1040-5488/92/6901-0015$03.00/0
OPTOMETRY AND VISION SCIENCE

Vol. 69, No. 1, pp. 15–24

symposium paper

Limitations of Image Enhancement for the Visually Impaired

ELI PELI*
Physiological Optics Unit, Eye Research Institute, and Departments of Ophthalmology, Tufts University School of Medicine and Harvard Medical School, Boston, Massachusetts

ABSTRACT

Image enhancement as an aid for the visually impaired may improve visibility of TV programs and provide portable visual aid. This paper describes the current techniques for image enhancement and their underlying models. The limitations of the various techniques and of potential methods of implementation are highlighted. Initial work in this area was based on a linear model. The finite dynamic range available in the video display and contamination of the enhanced image by high spatial frequency noise limited the model's usefulness. I propose a method to address some limitations of the original model that considers the nonlinear response of the visual system and requires enhancement of subthreshold spatial information only. This modification may increase the dynamic range available by decreasing the range previously used by the linear models to enhance visible details. However, for the modified technique to be most effective, the enhancement has to be continuously tuned, based on the patient's visual loss and the spatial frequency content of the displayed images. The implications of these limitations for the potential implementation in TV are discussed. Implementation of an image-enhancing visual aid in a head-mounted, binocular, full-field, virtual vision device may cause substantial difficulties. Patient adaptation may be difficult due to head movement and interaction of the vestibular system response with the head-mounted display. An alternate, bioptic design is proposed in which the display is positioned above or below the line of sight to be examined intermittently, possibly in a freeze-frame mode. Such implementation is also likely to be less expensive, enabling more users access to the device.

Presented at the Annual Meeting of the American Academy of Optometry in the Section on Visual Science at the Symposium entitled: Difficulties and Roadblocks in the Application of High-Tech to Low Vision, December, 1990, Nashville, Tennessee.
Received March 5, 1991; revision received June 3, 1991.
* Optometrist, M.Sc., Member of Faculty, F.A.A.O.

Key Words: head-mounted display, image processing, low vision, rehabilitation, video processing

Low-vision patients frequently report difficulties recognizing objects in continuous-tone images of natural scenes or in photographs. High-contrast photographs have been suggested to facilitate patient access to natural images.[1] Kenney[2] reported on the use of large, high-contrast photographic prints to aid visually impaired museum visitors. Peli and Peli[3] have described the use of digital image processing as an aid for the visually impaired. Digital image enhancement may be used to improve visibility of printed pictures and video images. For instance, TV programs can be enhanced either at a central broadcasting location or at the patient's receiver. Despite their difficulties, many low-vision patients continue to watch TV. To make it more accessible for the visually impaired audience, the Public Broadcasting System recently introduced a Descriptive Video Service (DVS) that broadcasts programs with a separate audio channel carrying a narrative description of the visual scene. As helpful as this service is, it does not improve TV visibility; instead, it substitutes auditory for visual information. Image enhancement may effectively supplement such a service. The same technology could be used to enhance images presented on the patient's closed-circuit TV magnifying system. It may even be possible to develop a portable system with a head-mounted, closed-circuit TV system to aid mobility.[4]

In the spirit of the Symposium, this paper is concerned with highlighting the limitations and difficulties associated with the application of image enhancement as an aid to the visually impaired. It presents the linear model that was applied initially to the design of such image-enhancement techniques and the limitations associated with that approach. Previous attempts to address the limitations are discussed and an alternative model is developed. The new model leads to a different method of enhancement, which has not yet been tested with patients. Even this improved method is limited. The limitations of the improved model and enhancement technique and their implications for potential implementation in various aids are described. The potential use of color contrast to supplement the enhancement of luminance contrast is discussed. Based on published work in this area, it appears that little benefit can be obtained from such an approach, although the natural color in the images can be of great contextual value. Because image enhancement also can be used in a portable, head-mounted device, the limitations and potential difficulties associated with such implementations are discussed as well.

PRE-EMPHASIS MODEL

A conceptual pre-emphasis model of image enhancement for the visually impaired has been proposed,[4,5] suggesting that images may be processed before presentation to the patient to compensate for the degradation caused by the patient's visual disability. The potential value of this approach was evaluated first using simulations with normal observers; photographs of enhanced images taken with a camera modified to simulate optically vision through cataracts[5] appear to provide more of the details needed for recognition. Isenberg et al.[6] found that for normal observers images enhanced using local histogram equalization reduced the contrast required for face discrimination by a full octave. However, for low-vision observers the performance was improved only for two of the three tasks. Rubin et al.[7] showed that increasing the contrast of low-pass-filtered text resulted in an increased reading rate for normal observers. For a few low-vision patients, Lawton[8,9] reported improved reading rates using band-pass-filtered text tuned for each patient's visual loss.

LINEAR PRE-EMPHASIS MODEL

The linear model[10] described in Fig. 1 implies that the contrast sensitivity function (CSF) can represent the modulation transfer function (MTF) of the visual system.[11,12] This approach ignores the well-known, highly nonlinear characteristics of the visual system. Regardless of the nonlinear nature of the visual system, it is possible to use the linear processing to analyze the appearance of images through a cataract. Linear analysis may be appropriate in this case because the cataract indeed can be represented as a linear optical filter. The ratio of a patient's CSF to a normal observer's CSF, called the Visual Degradation Transfer Function (VDTF), is assumed to measure the optical transfer function of the cataractous lens:[5]

$$\mathrm{MTF(f)_{cataract}} = \mathrm{VDTF(f)} = \frac{\mathrm{CSF(f)_{patient}}}{\mathrm{CSF(f)_{normal}}} \quad (1)$$

where f is the spatial frequency.

This ratio is used here as the linear filter applied to the image for the purpose of simulation of vision through the cataract. The inverse of this ratio, 1/VDTF(f), may be used in the linear implementation of the pre-emphasis model of enhancement (Fig. 1).

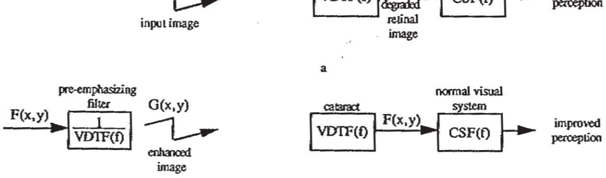


Figure 1. Linear pre-emphasis model. a: Imaging through the cataractous lens results in a degraded image on the retina. The cataract's MTF can be measured as the VDTF (equation 1). b: The pre-emphasis filter ($VDTF^{-1}$) is used to process the image before presentation to the patient. CSF, contrast sensitivity function; f, spatial frequency. (Reprinted from E. Peli, SPIE 1990;1382:49–59.[10])

PROBLEMS WITH THE LINEAR MODEL

Because most visual disabilities, including optical opacities (e.g., cataracts) and central visual field loss (e.g., macular disease), result in loss of sensitivity at high spatial frequencies, the pre-emphasizing filter should be a high-pass filter. Even when applying this model to a cataract, Peli and Peli[5] noted two difficulties: the appearance of substantial high-frequency noise in the processed image, and the limitations imposed on the pre-emphasis enhancement by the finite dynamic range of the display device. If all spatial frequencies have to be amplified by a factor of 5 to 10,[8,13] while the original image occupies most of the dynamic range available, the required enhancement cannot be attained. Rescaling of the image after such filtration back into the 0 to 255 range of gray levels available on the display will reduce the gain at all frequencies. Although this filtration will enhance spatial frequencies in proportion to the patient's visual loss, and in some cases can provide an absolute gain increase at some frequencies, it may not represent optimal use of the dynamic range of the display to aid in image recognition because the final image may not provide sufficient enhancement at any frequency.

Despite large differences in threshold contrast sensitivity for different frequencies at different eccentricities, the appearance of suprathreshold gratings of different frequencies is constant or almost constant.[14,15] Therefore, the CSF does not represent the apparent contrast of suprathreshold features in the image. This nonlinearity should be considered in the design of image enhancement for patients with central visual loss who represent the lion's share of the potential users for such an image-enhancement device. The differences in the appearance of images for patients with cataract and macular scotoma are simulated in Fig. 2.

The nonlinear characteristics of the visual system are such that different thresholds need to be applied at different spatial frequencies or scales.

Figure 2. Simulations of appearances of original images and enhanced face images (spanning 4° of visual angle) for patients with central scotoma and cataract. Both patients were assumed to have the same CSF. The left column represents what persons with normal vision see, the middle simulates vision with a central scotoma (nonlinear processing), and the right illustrates the appearance of the same images to a patient with a cataract (linear filtering). The top row is the original, unenhanced image; the bottom row is the adaptively enhanced image. Note the improvement in visibility of detail for both simulated patients and both types of enhancement.

Therefore, nonlinear simulations with a macular scotoma were based on the pyramidal structure of local band-pass-filtered contrast.[13] This pyramidal image-contrast structure enabled us to use threshold processing to simulate the appearance of images point by point and for every spatial frequency in the image. The images were partitioned in the frequency domain into 1-octave bandwidth sections. Contrast at each spatial position was calculated by dividing the band-pass-filtered value by the low-pass-filtered value at the same point. At each pyramid level, every point was compared with the appropriate contrast threshold for this level of the pyramid. If the contrast at that point was higher than the threshold, the amplitude of this point was not affected. If the contrast at the point was below the threshold, the amplitude was set to zero. Fig. 2 compares the simulated appearances of the original and enhanced facial images for patients with macular disease and cataract. The enhancement techniques used for these simulations are described below. The CSF used in both cases is the same. In one case, the simulation is of vision with a central scotoma, a nonlinear simulation. In the other case, the simulation is of appearance through a cataract, a linear optical filter, VDTF(f). The simulation suggests that face recognition may be substantially improved for both categories of patients and using both types of enhancement.

Previous Attempts to Address the Problems

When applying this model to text enhancement for patients with central field loss due to macular disease, Lawton[8] chose to reduce the high-frequency noise two ways: first, because she noted that the patients could not detect information above 8 or 10 cpd at any contrast, she limited the enhancement to the lower, visible frequency range. Second, within this range she added a high-frequency, noise-reduction factor, A, to the pre-emphasis filter, G(f):

$$G(f) = \frac{VDTF(f)}{VDTF^2(f) + A} \tag{2}$$

The proper value for the parameter A was explored empirically based on patient reading performance. Text enhancement using filters based on the individual patient's CSF was shown to reduce magnification demands modestly for patients reading with a central scotoma[8] and, for the same three patients, substantially increase the reading rate.[9] However, Lawton's text images should not have had any noise inasmuch as they were computer-generated and not digitized through a camera.

Peli and Peli[5] addressed both the high-frequency noise problem and the limited dynamic range heuristically via the application of the adaptive enhancement algorithm.[16] If the original unenhanced image occupies most of the display's dynamic range, the amplified, high-frequency component will necessarily exceed the available range. To provide the dynamic range required for this amplification, the local luminance level or the low-frequency content has to be modified as well. The adaptive filtering technique calculates a high-pass-filtered image, which may be modified locally based on local image brightness. The adaptability of the technique was used in that study only in processing the low frequencies to enable increased dynamic range for the high frequencies. The image was first separated into low- and high-spatial frequency components. The low-frequency component was obtained by calculating, for each pixel, the average brightness level found in a small window around it. The high-frequency component was obtained by subtracting the low-frequency component from the original image. The high-frequency component was then amplified. The ac portion of the low-frequency content was attenuated by a factor of 0.9, thus permitting an additional range for the amplified, high-frequency component (Fig. 3). The two modified components were then added to produce the final image.

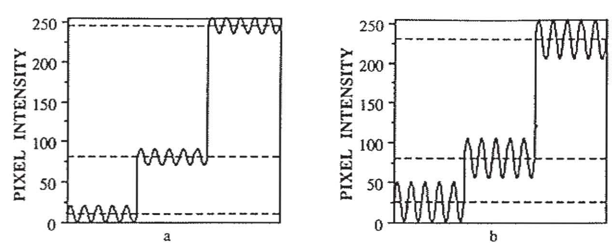


Figure 3. Extending the dynamic range in adaptive enhancement. Schematic illustration of the increase in local dynamic range obtained with the adaptive enhancement algorithm. a: The original unprocessed image. Here the high-frequency components cannot be amplified without saturation over both the low and high local luminance means. b: The processed image in which the extreme local luminance means were shifted toward the middle of the range, thus enabling high-frequency amplification without saturation (adapted from Peli and Lim[16]).

The high-frequency contrast over areas of high mean luminance is increased because both the amplitude is increased and the local mean luminance is decreased. Over areas of low mean luminance the amplitude is increased leading to increased contrast, but the increase in local luminance would counter these effects. However, Yang et al.[17] recently have shown that at very low luminance levels, increasing the luminance can result in an increase in apparent contrast at suprathreshold levels. Similarly, at threshold, contrast sensitivity increases (by a square root relation) with luminance.[18] The effects of adaptive enhancement on the appearance of images to low-vision patients are simulated in Fig. 2.

Peli et al.[4] suggested using adaptive thresholding to improve the utilization of the available dynamic range of the display. Thresholding is a method of transforming a gray-tone image into a binary one (i.e., an image with only two levels, black and white). Thresholding is not commonly considered as an enhancement technique, but may serve as such especially for the visually impaired. The binary image has inherent high contrast and, if it maintains the original image's information satisfactorily, may be useful as an enhancement technique. This technique was found to be effective with optically simulated cataracts.[4] In addition, binary display devices with high brightness and contrast may be much less expensive and thus provide the display for such an enhancement system.[19]

PATIENT RECOGNITION OF ENHANCED FACES

Although the simulation results showed considerable promise, the value of image enhancement in improving recognition of gray-scale images by visually impaired patients had to be demonstrated directly. It is generally difficult to evaluate whether the performance of an observer using enhanced images is improved. A study was carried out to determine if viewing enhanced images of human faces improved recognition for patients.[20] Faces were used to restrict the infinitely diverse range of possible image targets to a class of images for which a large body of knowledge about human performance is available.[21] Moreover, difficulty with face recognition is a frequent, early complaint of many patients with macular disease.[22]

Most of the patients (31 of 38) with a central scotoma demonstrated improved face recognition with the enhanced images as compared with the original, unenhanced images. The improvement in recognition was statistically significant for 9 of the 21 patients with macular disease tested with the adaptive enhancement ($p < 0.05$). For the two patients in this group whose recognition decreased with the enhancement, the difference was not significant. For 6 of the 17 patients with macular disease tested with the adaptive thresholding technique, a significant increase in recognition for the enhanced images was measured and 1 had a significant decrease in recognition. Patients reported that the enhanced images were significantly clearer, sharper, and easier to see. Although many patients noted that the binary thresholded images appeared distorted and cartoon-like, they were still considered clearer and easier to recognize by most patients. For many of these patients, the improvement was close to the maximum possible (Fig. 4).

These encouraging results, obtained using heuristic enhancement applied uniformly to all patients, indicate that image enhancement may be a beneficial technology for a visual aid. However, less than one-half of the patients tested had a significant improvement in face recognition. I believe that the uniform enhancement used was ineffective for so many of our patients due to one of two reasons. For patients with better vision, the enhancement was applied to details already visible to the patients. For those with poorer vision, it may have enhanced details that were not visible even when enhanced. Thus, it is now important to evaluate whether better results can be obtained by tailoring the enhancement to the individual patient's visual loss. A method for such individual enhancement, which differs from the method used by Lawton,[8,9] is proposed in the context of the modified pre-emphasis model below.

MODIFIED PRE-EMPHASIS MODEL

The main characteristic of the proposed modification to the original pre-emphasis model is its consideration of the nonlinearity of the visual response in the design of the individual compensation filters. Because suprathreshold features are perceived at their correct contrast, there is no need for enhancement. Indeed, enhancement of such fea-

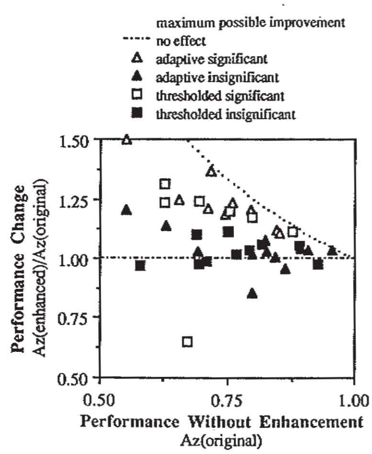


Figure 4. Enhancement results. Change in recognition for patients with central scotoma secondary to macular degeneration as a function of recognition without enhancement. Data points above dashed line represent improvement; those below represent a decrement in recognition with the enhanced images. Open symbols represent statistically significant change ($p < 0.05$); filled symbols represent nonsignificant changes. The dotted curve delineates the maximal improvement possible for each level of performance without enhancement.

tures should result in distorted appearance of the image. (The distortion will result in appearance similar to the appearance of high-pass or band-pass-filtered images to normal observers. Such images are clearly different than the original, though they are usually easily recognizable.) On the other hand, a large portion of the spatial frequency spectrum cannot be detected by visually impaired persons at any contrast; enhancing those frequencies will only waste precious dynamic range and should be avoided. Therefore, enhancement should be applied only to these spatial frequencies in the image that may still be detected despite a large increase in the threshold.

For such enhancement, the image should first be low-pass-filtered to remove all energy in frequencies not visible to the patient. The highest visible band then can be amplified by the ratio of the patient CSF at that band compared to the normal CSF (Fig. 5). The amplified enhanced band then should be added to the image and the complete new image rescaled to fit into the display range.

This analysis considers only the retinal spatial frequencies in cycles per degree without any consideration of the images' spatial content. Optimal enhancement should amplify the band of frequencies maximally in the range that is most important for recognition. These frequencies are generally discussed in terms of object spatial frequencies in cycles per object. For face images, Fiorentini et al.[23] have shown that for the midrange of spatial frequencies face recognition is better for the high-pass-filtered (above 5 cycles per face) images compared with low-pass-filtered (below 5 cycles per face) images. Hayes et al.[24] similarly have shown that the band of spatial frequencies most useful for face recognition straddles 20 cycles per face. Norman and Ehrlich[25] also found that higher spatial frequencies (above 28 cycles per picture) contribute more to the identification task (of toy tank models) than do low spatial frequencies (below 28 cycles per picture). In both studies,[24,25] the filtered images were photographed off the screen and displayed

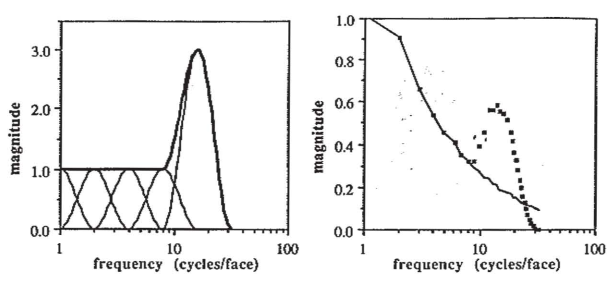


Figure 5. Enhancement of the highest band of frequencies. a: The proposed filter. The critical band (e.g., at 16 cycles per face) is amplified three times; the rest of the bands are unchanged. The resulting filter is illustrated by the thick line. b: The mean radially averaged face spectrum (solid curve) compared with the same spectrum after enhancement by the filter in (a) (■). (Reprinted from E. Peli, SPIE 1990;1382:49–59.[10])

with a slide projector. The nonlinear photographic process results in uncontrolled modification of spatial frequency content of the images. Analysis of the perception of high-pass-filtered images[26] suggests that some relevant low spatial frequency information may be obtained from high-pass-filtered images and nonlinear transformation, but the reverse, to the best of my knowledge, has not been demonstrated. In fact, Witkin[27] showed that the large-scale (low-frequency) content of an image does not contain information about the small-scale (high-frequency) content. However, large-scale information may be retrieved from small-scale image content. Thus, high frequencies (above 5 cycles per face) in terms of object spatial frequencies are more important to improvement of patients' recognition.

Although high-pass-filtered images may be sufficient for recognition, additional low-frequency information may help in presenting a more pleasing image. Aesthetics are an important consideration in our application. Patients may not want to use highly distorted images (such as the thresholded images), even if they could recognize more details from these images. Therefore, one should try to maintain as much of the low-spatial-frequency information as possible. To allow more low-frequency information, the gain of the enhancement of the high-frequency band has to be limited to the minimum required for recognition. This minimum may be at the detection threshold level or may be higher than threshold. Assuming that the minimum contrast needed for correct recognition is at the patient detection threshold, the minimum gain needed depends on the level of contrast in the original image. For example, if the contrast in the critical band, f_o, is five times the threshold level for this frequency for normal observers, the minimum gain needed for the patient is only $1/5\text{VDTF}(f_o)$. Therefore, the use of the full 1/VDTF value for the highest band may overestimate the required enhancement for this band. The cost of this extra contrast may be a reduction in the dynamic range and unneeded lowering of contrast in lower spatial frequencies. Fig. 6 demonstrates that the relevant information in the bands of 8 and 16 cycles per face appears to be at the level of 10 to 15% contrast. Because the normal observer's contrast threshold at these frequencies (assuming 4° per face) is better than 2%, a factor of 5 reduction in contrast to these bands may leave recognition unaffected. The proper reduction in gain can be determined by measuring the maximum reduction in contrast permissible without significantly affecting recognition performance. Using the results from such experiments, one can reduce the gain of the highest spatial frequency band by the appropriate factor.

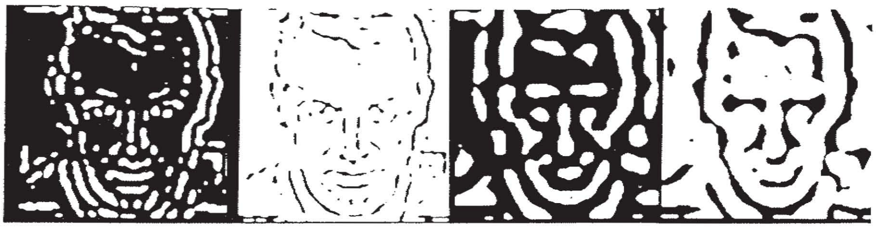
Partial Saturation to Improve Dynamic Range

Usually the enhanced image after high-pass-filtering needs to be rescaled into the range of the display. In many cases, a few points in the image that get highly enhanced may rob the rest of the image from the dynamic range. Therefore, it is common to saturate a few pixels in the enhanced image, leaving an improved dynamic range for the rest of the image. Fig. 7 demonstrates that for the bands of interest (8 and 16 cycles per face) substantial enhancement of the dynamic range may be obtained by saturating 5% of the points with little loss of relevant information. The level of saturation for face images at these bands that results in little or no degradation of image recognition should be determined. This level of saturation then can be incorporated into the enhancement algorithm, thus increasing the contrast available both for the highest visible band and the lower spatial frequency bands.

RESULTS

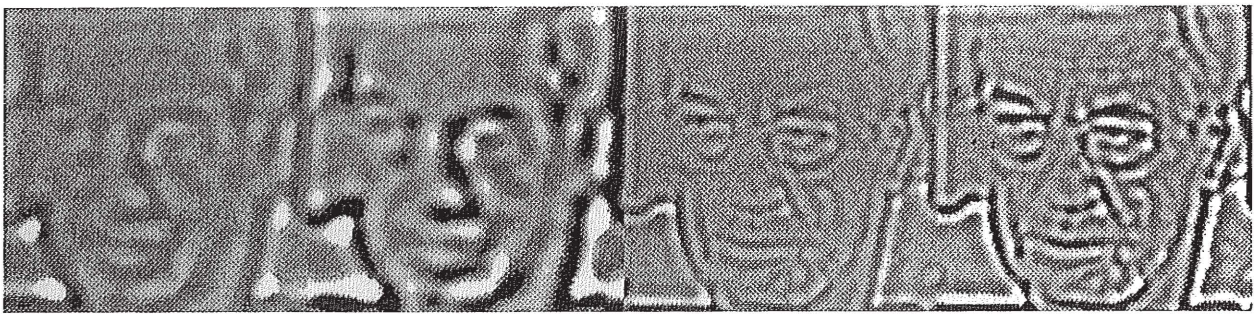
To evaluate the improvement in dynamic range that may be attained by attenuation of the lower frequencies, I have tested the effect with two face images. The images were enhanced using the filter described in Fig. 5a. The enhanced band was 1 octave wide centered at 4, 8, or 16 cycles per face. In each case, the band was enhanced by an amplification factor of 1, 2, 5, and 10, whereas the attenuation of lower frequencies implemented was 0.1, 0.5, 0.75, and 1.0. An amplification factor of 1 meant that the image was simply low-pass-filtered. The percent of image pixels that reached saturation (>255) or cutoff (<0) was calculated for each image and is presented in Table 1. The results indicate that for the face images used, saturation at the dark levels is more common than saturation at the bright levels. For amplification factors of 5 or more, the amount of saturation was substantial and generally more pronounced for the enhancement of lower frequencies. The amount of reduction of saturation obtained with attenuation of the low frequencies was relatively small for amplification factors of 5 or more, but did increase for an amplification of 2 and for enhancement of higher frequencies.

These results suggest that only low levels of enhancement are possible without substantial saturation. The benefit in dynamic range that can be attained by attenuation of the low frequency is modest and may be effective only at moderate levels of enhancement. Therefore, the enhancement should be optimally tuned to the critical band of frequencies that are just undetected by the patients. For features at these frequencies, a limited level of enhancement may be sufficient to make them visible and thus improve recognition. This specific frequency range depends both on the patient's visual loss (contrast sensitivity) and on the spatial content of the image. Therefore, on-line tuning of the device is needed both when the spectrum of the displayed image is changed and when the patient's visual sensitivity changes with progression of the



a b c d

Figure 6. Illustration of the level of a threshold contrast sensitivity needed to perceive most of the relevant information from the critical bands of spatial frequencies of a face: a: 16 cycles per face band-pass-filtered image thresholded at +10%; b: same image thresholded at −10%. Information depicted in white in a and black in b will be detected by an observer whose contrast detection threshold is 0.1 for the corresponding spatial frequency. c: 8 cycles per face band image thresholded at +10%; d: 8 cycles per face band image thresholded at −10%.



a b c d

Figure 7. Demonstrating the increase in contrast obtained by saturation of a small percent of the pixels in a band-limited contrast image: a: contrast image at 8 cycles per face; b: the same as a rescaled with 7% of lightest and darkest pixels saturated; c: contrast image at 16 cycles per face; d: same as c rescaled with 5% of pixels reaching saturation.

TABLE 1. Percent of pixels saturated due to enhancement at different levels and various spatial frequencies.[a]

			Low Frequency Attenuation															
			0.1				0.5				0.75				1.0			
			Amplification factor				Amplification factor				Amplification factor				Amplification factor			
			1	2	5	10	1	2	5	10	1	2	5	10	1	2	5	10
Image A	Frequency (cycles/face)	4	0	0	4	15	0	0	5	15	0	0	5	15	0	0	5	15
			4	15	32	40	3	15	32	40	5	16	32	40	8	18	33	40
		8	0	0	4	12	0	0	5	13	0	0	5	13	0	0	6	13
			3	11	28	38	2	11	30	39	3	14	31	40	11	21	33	40
		16	0	0	1	8	0	0	1	8	0	0	1	9	0	0	2	9
			1	6	19	31	1	7	21	32	1	9	23	34	13	20	29	36
Image B	Frequency (cycles/face)	4	0	0	7	19	0	0	7	19	0	0	7	19	0	0	7	19
			6	20	39	46	5	22	40	46	6	23	40	46	9	25	41	46
		8	0	0	2	10	0	0	2	11	0	0	3	12	0	0	3	12
			4	11	29	40	3	12	30	41	3	15	32	41	11	21	34	42
		16	0	0	0	3	0	0	0	4	0	0	0	4	0	0	0	5
			2	6	16	27	1	5	18	29	1	7	21	31	10	17	27	34

[a] In each entry the upper figure represents the percent of points of intensity >255 that had to be saturated. The lower figure represents the percent of pixels that were saturated because their filtered value became negative.

visual disability. Image spatial spectrum also changes when the patient varies his or her distance from the display. The effects of such distance changes on the spectrum and on the critical frequency for enhancement are illustrated in Fig. 8.

COLOR CONTRAST

Even though the experimental work and discussion were limited to monochrome (black and white), gray-tone images, it is clear that enhancement of color is required. Not only would color add important aesthetic values to the images, but color cues may be valuable as an additional channel of information. Enhancement of color could proceed in various ways: the images may be broken into their RGB components, and each component then can be enhanced separately and recombined. In another approach, the luminance and chromatic portions of the color video signal could be separate. The luminance (black and white) image can be enhanced as described above and then the original color signal can be added to it again. These yet untested enhancements would provide naturally colored images that are enhanced in the luminance domain.

Further improvement may be achieved by using color contrast as a method of increasing the dynamic range available in the display. Gur and Akri[28] recently demonstrated that adding very low subthreshold color contrast can increase contrast sensitivity by a factor of 2 to 3 over a range of frequencies from 0.3 to 20 cpd. These results may suggest that color contrast, phase-locked to the luminance contrast, could be used to further enhance the images. On the other hand, Legge et al.[29] found in normal observers that color contrast and luminance contrast act independently in their effect on reading. They could find no sign of additive interaction for this task. In addition, they reported no advantage of color contrast for low-vision reading; rather they found that low-vision reading is hampered by color contrast. Thus, even if color contrast facilitated luminance contrast detection, it may not be useful as an image-enhancement tool. Because in our application luminance contrast is initially stretched to a maximum, adding color contrast will necessarily reduce the available luminance contrast. Therefore, the facilitation obtained needs to be large enough to more than compensate for the loss of luminance contrast. Geisler[30] reported that although humans have similar efficiency for detecting color and luminance contrast, the addition of chromatic information increases edge detection only at low contrast levels. The results of Gur and Akri[28] and Switkes et al.[31] when considered in this framework show that the facilitation is not sufficient to provide a significant gain in detection.

IMPLEMENTATION OF IMAGE ENHANCEMENT

Implementation for Broadcast TV

The main application of image enhancement for the visually impaired is in broadcast TV. The same technology would apply to any method of transmitting regular TV programming, i.e., by cable, from VCR, or with closed-circuit TV systems. Implementing the enhancement could be performed centrally at the broadcasting station in a manner similar to the provision of closed captions for the hearing impaired or the DVS for the visually impaired. However, due to the high bandwidth requirements of the enhanced images, enhanced programs will have to be broadcast on a separate video channel, unlike the secondary channels used for captions and DVS. In addition, individual enhancement based on a patient's visual loss will not be possible with central broadcasting implementation.

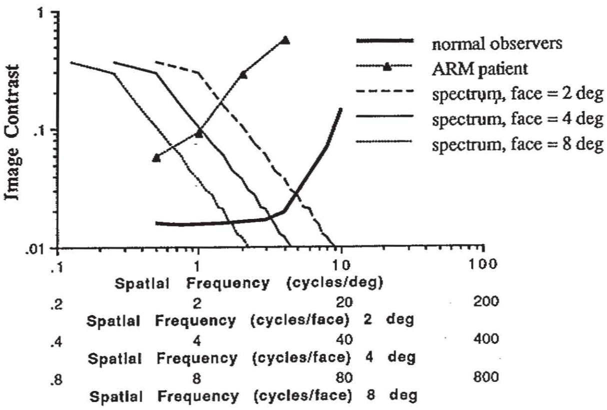


Figure 8. Relation between image spectrum and patient's contrast threshold that determines the critical frequencies. The averaged spectra from five faces are illustrated assuming three different observation distances. The face images are assumed to span 2, 4, and 8°, respectively. Frequencies higher than the point of intersection of the face spectrum and patient's threshold curve are not visible to that patient at the corresponding observation distance. The threshold curve for normal observers illustrates that the effect of distance change is less significant for normal observers. ARM, age-related maculopathy.

The requirements for a separate video channel are likely to reduce severely the availability of enhanced programming to be broadcast.

Local image processing at the patient's receiving monitor is more attractive. The advantages of local processing include an unlimited variety of programming and the possibility of enhancement tuned to patient's visual loss and/or patient-controlled tuning of the device with a knob or two. This variable tuning is important not only because a patient's visual condition may change over time, but also because the nature of various TV programs may require continuous changes in the enhancement parameters. Unfortunately, the cost of the equipment is much higher for local implementation. However, trends in this industry suggest that reasonably priced systems could be provided in the near future.

Implementation of a Head-Mounted Device

In addition to using image enhancement with a large stationary monitor for TV programs, this technology could be implemented in a portable visual aid using a miniature camera and a head-mounted display.[4,19,32]

One approach proposes using a binocular, wide-field, virtual environment display system with two cameras providing disparate images to both eyes' wide-angle displays.[33] The images presented to each eye will be enhanced, and patients with binocular vision will maintain their stereoscopic perception. The patient will see only the displayed images, and the natural view of the environment will be blocked.[34,35] This design suffers from shortcomings that will be discussed below.

A head-mounted, binocular, unit magnification enhancer display is similar to the night-vision goggles used by the military. These goggles are used despite difficulties caused by the small field and distortion of binocular depth perception at short distances, which result in poor space perception.[36] The distortions of depth perception are the result of the displacement of the imaging objective lens from the eye's pupil. The effect is prominent only at short distances. Near objects are perceived by the observers to be closer than they really are. In addition, observer's head movements result in perceivable movement of the image leading to loss of visual stability. A special optical design can reduce or eliminate these problems for the unit magnification device.[36]

Optical or electronic magnification is likely to be required in addition to image enhancement in a portable visual aid. Magnification will greatly improve the utility of the enhanced images. At the same time, magnification will complicate the use of a binocular virtual environment device. The binocular disparity will no longer be valuable, and head motion will result in amplified image motion. It is known that image motion of the magnitude anticipated will greatly reduce visual acuity and may limit the display's usefulness. Many low-vision patients successfully use optical magnification in the form of spectacle-mounted telescopes. However, these telescopes are almost always bioptic, mounted above the line of sight. Bioptic telescopes typically are used only for about 10% of the time to spot objects of interest.[37] Demer et al.[38] demonstrated that dynamic visual acuity is reduced if a 4.0× telescope is used centrally with the peripheral field occluded; however, with the peripheral view unobstructed dynamic acuity was significantly better. The design of a virtual environment aid display calls for a wide visual field. Current technology enables fields of about 50° for each eye. With magnification of 4.0×, for example, the display will provide an effective field of only 12.5°, similar to Keplerian design bioptic telescopes. Therefore, similar difficulties with a centrally mounted, constantly used device may be anticipated.

Head rotation during the wearing of telescopic spectacles with the peripheral view occluded was found to be a potent stimulus for motion discomfort.[39] Although discomfort was reduced with adaptation, individual susceptibility to motion sickness may limit the use of full-field, magnified devices for some visually impaired patients. Compensating for image motion associated with magnification and head motion may be possible, in part, if head and eye movements can be monitored accurately. Such a mechanism would further complicate the device and substantially increase its cost. In fact, even the $500 to $1000 cost of optical bioptic telescopes is too high for many elderly, low-vision patients. Appearance, field of view, weight, and cost of the visual aid were identified as the most important factors in the utilization of low-vision telescopic aids.[40]

As an alternative to the binocular, virtual environment aid, we proposed an image-enhancement aid implemented as a monocular bioptic device.[19] A head-mounted display placed above or below the line of sight may be used occasionally in the same way as the bioptic telescope. This device can combine the benefits of magnification with image enhancement without the psychological and functional drawbacks of the virtual environment device. The cost of this implementation can be reduced substantially because only one display is required. The display itself may be of a smaller field than the one required in the virtual environment inasmuch as the patient maintains his or her natural view of the environment. A larger field is required for safe navigation than is required for periodic investigation of objects of interest in a bioptic mode. A smaller field display device may be implemented in a smaller, lighter, and cosmetically more acceptable aid.

DISCUSSION

Image enhancement appears to be a viable technology for low-vision visual aids. Further improvements of the technique should be based on a more appropriate model than the original linear preemphasis model. Such improved enhancement may require individual tuning of the enhancement based on each patient's visual loss and spatial content of the image. Because of these requirements, the preferred implementation of image enhancement for broadcast TV will be local at the patient's receiver. Although central enhancement at the broadcast station may be less expensive, it will severely restrict the programs available to the visually impaired and the quality of the enhancement. A strategy for changing enhancement parameters with changes in image spectrum has not yet been developed. Manual control by the patient can be applied. However, the level of control and dexterity required for successful application may be beyond the ability of many elderly patients. Automatic image analysis and enhancement modification may be possible but will not be a trivial problem and will further increase the cost of the equipment.

ACKNOWLEDGMENTS

Supported in part by National Eye Institute Grant EY05957 (National Institutes of Health, Bethesda, MD) and a grant from the Teubert Trust Foundation.

REFERENCES

1. Covington G. Let your camera do the seeing: the world's first photography manual for the legally blind. Washington, DC: National Access Center, 1981.
2. Kenney AP. A range of vision: museum accommodation for visually impaired people. J Vis Impairment Blind 1983; 77:325–9.
3. Peli E, Peli T. Compensatory image enhancement for low vision publications. Am J Optom Physiol Opt 1982;59:41P.
4. Peli E, Arend LE Jr, Timberlake GT. Computerized image enhancement for visually impaired persons: new technology, new possibilities. J Vis Impairment Blind 1986;80:849–54.
5. Peli E, Peli T. Image enhancement for the visually impaired. Opt Eng 1984;23:47–51.
6. Isenberg L, Luebker A, Legge GE. Image enhancement of faces for normal and low vision. Invest Ophthalmol Vis Sci 1989;30(Suppl):396.
7. Rubin GS, Turano K, Schuchard RA. Effect of contrast stretching in spatial frequency range on reading grades. J Opt Soc Am Annual Meeting 1988;pg. Tuy-7.
8. Lawton TB. Improved word recognition for observers with age-related maculopathies using compensation filters. Clin Vis Sci 1988;3:125–35.
9. Lawton TB. Improved reading performance using individualized compensation filters for observers with losses in central vision. Ophthalmology (Rochester) 1989;96:115–26.
10. Peli E. New perspectives on image enhancement for the visually impaired. In: Casasent DP, ed. Intelligent Robots and Computer Vision IX: Neural, Biological, and 3-D Methods. Proceedings of SPIE 1990;1382:49–59.
11. Ginsburg AP. Visual information processing based on spatial filters constrained by biological data. 1978. Doctoral dissertation. Cambridge, England: Cambridge University.
12. Lundh BL, Derefeldt G, Nyberg S, Lennerstrand G. Picture simulation of contrast sensitivity in organic and functional amblyopia. Acta Ophthalmol (Kbh) 1981;59:774–83.
13. Peli E. Contrast in complex images. J Opt Soc Am A 1990;7:2032–40.
14. Cannon MW Jr. Perceived contrast in the fovea and periphery. J Opt Soc Am A 1985;2:1760–8.
15. Cannon MW Jr, Fullenkamp SC. Perceived contrast and stimulus size: experiment and simulation. Vision Res 1988;28:695–709.
16. Peli T, Lim JS. Adaptive filtering for image enhancement. Opt Eng 1982;21:108–12.
17. Peli E, Yang J, Goldstein R, Reeves A. The effect of luminance on superthreshold contrast matching. J Opt Soc Am A 1991;8:1352–9.
18. Banks MS, Geisler WS, Bennett PJ. The physical limits of grating visibility. Vision Res 1987;27:1915–24.
19. Peli E. Visual issues in the use of a head mounted monocular display. Opt Eng 1990;29:883–92.
20. Peli E, Goldstein RB, Young GM, Trempe CL, Buzney SM. Image enhancement for the visually impaired: simulations and experimental results. Invest Ophthalmol Vis Sci 1991;32:2337–50.
21. Ellis HD. Introduction to aspects of face processing: ten questions in need of answers. In: Ellis HD, Jeeves MA, Newcombe F, Young A, eds. Aspects of Face Processing. Boston: Martinus Nijhoff, 1986:3–16.
22. Rubin GS, Schuchard RA. Does contrast sensitivity predict face recognition performance in low-vision observers? In: Technical Digest of the Topical Meeting on Noninvasive Assessment of the Visual System, vol 3. Washington, DC: Opt Soc Am, 1990:130–3.
23. Fiorentini A, Maffei L, Sandini G. The role of high spatial frequencies in face perception. Perception 1983;12:195–201.
24. Hayes T, Morrone MC, Burr DC. Recognition of positive and negative bandpass filtered images. Perception 1986;15:595–602.
25. Norman J, Ehrlich S. Spatial frequency filtering and target identification. Vision Res 1987;27:87–96.
26. Peli E. Perception and interpretation of high-pass filtered images. Opt Eng, in press.
27. Witkin AP. Scale space filtering. In: Pentland AP, ed. From Pixels to Predicates: Recent Advances in Computational and Robotic Vision. Norwood, NJ: Ablex Publishing Corp, 1986:5–19.
28. Gur M, Akri V. Human contrast sensitivity is enhanced by color. Invest Ophthalmol Vis Sci 1990:31(Suppl):264.
29. Legge GE, Parish DH, Luebker A, Wurm LH. Psychophysics of reading XI: comparing color contrast and luminance contrast. J Opt Soc Am A 1990;7:2002–10.
30. Geisler WS. Ideal observer analysis of visual performance. Proceedings of the Applied Spatial Vision Models for Target Detection and Recognition. San Antonio, TX: Armstrong Laboratory Advisory Group Conference, US Air Force, 1991.
31. Switkes E, Bradley A, DeValois KK. Contrast dependance and mechanisms of masking interactions among chromatic and luminance gratings. J Opt Soc Am A 1988;5:1149–62.
32. Massof RW, Rickman DL. Obstacles encountered in the

development of the low vision enhancement system. Optom Vis Sci 1992;69:32–41.
33. Fisher SS, Tazelaar JM. Living in a virtual world. BYTE 1990;15(7):215–21.
34. Rubin GS. Low-vision enhancement with space-age technology. A compilation of papers presented at the Tenth National Science Writers Seminar in Ophthalmology. New York: Research to Prevent Blindness, 1988:57–8.
35. Schmeck HM Jr. Eyeglasses with T.V. to improve vision. New York Times; April 28, 1988.
36. Hadani Y. Night vision equipment. U.S. patent #4467190, 1984.
37. Jose R, Ousley BA. The visually handicapped, driving with bioptics—some new facts. Rehab Optom 1984;2:2–5.
38. Demer JL, Porter FI, Goldberg J, Jenkins HA, Schmidt K. Dynamic visual acuity with telescopic spectacles: improvement with adaptation. Invest Ophthalmol Vis Sci 1988; 29:1184–9.
39. Demer JL, Porter FI, Goldberg J, Jenkins HA, Schmidt K. Adaptation to telescopic spectacles: vestibulo-ocular reflex plasticity. Invest Ophthalmol Vis Sci 1989;30:159–70.
40. Greene HA, Pekar J. Bioptic telescope utilization survey. J Vis Rehab 1987;1:39–48.

AUTHOR'S ADDRESS:
Eli Peli
Eye Research Institute
20 Staniford Street
Boston, Massachusetts 02114