Computerized Image Enhancement for Visually Impaired Persons: New Technology, New Possibilities

E. Peli; L.E. Arend, Jr.; G.T. Timberlake

Abstract: Patients with age-related visual loss suffer reduced ability to recognize faces and other scenes in photographs and on television. Recently, progress has been made in image enhancement, using controlled distortion of digitally stored images that increases their usefulness in particular applications. Described are two approaches to image enhancement for the visually impaired. In one approach, the visual losses that characterize individual patients and disease classes are described using detailed measurements of visual degradation transfer functions, which are profiles of loss of image information at various spatial scales. The particular distortion used for image enhancement is then adjusted to the impairment of the individual patient or disease class. A second approach takes advantage of the resemblance between the visual losses of many patients and the degradation of picture information in other applications due to external limitations (e.g., fog and haze) on photography. Several enhancement algorithms have been found useful with such images and may also improve picture recognition by the visually impaired.

Interest in the importance and value of pictures to visually impaired people is increasing. For example, a photography manual for visually impaired people was written by a legally blind professional photographer (Covington, 1981); a conference on photography as an aid to the visually impaired was held in Cambridge, Massachusetts (Photography, 1983); and several museums have started using large black-and-white photographs to facilitate access to museums for visually impaired patrons (Kennedy, 1983). The production of black-and-white, high-contrast photographic prints of original artifacts or of color photographs is an example of image enhancement. These photographic techniques and other simple measures, such as cutting a photograph along major contour lines and mounting it on a contrasting background, have been applied for many years in the large print edition of *Reader's Digest*. However, most other large print publications avoid photographs and images of any kind.

Computerized image enhancement offers improved possibilities for those with low vision to perceive pictures in various settings and dramatically extends the limited attempts described above. For example, this technique could be used to enhance photographs before they are printed in publications; to enhance television images from either a central broadcasting location or the patient's own receiver, providing a service similar to captions for the hearing impaired; or to supplement video cassette recorders (VCRs) and closed-circuit TV (CCTV) systems. In the future, it should also be possible to process pictures from a head-mounted CCTV system to aid mobility. Further, image enhancement is considered essential for implementing cortical visual prostheses (Stockham, 1971), and the knowledge acquired in applications for low vision patients may be valuable for the design of such prostheses.

Computerized image enhancement techniques include various processes that can improve the visual appearance of the image or convert the image to a form more suitable for analysis by human observers. Usually, the image is actually distorted in a way that will facilitate extraction of some predetermined aspect of the image; for example, enhancing the edges of images makes them more pleasing to observers with normal vision (Pratt, 1978).

Representation of Images in Computers

Computer enhancement of images requires that the images be available to the computer as an array of numbers. Video photography is one common method of acquiring images. The image of the scene falls on the sensors of the camera, and local light intensities are converted into a series of analog electrical signals. The signals are sampled and digitized to produce an i x j two-dimensional array of numbers corresponding to light levels in the image. At each sampled location (called a picture element, or pixel), intensity is represented by one of 2^n numerical values corresponding to gray levels (where n is the number of computer "bits" required to represent all those values). The required i, j, and n depend on the intended use of the picture. Resolution that is too low results in lost information and objectionable distortion (Figures 1, 2). Unnecessarily high resolution wastes memory and processing time. For example, a 512 x 512 pixel image with 256 gray levels/pixel requires 262,144 bytes of storage, and each pixel-by-pixel calculation involved in enhancement must be done 262,144 times. A 128 x 128 image with 16 levels requires only 6,144 bytes of storage and 16,384 operations to do the same calculation.

For normally sighted observers, line drawings and graphs require high spatial

Figure 1. An image digitized at a resolution of 256 x 256 and at 256 gray levels. Compare with the same image at reduced resolution in Figure 2 and with coarser gray levels quantization in Figure 3.

The authors thank Tamar Peli for valuable discussions and suggestions. This work was supported in part by grant #EY05450 and #EY05957 from the National Eye Institute. Part of the imaging work was done at the Image Analysis Laboratory, Tufts-New England Medical Center, Boston.

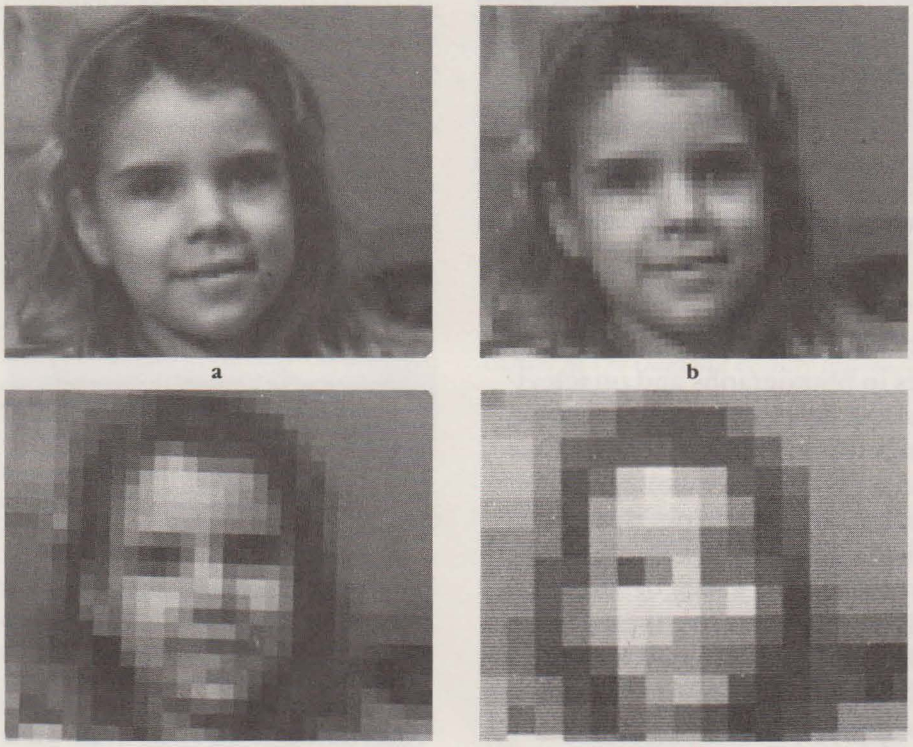

Figure 2. The image from Figure 1 displayed at reduced resolution. ***a:*** **128 x 128.** ***b:*** **64 x 64.** ***c:*** **32 x 32.** ***d:*** **16 x 16. Note that the girl can be identified in** ***d*** **if one blurs one's own vision by removing glasses or squinting eyes when looking at the image.**

resolution but only two intensity values. Satisfactory news photographs can be presented with as few as 128 x 128 pixels and 32 gray levels, but high-quality photographs for artistic or technical work may require 1,024 x 1,024 pixels and 1,024 gray levels (Figures 1, 3). On the other hand, under some conditions, photographs of faces can be recognized with remarkably little information (Harmon, 1973). Subjects can accurately identify one of a small set of faces from a picture with only 16 x 16 pixels and only 8 gray levels. Furthermore, severe optical blur, sufficient to obliterate all but the coarsest details of head, shoulder, and hairline shape, still allows far better than chance identification.

Image Enhancement Techniques

Image enhancement techniques may be classified into three major groups: contrast manipulation or rescaling, spatial filtering (usually by some form of high-pass filtering), and pseudocolor coding.

Contrast manipulation or rescaling

A common problem of photographic and video images is low contrast resulting from reduced brightness range. Image contrast may often be improved by rescaling the image to use the full range of the display medium. This can be done photographically (by changing the gamma), but is simpler and more reproducible by use of a computer (Pratt, 1978). Improving contrast is useful for observers with normal vision when they study such reduced-contrast images as medical x-rays or satellite images of earth. Increasing image contrast of better quality may make them more usable to partially sighted observers whose contrast sensitivity is generally reduced (Brown, 1981).

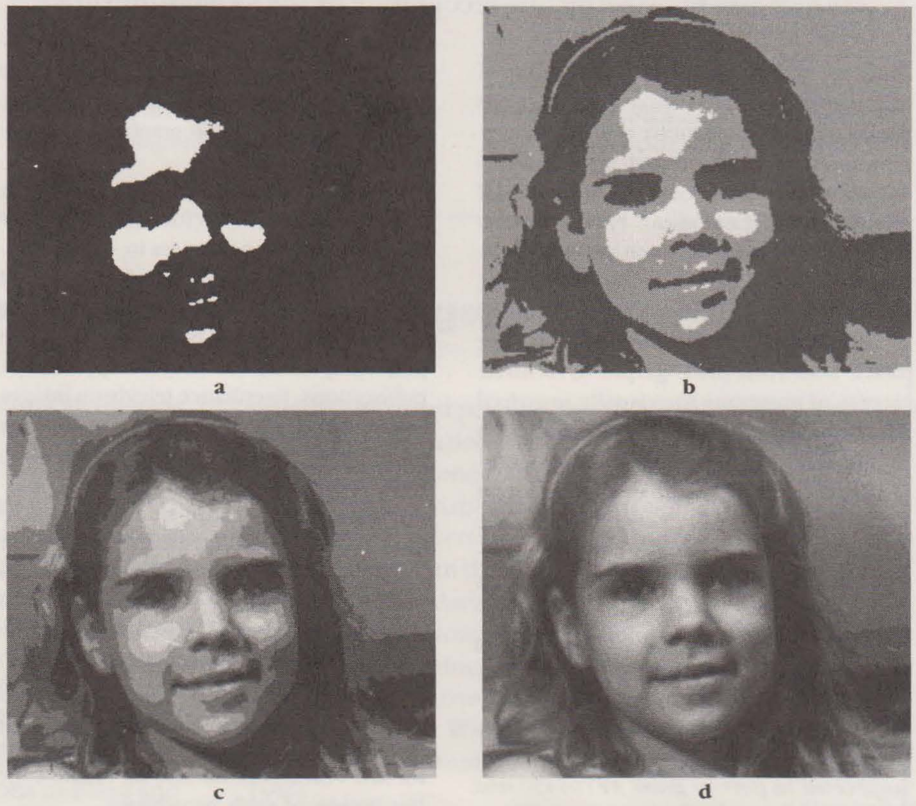

Figure 3. The image from Figure 1 displayed with reduced number of gray levels. ***a:*** **1-bit (2 levels) presentation.** ***b:*** **2-bits (4 levels).** ***c:*** **3-bits (8 levels).** ***d:*** **4-bits (16 levels). For 5-bits and more, the image is indistinguishable from the original 8-bit image.**

The simplest rescaling technique involves redistributing the image gray levels to cover the full gray-level range available in the display (Figures 4a, b). In a similar technique, the gray levels are stretched over a range larger than the display range, and the extreme brightness values are clipped to the maximum and minimum limits. The rescaling does not have to be a monotonically increasing function. For example, reversing contrast may enhance some images, whereas others may benefit from a "saw-tooth" transformation in which various subregions of the original images are mapped to the full display range, resulting in stronger local contrast variations as well as contrast reversals (Figure 7c). An even more nonlinear operation may include

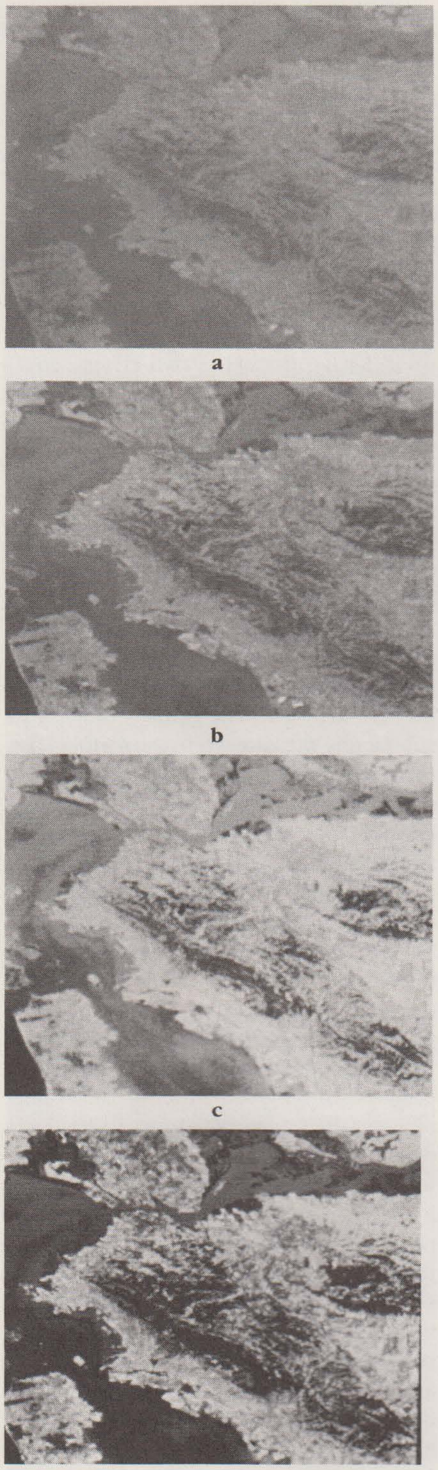

Figure 4. Histogram modification techniques. ***a:*** **An original satellite scan of San Francisco Bay. The image is flat and therefore has a narrow histogram of gray levels.** ***b:*** **Image rescaled to the full range of the display.** ***c:*** **Histogram equalization of the image in** ***a.*** **The gray level histogram of this image is almost flat.** ***d:*** **Histogram hyperbolization of the same image resulting in less blooming of bright areas and more visible details.**

Figure 5. Various enhancement techniques applied to the image of a girl's face. ***a:*** **The original continuous tone image.** ***b:*** **A two-level presentation of continuous tone image using adaptive thresholding. The improved visibility of the processed image may be demonstrated by looking at the two images through a thin tissue paper to simulate corneal or lenticular diffusion.** ***c:*** **The image in a process with extremum sharpening algorithm.** ***d:*** **The same image enhanced with adaptive filtering. The parameters were adjusted to enhance maximally at dark areas of the image to delineate the details of the girl's hair.**

selecting a specific subrange of gray levels and setting the pixels within this range to a new, usually extreme, value.

• *Histogram equalization and hyperbolization.* The gray levels histogram representing the frequency of appearance of each gray level in the image is an important measure of image quality (Pratt, 1978). The gray levels histogram of many photographs is not only narrow but is usually highly skewed toward dark levels. In such photographs, fine details are difficult to observe in the darker regions of the image. This kind of image may be enhanced by a different operation that provides an image with a histogram in a more desirable form. One such technique transforms the image's histogram into a uniform (flat) histogram (Hall, 1974). With this histogram equalization technique, bright regions in the original image often become extremely bright in the enhanced image, resulting in loss of detail and contrast (Figure 4c). Another technique, histogram hyperbolization, forces the enhanced image's histogram into a hyperbolic form. This modification reduces the possibility of blooming and loss of details in bright regions of the image. This technique was postulated to be matched to the logarithmic response of the retinal receptors in such a way that the histogram equalization will actually occur after the initial retinal processing of the image (Frei, 1977) (Figure 4d).

• *Thresholding.* Thresholding, an image processing technique that is not commonly considered an enhancement technique, may serve as such, especially for visually impaired individuals. Thresholding is the method of transforming an image into a binary image (i.e., an image with only two levels—black and white). These transformations usually serve as an information-compressing technique; reducing the number of gray levels in the image reduces the amount of information that must be transmitted over any medium of communication. Such information reduction increases the speed at which the image can be transmitted and significantly reduces the cost of the transmission. This property is especially important for video conferencing via satellite transmission. The binary image has high contrast, and if it maintains the original image's information in a satisfactory way it may be a useful enhancement technique for visually impaired viewers. Thresholding an image may be performed in various ways. The simplest way is to choose a single gray level and set any pixel with higher level to 255 (white) and every pixel with less than the threshold to 0 (black). However, in many cases no threshold value that will give a binary image of sufficient detail and clarity can be found. Therefore, techniques using variable and adaptive (Chow & Kaneko, 1972; Jarvis & Roberts, 1976) thresholds that change across the image as a function of local image properties are often needed. Figures 5a and b illustrate examples of images thresholded with such an adaptive technique.

Image enhancement by spatial filtering. Peli and Peli (1984) proposed a general preemphasis model of image enhancement for the visually impaired. This type of model is commonly used in analysis of communication systems. When the channel effect on signals (i.e., the channel transfer function) is known, the signal can be processed before being fed into the channel in a manner that will compensate for the effects of the channel (i.e., preemphasized) (Carlson, 1974).

In the visual system, the image presented to the eye is transformed by the contrast sensitivity function (CSF) to yield the perception (Figure 6a). In the low vision observer, the degradation of the image, whether optical, retinal, or neural in origin, may be described by another function of spatial frequency—the visual degradation transfer function (VDTF) (Peli & Peli, 1984). The VDTF represents the effect of the channel on the image before basic transformation by the CSF (Figure 6b). The preemphasis process therefore involves filtering the image with a filter that is the inverse of the VDTF (Figure 6c).

The threshold CSF data obtained from low vision patients may be converted to VDTFs. The VDTF of any patient is obtained by dividing the CSF of a normal eye into the measured CSF of the low vision eye (Peli & Peli, 1984). VDTFs may be classified into at least two groups: a low-pass filter VDTF (for cataract patients) (Hess & Woo, 1978; Hess & Carney, 1979), and a low-pass filter with significant intermediate frequency suppression (for macular disease) (Ginsburg, 1981; Loshin & White, 1984). Direct preemphasis approach will involve application of the appropriate filter either on a patient-by-patient basis or by patient classification in terms of the images presented. We are currently evaluating this approach. However, because most VDTFs that can be calculated from patients' CSFs represent high-frequency loss, a number of general high-pass image filtering techniques may also be valuable. These techniques include unsharp masking, extremum sharpening, and adaptive filtering. Their main advantage over direct preemphasis is that they require significantly fewer calculations per image and thus may be implemented more quickly on the computer. These techniques may be implemented in "real time" to process and enhance video images at the rate they are received at the television set with current technology.

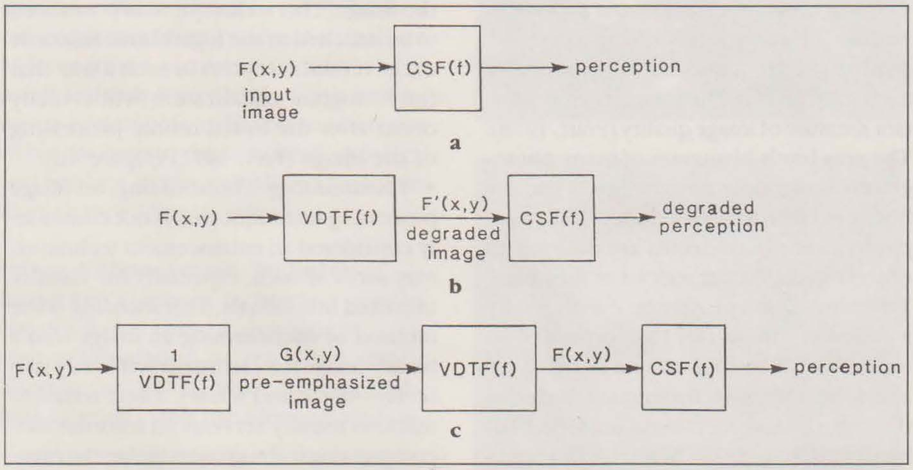

Figure 6. Schematic presentation of the spatial filtering preemphasis model for image enhancement. *a:* System diagram representing the role of CSF in spatial perception. F(x,y) represents the input image. *b:* Effect of the visual degradation transfer function (VDTF) on the spatial perception of low vision patients. The VDTF is the ratio of CSF of low vision eyes to CSF of normal eyes. *c:* Compensation for the effects of the VDTF is theoretically possible with application of a preemphasis filter (1/VDTF) to the image before it is presented to the patient.

- *Unsharp masking* is a technique commonly used in photography to enhance edges (Pratt, 1978). A positive, optically blurred transparency is made, and the final print is obtained by exposing the paper through both the original negative and through the blurred mask superimposed on it. This process has two effects: first, it results in overshooting edges (Mach Band-like effect) which make the image look sharper, and second, the masking reduces the overall brightness changes across the image, leaving more of the dynamic range for local high-frequency changes. Using the computer, this can be achieved easily either by simple calculations or by digitizing the image into two parallel memory planes, one with normal resolution and the other with low resolution (analog to the optical blur), and subtracting the latter from the former. The relative weight given to each component can be easily controlled, and the full display range may be used by rescaling the result of the subtraction.
- *Extremum sharpening* is a local operation that de-blurs or sharpens images. It has been applied successfully in processing biomedical images such as blood cells (Lester, Brenner, & Selles, 1980). The computer assigns to each image point a new value that depends on the neighboring points within a given window. Each point receives the value of the level of either extreme, the brightest or the darkest, in that window that is closer to the original gray level of the point. If the algorithm is applied repeatedly to the same image, it tends to consolidate regions of equal density and mask underlying texture and at the same time sharpen the border between such regions. Figure 5c illustrates an example of the result of a single application of extremum sharpening.
- *Adaptive filtering* was developed for images degraded by cloud cover or shadows (Peli & Lim, 1982). The adaptive nature of the algorithm is the result of changing the processing parameters at every image point on the basis of local image brightness. The image is first separated into spatial low- and high-frequency components. The low-frequencies component is obtained by calculating for each pixel the average brightness level found in a small window around it. The high-frequencies component is obtained by subtracting the

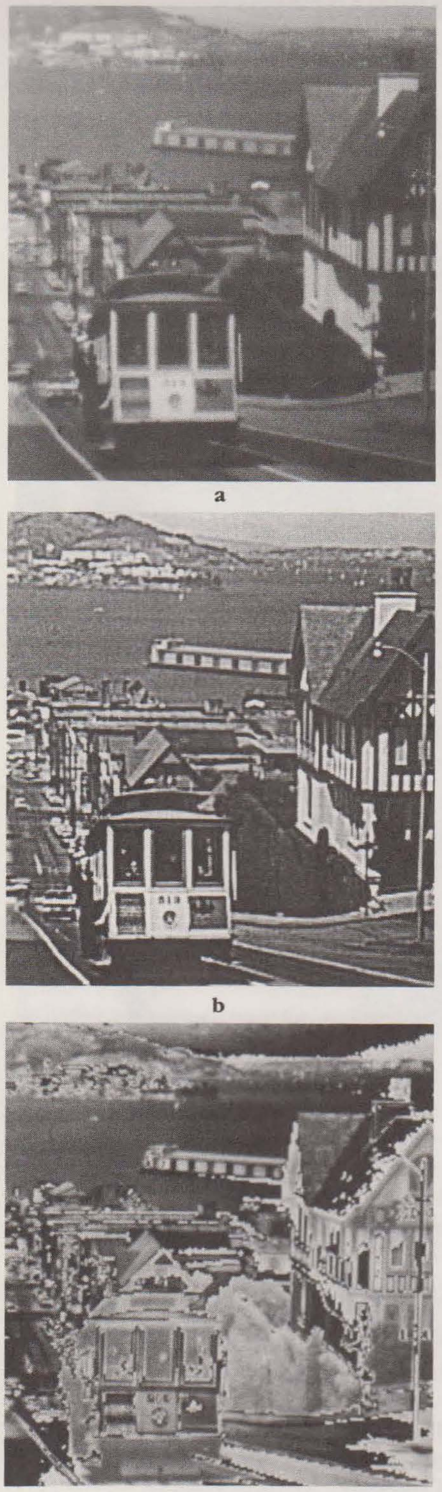

Figure 7. Enhancement of the San Francisco cable car scene. *a:* Original image. *b:* Adaptive filtering to enhance high-frequency content of the image in *a*. Note the visibility of the car number, the faces of the passengers, and the details in the far background. *c:* "Saw-tooth" enhancement of the same image by scaling the gray levels of the image four times into the full display range.

Figure 8. Optically simulated cataract demonstrates the benefit of image enhancement. The original photograph of an old city (left) was enhanced using the adaptive enhancement algorithm (right). The two images were then photographed together through a camera that was rendered "cataractous" by applying petroleum jelly to the lens.

low component from the original image. The high-frequencies component is amplified, the amount of amplification being dependent on the local low-frequencies content in a way chosen by the operator. For example, it is possible to enhance the high frequencies more in very bright sections of the image than in darker sections. The final local luminance level may also be modified in any way to achieve more or less overall change in brightness across the image. The two modified components are then added to produce the final image (Figures 5d, 7b). Peli and Peli (1984) suggested implementing this method of image enhancement for the visually impaired and found it helpful when tested with optically simulated cataractous vision (Figure 8).

Image enhancement by pseudocolor recoding. In pseudocolor coding, a widely used image enhancement technique, gradations of intensity are presented as gradations of hue, which takes advantage of the human visual system's sensitivity to hue differences. Small intensity differences not visible in black-and-white images are often visible when displayed as color differences in such pictures. Common applications are surface and satellite infrared photographs showing heat distributions. The usefulness of pseudocolor recoding for low vision applications is unknown. One disadvantage of the technique is the necessity of learning to see the recoded patterns. Gradient information is presented in relatively unfamiliar ways.

Summary

Video and computer technology applications for low vision patients have been concentrated on processing of textual data only (Stockham, 1971). Image enhancement is a fast-growing branch of the new field of computerized image processing. The techniques being developed in this field, in combination with analysis of spatial vision with the CSF, promise to offer new possibilities for the partially sighted. We are currently exploring these possibilities in a study testing the benefits of image enhancement based on CSFs in face and scene recognition.

References

Brown, B. (1981). Reading performance in low vision patients: Relation to contrast and contrast sensitivity. *American Journal of Optometry and Physiological Optics* 58, 218-226.

Carlson, A.B. (1974). *Communications systems,* 2nd ed. New York: McGraw Hill, 259-262.

Chow, C.K., & Kaneko, T. (1972). Automatic boundary detection of the left ventricle from cineangiograms. *Computers and Biomedical Research,* **5,** 388-410.

Covington, G. (cited by A. Ford) (1981). *Let your camera do the seeing: The world's first photography manual for the legally blind.* Washington, D.C.: National Access Center.

Frei, W. (1977). Image enhancement by histogram hyperbolization. *Computer Graphics & Image Processing,* **6,** 286-294.

Ginsburg, A.P. (1981). Spatial filtering and vision: Implication for normal and abnormal vision, in *Clinical applications of* visual psychophysics. Proenza, L.M., Enoch, J.M., & Jampolsky, A., eds. New York: Cambridge University Press.

Hall, E.L. (1974). Almost uniform distributions for computer image enhancement. *IEEE Transactions on Computers* C-23, 207-208.

Harmon, L.D. (1973). The recognition of faces. *Scientific American,* **229,** 71-82.

Hess, R., & Woo, G. (1978). Vision through cataracts. *Investigative Ophthalmology and Visual Science,* **17,** 428-435.

Hess, R.F., & Carney, L.G. (1979). Vision through an abnormal cornea: A pilot study of the relationship between visual loss from corneal distortion, corneal edema, keratoconus, and some allied corneal pathology. *Investigative Ophthalmology and Visual Science* 18, 476-483.

Jarvis, J.F., & Roberts, C.S. (1976). A new technique for displaying continuous tone images on a bilevel display. *IEEE Transactions on Communications,* COM-24, 891-898.

Kenney, A.P. (1983). A range of vision: Museum accommodation for visually impaired people. *Journal of Visual Impairment & Blindness,* **77,** 325-329.

Lester, J.M., Brenner, J.F., and Selles, W.D. (1980). Local transforms for biomedical image analysis. *Computer Graphics & Image Processing* 13, 17-30.

Loshin, D.S., & White, J. (1984). Contrast sensitivity: The visual rehabilitation of the patient with macular degeneration. *Archives of Ophthalmology,* **102,** 1303-1306.

Peli, E., & Peli, T. (1984). Image enhancement for the visually impaired. *Optical Engineering,* **23,** 47-51.

Peli, T., & Lim, J.S. (1982). Adaptive filtering for image enhancement. *Optical Engineering* 21, 108-112.

Photography seen as aid to low vision. (1983). *AOA News,* **22**(3), 4.

Pratt, W.K. (1978). *Image enhancement and restoration in digital image processing.* New York: Wiley, 307-407.

Stockham, T.G., Jr. (1971). The importance of non-linear image enhancement in visual prosthesis design, in *Visual prosthesis: The interdisciplinary dialogue.*

Sterling, T.D., Bering, E.A., Jr., Pollack, S.V., & Vaugham, H.G., Jr., eds. New York: Academic Press, 173-192.

Eli Peli, M.Sc., O.D., is assistant scientist at the Eye Research Institute of Retina Foundation, assistant professor in the Department of Ophthalmology at Tufts University School of Medicine, and director of Low Vision Service at the New England Medical Center. L.E. Arend, Jr., Ph.D., and G.T. Timberlake, Ph.D., are associate scientists at the Eye Research Institute of Retina Foundation. Address inquiries to the authors at Eye Research Institute, 20 Staniford St., Boston, MA 02114.