DCT-domain Scaling Algorithm with Application in Stereo Image Compression

Antoaneta A. Popova[1] and Alexander A. Krupev[2]

Abstract – **Our goal is to obtain a downsized version of an image in terms of its 8x8 block-DCT coefficients without passing through the spatial domain. DCT-domain scaling allows for fast, on-the-fly, resizing in DCT-based picture formats (like the widely used JPG) without the need for decompression, bilinear resizing and recompression. The ability of the human vision to compensate for resolution difference in the stereo image pair allows for smaller file sizes with preserved stereo perception.**

Keywords – **Stereo images, resize, compression, DCT-domain, JPEG.**

I. Introduction

In this work will be discussed some algorithms for resizing in the DCT-domain (Discrete Cosine Transformation) and their application for stereo image compression.

A known fact is that picture resolution directly influences the bandwidth requirements for its transmission, so a straightforward approach is to downsize it for purposes of image compression. Stereo pairs consist of two similar images intended for each eye in order to achieve depth perception. Human vision is able to compensate the lack of higher frequency information in one of the images if it is present in the other [9], [7]. This allows us to downsize one of the images before transmission and upsize it to its initial resolution again before stereoscopic visualization. Most visual materials in Internet and visual information systems, including stereo image pairs, are stored in a compressed form, DCT-based JPG format [10] being most common. So if resizing is performed directly in the DCT frequency domain instead of the spatial image domain, we don't have to perform inversed DCT (iDCT) transform, spatial resizing and DCT transform again. Without recompression, the process is much computationally simplified and allows for faster, on-the-fly resizing according to the user and/or bandwidth requirements.

When considering stereo image compression, it is important to measure how well the depth perception is preserved in the reconstructed images. The evaluation of the reconstructed compressed images is still an open problem. We offer objective measure (PSNR), but it doesn't take into account the peculiarities of the human visual system [8]. Subjective experiments are performed also.

[1]Antoaneta A. Popova is with the Faculty of Communications, Technical University - Sofia, Kliment Ohridski 8, 1000 Sofia, Bulgaria, e-mail: antoaneta.popova@tu-sofia.bg

[2]Alexander A. Krupev is with the Faculty of Communications, Technical University - Sofia, Kliment Ohridski 8, 1000 Sofia, Bulgaria, e-mail: asfalot@abv.bg

Works on DCT-resize have been published by Dugad and Ahuja [1], Mukherjee and Mitra [2], Merhav and Bhaskaran [3], Yim and Isnardi [4] and others. They all use implicit iDCT – DCT conversion. A scheme of Dugad and Ahuja's method for 1D signal is shown below (Fig. 1).

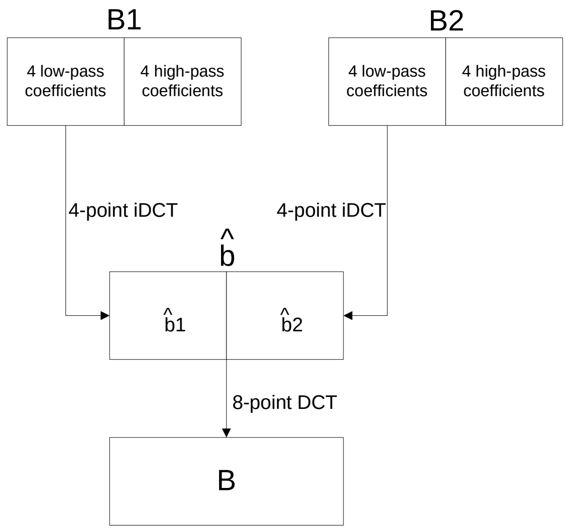

Fig. 1. 1D downsizing using Dugad and Ahuja's method

We see that iDCT is performed on 4 samples instead of 8 and low-pass filtering in the spatial domain is omitted thanks to the property of iDCT to smooth the edges i.e. we receive a low-passed and down-sampled version of the original 8-point block [5] with much lower computational complexity. Similarly the up-sampling is done by padding the DCT block with zeroes.

Our main goal is to suggest and implement an algorithm for DCT domain resize based on this principle and apply it on stereo pairs for compression purposes. The quality of the stereoscopic visualization before and after compression will be evaluated also.

II. DCT-domain resizing scheme

The developed scheme works in the following steps:

1. The stereo pair is split into two images left (L) and right (R).
2. The left image blocks $L_{i,j}$ are sent to the decoder without processing.

3. $l_{i,j}$ are obtained after iDCT transform in the decoder.

4. The right image DCT coefficient blocks $R_{i,j}$ are truncated to a smaller size, determined by the user. The new smaller nxn (n≤8, n≥1) blocks $\tilde{R}_{i,j}$ can have either standard square shape or triangle shape, proposed by the authors (Fig. 2) for energy compaction reasons. Zig-zag scanning is performed on the remaining coefficients. Block size and/or shape information (up to 4 bits) is sent to the decoder once for the whole image.

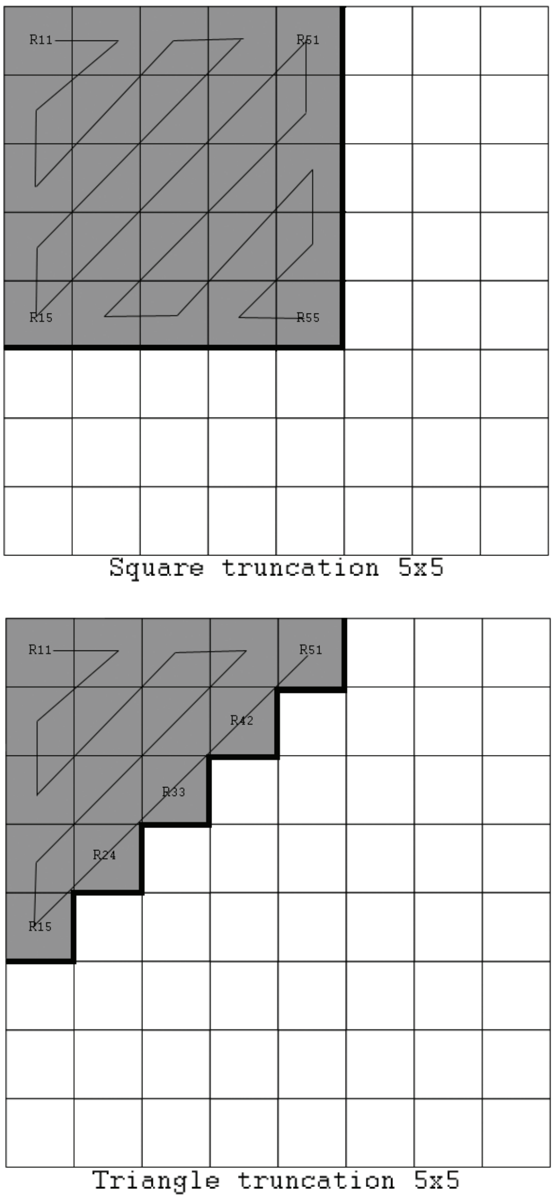

Fig. 2 DCT block square and triangle truncation with subsequent zig-zag scanning for n=5

5. The $\tilde{R}_{i,j}$ blocks are restored in the decoder after zig-zag scanning using the block size information. The truncated part is padded with zeroes – thus $\hat{R}_{i,j}$ blocks are obtained.

6. iDCT transform is applied on the restored to full size $\hat{R}_{i,j}$ DCT blocks and the reconstructed image blocks $\hat{r}_{i,j}$ to be displayed are obtained.

DCT and iDCT are implemented in the following form:

$$S(v,u) = \frac{C(v)}{2}\frac{C(u)}{2}\sum_{y=0}^{7}\sum_{x=0}^{7} s(y,x)\cos[(2x+1)u\pi/16]$$
$$\cos[(2y+1)v\pi/16]$$
$$s(x,y) = \sum_{v=0}^{7}\sum_{v=0}^{7}\frac{C(v)}{2}\frac{C(u)}{2} S(v,u)\cos[(2x+1)u\pi/16]$$
$$\cos[(2y+1)v\pi/16]$$
$$\begin{array}{llll} C(u)=1/\sqrt{2} & for\ \ u=0 & C(v)=1/\sqrt{2} & for\ \ u=0 \\ C(u)=1 & for\ \ u>0 & C(v)=1 & for\ \ u>0 \end{array} \quad (1)$$

$l_{i,j}$ and the reconstructed image blocks $\hat{r}_{i,j}$ can be displayed and/or joined and saved by the user. The scheme is shown on Fig. 3.

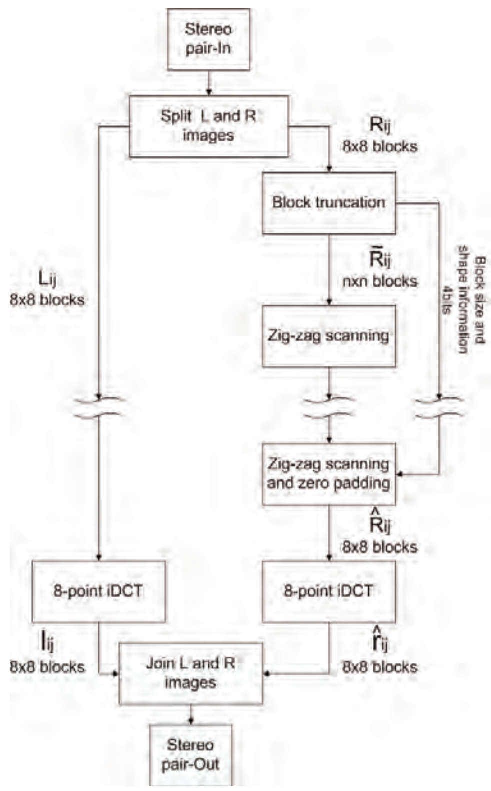

Fig. 3 Stereo pair DCT-domain resizing scheme

III. EXPERIMENTAL RESULTS

The scheme from Fig. 3 is implemented in Matlab 7.0.3 working environment. A PSNR evaluation module is added, comparing the processed image blocks $\hat{r}_{i,j}$ to the original $r_{i,j}$ right image according to Eq. 2:

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)[dB] \quad (2)$$

$$MAX_I = 255$$

Here MSE is:

$$MSE = \frac{1}{pq}\sum_{i=0}^{p-1}\sum_{j=0}^{q-1}[I(i,j) - \hat{I}(i,j)]^2 \quad (3)$$

p, q are the dimensions of the image, I and $\hat{I}$ - pixel intensity values before/after processing, ranging from 0 to 255.

A monochrome stereo pair of a juggler (resolution 784x600) has been processed (Fig. 4). 8 versions for square and triangle truncation have been obtained. Their PSNR values are shown in Table 1. The graphical representation is shown on Fig. 6 - PSNR as function of the DCT coefficients.

The resulting stereo pair for n=1 (1 DCT coefficient used - maximum compression, lowest quality) is shown on Fig. 5

Fig 4 Stereo pair of a juggler

Generally triangle truncation yields better results, up to 2.5dB in this case. But as we mentioned, PSNR doesn't take into account the subjective visual perception.

For the purposes of subjective evaluation the processed stereo pairs are visualized using the anaglyph method, based on color separation [6], and the shutter glasses method (we used 120 Hz ViewSonic VX2268wm LCD screen with NVidia 3D Vision stereoscopic glasses).

Fig. 5 Processed stereo pair from Fig. 4.

TABLE I

PSNR VALUES FOR THE JUGGLER IMAGE

DCT block size	**8x8**	**7x7**	**6x6**	**5x5**
PSNR (square trunc.)	∞	48.6737	42.2547	37.4448
PSNR (triangle trunc.)	44.4462	40.0701	36.6419	33.6821
DCT block size	**4x4**	**3x3**	**2x2**	**1x1**
PSNR (square trunc.)	33.3659	29.8392	26.5579	23.0081
PSNR (triangle trunc.)	30.9855	28.4413	25.9459	23.0081

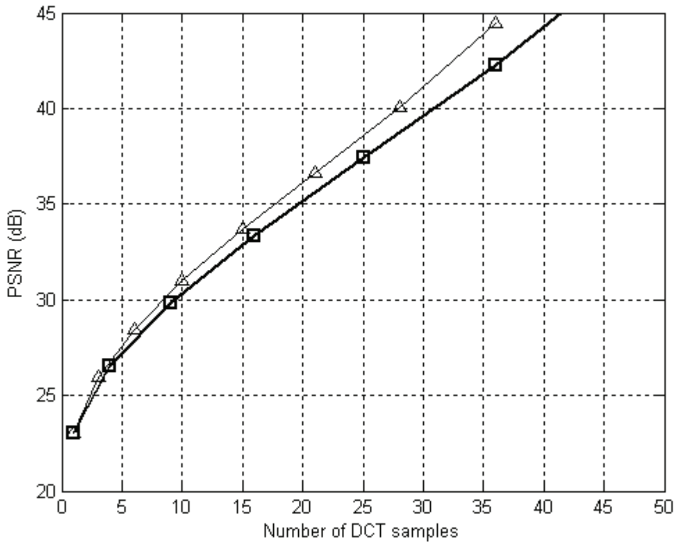

Fig. 6 The dependence between PSNR and the number of DCT samples. Thin line represents the proposed triangular truncation. Thick line – square truncation.

For n=1 stereo perception is lost and the image quality is very bad. For n=2 stereo perception is acceptable, image quality is still bad (noticeable pixelization). For n=3 and square truncation stereo perception is excellent, but image appears somewhat blurry. Triangle truncation gives noticeable artifacts and some pixelization, the stereo perception is good though. For n=4 stereo perception is excellent as well as image quality. Triangle truncation yields a bit blurrier image, but it uses only 10 coefficients as opposed to 16 for square truncation – the compression achieved is much better. For 5x5 DCT blocks (triangle or square) image quality is almost indistinguishable from the original stereo pair (and at n=8 images are identical and $PSNR \to \infty$).

Two of the visualized anaglyphs are shown below on Fig. 6

Fig. 6 Anaglyphs created from the stereo pairs from Fig. 4 (left-best quality) and Fig. 5 (right-best compression, worst quality)

IV. Conclusion

The DCT-scaling scheme can be applied in conjunction with most other schemes for stereo image compression like joint pyramid encoding, block based disparity map encoding, pioneering block-based predictive encoding and others. Optimal results are achieved for 4x4 and 5x5 truncation. Reducing block size further hampers image quality first and even stereo perception (for 2x2 block size and lower).

Acknowledgement

This work was supported by National Ministry of Education and Science of Bulgaria under contract BY-I-302/2007 “Audio-Video information and communication system for active surveillance cooperating with a Mobile Security Robots”.

References

[1] R.Dugad and N. Ahuja, “A fast scheme for image size change in the compressed domain,” IEEE Transactions on Circuits and systems for Video Technology, vol. 11, no. 4, pp. 461--474, 2001.

[2] J. Mukherjee and S. Mitra, “Image resizing in the compressed domain using subband DCT” IEEE Transactions on Circuits and systems for Video Technology, vol. 12, No. 7, July, pp.620-627, 2002

[3] N. Merhav and V. Bhaskaran, “Fast algorithms for DCT-domain image down-sampling and for inverse motion compensation,” IEEE Transactions on Circuits and systems for Video Technology, vol. 7, no. 6, pp. 468--476, 1997.

[4] C. Yim and M. Isnardi, “An Efficient Method for DCT-Domain Image Resizing with Mixed Field/Frame-Mode Macroblocks”, IEEE Transactions On Circuits And Systems For Video Technology, Vol. 9, No. 5, August 1999.

[5] K. N. Ngan, “Experiments on two-dimensional decimation in time and orthogonal transform domains”, Signal Processing, vol. 11, pp. 249-263, 1986.

[6] A. Krupev and A. Popova, “Flickering Reduction Algorithm in Anaglyph Stereoscopic Images”, ICEST’08 Proceedings of Papers, Vol. 1, pp.125-128, Nis, Serbia, 2008.

[7] M. Bax and A. Vitus, “Stereo Image Compression”, http://www.stanford.edu/~mbax/ee392c/report.html, Stanford, Dec. 1997.

[8] I. Dinstein, M.G. Kim, A. Henik, and J. Tzelgov, "Compression of stereo images using subsampling transform coding," Optical Engineering, vol. 30, no. 9, pp. 1359-1364, Sept. 1991.

[9] L. Kaufman, *Sight and Mind: An Introduction to Visual Perception”*, Oxford University Press, 1974.

[10] W. Pennebaker and J. Mitchell, “*JPEG still image data compression standard (3rd ed.)”*, Springer. p. 291, 1993.