MPEG-based image enhancement for the visually impaired

Jeonghoon Kim¹, ²
¹The Schepens Eye Research Institute, Harvard Medical School, 20 Staniford Street, Boston, MA 02114-2500, USA
²Shinheing College, Dept.of Electronic Communication #117, Howon dong, Uijeongbu-City, Gyeonggi-do, 480-701, Korea
E-mail: jhkim@vision.eri.harvard.edu

Avni Vora¹
¹The Schepens Eye Research Institute, Harvard Medical School, 20 Staniford Street, Boston, MA 02114-2500, USA
E-mail: vora@vision.eri.harvard.edu

Eli Peli¹
¹The Schepens Eye Research Institute, Harvard Medical School, 20 Staniford Street Boston, MA 02114-2500, USA
E-mail: eli@vision.eri.harvard.edu

Corresponding author:

Eli Peli
The Schepens Eye Research Institute, Harvard Medical School, 20 Staniford Street Boston, MA 02114-2500, USA
Telephone: 1 617 912-2597
Facsimile: 1 617 912-0111
E-mail: eli@vision.eri.harvard.edu

A previous conference proceeding that was published provided a brief overview of the concept and pilot experimental results. This paper provides a detailed explanation of the enhancement theories, methods, and provides complete experimental results with a 24 visually impaired subjects.
Abstract.

An MPEG-based image enhancement algorithm for people with low-vision is presented. Contrast enhancement is achieved by modifying the Inter and Intra quantization matrices in the MPEG decoder during the decompression stage. The algorithm has low computational complexity and does not affect the MPEG compressibility of the original image. We proposed and implemented an enhancement filter based on the visual characteristics of low-vision patients and report the results of image preference experiments with 24 visually impaired subjects. Subjects favored low to moderate enhancement for two video sequences but they favored only low enhancement and rejected higher enhancement for two other sequences that had fast motion.

Subject terms:

TV enhancement, MPEG decoding, spatial frequency, visual acuity, central field loss (CFL), motion compensation
1 Introduction

As the population ages, a growing number of people suffer from visual impairments. These impairments and their resulting disabilities greatly impact the quality of life of many elderly people. A Louis Harris survey found that vision impairment affects 17% of Americans 45 and older, and 26% of those 75 and older.\(^1\)

Traditionally, vision rehabilitation research was aimed at improving mobility and reading skills. More recently, efforts have been made to improve the ability of the visually impaired to perceive pictorial information with the use of high-contrast video-based devices. Compensating for loss of sensitivity by magnification improves the ability of the visually impaired to perceive visual information. But the magnification is not sufficient to restore some functions such as reading rate and face recognition.\(^2\)

The incorporation of computerized image enhancement to improve video images for the visually impaired was first proposed by Peli & Peli\(^3\) and Peli et al.\(^4\) Similar techniques were applied to the enhancement of text by Lawton\(^5\) and Fine & Peli.\(^6\) While image enhancement has been shown to modestly improve reading rate and may improve mobility as well, we see the main value of image enhancement in providing the visually impaired with better access to the growing flow of video images presented on monitors. Television is an important means of obtaining information and sharing in the culture. Since television is primarily a visual medium, visually impaired people have not had full access to its benefits. Yet, most do watch TV with their families, and prefer watching TV to other activities.\(^7,8\) Television use by the visually impaired has increased over the years and visually impaired people watch TV nearly as much as or more than normally-sighted people\(^7,9\). The Descriptive Video Service (DVS), which broadcasts programs with a
separate audio channel carrying a narrative description of the visual scene\textsuperscript{9}, is available for the visually impaired. Although it is helpful (particularly for the blind), it remains a limited service and substituting auditory for visual information takes away from the television-viewing experience. Therefore, image enhancement of video may be effective in helping the visually impaired enjoy television.

Methods of enhancement of images that were implemented in the past significantly and substantially increased face recognition for patients with central visual field loss and optical media opacities.\textsuperscript{10} Real time processing of live color video, using the adaptive enhancement algorithm, was made possible with the development of the DigiVision device.\textsuperscript{11} A pilot study using this device found increased recognition of details in the videos and almost 95\% preferred the individually tuned enhancement this device could offered.\textsuperscript{6} A different study, using a face recognition task (in static images), found that individually tuned enhancement improved recognition, but no better than uniformly applied adaptive enhancement.\textsuperscript{12} An additional study of motion video, found that subjects significantly preferred the enhanced images to the un-enhanced images, and that individual selection of parameters resulted in a greater affinity for enhancement over the lack of enhancement.\textsuperscript{13}

Most of the previous approaches were based on the filtering of analog (uncompressed) video even though digital signal processing was used. However, the use of digital video products applying MPEG standards (e.g. Digital TV, DVD player, and Digital Camcorder, etc.) is rapidly growing. The global sales of DVD players were estimated to reach about 41 million units in 2002 and 52 million units in 2003.\textsuperscript{14} The Federal Communications Commission (FCC) has adopted a plan that will give consumers access
to digital programming on television, by requiring off-air digital TV (DTV) tuners on nearly all new television sets.\textsuperscript{15} Thus, the need for video enhancement to aid visually impaired people should also shift to the new digital multimedia.

There are various approaches to enhancing digital compressed images. Images may be enhanced prior to compression and coding, after decoding, or within the coding/decoding process (which is the method we propose here). Pre-compression enhancement may affect the compressibility of the image and may require post decompression processing to maintain quality.\textsuperscript{16, 17} The post-decompression approach\textsuperscript{18-22} can be adopted without affecting the compressibility of the original image. But it tends to increase the severity of compression artifacts (e.g. blocking), making them clearly visible.\textsuperscript{18, 19, 23}

Tsai et al.\textsuperscript{20} have proposed an iterative algorithm for enhancing video sequences that are encoded at low bit rates. For MPEG sources the degradation of the picture quality originates mostly from the quantization function. Thus the iterative gradient-projection algorithm employed by these authors uses coding information such as quantization step size, macroblock types and forward motion vectors in its cost function. The algorithm shows promising results especially in enhancing decompressed low-bitrate video; however, its main disadvantage is its high computational complexity. Boroczky et al.\textsuperscript{21} proposed deriving a Usefulness Metric for Enhancement (UME) using compression information from the MPEG-2 bitstream to improve the performance of sharpness enhancement. The proposed algorithm was primarily developed for storage application at certain bit rates. Yang & Boroczky\textsuperscript{22} further improved the idea in\textsuperscript{21} by redefining the UME, as a quantitative value describing whether and how much a pixel can be enhanced without boosting coding artifacts. The new definition of UME correlates more precisely
with picture quality at various bit rates. Meanwhile, spatial features are taken into account to refine the UME. To improve the temporal consistency of enhancement, motion estimation and scene change detection are applied from I, P, and B pictures (see Appendix A). The new UME algorithm has relatively lower computational complexity than the iterative algorithm\textsuperscript{20} but it still has high computational complexity.

Various post-processing methods have recently been developed to remove the block artifacts that occur after image decompression.\textsuperscript{17, 23-28} While reduction of compression artifacts in still images has been studied extensively, little work has been done in improving the quality of compressed video.\textsuperscript{23} Martucci\textsuperscript{28} removed block artifacts in the compressed frequency domain, but the process required modification of the standard Discrete Cosine Transform (DCT) configuration in JPEG or MPEG. Konstantinides et al.\textsuperscript{29} implemented an image sharpening in the JPEG domain using the quantization matrix in the decoding stage, similar to our own method.\textsuperscript{30, 31} They took a degraded image (as produced by particular imaging systems, such as a color scanner or fax machine) and used a reference image to attempt recovery of the original image quality. The sharpening algorithm showed promising results but was applied only to static images where a high quality reference image was available.

We developed an MPEG-based video enhancement that operates in the decompression phase, which can reduce the block artifacts and is based on the visual characteristics of low-vision patients. To compensate for their reduced resolution and contrast sensitivity, low-vision patients tend to watch TV and PC monitors at very close distances. From such a short distance, they can easily note the effects of the enhancement as well as
severe block artifacts. Thus, the block artifacts are important considerations for low-
vision patients.

The MPEG enhancement is based on using the standard protocol for image compression,
which was also applied in the previously presented JPEG protocol for still images.\textsuperscript{30, 31}
The enhancement affected all DCT frequencies without considering the visual properties
or viewing distance typical of people with low-vision. The MPEG-based enhancement
approach presented here is constrained by two considerations. One constraint is the
compatibility of the processing with the properties of the low-vision patients’ visual
systems. The other is a requirement of compatibility with the current MPEG-2 standard
that handles digital TV. Because the enhancement is achieved simultaneously with
decompression and it only requires access to the quantization matrix, it has minimal
computational cost unlike conventional post-processing or pre-processing.

2 Image Enhancement in the MPEG-2 Domain

2.1 MPEG basics

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 of pixels. The
two-dimensional DCT is then computed for each block. The DCT coefficients are
quantized using the Intra quantization matrix. This compression takes advantage of
spatial correlations in the image. In the case of P (Predictive) and B (Bi-directionally
Predictive) picture modes, Inter-frame moving blocks are similarly processed using the
Inter quantization matrix. This step of the compression takes advantage of the temporal correlation between frames. 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 for storage or transmission. In the MPEG decoder, the compressed image is decoded. It is then dequantized by pointwise multiplication using the same Intra and Inter quantization matrices that were used during the encoding (these matrices are transmitted with the frame). Finally, the data is transformed using the inverse-DCT.32

2.2 Spatial frequency filtering in DCT domain

The DCT coefficients are arranged in the block, left to right and top to bottom, representing an order of increasing spatial frequencies. The properties of the DCT coefficients provide a very natural way for defining spatial frequency filters in the DCT domain.28 Effective image enhancement requires increasing the contrast in a specific range of frequencies. Increasing the contrast at spatial frequencies that are not at all visible is futile, while increasing the contrast of already-visible frequencies can cause distortions and is not particularly useful. Enhancement may be effective at frequencies that the viewer can detect only at high contrast levels. Figure 1 shows the contrast detection threshold as a function of spatial frequency adopted from Peli et al.10 The low-vision patients (Visual acuity 0.48 to 0.83 logMAR) could detect any targets in the spatial frequency range of 3 to 7 cycle/degree but required much higher contrast than normally
sighted observers. Most of them could not detect at all targets at frequencies higher than 8 cycle/degree. This range has to be mapped to the frequencies represented by the basis functions of the DCT. Figure 2 is an illustration of DCT basis functions for a 8x8 block commonly used in MPEG and JPEG coding. The top-left function represents the "DC" or zero spatial frequency. Along the top row the basis functions increase in horizontal spatial frequency content. Down the left column the functions increase in vertical spatial frequency, with an increase in both horizontal and vertical frequencies along the diagonals. The normalized spatial frequency, $f_n$ (cycles/pixel) of the corresponding basis functions in the DCT domain, is

$$f_n = K/2N, \quad K = 0,1,2,...N-1,$$  \hspace{1cm} (1)$$

where $K$ is the order of the coefficient and $N$ the size of the block. To relate the spatial frequencies or orders in the DCT domain to the low-vision patient’s contrast detection thresholds, the spatial frequency variable $f$ in cycles/degree of the contrast threshold function is converted to the normalized spatial frequency $f_n$ in cycles/pixel as follows.$^{33}$

$$f_n \text{ (cycles/pixel)} = f \text{ (cycles/degree)} / f_i \text{ (pixels/degree)},$$  \hspace{1cm} (2)$$

where the sampling frequency ($f_i$) depends on the viewing distance and the screen size. In a previous study in our lab, the median preferred viewing distance of low-vision subjects for watching a 27-inch TV Screen (720×480 pixels) was found to be 36 inches. The $f_i$ for this distance is approximately 22 (pixels/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 in the 8\( \times \)8 block. The conversion results for these conditions are given in Table 1. Therefore in the DCT domain, we enhanced the shaded frequency orders of \( K = 2 \) to \( K = 5 \) to achieve enhancement of the visual frequency range of approximately 3 to 7 cycles/degree. In Figure 2, the outlined area shows the basis functions we enhanced. We enhanced every component in the \( K = 2 \) to 5 bands except the two circled DCT basis functions. Removing the enhancement from these low frequency coefficients tended to decrease block artifacts.

### 2.3 Image enhancement using Quantization Matrices

Applying filtering in the DCT domain can be achieved in the MPEG decompression stage by manipulating the Q matrices available in the sequence header. In MPEG, unlike JPEG, there are two different Q matrices — Intra and Inter matrices — with different values for quantization of still and moving blocks, respectively. In our enhancement approach, both the Intra and Inter Q matrices may be multiplied, point by point, with pre-designed Intra and Inter enhancement filter arrays to obtain modified Q matrices. This technique requires only access to the Inter and Intra quantization matrices being decoded in the header, and the ability to modify them with the enhancement filter arrays. The filtration is applied as:

\[
\bar{q}_{ij} = e_{f_{ij}} \cdot q_{ij},
\]

where \( q_{ij} \) are the elements of Intra or Inter quantization matrices, \( e_{f_{ij}} \) are the elements of enhancement filters, EF, and \( \bar{q}_{ij} \) are the elements of the modified Intra or Inter
quantization tables, \( \bar{Q} \), which are then used in the MPEG decoding. The ‘MPEG Header Decoding’ and ‘New Header Encoding’ operations, shown in Fig. 3, were implemented here using the ReStream\textsuperscript{TM} software.\textsuperscript{34} In the preference study reported here, the same filters were applied to both Intra and Inter matrices.

\[
EF = \begin{pmatrix}
1 & 1 & 1 & \lambda & \lambda & \lambda & 1 & 1 \\
1 & 1 & \lambda & \lambda & a\lambda & \lambda & 1 & 1 \\
1 & \lambda & \lambda & \lambda & a\lambda & \lambda & 1 & 1 \\
\lambda & \lambda & \lambda & \lambda & \lambda & \lambda & 1 & 1 \\
\lambda & \lambda & \lambda & \lambda & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}
\]  

(4)

Equation 4 is the enhancement filter (EF) arrays applied in our study. The lambda (\( \lambda \)) parameter is an enhancement gain that might be modified by the user in real time from a remote control in the anticipated device. Enhancement of Interlaced TV format tends to increase line flickering artifacts. Previously, we used directional enhancement (vertical edge enhancement only) in an effort to reduce such flickering.\textsuperscript{30} Here, we used slight asymmetry with the placement of the arbitrary factor “\( a \)” in the filter structure emphasizing vertical edge enhancement, as shown in Eq. (4). This was sufficient to significantly reduce these artifacts. Figure 4 illustrates the coefficients of a standard “default” Intra Q matrix\textsuperscript{32} and it’s filtered Q matrix using the Eqs. (3) and (4) with (\( \lambda = 4 \), \( a = 1.5 \)).
3 The generation of enhanced test video sequences

Table 2 lists the four digital Standard Definition (SD)-grade MPEG-2 elementary test sequences we used in our study (Main Profile (MP) @ Main Level (ML)). SD grade is frequently used in DVD and other video sources. The Interlace scanned “Susie”, “Flowers (Flwer)” and “Table Tennis (Tennis)” sequences are available at 7 different bit-rates (40, 18, 12, 8, 6, 4, and 1.5 Mbps). We chose the 8 Mbps as it is the medium bit rate. For the “Lion” sequence, we used a lower bit-rate of 6 Mbps because the Lion sequence is a progressive sequence and thus has slightly higher quality. We carried out pilot tests with 9 visually impaired subjects to decide on a number of experimental parameters:

- Range and step size of enhancement levels.
- Mode of presentation - sequential vs side-by-side (split-screen).
- Sequence duration.
- Number of sequences and levels to be tested and repeated.

Based on the results of these pilot tests, we created enhanced MPEG-2 video sequences using a range of \( \lambda \) values (\( \lambda = 2, 3, 4, 5 \)) with a constant factor \( a = 1.5 \), for both the Intra and Inter matrices (see Appendix A).

For comparison of original and enhanced sequences, we used a side-by-side (split-screen) display as shown in Figure 5. To create the side-by-side display, we decoded original and enhanced MPEG sequences with the MPEG software decoder. After
decoding (including the required enhancement), we cut each sequence so that it was only half the original width, but maintained the center of the picture. We then merged the original and enhanced sequences using Matlab programs\textsuperscript{38} so that they played the scene simultaneously. We mirror-reversed the placement of original and enhanced video to enable side-by-side comparison of similar image areas. A total of 32 video sets (4 sequences x 4 gains x 2 sides) each 5-seconds in length was generated this way. Presentation of the 32 sequences took about 30 minutes. Experiments longer than 30 minutes might be too fatiguing for our mostly old subjects.

4 Experimental Evaluation

4.1 Subjects

Subjects were visually impaired, who did not use telescopes to view the screen. Twenty-four subjects (14 men and 10 women), ranging in age from 44.8 to 85.7 years (median age 71.0 years) participated in the study. The subjects’ log MAR visual acuity, measured using a BVAT (Model No. 22-4850, Mentor O&O Inc) ranged from 0.54 (20/70) to 2.10 (20/2500) (average 1.02 ± 0.35). All subjects had documented central field loss (CFL) in both eyes. Visual field was measured using a Bausch & Lomb Auto-Plot Tangent Screen (Cat. No. 71-54-41). The fields were measured monocularly, using a 6 mm white target at a distance of one meter.
4.2 Procedures

Subjects were asked to sit approximately 36 inches from the screen of a Dell Dimension 8250 computer with a 19-inch monitor (Dell P1130 Color Monitor) in a dimly lit room (3.6 foot-candles). Subjects were shown the 5-second video sequences which repeated until the subject responded. The subjects were asked to evaluate each side of the video sequence for “how clear the video was, how much details and information could be obtained from the video and the general quality of the picture.” Using these guidelines, they were asked to choose which side of the video (left or right) they preferred. If the subject could not see any difference in the two sides at the first test sequence (levels 4 or 5), they were allowed to move closer to whatever distance they chose. Subjects were forced to choose a side (i.e. they could not say the pictures looked the same). Once they chose a side, they were asked to rate the chosen side relative to the other side as “a little better,” “better,” or “much better” (responses were recorded as a score of 1, 2, or 3). If a subject selected the enhanced side sequence, a positive score was assigned. If the subject selected the original un-enhanced sequence, a negative score was assigned. The negative or positive score from the first question was combined with the second question to yield a score that ranged from −3 to 3 except 0 (zero). Two scores were derived from each level of enhancement for each sequence (one score from when enhancement was on the left side and one from the right). The two scores were averaged.

Measurements of their observation distances from the screen were taken at the end of the experiment. The average time for measuring visual acuity and visual field was about 45 minutes and the experiment took about 30 minutes.
5 Results

Figure 6 is another example of the mirror image from the experimental screen presented to the subjects. The left side is a frame from the enhanced sequences ($\lambda = 4$) and the right side is the corresponding frame from the original sequences. The subjects as a group preferred the 3 lower enhancement levels ($\lambda = 2$ to 4) for two of the sequences (“Susie” (Wilcoxon signed rank test, $Z_{23} > 2.26$, $p < 0.03$) and “Lion” ($Z_{23} > 2.55$, $p < 0.02$)) (Fig. 7). The small preference shown for the highest enhancement level ($\lambda = 5$) only approached significance ($p = 0.08$ and $p = 0.07$ for “Lion” and “Susie”, respectively). The two highest enhancement levels ($\lambda = 4$ and 5) for the two other sequences (“Flwer” and “Tennis”) were rejected ($Z_{23} > 2.48$, $p < 0.03$ and $Z_{23} > 2.36$, $p < 0.02$, respectively). The lower enhancement levels were not significantly different from the original although there was a slight preference for the low level of enhancement ($\lambda = 2$) and that effect was statistically significant for the “Flwer” sequence ($Z_{23} = 2.14$, $p = 0.032$).

During the experiments we noted that a few subjects seemed to have a clear preference for one side of the screen irrespective of the presentation of enhanced or original sequence. In each condition there were two presentations, one with the enhanced sequence on the right and one with the enhanced sequence on the left. We, therefore, tested to see for each patient if the selection was the same for the two presentations or different. For 11 of the 24 subjects, the preference was dependent on the side of the display (Paired t-test, $P < 0.05$) indicating a bias to one side.

Figure 8 shows the results from the thirteen subjects who showed consistent (unbiased) preference regardless of the position of enhanced sequences (i.e. on left or right side of
the screen). The results of these thirteen subjects are similar to those of the whole group. There is slightly higher preferences for the enhancement of the “Lion”, from $\lambda = 2$ to 4, ($Z_{12} > 2.05$, $p < 0.05$) and of the “Susie”, from $\lambda = 3$ to 5, ($Z_{12} > 2.15$, $p < 0.04$). For the other two sequences, subjects significantly rejected the two highest levels ($\lambda = 4$ to 5) for “Flwer” ($Z_{23} > 2.48$, $p < 0.03$) and the highest level ($\lambda = 5$) for “Tennis” ($Z_{23} > 2.36$, $p < 0.02$).

The median preferred viewing distance was found to be 20 inches (Min: 8 inches, Max: 32 inches) from the 19-inch PC monitor. Thus, the $f_s$ for this distance was approximately 16 (pixel/degree), and the maximum visual spatial frequency was 7 cycles/degree for the 8×8 block. Thus the frequency bands of $K = 2$ to $K = 5$ we enhanced corresponded in our case to approximately 2 to 5 cycles/degree.

6 Conclusion

We implemented and tested a new MPEG-based TV image enhancement for people with low-vision. The enhancement is applied during the MPEG decompression phase and requires only access to the quantization matrices. As such the computation load is minimal, it can be easily applied in real time and may be controlled and changed by the viewer using a remote control. Twenty-four visually impaired subjects favored a low to moderate level of enhancement for the “Lion” and “Susie” sequences which are likely to represent the majority of scenes in most TV programs such as drama or news. Most subjects favored only low-level enhancement for the “Flwer” and “Tennis” sequences, and clearly rejected the higher levels of enhancement for these two sequences. It is
possible that the enhancement for these sequences were rejected because these sequences contain more motion and the enhancement of fast motion sequences resulted in visible motion artifacts or led to too strong enhancement artifacts due to the combined enhancing effects of the Intra and Inter enhancement (see Appendix A). In view of these results, we experimented in reducing the enhancement level of Inter enhancement relative to the Intra enhancement (e.g. Intra enhancement level = 4, Inter enhancement level = 2 or 3). This resulted in reduced motion artifact. The information about motion is available within the MPEG video and could be used to adjust the enhancement levels adaptively for motion video segment or just for motion blocks. Others\textsuperscript{21, 22} have previously used motion estimation and scene change detection to ensure temporal consistency and to control the gain of enhancement of MPEG video. We plan to include such adaptive gain control based on motion in future studies.

MPEG-based image enhancement algorithm may provide an inexpensive and flexible way to deliver better visible digital video to elderly and visually impaired audiences, individually tuned by the user, applying only minimal modification to conventional MPEG decoders. This technology may have a wide market appeal for many elderly TV and PC viewers with moderate visual impairment who would appreciate the individual and controlled nature of the enhancement.

\textit{Acknowledgements}

This work was supported in part by NIH grants EY05957 and EY12890 to EP, and a postdoctoral fellowship program from Korea Science \& Engineering Foundation (KOSEF) to JK. We thank Robert B Goldstein for experimental programming help and Russell L Woods for advice regarding data analysis.
Appendix A

The effects of enhancement during decompression from Intra and Inter matrix

The MPEG coding and decoding apply different quantization matrices to different frames (I (Intra), P (Predictive), and B (Bi-directionally predictive) pictures). I Pictures use Intra Q matrix for all blocks. P and B pictures apply Intra and Inter Q matrices to still and moving blocks, respectively.\(^\text{32}\) While the Intra matrix is used to quantize actual image blocks, the Inter matrix is used to quantize the image difference of moving blocks. The enhancement can be applied to either matrix alone or to both. When both matrices are modified the modification can be identical or different.

### A.1 Intra-only enhancement

Intra-only enhancement is enhancing Intra macro blocks with the Intra Q matrix. Only Intra macro blocks are used in I pictures and few Intra macro blocks are used in static segments of P and B pictures. Thus, the Intra-only enhancement is enhancing still images or still blocks of images with motion.

\[
\begin{align*}
I & \rightarrow I', \\
P & \rightarrow I'_x + \Delta p, \\
B & \rightarrow \frac{I'_x + (I'_x + \Delta p)}{2} + \Delta b.
\end{align*}
\]

Equation (5) shows the effects of enhancing only the Intra matrix. If the current picture is an I picture, and the future picture is a P picture, the decompressed I picture will be enhanced only from Intra macro blocks with the enhancement filtering, applied to the
Intra Q matrix. The P picture will have some enhancing effects from the macro blocks forwarded from the previous I picture. However, the motion difference (\( \Delta p \)) blocks will be unmodified without Inter matrix enhancement. \( I'_i \) is an enhanced I picture from the Intra macro blocks and \( I'_p \) is an P picture partially enhanced with static forwarded macro blocks from the current I picture. The B pictures may have enhancing effects from forward, backward, and bi-directional macro blocks. Usually, the bi-directional averaged macro blocks are widely used and will be enhanced except for motion differences as shown in Eq. (5)

A.2 Inter-only enhancement

Inter-only enhancement is enhancing motion difference blocks by filtering the Inter Q matrix. Inter Q matrix is used for the forward, backward, and bi-directional macro blocks in P and B pictures. Thus, as shown in Eq. (6), the Inter-only enhancement can enhance the moving areas of P and B pictures. The \( \Delta p' \) and \( \Delta b' \) are enhanced motion difference components in both the P and B pictures.

\[
\begin{align*}
I & \rightarrow I'_i, \\
P & \rightarrow I'_p + \Delta p', \\
B & \rightarrow \frac{I'_i + (I'_p + \Delta p')}{2} + \Delta b'.
\end{align*}
\]  

(6)

A.3 Combined Intra and Inter enhancement

The enhancement filtering of combined Intra and Inter Q matrices will make combined enhancing effects of Intra and Inter Q matrices as shown in Eq. (7). This enhancement
will enhance all the macro blocks so both still and moving areas are enhanced together. While this is good for enhancement of the all areas, this combined enhancement may create too strong enhancement levels for the moving areas. This is because moving areas will be enhanced twice, once in the I picture and then again as a motion block, resulting in a double application of the enhancement to these blocks.

\[
\begin{align*}
I & \rightarrow I', \\
P & \rightarrow I' + \Delta p', \\
B & \rightarrow \frac{(I' + (I' + \Delta p'))}{2} + \Delta b'.
\end{align*}
\] (7)

Figure 9 illustrates the enhancement effect for a single video frame when applied to Intra and Inter matrices (as used in this study), and to the Intra, and Inter enhancement alone. These frames were captured from MPEG decoded/enhanced videos. Figure 9(a) is an original B (Bi-directional) picture decoded without enhancement. Figure 9(b) is the same picture enhanced with Intra and Inter enhancement. Figure 9(c) is the picture enhanced with Intra enhancement. Figure 9(d) shows the enhancing effect in moving area only resulting from modifying the Inter matrix.

Figure 10 illustrates the effects of each enhancement by presenting the differences between the enhanced frame and the original frame. It is evident that the Intra matrix enhancement (Figure 10(c)) enhances the whole image while the Inter matrix enhancement (Figure 10(d)) results only in enhancement of moving portions of the scene. The combined Intra and Inter enhancement in Figure 10(b) thus has a stronger enhancement effect in moving portions of the scene. We applied the combined
enhancement here under the assumption that the motion blur that results in reduced sensitivity to moving patterns\textsuperscript{39} would require stronger enhancement for such areas.


**Figure and Table captions**

Figure 1. The contrast detection threshold as a function of spatial frequency for low-vision patients and normally-sighted observers (modified from Peli et al., 1991). The bolded black line shows the average contrast detection threshold of people with normal vision. The fine lines curves show the contrast thresholds of 8 low-vision patients. The range of 3 to 7 cycle/deg is the range of frequencies that could be effectively enhanced for low-vision patients. Note patients could not see the frequencies higher than 8 cycle/degree.

Figure 2. The DCT basis functions for an 8 x 8 block. The basis functions inside the lined area represent the critical frequencies to be enhanced. We excluded the two functions circled because their enhancement increased block artifacts. In areas outside these bands the quantization matrix was not modified.

Figure 3. The flow of image enhancement in the MPEG decompression domain. Still blocks and motion blocks can be enhanced by filtering of Intra and Inter Q matrices, respectively. Note, \( \otimes \) is a point-by-point multiplication. The header decoding and new header encoding were implemented in software.\(^37\)

Figure 4. (a) Coefficients of “default” Intra Q matrix and (b) of its’ filtered/enhanced version \((\lambda=4, \quad a=1.5)\). The amplitude ratios between (b) and (a) provide the enhancement.
Figure 5. An example of side-by-side (split) screen view of the "Susie" sequences used in the experiment. Here the left side is an enhanced video ($\lambda=4$) and right side is an original video in mirror image. Only half the width of (352 pixels) the original and enhanced videos were used and merge into one video sequence.

Figure 6. An example of the “Table Tennis” sequence ($\lambda=4$). This sequence is an Interlace sequence with fast motion. Most subjects favored little enhancement for this and for the “Flwer” sequence.

Figure 7. The median values of the total 24 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 the “Lion” and “Susie” sequences. For the two other sequences, the subjects only favored low level of enhancement ($\lambda=2$). Note, the * indicate a significant effect at the $p < 0.05$ level.

Figure 8. The median values of the 13 subjects who did not have a bias to one side or another. The 13 subjects results show similar tendencies as the whole group of 24 subjects showd in Figure 7.

Figure 9. The effects of Intra and Inter, just Intra, and just Inter enhancement with $\lambda = 4.0$. (the preferred enhancement level used in the “Lion” enhancement) The original image (a) is an un-enhanced B (Bi-directional) picture. The Intra plus Inter enhancement
(b) shows good combining enhancing effects in moving and still areas. We used this Intra and Inter enhancement for our experiment. The Intra enhancement (c) shows the enhancement effects in all still areas. The Inter enhancement (d) shows the enhancement effects in the moving lion and the person’s trousers from the bi-directional or predictive motion of this image.

Figure 10. The differences between original and each enhanced image. Note, the Intra plus Inter enhancement (b) shows the strong combined enhancement. The Intra enhancement (c) shows the wide enhancement in all areas. The Inter enhancement (d) shows the enhancement effects in moving areas only. The static background in (d) is not enhanced at all.

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.

Table 2. The characteristics of the MPEG-2 sequences tested.
Figure 1.

Figure 2.
Figure 3.

Figure 4.
Figure 5.

Figure 6.
Figure 7.

Figure 8.
Figure 9.
Figure 10.
Table 1.

<table>
<thead>
<tr>
<th>$K$ (DCT order)</th>
<th>$f_n$ (cycles/pixel)</th>
<th>Visual frequency (cycles/degree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.44</td>
<td>9.6</td>
</tr>
<tr>
<td>6</td>
<td>0.38</td>
<td>8.3</td>
</tr>
<tr>
<td>5</td>
<td>0.31</td>
<td>6.8</td>
</tr>
<tr>
<td>4</td>
<td>0.25</td>
<td>5.5</td>
</tr>
<tr>
<td>3</td>
<td>0.19</td>
<td>4.1</td>
</tr>
<tr>
<td>2</td>
<td>0.13</td>
<td>2.8</td>
</tr>
<tr>
<td>1</td>
<td>0.06</td>
<td>1.4</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.

<table>
<thead>
<tr>
<th>Name</th>
<th>Lion.m2v</th>
<th>Susie.m2v</th>
<th>Flower.m2v</th>
<th>Tennis.m2v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profile/Level</td>
<td>Main</td>
<td>Main</td>
<td>Main</td>
<td>Main</td>
</tr>
<tr>
<td>Bit rate</td>
<td>6Mbps</td>
<td>8Mbps</td>
<td>8Mbps</td>
<td>8Mbps</td>
</tr>
<tr>
<td>Scan</td>
<td>Progressive</td>
<td>Interlace</td>
<td>Interlace</td>
<td>Interlace</td>
</tr>
</tbody>
</table>