

Editorial

Advanced Signal Processing Techniques for Bioinformatics

Xue-Wen Chen,¹ Sun Kim,² Vladimir Pavlović,³ and David P. Casasant⁴

¹Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045, USA

²School of Informatics, Indiana University, Bloomington, IN 47408, USA

³Department of Computer Science, Rutgers University, Piscataway, NJ 08854, USA

⁴Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received 5 January 2006; Accepted 5 January 2006

Copyright © 2006 Xue-Wen Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The success of bioinformatics in recent years has been driven, in part, by advanced signal processing techniques: estimation theory, classification, pattern recognition, information theory, networks, imaging, image processing, coding theory, and speech recognition. For example, Fourier analysis methods are used to elucidate the relationship between sequence structure and function; wavelet analysis methods have been applied in sequence comparison and classification; and various image processing methods have been developed to improve microarray image quality.

The development of advanced high-throughput technologies such as genome sequencing and whole genome expression analysis creates new opportunities and poses new challenges for the signal processing community. Analysis of data for life-science problems provides an interesting application domain for standard signal processing methods such as time series detection and prediction, casual modeling, and structure inference. At the same time, this increasingly important life-science domain draws the need for novel and computationally efficient analysis approaches. The goal of this special issue is to present the applications of cutting-edge signal processing methods to bioinformatics.

Eleven papers accepted for this special issue cover a broad range of topics, from RNA sequence analysis and gene expression analysis to protein structure predictions. The authors developed a variety of signal processing algorithms, such as artificial neural networks, decision trees, biclustering, matrix factorization, and FPGA reconfiguration methods, to tackle these central bioinformatics problems.

The issue starts with two papers on gene sequence analyses. Churkin and Barash developed a pattern recognition-based utility to perform mutational analysis and detect vulnerable spots within an RNA sequence that affect structures;

Babu et al. presented image processing/computer vision methods for automatic recovery and visualization of the 3D chromosome structure from a sequence of 2D tomographic reconstruction images taken through the nucleus of a cell.

The advent of microarray techniques that allow for measuring the expression levels of tens of thousands of genes simultaneously has drawn increased interest in signal processing community, covering a range of problems from microarray image processing and biomarker detection to genetic regulatory network reconstruction. Three papers in this special issue address microarray applications: Bajcsy provided an excellent overview on DNA microarray grid alignment and foreground separation approaches; Jin et al. proposed two automated methods for microarray image analysis; Tchagang and Tewfik described a novel biclustering algorithm for microarray data.

As gene products, proteins play an essential role in nearly all cellular functions. The remaining papers deal with issues in proteomics. Two papers focused on the prediction of protein-protein interactions (PPIs) based on domain information: Zhang et al. modeled the problem of interaction inference as a constraint satisfiability problem and solved it using linear programming; Chen and Liu developed neural network and decision tree-based approaches for predicting PPIs, and demonstrated that with decision trees, multiple domain interactions could be identified. The following three papers moved to protein structure related topics: Okun and Priisalu applied fast nonnegative matrix factorization methods to protein fold recognition; VanCourt et al. presented an FPGA reconfiguration method for alternative force laws with applications to molecular docking; Armano et al. developed a pattern recognition system for protein secondary structure prediction. Finally, Kolibal and Howard developed

a stochastic Bernstein approximation method for obtaining the baseline shift removal of matrix-assisted laser desorption ionization time-off-light mass spectrometry.

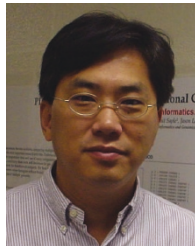
The guest editors would like to thank all the authors for their high quality work contributed to this special issue and all the reviewers for their hard work and expert comments in evaluating the manuscripts.

*Xue-wen Chen
Sun Kim
Vladimir Pavlović
David P. Casasent*

Xue-Wen Chen received the Ph.D. degree from Carnegie Mellon University, Pittsburgh, USA, in 2001. He then spent about one year as a postdoctoral fellow at the University of Illinois at Urbana-Champaign. He is currently an Assistant Professor of computer science at the University of Kansas, Lawrence, USA. He is also a Member in Kansas Masonic Cancer Research Institute. He is an IEEE Senior Member. His research interests include bioinformatics and machine learning. Much of his work addresses two core problems in learning: analyzing large-scale dataset and learning from high dimensions. His current research is focused on developing computational methods such as kernel-based classifiers and feature selection for genomic and proteomic data analysis.



Sun Kim is an Associate Director of Bioinformatics Program, an Assistant Professor in School of Informatics, an Associated faculty at the Center for Genomics and Bioinformatics, an Affiliated faculty at the Bio-complexity Institute at Indiana University - Bloomington. Prior to IU, he worked at DuPont Central Research as a Senior Computer Scientist from 1998 to 2001, and at the University of Illinois at Urbana-Champaign from 1997 to 1998 as a Director of Bioinformatics and postdoctoral fellow at the Biotechnology Center and a Visiting Assistant Professor of animal sciences. Sun Kim received B.S., M.S., and Ph.D. degrees in computer science from Seoul National University, Korea, Advanced Institute of Science and Technology (KAIST), and the University of Iowa, respectively. He is a recipient of Outstanding Junior Faculty Award at Indiana University in 2004, NSF CAREER Award DBI-0237901 from 2003 to 2008, and Achievement Award at DuPont Central Research in 2000.



Vladimir Pavlović is an Assistant Professor in the Computer Science Department at Rutgers University. He received his Ph.D. degree in electrical engineering from the University of Illinois in Urbana-Champaign in 1999. From 1999 until 2001, he had been a Member of research staff at the Cambridge Research Laboratory, Cambridge, Mass. Pavlović's research interests include modeling of time-series, statistical computer vision, machine learning, and bioinformatics.



David P. Casasent is a Full Professor at Carnegie Mellon University, Pittsburgh, Pennsylvania, in the Department of ECE, where he is the George Westinghouse Professor and Director of the Laboratory for Optical Data Processing. He is a Fellow of the IEEE, OSA, and SPIE and has received various best paper awards and other honors. He is the author of two books, editor of one text, editor of 70 journal and conference volumes, as well as contributor to chapters in 20 books and over 700 technical publications, on various aspects of data processing, image pattern recognition, and real-time signal processing. He originated and has organized the set of 6 to 11 annual SPIE Conferences on Intelligent Robots and Computer Vision. He has chaired the Intelligent Robots and Computer Vision conference for 22 years. He has also chaired and organized the Optical Pattern Recognition and Hybrid Image and Signal Processing SPIE conferences for 13 years. He is a Past President of SPIE and was on the Board of Directors of SPIE for 6 years. He received the 1996 SPIE President's Award. He is a Past Member of two Defense Science Board Task Forces (on image recognition and on automatic target recognition). He is presently a Faculty Advisor to Eta Kappa Nu among other such activities. His research interests include distortion-invariant pattern recognition, neural networks, Gabor and wavelet transforms, robotics, morphological image processing, and product inspection. He was on the Board of Directors of INNS (the International Neural Network Society) for 10 years, and was President on INNS in 1998.

