frequency domain as smearing together information from across the entire time domain. While mathematically elegant, such a technique is not appropriate for analyzing a signal with indeterminate future behavior [7].

To harness the power of a frequency representation without the need of a complete characterization in the time domain, one first obtains a time–frequency distribution of the signal, which represents the signal in both the time and frequency domains simultaneously. In such a representation the frequency domain will only reflect the behavior of a temporally localized version of the signal.

A. Adaptive Analysis

In time-frequency analysis, adaptivity is the possibility to conceive representations and operators whose characteristics can be modeled according to their input [13]. The adaptivity as the possibility to deal with different resolutions locally within a sound; then, a criterion to choose the best local resolution which provides for the adapted representation; and finally, the possibility to define a reconstruction method from the adapted analysis. The concept of adaptivity is closely related with the one of sparsity: an adaptive analysis must give a sparse representation of the signal, according to specific measures to be optimized, the optimal resolution being signal- and application-dependent. This is a highly prolific approach.

The analysis methods can be divided in two principal classes: parametric methods and non-parametric methods. The

*Corresponding author: hemabalan2001@gmail.com

parametric methods require a priori knowledge of the signal, and consist in adjusting the parameters of a model. The non-parametric models do not need any knowledge of the signal to be analyzed, but they often require a larger number of coefficients. In time-frequency analysis, adaptivity which is the possibility to conceive the representations and operates whose characteristics can be modeled according to their input.

3. System Analysis

A. Existing System

In the past for analyzing the signal, offer limited possibilities concerning the flexibility of their time-frequency precision [11]. Moreover, fundamental analysis parameters have to be set a priori, according to the signal characteristics and the quality of the representation required. Analyses with a non-optimal resolution lead to a blurring, or sometimes even a loss of information about the original signal, which affects every kind of later treatment.

B. Proposed System

In proposed system a novel approach, the adaptivity is the possibility to deal with different resolutions locally within a sound. A criterion to choose the best local resolution which provides for the adapted representation. The possibility to define a reconstruction method from the adapted analysis using the MEL – Filter Bank Approach with the parameters defined for the selection of coefficients.

4. Automatic Adaptation

The analysis parameters are the window size and oversampling varying which we analyze the sound. The adaptation parameters are entropy order, time-frequency region R and time-frequency step for region R.

The following fig. 1 shows the graphical representation of automatic adaptation.

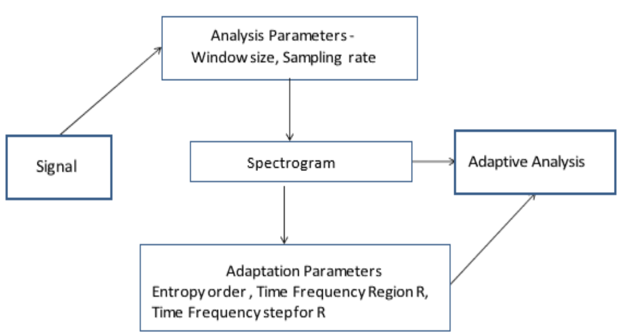

Fig. 1. Graphical representation of automatic adaptation

A. Filter Bank Approach

A novel approximation method using filter bank approach as shown in the fig. 2 is based on analyses with resolution changing in time and frequency, indicating theoretical bounds for the reconstruction error.

B. Coefficients Selection

The section of coefficients selection depends on the following parameters.

- Renyi's Entropy order.
- Time – Frequency Region.
- Time –Frequency step for the shifted R.

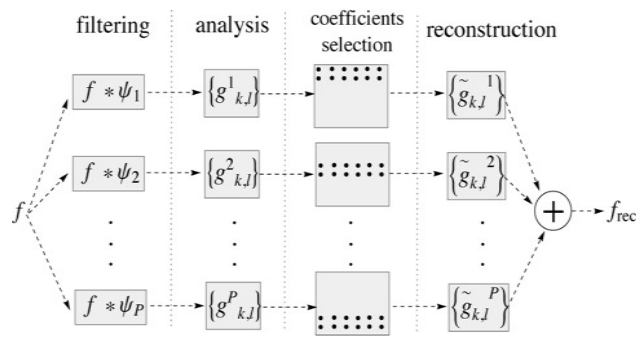

Fig. 2. Filter bank approach

C. Time Frequency Region

Fig. 3 shows the time frequency centers for a stationary Gabor frame G(g,a,b) with time and frequency step (a,b) respectively. Fig. 4 shows the time frequency for a stationary Gabor frame with a shift R.

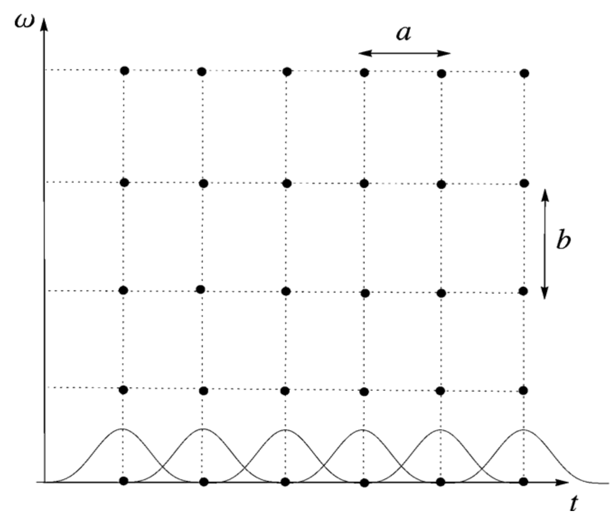

Fig. 3. Time frequency center for a stationary Gabor frame

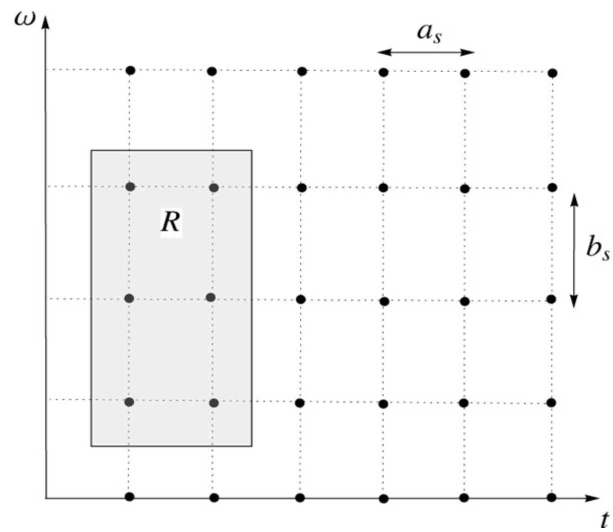

Fig. 4. Time frequency for a stationary gabor frame with a shift R

For a given spectrogram Ps and a time- frequency shift of the rectangle, the Rényi entropy of the coefficients within R is calculated.

5. Cepstrum Analysis

Speech is composed of excitation source and vocal tract system components. In order to analyze and model the excitation and system components of the speech independently and also use that in various speech processing applications, these two components have to be separated from the speech [9]. The objective of cepstral analysis is to separate the speech into its source and system components without any a priori knowledge about source and/or system.

Cepstrum Analysis is widely applied in signal processing. It is particularly useful in speech synthesis, speech recognition and, other audio applications.

A. Basic Principles of Cepstral Analysis

The block diagram representing the computation of cepstrum is given in the following fig. 5.

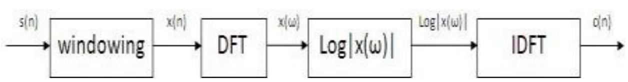

Fig. 5. Computation of cepstrum

The log operation transforms the magnitude speech spectrum where the excitation component and vocal tract component are multiplied, to a linear combination (summation) of these components i.e. log operation converted the "*" operation into "+" operation in the frequency domain. The separation can be done by taking the inverse discrete fourier transform (IDFT) of the linearly combined log spectra of excitation and vocal tract system components. It should be noted that IDFT of linear spectra transforms back to the time domain but the IDFT of log spectra transforms to quefrency domain or the cepstral domain which is similar to time domain.

B. Cepstrum

There are 2 basic types of cepstrum.

- Real cepstrum
- Complex Cepstrum

The cepstrum analysis discussed above falls under the real cepstrum category. As the real cepstrum is computed from the log magnitude spectrum, the phase part is ignored. This will not enable the reconstruction of the sequence from the cepstrum. However, the reconstruction can be done by preserving the Fourier phase and use it for reconstruction from the real cepstrum [15]. For the reconstruction of the sequence from the cepstrum, complex cepstrum is used. Instead of taking inverse Fourier transform of the log magnitude spectrum for the real cepstrum, the inverse Fourier transform of the logarithm of complex spectrum is used for computing complex cepstrum. As the logarithm of all the spectral values are used, the phase is preserved in the complex cepstral sequence which can be used for reconstructing back the sequence [19]. The methods for computing pitch and formant parameters from the complex cepstrum remain same as that of the real cepstrum as these parameters are obtained from the magnitude of the complex cepstral coefficients.

6. Mel Frequency Cepstral Coefficient (MFCC)

For speech recognition, the most commonly used acoustic features are MFCC. MFCC takes human perception sensitivity with respect to frequencies into consideration, and therefore are best for speech recognition.

A. MFCC Algorithm

1. Frame the signal into short frames.
2. For each frame calculate the periodogram estimate of the power spectrum.
3. Apply the Mel filter bank to the power spectra; sum the energy in each filter.
4. Take the logarithm of all filter bank energies.
5. Take the DCT of the log filter bank energies.

7. Experimental Results

Spectrogram of the FFT result and the reconstruction error given by the analysis- weight approach, on a sound sample are given as follows in fig. 6 and fig. 7.

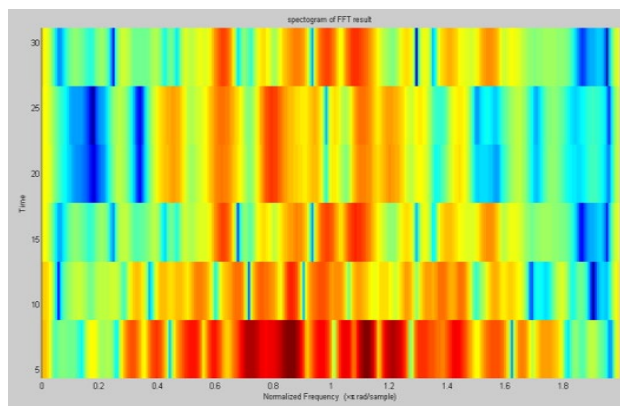

Fig. 6. Spectrogram of FFT result

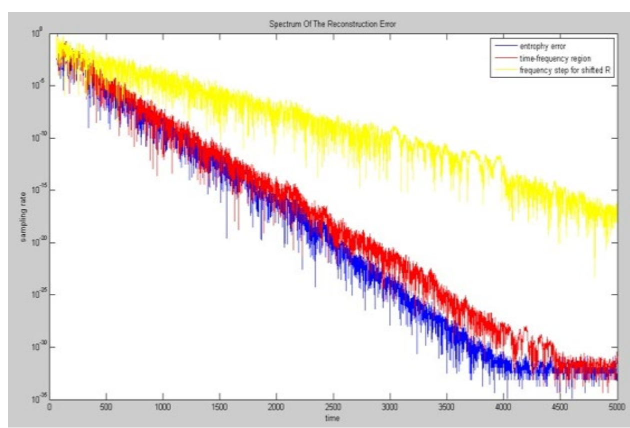

Fig. 7. Spectrum of the reconstruction error

8. Conclusion and Future Work

In this paper we introduced a technique for sound analysis and synthesis with local automatic adaptation of time-frequency resolution which is intended to provide a signal representation with optimal local time-frequency information with a high quality of analysis and synthesis process. The optimal local time- frequency resolution guarantees a solid ground to develop adaptive high-quality applications: sound object localization and separation with adapted time-frequency precision, as well as information retrieval with optimal local resolution, among several others.

The same can be extended for speaker recognition and verification which remains as the future work of this paper.

References

[1] Alexandra Carciun, Martin Spiertz, Adaptive Time Frequency Resolution for Blind Source Separation, IEEE Transactions on Signal Processing, Vol. 30, 2010.
[2] Balazs P, Jaillet, Theory, Implementation and applications of non-stationary Gabor frames, Elsevier Journal of Computational and Applied Mathematics, 2011.
[3] Baranuik Richard. R., Douglas. L. Jones, Measuring Time–Frequency Information Content Using the Rényi Entropies, IEEE Transactions on Information Theory, Vol. 47, No. 4, May 2001, pp. 435-441.
[4] Bruno Torresani, Time-Frequency Jigsaw Puzzle: Adaptive multiwindow and multilayered Gabor expansions, IEEE Transactions on Signal Processing, 2008.
[5] Davis Mike E, Sparse Representations in Audio and Music: From Coding to Source Separation, IEEE Transactions on Signal Processing, 2011.
[6] Dorfler M, Bruno Torresani, Representation of Operators in the Time-Frequency Domain and Generalized Gabor Multipliers, IEEE Transactions on Information Theory, 2008.
[7] Daniel Rudouy, Superposition Frames for Adaptive Time-Frequency Analysis and Fast Reconstruction, IEEE Transactions on Signal Processing, Vol. 58, No. 5, May 2010.
[8] Ewa Matusiak, Representation of Operators In The Time-Frequency Domain And Generalized GaborMultipliers, IEEE Transactions on Signal Processing, August 2012.
[9] Feichtinger H. G, Gabor Analysis and Algorithms—Theory and Applications, IEEE Int. Conf. Acoust. Speech Signal Process., 2011.
[10] Jones. Douglas L., Richard G. Baraniuk, A Simple Scheme for Adapting Time-Frequency Representations, IEEE Transactions on Signal Processing, Vol. 42, No. 12, December 1994.
[11] Jones Douglas L., Richard G. Baraniuk, An adaptive optimal kernel time frequency representation, IEEE Transactions on Signal Processing, Vol. 43, No. 10, October 1995.
[12] Jaillet Torresani, Non-stationary Gabor Frames, IEEE Transactions on Signal Processing, August 2009, pp. 543-551.
[13] Monika Dorfler, Prabahan Basu, Gabor Analysis for a Class of Signals called Music, IEEE Transactions on Signal Processing, Vol. 58, May 2010, pp. 431-439.
[14] Ladar Yonina C., Sub-Nyquist Sampling of Short Pulses, Yonina C. Eldar IEEE Transactions on Signal Processing, Vol. 60, No. 3, March 2009.
[15] Patrick J. Wolfe, Augustus J. E. M. Janssen, Adaptive short-time analysis-synthesis for speech enhancement, Multi-Gabor dictionaries for audio time-frequency analysis," in Proc. IEEE Workshop Appl. Signal Processing, 2010.
[16] http://www.cs.tut.fi/courses/SGN- 4010/ikkunointi_en.pdf
[17] http://mtg.upf.edu/mosart/papers/p40.pdf
[18] www.mathworks.in/help/matlab
[19] http://djj.ee.ntu.edu.tw/Tutorial_Wavelet%20for%20Music%20Signal%20Analysis%20_2.pdf
[20] https://www.math.ucdavis.edu/~strohmer/research/gabor/gaborintro/node4.html
[21] http://iitg.vlab.co.in/?sub=59&brch=164&sim=615&cnt=1