Sharpness Enhancement through Spatial Frequency Decomposition

Samira Bouzit and Lindsay MacDonald
Colour & Imaging Institute, University of Derby, United Kingdom

Abstract

An experimental study of image sharpness was conducted through adjustment of the image power spectrum. Each test image was decomposed into a set of spatial frequency bands, defined as octaves of the pixel sampling frequency. The Fourier power spectrum was derived, then amplitudes of selected bands were adjusted to enhance the desired spatial frequencies. A psychophysical experiment was performed to evaluate the sharpness of parameterised variants of the images displayed on a CRT television screen. The experimental results indicated that sharpness was perceived to be increased when certain spatial frequency bands were enhanced. Results were related to the standard observer's contrast sensitivity function (CSF).

1. Introduction

Sharpness is known to be one of the important factors relating to the perceived quality of reproduced images. Inoue and Tajima[1] derived an adaptive image sharpening method which estimated edge sharpness by a high-pass filter based on a difference of Gaussian functions. Interest in image sharpness has also grown because of its importance in cross-media reproduction. Hultgren[2] described an image processing system for automatic enhancement of images from various sources to be displayed on various destination media. MacDonald[3] proposed a framework based on the human visual contrast sensitivity function (CSF) and Modulation Transfer Function (MTF) of the input and output devices to determine an optimum correction to be made to the sharpness of an image. Recently Garrett and Fairchild[4] determined several rules for sharpness related to other image attributes such as contrast, noise and resolution.

The study described in this paper investigated the contribution of sharpness to overall image quality in colour reproduction, using a high quality television display under controlled viewing conditions. It followed on from our previous work[5] where we showed a good correlation between perceived sharpness and image quality metrics. Specifically we studied the relationship between adjustment of the power spectrum of an image, perceived sharpness, and contrast sensitivity as a function of viewing distance.

2. Image Spatial Frequency

2.1 Separating Frequency Bands

Enhanced versions of an original colour image were generated using the processing scheme shown in Figure 1. Because the luminance channel carries the majority of the sharpness information only the luminance component of the image was processed. The original *RGB* colour image was converted into the YC_bC_r colour space used in broadcast television, where Y is the luma component, and C_b, C_r are blue and red chroma components respectively. The latter were recombined with the enhanced luma and then converted back to *RGB* to produce an enhanced image.

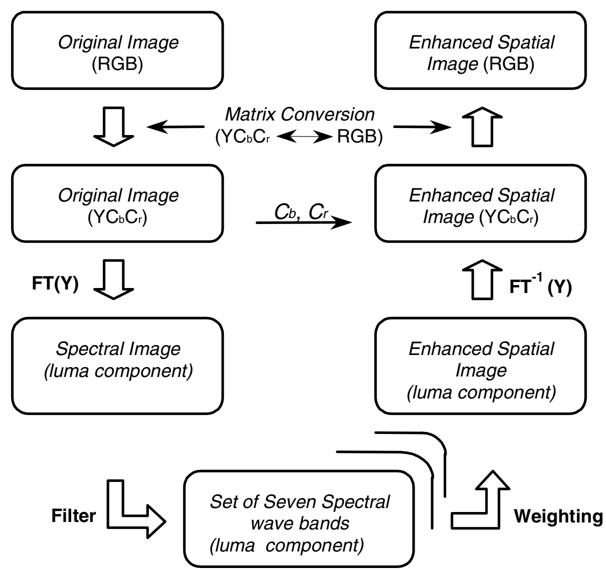

Figure 1. Image processing operations

The luma component was converted to the frequency domain by applying a Fourier Transform and then decomposed into a set of seven spatial frequency bands. Because the test images were 256x256 pixels in size, the Nyquist limit was taken to be 128 cycles per width, i.e. 2 pixels per cycle. Octave bands were then determined by successively halving the number of cycles per image width, i.e. halving the spatial frequency, as shown in Table 1. Band 7 contains all the low frequency components, including the DC (constant) term.

Table 1. Spatial Frequency Bands In Image Decomposition

Band	Cycles/width	Pixels/cycle
1	64 – 128	2 – 4
2	32 – 64	4 – 8
3	16 – 32	8 – 16
4	8 – 16	16 – 32
5	4 – 8	32 – 64
6	2 – 4	64 – 128
7	< 2	> 128

The power spectrum of each spatial frequency band was extracted by applying an annular filter to the spectral plane, so that all phase information was preserved, as shown in Figure 2. The amplitudes of a group of three adjacent bands were then adjusted using weighting parameters to enhance specific frequencies. The inverse Fourier Transform was applied to obtain a luma image in which the selected spatial frequencies were enhanced. The enhanced colour image was reconstructed by combining the original chroma components, and finally the image was converted to RGB signals for display.

2.2 Power Spectrum Weighting

Suppose that Y_0 is the luma component of the original image; and that F_0 its Fourier Transform is given by:

$$F_0 = FT(Y_0) \tag{1}$$

which can be expressed by:

$$F_0 = \sum_{i=1}^{7} P_i \tag{2}$$

where Pi is the power spectrum (amplitude plus phase) for one band, $i = 1..7$.

The four enhanced images were generated by weighting the amplitude of three adjacent wave bands:

$$\begin{aligned}
F_2 &= \alpha(k_1P_1 + k_2P_2 + k_1P_3 + P_4 + P_5 + P_6 + \frac{1}{\alpha}P_7)\\
F_3 &= \alpha(P_1 + k_1P_2 + k_2P_3 + k_1P_4 + P_5 + P_6 + \frac{1}{\alpha}P_7)\\
F_4 &= \alpha(P_1 + P_2 + k_1P_3 + k_2P_4 + k_1P_5 + P_6 + \frac{1}{\alpha}P)\\
F_5 &= \alpha(P_1 + P_2 + P_3 + k_1P_4 + k_2P_5 + k_1P_6 + \frac{1}{\alpha}P_7)
\end{aligned} \tag{3}$$

where:

F_j represents the power spectrum of the enhanced image;
$j = 2..5$ corresponds to the wavelength of the central band; k_1 and k_2 are the weighting coefficients;
α is a normalising factor, given by $\alpha = 7/(2k_1+k_2+4)$.

The lowest frequency components (Band 7) were left unchanged by the weighting procedure. The adjusted luma components were obtained by performing the inverse Fourier Transform on the weighted sums:

$$Y_j = FT^{-1}(F_j) \tag{4}$$

The enhanced image was reconstructed by recombination of the enhanced luma component with the original chroma components:

$$I_j = Y_j \oplus C_bC_r \tag{5}$$

where:

Y_j represents the enhanced luma component;
C_b, C_r represent the blue and red chrominance components;
I_j represents the enhanced colour image;
j = 2,3,4,5 corresponds to weighting applied.

In this study we used two sets of weighting factors (k_1=1.25, k_2 =1.5) and (k_1 =1.5, k_2 =2). Two corresponding sets of enhanced images were thus generated.

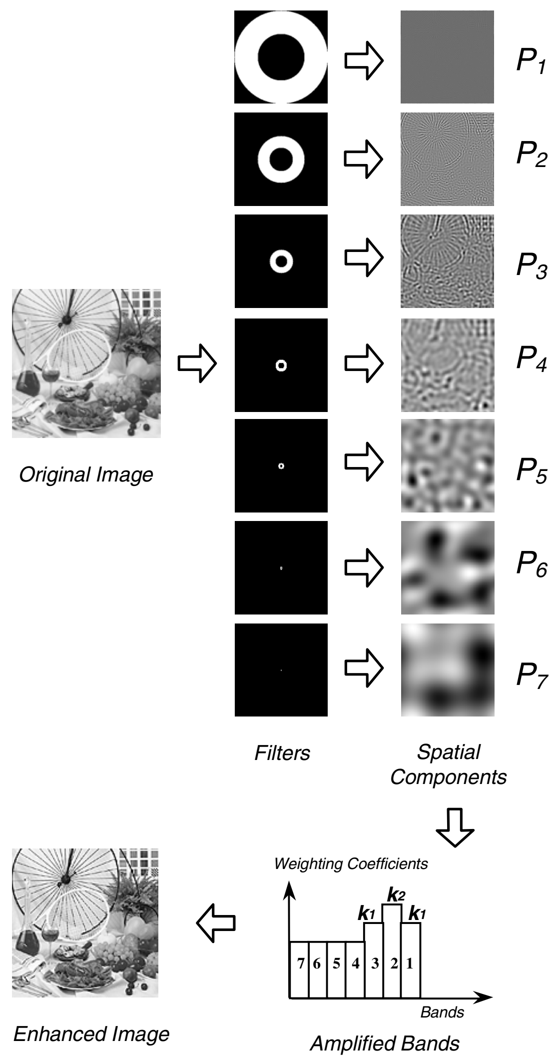

Figure 2. Ensemble of filters and weighting functions.

3. Experimental Design

3.1 Visual Contrast Sensitivity

The eye's ability to discriminate the contrast of a sinusoidal pattern at different spatial frequencies is described by the contrast sensitivity function (CSF)[6]. The shape of the CSF depends upon a number of factors, including image size (angle subtended at the eye), retinal illuminance and pattern orientation. The contrast sensitivity function is defined as the inverse of the minimum contrast threshold that an observer needs to detect a particular spatial frequency. Barten[7] developed a mathematical formula as an approximation of the contrast sensitivity function of the normal human visual system, which depends on the luminance level and angular display size:

$$\begin{aligned} CSF(u) &= au\exp(-bu)\sqrt{1+c\exp(bu)} \\ a &= 540(10+0.7/L)^{-0.2} \Big/ \left(1+\frac{12}{w(1+u/3)^2}\right) \\ b &= 0.3(1+100/L)^{0.15} \\ c &= 0.06 \end{aligned} \tag{6}$$

where:

- L = Average display luminance in cd/m^2,
- w = Angular display size in degrees,
- u = Spatial frequency of the pattern at the observer's eye in cycles/degree.

The parameter a controls the low-frequency behaviour of the contrast sensitivity function, and b and c control the high frequency behaviour. The parameter b is related to the visual acuity of the observer and is a function of the display luminance.

3.2 Experimental Setup

Observers were asked to judge a displayed test image of 256×256 pixels in size at four different distances. The test image dimensions on the display screen were 186 (H) x 160 (V) mm, corresponding to pixel dimensions of 0.73 x 0.63 mm. The horizontal angle subtended at the eye of the observer (in degrees) was calculated from the viewing geometry for the given pixel or image width D and viewing distance l:

$$w_{pixel} = \frac{D}{l}\times\frac{180}{\pi} \tag{7}$$

Table 2. Horizontal angular subtense and Nyquist limit of a test image viewed at different distances

Viewing distance (cm)	60	120	240	480
Angle subtended by one pixel	4.2'	2.1'	1.0'	0.05'
Angle subtended by image	17.8º	8.9º	4.4º	2.2º
Nyquist limit = 2 pixels (deg)	0.14	0.07	0.035	0.017
Nyquist limit (cycles/degree)	7	14	28	56

The values of horizontal angular subtense corresponding to the four viewing distances are summarised in Table 2. The horizontal angular spatial frequency at the observer position corresponding to the Nyquist sampling limit of the image is calculated from the dimension of 2 pixels.

Figure 3 shows the CSF calculated from Eq. (6) at different viewing distances. It can be seen that the peak sensitivity at 60 cm falls to about half at 480 cm, while the corresponding angular frequency increases from 3.5 cpd to 5 cpd. The width of the curves, at 50% of peak, is a spatial frequency ratio of about 8, i.e. three octaves.

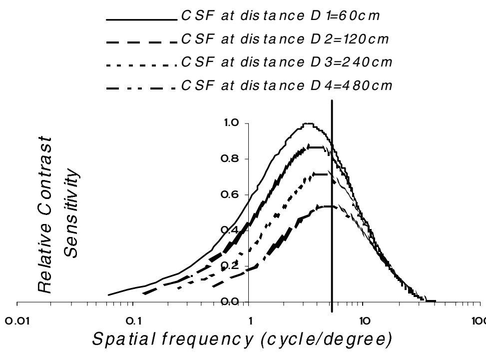

Figure 3. Calculated spatial frequency sensitivity function

3.3 Hypothesis

Preferred image sharpness should be related to observer contrast sensitivity. By asking observers to judge variants of images in which different spatial frequency bands have been enhanced, we expect to find that overall perceived image sharpness is greatest for images in which spatial frequencies have been enhanced which are close to the peak of the standard observer CSF. By repeating the experiment at different viewing distances, moreover, we expect to find a corresponding shift in the spatial frequency dependence.

Figure 4 shows how the seven spatial frequency bands described in Sec. 2 correspond to the normalised observer CSF at each of the four viewing distances. Band 1, the highest spatial frequency occupies the octave below the Nyquist limit. All bands move right one octave (double spatial frequency) as the viewing distance is doubled. The peak observer CSF is expected to correspond to Band 2 at 60 cm and to Band 4 at 480 cm, a difference of two (not three) octaves because of the shift to the right of the CSF peak shown in Fig. 3.

3.4 Test Images

Four different test images were selected from the standard SCID image set[8], as shown in Figure 5. The first scene with flowers in a transparent glass vase has sharp contrast of the foreground against both a white dish and a dark defocused background. The second contains fine details of a bicycle wheel, sine patterns and different shapes with high chroma colours. The third contains metallic objects against a plain grey background and has many highlight details and curved edge gradients. The fourth shows a young woman with clear skin and contains shadow detail (in the hair) and gentle tone and colour gradations in the skin. Each image was adjusted by two different degrees with ($k_1=1.25$, $k_2=1.5$) and ($k_1=1.5$, $k_2=2$), i.e. the second had double the degree of enhancement of the selected three spatial frequency components relative to the first. Four enhanced versions of each individual image (I_2 .. I_5) were generated.

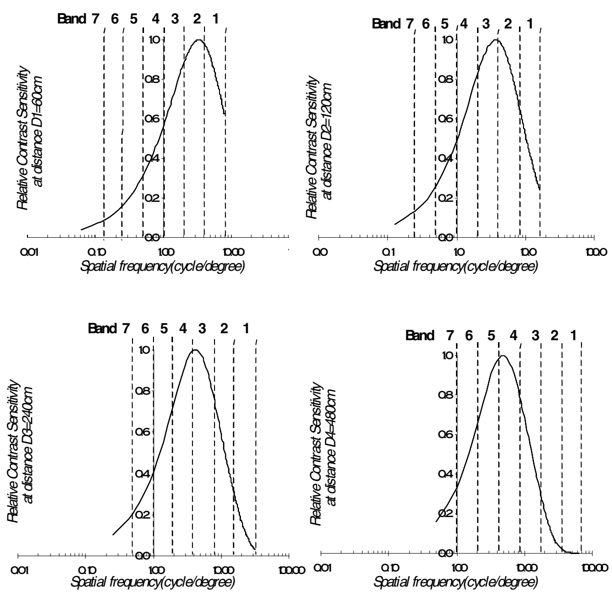

Figure 4. Relative contrast sensitivity at different distances and wave bands

3.5 Observers, Environment and Task

The images were presented on a 29" diagonal CRT display, a Bang & Olufsen *Beocentre AV5* television, driven by a graphics card in the host PC. The colour characterisation of the television employed the gain-offset-gamma (GOG) display model developed by Berns[9]. The television white point was 6500K. The peak white luminance was measured to be 108.7 cd/m^2 and the black background luminance was 1.0 cd/m^2 in the dark environment of the experimental room.

Nine subjects, 5 females and 4 males, with normal or corrected-to normal vision took part in the experiment. They had no deficiencies in colour vision and their ages ranged between 26 and 50.

The experiment was conducted in a darkened room. Using the pair-comparison technique, each subject was successively presented with 80 pairs (original image plus four enhanced versions giving 10 combinations, times 2 degrees of adjustment, times 4 images) in randomised order. Pairs of images were presented side by side on a black background, separated by 15 mm, and were shown with flexible timing until the subjects gave a response. For each pair, the subject was asked to make a judgement as to which image appeared sharper, *Right* or *Left.* Subjects were forced to make a choice in this experiment, even when the two displayed images appeared equally sharp. The experiment was repeated four times, with the subject seated at four different viewing distances, 60, 120, 240 and 480 cm, measured from the subject's forehead to the centre of the television screen. To reduce subject fatigue, the experiment was conducted in two sessions of about 45 minutes each, with two distances in each session.

Figure 5. ISO 12640 (SCID) images used in experiment

4. Results and Data Analysis

The observers' raw numerical judgements were transformed into a subjective sharpness scale in terms of *z*-score, according to Thurstone's 'Law of Pair Comparison'[10]. The decisions were converted to an interval scale where the 95% confidence interval for *N* observations was calculated as:

$$95\%CL = \pm 1.96\frac{(1\sqrt{2})}{\sqrt{N}} \tag{8}$$

Since the number of observers was N=9 in this experiment, the confidence interval around each scale value was 0.46. Hence if the mean scale values were within 0.46 of each other, there would be no significant difference between the mean perceived sharpness. The results of the experiments, averaged over all observers, are summarised in Fig. 6, in which the upper graph is for the lesser degree of adjustment.

At each distance in Fig. 6 are shown the results for the four enhanced images I_2 .. I_5 plus the original image I_0 for reference. Although there are small differences among the four test images, the trends are clearly defined. The images in which the highest spatial frequencies had been enhanced (I_2) were preferred in all cases to the original. The images in which the lower spatial frequencies had been enhanced were given successively lower sharpness ratings. For the first three viewing distances (60, 120 and 240 cm) the trend was monotonically upward to Band 2. At the longest distance (480 cm) the preferred sharpness seems to peak at Band 3. The results for the double strength enhancement (bottom row of Fig. 6) show similar trends to the standard enhancement (top row), except that the lower frequency results (especially I_4 and I_5) were rated worse.

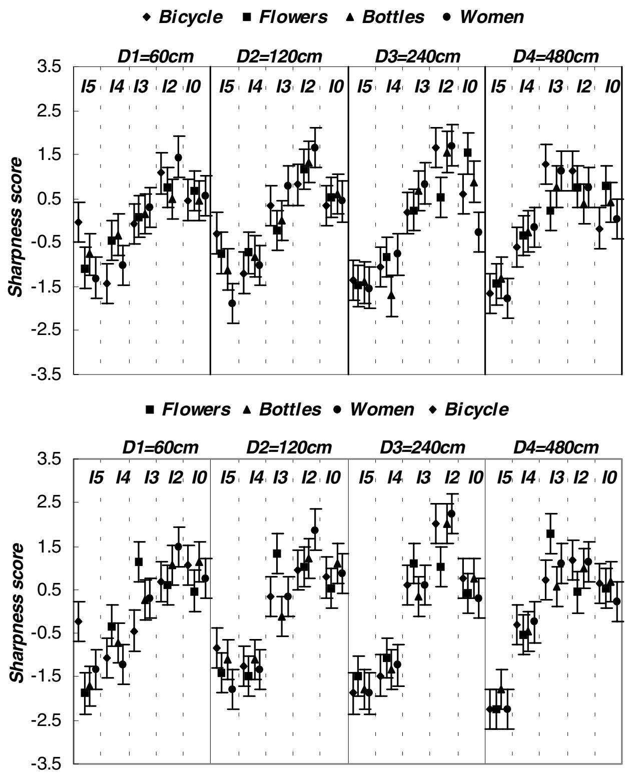

Figure 6. Perceived sharpness (95% confidence interval). Top: k_1=1.25, k_2=1.5; Bottom: k_1=1.5, k_2=2; (I_0 = original image; I_2, I_3, I_4, I_5 = enhanced images)

Do the results support the hypothesis? It appears that there is definitely a relationship between image spatial frequency and preferred image sharpness. But the peak spatial frequencies are higher than anticipated (see Fig. 4). At 240 cm the peak is in Band 2 whereas it was expected between Bands 3 and 4, and at 480 cm the peak is between Bands 2 and 3 whereas it was expected between Bands 4 and 5. This suggests a shift of about two octaves in the peak of the observer contrast sensitivity in this experiment.

There is a fundamental issue at the heart of this experiment. Contrast sensitivity functions are based on threshold experiments in which observers make judgements on just-perceptible changes in simple stimuli. Here we asked observers to make super-threshold judgements on complex images. There is no guarantee that standard CSFs should apply under these conditions. It would seem that the observers actually based their judgements on higher spatial frequencies in the images than would be expected from CSF performance data alone.

In conclusion, it was shown in this study that the perception of sharpness in colour images could be influenced by separating the luma component into spatial frequency bands and manipulating the bands individually. In effect this redistributes the power of the image across different spatial frequencies. The preferred sharpness depended on viewing distance but was largely independent of image content. Further experiments are planned with higher resolution images and higher resolution graphic displays to investigate the relationship between viewing distance and spatial frequency enhancement.

References

1. Inoue A. and Tajima J. (1993) Adaptive Image Sharpening Method Using Edge Sharpness, *IEICE Trans. Inf. & Syst.*, E76-D/**10**, 1174-1179.
2. Hultgren B.O. (1998), The Role of a Virtual Observer in an Automatic Image Processing System, *Proc. IS&T's PICS Conf.*, 150-153.
3. MacDonald L.W. (1999), Framework for an Image Sharpness Management System, *Proc. 7th IS&T/SID Color Imaging Conf.*, 75-79.
4. Garrett M. J. and Fairchild M.D (2000), Sharpness rules, *Proc. 8th IS&T/SID Color Imaging Conf.,* 24-30.
5. Bouzit S. and MacDonald. L.W (2000), Image Sharpness and colour difference metrics, *Proc. 8th IS&T/SID Color Imaging Conf.,* 262-267.
6. Daly S. (1993), The visible difference predictor: An algorithm for the assessment of image fidelity. In Watson A.B. (Ed.), *Digital Imaging and Human Vision*, MIT Press, Cambridge, MA, 179-206.
7. Barten, P.G.J. (1990), Evaluation of subjective image quality with the square-root integral method, *J. Opt. Soc. Am.*, **7**(10), 2024-2031.
8. ISO Standard 12640 (1997), *Graphic Technology – Prepress Digital Data Exchange CMYK standard Colour Image Data* (CMYK/SCID).
9. Berns R.S. (1996), Methods for Characterising CRT Displays, *Displays,* **16**: 173-182.
10. Torgerson W.S. (1958), *Theory and methods of scaling*, John Wiley, New York.

Biography

Samira Bouzit is a research student at the Colour & Imaging Institute, where she is studying the relationship between colour appearance and sharpness of images reproduced in a variety of media, including displays and print. She has a B.Sc. in Electrical Engineering and and an M.Sc. in Robotic Systems from the University of Versailles, France.