REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
09-10-2008	Final Report	1-Jun-2005 - 31-Aug-2008

4. TITLE AND SUBTITLE
ADVANCED MATHEMATICAL METHODS FOR PROCESSING LARGE DATA SETS

5a. CONTRACT NUMBER
W911NF-05-1-0227

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Ronald A. DeVore, Peter G. Binev, Robert C. Sharpley

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of South Carolina
Office of Sponsored Programs & Research
James F. Byrnes International Center
Columbia, SC 29208 -

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
48711-MA.1

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Research on this contract was directed towards areas of mathematics and numerical computation which have applications to image/signal processing. The research can be broadly classified into the following areas: (1) compressed sensing, (2) sparse representation and encoding for digital elevation maps, (3) learning theory, and (4) high dimensional approximation. In addition to solving several fundamental mathematical questions in these areas, this work has developed numerical algorithms and software for encoding digital elevation maps which perform at the highest compression rates. Additionally, ground has been broken on new methods in the emerging field of compressed sensing.

15. SUBJECT TERMS
compressed sensing, sparse representation, DEM encoding, learning theory, high dimensional approximations

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Ronald DeVore
U	U	U	SAR		19b. TELEPHONE NUMBER 803-777-0086

Report Title

ADVANCED MATHEMATICAL METHODS FOR PROCESSING LARGE DATA SETS

ABSTRACT

Research on this contract was directed towards areas of mathematics and numerical computation which have applications to image/signal processing. The research can be broadly classified into the following areas: (1) compressed sensing, (2) sparse representation and encoding for digital elevation maps, (3) learning theory, and (4) high dimensional approximation. In addition to solving several fundamental mathematical questions in these areas, this work has developed numerical algorithms and software for encoding digital elevation maps which perform at the highest compression rates. Additionally, ground has been broken on new methods in the emerging field of compressed sensing.

List of papers submitted or published that acknowledge ARO support during this reporting period. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Binev, P.; A. Cohen, W. Dahmen, and R. DeVore, "Universal Algorithms for Learning Theory. Part II: Piecewise Polynomial Functions," Constr. Approx. 26 (2007), 127-152.

Daubechies, I; R. A. DeVore, C. Sinan Gunturk, and V. A. Vaishampayan, "A/D Conversion with Imperfect Quantizers," IEEE Transactions on Information Theory, Vol. 52, No. 3 (2006), 874-885.

Number of Papers published in peer-reviewed journals: 2.00

(b) Papers published in non-peer-reviewed journals or in conference proceedings (N/A for none)

N/A

Number of Papers published in non peer-reviewed journals: 0.00

(c) Presentations

Plenary talk (DeVore) at the workshop 'Approximation and Learning in High Dimension' at Texas A&M University, College Station, TX, October 19-21, 2007.

Plenary talks (DeVore and Binev) at the workshop 'Nonlinear and Adaptive Approximations in High Dimensions' at Physikzentrum Bad Honnef, Bonn, Germany, Dec. 10-15, 2007.

Invited talk (Binev) at the workshop 'Adaptive numerical methods for PDE's' at Wolfgang Pauli Institute (WPI) Vienna, Austria, January 21-25, 2008.

Number of Presentations: 3.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

DeVore, R., "Optimal Computation," Proceedings of the International Congress of Mathematicians, Plenary Address, Madrid, Spain (August 2006).

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 1

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Binev, P.; A. Cohen, W. Dahmen, and R. DeVore, "Universal Piecewise Polynomial Estimators for Machine Learning," Curve and Surface Design: Avignon 2006, (P. Chenin et al., eds.), Nashboro Press (2007), 48-77.

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts): 1

(d) Manuscripts

DeVore, R.; G. Petrova, and P. Wojtaszczyk, "Instance-optimality in Probability with an l_1-Minimization Decoder"

Barron, A.; A. Cohen, W. Dahmen, and R. DeVore, "Approximation and learning by greedy algorithms"

DeVore, R.; P. Binev, M. Hielsberg, L. S. Johnson, B. Karaivanov, B. Lane, and R. Sharpley, "Geometric encoding of natural and urban terrains," preprint.

Cohen, A.; W. Dahmen, and R. DeVore, "Near optimal approximation of arbitrary vectors from highly incomplete measurements," preprint.

Binev, P.; W. Dahmen, R. DeVore, and P. Lamby, "Sparse tree approximation in high dimension," preprint.

Cohen, A.; W. Dahmen, R. DeVore, and Clay Scott, "Classification using Reliable Sets," in preparation.

Number of Manuscripts: 6.00

Number of Inventions:

Graduate Students

NAME	PERCENT_SUPPORTED
Kanadpriya BASU	0.25
Pavel ZHELTOV	0.25
Entao LIU	0.05
John WEBB	0.05
Luke OWENS	0.25
FTE Equivalent:	**0.85**
Total Number:	**5**

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
Brendan LANE	0.83
FTE Equivalent:	**0.83**
Total Number:	**1**

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Ronald A. DeVORE	0.21	No
Robert C. SHARPLEY	0.17	No
Peter G. BINEV	0.21	No
FTE Equivalent:	**0.59**	
Total Number:	**3**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED
n/a	
FTE Equivalent:	
Total Number:	**1**

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period:	0.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:......	0.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:......	0.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):	0.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:	0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense	0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:	0.00

Names of Personnel receiving masters degrees

NAME	
Elizabeth PEREZ	
Total Number:	**1**

Names of personnel receiving PHDs

NAME	
Vesselin VATCHEV	
Luke OWENS	
Total Number:	**2**

Names of other research staff

NAME	PERCENT_SUPPORTED	
Matthew HIELSBERG	0.37	No
Jonathan WINDERS	0.17	No
FTE Equivalent:	**0.54**	
Total Number:	**2**	

Sub Contractors (DD882)

Inventions (DD882)

ADVANCED MATHEMATICAL METHODS FOR PROCESSING LARGE DATA SETS

ARO Contract W911NF-05-1-0227

Principal Investigator: Ronald A. DeVore
Co–Principal Investigators: Robert C. Sharpley and Peter G. Binev

Final Report
Scientific Progress and Accomplishments

STATEMENT OF THE PROBLEM STUDIED

Research on this contract was directed towards areas of mathematics and numerical computation which have applications to image/signal processing. The research can be broadly classified into the following areas: (1) compressed sensing, (2) sparse representation and encoding for digital elevation maps, (3) learning theory, and (4) high dimensional approximation.

SUMMARY OF THE MOST IMPORTANT RESULTS

1. Encoding Signals: Compressed Sensing

The classical paradigm for encoding signals is to model signals as bandlimited functions. This leads to the Shannon sampling theorem which says that a signal can be captured from equally spaced time samples provided the sampling is done at points at most $1/2\lambda$ apart where λ is the bandwidth. Many encoders and decoders are built based on this theory and in many cases sample at even faster rates (e.g. the Sigma-Delta modulation schemes). The problem with these encoders is that sensors cannot physically sample broadbanded signals at the necessary rate. On the other hand, most signals we are attempting to capture have much less information content than a general bandlimited signal. The question then arises whether we can build encoders that sample closer to the information rate of the signal than at their Nyquist rate.

The new field of compressed sensing is indeed addressing this very point. It shows that when a signal is sparse (when represented with respect to some basis) or more generally compressible, then it is sufficient to sample at the sparsity rate. This field is only now emerging and there are many analytic questions centering around how to do this sampling how to do the decoding, and what is the provable performance of such systems.

Our contributions to compressed sensing centered on developing precise measurements of performance of sensing systems and then finding which systems are optimal in performance. In [CDD] we have proven the best bounds for the performance of compressed sensing systems by comparing the performance with best k term approximation. We describe precisely when such a system can perform comparable to best k term approximation. This is determined by the number of samples (it must be slightly larger than k) and the types of samples (the compressed sensing matrix should satisfy a Restricted Isometry Property (RIP)). We also define a concept of best performance (instance optimality) in probability.

It is now well understood that the optimal matrices for compressed sensing are given by random processes such as Gaussian or Bernouli. It has been an outstanding question as to which decoders perform optimally when used in conjunction with random matrices. In [DPW1] we prove that ℓ_1 minimization is an optimal decoder for very general (subgaussian) random matrices. In [CDD1], we show that greedy algorithms provide almost optimal decoders for random matrices.

While random matrices are optimal for compressed sensing, they do not always merge well with applications where randomness may not be implementable. Therefore, there is great interest in deterministic constructions of compressed sensing matrices. In [D], we use number theory to construct what are the best performing deterministic systems that are known.

In another work, [BDDW], we have given a simple proof that certain random processes generate matrices satisfying RIP. This gives the most accessible verifications of the RIP for classical random matrices such as Gaussian or Bernouli ensembles.

2. Digital Elevation Maps

One of the focal applications of this research is the compression of Digital Elevation Maps (DEMs). DEMs are usually rendered as 3-D surfaces and image processing techniques are not appropriate for processing these maps. We have stressed the importance of developing data compression in the framework of new metrics (such as the Hausdorff metric) which incorporate the geometry in DEMs and are also more pointed to their intended applications.

Our research has been directed at two fronts. The first is to determine the Kolmogorov entropy in the Hausdorff metric for various model classes for DEMs. This has led to many interesting results [CDDD] for classes such as BV or piecewise smooth functions. To complete this direction, we want to incorporate more geometry into the model classes since we feel this captures the spirit of DEMs.

The second front of our research in DEM compression is directed at the development of algorithms and encoders for DEMs. Some of the desired features of the algorithms under development are: (a) high compression, (b) robust error handling, (c) progressive transmission of the data, (d) quick rendering, and (e) burning in (tunnelling) and line of sight display. Since almost all graphic hardware uses triangular polygonal patches as building blocks for object description, we focus our attention to algorithms utilizing meshes of polygonal elements. We have investigated several types of algorithms and encoders:

- Nonlinear approximation algorithms based on adaptive multiresolution analysis;
- Greedy (insertion or removal) algorithms for mesh construction which utilize Delaunay triangulation;
- Progressive encoding based on level sets.

The first algorithms include: (a) initial coarse adaptive triangulation which allows a low resolution good approximation, (b) wavelet decomposition of the function for achieving sparse representation of the function (surface), (c) conversion to hierarchical B-spline representation and application of the nonlinear uniform approximation scheme from [DJL] and [DPY], and (d) compression and progressive transmission of the data using the hierarchical representation.

The greedy removal algorithm is a recursive procedure with the following basic elements: (a) determination and updating of the significance table for the grid points, (b) removal one by one of the least significant points, and (c) mesh updating after each removal with Delaunay triangulation algorithm. The greedy insertion algorithm utilizes the same elements but in a reverse order. We pay special attention to the data structure that enables us to compress and transmit the data progressively.

The level set method seeks first to give a progressive description of the surface in terms of level curves and Morse trees. We prioritize the level curves and ridge curves and then encode each of them in a progressive manner. The remainder of the surface is then extrapolated from this information by blending or interpolation.

The level curves are compressed using a multiscale decomposition as described in [BDDD]. This paper also proves various theorems which prove the efficiency of this method of representing and encoding curves.

A major question that we are studying in detail is to understand which surfaces can be compressed well using level set methods. In this direction, we have introduced new anisotropic spaces of functions in [DPW] and shown that surfaces which are graphs of these functions can be compressed well with level set methods. These new anisotropic spaces are completely different from the anisotropic spaces usually studied in harmonic analysis and PDEs. For example functions with large gradients are in a certain sense nice functions with respect to this family of spaces. The correct description of these spaces when measuring higher order smoothness is still to be completely worked out. We believe that these spaces will play an important role in analysis, not only for surface compression but also for the study of nonlinear evolution equations.

The theory and algorithms behind our methods for surface compression are developed in [BDHJKLS]

3. Learning Theory

A typical application of surface processing is to generate a faithful representation of noisy point cloud data associated to a given surface. This can be viewed as a regression problem in learning theory where the unknown underlying probability distribution corresponds to the noisy data. The noise arises from sensor noise, sensor jitter, error in global positioning, misclassification of points on the surface, etc. We have developed in [DKPT] a general theory which describes when learning algorithms are optimal and gives the theoretical framework for creating optimal algorithms. In [BCDDT, BCDD, BCDD1] we have developed an adaptive algorithm (an alternative to model selection) which is shown to be optimal (in a certain sense) for learning the regression function from a given data set. This technology has been applied to learning surfaces generated in real time in the autonomous navigation of Micro Air Vehicles (MAVs) (see [KNPDBDS]).

A major question in the development of learning algorithms is the computational speed and whether they can handle streaming data. This is especially true in high dimensions where the curse of dimensionality can have a debilitating effect. Directed at this problem we have constructed and analyzed greedy algorithms for learning in [BaCDD] which are provably optimal in performance and computational speed.

A second area of learning theory that is important in many applications of signal and image processing is classification. We have developed a new mathematical theory for binary classification using reliable set in [CDDS]. The algorithms build a classifier from training sets by using set partitioning. We give bounds in probability on the performance of the classifier as compared to the Bayes classifier. We are now building practical classifiers based on sparse tree approximation and other recursive partitioning algorithms.

4. Beating the Curse of Dimension

One of the drawbacks of adaptive methods in learning and other application domains is that they are computationally expensive for high dimensional problems. For example if the Euclidean space dimension is d then partitioning just one cell into its children results in 2^d cells. So this is impossible to implement when d is larger than 20 or so. The usual method for circumventing this difficulty is to use kernel methods such as the Mercer kernels or support vector machines. We find these unsatisfactory on many problems since the representations are not local. We have tried to develop alternative methods to retain localness of the representation and to still treat high dimensions. Our results are in two directions: greedy algorithms and sparse tree approximation.

In greedy algorithms one seeks a representation of a function (signal/image etc.) as a linear combination of a few elements from a redundant family (called a dictionary) of waveforms. There is a long history to such algorithms. Our recent work [BaCDD] has identified the performance of such algorithms and showed how our analysis can be applied to learning problems to significantly cut down on the computational complexity of generating an approximation to the regression function.

The rough idea behind sparse trees is to only look at children in an adaptive partition that have data points. When the ambient space dimension is large there are only a few cells (determined by the size of the data set) which contain data points. We are developing [BDDL] a theory for the performance of sparse tree approximation and commensurate algorithms for their implementation. We are applying this technology to problems in meteorology together with scientists at the University of Maryland. In this application to long term weather forecasting the ambient space dimension is $d > 200$.

Bibliography

[BDDW] R. Baraniuk, M. Davenport, R. DeVore and M. Waken, *A Simple Proof of the Restricted Isometry Property for Random Matrices*, Constructive Approximation, to appear.

[BaCDD] A. Barron, A. Cohen, W. Dahmen, and R. DeVore, *Approximation and learning by greedy algorithms*, Annals of Statistics, **36** (2008) 64–94.

[BCDD] P. Binev, A. Cohen, W. Dahmen, and R. DeVore, *Universal Algorithms for Learning Theory Part II: Piecewise Polynomials*, Constructive Approximation **26** (2007), 127-152.

[BCDD1] P. Binev, A. Cohen, W. Dahmen, and R. DeVore, *Universal Piecewise Polynomial Estimators for Machine Learning*, Curve and Surface Design: Avignon 2006, (P. Chenin et al., eds.), Nashboro Press (2007), 48-77.

[BCDDT] P. Binev, A. Cohen, W. Dahmen, and R. DeVore, and V. Temlyakov, *Universal Algorithms for Learning Theory, Part I: Piecewise Constant Functions*, J. Machine Learning Research 6 (2005), 1297-1321.

[BDDD] P. Binev, W. Dahmen, R. DeVore, and N. Dyn, *Adaptive Approximation of Curves*, in APPROXIMATION THEORY: A volume dedicated to Borislav Bojanov, (D. K. Dimitrov, G. Nikolov, and R. Uluchev, Eds.) Marin Drinov Academic Publishing House, Sofia (2004), 43–57.

[BDDL] P. Binev, W. Dahmen, R. DeVore, and P. Lamby, *Sparse tree approximation in high dimension*, preprint.

[BDHJKLS] P. Binev, R. DeVore, M. Hielsberg, L. S. Johnson, B. Karaivanov, B. Lane, and R. Sharpley, *Geometric encoding of natural and urban terrains*, preprint.

[CDDD] A. Cohen, W. Dahmen, I. Daubechies, R. DeVore, *The Kolmogorov entropy of model classes of surfaces in the Hausdorff metric*, in preparation.

[CDD] A. Cohen, W. Dahmen, and R. DeVore, *Compressed Sensing and best k-term approximation*, JAMS, to appear.

[CDD1] A. Cohen, W. Dahmen, and R. DeVore, *Instance optimality in probability for greedy decoders*, preprint

[CDDS] A. Cohen, W. Dahmen, R. DeVore, and Clay Scott, *Classification using Reliable Sets*, in preparation.

[D] R. A. DeVore, *Deterministic constructions of compressed sensing matrices*, Journal of Complexity, **23** (2007) 918–925.

[DJL] R. DeVore, B. Jawerth, and B. Lucier, *Surface compression*, Computer Aided Geometric Design, **9**, (1992), 219–239.

[DKPT] R. DeVore, G. Kerkyacharian, D. Picard, and V. Temlyakov, *On Mathematical Methods for Supervised Learning*, J. FoCM **6** (2006), 3-58.

[DPW] R. DeVore, G. Petrova, and P. Wojtasczek, *Anisotropic Smoothness via Level Sets*, Communications in Pure and Applied Mathematics, **61**(2008), 1264–1297.

[DPW1] R. DeVore, G. Petrova, and P. Wojtasczek, *Instance optimality in probability with an ℓ_1-minimization decoder*, preprint.

[DPY] R. DeVore, P. Petrushev, and X. M. Yu, *Wavelet approximation in the space C*, Progress in Approximation Theory, Springer Verlag, New York (1992), 261–283.

[KNPDBDS] A. Kurdila, M. Nechyba, R. Prazenica, W. Dahmen, P. Binev, R. DeVore, and R. Sharpley, *Vision-Based Control of MicroAirVehicles: Progress and Problems In Estimation*, 43rd IEEE Conference on Decision and Control (2004), 1635–1642.