Compressed Sensing - A New Mode of Measurement

Research · December 2015
DOI: 10.13140/RG.2.1.3697.5449

1 author:

Thomas Vogt
University of South Carolina

To be published in “Reasoning in Measurement”
ISBN 9781848936027 Pickering & Chatto May 2016

Compressed Sensing – A New Mode of Measurement

Thomas Vogt
University of South Carolina

Abstract

After introducing the concept of compressed sensing as a complementary measurement mode to the classical *Shannon-Nyquist* approach, I discuss some of the drivers, potential challenges and obstacles to its implementation. I end with a speculative attempt to embed compressed sensing as an enabling methodology within the emergence of data-driven discovery. As a consequence I predict the growth of non-nomological sciences where heuristic correlations will find applications but often bypass conventional pure basic and use-inspired basic research stages due to the lack of verifiable hypotheses.

Introduction

Making digital representations of physical objects has been approached with a pessimistic attitude demanding a very high rate of regularly-spaced measurements without taking into account that the object itself might have sparsity. In this text sparsity is used as an operational gauge of an object’s complexity rather than a well-defined mathematical property. In mathematics we define a sparse matrix in contrast to a dense one as containing mostly zeroes. Compressed sampling takes into account the sparsity of an object and is able to successfully reconstruct images even after dramatically reducing the number of measurements required without loss of reconstruction fidelity. If one defines ‘sampling’ as the act of performing measurements of different objects such as pixels in an image, one can conservatively measure one at a time or group several objects and measure groupings. Compressed sensing allows us to optimize measurements of such groupings and thereby perform significantly fewer measurements. This is related to the 12-ball problem where one is tasked with finding 1 lighter or heavier ball out of a set of 12 balls by only 3 comparative weighing of groupings.[1]

Compressed sampling will not replace conventional sampling but complement it by making measurements possible which before were prohibitively costly. Cost in this terminology refers to measurements that are too expensive, take too long, require too much energy or subject the object to damage when measuring in the conventional linear fashion.

Shannon-Nyquist measurements and compressed sensing.

Harmonic analysis has shown that signals we measure can be described as convolutions of mathematical basis functions with frequency and intensity as variables. The best known example is the *Fourier series* where a signal can be decomposed into a sum of simple oscillating functions such as sines, cosines or complex exponentials multiplied with a weighting function. Signals can then be represented by their Fourier coefficients. Mathematics provides us a plethora of basis functions such as wavelets, ridgelets, curvelets and contourlets to efficiently approximate signals.[2] The conventional approach when measuring signals such as sounds and images relies on Shannon's theorem which states that the sampling rate with which one should record or image must be at least twice the maximum frequency present in the signal in order to record, transmit and reconstruct with high fidelity. This frequency is called the *Nyquist rate*[3] and the *Shannon-Nyquist theorem*[4] assures us that our sampling is dense enough to allow us to reconstruct original analog signals such as a Coltrane saxophone solo or a lake view. This immensely powerful theorem is used in most consumer audio and video electronics, in conventional analog-to-digital (ADC) converters and in medical imaging i.e. ultrasound and magnetic resonance imaging (MRI). The latter points to the importance of the quality of reconstruction since overlooking or altering even minute features in images might have important consequences as they are often the goal of such measurements. Intuitively one would think that measuring faster and omitting data from the reconstruction will always result in a less faithful reconstruction of the original signal and should therefore be avoided at all cost. However, as I will show below, the amount of data that needs to be measured, stored and transmitted impacts the feasibility of a measurement and is often constrained by external 'cost functions' such as the speed with which one can image and reconstruct and therefore actively control a process, the time a patient is scanned in a computer-aided tomography (CAT) or MRI procedure or the amount of energy needed for a measurement, to name just a few. These cost functions drove the exploration of new modes of measurement to circumvent the shortcomings mentioned above and others discussed later. Compressed sensing allows measurements that optimize external cost functions by measuring faster and less without a commensurate loss of reconstruction fidelity.

In the *Shannon-Nyquist* measurement mode we uniformly and linearly measure our signal often by sequential line scans ('raster scans'). This results in a lot of data points N as we all realize when confronted with the storage capacities of our digital cameras. Therefore, we use programs based on compression algorithms which extract a subset $K << N$ that then stores our soundscape as MPEG and our images as JPEG files. Compression is a non-linear process and relies on the fact that many of the Fourier or wavelet transform coefficients with which we can describe our signals are small or close to zero. In the case of JPEG or JPEG-2000 the signal is represented in a mathematical basis space different than the pixel basis, relying on the fact that most images have many zeroes in the discrete cosine (JPEG) or wavelet (JPEG-2000) basis. The conceptual leap compression algorithms are based on is that we do not need to store zeroes or minute coefficients in order to be able to reconstruct our signal to a certain degree of fidelity. This shows that not all frequencies and all pixels or voxels have the same importance when we reconstruct the original signal. Some are more important than others and sparsity might provide us with an opportunity to escape the constraints of the *Shannon-Nyquist* theorem and reduce the number of coefficients we need to store and be able to reconstruct our original signal. In other words: some data sets that we measured uniformly at the *Nyquist rate* turn out to be sparse. We can therefore compress them for data transmission and storage purposes. Sparsity is thus a working measure of the complexity or lack thereof in a signal representation. Many signals are compressible by using some known transform coding scheme such as JPEG or JPEG-2000 which creates sparse representations in the transform basis. It is important to realize that a Mark Rothko color field painting is sparse whereas a totally random screen test image is not. There is a lot more information that cannot be compressed in the screen test image since there are significant and random pixel to pixel variations.

With this coarse understanding we can summarize the conventional paradigm for digital data acquisition:

We uniformly sample data at the *Nyquist rate* and obtain N data points, which we then compress to K data points with K<<N using some threshold value of the data intensity and subsequently transmit or store this sparse data set. Using appropriate algorithms we can decompress the reduced data sets with K data points back to N data points when we reconstruct the signal. This works quite well. There are well-known examples that show how close we get to reconstructing our originally sampled image after we omit 97.5% of all wavelet coefficients in the compression step.[5] However, as Mark Rothko and John Coltrane aficionados will tell you there are differences in perception and quality when omitting wavelets coefficients with small values and subtle differences of hue and sounds can and will be noticed by the trained eye and ear. Humans can distinguish between 2.3 and 7.5 million colors [6] and about 340,000 tones.[7]

Apart from desires to remain as close as possible to a uniformly sampled signal the question begs in particular for most commoditized images and sound recordings if instead of measuring lots of data and then compressing them, one could not attempt to measure only the data points which have significant wavelet coefficients above a pre-set threshold and omit the rest. The concept of compressed/compressive sampling/sensing [8] follows this approach and has proven that one *can* measure only a compressed data subset or something close to it and then find a way to reconstruct the signal N>>K. One is no longer measuring at or above the Nyquist rate but significantly below it since K<<N. This represents a radical departure from the traditional mode of measurement. While on first sight it appears to be audacious, it has proven to be mathematically sound in the sense that it is highly probable given a sparse data set that one can reconstruct the image based on random subsets of data. These two pillars of

compressed sensing, namely sparsity and random measurements, are crucial for the method to work. In compressed sampling one thus directly measures compressed data M<<N and thereby drastically reduces the amount of data needed to measure a signal *if* one finds a way to reconstruct the total signal. More simplistically formulated: why measure the zero or below-the-threshold wavelet coefficients if "all" we have to do is measure the non-zero ones and then find a way to "add the zeroes back in". This appears very counterintuitive: how can one know what the compressed subset (the JPEG or MPEG) is before measuring it? Isn't the compression to a JPEG an analysis step one needs to do before one can omit data points with low or zero wavelet coefficients? How do I know which pixel is important and which is not? How do we find the "non-zero coefficients"? Very complex mathematical details show that random measurements allow the acquisition of "compressed" data followed by a recovery relying on a technique called the *L1 minimization*. This procedure will provides us with the sparsest solution at a very high probability. The sparser the data are the higher the probability that we will recover the "uncompressed data" – without ever measuring them! The *L1 minimization* is a linear process and we have thus turned the conventional *Shannon-Nyquist* measurement philosophy on its head: we now measure random, non-linear subsets of the data and recover the uncompressed data using a linear process (*L1 minimization*). Of course this all hinges on the sparsity of our data since the mathematical core concept of compressed measurements is based on the realization that sparse signals can be recovered exactly with a high probability despite the fact that we are dealing with underdetermined linear systems of observations due to measuring below the *Nyquist rate*. Despite its counter-intuitive nature compressed sensing is based on sound and uncontested mathematics. While accepted in the sciences and engineering communities this new measurement mode will be difficult to explain to a lay audience and therefore be contested in the public arena in particular when there are important consequences of such measurements. The early detection of cancer and other diseases based on measurements of only a subset of what can be measured will require justification and reassurances based on, as I will show below, non-fiscal cost functions such as radiation exposure or increased accuracy when imaging moving objects.

'*Measure what can be measured and make measurable what cannot be measured*' is a quote attributed to Galileo Galilei and has become the paradigm of the conventional *Shannon-Nyquist* measurement philosophy. Generations of scientists have and will continue to improve measurement devices such as telescopes, analog and digital cameras and electron microscopes to 'make measurable what cannot be measured'. Minute signals indicating new phenomena are being confirmed after long periods of data acquisition and careful analysis. This mode of measurement will not be replaced by compressed sensing. Rather compressed sensing will complement the *Shannon-Nyquist* mode of measurement and allow new types of measurements previously not possible to be done. "*Measure what should be measured*", a phrase coined by Strohmer[9] describes the operating philosophy of compressed sensing. In the conventional approach all data points are measured without taking into account the sparsity of the signal whereas compressed sensing advocates to measure only the most important subset using random measurements in order to increase data flow, reduce storage requirements or required detector coverage. The need for compressed sensing and how it can be accomplished will be illustrated below using the paradigmatic example of the single-pixel camera.

How sparse, how good?

The sparsity of signals allows the freedom to set threshold values in order to define what the cutoff intensity of the basis function is. This brings up the question of who needs or wants uniform sampling above the Nyquist limit and who will get to listen to a Coltrane saxophone solo recorded using compressed sensing or looks at a Mark Rothko painting with altered hue, or has his brain scanned less and faster? The important word in Strohmer's dictum is 'should'. We need to remind ourselves that compressed sensing works best with sparse data. In the case of imaging a screen test image which does not have a high degree of sparsity, a conventional measurement at the *Nyquist frequency* should be the choice of measurement mode. Many natural images are quite sparse and standard wavelet decomposition reveals often that most coefficients are actually very small. "Natural" images are highly structured and can be very efficiently represented using sparse representations.

Human perception is a very complex process with its own ‘biological’ threshold values and in many cases the aspect of quality is difficult or even impossible to parameterize. The perception of color in a Mark Rothko painting serves as an example: the human eye’s response to light spans from about 400 to 750 nm and varies with age. We have a standard for colorimetry that dates back to 1931 and was devised by the Commission Internationale de l’Eclairage (CIE) which assigns two coordinates, called CIE coordinates to a particular color based on a photoluminescence measurement. In stark contrast the standard method of characterizing the *quality* of light is to rely on a color rendering index which specifies how well a light source can illuminate, or render, the true color of an object. The color rendering index is determined by *comparing* the differences in the CIE color coordinates of an object illuminated first with a test source and then separately with a black body having the same correlated color temperature[10]. This complex process reveals that the parameterization of quality often relies on comparison using human perception and can become a question of preference as in the case of the audiophile and art connoisseur, who resists the notion of a particular signal as being sparse. In the lighting industry lamps have different color rendering indices depending on if they are targeted for the US or Asian market as the interaction with the skin complexion is an important quality factor. The US initial response to fluorescent lighting was that it was “seen as cold” in comparison to the more energy inefficient incandescent light. There are thus external factors to a measurement that can play an important role in the decision which mode to use. These factors can often be best understood as external cost functions. In order to describe some of them and understand why they play such an important role in the implementation and use of compressed sensing I will now describe the single pixel camera which is a paradigmatic compressed sensing device.

The single pixel camera – a compressed sensing device

An important aspect in the future design of sensing and measurement devices will be the hybridization of hard ware and reconstruction algorithm as epitomized in the so-called single pixel camera developed by Richard Baraniuk from Rice University.[11]

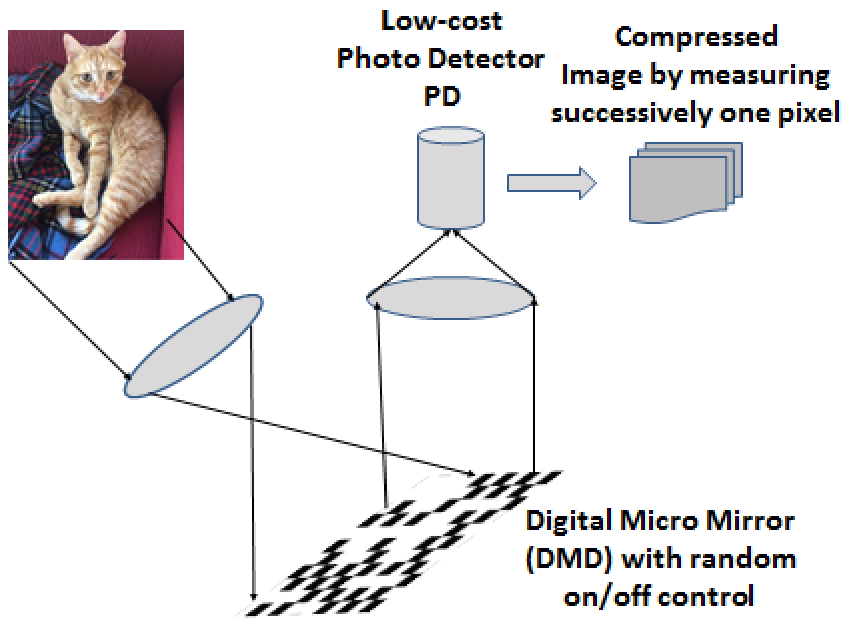

Figure 1: Operation of a single pixel camera: An image is projected onto a digital micro mirror (DMD), half of whose mirrors are switched randomly off. This image is then directed towards a second lens behind which a single photo detector (PD) registers the measurement. The mirrors are switched randomly and each time the PD measures one pixel. The single-pixel camera thus creates a picture by measuring just one pixel over and over again which allows the recovery of the image.

Image from http://phys.org/news143210026.html

The mathematical concept of random measurements needed for compressed sensing is incorporated in this new type of measurement device as follows: the light making up an image is focused from an array of mirrors and from there into a single detector. This arrangement resembles a projector running backwards. The mirror array consists of over 7000 little mirrors the size of a sand grain that can each be addressed in less than 10 microseconds either to reflect light into the detector or not. If one now randomly addresses groupings of half of the mirrors to reflect light into the mirror and the other half not and measures the intensity in the single detector a few hundred thousand times one is able to reconstruct the image from which the light comes. Each measurement contains information coming from half of the mirrors. A one megapixel image would therefore require approximately 100,000 single pixel measurements representing only about 100 kilobytes of memory – a reduction of 90% in the number of measurements and with that measurement time as well as storage space.

The 'cost function' driving the use of compressed sensing

Compressed sensing has become *the* way to record data when an external 'cost function' prevents or severely limits measurements in the classic *Nyquist-Shannon mode*. One way this 'cost function' manifests itself is by the fact that certain detectors can cost upwards of a million dollar. To measure i.e. infrared images using compressed sensing one can thus use only a very small single pixel camera which enables measurements that would have prohibitive costs if measured in the *Nyquist-Shannon mode* where one would use a large detector to record as much signal as possible. Another way a cost function can promote this new type of measurement mode is due to the fact that in medical imaging one needs to minimize exposing patients to harmful ionizing radiation such as electrons and x-rays. Current estimates suggest that 3-5% of all new cancers diagnosed in the US are due to CAT scans.[12] Drivers for this extensive use are the need to pay for these expensive instruments and the still prevalent philosophy of "measure what can be measured" endorsed by doctors and patients alike. Less exposure also reduces the data collection time and can thereby enable certain measurements: reducing the time patients need to hold their breath or remain still during an MRI scan. Compressed sensing will provide less perturbed images and is therefore already being used in pediatric medicine since it significantly reduces artifacts due to uncontrolled patient movement.[13] The amount of data reduction in MRI imaging by a factor of 2-10 comes with little or no impact of the quality of images. [14] In the case of brain imaging 3D images can be obtained in as little as 32 milliseconds.[15]

Measuring only 10% or less of the signal will consume less energy, an issue important in space exploration and the deployment of autonomous sensor arrays for environmental or nuclear non-proliferation monitoring since it will increase the time period a probe can operate and send signals. In the short term many new applications in the design of active and autonomous sensing devices will drive the implementation of this new mode of measurement. Less energy use allows for significant further miniaturization and cost reductions. The 'SeeChange' technology described in Dave Egger's novel[16] *The Circle* is technologically plausible using compressed sensing as a measurement mode.

Challenges to compressed sensing

With these examples of cost functions in mind we can now discuss some challenges to the use of compressed sensing. Initially the counterintuitive nature of compressed sensing can lead to a technically unfounded but very strong dismissal of this new mode of measurement. In the February 2010 edition of Wired Magazine compressive sampling was introduced under the title "Fill in the Blanks: Using Math to Turn Lo-Res Datasets into Hi-Res Samples".[17] Certain reader comments on the website display a lack of understanding of the general concept by assuming that common interpolation and not sophisticated and well established mathematical principles as outlined above are the basis of this method. This rejection is not so much a concern for the general public if one is imaging objects in intergalactic space or storing pictures of a birthday party. However, when using MRI which is now the premier diagnostic tool to detect cancer and vascular disease the '*measure what can be measured*' attitude is passionately evoked. Two quotes from the "Wired" website express this sentiment:

> "*Can you believe that someone would have such a fundamental misunderstanding of basic mathematics and information theory that they would base a medical diagnosis on features produced by data interpolation? I hope it's not my doctor doing it. Prettying up pictures is great. Looking for tumors, etc. is insane. By definition, you're looking for an aberration, which, by definition, this algorithm would not produce." (Iamorpa)*

> *"Would you be willing to gamble your life on interpolated data where the shortcoming would be missing clinical pathology? A MRI image contains many subtle shades of gray in abstract shapes. For an artistic image clear, sharp edges and vivid colors may enhance, but to render a line on a*

diagnostic image that appears to abruptly end and restart may draw a complete artery where there is really an occlusion or to smooth out faint variations representing a brain tumor can be deadly. I'll pay for the long scan please" (Grego)

It is important to recognize that compressed sensing can be rejected even by highly trained specialists who will not immediately take advantage of better or faster measurements available. As an example take a radiologist's unique skill set that allows him to scan huge amounts of data very fast and recognize and identify minute deviations. When varying the contrast in order to measure faster and expose the patient to less radiation such a highly trained specialist will no longer be able to identify anomalies and will insist on looking at the images with the same contrast settings they were used when he was trained – faster imaging and less radiation exposure is not always better when a significant amount of training and tacit knowledge to identify anomalous features relies on a specific contrast or image quality.

Clearly there will be early and late adopters of compressed sensing. The cost function be it money, radiation damage or speed with which one can measure will incentivize many to use this transformative imaging mode. In medical imaging using MRI, CAT or ultrasound devices this shift to taking advantage of compressed sensing is currently in progress. In the future this new mode of imaging will become the new standard in particular when it is also used to train the tacit skills of medical specialists.

'Big Data Science'

As highlighted when describing the "single pixel camera" a significant reduction in measurements required to reconstruct an image by combining measurement and analysis modes will enable us to continue to cope with large data flows and storage requirements which do not increase commensurately with the data flow. Using compressed sensing we are able to measure a larger portion of the information content available by pushing back the limits of Shannon-Nyquist measurements.

In large scale, complex scientific experiments which produce enormous amounts of data compressed sensing is already the mode of choice due to the need to "pre-select" events. The bottleneck in many measurements used to be at the detector level; now the commoditization of detectors which can be assembled into large arrays has moved the bottleneck to the processing/transmission/storage part of the measurement. The compact muon solenoid (CMS) detector at the Large Hadron Collider at CERN, the European nuclear science laboratory, produces 320 terabits per second of data.[18] Using a hardware-based triage, "triggers" select only about 800 gigabytes per second which will be characterized as "interesting events" and subsequently analyzed.

Due to the appealing conceptual simplicity of compressed sensing's 'second pillar', random measurement capabilities can directly be implemented in the design of the detector as described above for the single-pixel camera. This hybridization of measurement and analysis is needed to be able to measure the gargantuan amounts of data without ending up with data flow bottlenecks and not enough storage space. Integrating sensing and data processing this way is "to bring mathematics into the lens" as Candes and Tao point out.[19] An example highlighting the tremendous advances being made in data flow is the Sloan Digital Sky Survey[20], which started in the year 2000 and in its first few weeks collected more data than had been amassed during the entire history of astronomy. It will be surpassed by the Large Synoptic Survey Telescope[21] in Chile which will measure 140 terabytes of data every 5 days.

Building on the advances and lessons learned in high energy and astronomy experiments compressed sensing will be employed in many other areas where the cost function is driven by the size of the already existing and growing data avalanche. The use of compressed sensing is growing rapidly and new devices are being designed and built. In analog-to-digital (ADC) conversion technology a receiver/ADC chip, the "Random Modulator Pre-Integrator" (RMPI) has been built as one of the first electronic hardware devices based on the compressed sensing paradigm which will replace the conventional ADC in appropriate applications.[22]

A consequence of this new mode of measurement is that data are no longer autonomous and complete because compressed sensing measurements rely on the premises of sparsity and measure only random subsets of all possible measurements. This will rekindle important discussions in the epistemology of measurement and change our view of what a discovery is. Discovery is no longer a 'eureka moment' but a statistical war of attrition and the unease about increasingly theory-laden experiments will require a modification of what it means to discover something new. Serendipitous discoveries from very large data sets will in many cases be almost impossible since these particular events might be excluded from analysis precisely due to the fact that they do not exist within the current framework of knowledge at the time of measurement. One such framework of knowledge is the Standard Model in Physics describing the occurrence and relationships of elementary particles. This model predicted the existence of the Higgs boson in order to explain the particular masses of elementary particles. Its discovery was based on looking for certain decay pattern allowed with the Standard Model of Physics and serves as a paradigm changing example of a theory-laden, large dataflow discovery. This type of measurement will become dominant in high energy physics and astronomy where data intensive experiments rely on a library of known and simulated processes to trigger further data analysis.

The continuous sampling of physical objects will amass enormous amounts of digital data as representations which can be re-analyzed at later stages using new algorithms and methodologies. We can learn something new using existing data and new algorithms. Antonioni's 1966 movie *Blowup* portrays the discovery of a body and a person with a gun after progressive enlargements of an existing picture. We need to be mindful, however, that there is a difference between having compressed data or used Shannon-Nyquist based measurements and simple experiments with Photoshop™ using raw and JPEG representations reveal striking differences when subsequently manipulating contrast or other image parameters. Software epistemology will need to be better understood as it can redefine what discovery means in such a context.

Compressed Sensing and Data-Driven Discovery – A New Mode of Inquiry

Stored data and information are now almost completely digitized: in 2013 the amount of globally stored information is estimated to be about 1,200 exabytes, of which 98% is in digital form. In 2007 the amount of data generated worldwide was larger than what we could store and transmit. In 2011 we measured twice as much data as we could store. The amount of data generated globally is growing by about 58% a year. [23]

At the same time a new strand of scientific inquiry, data-driven discovery, is establishing itself next to the traditional hypothesis-based mode. The measurement mode used to collect data is an important meta-property that needs to be taken into account when discussing results of such data-driven discoveries. There are reasons to suggest that this data-driven discovery mode will in the long run lead to fundamental changes in the daily practices and philosophical grounding of many sciences. This type of 'big data' analysis will be, at least initially based on the discovery of correlations which in many cases cannot be established unequivocally as causalities. A phenomenon identified by correlations might be a useful proxy for a highly probable prediction (i.e. flu tracking by Google) but shed no light on the reason for its existence or inner workings. A resurgence of devising heuristic rules and proxies in the physical sciences to correlate with desired physical (i.e. superconductivity, thermoelectricity) or chemical (i.e. catalytic) properties to explore vast parts of parameter space in chemical compound space are used in materials science ('materials genome initiative'). Simple proxies for properties are sought that can circumvent the traditional first-principles route using density functional theory or molecular dynamics in order to accelerate and enable new discoveries.

Filling Peterson's Quadrant – the growth of non-nomological science

This increased use of heuristics might result in the emergence of new non-nomological research in various disciplines in particular the social sciences and economics. A convenient way to categorize types of scientific knowledge employs the scheme developed by Stokes[24] which assigns individual quadrants for pure basic (Bohr), use-inspired (Pasteur) and pure applied (Edison) research. The first quadrant contains research with little considerations for use and fundamental understanding. As Stokes writes:

> 'This quadrant includes research that systematically explores particular phenomena without having in view either general explanatory objectives or any applied use to which the results will be put…'(Stokes[21])

Stokes named this quadrant Peterson's quadrant to point out that Peterson's *Guide to the Birds of North America* exemplifies such systematic research.

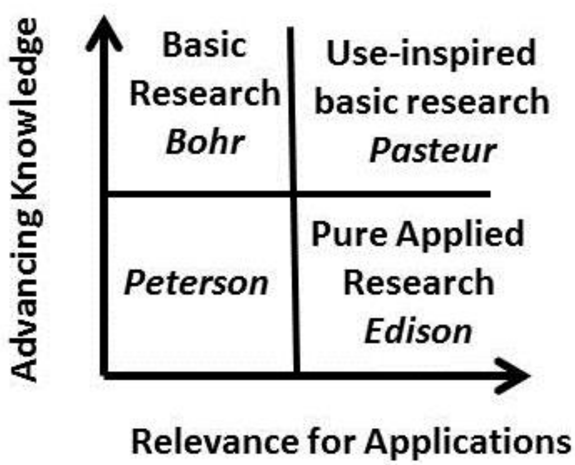

Figure 2: Stokes' Quadrants (18)

Within the framework of this classification scheme one can describe dynamic pathways in which research progresses from one quadrant to the next. The progression from Bohr's over Pasteur's to Edison's quadrant can be used in retro perspective to describe various phases in the evolution of modern electronics from quantum mechanics to the modern computer chip. An important transition from Peterson's to Bohr's quadrant is typified by Charles Darwin's theory of evolution described in *The Origin of Species*. Darwin relied on a vast amount of not understood correlations as well as many disparate observations. These collations of observations in Peterson's quadrant can thus lead to important pre-cursors of theories, proxies and correlations that can find applications in pure applied research (Edison's quadrant). Another important role of this quadrant as pointed out by Stokes is education and the development of skills. The availability of large amounts of digitized data triggered the development of *Google*'s page ranking algorithm, unstructured data bases and many other problems needed to be solved to correlate observations using large heterogeneous data sets. These methodological problems are in themselves worth pursuing as they can sharpen our tools to "dig deeper" into certain types of data. In analogy to the above mentioned scenario from the film *Blowup* the existence of large data bases will allow us to develop new algorithms and methodologies that could lead to the discovery of new features and correlations in already existing data and allow their benchmarking.

Subsequently a complementary hypothesis-driven inquiry mode might be able to devise measurements to verify whether the heuristic rules and proxies developed in Peterson's quadrant are useful and can be reduced to causal chains. While large data correlations might point towards some new, unforeseen and unexpected hypotheses many such correlations are likely not to be reducible to discipline-specific models and causation chains. Nevertheless, these correlations might still be useful as rules and become test cases for meta-theories used for instance in complexity science where strong causality is often suspended in favor of a more heuristic and empirical view of the sciences.

Some correlations might be used to develop 'toy models' rather than complex and very detailed models with too many parameters. An example for a 'toy model' theory is the theory of 'self-organized criticality' developed by Per Bak.[25] The power of these oversimplified models lies not in the prediction of individual events but in the description of complex systems such as financial markets, forest fires, earthquakes and the size of avalanches in a sand pile. These types of theories might undergo a renaissance as our data bases grow. As Peter Norvig[26] writes: 'Simple models and a lot of data trump more elaborate models based on less data'.

Understanding the world without relying on hypotheses will become a new type of exploration based on probabilities and correlations. Non-causal analyses will help us see the world within a context of 'what' and not 'why'. While we have an intuitive desire for causal connections many fields of exploration in particular in the 'soft' and idiographic sciences will be driven by the discovery of correlations. It is too early to assess how the shift from causation to correlation will change the operational pragmatism in different scientific fields.

In conclusion, compressive sensing allows us to keep up with the in the foreseeable future continuing massive flow of data and enables, as a complementary tool the measurement of sparse data when an external cost function (i.e. price, speed, stability) makes measurement, transmission and data storage in the classical Shannon-Nyquist mode no longer feasible. It will therefore make a significant contribution to our digital universe and provide ample research data for Peterson's quadrant which can fuel Bohr's, Pasteur's or Edison's quadrant by classical hypothesis-driven inquiries or use heuristic rules and proxies to enable pure applied research where correlation can be enough to predict the behavior of large complex systems without an underlying causal model.

Thomas Vogt is the Educational Foundation Distinguished Professor of Chemistry and Biochemistry at the University of South Carolina. His research focuses on the structural characterization of new materials using x-rays, neutrons and electron diffraction and direct imaging using electron microscopy. His philosophical interests are in the history and philosophy of science, epistemology of measurements and philosophy of chemistry. He is a Fellow of the American Physical Society and the American Association for the Advancement of Science.

---

[1] http://en.wikipedia.org/wiki/Balance_puzzle

[2] Another approximation often used to approximate and create images, particular mountain- and other landscapes is based on self-affine fractals. This is used is animation and virtual reality images. R.F. Voss "Fractals in Nature: from characterization to simulation" in H.-O. Peitgen and D.Saupe: The Science of Fractal Images, Springer New York 1988.

[3] H. Nyquist "Certain topics in telegraph transmission theory" Trans. AIEE 47, 617-644 (1928)

[4] C.E. Shannon "Communication in the presence of noise" P. IRE 37, 10-21 (1949)

[5] See figure 1 in http://www.ams.org/samplings/math-history/hap7-pixel.pdf
[6] M.R. Pointer, G.G. Attridge " The Number of Discernable Colors" Color Res. Appl. 23, 52-54 (1998) & D. Nickerson, S.M. Newhall "Number of discernible object colors is a conundrum " J. Opt. Soc. Am. 33, 419 – 422 (1943)
[7] S.S. Stevens, H. Davis, Hearing, Its Psychology and Physiology Wiley, New York, 1938 pp 152-154
[8] The terms compressive sensing, compressed sensing or sampling are used interchangeable in the scientific literature.
[9] T. Strohmer, "Measure What Should be Measured: Progress and Challenges in Compressive Sensing" IEEE SIGNAL PROCESSING LETTERS, Vol. 19, No. 12, 887 – 893 December 2012
[10] A light source's color temperature is the temperature an ideal black body would have that is closest in hue to that of the light source in question. Diffuse sunlight has a color temperature between 5,700 and 6,500 degrees Kelvin, whereas a candle's color temperature is near 2,000 degrees Kelvin.
[11] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-pixel imaging via compressive sampling", March 2008, IEEE Signal Processing Magazine, 25(2), pp. 83 - 91.
[12] "We are giving ourselves cancer" New York Times Opinion pages by Rita F. Redberg and Rebecca Smith-Bendmann, January 30, 2014 http://www.nytimes.com/2014/01/31/opinion/we-are-giving-ourselves-cancer.html?_r=0
[13] S.S. Vasanawala et al "Improved pediatric MR imaging with compressed sensing" Radiology 256, 607-616 (2010)
[14] M. Lustig, D. Donoho, J. M. Pauly "Sparse MRI: The application of compressed sensing for rapid MR imaging" Magn. Reson. Med. 58, 1182-1195 (2007)
[15] T. Hugger et al "Fast undersampled functional magnetic resonance imaging using nonlinear regularized parallel image reconstruction" PloS one 6, e28822 (2011)
[16] Dave Eggert 'The Circle' , McSweeney's, October 2013
[17] http://www.wired.com/magazine/2010/02/ff_algorithm/
[18] http://cms.web.cern.ch/
[19] E. J. Candes, T. Tao " reflections on compressed sensing" , IEEE Information Theory Society Newletter p14-17, December 2008
[20] http://www.sdss.org/
[21] http://www.lsst.org/lsst/

[22] “Practical Compressed Sensing: modern data acquisition and signal processing” PhD thesis by Stephen R. Becker, California Institute of Technology, Pasadena, California, 2011

[23] J. Gantz, D. Reinsel, “The Digital Universe Decade – Are You Ready?” IDC White paper, May 2010

[24] Donald E Stokes “Pasteur’s Quadrant: Basic Science and technological Innovation” page 74 Brookings Institute Press (June 1997)

[25] Per Bak, *How Nature Works: The Science of Self-Organized Criticality*, New York: Copernicus. 1996

[26] A. Halevy, P. Norvig, and F. Pereira “The Unreasonable Effectiveness of Data” IEEE Intelligent Systems March/April 2009 pp 8-12