Speech Compression Based on the Discrete Shapelet Transform

Rodrigo Capobianco Guido, Sylvio Barbon Junior, Lucimar Sasso Vieira, Paulo Cesar Fantinato, Fabricio Lopes Sanchez, Joao Paulo Lemos Escola, Enio Bolognini, Leonardo Mendes Souza, Paulo Ricardo Franchi Zulato, Jussara Ribeiro, Michel Alves Lacerda
Depto de Física e Informática, IFSC, USP
13566-590, São Carlos, SP

ABSTRACT

The paradigm of time-frequency-shape joint analysis of discrete-time signals, that is possible by means of a novel transform recently introduced, called Discrete Shapelet Transform (DST), is used in this paper for audio compression, improving the previous results obtained with wavelet bases. Both perceptual and rate-distortion criteria are considered, assuring the efficacy of the proposed approach.

An considerable number of high-quality algorithms for audio compression [1] has appeared in the last years. Some propose the use of a specific wavelet tree [2] wav1based on a rate-distortion or a perceptual criterion [3], while others use different types of filter-based algorithms [4], lapped orthogonal transforms [5], and so on.

Extending our previous results on the best wavelet basis for high quality audio compression in a perceptual sense, recently published in [6], and our study which considers the perceptual and the rate-distortion criteria separately, published in [7], we propose a scheme where the compression is based on the Discrete Shapelet Transform (DST) [8], instead of the traditional Discrete Wavelet Transform (DWT). In the previous papers, we concluded that almost linear phase wavelet filters, such as Symmlets, are the best ones for audio compression in a restricted sense, while their support-size should be found adaptively, according to each particular audio frame being analyzed. In this paper, we show that the DST, which is a new promising transform for joint time-frequency-shape analysis of signals, can be used to improve the previous results in terms of compression rates for speech data.

References

[1] Bosi, M., Goldberg, R. Introduction to Digital Audio Coding and Standards. 2ed., Kluwer Academic Publishers, Massachusetts, 2003.

[2] Strang, G.; Nguyen, T., "Wavelets and Filter Banks", Wellesley-Cambridge Press, Wellesley, 1997.

[3] M. Erne, G. Moschytz, C. Faller, ``Best Wavelet-Packet Bases for Audio Coding Using Perceptual and Rate Distortion Criteria", in Proc. of IEEE ICASSP, v.2, 1999, pp. 909 - 912.

[4] K. Ramchandran, M. Vetterli, ``Best Wavelet Packet Bases in a rate-distortion Sense", IEEE Trans. on Image Processing, v.2, n.2, apr-1993, pp.160-175.

[5] Malvar, H.S., Signal Processing with Lapped Transforms, Artech House, Norwood, 1996.

[6] Guido, R.C., Maciel, C. D.; Monteiro, M., Fonseca; E.S., Panchapagesan, S.; Pereira, J. C.; Vieira, L.S.; Barbon, S.Jr.; Sanchez, F.L.; Guilherme, M.B.A.; Sergio, K.I.C.; Scarpa, T.L.; Fantinato, P.C.; Moura, E..J.R. A study on the best wavelet for audio compression. 40th IEEE Asilomar Int. Conference on Signals, Systems and Computers, Pacific-Grove, CA, USA, 2006, v. 1. p. 2115-2118.

[7] Guido, R. C. ; Vieira, L. S. ; Barbon Jr, S. ; Sanchez, F. L. ; Guilherme, M. B. A. ; Sergio, K. I. C. ; Escarpa, T. L. ; Fonseca, E. ; Pereira, J. C. ; Monteiro, M. . A Fractal and wavelet-based approach for audio coding. In: 8th IEEE International Symposium on Multimedia 2006, San Diego, CA, USA, 2006. v. 1. p. 253-256.

[8] Guido, R. C. ; Barbon Jr, S. ; Vieira, L. S. ; Sanchez, F. L. ; Maciel, C. D. ; Pereira, J. C. ; Scalassara, P. R. ; Fonseca, E. Introduction to the Discrete Shapelet Transform and a New Paradigm: joint time-frequency-shape analysis. In: IEEE International Symposium on Circuits and Systems (IEEE ISCAS 2008), Seattle, CA, USA, 2008 (accepted, in--press).