APPLICATION ORIENTED INSIGHTS INTO THE GABOR TRANSFORM FOR ACOUSTIC SIGNALS PROCESSING

Ewa £ukasik
Institute of Computing Science, Poznañ University of Technology
60-965 Poznañ, Piotrowo 3a, Poland,
Tel: +48 61 782373; fax: +48 61 771525
e-mail: LUKASIK@POZN1V.TUP.EDU.PL

ABSTRACT

The paper presents results of analysis of certain quasi stationary and non stationary signals using Gabor transform and Gabor spectrogram. Initial results are based on the original programs realising Gabor transform, whereas the main part of the work - the comparative analysis of signals by Gabor spectrograms of higher orders and other time-frequency distributions was performed using the commercially available software package: Joint Time Frequency Analysis (JTFA) Toolkit from National Instruments. In most cases the experiments showed superiority of Gabor spectrogram over other methods mainly due to better time and frequency resolution and elimination of some of the cross terms inherent e.g. for Wigner-Ville or Choi-Williams Distributions.

1. INTRODUCTION

Gabor Transform has been known since 1946 [1] but only lately has found recognition in various branches of signal processing. The signal can be reconstructed from the sliding window spectrum, when we know its values at the points of a certain time-frequency lattice. Gabor's signal representation expresses the signal as a superposition of properly shifted and modulated versions of the window. Shie Quian and Dapang Chen evaluated the efficient algorithms for finding the reciprocal analysis and synthesis windows pair for sliding window signal representation and computing the Gabor coefficients of infinite time series [3-5]. They also gave the theoretical background to the new tool for time-frequency analysis - Gabor spectrogram, which is based on the Wigner-Ville representation of Gabor coefficients [6]. It is commercially available as a Joint Time-Frequency Analyser (JTFA), a part of National Instruments graphical programming package LabView. However up to now very little reports have been published on the actual applications and results of using this new tool for signal analysis.

The aim of this work is to perform the comparative time-frequency analysis of various non-stationary acoustical signals using the JTFA package with the emphasis on Gabor Spectrogram. In order to get better understanding of its features certain practical insights into the bases of the method were performed following the works [3-6].

2. DISCRETE GABOR EXPANSION AND GABOR SPECTROGRAM

2.1. Definitions

The Gabor Transform pair is defined as:

$$s(i) = \sum_{m=0}^{\infty} \sum_{n=0}^{N-1} C_{m,n} h(i - m\Delta M) \cdot exp(j2\pi in / N)$$

$$C_{m,n} = \sum_{i=0}^{\infty} s(i) \gamma^*(i - m\Delta M) \cdot exp(-j2\pi in / N) \quad (1)$$

s(i) - time series, $C_{m,n}$ - Gabor coefficient, m - time index, n - frequency index, *h(i)* - synthesis window, $\gamma^*(i)$ - analysis window, N - number of frequency indices, ΔM,. ΔN - sampling intervals in time and frequency domains.

According to [6] for a discrete normalised Gaussian signal it is possible to precompute interference free auto Wigner-Ville Distribution and apply it to the Gabor coefficients representing the signal. In this way cross terms free Gabor spectrogram of zero order is received:

$$GS(i,k) = \sum_{m=0}^{\infty} \sum_{n=0}^{N/2-1} C_{m,n}^2 \, exp\{-[(i - m\Delta M)^2 / \sigma^2\} \cdot \quad (2)$$
$$\cdot exp\{(2\pi\sigma / L)^2 (k - n\Delta N)^2]\}$$

L- window length, $0 \le i \le \infty$, $0 \le k < L/2 - 1$.

The above equations were implemented for the initial calculations of Gabor spectrogram for critical sampling, as well as double and quadruple oversampling and Gaussian window of the length 128 and 256. The results are presented in next subsection.

2.2. Practical results

The main problem is to find $\gamma^*(i)$ biorthogonal and close in shape to h(i). In our case it is standard Gaussian function. Its shape depends on the sampling pattern. For oversampling it is better localised, whereas for critical sampling it is localised either in time or in frequency. Fig. 1. represents analysis and synthesis windows for infinite data, double oversampling and critical sampling, computed according to [3,4,5]. The resulting pairs are in accordance with those predicted theoretically by Bastiaans in [2]. However other shapes of the windows are still under the consideration, see recently published work of B.Friedlander and A. Zeira [9] where the one-sided exponential window was successfully applied for detection and analysis of transient signals in oversampled cases and [10].

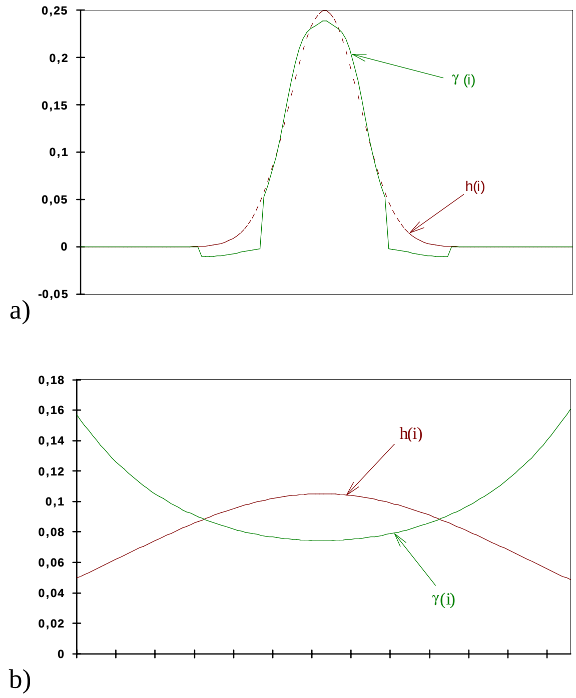

Fig.1. Examples of reciprocal optimal Gaussian windows of the length L=128 for infinite series a) double oversampling, b) critical sampling

We examined in more details the performance of the Gaussian window pairs for various samplings model, as they were used in NI JTFA toolkit [11]. The reconstruction error (MSE) was computed for various random and acoustical signals. Results for some of the signals are presented in table 1.

The MSE parameter is important not only for signal analysis, but also if signal coding using Gabor coefficients would be considered.

Since the Gabor expansion algorithm has been claimed as computationally effective, we realised it using the standard FFT and also WFTA algorithm [7]. The computation time was satisfactory, however further calculations of Gabor spectrogram implementing the cross-terms deleted Wigner Ville transform directly to Gabor coefficients using (2) was time consuming.

The obtained spectrogram resolution in frequency domain was far from ideal [8]. In fact the same effect was also reported by the authors of the method and led them to design a higher order Gabor spectrograms implementing certain of cross-terms rejected previously [6]. Those solutions are implemented in JTFA toolkit [11].

3. JTFA TOOLKIT OF NI LabView

The Joint Time-Frequency Analyser (JTFA) [11] allows to use any of six joint time-frequency algorithms to analyse stored data files as well as data from DAQ. Those are:

- Wigner-Ville Distribution
- Cone-Shaped Distribution
- Choi-Williams Distribution
- Short-Time Fourier Transform
- Gabor Spectrogram
- Adaptive Spectrogram

The last one adjusts the variance, time and frequency centres of the Gaussian functions to best match the analysed signal. The references for the above distributions are given in [11].

The analyser enables processing data stored in ASCII files, apart from spectrograms displays the waveform and spectrum or instant spectrum (a kind of disadvantage is that the instant spectrum is displayed only in linear scale). The signal block length can be chosen deliberately.

4. EXPERIMENTS RESULTS

Our initial experiments with Gabor Spectrogram of zero order [8] showed, that the frequency resolution is rather poor, comparable with STFT. Higher order Gabor

Reconstruction MSE for various sampling models

Sampling model Data	Quadruple $\Delta N=4$, $\Delta M=8$ L=128	Double $\Delta N=4$, $\Delta M=16$ L=128	Critical $\Delta N=1$, $\Delta M=128$ L=128
speech	$1.30 \cdot 10^{-6}$	$4.1 \cdot 10^{-6}$	$0.93 \cdot 10^{-12}$
orchestra	$1.10 \cdot 10^{-6}$	$3.4 \cdot 10^{-6}$	$0.84 \cdot 10^{-12}$
rock music	$1.20 \cdot 10^{-6}$	$3.5 \cdot 10^{-6}$	$0.85 \cdot 10^{-12}$
bird’s call	$1.80 \cdot 10^{-6}$	$4.5 \cdot 10^{-6}$	$1.40 \cdot 10^{-12}$
random noise	$4.61 \cdot 10^{-6}$	$14.4 \cdot 10^{-6}$	$3.10 \cdot 10^{-12}$

spectrogram seemed to be superior. We have prepared the initial database with various acoustical signals including speech, string instruments sounds, symphony and rock orchestra, birds calls and other nature sounds with intensive non-stationary features. The samples were captured using Volyzer - real time signal analyser built in the home Institute of the author [12]. From Volyzer's format they were converted into ASCII. The sampling frequency was 16.8 kHz.

Fig. 2. and 3. represent spectrograms of some of the registered signals. Those are the and bird's call guitar sound. The features of various distributions in time-frequency domain are quite different. The Wigner-Ville and Choi Williams distribution based spectrograms are difficult to read because of their cross terms. For some applications the adaptive and Cone-Shaped distributions would be advantageous since they smooth the cross terms and enhance the spectrum in the frequency direction while maintaining finite time support in the time direction. They rather show the enhanced energy concentration of the components of the signal. Their performance could be comparable with the lateral inhibition effect. However the resulting time-frequency distribution of fast changing features might be rather difficult to interpret (see the guitar sound). Gabor spectrograms of 3rd, 4th orders, including some of the cross-terms of the Wigner-Ville distribution seem to give best resolutions in time and in frequency. However one needs some practice in interpreting those spectrograms. For the lack of space the speech example is not presented here.

5. CONCLUSIONS

The main task of this work was experiencing the Gabor spectrogram, claimed to be best suited for non stationary signals analysis. The extensive comparative experiments were possible thanks to the use of the JTFA package of NI LabView - the first commercial tool for joint time frequency analysis using various methods of computing time frequency distributions including Short-Time Fourier Transform, Wigner-Ville, Cone-Shaped, Choi-Williams Gabor and Adaptive Distributions.

The experiments were performed on the database of non-stationary and quasi stationary signals using all of above methods. They showed that in multiple cases higher order Gabor Spectrogram (especially order 3 and 4) is indeed superior to other methods, give better resolution than traditional STFT and eliminate the cross terms characteristic for Wigner Ville and Choi Williams distribution.

However one practical remark should be made here: there is too little details known about the parameters of the time frequency distributions used in JTFA toolkit, e.g. those concerning window length and its exact shape, sampling model etc. They would be useful for more precise interpretation of the results. Despite those disadvantages further experimental work is planned for various signals, including speech, especially the automatic feature extraction using Gabor representation.

6. REFERENCES

1. D. Gabor, "Theory of communication", J. IEE, vol.93, No III, 1946, pp. 429-457
2. M.J. Bastiaans, "On the sliding window representation in Digital Signal Processing", IEEE Trans. on ASSP, No 4, Aug. 1985, pp. 868-873.
3. S. Qian, D. Chen, "Discrete Gabor Transform", IEEE Trans. on S.P. July 1993
4. S. Qian, D. Chen, "Orthogonal - Like Discrete Gabor Expansion", 26th Conference on Information Sciences and Systems, Princeton Univ., March 1992.
5. S.Qian, D.Chen, "Optimal biorthogonal analysis window functions for Discrete Gabor Transform", IEEE Trans. Signal Proc., vol 42, no 7, 1994, pp.694-697
6. S.Quian, D.Chen, "Time-frequency Distribution Series", Proc. ICASSP-94, vol.III., pp.29-32
7. E. £ukasik, R. Stasiñski, Certain aspects of fixed-point FFT and WFTA rounding error comparison, Proc. EUSIPCO 92, pp.941-944.
8. E. £ukasik: "Comments on Using Gabor Spectrogram for Signal Analysis", Proc. XVIII KKTOiUE, Polana Zgorzelisko, pp. 509-514, 1995.
9. B.Friedlander, A.Zeira, "Oversampled Gabor Representation for Transient Signals", IEEE Trans. Signal Proc. vol.43, no 9, 1995, pp. 2088-2094
10. S.Qiu, H.G.Fechtinger, "Discrete Gabor Structures and Optimal Representations", IEEE. Trans. Signal Proc., vol. 43, no 10, pp.2258-2268
11. "Joint Time-Frequency Analysis Toolkit Reference Manual", National Instruments Corporation, March 1995
12. S.Grocholewski, E.£ukasik, J.Ogórkiewicz, "Low Cost Real Time Digital Spectrograph", Proc. ICDSPAT, Berlin 1991, pp. 590-597

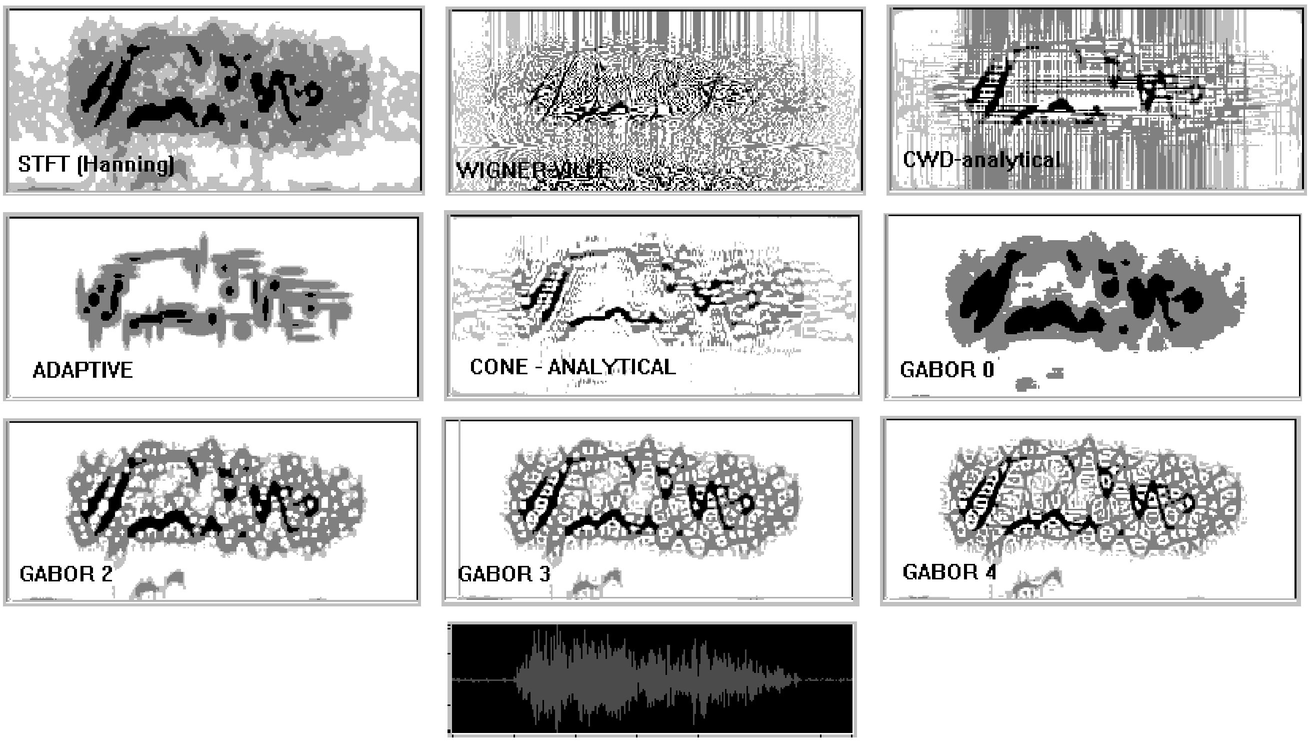

Fig. 2. Various spectrograms from NI JTFA of the bird's call and the waveform (80 ms, sampl. freq. 16.8 kHz)

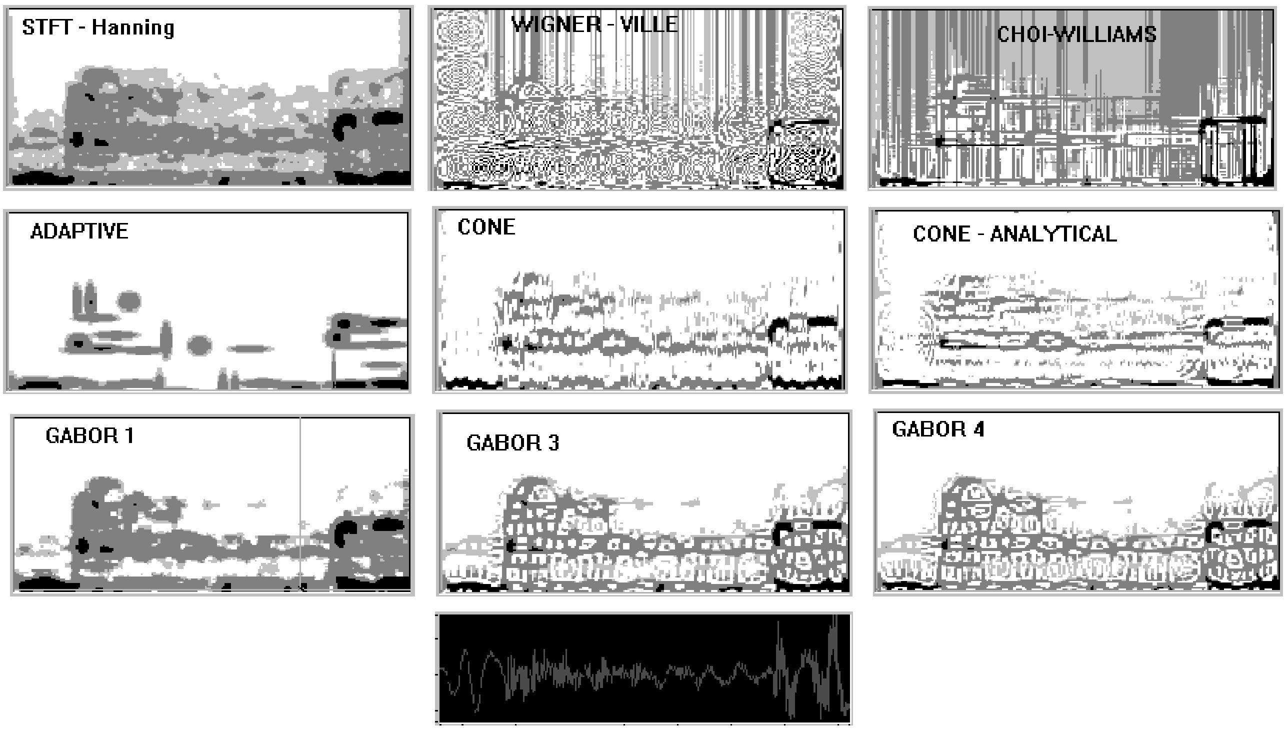

Fig.3 Various spectrograms of the guitar sound from NI JTFA and the waveform (50 ms, sampl. freq. 16.8 kHz)