Digital Signal Processing for Analytics in Biostatistics & Biometric Applications

Ehsan Sheybani* and Giti Javidi

University of South Florida, College of Business, USA

Abstract

All the biostatistics and biometric applications suffer from the effects of added noise due to their data dependency. The quality of data and impurities due to noise could affect the decisions made based on these datasets. Detecting anomalies caused by noisy datasets requires special preprocessing techniques that do not hurt the integrity of data. The authors have developed computationally low power, low bandwidth, and low-cost filters (DMAW) that will remove the noise, compress the dataset, and decompose the dataset so that a decision can be made by looking at different layers of data. This wavelet-based method is guaranteed to converge to a stationary point for both uncorrelated and correlated data. Presented here is the theoretical background with examples showing the performance and merits of this novel approach compared to other alternatives.

Keywords: Biostatistics; Biometric; Discrete wavelet transform; Discrete meyer mdaptive wavelet

Abbreviations: FFT: Fast-Fourier Transform; IFFT: Inverse Fast-Fourier Transform; FT: Fourier transform; CWT: Continuous Wavelet Transform; DWT: Discrete Wavelet Transform; MWT: Multi-Resolution Wavelet Transform; DMAW: Discrete Meyer Adaptive Wavelet

Introduction

Biostatistics and biometric (bio-systems) are exciting fields that try to answer a broad range of questions in the interface of medicine, science, statistics, and engineering. It is hard to imagine where medicine and science would be today without advanced bio-systems, wireless sensor networks, robotics, 3D printers, nanoparticles, and big date analytics, to name a few. Technology has had a dramatic effect on the trend of biostatistics and biometric, producing what may be the most challenging period in the history of human beings. One form of the convergence of technology in science and medicine is advanced bio-systems. Bio-systems provide alternative means for analysis and visualization of the scientific and medical information and services, providing patients and researchers with an extraordinary new range of options.

Traditionally, standard methods such as Fast-Fourier Transform (FFT) and Inverse Fast-Fourier Transform (IFFT) have been used to process datasets for medical and scientific purposes. Due to non-parametric features of these methods and their resolution limitations and observation time dependence, use of spectral estimation and signal pre and post-processing techniques based on wavelets to process data has been proposed. Multi-resolution wavelet transforms and advanced spectral estimation techniques have proven to offer efficient solutions to this problem [1]. This paper discusses innovative wavelet-based filter banks designed to enhance the analysis of dataset using parametric spectral methods and signal classification algorithms.

Finally, the volume, depth, and breadth of data acquired from advance imaging systems has resulted in the recent out of bounds growth of data, thus, requiring modern processing and visualization techniques to extract useful information for decision-making. Finding correlation among thousands of variables in big datasets to determine their relative importance is not a simple task. Advanced analysis and scientific data visualization have proven to be effective techniques in discerning information from big datasets. Using proven, fast, and sophisticated filtering techniques, this article also aims at extracting information, showing patterns, and allowing mining of big datasets in real-time for faster and more effective decision making. Given the unique challenges of scientific big data visualization, the research presented in this paper covers

some potential solutions and offers a means of setting standards for this new and evolving field.

Discussion

This study aims at reducing the dimensionality of dataset to reduce computational load in further processing [2]. The proposed method ranks features for learning a distance function in order to capture the semantics of the dataset [3]. It also uses the orthogonality properties of wavelets to decompose the dataset into spaces of coarse and detailed signals. Depending on whether a given function is analyzed in all scales and translations or a subset of them, the continuous (CWT), discrete (DWT), or multi-resolution wavelet transform (MWT) can be applied. The DWT can be used to suppress noise and reduce order of data in a wireless sensor network. Due to its ability to extract information in both time and frequency domain, DWT is considered a very powerful tool. The approach consists of decomposing the signal of interest into its detailed and smoothed components (high-and low-frequency).

The detailed components of the signal at different levels of resolution localize the time and frequency of the event. Therefore, the DWT can extract the coarse features of the signal (compression) and filter out details at high frequency (noise). DWT has been successfully applied to system analysis for removal of noise and compression [4]. DWT-based filters can be used to localize abrupt changes in signals in time and frequency. The invariance to shift in time (or space) in these filters makes them unsuitable for compression problems. Therefore, creative techniques have been implemented to cure this problem [5]. These techniques range in their approach from calculating the wavelet transforms for all circular shifts and selecting the "best" one that minimizes a cost function [6], to using the entropy criterion [7] and adaptively decomposing a signal in a tree structure so as to minimize the entropy of the representation. In this paper a new approach to cancellation of noise and compression of data has been proposed. The discrete Meyer adaptive wavelet (DMAW) is both translation- and scale-invariant and can represent a signal in a multi-scale format. While DMAW is not the best fit for entropy criterion, it is well suited for the proposed compression and cancellation purposes [8-10].

Conclusion

As expected, the DMAW filters perform well under noisy conditions in an imaging environment. The decomposed signal could be easily freed up from noise and reduced down to its coarse component only. This could be reduction by several orders of magnitude in some cases. Future plans include the application of these filters to fused datasets and comparison between different approaches. Additionally, the results of these study can be used in the decision making stage to realize the difference this approach can make in speed and efficiency of this process. Future work will address issues such as characterizing the parameters for simulation and modeling of the proposed filter; showing how complex examples with correlated data will be filtered for redundancy; comparing the proposed approach with other similar approaches and giving comparative results to support the claimed advantages, both theoretically and experimentally.

Acknowledgement

None.

Conflict of Interest

No conflict of interest.

References

1. Sheybani E, Garcia Otero S, Adnani F, Javidi G, Deshpande M (2012) An Algorithm for Real-Time Blind Dataset Quality Comparison and Assessment. IJECE 2(1): 120-129.

2. Sheybani EO, Mengshoel OJ, Poll S (2010) Removing spikes while preserving data and noise using wavelet filter banks. Proceedings of IEEE Aerospace Conference, USA.

3. Aparna Varde, Elke Rundensteiner, Giti Javidi, Ehsan Sheybani, Jianyu Liang (2007) Learning the Relative Importance of Features in Dataset Data. Proceeding of IEEE ICDE's DB Rank, USA.

4. Sheybani E, Sankar R (2002) A Survey of Telemedicine/Teleradiology/Telemammography Network Architectures. SPIE Journal of Electronic Imaging.

5. Sheybani E (2011) Dimensionality Reduction and Noise Removal in Wireless Sensor Networks. New Technologies, Mobility and Security (NTMS) Proceedings of 4th IFIP International Conference, Japan.

6. Hanqi Guo, He Xiao, Xiaoru Yuan (2012) Scalable Multivariate Volume Visualization and Analysis Based on Dimension Projection and Parallel Coordinates. Visualization and Computer Graphics IEEE Transactions on 18(9): 1397-1410.

7. Samuel K, Jaakko P (2011) Dimensionality Reduction for Data Visualization. IEEE Signal Processing Magazine 28(2): 100-104.

8. Sheybani E, Javidi G (2009) Dimensionality Reduction and Noise Removal in Wireless Sensor Network Datasets. Computer and Electrical Engineering Second International Conference on ICCEE.

9. Varde A, Rundensteiner E, Javidi G, Sheybani E (2007) Learning the Relative Importance of Features in Dataset Data. Data Engineering Workshop IEEE 23rd International Conference, Istanbul.

10. Sheybani E, Javidi G (2012) Multi-resolution filter banks for enhanced SAR imaging, Systems and Informatics (ICSAI) International Conference.