TECHNIQUES FOR REPRODUCING IMAGES IN DIFFERENT MEDIA: ADVANTAGES AND DISADVANTAGES OF CURRENT METHODS

Tony Johnson*

Abstract: Issues which currently prevent a fully objective definition of colour reproduction between different media are discussed and how algorithms used in Graphic Arts overcome these problems. Advantages and disadvantages of this approach are compared and a brief review of work being undertaken to produce more objective methods given. Issues raised by this work are discussed.

INTRODUCTION

The Graphic Arts Industry has faced the problems of reproducing colour in different media for many years. Photographic transparencies, which are produced with a gamma greater than unity to take account of the effect of viewing, have long been the dominant type of original. Since they generally have to be reproduced on paper, which is viewed quite differently, it is clear that a direct reproduction of the tristimulus values of the original will not produce a colour match because the conditions of viewing have a significant effect upon colour appearance.

In the last 30 years the problem has become even more complex as we have witnessed the accelerating trend toward 'soft' proofing on colour monitors. Again the conditions of viewing influence appearance and matching tristimulus values to those of transparency or print will not provide a match to either. Put another way, if we require colour monitor, colour transparency and print to match the tristimulus values required for each would be different.

However, the difference in viewing conditions is not our only problem. Generally the colour gamut of the media varies. Thus, each reproduction made will require a different gamut mapping. The criteria for this are not simple and a gamut compression that is acceptable for one image may not suit another. This does not mean that a single set of rules cannot be found which will prove acceptable for all images; we just have not found it yet! I believe that the issue of what makes a 'pleasing' reproduction has not yet been fully explored, when taking different image content into account, despite the published work, and until this has been resolved any gamut mapping algorithm can only be a compromise.

* Crosfield Electronics Ltd.

Other practical difficulties include metamerism between the media and accurate modelling of colorant behaviour. Both problems can be resolved but not always with the simplicity we would like.

The approach to colour reproduction which has stood us in good stead over the past 30 years evolved from a considerable amount of practical expertise and the theoretical studies of a relatively small number of workers. Notable among them were Yule, Clapper, and Preucil. The approach they took can be summarised as "measure the amount of each of the three colorants present in a photograph and reproduce that by printing the same amount of colorant on paper". (Like many summaries this is a little simplistic, but is broadly correct).

This approach inevitably leads to the view that a measuring system is preferable which produces a reasonably simple correlation with colorant concentration and/or thickness. Density provides such a measure. Colorimetry has a relatively limited application in such a model; hence the quote from Yule (1967) "Conventional colorimetric parameters are not very suitable, [for the measurement of ink colour in Graphic Arts] and methods involving a greyness parameter which is almost independent of ink film thickness are the most satisfactory". Whilst Yule certainly appreciated the benefits of colorimetry very well, (he subsequently assisted in the design of a colorimetric scanner) he was all too well aware of its' shortcomings. Many of these still exist today.

Despite many attempts during the past 20 years to develop systems and models entirely based on colorimetry (see, for example, Pobboravsky, Pearson and Yule (1968a and b, 1971a and b, 1973), Korman and Yule (1971), Birkenshaw (1977a and b), Schlapfer (1981) and Johnson (1982)) few have found substantial commercial acceptance. One of the more obvious approaches is to use colorimetric density but this has its' problems. Thus, with the notable exception of the Eikonix Designmaster and it's derivatives most graphic arts images are still produced on systems which rely on the early work mentioned earlier. This raises an important question. Is the failure to convert to systems based on colorimetry due to a natural reluctance to 'change a winning team' or because there are inherent limitations to a such an approach which are as severe as those in the existing one? I believe the answer is a little of each and plan to briefly explore why in this paper.

Algorithms which are currently in use to achieve good quality reproduction have generally evolved without any objective solutions existing to the problems of gamut mapping and appearance modelling. They have therefore been developed empirically. As objective solutions start to become available the algorithms are changing. However, the issue of backward compatibility with existing systems has to be addressed. Purely objective procedures which do not provide the flexibility of current, empirical, methods may not provide experienced users with the creative tools they require. Getting to the point where we can reproduce an original perfectly may not be adequate since many originals are not always pleasing. However, it is clearly a useful starting point and one we should strive for. It can then be developed to include the flexibility which currently exists.

The structure of this paper will be to provide a short introduction to the type of algorithms which are commonly used in current graphic arts equipment and review how they have been used to overcome the difficulties described above. By comparing these with the more obvious alternative algorithms the advantages and disadvantages of each soon start to become clear.

Current algorithms

During the following discussion I will frequently refer to colour processing being part of a scanner. This is partly for convenience and partly because of historical precedent. The same principles apply if the scanner is a simple digitiser and the colour processing is undertaken off-line.

If we make the assumptions that printing inks and photographic dyes are isomeric, using density as the measurement domain provides a linear system and appearance and gamut problems do not arise the process can be described by the simple first order masking equations described by Yule (1938). The fact that the system is not linear can be partially compensated for by the proposals of Clapper (1961) or Neugebauer (1937). Furthermore, these models can be used equally well in a colorimetric scanner to compensate for spectral differences between inks and dyes. Such algorithms form the basis of those used in some scanners but are generally somewhat limiting because of differences in media. They are most efficacious when photographic prints are being reproduced as ink on paper.

Because of the need to take account of gamut mapping and appearance effects when media are dissimilar most current algorithms are based on the premise that it is necessary to optimise the following parameters:

a) Tone reproduction
b) Grey balance
c) Colour correction

Virtually all existing graphic arts scanners have some means of adjusting these parameters in order to achieve the reproduction required. As a general rule calculations are based on density units which, as already described, produce a reasonably linear relationship to the colorant amount, particularly for non-turbid media such as photographic dyes. Furthermore, measuring with narrow band filters is useful in improving the approximation to a linear system due to the primaries being unstable. (This means that the primary absorption band changes width as the concentration varies).

The algorithms used to optimise these parameters are quite straightforward. Although the specific details of each vary from one manufacturer to another the final effect is similar. For this reason I will not attempt to define each of them but simply provide some examples.

Tone reproduction simply requires a mapping to be established between original and reproduction densities which produces pleasing results. This is then applied to all original densities. A variety of methods are used to define this function and there are conflicting views on the precise form which it should take. The general requirement is to compress the dynamic range of the original into that of the print with a function which compresses shadow more than highlight densities.

However, in order to be adaptable for different types of originals what is generally provided is a function which contains a number of terms, each of which has a greater influence on a particular region of density than the others. A simple example is:

$$y = ax - bx^2 - cx^3 \qquad \text{-eq.1}$$

where x and y are the original and reproduction densities respectively. (a, b and c must be chosen to provide functions which are correctly scaled and increasingly monotonic). The expressions used in practice tend to be more complex to provide greater flexibility and on digital scanners may be provided as look up tables with simple functions to provide local tone modification. However, the example given here could be used as a basis for development.

Achieving optimum tone reproduction is one of the most fundamental tools used by operators to achieve both gamut compression and appearance matching. It is for this reason that there seems to be different opinions between researchers as to what constitutes optimum tone reproduction and why the flexibility needs to be provided to alter it in practice. I will return to the subject in more detail when discussing how current algorithms meet the requirements of matching across different media.

Grey balance is achieved by determining the relative amounts of each ink which are required to produce a neutral and ensuring that when such a colour is measured by the scanner it has the appropriate proportions of each ink defined. Clearly the absolute amount of ink is that defined by the tone reproduction requirement. This can be achieved by using equation 1 above, but with different coefficients, for each colour. Again, in practice, to achieve the control required more complex functions or look-up tables are generally used.

It is also necessary to ensure that a neutral is well defined by the input scanner so that the algorithm can be properly interpreted. In general, this is achieved by ensuring that R=G=B for a neutral by modifying the signals obtained from the detector amplifiers. However, so long as the relationship is known any other form can be compensated for.

Colour correction is generally very loosely based on the masking equations of Yule and/or Clapper mentioned earlier. However, metamerism, gamut compression, appearance matching and editorial requirements mean that additional complexity, or a more 'user friendly' form, is needed. This has resulted in a number of different methods being used, almost all of which rely on dividing the 'colour' space into a number of regions. For example, ranking the density values obtained for each pixel enables us to define six (overlapping) regions each of which can be modified separately. Thus if R>G>B the pixel has a 'red' component ($\underline{R}$) defined by R-G and a 'yellow' component ($\underline{Y}$) defined by G-B. (It also has a neutral component defined by B). Other rankings give rise to 'green ($\underline{G}$),' 'blue ($\underline{B}$),' 'magenta ($\underline{M}$)' and 'cyan ($\underline{C}$)' components. Each of these can be modified separately in the colour transformation by quite complex functions. Thus, colour correction may be achieved by using

equations of the form:

$$y' = y + f_{11}(\underline{R}) + f_{12}(\underline{Y}) + f_{13}(\underline{G}) + f_{14}(\underline{B}) + f_{15}(\underline{C}) + f_{16}(\underline{M})$$

for:

$$(\underline{R}),(\underline{Y}),(\underline{G}),(\underline{B}),(\underline{M}),(\underline{C}) > 0 \qquad \text{-eq.2}$$

where y represents the amount of ink prior to correction and y' the corrected amount. Different equations can exist for each of the inks with functions f_{2i}, f_{3i} and f_{4i}.

Certain interesting properties are evident from these equations. Firstly, neutrals (in which R=G=B) are not corrected. Secondly, for each pixel, only two (at most) of the terms can be positive and so only colours with certain rankings will be modified if any of f_{ii} are changed.

Thirdly, the magnitude of the terms (R), (Y), etc are crude measures of chroma and so according to the form of the functions f_{ii} correction may be biased to more or less saturated colours. Finally, it is clear that the same sort of technique can be used to extend the number of colour regions and thereby make each of the functions even more selective.

As will be easily understood when considering the above equations tone reproduction and grey balance are not, in general, independent of colour correction. However, neither should they be. If the tone reproduction is altered colour correction must be also to maintain the correct hue, Yule (1967).

If we consider the equations of Yule, Clapper or Neugebauer it is clear that when any coefficients are changed it will affect all colours, including neutrals. Thus, grey balance is not relevant to such a model although, because of the critical nature of neutral reproduction it would normally be used to modify the parameters of such models empirically to ensure that greys remain neutral.

With equation 2, however, this is not the case. Any change to the colour correction coefficients will not affect neutrals. Therefore, the equations are not achieving colour correction in the conventional sense (i.e. correcting for the unwanted absorptions of the inks). When combined with tone reproduction and grey balance they achieve far more and thereby enable us to accommodate gamut compression and media differences as discussed in subsequent sections.

Black printer generation is a fundamental part of the colour conversion process of any of the current algorithms. Again, the way that is achieved varies quite significantly from system to system and has greater potential for variation since the addition of black ink means that for most colours in the original there is no unique solution.

The black printer has two roles in the reproduction. It extends the gamut of colours achievable in the darkest regions of the image and in addition can be used to replace appropriate amounts of the three coloured inks to ease control of colour variation on the press. When the latter is constrained to near neutral areas it is called UCR, when applied to all colours it is called GCR.

The calculations are often based on the simple addition of density. A printed test image can be used to find the density (or lightness) achieved when different levels of black ink are used to extend the colour gamut. However, it is invariably found that the densities of the three colour grey and black do not combine additively. Any model for calculating black must take account of this.

Equations which can be used for this calculation have been described in an earlier paper, Johnson (1985a), and will not be discussed in detail here. Suffice it to say that if the tone reproduction requirement defines a specific density value c for a pixel then the density of black ink required to extend the gamut is given by:

$$y=(c-a)/(1-a/K)$$

where c is the total density, a is the three colour density and K is the convergence point of the additivity diagram as described by Yule (1967).

To achieve GCR some fraction m of the three colour density a is replaced by an equivalent density of black. The total black is then given by:

$$z=y+ma(1-y/K)$$

If m<1 it is necessary yet again to correct for additivity failure. This can be achieved by the same sort of function as those above and can be applied either to the residual three colour grey, replacement black or both. In practice it is quite convenient to correct the three colour component since any residual chromatic component also needs correcting. One technique suggested by Otschik (1981), on which much of the above is based, is to multiply the residual colour components by a correction factor which was published in the form of a graph. We found this to be fitted well by the expression:

$$b=10^{x^2}$$

where b is the correction factor and x is the fractional area of black replaced.

The form of black generation described above is the basis of that used by Crosfield, although some additional functionality is provided to enable the user to 'trade-off' three colour grey and black selectively in different tonal regions and correct for the approximations of Yule's additivity failure model. Other vendors have slightly different approaches but all achieve similar ends by providing the user with a range of controls to obtain the ratios of black to three colour grey required in each tonal region.

CMYK to RGB conversion is necessary for converting the images from device dependent CMYK form, which is commonly used in current systems, to the RGB required for display. I am aware of three methods which are used to achieve this conversion. Clearly, the primary requirement is to match the print or proof on which the image will be printed, since we are dealing with device specific images and simply wish to proof them. Generally, therefore, gamut mapping is not a major problem since the monitor can render most of the colours obtainable with even the best four colour processes providing that sensible viewing conditions are used and lightness relative to the white point of each is accepted as the matching criterion. Appearance matching is the major issue in this transformation.

One method is to provide the operator with a set of colours, of known CMYK combinations which have been printed on the printing (or proofing) press to be matched by the monitor. He is also provided with a set of 'controls' which enable him to modify the RGB drive values. Each colour is presented in turn and the simulation on the monitor is modified by alteration of the RGB values until a match is obtained. Once all the colours have been matched a least squares fit to some pre-defined funtion, such as a polynomial, is computed. This then defines the colour conversion to be carried out on each pixel.

An alternative method, which is also based on viewing printed samples, is to provide functions which are similar to those described earlier for obtaining the correct colours on a scanner. The operator then adjusts the coefficients to obtain a match in much the same way.

The third method, which attempts to be rather more objective, is to convert the colours using tristimulus values as an intermediary. By using one of the models described earlier, such as those defined by Clapper or Neugebauer (or some derivative of them), the tristimulus values resulting from the various ink combinations may be computed. These may then be transformed directly into RGB values by means of a 3x3 matrix and gamma correction. See, for example, Sproson (1983). To be really precise the errors in additivity and gamma of the monitor can be corrected also. Such a procedure has been described by Luo et al (1990).

The problem with an approach which matches tristimulus values is in compensating for conditions of viewing in obtaining an appearance match. As already stated matching tristimulus values will not suffice. In practice, this is achieved empirically by modifying the gamma correction tables as will be discussed subsequently.

Appearance Matching

The fact that matching tristimulus values across different media does not produce a colour match has already been described. The same would be true of conventional density units even were dyes and pigments isomeric. Clearly such a problem gives rise to significant difficulties when trying to define colour reproduction in an objective way. For this reason Crosfield have been collaborating with Coats-Viyella Ltd. and Loughborough University, with the support of the British Government, to develop a colour appearance model appropriate to the Graphic Arts industry. We are using the model developed by Hunt (1987, 1991) for this work and he has been assisting us in the study.

Various papers have been given describing this work. See, for example, Luo et al (1991a and b), Johnson (1989) and Johnson and Luo (1990, 1991). An earlier discussion reviewing the literature to date was given by Johnson (1982). However, it is not yet being used in practical systems. Whilst we are satisfied that the method is viable the computation required is quite intensive. This requires new equipment architectures to make it practical which will start to appear in the near future. However my task in this paper is not to discuss the future; that is for other speakers at this conference. I will therefore revert to the methods which are used currently to overcome this problem although a brief summary of where the work will lead is given later.

In general, appearance matching is currently achieved by trial and error. The procedures used for matching monitors to prints, by making visual comparisons and modifying colours on the monitor to achieve a match, has already been described. It was noted that even when using tristimulus values as the basis of the transformation some empirical correction is normally required because of viewing differences.

Some attempts have been made to use quantitative models to predict this change, with some degree of success. The equation derived by Bartleson and Breneman (1967) has been used to predict lightness requirements in different viewing conditions. Thus, if the monitor is viewed in a darkened room, the change in luminance required to match the print can be computed directly. Care has to be taken to ensure that changes in hue and chroma are minimised and for a first approximation simply modifying X and Z to follow the changes in Y seems to suffice. Use of the Hunt model would improve on this.

Unfortunately, in practice, the conditions of viewing are not easily controllable because of issues such as safety, comfort of working and cost of dividing work areas. What has to be provided, therefore, are facilities to enable the user to modify an objective calibration empirically. Changing the appropriate parametric coefficients of the Bartleson-Breneman equation is one such method, providing excessive glare is minimised.

The same Bartleson-Breneman equation has also been used to good effect in calculating the tone reproduction changes required to produce an appearance match between transparencies and printed reproductions. The equation, which has parametric constants which vary with viewing conditions, is specified below:

$$\log B = a + b\log L - c\exp(d\log L) \qquad \text{-eq.3}$$

where B= brightness, L=luminance (mL) and a,b,c and d are parametric constants. The values of a and b were found to be almost constant and were set at 2.037 and 0.141 respectively. The values of c and d can be derived from the following expressions derived by Yule, see Johnson and Birkenshaw (1977):

$$c = 0.99 + 10^{(0.25\log(L_0 f) - 0.372)}$$

and

$$d = -0.12 - x10^{(0.093\log(L_0 f) - 0.961)}$$

where f varies from 0.1 for dark surround to 1.0 for light surround.

If brightness is computed, from density, for dark surround transparency viewing conditions the density to provide a match can then be computed for light surround print conditions, from the inverse of equation 3, if the dynamic range is the same in both. Of course attempting to match brightness is not realistic; in practice we normalise relative to white to match lightness. Generally, the reference white is paper for the print and the lightest catchlight for the transparency. However, if the operator has selected the highlight point it is better to use the value of these in both cases.

Clearly, such a procedure does not correct for the reduction in chroma which comes from a dark surround viewing condition, Pitt and Winter (1974) and Bartleson (1979). There are three difficulties in achieving this. Firstly, with the exception of Eikonix, no major systems use tristimulus data; density is computed using spectral sensitivities which are not colour matching functions. As already stated this permits a relatively simple transformation into ink amounts and there would be little value in converting into tristimulus values purely for this correction. It would be possible to compute the necessary modification to the transformation, for each material type, using the appropriate appearance corrections but other complications associated with gamut compression make this an academic suggestion. For lightness the correction to the transformation, as described above, is far easier to determine and, as we shall see, gamut compression in lightness is reasonably well understood. For this reason an appearance correction based on lightness is worthwhile.

Another problem is the complication introduced by gamut compression. In general the colour changes required for achieving this are far greater than those required for appearance matching. Thus any change needed would need to be combined with the gamut compression transformation. If this is not well defined, and in my view it is not, there is little point in undertaking the corrections needed for appearance matching of chromatic colours objectively when such major empirical changes are required for gamut compression.

Measurement effects also contribute to the difficulties of achieving a match numerically. Even with a good appearance model a match will be difficult to predict when significant flare or gloss differences between the media exist.

Clearly existing Graphic Arts practitioners have some success in matching colour appearance across different media. Trial and error using visual comparison has its advantages since issues of measurement do not arise. The parameters available to modify tone reproduction, grey balance and colour correction provide a significant degree of flexibility. Tone reproduction and grey balance enable a exact match of the grey scale to be achieved and then the colour correction controls (if they are derivatives of the sort of functions described earlier) can be used to optimise the colour match. Since each of the functions f_{ii} only affects fairly selective regions of colour the degree of control is quite flexible.

The disadvantages of such a system are threefold. Firstly, the subjective nature of the matching procedure means that different operators are likely to achieve different results. For neutrals the flexibility normally afforded by grey balance and tone reproduction controls do permit this to be measured and set rather more objectively if the Bartleson-Breneman compensation for viewing is used. It is also feasible to use measurement to attempt to minimise differences between operators for colours other than grey but this does not have the advantage that differences can be translated directly into corrections as is the case for a grey scale. Nevertheless, it is possible to ensure that differences are minimised if necessary.

Clearly, the limit to which two different systems can be matched depends upon the degree of control afforded to the operator. If enough colour regions are provided and the functions f_{ii} are sufficiently complex the final differences can be very small. However, this makes them very difficult to use. Normally a compromise is achieved in which the degree of flexibility is limited to keep the set-up within reasonable bounds and some degree of colour error is accepted. (This is comparable to the acceptance of errors in a device independent system in order to limit the measurement and correction procedures required). The reason that we can accept such errors is due to the far greater differences introduced by gamut compression.

The other major difficulty with such systems is the need to carry out this exercise for each original type and printing condition. Whilst this sounds fairly horrendous, in practice it is far less significant since the range of both is rather limited. The 'standardisation' which has been evolving in the printing industry over the past 20 years has greatly reduced the number of printing characteristics which have to be considered. Many processes use inks which are similar in colour; the main variation when they are printed is the gloss and ink film thickness which can be achieved and substrate used. In practice the effect of this on the overall colour transformation is rather small. The primary effect is to modify the gamut but since the conversion to CMYK dot areas has effectively 'ranged' the data the resultant gamut mapping is quite efficient without any changes to the ink amounts. The main requirement is to correct for dot gain which is trivial.

Where ink colours are significantly different a new transformation is required. For most users this means empirical correction of the controls described earlier. However, systems have evolved over the past few years which automate this using a quite simple procedure. A colour chart is printed by the two processes under consideration and tristimulus values are determined for each colour patch. This enables a transformation to be established between each CMY(K) and the tristimulus values (or vice versa). The two can be convolved together to provide a conversion from one to the other. The major source of error which can arise with such a system is that attributable to measurement, when the gloss of the two prints is very dissimilar, providing that the gamuts are similar and sufficient colours are measured. If the gamuts differ then a mapping must also be defined which introduces subjectivity again.

The problems of metameric originals will be discussed in a later section.

Gamut Compression

The controls used for appearance matching are also useful to achieve the gamut compression so often required in printing. Again a trial and error approach is used and, of course, in practice both tasks are undertaken simultaneously.

The Bartleson-Breneman equation can be used to good effect in defining, objectively, compression of the lightness axis of the colour transformation. Earlier studies have shown that if the preferred tone reproduction data obtained by Yule (1964) is converted to lightness, as defined by this model, optimum compression is obtained by a linear scaling defined by the respective lightness ranges, Johnson and Birkenshaw (1977). Thus, adding such a routine to that described above for obtaining a lightness match establishes an objective transformation for the preferred reproduction of neutrals when a 'facsimile' reproduction is required. This is discussed in more detail in Johnson (1985b).

It must be said that not all researchers agree that a linear scaling provides the optimum tone reproduction. Jorgensen (1976) suggests that this is achieved by selecting an area of interest within the image, reproducing the range of lightnesses within that area with an equal range of lightness on the print and compressing elsewhere. Archer (1985), on the other hand, adopted the approach recommended by Ovchinnikov et al (1973) and produces a histogram of the density differences between adjacent pixels of the image. By modifying the histogram to some pre-defined rules a tone reproduction characteristic is defined. In particular, it is claimed that such techniques are better at reproducing high and low key subjects.

Compression of chromatic colours is not so simple. In a report published some time ago, Johnson (1982) a gamut compression algorithm was proposed which would

a) scale lightness linearly

b) maintain hue

c) scale chroma linearly

It was assumed that appearance modelling would be used to obtain these parameters. In practice, however, such a procedure has had limited success. There seem to be a number of reasons for this including the fact that the two scalings are not commutative, the two gamuts are not, generally, of the same shape so a decision on the scaling factor (for chroma) is not simple and maintaining hue is not always desirable. Laihanen (1987) concluded that a different compression is required depending upon image content; it is not possible to measure the gamut of the original and reproduction media and produce a generic compression for all images. This is not surprising when we consider lightness scaling since it is known, as discussed earlier, that lightness should be scaled by mapping the lightness range of the original linearly into that of the print. This does change on an image by image basis. However, it is rather more surprising for chroma since it is not normal practice for users in conventional colour reproduction to substantially modify the colour correction for each image.

On existing systems, even if linear scaling of lightness is used for the neutral axis, the compression of chromatic colours is achieved by trial and error. Using the selective controls described in the previous sections enables the user to achieve local modifications of hue, chroma and lightness in some regions of the colour transformation whilst leaving others unaffected. Such flexibility can permit a whole variety of ways of achieving gamut compression. Thus the user can choose that best suited to his needs.

The disadvantages of such an approach are the same as those already discussed; subjectivity makes it user dependent, complex functions which are difficult to optimise are required to meet all requirements and different compressions required for each process and original type mean a number of such transformations may be necessary.

Metamerism

Few Graphic Arts scanners have the wide-band spectral sensitivities necessary to define colour matching functions. The reasons for this have already been discussed; the simple reproduction theory which is used as the basis for most scanners (albeit adapted to overcome the inherent limitations) is based on isomeric matching. The fact that this is a reasonable approximation means that using colour matching functions to provide metameric matching capability would degrade signal-to-noise performance. Changing algorithms may be sufficient to compensate for this but is as yet largely unproven.

Such an approach means that a different colour transformation is required for each original type. For transmission scanning this is not a significant problem. Only 4 dye sets are encountered by most users and producing 4 transformations is hardly onerous. For reflection copy scanning, however, the situation is less simple. The number of photographic print dye sets is still fairly limited but users can also be confronted with a range of dyes and pigments, which are vary variable, on non photographic media. We cannot expect narrow band filters to cope so well with this range.

Thus, the risk always exists that a new range of pigments or dyes may be encountered by the user which do not reproduce well with any existing calibrations. In such a situation the controls already described are employed to obtain the correct colour rendition.

This is a relatively minor problem, in practice, because the range of original types encountered is normally quite limited. At least 80% of originals for graphic arts are 35mm and 60mm transparencies and most of the remainder are larger format transparencies and photographic prints. By the time originals which do not lend themselves to scanning (due to size or rigidity), and negatives, have been removed the number of possible problem originals is quite small and can be accommodated by a trial and error approach if necessary. Because of this the advantages of narrow band filters are generally employed although two approaches are used to minimise the problem. One is to compromise by having a spectral sensitivity which is neither wide nor narrow. The other is to provide separate filter sets for each mode. Both are used in practice.

Modelling colorant behaviour

As already described the theory underlying conventional graphic arts reproduction does not require complex modelling to be carried out. The use of narrow band filters, density measurement and a limited range of originals, combined with the controls provided to empirically take account of the various problems described, makes for a simple set of algorithms. If we do not employ such an approach but define the tristimulus values of an original prior to computing the ink amounts it becomes necessary to provide a characterisation of the device to enable the transformation to be carried out. Determining this to a reasonable level of accuracy is not particularly simple.

Two approaches offer themselves. One is to model the process analytically from a limited number of colours. The models of Yule, Clapper and Neugebauer, or the derivatives of them which have appeared in the literature since, all provide possible options. It is not appropriate to review all of these in detail here. Anybody looking for further information may refer to Johnson (1982) where further details and references are provided or texts such as Yule (1967) or Field (1990).

Using such models requires only a small number of colours to be measured; from 5 (three for the first order masking equations and two for black addition) to a hundred or more for fitting polynomials. The choice depends upon the accuracy required.

The alternative is to measure a large number colours and develop simple look up tables with interpolation for any intermediate colours. The more colours measured the more accurate will be the transform. Providing that the colours are suitably chosen (i.e. they are well 'spaced'), 700 colours and upwards is the typical sort of number used. To many users such a measurement task is not attractive and so methods which provide a basic characterisation made from a large number of colours which is then modified by analytical modelling as the device characteristics alter have some attraction. This is discussed in more detail in Johnson and Luo (1991).

Such a procedure is only attractive when the problems of appearance matching and gamut compression have been resolved. A system based on such an approach will be briefly described in the next section.

The Future?

As described earlier Crosfield are involved in an extensive project which has the objective of providing an appearance model appropriate to the Graphic Arts industry and defining methods of gamut compression. We are also involved in another project, partly funded by the European Community, in which a number of vendors, printers, publishers and others will produce a system for multi-media publishing based on CIE data. CIEL*a*b* has been selected as the interchange colour space.

Providing such projects prove viability of the hypothesis, and there is no evidence to date that they will not, figure 1 provides a possible model for colour reproduction for the future. There is nothing particularly new in this, similar models have been suggested many times in the past. What has changed in the past 10 years, and may just make such a system viable in the future, are four things:

a) Availability of cheap and easy to use colour measuring equipment

b) Development of suitable colour appearance models

c) A significant increase in computational speed per unit of cost

d) Publishing being undertaken on a range of media with very different characteristics

The sequence shown in figure 1 is not the only one which may be used. However, it shows the important transformations which must be undertaken. Clearly some of the steps may be omitted if gamut compression is not necessary and/or media have the same appearance characteristics. Furthermore, once the data has been converted to LHC (Lightness, Hue and Chroma) the subsequent steps may be undertaken a number of times, once for each device.

In the light of our findings to date a number of issues have been raised concerning such a model. I believe the following suggestions will assist in resolving these and determine whether such a system will replace device dependent methods which are inevitably very efficient for the process they are optimised for.

<u>Gamut compression</u> is one of the key issues. The conventional wisdom is that to be device independent the tristimulus values of the <u>original</u> must be the basis of the data transmitted. This gives rise to issues of data precision for originals with a large colour gamut. I would argue that such a premise is really not necessary, it is far more efficient to transfer data which has been mapped into the gamut of the <u>device which is likely to be the most demanding and dominant in terms of volume and/or quantity.</u> Currently this would be the printing press. Data is called device independent if transferred as tristimulus values which can be rendered when produced on a hypothetical device (i.e. one which can reproduce the original perfectly). Why is it any less independent if appropriate to a specific, well-defined, process? If necessary the compression could be reversed to recreate the original tristimulus values since gamut mapping is fundamental to such a system.

If we follow such a suggestion we get two significant advantages. The first, as already stated, is in data precision. In the multi-media project currently being undertaken we have chosen, after much debate, to transfer data in gamut compressed form. By doing this it has been demonstrated that 8 bits per channel will suffice although this has yet to be verified with the addition of any retouching. The second advantage is that most devices produce gamuts which are closer to that of a print than many originals. Using data appropriate for printing thereby brings the advantage that the complexity of gamut mapping for most devices will be quite low; the more difficult one is only undertaken once. This significantly reduces the degree of subjectivity which is likely to intrude otherwise. Clearly the main exceptions are transparency recorders and devices using additional colorants but the amount of publishing undertaken on these is fairly minimal. (Under idealised viewing conditions some colour monitors can produce quite large gamuts also but the use of such conditions is rare).

To make such a situation work it is necessary to produce a model of a 'standard' device (printing press). For our multi-media work we are currently using an approximation to the gamut defined by the FOGRA/BVD printing specifications for coated paper. I believe this would make an excellent standard for such a procedure.

Measurement effects need to be included in the appearance modelling. Much of the data derived in our studies have been obtained using telespectroradiometry which does not correlate well with other, more common, measuring devices. We are in the process of investigating how to transform from one to another and find the suggestions of Berns and Petersen (1988) helpful in this. We anticipate publishing our work on this in the future.

Even if measurement conditions are standardised problems still arise because of gloss differences between media. When such a situation exists it is necessary to undertake prior cross-calibration.

The Hunt model, as it is evolving, seems to provide an excellent basis for computing appearance. The latest version seems to provide us with reasonable expressions for the main viewing conditions encountered in Graphic Arts when the appropriate coefficients are chosen.

Computing colorant amount procedures are likely to require additional functionality to those described above, at least until gamut compression algorithms are well defined. The need to produce 'pleasing,' rather than colour accurate, reproductions has long been the demand of the print buyer. Whilst we may argue that this simply tells us that our gamut compression algorithms are inadequate I am doubtful if this is the full story. Neither am I convinced that we can make the image pleasing at our work station and then render it on a range of devices. I believe it is going to require some 'tweaking' at each device to produce the most pleasing results in order to overcome the approximations inevitable in our algorithms and also to resolve calibration issues. There may well be a role for the algorithms described earlier to permit local editing of colour transformations defined automatically.

There is a belief, which is quite widespread, that the operation of current graphic arts systems is difficult and that transferring to a device independent system will remove this difficulty. It is my contention that this is not true. What requires skill, currently, is defining the colour transformation at the outset. Once this is achieved scanning is a very mundane business. It is only when editorial corrections or some very specific, image dependent, gamut compression is required that skill and experience become necessary. For many reproductions this does not arise. Systems using colorimetric data will be exactly the same.

The advantages of such data seem obvious in a multi-media environment but the advantages of direct encoding of device colorant requirements are also clear to many printers. I would not be confident in asserting that the sort of system shown in figure 1 is inevitable for all colour encoding. However, there are two things of which I am confident. In future, where images are encoded in a device dependent form the colorimetric properties of that process will be understood to enable conversion of the image to a different encoding if required. Furthermore, appearance modelling and gamut compression studies will enable existing systems to be improved. There is little doubt that a better definition of how to achieve a colour match on different media will be of great value to all involved in colour reproduction.

SUMMARY

Objectives in colour reproduction are as yet not well defined. Some of the issues in attempting to provide a fully objective system have been discussed and the main issues to be resolved shown to be the problems of appearance matching across different media, gamut compression, producing 'pleasing' reproductions and measurement geometry. Solutions are being found for some of these but not yet all.

Current techniques have been described and shown to be a mixture of empirical and objective methods which have evolved as a means of circumventing the unresolved problems. The advantages of such a procedure are shown to be simplicity, adaptability to whatever reproduction objective is desired and efficiency of encoding for printing. The disadvantages are the level of skill required, the subjective nature of the result, the limited capability of the algorithms employed in order to achieve simplicity and the need to develop transformations for every original type and printing process. However, it is shown that the latter problem is less severe than frequently stated.

REFERENCES

Archer, H. B.
1985 "Automatic analysis of reproduction copyfor optimum tone reproduc tion," TAGA proc.p546

Bartleson, C. J.
1979 "Predicting corresponding colours with changes in adaptation," Col. Res. Appl.,4, 3, p143

Bartleson, C.J. and Breneman, E. J.
1967 "Brightness perception in complex fields," J. Opt. Soc. Am., 57, 7, p953

Berns, R. and Petersen K.
1988 "Empirical modelling of systematic spectrophotometric errors," Col. Res. Appl., 13, 4, p243

Birkenshaw, J. W. B.

1977a "A study of the tone reproduction capabilities of colour reproduction systems," Pira report PR153
1977b "The Black Printer," Pira report PR154

Clapper, F. R.
1961 "An empirical determination of halftone colour reproduction require ments," TAGA proc. p31

Field, G. G.
1990 "Colour scanning and imaging systems," GATF, Pittsburg

Hunt, R. W. G.
1987 "A model of colour appearance for predicting colour appearance in various viewing conditions," Col. Res. Appl., 12, p297
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
1985a "Polychromatic colour removal-Evolution or revolution?" TAGA proc. p1
1985b "Designing scanner software to meet the press requirements," TAGA proc. p135
1989 "Defining optimum photomechanical colour reproduction," TAGA proc. p350

Johnson, A. J. J. and Birkenshaw, J. W. B.
1977 "Influence of viewing conditions on colour reproduction objectives," 14th IARIGAI proc. p48

Johnson, A. J. J. and Luo, R
1990 "Colour appearance modelling," TAGA proc. p144
1991 "Optimising colour reproduction," 21st IARIGAI proc. (to be pub lished)

Jorgensen, G. W.
1976 "Improved black and white halftones," GATF Research Project Report No. 105

Korman, N. I., and Yule, J. A. C.
1971 "Digital computation of dot areas in a colour scanner," 11th IARIGAI conf. proc. p93

Laihanen, P.
1987 "Colour reproduction theory based on the principles of colour science," 19th IARIGAI conf. proc. p1

Luo, R., Xin, J. H., Rhodes, P.A., Scrivener, S. A. R. and MacDonald, L. W.
1990 "Studying the performance of high resolution colour displays," Proc. of CIE 22nd session-Division 1, p97

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 perfor

mance using LUTCHI colour appearance data," Col. Res. Appl. 16, 3, p181

Neugebauer, H. E. J.

1937 "Die theoretischen grundlagen des mehrfarbendruckes," Z.Wiss.Photogr.,36,73

Ovchinnikov, Y., Fainberg, I., Litvan, R., Solntsev, I. and Avatkova, N.

1973 "A new approach to programming in photo-mechanical reproduction," 12th IARIGAI conf. proc.

Pitt, S. T. and Winter, L. M.

1974 "Effect of surround on perceived saturation," J. Opt. Soc. Am.,64, 10, p1328

Pobboravsky, I., Pearson, M. and Yule, J. A. C.

1968a "Computation of the halftone colour gamuts of process inks," TAGA proc. p331

1968b "The colour gamut of process inks computed by the Neugebauer equations," Print. Tech. 12, 3, p150

1971a "The relationship between photomechanical colour reproductions and the original copy" ISCC conf. proc., p159

1971b "Comparisons of printed originals and colour reproductions," 11th IARIGAI conf. proc. p73

Pobboravsky, I. and Pearson, M.

1973 "Colour transformations from psychophysical parameters to photome chanically printed dot areas," 12th IARIGAI proc. p181

Schlapfer, K

1981 "Optimum conditions for colour separation in graphic arts reproduc tion," 16th IARIGAI conf. proc.

Sproson, W.N.

1983 "Colour Science in television and display systems," Adam Hilger, Bristol, UK

Yule, J. A. C.

1938 "The theory of subtractive colour photography; I, The conditions for perfect colour rendering," J.Opt.Soc.Am.,28, p419

1964 "7th Int. Conf. Print. Res. Inst.",Pergamon Press

1967 "Principles of Colour Reproduction," John Wiley, New York