This paper is SAB Advisory only and does not reflect the opinions or policies of NOAA

Compressive Sensing

Introduction

NOAA's satellite and radar systems collect and transmit over 1TB of data on a daily basis. Processing increases the data volume to over 10TB daily. NOAA's National Marine Fisheries Service conducts optical fish surveys that acquire 10s of millions of images per year and 100s of hours of video.[1] NOAA's Unmanned Aircraft Systems Program conducts research on aerial sensor platforms that could greatly expand NOAA's collection and processing of environmental imagery. New techniques in data compression may offer opportunities at cost savings and might accelerate the introduction of emergent technologies that greatly enhance measurement capabilities.

Compressive sensing (CS) is a signal processing technique for efficiently acquiring and reconstructing a signal by finding solutions to under-determined linear systems. Its use is quickly emerging in signal/image processing for the purpose of data compression/recovery. It has been theoretically guaranteed that recovery of the information can be possible if the original signal and the measurement matrix satisfy some mathematical conditions.

A simple conceptual example of compressive sensing is the use of a balance scale to find one ball among a set of balls that is slightly lighter (heavier) than all the other balls. Instead of measuring one ball at a time, random combinations of balls can be measured. If each ball is numbered 1-12 the following 3 measurements could be made:

Measurement	Combination on Left Side of Scale	Combination on Right Side of Scale	Possible Result: 1 is light	Possible Result: 2 is light
1	1 2 3 4	4 6 7 8	Up	Up
2	1 4 8 9	2 3 11 12	Up	Down
3	3 7 9 12	1 2 5 10	Down	Down

In this example the number of measurements needed to infer the answer is reduced by over 70%. In image processing a similar conceptual approach can be taken. Although images don't typically have one unique pixel, random sets of pixels can be measured and analyzed to recover the image. The number of measurements required is far fewer than observing and analyzing each pixel saving cost on instrument design, power and data storage needs.

Key Questions

Should compressive sensing be in the mix as NOAA contemplates future observing systems?

On-board processing and data compression are among key requirements that will be important for near to long-term remote sensing needs at NOAA. Given these potential benefits of CS NOAA should have a strategy for investigating the technology further, determining which of its observing needs now and in the future offer the greatest opportunity (in terms of science, cost, complexity, risk etc) for application of this technology, and what further research and development would be needed.

Some advanced machine learning methods are alternative approaches for data compression including online non-negative matrix decomposition, sparse coding and online dictionary learning. In contrast to the traditional dimension-reduction methods such as principle component analysis (PCA), the data vector can be generally represented as a weighted linear combination of a small number of (unknown) basis vectors , and the over-complete basis set can also be learned from the historical data. Algorithms of such methods are very efficient for dealing with large-scale data since these algorithms can process the streaming real-time data. These include:

Online non-negative matrix decomposition/factorization (ONMF): unlike conventional NMF solutions which require the entire data matrix to reside in the memory, the ONMF algorithm proceeds with one data point or one chunk of data points at a time. ONMF has been successfully applied in document clustering, image representation, and other domains to efficiently handle very large-scale and/or streaming datasets.

Sparse coding: provides a class of algorithms for finding succinct representations of stimuli; given only unlabeled input data, it learns basis functions that capture higher-level features in the data. When a sparse coding algorithm is applied to natural images, the learned bases resemble the receptive fields of neurons in the visual cortex.

Sparse dictionary learning : a method which aims at finding a sparse representation of the input data (also known as coding) in the form of a linear combination of basis elements as well as those basis elements themselves. This method is applied to data decomposition, compression and analysis.

How mature is the technology?

CS is a relatively mature technology in terms of its theoretical development. While there is a growing list of applications, more opportunities remain. Further sensor development and work on complex decompression methods are needed.

The acquisition and processing of signals have until recently been governed by the Nyquist Theorem. This theorem states that given a dense array of sensors a sampled signal can be exactly recovered if the sampling rate is equal to at least twice the highest frequency present in the signal. Although this has enabled important advances in signal processing such as digitization, sampling at the Nyquist rate limits further innovations because a large amount of data has to be acquired to accurately capture the signal before it can be compressed (Baraniuk et al 2014).

Although the theoretical roots for CS can be traced back to the 1800's, it was not until 2004 that researchers (e.g., Cande's et al 2004) introduced a workable theoretical framework for CS. This research established that accurate signal recovery is possible for signals that have the property of sparsity. This requires that the undetermined matrix equations used to solve or recover the signal from compressed measurements has a small number of nonzero elements. Conceptually sparsity refers to systems with elements that are loosely coupled within a given domain. The key result of Cande et al 2004 was that under certain mathematical conditions (e.g., sparsity), signal acquisition and reconstruction could be achieved at far less rates (or measurements) than required by the Nyquist theorem. Since then research has mostly centered on how many simultaneous or multiplexed random measurements are necessary to reconstruct the original signal as well as work on the solvers needed to do so. Effort has also focused on sensor design. CS offers the greatest advantage where the sensor arrays are difficult or expensive to fabricate.

The following are applications where CS has had an impact.

- **Mobile Devices**: Compressed sensing is used in mobile device imaging to reduce the number of images required (based on the Nyquist rate) for cell phone videos. This results in a reduction of the A/D conversion required as part of the standard digital compression process, and thus the energy requirements/battery usage. The energy per image reduction is as much as a factor of 15. However complex decompression algorithms are needed and may require an off-device implementation because of the computational demands (Schneider 2013).

- **Medical Imaging**: CS had significantly increased imaging (x7) while not compromising diagnostic quality. It has done so by measuring fewer Fourier coefficients (Lustig et al 2007).

- **Single-pixel cameras from Rice University and Bell Labs (a lensless single-pixel camera)**: This technology enables still images using repeated snapshots of random gridded exposing a single sensor.

Other areas where CS is used includes holography, tomography, facial recognitions, and radio astronomy (Liutkus et al 2014)

Are there parts of NOAA's mission for which compressive sensing is more likely to provide significant opportunities than others?

Systems that are based on infrared imaging radiometry are low hanging fruits for CS application at NOAA. This includes sensors that provide information on the vertical structure of temperature and constituent gases such as moisture. Interferometry is fundamental to these sensors, and has the natural advantage of possessing characteristics that are important for CS. Some of these characteristics include multiplexing which allows for reduced data acquisition or the acquisition of a compressed signal and sparsity in the signal, both of which are essential conditions for recovery of the signal.

A challenging issue for application in NOAA observations is how to appropriately incorporate the specific domain knowledge of the satellite remote images into established CS methods. Specifically, how to ensure the potential losses due to the CS compressing processes do not compromise other specific or broad scientific applications. A hybrid approach could be taken in initial adoptions of CS where both traditional sensing approaches are combined with CS.

Is NOAA taking advantage of related technologies and approaches such as pre-processing at the site of data collection for reducing the cost of data transmission?

Currently there is not obvious evidence of any intentional application of CS at NOAA. However our research was not sufficiently exhaustive to determine what if any techniques are being employed to reduce the cost of data transmission.

Potential Speakers

Mark Baraniuk

Emmanuel Candès

Justin Romberg

Terence Tao

David Donoho

References

Baraniuk, M. A. Davenport, Marco F. Duarte, Chinmay Hegdehttp://legacy.cnx.org/content/col11133/1.5/

Cande`s, E. J., Romberg, J. & Tao, T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE T. Inform. Theory 52, 489 (2006).

Liutkus, A. et al. Imaging With Nature: Compressive Imaging Using a Multiply Scattering Medium. Sci. Rep. 4, 5552; DOI:10.1038/srep05552 (2014)

Lustig, M., and Donoho, D. and Pauly, John, Magnetic resonance in medicine, 58(6), 1182-1195 doi:10.1002/mrm.21391 (2007)

Schneider, D. "New Camera Chip Captures Only What It Needs". IEEE Spectrum. Retrieved 2013-03-20 (2013).

Additional Reading

Nuit Blanche A blog on Compressive Sensing featuring the most recent information on the subject (preprints, presentations, Q/As)

Wiki on sparse reconstruction

SigView, the IEEE Signal Processing Society Tutorial Library