Final Report for Period: 08/2007 - 07/2008   **Submitted on:** 11/02/2008

Principal Investigator: Vidakovic, Brani D.   **Award ID:** 0505490

Organization: GA Tech Res Corp - GIT

Submitted By:

Vidakovic, Brani - Principal Investigator

Title:

Collaborative Research: Analysis of Functional and High-Dimensional Data with Applications

Project Participants

Senior Personnel

>**Name:** Vidakovic, Brani

>**Worked for more than 160 Hours:**   Yes

>**Contribution to Project:**

Post-doc

>**Name:** Kim, Jongphil

>**Worked for more than 160 Hours:**   No

>**Contribution to Project:**

>Jongphil Kim is member of BESTA, BIOoengineering STAtistics group at the school of Biomedical Engineering at Georgia Tech. He is involved in a research project under grant's umbrella that deals with assessment of multifractal spectra of protein mass spectra and its use in classification of patients. He is supported by his teaching activity at Georgia Tech and no grant funds are aloted to his work.

>**Name:** Nicolis, Orietta

>**Worked for more than 160 Hours:**   No

>**Contribution to Project:**

>Dr Orietta Nicolis, visiting scholar from University of Bergamo, Italy is involved in a project that uses wavelets and scaling for analysis of geophysical satelite images and nano-surfaces.
>The images produce wavelet summaries-descriptoirs amenable to classification and discrimination.
>Dr Nicolis is supported by Italian CNR, Industry Grant and University of Bergamo and no funds have been aloted.

Graduate Student

>**Name:** Lee, Kichun (Sky)

>**Worked for more than 160 Hours:**   Yes

>**Contribution to Project:**

>Kichun Lee is Graduate RA at the ISyE fully supported by this grant. His support was 1/3 of the time for 1 year now.
>Kichun is involved in two research projects related to the grant: (1) potential use of semi-supervised learning (Laplacians on manifolds) for inteligent wavelet shrinkage and (2) wavelet-based classification of unprecessed protein mass spectra. He also is busy with the course-work and takin qualifying exam, but first manuscript related to his research problems is expected at the end of this calendar year.

>**Name:** Ramirez, Josefa Pepa

>**Worked for more than 160 Hours:**   No

>**Contribution to Project:**

>Pepa Ramirez is Fulbright Student Scholar visiting BESTA-BME from University Carlos Tercero -- Spain. She is involved full time in a project with Emory University on use of wavelets to classify mammogramy images. This project is under grant's umbrella and is ongoing collaboration with Dubois Bowman, Gordana Derado from Emory's Biostatistics and Dr Mary Newell, MD from the Radiology Department at Emory.
>Pepa Ramirez uses sumaries of multifractal spectra of mammogramy background images to classify cases as benign or malignant.

Undergraduate Student

Technician, Programmer

Other Participant

Research Experience for Undergraduates

<u>**Organizational Partners**</u>

University of Central Florida

This Award was joint endeavor between Georgia Tech and University of Central Florida. The collaborative project was headed at UCF by Dr Marianna Pensky and at Georgia Tech by Brani Vidakovic.

<u>**Other Collaborators or Contacts**</u>

<u>**Activities and Findings**</u>

Research and Education Activities:

The underlying theme of the awarded project is the use of multiscale methodology in massive data sets. The main thrust of our research concerns summaries of multidimensional data that involve regular and possibly irregular scaling. Fractal and multi-fractal spectra of high-dimensional data sets are calculated using wavelet-based approaches and the summaries of those spectra are used in statistical inference. Although the topic seem to be narrow, the field of applications is huge.

First, a few descriptors well summarize massive data sets, and second, no traditional statistical methods are efficient when the data contains fractal (self-similar) features and when smooth organized trends carry no information or are simply irrelevant to the phenomenon of interest. In the regression context, convolution of two fractal signals can be satisfactorily resolved in the wavelet domain by looking at the scaling descriptors of the convolution.

We applied spectral summaries to massive bioresponse signals, high resolution medical images, geophysical data and their theoretical surrogates (fractals, fBm, 2D fBm, etc).

Findings:

The key finding is that wavelet based summaries of mnassive data sets well capture relevant information and are useful in feeding low-dimensional statistical models for further inference. The main statistical tool applied to such summaries are various classification and discrimination methods, regression, and linear models.

Training and Development:

Some findings of this grant will be part of the PhD thesis of 2nd year graduate student Kichun Sky Lee. He will be graduating likely in 2010. Also, Fulbright PhD scholar from Spain, Pepa Ramirez is getting proficient in use of wavelets and related inference.

Outreach Activities:

<div align="center">

Journal Publications

</div>

Pensky, M., Vidakovic, B. and De Canditiis, D., "Bayesian Decision Theoretic Scale Adaptive Estimation of Spectral Density", Statistica Sinica, p. 635, vol. 17, (2007). Published,

Nicolis, O., Garutti, C., and Vidakovic, B., "2-D Wavelet-Based Spectra with Applications in Analysis of Geophysical Images. ", Journal of Computational and Graphical Statistics, p. , vol. , ( ). Submitted,

Derado, G; Lee, K; Nicolis, O; Bowman, FD; Newell, M; Rugger, FF; Vidakovic, B, "Wavelet-based 3-D multifractal spectrum with applications in breast MRI images", BIOINFORMATICS RESEARCH AND APPLICATIONS, p. 281, vol. 4983, (2008). Published,

Yi, J.-S., Jung, Y.-Y., Jacko, J., Sainfort, F., and Vidakovic, B, "Parallel Wavestrap: Simulating Acceleration Data for Mobile Context Simulator", Current Development in Theory and Applications of Wavelets, p. 251, vol. 1, (2007). Published,

Jung, Y. Y., Park, Y., Jones, D., Ziegler, T., and Vidakovic, B., "Self-Similarity in NMR Spectra: An Application in Assessing the Level of Cysteine", Journal of Data Science, p. , vol. , (2008). Accepted,

Cutillo, L., Jung, Y.-Y., Ruggeri, F., and Vidakovic, B., "Larger Posterior Mode Wavelet Thresholding and Applications", Journal of Statistical Planning and Inference, p. 3758, vol. 138, (2008). Published,

Lavrik, I., Jung, Y.-Y., Ruggeri, F., and Vidakovic, B., "Bayesian False Discovery Rate Wavelet Shrinkage: Theory and Applications", Communications in Statistics, p. , vol. 37, (2008). Accepted,

<div align="center">

Books or Other One-time Publications

</div>

Paul Kvam and Brani Vidakovic, "Nonparametric Statistics for Engineers and Scientists", (2007). Book, Published
Bibliography: Wiley & Sons, Hoboken, NJ
420pp

<div align="center">

Web/Internet Site

</div>

<div align="center">

Other Specific Products

</div>

<div align="center">

Contributions

</div>

Contributions within Discipline:
Novel informative summaries of multifractal spectra. A multifractal
spectra by its nature is a density, thus infinitely dimensional.
Descriptors such as maximum, left or right slope, left or right
tangent and ``broadness'' describe the dominant scaling exponent and
assess the deviation from monofractality.

Contributions to Other Disciplines:
Novel methodology to look at the scaling of unproicessed protein mass

spectra to make inference about deficiency of important metabolites in humans.

Applications in Cancer Diagnostics via assessing Multifractal summaries of mammograms.

Contributions to Human Resource Development:

Contributions to Resources for Research and Education:

Contributions Beyond Science and Engineering:

<u>**Categories for which nothing is reported:**</u>

Activities and Findings: Any Outreach Activities

Any Web/Internet Site

Any Product

Contributions: To Any Human Resource Development

Contributions: To Any Resources for Research and Education

Contributions: To Any Beyond Science and Engineering