37.4: MPEG based Image Enhancement for People with Low-Vision

Jeong-Hoon Kim+ and Eli Peli**

*The Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA

+Dept.of Electronic Communication Eng., Shinheung College, Euijungbu, Korea

Abstract

A new MPEG based image enhancement algorithm for people with low-vision is presented. Contrast enhancement is performed by modifying the Inter and Intra quantization matrices in the decoder. The algorithm has low computational complexity and does not affect the existing MPEG compressibility of the original image.

1. Introduction

Millions of Americans are visually impaired, and their number is expected to grow rapidly [1]. People with reduced visual acuity have difficulties in reading small print, driving, and watching TV. With the increasing importance of MPEG compression as a source of information and entertainment like digital TV, DVD, and digital Camcorder, the need for image enhancement to aid these people in accessing new digital multimedia is growing. There are various approaches to enhance MPEG compressed images. The first approach is to enhance the image before compression. One disadvantage of this approach is that the pre-enhancement may affect the compressibility of the original image and the enhancement will affect the receivers. A second approach is to enhance the image after decompression. The post-decompression approach can be adopted without affecting the compressibility of the original image. However, the post-decompression approach requires much additional signal processing circuitry and it can also increase the severity of compression artifacts by making them clearly visible [2]. Previous work on image enhancement for low-vision work was carried out in an uncompressed domain [3-5]. This paper describes an image enhancement concept tuned to the spatial visual frequencies critical for people with low-vision and applied directly in the compression domain. Last year at SID we presented an image enhancement algorithm in the JPEG domain using the standard Discrete Cosine Transform (DCT) quantization matrix [6]. We were only able to process static images and the enhancement was affected in all DCT frequencies without considering the visual properties and viewing distance typical of people with low-vision. Here we propose an MPEG based enhancement of moving video tuned to those parameters to provide better TV and PC monitor visibility for the various digital multimedia. The enhancement we implemented with MPEG-2 is compatible with all the standard MPEG video protocols.

2. Image Enhancement in MPEG-2 Domain

An MPEG system is composed of an encoder and a decoder. In the I (Intra) picture mode of the encoder, the image is first divided into non-overlapping 8×8 blocks. The two-dimensional DCT is then computed for each 8x8 block. After the DCT coefficients are obtained, they are quantized using the Intra quantization matrix. In the case of P (Predictive) and B (Bi-directional) picture modes, inter-frame moving blcks are processed using the Inter quantization matrix. Quantization of the DCT coefficients is a lossy process. Many small coefficients are quantized to zeros in this step. The zig-zag scan of the DCT matrix and entropy coding make use of this property to lower the bit rate required to encode the coefficients. In the MPEG decoder, the compressed image is decoded. It is then dequantized by pointwise multiplication using the same Intra and Inter quantization matrices as used during the encoding. Finally the data is transformed using the inverse-DCT [7]. The DCT entries coefficient are arranged in a block left to right and top to bottom, in the order of increasing spatial frequencies. The properties of the DCT coefficients provide a very natural way for us to define spatial frequency filters in the DCT domain.

2.1 Spatial frequency filtering in DCT domain

Effective image enhancement requires increasing the contrast in a specific range of frequencies. Increasing the contrast at spatial frequencies that are not visible at all is futile and increasing the contrast of already-visible frequencies can cause distortions. Therefore, we needed to determine the critical range of frequencies as it relates to the DCT coefficients. Figure 1 is an illustration of conventioanl DCT basis functions. Where K is a frequency order and the lined area inside shows the band of critical frequencies to be enhanced. The specific range to be used with our target low-vision population was determined as follows: We assume a population of people with low-vision who have the visual acuity from 20/100 to 20/200. To relate the observers Contrast Sensitivity Function (CSF) to the spatial frequencies or orders in the DCT domain, some conversions are needed. The spatial frequency variable f in cycles/degree of the CSF was converted to the normalized spatial frequency f_n in cycles/sample or cycles/pel as follows [8]:

$$f\,\text{(cycles/degree)} = f_n\,\text{(cycles/pel)} \bullet f_s\,\text{(pel/degree)} \quad (1)$$

Where f_n, the corresponding frequency in the DCT domain,

 ISSN/0003-0966X/03/3402-1156-$1.00+.00

is $$f_n = k/2N, \quad k=0,1,2,\ldots N-1 \tag{2}$$

The sampling (f_s) depends on the viewing distance. In our experiment, the mean favorite viewing distance of 24 low-vision subjects for watching a 27-inch TV Screen (720×480 pixels) was found to be about 36 inches (Fig. 2). The f_s for this distance is about 22 (pel/degree). By substituting k = 7, N = 8, and f_s = 22 into equations (1) and (2), the maximum visual spatial frequency is 9.6 cycles/degree for the 8×8 block. The conversion results for this observation distance are given in Table 1. Peli et al. [5] found that the critical frequencies ranged approximately from 3 to 6 cycles/degree in people with low-vision. The shaded area in Table 1, from k = 3 to 5, thus represents frequencies that have to be enhanced in the DCT domain.

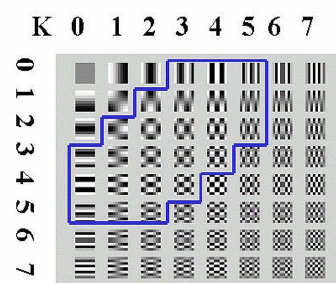

Figure 1. DCT basis functions. The functions inside the lined area represent the band of critical frequencies to be enhanced.

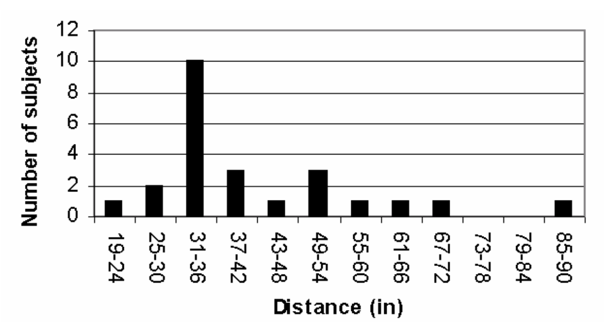

Figure 2. The distribution of viewing distances from a 27-inch TV selected by 24 people with low-vision.

Table 1. Visual frequencies corresponding to the DCT orders of basis functions for a viewing distance of 36 inches and a 27-inch TV monitor.

K (DCT order)	Visual frequency (Cycles/degree)
7	9.6
6	8.3
5	6.8
4	5.5
3	4.1
2	2.8
1	1.4
0	0

2.2 Image enhancement using Quantization Matrices

The basic idea of MPEG based enhancement method is to enhance the image by manipulating the Q tables in the MPEG sequence header. As shown in Figure 3, the image enhancement algorithm operates in the decompression stage by modifying the Intra and Inter Q matrices stored in the header with pre-designed Intra and Inter enhancement filter arrays to obtain new Q matrices. The new matrix includes the enhancement function. This frequency-domain filtering can be done by multiplying the entries in the quantization matrices decoded in the MPEG header, point by point, with new filter arrays. This results in the enhancement of the corresponding DCT frequency components by increasing the weighting of the critical high frequencies ranges during normal MPEG decoding. This technique only requires simple access to the Inter and Intra quantization matrices being decoded in the header, and the ability to modify them with the proposed enhancement filter array.

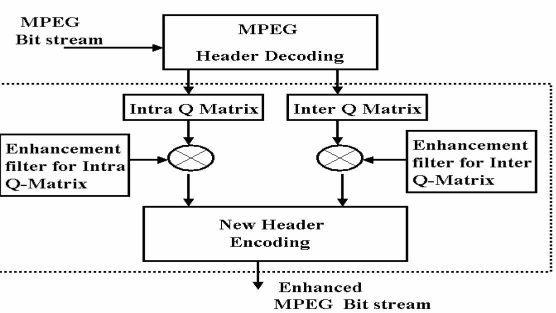

Figure 3. The flow of image enhancement in the MPEG domain. Note, Ä is a point-by-point multiplication.

Equation (3) is an example of a proposed enhancement filter (EF) array used here. The enhancement is applied only to the critical visual frequencies range identified above. The lambda (λ) in equation (3) is an image enhancement gain parameter that might be modified by the user.

$$EF = \begin{pmatrix} 1 & 1 & 1 & \mathrm{l} & \mathrm{l} & \mathrm{l} & 1 & 1 \\ 1 & 1 & \mathrm{l} & \mathrm{l} & a\mathrm{l} & \mathrm{l} & 1 & 1 \\ 1 & \mathrm{l} & \mathrm{l} & \mathrm{l} & a\mathrm{l} & \mathrm{l} & 1 & 1 \\ \mathrm{l} & \mathrm{l} & \mathrm{l} & a\mathrm{l} & \mathrm{l} & 1 & 1 & 1 \\ \mathrm{l} & \mathrm{l} & \mathrm{l} & \mathrm{l} & 1 & 1 & 1 & 1 \\ \mathrm{l} & \mathrm{l} & \mathrm{l} & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \tag{3}$$

Enhancement of interlaced TV format tends to increase line flickering artifacts. Previously, we used one-directional enhancement (vertical edge enhancement only) in JPEG domain to reduce the line flickering [6]. For MPEG based enhancement, we found that slight asymmetry in the filter structure, as shown in equation (3), was sufficient to control these artifacts. In MPEG, unlike JPEG, there are two different Q matrices — Intra and Inter matrices — with different values for still and moving blocks respectively. Thus, the above processing requires separate multiplications for Intra and Inter quantization matrices respectively. One can also use different filters for each matrix. The filtration is applied as

$$\bar{q}_{ij} = ef_{ij} \cdot q_{ij} \quad (4)$$

where q_{ij} are the elements of Intra or Inter quantization matrices, ef_{ij} are the elements of enhancement filter table, EF, and $\bar{q}_{ij}$ are the elements of the modified Intra or Inter quantization tables, $\bar{Q}$, being used for the actual enhancement in MPEG decoding. For initial testing, we used the four original MPEG-2 encoded sequences [9] shown in Table 2. The 'MPEG Header Decoding' and 'New Header Encoding' configurations, shown in Figure 3, were each processed by ReStream^{TM} software [10].

Table 2. MPEG-2 sequences tested.

Name	Lion.m2v	Susie.m2v	Flwr.m2v	Tennis.m2v
Profile/Level	Main	Main	Main	Main
Bit rate	6Mbps	8Mbps	8Mbps	8Mbps
Scan	Progressive	Interlace	Interlace	Interlace

3. Results

Initial experiments were performed on MPEG-2 test sequences (shown in Table 2) using the method described in section 2.2. We controlled the level of enhancement by manipulating the gain factor lambda (λ) in equation (3). We created various enhanced MPEG-2 video sequences using a range of λ from 2.0 to 5.0 with a constant factor a =1.5, for the Intra, Inter, and Intra plus Inter matrices. The enhancement effects were evaluated by low-vision people inspecting side-by-side comparisons of enhanced with original sequences, as shown in Figure 5. We performed a preliminary MPEG-2 enhancement evaluation with 13 people with low-vision (median VA 20/153, range 20/64 to 20/1257) 7 of whom had Central Field Loss (CFL). In a pilot test (N=9), we showed three enhanced segments (Lion, Susie, and Flwr (Flowers)) using the side-by-side presentation paradigm. The subjects significantly favored 3 levels of enhancement (λ=3.0, 4.0 for Lion, λ= 3.0 for Susie) (Z_8>2.04, p<0.04). In the next study (N=4 so far), we showed all the four segments in Table 2 in a more elaborate presentation method randomizing the position of the original and enhanced segments. The subjects could note obvious enhancement and improved perception of details for face and for moderate motion segments (Fig.6). Segments such as the Lion and Susie sequences represent the majority of scenes in most TV programs such as drama and news. The results revealed a trend showing that as enhancement level is increased, preference increases within a certain range (from λ = 2.0 to 4.0), after which it tends to decrease. We found that the preference may depend on image content. Most subjects did not favor the enhanced Flwr and Tennis sequences containing fast moving high frequency details. When subjects saw the enhanced Flwr and Tennis sequences, they noted motion and compression artifacts as well as enhancement at their short viewing distance. These image characteristics are available within the MPEG coding and can be used to modify enhancement.

4. Impact

A new MPEG based image enhancement algorithm may provide an inexpensive and flexible way to deliver better TV and PC monitor visibility for the digital video, individually tuned by observers, using the minimal modification of conventional MPEG decoders. The technology may have a wide market appeal for many elderly TV and PC viewers with minor visual impairment who would appreciate the individual and variable nature of the enhancement.

5. Acknowledgements

Supported in part by NIH grants EY05957, EY12890, and a postdoctoral fellowship program from Korea Science & Engineering Foundation (KOSEF) to Jeong-Hoon Kim.

6. References

[1] http://www.nei.nih.gov/nehep/faqs.htm.

[2] M.Y.Shen.et al.: Review of Post processing Techniques for compression Artifact removal, J Visual Communication and Image Representation 9(1), pp.2-14(1998).

[3] Peli E: Perceived Quality of Video Enhanced for the Visually Impaired. Technical Digest on Vision Science and its Applications, Technical Digest Series Vol. 1, pp.46-48 (1999).

[4] Ofer hadar, Adrian Stern: Enhancement of an image compression algorithm by pre- and post-filtering, Opt.Eng.40(2), pp.193-199(2001).

[5] Peli E. et.al.: Image enhancement for the visually impaired: simulations and experimental results. Invest Ophthalmol Vis Sci; 32, pp.2337-2350(1991).

[6] J.Tang, Peli E:Image Enhancement in JPEG domain for low-vision patients. SID 02 digest, pp.990-993(2002).

[7] B.G.Hasket.et al.: Digital Video: An introduction to MPEG-2, International Thomson Publishing(1997).

[8] Bowonkoon C and K.R.RAO: Human Visual Weighted Progressive Image Transmission, IEEE Trans on Communication, 38(7), pp.1040-1044(1990).

[9]ftp://ftp.tek.com/tv/test/streams/Element/MPEG-Video/525/.

[10] http://shh.dvdboard.de/.

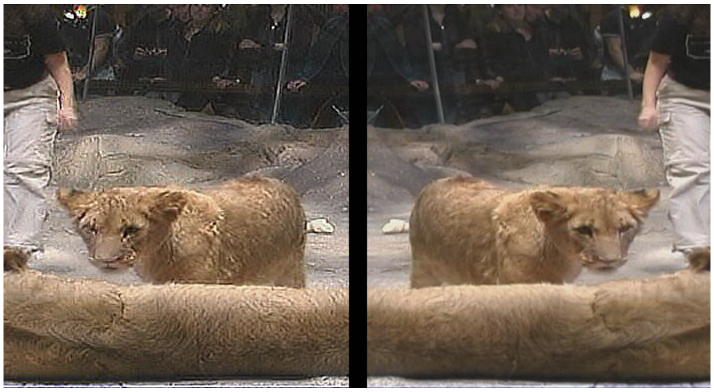

Figure 5.a. The example of Intra matrix only enhanced image (left) with λ = 4.0 and original decoded image (right).

Note, Intra only enhancement shows good enhancement in all parts of this image.

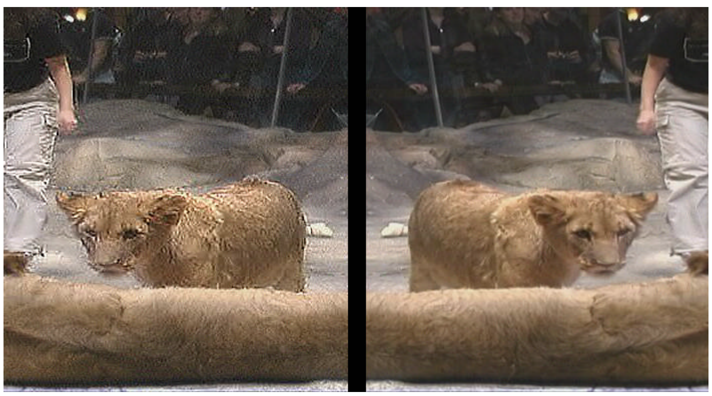

Figure 5.b. The example of Inter matrix only enhanced image (left) with λ = 4.0 and original decoded image (right).

Note, Inter only enhancement shows good enhancement in moving parts (moving tiger and person's trousers) from bi-directional or predictive motion of this image.

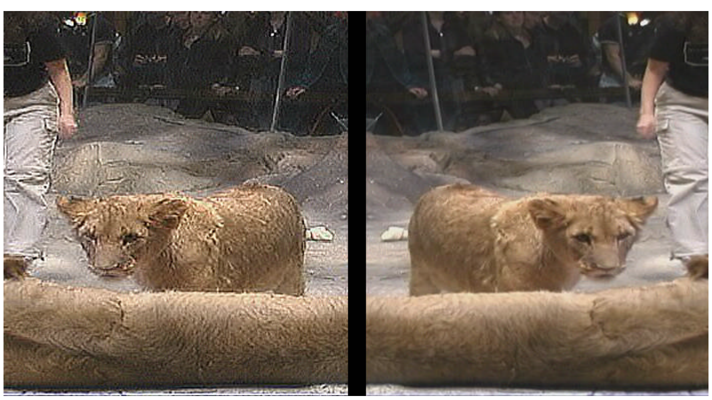

Figure 5.c. The example of Intra plus Inter enhanced image (left) with λ = 4.0 for both enhancements and original decoded image (right).

Note, the combination of Intra and Inter enhancement enhances all parts including bi-directional or predictive motion of this image. This enhancement shows good combined enhancing effects on both still and moving parts of this image. This is the enhancement we used for the experiment.

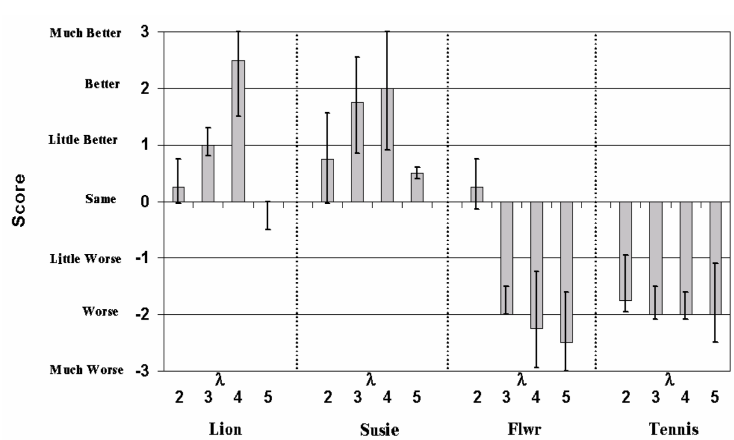

Figure 6. The median values of the 4 subjects' responses for the different sequences and levels of enhancement.

The error bar shows the range from first quartile (25%) to third quartile (75%). The subjects noted obvious enhancement for face and for moderate motion sequences such as the Lion and Susie. For the sequences with fast motion, the subjects rejected the enhancement. Information about the motion is available within the MPEG and can be used to adapt the enhancement.