

Colour Difference Metrics and Image Sharpness

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

Abstract

We describe experiments to assess differences in the perceived sharpness of images, in which a Gaussian filter of variable width was convolved with a set of test images. Differences between the original and filtered images were computed pixel by pixel using three colour metrics: (a) mean square error (MSE) computed in RGB values; (b) CIELAB colour difference; and (c) s-CIELAB spatial colour difference. A psychophysical experiment was performed in which observers assessed the sharpness of the filtered images on a numerical category scale. The modulation transfer function (MTF) for the display was measured and the square root integral (SQRI) was evaluated as a predictor of perceived image sharpness. The results showed good agreement between the mean observer judgements and the prediction of both the colour difference metrics and the image quality metric across the full range of widths of the Gaussian filter.

Introduction

Sharpness is known to be one of the important factors relating to the perceived quality of reproduced images. The sharpness of an image is dependent upon the image formation and capture process (optics and sensor), the digital encoding and intermediate image processing (including resizing and compression) and the reproduction process (display or printer). It is also influenced by the visual acuity and state of adaptation of the observer. When an image is sharp, more detail can be discerned – sharp edges permit the observer to discriminate features more clearly, and sharp details permit the observer to recognise object surface characteristics more accurately.

Interest in image sharpness has grown because of its importance in cross-media reproduction. MacDonald[1] recently proposed a framework based on the human visual contrast sensitivity function (CSF) and the Modulation Transfer Function (MTF) of the input and output devices to determine an optimum correction to be made to the sharpness of an image. Such a system would support the semi-automated sharpness enhancement of images from any source device to any destination device.

A different approach was taken by Bech et al.,[2] who developed the Rapid Perception Image Description (RaPID) methodology for determining the overall perceived image quality of an imaging system. This permits engineers to describe and quantify the perceptual factors that substantially contribute to image quality, and then to develop perceptual models based on those factors. The RaPID model was successfully used by Nijenhuis et al.[3] where sharpness was the first perceptual attribute tested. The results showed linear relationships between perceived detail rendering, modulation depth and contour rendering. With so many variables affecting perceived image sharpness, it was concluded that future imaging systems would benefit from more optimised designs.

In conjunction with the growth in digital imaging technology, many techniques have been developed to improve the appearance of sharpness of digital images. One of the most widely used is Unsharp Masking (USM),[4] which offers great potential for improving image quality. A family of linear filters for image sharpening without altering image colour content was described by Mintzer.[5] For medical images such as X-rays, Ogoda et al.[6] made a significant improvement to standard unsharp masking by selectively controlling the degree of enhancement for each spatial frequency band. Deng[7] used the Logarithmic Image Processing (LIP) model, with a non-linear mapping function to enhance both contrast and sharpness. The comparison of different image enhancement algorithms is difficult, because a given method may work well in one application but completely fail in another.

The study described in this paper was part of a programme of research to investigate the perceived sharpness of colour images on a television display after processing with filters of various types. In order to design the optimum image sharpening filter, we need to understand how perceived image sharpness varies with different filter parameters. Ideally, one would perform a psychophysical experiment with human observers to assess the perceived image quality for a large set of images for every possible combination of filter parameters. Unfortunately the time required to perform such an extensive programme of experiments makes this approach unrealistic. Instead one must supplement and reinforce the subjective (psychophysical) assessment with objective (computational) methods for estimating image quality.

Quality metrics fulfil this need by encapsulating our understanding of human visual processing.[9] By making suitable measurements of device and image characteristics and viewing conditions, and computing metrics based on

these measurements, we aim eventually to be able to predict the visibility of sharpness. As a first step we examined the correlation between colour difference metrics and observations of image sharpness. In the present study four metrics were tested:

- Mean square error (MSE) computed pixel by pixel in RGB values;
- CIE 1976 ($L^*a^*b^*$) colour difference formula ΔE^*_{ab} [10];
- s-CIELAB spatial colour difference metric[11];
- Square root integral (SQRI) image quality metric[12, 13.]

The s-CIELAB metric is particularly interesting because it makes an attempt to model the response of the human visual system to different spatial frequencies, and one would therefore expect it to produce a better prediction of visual judgements than the other metrics. The algorithm comprises three stages:

(1) The image is converted into a device-independent representation consisting of one luminance and two chrominance components;
(2) Each component image is convolved with a spatial filter tuned to the corresponding spatial sensitivity of the human eye;
(3) The filtered image is transformed via CIE-XYZ into CIELAB co-ordinates, from which the colour difference is calculated.

The SQRI metric, which was developed by Barten[12] for evaluation of image quality, can be applied to images on CRT displays. The human visual response is taken into account by making use of the modulation threshold function of the eye.

Experimental Approach

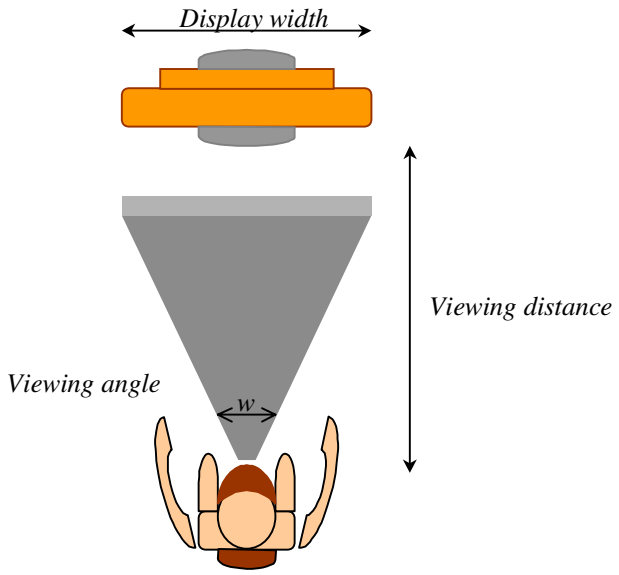
Experimental Setup

In the experiment we used a standard Bang & Olufsen *Beocentre AV5* television (29-inch diagonal, 4:3 aspect ratio) to display a still image generated in PAL format from a video card in the PC host computer.

The colour characterisation of the television employed the gain-offset-gamma (GOG) display model developed by Berns,[14] so that the red, green and blue digital pixel values needed to produce any desired colour could be determined. This model was used to process the test images. The peak white luminance was measured using a Minolta CS-1000 telespectroradiometer as 108.7 cd/m^2 and the lowest black-level was less than 1 cd/m^2 in the dark environment of the experimental room with black-out blinds over the windows and lights off. The television white point was 6500K.

Spatial Characterisation of the Display

The viewing environment provided to the observer was as shown below:



The *BeoCenter AV5* television had dimensions 47×37 cm (4:3 aspect ratio) corresponding to a display area of 720×576 pixels. At the fixed viewing distance of 180 cm, the dimensions of a pixel on the screen were 0.64×0.65 mm.

The angular subtense of one pixel from the observer's viewing position was calculated as:

$$w_{pixel} = \frac{d}{l} \times \frac{180}{\pi} = \frac{1}{1800} \times \frac{180}{\pi} = 0.032\,\text{deg}$$

which corresponds to 15 cycles/degree.

The angular subtense of the display from the observer's viewing position was calculated for vertical and horizontal directions as:

$$w = \frac{d_v}{l} \times \frac{180}{\pi} = \frac{370}{1800} \times \frac{180}{\pi} = 11.78\,\text{deg}$$

$$w = \frac{d_h}{l} \times \frac{180}{\pi} = \frac{470}{1800} \times \frac{180}{\pi} = 14.96\,\text{deg}$$

The modulation transfer function (MTF) of an imaging system or device describes the its ability to reproduce spatial frequencies. In this study, the modulation transfer function of the television display was determined by using a still digital camera.

Firstly, the monochrome MTFs of the display-camera system were evaluated for both horizontal and vertical display orientations by using black-and-white bar patterns. For each individual spatial frequency the modulation was calculated using eqaution (1). The MTFs were determined by the ratio of the output to the input modulation.[15]

$$Modulation = \frac{V_{max} - V_{\min}}{V_{\max} + V_{\min}} \qquad (1)$$

Secondly, the MTF of the camera alone was determined by using the Noise Power Analysis technique.[16]

Finally, the MTF of the television for each orientation was determined:

$$MTF_{Television} = MTF_{System} / MTF_{Camera} \quad (2)$$

Figure 1 illustrates the MTF curves for both display orientations as a function of spatial frequency. The curves do not show significant differences; by giving greater low frequency responses and poorer high frequency responses, with the vertical response being relatively higher.

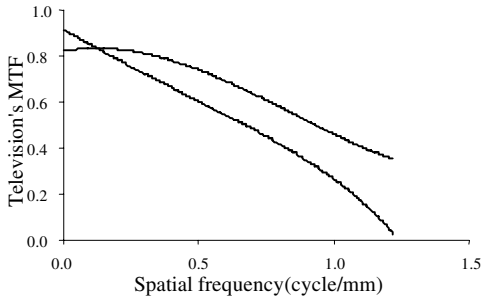


Figure 1. MTFs of the television using bar patterns target for vertical orientation (top) and horizontal orientation (bottom)

Test Images

A set of four test images was selected from the ISO 12640 (SCID) set,[17] denoted *Bicycle*, *Flowers*, *Woman* and *Bottles*. The first scene contains fine details of a bicycle wheel, sine patterns and different shapes with high chroma colours. The second is a flower in transparent glass vase producing sharp contrast against both a white dish and a dark defocused background. The third is a young woman with clear skin tones and fine details in the hair against a neutral grey background. The fourth is an arrangement of silver objects and glassware with predominantly high-key tones and many specular highlights. The images were converted from CMYK to RGB format at 8 bits/pixel, and scaled to 720×576 pixels in size (full screen for television display).

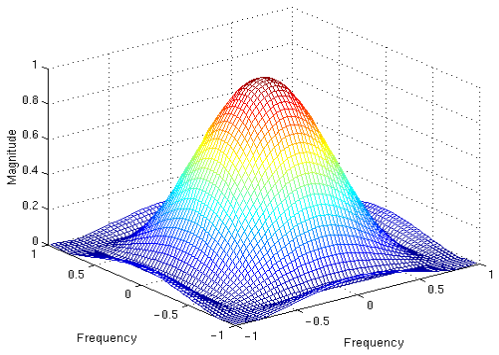


Figure 2. Two-dimensional Gaussian filter of 0.5-pixel width.

Image Processing

Blurred versions of each original image were generated by convolution with a low-pass Gaussian filter, as shown in Figure 2. The filter width was varied to produce different degrees of blur in the image. Eleven filtered versions were computed for each test image, with the standard deviation (width) of the spatial filter ranging from 0.1 to 1.0 pixels in steps of 0.1 for the first ten and 1.5 pixels for the last. Thus there were 12 versions of each image (the original plus 11 filtered images), making a total of 48 test images for the four scenes.

Because the sharpness of an image depends much more on the luminance than the chrominance content,[18] and for reasons of computational efficiency, we wished to investigate the effects of processing only the image luminance component. Various colour spaces that separate luminance from chrominance could be used for this purpose, such as CIE-L*a*b*, CIE-L*u*v*, YIQ and HSV, but for this study we chose the YC_bC_r colour space transformation defined by NTSC[19]:

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ C_b &= -0.168R - 0.331G + 0.5B \\ C_r &= 0.5R - 0.418G - 0.081B \end{aligned} \quad (3)$$

where Y is the luma component, and C_b, C_r are blue and red chrominance components respectively. Following television convention, we use the term 'luma' for Y rather than the usual 'luminance' to avoid confusion with CIE luminance. Figure 3 shows the image processing procedure used in this study. The original colour image was converted into YC_bC_r colour space. Only the luma component was filtered, then recombined with the chrominance components. The filtered image was transformed to CIE-*XYZ* co-ordinates and then converted to CIELAB using the standard formulae. Finally, the three colour difference metrics defined above were computed on a pixel-by-pixel basis and their frequency distributions examined.

Procedure

The images were evaluated using the category judgement technique by 10 observers from the Colour & Imaging Institute, 5 female and 5 male, with ages ranging from 22 to 49. All had normal colour vision and normal or corrected-to-normal visual acuity. The images were displayed on the television in a darkened room. The viewing distance from the observer's eye position to the centre of the screen was 180 cm. Observers were asked to read the instructions, then allowed to adapt to the viewing conditions for 5 minutes. In order to help them to establish an internal judgement scale, they were trained by displaying a series of images different from those in the real experiment. Observers were asked to make their judgements using an eleven-point numerical category scale ranging from 0 (the lowest degree of perceived sharpness) to 10 (the highest degree of perceived sharpness), by answering the following question:

"Please assess the degree to which you consider that the sharpness of the image is optimum."

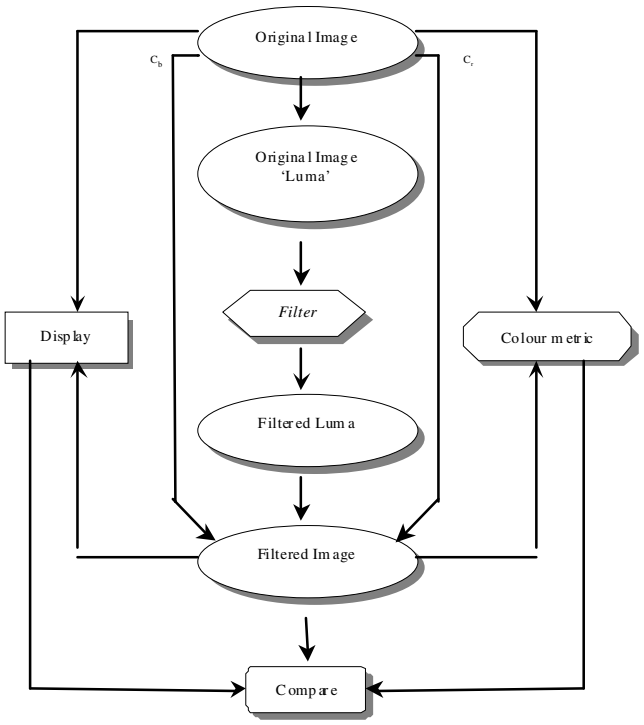


Figure 3: Flow chart of the image processing procedure.

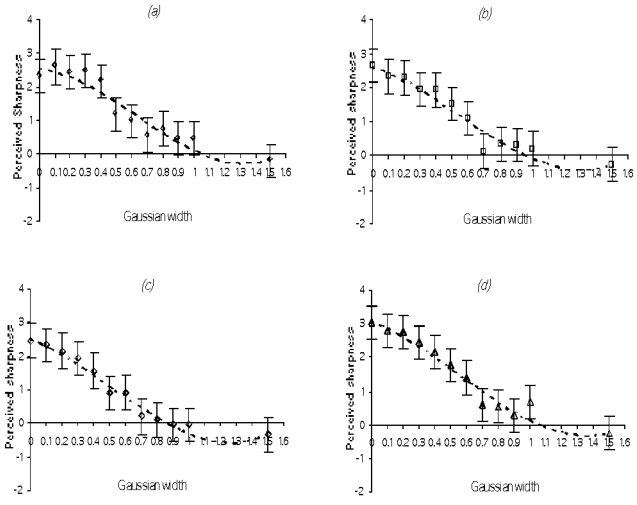


Figure 4. Perceived sharpness (95% confidence interval) versus Gaussian filter width for: (a) Bicycle (b) Flowers (c) Woman (d) Bottles

This technique has previously been used successfully for assessment of other perceptual attributes of images on a television display.[20] The observers were allowed as long as they wished to make the judgement of each image. The sequence of 48 test images was presented in random order separately for each observer.

Results

Observer Estimation

As is well known, low-pass Gaussian filtering reduces the high frequency components in an image, and thus results in blurring of the image (and thus a reduction in fine detail associated with noise). In our processing, after filtering the luma component, most of the fine texture had been filtered away, and the gradients of edge boundaries were reduced. The overall balance of light and dark tones was unaffected, because their low spatial frequencies were well within the pass-band of the filter.

The observers' raw numerical judgements were transformed into a subjective sharpness scale, according to Thurstone's 'Law of Categorical Judgement'.[21] The decisions were converted to an interval scale where the 95% confidence interval for N observations was calculated as:

$$95\%CL = \pm 1.96\frac{(1\sqrt{2})}{\sqrt{N}} \tag{4}$$

Since the number of observations was N=10 in this experiment, the confidence interval around each scale value was 0.438. Hence if the mean scale values were within 0.438 of each other, there was no significant difference between the sharpness judgements for the test images. Figure 4 shows the mean perceived sharpness, in terms of z-score calculated from Equation (4), as a function of Gaussian filter width. The four graphs show similar tendencies: for low filter width there was high perceived sharpness, but for larger filter widths, noticeably above 0.5 pixels, the perceived sharpness became insignificant.

Colour Difference Metrics

The mean square error (MSE) was calculated as a point-by-point vector between the original and filtered images:

$$MSE_{ij} = \sqrt{(\Delta R_{ij}) + (\Delta G_{ij}) + (\Delta B_{ij})} \tag{5}$$

where ΔR, ΔG and ΔB represent the differences in the R,G,B channels respectively.

The standard CIELAB colour difference formula was used to calculate ΔE^*_{ab} pixel by pixel between the original and the filtered images. The s-CIELAB metric was computed using the method described by Zhang & Wandell.[11]

Figure 5 shows the relationship between the mean colour difference (metrics) and mean perceived sharpness (observers) for all the images, computed across the full range of Gaussian filter widths. The results, using the best fitting line by linear regression, indicate a strong correlation between the predictions of the metrics and the sharpness perceived by the observers. The average errors for the RMS, ΔE^*_{ab} and s-CIELAB metrics were 0.024,

1.67 and 0.557 respectively, which are of the order of average perceptibility tolerances for complex images.

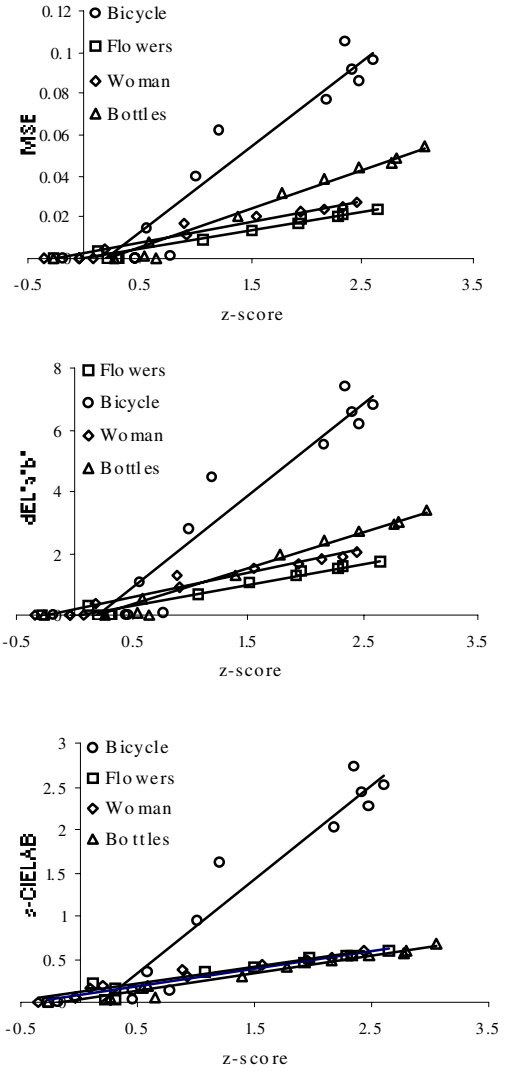


Figure 5. Observer z-score versus colour difference for three metrics

Image Quality Metric

The square root integral (SQRI) image quality metric proposed by Barten[11,12] was calculated as:

$$SQRI = \frac{1}{\ln 2}\int_0^{U_{\max}} \sqrt{\frac{M(u)}{m_t(u)}}\,\frac{du}{u} \qquad (6)$$

where u is the angular spatial frequency at the eye of the observer in cycles/degree, u_{max} is highest frequency to be displayed. $M(u)$ is the MTF of the display as shown in Figure 1, $m_t(u)$ is the modulation threshold function of the eye calculated using Barten's model.[12] In the SQRI metric, the non-linear behaviour of the eye has been taken into account by using the square root of the ratio of $M(u)$ and $m_t(u)$ in the integrand.

SQRI expresses the image quality in just noticeable differences, due to the normalising factor of *1/ln2*. This factor was included on the basis[12] that the value of the metric increases by 1 when the square root of the ratio of the MTFs increases by 1 in a spatial frequency octave (i.e. doubling of frequency. SQRI values were related to the subjective image sharpness data for each individual image.[13]

In Figure 6, the subjective quality of the set of four display images is plotted as a function of the calculated SQRI. The subjective image sharpness appears to be highly correlated with the calculated SQRI value. The coefficient R^2 for the correlation between subjective image sharpness and calculated SQRI value was 83.79%, 91.28%, 91.71%, 91.28% for the Bicycle, Flowers, Woman and Bottles images respectively.

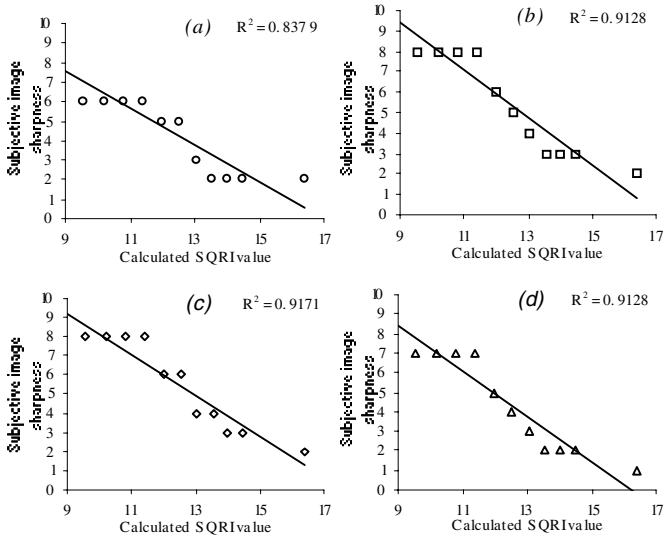


Figure 6. Subjective image sharpness as a function of calculated SQRI values: (a) Bicycle (b) Flowers (c) Woman (d) Bottles

Conclusion

The results of this study showed that the mean observer judgement of perceived image blur was well correlated with the mean values of various colour difference metrics and that it was dependent on the scene content. The results show a good correlation with the perceived sharpness at different standard deviation. The perceived overall image sharpness shows a high correlation with the calculated SQRI values.

References

1. MacDonald L.W., Framework for an Image Sharpness Management System, *Proc. 7th IS&T/SID Color Imaging Conference*, 75-79, 1999.
2. Bech S., Hamberg R., Nijenhuis M., Teunissen C., Looren de Jong H., Houben P., and Paramanik, S.K., The RaPID Perceptual Image Description Method (RaPID), *Proc. SPIE*, 2657: 317-323, 1996.
3. Nijenhuis M, Hamberg R., Teunissen C., Bech. S, Looren de Jong H., Houben H., and Pramanik S.K., Sharpness related attributes and their physical correlates, *Proc. SPIE*, 3025: 173-183, 1997.
4. Gonzalez R.C. and Woods R.E., *Digital Image Processing*, 3rd Ed., Addison-Wesley, 1993.
5. Mintzer F. and Braudaway G.W., A Family of Linear Filters for Image Sharpening, *Proc. 3rd IS&T Technical Symposium on Prepress*: 144-143, 1993.
6. Ogoda M., Unsharp Masking Technique Using Multi-Resolution Analysis for Computed Radiography Image Enhancement, *Journal of Digital Imaging*: 10/3: 185-189, 1997.
7. Deng G., Cahill L. W., and Tobin G. R. The study of Logarithmic Image Processing Model and Its Application to Image Enhancement, *IEEE Transaction on Image Processing,* **4**/4:506-512, 1995.
8. International Commission on Illumination (CIE), *Recommendations on uniform color spaces, color difference equations, and psychometric color terms*, Publication CIE 15, Supplement No. 2, Bureau Central de la CIE, Vienna, 1971.
9. Farrell J.E., Image Quality Evaluation, *Colour Imaging: Vision and Technology*, in MacDonald L.W. and Luo M.R. (Eds.), John Wiley, Chichester: 285-313, 1999.
10. International Commission on Illumination (CIE), *Industrial Colour-Difference Evaluation*, Publication CIE 116-95, Bureau Central de la CIE, Vienna, 1995.
11. Zhang X.M. and Wandell B.A., A spatial extension to CIELAB for digital color image reproduction, *SID Digest,* 27:713-734, 1996.
12. Barten, P.G.J., The SQRI: A new method for the evaluation of visible resolution on a display. *Proceedings of the Society for Information Display*, 28, 253-262, 1987.
13. Barten, P.G.J., Evaluation of subjective image quality with the square-root integral method, *Journal of the Optical Society of America*, 7(10), 2024-2031, 1990.
14. Berns R.S., Methods for Characterising CRT Displays, *Displays,* 16, 173-182, 1996.
15. Holst G. C. (1996) *CCD Arrays, Cameras and Display*, SPIE Press, JCD Publishing
16. Van Metter R., Measurement of MTF by Noise Power Analysis of one dimensional white Noise Patterns, Journal of Photographic Science, 38, 144-147, 1990.
17. ISO Standard 12640, *Graphic Technology – Prepress Digital Data Exchange CMYK standard Colour Image Data* (CMYK/SCID), 1997.
18. Hunt R.W.G., Why is Black-and-White so Important in Colour?, *Colour Imaging: Vision and Technology*, in MacDonald L.W. and Luo M.R. (Eds.), John Wiley, Chichester: 3-15, 1999.
19. Jain, A. K., *Fundamentals of the Image Processing*, Prentice Hall, London, 1989.
20. MacDonald L.W., Yendrikhovskij S., Bech S., Jensen K., Toward Optimal Colour Image Quality of Television Display, *Proc. SPIE,* 3963: 260-270, 2000.
21. Torgerson W.S., *Theory and methods of scaling*, John Wiley, New York, 1958.