

Special Issue on Advanced Cognitive Systems Based on Nonlinear Analysis

Carlos M. Travieso · Jesús B. Alonso

Received: 4 July 2013 / Accepted: 7 November 2013 / Published online: 17 November 2013
© Springer Science+Business Media New York 2013

Cognitive systems are being addressed to different computational tasks related to the analysis, processing, interpretation and evaluation of information. Since a relatively short time, the strategies implemented with classic linear models are being reviewed in order to evaluate their potential in the study of nonlinear modeling problems. Speech signal is a clear example of a situation where this is possible.

Many aspects of the signal analysis are not well addressed by the conventional models currently used in the field of signal processing. The purpose of the special issue is to present and discuss novel ideas, works and results related to alternative techniques for nonlinear processing, which come from mainstream approaches.

Nowadays, the studies based on complex systems have opened new doors in the research field, allowing the improvement of the quality and the results of diverse applications. Nonlinear analysis has made this task easier in areas like audio analysis, voice evaluation, quality of voice, blind source separation, expression detection and voice identification. We summarize the papers as follows, in seven main topics.

1. *Audio analysis*. Gomez-Vilda et al. developed a methodology that has benefited from the advances achieved during the last years in detecting and assessment of organic pathologies in phonation. The paper hypothesizes that some of the underlying neurological mechanisms affecting phonation produce

observable correlates in vocal fold biomechanics and that these correlations behave different in neurological diseases than in organic pathologies. Cadone et al. explored an alternative to the morphological filtering for speech enhancement and noise compensation scheme. They used auditory-inspired anisotropic structuring elements applied to grayscale spectrograms not only for speech enhancement but also for automatic speech recognition.

2. *Voice evaluation on nonlinear models*. Drugman and Dutoit showed an approach which relies on oscillating statistical moments. Such moments exhibit the property of having a phase shift which depends on the speech polarity. This dependency arises from the higher-order statistics in the moment calculation. The characterization of hypernasal vowels and words using nonlinear dynamics was presented by Orozco et al., considering different complexity measures that are mainly based on the analysis of the time delay embedded space. Blanco et al. showed that the usage of automatic tools based on speech technologies is indeed a reliable, cost-effective approach for the early detection of OSA patients at severe stages of the syndrome. The introduction of nonlinear measures describing the underlying dynamics in the production of sustained vowels has enhanced our characterization of patients' acoustic space and resulted in an increase in the overall OSA detection accuracy.
3. *Quality of voice*. In audio analysis, Calzada and Socoró showed an adaptation of the adaptive pre-emphasis linear prediction technique (APLP) to the HMM for modifying the vocal effort. The proposed transformation methodology is validated using a Copy Re-Synthesis strategy on a speech corpora specially designed for vocal effort researches.

C. M. Travieso (✉) · J. B. Alonso
Institute for Technological Development and Innovation in
Communications, University of Las Palmas de Gran Canaria,
Las Palmas de Gran Canaria, Spain
e-mail: ctravieso@dsc.ulpgc.es

J. B. Alonso
e-mail: jalonso@dsc.ulpgc.es

4. *Blind source separation.* Sole-Casals and Caiifa proposed a simple method to reduce the computation time for the inversion of Wiener systems or the separation of post-nonlinear mixtures, by using a linear approximation in a minimum mutual information algorithm, applied to voice signal. Zdunek demonstrated that imposing MRF smoothing on the power spectrograms of audio sources estimated from under-determined unmixing problems may improve the quality of estimated audio sounds considerably. That study addressed the application of MRF smoothing in the EM-NMF algorithm, but this type of smoothing could be applied to many other related BSS algorithms based on feature extraction from power spectrograms. About blind source separation and noisy signals, Rotili et al. presented advanced real-time speech processing front-end aimed at automatically reducing the distortions introduced by room reverberation in distant speech signals, also considering the presence of background noise, and thus to achieve a significant improvement of speech quality for each speaker.
5. *Expression detection.* The application of measures based on nonlinear dynamics for emotional speech characterization is applied by Henriquez et al. Measures such as mutual information, dimension correlation, entropy correlation, Shannon's entropy, Lempel–Ziv complexity and Hurst exponent are extracted from the samples of a database of emotional speech. Experiments were conducted on the Berlin emotional speech database for a three-class problem (neutral, fear and anger as emotional states). Planet and Iriondo developed a classification of children's affective states in a real-life non-prototypical emotion recognition scenario by a wrapper method. They aimed to reduce the acoustic set of features and feature-level fusion to merge them with the set of linguistic parameters. After the classification system based on a Naïve Bayes classifier, a support vector machine and a logistic model tree, the approach showed that the linguistic features improve the performances of the classifiers that use only acoustic datasets.
6. *Voice identification.* Alam et al. researched low-variance multi-taper spectrum estimation methods to compute the mel-frequency cepstral coefficients (MFCC) features for robust speech and speaker recognition systems on the AURORA-2 and AURORA-4 corpora and the NIST 2010 speaker recognition evaluation (SRE) corpus (telephone as well as microphone speech), respectively. They concluded that the multi-taper methods perform better compared to the Hamming method. Ezeiza et al. showed that the nonlinear nature of speech suggests that extra information provided by some nonlinear features could be especially useful when training data are scarce, or when the ASR task is very complex. The fractal dimension (FD) of the observed time series was combined with the traditional MFCCs in the feature vector in order to enhance the performance of two different ASR systems.
7. *Other applications to nonlinear, which can be applied on voice.* Vásquez et al. developed a fixed point algorithm, designed on a field programmable gate array (FPGA), based on artificial neural networks for temperature prediction. This method can be applied on the voice processing field for future applications. Travieso et al. proposed the development of a kernel based on MFCC features in order to extent its discriminative information. Although it was originally applied to apnea detection, it can easily be extended to voice signals thanks to the periodicity information of both signals; they are different, but the adjustment of variables can make it possible.

We wish to thank all the people that enabled the publication of this special issue. First of all, we wish to thank Prof. Amir Hussain, Editor-in-chief, for the special issues of this journal, for accepting the idea and for his support, patience and motivation. Our gratitude also goes to the Journal Manager and to all the staff from Elsevier, in particular to Ms. Roshna Mohan, for the impeccable and timely logistical support. The papers in this issue were reviewed on two or three rounds of reviews. We wish to equally thank the authors and the reviewers for all their hard work and contribution for the excellence of this special issue.