# Image Enhancement in the JPEG Domain for People With Vision Impairment

Jinshan Tang\*, Senior Member, IEEE, Jeonghoon Kim, and Eli Peli

Abstract—An image enhancement algorithm for low-vision patients was developed for images compressed using the JPEG standard. The proposed algorithm enhances the images in the discrete cosine transform domain by weighting the quantization table in the decoder. Our specific implementation increases the contrast at all bands of frequencies by an equal factor. The enhancement algorithm has four advantages: 1) low computational cost; 2) suitability for real-time application; 3) ease of adjustment by end-users (for example, adjusting a single parameter); and 4) less severe block artifacts as compared with conventional (post compression) enhancements. Experiments with visually impaired patients show improved perceived image quality at moderate levels of enhancement.

*Index Terms*—DCT, DCT filtering, image enhancement, JPEG, quantization table, television, vision impairment.

## I. INTRODUCTION

ILLIONS of people are visually impaired, with the number of people with disabling visual problems increasing with the growing aging population. A Louis Harris survey found that vision impairment affects 17% of Americans 45 and older, and 26% of those 75 and older [1]. Visually impaired people have difficulties reading small print, watching television, recognizing faces, etc. While much research and rehabilitation effort has been aimed at improving the reading ability of low-vision patients [2], [3], the increasing use of television and personal computers heightens the need for image enhancement particular to these domains as well. Previous work on image enhancement as a low-vision aid [4]-[9] has been carried out with uncompressed images. However, many images are now handled in compressed formats, e.g., in computers and digital television, and the expected growth of such applications is likely to increase the need for enhancement that can perform well without access to the uncompressed original.

J. Kim and E. Peli are with the Schepens Eye Research Institute, Harvard Medical School, Boston, MA 02114-2500 USA.

Digital Object Identifier 10.1109/TBME.2004.834264

This paper describes an image enhancement approach for low-vision viewers applied directly within the compression domain, based on aspects of the JPEG standard compression protocol that are also applicable to MPEG compression for moving images [10]. Images delivered ultimately in a JPEG format may be enhanced: prior to compression, after decompression, or within the JPEG domain (the method chosen here). Precompression enhancement has the disadvantage of reducing the amount of compression subsequently possible as compared to the unenhanced original. For example, Hader et al. [11] proposes to low-pass filter the image before compression and then enhance it after decompression. In addition, as it is designed to reduce high-frequency content, any amount of compression will counteract the enhancement effect. On the other hand, postcompression enhancement is more likely to increase block artifacts: compressing an image creates block artifacts above, near-to, or (ideally) below their visibility thresholds, and enhancement is likely to increase the visibility of the artifacts as well. Our approach of applying image enhancement within the JPEG domain helps to reduce this problem. Since block artifacts are mainly affected by the quantization of low-frequency coefficients, keeping these (and the DC) coefficients unmodified or minimally adjusted should reduce the severity of artifacts. The algorithm implemented here has this property and indeed reduces the appearance of block artifacts.

To apply and assess the visual effects of image enhancement, one requires a visually meaningful definition of contrast (See, for example [12].) Various contrast measures have been proposed, and those defined in the spatial frequency domain may be considered applicable for use in our proposed image enhancement application. In particular, Peli [13] defined contrast for natural or complex images as the ratio of the bandpass filtered image at a given frequency band to the low-pass filtered image one octave below it. Similarly, Toet [14] defined contrast as the ratios of low-pass versions of the image. The measure we propose is based on ratio of band pass versions of the image redefined within the discrete cosine transform (DCT) domain that is used in JPEG compression [10].

As our goal is to improve everyday image viewing for low-vision patients, we chose to use TV-type monitors (NTSC, interlaced) in our subjective evaluation experiment rather than computer progressive displays; although use of computer displays is also increasing, individuals in the aging population view television screens more often and for longer periods than computer screens. Our efforts as described here for still images are expected to be applied in a future study to moving images using the (JPEG-like) MPEG format.

Manuscript received October 6, 2003; revised February 29, 2004. The work of E. Peli was supported in part by the National Institutes of Health (NIH) under Grant EY05957 and Grant EY12890. The work of J. Kim was supported by the Korea Science and Engineering Foundation (KOSEF) through a postdoctoral fellowship program. *Asterisk indicates corresponding author.* 

<sup>\*</sup>J. Tang is with the Schepens Eye Research Institute, Harvard Medical School, 20 Staniford Street, Boston, MA 02114-2500 USA, and also with the Virginia Image and Video Processing (VIVA), Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904-4743 USA (e-mail: jt6w@virginia.edu).

#### II. IMAGE ENHANCEMENT IN JPEG DOMAIN

## A. JPEG Basics

JPEG is an image compression and decompression standard that is based on the DCT [15], [16]. In the compression stage, a given image is first divided into nonoverlapping blocks of  $8 \times 8$ pixels. The two-dimensional (2-D) DCT is then computed for each block. The 64 DCT coefficients are subsequently quantized using a quantization table (a lossy operation), and thereafter they are losslessly coded and transmitted or stored together with the quantization table. In the decompression stage, each block of the received compressed data is decoded, dequantized using the quantization table, and inverse DCT transformed into an image block. The specific design of a quantization table is important because of its influences on both the compression ratio and reconstructed image quality. A basic quantization table (see [16, Fig. 5 ]) is often used in JPEG-based image compression, but other quantization tables can be derived from it by adjusting a quality factor [16]. Any other quantization table is acceptable within the JPEG standard since the table is transmitted or stored with the coded image.

#### B. Contrast Measure of Images in DCT Domain

Image enhancement methods may be classified into those that enhance contrast directly and those that enhance contrast indirectly. Direct contrast enhancement methods [17]–[19] measure the image contrast before enhancement. In this paper, we introduce a new direct contrast enhancement method based on a definition of image contrast in the DCT domain.

Let D be an  $8 \times 8$  array of DCT coefficients of an image block



The DCT coefficients represent the spatial frequency content of the image in a similar way to the coefficients in one quadrant of the 2-D Fourier domain. The  $d_{00}$  coefficient represents the DC level of the block, and the other coefficients represent spatial frequencies that increase with their distance from  $d_{00}$ . For instance, coefficients  $d_{04}$  and  $d_{40}$  represent a spatial frequency of 4 cycles/block in the horizontal and vertical directions, respectively. A band limited contrast measure in the DCT domain can be defined by

$$C_n = \frac{E_n}{E_{n-1}} \quad (1 \le n \le 14, n \in Z)$$
 (2)

where

$$E_n = \frac{\sum_{i+j=n} |d_{i,j}|}{N_n} \tag{3}$$

is the average amplitude over a spectral band enclosed by an ellipse in (1) and

$$N_n = \begin{cases} n+1, & n<8\\ 14-n+1, & n \ge 8 \end{cases}.$$
 (4)

Note that the 14 bands defined in this way represent approximately equal spatial frequencies and are consistent with the zigzag structure of coding the blocks within the JPEG standard.

#### C. Image Contrast Enhancement in the DCT Domain

Let the contrast of the various bands defined in the compressed/quantized DCT block be  $C = (c_1, c_2, \ldots, c_{14})$  and the contrast of the enhanced image be  $\overline{C} = (\overline{c}_1, \overline{c}_2, \ldots, \overline{c}_{14})$ . If we enhance the contrast by the same constant enhancement factor,  $\lambda$ , for all bands of frequencies, then the relationship between them can be described by

$$\bar{c}_n = \lambda c_n. \tag{5}$$

Thus, we have

$$\frac{\bar{E}_n}{\bar{E}_{n-1}} = \bar{c}_n = \lambda c_n = \frac{\lambda E_n}{E_{n-1}} \tag{6}$$

that can be expanded as follows:

$$\bar{E}_n = \frac{\lambda E_n}{E_{n-1}} \bar{E}_{n-1} = \frac{\lambda E_n}{E_{n-1}} \frac{\lambda E_{n-1}}{E_{n-2}} \bar{E}_{n-2}$$
$$= \lambda^2 \frac{E_n}{E_{n-2}} \bar{E}_{n-2} = \cdots \lambda^n \frac{E_n}{E_0} \bar{E}_0 = \lambda^n E_n \qquad (7)$$

where  $E_0 = \overline{E}_0$ .

Using (3) and (7), we can obtain the enhanced DCT coefficients  $\bar{d}_{ij}$  as

$$\bar{d}_{ij} = \lambda^{i+j} d_{ij} \tag{8}$$

that can be realized by weighting the dequantization table.

The enhancement algorithm (Fig. 1) shows that the modified dequantization table  $\bar{Q}$  is obtained by weighting the quantization table Q, transmitted with the compressed image, by the following equation:

$$\bar{Q} = \Lambda * Q \tag{9}$$

where the notation "\*" is point-wise multiplication of two matrices and

$$\Lambda(i,j) = \lambda^{i+j} \tag{10}$$

is the filtering matrix.



Fig. 1. Image Enhancement in JPEG Domain is achieved by weighting the quantization table with an appropriate filter (weighting array).  $\overline{Q}$  is the modified quantization table obtained by multiplying (point by point) the weighting array with the quantization table Q, which can be accessed from the JPEG bit stream.

The posttransmission enhancement of an image in the DCT domain provides distinct advantages over image enhancement methods that either utilize spatial filtering after application of an inverse discrete cosine transform (IDCT) or filter the image before transmission. The present method has minimal computational cost (only 64 multiplications) and uses the IDCT operation already performed as part of the decompression. Implementation merely requires access to the quantization table employed to decode the image. Futhermore, the method allows a user to choose the desired filter interactively. For example, the user may continuously vary  $\lambda$ , and view the result in real-time, in order to select the appearance meeting their individual requirements. This can be easily accomplished in real time.

## D. Directional Contrast Enhancement of Images in the DCT Domain

Applying enhancement in the interlaced video domain normally results in a substantial increase in interlace artifacts. Such artifacts may be reduced by enhancing the horizontal direction contrast (within a scan line) more than the vertical direction contrast. Using the same formulation as above, more contrast enhancement in the vertical direction frequency in the DCT domain (the horizontal space domain) may be achieved by limiting the enhancement to the upper-right segment of the filtering matrix using

$$\bar{d}_{ij} = \begin{cases} \lambda^{i+j} d_{i,j}, & i \le j \\ d_{i,j}, & i > j \end{cases}.$$
 (11)

#### **III.** METHODS

### A. Choice of Quality Factor for JPEG Compression

For our experiments we needed a level of compression appropriate for good quality TV pictures, one in which the compressed images were almost indistinguishable from the originals. In computer screen JPEG-based compression applications [16] the quality factor, q, is often set at 50. We tested four candidate quality factors, one of which is above and two of which are below that value, q = 15, 35, 50, and 60, and applied subjective testing by normally sighted observers to compare the quality of the variously compressed versions with their uncompressed originals. Thirty analog images were captured randomly from cable television broadcasts, and each image was compressed using the four different quality factors.

The decompressed images and the uncompressed original images were displayed on a 27-in television in random order. Subjects with normal vision were asked to rate the display images in terms of their quality as compared to a standard TV image using a graphics tablet and mouse. The distance from the TV monitor to the subject was 36 in (such short viewing distance is frequently used by visually impaired observers). Subjects were asked to rate the decompressed images as "very bad," "bad," "acceptable," "good," or "excellent," and their ratings were scored on a corresponding scale of 1 to 5.

Eight subjects (20 to 40 years old) with normal, or corrected to normal, vision participated in the experiment. The subjects' quality rating scores (1-5) were used to calculate receiver operating characteristic curves (ROCs)1; a nonstandard application of ROC analysis was used for a comparison between the ratings for the variously compressed versions and the uncompressed original. (Typically, in ROC analysis, performance is compared against a known ground truth—see Section III-G). The ROC curve was obtained as follows: For the first data point on the curve, the fraction of the subjects giving a rating of "very bad" (corresponding to a score of 1) for the original image was plotted against the fraction of the subjects giving the same rating for a particular compressed image. The corresponding cumulative fractions were calculated similarly for the other points of the curve, and the data were fitted using a binormal model.<sup>1</sup> The area under the ROC, Az, was taken as a measure of the relative quality of the compressed images. The level of correlation between the responses for the two compared conditions was used to determine the statistical significance of the difference between the areas under the two ROC curves. (A p-value of less than 0.05 level was considered to be significant.<sup>1</sup>) The ROC curves are shown in Fig. 2. For q = 15, 35, and 50, the quality of the unenhanced compressed images was inferior to that of the uncompressed images as the  $A_z$  was < 0.5 (p < 0.013). For q = 60, the perceived quality difference between the original images and the compressed images was not statistically significant (p = 0.26). Therefore, in the following experiments, q = 60 was used as the quality factor in the JPEG compression.

#### B. Directional Enhancement

In pilot experiments we noted that when enhanced images obtained by the nondirectional enhancement method were displayed on the television monitor they produced a significant flickering artifact. This flickering did not occur when the same images were displayed on a computer display. The flickering was a result of field interlacing. While flickering artifacts in a single interlaced frame are well known to occur due to image

<sup>&</sup>lt;sup>1</sup>The software used in the ROC analysis is available at: http://www-radiology.uchicago.edu/krl/toppage11.htm.



TABLE IThe 16 Image "Enhancement" Levels Used in Procedure 1. The First6 Levels Are Images Processed With Values Chosen to ProduceDegraded Images ( $\lambda < 1.0$ ). The 7th Level Is the Original Imageand the Other 9 Enhancement Levels Are Processed Usingthe JPEG Enhancement Algorithm Described HereWith the Indicated Gain Factors ( $\lambda > 1.0$ )

Level No.	Scale				
	Factor or $\lambda$				
1	0.1				
2	0.3				
3	0.5				
4	0.7				
5	0.8				
6	0.9				
7	1.0				
8	1.1				
9	1.2				
10	1.3				
11	1.4				
12	1.5				
13	1.6				
14	1.7				
15	1.8				
16	1.9				

Fig. 2. ROC analysis for the choice of quality factor, q, of compressed images as compared to uncompressed originals. The area under the ROC,  $A_z$ , is taken as a measure of the relative quality of the compressed images. The p-value indicates the statistical significance of the difference between two areas. Only for q = 60 the perceived quality difference between the original images and the compressed images was not significant; we used images with this quality factor to evaluate our enhancement technique.

motion occurring between the two fields, this artifact was also seen in images with minimal or no motion. The flickering artifact occurred when the enhancement resulted in two abutting raster line segments with a large difference in brightness. As these segments were alternately refreshed (repeating at 30 Hz), they appear to move or flicker. The effect is even stronger from the shorter observation distance typically used by low-vision persons (approximately 36 in, 91 cm). Since the artifact appeared to be associated only with the enhancement of vertical contrast (in the spatial domain), we applied the directional enhancement method (Section II-D), which substantially reduced these flickering artifacts.

## C. Image Processing

The static TV images (see Section III-E) were compressed using a standard JPEG compression algorithm with the quality factor of 60. The corresponding quantization table is

Γ13	9	8	13	19	32	41	49		
10	10	11	15	21	46	48	44		
11	10	13	19	32	46	55	45		
11	14	18	23	41	70	64	50		(12)
14	18	30	45	54	87	82	62	•	(12)
19	28	44	51	65	83	90	74		
39	51	62	70	82	97	96	81		
158	74	76	78	90	80	82	79		
	$\begin{bmatrix} 13\\ 10\\ 11\\ 11\\ 14\\ 19\\ 39\\ 58 \end{bmatrix}$	$\begin{bmatrix} 13 & 9\\ 10 & 10\\ 11 & 10\\ 11 & 14\\ 14 & 18\\ 19 & 28\\ 39 & 51\\ 58 & 74 \end{bmatrix}$	$\begin{bmatrix} 13 & 9 & 8 \\ 10 & 10 & 11 \\ 11 & 10 & 13 \\ 11 & 14 & 18 \\ 14 & 18 & 30 \\ 19 & 28 & 44 \\ 39 & 51 & 62 \\ 58 & 74 & 76 \end{bmatrix}$	$\begin{bmatrix} 13 & 9 & 8 & 13 \\ 10 & 10 & 11 & 15 \\ 11 & 10 & 13 & 19 \\ 11 & 14 & 18 & 23 \\ 14 & 18 & 30 & 45 \\ 19 & 28 & 44 & 51 \\ 39 & 51 & 62 & 70 \\ 58 & 74 & 76 & 78 \end{bmatrix}$	$\begin{bmatrix} 13 & 9 & 8 & 13 & 19 \\ 10 & 10 & 11 & 15 & 21 \\ 11 & 10 & 13 & 19 & 32 \\ 11 & 14 & 18 & 23 & 41 \\ 14 & 18 & 30 & 45 & 54 \\ 19 & 28 & 44 & 51 & 65 \\ 39 & 51 & 62 & 70 & 82 \\ 58 & 74 & 76 & 78 & 90 \end{bmatrix}$	$\begin{bmatrix} 13 & 9 & 8 & 13 & 19 & 32 \\ 10 & 10 & 11 & 15 & 21 & 46 \\ 11 & 10 & 13 & 19 & 32 & 46 \\ 11 & 14 & 18 & 23 & 41 & 70 \\ 14 & 18 & 30 & 45 & 54 & 87 \\ 19 & 28 & 44 & 51 & 65 & 83 \\ 39 & 51 & 62 & 70 & 82 & 97 \\ 58 & 74 & 76 & 78 & 90 & 80 \end{bmatrix}$	$\begin{bmatrix} 13 & 9 & 8 & 13 & 19 & 32 & 41 \\ 10 & 10 & 11 & 15 & 21 & 46 & 48 \\ 11 & 10 & 13 & 19 & 32 & 46 & 55 \\ 11 & 14 & 18 & 23 & 41 & 70 & 64 \\ 14 & 18 & 30 & 45 & 54 & 87 & 82 \\ 19 & 28 & 44 & 51 & 65 & 83 & 90 \\ 39 & 51 & 62 & 70 & 82 & 97 & 96 \\ 58 & 74 & 76 & 78 & 90 & 80 & 82 \end{bmatrix}$	$\begin{bmatrix} 13 & 9 & 8 & 13 & 19 & 32 & 41 & 49 \\ 10 & 10 & 11 & 15 & 21 & 46 & 48 & 44 \\ 11 & 10 & 13 & 19 & 32 & 46 & 55 & 45 \\ 11 & 14 & 18 & 23 & 41 & 70 & 64 & 50 \\ 14 & 18 & 30 & 45 & 54 & 87 & 82 & 62 \\ 19 & 28 & 44 & 51 & 65 & 83 & 90 & 74 \\ 39 & 51 & 62 & 70 & 82 & 97 & 96 & 81 \\ 58 & 74 & 76 & 78 & 90 & 80 & 82 & 79 \end{bmatrix}$	$\begin{bmatrix} 13 & 9 & 8 & 13 & 19 & 32 & 41 & 49 \\ 10 & 10 & 11 & 15 & 21 & 46 & 48 & 44 \\ 11 & 10 & 13 & 19 & 32 & 46 & 55 & 45 \\ 11 & 14 & 18 & 23 & 41 & 70 & 64 & 50 \\ 14 & 18 & 30 & 45 & 54 & 87 & 82 & 62 \\ 19 & 28 & 44 & 51 & 65 & 83 & 90 & 74 \\ 39 & 51 & 62 & 70 & 82 & 97 & 96 & 81 \\ 58 & 74 & 76 & 78 & 90 & 80 & 82 & 79 \end{bmatrix}$

The images were subsequently decompressed. In the decompression stage, the directional enhancement method (Section II-D) was applied to enhance the images. Only the luminance component was enhanced; the color components were not modified. For each JPEG image, we implemented 16

different processing levels, each corresponding to a different value of  $\lambda$ , ranging from 0.1 to 1.9 (Table I). The image with  $\lambda = 1$  reproduced the original decompressed unenhanced image. The levels of  $\lambda = 1.1$  to  $\lambda = 1.9$  produced enhanced images while  $\lambda = 0.1$  to  $\lambda = 0.9$  levels resulted in lowpass filtering of the images and, thus, produced degraded images. Degraded images were necessary to verify that the subjects were responding to contrast enhancement, and not just contrast modification. The inclusion of degraded images also prevents the original images from always being the lowest contrast images presented. Images were preprocessed and stored for presentation during the experiment. Fig. 3 shows samples of enhanced images. The printed images are not a valid representation of the displayed images. In particular, the printed image cannot present the flickering effect associated with the interlaced video used in the presentation (this effect also cannot be seen when the same image is presented on a progressive display).

#### D. Enhancement Evaluation by Low-Vision Patients

Two procedures were used to examine the low-vision patients' appreciation of the JPEG enhanced images. In the first procedure, a subject was instructed to select the "level" (one of sixteen choices of  $\lambda$ ) that they considered to "look the best." In the second procedure, the subject compared the original images (The original images here are unenhanced decompressed images. Because we are interested in enhancing compressed images, we do not use the images before compression in the following experiments for low-vision patients.) to the images that were processed using the median of the levels selected in procedure 1, and the subject ranked the images on a scale of perceived quality (see detailed description of procedures 1 and 2 below).

Prior to testing, each subject was asked about the size of his or her television at home, and about how close to it he or she usually sits while watching a program. Subjects were then seated







Fig. 3. Examples of processed images. (a)–(c) Samples of enhanced images; (d)–(f) difference images between the decompressed original image and the enhanced image in (a)–(c), respectively. Enhanced images in (a) and (b) were obtained using  $\lambda = 1.9$ . The enhanced image in (c) was obtained using  $\lambda = 1.5$ . Note, the effect of the enhancement as seen on the TV screen is more dramatic than it appears in print.

at a distance from our 27-in television test monitor that approximated the visual angle they were accustomed to viewing their own set at home. This distance was reduced if the subject could not discern image changes as the enhancement level varied. For our low-vision subjects, the average seating distance was  $39 \pm 14$  in from the television, while the standard viewing distance of a 27-in television would have been 105 in [20].

Because many of the subjects were elderly, with little or no computer experience, before the actual experimental session they sometimes needed a practice session to become comfortable with the graphics tablet and mouse. The images used in these practice sessions were different from those used to collect the image quality data. The room was dimly lit by recessed overhead incandescent lamps, and the luminance at the monitor surface was measured approximately 1 fc.

1) Procedure 1: Selecting the Preferred JPEG Enhancement Level: Subjects were shown a static image (drawn from a set of ten different images) on the TV screen. By moving the mouse up and down on the blank graphics tablet, they could select which of the 16 predetermined levels of contrast adjustment were applied to the image.

Each subject was asked to find the spot on the tablet corresponding to the image adjustment level where, "you like the picture the best, where it is clearest for you, and where you got the most detail out of the picture." Once subjects found an image that looked the best to them, they recorded that setting by clicking on a mouse button. After a response, the next image from the set of ten was displayed. For each trial, the mapping of the active region of the graphics tablet to the enhancement level presented was randomly shifted so that the subjects were unable to associate a fixed mechanical position with their choices.

2) Procedure 2: Perceived Image Quality: The rounded median level [21] of procedure 1 was chosen as the individually preferred enhancement level for use in procedure 2. Four versions of each of 50 images (a total of 200 images) were shown to subjects in a randomized sequence. The four versions for each

TABLE II Available Individually Chosen Enhancement Levels and Their Corresponding Second Arbitrarily Selected Enhancement Levels

Individually chosen enhancement	Second arbitrarily selected enhancement					
	10					
Level 1	10					
Level 2	10					
Level 3	10					
Level 4	11					
Level 5	12					
Level 6	12					
Level 7	12					
(Original)						
Level 8	12					
Level 9	12					
Level 10	13					
Level 11	13					
Level 12	14					
Level 13	16					
Level 14	12					
Level 15	12					
Level 16	12					

image included: 1) original image; 2) individually chosen enhancement (based on procedure 1); 3) a degraded image ( $\lambda = 0.8$ ); and 4) an image enhanced by a second arbitrarily selected enhancement level. The 50 images used in procedure 2 did not include the ten used in procedure 1.

The second arbitrarily selected enhancement level was selected to supply another enhancement which had a different appearance from the level selected by the patient. The second arbitrarily selected enhancement level was chosen to be several levels above the individually selected enhancement level for those who selected a low level of enhancement, and several levels below the individually selected enhancement level for patients who selected a high level of enhancement. Table II lists the possible individually chosen enhancement levels and the corresponding second selected enhancement levels.

By moving the mouse vertically on the graphics tablet, the subjects rated the quality of each enhanced image against an unenhanced version of the same image. The subjects were asked to rank each image as "better," "slightly better," "typical," "slightly worse," or "worse" than the original image, with these rankings printed in large font on the graphics tablet, with "better" near the top of the tablet, and "worse" near the bottom. Before the computer accepted their score, the subjects were required to compare the test image to the original image at least once by moving the mouse to a designated section at the right edge of the tablet marked by a black stripe, thereby displaying the original image. Once this was viewed, the subject was allowed to grade the test image. If desired, subjects could view the original image and compare it with the test image multiple times.

#### E. Image Acquisition

Single video frames (static images) were randomly grabbed from various shows on cable television channels in Boston, MA on June 26, 2000. The frames were captured using a Video Toaster card [22] as  $480 \times 720 \times 3$  RGB bitmap images. They were converted using Matlab [23] to RTV format for presentation on a TV monitor using a SpeedRazor graphics card [24]. Of the 200 digitized images acquired, 127 were judged by two normally sighted observers to contain little or no apparent motion due to differences between the two interlaced fields. Fifty of these were selected randomly for the study.

## F. Apparatus

All processing, experiment control, and analysis were done using an Intel PC running Windows NT 4.0 (Service Pack 6). Images were displayed on a 27-in (diagonal) Sony Trinitron NTSC format television monitor using a Video Toaster image processing system [22], under the control of programs written in Microsoft Visual Basic and Matlab [23]. In procedure 1, subjects moved the mouse over a 12-in SummaSketch III [25] graphics tablet device to select enhancement levels. The same tablet was used to grade the images in procedure 2, as described above. In both experiments, subjects designated their final choice by pressing the mouse button.

#### G. Data Analysis

Data from procedure 2 was analyzed using the ROC signal detection approach [26] described above. The Rockit program,<sup>1</sup>[27] was used to determine the area under the fitted ROC curve,  $A_z$  [28]. Paired comparisons were made between responses to the original images and a set of corresponding processed images. As there were three sets of processed images for each subject, three ROC curves were computed (see Fig. 4), representing the perceived image quality of each of the processing options as compared with the original.

In traditional ROC analysis, system (e.g., subject) responses to "noise" presentations and to "noise-plus-signal" presentations are compared. In our paper, the original images are treated as the noise presentations, and the processed versions are treated as the noise-plus-signal presentations. As can be seen in Fig. 4, our raw data consisted of multiple distributions along the perceived image quality dimension (for simplicity, Fig. 4 only shows data for 3 of the 4 image sets). When the perceived image quality of the processed images was higher than the original images [level 9 image set in Fig. 5(a)],  $A_z$  was greater than 0.5. For the degraded image set, the perceived image quality distribution was lower than that of the original images, creating an Az of less than 0.5. Our ROC analysis measures perceived relative image quality, and not enhancement detection, as might be done in another application. Consequently, the traditional labels of the axes of the ROC figure (e.g., true-positive fraction, or "hit" rate) do not represent our situation. In our analysis, the true-positive fraction dimension is the proportion of the processed image set with a higher perceived image quality than the original images, while the false-positive fraction ("false-alarm" rate) dimension is the equivalent proportion for the original images perceived as having higher quality than the processed images (a "higher" quality being relative to the criterion used for the particular point on the ROC curve). However, we use the traditional axis labels as shown in Fig. 5.

While the graphics tablet gives a continuous response measure, for some subjects, the responses were multimodal, a consequence of the large-font guide words on the tablet (many sub-



Fig. 4. One patient's distribution of responses to the original, chosen enhancement, and degraded images. For simplicity, the responses to the second enhancement are not included. The responses are taken from the graphics tablet readings and quantized to the range 0 to 10, where 0 represents a score of "worse" and 10 represents "better" in comparison to the original image. These data were for a patient who clearly preferred the enhancement; the three distributions are clearly separated. These data were used to construct the ROC curves shown in Fig. 5(a).



Fig. 5. The ROC fitted curves for two patients. The thick lines are the fits to the filled triangular symbols (the chosen enhancement level) and the thin lines are the fits to the open square symbols (the second enhancement level tested). The dotted lines are the fits to the filled diamond symbols (degraded image). (a) A patient with Optic Atrophy (visual acuity 20/250) who clearly favored the chosen enhancement (level = 9,  $A_z = 0.83$ ), and showed no preference for the second enhancement (level = 12,  $A_z = 0.49$ ). This patient, as did all others, clearly rejected the degraded images ( $A_z = 0.01$ ). (b) A more typical example in which only slight and not statistically significant preference was found for the chosen enhancement: a patient with Retinitis Pigmentosa (visual acuity 20/83). Here, preference for the chosen enhancement (level = 8,  $A_z = 0.61$ ) was slightly but not significantly higher than for the original. This subject showed no significant preference for the second enhancement level =  $12(A_z = 0.37)$ , and also rejected the degraded image (level =  $5, A_z = 0.002$ ). Note that the ROC data shown in (a) is constructed from distribution of data shown in Fig. 4.

jects did not interpolate between the five words). The data shown in Fig. 4 has a slight tendency toward this multimodal response pattern. In addition, often the response distributions were not normally distributed. Even so, the Rockit program appeared to give a reasonable fit to our data in most cases (Fig. 5). The Rockit program provides confidence limits for each ROC curve area [28], and these were used to determine the significance of the responses of individual subjects to a particular type of image processing.

Since the image enhancement levels used in procedure 1 were ordered but the perceptual intervals were not necessarily equal, nonparametric statistical tests were used for these comparisons.  $A_z$  data distributions from procedure 2 were found to be approximately normally distributed and, thus, parametric statistical tests were used for these comparisons.

#### H. Subjects

Patients were recruited from clinical practices that concentrated on retinal diseases, and most had central retinal dysfunction from such diseases as age related macular degeneration. All patients signed a subject consent form, approved by IRB committee. All included subjects were at least 18 years of age, able

TABLE III THE CLINICAL CHARACTERISTICS OF THE PARTICIPATING PATIENTS. DUE TO SCHEDULE CONSTRAINTS AND OTHER FACTORS ONLY SOME OF THE PATIENTS WHO COMPLETED PROCEDURE 1 (GROUP A) ALSO COMPLETED PROCEDURE 2 (GROUP B). N IS THE NUMBER OF SUBJECTS IN EACH GROUP, AND CFL IS THE NUMBER OF SUBJECTS WHO HAD DOCUMENTED CENTRAL VISUAL FIELD LOSS IN BOTH EYES

Group	N	Age (years) Average ± SD Median (Range)	VA (Log MAR) Average ± SD Median (range)	CFL
A	47	61± 20 68 (19-86)	$\begin{array}{c} 0.93 \pm 0.31 \\ 0.88 \ (0.50\text{-}2.10) \end{array}$	32
В	27	58±20 59(21-86)	0.99 ± 0.37 0.96 (0.50-2.10)	15

to read and understand the consent form, able to follow verbal instructions in English, and were not suffering from a condition, such as arthritis, that might inhibit their ability to control the mouse. We did not recruit low-vision subjects who use a telescope device to view television. Subjects viewed the TV images with both eyes.

Visual acuity was measured using a BVAT (Model no. 22-4850, Mentor O&O Inc). Visual fields were measured using a Bausch & Lomb Auto-Plot Tangent Screen (Cat. no. 71-54-41) to document central field loss (CFL). Visual fields were measured monocularly, using a 6-mm target at 1 meter, with the subject wearing habitual distance correction (e.g., glasses). Some of the patients did not undergo the visual field tests but had a clear diagnosis of macular lesions accounting for their acuity loss and, thus, were presumed to have CFL. One subject of the 48 total subjects who were referred did not meet the study inclusion criteria. The remaining subjects (Group A, N = 47) completed procedure 1. Due to clinical schedules and physical constraints (such as age-related stamina), fewer subjects also completed procedure 2 (Group B, N = 27). Table III shows the characteristics and numbers of subjects that completed the two portions of the experiment.

#### **IV. RESULTS**

## A. Reduction of Block Artifacts by Image Enhancement in JPEG Domain

Fig. 6 provides a comparison of JPEG-based enhancement with standard postcompression enhancement of a JPEG compressed image. To better illustrate the effects in print, we show an enlarged partial image of the familiar "Lena" image. Fig. 6(a) shows the original (uncompressed) image, while Fig. 6(b) is a JPEG decompressed image without enhancement with a peak signal-to-noise ratio (PSNR) of 35.4 dB (q = 60) [20]. Only minor degradation is evident with this level of compression. Fig. 6(c) shows the decompressed image enhanced using conventional postcompression enhancement with Paint Shop Pro<sup>TM</sup>,<sup>2</sup> applying the Image-Sharpen tool four times recursively. This produced similar contrast enhancement to the  $\lambda = 1.35$ comparison images. As seen in Fig. 6(c), this processing results in a high pass filtering contrast enhancing effect but also causes

<sup>2</sup>[Online]. Available: http://www.jasc.com/products/psp/.

an obvious increase in block artifacts. Fig. 6(d) shows the result of the JPEG-based enhancement applied in the decoding stage. Enhancement with  $\lambda = 1.35$  was used to produce a similar enhancement effect to that of Fig. 6(c), but, as is evident in the figure, that level of enhancement is achieved with less severe block artifacts.

#### B. Experimental Results for Low-Vision Patients

Fig. 7 is the histogram of the selected preferred enhancement level in procedure 1 (Group A, N = 47). The median preferred level selected was 8, corresponding to  $\lambda = 1.1$  (25% quartile: level 7, 75% quartile: level 10), which was significantly different from the original image level of 7 (Wilcoxon signed-rank test,  $Z_{34} = 3.831, p < 0.0001$ ). The enhancement levels that the patients selected were not correlated with their visual acuities (r = 0.030, p = 0.842). Seven patients preferred the original image without enhancement ( $\lambda = 1$ ). Eleven of the 47 patients actually selected degraded images, although only mildly degraded images were selected (most chose  $\lambda = 0.9$  with only two patients selecting  $\lambda = 0.8$ ). It is likely that the patients could not differentiate these low-level degradations from the original images. In response to our questions and often spontaneously, all of the patients who selected degraded images reported that the enhanced images appeared the same or were not as clear as the original images. Most of the 29 (62%) subjects who selected enhanced images reported, in response to questions or spontaneously, that the enhanced images were clearer, sharper, and easier to see than the original ones.

Twenty of the 27 subjects who completed both procedures 1 and 2 repeated procedure 1 after completing procedure 2. For this group, the median score of their first selection was 9 and the median selection on repeat was 8.0, this difference was not significant. The individual selections in the two repeats were highly correlated (r = 0.764, p < 0.0001).

In procedure 2, patients viewed 50 images, each with their individually chosen level of enhancement, with an arbitrarily selected level of enhancement (Table II) and a degraded version of the image (level 5,  $\lambda = 0.8$ ). The patients compared each of these images to an unenhanced version and indicated a comparative perceived image quality using the graphics tablet. The measurements for each of the three image versions were quantized to eleven levels and converted to ROC curves each with an associated area, Az. Fig. 5 shows the results for two subjects. One patient clearly favored the images with the chosen enhancement (this was the only patient who had such a clear appreciation of the enhancement). The other patient only slightly preferred the enhanced image and that effect was not statistically significant. These latter responses are similar to those of most of the patients. The degraded ( $\lambda = 0.8$ ) images were clearly rejected by all patients in this procedure.

If the quality of enhanced images was judged to be superior to that of the original images,  $A_z$  (the area under the ROC curve) would be larger than 0.50. Seven subjects of 27 had  $A_z$  greater than 0.5 for their individually selected enhancement level, but the difference was statistically significant for only one of these subjects ( $A_z = 0.83$ , Asymmetric 95% Confidence Interval (0.74, 0.90)). The average  $A_z$  for group B was  $0.38 \pm 0.19$ .



Fig. 6. The effect of enhancement on block artifacts. (a) The original "Lena" image. Note that the image is a magnified partial face image. (b) The JPEG decompressed image without enhancement (PSNR = 35.4 dB, q = 60). Only minor degradation is evident with this level of compression. (c) The compressed image of (b) enhanced with conventional graphics software (Paint Shop Pro<sup>TM</sup>) (PSNR = 19.9 dB). Note the clear visibility of block artifacts in the image. (d) Image enhanced with the proposed JPEG-based enhancement,  $\lambda = 1.35$  (PSNR = 21.1 dB). Note the similar contrast enhancement compared to (c) with fewer block artifacts.



Fig. 7. Preferred enhancement level distribution found in procedure 1 for Group A (N = 47) and Group B (N = 27). The two groups did not significantly differ in their selections of enhancement level (25% level = 7 and 6.5, respectively, median level = 8 for both, 75% level = 10 for both). Note that few patients selected degraded images and then only images with slight degradation were selected.

Although most subjects indicated a preference for a particular JPEG enhancement in procedure 1, most did not find individually selected enhancements to be much better than the original images. In procedure 2, subjects also viewed a second arbitrarily selected enhancement along with intentionally degraded images. This allowed us to investigate the validity of our psychophysical method. If our method was flawed, we might expect that the subjects would not report a difference in image quality for these other image sets. All of the subjects did indicate that the second enhancement set had less image quality  $(A_z = 0.20 \pm 0.16)$ , and all indicated that the degraded images had lower quality scores ( $A_z = 0.18 \pm 0.15$ ) than the original images, even though a very modest level of degradation was used. The  $A_z$  for chosen enhancement was statistically significantly different than the Az with the second enhancement (paired sample T test, t = 6.745, p < 0.0001) and the degraded enhancement (t = 4.870, p < 0.0001), but the second enhancement was not significantly statistically different than the Az with the degraded enhancement (t = 0.802, p = 0.430).

## V. CONCLUSION

An image enhancement algorithm for low-vision patients in the JPEG domain has been proposed and implemented. The algorithm, which was tested here on static images, is intended for use with moving video sequences and can be easily applied to MPEG video formats that include a JPEG-like coding.

The proposed algorithm has numerous advantages. The computational cost is very low, since it needs only to filter the quantization table in the decompression stage (64 multiplications), allowing for a real-time implementation. Because the enhancement is postcompression and could be implemented in the user's TV receiver, it could be adjusted manually by low-vision patients via a remote control unit. This would permit individual enhancement tuned to the patients' visual loss and would allow adjustment in response to the differing spatial content of images.

Experimental results have shown that most low-vision patients select a moderate level of enhancement when viewing still images displayed on a television monitor. Surprisingly in procedure 2, when the patients compared their individually selected enhancement to the original, only one subject showed a significant level of preference. The reasons for this dichotomy between the results of the two procedures are not clear and will require further investigation. Patients remarked that they preferred to see natural-looking images, and that the enhanced images were, to some extent, distorted. Whenever the patients could notice the distortion as such, they rejected it. The specific filter implemented in this paper (11) provides a uniform contrast enhancement for all frequencies (though anisotropic). While this concept is simple and the resulting filter has an elegantly simple (single parameter) structure, it may be a less than optimal way of enhancing images for the visually impaired; a limited band enhancement filter might be more effective since it reduces the distortions [29]. In addition, since most low-vision patients are completely unable to see very high frequencies, it might be better to actually suppress these frequencies as they can cause visible quantization artifacts, potentially without benefit. In addition the patients found the interlace artifacts in our static images to be particularly bothersome. When compared to frozen single frames, interlace artifacts are much less noticeable in moving images; moving videos might not be as objectionable with enhancement. Implementation and testing of this concept with moving video will be the topic of further study.

#### ACKNOWLEDGMENT

The authors would like to thank R. B Goldstein for programming help and A. Vora for helping to run the patient experiments.

#### REFERENCES

- L. Harris and I. Associates, *The Lighthouse National Survey on Vision Loss: The Experience, Attitudes, and Knowledge of Middle-Aged and Older Americans.* New York: The Lighthouse, Inc., 1995.
- [2] E. Peli, E. M. Fine, and K. Pisano, "Video enhancement of text and movies for the visually impaired," in *Low Vision: Research and New Developments in Rehabilitation*, A. C. Kooijman, P. L. Looijestijn, J. A. Welling, and G. J. van der Wildt, Eds, Amsterdam, The Netherlands: IOS, 1994, pp. 191–198.
- [3] T. B. Lawton, "Image enhancement filters significantly improve reading performance for low-vision observers," *Ophthalmic Physiological Opt.*, vol. 12, pp. 193–200, 1992.
- [4] E. Peli, "Perceived quality of video enhanced for the visually impaired," in *Vision Science and Its Applications*. Santa Fe, NM: OSA, 1999, pp. 46–48.

- [5] E. Peli and T. Peli, "Image enhancement for the visually impaired," *Opt. Eng.*, vol. 23, pp. 47–51, 1984.
- [6] E. Peli, "Simple 1-D image enhancement for head-mounted low-vision aid," *Visual Impairment Res.*, vol. 1, pp. 3–10, 1999.
- [7] L. Isenberg, A. Luebker, and G. Legge, "Image enhancement of faces for normal and low vision," *Investigat. Ophthalmol. Vis. Sci.*, vol. 30 (Suppl.), p. 396, 1989.
- [8] E. Peli, R. Goldstein, B. Young, C. Trempe, and S. Buzney, "Image enhancement for the visually impaired: Simulations and experimental results," *Investigat. Ophthalmol. Vis. Sci.*, vol. 32, pp. 2337–2350, 1991.
- [9] E. Fine and E. Peli, "Enhancement of text for the visually impaired," J. Opt. Soc. Amer. A, vol. 12, pp. 1439–1447, 1995.
- [10] J. Kim and E. Peli, "MPEG-based image enhancement for people with low vision," in *Proc. SID 2003*, Baltimore, MD, 2003, pp. 1156–1159.
- [11] O. Hader, A. Stern, and R. Koresh, "Enhancement of an image compression algorithm by pre- and post-filtering," *Opt. Eng.*, vol. 40, pp. 193–199, 2001.
- [12] Vis. Res. (Special Issue), vol. 37, no. 23, 1998.
- [13] E. Peli, "Contrast in complex images," J. Optical Soc. Amer. A, vol. 7, pp. 2032–2040, 1990.
- [14] A. Toet, "Multiscale contrast enhancement with application to image fusion," *Opt. Eng.*, vol. 31, pp. 1026–1031, 1992.
- [15] V. Bhaskaran and K. Konstantinides, *Image and Video Compression Standards: Algorithms and Architecture*. Norwell, MA: Kluwer Academic, 1995.
- [16] Handbook of Image and Video Processing. New York: Academic, 2000, pp. 513–526. R. Ansari and N. Memon, "The JPEG lossy image compression standard".
- [17] J. Tang, E. Peli, and S. Acton, "Image enhancement using a contrast measure in the compressed domain," *IEEE Signal Processing Lett.*, vol. 10, pp. 289–292, Oct. 2003.
- [18] A. Beghcladi and A. L. Negrate, "Contrast enhancement technique based on local detection of edges," *Comput. Vis., Graphics, Image Processing*, vol. 46, pp. 162–274, 1989.
- [19] W. M. Morrow *et al.*, "Region-based contrast enhancement of mammograms," *IEEE Trans. Med. Imag.*, vol. 11, pp. 392–406, 1992.
- [20] M. Ardito et al., MOSAIC Handbook. Sebastopol, CS: O'Reilly, 1996, pp. 165–181.
- [21] J. Freund and B. Perles, "A new look at quartiles of ungrouped data," *Amer. Statistician*, vol. 41, pp. 200–203, 1987.
- [22] VideoToaster. San Antonio, TX: NewTek, Inc., 1998.
- [23] MatLab Image Processing Toolbox. Natick, MA: MathWorks, 1997.
- [24] Speed Razor. Bethesda, MD: In:Sync, 1999.
- [25] SummaSketch III. Columbia, MD: GTO CalComp, Inc., 1989.
- [26] N. Macmillan and C. Creelman, Detection Theory: A User's Guide. Cambridge, U.K.: Cambridge Univ. Press, 1991.
- [27] C. Metz, P. Wang, and H. Kronman, "A new approach for testing the significance of differences between ROC curves measured from correlated data," *Proc. 8th Conf. Information Processing in Medical Imaging*, 1983.
- [28] J. Hanley and B. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, pp. 29–36, 1982.
- [29] E. Peli, "Limitations of image enhancement for the visually impaired," *Optometry Visual Sci.*, vol. 69, pp. 15–24, 1992.



**Jinshan Tang** (M'00–SM'03) received the Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, P.R. China, in 1998.

From September 1998 to June 2000, he worked as an invited researcher in ATR Media Integration and Communications Research Laboratories (MIC), Kyoto, Japan. In June 2000, he went to Harvard Medical School to work on image enhancement technology. Since Dec. 2001, he has been a Research Scientist in Department of Electrical and Computer Engineering, University of Virginia, Charlottesville.

His research interests are image enhancement for low-vision, content-based image retrieval, cell detection, and tracking, and face detection and recognition.



**Jeonghoon Kim** received the M.S. and Ph.D. degrees in electronics from the Yonsei University, Seoul, Korea in 1989 and 2000, respectively.

Since 1989, he has been a Senior Research Engineer with the Signal Processing center at Samsung Electronics, Suwon, Korea. From 2002 to 2003 he worked as a Postdoctoral Researcher at Schepens Eye Research Institute, Harvard Medical School, Boston, MA. He is an Assistant Professor in the department of Electronic Communication at the Shinheung College, Euijungbu, Korea. His interests are vision reha-

bilitation, digital image processing, 3-D processing, and digital ASIC design.



**Eli Peli** received the B.S.E.E., (*cum laude*) and the M.S.E.E. degrees from the Technion-Israel Institute of Technology, Haifa, in 1976 and 1979, respectively. He received the Ph.D. degree from New England College of Optometry, Boston, MA, in 1983.

He is a Senior Scientist and the Moakley Scholar in Aging Eye Research at The Schepens Eye Research Institute, and Professor of Ophthalmology at Harvard Medical School, Boston. He also serves on the faculty of the New England College of Optometry and Tufts University School of Medicine. Since 1983, he

has been caring for visually impaired patients as the Director of the Vision Rehabilitation Service at the New England Medical Center Hospitals, Boston. His principal research interests are image processing in relation to visual function and clinical psychophysics in low-vision rehabilitation, image understanding and evaluation of display-vision interaction. He also maintains an interest in oculomotor control and binocular vision.

Dr. Peli is a Fellow of the American Academy of Optometry, the Optical Society of America, and the Society for Information Display (SID).