BIOLOGICALLY REASONED MACHINE VISION: RLE VS. ENTROPY-CODING COMPRESSION OF DWT-TRANSFORMED IMAGES

Michal Podpora
Doctoral Degree Programme, Opole University of Technology
E-mail: ravyr@klub.chip.pl

Supervised by: Jan Sadecki
E-mail: jsad@po.opole.pl

ABSTRACT

In this article author compares two compression methods in a specific computer vision approach. Images (acquired locally by the mobile robot) are sent to a remote mode (i.e. a computer cluster) for non-time-critical AI processing tasks. The main problem of transmission – lack of efficient compression algorithm – is solved.

1. INTRODUCTION

Nowadays, robots can easily navigate in our natural environment. They can avoid obstacles, move towards a predefined/marked object. Nevertheless, their “brain” is still not flexible enough to “understand” our world in our way. Many years of evolution made our environment so ergonomic that any being (or any machine) which is not equipped with two legs and two arms can hardly exist. Despite packing robots with numerous sensors for better human-machine interaction, the improvement is minimal and the control algorithm’s complexity is formidable.

There is one entity in this world – a human being that is perfectly accommodated to our life conditions. It has no sophisticated sensors, though it operates almost everywhere on this planet. Replicating its algorithms is unfortunately still impossible but we’re constantly studying it to find out by what method it processes the acquired information to interact with its environment. [1]

2. SEARCHING FOR THE POWER

Some of the recent investigations show that simulating human brain in some tasks is possible, but extremely complex for any “non-biological” mind. A brain processes information in a naturally parallel manner. Computers are usually sequential machines, and even parallel supercomputers cannot compete with the most primitive animal brains. Although processors are getting increasingly powerful, the most problematic issue of a parallel algorithm is the communication time.

Simulating the “algorithm” of the neocortex would be much more efficient, if a computer cluster could be used, but using a computer cluster as an additional CPU-power for a mobile robot seems to be pointless due to extremely heavy traffic. It is true, unless the amount of data transmitted between nodes is reduced.

3. THE EYE AND THE BRAIN

Some situations require a real-time reaction (e.g. when an object is too close), while others don't. "Mental" tasks like "associating" or "memories" usually do not require real-time processing. Human brain is also capable of reacting "immediately" to some stimuli, while recalling some memories can take much time. The borderline between real-time and non-real-time reactions can easily be described, but it is far more difficult to develop an application implementing such a borderline.

Many scientists presume that splitting image acquisition from processing is possible only between the eye (camera) and the very first image processing procedures. However, this not only excludes partitioning processed data among supercomputer's nodes, but also hinders processing on a single multiprocessor machine.

Surprisingly, the "biological algorithm" in the human brain does not need to increase the "amount of data". The acquired information is processed by some visual brain centres – the meaning of the information changes, but it doesn't mean that it engages all nerve cells in the brain. The idea of the visual brain centres suggests that there might be a possibility of transmitting the data (to a remote node) in other part of the algorithm.

4. SPEEDING UP THE COMMUNICATION

Even though a part of the algorithm to be moved to a remote node could be found, the main problem of parallel algorithms is not solved – the communication time must be minimized. Again, the biological model gives the answer – the human eye doesn't acquire every part of the image with the same quality. This makes the term of "resolution" improper. "Yellow spot" – a part of the retina, containing the most packed cones – offers the best image resolution. Although this sounds like a promise of an efficient compression possibility, the existence of the yellow spot has seemed to be useless for scientists. Transforming the image using a lossy compression meant loosing the most valuable part of the acquired information – the details. Classical image processing is highly inefficient when processing low quality or noisy images. In fact, there was an idea (called "active vision") of analysing a fragment of image for better throughput of information, but usually a strictly selected area (like car plates) or the whole image was scanned line-by-line using this active area. None of these ideas is useful for a human-interacting mobile robot.

Improving the compression and minimising the communication time can be done only if in spite of classical image processing a "yellow spot –based" processing is implemented. Analysis of the acquired image becomes possible by dint of HTM (hierarchical temporal memory) networks. HTM networks – modelled on a human brain's neocortex algorithm – seem to fit this problem. The input is a set of only a few pixels, and the whole image is assembled in higher layers as an abstract term. [2]

Redefining yellow spot's coordinates or moving the camera is a separate issue. [3]

5. TRANSFORMATION

To prepare the image for the HTM network (i.e. to preserve the best possible quality in a specified point and its neighbourhood), the image was DWT-transformed (discrete waveform transformation) using a variable (gaussian-based) threshold level (1). [4]

Figure 1: Lena – a DWT-transformed image with a variable threshold level.

DWT transformation is used e.g. in the Jpeg2000 standard – it is known to produce easy-to-compress output data.

6. COMPRESSION

Two compression algorithms were implemented for comparison.

6.1. RLE

First algorithm was a modified version of RLE (run-length encoding). The modification results in RLE-encoding of bytes containing "0" only.

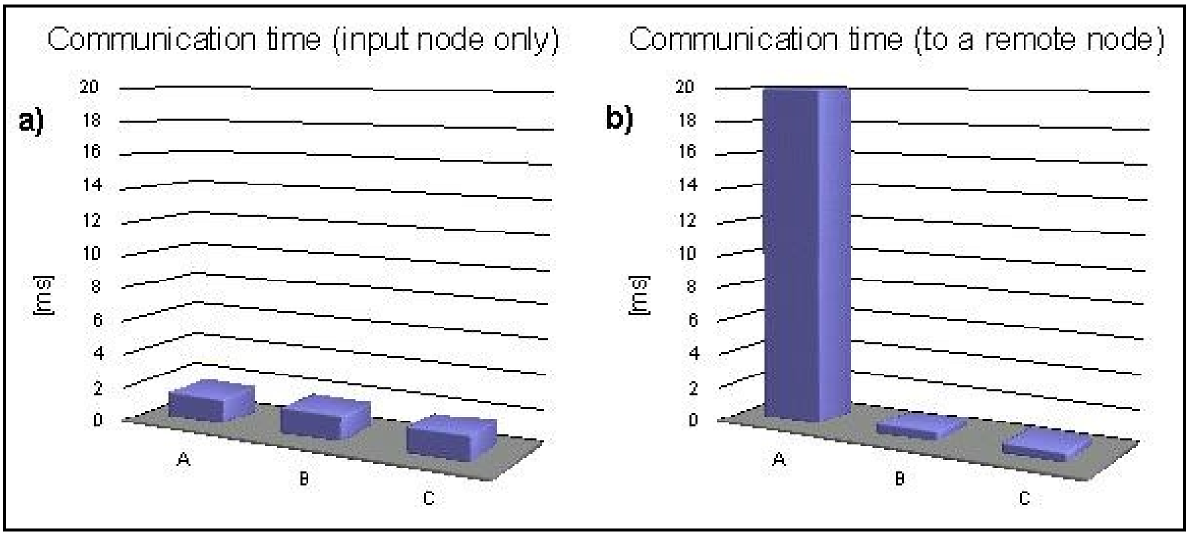

Figure 2: Average communication time (250 frames sent) : a) within input node, b) between two nodes ; A- uncompressed image, B- DWT-transformed and RLE-encoded image, C- DWT-compressed and RLE-encoded and entropy-coded image

Figure (2a) shows that within the input node both – compressing (2a:B) and not compressing (2a:A) the images – took the same communication time. Figure (2b) shows that the difference between sending an uncompressed frame (2b:A) and a compressed one (2b:B) is gigantic.

6.2. ENTROPY-CODING

Although all series of “0” values are substituted with a specific run-length code, the 0’s are still the most frequent values. As shown in Fig.(3), RLE with entropy-coding (3a:C , 3b:C) gives better results than RLE alone (3a:B , 3b:B).

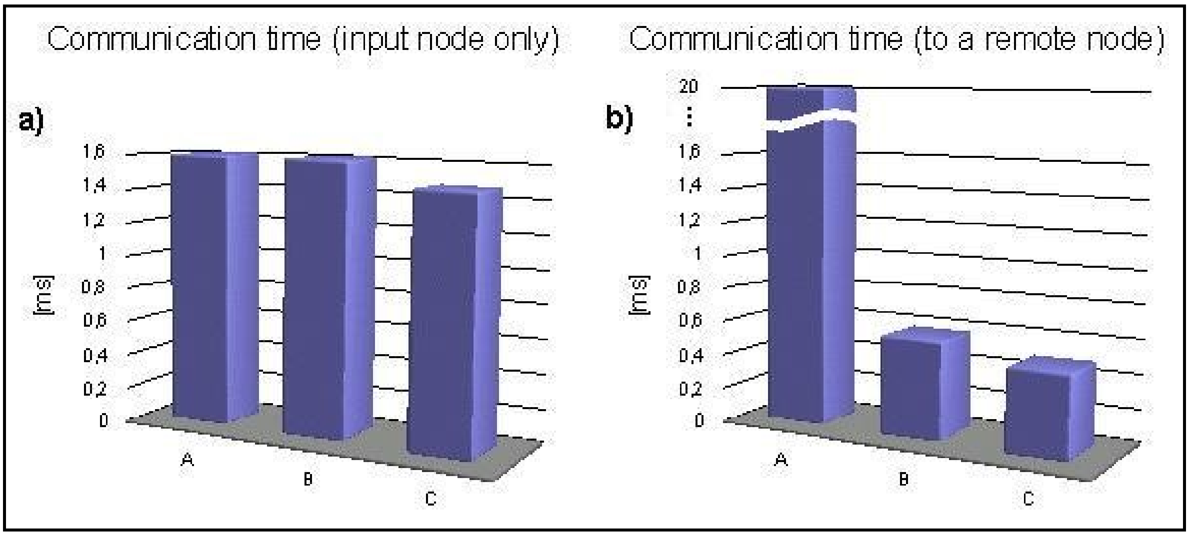

Figure 3: The same values as on Fig.(2) – scaled to show the differences; a) A=1.60 , B=1.61 , C=1.47 ; b) A=19.97 , B=0.59 , C=0.49

7. CONCLUSIONS

Figure (2b:A) shows that sending an uncompressed frame to a remote mode is a big mistake. Transmitting 20 frames per second means almost 40% of time spent on sending acquired images. Compressing frames with RLE combined with entropy-coding, notably decreases the communication time.

REFERENCES

[1] Hawkins, J., Blakeslee, S.: On Intelligence, Times Books, 2004, ISBN0805074562,

[2] Hawkins, J., Dileep, G.: Hierarchical Temporal Memory - Concepts, Theory, and Terminology, available: www.numenta.com/Numenta_HTM_Concepts.pdf , 2007, Numenta Inc.,

[3] Podpora, M.: Computer vision in parallel computing – reducing communication time, Proceedings of the 13th Conference Student EEICT 2007 vol.4, ISBN 978-80-214-3410-3, Brno 2007,

[4] Podpora, M.: Selected compression algorithms for mobile robots’ vision systems, InterTech – I International Interdisciplinary Technical Conference of Young Scientists, Poznań 2008, in press