

The Influence of Image Gamuts on Cross–Media Colour Image Reproduction

Ján Morovic and Pei–Li Sun
Colour & Imaging Institute, University of Derby
Derby, United Kingdom

Abstract

The present paper aims to investigate the influence of image gamuts on cross–media colour image reproduction by creating a set of images that have this image characteristic perturbed in a known way. This is done using the following two approaches: First, by selecting a single image and obtaining variations of it which have different colour gamuts. Second, by creating a pair of image sets whereby the images in the first set differ in colour gamut and the images of the second set are transformations of the first set of images so that all the images have the same gamut. Reproducing these images using a range of gamut mapping algorithms (GMAs), it can be seen whether variations in gamut and the difference between multi– and equi–gamut sets result in difference of performance. The results of these experiments then show that image gamuts have no significant effect on colour gamut mapping.

Introduction

Image gamuts have been extensively used in cross–media colour image reproduction systems as parameters of the gamut mapping stage and this has been widely shown to result in significant improvement.[1-4] However, the impact of the image gamut as an image characteristic has not been studied before in isolation. To do this, one needs to create such conditions where the image gamut characteristic is the variable in which differences lie predominantly rather than just being one of many dissimilar characteristics in terms of which test images differ, as is normally the case. One way of doing this is by perturbing image gamuts in a known way and having sets of images which differ in image gamut and corresponding sets of images which do not. It is this idea, which was previously introduced as a general framework for investigating the influence of image characteristics,[5] that is the basis of the experimental work presented in this paper.

The aim of this paper is to provide details of what considerations went into obtaining these sets of images and to present the experiment in which they were evaluated. The results will provide information about how significantly image gamuts influence the performance of colour reproduction systems using the given GMAs.

Problems of Image Gamut Description

For GMAs that take the image gamut as a parameter, the image gamut boundary descriptor (IGBD) is one of the key elements for making good colour reproductions. However, IGBDs may produce different results when they are obtained using different processes and according to different rules. Three potential problems of calculating IGBDs, resulting in a mismatch between calculated and perceived image colour gamuts and influencing the performance of GMAs are discussed next:

Colour Spatial Frequency

The human visual system cannot detect fine detail in an image due to optical blur and resolution limitations.[6] Hence, the colours in high spatial frequency parts of images (e.g. frequencies higher than 60 cycles per degree for luminance variation) might not be suited for inclusion in the process of calculating representative image gamut boundaries. Two techniques, image sub–sampling and the s–CIELAB model,[7] can be used to alter images and thereby remove high spatial frequency colours. In the image sub–sampling method (where the extent of sub–sampling should be determined on the basis of the contrast sensitivity function), local colours are averaged by reducing image size while s–CIELAB uses a Gaussian filter to blur an image to an extent simulating the characteristics of the human visual system. Both these techniques reduce the impact of very high spatial frequency colours on the calculation of image gamut boundaries.

Sample Selection

If an image contains only a few saturated colours, should those colours be regarded as samples from the image gamut boundary? If the answer is no, what kinds of samples should be used? It will be assumed in this study that colours occurring only very infrequently in an image should not be regarded as the samples on which the IGBD is based. To avoid containing very low probability colours in IGBDs, two techniques – high–pass filtering and cumulative colour histogram clipping[8] – can be used. The high–pass filtering method suggested here removes the colours whose pixel–frequency in an image is very low whereby a threshold is needed for this operation. Cumulative colour histogram clipping is a method where a cumulative colour histogram is first created for each centre–radial sector and then colours which correspond to the 95% cumulated frequency are used for calculating the IGBD. A disadvantage of this method is that if most of the image colours are in one sector, some saturated colours whose overall probability is not very low will be removed.

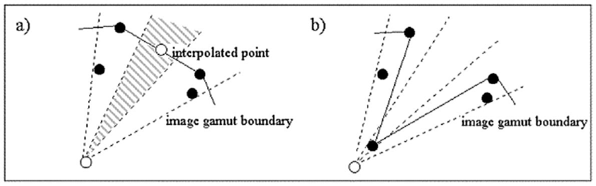

Figure 1. Image gamut boundaries when a segment have a) no colour or b) a low–chroma colour.

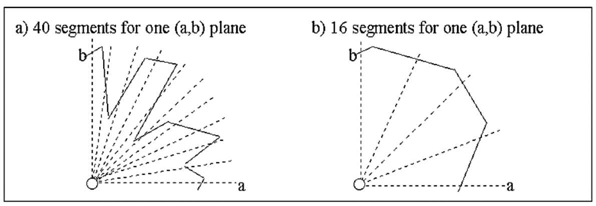

Figure 2. Image gamut boundaries using a) large number of segments and b) small number of segments

Smoothness of Gamut Boundary

The segment maxima GBD method[9] is a technique for describing gamut boundaries which first divides a gamut into many centre–radial segments and then finds the maximum radius (*r*) for each segment. When using this method for image gamut description and if a segment contains no colour, then the maximum *r* for the segment will be obtained by interpolating the maxima from neighbouring segments. However, quite different results will be obtained if a low–chroma colour is located in the segment (Figure 1). The size of the segments may also influence image gamut description. Normally, the larger the segment–size, the smoother, but also less accurate, the image gamut boundary (Figure 2). While a gamut boundary which has been smoothed does not represent the actual gamut boundary of an image, smoothing might be necessary for obtaining smooth and artefact–less colour reproduction when original and reproduction gamut boundaries are quite different in shape and especially when they are locally polymodal.

To smooth the image gamut boundary, four solutions are proposed here – segment reduction, IGBD averaging, Gaussian smoothing and 2D gamut function modelling. In segment reduction, the number of segments is simply reduced. It can reduce the likelihood of this problem caused by a segment containing low–chroma colours. A smoother image gamut boundary can also be obtained by averaging the IGBDs obtained using both large and small numbers of segments. Using the Gaussian smoothing technique, the maximum *r* of a segment is replaced by a weighted average of the maximum radii of neighbouring segments and weights are derived using the Gaussian distribution function. This method provides a gamut boundary which is smooth and approximately close to the real image gamut. Alternatively, a 2-D gamut functions[10] which use a least-squares method for curve-fitting to saturated image colours can also smooth the image gamut boundary. Further work is necessary for evaluating the performance and parameters of all the methods mentioned above for producing optimal image gamuts.

In this study, the CAM97s2 JCh colour space[11] was used and image gamut boundaries were produced as a 2D LUT using the following steps: (a) image sub–sampling: blur very high spatial–frequency colours; (b) high–pass filtering: remove very low probability colours; (c) segment maxima GBD method: find gamut boundary colour for each of 16×16 sectors.

The gamut smoothing process was not used in this study because it does not describe actual image gamuts, instead images where smoothing would have been necessary were not used. In this experiment the aim is to provide particular sets of test images rather than be able to use any image in such sets.

Perturbing Image Gamuts

To investigate the influence of image gamuts on gamut mapping, two experiments were conducted: the equi–gamut test and the equi–image–content test:

Equi–gamut Test

This test aims to evaluate the performances of GMAs for images having the same image gamut but different image contents. The steps of the test are as follows. First, chose four images (O_{m1-m4}) where each image is of a different type and has a different histogram, different spatial characteristics, different image gamut, etc. Second, use the steps mentioned in the previous section for obtaining image gamuts (G_{m1-m4}) and then average the four, resulting in an averaged gamut G_e. The reason for using the image–set average gamut is to approximately equally change the gamut of each of the original images. Third, based on the image gamuts (G_{m1-m4}), colours of the four images (O_{m1-m4}) were mapped (both compressed and extended, depending on gamut boundary conditions along a given line) to the target gamut G_e relative to the centre of the colour space. The mapping resulted in a set of original equi–gamut images (O_{e1-e4}) (Figure 3). Fourth, all reproductions were made of the two sets of originals using the given GMAs. Fifth, two sets of psychophysical experiments were conducted whereby the experiment on the multi-gamut images (O_{m1-m4}) and their reproductions (R_{m1-m4}) resulted in scores S_{m1-m4} (range of the scores is SR_m) and the experiment on equi–gamut images (O_{e1-e4}) and their reproductions (R_{e1-e4}) resulted in scores S_{e1-e4} (having a range of SR_e).

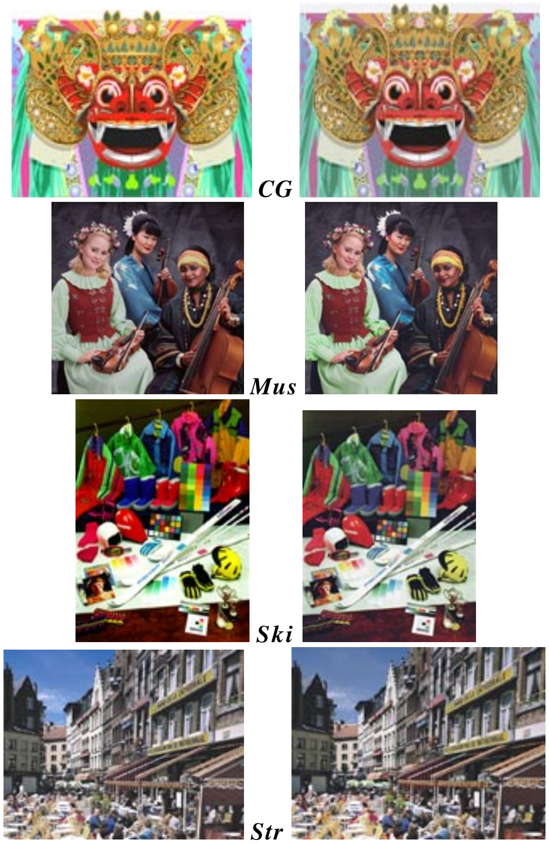

Figure 3. Multi–gamut images (left column) and equi–gamut images (right column).

If the results coming from the equi–gamut set produce smaller differences (i.e., $SR_e < SR_m$) then this will suggest that image gamuts indeed influence the performance of GMAs. Alternatively the difference in performance is due to some other image characteristic.

Equi–image–content Test

This test aims to evaluate the performance of GMAs when the image gamut changes while the image contents are kept the same. A high chroma image Ski (Figure 3) was used as the standard original and the CAM97s2 lightness and chroma of the image were transformed using Equations 1 and 2 respectively and three compression ratios – 0.7, 0.85 and 1.0 – were used. This resulted in a set of 9 (3×3) original images showing the same scene but having different colour gamuts. Some colours moved out–of–gamut after the compression and these colours were mapped to the gamut boundary of the CRT towards the gamut centre (J,a,b) = (50,0,0). The results of this mapping were regarded as the originals for the equi–image–content test.

Psychophysical experiments were conducted whereby the evaluation of originals and their reproductions resulted in scores S_{JC} (having a range of SR_{JC}). As the main difference between the originals was their image gamut, a similar range (SR_{JC}) for each original would suggest that the image gamut has little influence on the performance of GMAs.

$$L_{compressed} = Comp.Ratio_L \times (J - 50) + 50 \tag{1}$$

$$C_{compressed} = Comp.Ratio_C \times C \tag{2}$$

Experimental Setup

The accuracy of reproduction between the various original images and corresponding reproductions discussed in the previous section was evaluated by 12 observers using the pair comparison method[12] in a simultaneous binocular viewing setup under a D50 simulator. The originals were displayed on a calibrated CRT monitor (an Apple 21" Studio Display), and the reproductions were obtained with a Canon BJC–6100 bubble–jet printer on Canon HR-101S coated paper. The CRT monitor was characterised using the GOG model[13] and the mean and maximum errors were 0.89 and 1.80 ΔE_{97s2} units respectively. The printer was characterised using an inverse 10^3 3D LUT with tetrahedral interpolation[14] and the mean and maximum errors were 2.08 and 5.81 ΔE_{97s2} units respectively.

Gamut Mapping Algorithms

The reproductions of the various images used here were obtained using the following four GMAs:

(a) CARISMA:[9] first compress lightness, shifts hues based on the six primaries of the two media and maps colours depending on the gamut shapes at corresponding hues.

(b) GCUSP:[9] maps lightness in a chroma–dependent way followed by linear compression towards the point on the lightness axis having the lightness of the cusp.

(c) WCLIP: maps out–of–gamut colours to the colour on the gamut boundary with the smallest colour difference obtained using a weighted colour difference formula[15] ((K_J:K_C:K_H) = 1:2:1 was used in this study).

(d) SKNEE:[16] first compresses lightness using a sigmoidal function and then compresses towards the cusp using a non-linear knee function. The image independent version for normal tonal image was used in the study.

Note, that the choice of GMAs (which were all image independent) was guided by two principles. Firstly, to chose GMAs which could perform well and secondly, to choose such a set of GMAs where the individual members are significantly different from each other. Hence there is a range of lightness mappings – linear, chroma–dependent, sigmoidal and clipping – as well as a range of chroma compressions – clipping, knee and linear – and a range of compression directions. This latter criterion is important for the purposes of this study, which is not the evaluation of GMAs but of the image gamut image characteristic.

Can Accuracy Score Ranges Be Compared?

As the present study intends to compare the ranges of scores from different experiments and as this is not a comparison normally done with the data obtained from pair comparison experiments, it is necessary to understand whether this comparison is valid.

The relative quality (i.e. in terms of accuracy here) of any two stimuli can be represented by a Z–score obtainable from raw observer responses in the pair comparison experiment. The unit of the relative quality scale is the unit of the Z–score scale, which is an interval scale (i.e., a scale that has equal intervals). Let us, for example have a pair of stimuli on an interval scale separated by three units and a second pair on another interval scale also separated by three units. The differences between the pairs will be perceptually equal given that both scales have the same units and that both scales represent the same stimulus quality. However, as there is no meaningful zero point on an interval scale, absolute quality cannot be shown on them. The zero point on a particular Z–score scale represents the mean quality of stimuli in a given experiment only. Different experiments normally have different mean stimuli (i.e. zero–points) on an absolute quality scale (Figure 4) and their Z–scores therefore cannot be compared directly. However, all the individual experiments that will be carried out here will have the same units and will represent the same perceptual quality (accuracy of match). The differences (ranges) of Z–scores between two sets of experiments can hence be compared in terms of ratio (e.g. one can say that the range of one experiment is twice that of another) and they represent difference of perceptual quality.

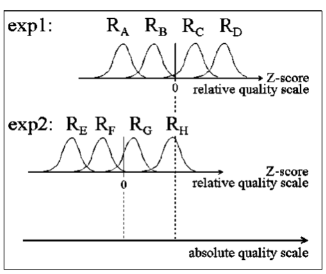

Figure 4. Comparison of Z–scores on an absolute quality scale.

Results

Equi-gamut Experiment

The performances of GMAs in the equi-gamut experiment in terms of Z-scores are shown in Table 1. In the table, each row represents the result of an independent pair-comparison experiment. The 95% confidence interval is ± 0.4 on the Z scale. The range of Z-scores (SR) is the difference between maximum and minimum Z-scores.

The assessment of the impact of making each image have the same gamut will be done in terms of looking at the relationship between image gamut volumes (IGVs) and accuracy score ranges. Even though the following analysis is based on IGVs, its conclusion extend to image gamuts as such as the images that had the IGVs of the average gamut also had gamut shapes that matched the gamut shape of the average. Hence claiming that the conclusions of this analysis only pertain to image gamuts is unjustified as the equi–gamut test images also agreed with each other in terms of all their other image gamut characteristics.

The image gamut volumes of the originals in this experiment were calculated by summing up sub-tetrahedral volumes of the image's three-dimensional gamut boundaries and the gamut volume of the original Ski image was 739,532 cubic CAM97s2 *Jab* units. The IGVs dealt with here are normalised to it are also shown in Table 1 for further data analysis.

Table 1. GMAs performances in equi-gamut experiment.

Images vs. GMAs		Accuracy (z-scores)				Stdev.	SR	IGV
		CARISMA	GCUSP	SKNEE	WCLIP			
Multi-gamut set	CG	-0.38	-0.46	-0.60	1.43	0.96	2.02	0.72
	Mus	-0.28	0.51	-0.46	0.22	0.44	0.96	0.35
	Ski	-1.18	-0.26	-0.14	1.57	1.15	2.75	1.00
	Str	-0.83	-0.58	0.58	0.83	0.83	1.66	0.33
Equi-gamut set	CG	-1.19	-1.44	1.08	1.55	1.53	2.99	0.49
	Mus	-0.59	0.04	0.20	0.35	0.41	0.94	0.56
	Ski	-1.78	-0.24	0.94	1.08	1.32	2.85	0.58
	Str	-0.81	-0.27	0.33	0.75	0.69	1.57	0.53
Overall		-0.88	-0.34	0.24	0.97	0.92	1.97	0.57

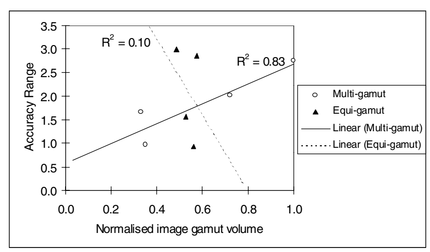

Figure 5. Accuracy range (SR_m & SR_e) vs. normalised image gamut volume (IGV).

Figure 5 illustrates the correlation between the accuracy range and the image volume. In the multi-gamut set, the correlation was strong (R^2 = 0.83). However, the correlation was very low (R^2 = 0.10) in equi-gamut set. This suggests that there is no significant relationship between SR and IGV.

Accuracy rankings of the four GMAs for the four multi-gamut images are shown in Figure 6. As can be seen, in the original set (left figure), it is significant that the rankings were image dependent. However, in the equi-gamut set, the agreement between the four equi-gamut images was much stronger. This suggests that GMA performance order is correlated with the image gamut characteristic.

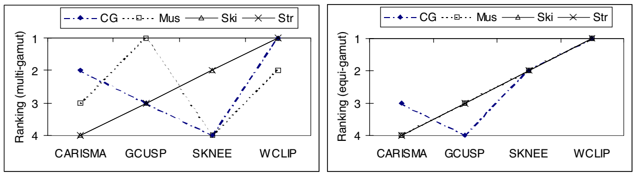

Figure 6: Image vs. GMA rankings (left: multi-gamut sets, right: equi-gamut sets).

The correlation between combined and individual image accuracy scores for the equi-gamut set was much higher and had a much lower standard deviation (mean = 0.96, stdev. = 0.03) than in the multi-gamut set (mean = 0.78, stdev. = 0.23). This suggests that GMA performance is image dependent but that performance becomes similar when the original images have similar image gamuts.

Equi-image-content Experiment

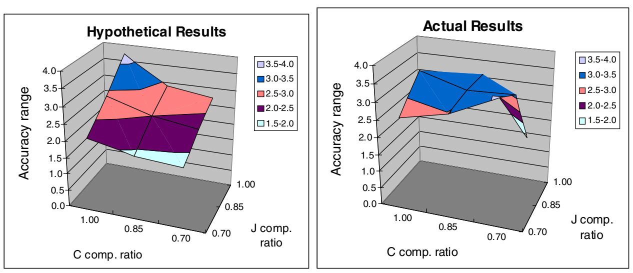

Figure 7. Accuracy range (SR_{JC}) vs. gamut compressing ratio (left: hypothetical results for IGV effect, right: actual results).

In the equi-image-content experiment, the averaged accuracy rankings of the four GMAs from the best to the worst again were WCLIP, SKNEE, GCUSP and CARISMA. If the SR_{JC} is correlated to the compression ratio, the ($J_{comp.ratio}$, $C_{comp.ratio}$, SR_{JC}) three-dimensional diagram is expected to be similar to Figure 7-left. The actual result (Figure 7-right), however, was quite different to what would be expected if there were a correlation between IGV and SR. No trend can be found between SR_{JC} and the compressing ratio. This means that neither lightness nor chroma ranges are correlated to the accuracy range. There is virtually no correlation ($R^2 = 0.01$) between the accuracy range (SR_{JC}) and the image gamut volume (IGV) and this backs up the idea that image gamut volume is not a factor influencing the performance of GMAs.

Correlation ($Corr_{JC}$) between averaged image accuracy scores and individual image accuracy scores were calculated and then plotted in ($J_{comp.ratio}$, $C_{comp.ratio}$, $Corr_{JC}$) three-dimensional space. The mean and standard deviation of the $Corr_{JC}$ values were 0.97 and 0.04 respectively. The high correlation suggests that the order of GMA performance is unrelated with image gamut.

Discussion

Two experiments were carried out for evaluating how significantly image gamut influences the performance of GMAs. The results of the two experiments are summarised as follows:

- Accuracy range is not correlated with image gamut.
- Accuracy ranking order changes with gamut change in an image dependent way and it hence depends upon characteristics other than image gamut.
- Variation of performance between algorithms for given test images is not due to their image gamuts

The possible reasons why no image gamut effect can be observed are as follows:

1. *There is no effect*
   Image gamut only represents extreme colour information about an image. Other colour related information such as probability, spatial frequency and local contrast are not included and it could be those, rather than the image gamut, that affect how images are to be reproduced.
2. *Properties of experimental setup do not give rise to effect*
   (a) GMAs: all the chosen GMAs were reported to perform well and recommended by various papers so that their performances did not differ much in this study. However, there are significant differences between their scores (e.g. see Ski image, Table 1).
   (b) Media: The difference between the two media used in this might not have been large enough for GMAs to show their differences. It might have been due to this that WCLIP, which made the smallest colour shifts, was judged to be superior to the other GMAs. However, accuracy ranges (SR) of the results had a range from 0.94 to 3.61. This indicates that the gamut difference of the two media and the differences between GMAs were sufficient.
   (c) Gamut boundary description: the way gamut boundaries are calculated affect the results of gamut mapping. However, this gives the same errors for all the GMAs and thus can have only little impact on the results of the study.
   (d) Characterisation: characterisation errors have been reported in pervious section. The errors were equal for all the GMAs and were not large enough to affect the results. If characterisation errors had been larger than differences between GMAs then the reproductions obtained using various GMAs would have been judged to be not significantly different.
3. *Inappropriate choice of images*
   The results for the multi-gamut set of images in the equi-gamut experiment were image dependent. As both multi-gamut and equi–gamut images show an accuracy range effect which is not removed by removing their gamut differences, it seems that the

effect was not due to their image gamuts. Furthermore, it is sufficient to show that there is no image gamut effect for the present set of images to show that there is no such effect for all images.

While the findings of this study, which suggest that image gamuts are not the reason for differences in GMA performances for different images, might seem to contradict earlier work, which has shown that using image gamuts as a parameter for gamut mapping results in more accurate reproductions,[1,4] this disagreement is only apparent. What the earlier papers have shown is merely that for a given GMA compression ratios that are determined on the basis of image gamuts rather than the original medium gamut give more accurate results. The principal claim of this study, on the other hand, is that the differences between how different GMAs perform for different images are not due to the image gamuts of these.

In summary, the combined conclusions of previous work and this paper are that if image gamuts are used to determine compression ratios, a GMA's performance improves but that it is not the image gamut characteristic that makes GMAs perform differently for different images. What this means is that knowing an image's gamut is not sufficient for determining how best to reproduce it but that this knowledge can help to improve the performance of a given gamut compression algorithm used for reproducing it.

Conclusions

Two methods for evaluating the importance of image gamuts on colour image reproduction were described in this paper. The first results in two sets of originals which are based on the same set of images and where the images in one set differ in gamut and the images in the other set do not. The second method perturbs the colour gamut of a single image by various degrees in terms of lightness and chroma. Reproductions of all these original images are then made using four different GMAs and their accuracy is then evaluated using a psychophysical technique. The ranges of accuracy score ranges obtained in this way for the different original image sets then indicate the importance of the image gamut image characteristic. The overall results of this study show that the image gamut characteristic has no significant effect on the performance of GMAs. Since the results were image dependent, other image characteristics (e.g., colour histogram, local colour contrast) have to be evaluated next.

Acknowledgements

The authors would like to thank Tadashi Yoshida and Todd Newman (Canon R&D Center Americas, Inc.) for supporting this study.

References

1. R. S. Gentile, E. Walowitt and J. P. Allebach, A comparison of techniques for color gamut mismatch compensation, *Journal of Imaging Technology*, **16**, pg.176–181, (1990).
2. L. W. MacDonald and J. Morovic, Assessing the Effects of Gamut Compression in the Reproduction of Fine Art Paintings, *Proc. IS&T/SID 3rd Color Imaging Conf.*, pg. 194-200, (1995).
3. E. D. Montag and M. D. Fairchild, Psychophysical Evaluation of Gamut Mapping Techniques Using Simple Rendered Images and Artificial Gamut Boundaries, *IEEE Trans. Image Proc.,* pg. 977- (1997).
4. R. Y. C. Wei, M. J. Shyu and P. L. Sun, A New Gamut Mapping Approach Involving Lightness, Chroma and Hue Adjustment, *TAGA Proc.*, pg. 685–702, (1997).
5. J. Morovic and P. L. Sun, Methods for Investigating the Influence of Image Characteristics on Gamut Mapping, *Proc. IS&T/SID 7th Color Imaging Conf.*, pg. 138–143, (1999).
6. P. Barten, MTF, CSF and SQRI for Image Quality Analysis, *SPIE Tutorial*, San Jose, pg. 34–44, (1997).
7. X. M. Zhang and B. A. Wandell, A Spatial Extension to CIELAB for Digital Color Image Reproduction, *SID Symposium Proc.*, pg. 731-734, (1996).
8. G. J. Braun and M. D. Fairchild, Gamut Mapping for Pictorial Images, *TAGA Proc.*, pg. 645–660, (1999).
9. J. Morovic and M. R. Luo, Developing Algorithms for Universal Colour Gamut Mapping, *Colour Imaging - Vision and Technology*, John Wiley & Sons Ltd., pg. 253-284, (1999).
10. H .S. Chen, O. Minoru and K. Hiroaki, Gamut Mapping Method Adaptive to Hue-divided Color Distribution, *Proc. IS&T/SID 7th Color Imaging Conf.*, pg. 289-294, (1999).
11. C. J. Li, M. R. Luo and R. W. G. Hunt, The CAM97s2 Model, *Proc. IS&T/SID 7th Color Imaging Conf.*, pg. 262–263, (1999).
12. L. L. Thurstone, A Law of Comparative Judgement, *Psych. Review*, pg. 273–289, (1927).
13. R. S. Berns, Methods for Characterising CRT displays, *Displays. Special Issue: 'To Achieve WYSIWYG Colour'*, pg. 173–182, (1996).
14. P. C. Hung, Colorimetric Calibration in Electronic Imaging Devices Using a Look–Up–Table Model and Interpolations, *Journal of Electronic Imaging*, **2**/1, pg. 53–61, (1993).
15. N. Katoh and M. Ito, Gamut Mapping for Computer Generated Images (II), *Proc. IS&T/SID 4th Color Imag. Conf*, pg. 126–129, (1996).
16. G. J. Braun and M. D. Fairchild, Image Lightness Rescaling Using Sigmoidal Contrast Enhancement Functions, *Journal of Electronic Imaging*, vol. 8, no. 4, pg. 380–393, (1999).