

Compression Method Reduces Big Data Bottleneck

Matthew Chin, UCLA



In creating an entirely new way to compress data, a team of researchers from the UCLA Henry Samueli School of Engineering and Applied Science has drawn inspiration from physics and the arts. The result is a new data compression method that outperforms existing techniques, such as JPEG for images, and that could eventually be adopted for medical, scientific and video streaming applications.

In data communication, scientific research and medicine, an increasing number of today's applications require the capture and analysis of massive amounts of data in real time.

But "big data," as it is known, can present big problems, particularly in specialized fields in which the events being studied occur at rates that are too fast to be sampled and converted into digital data in real time. For example, in order to detect rare cancer cells in blood, researchers must screen millions of cells in a high-speed flow stream.

To help improve the process, the UCLA group, led by Bahram Jalali, holder of the Northrop Grumman Opto-Electronic Chair in Electrical Engineering, and including postdoctoral researcher Mohammad Asghari, created an entirely new method of data compression. The technique reshapes the signal carrying the data in a fashion that resembles the graphic art technique known as anamorphism, which has been used since the 1500s to create optical illusions in art and, later, film.

The Jalali group discovered that it is possible to achieve data compression by stretching and warping the data in a specific fashion prescribed by a newly developed mathematical function. The technology, dubbed "anamorphic stretch transform," or AST, operates both in analog and digital domains. In analog applications, AST makes it possible to not only capture and digitize signals that are faster than the speed of the sensor and the digitizer, but also to minimize the

Compression Method Reduces Big Data Bottleneck

Published on Scientific Computing (<http://www.scientificcomputing.com>)

volume of data generated in the process.

AST can also compress digital records— for example, medical data so it can be transmitted over the Internet for a tele-consultation. The transformation causes the signal to be reshaped in such a way that "sharp" features— its most defining characteristics— are stretched more than data's "coarse" features.

The technique does not require prior knowledge of the data for the transformation to take place; it occurs naturally and in a streaming fashion.

"Our transformation causes feature-selective stretching of the data and allocation of more pixels to sharper features where they are needed the most," Asghari said. "For example, if we used the technique to take a picture of a sailboat on the ocean, our anamorphic stretch transform would cause the sailboat's features to be stretched much more than the ocean, to identify the boat while using a small file size."

AST can also be used for image compression, as a standalone algorithm or combined with existing digital compression techniques to enhance speed or quality or to improve the amount images can be compressed. Results have shown that AST can outperform standard JPEG image compression format, with dramatic improvement in terms of image quality and compression factor.

The new technique has its origin in another technology pioneered by the Jalali group, time stretch dispersive Fourier transform, which is a method for slowing down and amplifying faint but very fast signals so they can be detected and digitized in real time.

High-speed instruments created with this technology enabled the discovery of optical rogue waves in 2007 and the detection of cancer cells in blood with one-in-a-million sensitivity in 2012. But these instruments produce a fire hose of data that overwhelms even the most advanced computers. The need to deal with such data loads motivated the UCLA team to search for a new data compression technology.

Jalali said the discovery is rooted in— and inspired by— both physics and the arts.

"Reshaping the data by stretching and wrapping it in the prescribed manner compresses it without losing pertinent information," he said. "It emulates what happens to waves as they travel through physical media with specific properties. It also brings to mind aspects of surrealism and the optical effects of anamorphism."

Source URL (retrieved on 04/16/2014 - 9:24pm):

<http://www.scientificcomputing.com/news/2013/12/compression-method-reduces-big-data-bottleneck>