Neuronal Communication Process Opens New Directions in Image and Video Compression Systems

by Effrosyni Doutsi (FORTH-ICS), Marc Antonini (I3S/CNRS) and Panagiotis Tsakalides (University of Crete and FORTH/ICS)

The 3D ultra-high-resolution world that is captured by the visual system is sensed, processed and transferred through a dense network of tiny cells, called neurons. An understanding of neuronal communication has the potential to open new horizons for the development of ground-breaking image and video compression systems. A recently proposed neuro-inspired compression system promises to change the framework of the current state-of-the-art compression algorithms.

Over the last decade, the technological development of cameras and multimedia devices has increased dramatically to meet societal needs. The significant progress of these technologies has ushered in the big data era, accompanied by serious challenges, including the tremendous increase in volume and variety of measurements. Although most big data challenges are being addressed with paradigm shifts in machine learning (ML) technologies, where a limited set of observations and associated annotations are utilised for training models to automatically extract knowledge from raw data, little has been done about the disruptive upgrade of storage efficiency and compression capacity of existing algorithms.

The BRIEFING project [L1] aims to mimic the intelligence of the brain in terms of compression. The research is inspired by the great capacity of the visual system to process and encode visual information in an energy-efficient and very compact yet informative code, which is propagated to the visual cortex where the final decisions are made.

If one considers the visual system as a mechanism that processes the visual stimulus, it seems an intelligent and very efficient model to mimic. Indeed, the visual system consumes low power, it deals with high resolution dynamic signals (109 bits per second) and it transforms and encodes the visual stimulus in a dynamic way far beyond the current compression standards. During recent decades, there has been significant research into understanding how the visual system works, the structure and the role of each layer and individual cell that lies along the visual pathway, and how the huge volume of visual information is propagated and compacted through the nerve cells before reaching the visual cortex. Some very interesting models that approximate neural behaviour have been widely used for image processing applications, including compression. The biggest challenge however, is that the brain uses the neural code to learn, analyse and make decisions without reconstructing the input visual stimulus.

There are several proven benefits to applying neuroscience models to compression architectures. We developed a neuro-inspired compression mechanism by using the Leaky Integrate-an-Fire (LIF) model, which is considered to be the simplest model that approximates neuronal activity, in order to transform an image into a code of spikes [1]. The great advantage of the LIF model is that the code of spikes is generated as time evolves, in a dynamic manner. An intuitive explanation for the origin of this

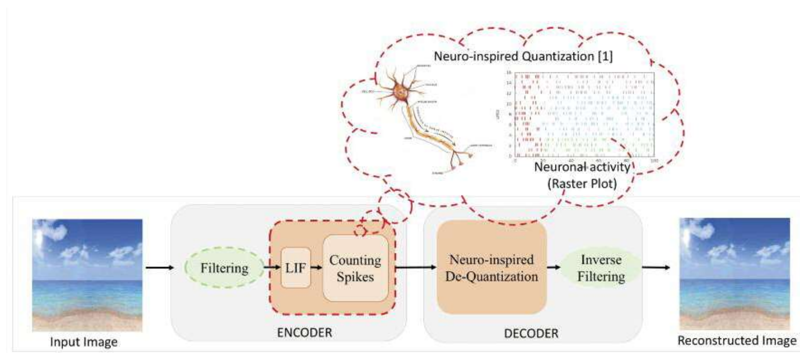

Figure 1: An illustration of the neuro-inspired compression mechanism that enables efficient reduction of the number of bits required to store an input image using the Leaky Integrate-and-Fire (LIF) model as an approximation of the neuronal spiking activity. According to this ground-breaking architecture, an input image can be transformed into a sequence of spikes which are utilised to store and/or transmit the signal. The interpretation of the spike sequence based on signal processing techniques leads to high reconstruction quality results.

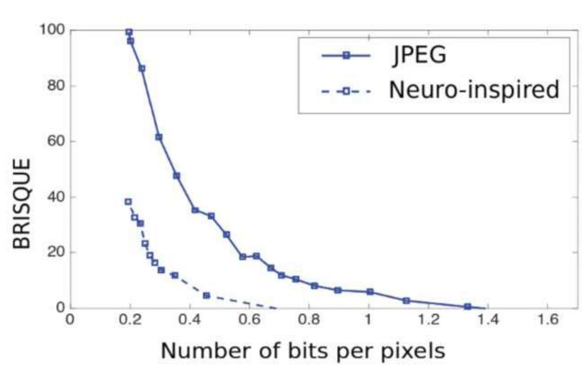

Figure 2: This graph shows that the BRISQUE algorithm that has been trained to detect the natural characteristics of visual scenes is able to detect far more of these characteristics within images that have been compressed using the neuro-inspired compression than the JPEG standards. The BRISQUE scores are typically between 0 and 100, where the lower the score the better the natural characteristics of the visual scene.

performance is that the longer the visual stimulus exists in front of a viewer, the better it is perceived. Similarly, the longer the LIF model is allowed to produce spikes, the more robust is the code. This behaviour is far beyond the state-of-the-art image and video compression architectures that process and encode the visual stimulus immediately and simultaneously without considering any time parameters. However, taking advantage of the time is very important, especially when considering a video stream that is a sequence of images each of which exists for a given time.

Another interesting aspect is that a neuro-inspired compression mechanism can preserve important features in order to characterise the content of the visual scene. These features are necessary for several image analysis tasks, such as object detection and/or classification. When the memory capacity or the bandwidth of the communication channel are limited it is very important to transmit the most meaningful information. In other words, one needs to find the best trade-off between the compression ratio and the image quality (rate-distortion).

We have proven that neuro-inspired compression is much more valuable than the state-of-the-art such as JPEG and/or JPEG2000, which both cause drastic changes in these features [2]. More specifically, we evaluated the aforementioned models using the BRISQUE algorithm [3], a convolutional neural network that has been pre-trained in order to recognise natural characteristics within the visual scene. As a first step, we compressed a group of images with the same compression ratio using both the neuro-inspired mechanism and the JPEG standard. Then, we fed the CNN with the compressed images and we observed that it was able to detect far more natural characteristics within images that had been compressed by the neuro-inspired mechanism than JPEG standard.

This project, funded by the French Government within the framework of the "Make Our Planet Green Again" call, is a collaboration between the Institute of Compute Science (ICS) at Foundation for Research and Technology – Hellas (FORTH) and the Laboratory of Informatics, Signals and Systems Sophia Antipolis (I3S) at French National Centre for Scientific Research (CNRS)

Link:
[L1] https://kwz.me/h57

References:
[1] E. Doutsi, et al.: "Neuro-inspired Quantization", 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, 2018, pp. 689-693.
[2] E. Doutsi, L. Fillatre and M. Antonini: "Efficiency of the bio-inspired Leaky Integrate-and-Fire neuron for signal coding", 2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2019, pp. 1-5.
[3] A. Mittal, A. K. Moorthy, and A. C. Bovik: "No-reference image quality assessment in the spatial domain", IEEE Transactions on Image Processing, vol. 21, no. 12, pp. 4695–4708, Dec 2012.

Please contact:
Effrosyni Doutsi, ICS-FORTH, Greece
edoutsi@ics.forth.gr

Fulfilling Brain-inspired Hyperdimensional Computing with In-memory Computing

by Abbas Rahimi, Manuel Le Gallo and Abu Sebastian (IBM Research Europe)

Hyperdimensional computing (HDC) takes inspiration from the size of the brain's circuits, to compute with points of a hyperdimensional space that thrives on randomness and mediocre components. We have developed a complete in-memory HDC system in which all operations are implemented on noisy memristive crossbar arrays while exhibiting extreme robustness and energy-efficiency for various classification tasks such as language recognition, news classification, and hand gesture recognition.

A cursory examination of the human brain shows: (i) the neural circuits are very large (there can be tens of thousands of fan-ins and fan-outs); (ii) activity is widely distributed within a circuit and among different circuits; (iii) individual neurons need not be highly reliable; and (iv) brains operate with very little energy. These characteristics are in total contrast to the way traditional computers are built and operate. Therefore, to approach such intelligent, robust, and energy-efficient biological computing systems, we need to rethink and focus on alternative models of computing, such as hyperdimensional computing (HDC) [1][2].

The difference between traditional computing and HDC is apparent in the elements that the machine computes with. In traditional computing, the elements are Booleans, numbers, and memory pointers. In HDC they are multicomponent vectors, or tuples, where neither an individual component nor a subset thereof has a specific meaning: a single component of a vector and the entire vector represent the same thing. Furthermore, the vectors are very wide: the number of components is in the thousands. These properties are based on the observation that key aspects of human memory, perception, and cognition can be explained by the mathematical properties of hyperdimensional spaces comprising high-dimensional binary vectors known as hypervectors [1]. Hypervectors are defined as d-dimensional (where $d \geq 1,000$) (pseudo)random vectors with independent and identically distributed (i.i.d.) components. When the dimensionality is in the thousands, a huge number of quasi-orthogonal hypervectors exist. This allows HDC to combine such hypervectors into new hypervectors using well-defined vector space operations, defined such that the resulting hypervector is unique, and with the same dimension.