

Efficient Spectral Estimation of Non-Stationary Harmonic Signals Using Harmonic Transform

Michal Trzos and Hasan Khaddour

Abstract—Non-stationary harmonic signals cannot be accurately represented by Fourier transform (FT). Fortunately, several methods for representing non-stationary harmonic signals exist including Fan-Chirp transform (FChT) or Harmonic transform (HT). This article is focused on the Harmonic transform and its computation. Estimation of slope of fundamental frequency change between analysed segments is essential for computation of the HT. The slope is estimated using several realisations of HT and comparing the spectral flatness. Optimisation of this procedure is presented in the article.

Keywords—Harmonic transform, non-stationary harmonic signals, time-frequency representations.

I. INTRODUCTION

The FT is generally able to represent frequency content of a signal, when the signal is composed of components with invariant frequency. Such signals can be called stationary harmonic signals and by using FT we can get their frequency representation with sufficient resolution in a specified frequency band. The ability of the FT to represent frequency content of a signal diminishes if the signals contains components with varying frequency [1], [2].

One solution of this problem is to use Warped Fourier Transform (WFT) [3], where the signal is frequency warped before applying the FT. This operation can be interpreted as change of the signal's scale for the conversion of time-varying frequency components to frequency invariant components. The scaling operation can be generalized using the Scale Transform (ST) [4], [5], [6], where the scale is taken as a physical property of the signal, or the scaling operation can be integrated into the definition of transformation, as in Harmonic Transform [7].

There are other transforms for representation of nonstationary harmonic signals, which are suitable for specific applications. The Fractional Fourier Transform (FrFT) [8], [9] and Chirp Transform (CT) [10], [11] are suitable for linearly changing frequency components, whereas the Fan-Chirp Transform [12], [13] is suitable for signals with frequency components varying linearly on a fan geometry.

II. HARMONIC TRANSFORM

Harmonic transform has been introduced in [7] and its main difference from Fourier transform is the integrated time-warping function. It is defined as

$$S_{\phi_u(t)}(\omega) = \int_{-\infty}^{+\infty} s(t) \phi'_u(t) e^{-j\omega \phi_u(t)} dt, \quad (1)$$

M. Trzos and H. Khaddour are with the Department of Telecommunications, Brno University of Technology, Brno, Czech Republic, e-mail: trzos@phd.feec.vutbr.cz and xkhadd00@stud.feec.vutbr.cz.

Manuscript received November 15, 2012; revised January 11, 2007.

where $\phi_u(t)$ is unit phase, which is the phase of the fundamental harmonic component divided by its instantaneous frequency [7], and $\phi'_u(t)$ is first derivation of $\phi_u(t)$. Linear change of fundamental frequency in a given segment is presumed, which is sufficiently satisfied by selecting an analysis window of appropriate length.

Instantaneous phase $\varphi(t)$ of a sinusoid with linear change of frequency [14] is defined as

$$\varphi(t) = 2\pi \left(f_0 t + \frac{\epsilon t^2}{2} \right), \quad (2)$$

where f_0 is fundamental frequency and $\epsilon = \Delta f_0 / T$ is the change of fundamental frequency divided by length of the segment. Assuming discrete signal segment of the length N , where $T = N / F_s$, the discrete phase $\varphi(n)$ of a sinusoid with linear frequency variation [15] can be written as

$$\varphi(n) = 2\pi \left(\frac{f_0 n}{F_s} + \frac{\Delta f_0 n^2}{2N F_s} \right), \quad (3)$$

where f_0 is discrete instantaneous frequency, N is length of the analysis window, and F_s is sampling frequency (initial phase is disregarded for simplicity).

Initial fundamental frequency in a given segment can be written as

$$f_0 = f_c - \frac{a f_c}{2}, \quad a = \frac{\Delta f_0}{f_c}, \quad (4)$$

where f_c is the central fundamental frequency within a segment and a is the slope of fundamental frequency change within the segment. Substituting (4) to (3) we get [15]

$$\varphi(n) = \frac{2\pi}{N} \alpha(n), \quad \alpha_a(n) = n \left(1 - \frac{a}{2} + \frac{an}{2N} \right). \quad (5)$$

Frequencies of spectral lines of the Fourier transform are given as

$$f_c = \frac{F_s}{N}, \quad (6)$$

and from the equation (5) it is obvious that the instantaneous phase is

$$\varphi(n) = \frac{F_s}{N} \alpha(n). \quad (7)$$

Discrete harmonic transform (DHT) of signals with linear fundamental frequency variation [15] is defined as

$$s(n) = \frac{1}{N} \sum_{k=0}^{N-1} s(n) \alpha'(n) e^{j \frac{2\pi k}{N} \alpha(n)}, \quad (8)$$

where

$$\alpha'_a(n) = 1 - \frac{a}{2} + \frac{an}{N}, \quad (9)$$

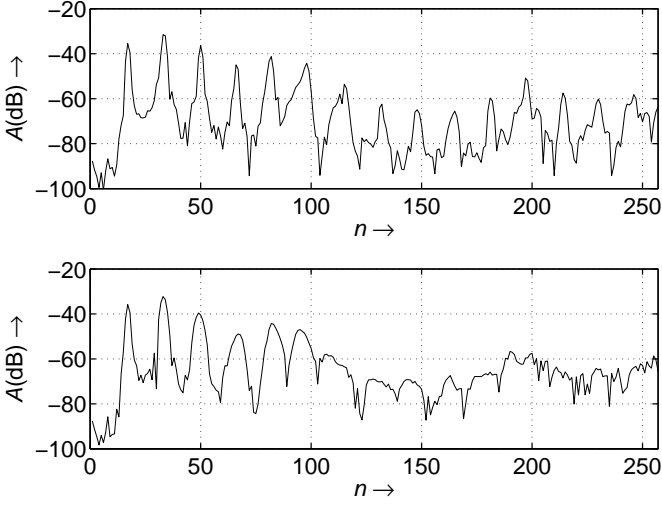


Fig. 1. Fourier transform spectrum (bottom); Harmonic transform spectrum with parameter $a = 0.2$ (top).

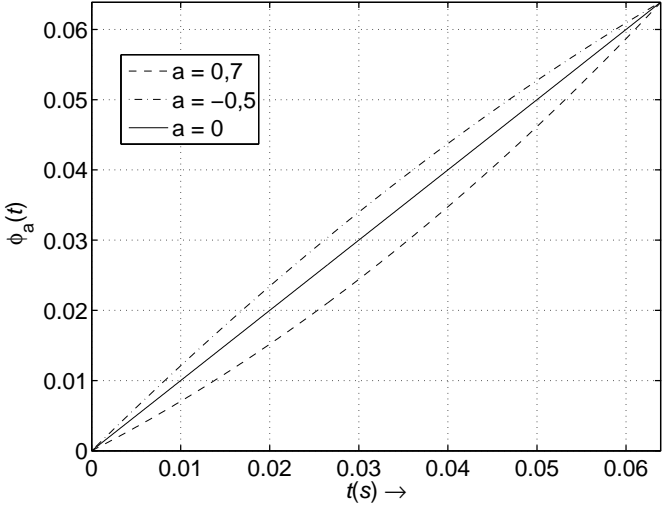


Fig. 2. Influence of parameter a on the phase function $\phi_a(t)$.

is first-order derivation of (5).

The difference between DFT and DHT at analysis of non-stationary harmonic signals can be seen on a part of speech uttering from the PTDB-TUG database with frequency modulation. On Fig. 1 (bottom) we can see that the higher frequencies are smoothed due to frequency modulation. On Fig. 1 (top) even high frequency peaks are clearly visible.

III. ESTIMATING PARAMETER a

As can be seen from equation (4), a determines the change of fundamental frequency in the analysed signal, with the assumption of linear fundamental frequency change. Variation of a results in time-warping of the analysed signal, as can be seen on Fig. 2.

The parameter a is usually estimated by grid search. It consists of performing several realisations of HT with different a [7]. Afterwards, the realisation which minimises spectral

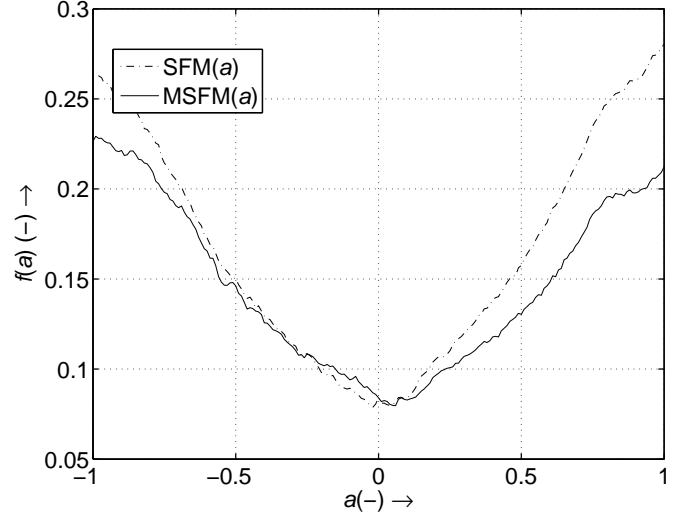


Fig. 3. Comparison of spectral flatness measure and modified spectral measure for the analysed segment.

flatness measure (SFM)

$$\arg \min_a \text{SFM}(a) = \frac{\sqrt{\prod_{k=0}^{N-1} |\text{DHT}(a, k)|}}{\frac{1}{N} \sum_{k=0}^{N-1} |\text{DHT}(a, k)|}, \quad (10)$$

where DHT is discrete harmonic transform and $|\cdot|$ denotes absolute value, is selected as the correct value. This approach is computationally extensive and is usually simplified restricting values a can have and by following the change of fundamental frequency in time [15].

When using (10) the SFM has several minimums and if a search algorithm was used, it could fall into local minimum. It is also noteworthy that it is possible the harmonic transform $|\text{DHT}(a, k)|$ will be equal to zero for some values of k , which could mean that the spectral flatness will be zero for all a . Removing zero values solves this problem and leads to band-limited spectral flatness measure.

The harmonic spectrum is not complex conjugated even for real signals (which is true for Fourier transform). From the frequency axis point of view, the unit phase function $\phi_u(t)$ shifts the spectrum towards lower frequencies if a is positive, and to higher frequencies if its negative. Using the formula (8) we get only one-sided spectrum, the right part will not represent harmonic components of the analysed signal. When estimating a , (10) has two minimums (see Fig. 3). For harmonic signal analysis, only left side of the spectrum is useful, because it appropriately represents non-stationary harmonic signal. Using the modified spectral flatness measure (MSFM)

$$\arg \min_a \text{MSFM}(a) = \frac{\sqrt{\prod_{k=0}^{N/2} |\text{DHT}(a, k)|}}{\frac{1}{N/2+1} \sum_{k=0}^{N/2} |\text{DHT}(a, k)|}, \quad (11)$$

we can get function of a which has clearly defined minimum. This is caused by using only left side of the spectrum and it consequently leads to reducing the number of operations needed to compute spectral flatness by $\frac{N}{2} - 1$.

IV. CONCLUSION

Harmonic transform with some of its features and parameters has been described. Furthermore, the parameter a and its influence on non-stationary signal analysis was analysed. By reducing the number of operations in computation of spectral flatness measure, the number of operations needed for computation of Harmonic transform is reduced.

Future work aims at further improving the performance of harmonic transform computation using parallelisation and using machine learning for estimation of parameter a . Sensitivity to noise will also be measured.

ACKNOWLEDGMENT

This work was supported in part by the Faculty of Electrical Engineering and Communication, Brno University of Technology internal grant no. FEKT-S-11-17.

REFERENCES

- [1] M. Goodwin and M. Vetterli, "Time-frequency signal models for music analysis, transformation, and synthesis," in *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, 1996, pp. 133–136.
- [2] G. Wakefield, "Time-pitch representations: acoustic signal processing and auditory representations," in *Time-Frequency and Time-Scale Analysis, 1998. Proceedings of the IEEE-SP International Symposium on*, Oct. 1998, pp. 577–580.
- [3] A. Makur and S. Mitra, "Warped discrete-fourier transform: Theory and applications," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 48, no. 9, pp. 1086–1093, Sep. 2001.
- [4] L. Cohen, "The scale representation," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3275–3292, Dec. 1993.
- [5] T. Irino and R. D. Patterson, "Segregating information about the size and shape of the vocal tract using a time-domain auditory model: the stabilised wavelet-mellin transform," *Speech Commun.*, vol. 36, pp. 181–203, Mar. 2002.
- [6] A. De Sena and D. Rocchesso, "A study on using the mellin transform for vowel recognition," in *SMC05, International conference on Sound Music and Computing*, Salerno, Italy, 2005.
- [7] F. Zhang, G. Bi, and Y. Chen, "Harmonic transform," *Vision, Image and Signal Processing, IEE Proceedings*, vol. 151, no. 4, pp. 257–263, Aug. 2004.
- [8] D. Bailey and P. Swartztrauber, "The fractional fourier transform and applications," *SIAM Review*, vol. 33, no. 3, pp. 389–404, Sep. 1995.
- [9] R. Tao, Y.-L. Li, and Y. Wang, "Short-time fractional fourier transform and its applications," *Signal Processing, IEEE Transactions on*, vol. 58, no. 5, pp. 2568–2580, May 2010.
- [10] L. Weruaga and M. Képesi, "Speech analysis with the short-time chirp transform," in *Eurospeech*, Geneve, Sep. 2003, pp. 53–56.
- [11] M. Képesi and L. Weruaga, "Speech analysis with the fast chirp transform," 2004.
- [12] L. Weruaga and M. Képesi, "The fan-chirp transform for non-stationary harmonic signals," *Signal Processing*, vol. 87, no. 6, pp. 1504–1522, 2007.
- [13] M. Képesi and L. Weruaga, "Adaptive chirp-based time-frequency analysis of speech signals," *Speech Communication*, vol. 48, no. 5, pp. 474–492, 2006.
- [14] P. Zubrycki and A. Petrovsky, "Accurate estimation of harmonic amplitudes in voiced speech based on harmonic transform," in *Signals and Electronic Systems, 2008. ICSES '08. International Conference on*, Sep. 2008, pp. 47–50.
- [15] —, "Accurate speech decomposition into periodic and aperiodic components based on discrete harmonic transform," in *15th European Signal Processing Conference (EUSIPCO 2007)*, Poznań, Poland, 2007, pp. 2336–2340.

Michal Trzos received B.Sc. degree in teleinformatics and M.Sc. degree in telecommunications and informatics from the Faculty of Electrical Engineering, Brno University of Technology (BUT), CZE, in 2007 and 2009 respectively. Since 2009, he is a Ph.D. candidate at the Department of Telecommunications. His main task is to explore new methods of time and frequency warping of audio signals. His research interests include: general-purpose computing on graphics processing units, audio algorithm parallelization, time-frequency transformations, and speech processing.

Hasan Khaddour received his Eng. title from the Department of Telecommunications and Electronics, Faculty of Mechanical and Electrical Engineering, Tishreen University, Syria, in 2007. Since 2009, he is a Ph.D. candidate at the Department of Telecommunications, Faculty of Electrical Engineering, Brno University of Technology (BUT), Czech Republic. His current research is focused on sound source localization, acoustical zooming, and sound rendering methods.