I International Interdisciplinary Technical Conference of Young Scientists
17-18 April 2008 , Poznań, Poland

Selected compression algorithms for mobile robots' vision systems

Michał Podpora, Opole University of Technology

ravyr@klub.chip.pl

Abstract: **In this article author describes three approaches to the compression of mobile robot's vision system's specific DWT-transformated images.**

Keywords – machine vision, active vision, image coding, robot vision systems

I. INTRODUCTION

Nowadays, "truly intelligent" machines (capable of understanding the surrounding and the reality) exist only in science-fiction. Our everyday life is still too complex for them – even the simplest commands like "give me a cup of tea" or "close the door" demand understanding the surrounding. Of course, some commands can be pre-programmed, but real human-machine interaction can be achieved only if the machine could "learn" just like a human. [1]

Pattern analysis and object recognition are not enough if the robot is expected to learn new terms or skills. (Besides, its knowledge should not be based on a limited number of objects.) Human brain is not analyzing/processing the gathered information the way robots do.

HTM (hierarchical temporal memory) networks imitate the "algorithm" of the neocortex. Sensor nodes collect elementary data and higher levels of nodes build abstract "terms" describing the input data. [2] This idea gives new possibilities, but also causes some serious problems in practical implementation. The most problematic issue is lack of computation power. While results of tasks like "memories" or "associations" are not real-time-critical, a use of a computer cluster is worth considering. Although clusters seem to be too slow (due to the communication time), there are some ways of speeding up the communication between nodes of a cluster. [3]

The most effective part of the reduction deals with the image representation of the robot's vision system.

II. IMAGE REPRESENTATION

In HTM networks it is absolutely natural to use spatial fragments of acquired data and to send it to the input of the network in a time sequence [2]. Of course, the image can be scanned line-by-line (this also gives spatial fragments to the input), but better results might give feeding the network only with the important part of visual data. A human eye does not acquire the whole image at the same detail level – the most detailed fragment comes from *fovea centralis* (small pit on retina packed with cones). [4]

Image compression is especially useful if robot's embedded computer connects wirelessly to a remote processing unit.

Bringing the idea of yellow spot to the robots' realm, the image can be compressed using variable compression level. In currently implemented practical realization author uses DWT transform (discrete wavelet transform) with variable threshold value. The algorithm of the robot's vision system chooses coordinates of a specified point-of-interest in the field of view, and the embedded computer uses these coordinates to mark the point of the best image quality.

Threshold value changes (and the image quality becomes lower) with the distance from this specified point. Figure (1) shows the transformed image.

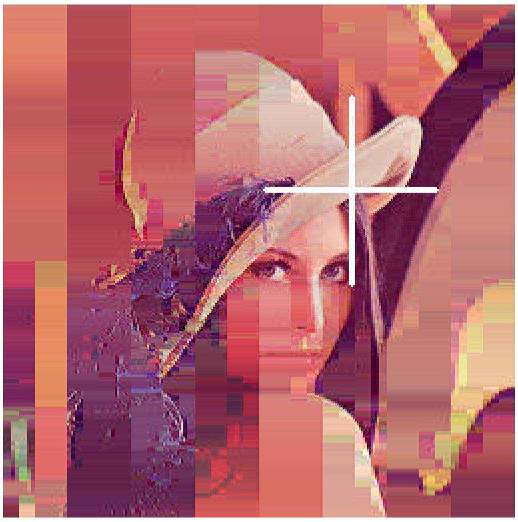

Figure 1. Lena – DWT & IDWT transformed image (DWT - using variable threshold); the "yellow spot" is marked as the white cross

III. –COMPRESSION

Although sending uncompressed video frames is possible, it decreases the robustness of the whole system by generating massive traffic. Using an additional computer (or e.g. computer cluster) as a processing unit requires connecting it wirelessly to the mobile robot's computer. This connection should also be reliable enough to minimize error rate and reduce the communication delay.

Additional processing power becomes a valuable advantage of a system, if only the traffic is not too heavy. Therefore, designing a robust and scalable system becomes a challenging issue, and for this reason sending uncompressed frames is unacceptable.

Memory or disk usage for the presented in Fig. 1 image (after the DWT transformation), is not lower, but compressing the transformed image gives better results than compressing the acquired image.

The basic difference between the acquired image and the transformed image is that the latter one contains many zeros in its coefficient matrixes. This fact allows implementing some elementary compression methods.

The first compression algorithm implemented by the author was a modified version of RLE (run-length encoding) compression. The RLE was simplified so that only repeating "0" values were substituted with encoded value, i.e. two bytes: "0" and "quantity" (Fig. 2). If more than 255 zeros in a row were found, the algorithm would repeat these two bytes with proper "quantity" value. This simple procedure compresses an 192kB image to e.g. 45kB (Fig. 1) – depending on the threshold function parameters.

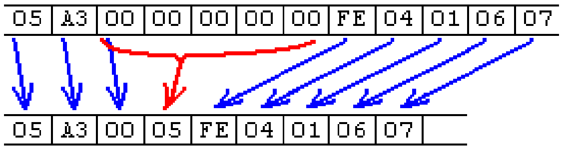

Figure 2. Modified RLE compression – compressing only zeros.

Even this unsophisticated algorithm gives interesting results, because after DWT transform 93% of bytes describing the image contain zeros.

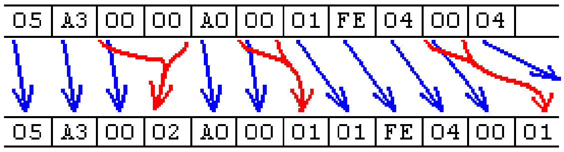

Figure 3. Modified RLE – not always giving the desired result (i.e. compression of data)

The author was trying to find a better compression algorithm, which could use this fact more effectively. Although short series or single values could make the file even bigger instead of compressing it (Fig. 3), the compressed files were about 77% smaller than the original. The hexadecimal preview of the compressed image show, that despite compressing the image, still 42% of the data are bytes with zero value.

The next step was to improve the algorithm basing on the entropy coding scheme:

- Value "0" represented as an 8-bit number, was changed to 1-bit representation,
- Bytes representing the quantity of zeros in a current series were not changed,
- Bytes containing other values were preceded with 1-bit "1" value (their representation changed from 8-bit to 9-bit).

Image shown in Fig. 1, compressed using this method, was 85% smaller than the original (i.e. 29kB instead of 192kB).

IV. Conclusions and Future Work

HTM networks have brought to life an old idea called “active vision”, where only a part of an image was analyzed to speed up vision systems’ computation time. In active vision, the region of interest was usually a rectangle area of acquired image, transmitted at the maximum detail level.

The DWT transformation of the acquired image with variable threshold level gives a very interesting and useful result: The final image not only contains data of variable quality (depending on the distance from the point of interest), but also preserves all advantages of a DWT-transformed image (e.g. fast changes of color, i.e. details of a scene/object).

The most important advantage of proposed solution is compressing the neighborhood of the point of interest with the best quality, which not only makes the processing of the image possible, but also enables analyzing objects placed “far” from the point of interest. Vision system can “notice” existence of an interesting object, which entered the field of view, whereas vision systems based on active vision completely ignore the rest of image (the image must be scanned periodically in order to detect new objects).

The DWT image transformation algorithm and described compression algorithms are currently being implemented in distributed vision system built on a mobile robot and a computer cluster.

V. References

[1] J. Hawkins, “Learn Like a Human”, IEEE Spectrum, Apr.2007

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

[3] M. Podpora, “Computer vision in parallel computing – reducing communication time”, Proceedings of the 13th Conference Student EEICT 2007 vol.4, Brno 2007

[4] M. Podpora, “Biologically reasoned point-of-interest image compression for mobile robots”, Conference on Human System Interaction, Kraków 2008, unpublished