**REPORT DOCUMENTATION PAGE**

<table>
<thead>
<tr>
<th>1. AGENCY USE ONLY (Leave Blank)</th>
<th>2. REPORT DATE</th>
<th>3. REPORT TYPE AND DATES COVERED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dd-mm-yyyy</td>
<td>Final 01-08-1999 to 31-12-2002</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE AND SUBTITLE</th>
<th>5. FUNDING NUMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Report on Nonparametric Function Estimation and Visualization Application to C2</td>
<td>DAAD19-99-1-0314</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. AUTHOR(S)</th>
<th>8. PERFORMING ORGANIZATION REPORT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edward J. Wegman</td>
<td>Final 2003-01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</th>
<th>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</th>
</tr>
</thead>
<tbody>
<tr>
<td>George Mason University, MS 4A7</td>
<td>U.S. Army Research Office</td>
</tr>
<tr>
<td>4400 University Drive</td>
<td>P.O. Box 12211</td>
</tr>
<tr>
<td>Fairfax, VA 22030-4444</td>
<td>Research Triangle Park, NC 27709-2211</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. SUPPLEMENTARY NOTES</th>
<th>12 a. DISTRIBUTION / AVAILABILITY STATEMENT</th>
<th>12 b. DISTRIBUTION CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.</td>
<td>Approved for public release; distribution unlimited.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. ABSTRACT (Maximum 200 words)</th>
<th>14. SUBJECT TERMS</th>
<th>15. NUMBER OF PAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>This project focused on the development of fast, accurate density estimation procedures. The methods raised basic research issues as to the implementation, computational complexity, visualization, and optimization of estimators in this class. In addition to being useful in a direct role, it is argued that density estimation plays a crucial role in clustering algorithms, discriminant methods and pattern recognition. All of these methods are used extensively in Situation and Informational Awareness and Understanding and in Monitoring and Discovery Processes. In addition, because of their intuitive appeal and ease in understanding, visually rendered density and function estimators provide a natural format for human-computer interactions with decision makers. This report describes results related implementation, computational complexity, visualization, optimization and application of recursive orthonormal density estimators.</td>
<td>visual data mining</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>quantization</td>
<td></td>
</tr>
<tr>
<td></td>
<td>clustering</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nonparametric density estimation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED</th>
<th>18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED</th>
<th>19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED</th>
<th>20. LIMITATION OF ABSTRACT UL.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18

298-102

20030806 022
SF 298 Continuation Sheet

FINAL REPORT

ARO PN 40276-MA

CONTRACT NO. DAAD19-99-1-0314

Period of Performance: 1 August 1999 to 31 December 2002

Title: Nonparametric Function Estimation and Visualization with Applications to C2

PI: Edward J. Wegman, George Mason University

Papers Published Under ARO Sponsorship

Special Issues or Books


National Research Council, Naval Forces' Capability for Theater Missile Defense, (Wegman was a member of NAS-NRC authoring committee), Washington, DC: National Academy Press, 2001

Manuscripts Submitted but not yet Published

Wegman, E. J. and Dorfman, A. H. “Visualizing cereal world,” to appear Computational Statistics and Data Analysis


Solka, J. L., Wegman, E. J., and Marchette, D. J. “Data mining strategies for detection of chemical warfare agents,” to appear Statistical Data Mining and Knowledge Discovery


Papers Published in Peer Review Journals


Papers Published in Non-Peer Reviewed Journals or Conference Proceedings


Papers Presented at Meetings but not Published in Conference Proceedings

Wegman, E. J. (1999) Distinguished lecture series: Two lectures on data mining and visualization, Utah State University, Logan, UT


Wegman, E. J. (2001) “Data Reduction by Quantization,” Nonparametrics in Large, Multidimensional Data Mining Conference, Dallas, TX, January, 2001


Wegman, E. J. (2001) Five Lectures on Geometry, Visualization and Data Mining, University of Aalborg, Denmark, May, 2001


Wegman, E. J. (2001) Short Course on Statistical Data Mining, Interface ‘01, Orange County, CA, June 2001


Wegman, E. J. (2002) “Collaborative Visualization Environments,” (Best of JCGS Session), Interface 2002, Montreal, Quebec, Canada, April 2002


Scientific Personnel Supported and Awards

Edward J. Wegman, Awarded Army Wilks Award
Nkem-Amin "Martin" Khumbah, earned Ph.D.
Rida E. A. Moustafa, earned Ph.D.

Dr. Wegman was selected as a NSF/ASA/BLS Senior Faculty Fellow, Spring 2000. He was also elected Chair, Statistical Graphics Section of the American Statistical Association for calendar year 2000.

Scientific Progress and Accomplishments

During the period of performance, progress was made on several fronts. The “Grand Tour” algorithm was extended to hybrid data sets generated by image analysis algorithms. This will extend the possibilities of automatic data analysis and target hunting on a real time basis. It will enlarge the capabilities of the commander in the field when attempting to coordinate and organize incoming data of several types. The image to the right illustrates the ability to use this image grand tour technique to combine images from distinct spectral bands to identify minefields (the three near vertical rows of dots on the left hand side of the image), which are not visible in the individual images.

The utilization of D-optimal designs was made possible for robust estimation of multivariate location and scatter. The mathematical basis involves a penalty function approach for simplifying adaptive mixtures and density estimates. This has been completed and programming is being completed at the present time.

Finally, the new algorithms have been tested on several real data sets including the detection and enhancement of the behavior of dust behind moving vehicles in the field, and the effect of dust particles on in a networked virtual environment. See image at left. This is highly appropriate for soldier training in a virtual environment.

The most significant work, which represents the culmination of a long series of research efforts was our paper entitled “Visual data mining.” This paper describes a combination of tools developed through ARO sponsorship over an extended period of time and describes techniques for using these tools, including visual techniques for density estimation, rapid data editing, inverse regression, tree-structured decision rules, variable selection and dimension reduction, clustering, classification and discrimination, and outlier and unique event detection.
Parallel Coordinate Display used for Dimension Reduction in the Rapid Detection of Chemical Warfare Agents

Several of the most important papers mentioned above are reproduced on the accompanying CD.