DEVICE INDEPENDENT COLOUR . . . IS IT REAL?

Tony Johnson*

Abstract: The requirements for device independent colour are discussed and the problems of achieving successful reproduction by this means are reviewed. Additional experimental work to that described at previous TAGA conferences, to enable the modelling of appearance, is described. A straightforward, but reasonably effective, procedure to enable colour reproduction using CIELAB encoding is described and some of the errors analysed. This work was based on tools developed by the ANSI IT8 standards committee.

Introduction

As images are transmitted between computer systems with increasing frequency it is desirable that they can be interpreted unambiguously by the receiving system with the minimum of effort. In order to achieve this certain conventions have evolved to describe images, pages and documents which can be understood by any system which acknowledges them. The Postscript language from Adobe has become the best known computer language for describing documents in such a way. Because it defines the attributes of a page without specific reference to any device used to render it the language is referred to as being device independent. When the page is sent to a device for rendering the attributes are interpreted by the RIP (Raster Image Processor) associated with it so that the page may be printed.

Clearly describing the colour of a pixel is an important attribute which is needed to render it; hence the term device independent colour. It refers to a method for encoding data which permits it to be rendered in the correct colour by any device receiving it. The question I wish to address in this

*Crosfield Electronics Ltd.

paper is what constraints are necessary to make such a system feasible at the present time and what developments are needed to make it more general. I will then review some work we are undertaking which is relevant to this.

What is Device Independent Colour?

The question above may seem a little bizarre. The answer seems obvious. A definition of colour which takes account of human perception must achieve what is required and such a system was defined in 1931 by the CIE. Therefore any encoding system derived from this should provide a suitable definition. For a colour reproduction system in which each pixel of an image is measured according to the CIE method, and is made to produce the same tristimulus values when reproduced, it may be expected that a colour match will result. What will have been achieved may be described as colorimetric equivalence. However, such a system has somewhat limited application.

The CIE system was originally defined to measure the visual equivalence of coloured lights. It was later applied to the measurement of colorant systems but was primarily used for comparing the colour of two pigment, or dye, mixtures on similar media. Thus, it proved very efficacious for comparing, for example, two paints applied to a similar surface. It can provide an equally successful metric for colour reproduction if we constrain our system to meet the same limitations.

However, if we cannot meet these constraints with our system colorimetric equivalence does not, in my view, provide a device independent colour system. It is my contention that what is really required for device independent colour is that any image stored in a system should look the same when rendered on a variety of devices. Colorimetric equivalence does not guarantee this, not because of any inaccuracy of the CIE system but simply from endeavouring to use it for purposes for which it was not designed. (There are also limitations in the way in which the data is generally obtained).

There are five issues which can create difficulties and it is only if these are avoided that colorimetric equivalence can prove acceptable.

These issues are:

a) White point equivalence
b) Gamut mapping
c) Measurement geometry
d) Media differences
e) Colorant quantification

In other words a reproduction will only look the same as the original image if both have the same tristimulus values for the white point; the reproduction media can actually produce <u>all</u> the colours required and the amount of each colorant can be computed accurately, and the two media have similar surface characteristics and are seen in similar viewing conditions. In general, for Graphic Arts, this is only achieved when a printed image is itself being reproduced, a relatively rare event.

Of course, some of these issues do lend themselves to reasonable engineering solutions; it is simply a question of implementing them. They have been discussed in earlier TAGA papers; Johnson (1989 and 1992a) and Johnson and Luo (1990), so do not require a great deal of elaboration here. However, for the convenience of the reader they are reviewed in a little more detail below together with the solutions required where they are known.

a) White point equivalence:- If the original has a brighter white than the reprodution can achieve how should it be reproduced? In general, the best solution is to reproduce lightness <u>relative to the white point for each</u>. In other words both sets of image data are normalised to their respective white points. This procedure will also accommodate small differences in colour between the white points. It is only when the differences become quite large that this technique fails.

b) Gamut differences:- This is the most difficult problem of those listed. Whilst certain algorithms have been proposed (see, for example, Johnson (1982)) they clearly do not work for all images. The suspicion still exists that optimum gamut compression is image dependent particularly when the 'shape' of the gamuts is radically different.

c) Measurement geometry:- This becomes a significant issue when the surface characteristics of the images are different. Most colorimeters and spectrophotometers have measurement geometries which are chosen quite deliberately to either eliminate or totally include first surface reflections. Unfortunately, typical viewing conditions cannot achieve either and so colours which look alike in a practical viewing condition may measure differently. This is likely to be most significant for dark colours. The only practical solution is to use telecolorimetry for the measurement. By siting the measuring device in the same position as the observer such issues are avoided. Unfortunately, these conditions are not then easily reproducible elsewhere unless viewing conditions are well controlled. A standard such as ISO 3664 should be specified.

d) Media differences:- Quite apart from issues of gloss discussed in the earlier paragraph another restriction of colorimetric equivalence is that the CIE system has limited capability to take account of changes in appearance which arise from a change in viewing conditions. It is an excellent tool for defining the similarity of appearance of two stimuli which have similar surface characteristics and are seen in the same conditions of viewing. However, for the same stimuli seen in different conditions of viewing it is less successful. Whilst it will tell us that two metameric samples will no longer match as the illuminant changes it does not describe what the colours look like. This means that we cannot describe what happens to the colour or the match as surround conditions change. Even isomeric samples will look quite different as these are altered for one of the samples but the CIE system cannot predict this. Colour appearance models are being developed to resolve this problem.

e) Colorant quantification:- Calculating the amount of ink required to reproduce specific tristimulus values is not a simple task. Two approaches are feasible. One is to measure a small number of colours and model the process, the alternative is to measure a larger number of colours and use a look-up table with interpolation to compute the colorant quantities. Essentially it is a question of trading off accuracy against measurement time. This becomes particularly significant if the output process is inconsistent.

It is clear from the above discussion that the doctrine of colorimetric equivalence can be freely applied to images seen in the same conditions of viewing if telecolorimetry is used to avoid gloss issues, the gamuts match and sufficient attention is paid to the colorant calculation. Furthermore, it can be extended objectively by using appearance modelling and white point normalisation so that, although colorimetric equivalence no longer holds, the hue, chroma and lightness values encoded are uniquely related, for specified viewing conditions, to the tristimulus values measured with the telecolorimeter. By ensuring that hue, chroma and lightness are transformed into the tristimulus values appropriate to the output viewing condition the colorant amounts required may then be calculated.

However, that still leaves one unresolved issue. how do we achieve gamut compression (or expansion)? Even if an image is encoded by it's appearance parameters, unless we have a standard method for gamut mapping that image could still render differently, even on two devices with the same gamut, when compression or expansion is required. As already suggested this really does not meet the user's requirements of device independent colour. Clearly, if the encoding and device gamuts are very different the likelihood of such differences arising are very much greater. I believe this can be minimised by agreeing a gamut in which all images should be encoded which is intermediate to that of transparency materials and real surface colours published by Pointer (1980) and the really low gamut devices such as newspaper presses. A gamut such as that achieved by printing to the FOGRA specification for coated paper, Rech et al (1981), would appear to achieve such a goal. This means that much of the compression takes place at the time of encoding which has the advantage that image interpretation is undertaken primarily by the originator of the image and far less by the various renderers, each of which may interpret it differently.

Thus, the answer to the question posed in the title of this paper - Is device independent colour real? - is 'yes, but . .'. Clearly if unreasonable constraints are imposed simple colorimetric equivalence (using a definition such as CIELAB) provides a satisfactory method. However, for the general case this needs to be extended to appearance equivalence, based on telecolorimetry, if we require images to look the same. When matching an encoded image across devices agreement on gamut mapping is required and this seems the most intractable problem at present. It is suggested that the problem may be minimised by agreeing an interchange gamut based on such a device as a printing press meeting the requirements of the FOGRA specification for coated papers.

Using such a procedure device independent colour is almost a reality. Unfortunately, there is no general agreement among users to work in this way. Most seem to want to obtain simple colorimetric equivalence (possibly with a white point correction). One of the exceptions is Xerox who provide for what are known as appearance hints in their encoding, Buckley (1990). Until there is such agreement I have few doubts that defining colorant amounts is a preferable way of encoding colours since fairly efficacious algorithms have evolved for overcoming all the issues above empirically. This has been discussed at some length by Johnson (1992a). It is a relatively simple task to characterise such data to enable it to be transformed into colorimetric encoding if required. This is the objective of the standardisation work of the ANSI IT8 committee now nearing completion.

Appearance modelling

In previous TAGA papers, Johnson (1989) and Johnson and Luo (1990), we discussed on-going work, which is being undertaken at Loughborough University with the support of Crosfield and Coats-Viyella plc, to evaluate colour appearance models. Since then Ronnie Luo has undertaken futher work which will be summarised below.

The 1989 paper showed that neither the Hunt or Nayatani models performed particularly well in predicting perceived lightness although by changing parametric constants in the Hunt model a reasonable prediction was obtained for both prints and colour monitors. However, this produced certain difficulties in that it conflicted with brightness data alredy published and Hunt, therefore, proposed a modification to the model to overcome this. This was described in the 1990 paper. Since then the information has been published in more detail by Luo et al (1991a and b) and Hunt has published a full revised model, incorporating these findings, Hunt (1991).

Since this data was published the work has been substantially extended. We have now evaluated transmissive samples, both in cut-sheet form and as projected 35mm transparancies, as well as extending the range of luminance levels used for evaluating reflection copy. Apart from further experiments to investigate the effects of simultaneous contrast, with both chromatic and achromatic surrounds, the experimental work related to appearance determination for this project is now completed.

Some of the results obtained in these studies are summarised below. It has to be emphasised that these results are somewhat preliminary. There is a substantial amount of data to analyse and much of this has yet to be done. Nevertheless, there are already some interesting findings which are worthy of discussion at this time although we do anticipate additional revelations as the analysis proceeds. Hopefully, a more comprehensive review will be presented at a future conference.

a) Cut-sheet Transparencies:- As for our previous studies an arrangement of colours was used in the experiments to simulate a complex viewing pattern. The field consisted of a background area, of variable lightness, which surrounded the sample being evaluated. Close to it were two reference samples (one white and one coloured) which had been given attributes at the outset of the experiment. The white was designated a lightness of 100 and the subject chose a value for colourfulness of the coloured sample which was later used for normalising the data. (This is necessary since colourfulness is an open-ended scale). Finally, some additional colours were included, at the outer part of the field of view, to simulate the complex field. The pattern is shown in figure 1.

The subject was presented with each of the colours under test sequentially and asked to specify the hue, colourfulness and lightness of the sample. This procedure was then repeated under different conditions of viewing with a random order to the samples. Finally, the tristimulus values of the sample were obtained from measurements made with a telespectroradiometer.

The experimental conditions used for each phase of the experiment are listed in Table 1. It can be seen that the whole image was surrounded by both white and black surrounds as well as the sample being backed by two levels of grey background. These were selected to try and provide the same perceived lightness under the two different viewing conditions. (Cross-over experiments were then carried out to separate the effect of surround from background). The experiments with frontal flare falling on the transparency were to investigate whether such a viewing condition provided significantly different perceptual attributes to those obtained with the viewing condition currently specified in ISO 3664.

Comparison of the different viewing conditions leads to the following conclusions:

a) As the luminance level increases colours appear more colourful and lighter. The latter effect is more noticeable for darker colours. An example is shown in figure 2.
b) As the background is made darker similar effects occur. See figure 3.
c) As the surround is changed from light to dark similar effects occur for lightness. See figure 4.
d) Adding veiling flare, as opposed to a white surround, has the effect of reducing the lightness of dark colours. No significant effect on colourfulness is seen. An example is given in figure 5.

When evaluating the performance of the various models we obtained the data shown in table 2 which are clearly not encouraging. They show far poorer correlation and scatter than any of our earlier studies, particularly for the colourfulness predictions of the appearance models. Even the reasonable fit to the Hunt91 lightness data was only obtained by significantly deviating from values for parameters in the model which are recommended by Hunt for these viewing conditions. (The z factor in the computation of lightness was arbitrarily reduced from 1.45 to 1.0 for the white surround and 0.85 for the dark). Further work will be undertaken to develop the proper alteration to the model to accommodate such a change.

When considering colourfulness some other modifications to the models are clearly required. One of the changes which has been attempted is to investigate the chroma function and this provided a much better model of the perceptual data. The results of this can be seen in table 3. A similar result will be found in the next section. This suggests that the colourfulness model may have some limitations. We intend to review our earlier data to see whether any similar effects exist.

It seems clear that cut-sheet transparencies require some modifications to the parameters of the model proposed by Hunt to provide a better fit to the data. The next phase of experimental work should show whether this is a general requirement for transparency viewing conditions or is specific to these.

b) 35mm transparencies:- For this study the subject was asked to scale hue, colourfulness and lightness, for samples projected by a normal 35mm slide projector, onto a diffuse white screen. The sample was seen in a pattern similar to that of figure 1, with a mid-grey background, and subtended an angle of approximately 1 degree at a viewing distance of about 12 feet. Each of six observers scaled 99 samples for each of four phases.

Two spectral power distributions were chosen for study; the normal tungsten halogen source and a filtered version simulating xenon. The halogen source was viewed with luminance levels of the reference white of 113 and 46 cd/m^2 (phases 1 and 3) and the filtered version had a luminance level of 47 cd/m^2 (phase 2). The high illumination level was repeated to investigate consistency of the data (phase 4).

One thing became very clear as this experiment progressed; there was a phenomenon present which had not occurred to the same extent in any of our previous studies. This can be seen in figure 6 which shows lightness predictions against those of L* for all 4 experimental phases. All models (except CMC) showed the same effect. It is clear that for lighter colours the models over-predict lightness. We thought that this may be due to having the reference samples too close to the test sample (they were closer than for previous studies to minimise the effect of screen/gate non-uniformity) so the experiment using the high illumination level was repeated with a distance similar to that for previous experiments. However, the effect was unaltered. We finally repeated the experiment with a sample matching the reference white being shown, randomly, three times during an experiment. Since this was always scaled with the same value as the reference sample it showed that the effect was not due to any non-uniformity of the field.

Table 4 shows the results of evaluating the models using this experimental data and they are not encouraging. They show far poorer correlation and scatter than any of our earlier studies, even considering the cut-sheet experiment. Some changes have been made to the parameters of the Hunt model and the improvements from these can be seen in table 5. Removing the Helson-Judd factor clearly improves the hue predictions which suggests that despite the low colour temperature of the source adaptation is effectively complete. Neutrals are seen as neutral. (This contrasts with normal viewing conditions using such a source in which neutrals appear somewhat chromatic). This is probably due to the wholly darkened viewing condition with no clues as to the source, apart from the image. However, it has no significant effect on colourfulness but like the cut-sheet experiment the perceptual data is better modelled by the chroma function than the colourfulness one.

It is clear from figure 6 that some improvement can be made to the 'fit' of the lightness data by simply modifying the model predictions by a function of the form y=mx+c. Doing this to the Hunt model virtually halved the CV values. However, it looks as though a power function would be even better and so we are reviewing where such a change would be appropriate to the model without causing detrimental consequences to our data for other viewing conditions.

Table 4 shows that the Nayatani model performs very badly for this viewing condition. The Hunt model and CIELAB perform very similarly for chroma and lightness but the performance of the Hunt model is far poorer than for any viewing conditions studied previously. The hue predictions from the Hunt model are considered good.

In the light of these results it is interesting to look again at the cut-sheet experiment results. The two sets of data show certain similarities. Both are better modelled by chroma than colourfulness and both have good agreement with the hue prediction obtained from the models. Both give very poor lightness prediction using the values for the parameters recommended by Hunt. However, what is interesting is that if figures 6 and 7 are compared it can be seen that both show a degree of similarity in how they deviate from the model. We need to work on defining the parameters for this condition.

c) Reflection copy:- In our previous experiments observers were only asked to scale lightness. Whilst this represents the more useful parameter for evaluating colour reproduction, when compared to brightness, it is nevertheless of interest to know what the effect on the absolute attribute is when viewing conditions are changed. Thus we have now gathered data to enable evaluation of the colour appearance models for brightness changes.

Six observers each scaled 40 samples, in a random order, twelve times. Each of the six luminance levels selected was used twice; on one occasion lightness, colourfulness and hue were scaled and on the other brightness, colourfulness and hue. The six luminance levels selected were chosen to cover a wide range. The measured values of the reference white were:

843.1, 200.3, 61.9, 16.6, 6.2 and 0.4 cd/m^2

The samples were placed in a complex field as shown in figure 1; the background was a mid-grey of approximately 50% lightness. For brightness assessments the reference white was removed; observers were asked to memorise a sample at the end of each phase and then, following adaptation to the new phase, estimate colourfulness and brightness for a different reference. This was then left in the viewing field for the remainder of that phase.

Analysis of the individual observers results showed one interesting anomaly. Some of them were not consistent in their estimates of colourfulness between the two phases of comparable luminance. This is probably due to the change in reference sample for each phase of the brightness estimation. It could suggest that memory is poor or the concept of colourfulness changes when associated with brightness. However, I believe it more likely to be associated with the high scatter we always find for this attribute. The initial selection of the reference sample is therefore somewhat uncertain. For any phase the relationship between the test and reference colours was reasonably consistent; essentially the difference between phases of the same luminance level was a scaling factor for colourfulness but with no consistent value across observers.

Figure 8 and table 6 show the effect of luminance level on brightness, colourfulness and hue. Despite two anomalies, it may be seen that colourfulness and brightness both increase with luminance level as expected. It is also interesting to note how much the scatter of the data increases with decreasing luminance level, particularly for the colourfulness data. It should also be noted how consistent the hue data remains with very substantial changes in illumination level.

Figure 9 and table 7 show the effect of luminance level on lightness, colourfulness and hue. (All data for this table is only for the phases in which lightness was assessed).

When comparing tables 6 and 7 there are some striking similarities and differences for the colourfulness and hue data. (Both should, of course, be identical). The consistency of the hue data is particularly noticeable. The colourfulness data shows the same trend in both cases in that it increases substantially with luminance level but the gradients are rather different in some instances. This stems from the single observer uncertainty discussed earlier. Again the increasing scatter as luminance level falls is marked.

For lightness there is an unexpected change in gradient for the highest and lowest levels. This is probably explained by being in the mesopic region in one case and nearing saturation in the other.

When evaluating appearance models we again confirmed that the Hunt91 model provided the best predictions of appearance. This is shown in table 8. (Note that for colourfulness and hue the data used was obtained by combining that from the two phases). It clearly outperforms all of the other models for predicting lightness. Furthermore, it does so for all luminance levels, except the highest where L* and Nayatani compare. For colourfulness it performs as well as any other but no better. (However, it should be remembered that the others do not account for the change that occurs as surround conditions change). For hue and brightness, where only the Hunt and Nayatani models are relevant, the former is far better predicted by Hunt. For the latter there is nothing to choose between them.

Overall our experiments over the past few years have enabled substantial improvements to be made to the Hunt model. We also anticipate that with more analysis it may be possible to improve it further. We believe that even now it provides a useful model to assist in quantifying colour reproduction across different media and have demonstrated this to our own satisfaction. The next stage is to incorporate it into products but to achieve that we need to resolve the gamut compression and measurement issues described earlier. This will be the objective of our future research. In the meantime we are working in a consortium to demonstrate the feasibility of using CIE data encoding in a publishing environment. To achieve this, in a practical system, we have made a number of simplifications which will be briefly described in the next section.

L*a*b* data exchange

In 1989 a consortium was put together to demonstrate the viability of a Distributed Integrated Multi-media Publishing Environment (DIMPE). This required agreement on an encoding format for communicating images. For this part of the work the relevant members of the consortium were Crosfield, Linotype-Hell, Scitex, Burda and Maxwell Communications Corporation.

After much debate it was decided that we should use CIELAB as the encoding domain, quantised at 8 bits per channel. Some early studies demonstrated that it was a fairly marginal level of precision but would probably suffice if no significant colour changes were made to encoded images. This decision is bolstered if only prints and not separations are evaluated since most artefacts are only visible on separations. (This is primarily due to 'dot gain'- both geometric and optical-and the more limited dynamic range when viewing prints).

It was also decided that we should use the tools for colour calibration currently under development by the ANSI IT8.SC4 working group. The first of these is a colour chart (based on the Kodak Q60) containing 250 colours. This will be manufactured on the main colour film and paper products available from the major photographic materials suppliers and may be supplied with colorimetric data (including CIELAB) if required. Characterising printing processes is achieved by a procedure whereby the values of combinations of halftone dot percentages are defined and after printing may be measured colorimetrically. This enables the conversion between CMYK and L*a*b* to be computed.

The Crosfield role in this work was to scan the images using a conventional drum scanner but with the colour converter loaded to obtain CIELAB data directly. This was then converted into the CMYK values required for reproduction by Burda and Maxwell Communications. In addition we output the images on a range of proofing devices. 6 images were selected for demonstration purposes; the Kodak Q60 (since the final IT8 target was not available) and a series of vignettes generated by Linotype-Hell plus 4 standard Crosfield images.

In order to simplify the process we decided to use the approach advocated at the beginning of this paper. We captured the data in gamut compressed and appearance rectified form. This was easy to achieve. We simply took the Q60, scanned it using a conventional scanner set-up and made a proof. (This, of course, has already achieved the gamut compression and appearance matching). We then measured the tristimulus values of the Q60 proof and used this data to define the mapping between RGB densities and proof L*a*b*. With such a limited number of colours as the Q60 we decided to develop a modelling procedure for the transformation using polynomial functions as described by Clapper (1961). A look-up table with interpolation was a possibility but the modelling approach was somewhat simpler.

For the output we had a greater choice since we used the data file defined by ANSI IT8.7/3 to produce a proof on each of the devices. We therefore had the choice of setting up a 6x6x6 matrix and adding a black calculation based on the remaining data of the IT8 file or modelling the process from a smaller number of colours. For simplicity we decided to model the process from the limited number of colours in the default set of the IT8.7/3 data and used the same sort of procedure as for the input conversion, with a black calculation similar to that described by Johnson (1985).

Thus, the input polynomial was based upon the 236 colours of the Q60 and the output polynomial on the 107 colours of the IT8 basic data file which contain no black. For our black algorithm we only use a small number of the remaining colours; those necessary to compute additivity failure and relative lightness of black to grey. In principle we can do this from 5 grey patches, 5 black patches and 1 4-colour overprint. However, in practice we use 3 to 5 4-colour overprints to compute the additivity corrections.

The results of this exercise were assessed visually and considered to be quite acceptable. Certain differences were noted but these were only of the order we expect to see between different scans of the same image undertaken by more than one operator. In an attempt to quantify this the Q60 proof made by this route (i.e. RGB to L*a*b* to CMYK) was measured and compared to the original which had been made by regular scanning and which had been used to set the system up. The average DeltaE(ab) and variance was computed for various stages of the process as shown in table 9.

The first line in this table compares the L*a*b* data for the proof used to determine the input polynomial with that predicted by the polynomial itself. It is thus a measure of the error of the model. The next line compares scanner L*a*b* values predicted by the input model with results measured on the proof. It may be thought of as a measure of the accuracy of the output model though it is more complex than that. (We shall return to this point). The third line compares the measurements of the final proof with those used to set up the system.

The overall average error is just over 5 DE(ab) units for the 236 colours. To many colour physicists such an error would seem fairly horrendous but in fact for complex images is not unusual. As a comparison we had the same original scanned again as a conventional CMYK image by another scanner operator using his own data set, which had been defined for the same proofing system. The average error when comparing the two CMYK scans is significantly greater, at 7.25, as shown in line 4 of table 9. Whilst I would not claim this error would be typical of a large number of scans of the same image the value of 5 obtained in the L*a*b* scanning would not surprise me.

In order to indicate where the error is arising it is necessary to consider the other lines of table 9. We can see that errors comparable to the overall error are obtained at each of the input and output stages separately (lines 1 and 2). One interpretation of this would be that there is a reasonable cancellation of error from one stage to another as we might expect given that the error is not signed. The conclusion would be that neither input nor output model is particularly accurate but overall the two combine in a reasonable way. However, such a conclusion is probably erroneous.

The last line of table 9 shows some values obtained in an earlier study, Johnson (1992b). It shows two DeltaE values obtained by the same sort of technique; the smaller was obtained by using 729 colours to set up the polynomial and the larger value by using only 27. The average values are those obtained for all 729 colours. It is clear from this that much smaller errors should be expected on output than is suggested by line 2. How do we reconcile this?

Since the low error was obtained with a test set that only included cyan, magenta and yellow one possible source of error could have been colours with black. However, this was quickly ruled out. By assessing the data obtained for columns 3, 4, 5 and 8 of the Q60, together with the top 6 steps of the grey scale, we obtained an average DeltaE of 3.92. (The average values for each column were 5.07, 4.42, 2.95, 4.03 respectively and 2.3 for the grey steps).

When looking at the data in detail it soon became apparent where the biggest errors were. Yet again the culprit was gamut compression! When comparing the average errors for the cyan, magenta and blue scales of the Q60 (columns 12, 13 and 18), between the encoded CIELAB values from the scanner and the resultant colours on the proof, they were 15.4, 10.61 and 11.0 respectively. The biggest errors arose in the middle to solid tone values. Obviously the blue, cyan and magenta colours in the rest of the Q60 were similarly affected. Similar values were obtained when comparing the encoded L*a*b* values against those on which the process was modelled. However, when comparing the original CMYK proof against that made via L*a*b* these had reduced somewhat, particularly for magenta and blue, to 12.37, 7.95 and 3.71 respectively.

Visual comparison showed the cause. The proof which had been used to set up the RGB to L*a*b* conversion had very high contrast in these scales. In fact they were 'clipping' in some of the steps. In setting up for this reproduction the operator had sacrificed differentiation in the saturated colours to produce colourful mid-tone colours.The relatively abrupt changes of slope had not been well modelled by the input transformation. This had effectively pushed many of the more saturated cyan, magenta and blue colours out of gamut. However, the output transformation had brought them back again to produce overall results which were reasonably acceptable. The colours which were in gamut were reproduced quite well by the output transformation; the major errors were largely introduced by the input sending colours out of gamut.

Clearly there are various ways of improving this. More complex functions or more colours with interpolation would do so. However, I am somewhat doubtful about the usefulness of this. Since so many images are gamut compressed, on an image by image basis, a good approximation is all that is required as a starting point. In that context the relatively straightforward procedure described here is quite effective.

Conclusions

It is suggested that device independent colour requires colours rendered on different devices to look the same, not simply exhibit colorimetric equivalence. This can only currently be achieved by imposing constraints on the colour reproduction process which are not always attainable. However, some work to assess and improve current techniques for appearance modelling is summarised which can go part of the way to improving this. A relatively straightforward procedure can be used for computing gamut compressed CIELAB data on a scanner, and subsequently rendering it for printing, which only requires a limited amount of measurement. The errors obtained using such a system are shown to be comparable to those often obtained by current practice. For many applications such a procedure is quite adequate and could possibly be based on even fewer samples.

Acknowledgements

I would like to acknowledge my colleagues at Crosfield who did all the work which enabled me to write this paper and will, hopefully, subsequently publish some of it in more detail. Ronnie Luo undertook all the colour scaling and apperance modelling evaluation, Jacquie Deane produced the colour transformation algorithms and software for the L*a*b* conversions and Melanie Barnes did much of the colour measurement work.

References

Buckley, R. R.
1990 "Specifying colour in office documents," TAGA proc. p30

Hunt, R. W. G.
1991 "Revised colour appearance model for related and unrelated colours," Col. Res. Appl., 16, p146

Johnson, A. J. J.
1982 "Defining optimum photomechanical colour reproduction," Pira report PR170
1985 "Polychromatic colour removal - Evolution or revolution?" TAGA proc. p135
1989 "Defining optimum photomechanical colour reproduction," TAGA proc. p350
1992a "Techniques for reproducing images in different media: Advantages and disadvantages of current methods" ISCC/ TAGA conf. proc. (to be published)
1992b "Calibration of colour pre-press systems," SPIE/IS&T conf. proc. (to be published)

Johnson, A. J. J. and Luo, R.
1990 "Colour appearance modelling," TAGA proc. p144

Luo, R., Clarke, A. A., Rhodes, P. A., Schappo, A., Scrivener, S. A. R. and Tait, C.
1991a "Quantifying colour appearance. Part 1. LUTCHI colour appearance data," Col. Res. Appl. 16, 3, p166
1991b "Quantifying colour appearance. Part 2. Testing colour models performance using LUTCHI colour appearance data," Col. Res. Appl. 16, 3, p181

Pointer, M. R.
1980 "The gamut of real surface colours" Col. Res. Appl., 5, 3, p145

Rech, H., Lospichl, R. v. and Werner, G.
1981 "Standardisation in the field of offset printing - instruction material for copying and printing," BVD/FOGRA, Wiesbarden/Munich

TABLE 1 - Viewing conditions for the cut-sheet transparency experiments

Phase	Illum.	Lightness of background	Luminance of ref. white	Surround	No. of colours	No. of observers	No. of estimates
1	D50	15.9	2259 cd/m^2	white	98	7	2058
2	D50	17.1	689	white	98	6	1766
3	D50	16.7	325	white	98	7	2058
4	D50	17.4	670 (+flare)	white	98	7	2058
5	D50	9.6	1954	black	98	8	2350
6	D50	9.5	619	black	98	8	2350
7	D50	9.8	319	black	98	8	2350
8	D50	9.4	642 (+flare)	white paper	98	8	2350
9	D50	9.6	658	white	98	7	2058
10	D50	17.5	680	black	98	7	2058

100

TABLE 3 - Coefficients of correlation and variation for chroma predictions of Hunt and Nayatani

		Experimental phase (See Table 1)									
Model		1	2	3	4	5	6	7	8	9	10
Nayatani											
Chroma	(r)	0.90	0.88	0.84	0.81	0.90	0.91	0.90	0.89	0.90	0.92
	(CV)	17	19	24	25	18	16	18	18	17	16
Hunt91											
Chroma	(r)	0.93	0.92	0.91	0.90	0.90	0.92	0.93	0.93	0.94	0.94
	(CV)	15	16	20	19	19	14	16	16	13	14

TABLE 2- Coefficients of correlation and variation for different colour models

Model		Experimental phase (See Table 1)									
		1	2	3	4	5	6	7	8	9	10
CMC											
Lightness	(r)	0.97	0.97	0.97	0.98	0.96	0.96	0.98	0.97	0.97	0.96
	(CV)	13	23	19	20	9	1	11	19	17	16
Chroma	(r)	0.86	0.88	0.88	0.94	0.87	0.86	0.90	0.90	0.92	0.85
	(CV)	20	19	21	15	21	17	19	18	15	21
CIE L*											
Lightness	(r)	0.96	0.96	0.97	0.96	0.95	0.96	0.97	0.96	0.96	0.96
	(CV)	22	16	16	18	28	22	23	18	19	21
L*a*b*											
Chroma	(r)	0.83	0.86	0.86	0.93	0.86	0.85	0.85	0.85	0.88	0.81
	(CV)	24	22	23	16	22	22	22	21	20	25
L*u*v*											
Chroma	(r)	0.79	0.81	0.84	0.83	0.76	0.79	0.80	0.86	0.84	0.80
	(CV)	27	26	24	25	30	27	27	23	24	21
Nayatani											
Lightness	(r)	0.95	0.96	0.97	0.96	0.95	0.96	0.96	0.96	0.97	0.96
	(CV)	22	16	16	18	28	22	23	17	18	21
Colourfu.	(r)	0.41	0.55	0.65	0.61	0.40	0.56	0.68	0.59	0.59	0.50
	(CV)	40	24	29	27	45	20	27	19	18	18
Hue	(r)	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	(CV)	8	7	9	8	8	9	10	9	8	7
Hunt91											
Lightness	(r)	0.96	0.96	0.97	0.97	0.96	0.96	0.97	0.97	0.97	0.97
	(CV)	10	10	9	12	10	9	8	11	10	12
Colourfu.	(r)	0.78	0.80	0.83	0.86	0.69	0.70	0.72	0.65	0.81	0.71
	(CV)	27	25	25	22	33	31	31	25	25	31
Hue	(r)	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	1.00	0.99
	(CV)	8	5	9	7	7	8	8	8	6	7

Table 4:- Comparison of models

Model	Experimental phase 1	2	3	4
CMC				
Lightness(r)	0.97	0.97	0.98	0.96
(CV)	37	41	40	32
Chroma (r)	0.89	0.92	0.93	0.92
(CV)	18	17	16	17
CIE				
Lightness(r)	0.95	0.95	0.96	0.95
(CV)	20	20	19	17
L*a*b*				
Chroma (r)	0.88	0.91	0.92	0.91
(CV)	19	18	17	17
L*u*v*				
Chroma (r)	0.85	0.91	0.86	0.87
(CV)	27	20	26	25
Nayatani				
Lightness(r)	0.95	0.95	0.95	0.95
(CV)	21	22	21	19
Chroma (r)	0.61	0.74	0.54	0.58
(CV)	34	28	35	35
Hue (r)	0.98	0.99	0.98	0.98
(CV)	12	11	12	12
Hunt91				
Lightness(r)	0.94	0.95	0.95	0.94
(CV)	20	19	19	18
Colourfu.(r)	0.76	0.79	0.80	0.79
(CV)	30	29	27	27
Hue (r)	0.98	0.99	0.98	0.98
(CV)	11	10	12	12

Table 5:-Modifications to the Hunt91 model

Hunt91-(No Helson-Judd correction)				
Colourfu.(r)	0.76	0.74	0.80	0.81
(CV)	31	33	29	28
Hue (r)	0.99	0.99	0.99	0.99
(CV)	7	7	7	7
Chroma (r)	0.91	0.89	0.93	0.92
(CV)	17	18	16	16

Table 6:-Scaled attributes relative to the highest luminance level (843.1 cd/m^2)

Luminance level	Brightness r	CV	m	Colourfulness r	CV	m	Hue r	CV
200.3	0.98	5	1.04	0.98	8	0.97	1.00	2
61.9	0.98	5	0.84	0.97	10	0.79	1.00	3
16.6	0.98	6	0.80	0.96	12	0.71	1.00	4
6.2	0.98	6	0.71	0.94	14	0.73	1.00	4
0.4	0.96	8	0.48	0.86	23	0.47	0.99	8

(where r=correlation coefficient, CV= coefficient of variation and m=gradient)

Table 7:-Scaled attributes relative to the highest luminance level (843.1 cd/m^2)

Luminance level	Lightness r	CV	m	Colourfulness r	CV	m	Hue r	CV
200.3	0.98	5	1.08	0.98	9	1.06	1.00	3
61.9	0.98	5	1.09	0.96	11	0.93	1.00	3
16.6	0.99	4	1.09	0.94	14	0.94	1.00	4
6.2	0.98	6	1.12	0.92	18	0.88	1.00	4
0.4	0.98	6	1.22	0.83	30	0.57	0.99	9

(where r=correlation coefficient, CV= coefficient of variation and m=gradient)

Table 9 - Analysis of errors in RGB to L*a*b* to CMYK conversion

Comparison	Mean DE(ab)	Variance
Original proof v scan L*a*b*	5.87	11.44
Scan L*a*b* v final proof	6.32	17.49
Original proof v final proof	5.26	11.97
Different CMYK proofs	7.25	25.78
Previous modelling experiments	1.56 to 1.87	

Table 8:- Comparisón of models

Model	Luminance level (cd/m^2) 843.1	200.3	61.9	16.6	6.2	0.4
CMC						
Light.(r)	0.95	0.96	0.96	0.98	0.97	0.94
(CV)	29	37	38	38	39	48
Chroma(r)	0.96	0.94	0.95	0.94	0.91	0.86
(CV)	18	20	18	20	25	36
CIE						
Light.(r)	0.95	0.97	0.96	0.98	0.97	0.94
(CV)	13	17	18	18	17	25
L*a*b*						
Chroma(r)	0.94	0.91	0.94	0.92	0.89	0.84
(CV)	19	23	20	21	27	34
L*u*v*						
Chroma(r)	0.93	0.92	0.93	0.91	0.93	0.88
(CV)	22	22	21	24	22	30
Nayatani						
Light.(r)	0.94	0.96	0.95	0.97	0.97	0.93
(CV)	15	20	21	21	19	27
Brigh.(r)	0.95	0.97	0.98	0.99	0.97	0.94
(CV)	9	11	11	11	12	17
Chroma(r)	0.94	0.93	0.90	0.85	0.83	0.73
(CV)	20	20	24	29	32	44
Hue (r)	0.99	0.99	0.98	0.95	0.95	0.94
(CV)	6	7	13	18	17	17
Hunt91						
Light.(r)	0.94	0.96	0.95	0.97	0.97	0.93
(CV)	14	12	13	11	12	16
Brigh.(r)	0.97	0.97	0.98	0.99	0.97	0.94
(CV)	7	11	11	10	12	17
Colou.(r)	0.95	0.92	0.95	0.93	0.91	0.69
(CV)	16	21	17	19	24	46
Hue (r)	1.00	0.99	0.99	0.99	0.99	0.97
(CV)	6	6	6	8	7	13

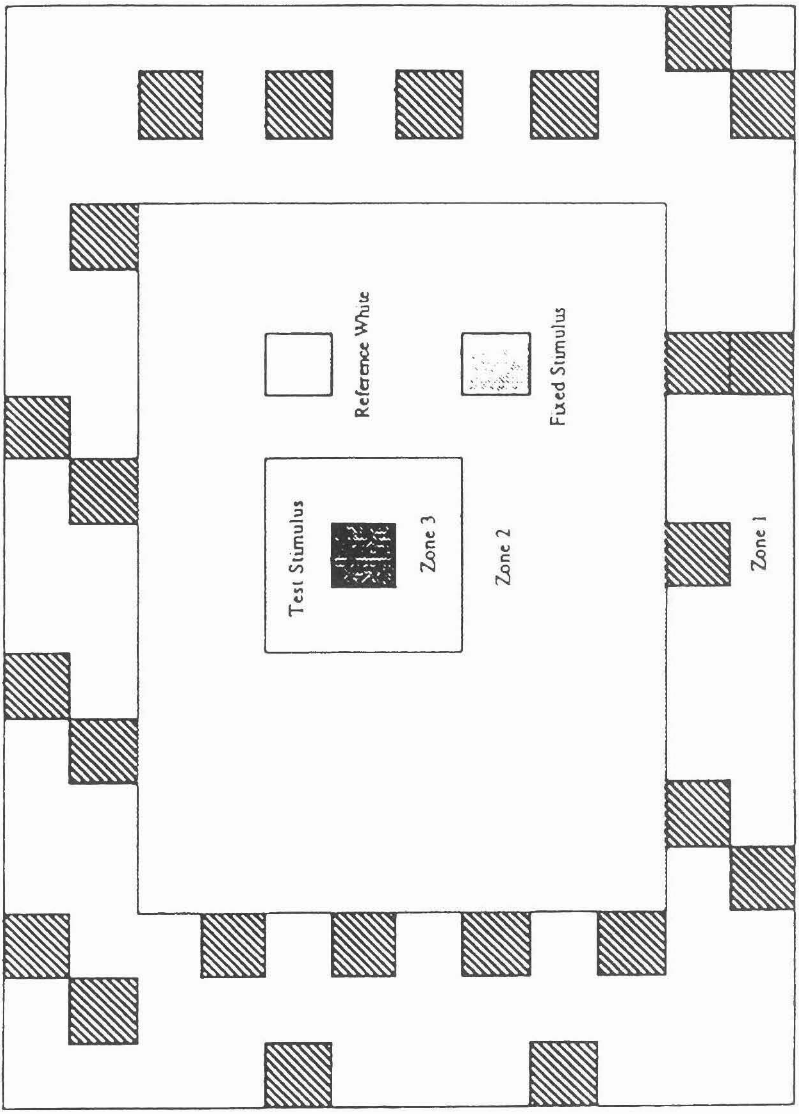


Figure 1 - Viewing Pattern

Figure 2 - Perceptual attributes at high level of
illumination as a function of low level

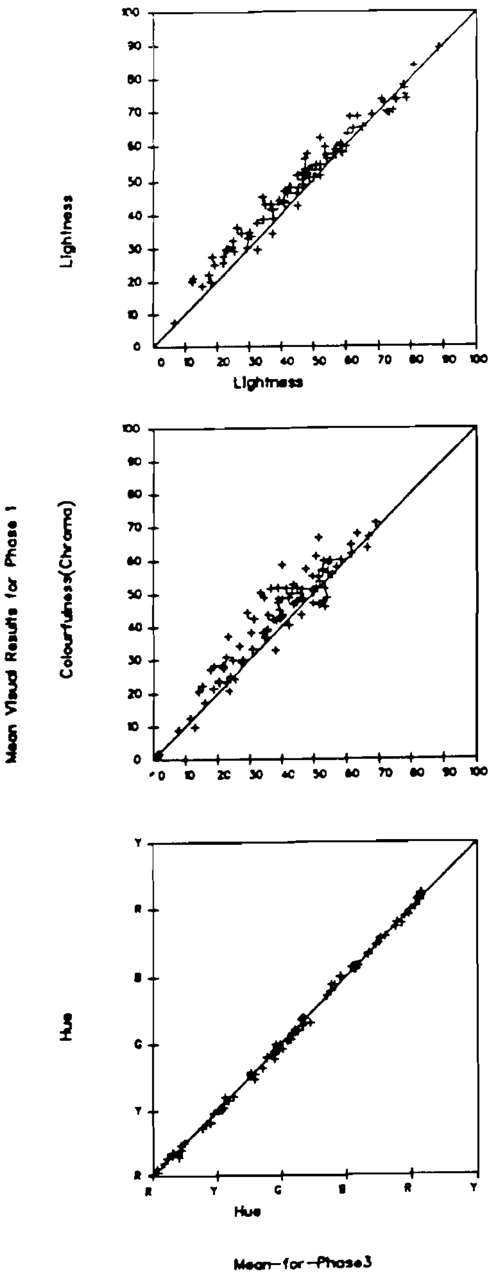


Figure 3-Perceptual attributes with dark background as a function of those with light

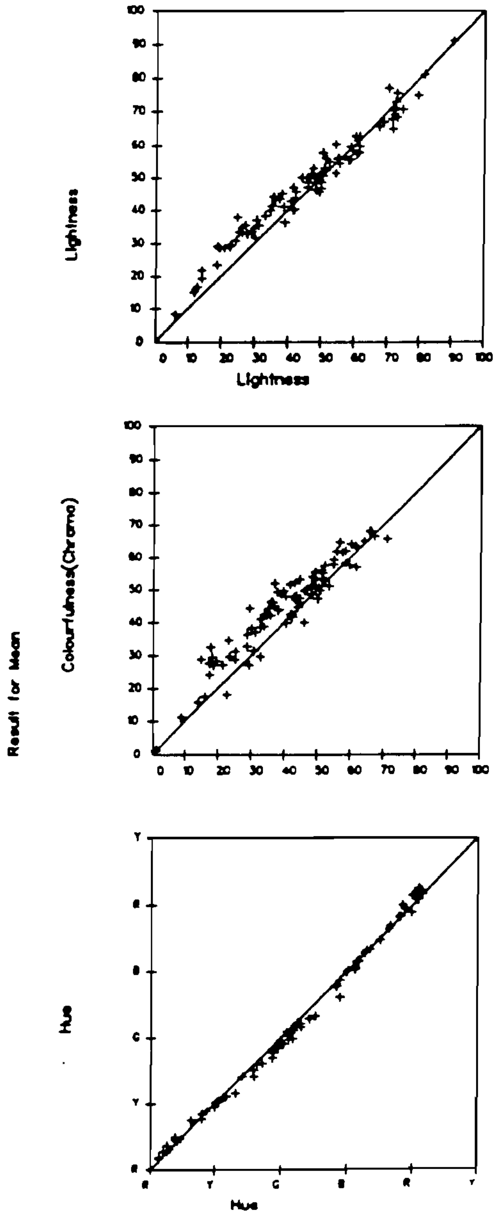


Figure 4-Perceptual attributes with black surround as a function of those with white

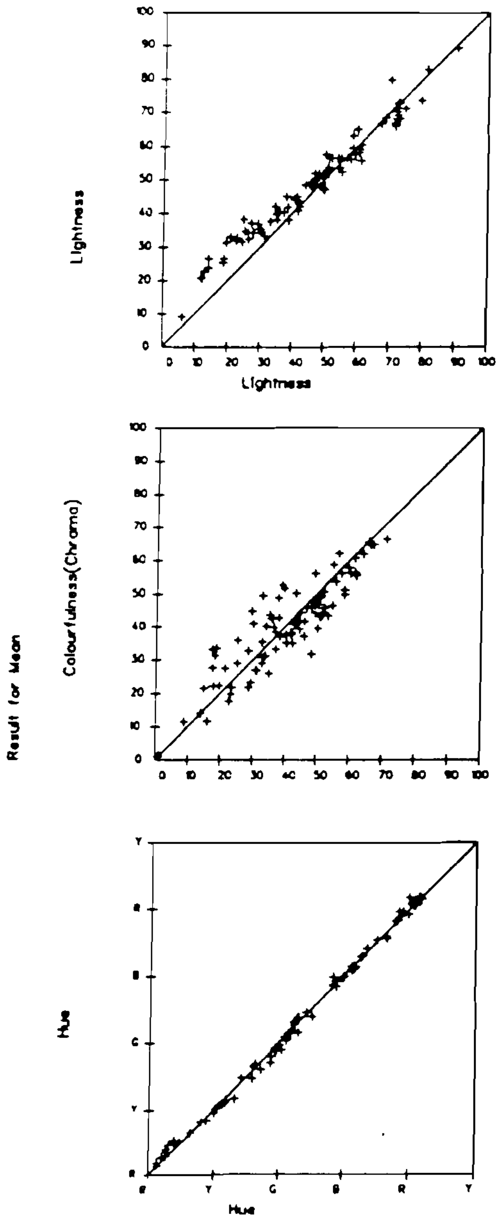


Figure 5-Perceptual attributes with veiling flare as a function of those with no flare

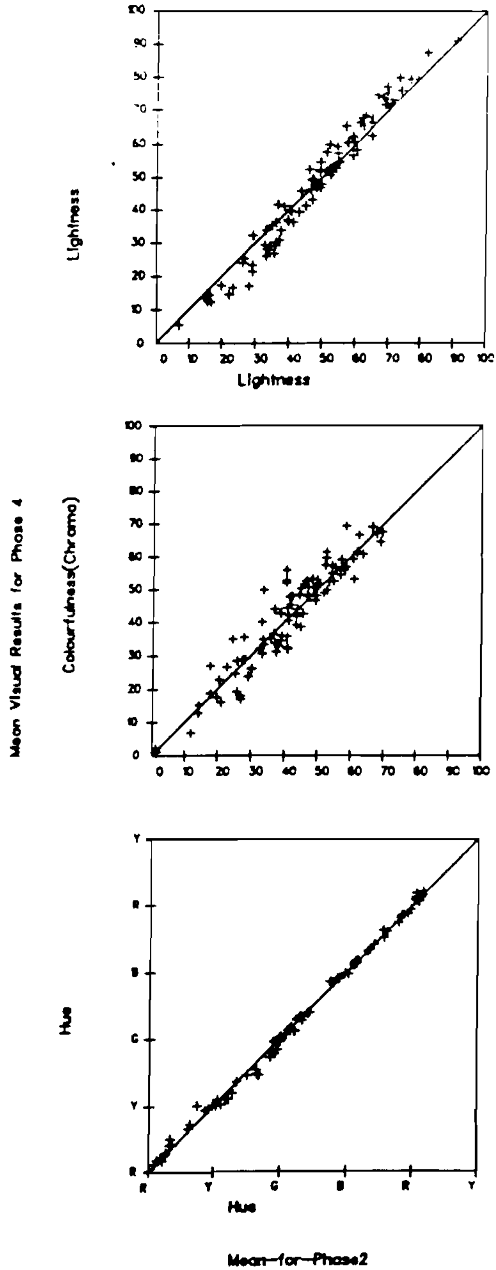


Figure 6-Perceived lightness as a function of L* predictions

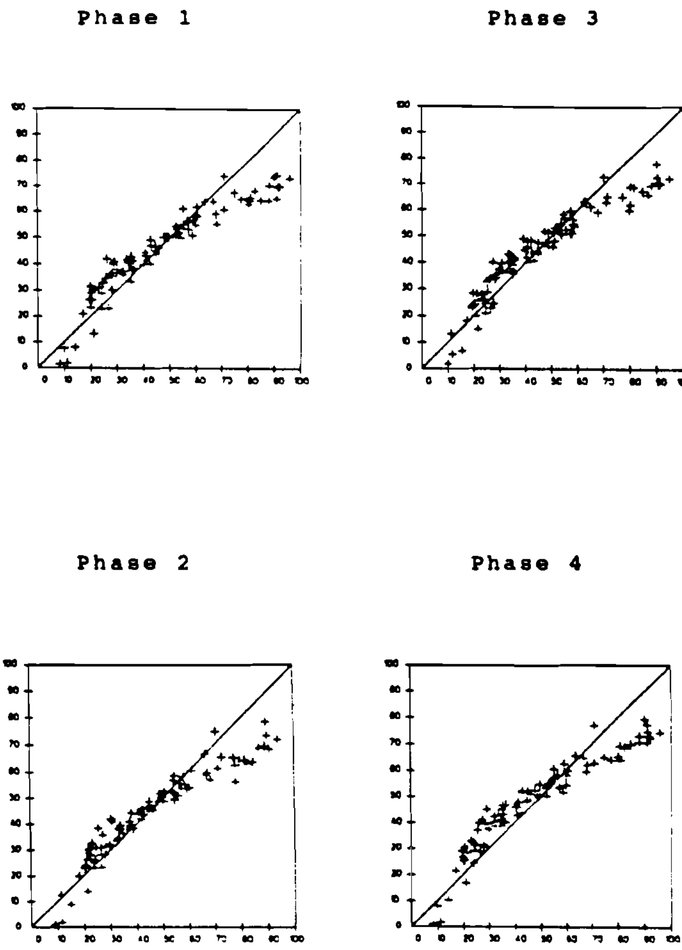
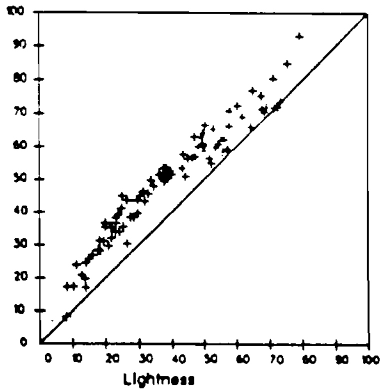


Figure 7-Perceived lightness as a function of Hunt91 predictions for cut sheet transparencies (phases 7/8)



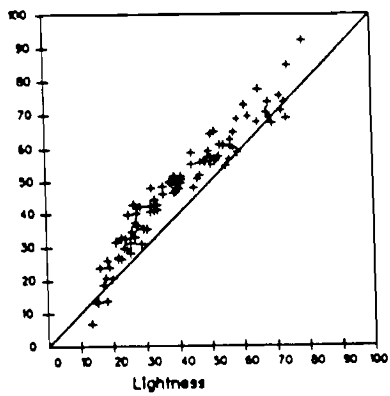


Figure 8-Brightness, colourfulness and hue for each phase as a function of that for highest level of ill.

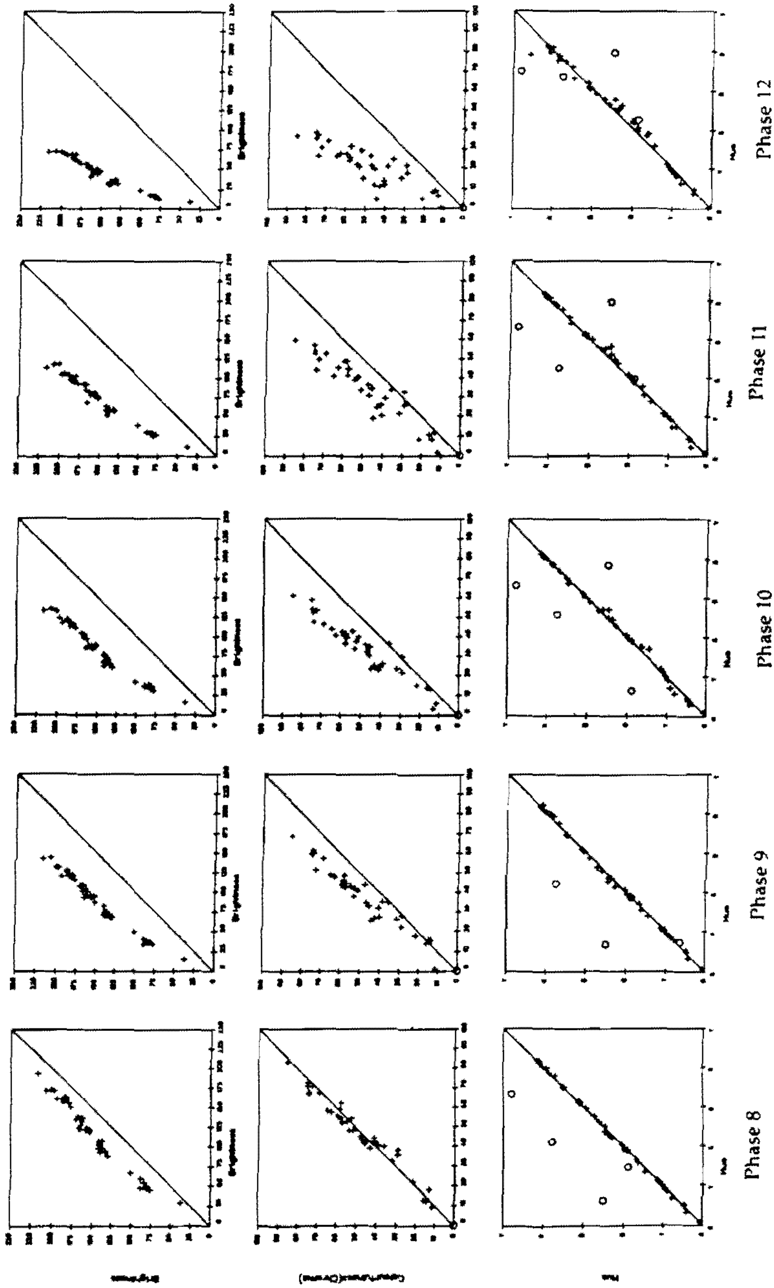


Figure 9-Lightness, colourfulness and hue for each phase as a function of that for highest level of ill.