ECMCS'99
EURASIP Conference
DSP for Multimedia Communications and Services
Kraków, 24-26 June

OBSERVER DEPENDENT WAVELET IMAGE COMPRESSOR

Piotr Augustyniak
Institute of Automatics, AGH Technical University[♣]
al. Mickiewicza 30, 30-059 Kraków, Poland e-mail: august@biocyb.ia.agh.edu.pl

ABSTRACT

The image compression by means of wavelet technique is considered in this paper. The proposed solution takes advantages of the current observer eye position rather than the image content. The noninvasive eye-pursuit system determines the zone of observer's interest and sends the information back to the image capturing system, which maximises the image resolution inside the zone of interest, while the remaining part of the image is sent at lower resolution. This method, however limited to the real-time observations, makes possible an up to 12-times compression of a single image without any loss of quality (e. g. colour definition) inside the observer specified area.

1. INTRODUCTION

This paper deals with widely discussed subject of image compression [1], [2], [3], [4]. However the information about the actual observer focus attention is considered rarely. This approach limits the area of application to the real-time observations and is not suitable for recorded images. However for widely used remote-inspection systems the considerable reduction of information without affecting the image quality may be interesting. Several demands for remote-inspection systems using narrow information channel come from medical diagnostics (e. g. swallowed stomach probes), traffic controllers, airport or supermarket security, video-telephone and other interactive systems or even space pilotless probes.

2. THE FOCUS ATTENTION CONCEPT

The general focus attention concept is physiologically merited by the irregular placement of photoreceptors on the surface of retinea [5]. The observer independent image systems are supposed to assure the equal geometric resolution on the whole area, while an "interactive" compressor takes advantage of the fact that the demanded density of information varies on the image area [6]. The compression can be achieved by colour or resolution (or both) data reduction outside of the zone of observer's interest. It results in loss of quality, but due to the human eye features, it is hardly remarkable. Additionally, the observer's eye position is captured every 40 ms (25 times per second), thus the position of the zone of interest is updated almost continuously. The following issues concerning the zone of interest should be considered additionally:

- Shape of the zone of interest. From the physiological point of view the circular zone of interest seems to be the most appropriate, however changing the resolution in the rectangular zone is more reasonable from technical point of view.
- Reducing the data outside the zone of interest. The geometric resolution and colour depth can be equally reduced in an non-focused area of image. In our proposal only the geometric features were affected, because of the high efficiency. Although the colours depth decreasing allows further reduction of data stream.
- The size of the zone of interest. The size of area where the image contents remains unaffected by the data loss determines theoretically the compression ratio. To fulfil the physiological and optical dependencies the zone diameter should correspond to the projection of the most sensitive region of retinea through the eye-lens. It is a linear function of the distance between the monitor and observer's eye, modified by an intra-observers variable coefficient [7], [8], [9]. It is unnecessary to update the zone position along with every eye position measurements.

3. DATA REDUCTION TECHNIQUE

Maintaining the main image features outside the zone of interest is the principal goal of intelligent data elimination. How could the observer focus on the other target if the image were too distorted. It is very important that the loss of quality were as low as possible. Thus some advanced techniques of downscaling rather than a simple subsampling should be used in order to maintain as far as possible the image features outside the zone of interest. For that reason we used orthogonal wavelet filters performing

[♣] This work was supported by AGH Technical University grant no. 10.10.120.39

the bandwidth (resolution) split without any loss of data [10]. Using wavelets makes sure that only the high frequency (resolution) information is lost from the compressed signal. Supplementary advantage of the orthogonal wavelet filters is their reconstruction capability useful while restoring the original image size. Like for other image (and short 1-D signals) processing tasks, the compact support wavelets (e. g. Daubechies) seem to be the most suitable because of the suppression of the distortion at the edges. On the other hand, the use of very short time support makes wavelets less precisely defined in frequency domain, what in case of images may lead to emphasise of some less important details. Considering all points mentioned above, we applied the 6-th order Daubechies orthonormal decomposition filter [11].

4. EYE TRACKING SYSTEM.

The main feature the eye tracking system must have is the noninvasivity. While eye tracking is performed on open eyes, during the active target pursuit, any technique not influencing the sight is applicable. Additionally the precision in tracking and speed is not crucial, so the simplest video or infrared 2-dimensional trackers are sufficient. Finally the infrared eye tracking system was chosen due to its low cost and high reliability (fig. 1) [12], [13] [14]. Its principle is the differential measurement of the pair of infrared beams reflected by iris (dark) and white (light). The zero value represents similar reflection conditions for both beams, what takes place when looking straight on. The positive value represents better reflection on the left beam, when the eye-globe is turned to the right, and similarly the negative value represents the turn to the left. The eye position is sampled on both eyes at 50 Hz and then averaged. 25 pairs of bytes per second (respectively horizontal and vertical eye co-ordinates) are sent back to the compressor and used as a centre of zone of observer's interest.

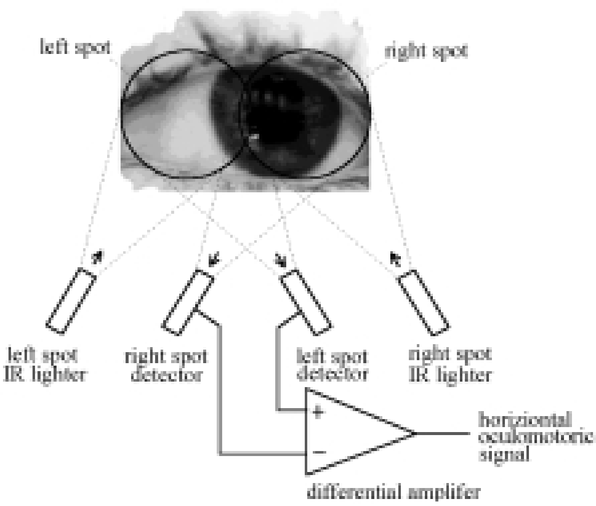


Fig. 1. The principle of an Infra-Red Eye Tracking System.

5. SYSTEM STRUCTURE AND DATA STREAMS

The basic system structure consists of:

- video capture unit (high resolution camera)
- adaptive wavelet encoder
- bi-directional transmission channel
- wavelet decoder and video-monitor
- eye positioning subsystem

The proposed system diagram is shown in the fig. 2.

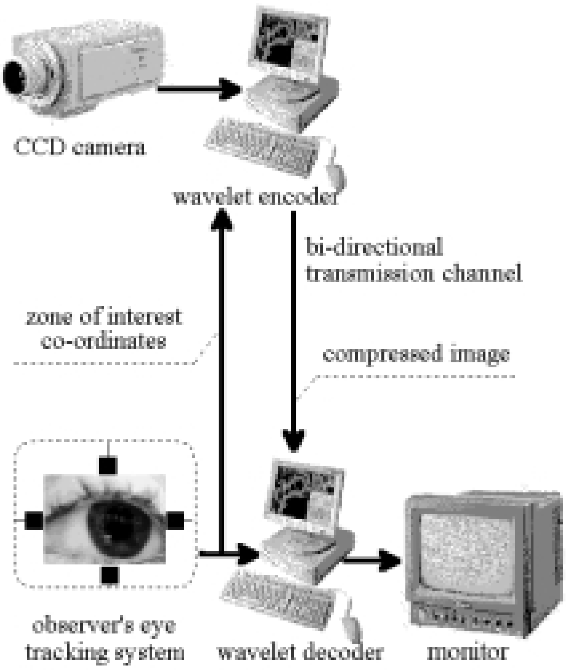


Fig. 2. Schematic diagram of an observer dependent wavelet image compressor system.

In the figure 2 data streams are represented by arrows. A question may arise, what an interest is in replacing a uni-directional information channel (typical for a video system), by a bi-directional one? The answer comes when comparing the amount of transmitted data. As an example we take a typical high-resolution image of about 256k pixels. Since a large part of the image is transmitted at a four times lower resolution, the reduction of data stream of about 16 times is assumed. In practice, depending on the size of the zone of interest, the compression factor of 12 is easy to achieve. On the other hand, the send back data stream is extremely weak, only 2 bytes per frame. For the example above (256 level grey scale image at a resolution of 512 x 512 pixels, 25 frames per second) the following data flow is expected:

- without compression: 6553600 bps,
- with an observer dependent wavelet image compressor: 546100 bps (forward) + 100 bps (backward)

6. CONTENTS OF A COMPRESSED IMAGE

The original image of assumed resolution 512 x 512 points is downscaled twice resulting two simplified

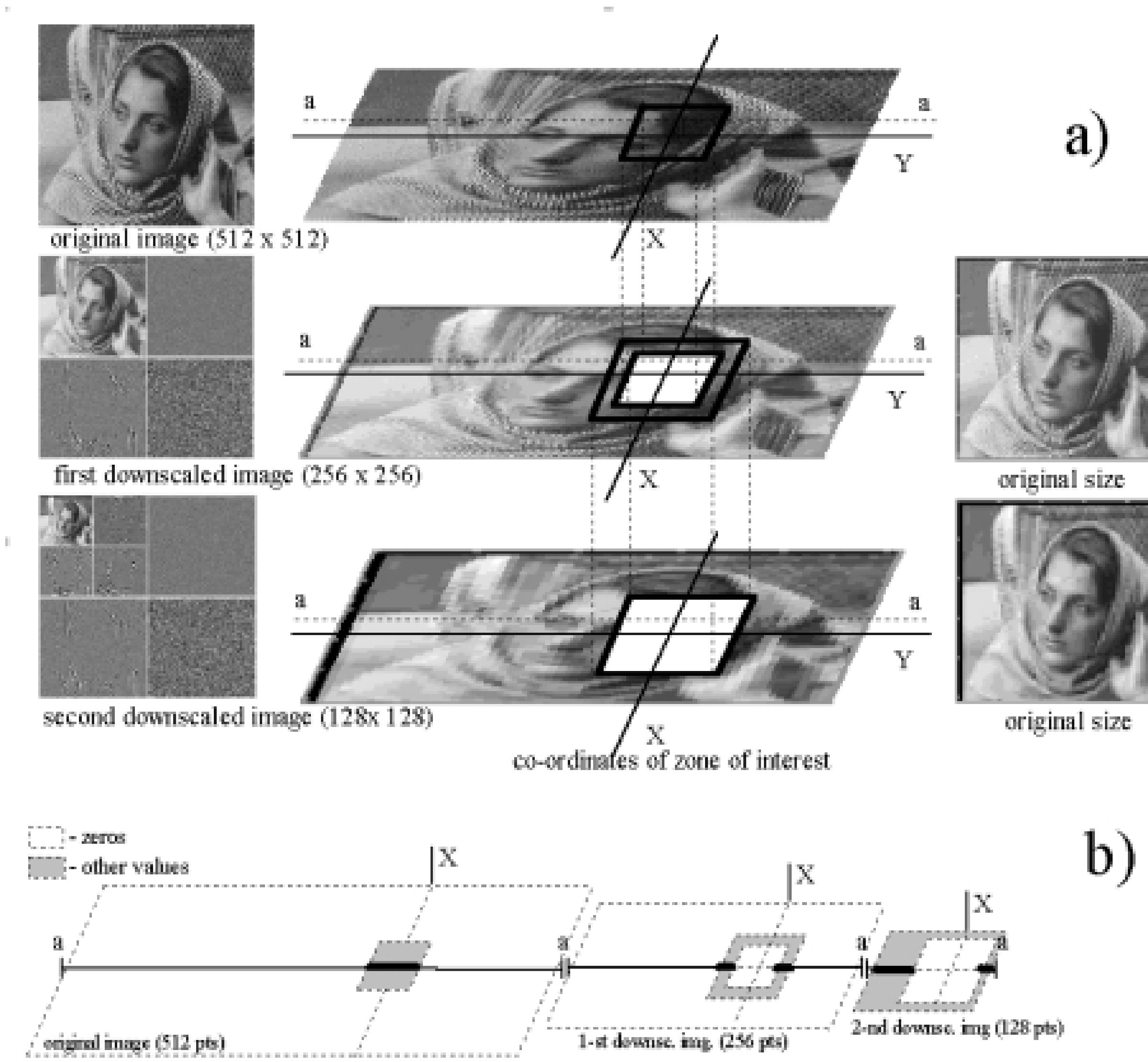


Fig. 3. a) Details on compression procedure and the contents of compressed image; b) sample wavelet coefficient vector for a line a-a (first 3 levels), a thick black line indicates where non-zero wavelet coefficients may appear.

versions of resolutions 256 x 256 and 128 x 128 respectively (fig. 3.). The downscaling procedure uses pyramid decomposition scheme and bi-orthogonal filters in order to provide best possible quality of simplified images. Then the 2-dimensional image is scanned vertically like in standard television. Depending on the position of zone of interest in the image:

- a single pixel value of the 2-nd simplified version (128 x 128),
- four pixel values of the 1-st simplified version (256 x 256),
- sixteen pixel values of the original image (512 x 512)

are sent through the transmission channel.

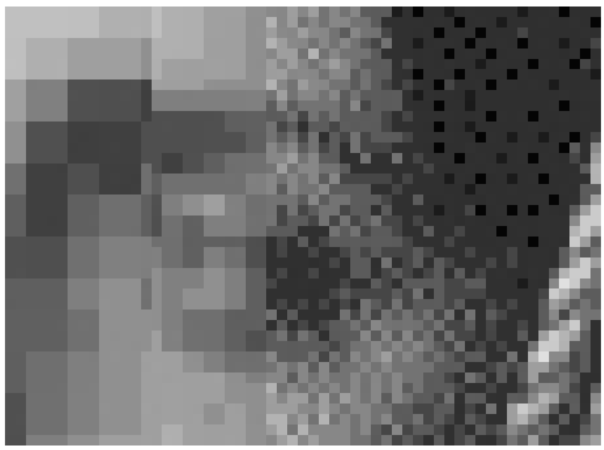
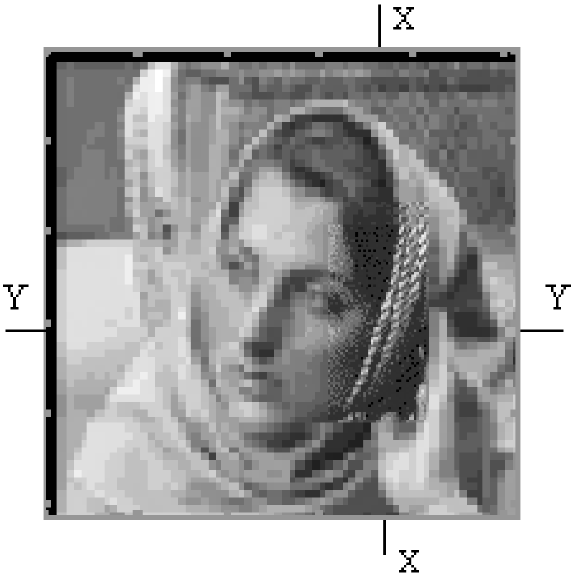


Fig. 4. Details on the variable pixel size at the edge of a zone of interest. From the left to the right: the lowest resolution outside of the zone of interest (corresponding to 128 x 128), the middle intermediate resolution (256 x 256) and the highest resolution (512 x 512) inside the zone of interest.



original image (512 x 512)
262144 pixels

compressed image
4851 pixels of original resolution
960 pixels of 1-st downscaled image
14655 pixels of 2-nd downscaled image
total: 20466 pixels

compression ratio 12,8

Fig. 5. Comparing of original and compressed picture with detailed calculations of pixels amount and compression ratio. X and Y are co-ordinates of centre of the zone of interest. For the compatibility with an average human visual perception the compressed image should be seen from a distance of about 7 cm.

7. DISCUSSION

The main goal of this work was a theoretical verification of possibilities offering by an observer dependent wavelet image compressor. Considering two level decomposition and only resolution data reduction the compression factor of about 12 was easy to obtain.

The proposed system configuration consist of desktop PCs as wavelet encoder and decoder, but there is a motivated hope to have a specialised DSP chip for this task in the future. Also the properties of the zone of interest should be proposed more precisely.

The compressor does not take any information about the picture itself nor the picture sequence. The future works should consider all these three aspects of image compression.

REFERENCES

[1] J. W. Woods S. D. O’Neil „Subband Coding of Images” IEEE Trans. on ASSP, Vol 32 No. 5, Oct 1986.

[2] S. Lewis G. Knowles „Image Compression Using 2-D Wavelet Transform”, IEEE Trans on Image Processing Vol 1. No. 2, pp 244-250, April 1992

[3] N. Ranganathan S. G. Romaniuk K. R. Namuduri „A Lossless Image Compression Algorithm using Variable Block Size Segmentation”, IEEE Trans on Image Processing Vol 4. No. 10, pp 1396-1406, October 1995

[4] Y. Zhuang „Adaptive Wavelet System for Image Representation” Proc. of ICSP’96 Beiging, Vol. 1 pp. 291-294, 1996

[5] R. W. Gubisch „Optical Performance of Human Eye”. J. Opt. Soc. AM 57, pp 407-415, 1967

[6] R. Smirnoff „Simulated Bipolar Cells in Fovea of Human Retinea” Biol. Cybern. 66, 123-135, 1991

[7] S. French A. W. Snyder D. G. Stavenga „Image Degradation by an Irregular Retinal Mosaic” Biol. Cybern 27 pp 229-233, 1977

[8] W. Snyder T. R. Bossomaier A Hughes „Optical Image Quality and the Cone Mosaic” Science NY 321 pp 499-501, 1986

[9] B. Fisher H. Weber „Express Saccades and Visual Attention” Behavioral and Brain Sciences Vol. 16, pp 553-610, 1995

[10] S. G. Mallat, „A Theory for Multiresolution Signal Decomposition: The Wavelet Representation”, IEEE Trans on PAMI, Vol. 11 No. 7, July 1989

[11] I Daubechies „Ten Lectures on Wavelets” SIAM CBMS-NSF Regional Conf. in Appl. Math., 1992

[12] J. L. Barbur W. D. Thomson P. M. Forsyth „A New System for the Simultaneous Measurements of Pupil Size and Two-Dimensional Eye Movements” Clin. Vision Sci., 2(2), pp 131-142, 1987

[13] J. Ober J. Hajda M. Jamicki J. Loska „System Pomiarowy Ruchu Oka Ober2 w Medycynie. Proc. Techn. Inform. w Med. UŚ, pp 123-132, 1997

[14] Y. Yamanobe S. Taira T. Morizono „Eye Movement analysis system Using Computerized Image recognition” Arches of otolayngology Vol. 116 pp 338-341, 1990