Image enhancement for the visually impaired: the effects of enhancement on face recognition

Eli Peli

The Schepens Eye Research Institute, Harvard Medical School, Boston, Massachusetts 02114, and The New England Eye Center, Tufts University School of Medicine, Boston, Massachusetts 02111

Estella Lee, Clement L. Trempe, and Sheldon Buzney

The Schepens Eye Research Institute, Harvard Medical School, Boston, Massachusetts 02114

Received September 27, 1993; revised manuscript received January 13, 1994; accepted February 1, 1994

Image enhancement has been shown to improve face recognition by visually impaired observers. We conducted three experiments in an effort to refine our understanding of the parameters leading to this effect. In experiment 1 we found that the band of spatial frequencies between 4 and 8 cycles/face is critical for face recognition. In experiment 2 we found that enhancement of these frequencies and the resulting image distortion actually reduced recognition performance for normal observers. Since the degradation of performance by low vision is larger than the effect of distortion, the enhancement that reduces performance for normal observers may still be beneficial for the visually impaired observer. Experiment 3 found that patients tend to prefer images enhanced at frequencies higher than the critical frequencies found in experiment 1. Such individually selected enhancement did not improve recognition in comparison with uniformly applied enhancement. The lack of an enhancement effect may be due to the small variability in enhancement frequencies selected by our subject population.

INTRODUCTION

A selective loss of sensitivity at high spatial frequencies is often associated with visual impairment. This sensitivity loss, and the difficulties that many low-vision patients have in recognizing faces, led us to propose, test, and demonstrate that high-pass filtering (adaptive enhancement) of images improves the recognition of faces.[1] In this paper we describe three experiments aimed at better understanding the effects of enhancement on face recognition, for the purpose of designing improved image-enhancement methods for the visually impaired.

Initial work in this area was based on a linear pre-emphasis model.[2] In this model the image is processed with a linear filter designed to compensate for a patient's contrast-sensitivity loss. However, the finite dynamic range available in the video display and the contamination of the enhanced image by high-spatial-frequency noise limited the model's usefulness. Peli[3] recently proposed an approach that addresses some limitations of the original model by considering the nonlinear response of the visual system (contrast constancy) and requiring enhancement of subthreshold spatial information only. The linear model of Peli and Peli[2] requires enhancement of information at all frequencies at which patients have sensitivity loss. This results in enhancement of details at relatively low frequencies that, because of their original high contrast, were visible to the patients. This unnecessary enhancement uses precious dynamic range needed for the enhancement of originally invisible details at higher frequencies. As an additional improvement, the high-frequency noise resulting from the enhancement of high frequencies could be reduced by elimination of the enhancement of frequencies that are beyond the patients' acuity limits.

Results of measurements suggested that only low levels of enhancement were possible without substantial saturation (see Table 1 of Ref. 3). The reduction in saturation that can be attained by attenuation of the low frequencies is modest and may be effective only at moderate levels of enhancement.[3] Therefore it was concluded that for the enhancement to be more effective, it should be optimally tuned to a critical band of frequencies that are just undetectable by the observer. For features at these frequencies, a limited level of enhancement may be sufficient to make them visible. If such features are critical for recognition, this processing would improve performance.

The first experiment was designed to determine whether there is a band of frequencies that is critical for face recognition. The second experiment evaluated, for elderly normal observers, the effects on face recognition of enhancing a single band of frequencies. In the third experiment individual tuning of the enhancement by visually impaired observers was explored. Patients selected the best enhancement by comparing images of the same face enhanced in different bands of frequencies and at various levels of enhancement. For some of these patients the individually selected enhancement was used to process a large set of face images presented for further testing. The effect of this individually tuned enhancement on face recognition was compared with the improvement in recognition attained with uniformly applied adaptive enhancement.

EXPERIMENT 1: THE CRITICAL FREQUENCY FOR FACE RECOGNITION

Over the past decade there has been a debate in the literature about the existence of a critical spatial-frequency

0740-3232/94/071929-11$06.00

range for face recognition. Ginsburg[4] argued that the low spatial frequencies (1–4 cycles/face) were sufficient for recognition. On the other hand, Fiorentini *et al.*[5] found that recognition of face images limited to spatial frequencies below 5 cycles/face was worse than recognition of images filtered to contain only higher frequencies.

The exact nature of the recognition task has been found to be important in determining the critical frequency. Rubin and Siegel,[6] using only one pose of each face, showed that discrimination was possible even when the faces were low-pass-filtered to 1.12 cycles/face, but accurate recognition of numerous poses of the same face, including changes in expression, required information above 4.12 cycles/face. Sergent[7] similarly concluded that results of matching tasks may be equal with high or low frequencies but that an identification task benefits from the presence of high frequencies.

The need for high spatial frequencies for face recognition is indicated indirectly from a number of studies. Bandpass-filtered images judged most similar to the mental image of a recently memorized face had a center frequency of 3.8 cycles/deg.[8] This frequency typically would correspond to 12–16 cycles/face. Hübner *et al.*[9] also found that, in face images composed of the high-frequency content derived from one person and the low-frequency content from another, the high-frequency content dominated recognition. The requirement of high spatial frequencies for recognition, and the difficulties that many low-vision patients have in recognizing faces, led us to the proposal that high-pass filtering of images may improve recognition of faces.[10]

Schuchard and Rubin[11] concluded that no critical spatial frequency could be enhanced to improve face recognition for low-vision observers. Using bandpass-filtered face images, they compared face-recognition performance of bandpass-filtered images with center frequencies at 4.0, 11.3, and 32.0 cycles/face width. They found that the performance of normal observers did not depend on the center frequency of the bandpass-filtered image. In contrast, Hayes *et al.*,[12] using an image-processing paradigm similar to that of Schuchard and Rubin, found decreased performance with decreased center frequency, down to a chance level at 3.2-cycles/face width. In addition, Hayes and colleagues showed that performance depended only on spatial frequencies in terms of cycles/face, not on the retinal spatial frequencies (in cycles per degree). It should be noted, however, that Hayes *et al.* used photographic slides of their processed images for the actual testing. The nonlinear response of the film medium makes the spectral content of the images difficult to ascertain.[13]

To verify the existence of critical spatial frequencies for face recognition, which could then be enhanced to improve performance in that task, we measured the degradation in recognition of familiar faces, i.e., celebrities, by older observers. We used low-pass-filtered images rather than bandpass-filtered images, because the former more accurately represented the appearance of images to low-vision patients.[2] The testing paradigm was identical to the one that we used in evaluating the improved performance of low-vision patients with adaptively enhanced images.[1]

Methods

Subjects. Fifty adult subjects with good visual acuity in at least one eye (>20/40) were selected. Most were patients with uniocular age-related maculopathy (ARM) or were spouses of patients. The other subjects were volunteers with normal visual acuity and in the same age range. Mean age of the subjects was 50 years (range 23–82); the median age was also 50. Informed consent for participation in the study was obtained from each subject before testing.

Images. Photographs of 50 celebrities and 40 unfamiliar people were used. The celebrity photographs were expected to be familiar to most subjects. Transparencies of both sets of photographs were digitized at a resolution of 256 × 256 and at 256 gray levels. Illumination was adjusted to produce good dynamic range and clear visibility of all images. All images were digitized under the same magnification and illumination conditions. The face images were low-pass filtered with a bank of 1-octave bandpass filters[14]; when added in full magnitude, the filters summed to unity. To produce low-pass filters, the highest bands were eliminated and the remaining highest filter was set at a fraction of its magnitude. Images that were filtered the most in this set contained energy only as high as the 4-cycles/face component, decreasing to half-maximum amplitude at 6 cycles/face.[15] Other filters used included 25%, 50%, 75%, and 100% of the 8-cycles/face-height component (Figs. 1 and 2); in total, therefore, five different filters were used.

Equipment. Stimuli were generated on a MicroVAX II computer and presented with an Adage 3000 image processor with the use of 10-bit digital-to-analog converters on a U.S. Pixel monochrome, 60-Hz noninterlaced monitor. A calibrated lookup table was used to correct for the nonlinear response of the display.[13]

Procedure. The images were presented to the subject, who was sitting in a dimly lit room. The image sizes were adjusted to 4° × 4° on the display. Original (unfiltered) and filtered images were presented in random order by the computer, except that the low-pass version of each image was always presented before the original version. Subjects indicated on a scale of 1 to 6 their level of confidence in recognizing a face as a celebrity. A rating of 1 meant that the subject was positive that the face belonged to a celebrity, whereas 6 meant that the face was clearly visible but not recognizable as a celebrity; 2 indicated that the subject was quite sure but not positive that the face was a celebrity; 5 signified that the subject was quite sure but not positive that the face was not a celebrity; 3 and 4 were used when facial features were difficult to discern. A score of 3 meant that the subject had an inkling that the image was a celebrity; 4 signified that the image was not clear but was judged not to be that of a celebrity. An even number of ratings was used to reduce the tendency of subjects to select the midpoint and to force a choice in each case. If subjects could not recognize a particular celebrity from the original unfiltered image and rated it as 5 or 6, we reclassified that celebrity as a person unfamiliar to these subjects in our analysis of their responses.

We used these ratings to calculate receiver operating curves (ROC's), plotting the probabilities of true celebrity versus false celebrity. Separate curves were calculated

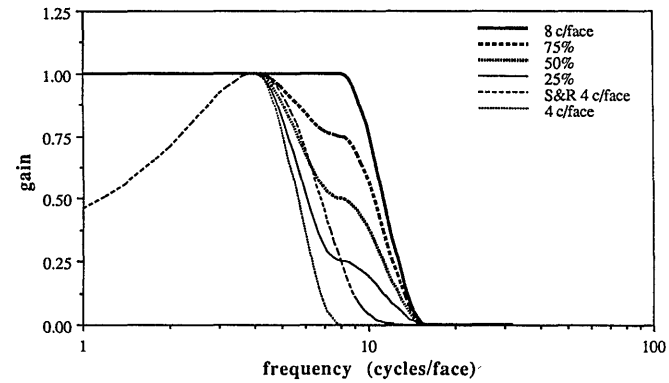


Fig. 1. Five filters used in our study compared with Schuchard and Rubin's 4-cycles/face (c/face) filter.

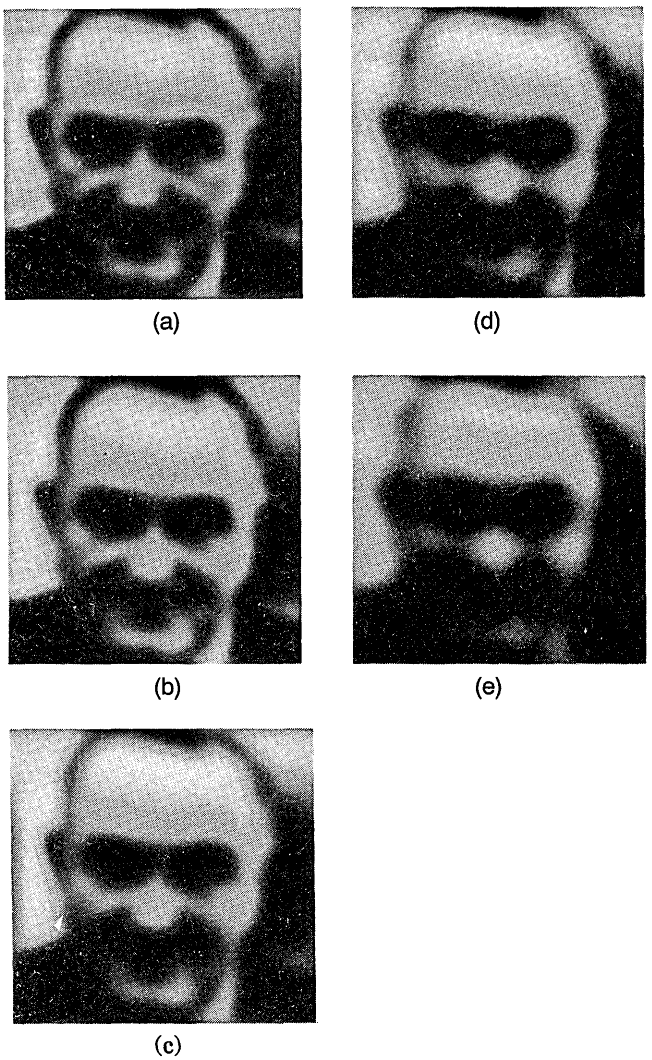


Fig. 2. Examples of the images used: (a) image containing 100% of 8 cycles/face, (b) image containing 75% of 8 cycles/face (see Fig. 1), (c) image containing 50% of 8 cycles/face, (d) image containing 25% of 8 cycles/face, (e) image containing 0% of 8 cycles/face or 100% of 4 cycles/face.

for original and filtered images. The area under the ROC (A_Z) was taken as a measure of recognition.[16] If filtration reduces face recognition, the area under the ROC for the filtered images should be less than that for the original image. Because the same faces were presented in both forms, the responses for each face were assumed to be correlated, requiring a correlated ROC analysis.[17] The level of correlation was used in determining the significance of the difference between the two areas.

Results

The recognition performance for unfiltered images varied substantially. Therefore we normalized the data by calculating the ratio of the area under the ROC obtained from the filtered images to the area under the ROC obtained from the unfiltered images. This ratio is presented in Fig. 3 as a function of the fraction of the 8-cycles/face band used in the low-pass filter. The ratio was expected to be close to 1 if filtration did not affect recognition, larger than 1 if filtration improved face recognition, and near 0.5 if filtration substantially reduced recognition performance to near chance level. Open symbols indicate cases in which the difference between the two areas under the ROC curves was significant; filled symbols indicate cases in which the difference was not significant.[17]

Analysis of variance showed that there was a significant difference in performance for the five filters [$F(4, 45) = 10.13$, $p < 0.0001$]. Using Scheffe's *post hoc* analysis, we found that performance with the four filters containing 25% or more of 8 cycles/face did not differ significantly among one another, but performance with these filters was significantly better than performance without the 8 cycles/face component (see Fig. 3). These results indicate that a band of frequencies higher than 4 cycles/face, and centered at 8 cycles/face, is critical for recognition, since even a small portion of the energy at these frequencies resulted in a significant increase in recognition performance.

It is interesting to note that for this group of subjects, despite the limited acuity range, we did find a significant correlation between acuity expressed as the log minimum angle of resolution (logMAR) and face recognition performance with the original unprocessed images ($r = -0.50$, $p = 0.0002$). We did not find this correlation

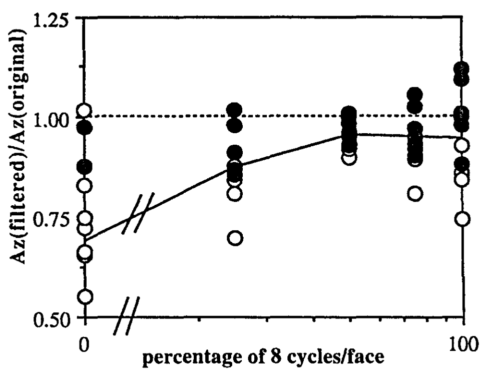


Fig. 3. Face-recognition performance as a function of the percentage of the 8-cycles/face content. The ratio of the area under the ROC for the filtered images to the area under the ROC for the original unfiltered images for each subject is represented by one data point. Filled circles, nonsignificant difference between degraded and original images; open circles, significant ($p < 0.05$) difference between degraded and original images.

in our study of low-vision patients reported below or in our previous study of low-vision patients.[1]

Discussion

Our results show that the spatial-frequency content at the 1-octave-wide band of frequencies centered at 8 cycles/face is critical for recognizing familiar faces. Although spatial-frequency content in the range of 4–6 cycles/face may provide sufficient information for face recognition above chance level in matching tasks,[11] recognition performance even in this task is virtually impossible with filtration below this level.[12] The frequencies at 8 cycles/face are critical in that they are sufficient for recognition. However, many studies have shown that they are not necessary, as observers easily recognize face images that have been filtered with band-pass filters centered at 16 cycles/face.[4,5,11,12]

We have shown that many patients with low vision that is due to central scotoma lose contrast sensitivity to the point that they cannot see at all above 8–10 cycles/deg.[1] Although most of these patients can detect gratings between 1 cycle/deg and 2 cycles/deg (corresponding to 4 cycles/face and 8 cycles/face, respectively, for our 4-deg face image), the contrast of the face content at these frequencies is usually well below the low-vision observers' thresholds. The enhancement of the band above 4 cycles/face, therefore, could render this information visible to such patients and thus aid them in face-recognition tasks.

The apparent discrepancies between our conclusions and those of Schuchard and Rubin[11,18] may be reconcilable on the basis of differences in the testing paradigm used (recognition of familiar faces versus matching from a small set of images). Further, as explained by Peli *et al.*,[1] some of the quantitative differences can be accounted for by differences in the measuring units of cycles per face. Schuchard and Rubin used the ear-to-ear measurement as the face width. We used face height, i.e., chin to the beginning of the hair line, as our unit. Thus their 4-cycles/face-width images actually contained 8-cycles/face height in our units. (For the remainder of this paper we will use cycles/face to represent cycles/face height.)

EXPERIMENT 2: THE EFFECT OF ENHANCEMENT-CAUSED DISTORTIONS ON FACE RECOGNITION

The results of experiment 1 and those of Hayes *et al.*,[12] viewed together and compared with those of Schuchard and Rubin,[11,16] suggest that spatial frequencies above 4 cycles/face are indeed critical for face recognition. Our results further demonstrate that even partial detection of this frequency range can substantially improve face recognition.

The enhancement of a single band of frequencies results in visible image distortion for a normal observer. Furthermore, since even a moderate level of enhancement results in values outside the display range, clipping of the extreme values (saturation) causes additional unintended image distortions.[3] The goal of experiment 2 was to determine whether and to what extent the distortion of the image resulting from enhancement, as well as the associated saturation, affects face recognition. We enhanced face images by amplifying a 1-octave-wide band of frequencies centered at 8 cycles/face or at 16 cycles/face and removed all energy at higher frequencies. The higher band of 16 cycles/face was included because this range can be enhanced with less saturation than can lower frequencies, and some role for these higher frequencies was suggested by the literature reviewed above in the introduction to experiment 1.

Methods

Subjects. Thirty-one adult subjects were selected, as in experiment 1. The mean age of these subjects was 61 years, with a range of 24–86; the median age was 67. Informed consent for participation in the study was obtained from each subject before testing.

Images. The same celebrity images used in experiment 1 were used here. Images were filtered to enhance the bands of 8 cycles/face or 16 cycles/face by a factor of 2 or 5 (see the central 4 faces in Fig. 8 below). These frequencies and amplifications were selected because in our previous study this was the range that we applied, using adaptive enhancement, that was shown to improve recognition.[1] We used the same bank of 1-octave band-pass filters as those used in processing the images in experiment 1. To produce the enhancement, one band (centered at 8 or 16 cycles/face) was amplified by a factor of 2 or 5, and the higher bands were eliminated (Fig. 4). Following filtration in the frequency domain by means of fast Fourier transforms, the images were transformed back to the space domain. In many cases this filtration resulted in values outside the display range (i.e., higher than 255 or negative values). The filtered images were clipped at both ends of the range rather than scaled back into that range. Rescaling would have reduced the amplitudes at all frequencies by the same amount, whereas saturation resulting from clipping distorted only the information enhanced by the filtering. Comparison of the designed (presaturation) and the actual (postsaturation) mean radial amplitude (contrast) spectra averaged from five faces is illustrated in Fig. 5. The averaged spectrum of our previously used adaptive enhancement images[1] is also shown for comparison. For two randomly selected face images from our set, we found that 3–30% of the pixels were saturated for the parameters used here.

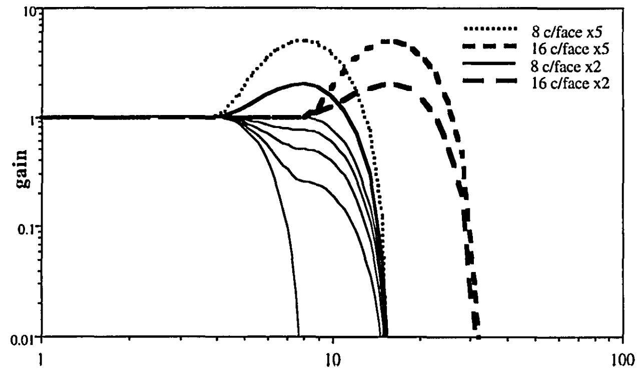


Fig. 4. Four band-enhanced filters used in experiment 2 (see legend) compared with the five low-pass filters used in experiment 1 (thin solid curves); c/face, cycles/face.

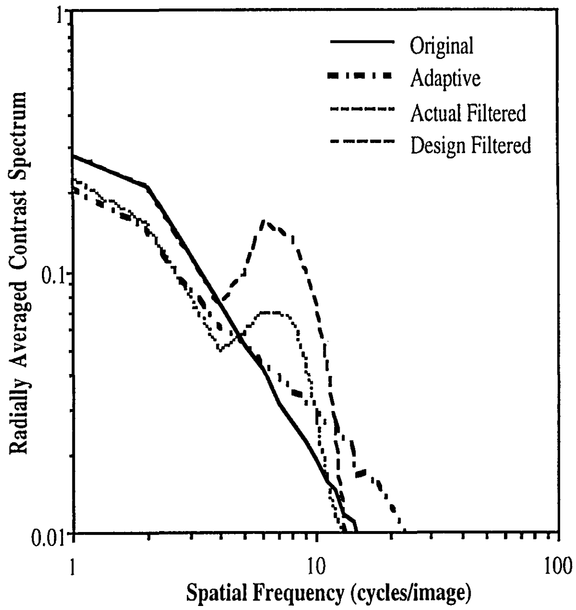


Fig. 5. Averaged contrast spectra (radially averaged amplitude spectrum normalized by mean luminance) from five faces. For the filter enhancing the 8-cycles/face band by a factor of 5 (solid curve), the spectrum of the designed filtered image (long-dashed curve) is compared with the actual spectrum resulting from saturation (short-dashed curve) and with the measured spectrum of the corresponding images enhanced by the nonlinear adaptive enhancement (dotted–dashed curve).

Procedure. The same experimental procedure and data analysis as those used in experiment 1 were applied.

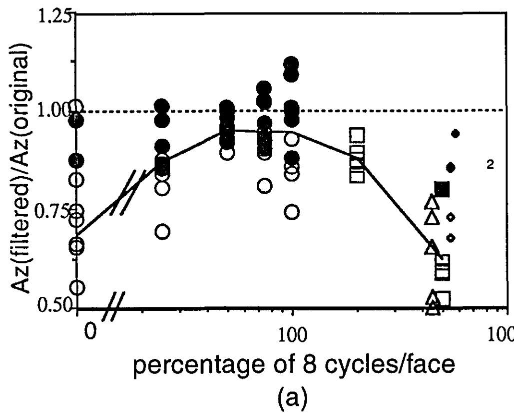
Results

The ratio of the area under the ROC obtained from the filtered images to the area under the ROC obtained from the original, unfiltered images was calculated. This ratio is presented in Fig. 6 as a function of the percentage of the 8-cycles/face and the 16-cycles/face band used in the enhancement filter. The data from this experiment are shown with the results of experiment 1.

Increasing the amplitude of the 8-cycles/face band by a factor of 2 (200%) resulted in distortion of the image, which led to a reduction in face recognition. There was also a reduction in recognition with an amplification factor of 5 (500%) for the same band. In fact, the reduction in recognition was significantly greater for the images amplified by 500% than for those amplified by 200% [$F(1,8) = 26.312$, $p < 0.001$]. With 500% amplification and the resulting distortions, most subjects performed almost at chance [Fig. 6(a)]. These results were somewhat surprising since we, the experimenters, recognized most of the enhanced (distorted) images. The subjects recognized far fewer of the celebrities from the enhanced images than from the unprocessed images. However, we were familiar with the specific set of images under a variety of processing modifications, whereas the subjects were exposed to those specific images for the first time.

As with the band at 8 cycles/face, the enhancement of the higher band (16 cycles/face) resulted in reduced recognition [Fig. 6(b)]. In addition, a 2 (frequencies) × 2 (200% or 500% amplification) analysis of variance showed a significant main effect for both frequency and amplification. There was also a significant interaction between these two variables, showing that the effect of the change in amplification was much greater for the 8-cycles/face stimuli.

Masking effects resulting from the spurious frequencies generated by saturation could result in reduced recognition. We tested for this effect in a control experiment by comparing the recognition of low-contrast images (which are easily recognized) with the recognition of the filtered low-contrast images, which result in minimal saturation. The images were first reduced in contrast (linearly rescaled) by a factor of 5. These images were then enhanced by a factor of 5, with the 1-octave filter centered at 8 cycles/face. Thus the same enhancement was applied but without the saturation (and the associated distortions) that accompanies enhancement of the full-contrast image. The enhanced images in this case correspond to rescaling rather than to the saturation clipping applied to the original images. The results [diamonds in Fig. 6(a)] illustrate that the reduction in recognition for the low-contrast, unsaturated images was less than the reduction for the saturated images [$F(1,8) = 7.9$, $p < 0.02$]. However, since there was still a substantial reduction in performance with the unsaturated filtered images (even in comparison with the low-contrast originals), saturation alone could account for only part of the reduction in recognition.



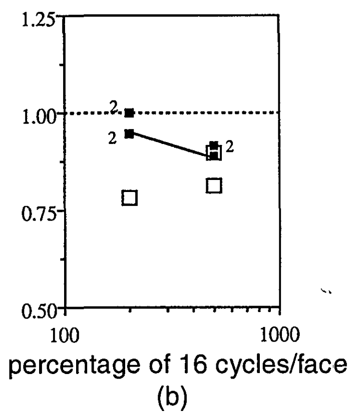


Fig. 6. Degradation in face-recognition performance as a function of the percentage of the content at (a) 8 cycles/face and (b) 16 cycles/face. The notations follow the same convention as in the caption for Fig. 3. Data from experiment 1 (circles) are compared with the data of experiment 2: squares indicate 1-octave bandwidth, triangles indicate 2-octave bandwidth, and diamonds indicate low-contrast images. The solid curves connect the mean values for the 1-octave conditions.

In a second control experiment we tested for the recognition of faces filtered with 2-octave-wide filters by enhancing the 8-cycles/face and the lower 4-cycles/face band, both by the same factor of 5. The results of this control experiment [triangles in Fig. 6(a)] indicated no significant effect for the difference in bandwidth without a change in high-frequency cutoff [$F(1, 8) < 1$, n.s.]. Therefore the 1-octave bandwidth that we used, narrower than those used in previous studies, could not account for the reduction in recognition with filtration that we found.

Discussion

Schuchard and Rubin[11,18] found that the performance of normal observers did not depend on the center frequency of the bandpass-filtered image. Hayes *et al.*,[12] using a paradigm similar to that of Schuchard and Rubin, found decreased performance with decreased center frequency, down to chance level at 3.2-cycles/face width. However, for higher filtration frequencies the data of Hayes *et al.* did not show the substantial decrement of recognition performance that we found. Our bandpass-filtered images differed from those used in the two other studies[11,12,18] in three ways, which may account for the differences between our results and theirs. First, our images contained low frequencies, whereas the other studies used images in which both higher and lower frequencies were removed (Fig. 1); second, our images had significant distortions that were due to saturation; and third, our filter's bandwidth was narrower than the bandwidths used in the other studies.

Our images retained all the energy at frequencies lower than those of the enhanced band. This is unlikely to cause a reduction in recognition, as it adds relevant information. Furthermore, similar levels of low-frequency information were available in our adaptively enhanced images (see Fig. 5) that were found to improve recognition.[1]

The images in the other studies[11,12] were rescaled after filtration and therefore did not suffer from saturation-related distortions. The distortion resulting from saturation was found here to contribute to the reduction in recognition. However, despite the relative improvement, the recognition of the filtered low-contrast images was still substantially reduced, indicating that the saturation could not account for the whole effect.

Finally, our enhancement filter's bandwidth was narrower (1 octave versus 1.5–2 octaves in other studies), resulting in more ringing artifacts, which may have contributed to the distortion and the reduced recognition that we found. Visual inspection of such images shows that the appearance of the 2-octave filtered images is usually more natural than that of 1-octave images. The results of the second control experiment, however, indicated no significant effect of the difference in bandwidth.

All these factors combined cannot account for the decrement in performance that we found with the 16-cycles/face enhancement compared with the findings of previous studies.[11,12] This difference may be related to the recognition task that we used. We employed the celebrity-recognition paradigm, which has been shown to differ in visual requirements from discrimination-type tasks.[19]

Since image enhancement in the range of 8 cycles/face and up improves recognition in visually impaired patients,[1] a similar enhancement was not expected to reduce recognition by normal observers. The adaptively enhanced images that were found useful in our previous study are visually, and by spectral analysis, similar to images filtered with a 2-octave filter centered at a frequency of 16 cycles/face and with an amplification factor of 5. Yet the performance of normal observers with these images was significantly reduced. This apparent contradiction between the effects of enhancement on the performance of normal observers compared with the performance of people with low vision may be accounted for by the normalization that we used to compare performance. While normal observers' performance with the enhanced images was reduced from its excellent level with the original images, the observers' absolute level of recognition was similar to that attained by patients with the same images. The patients' performance with the unenhanced images was so poor that the level of recognition afforded by the enhanced images constituted an improvement in performance for them.[1] For the normal observers the area under the ROC curves for the original images was very high (average $A_Z = 0.919$). Their performance with the enhanced distorted images (1-octave, 16-cycles/face amplified by 5) was similar to the performance of the low-vision patients with the adaptively enhanced images (average $A_Z = 0.813 \pm 0.08$ and average $A_Z = 0.857 \pm 0.09$, respectively). Since the level of distortion resulting from saturation increases with reduced spatial frequency of the filter, the enhancement at lower frequencies may be severely restricted by the available dynamic range of the display.

EXPERIMENT 3: INDIVIDUALLY TUNED ENHANCEMENT

Individually tuned enhancement may depend both on the patient's contrast-sensitivity function (CSF) and on the spatial-frequency content of the image. We postulated that the enhancement should be tuned to the critical band of frequencies that are just undetectable by the patient,[1] since for features at these frequencies a limited level of enhancement may be sufficient to make them visible and thus improve recognition. For the third experiment we designed a method that allows the patient to choose the kind of enhancement that he or she prefers for improving face recognition. Subjects with visual impairments were presented with a matrix of images processed to enhance different ranges of spatial frequencies at different levels. After reviewing all the available options, the subjects selected the image that appeared to provide the most distinct, visible, and recognizable face. For some of these subjects a second part of the experiment applied the individually selected enhancement to the full set of face images. These images were used to test the effect on recognition of the individually tuned enhancement in comparison with our previously uniformly applied adaptive enhancement.[1]

Methods

Subjects. In the first part of this experiment we tested a total of 93 patients. Of these, 48 had documented central scotoma in the tested eye resulting from ARM (32 patients), diabetic retinopathy (6), central retinal vein occlusion (CRV) (3), and other conditions (7). Seventeen patients had visual impairment with intact central field resulting from ARM (four patients), diabetic retinopathy (two), CRV (one), detached retina (two), and other conditions (six). The remainder were not clearly documented with regard to their central field status; their impairments were due to ARM (5), diabetic retinopathy (7), CRV (3), macular holes (2), and other conditions (6). The mean age of the subjects was 67 years (range 13–89). Average logMAR was 1.105 (range 0.544–1.301). There was no difference in acuity among the scotoma groups. Most of these patients were tested binocularly; each used, in effect, the better eye. However, those with vision better than 20/70 in the better eye were tested with the better eye covered. For the second part of the experiment, 18 of the patients were available. Their mean visual acuity was 0.791 (range 0.544–1.079). Fourteen of the eighteen had documented central scotoma that was due to ARM (nine), diabetic retinopathy (two), CRV (one), and other conditions (two). Two of these patients had intact central fields. Informed consent for participation in the study was obtained from each subject before testing.

Images. The same face images that were used in experiments 1 and 2 were used here. Ten additional face images from the same set were included in the first part of the experiment. The ten images were filtered to enhance a 1- or 2-octave-wide band by a given factor. The face images were filtered with a bank of 1-octave bandpass filters; when added in full magnitude, the filters summed to unity. To obtain the enhancement, we amplified the appropriate band by a factor of 1, 2, 5, or 15 and eliminated higher bands. Following filtration in the frequency domain by means of fast Fourier transform, the images were transformed back to the space domain. Clipping of out-of-range pixel values was applied, as in experiment 2.

Two 4×4 grids of enhanced images were created. Each square on the grid contained a different enhancement of the face. The spatial frequencies enhanced and the level of enhancement for both grids are presented in Fig. 7. In one grid [Fig. 7(a)] a 1-octave band of frequencies centered at 4, 8, 16, or 24 cycles/face was enhanced by factors of 1, 2, 5, or 15 (a sample image set is presented in Fig. 8). For the other grid [Fig. 7(b)], 2-octave-wide bands were enhanced by the same factors (images are presented in Fig. 9). At any time during the selection process, only one image was visible. The patient could display every image by moving a bit pad into the corresponding place on the grid and choose the preferred image by pressing a bit-pad button.

Procedure. The images were presented on a 60-Hz, noninterlaced video monitor to the subject, who was sitting in a dimly lit room. The image sizes were adjusted to $4° \times 4°$ on the display. In the first part of the experiment the patient selected the one enhanced image that appeared to be the most distinct, visible, and recognizable face. Selection was made after free exploration of all 32 images (two of which were the original, unprocessed image). Patients were required to view each of the 32 images before making a selection. The selection process was repeated for 10 different faces. We then calculated the mean frequency and level of enhancement selected.

The second part of the experiment tested patients' face-recognition performance by using their individually selected enhancements. Only 18 of the subjects were available for the length of time required for this part of the experiment. Using the mean frequency and level of enhancement selected by each patient for the 10 face images that were presented in the first part, we enhanced 50 celebrity and 40 noncelebrity face images with the corresponding filter. The enhanced images were then presented to the patient for testing of celebrity face recognition with use of the same paradigm as in experiments 1 and 2. Two comparisons were made: recognition with individually tuned images versus the original, unenhanced images (11 patients) and recognition with individually tuned images versus our previously used,[1] adaptively enhanced images (7 patients).

Results

In testing of the first 56 patients, almost all selected the 2-octave-wide filter over the 1-octave-wide filter (1-octave-filtered faces had more distortions because of the ringing associated with sharp filtering). We therefore discontinued the presentation of the 1-octave images for the rest of the study.

Relatively high spatial frequencies and low levels of enhancement were selected. The 48 patients with central scotoma selected faces filtered with a mean center frequency of 17.36 ± 5.0 cycles/face and an amplification factor of 2.4 ± 1.5. The 17 patients with intact central fields selected a mean frequency of 15.88 ± 4.0 cycles/face and

1-octave

original	24 **2**	24 **5**	24 **15**
16 **1**	16 **2**	16 **5**	16 **15**
8 **1**	8 **2**	8 **5**	8 **15**
4 **1**	4 **2**	4 **5**	4 **15**

(a)

2-octave

original	16+32 **2**	16+32 **5**	16+32 **15**
8+16 **1**	8+16 **2**	8+16 **5**	8+16 **15**
4+8 **1**	4+8 **2**	4+8 **5**	4+8 **15**
2+4 **1**	2+4 **2**	2+4 **5**	2+4 **15**

(b)

Fig. 7. Virtual grids of processed images presented for selection: (a) 1-octave, (b) 2-octave. In each square the top numbers represent the center spatial frequency (in cycles/face) of the bands that were enhanced and the lower numbers (bold) the amplification applied to that band.

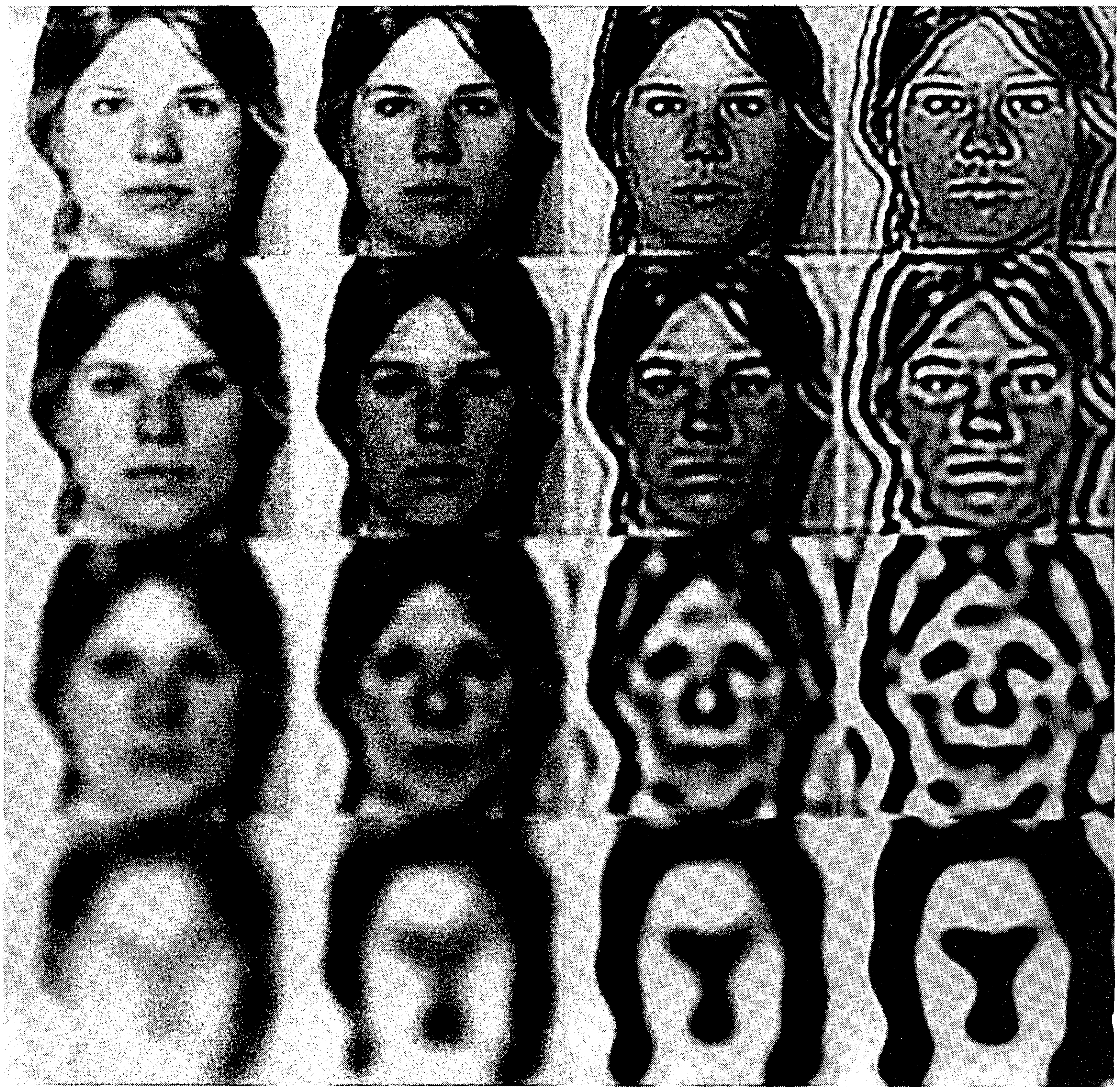


Fig. 8. Examples of the face images enhanced with a 1-octave-wide filter. The filter-center frequencies and the amplification factors used are illustrated in Fig. 7(a). The central four images are the same as those used in experiment 2.

an amplification of 1.9 ± 1.0. Those with undocumented central field status ($n = 28$) chose a mean center frequency of 15.96 ± 5.8 and an amplification of 2.4 ± 1.5. The differences among these groups were not significant. For all subjects there was a significant correlation between acuity (logMAR) and the selected frequency ($r = -0.30$, $p = 0.0043$) and between the frequency and the amplification selected ($r = -0.50$, $p < 0.0001$).

For the subjects tested in the second part, we normalized the data as in experiments 1 and 2. We calculated the ratio of the area under the ROC obtained from the individually tuned filtered images and the area under the ROC obtained from the original and or the area under the ROC obtained from the adaptively enhanced images. This ratio is presented in Fig. 10. The ratio was expected to be close to 1 if individual filtration did not affect recognition, to be larger than 1 if individual enhancement improved recognition, and to be less than 1 if it reduced recognition in comparison with the other condition.

As in our previous study and in other studies, recognition of unprocessed face images was not significantly correlated with acuity. Six of the eleven patients did not improve their recognition with the individually tuned enhancement, because they had good recognition with the original images and thus no real room for improvement. The enhancement with the individually chosen filters improved recognition for only three of the remaining five patients. The level of improvement was similar to the effect found in our previous study,[1] although the difference in recognition in the two conditions was not statistically significant for any of the subjects. We did find statistically significant improvement with adaptive enhancement for almost half of the subjects in our previous study.[1]

Performance with individual enhancement did not differ from performance obtained with adaptive enhancement. This finding might be expected, since many patients selected enhancement that resulted in images similar in appearance to those obtained by the adap-

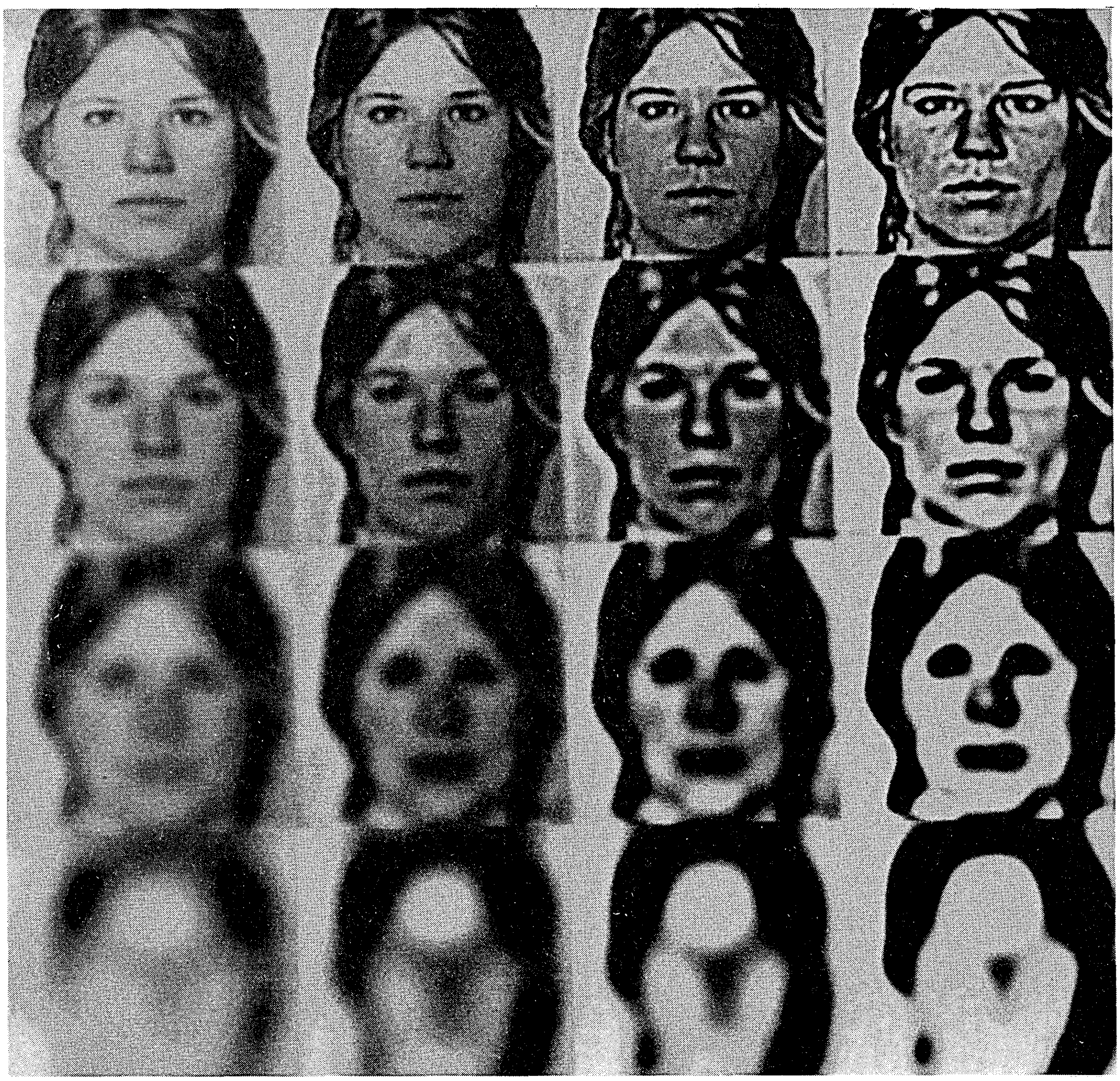


Fig. 9. Examples of the face images enhanced with a 2-octave-wide filter. The filter-center frequencies and the amplification factors used are illustrated in Fig. 7(b).

tive enhancement. For both groups tested in the second part of the experiment, performance with the individually selected enhancement was strongly correlated with the frequency selected ($r = 0.835$, $p = 0.016$ and $r = 0.618$, $p = 0.041$, for the two groups). This correlation indicates that patients who select higher frequencies (i.e., who have better acuity) tend to perform better with the individually tuned enhancement than do those who select lower frequencies.

Discussion

These results may appear puzzling at first. The frequencies selected for enhancement (15.9–17.4 cycles/face) were higher by at least 1 octave than those that we anticipated on the basis of our and others' studies of face recognition. We found that the band of frequencies at 8 cycles/face is critical for face recognition (experiment 1), whereas the removal of higher frequencies has little effect on face recognition.[11,12,16] Furthermore, the patients' sensitivity at the selected range of frequencies (16 cycles/face corresponding to 4 cycles/deg in our images) is very low.[1]

We have noted that the adaptively enhanced images that were useful in our previous study[1] appear to be similar to images filtered with a 2-octave filter centered at a frequency of 16 cycles/face and with an amplification factor of 5. These images also have very similar amplitude spectra (Fig. 11). This similarity may account for the lack of improvement in recognition with the individually tuned enhancement compared with the effect of the adaptive enhancement. Further, it suggests that the simulations that we used in tuning the parameters of the adaptive enhancement[1] indeed led us to the enhancement that is close to optimal for most of our patient population.

We may account for the apparent discrepancies with previous results, including those of our own experiment 1, which suggest that only lower frequencies are important for face recognition, by noting that because of our use of

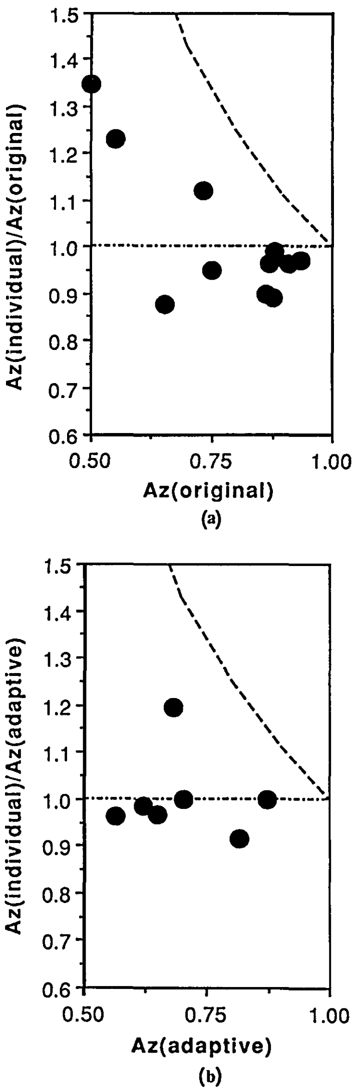


Fig. 10. Results of celebrity-recognition experiments (second part of experiment 3). The improvement ratio with the individually tuned enhancement as a function of face-recognition performance is shown for the two comparison conditions: (a) individual enhancement with the original unenhanced image, (b) individual enhancement with the adaptive enhancement.

2-octave-wide filters the selected images were enhanced in the band of 8 cycles/face, although by a lesser amount than those enhanced by filters at 8 cycles/face (see Fig. 11). This enhancement may have been sufficient to boost the critical facial information above patients' detection level. The enhancement of the higher frequencies was still within the range detected by most patients of this population, as was previously shown.[1] Although the enhancement of these higher frequencies may add only marginally to face-recognition performance, the visibility of such high-frequency detail may add to the esthetic appearance of our otherwise quite distorted images. Preference, as determined by esthetic appeal, may be important in our application even if it does not indicate improvement in performance.

GENERAL DISCUSSION

The results of this series of experiments clarify many of the questions that were raised by our previous report of improved face recognition with image enhancement.[1] We have shown here that there is a range of critical frequencies that, when lost, make face recognition difficult or impossible. The existence of such critical frequencies suggests that enhancement of these frequencies may aid patients in face recognition. Such improvement is possible if the enhancement renders the critical frequencies visible to the patient. Unfortunately, such enhancement of a band of frequencies results in image distortion for normal observers, with and without the effects of saturation. Although the distortion associated with enhancement reduced the recognizability of the face images (as shown in experiment 2), the reduction in performance is not so large as the performance degradation that is due to low vision. Thus, if the enhancement renders the image more visible, it leaves room for improvement in recognition. Such improvement has been demonstrated previously.[1]

While we and many others in the field addressed image filtration in terms of object spatial frequency (i.e., cycles per face), Lawton[20] insisted that only retinal spatial frequencies (i.e., cycles per degree) should be considered. Peli's[3] modified model proposing the enhancement of the just-invisible band of frequencies requires consideration of the interaction of both. Whereas the critical frequencies for image recognition are determined in terms of cycles per image, the just-invisible band in a particular image will depend on both patient CSF and the image spectrum and size.

If we assume that most images on television, for example, have a fairly similar spectral content, then the frequencies that can be effectively enhanced depend mostly on patients' CSF's. As the scene changes, resulting in a change in the displayed size of objects, the enhancement may become more or less effective. For example, if, for a specific patient, the enhancement of a scene containing two faces significantly increases recognition, the same

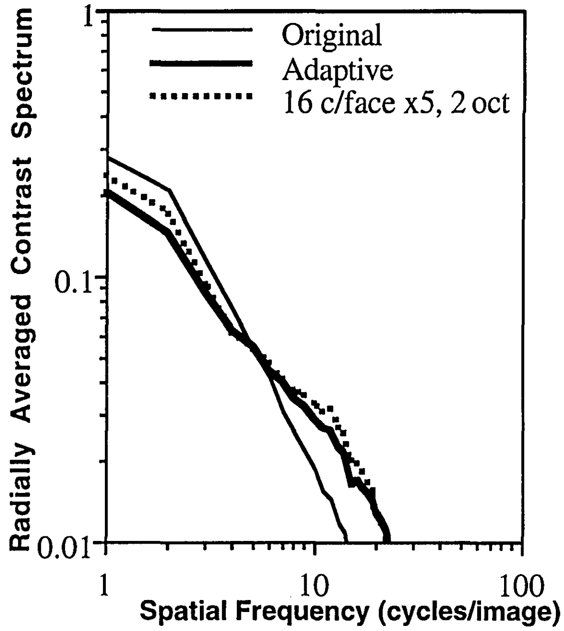


Fig. 11. Radially averaged contrast spectrum of unenhanced face images compared with the spectra of the images enhanced with the adaptive enhancement and with a 2-octave-wide filter centered at 16 cycles/face and amplified by a factor of 5: c/face, cycles/face.

enhancement may be less effective when only one face is displayed on the screen, simply because that larger face may be recognizable even without the enhancement. Reducing the magnification to display eight faces on the same screen may cause the faces to be unrecognizable even with the enhancement. A change in the enhancement to correct for this may not be possible for this specific patient if the critical frequencies (in cycles per face) are now at a range of completely invisible frequencies (in terms of cycles per degree). The enhancement, nevertheless, may be helpful in providing the patients with useful information regarding the relative position and movements of the eight characters whose faces are no longer recognizable.

We have previously postulated that individually tuned filtration may provide an advantage over a uniformly applied enhancement.[1-3] We have also argued that the benefit of individual enhancement should be substantial to justify the complexity and expense of this approach. Lawton,[20,21] in her study of enhancement of text using filters based on the patient's CSF, has insisted that the individual tuning of the filters was critical. Using calculations of the filters from CSF's and formulas published by Lawton,[21] Engel[22] claims that, although the CSF's of Lawton's three patients differed substantially, the filters that she designed greatly reduced the difference in the enhancement filters that were applied for each patient.

We have failed to demonstrate substantial improvement in face recognition with the use of individually tuned enhancement compared with that from uniformly applied adaptive enhancement. However, this failure may have stemmed from the fact that, for many of the patients in the group tested in the second part of experiment 3, the individually selected enhancement was essentially the same as the applied adaptive enhancement. Since in both studies (ours and Lawton's[21]) the applied enhancement differed only slightly between patients, few conclusions can be drawn regarding the effect of individual tuning. For other patients, and even for the same patients given different images, it is still possible that some improvement with individual enhancement may be shown. However, we are no longer confident that the benefit will be substantial.

ACKNOWLEDGMENTS

This research was supported in part by National Eye Institute grant #EY05957 and by grants from the Teubert Charitable Foundation, the Ford Motor Company Fund, and DigiVision Inc. We thank George M. Young, Jenny Yu, and Jill Strauss for valuable help in data collection; Robert B. Goldstein for important programming help; and Charles Metz for the use of his software. Special thanks to Elisabeth M. Fine for many helpful comments and suggestions that greatly improved the quality of this paper.

REFERENCES AND NOTES

1. E. Peli, R. B. Goldstein, G. M. Young, C. L. Trempe, and S. M. Buzney, "Image enhancement for the visually impaired: simulations and experimental results," Invest. Ophthalmol. Vis. Sci. **32,** 2337–2350 (1991).
2. E. Peli and T. Peli, "Image enhancement for the visually impaired," Opt. Eng. **23,** 47–51 (1984).
3. E. Peli, "Limitations of image enhancement for the visually impaired," J. Opt. Vis. Sci. **69,** 15–24 (1992).
4. A. P. Ginsburg, "Specifying relevant spatial information for image evaluation and display design: an explanation of how we see certain objects," Proc. Soc. Inf. Disp. **21,** 219–227 (1980).
5. A. Fiorentini, L. Maffei, and G. Sandini, "The role of high spatial frequencies in face perception," Perception **12,** 195–201 (1983).
6. G. S. Rubin and K. Siegel, "Recognition of low pass filtered faces and letters," Invest. Ophthalmol. Vis. Sci. Suppl. **25,** 28 (1984).
7. J. Sergent, "Microgenesis of face perception," in *Aspects of Face Processing,* H. D. Ellis, M. A. Jeeves, R. Newcombe, and A. Young, eds. (Nijhoff, Boston, Mass., 1986), pp. 17–33.
8. L. O. Harvey and G. P. Sinclair, "On the quality of visual imagery," Invest, Ophthalmol. Vis. Sci. Suppl. **26,** 281 (1985).
9. M. Hübner, I. Rentschler, and W. Encke, "Hidden face recognition: comparing foveal and extrafoveal performance," Human Neurobiol. **4,** 1–7 (1985).
10. E. Peli, R. Goldstein, C. Trempe, and L. Arend, "Image enhancement improves face recognition," in *Noninvasive Assessment of the Visual System,* Vol. 7 of 1989 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1989), pp. 64–67.
11. R. A. Schuchard and G. S. Rubin, "Face identification of band pass filtered faces by low vision observers," Invest. Ophthalmol. Vis. Sci. Suppl. **30,** 396 (1989).
12. T. Hayes, M. C. Morrone, and D. C. Burr, "Recognition of positive and negative band pass-filtered images," Perception **15,** 595–602 (1986).
13. E. Peli, "Display nonlinearity in digital image processing for visual communication," Opt. Eng. **31,** 2374–2382 (1992).
14. E. Peli, "Contrast in complex images," J. Opt. Soc. Am. A **7,** 2032–2040 (1990).
15. Note that the high-frequency half-maximum point of the 1.5-octave Gaussian filter centered at the 4 cycles/face that were used by Schuchard and Rubin is at 6 cycles/face.[11,18]
16. J. A. Swets and R. M. Pickett, *Evaluation of Diagnostic Systems: Methods from Signal Detection Theory* (Academic, New York, 1982), pp. 15–45.
17. C. E. Metz, P. Wang, and H. B. Kronman, "A new approach for testing the significance of differences between ROC curves measured from correlated data," in *Proceedings of the Eighth Conference on Information Processing in Medical Imaging,* F. Deconinck, ed. (Nijhoff, The Hague, 1983), pp. 431–445.
18. R. A. Schuchard and G. S. Rubin, "Effect of band width on discrimination and recognition of band pass filtered faces," in *Annual Meeting,* Vol. 18 of OSA Proceedings (Optical Society of America, Washington, D.C., 1989), p. 161.
19. V. Bruce, "Recognizing familiar faces," in *Aspects of Face Processing,* H. D. Ellis, M. A. Jeeves, F. Newcombe, and A. Young, eds. (Nijhoff, Boston, Mass., 1986), pp. 107–117.
20. T. B. Lawton, "Image enhancement filters significantly improve reading performance for low vision observers," Ophthalmol. Physiol. Opt. **12,** 193–200 (1992).
21. T. B. Lawton, "Improved reading performance using individualized compensation filters for observers with losses in central vision," Ophthalmology **96,** 115–126 (1989).
22. F. L. Engel, Institute for Perception Research, Eindhoven University of Technology, The Netherlands (personal communication, 1993).