USING PHASE OF SHORT-TERM FOURIER TRANSFORM FOR EVALUATION OF SPECTROGRAM PERFORMANCE

KRZYSZTOF CZARNECKI[*/**], MAREK MOSZYNSKI[*], ANDRZEJ CHYBICKI[*]

[*]Gdansk University of Technology, Department of Geoinformatics
11/12, G. Narutowicza St., 80-952 Gdansk, Poland
[**]Medical University of Gdansk, Department of Informatics and Statistics
3a, M. Skłodowskiej-Curie St., Gdansk, Poland
czarnecki.engineer@gmail.com

The concept of spectrogram performance evaluation which exploits information on phase of short-term Fourier transform (STFT) is presented. A spectrograph which is a time-frequency analyzing tool, is compared to a filter bank that demultiplexes a signal. Local group delay (LGD) and channelized instantaneous frequency (CIF) is obtained for each filtered component signal. In presented solution the performance is evaluated using so-called length of weighted average of reassignment vectors (WARV length). Orthogonal components of each reassignment vector are calculated using above mentioned parameters.

INTRODUCTION

One of the most commonly used methods for time-frequency analysis of non-stationary signals is the spectrograph. It is used for analysis and processing of various types of signals including telecommunications, acoustics, mechanics, seismic, biomedical and speech. To calculate spectrogram the well known Gabor type short-term Fourier transformation is used [1]. In this case the STFT is considered to be an energy distribution on the time-frequency surface. In overall, the Heisenberg-Gabor rule says that the ambiguity of spectrogram in one dimension is inversely proportional to ambiguity in second dimension. This makes impossible to achieve maximal ability of details separation in both time and frequency dimensions simultaneously. This dependency is strictly related to the window width of spectrograph.

To solve this problem, some methods for energy concentration evaluation were presented in [2], where the energy concentration is used to evaluate the ability of details separation, similarly as the image sharpness helps to take stock of image performance. The presented concept was based on log magnitude of STFT only, however it was shown that optimal windows width for maximal energy concentration can be estimated. Moreover, proposed spectrogram performance evaluation parameter, unlike energy concentration index [2], is based on processing STFT phases considered as mean reassignment vector lengths.

The concept of reassignment vector originates from work of Kunikiko Kodera [3] in 1978. The author introduced 'The Modified Moving Window Method' that was used for study of reassign spectrogram by other authors like: A. Rihaczek [4], P. Flandrin, F. Auger [5], D. Nelson [6], S. Fulop, K. Fitz [7] and D. Friedman [8].

In general, the short-term Fourier transform of a discrete signal using a filter bank, is defined as:

$$X[l,k] = STFT\{x[n], w[m]\} = \sum_{m=0}^{M-1} x[l+m]w[-m+\lfloor M/2 \rfloor]\mathrm{e}^{-j2\pi mk/K}\ , \tag{1}$$

where $X\,[l,k]$ denotes a complex value of short-term Fourier transform in the point $[l, k]$, $w[m]$ represents the M-samples wide spectrograph window. The symbol $\lfloor\rfloor$ represents the rounding a real number down to the closest integer. Sizes of the transform are marked by L and K.

In following section of the paper the definition of the reassignment vector using phasors of STFT points was described. The new parameter of spectrogram performance evaluation (mean of reassignment vectors length) was introduced in section 3. Results of test signal analysis, time-frequency images and conclusions were presented in the summary.

1. REASSIGNMENT VECTOR

The reassignment vector located on time-frequency surface in point $[l, k]$ is constructed using two orthogonal components (time and frequency):

$$\vec{V}[l,k] = [V_t[l,k], V_{\mathrm{f}}[l,k]] \tag{2}$$

Both components can be calculated using so-called interframe method or so-called intraframe method. In this paper the interframe method is applied and 3-point estimators of LGD and CIF are used. Thus for each non-marginal element of STFT can be obtained only one reassignment vector. The reassignment vector component of time can be defined as LGD [9]:

$$V_{\mathrm{t}}[l,k] = \mathrm{LGD}[l,k] = \frac{-K}{4F_s}(\arg(-X[l,k+1]X^*[l,k]) + \arg(-X[l,k]X^*[l,k-1])) \tag{3}$$

where arg( ) denotes the argument of complex number, X^* marks the conjugation of complex number and F_{s} is the sampling frequency. Respectively, the reassignment vector component of frequency can be defined as:

$$V_{\mathrm{f}}[l,k] = \mathrm{CIF}[l,k] - 2\pi k/K \tag{4}$$

where CIF is the channelized instantaneous frequency which can be introduced by the following formula:

$$\mathrm{CIF}[l,k] = (\arg(X[l+1,k]X^*[l,k]) + \arg(X[l,k]X^*[l-1,k]))/2 \tag{5}$$

CIF shows a new position of STFT element on frequency axis.

The definitions introduced by formulas (3) and (5) are modified as compared to suggestions of Kodera [3]. Modifications were introduced in order to convert these components to common unit. Consequently, it is possible to compare both components and obtain a normalized length of the reassignment vector.

2. SPECTROGRAM PERFORMANCE EVALUATION

According to (3) and (4), the definition of two dimensional vector field can be derived. This definition is useful during the calculation of a reassignment spectrogram for STFT energy redistribution in time dimension and frequency dimension. This procedure is applied to make the spectrogram independent of the Heisenberg-Gabor uncertainty rule. In most of the cases it results in better energy concentration. However, it is worth mentioning that only both factors: amplitude and phase, contain the full information about the signal.

Unfortunately, during signal transformation, energy of various independent signal components are usually dispersed and mixed. That results in arising errors and uncertainties during estimation of reassignment vectors. The calculated vectors do not point to right localizations of spectrogram energy but to some resultant position of all signal components that energy is contained in a considered spectral line of the spectrogram.

In presented approach of image performance evaluation, it is assumed that if energy in spectrogram is more concentrated, an average length of reassignment vectors should be shorter. In order to decrease the influence of low-energy signal components (usually generated by noise) the weighted average is introduced. The weight should be proportional to the energy of STFT which is linked with the respective vector. The formula of the weighted average of reassignment vectors (WARV) length is defined as follows:

$$\overline{V} = \frac{1}{S}\sum_{l=0}^{L-1}\sum_{k=0}^{K-1}\left\|\vec{V}[l,k]\right\| \cdot \left|X[l,k]\right|^2 , \tag{6}$$

where

$$S = \sum_{l=0}^{L-1}\sum_{k=0}^{K-1}\left|X[l,k]\right|^2 \tag{7}$$

denotes total energy which is contained in the spectrogram, $\left\|\vec{V}\right\|$ means a length of reassignment vector and $\left|X\right|$ denotes absolute value of complex number (of STFT element).

3. EXPERIMENT DESCRIPTION AND SUMMARY

The test signal x_t was constructed in order to show potential application of using WARV length and is defined as:

$$x_t = \begin{cases} c_1[n] + c_2[n] & \text{for } n = 1,2,\ldots,950 \\ c_1[n] + g[n]c_2[n] & \text{for } n = 951,\ldots,1000 \\ c_1[n] & \text{for } n = 1001,\ldots,2000 \end{cases} , \tag{8}$$

where

$$c_1[n] = \exp(j\pi(0.5n^2 - n)), \quad (9)$$

$$c_2[n] = \exp(j\pi(0.9 + 0.1F_s^{\ 4} / (n - 1630)^3)), \quad (10)$$

$$g[n] = 0.5 + 0.5\cos(0.02\pi(n - 950)). \quad (11)$$

The dependence of the WARV length as a function of the window width for the test signal (8) were calculated for various window types (Fig. 1). The signal was analyzed using transform defined by Eq. (1) with rectangular, Hamming and Blackman-Harris windows respectively. Resulted spectrograms of test signal are shown in Fig. 2-4.

The shortest WARV length was obtained for the 84-samples wide Hamming window. This is the window which guarantees the best ability for distinguishing details of a classical spectrogram in both dimensions simultaneously. The same result was achieved in [2], where the same test signal was analyzed using the energy concentration index. It is in line with conclusions of [2] and suggests that defined WARV length is a reliable spectrogram performance evaluation.

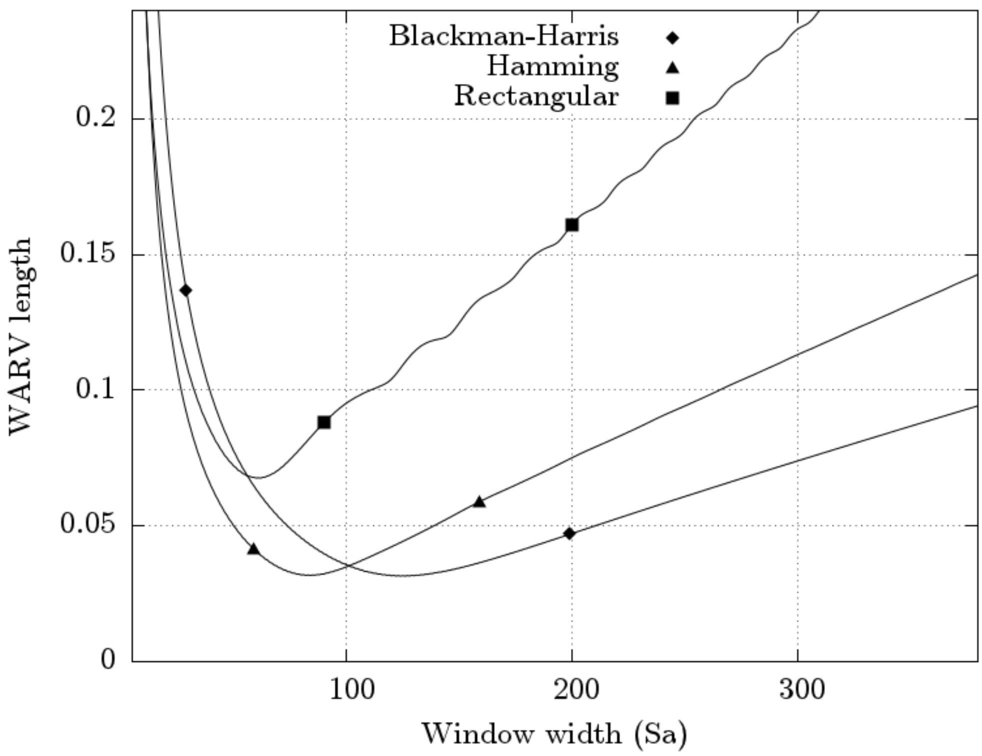

Fig. 1. WARV length curves as a function of the window width

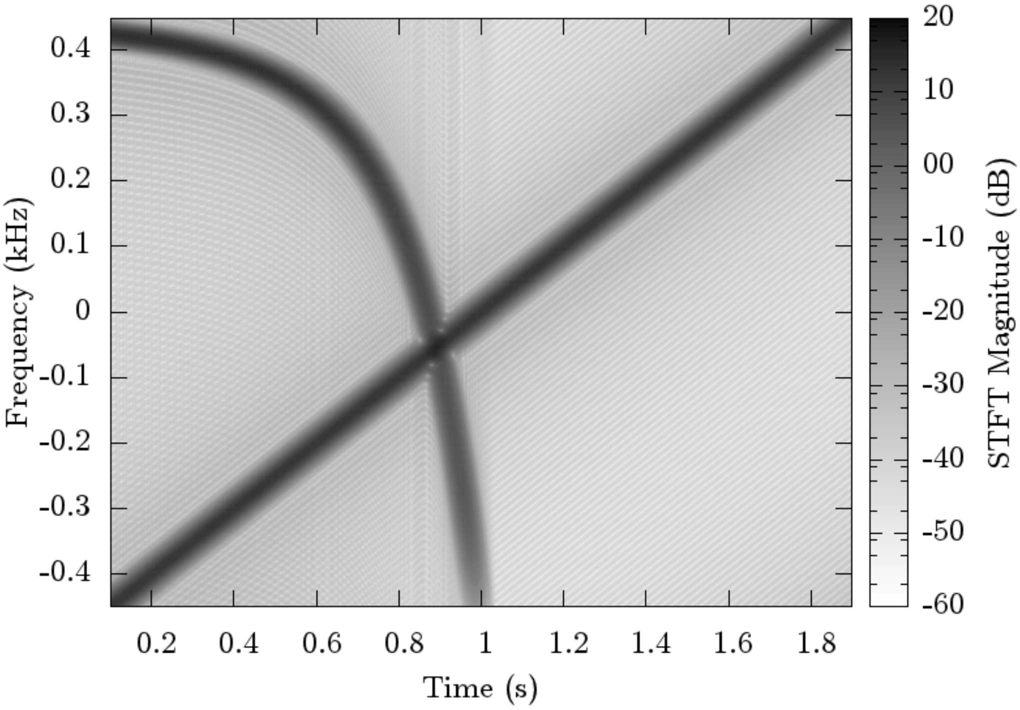

Fig. 2. Spectrogram of test signal (8) using 84-samples wide Hamming window

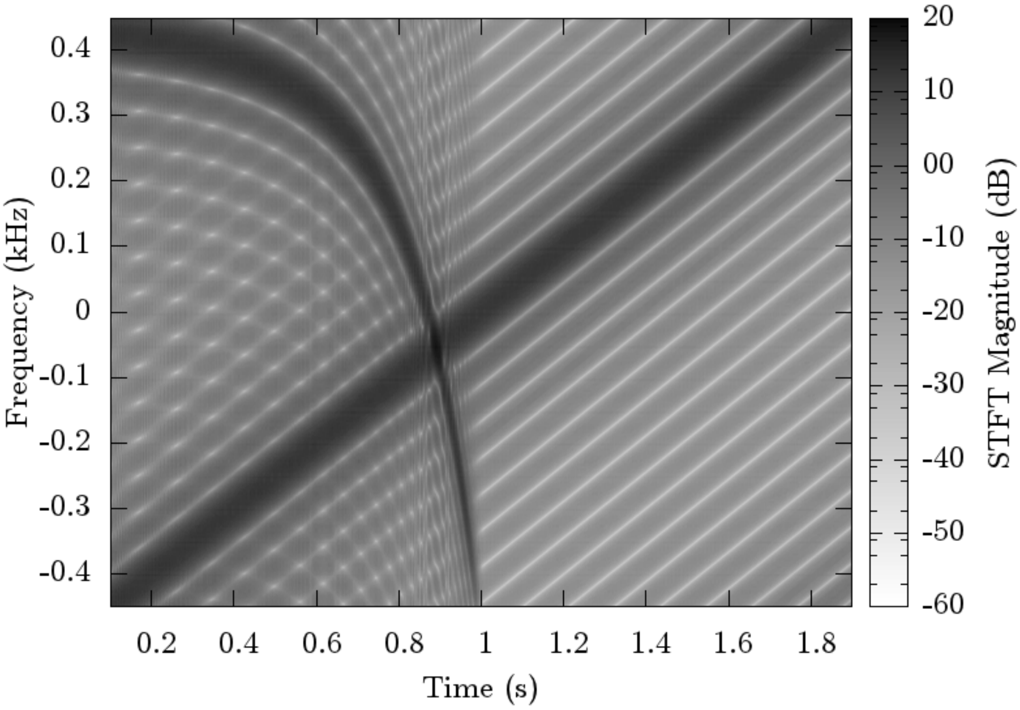

Fig. 3. Spectrogram of test signal (8) using 18-samples wide Rectangular window

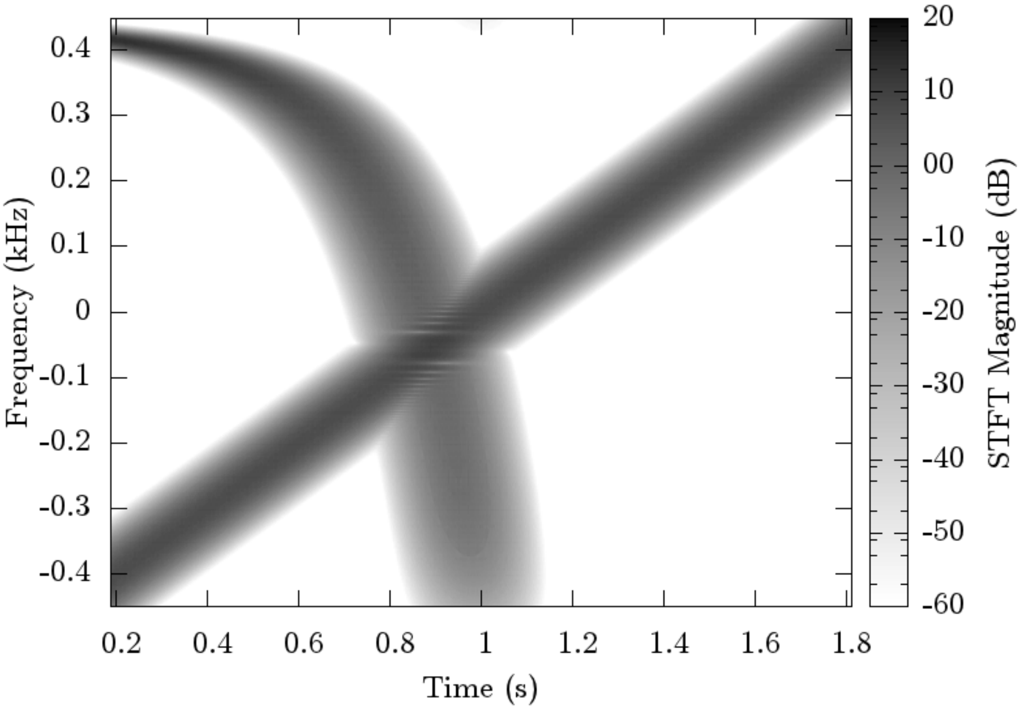

Fig. 4. Spectrogram of test signal (8) using 373-samples wide Blackman-Harris window

REFERENCES

[1] Gabor D., Theory of Communication, J. IEE (London), Vol. 93, November, 1946, pp. 429.
[2] Czarnecki K., Moszyński M., Rojewski M., An Objective Focussing Measure for Acoustically Obtained Images, Proceedings of 31st International Acoustical Imaging Symposium, Warsaw, April 2011.
[3] Kodera K., Gendrin R., Villedary C., Analysis of Time-Varying Signals with Small BT Values, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-26, No 1, February 1978.
[4] Rihaczek A., Signal Energy Distribution in Time and Frequency, IEEE Transaction on Information Theory, Vol. IT-14, No 3, May 1968.
[5] Auger F., Flandrin P., Improving the Readability of Time-frequency and Time- scale Representations by the Reassignment Method, IEEE Transaction on Signal Processing, Vol. SP-43, No 5, 1995.
[6] Nelson D., Cross-spectral Methods for Processing Speech. J. Acoust. Soc. Am., Vol. 110, November 2001.
[7] Fulop S., Fitz K., Separation of Components from Impulses in Reassigned Spectrograms, J. Acoust. Soc. Am., Vol. 121, March 2007.
[8] Friedman D., Instantaneous-frequency Distribution vs. Time: An Interpretation of the Phase Structure of Speech, IEEE Transactions on Acoustics, Speech and Signal Processing, 1985.
[9] Rojewski M., Unpublished notes, 2010.