Enhancing the soft proofing paradigm

Alexis Gatt

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy
University of Leeds
School of Design

October 2007

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.
This copy has been supplied on the understanding that it is copyrighted material and that no quotation from the thesis may be published without proper acknowledgement.

Acknowledgement

I wish to thank first of all the Xerox Foundation for the financial support they provided throughout this study. More specifically, I would especially like to acknowledge Dr. Raja Bala and Dr. Karen Braun for their insightful comments as well as their kindness.

I would also like to thank Professor Stephen Westland and Dr. Phil Green for their continuous help and support.

I would like to thank Françoise Viénot and Nacim Ladjouze from the *Centre de Recherches sur la Conservation des Documents Graphiques*, Muséum National d'Histoire Naturelle (Paris) for allowing access to the EZContrast, their tremendous help in performing the measurements as well as the very kind reception, Gaël Obein for the pertinent advices on correctly implementing his method, and Chris Grawe from Keene Repro Ltd (London) for supplying a superb sets of samples.

I would also like to salute the many friends and colleagues that somehow managed to bear my person so stoically for so long.

Last but not least, to Yazhu Ling. Thanks for enlightening my life.

Abstract

The first part of this study is devoted to assessing the suitability of soft proofs as a surrogate for the final print in judging colour-reproduction quality. Two sets of psychophysical experiments were carried out. The first investigated whether the type of medium used to generate a stimulus influenced its appearance, while the second attempted to determine whether judgments made on the basis of hardcopy simulation are transferable to prints. The second part of this study is devoted to improve the accuracy of soft proofing by increasing the amount of information rendered on screen. We present a workflow that extends traditional colorimetry beyond simple colour appearance prediction to a higher dimensional representation, which includes gloss. The intention is to provide a full goniometric simulation of the spatial distribution of light reflected by a print based upon data gathered with simple and inexpensive instruments, glossmeters and spectrophotometers.

Table of Contents

List of figures

List of tables

Chapter 1: Introduction

While soft proofing is not a new concept to printers, the emergence of a global network has radically modified their production workflow (Green, 1999). Great benefits, both in terms of time and money, emanated from the possibility of performing the proofing task remotely. Although remote soft proofing for content has matured to the point of becoming almost a standard in the production workflow (Marin & Bassinger, 2003), proofing for contract colour dropped behind due to diverse technical limitations. Despite the advances in colour appearance models and colour management techniques, customers' final approval is still very subjective, and depends on their willingness to trust the soft proof as an accurate representation of the appearance of the real print. The precision of the colour match that can be achieved is indeed subject to the accuracy of the colour reproduction media involved. Their colorimetric characterisation is furthermore complicated by the fundamentally different principles used to generate colour information, as simple measurement devices experience problems to evaluate prints and monitors in a consistent manner. In addition, these conceptual differences might account for more than just troubling simple colorimeters, and be responsible for a more fundamental difference in stimuli perception. Results recently obtained (MacDonald, Morovic, & Xiao, 2002) at the late Colour and Imaging Institute challenged a very common practice in both research and industrial settings, i.e. the use of a soft proof as a surrogate for the final print in judging colour quality. While studying the quality of diverse gamut mapping algorithms, it was found that their relative performance was very different when a softcopy simulation of the print was used instead of a real hardcopy, despite a careful characterisation procedure. This finding challenged the widely held belief that a colorimetric match is equivalent to a visual match under controlled viewing conditions (Alessi, 1994). Indeed, the appearance of a stimulus is not only determined by its colorimetric attributes, but also by additional factors, such as the state of chromatic adaptation, mode of viewing, luminance level, background, surround, etc. If all these factors are carefully equated, it is often assumed that two stimuli exhibiting the same colorimetric attributes will be perceived as matching visually. However, the validity of this statement is limited in practice, as the precision of a colorimetric match is subject to the accuracy of the devices involved. Obtaining a perfect match across a wide

range of complex stimuli such as pictorial images is probably impossible. A series of experiments was therefore designed to address the viability of soft proofing, as well as establishing means of enhancing the realism of current soft proofing simulations.

1.1 Aims of the study

The aim of this project consisted in assessing the suitability of soft proofs as a surrogate for the final print in judging colour-reproduction quality, as well as identifying potential strategies for the development of next-generation soft proofing workflows. This includes the following objectives:

- To design a complex viewing apparatus to ensure that hardcopies and softcopies are assessed in identical viewing conditions.
- To investigate whether the type of medium used to generate stimuli has an influence on their appearance.
- To investigate whether judgements made on the basis of soft proofs are transferable to prints.
- To identify appearance attributes not currently taken into account that would significantly enhance the realism of softcopies.
- To develop a soft proofing model based on sound physics foundations that incorporate such attributes.
- To develop a feasible solution, i.e. one that could realistically be implemented in actual industrial settings.

1.2 Thesis outline

This thesis is organised into the following sections:

Chapter 1: Introduction

This chapter presents some background information on the topic studied and introduce the main objectives of this project.

Chapter 2: Literature Survey

This chapter provides a general overview of the diverse factors that determine the appearance of a stimulus, while focusing on what kind of influence the disparities in viewing conditions between hardcopies and softcopies may have on the appearance of such stimuli. A brief summary of the diverse colorimetric, psychophysical and statistical tools, methods and techniques employed in this study is also given.

Chapter 3: Experimental Apparatus

This chapter provides the design motives and a detailed description of the settings employed for all aspects of the experiment setup, and more specifically the complex viewing apparatus developed to equate the viewing conditions for hardcopies and softcopies.

Chapter 4: Colour Reproduction Framework

This chapter provides of a detailed analysis of the abilities of all the devices involved in the experiment, and discusses the overall accuracy of the colour-reproduction framework implemented.

Chapter 5: Experiment - Intrinsic differences between hardcopy and softcopy

This chapter presents the diverse experiments carried out to evaluate the viability of soft proofing. The first one investigates whether the type of medium used to generate stimuli has an influence on their appearance. The next two investigates whether judgements made on the basis of soft proofs are transferable to prints.

Chapter 6: Transition

This chapter discusses further the results presented in the previous chapter and identifies the most suitable appearance attribute that may enhance the realism of soft proofing: gloss.

Chapter 7: Physics of light and matter interactions

This chapter provides a detailed review of the physics behind colour and gloss, as well as means of accurately simulating the latter.

Chapter 8: Empirical analysis

This chapter provides an analysis of the optical properties of a large set of printed papers exhibiting varying levels of gloss. The emphasis is on the scattering of light occuring in the hemisphere above the surface of prints.

Chapter 9: BRDF Modeling

This chapter presents a soft proofing workflow that is able to generate accurately a full goniometric simulation of the spatial distribution of light reflected by a print. The best-fitting parameters for the samples set introduced in the previous chapter are derived, and the overall accuracy of the proposed workflow is evaluated.

Chapter 10: Conclusion

This chapter summarises all the findings and further directions of work are suggested.

1.3 List of publications

- Gatt, A., Westland, S. & Bala, R. (2004). *Testing the soft proofing paradigm*. Paper presented at the IS&T/SID Color Imaging Conference.

- Gatt, A., Westland, S. Bala, R. & Ling, Y. (2005). *Testing the soft proofing paradigm II*. Paper presented at the CGIV Conference.

- Gatt, A., Bala, R., Westland, S. & Green, P. (2006). *Increasing the Dimensionality of Soft Proofing: Gloss*. Paper presented at the IS&T/SID Color Imaging Conference.

Chapter 2: Literature Survey

The spread of imaging technology over the past century has fueled the emergence of colour science. The complexity involved in achieving a successful reproduction of colour across disparate devices required the establishment of theories and techniques based on firm scientific grounds. This section will provide a brief introduction to the science of colour, followed by a thorough review of the specific elements relevant to this project. Finally, a review of the diverse methods of psychophysical assessment will conclude this chapter.

2.1 Colorimetry

Colour vision emanates from the interactions between photons and photoreceptors in the human retina. There are two types of photoreceptors in persons with normal colour vision. *Rods* provide monochromatic vision mainly at scotopic levels of light, whereas *cones*, of which there are three different types, are responsible for trichromatic vision at photopic illumination levels. Each type of cone photoreceptors is sensitive to different portions of the visible spectrum, and overall the visible spectrum ranges approximately from 380 to 780 nm for humans with normal colour vision. Any ray of light incoming onto the retina can therefore be characterized by a unique combination of pure spectral colours, and the relative set of weights representing this combination is termed the spectral power distribution. Alternatively, the perception of colour engendered by this ray of light may also be described by three numbers, which represent the extent to which each of the three types of cones is stimulated. These two numerical representations depict the same phenomenon, i.e. certain properties of the light ray, but from two different perspective, i.e. physics and perceptual. The link between the two representations is provided by the colour matching functions, which characterise numerically the chromatic response of an observer. They are not a direct description of the cones' spectral sensitivity, but rather a linear combination of them derived for convenience. The *Commission Internationale de l'Eclairage* (CIE) has defined two sets of colour matching functions, the 1931 *Standard Colorimetric Observer* (Figure 1) and the 1964 *Supplementary Standard Colorimetric Observer*, to take into account the non-uniformity of the cones distribution in the human retina. As can be easily envisaged, the transformation from spectral power distribution to chromatic answer is surjective, which

is a phenomenon referred to as metamerism. Metamerism occurs because each type of cone responds to the cumulative energy from a broad range of wavelengths, so that different spectral power distributions can produce equal receptors response. Equivalently, variability between the colour matching functions of individuals will also engender metamerism. To tackle this issue, the CIE defined the *Standard Deviate Observer* to provide a model of the uncertainty of colour matching.

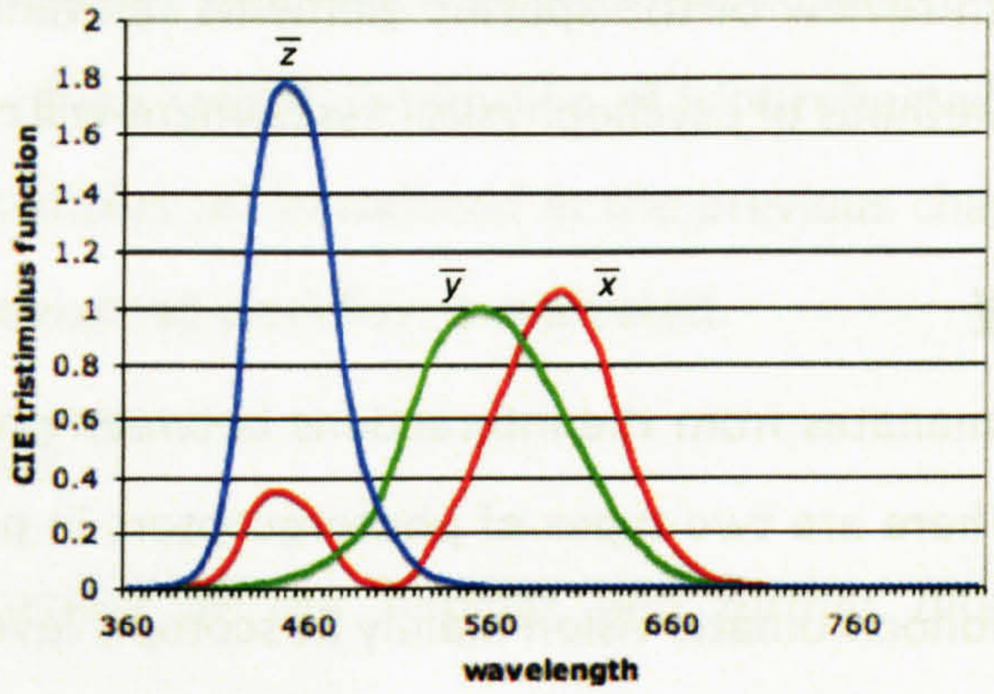


Figure 1 - CIE 1931 Standard Colorimetric Observer's colour matching functions

As for the current paragraph, the rest of this colour science review will only provide an overview of the major concepts involved. For in-depth information, interested readers are invited to turn to the reference textbook published by Hunt (Hunt, 1998).

2.2 Colour spaces

A colour space is a space whose system of coordinates map to colour perceptions. Since the human visual system is trichromatic, colour spaces are usually three-dimensional, although there are exceptions. The first colour space defined by the CIE is the *CIE 1931 colour space*. Colour coordinates are termed tristimulus values (X, Y and Z) and are obtained by factoring the spectral power distribution $P(\lambda)$ of the light emitted from a coloured object by the colour matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$ and $\bar{z}(\lambda)$ representing the sensitivity of human observers. This space forms the basis for all other colour spaces defined by the CIE.

$$X = k\int P(\lambda)\bar{x}(\lambda)d\lambda$$
$$Y = k\int P(\lambda)\bar{y}(\lambda)d\lambda \qquad \text{where } k \text{ is a scaling constant} \qquad (1)$$
$$Z = k\int P(\lambda)\bar{z}(\lambda)d\lambda$$

2.2.1 Uniform colour spaces

Despite its usefulness at quantifying colour perceptions, the *CIE 1931 colour space* lacks uniformity, i.e. equal distances in colour space do not necessarily correspond to equal differences in perceived colour. To remedy this state of affairs, two other spaces were introduced by the CIE in 1976, the *CIE 1976 L*u*v* colour space* and *CIE 1976 L*a*b* colour space*. Unlike the *CIE 1931 colour space*, these models do not provide an estimate of certain physical properties of stimuli, but rather attempt to describe its appearance by taking into account certain aspects of the viewing conditions, namely the adapting white point. Thus, colour coordinates in these spaces are correlates of the perceptual attributes of lightness, chroma and hue. Means of quantifying colour differences are also provided. While they offer significant improvements over the pre-existing model, they are still imperfect attempts at defining a uniform colour space.

*CIE 1976 L*a*b* equations*

$$\begin{aligned} L^* &= 116 f(Y/Yn) - 16 \\ a^* &= 500\left[f(X/Xn) - f(Y/Yn)\right] \\ b^* &= 200\left[f(Y/Yn) - f(Z/Zn)\right] \end{aligned} \qquad (2)$$

where $f(Y/Yn) = (Y/Yn)^{1/3}$ for $Y/Yn > 0.008856$ or $f(Y/Yn) = 7.787\,Y/Yn + 16/116$ otherwise.

*CIE 1976 L*u*v* equations*

$$\begin{aligned} L^* &= 116(Y/Y_n)^{1/3} - 16 && \text{for } Y/Yn > 0.008856 \\ L^* &= 903.3(Y/Y_n) && \text{otherwise} \\ u^* &= 13L^*(u' - u'_n) \\ v^* &= 13L^*(v' - v'_n) \end{aligned} \qquad (3)$$

where $u' = \dfrac{4x}{-2x + 12y + 3}$ and $v' = \dfrac{9y}{-2x + 12y + 3}$.

2.2.2 Colour appearance models

As previously mentioned, the appearance of a stimulus does not only depend on its physical characteristics, i.e. its spectral power distribution, but also on the conditions under which it is observed. In other words, a match between tristimulus values does not necessarily imply a match in appearance. Several other appearance phenomena should be included into the equation to describe precisely how changes in viewing conditions affect the appearance of a stimulus. An extensive review of such phenomena is given in Fairchild's book (Fairchild, 2004). In order to solve this conundrum, colour appearance models were developed to provide mathematical formulae transforming physical measurements of the stimuli and viewing environment into correlates of perceptual attributes of colour. The standard at the time of this writing is the CIECAM02 published by the CIE Technical Committee 8-01 (Moroney et al., 2002). The CIECAM02 model takes numerous inputs: the tristimulus values of the stimulus, the tristimulus values of the adapting white point, adapting background, and surround luminance information, and the degree to which observers are discounting the illuminant. The inclusion of several different areas of the visual field necessitates a precise geometrical definition of each of them, and their exact definitions will be given in the next chapter. Since it would be totally impractical to obtain a complete representation of the viewing environment, Hunt simplified the situation by defining four components of the visual field that play an important role in determining the appearance of colour stimuli (Hunt, 1998):

- *Stimulus*: its size is dependent on the application considered. Typically 2° for uniform colour patches.
- *Proximal field*: immediate environment of the stimulus extending about 2° from the edge of the stimulus in all directions.
- *Background*: environment of the stimulus extending about 10° from the edge of the proximal field in all directions.
- *Surround*: field outside the background.

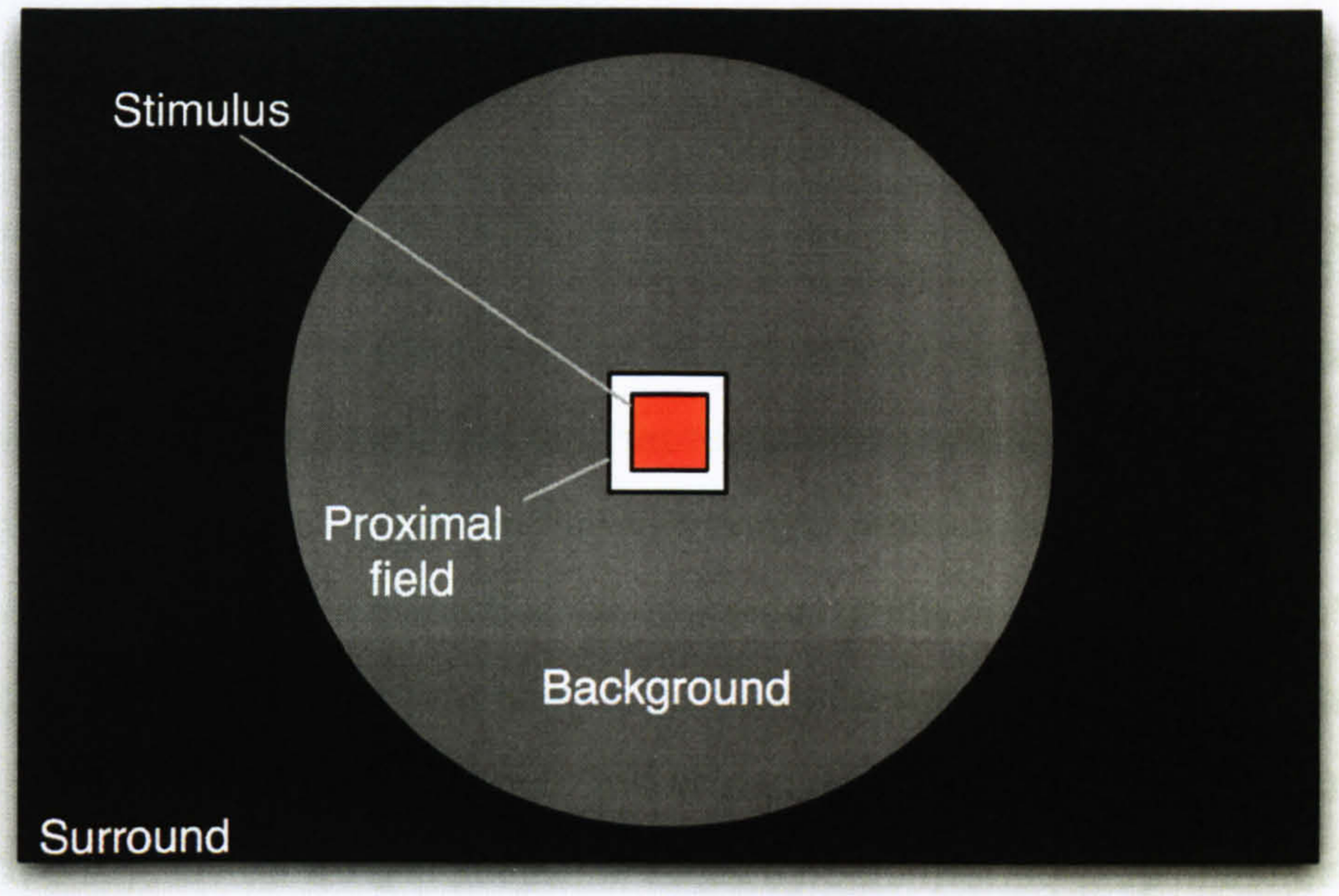


Figure 2 – Critical zones of the visual field

Their respective impact on the appearance of a stimulus will also be discussed further later on in this literature survey. The outputs of the model are correlates of the perceptual attributes: brightness, lightness, colourfulness, chroma, saturation, and hue, as well as hue composition.

2.2.3 Device characterisation

The diverse colour spaces described until now were specifically designed to provide numerical descriptions of human colour perception. While they achieved varying degrees of success, most of them are too complex for embedded industrial applications, where resources and computational power are scarce. Moreover, it is far easier to control the outputs of a colour reproduction device by manipulating directly the parameters of its underlying physical process. For instance, the outputs of halftone printers are specified pixel-wise by the coverage of the four primary colorants, while electrical currents control the amount of light emitted by monitors. This lead to the definition of *device-dependent* colour space, whereby colour coordinates are specified directly in units that are particular to a specific device. Transferring colour information between disparate devices therefore requires converting the information to a *device-independent* colour space. One of the perceptual colour spaces recommended by the CIE is normally used for that purpose. The

process of mapping device-dependent colour coordinates into device-independent ones is referred to as *device characterisation*, and it constitutes the heart of colour-management systems. The diverse colour characterisation models that were developed to characterise the devices employed during the course of this project will now be briefly introduced.

Device characterisation models can be broadly divided into two distinct categories. First of all, empirical models are derived from the combination of *"a large set of colour samples used in conjunction with some type of mathematical fitting or interpolation technique"* (Bala, 2003). An example of such models is polynomial regression, whereby the characterisation function is approximated by a series of polynomial expressions, whose coefficients are obtained by least-square fitting. Another common approach consists of interpolation techniques performed on a set of data sampling the colour space considered in a uniform manner. Among the diverse interpolation strategies available, tetrahedral interpolation is often considered as particularly efficient and accurate, since tetrahedrons are the smallest poyhedrons that can be defined on a three-dimensional regular lattice. Neural networks also offer another generic approach to characterisation, although their complexity forbids their use for computationally intensive applications (Masters, 1993).

The second category of characterisation models is physical models, which are based on a mathematical formulation describing the physical process by which the device generates colour. For instance, the phosphors luminance of CRT monitors can be accurately predicted by three parameters representing the gain, the offset and gamma of each electron gun (CIE, 1996). Similarly, halftone colour printers can be characterized thanks to the Neugebauer model, which predicts the reflectance of a colour halftone as a weighted average of the reflectances of all the combinations of primary colorants. The magnitude of the weights are derived from Demichel equations, whose inputs are the fractional area coverage of each individual colorants (Yule, 1967). The whole formulation of these models may be found in the cited references. To conclude, physical models are usually more advantageous than empirical techniques as *"they require fewer measurements and are thus less laborious and time consuming than empirical methods.* [...] *Model-based approaches generate relatively smooth characterization functions, whereas empirical techniques are subject to additional noise from measurements"* (Bala, 2003).

2.3 Colour difference formulae

The last major aspect of colour science that will be presented here is colour difference formulae. While providing an unambiguous and accurate representation of colour perception is important to ease its reproduction and communication, an equally critical challenge in colour science consists in estimating accurately the differences between two colour stimuli. It would otherwise be impossible to assess the quality of device characterisation and colour management systems, or to check whether industrial processes are behaving according to standards. The first two colour difference formulae recommended by the CIE are denoted ΔE^*_{ab} and ΔE^*_{uv}, which are simply Euclidian distances in respectively the *CIE 1976 L*a*b** and the *CIE 1976 L*u*v** colour spaces. As previously mentioned, both spaces exhibit a serious lack of perceptual uniformity, and new formulae were derived to rectify this state of affairs. Both the *CMC(l:c)* and the *CIE94* colour difference formula are modified ΔE^*_{ab} formula whereby the contribution from each channel is weighted differently.

CMC(l:c) colour difference formula

$$\Delta E_{CMC(l:c)} = \sqrt{\left(\frac{\Delta L^*_{ab}}{lS_L}\right)^2 + \left(\frac{\Delta C^*_{ab}}{cS_C}\right)^2 + \left(\frac{\Delta H^*_{ab}}{S_H}\right)^2} \quad (4)$$

where $S_L = \dfrac{0.040975L^*_{ab,l}}{\left(1+0.01765L^*_{ab,l}\right)}$ unless $L^*_{ab,l} < 16$ when $S_L = 0.511$

$$S_C = \frac{0.0638C^*_{ab,l}}{\left(1+0.0131C^*_{ab,l}\right)} + 0.638$$

$$S_H = S_C\left(T_f + 1 - f\right)$$

$$f = \sqrt{\frac{C^*_{ab,l}{}^4}{C^*_{ab,l}{}^4 + 1900}}$$

$$T = 0.36 + \left|0.4\cos\left(h_{ab,l} + 35\right)\right|$$

unless $h_{ab,l}$ is between 164° and 345° when $T = 0.56 + \left|0.2\cos\left(h_{ab,l} + 168\right)\right|$

CIE94 colour difference formula

$$\Delta E_{94} = \sqrt{\left(\frac{\Delta L^*_{ab}}{K_L S_L}\right)^2 + \left(\frac{\Delta C^*_{ab}}{K_C S_C}\right)^2 + \left(\frac{\Delta H^*_{ab}}{K_H S_H}\right)^2} \quad (5)$$

$$S_L = 1$$
$$\text{where } S_C = 1 + 0.045 C^*_{ab,1}$$
$$S_H = 1 + 0.015 C^*_{ab,1}$$

A major milestone was reached in 2000, when the *CIE* ΔE^*_{00} was published. Unlike its predecessors, the *CIE* ΔE^*_{00} is not a simple function of *L*a*b**, since colour coordinates undergo complex transformations before being passed on to the actual colour difference equation. This series of extra steps compensate for the lack of uniformity, which allows the *CIE* ΔE^*_{00} to outperform other colour difference formulae. Of interest, the CIECAM02 colour appearance model recommended by the CIE also includes a colour difference formula. It is once again a simple Euclidian distance between colour coordinates, but it offers a better prediction of perceptual differences as the underlying colour space is much more uniform.

2.4 Colour reproduction intents

The way success is defined in the imaging field depends greatly on the application involved. Hunt has identified several objectives of colour reproduction (Hunt, 1998). Whereas the first two, colour and pleasantness, do not make use of some recent findings in the colour imaging field, the third one, colorimetric, does. It aims at accurately duplicating the CIE XYZ colorimetric values across media by precisely calibrating and characterising them. That is how Hunt defines an exact reproduction (Hunt, 2004), i.e. where the reproduced colour image faithfully reproduces the absolute luminance and chromaticities of the original picture. However, this criterion is not sufficient to ensure an appearance match. The viewing conditions, i.e. the chromaticity of the white point, the luminance level, the characteristics of the surround and the size of the stimuli, have to be taken into account in order to achieve an appearance match. Colour appearance models have been designed to meet the needs defined by this next objective, appearance match. Unfortunately, they are not yet generic enough to be able to handle many practical situations. Indeed, they suffer from several shortcomings, such as a single adaptation model. Thus they do not predict well the appearance of a softcopy image seen in a typical office environment, because the adapted white point will be a mixture of the monitor's and the ambient light's (Katoh, 1995). It might even be necessary to alter the colorimetric content of the image even more, for aesthetic reasons or to compensate for various mismatches between the colorimetric

characteristics of the media, such as the colour gamuts they can reproduce. This last level is defined as preferred colour reproduction. Most of the research in cross-media colour reproduction, especially on the devices that are of interest in this study, has concentrated on either the differences in viewing conditions or the various mismatches existing between the devices' characteristics. But none of them have really attempted to verify whether there are any intrinsic differences when a human observer views an original image on one medium and its reproduction on another. In other words, is a difference perceived if identical stimuli are viewed on two distinct media, everything else being equal? There are some clues that it might not be the case. Cognitive information may, for instance, help the human visual system to adapt to the white point chromaticity of a medium (Berns & Gorzynski, 1991). But a more systematic research will be needed to provide an accurate answer to that question, which is the goal of the present project. More specifically, it will concentrate on the requirements of the Graphic Arts industry, where the need for an accurate simulation of a print on a monitor increases continuously. The rest of this literature survey will be devoted to the review of articles that may provide answers to these questions.

2.5 Cross-media colour reproduction

2.5.1 Spectral methods

Although most of the work performed in the cross-media reproduction field is based on trichromatic values, some researchers prefer to go beyond and use spectral data. For instance, Usui published the *Extended Phong Model* which aims at rendering printed images realistically on a monitor (Usui & Imamura, 1995). The original Phong model is an empirical model, whereby specular highlights are simulated with a cosine function raised to a power. For glossy surfaces, the magnitude of the power is large, hence the highlight is small and the intensity fell off rapidly. This model is widely used in the computer graphics field on RGB data, and the *Extended Phong Model* is simply a spectral extension. Based on the knowledge of the spectral reflectance data of the original image that is to be simulated (e.g. inks on paper), Usui proposed a model that combines the chromaticity of the specular and the diffuse reflection in order to obtain the final chromaticity of the stimulus. Using a ray-tracing method, the image can then be rendered on a monitor. In a similar way, Greenberg (1996) introduces a project that intends to generate synthetic images that are visually and

measurably indistinct from real-world images. A model of how light scatters when it strikes a surface has been developed. An attempt to enhance a rendering algorithm (radiosity) was made using statistical methods that reduce the amount of data that needs to be computed as this is usually tremendous. As accurate as these methods may be, they are not of practical interest in the framework of the present project, as their requirements, i.e. spectral data and computing power, are not usually available in a standard Graphic Arts environment.

2.5.2 Spatial factors

Among the most obvious differences between a monitor and a printer is the resolution at which they can render an image. Their maximum is obviously different, but that is not interesting for our study, which aims at comparing stimuli that have characteristics as close as is feasible. Laihanen (1994) studied the influence of spatial factors on the appearance of an image using a monitor and a printer. More specifically, he analysed how high-frequency and geometrically repetitive patterns affects integrated colorimetric attributes, i.e. values obtained not from a very small region of the image, but from a much larger zone, such as the whole image. Taking into account the spatial non-uniformity of the devices considered, he produced images containing horizontal and vertical lines as well as chess-board patterns at different spatial frequencies. The stimuli were shown under viewing conditions that did not correspond to ISO standard recommendations (ISO, 1999, 2000a), but that were closer to real life environments. His first finding was that the integrated luminance varies with the frequency of the pattern. But he also showed that there is a spatial orientation dependency, especially for the CRT monitor, where patterns containing horizontal lines do not behave at all like the rest. On the other hand, the printer behaves more or less the same for all patterns. But the most interesting finding was that the printer results are quite different from those of the monitor. A similar study on colours showed the same results. Spatial variations may greatly affect the average chroma and lightness, and even the hue of the monitor in some cases. Since all the results were obtained with geometrically repetitive patterns, Laihanen then attempted to evaluate the influence of spatial variations on pictorial images. Monochrome images showed that spatial factors may cause a small but significant luminance differences, but colour images were unfortunately not tested. The most significant discrepancy observed between the two devices was the spatial orientation dependence of the monitor, and an algorithm was implemented to

address that issue. Although it corrected the differences for images containing patterns, it is very interesting to note that it changed pictorial images so little that no visual differences were observed. So although the results deduced from this study are very interesting, their importance needs to be relativised. Indeed, this article was written in 1993, and the technology evolved very quickly. It might however be worth performing an approximate verification of the impact that spatial factors have on state of the art devices. On a more conceptual ground, this study concentrated on integrated colour attributes, whereas the practice of soft proofing requires the ability to resolve even the finest details. Furthermore, most of the images generated in the Graphic Arts industry are quite different from repetitive geometrical high frequency patterns, and the implemented algorithm almost did not affect such images. Such spatial factors should therefore not have a significant impact on the context of this project. On the other hand, Fairchild *et al.* (1996) studied the uniformity of CRT displays and its effect on the appearance of stimuli. Although the discrepancy between two regions of the screen can be different by as much as 20%, the human visual system is very insensitive to such differences, because they mainly occur in the luminance channel and their frequency is very low. The authors also mention that human sensitivity to low-frequency changes in colour is much higher. Although they chose to ignore them, the magnitudes of differences reported imply they are of a greater concern for the present project. This issue is not only associated with the self-luminous display (SLD), but also concerns the hardcopy image, since its illumination, though as uniform as a viewing booth can be, is very probably not perfectly uniform. Spatial factors, especially uniformity, will not only have to be analysed, but also critically compared for both media in order to ensure the accuracy of the reproduction.

2.5.3 Veiling reflections and gloss

Gloss and veiling reflections might play a much more important role in the present project. Veiling reflections are defined as *"the reflection of light by an image-bearing surface, that reduces the apparent contrast of the image"* (ASTM, 2003a). Such reflections are also often referred to as flare, i.e. *"light falling on an image, in an imaging system, which does not emanate from the subject point"* (ISO, 2000a). Mason reviewed the influence of surface properties on the appearance of colour (Mason, 1992). The imaging media that are used to display images do not unfortunately transmit only information about the stimulus

itself, but some unwanted noise may also appear due to the reflections occurring at the surface of the media. All imaging media suffer from surface reflections, which are characterised by two main features: directionality and wavelength independence. The former affect especially glossy prints illuminated by a directional light source, whereby the specular reflection will generate a mirror-like image of the source, strongly affecting the appearance at that specific angle. On the other hand, glossy surfaces are essential to preserve the full contrast and the saturation of colours. Although matt surfaces (and glossy prints illuminated by a diffuse light source) eliminate gloss, the noise from unwanted light is still present, in the form of flare created by ambient illumination. This results in a reduction of the achievable dynamic range, mostly in the dark parts of the colour gamut. Similarly, Johnson investigated the effects of gloss and veiling reflections on the colour appearance of prints (Johnson, 2002). Although modern colour appearance models greatly enhance the accuracy of colour reproduction when stimuli are viewed under two distinct viewing conditions, they do not take into account attributes of appearance that are associated with the surface properties of stimuli. Johnson's goal consists of performing a detailed analysis of that phenomenon, and incorporates such parameters in the next generation of colour appearance models if the need arises. These parameters being extremely dependent on the geometrical viewing conditions, as well as the illuminant configuration, a rigorous analysis of this phenomenon would require a tremendous effort. A pilot experiment was thus first performed by Johnson to assess the importance of the issue considered. A small number of observers were asked to assess the magnitude of appearance attributes of coloured samples produced on three different substrates with glossy, satin and matt finishes. The outcome is unambiguous: the perceived colour difference is much lower for glossy samples than for matt ones, especially for dark colours. Viewing conditions do have some influence as well. An important issue has thus been raised by this study, and the design of the experiments to perform would have to pay special care to that issue. More specifically, the choice of paper gloss for the hardcopy will be crucial. Glossy papers are obviously the most appropriate choice, but they make the design of the illumination more complex. It first has to be made of directional light sources, which reduces the uniformity of the illumination, and it also requires the walls to be covered with low-reflectance surfaces in order to avoid flare. The design for the monitor is much easier in comparison. Enclosing it in a volume covered with low reflectance surfaces eliminate both flare and gloss. However, in real conditions, a certain

amount of flare will always remain. In the best cases, that can be easily treated by a simple addition of the tristimulus values of the flare, whereas more sophisticated treatments are required for worse cases (Fairchild, 2004). Similarly, fluorescence is another factor that might addle the accuracy of the reproduction. This set of issues will certainly offer some interesting challenges before achieving a close match. Dark colours would need to be given special attention during the preparation of the experiments.

2.5.4 Background and surround conditions

Although the influence of gloss and veiling reflections has not been rigorously studied yet, the effect that surround and background have on the colour appearance of stimuli has been well investigated over the years. Fairchild provided a comprehensive review of the work that has been done on this topic (Fairchild, 1995). Most of the research performed in the imaging field on the influence of the surround concentrated on the perception of luminance, or more specifically on the difference of the optimal tone reproduction curve needed to reproduce a print and a transparency. More recent research performed by Fairchild tested the influence of surround on colour reproduction between a hardcopy image and a softcopy reproduction (Fairchild & Johnson, 1999). Observers were asked to memorise the colour of a printed pictorial image illuminated by a D50 illuminant with a high luminance level. They were then asked to move to a totally darkened area containing a D50- or D65-calibrated CRT, and to adjust the colorimetic attributes of the image using Photoshop. They could switch back and forth as desired, but needed to wait one minute to adapt to the new viewing conditions. Although the luminance and surround conditions were very different, the relative luminance of the softcopy plotted relatively to the relative luminance of the hardcopy follows closely the perfect linear tone reproduction curve, suggesting that there was very little surround or luminance level effect. Since the experimental viewing conditions will be equated as carefully as possible for the experiment in this study, the surround should not have any particular impact on the experiment's outcome. Bearing in mind that it mostly has an influence on contrast perception would nevertheless be wise.

Before background effects were included in the latest colour appearance model, Brainard performed a series of experiments on that topic, focusing his study on their influence on CRT monitors (Brainard, 1998; Brainard & Ishigami, 1995). Two observers were

asked to adjust the chromaticity of a test patch until it appeared achromatic. In order to test the influence of the context used to generate the background, the experiment was condicted for surface colours as well as computer-generated simulations. Both the luminance and chromaticity of the illuminant and the background were varied. The stimulus being a small rectangular patch, its background was provided by the monitor itself. The results showed clearly that the background has a strong effect on the location of the achromatic loci, but the effect of the illuminant, although being more modest, is significant enough not to be ignored. What is even more striking is the fact that this finding contradicts results obtained previously by Brainard under the same experimental conditions, but where the stimuli used were reflective surfaces. The influence of the background was then much smaller. In a second set of experiments, the test stimulus was enlarged to fill the entire monitor screen, and the rest of the setup remains the same. The effect of the illuminant being now much larger, the author concluded that the size and location of the patch are of primary importance in order to predict the influence of the illuminant on the colour appearance of the stimuli. Using carefully equated experimental conditions, softcopy stimuli were finally replaced by reflective surfaces having approximately the same size. The achromatic loci were then extremely close to the ones previously obtained.

From the last two experiments, it might be concluded that CRT stimuli and reflective surfaces are processed by the same visual mechanisms. However, the results of the first experiment contradict that conclusion, since the background effects obtained for a CRT softcopy are very different from those for reflective surfaces. Choh *et al.* performed an experiment very similar to Brainard's (Choh, Park, Kim, & Seo, 1996). The main difference was that the background was composed of random colour dots in order to mimic a pictorial image. Moreover, they specifically investigated the effect of the size of the background on the adaptation state. They concluded that the achromatic loci slowly converge towards the monitor white point as the background increases in size, starting from 5 degrees of visual angle, and reaches a stable point at approximately 12.5 degrees, a value that was also confirmed by Fairchild (1993). In Brainard's first experiment, the background was limited to 6 to 8 degrees. His results might therefore be explained by Choh *et al*'s conclusion. Or it might be due to the fact that the background, as defined by the CIE recommendations (1994), was then part of the stimulus itself, given the large size of a typical screen. Whereas

the luminance of the background was limited to 20cd.m^{-2} in the first experiment, the luminance of the stimuli tested in the second experiment ranged from 50 to 90 cd.m^{-2}. However, it is important to note as well that the illuminant conditions used by Brainard and Choh *et al.* (i.e. correlated colour temperature of 2500K, 10000K, and 9300K and several fluorescent lamps respectively) are very different from what is often used in the Graphic Arts industry.

Furthermore, Braun and Fairchild assessed the influence of a change of the background on the appearance of images while testing the ability of colour appearance models to effectively handle changes in viewing conditions for cross-media reproduction (Braun & Fairchild, 1997). The reproduction of six complex pictorial images, each subtending an angle of 16 degrees in the observers' field of view, were compared. Two identical viewing configurations (except that the background was either white or black) were employed. Although the conclusions mainly concern the performance of appearance models, Braun and Fairchild note that *"the effect of the background may be less for images than patches, and is certainly less as the area of the background decreases and distance from the center of the image increases"*. It would therefore be interesting to design a pilot experiment that estimates the significance of all these conclusions. Its outcome would help to specify the parameters for the present experiments.

Oskoui performed an experiment where standard illumination was used (Oskoui & Pirrota, 1998). She studied the state of adaptation of observers viewing stimuli on a CRT with different background configurations under D65. Although her ultimate goal was to find a background that did not affect chromatic adaptation in order to determine the adapted white point of a CRT in typical office conditions, the results are nevertheless directly relevant to this study. Indeed, chromatic adaptation for reflective surfaces viewed under D65 illuminant is complete. Oskoui also attempted to determine the best background conditions that would yield the same results for a softcopy stimulus. Several backgrounds were generated with different characteristics – pixel size, chroma and lightness range – in order to mimic an ordinary computer background. However, the average lightness for all backgrounds was the same. The observers were asked to perform an achromatic matching experiment through an iterative process. However, unlike Brainard's experiments, multiple uniform patches were displayed at the same time, and the background generated did not fill

the entire screen. The results clearly show that adaptation is never complete, although it is very close (98%) for achromatic backgrounds. This result is independent of the pixel size used to generate the background, so a solid colour or a textured background should behave in the same way. However, chromatic backgrounds, even if they average to grey, give a much lower chromatic adaptation (93%). Furthermore, the perceived appearance of an object does not only depend on its immediate background, but also to a certain extent on remote chromaticities. The study of the magnitude of induction generated from remote areas of the visual field has been an active area of research in recent years. *"In a color matching task, Tiplitz Blackwell and Buchsbaum (1988) found induction effects when a colored surround was not adjacent to the test field. Induction by remote inducers is qualitatively different from induction by adjacent inducers (Wesner & Shevell, 1992). Brenner, Cornelissen, and Nuboer (1989) used remote inducers of different geometries and reported that the magnitude of induction was essentially unrelated to distance, surface area or total edge length of the inducers. Jenness and Shevell (1995) showed that even very sparse chromatic information could alter the appearance of a test field. Nonlocal interactions and nonlinear spatial summation have also been reported for the induction of color contrast (Singer & D'Zmura, 1994; Zaidi, Yoshimi, Flanigan, & Canova, 1992)"*. More recently, Wachtler et al. (2001) reported interesting results on that topic. The authors studied how the presence of colour fields not adjacent to the test stimulus influences the magnitude of the induction generated by a change in the chromaticity of the background. *"The magnitude of induction by the background color was modulated on average by 7.6% by chromaticity changes in the remote color fields.* [...] *The spatial range of these chromatic interactions extended over at least 10 degrees from the fovea"*. However, this is unlikely to have an impact in our case , as chromaticity changes in the remote fields had virtually no inducing effect when they occurred without a change in background color. The background therefore does influence the state of chromatic adaptation, and moreover the adaptation seems to be different depending on the medium used. Since the main purpose of this study consists of checking whether differences in appearance between media come from the media themselves, equating the background and surround conditions should be sufficient for that purpose. That might be easily feasible for the surround conditions, but a very careful approach will be needed to specify the parameters of the background configuration.

2.5.5 Chromatic adaptation

The Graphic Arts industry has greatly benefited from the advances made in the field of colour appearance. Modern colour appearance models allow users to achieve accurate predictions of colour across different devices and a wide range of viewing conditions (Fairchild, 2004; Hunt, 1998; Sueeprasan, Luo, & Rhodes, 2001). Two of the three parameters related to the viewing conditions have been introduced so far in the studied context: the luminance of the surround, and the colorimetric attributes of the background. The last one required by colour appearance models consists of the colorimetric values of the adapted white, which will be the main focus of this section.

Modern colour appearance models were mostly derived from studying uniform reflective surfaces. And, as was mentioned in the introduction, that inheritance entails some serious drawbacks on their current capabilities. According to the CIE guidelines for coordinated research to test colour appearance models for cross media colour reproduction (Alessi, 1994), it is assumed that observers have a single steady-state of adaptation. But that does not correspond to realistic situations. Softcopy images are most often viewed in an illuminated room, which implies that observers' state of adaptation is affected by both the monitor's white point and the ambient illumination. Furthermore, observers do not strictly stare at the screen all the time. The state of adaptation commonly found in practical situations is therefore not steady, and it is not known whether it is single or not. It therefore does not comply with the requirements for applying colour appearance models. In order to address this issue, the CIE Technical Committee 8-04 was formed, and its terms of reference are *"to investigate the state of adaptation of the visual system when comparing soft-copy images on self-luminous displays and hard copy images viewed under various ambient lighting conditions"*. Although that does not correspond exactly to the requirement of the present project, the topic is nevertheless closely related, the only difference being the illuminant configuration. Whereas the white point of the monitor and of the ambient illumination were different for the studies performed by the members of this committee, the viewing conditions will have to be carefully equated for the present study – the monitor might not even be illuminated. However, the ultimate goal is the same, i.e. achieving a perfect reproduction of colour between a CRT monitor and a print. Each pathway will surely benefit from the experience and results accumulated by the other, such as the experimental

design or the psychophysical methods employed. Furthermore, the present project would very probably need to use results obtained by this committee at a later stage. An extensive literature survey of the articles published related to mixed adaptation was therefore performed.

Results relating to this section have partly been introduced in the previous section. Indeed, in Brainard's experiments (1995), both the illumination and the background configurations were varied. Although it was concluded that the adaptation ratio was very dependent on the background physical size, the effect of the ambient illumination was nevertheless significant. For the small background, a shift of approximately 10 to 20 % in CIE xy chromaticity coordinates was observed towards the chromaticity of the illuminant, and this shift was much larger when the background was enlarged to fill the entire screen. Similarly, Choh *et al.* also investigated the effect of ambient illumination on the state of adaptation of observers looking at a CRT (Choh, Park, Kim, & Seo, 1996). Once again, the shift ratio was between 10 and and 20 %, and was dependent on the luminance level of the ambient light. A model of the shift ratio was even provided.

If the state of adaptation when viewing a softcopy image in normal office conditions is a mixture of the monitor and the ambient illumination white points, colour appearance models should clearly be modified to take this fact into account. Several studies therefore attempted to investigate the benefit of incorporating mixed adaptation in existing colour appearance models, or to build new models from original ideas. Katoh, the chairman of the CIE Technical Committee 8-04 and the most prolific author on this topic, proposed a new chromatic adaptation model called S-LMS (Katoh, 1995; Katoh & Nakabayashi, 2001; Katoh, Nakabayashi, Ito, & Ohno, 1998). Chromatic adaptation transforms are the first component of modern colour appearance models, and his approach consists of creating a compatible method that could ultimately replace the existing chromatic adaptation model of CIECAM97s. It contains three stages, the first one being the most interesting in the context presently studied. Indeed, it aims at compensating for the contrast difference caused by the ambient light, or more specifically its reflection, on the CRT screen. This reflected light defines the darkest black the monitor can reproduce. If the black of the CRT display is not dark enough, the perceived contrast of the softcopy will look weaker than its hardcopy equivalent, since the human visual system is more sensitive to dark areas and less to light

ones. In order to overcome this issue, his proposed solution simply consists of adding the reflection of the ambient light to the colours produced by the phosphors of the display, and incorporate it in the characterisation model. The rest of the model is relatively straightforward. The device dependent colorimetric values are transformed in tristimulus values, which are then transformed into cone signals, and a simple Von Kries adaptation is finally performed. The only notable difference with other common chromatic adaptation transforms is that mixed adaptation is now added to the incomplete adaptation formulae to obtain the final adaptation white point used by the Von Kries transform. In order to find the optimal adaptation ratio, a paired-comparison experiment was performed to compare how hardcopy pictorial images accurately reproduce an original softcopy. The ambient illumination was provided by a fluorescent F6 light source at two different luminance levels, and three different white points were tested for the CRT. In all cases, the best reproduction corresponded to an adaptation ratio between 40 and 60%. It is interesting to note that single models of adaptation, whether that was to the monitor (CIE XYZ match) or to the ambient light (CIE L*a*b*), had the two lowest scores. The results also suggested that the adaptation ratio was independent of the monitor white point and the level of luminance. Similar results were also found when comparing two softcopy images illuminated under different illumination conditions. However, the optimal adaptation ratio, approximately 50%, was significantly higher than what was previously found by Brainard and Choh *et al.*

Berns and Choh, using an adapted version of the RLAB model, also found that mixed adaptation significantly improved the accuracy of colour reproduction (Berns & Choh, 1995). Hardcopy reproduction of a D65 balanced softcopy pictorial image viewed under a F2 illuminant were always preferred when an adaptation ratio of approximately 50% was used to generate them. Furthermore, they also verified that the most accurate reproductions were preferred to the rest of the reproductions when viewed without the original softcopy. Although the accuracy of the reproduction is vital, it is indeed important as well to test whether the images look appealing, since their final uses are mostly as stand-alone images. The adaptation ratio they suggest is therefore once again different from what Brainard and Choh *et al.* obtained. Katoh provides a satisfying explanation, which uses the time course of the chromatic adaptation, studied by Fairchild and Reniff (1995), as the main argument. Indeed, this mechanism is very quick, and the human visual system only requires few

seconds to reach 60% of adaptation to a change in the white point chromaticity. Therefore, when judging reproductions in a paired-comparison experiment, the state of adaptation of an observer would quickly shift between the CRT and the ambient light white point, which explains why the most accurate reproduction correspond to the image generated with an average adaptation ratio. The adaptation state is not fixed and thus less complete, whereas it is the opposite for the achromatic matching studies Brainard and Choh *et al.* performed, which explains why the magnitude of the observed shift was smaller.

Henley and Fairchild also undertook an extensive experiment in order to assess how mixed adaptation influences the quality of colour reproduction across disparate viewing conditions (Henley & Fairchild, 2000). The experiment was performed in accordance with the CIE guidelines for coordinated research on colour appearance models for reflection print and self-luminous display image comparisons, and the quality of reproduction of softcopy images displayed on a 9300K balanced CRT monitor was compared to an original print illuminated by a standard D50 illuminant. Four different chromatic adaptation methods were tested, and once again, the results confirmed that mixed adaptation improves the quality of reproduction for all the models tested, CIECAM97s included. Moreover, the results show that a simple Von Kries adaptation provides the best predictions of all, implying that simplicity does not mean lack of accuracy. This is especially relevant regarding the results achieved by the CIE recommended colour appearance model, CIECAM97s. Indeed, the introduction of mixed adaptation improved moderately its performance, whereas optimising the D factor, which determines the degree to which the illuminant is discounted, entailed results that were almost better than the Von Kries method. Katoh himself assessed the importance of the incomplete adaptation factor D in one of his experiments (Katoh & Nakabayashi, 2001). It is interesting also to note that the incorporation of mixed adaptation was more effective with simultaneous comparisons, confirming once again Katoh's suggestion.

Sueeprasan and Luo went further and tested the performance of Katoh's model vs. the standard CIE chromatic adaptation transforms (Sueeprasan & Luo, 2001). The originality of their approach is that they did not modify the transforms to incorporate mixed adaptation, but evaluated the influence of degree of adaptation by varying the incomplete adaptation factor (D for CIECAM97s for instance). The experimental design once again

followed the guidelines previously introduced, and six illuminant conditions (illuminant D50, A and fluorescent, each at two luminance levels) were used to illuminate the printed original while the monitor's white point was set to 9300K. Quite remarkably, it was found that the human visual system was adapted between 40 and 60% to the monitor white point. Recall that Katoh found the same adaptation ratio between the monitor and the ambient illuminant's white point. Furthermore, the best performances were obtained by the latest adaptation transform, CMCCAT2000, and the S-LMS model, although specifically designed for that task, did not achieve results as good as that. Therefore, by choosing the incomplete adaptation factor carefully, a chromatic adaptation transform may be sufficient to handle mixed adaptation scenarios. It was also confirmed that the state of adaptation is independent of the illuminant and luminance level used. Similar results were obtained by Laine in an experiment aiming at finding the optimal parameters for CIECAM97s (Laine, 2002). A much smaller shift of the adaptation degree was observed for a set of viewing conditions, but that is probably only due to the fact that the luminance level used was as low as 2 $cd.m^{-2}$. The importance of the surround parameter *Nc* was also highlighted.

As the results obtained by Henley and Fairchild suggested, a simple matrix transformation such as Von Kries could approximate very closely the performance of the best colour appearance model, even the performance of their finely tuned version. Braun and Fairchild confirmed this fact (Braun & Fairchild, 1996). They generated a set of colour appearance image data independent of colour appearance models. Experienced observers were asked to adjust images on an SLD in order to match to original hardcopy as accurately as possible under several different viewing conditions. From the adjustments made by the experienced observers, a 3×3 matrix that best converted the image from the original to the reproduction viewing condition was determined using multiple linear regression. When observers compared the accuracy of the reproduction produced by this matrix with the predictions of colour appearance models, the images produced by that matrix were systematically preferred. It is however of importance to note that the tested colour appearance models are now well outdated. The task was further pursued by Fairchild and Johnson (1999), accumulating more data and testing newer colour appearance models. The best-fit matrices still obtained the best results, but now, the recommended CIECAM97s performed as well as them. However, the results of this model were still very dependent on

the choice of its parameters, as the surround compensation had to be turned off to achieve those performances. But very interesting conclusions emerged from a detailed analysis of the degree of adaptation of the experienced observers. Indeed, a perfect match in appearance implies that the CIELAB a* and b* value should match. If the CIELAB a* values of the original image are plotted against their corresponding value for the reproduction, the data should therefore approximately lie along the major diagonal of the plot, i.e. the linear function defined as $y = ux + v$ with $u=1$ and $v=0$. Any other behaviour would indicate a difference in the degree of adaptation. The plot of the CIELAB a* do not show any significant deviation, but a very specific trend can be observed in the plot of the CIELAB b* data. Indeed, the D50 data deviates quite significantly from the diagonal line, indicating that a CRT with a D50 white point retains a strong yellowish appearance. Such a colour balanced monitor should therefore not be used for critical colour comparisons. On the other hand, the D65 data still deviates from the diagonal line, but in the other direction and more slightly. This indicates that the D65 CRT retains a small amount of bluish appearance when compared to a print viewed under a D50 illuminant, which is assumed to look neutral. However, the observers mentioned that D65 CRT looked neutral, whereas the D50 balanced clearly appeared yellowish. In another experiment (Fairchild, 1993), although this fact is not mentioned in his analysis, the data reported by Fairchild corroborated this fact. Therefore, the popular assumption that adaptation is almost complete when a stimulus is viewed under a D65 balanced illuminant does not seem to precisely hold when the medium used is a CRT display. A similar behaviour is observed under illuminant A. Although the adaptation is never complete for that particular viewing condition, the adaptation is even less complete when the stimuli are shown as softcopy images (Berns & Gorzynski, 1991; Fairchild, 1993). It is as well interesting to note that this conclusion was obtained both from experiments using colour patches (Fairchild, 1993) and pictorial images (Fairchild & Johnson, 1999). Berns reports an interesting observation he made in a similar experiment (Berns, 1991). Observers tend to fixate on a CRT image, whereas their eyes move more freely when they deal with surface colour. However, that would mean that the adaptation to a SLD should be more complete.

2.5.6 Identical viewing conditions

The emergence of colour appearance models and the recent advances made in the studies of mixed adaptation are leading towards a better accuracy in cross-media reproduction, especially with the devices that are of interest to us – SLD and prints. Nevertheless, the latest results introduced in the previous section suggest that some discrepancies remain. This section will attempt to analyse the work that has been done concerning this issue, by introducing studies performed on cross-media reproduction when both viewing environments were equated.

The experiments introduced in the previous section suggest that the state of adaptation is not exactly the same depending on which media render the stimuli. Similar conclusions were deduced by Abe (Alessi, 1996). He performed an image matching experiment where the stimuli were composed of several colour patches presented on a neutral gray background. Two similar viewing conditions with D50 and D65 white point chromaticities were tested. The obtained data clearly show a significant difference between the CIELAB a* and b* values for some of the patches. It might be concluded that a colorimetric match, even performed under the same viewing conditions, does not imply a visual match when comparing SLD and prints. However, Abe does not report whether the luminance level matched for the two devices. That might not be the case, since the given the tristimulus value Y neutral gray background did not match, which might provide an explanation for the phenomenon discussed. This was also the case for the Fairchild and Johnson experiment (1999).

On the other hand, Lo *et al.* undertook an experiment that suggested different results (Lo, Luo, & Rhodes, 1996). Performed according to the guidelines specified by the CIE TC 1-27 committee, the performances of several colour appearance models were tested when reproducing six complex pictorial images. A white border surrounding the images and a neutral gray background were used for media, as well as identical white point chromaticities (both D50 and D65) and luminance levels. A forced-choice pair comparison psychophysical method was used to assess the tested models. All of them obtained the same rank order, but more interestingly, the errors associated were considerably higher than the ranking range. That clearly indicates the observers were not certain when making a choice, because the images were very close in appearance. Moreover, although all the methods obtained the

LEEDS UNIVERSITY LIBRARY

same rank, the model based on CIE XYZ matching obtained the best performance. When reviewing those results, Alessi therefore concludes that a colorimetric match is an appearance match when the viewing conditions across media are identical (Alessi, 1994), which seems to be confirmed by another article from Alessi (1992), or at least the introduction of that article made by the editor, since the article itself is not included in the proceedings.

As previously introduced, Braun and Fairchild undertook an experiment where experienced observers were asked to adjust a reproduction until it appeared as identical as possible to the original (Braun & Fairchild, 1996). More precisely, the reproduction was a softcopy in all cases, whereas the original was either a softcopy or a hardcopy. However, the viewing configurations were identical. D65, with the same level of luminance, was the illuminant employed, and the pictorial images were presented with a white border against a gray background, uniform for the hardcopy and composed random white and black dots for the softcopy. The images subtended an angle of approximately 17 degrees in the observer's field of view. The average CIE ΔE_{ab}^{*} for the softcopy-softcopy experiment was about 3.0 colour difference units, i.e. just as the threshold for the detection of a difference in colour appearance when comparing complex pictorial images (Stokes, Fairchild, & Berns, 1992). The difference was slightly higher, 3.7, for the softcopy-hardcopy comparison. However, such results cannot be used to critically compare the hardcopy and the softcopy reproductions, since different psychophysical methods (simultaneous and memory matching) were used for the two configurations. Furthermore, there was a small difference in white point chromaticities between the CRT and the viewing booth used to illuminate the images. A statistical analysis of the reproductions vs. their original shows that observers adjusted the lightness very accurately, whereas the coefficient of regression for chroma indicates a stronger dispersion. It is also necessary to stress the fact that the presented results were obtained by using only two images.

The accuracy of colour reproduction across media was tested also for uniform patches. Hence Herbert *et al.* aimed at achieving *"a high degree of color accuracy perceptually for a single color representation or a single stimulus on a display unit"* (Herbert, Kirkenaer, & Ladson, 2003). The original stimuli in this case were small plastic chips, both opaque and translucent, whose colours sample the gamut of the display device

approximately uniformly. Several illuminant configurations, including one with D65 white point chromaticities, were tested. The performance of the characterisation model of the display was assessed by measuring the CIELAB values of the chips and comparing them to the corresponding values on the SLD. The Pearson correlation coefficient was higher than 0.99 in each case, and the average ΔE_{ab}^{*} for D65 was 2,57. The simulated colours thus approximated very closely the original chips, and that was confirmed by the visual experiment. In two different locations, 85% of observers who were asked to judge the quality of the colour reproduction answered that there were no perceived colour difference between the object and its simulation. In order to achieve such results, the authors emphasized the care required in settings up the experiment. The SLD is obviously a very important component of the chain, and its luminance was identified as a crucial parameter. An LCD was therefore chosen, and its characterisation was done with extreme care to approximate very closely both the white point of the viewing booth and its brightness. No coloured areas were placed in the field of view of the SLD, and the whole experimental room was a neutral gray to "[minimize] *any effect disrupting the vision of the viewer*". Furthermore, providing a white reference (a frame around the chip) in the field of view of both media was identified as a crucial parameter, in order to help the observer to adapt chromatically to the given white point.

Rich *et al.* also attempted to visually quantify the accuracy of the CRT simulation of colour surfaces (Rich, Alston, & Allen, 1992). Observers were asked to estimate the magnitude of the colour difference between the original colour patch (taken from the NCS colour atlas) and its softcopy reproduction. A gray scale mounted in the viewing booth provided the scaling units. Both the surround luminance and the chromaticities of the white point (D65) were matched between the devices used. The accuracy obtained was 2.8 CIELAB ΔE^{*}_{ab}, i.e. identical to what Braun and Fairchild reported. This is very close to the just-noticeable difference threshold for complex pictorial images, and furthermore close to the theoretical limit that a CRT can achieve. Indeed, the differences between each step of the 8-bit, 256 addressable steps of a CRT are between 0.6 and 1.8 ΔE_{ab}^{*} colour difference units. The characterisation of the CRT is therefore crucial to achieve such a performance. But if it is, the author concludes that such a CRT can well be used to approximate the colour appearance of colour surfaces. However, a plot of the measured chromaticity of both the

original surface samples and their softcopy reproduction in the CIE u'v' diagram clearly shows a difference between the CRT and the sample, which leads the author to suggest that *"the observer is never quite fooled, that is, adapted, to source neutral point on CRT"*. The accuracy varied as well as the colour varied, but more especially between observers. It is especially surprising since all of them were experienced in taking part in industrial colour matching experiments. But the author points out that this is consistent with what he already observed and reported, i.e. the experimental noise due to observers is much higher for judgements involving SLD-based simulations.

Although obtaining an accurate colour reproduction is highly desirable, the accuracy of colour differences need as well to be reliable for a reproduction to be qualified as "successful". In that context, Rich *et al.* also attempted to estimate the accuracy of colour difference reproduction, using the same experimental design. The conclusions do not differ from what was previously stated. A simulation on a CRT of the colour differences of pairs of colour surfaces can also be reasonably well approximated by a softcopy reproduction. Berns undertook an experiment that aimed at the same goal, i.e. investigating the feasibility of simulating the perceptual difference of pairs of surface colours on a CRT (Berns, 1991). Observers were first asked to quantify the difference of simulations on a CRT viewed with and without ambient illumination, and finally to repeat the same experiment with the real surfaces, which were acrylic-lacquer samples. All the pairs were generated around a single green center and in seven directions in CIELAB. A gray anchor pair of CIELAB ΔE^*_{ab} 1.0 was used to estimate the magnitude of the differences. Observers were simply asked whether the colour difference of the pair of stimuli was greater or less than the anchor difference. The chromaticities of the light booth used to illuminate the physical samples and of the CRT were both D65. A statistical analysis concluded there was no significant difference when judging CRT simulated or surface samples colour differences. However, it indicates as well that observers were more uncertain when making judgements of the simulation. Factors explaining this conclusion range from gloss, since the softcopy does not contain any gloss information, sharpness, the resolution of the CRT used being quite poor, or viewing angle, which was not perfectly matched, or even cognitive issues, which will be discussed later. All in all, the feasibility of simulating colour difference of surface samples using a CRT monitor was confirmed. Similarly, Choi pursued an experiment that investigated the efficiency of

colour difference equations for the prediction of moderate and large perceptual differences (Choi, 1994). Previous work was performed on hardcopy images, more specifically photography prints, and Choi mimicked the experimental setup using SLD to display the stimuli instead. The experiment was performed in a dark environment, and the white point of the display approximated the D50 standard illuminant with an average luminance level. Observers were asked to quantify the colour difference of two chromatic test patches relative to a reference set which was displayed alongside them. 320 test patch pairs were divided into eight main colour centers, which included blue sky and skin tone. An extensive analysis of the collected results was compared with the results obtained during the previous experiment. Once again this did not show any significant difference in the perception of the colour difference.

MacDonald *et al.* also reported intriguing results concerning the simulation of hardcopy images (MacDonald, Morovic, & Xiao, 2002). They studied the influence of the medium used to generate a reproduction on the performance of several gamut mapping algorithms. In the first experiment, observers were asked to choose from among hardcopies the one that reproduced most accurately an original image viewed on a CRT. The second experiment was similar, except that the hardcopy reproductions were now simulated on another display. The viewing conditions used were standard, and special attention was given to carefully equate them. Furthermore, the gamut of the print was not very different from the CRT's used to generate the reproduction, even at both extremes of the lightness range. It would then be reasonable to expect similar results for each experiment. But that is not the case. The results obtained are very different. The correlation of the scores obtained by each gamut mapping algorithms between both experiments is very low (ρ=0.19). Unfortunately, no measurement of the colorimetric data was performed, which forbids the quantification of the difference between the print and its simulation. Furthermore, the authors report as well the significant inconsistency of results published in the gamut mapping field. Indeed, a clear division may be seen between the studies that used real media and the ones using simulation. Although many other factors may as well explain such a discrepancy, this is still striking.

2.5.7 Cognitive Issues

Of the differences between SLD and reflective surfaces, the most obvious is their different mode of appearance (Jones, 1963). A mode of colour appearance is "*the various kinds of context or setting in which color is perceived*". The usual descriptions or definitions of different modes simply consist in referring to the physical or psychological conditions under which they are experienced. For instance, a print is a reflective surface being illuminated, and is thus viewed in the surface mode of appearance. On the other hand, a SLD corresponds more to the illuminant mode of appearance, because the colour is perceived as belonging to a primary or secondary source of light. Three other mode of appearance exist, aperture or film when a colour is seen through an aperture (i.e. in isolation from others), volume when transparency occurs and illumination for elements such as shadows and smokes. Modes of colour appearance are of fundamental importance for the specification of colour appearance, because a change in mode entails a change in appearance. For instance, predictions made by modern colour appearance models are only valid for a certain subset of those modes. Two main subsets were identified as being crucial for colour appearance: related and unrelated colours (Hunt, 1998), i.e. colours perceived to belong to an area or object seen respectively in relation or in isolation to other colours. For instance, the CIECAM97s model was only designed to handle the appearance prediction of related colours, of which the surface mode of appearance provides a good example. Unrelated colours, perceptions belonging to the aperture mode of appearance for instance, required a difference model of colour appearance specially adapted for this task, such as the CAM97u proposed by Hunt.

Since the stimuli used in the experiment related to this project will be presented through an aperture, it might be thought that their mode of appearance will shift to the aperture mode. However, this mode and its corresponding model specifies clearly that it is restricted to the appearance of small and simple colour patches viewed in very low luminance level environment, usually completely dark fields. Those constraints are not compatible with the requirements for the present experiment, which specify fairly large images and achromatic background and surround. The stimuli will thus not be seen in the aperture mode.

According to the experimental guidelines (CIE, 1994, 1995), everything will be made to remove any cognitive cues that could discriminate the media. That should not affect very much the appearance of the hardcopy, which will be viewed in an environment very closely related to its natural one. It is thus to be expected that its mode of appearance will probably remain stable. However, it is likely that this will not be the case for the softcopy. Indeed, the ambient illumination used to generate the background and surround stimuli will have the same chromaticity as the white point of the SLD. Furthermore, the SLD will be protected from the viewing flare created by the ambient illumination by the screen used to generate the surround, which will also hide the SLD border. Therefore, most of the information (or at least as much as can be estimated at this early stage) that indicates the medium is a SLD will be removed. Such a drastic change in the external conditions will probably entail a shift or conversion of the mode of appearance, from illuminant to surface. It might be argued that this change will alter the intrinsic essence of the stimuli generated by the SLD, i.e. as if they were not produced by a SLD, and therefore seriously impair the pertinence of the comparison between the two media. On the other hand, it has the advantage of removing as many parameters as possible from the already complicated equation. Providing that the general framework generates a close match between the two tristimulus values, an answer to the considered problem should be easily reached. Furthermore, if a match in tristimulus values is an appearance match, tracking the influence of every parameter(s) on the accuracy of the appearance match can be performed in a mode controlled way, since the cognitive cues can be added or removed at will.

The mode of colour appearance does not in general depends only on external conditions, but also on the internal mental set of the observer. *"When the mode is doubtfully or ambiguously determined by the external context, the internal factor becomes especially important"* (Jones, 1963). The stimuli generated using the softcopy will be likely to beget such confusion among observers. This fact will thus highlight individual variability and maximise the inter-observer repeatability, providing that the match between the tristimulus values between media is close enough. The set of instructions given to observers will therefore play a crucial part in the outcome of the experiment, as was demonstrated by Fairchild (1993). In his experiment, observers were instructed to modify the softcopy reproduction of an object, which looks achromatic under a tungsten light, until it looks like

the object itself. The chromaticities of the resulting patch were very close to those of the original object. But the reproduction did not look achromatic at all. Similarly, a CRT was divided vertically in two parts, one simulating a D65 illuminant and the other a tungsten source. When asked to match a patch from one side of the display to the other, observers performed a simple colorimetric match. But when asked to match the appearance of the patch, the observers behaved as if they were now adapted to each part of the display. The outcome of the experiment was thus determined by the mental set created by the instructions. For the present experiment, their design will have to be done in a very careful manner in order to avoid any influence on the observers' mindset. It is indeed very important that her/his approach towards the experiment is very candid.

The mode of colour appearance can be influenced by changes in the set of external conditions, for instance by removing cognitive cues as the experimental setup states. It is also possible to modify the mode by doing the exact opposite, adding cognitive cues to the context in which a stimulus is viewed. Fairchild provides an interesting description of such a phenomenon (Fairchild, 1993). As was previously stated in previous sections, the state of chromatic adaptation is different depending on the medium involved. Whereas it is complete for a reflective surface illuminated by a D65 illuminant, this is not the case for a similarly colour balanced SLD, and this fact also holds for other illuminant such as standard illuminant A. Fairchild set up an experiment to compare the state of adaptation when hands are added to the stimulus background in such a way that they seem to hold it. The results obtained with illuminant A colour balanced CRT clearly indicate that, although the adaptation is still not totally complete, it is nevertheless much more complete than when the hands are not present. Similarly, the state of adaptation changes from mostly complete to totally complete when using a D65 colour balanced CRT. Berns also reports that colour difference pairs are judged differently depending on whether a light booth is displayed on the CRT or not (Berns, 1991). Rich further supports this idea by stating that results discrepancies obtained for the two considered media can be explained by the fact that "[...] *the mode of viewing of the CRT simulation never fully approaches surface mode*" (Rich, Alston, & Allen, 1992). Please note that the term "mode of colour appearance" is used in its strict sense in the present paragraph, which states that a change in the external conditions, i.e. the very definition of mode of colour appearance, modifies the state of chromatic

adaptation. That does not mean the mode of appearance itself is changed. A change in appearance can be explained by a shift or conversion of mode of appearance, but also by a shift in the same mode of appearance.

Chromatic adaptation plays a crucial role in the prediction of the colour appearance of stimuli. Having been widely studied, it is well established that it is influenced by two types of mechanisms, sensory and cognitive. The former corresponds to mechanisms that respond automatically to the stimulus energy, whereas the latter corresponds to responses that are based on the observer's knowledge of the scene content, as illustrated by the previous paragraph. Under typical viewing conditions, sensory mechanisms are not capable of complete chromatic adaptation, and thus cognitive mechanisms are necessary to complete the process. Observers perceive the colour of objects as if the adaptation to the chromaticity of the illuminant were almost complete, a process known as *discounting the illuminant* (von Helmholtz, 1866). However, this does not occur when viewing softcopies since there are no illuminated objects. Modern colour appearance models therefore include this cognitive effect in their set of parameters. As a matter of fact, CIECAM97s contains a parameter D that controls to which extent the chromaticity of the illumination is discounted. In typical office scenarios, tuning this parameter can provide appearance predictions as accurate as those given by appearance models including a mixed chromatic adaptation, as the results presented in the previous sections demonstrated. Though it seems clear that the human visual system is not completely adapted to the chromaticity of a SLD in a typical office environment, it is still not very clear whether its state of adaptation is simply incomplete or a mixture of both the SLD and the ambient illumination. Given the characteristics of the viewing conditions that will be used for the present experiment, it seems safe to assume that the chromatic adaptation of observers will be totally complete. Bearing in mind that this might not be exactly true would thus be wise depending on the outcome.

2.5.8 Metamerism

Metamerism is an important and difficult issue to deal with in colour reproduction. Fortunately, the configuration adopted for the present experiment will greatly reduce it. Indeed, the similarity imposed on the viewing conditions entails that both the illuminant and the size of the stimulus will have no effect. But observer metamerism remains. Fairchild

carried out an experiment that aimed at estimating the importance of observer metamerism in cross-media reproduction (Fairchild & Alfvin, 1995). Observers were asked to adjust the chromaticities of a patch displayed on a CRT until it matches the colour of an original, which was either a print or a transparency. Seven different colours were used for each original medium, and they were illuminated with a standard D50 illuminant. The inter-observer variations reported could be as high as 19.7 colour difference units. But what is even more striking is that the 95% confidence interval of colour mismatches obtained is about three times greater than what the CIE Standard Deviate Observer predict. The author notes that the method used by the CIE may compress too much the variance. However, Fairchild concludes by suggesting that the CIE recommendations on observer metamerism may need to be reviewed for cross-media applications. But the medium difference could once more provide an explanation, as for chromatic adaptation. Indeed, the CIE predictions were obtained from observations of reflective surfaces. Furthermore, Rich reported that the level of observer metamerism is increased when the task consists of simulating a colour on a CRT monitor (Rich, Alston, & Allen, 1992). On the other hand, more recent results obtained by Oicherman *et al.* suggest that "*in soft-proofing conditions, colour discrimination mechanisms are more dominant than colour matching ones, rendering the phenomenon of observer metamerism less significant than in the conditions of a classical colour matching experiment. The consequence is that any model of uncertainty of colour matching based on variability of colour matching functions (such as the CIE Standard Deviate Observer) would fail when applied to soft-proofing*" (Oicherman, Luo, & Robertson, 2006). While the influence of observer metamerism appear to be lessened, they also report on another issue that is of critical importance to soft-proofing, the failures of colorimetric additivity. "[W]*hen colour matching functions measured with narrow-band lights are applied to the prediction of metameric matches between narrow-band and broadband stimuli, the calculated "blue" tristimulus values of the mixture of narrow-band stimuli are smaller than the ones set by visual colour matching of the broad-band stimulus. The light emitted by the LCD and CRT displays is in part narrowband due to the red primaries in CRT and red and green primaries in LCD. Therefore, if this assertion is correct, they are affected by this phenomenon*". However, such discrepancies are mostly limited to the bluish part of the colour gamut, and have a "*minimal or non-existent*" effect elsewhere.

2.5.9 Other devices

Although most research on cross-media colour reproduction in the last decade have focused on hardcopy-softcopy reproduction, several contributions were also performed with other devices. For instance, Fairchild *et al.* carried out an experiment to test the accuracy of reproducing a colour stimulus from a CRT monitor to projected 35mm slides (Fairchild, Berns, & Lester, 1996). This is especially difficult since actually three stages are required: CRT, film printer and finally projection system. The fitness of several colour appearance models were tested to account for changes across quite different conditions, from the dim conditions typical of a CRT to the dark environment of a projection system. Although the results concerning the performance of the colour appearance models are outdated, the key point stressed by the authors is the importance of the characterisation of the three processes, without which any decent results cannot be obtained. This is likely to also be the case in the present experiment. Unfortunately, no information regarding the intrinsic differences between the media involved is provided.

On the other hand, a study performed by Kwak *et al.* focused more on that specific topic (Kwak, MacDonald, & Luo, 2001). The colour appearance difference caused by media difference, LCD projectors and LCD monitors as it happens, were investigated alongside the performance of several colour appearance models for changes in the luminance level of backgrounds. The luminance level was set to a higher value for the projector, but the rest of parameters, including white point chromaticities, were identical for both media. Observers were asked to assess the colour appearance of a series of colour patches displayed in a complex viewing field formed by a random pattern of colour patches. Results showed that there is no strong evidence that the appearance depends on the device used to generate the stimuli. The only noticeable difference concerns the colourfulness, which was perceived as higher for the projector. Its luminance settings were set to a value higher than what was used for the monitor, which can explain the discrepancy observed. However, it was found once more that the effect of the background on the colour appearance was different from one device to another. The authors suggest that the overall size of the screen might have such an effect. It is not clear though whether the visual angle subtended by the projector was the same to the monitor.

2.6 Psychophysics

Psychophysics is defined by Fairchild as *"the scientific study of the relationships between the physical measurements of stimuli and the sensations and perceptions that those stimuli evoke"* (Fairchild, 2004). Credit to its importance to the field of colour science, the large majority of the research work introduced previously employed in some ways experimental methods and tools emanating from the field of psychophysics. The classical framework to perform psychophysics is embodied by the *threshold theory*, whose main historical focus was the study of sensory systems. In this context, the sensitivity of human observers to a certain class of stimuli is expressed in terms of a threshold, be it absolute or differential. An absolute threshold is the level of intensity of a stimulus at which the subject is able to detect the presence of the stimulus some proportion of the time, typically 50%, whereas a differential threshold corresponds to the difference in stimuli that subjects notice. Since the first objective of this project consists in detecting the potential appearance differences generated by disparate colour reproduction media, the threshold theory appears to be a perfectly appropriate choice. However, it does suffer from certain limitations, which will be briefly mentioned below, and another framework, the *signal detection theory*, has now superseded the traditional approach. The major concepts of the signal detection theory will be introduced in the following subsection. Another interesting aspect of psychophysics is the ability to produce scales of human perception of certain features of physical stimuli, which will also be presented next.

2.6.1 Signal detection theory

The signal detection theory is *"a general psychophysical approach to measuring performance"* (Macmillan & Creelman, 2004). It consists in a collection of psychophysical procedures and analytical techniques to perform diverse experimental tasks, and it was derived to overcome the limitations of the traditional threshold theory. Central to the theory are the concepts of *signal* and *noise*. Any stimulus, i.e. signal, is always perceived in the presence of noise, no matter how precisely the experimental conditions are controlled. Noise has two fundamental properties: it varies randomly, and can be either internal (within the observer's sensory system) or external (any other factors affecting the stimulus). The superiority of the SDT actually comes from the specific treatment reserved to noise. The

signal (or stimulus) also varies randomly, as what observers actually perceive is a combination of the stimulus itself and to which the noise signal is added. The task of the observer is therefore simple, i.e. decide whether an observation is due to noise (N) only or to a combination of signal and noise (SN). Since both events vary randomly, the whole concept can be represented graphically by two probability distributions (Figure 3).

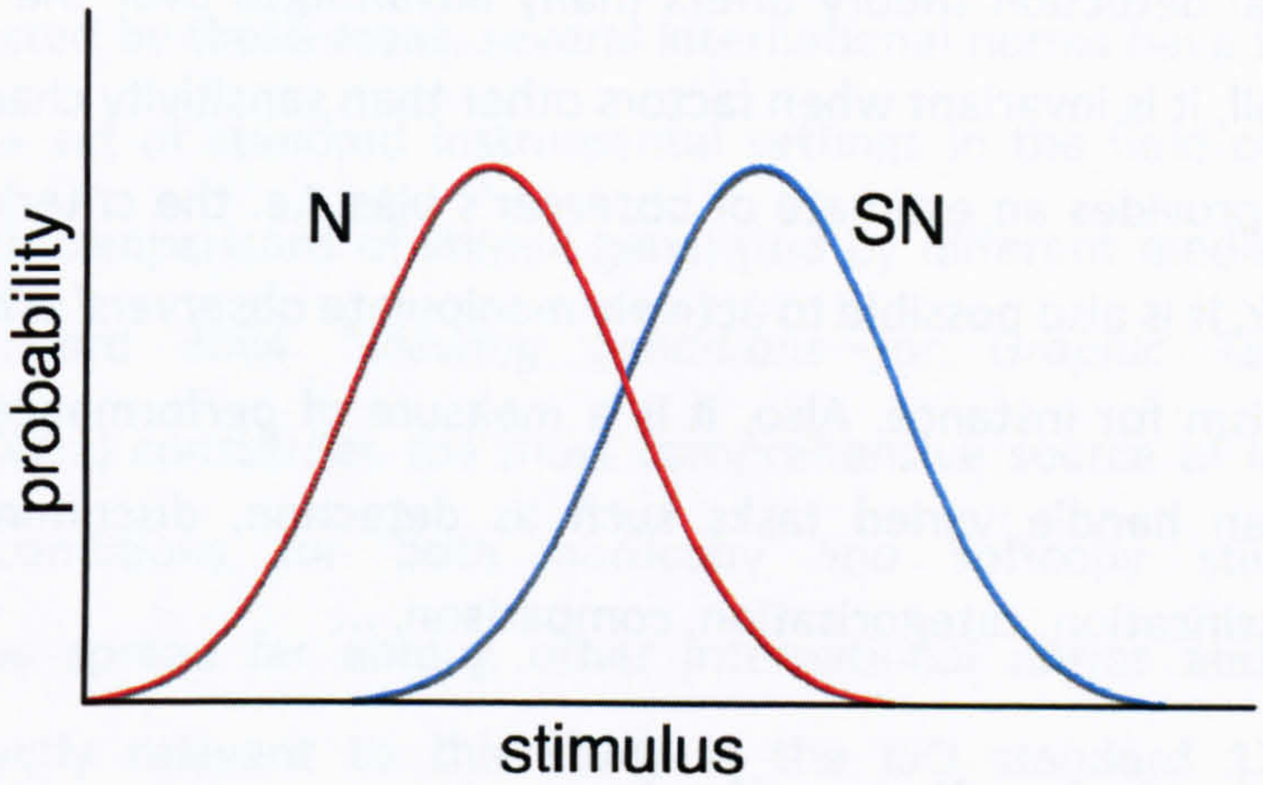


Figure 3 - Signal detection theory: decision space.

For instance, say an experiment is carried out to determine the absolute sensitivity of human observers to sound. Observers are asked to put on headphones, and report when prompted whether they detected sound. The source of sound will be switched on for certain trials and switched off for the rest, and the intensity of the sound will also be varied. Therefore, this type of procedure involves two stimuli classes (sound or no sound) and two possible responses (detected or not). Conventionally, the four possible outcomes are described as follow. Correctly detecting the presence of a stimulus is termed a *hit* (*H*), and failing to detect it a *miss*. Mistakenly detecting a stimulus where none was presented is a *false alarm* (*F*), and correctly detecting the absence of stimulus is a *correct rejection*. Observers' answers can thus be summarized into a 2x2 matrix, whereby the rows represent the stimuli classes and the columns observers' responses, and each cell corresponds to one of the four possible cases.

The true performance is given by the difference between the two distributions S and SN, i.e. to obtain a performance index, the means of the two distributions simply need to be subtracted. In other words, such a performance index expresses how close or far observers perceive the signal to be from the noise. It can be obtained computationally from the tally of

hits and false alarms. These data first need to be transformed into proportion rates by a simple normalisation with the total number of trials. The canonical performance index is termed d' and is defined in terms of the inverse of the normal distribution function, z:

$$d' = z(H) - z(F) \tag{6}$$

The signal detection theory offers many advantages over the traditional threshold theory. First of all, it is invariant when factors other than sensitivity change, such as observer bias. It actually provides an estimate of observer's bias, i.e. the criterion they employed to perform the task. It is also possible to actively manipulate observers' criterion, by providing a reward mechanism for instance. Also, it is a measure of performance in general, not just sensitivity. It can handle varied tasks such as detection, discrimination, identification, recognition, classification, categorisation, comparison, ...

2.6.2 Psychophysical scaling methods

Another interesting aspect of psychophysics is the ability to produce scales of human perception of certain features of physical stimuli. Such methods have been used extensively in the field of colour science. Out of the three principal methods available (Engeldrum, 2000), two are of particular interest to this study. The most intuitive one is Togerson's law of categorical judgement, whereby observers are asked to rate certain qualities of a stimulus into pre-defined categories. Observers' answers can then be transformed into an interval scale in which each stimulus class is given a unique score based on the relative position of the stimuli with respect to category boundaries (Togerson, 1954). Similarly, Thurstone's law of comparative judgement is equally trivial to set up, as observers are simply required to choose among two stimuli which one exhibits more of the quality being evaluated. As for the law of categorical judgement, their answers can also be transformed into an interval scale, with the difference that scores for stimulus classes are relative to the mean score of the whole set of stimuli (Thurstone, 1927). An interesting alternative of Thurstone's law is the Bradley-Terry model, whereby dependence on the normal density function is replaced with a simple logistic function (Bradley & Terry, 1952).

Chapter 3: Experimental Apparatus

Modern colour appearance models (CAMs) highlight the need to incorporate colorimetric information about several areas of the viewing field in order to achieve an accurate prediction of the appearance of colour stimuli. Since the appearance of colour stimuli is strongly affected by these areas, several international norms have been published in order to establish a set of standard instrumental settings in the field of colour image reproduction. Although comparisons of stimuli generated by different media are out of its scope, the ISO standard 3664 "*Viewing conditions—for Graphic Technology and Photography*" (ISO, 2000a) constitutes the most comprehensive source of information for specifying viewing conditions for both hardcopy and softcopy stimuli, and its recommendations thus spread far among other international norms and experimental guidelines. More directly relevant to this study is the ISO standard 12646 *"Graphic technology—Displays for colour proofing—Characteristics and viewing conditions"* (ISO, 1999), which concentrates on soft proofing. Similarly, several CIE technical committees (TC) have also published guidelines to outline experimental settings involved in generating visual data that can be used to evaluate the performance of colour reproduction systems. Despite their lack of interest in self-luminous colours and complex stimuli, the guidelines diffused by TC 1-34 *"Testing Colour-Appearance Models"* nevertheless include interesting information concerning the reproduction of colour across different viewing conditions (CIE, 1995). The recommendations of TC 1-27 *"Specification of Colour Appearance for Reflective Media and Self-Luminous Display Comparisons"* are directly relevant to this study (CIE, 1994). Although their research target some very specific aspects of the colour reproduction chain, the guidelines published by TC 8-03 *"Gamut Mapping"* and TC 8-04 *"Adaptation under Mixed Illumination Conditions"* can also be taken into account since their investigation involves the reproduction of colour across different media and/or various viewing conditions (CIE, 2000, 2004).

This section will provide the design motives and a detailed description of the settings adopted for all aspects of the experiment setup. The general consensus among these recommendations was followed as closely as possible as most of these guidelines largely overlap. Any significant departures required to satisfy the specific needs of this set of

experiments will be clearly indicated. In order to facilitate the understanding of the issues discussed thereafter, schematic diagrams and photographs of the final setup are given in the figures below.

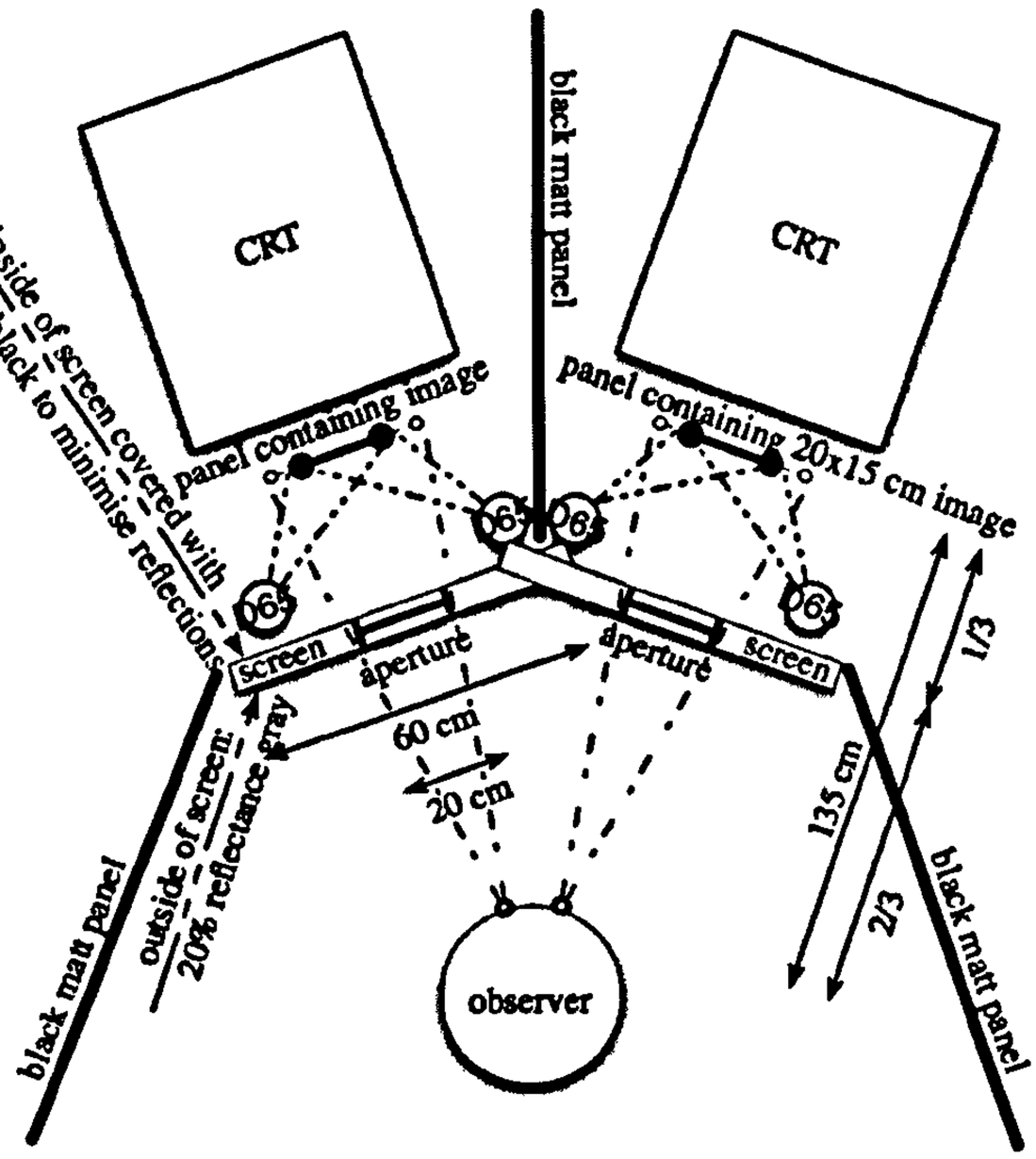


Figure 4 – Schematic diagram of the experimental setup: top view

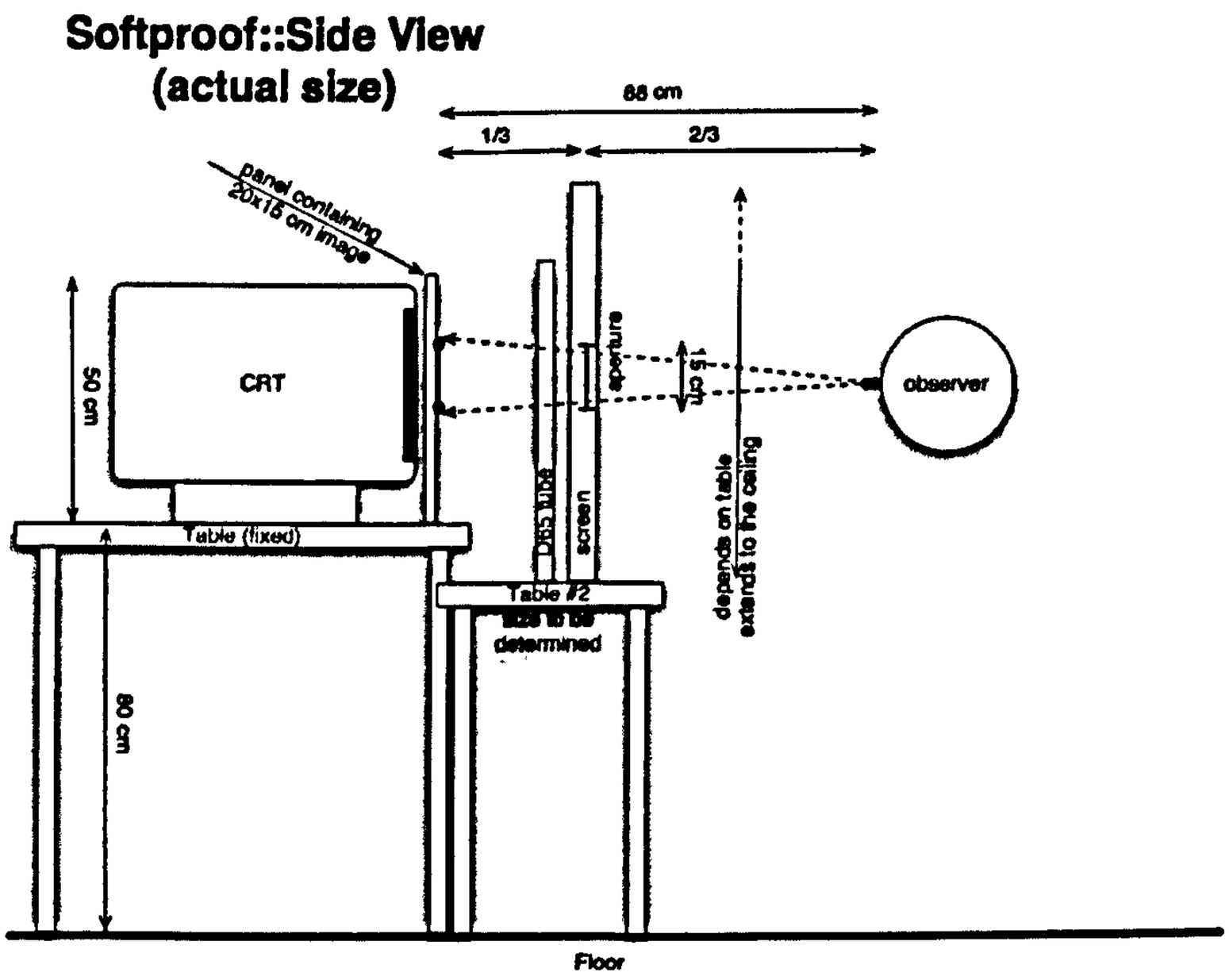


Figure 5 – Schematic diagram of the experimental setup: side view

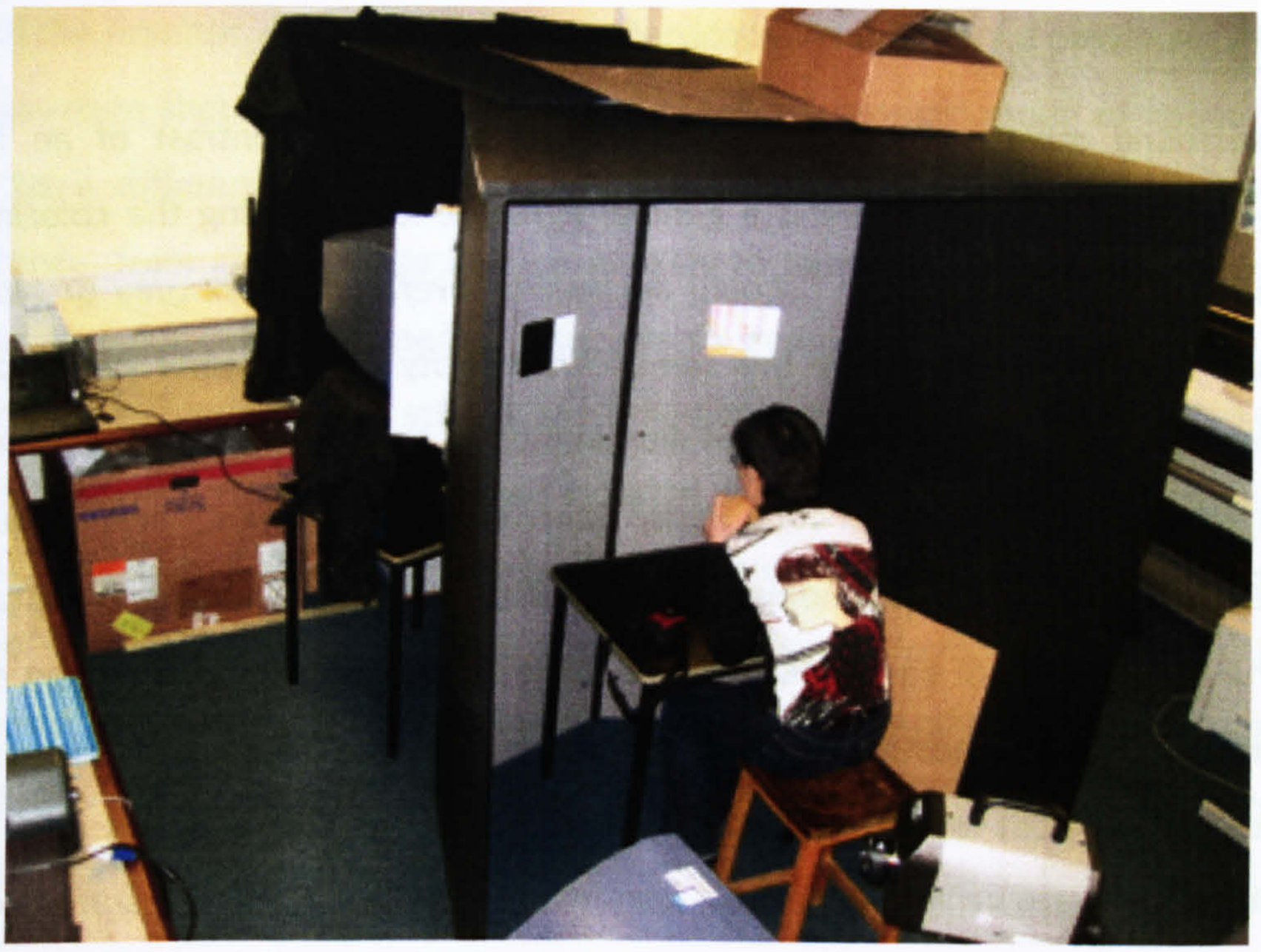


Figure 6 – Actual picture of the experimental apparatus

3.1 Viewing Technique

For the visual assessment of colour, several different methods can be employed to present the stimuli to observers. Successive viewing techniques, whereby a single stimulus is displayed at a time, are mostly used when the viewing configurations are significantly different for the two media (CIE, 1994). They are, however, not the most suitable for the task considered, as they require the observers to memorise the appearance of the stimuli. The lapse of time between viewing the original and the reproduction may degrade observers' memory of the original and thus alter their judgements (Braun, Fairchild, & Alessi, 1996). Similarly, using an haploscopic viewing technique, whereby each eye is completely adapted to a different viewing condition, is also not reasonable since the viewing conditions will be equated as carefully as possible for each both media. On the other hand, the simultaneous binocular viewing technique, whereby both stimuli are present in the same field of view and observed simultaneously, will be perfectly appropriate in this case. The adaptation state of observers will be single and steady and their discrimination ability will be maximised as they will be able to freely look back and forth between both stimuli. It will therefore be necessary for both stimuli to be in the same field of view.

3.2 Surround

The surround mostly influences the perceived brightness contrast of an image (Bartleson & Breneman, 1967). There is a general agreement concerning the colorimetric attributes for this visual area, which *"should be neutral, preferably dark grey or black to minimize flare, and of approximately the same chromaticity as the white point* [of the adapting stimulus]" (ISO, 2000a). The level of illumination is, however, left to the discretion of each experimenter, and two conditions, dim and dark levels, are frequently recommended. The former, corresponding to typical office conditions, has the advantage of preserving the full contrast of hardcopies, whereas the latter results in a reduction of the achievable dynamic range, mostly in the dark zone of the colour gamut. However, the situation is to a large extent reversed for stimuli generated by monitors. The presence of an ambient light will generate veiling reflections, i.e. unwanted reflections at the surface of the media coming from the ambient illumination, which will *"lighten and reduce the apparent contrast of the darker parts of an image"* (ISO, 2000a). Although each situation demands opposite solutions, the effects of ambient illumination affect more strongly the appearance of stimuli displayed on monitors. A totally dark environment will minimise the possible impact of external illumination on both media and thus maximise their colour reproduction ability. The experimental room will not contain any ambient lighting, and panels will be erected between the monitors and the hardcopies in order to avoid stray lights from one side affecting the other.

The presence of any cognitive cues that might help observers to discriminate the type of medium used to generate the stimuli and could seriously impair the pertinence of the results obtained. The medium type used would become obvious if observers are given a directly unobstructed view of the experimental setup, which will certainly not satisfy the strict exigencies necessary. A first solution could consist of modifying the optical pathway of stimuli's light-beams by the use of a combination of lenses, mirrors or prisms. This proposition has, however, to be discarded because it might modify the mode of appearance of stimuli resulting in a fundamental change in the way observers perceive them. The proposed solution consists in placing a baffle between observers and stimuli that contains two apertures through which the stimuli will be displayed. The size of the apertures will be adjusted in order to hide anything that is not part of the stimuli themselves, such as the

bezel of the monitor, or the backing of the hardcopy. As mentioned previously, the presence of an aperture should not significantly modify the hardcopies' mode of appearance, but it might be a different matter in the case of softcopy stimuli. A shift in their mode of appearance, from *illuminant* to *surface*, is likely to happen, but since we are interested in whether a medium's fundamental principle to generate colour information affects the appearance of the generated stimuli, the influence of external conditions can be neglected in a first instance. Furthermore, these external conditions should be equated as closely as possible in order to make the comparison between media fair. The baffle will therefore have the double advantage of hiding the experimental setup from the observers, but also to equate the surround conditions as closely as possible. Several panels will also be added to enclose each media in order to avoid any light leaks, and they will be painted in matt black in order to create a totally dark environment.

3.3 Background and Proximal Field

Although the surround effects are somehow limited to the perceived contrast of images, the background has a much wider range of influences on the appearance of colour stimuli, since it defines - along with the proximal field - the adapting stimulus that determines the state of adaptation of observers. Tribute to its extensive effect, most guidelines provide recommendations that leave no margins for originality, unless its specific influence is being tested (CIE, 1994, 1995, 2000, 2004). The standard condition consists of a *"white border and a uniform, spectrally nonselective neutral-gray background"* and *"the gray background luminance level should be a proportion of that white with a luminance factor of 0.2"* (CIE, 1994). The white border corresponds to the proximal field defined by modern CAMs, and its purpose is to provide a reference-white adapting stimulus in order to steady the observers' state of adaptation. Its luminance level defines the luminance level of the adapting stimulus, and its chromaticities should match those of the gray background. Overall, the stimuli should be *"situated so there are no strongly coloured areas (...) in the field of view (...) which may affect the vision of the viewer"* (CIE, 2004).

The physical form of the background is also critical to the success of these experiments. It cannot be included in the baffle introduced in the previous section, because the baffle would need to be illuminated in order to provide the level of luminance required

of the background. As previously stated, any ambient illumination is to be avoided in order to limit the effects of veiling glare. Moreover, in the case of softcopy, generating the background with anything other than the monitor itself implies to indirectly illuminate the monitor, which would once again allow the creation of veiling glare. Also, from a cognitive point of view, generating the background using a material other than the medium itself could potentially give observers enough cues to discriminate between the types of devices used. It would indeed be extremely difficult to entirely hide the presence of sharp edges at the stimuli boundaries. That might be attainable with a careful setup, but changing stimuli between trials would then require a precision so high that the time required to perform the experiment would increase exponentially. Therefore, the most appropriate solution consists of generating the background – proximal field included - using the same medium as for displaying the stimuli themselves, which makes perfect sense since the background is a full-fledged part of the adapting field.

3.4 Illumination

Due to their fundamentally different methods for producing colours, softcopies and hardcopies are generally observed under very different conditions (Hunt, 1998). Being self-luminous, softcopy images can be viewed almost anywhere, standard office conditions being most typical. Except for the presence of viewing glare, the colorimetric attributes of a softcopy are entirely specified by the characteristics of the monitor employed. On the other hand, a hardcopy is part of a binary system, whereby an illuminant is necessary for the print to be viewed. The lighting conditions specify not only the appearance of a hardcopy, but also its colorimetric attributes. Accurately controlling these parameters require to precisely control the illuminant conditions, and the use of a viewing booth is generally recommended in such cases (CIE, 2004; ISO, 2004). Since they are designed in conformity to international norms, they offer a convenient way to generate a controlled viewing environment, but they generate a very diffuse illumination, which might affect the apparent contrast of hardcopies. A directional light source can on the other hand preserve the full contrast and the saturation of colours, although at the cost of creating gloss caused by specular reflections, i.e. mirror-like image of the source, which strongly affects the appearance at a specific angle (Hunter & Harold, 1987). Providing that such problems are adequately handled, this type of

illumination is therefore more advantageous. In order to overcome such issues when assessing the appearance of colour stimuli, the CIE defined several standard geometries of illumination and viewing for reflective surfaces (Hunt, 1998). Among them the one named *45/0*, whereby *45* refers to the angle between the central ray of the beam and the normal to the observation plane, and *0* to the angle between observers' viewing axis and the normal to the observation plane, was specially designed in order to avoid the negative effects that gloss may have. In order to obtain optimal viewing conditions, it is therefore proposed to set up an illumination configuration using directional light sources and the *45/0* geometry of illumination and viewing rather than a viewing booth. In order to maximise the uniformity of the illumination across a large area, fluorescent tubes will be employed.

From a geometrical point of view, replacing the viewing booth by directional light sources also allows to increase the flexibility of the spatial arrangement of the experimental room. In order to provide a uniform illumination, most viewing booths are relatively cumbersome, which might entail certain troubles for fitting them in the tight design specified so far. Also, it would be extremely difficult to interchange the media's location, as only one of them could be seen at a time through the apertures. However, protecting each medium from the other's influence becomes even more important in this case, as the light sources will spread all their energy in every direction, without any restraint. As previously mentioned, such illumination might seriously affect the appearance of softcopy stimuli, as monitors are very sensitive to veiling glare. Moreover, the illumination configuration considered could also provide enough cues to deduce the type of medium employed. The luminance level required is likely to be able to illuminate not only the stimuli themselves, but also parts of the experimental room, and thus creates differences in the brightness perceived in observers' field of vision. Enclosing both media and also the observers is therefore of primary importance. Additional panels were added to the baffle in order to cover the entire field of view of observers and prevent any stray light to be perceived. Several panels were also placed between both media for the same reason. The directional light sources were fixed directly to the baffle, next to the apertures, in order to hide them from observers. As previously stated, the baffle and panels were painted in matt black in order to create a totally dark environment.

The chromaticity of their respective white points is also a source of discordance between prints and monitors. The CIE standard D50 illuminant is the preferred choice for assessing the quality of hardcopy stimuli (ISO, 2000a). However, the correlated colour temperature of typical monitors under standard office environments is usually close to 9300K (CIE, 2000). Nonetheless, the ISO 3664 Standard recommends the use of the CIE standard D65 illuminant for appraisal of images displayed on monitors because *"[t]here is some evidence that, at the low luminance levels obtained with monitors, a chromaticity close to that of D65 provides a better evocation of white"* (ISO, 2000a). Furthermore, *"[t]he majority of research has been performed using sources that simulate CIE Illuminant D65"* (CIE, 2000), as their high colour rendering index make them an appropriate choice, and is thus recommended by all experimental guidelines. This intermediary illuminant will thus be the target to specify the chromaticity of the adapting stimulus for both the prints and the monitors. As for the level of illumination, two were found to be necessary in order to assess the quality of hardcopy, 2000 and 500 lx (ISO, 2000a). However, although the high one is recommended for the critical comparison of several reproductions of an image, it *"can give a misleading impression of the tone reproduction curve and colourfulness of an image which will ultimately be viewed by the consumer in much lower levels of illumination"* (ISO, 2000a). Moreover, physical limitations of current CRT monitors impose limits on the level of luminance of their white, which approximately corresponds to the level specified in the lower condition. The luminance level of the adapting stimulus will therefore be set according to the low luminance level, more precisely the level achievable the monitors, and its chromaticity will be set to approximate that of the CIE standard D65 illuminant. The colorimetric attribute of the background will be specified according to those values.

3.5 Geometrical set-up

In order to maximise the apparent similarity of both type of stimuli, it is important to accurately reproduce not only their colorimetric attributes, but also their spatial characteristics. The spatial resolution of both media being clearly different, they should be viewed at a distance where individual dots cannot be distinguished. Since CRTs exhibit the lowest resolution of the two types of media, the minimum distance should be specified according to their abilities. The maximum acuity of the human visual system, as measured by

the contrast sensitivity functions for the luminance channel, corresponds to approximately 60 cycles per degree of visual angle. Given the size of a pixel on the monitor, the minimum distance can be estimated to be approximately 100 cm, which concords to most recommendations and also to typical office conditions. However, such distances would create many problems in the present case. Indeed, it imposes restrictions on both the observer-aperture and aperture-stimuli distances. The latter is dependent on the location of the illuminant tubes, which needs to be at a 45 degrees angle from the axis defined by the observer's viewing axis to be compliant with the CIE standard geometry of illumination and viewing *45/0*. However, their location, more specifically the distance between the left and right tubes, is also limited by the dimensions of the aperture, which in turn is dependent on the stimulus size and the aperture-observer distance. Similarly, the aperture-observer distance is also dependent on the stimulus size and the aperture dimensions, but also on the physical dimensions of the printing-paper sheets and the monitors, as the viewing configuration needs to be specified in such a way that observers are not given a direct view of the medium casing and boundaries. The stimuli and their background, both provided directly by each medium, therefore need to cover the entire area that observers can see through the aperture. Moreover, despite the fact that observers will use a chinrest in order to precisely control what they can see, the various faces shapes can be responsible for a vertical shift of up to 6 cm in what observers can perceive through the apertures. Adjusting all those parameters for the distance observer-stimuli to be 80 cm proved to be impossible. Rather, the optimal solution to this intricate problem consists in having the stimuli located at 45 cm from the baffle, and to set the distance baffle-observer to twice more. Observers would definitely not be able to resolve individual pixels with this configuration. Setting up the distance observer-aperture and aperture-stimuli to be in a *2:1* ratio also has the advantage of reducing by half the margins necessary to handle the differences in the shape of observers' faces.

Another important geometrical characteristics is *"for the observer's view to be equally distant from the hardcopy and softcopy images. This can be accomplished in either of two ways. Both the illuminated reflection print and the self-luminous display images can be on the same radius with regard to the observer's view or both images can be coplanar"* (CIE, 1994). The latter arrangement has the advantage of being easy to set up. However, some of

the SLDs that might be used during the course of this experiment have a rather strong angular dependency, implying the fact that viewing them at an angle, as having coplanar stimuli would force, might affect their colour reproduction performance. On the other hand, setting both hardcopies and softcopies to be on the same radius with regard to the observer's point of view, and thus also having their normal axis parallel to the observer's viewing axis, would not have such a negative effect. Furthermore, this configuration is also compliant with the CIE standard geometry of illumination and viewing 45/0 and was therefore adopted.

Despite a significant increase in viewing distance, the physical size of stimuli were also increased for the visual angle to be consistent with what is usually used in other experiments (CIE, 1994). As previously mentioned, the ability of changing the type of stimuli seen through an aperture was crucial for this set of experiments. A slot mechanism was thus built below the monitors in such a way that removable cardboard, onto which hardcopy stimuli were mounted, could be easily placed and removed just 1 cm in front of the monitor screen. This difference in the distance observer-stimuli actually corresponded to less than 0.1 degrees of visual angle, and was therefore considered negligible. The whole visual field that can be seen through the aperture, i.e. the stimulus plus its border and background, subtended an angle of 13 degrees horizontally and 10 degrees vertically. The stimuli themselves subtended an angle of 8.5 (20 cm) degrees horizontally and 6.5 degrees (15 cm) vertically. The size of the proximal field was specified to be approximately 10% of the size of the stimuli. Once again, schematic diagrams and photographs of the actual setup are given in Figures Figure 4, Figure 5 and Figure 6.

Chapter 4: Colour Reproduction Framework

The purpose of this study consists of investigating whether identical stimuli generated by two different types of colour reproduction media are perceived differently, and whether these discrepancies are responsible for different judgements when assessing the quality of colour stimuli. In order to make the comparison pertinent, the viewing conditions obviously need to be equated very precisely to maximise the apparent similarity of both types of stimuli, and the setup produced was described in details in the previous section. Another major factor that will determine the reliability of the findings is the overall colorimetric accuracy between monitors and prints. Obtaining a visual match obviously imply that the intrinsic inaccuracies of colour reproduction devices are kept under a certain threshold, determined by the observers' sensitivity to colour stimuli. This section will analyse in details the abilities of both medium to reproduce colour, both from a theoretical and a practical point of view. This section will provide a detailed analysis of the abilities of each component in the colour reproduction chain before considering the overall accuracy of the implemented framework.

Before broaching the setup details, a general overview of the whole colour reproduction chain will be given first for reference. The workflow studied here, i.e. monitor to print, is among the most commonly used in the Graphic Arts industry (CIE, 2004). Its general principle is very simple: *"The images originate in RGB space (i.e. these images either have been created on the CRT, or converted into screen RGB space for processing)(...). Finally the image is sent to output products such as photographic paper prints or inkjet prints"* (CIE, 2004). If any kind of colour management is involved, many other steps are also comprised between the two mentioned to accommodate the diverse ways media employ to generate colour information. A general overview of such a framework is outlined in the following figure. Please note that the presence of a second pathway is specific to the concept studied here, which involves simulating the appearance of a print on a monitor.

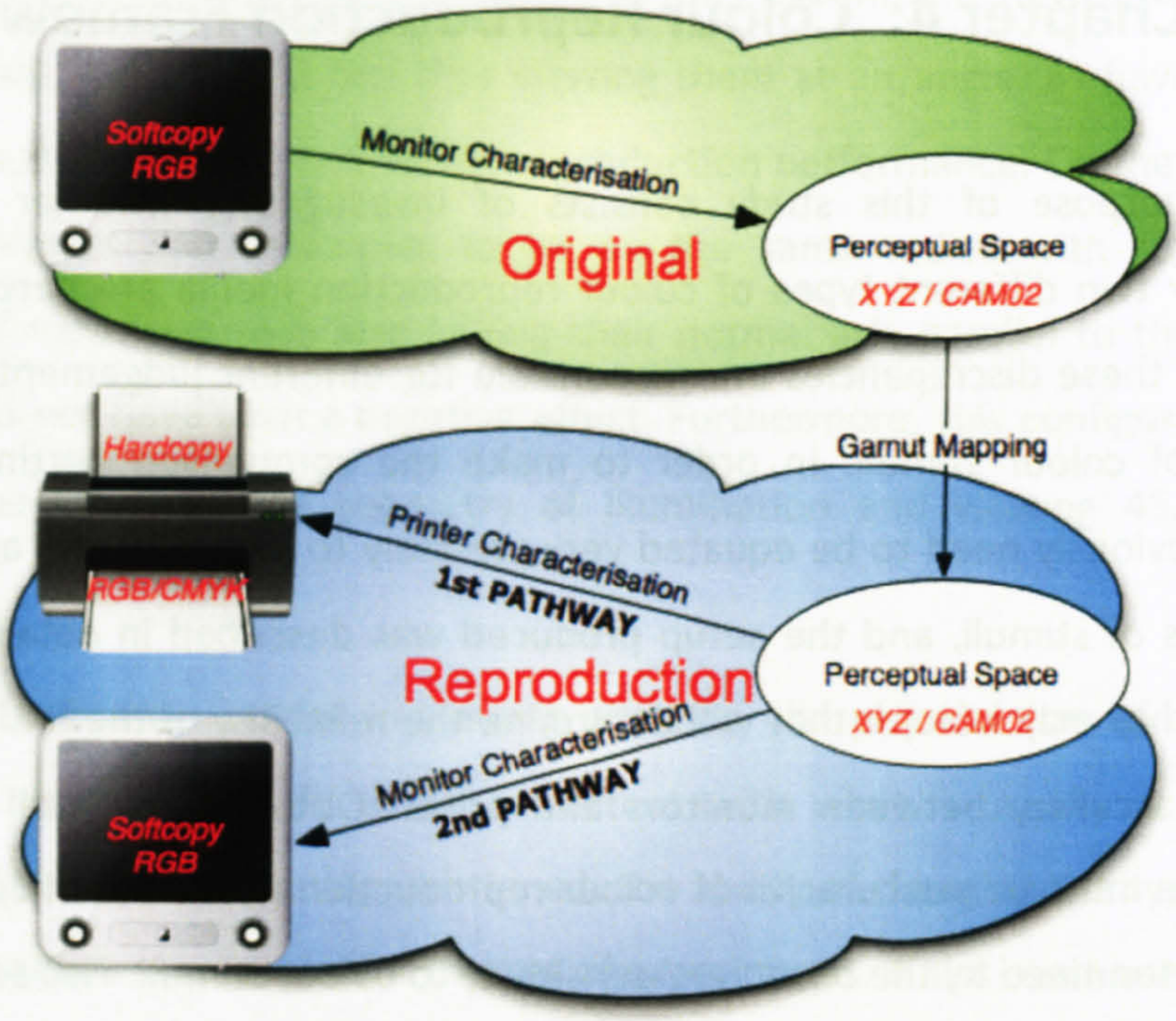


Figure 7 - Workflows overview

4.1 Miscellaneous Settings

Both the setup and the experimental phases were performed in a single room. The windows contain opaque matt black blinds that could totally block out daylight. The ambient lighting was also switched off while measuring and carrying out the experiments. No external light sources could have tempered the instruments calibration. Furthermore, the temperature was monitored and kept constant at 22˚C by the air-conditioning system.

4.2 Metrology

The ultimate goal of the experimental setup would be to achieve a perfect match of the colorimetric attributes of the hardcopy and softcopy stimuli when observed under experimental conditions. However, colour reproduction devices are intrinsically inaccurate, which implies that the degree of their inexactness has to be assessed, first independently and then as a whole, in order to be able to evaluate the quality of the colour reproduction workflow implemented. Furthermore, the worth of this evaluation also very much relies on the quality of the measurements, and it is therefore crucial to employ the most appropriate instruments and procedures. According to the experimental guidelines published by the CIE

TC 1-27 *"Specification of colour appearance for reflective media and self-luminous display comparisons", "[t]he primary instrument for objectively characterizing the colour and luminance of the self-luminous display and the illuminated reflection print samples (...) is the telespectroradiometer (TSR)"* (CIE, 1994). Therefore, in order to achieve the maximum accuracy, a telespectroradiometer was the reference instrument used to characterise the components of the colour reproduction workflow.

Before broaching the descriptions of the measurement instruments employed, few words need to be said about the kind of measurements that are needed here. The crucial issue here concerns more the level of match of the colorimetric attributes of the hardcopy and softcopy stimuli, regardless of the instrument employed. Using a unique instrument as reference is thus the most appropriate course of action, as this simplifies the process by eliminating a potential source of variation.

However, building an accurate characterisation model of printers involves measuring hundreds of samples (Bala, 2003). The suitability of TSRs for such cases is doubtful, as several spectrophotometers (SP) have the ability to perform automatic measurements of reflection samples, which reduce very significantly the amount of time and effort necessary to perform this task. On the other hand, employing another instrument re-introduces the need for inter-instrument calibration, i.e. relating all the measurements to what the reference instrument would have given. Despite this drawback, it was felt that the profit gained from using SPs was worth it overall. A SP was therefore used to build a characterisation model of the printer and assess its colorimetric characteristics. All the measurements made of the monitors were obtained with a TSR. Being the reference instrument, it was also used to evaluate the overall accuracy of the whole reproduction chain.

4.2.1 Measuring Instruments Calibration

As previously mentioned, calibrating the instruments so that they conform to reference laboratory settings was not consequential in this case. The calibration process did not therefore need to follow the strict recommendations given by standardisation bodies concerning photometric and wavelength scales for instance (ASTM, 1991). The process was

simplified and concentrated on the main issue, the consistency of results. Similar procedures were used for both the TSR and the SP.

First of all, the approximate correctness of the measurements was tested by measuring the 24 colours of the Macbeth ColorChecker Chart under a standard daylight illuminant. The correlation between the obtained colorimetric data and the chart manufacturer specifications was found to be satisfactory. In order to guarantee the traceability of the results, the calibration state of each device was also regularly assessed during the setup phase by measuring the spectral reflectance factor - or spectral radiance depending on the medium - of a white tile simulating a perfect diffuser. The variation of the obtained spectral data in both cases never exceeded 5 % and was judged to be acceptable.

4.2.2 Telespectroradiometers

The performance of two TSRs was evaluated in order to select the most appropriate one. The manufacturers' specifications, given in the next Table, were checked for conformance with the minimum requirements specified in standard measurement practice. The performance of both instruments are traceable to the National Physical Laboratory (UK).

Table 1 - Manufacturer specifications of the different telespectroradiometers tested

	Minolta CS-1000	Photo Research PR-650
Wavelength range	380-780 nm	380-780 nm
Spectral bandwidth	5 nm	8 nm
Wavelength resolution	0.9 nm/pixel	< 3.5 nm/pixel
Luminance range	0.1 – 8000 $cd.m^{-2}$	3.4 – 17000 $cd.m^{-2}$
Spectral accuracy	± 0.3 nm	± 2 nm
Chromaticity accuracy	± 0.0015 x ± 0.01 y	± 0.0015 x ± 0.01 y
Luminance accuracy	± 2 % ± 1	± 2 %

The specifications of both instruments match each other quite well. The major discrepancy comes from the luminance range, where the Minolta TSR has a clear advantage,

as it is able to measure very dark colours. The parts of the colour gamut near the black point are usually very prone to characterisation mistakes because of the presence of noise. Being able to precisely measure this zone is therefore primordial. The Minolta CS-1000 was therefore chosen as the reference telespectroradiometer.

4.2.3 Spectrophotometers

Similarly, the performances of the available SPs were assessed to find the most suitable one. The manufacturers' specifications are given in the next Table.

Table 2 - Manufacturer specifications of the different spectrophotometers tested

	GretagMacbeth Spectrolino	GretagMacbeth CE-7000A	X-Rite 938
Wavelength range	380-730 nm	360-750 nm	400-700 nm
Spectral bandwidth	10 nm	10 nm	20 nm
Photometric range	0 – 120 %	0 – 200 %	0 – 200 %
Photometric resolution		0.001 %	0.01 %
Illuminant type	D_{50}, D_{65}, A, C	Pulsed xenon	D_{50}, D_{65}, A, C, F2
Measuring geometry	45°/0°	d/8°, integrating sphere	0°/45°

The technical specifications of the X-Rite SP do not reach the standard set by the other two instruments, and it was thus discarded. Choosing between the two remaining was more difficult, as their respective advantages and disadvantages cancel each other. The spectral specifications of the CE-7000A are of higher quality, although those of the Spectrolino are more than sufficient, but its measuring geometry does not match the experimental setup, whereas the Spectrolino's geometry does. Since the consistency of the measurements is of primary importance, the choice between them was done by analysing their repeatability. The 24 colours constituting the MacBeth ColorChecker chart were measured repeatedly over a period of 4 days. Three measurements were done sequentially the first day, and a single measurement was performed for each of the three remaining days. The spectral reflectance factor given by each instrument was transformed into CIE

1931 XYZ data by using the spectral power distribution of the CIE D_{65} illuminant. The differences between each measurement and the average of the first day measurements are displayed in the next three figures.

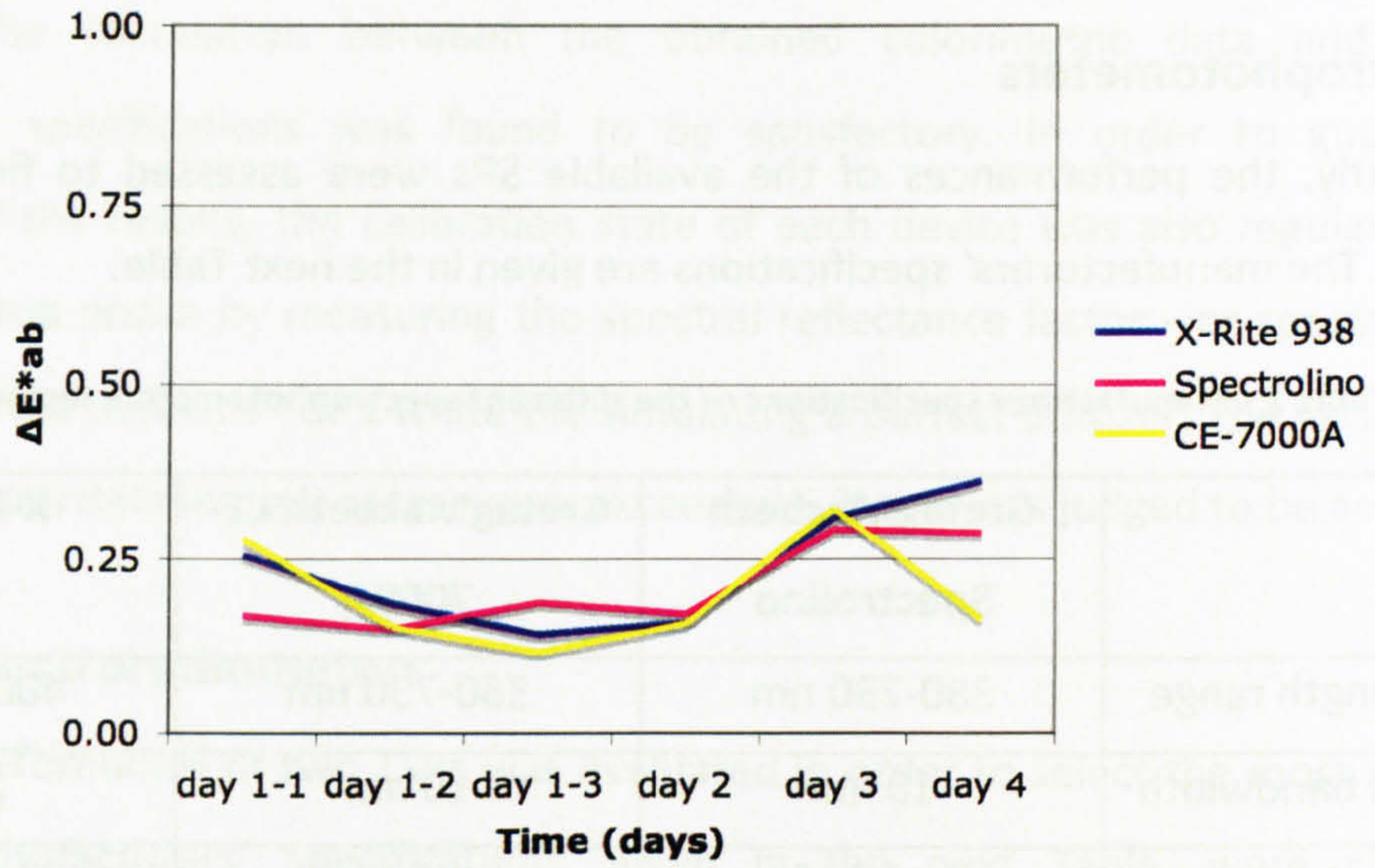


Figure 8 – Temporal repeatability of the diverse spectrophotometers available: average ΔE^*_{ab}

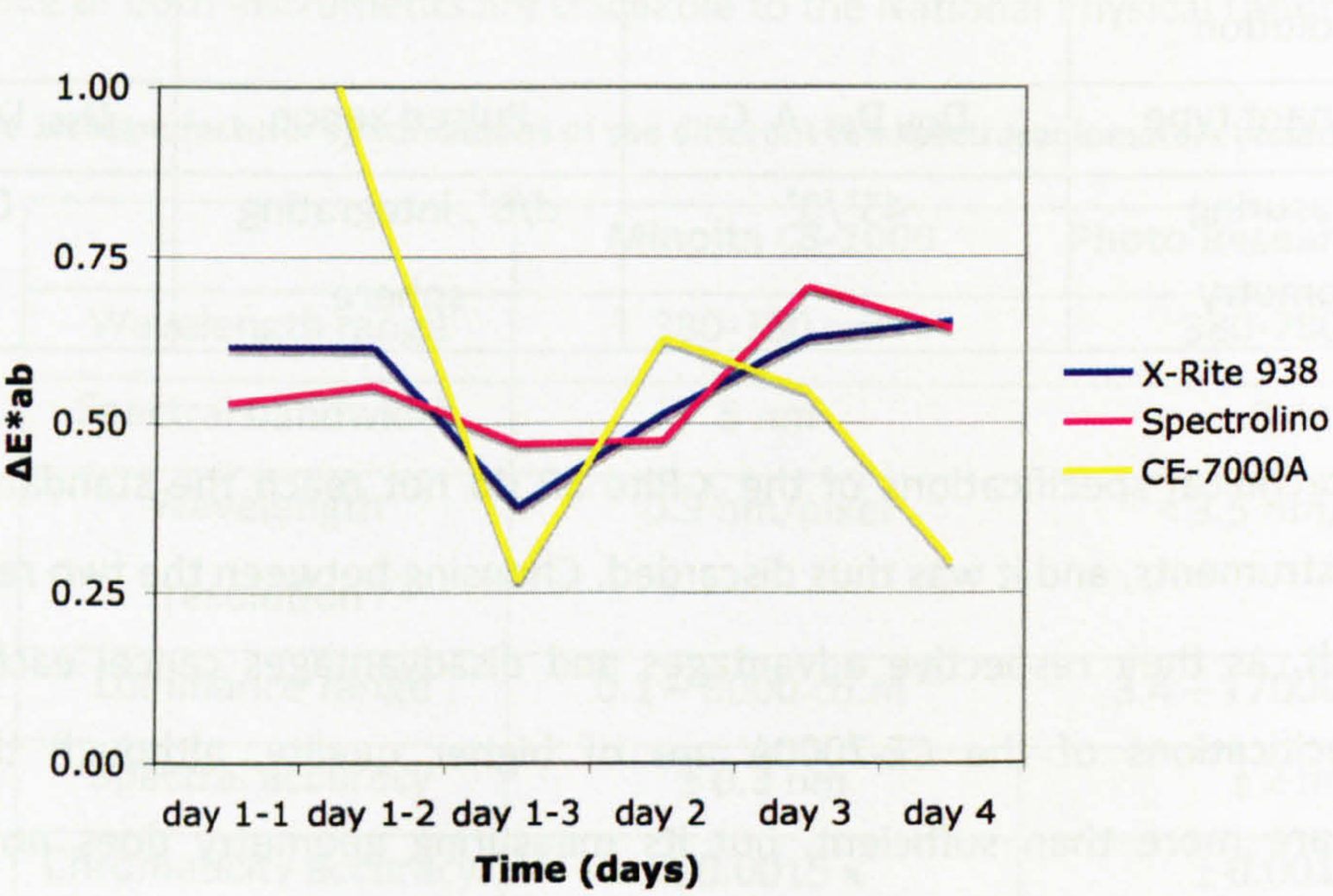


Figure 9 - Temporal repeatability of the diverse spectrophotometers available: maximum ΔE^*_{ab}

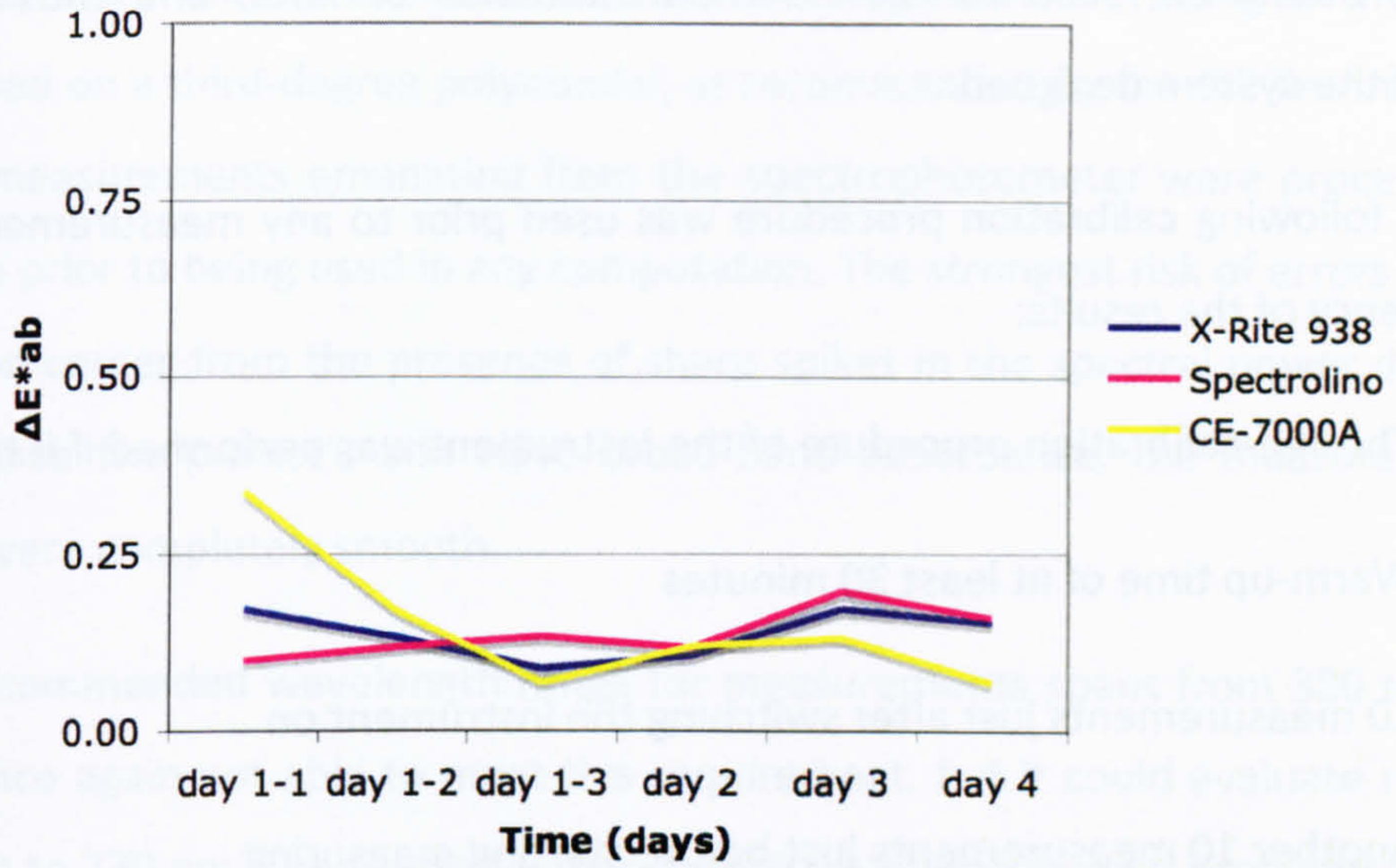


Figure 10 - Temporal repeatability of the diverse spectrophotometers available: standard deviation ΔE^*_{ab}

The average error obtained for all three devices indicate a very slight advantage for the CE-7000A, but the last figure shows that the maximum error for this device has wide fluctuations. On the other hand, the results given by the Spectrolino are the most stable, and it was thus chosen for evaluating the reflectance factors of reflective samples.

4.3 Measurement Procedure

The spectrophotometer uses the CIE 45/0 geometry of measurement and viewing with specular reflection excluded, and samples are illuminated by a fluorescent illuminant simulating the CIE D65 reference illuminant. The prints to be measured were not placed directly on the black surface of the SP, but another blank print was first put directly in between the SP and the stimuli in order to provide a white backing, which helps reducing the amount of noise in the data (Bala, 2003). The measurements made with the TSR, i.e. for both monitors and also for the prints used to derive the inter-instrument calibration procedure and to assess the overall accuracy of the colour reproduction framework, corresponded to the real experimental conditions. The instrument was placed at the exact position where observers' eyes would be, and the lighting conditions were replicating very precisely the true experimental conditions. This had the advantage to take into account directly into the device characterisation data any parameters that may have been

overlooked otherwise, such as the potential influence of flare, and thus increase the accuracy of the system designed.

The following calibration procedure was used prior to any measurements to ensure the consistency of the results:

- The self-calibration procedure of the instrument was performed if it had one
- Warm-up time of at least 30 minutes
- 10 measurements just after switching the instrument on
- Another 10 measurements just before starting measuring
- The self-calibration procedure was performed once again if the instrument had one

In order to reduce the bias of instrumental measurements, every sample was measured three times sequentially (ASTM, 1990b). The true estimation of the spectral data given by the measuring device was then specified to be the arithmetic mean of the three quantities, whether the averaging process was directly performed by the device itself or processed later. All the results presented from now on are based on average of multiple measurements.

4.4 Computational procedure

Two different instruments were used to measure the colorimetric attributes of stimuli. Since they have different characteristics and were not calibrated to concur with reference instruments, it is not desirable to directly obtain tristimulus values from their readings. This section will introduce the computational method implemented to provide a meaningful and direct comparison between data measured for hardcopy and softcopy.

4.4.1 Spectral data

In order to obtain tristimulus values, it is recommended to use spectral data at wavelength intervals equal to 1nm (CIE, 2003b). The TSR provides such kind of data, but not the SP, as its wavelength interval is 10nm. A method to produce the needed but

unmeasured data is therefore needed. The preferred way consists of using an interpolation technique based on a third-degree polynomial, as recommended by the CIE Technical Report 15.3. All the measurements emanating from the spectrophotometer were processed using this technique prior to being used in any computation. The strongest risk of errors generated by this method comes from the presence of sharp spikes in the spectral power distribution curves, but since the printer's inks have broad-band absorbance, the measured spectral distributions were completely smooth.

The recommended wavelength range for measurements spans from 380 to 780 nm. The SP was once again not able to meet this requirement, but it could evaluate reflectance distribution up to 730 nm. The recommended technique to handle such problems consists of extrapolating the missing data by setting them to the nearest measured value (CIE, 2003b). Despite the crudeness of this approach, the potential errors it could generate are minimal, as the magnitude of the colour-matching functions at that wavelength range imply. All the reflectance estimated from the SP were processed to obtain spectral data spanning a range from 380 to 780 nm at 1 nm wavelength interval in order to fully comply with the recommendations on tristimulus values computation.

4.4.2 Inter-instrument agreement

As previously mentioned, different instruments were used to characterise the printer and the monitor. Given the intrinsic differences between the instruments, it is unlikely that a set of stimuli measured by both would provide exactly the same readings. It was therefore necessary to develop modeling techniques for correlating colour measurements results between different instruments. Such a model was implemented using techniques based on multiple linear regression described by Morovic *et al.* (1999). It includes a correction equation for each wavelength containing seven coefficients to model three types of systematic errors: photometric (zero offset, linearity, non-linearity), wavelength (linearity, quadratic non-linearity, sine wave non-linearity) and bandwidth. The parameters at each wavelength were estimated using multiple linear regression techniques based on 35 samples sampling uniformly the printer device-dependent space. The spectral reflectance curves of these colours were first evaluated with the SP and then with the TSR under the real experimental conditions. That had the advantage of taking into account into the model any

external factors, such as veiling flare that might have occurred despite the careful setup. A standard white tile, approximating a perfect diffuser, was also measured under both conditions in order to normalise the reflectance data before processing. The performance of the derived model was tested by assessing the accuracy of the predictions for the 24 colours constituting the MacBeth ColorChecker chart. The following table shows the average differences between the reflectance curves obtained for both instruments for corrected and uncorrected data.

Table 3 - Inter-instrument reflectance correction model: the errors are expressed in terms of percentage of reflectance

% difference	Average	Standard deviation	Maximum
Corrected	1.19	0.38	5.30
Uncorrected	1.43	0.61	16.59

The benefits from this model are obvious, since the maximum error is reduced by a factor of three. It was therefore employed to modify all the measured spectral reflectance.

4.4.3 Tristimulus value computation

Tristimulus values are calculated from the integration over the specified wavelength range of a constant multiplied by the inner product between the CIE colour-matching functions with the colour stimulus function, defined as the product between the spectral reflectance function and the spectral power distribution of the illuminant (CIE, 2003b). Several choices can be made regarding the diverse parameters of this computation. Firstly, two different sets of colour-matching functions can be employed depending on individual needs, the *CIE 1931* and *1964 Supplementary Standard Colorimetric Observers*. The former is better suited for the colorimetric specifications of viewing fields with small visual angles – between 1° and 4 °- while the latter takes into account larger fields. Since complex images, both pictorial and business graphics, constituted the stimuli employed in this set of experiments, the focus need to be put on small details, indicating that the former set of functions is the most appropriate. Secondly, the value of the constant can be specified in two different ways, resulting in two types, i.e. absolute and relative, of colorimetric data. In the case of hardcopy stimuli, it is more customary to use the latter whereby the value of the

constant is specified so that *Y* is equal to 100 for the paper white. These choices were employed to calculate the tristimulus values for hardcopy stimuli.

The computation for softcopy stimuli is more awkward as the previous calculation method, derived from reflecting object colour data, doesn't translate very well to monitors data. Since the constant is computed from the illuminant-object pair, it is indeed not possible to estimate its value in this case, as the monitor is both the object and the source of light. The preferred method to overcome this issue consists in specifying the constant as being the ratio between the maximum spectral luminous efficacy and the integration of the $\bar{y}$ function. This is however unlikely to produce satisfactory results as, even if two exact spectral colour stimulus functions emanate from an illuminated print and a monitor, the colorimetric attributes of both stimuli are not guaranteed to produce the same tristimulus triplet, since the constants may differ for the devices. The computational procedure therefore needs to be adapted to handle such cases. The ultimate purpose of the setup is to generate stimuli that are identical in colorimetric terms or, more precisely, to produce colour stimulus functions that are metameric but which produce the same colorimetric attributes when integrated with the colour-matching functions. The presence of the constant is actually more for convenience rather than anything else. The colour stimulus functions in the hardcopy case were obtained from spectral data, both reflectance and spectral power distribution, measured under real experimental conditions, and similarly for the monitor, whereby the colour stimulus functions generated were had under real experimental conditions. Since the ultimate purpose is to generate metameric colour stimulus functions, estimating tristimulus values by multiplying both of them by the same constant makes perfect sense. The colorimetric attributes of softcopy stimuli were thus computed directly from the spectral measurements made with TSR multiplied by the constant employed in the hardcopy case. Other arbitrarily defined constants could also be used. The results would be linearly related with the ones obtainer here. The important issue is actually to be consistent for both media in order to obtain pertinent results.

4.4.4 Colour difference estimation

Evaluating the performance of several devices in terms of colour reproduction can be done visually, but a more objective metric is better suited for comparisons. Credit to the

importance of this type of evaluation, colour differences formulae have been published in abundance over the years, ΔE^*_{ab}, ΔE^*_{94} and CIE ΔE_{00} to name the principal ones. A recent study comparing the performance of several formulaes concluded that the ΔE^*_{ab} was not statistically different from the best when estimating differences in the range of one to five units. Given its simplicity, it was adopted to estimate all the differences in colour appearance in the present study.

4.4.5 Colour appearance models

Gamut mapping algorithms, which were an important component of the colour reproduction framework implemented, need to be executed in a perceptually uniform colour space. The latest version of the CIE colour appearance model, CIECAM02, was employed for that task (CIE, 2003a). Given the viewing conditions adopted, the parameter settings adopted corresponded to the dark surround set. More specifically, the surround ratio S_r was set to nil, and the luminance of the adapting field and of the background, L_a and Y_b, were both calculated from the absolute luminance of the adapting white point. The c, N_c and F parameters corresponded to the recommendations for dark surround conditions.

4.5 Subtractive colour reproduction system

As previously mentioned, hardcopy stimuli consist of two principal components; an illuminant is necessary for the print to be observed. The complexity of the situation is increased further when hardcopy generation is considered. Another component, the printer and its inks, are added to what becomes a ternary system. This section will present the assessment of these three components as colour reproduction devices.

4.5.1 Illumination

a) Colorimetric characteristics

In the present case, the hardcopy illumination was provided by daylight simulators approximating the spectral power distribution of the CIE Standard Illuminant D_{65} manufactured by *Verivide*. The spectral power distribution (SPD) of the test source was estimated by measuring with a telespectroradiometer under real experimental conditions the reflectance of a perfect diffuser. The relative SPD – i.e. normalized such as the power at

560 nm is 100 – of the daylight simulators is compared to the relative SPD of the CIE Standard Illuminant D_{65} in the following figure.

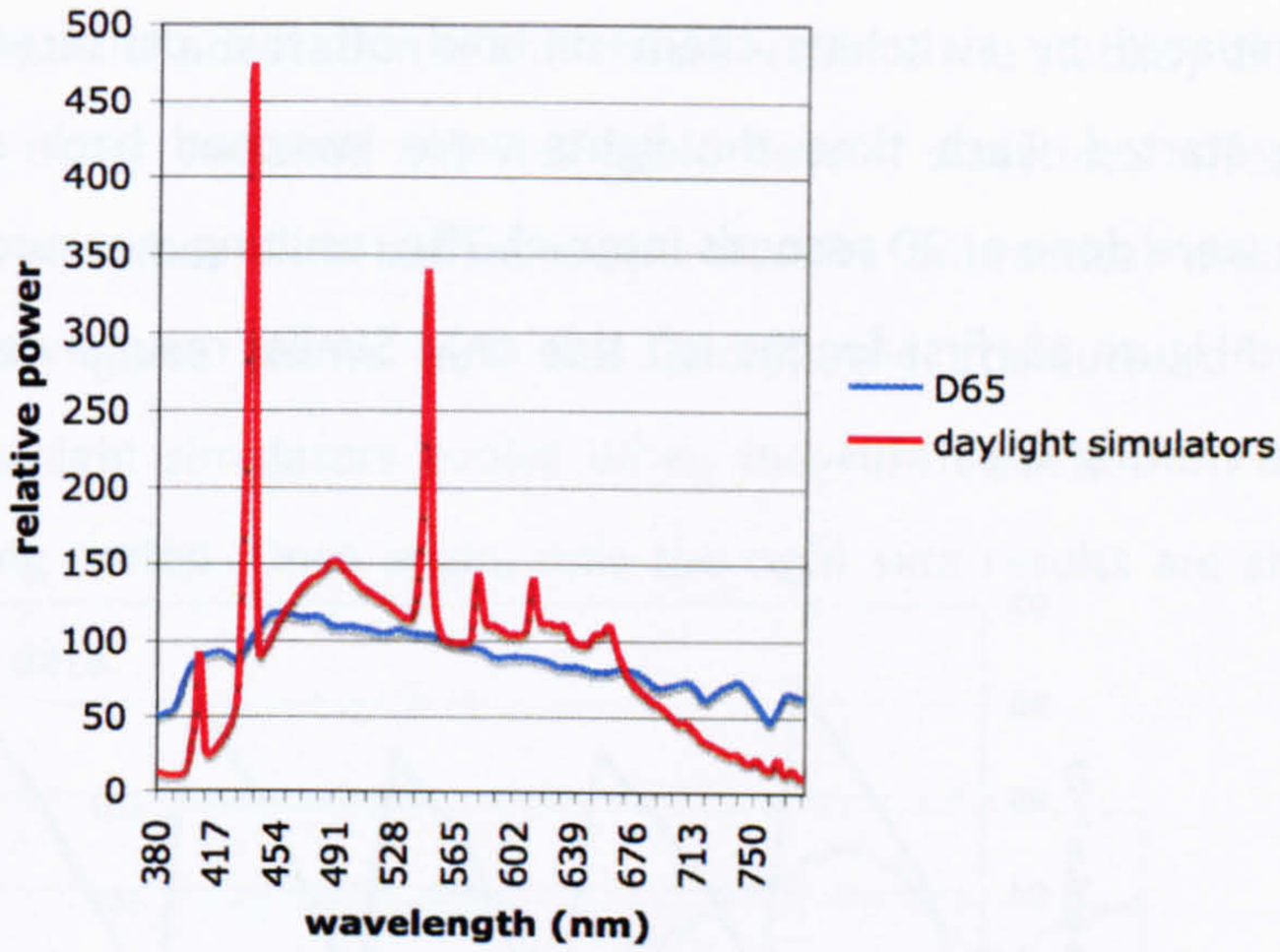


Figure 11 - Relative spectral power distribution of CIE Standard Illuminant D65 and daylight simulators.

The ability of the test source to simulate the CIE Standard Illuminant D_{65} can be assessed by computing its CIE General Colour Rendering Index (CRI) (CIE, 1998). The CRI is a unit of measure defining how well colours are rendered by different illumination conditions in comparison to a standard. It is a linear transformation of the distance in the CIE U*V*W* between eight specific Munsell colours under two types of illuminant, the test source considered and a reference source, i.e. a black-body radiator or a phase of daylight depending on the colour temperature considered. The CIE D65 reference illuminant was used here to perform the computation. The obtained CRI for the daylight simulators was 95.9, indicating that they and are a very close simulation of the real CIE Standard Illuminant D_{65}, and have excellent colour rendering properties.

b) Temporal Stability

In the present case, four tubes were used, two for each side of the viewing apparatus since the exigencies of the present experiment required the ability to alternate the location of the hardcopy and softcopy stimuli. When a softcopy needs to be displayed on a certain side, the illumination on this side obviously needs to be switched off. The light shining on the monitor would otherwise seriously affect its appearance. Repeatedly switching on and off

the will certainly have an influence on the stability of the illumination. The following figures shows how the colorimetric attributes of a blank print evolves when the lights are constantly switched on and off for periods of two minutes. After a warm-up time of several hours, their cycles were initiated by switching them on and off five consecutive times, and the measurements started. Each time the lights were switched back on, a series or five measurements were done at 20 seconds interval. The resulting measurements are displayed in Figure 12 and Figure 13, first for the left side only. Similar results were obtained with the right side of the viewing apparatus.

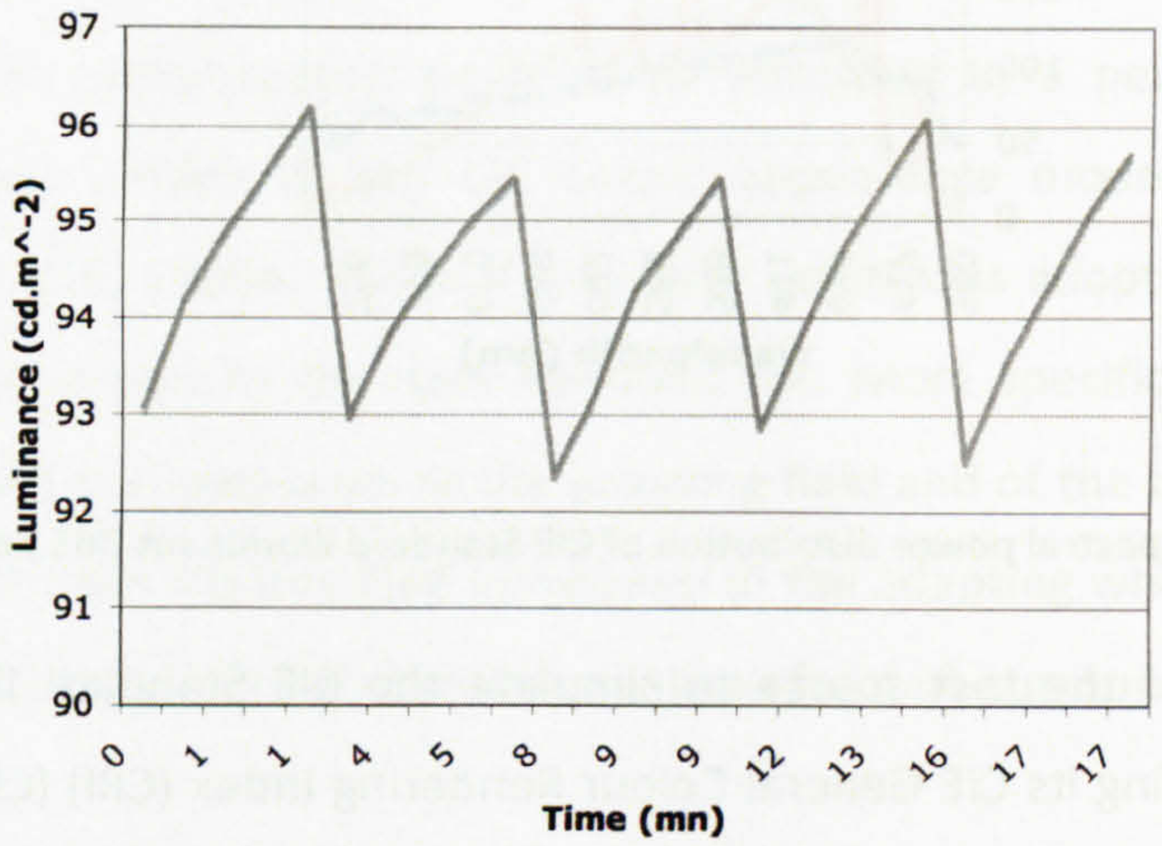


Figure 12 - Viewing apparatus: influence of the switching process on the stability of the luminance of the daylight simulators

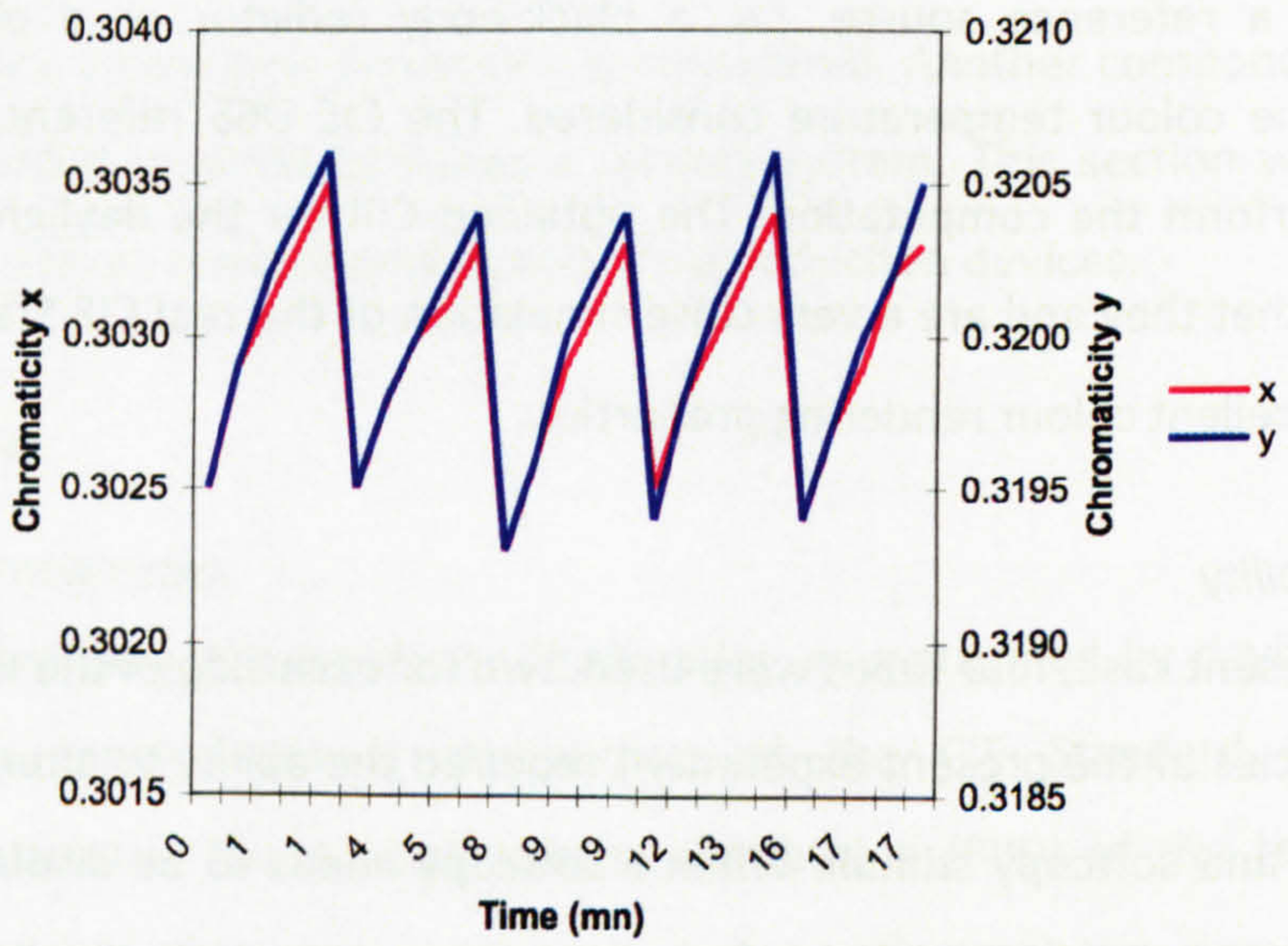


Figure 13 - Viewing apparatus: influence of the switching process on the stability of the chromaticity of the daylight simulators

The previous figures clearly show both the luminance level and the chromaticity of the illumination changing as the tubes warm up again after being switched off. Constantly switching on and off the lights seriously affect the stability of the illumination, which could then also affects the state of adaptation of the observer. Since the stability of the monitor is much better, as it is never switched off, this could have a disastrous effect. Temporally masking the fluorescent tubes with baffles would also be impractical while running the experiment. The next figures shows how the colorimetric attributes of a blank print illuminated by the daylight simulators evolve when they are switched on and remain so permanently for a long period. Once again, only the right side results are shown, the left obtaining very similar data.

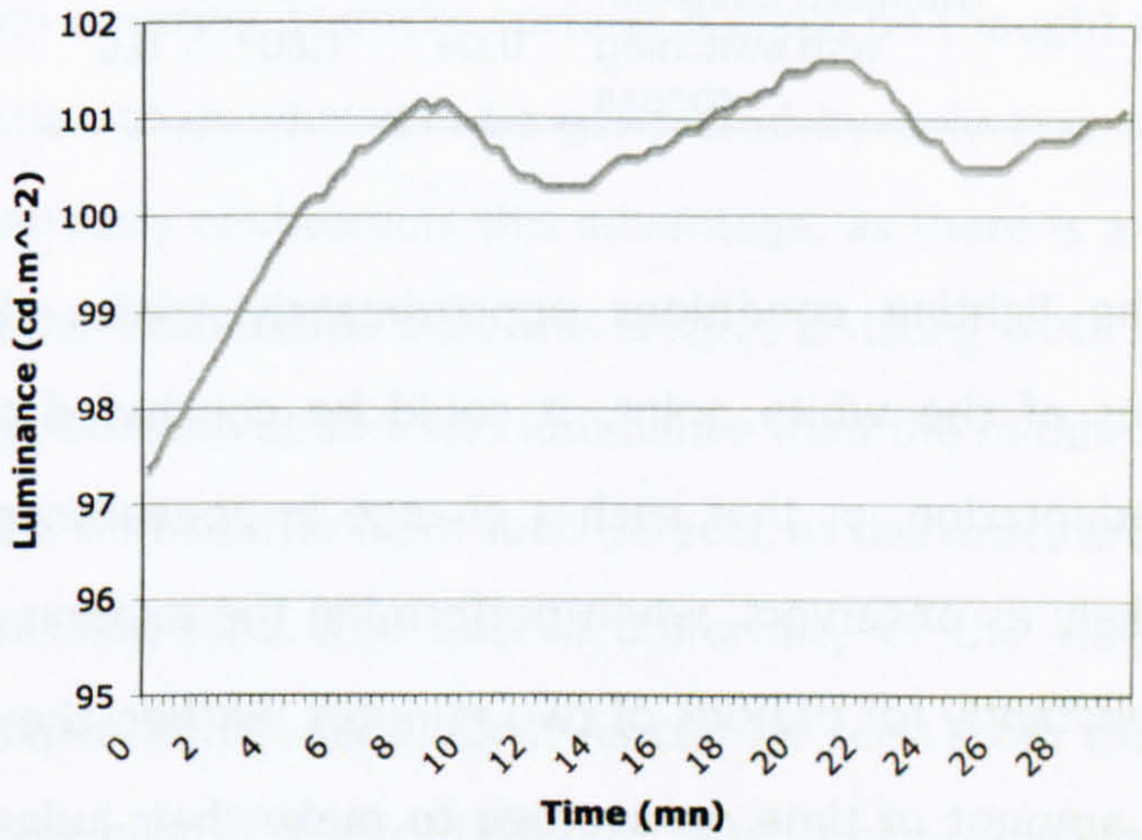


Figure 14 - Viewing apparatus: inherent stability of the luminance of the daylight simulators

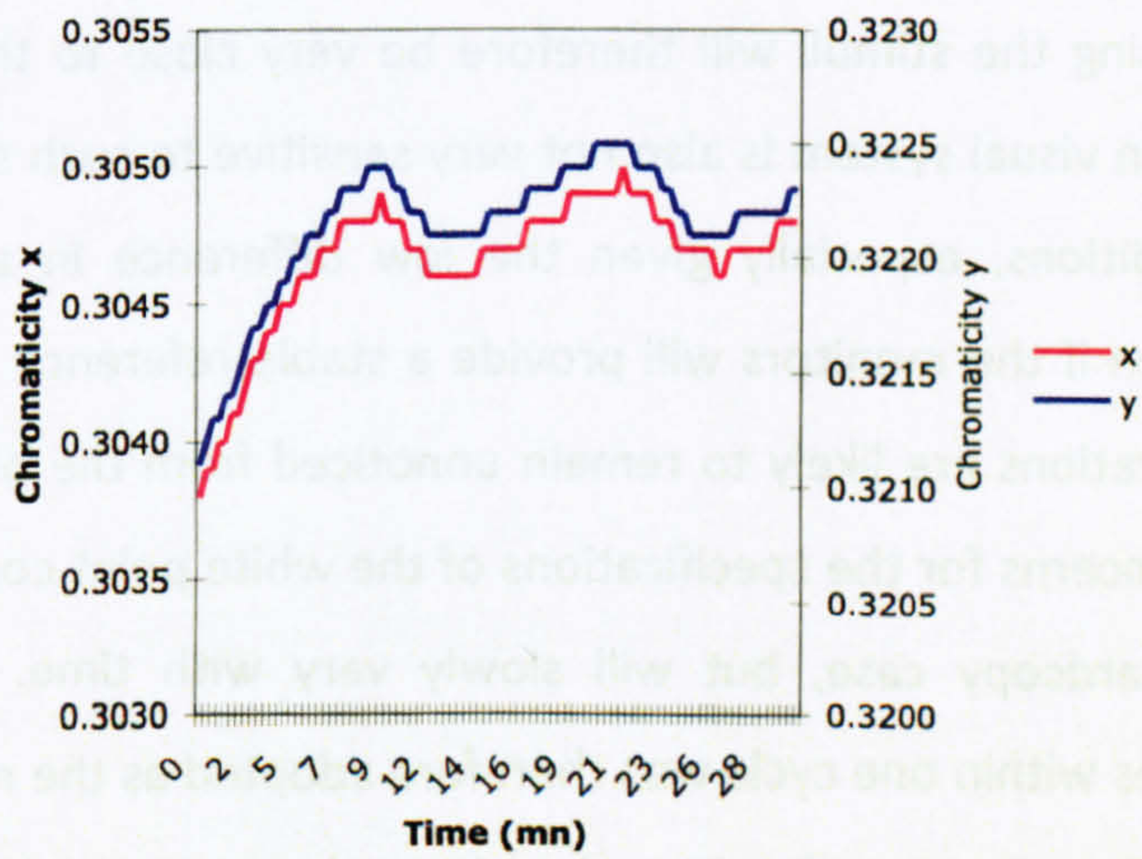


Figure 15 - Viewing apparatus: inherent stability of the chromaticity of the daylight simulators

It can be seen from the figures that a stable state is reached after about 10 to 15 minutes. However, this state is definitely not constant, but would better be described by a sinusoid function with an approximately 12 minutes long cycle. Thus, even under optimal conditions for stability, it would be impossible to achieve an illumination configuration that remains constant over time. The following table shows how the variability due to the switching process compares with the intrinsic variability of the illumination.

Table 4 - Comparison of the variability of the colorimetric attributes of the white point for both sides of the viewing apparatus

	L*	a*	b*
Standard deviation without switching process	0.01	0.67	0.01
Standard deviation with switching process	0.04	1.80	0.0

Alternating the lighting conditions approximately triples the variability of the colorimetric attributes of the white point. It could be concluded that this could affect observers' state of adaptation, or that such a change in appearance could be perceived. However, this is unlikely as observers, when performing the experiments, will not stare at the diverse stimuli constantly for periods of two minutes. Rather, they probably will need a considerably shorter amount of time on average to make their judgements. Furthermore, the average colour differences between two successive measurements made at 20 seconds interval was found to be 1.22 ΔE^*_{ab}. The change in appearance that will occur while observers are analysing the stimuli will therefore be very close to their colour difference threshold. The human visual system is also not very sensitive to such slow temporal change of white point conditions, especially given the low difference in appearance observed (Fairchild, 1993). Even if the monitors will provide a stable reference for comparison, such slow and minor alterations are likely to remain unnoticed from the observers. Rather, this will cause greater concerns for the specifications of the white point conditions, as it will not be unique in the hardcopy case, but will slowly vary with time. The average of the colorimetric attributes within one cycle was therefore adopted as the reference colorimetric attributes for the white point. Another important issue that emerges from this analysis is the colorimetric characteristics of the variation. The lightness and blue-yellow channels remain

virtually unaffected by the temporal changes, which only influence the chromaticity of the red-green channel. The chroma of colours will therefore be the most affected dimension, especially along the red-green axis, although the hues for yellowish and bluish colours might also suffer from some shift. The overall stability of the system can be considered, if not optimal, then satisfactory.

c) Uniformity of illumination

The appearance of a print is mostly determined by the colorimetric attributes of the illumination, but also by its geometry. Its influence on the amount of veiling flare present has already been mentioned, but the spatial arrangement of the tubes also affect the uniformity of the illumination, hence the stimuli appearance. Since viewing booth often contain an array of tubes covering its entire surface, it could be thought they provide a much more uniform illumination than what can be generated by only two tubes. However, the geometry of viewing seriously counteracts this advantage, as there is a 90° angles between observers' viewing and the illumination axis. This creates a strong vertical lack of uniformity, as the bottom of the booth receives 10% less luminance than the middle. On the other hand, the designed apparatus will not suffer from such defects, as the tubes are mounted vertically and span the entire viewing field. The overall uniformity of the viewing apparatus was measured under real experimental conditions. As can be seen from the next figures, both sides exhibit only a slight lack of horizontal uniformity, which however remains perfectly acceptable, as the average ΔE^*_{ab} colour difference was less than 2 units. It is half of what a typical viewing booth can achieve.

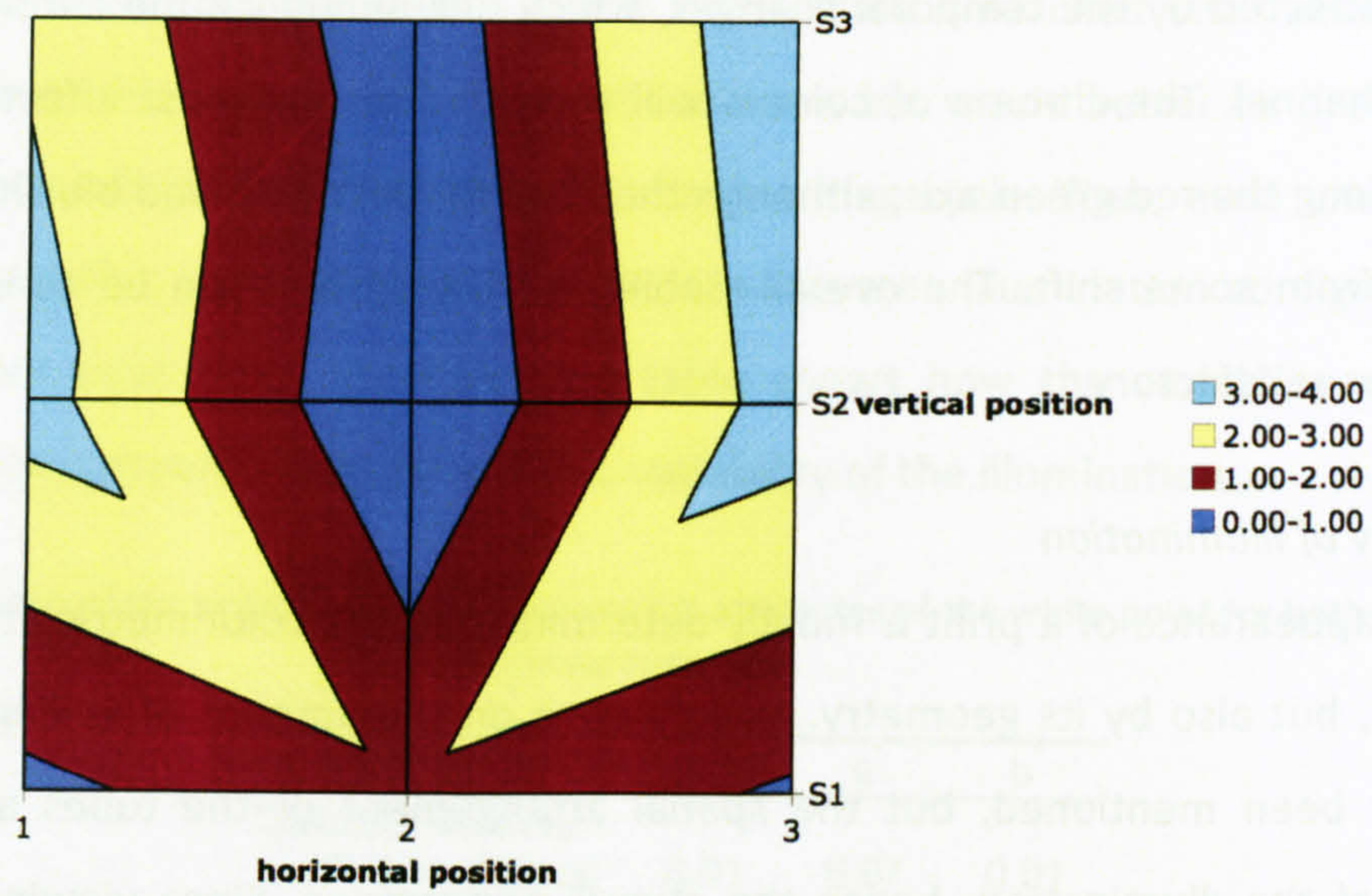


Figure 16 - Viewing apparatus: uniformity of right side (expressed in terms of ΔE^*_{ab} from the center)

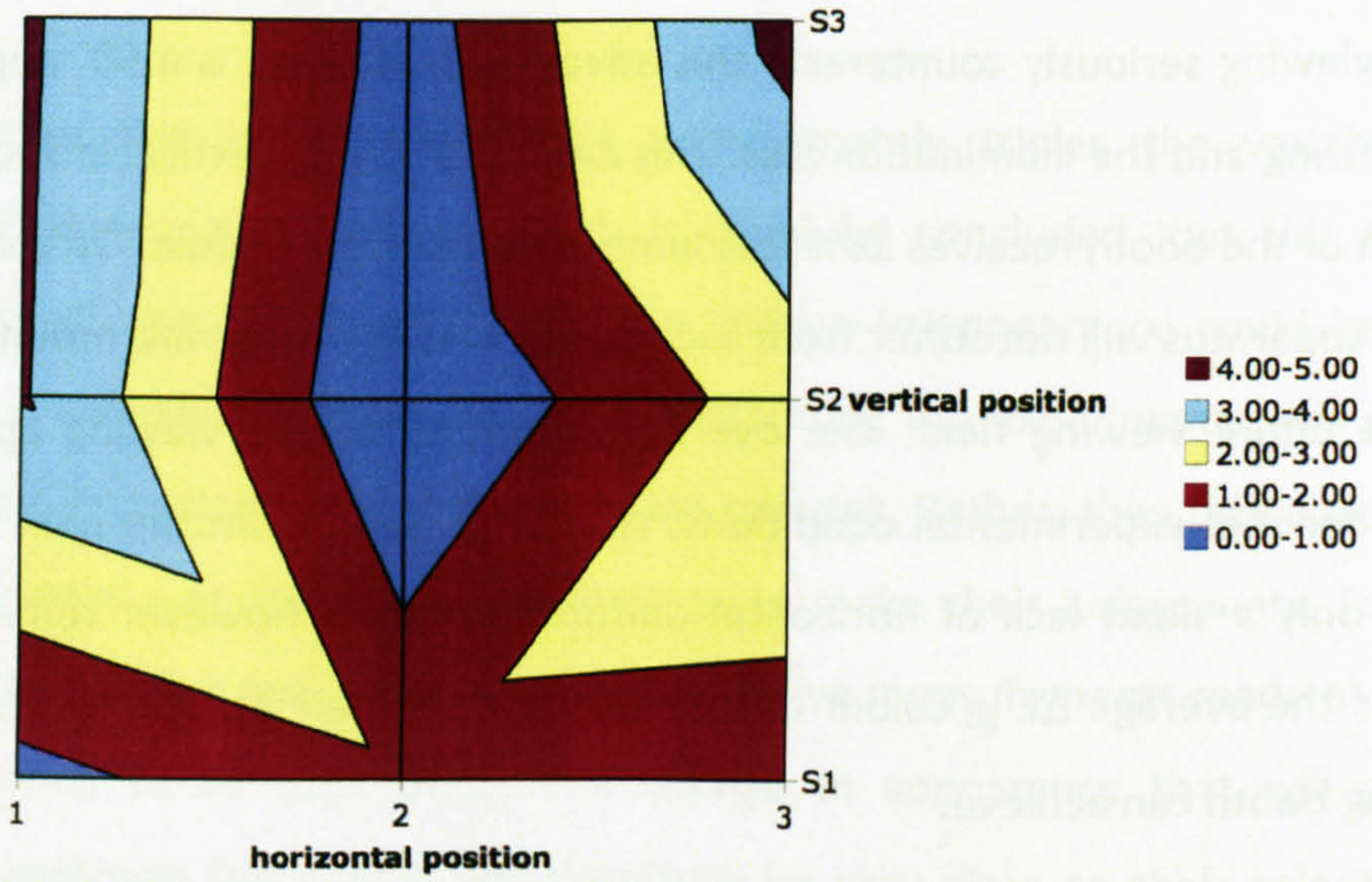


Figure 17 - Viewing apparatus: uniformity of left side (expressed in terms of ΔE^*_{ab} from the center)

d) Calibration

"Device calibration is the setting of the imaging device to a known state" (Fairchild, 2004). The daylight simulators were actually the components of the colour reproduction framework whose calibration required to be overseen the most. Although they were never switched off throughout the whole experimental phase, that was not sufficient enough to ensure their state remain constant. Their inherent variability certainly did not simplify the calibration procedure, which only fortunately targeted their luminance level, the chromaticities of the illumination being obviously fixed. The optimal method for the task

consisted in adjusting appropriately the light intensity until five measurements made at one minute interval produced results consistent with the magnitude of the cyclic behaviour previously described. This procedure was performed daily throughout the experimental phase.

4.5.2 Printer

Hardcopy reproductions of original images displayed on the monitors were made on a *Xerox© Phaser 7300 DN* LED printer at a resolution of 300 dots per inch. A single batch of colorants was used throughout the experimental setup and the stimuli generation process. Prints were obtained by sending the corresponding data to the printer via *Xerox*'s software driver from an *Apple iBook©* computer running *Adobe Photoshop©* version 7.01 under *MacOS X* version 10.2. The built-in colour management system of the software, the operating system and also the printer driver were disabled.

4.5.2.1 Spatial characteristics

The discussion has mainly concentrated until now on the colorimetric aspect of the reproduction framework. This is right, as it is likely to generate the greatest difficulties, but a rigorous analysis need to consider every aspects of the reproduction. Obtaining a faithful reproduction of an original involves first and foremost to precisely duplicating the spatial information of the stimuli. Owens and VanSant provide a clear overview of the difference image quality characteristics, focusing on spatial properties, used to assess the worth of a printing system (Owens & VanSant, 1999).

A first group includes considerations that are unique to digital system, and thus are directly relevant here. The spatial artifacts mentioned are all related to sampling effects, which will not occur in this case as the size of the image sent to the printer will match its native resolution, and will thus not need to be down-sampled. The second group of image quality characteristics affects both digital and conventional printing systems. The first type consists of geometrical distortions to which the human visual system is particularly sensitive, as *"a distorsion of 2% is readily detected in the image of a face"* (Owens & VanSant, 1999). The resolution chart provided by the ISO Standard 12233, although not primarily intended for this kind of applications, can nevertheless be used to precisely assess the presence of such artifacts in prints (ISO, 2000b). The vectorial image of this chart was thus sent to the

printer at the proper resolution. Several distances between different items of the chart measured on the obtained print were compared to the reference distances given by *Adobe Photoshop*, and showed an accuracy level down to the millimeter, which, given the viewing distance, should be totally satisfactory for the current needs.

Another type of image quality characteristics is noise, defined by the authors as *"those artifacts that are not related to image content. They may be periodic or random, caused by the capture device or the printer, and include banding, rastering, streaking, mottle and grain*[...]. *All of them are normally measured using a target consisting of large neutral patches at* [specific] *densities* [...] *scanned with a microdensitometer"* (Owens & VanSant, 1999). Large uniform gray patches with a luminance factor of approximately 0.5 were thus printed, but the lack of appropriate equipment did not allow quantifying the extent of the noise. The analysis was instead carried visually, following the authors' recommendations concerning this type or artifacts: *"Our overall specification was that they could be above the threshold of perception in flat fields, but not above the threshold of objectionability"*. The visual inspection of the uniform patches revealed a fair amount of banding and graininess present in the resulting print. Only large uniform colour areas were affected, leaving pictorial images free from these defects. It could however pose a great threat to the reproduction quality of business graphics images, as well as the stimuli background, since both contain large uniform zones. In spite of this, the experimental viewing configuration, more specifically the dark surround conditions, hid these defects of the hardcopy, as the luminance difference between the adapting field and the surround was extremely high, effectively making all the small colour differences caused by graininess invisible. Observers, once they completed all the experiments sessions, were asked all casually whether they could detect any lack of uniformity present in the print background, and their answers were firmly negative.

Sharpness is the last important spatial characteristics affecting image quality according to Owens and VanSant. Obtaining sharp edges free of any fringing or flare is clearly a desirable feature, especially for business graphics images where reproducing high-quality text is primordial. The analysis was once again performed visually by assessing the quality of "a *step tablet containing white lines of 1, 2, 4 and 8 pixels width running through the steps*" (Owens & VanSant, 1999). Some 4-point text was also added to the image in order

to check its readability. Despite the presence of the noise artifacts previously described, the resulting print show crisp edges visible for all the different lines width, without any fringing. The 4-point text was furthermore very readable. The general conclusion that can be made about the spatial quality of the printer is that, although some noise could be detected under close examination, it did not affect the quality of the print enough to raise objections. It was furthermore invisible under experimental conditions, which lead to the conclusion that the spatial characteristics of this printer are sufficient for the intended purpose.

4.5.2.2 Colorimetric characteristics

a) Device-dependent colour space

Colorimetric data in the device-dependent space can be represented in two different colour spaces. CMYK is classically employed in the Graphic Arts industry, as it is a direct quantitative representation of the colorants transferred onto the substrate, but RGB has become popular with the advent of desktop publishing. Since the main concern with regard to the characterisation process is to maximise the colour reproduction accuracy, the suitability of each colour-representation mode was first investigated.

Defining a characterisation model using CMYK is intrinsically more complex than with RGB, as the former has a higher dimensionality than the device-independent space. An inverse transform, whereby colorimetric or perceptual data are transformed back into device-dependent ones, cannot be defined in a unique manner as the functional representation is non-injective. Constraints therefore need to be added to the set of equations in order to achieve the unique mapping desired. They usually take the form of an additional function whose purpose is to define a bijective transform between CMY and CMYK data. The problem can then be solved by injecting this function in the inverse transform, which would then consist first of a transform between two three-dimensional spaces and finally of the function transforming CMY data into CMYK. Diverse methods are used to derive such functions, and the one implemented in this case is named *gray component replacement*. It consists in quantifying the amount of reduction in the CMY primaries to compensate for the addition of black colorant. Conversely, characterising a printer using RGB as the device-dependent colour space is straightforward, as the input and output colour spaces have the same dimensionality.

Two characterisation models, based on a polynomial interpolation and a neural network, were implemented in order to test the appropriateness of both device-dependent colour spaces. Their parameters were derived from the spectral measurements made of a standard ANSI IT8.7/3 (ANSI, 1999) printed reflection target and tested with 81 randomly chosen colours sampling the CMYK colour space approximately uniformly. The results, shown in the next table, clearly demonstrated the superiority of the RGB colour space, since its average error for all tested models was 28% less than the CMYK space. Although using RGB will probably limit the colour gamut achievable by the printer to the standard sRGB colour space (IEC, 1999), it was nevertheless chosen since the main concern in this experiment is colour accuracy.

Table 5 - Descriptive statistics of ΔE*ab distribution generated by characterisation models for both device-dependent colour space

	ΔE^*_{ab}	average	maximum	standard deviation
RGB	5^{th} order polynomial interpolation	3.07	9.14	1.39
	50 nodes neural network	3.11	9.56	1.34
CMYK	5^{th} order polynomial interpolation	3.68	9.45	1.39
	50 nodes neural network	3.59	10.15	1.37

b) Uniformity

Spatial uniformity can be considered both as a spatial or a colorimetric characteristic of printers. Most of the following discussion can therefore directly be juxtaposed to the previous analysis on the spatial characteristics of the printer. Analysing this characteristic from a colorimetric point of view simply consists in measuring at different locations along a grid the spectral reflectance of a uniform gray print with a reflectance factor of 0.5. The uniform grid consisted of 15 sampling points along the length of the print and 12 across its width. The colour differences between the average of the six central measurements and each point on the grid are displayed shown in the following figure.

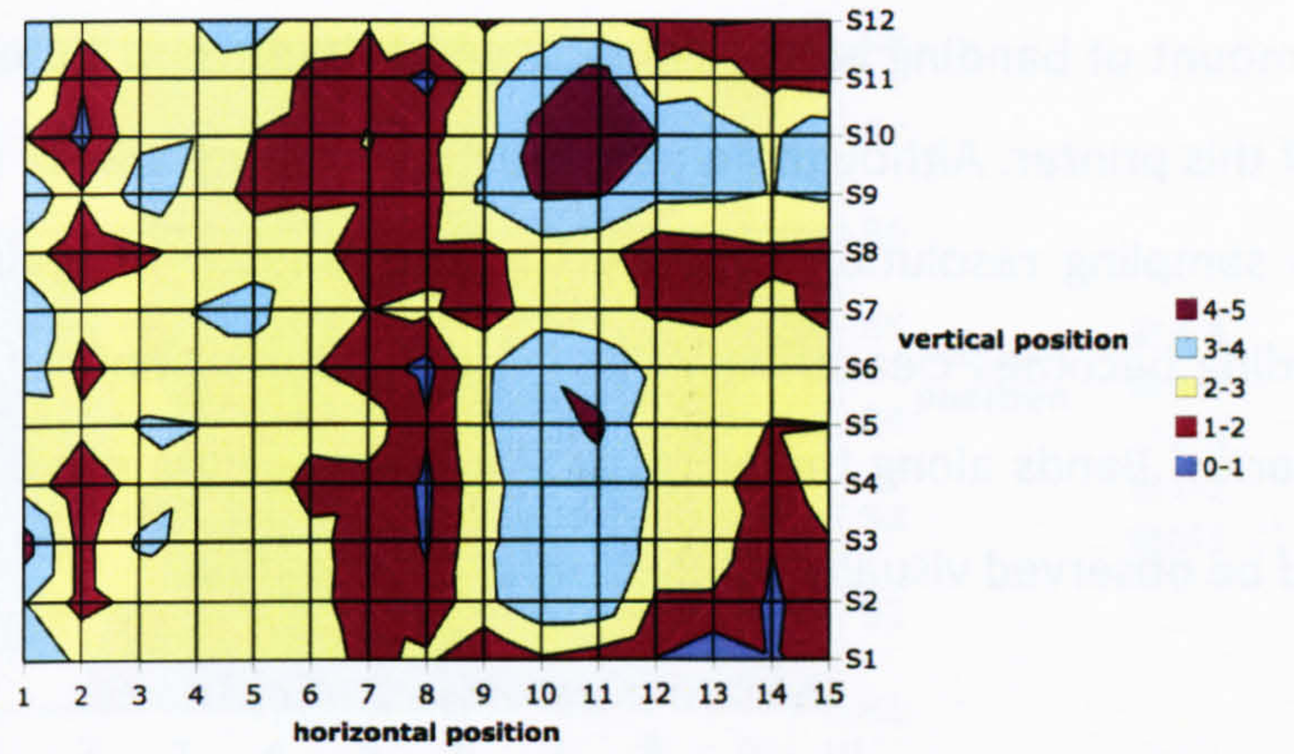


Figure 18 – Chart of Phaser 7300 uniformity for uniform gray print (expressed in terms of ΔE^*_{ab} from the center)

Table 6 - Phaser 7300 uniformity: descriptive statistics of ΔE*ab distribution generated by the different points on the sampling grid for a uniform gray hardcopy

	Average	maximum	standard deviation
ΔE^*_{ab}	2.38	4.48	0.88

The uniformity achieved by the printer can be described as sub-standard. The average ΔE^*_{ab} recorded in Table 5 is high, and the histogram of the ΔE^*_{ab} population shown in Figure 19 shows that differences less than 2 ΔE^*_{ab} units account for less than 10% of all the errors.

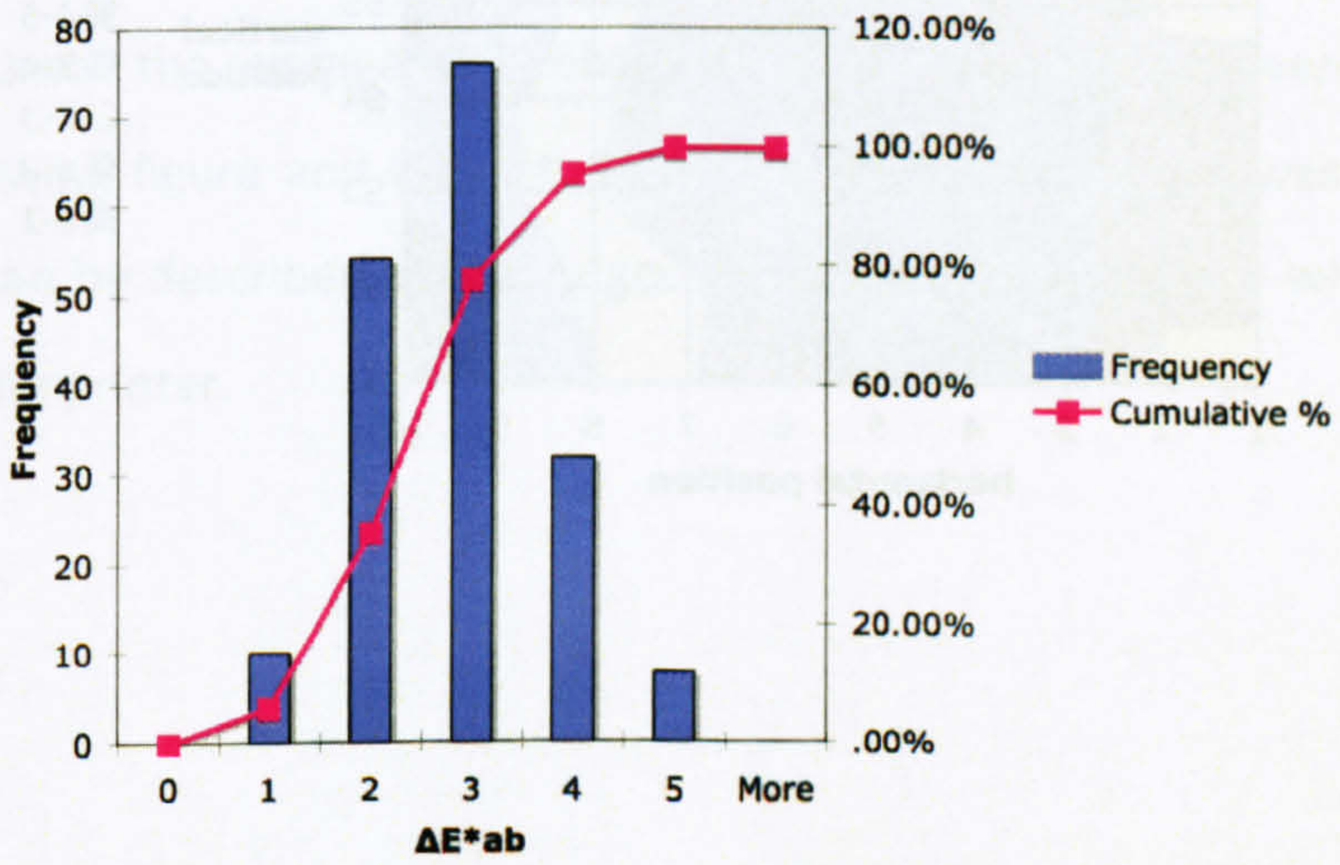


Figure 19 - Phaser 7300 uniformity: histogram of the colour differences population expressed in terms of ΔE^*_{ab}

This was to be expected, as the visual inspection previously carried out revealed a relatively high amount of banding and graininess, which both contribute to severely reduce the uniformity of this printer. Although no real structure is detectable in the figure, probably due to the high sampling resolution combined to the amount of graininess present, the presence of banding becomes clearer when analysing uniform prints of the cyan, magenta and yellow primaries. Bands along the width of the hardcopy are more readily observable, and could indeed be observed visually with a careful examination.

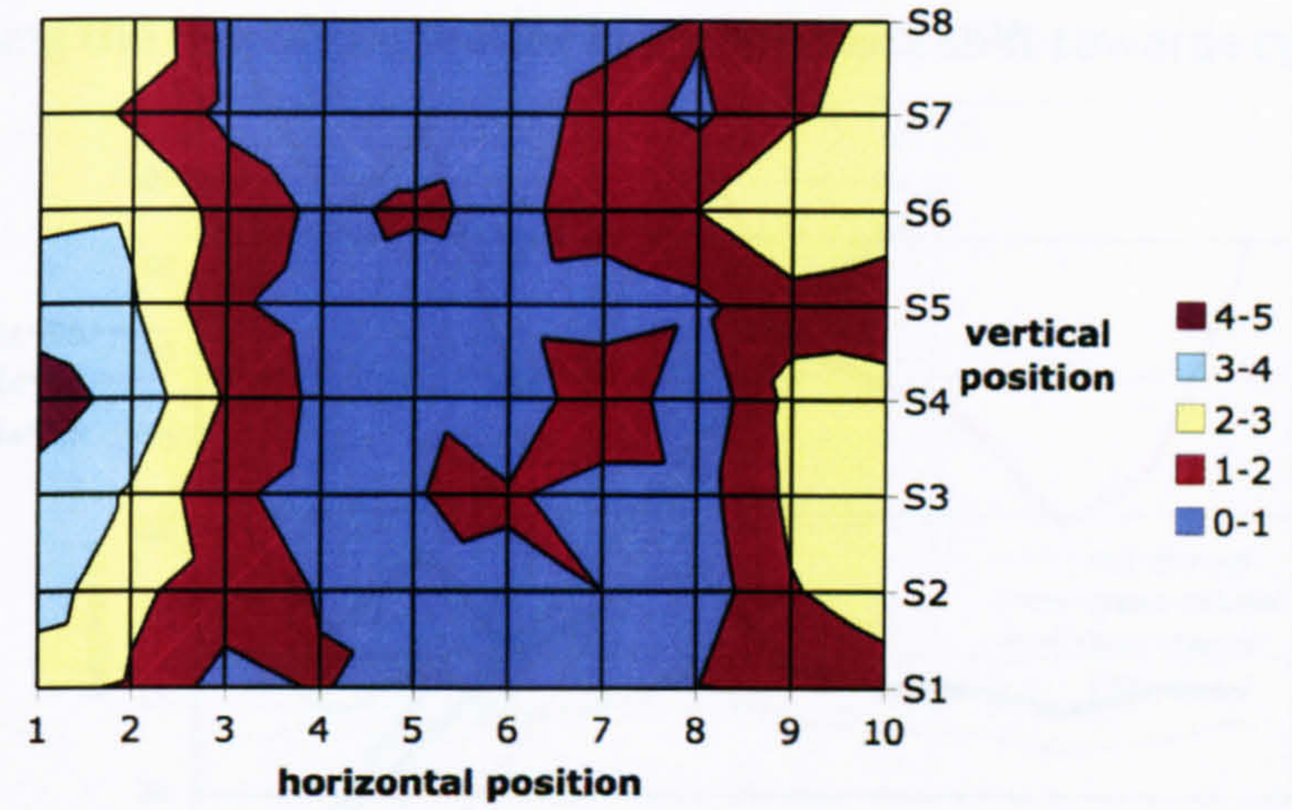


Figure 20 - Phaser 7300: uniformity of CMY primaries, cyan (top), yellow (middle) and magenta (bottom)

Despite the relatively high magnitude of non-uniformity exhibited by this printer, the verdict concerning its ability to reproduce colours still remains favorable. As previously mentioned, spatial artifacts, although they can be detected, are not strong enough to make the print objectionable, especially under the unnatural experimental conditions designed. It is however likely that this non-uniformity will significantly reduce the overall accuracy of the printer in terms of colour reproduction, and thus also of the whole reproduction framework.

c) Repeatability

Repeatability is another important characteristic of printers. The ANSI IT8.7/3 chart was printed repeatedly over a period of twelve weeks, and the colour differences between the different weeks and the average ΔE^*_{ab} obtained for the first day sets were computed. As shown in the following figure and table, the average colour difference was slightly over 1 ΔE^*_{ab} unit, which can be described as a very good performance especially when considering the uniformity of the printer.

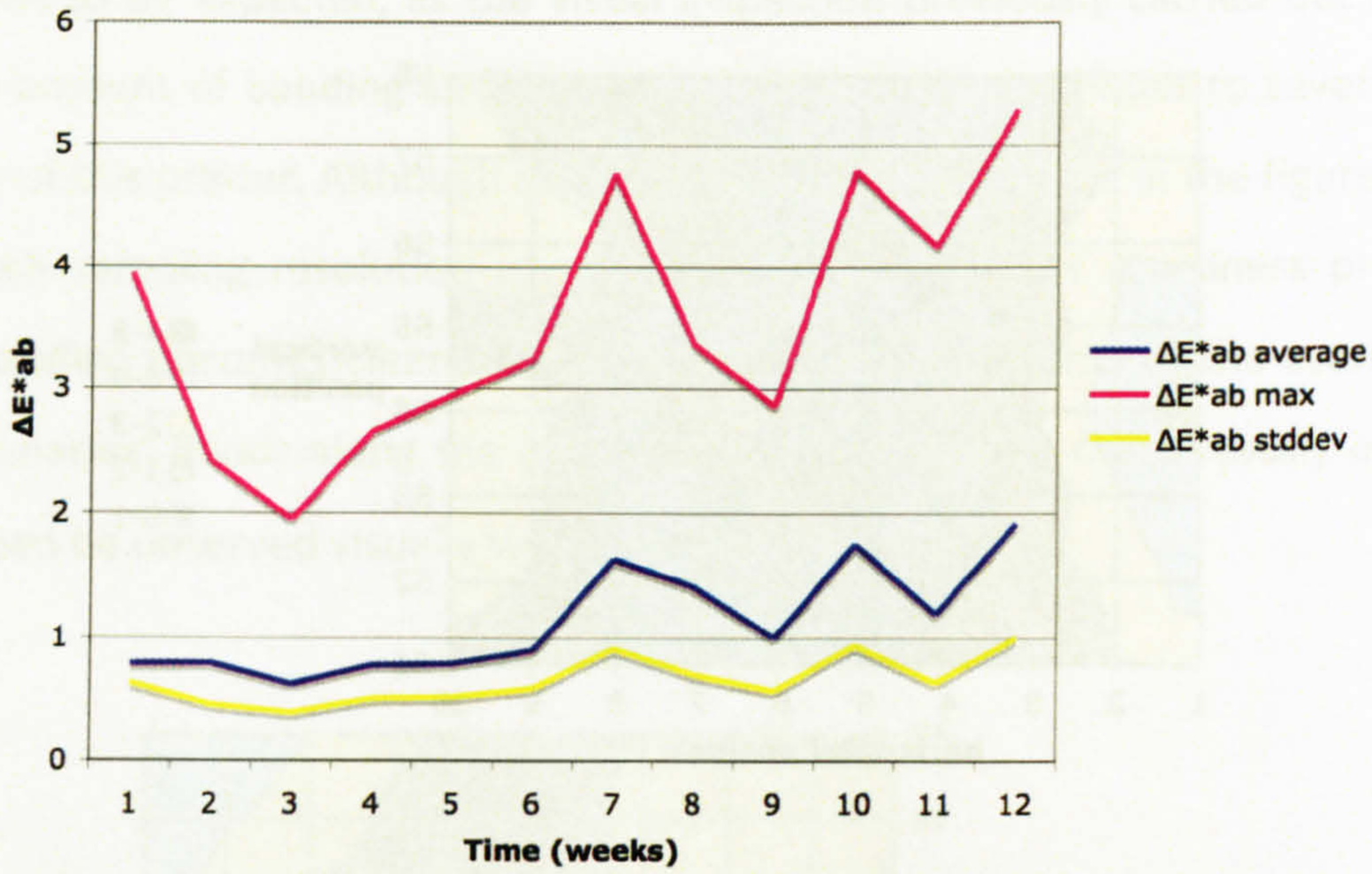


Figure 21 - Phaser 7300 repeatability

Table 7 - Phaser 7300 repeatability: descriptive statistics of ΔE*ab distribution generated by the sets corresponding to different weeks

ΔE^*_{ab}	average	maximum	standard deviation
	1.13	3.53	0.65

d) Tone reproduction and addressability

As previously mentioned, the ability of a printer to reproduce accurately large areas of uniform colours is important, but the capacity to generate smooth and continuous transitions between colours is also crucial. Since the human visual system is most sensitive to spatial information in the luminance channel (Hunt, 1995), the visual attribute most commonly employed to assess this characteristics is lightness, or more specifically whether it is reproduced smoothly from minimum to maximum, which Owens and VanSant refers to as "*achieving D-min to D-max with grace and robustness*" (Owens & VanSant, 1999). The set of RGB triplets necessary to obtain neutral colours, i.e. defined as $a^*=b^*=0$, across the lightness range was derived empirically and is shown in Figure 22. The tone reproduction can be considered as satisfactory, as the curves obtained for each channel are very smooth. The addressability of the printer, i.e. its ability to reproduce small variations at the extremes of the input data range, can also be assessed from this figure. The darkest neutral colour that can be obtained with this printer has a CIE 1976 lightness L* of approximately 14 units. The red channel is the limiting factor, as digital counts below 30 do not generate any colour

variations. The two remaining channels can however still reproduce small transitions below that value, indicating the hue of the printer black point will shift towards cyan.

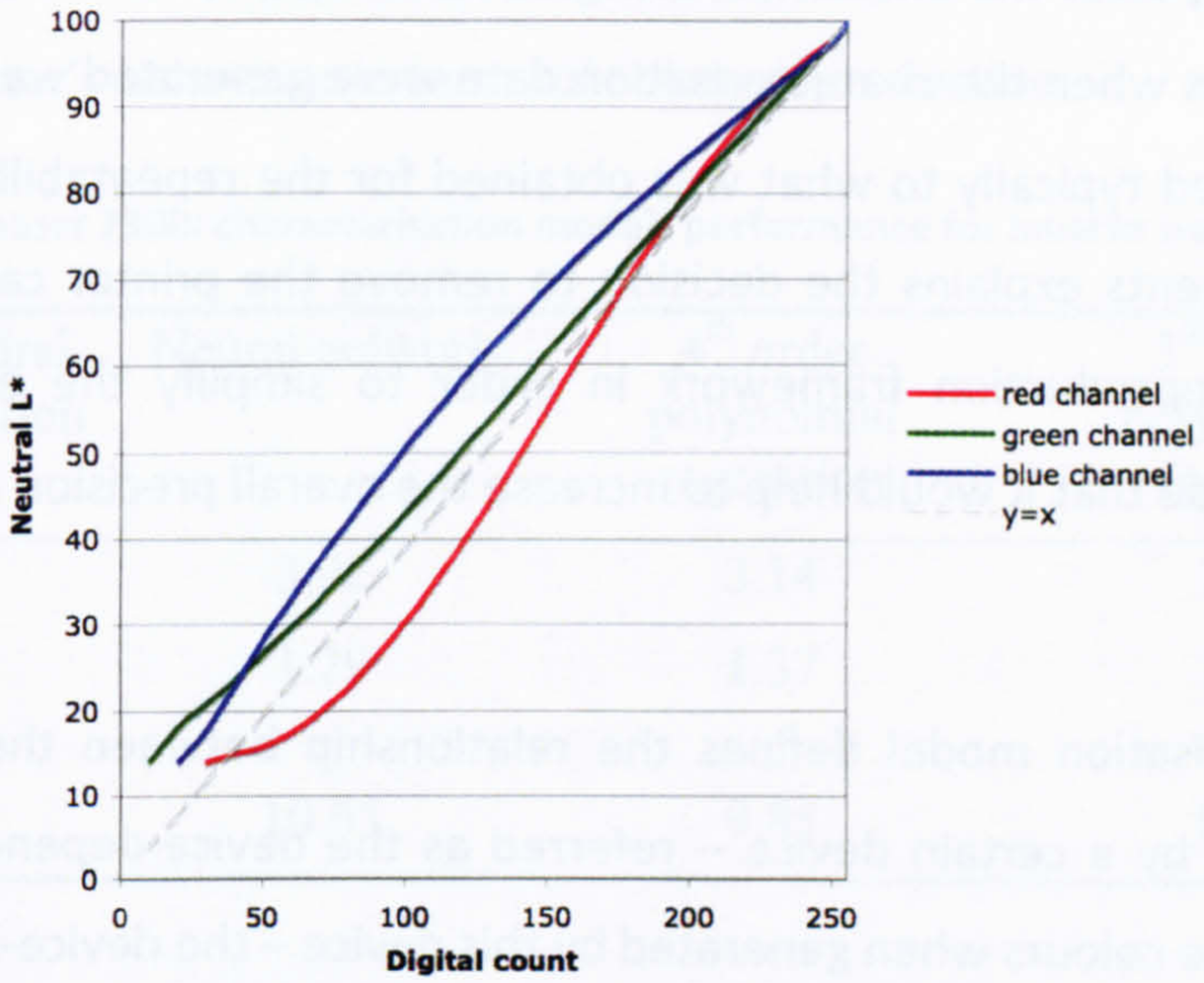


Figure 22 - Phaser 7300 tone reproduction curves

e) Calibration

The process of calibrating a colour reproduction device usually consists in setting it up in a specific and known state in order to obtain consistent results. The situation is slightly more complex for printers, as the procedure contains another phase destined to derive a calibration function. Its purpose consists usually in transforming the device-dependent data before they are sent to the printer in order to obtain a certain linear relationship between them and the colorimetric attributes of the hardcopy. A *gray-balancing calibration* technique is typically used, whereby a transform is defined "*so that equal amount of C, M, Y processed through the calibration results in a neutral (i.e. $a^*=b^*=0$) response*" (Bala, 2003). The associated procedure consists in finding empirically the combinations of colorants that generate neutral colours for different lightness levels, and then derive calibration functions for each channel by simply inverting inputs (digital counts) and outputs (scaled lightness levels) of the previous relationship. The relationship between digital counts and lightness for each channel is displayed in Figure 22. The automatic gray-balancing performed by the printer software can be easily observed, as the relationship for each channel is approximately linear for most of the input range, except for dark colours where higher order polynomials are needed to precisely describe the relationship. The use of calibration

functions in this case would thus be superfluous. Moreover, the repeatability of the printer was also found to be of satisfactory accuracy, and the printer was only used once, to produce the stimuli, after the end of the characterisation process. Assessing whether it was in the same state as when the characterisation data were generated was sufficient, and the results corresponded typically to what was obtained for the repeatability assessment. This amalgam of arguments explains the decision to remove the printer calibration procedure from the colour reproduction framework in order to simplify the already complicated equation, in the hope that it would help to increase the overall precision obtained.

f) Characterisation

A characterisation model defines the relationship between the representation of colour information by a certain device – referred as the device-dependent space and the appearance of these colours when generated by this device – the device-independent space.

As shown in the outline of the reproduction framework, the role of the printer is to accurately reproduce a set of given colours, defined either by their colorimetric or appearance attributes. An inverse transform, whereby device-independent colorimetric data are transformed back into the printer's internal representation of colour information, is required. Such transforms are usually derived from two types of characterisation models, "*physical models and empirical techniques* [that] *offer different trade-offs between effort and accuracy*" (Bala, 2003). Due to the complex non-linear relationships between colorants, the performance of physics-based characterisation models is generally considered worse than empirical methods. The absolute precision of the colour reproduction being the main concern, only generic models were developed. Several techniques were tested: tetrahedral interpolation, fourth and fifth order polynomial regression and a standard back-propagation neural network trained with a Conjugate Gradient algorithm (Masters, 1993). The parameters for each model were derived from empirical data sampling uniformly the device-dependent space. Sample sets with up to 25 points per channel were laboriously obtained, but it was found that no significant improvements were made by using more than eleven sampling points per channels. Once the parameters of each models were derived, an independent testing set was formed by choosing randomly 189 colours which samples the device-independent colour space approximately uniformly. They were transformed using the four models into RGB data that were printed and measured. The performance of each model

was finally assessed by computing the colour differences between the actual colours and their predictions. The results are shown in Table 8. The lowest average and maximum errors were obtained by the 5th order polynomial regression, which was consequently chosen for representing the printer's behaviour in terms of colour reproduction.

Table 8 - Phaser 7300: characterisation models performance for inverse transform

ΔE^*_{ab}	Tetrahedral interpolation	Neural network	4th order polynomial regression	5th order polynomial regression
Average	3.28	3.02	3.14	2.97
Standard deviation	1.63	1.29	1.37	1.32
Maximum	10.91	10.03	9.85	8.77

It is important to note that the resulting colour differences do not correspond to the intrinsic performance of the printer usually represented. Since the output of the implemented transform is in the device-dependent space, it is mandatory to generate another print in order to colorimetrically quantify the accuracy of the transform. This quantification is therefore also subject to the inaccuracies of the printer and the measuring instrument. A more pertinent overview of the printer performance can be had from the precision of the forward transform, i.e. from device-dependent to device-independent space, which is shown in Table 9. The best model is once again the 5th order polynomial regression, which obtains results that can be described as average.

Table 9 - Phaser 7300: characterisation models performance for inverse transform

ΔE^*_{ab}	Tetrahedral interpolation	Neural network	4th order polynomial regression	5th order polynomial regression
Average	2.77	2.61	2.76	2.44
Standard deviation	1.53	1.17	1.21	1.22
Maximum	7.03	6.81	6.80	6.46

4.5.3 Print

The last part of the subtractive colour reproduction system to be analysed is the substrate itself. The one used throughout the diverse experimental phases was *Xerox*'s *Glossy Coated Paper.*

a) Temporal Stability

Although the temporal repeatability of the printer had a mild importance, the manner with which the colorimetric attributes of the print decay with time is crucial. Given the great length of time required to complete the data collection phase, the temporal stability needs to be very steady, as any changes in the appearance of stimuli is likely to impair the pertinence of the results. Alternatively, other prints could be produced to replace the originals, providing that the printer is still in the same state as when producing the original stimuli, which is likely since its repeatability was found to be very good on average. However, the maximum error due to repeatability issues was significant enough to seriously affect the results if it occurs in a critical zone, such as well-known natural colours - sky, grass or skin for instance. It would therefore be much safer to use the same prints throughout the experimental phase, providing that their colorimetric attributes do not decay beyond a reasonable threshold. Ensuring the stability of stimuli, and incidentally observers' judgements, is essential.

The temporal stability of the print was investigated in a very similar way as its repeatability, i.e. by measuring over a period of sixteen weeks the colorimetric attributes of each patch on three ANSI IT8.7/3 charts printed the same day. In between measurements, the prints were placed in envelopes to protect them from any potentially disturbing environmental factors, such as light, and were kept in a temperature-controlled room. Figure 23 shows the colour difference obtained between the measurements performed at different weeks and the average ΔE^*_{ab} obtained for the first day set.

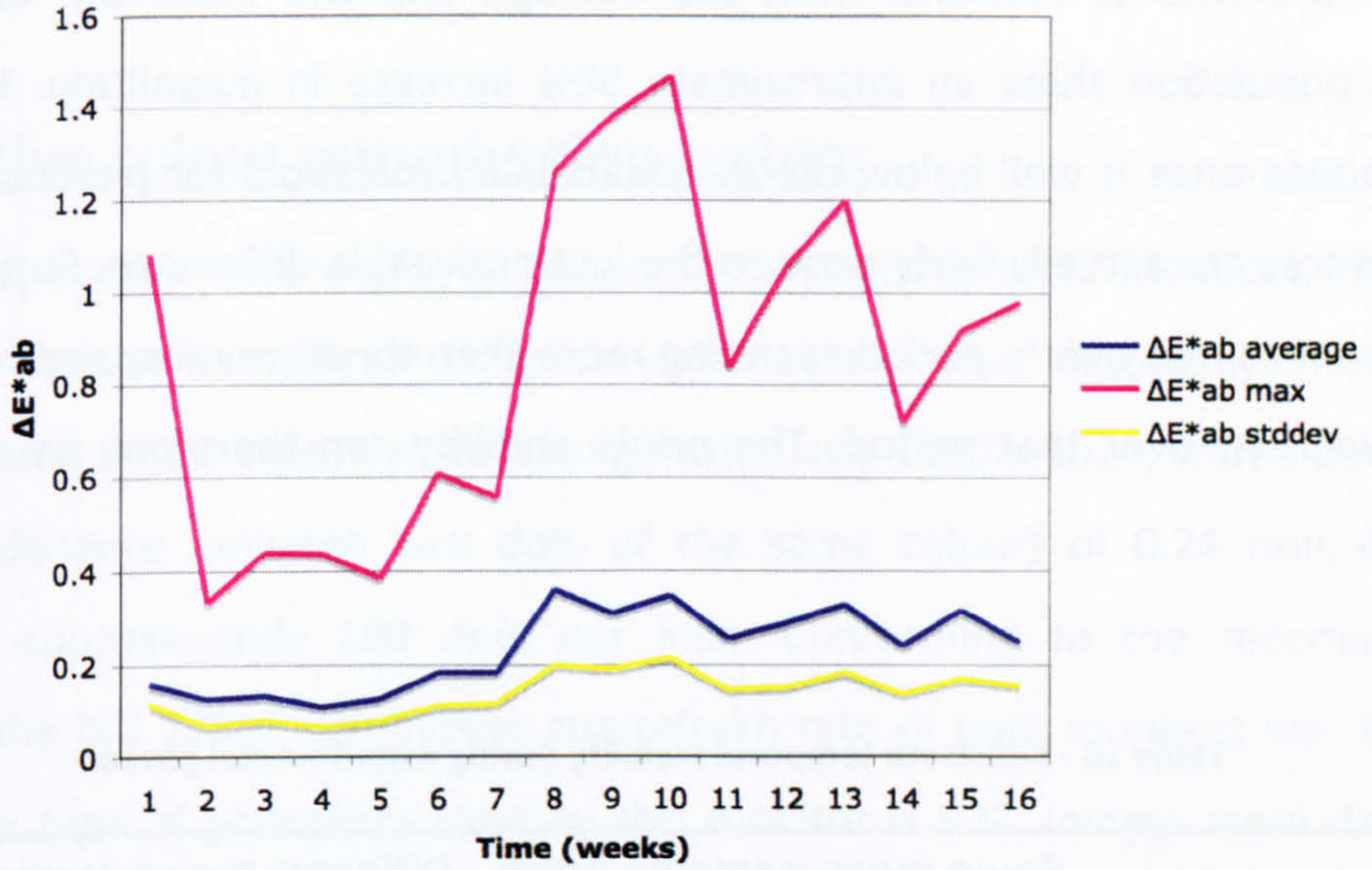


Figure 23 - Substrate temporal stability

Despite the fact that the evolution of the errors seem to be bi-modal, as they suddenly doubled after the eighth week, their magnitudes indicate the changes occurring are not significant at any time. This however represents the best possible performance, as the prints were protected in optimal conditions. Manipulating them while performing the experiments, and exposing them to strong lighting conditions, even if for short amounts of time, may considerably change these figures. It was therefore more appropriate to assess the stability of the print *while* performing the experiments. Several zones of uniform colour were chosen in three of the experimental stimuli to perform this investigation. Since each image used in the experiments was printed twice, the same zones were measured in both reproductions for the measuring instrument bias to be taken into account. The measurements were indeed performed manually rather than automatically which, given the relatively poor uniformity of the printer, could increase the variability of the results. Each predefined zones were measured at different times throughout the experimental phase, and

Table 10 shows the colour differences obtained between the measurements made at the very beginning of the experiment and at the very end. The differences between the two reproductions, whether the measurements were made at the beginning or at the end of the experimental phase, are higher than the printer repeatability and the substrate stability, probably reflecting the influence of the non-uniformity on manual measurements. The main observation that can be had from this table is the increase in the colour differences between

the two measurements sessions. Both the average and the maximum of the colours differences population show an approximate 50% increase in magnitude. However, this change in appearance is well below the discriminability threshold for pictorial images, and most differences are actually fairly close to the just-noticeable difference. Furthermore, this change was measured over a period spanning more than three months, and represents the maximal evolution over that period. The prints stability can therefore be considered as satisfactory.

Table 10 - Substrate temporal stability during experimental phase

ΔE^*_{ab}	Same measurements batch	Different measurements batch
Average	0.91	1.32
Maximum	1.77	2.28
Standard deviation	0.07	0.18

b) Gamut

The gamut of the colours addressable by the print is displayed in the next figure. The gamut is maximised thanks to the good CRI obtained by the illuminant, but also minimised by the colour space used for the device-dependent representation of colours.

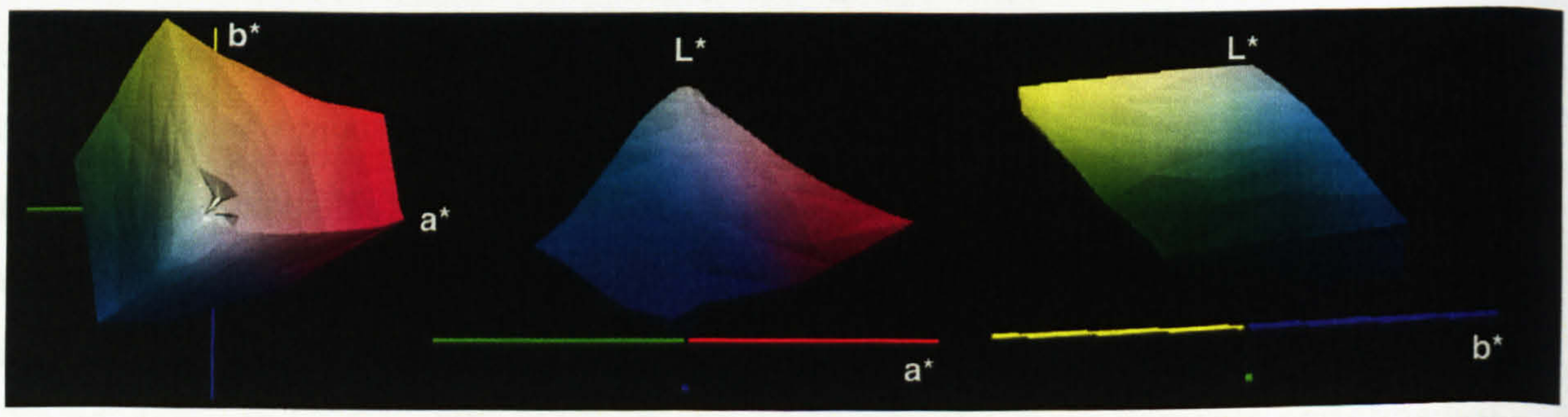


Figure 24 - Print: Gamut

4.6 Additive colour reproduction system

Softcopy stimuli were displayed on two *Lacie electronblue IV* monitors. The diagonal dimension of the tube was 19", resulting in a visible area of 18.1" diagonally. The number of addressable pixels in the experimental configuration was 1280 x 1024, with a dot pitch (i.e. the shortest distance between two dots of the same colour) of 0.24 mm, entailing a resolution of approximately 100 dots per inch. Conforming to the recommendations published by the ISO 12646 committee, the refresh rate of both monitors was 85 Hz non-interlaced. The type of phosphors used by this monitor is *P22*. Images were displayed on these screens using *Adobe Photoshop©* version 7.01 under *Windows NT* version 4. The built-in colour management system of both the software and the operating system were disabled. From now on, the monitor on the left side of the viewing apparatus is referred as *212* and the right side one as *214*.

4.6.1 Spatial characteristics

Although the spatial characteristics of printers are essentially static in nature – only the manufacturer can really interact with them, those of monitors are very much dynamic. In order to account for the gun decrease in spatial accuracy associated with aging, a wide range of geometrical parameters are adjustable: size, position, pincushion, tilt, trapezoid, parallelogram to name just a few. Evaluating the spatial characteristics of monitors does not just consist of verifying their geometrical configuration, but performing an active calibration of these properties.

Given the large variety of settings to take into account, the use of specific software (Nokia, 2005) containing all the appropriate test charts and patterns is highly recommended. The overall goal then consists in maximising the *orthogonality* of the image, i.e. making all lines look straight and all right angles appear perfectly perpendicular when observed from the viewing distance specified by the experimental setup. Furthermore, the dimensions of stimuli displayed on screen also need to be calibrated by adjusting properly the available geometrical properties. More specifically, the resolution chart recommended by ISO 12646 (ISO, 1999), which served for the evaluation of the printer spatial characteristics, was

displayed on both screens, and the geometrical properties were iteratively modified so as to obtain a perfect match between the diverse distances measured on the print and the reference ones read from *Adobe Photoshop*. This last step was found to be of fundamental importance, as the width settings needed to be changed very significantly from a typical setup – leaving a gap of more than 1 inch on both sides – in order to obtain the required dimensions match.

a) Convergence

Convergence is the quality representing how well three coloured lines, red, green and blue overlay each other to produce a white line. If there is misconvergence, colour fringing will occur as the primaries lines will not perfectly overlay each other, which will influence the image sharpness. Since the tubes are based on the Diamondtron technology, the convergence was perfect for most parts of the screen, except on the very edges where small amounts of misconvergence could be observed under very careful examination, while still remaining absolutely acceptable.

b) Focus and Readability

Focus also influences the image sharpness. *"A badly focused monitor gives rise to beam spread, especially in the corner"* (Nokia, 2005). Using the patterns present in the software previously mentioned, a visual analysis revealed that the focus of this monitor was nothing short of perfection, even in the corners. Similarly, a 4-point text was very readable on all parts of the screen.

c) Moiré

"Moiré is a natural interference phenomenon that appears on all colour CRT monitors [...] due to an interference between the shadow mask (or aperture grill) and the resolution used" (Nokia, 2005). Although some ripples and waves could be easily detected while visualising an appropriate high-frequency pattern, the effect of Moiré on typical images such as those used for the experiments were nil, as none of them contain the frequency high enough to initiate the phenomenon.

Overall, the spatial characteristics of both monitors could be considered as excellent in every aspects.

4.6.2 Colorimetric characteristics

As for geometrical properties, monitors allow users to adjust a wide range of colorimetric properties, which will now be introduced. *"The brightness control is used to adjust the video signal's zero-level to equal black on the screen"* (Nokia, 2005). Properly adjusting this setting is therefore crucial to obtain a true black on screen. A chart containing several patches with different digital counts, from 1% (of the digital count range) to 9%, under a totally black background is used in this case. The brightness controls, which corresponds to the *offset* parameter defined in the GOG characterisation model (Berns, 1996) for monitors, is first increased until every patch become clearly visible, and finally decrease until the 1% to 3% patches disappear. Another important parameter is the contrast, which determines the amplification of the video signal, and corresponds to the *gain* parameter in the GOG model. Whereas only a single *offset* can be specified for all three channels, the *gain* on the other hand can be adjusted independently for all three channels. That allows specifying accurately the colour temperature, i.e. the white point's colorimetric attributes, of the screen. As previously mentioned, the goal was here to reproduce very accurately the white point of the hardcopy substrate when illuminated and measured under real experimental conditions, while maximising the luminance level emitted by both monitors.

a) Internal and ambient flare

"Because of the glass faceplate on all CRT displays, inter-reflection flare can occur between neighbouring pixels; this can cause a discrepancy between the measured and emitted flux at a given pixel location.[...] *The internal flare caused by the neutral background was measured by setting the central stimulus to 0* [digital counts] *and the background to both 0* [digital counts] *and the* [digital counts] *necessary to yield a relative luminance factor of 0.2"* (CIE, 1996). The luminance level of the black patch in both cases were low enough for the TSR to display a warning indicating that an insufficient amount of light was used to perform the measurements. Since the corresponding colour difference was less than 0.3 ΔE^*_{ab}, the internal flare parameter was not included in the subsequently built characterisation model.

Similarly, "[W]*hen a display is situated in an environment with ambient illumination, the flux reflecting from the faceplate to the observer also must be taken into account*" (CIE,

1996). Although the experimental room was totally darkened, the measurements were made under true experimental conditions, i.e. with the daylight simulators switched on on the side opposite to the monitor under examination. The measured luminance level on the monitor's faceplate was less than 0.1 $cd.m^{-2}$. The effect of the ambient flare was therefore also ignored during the implementation of the characterisation model.

b) Channel independence

"An important property of a display [...] *is that the output of one electron gun and therefore the exitance of one channel at a given pixel is not dependent on the other channels at that pixel"* (CIE, 1996). The influence of one channel on the output of the others can be easily tested by measuring the peak tristimulus values of each channel and compare the resulting sum with the tristimulus values of the peak white. The results, shown in Table 11, indicates the excellent channel independence of both monitors.

Table 11 - Lacie monitors: channel independence

monitor	ΔE^*_{ab}
212	0.45
214	0.57

c) Spatial independence

"When characterizing a portion of the display, it is assumed that the luminance and chromaticity at that portion is stable and invariant to colours displayed in other portions of the display" (CIE, 1996). The evaluation of this effect consists in comparing the tristimulus values of the peak white displayed in the center of the screen when the background is changed from total black to peak white. The spatial dependence of both monitors can be considered as satisfactory, as can be seen from Table 12.

Table 12 - Lacie monitors: spatial independence

monitor	ΔE^*_{ab}
212	1.34
214	1.39

d) Uniformity

The uniformity of both displays was assessed by measuring at twelve different locations the colorimetric attributes of a peak white image covering the entire screen. The level achieved by the two monitors can be considered very satisfactory, as can be seen from Figure 25 and Figure 26. The average ΔE^*_{ab} at the diverse locations were around 1.5 units in each case. The only discrepancy originated from the *214* monitor, as its upper left corner exhibited a significant lack of uniformity from the rest of the screen. Fortunately, this zone was masked from observers' view by the viewing apparatus. On the other hand, the central area, where stimuli were displayed during the experiments, achieved an excellent overall uniformity, since the average ΔE^*_{ab} colour difference in this case was approximately 1.1 units. The luminance level was also very well distributed over the whole screen, as all the diverse measurement locations were all well within 5% of the luminance at the centre as recommended by ISO 12646 (ISO, 1999). A visual inspection of both monitors confirmed the empirical measurements, i.e. the uniformity of both monitors was excellent.

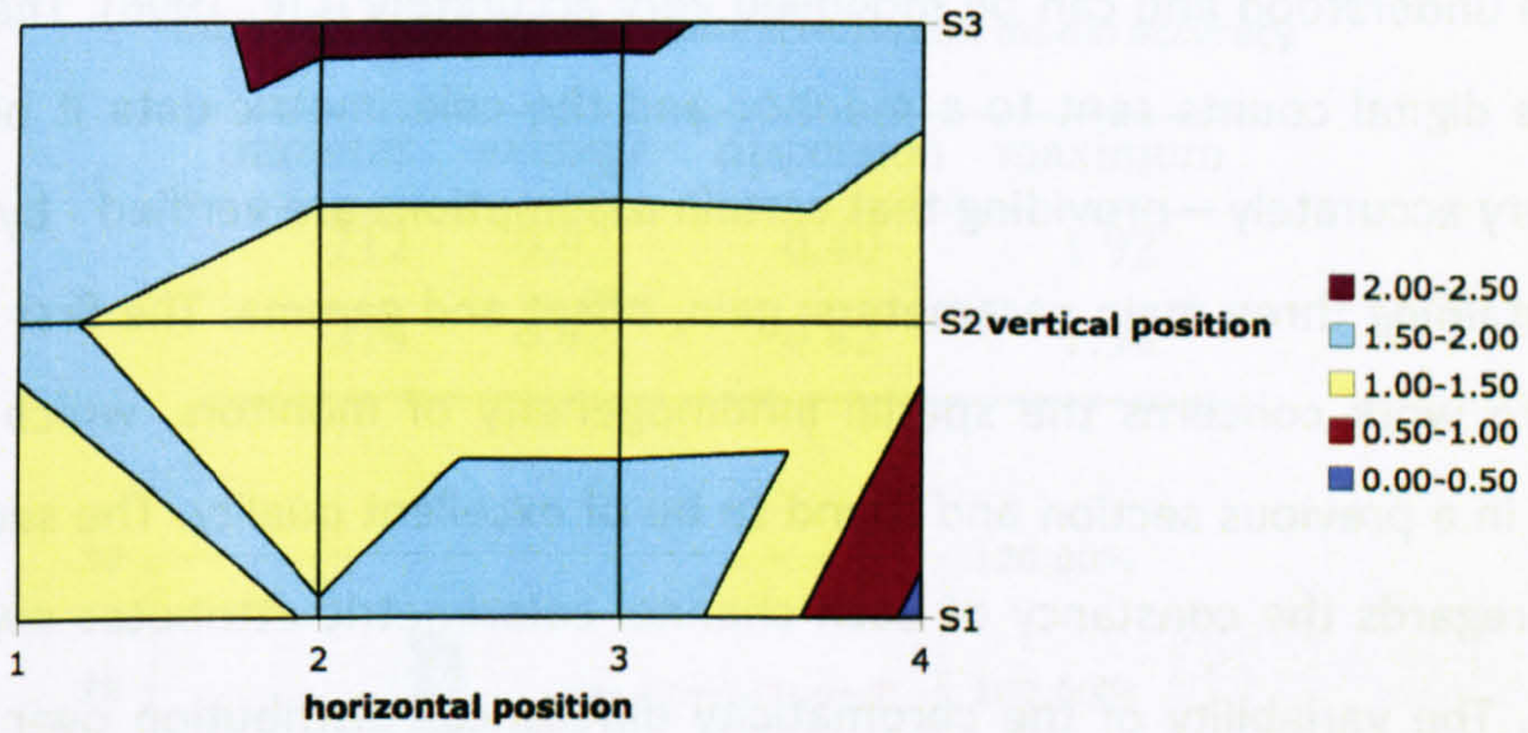


Figure 25 - Lacie 212: spatial uniformity

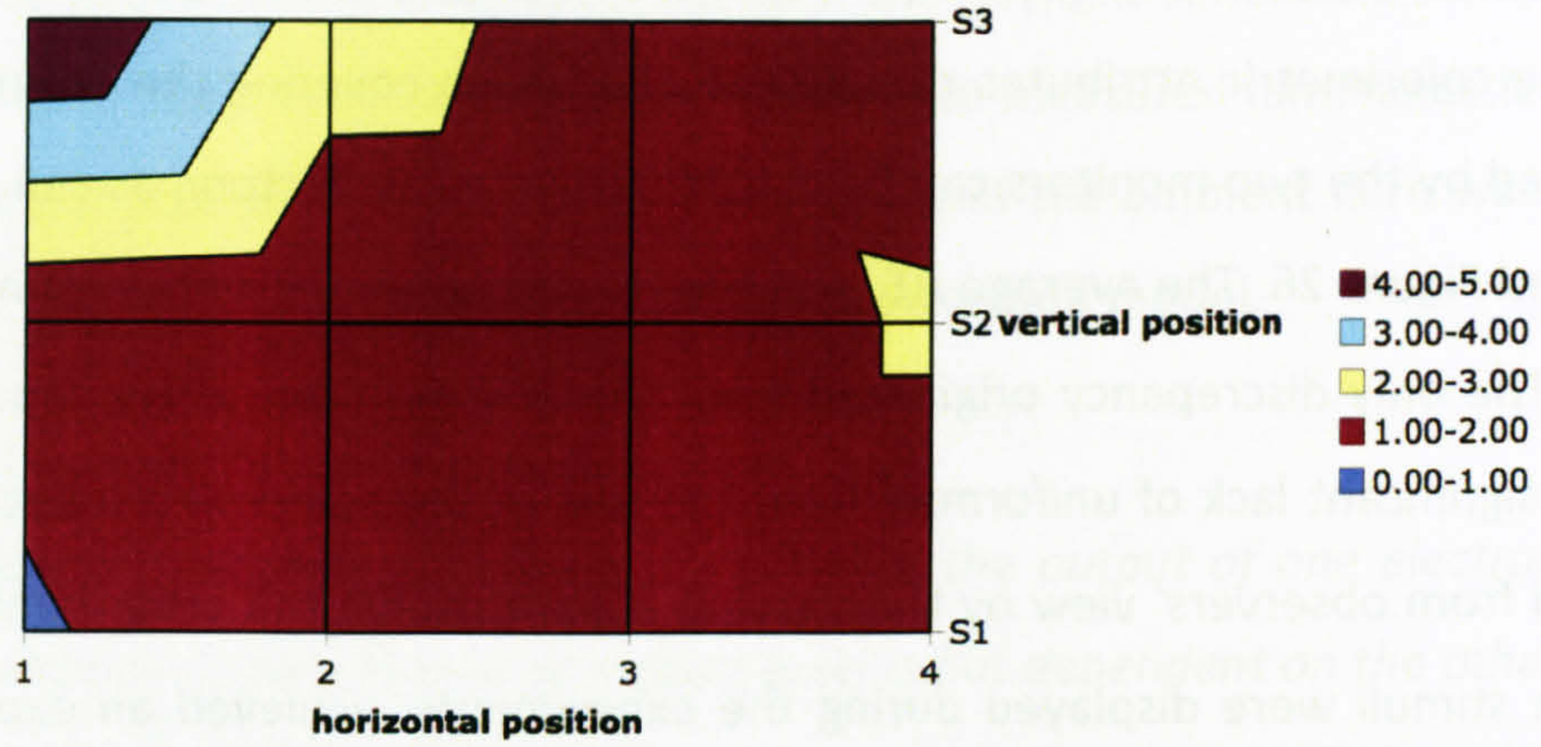


Figure 26 - Lacie 214: spatial uniformity

e) Characterisation

Whereas the complexity of the physical process associated with printing limits the printer's accuracy in terms of colour reproduction, the process underlying CRT monitors is simpler, well understood and can be modelled very accurately (CIE, 1996). The relationship between the digital counts sent to a monitor and the colorimetric data it output can be described very accurately – providing that certain assumptions are verified - by a power-law function containing three main parameters: gain, offset and gamma. The first condition for this model to work concerns the spatial inhomogeneity of monitors, which was already investigated in a previous section and found to be of excellent quality. The second and last assumption regards the constancy of each channel colorimetric attributes over the digital count range. The variability of the chromaticity differences distribution over this range is illustrated in Table 13. The variation is very small, confirming the last assumption required to assert the validity of the model.

Table 13 - Lacie monitors: constancy of channel chromaticities

monitor	Δx	Δy
212	0.0062	0.0066
214	0.0093	0.0066

As previously mentioned, the internal and ambient flare did not have any effect on the monitors under the designed experimental conditions, and the model was further simplified by removing them from the equation. Deriving its parameters can be then performed by simply measuring the peak colorimetric attributes of each channel and also the tristimulus values of fourteen neutral colours sampling the achromatic axis uniformly. Two choices are available to define the non-linear part of the relationship between colorimetric data and digital counts. Either an optimisation procedure can be performed to estimate the best-fitting gain, offset and gamma parameters, or a simple piece-wise linear interpolation can be used. Both options were tested by estimating the differences between the tristimulus values predicted by the model with measured data for 64 random colours sampling the device-dependent space approximately uniformly. The latter model gave noticeably better results, which are summarised in the following table. The histogram of the ΔE^*_{ab} colour differences are displayed for both monitors respectively in Figure 27 and Figure 28. Overall, the precision achieved by both characterisation models can be considered as very good.

Table 14 - Lacie monitors: characterisation model accuracy

monitor	average	dispersion	maximum
212	0.92	0.40	1.92
214	0.92	0.42	1.94

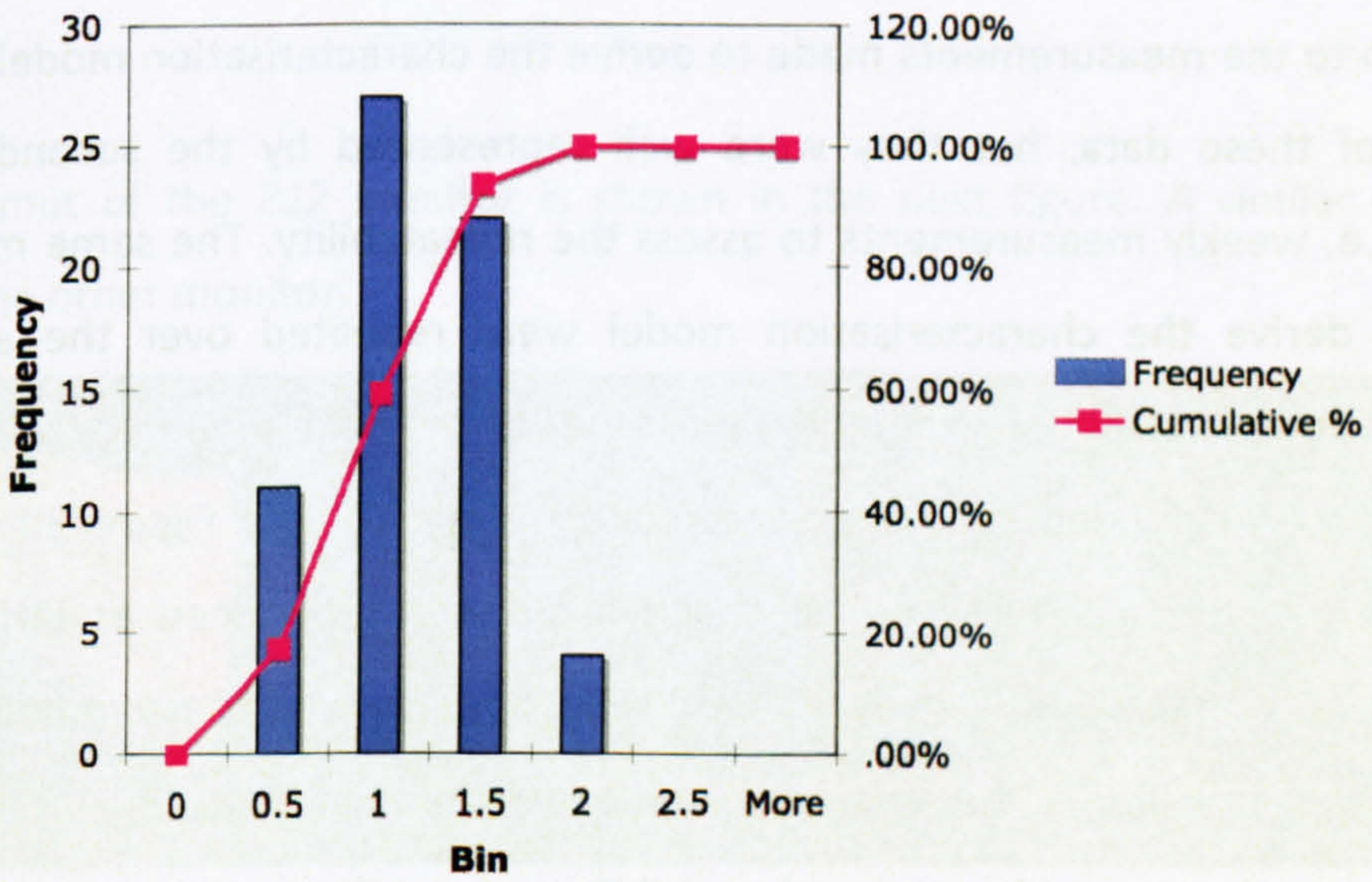


Figure 27 - Lacie 212: histogram of the colour differences population obtained by the characterisation model

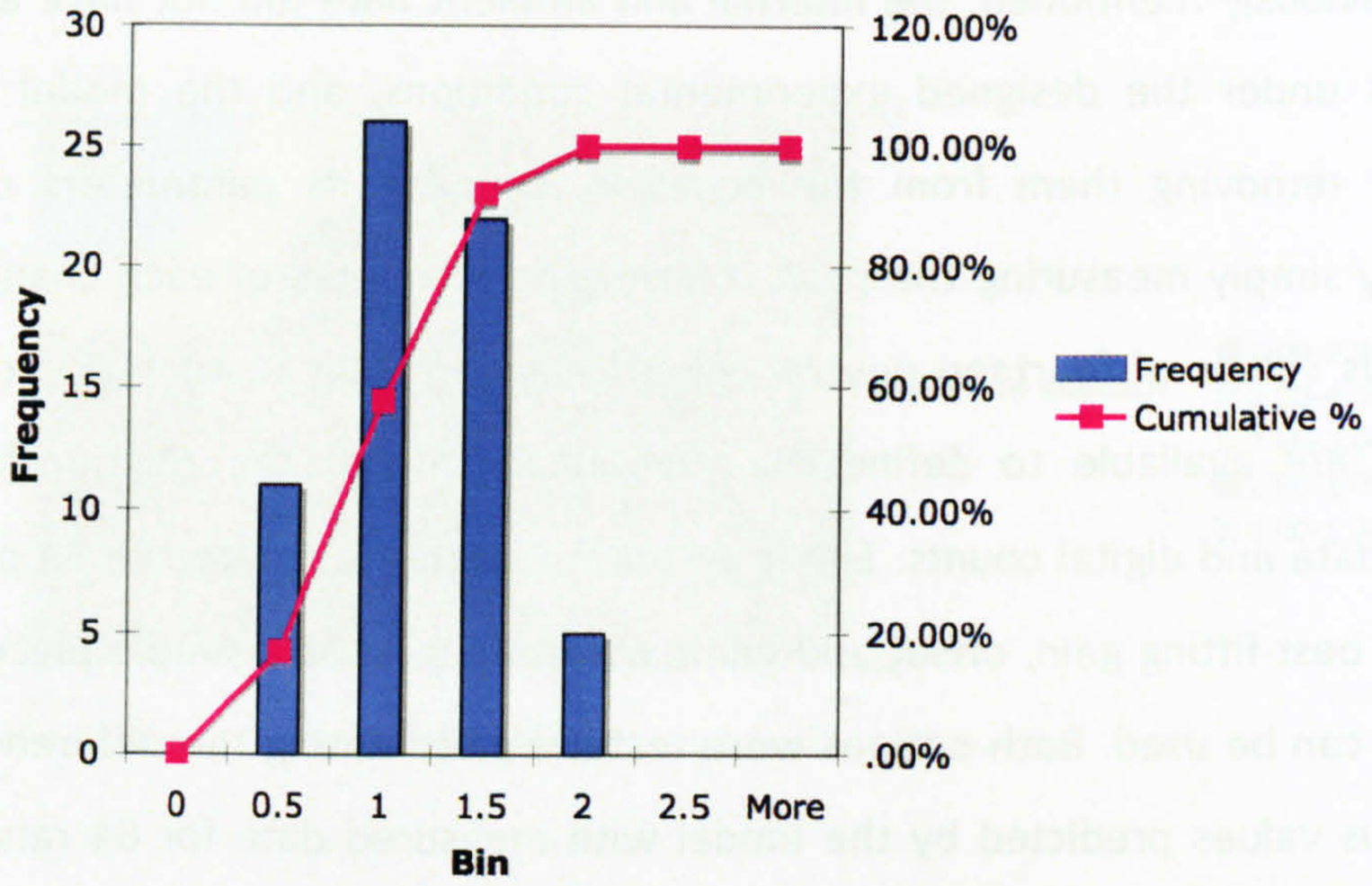


Figure 28 - Lacie 212: histogram of the colour differences population obtained by the characterisation model

f) Repeatability and Calibration

Unlike the printing system, where the repeatability and the temporal stability are actually distinct since they affect different components of the system, they correspond to the same issue for monitors. They are furthermore intimately linked with the calibration procedure, as the quality of the repeatability level will determine whether the calibration states, and thus the characterisation model, are still valid. They were both evaluated as a single entity while performing the experiments rather than at any another arbitrary time. The procedure was divided into two parts. Firstly, the white points of both monitors were evaluated daily prior to performing any experiments to check whether they corresponded *grosso modo* to the measurements made to derive the characterisation models. No records were kept of these data, but they were well represented by the second part of the procedure, i.e. weekly measurements to assess the repeatability. The same measurements required to derive the characterisation model were repeated over the eleven weeks necessary to complete all the experiments in order to evaluate whether the state of both monitors still corresponded to the original state. Figure 29 and Figure 30 illustrate the evolution of the colour differences for the measurements obtained at different weeks. Although the ΔE^*_{ab} magnitudes slowly increase with time, the worst case remains well within the range of acceptability, as the average error is less than half one ΔE^*_{ab} unit in that case. The monitors' settings were thus never changed throughout the experiments.

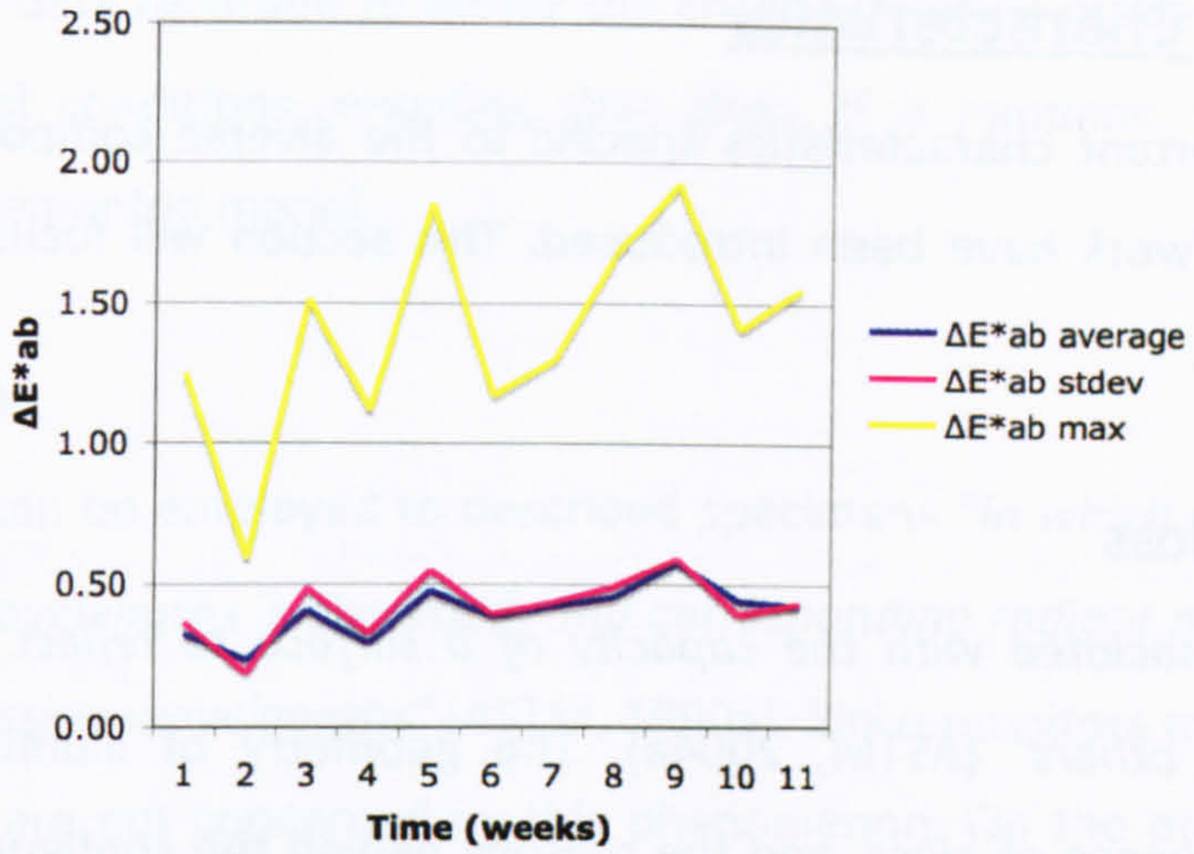


Figure 29 - Lacie 212: repeatability

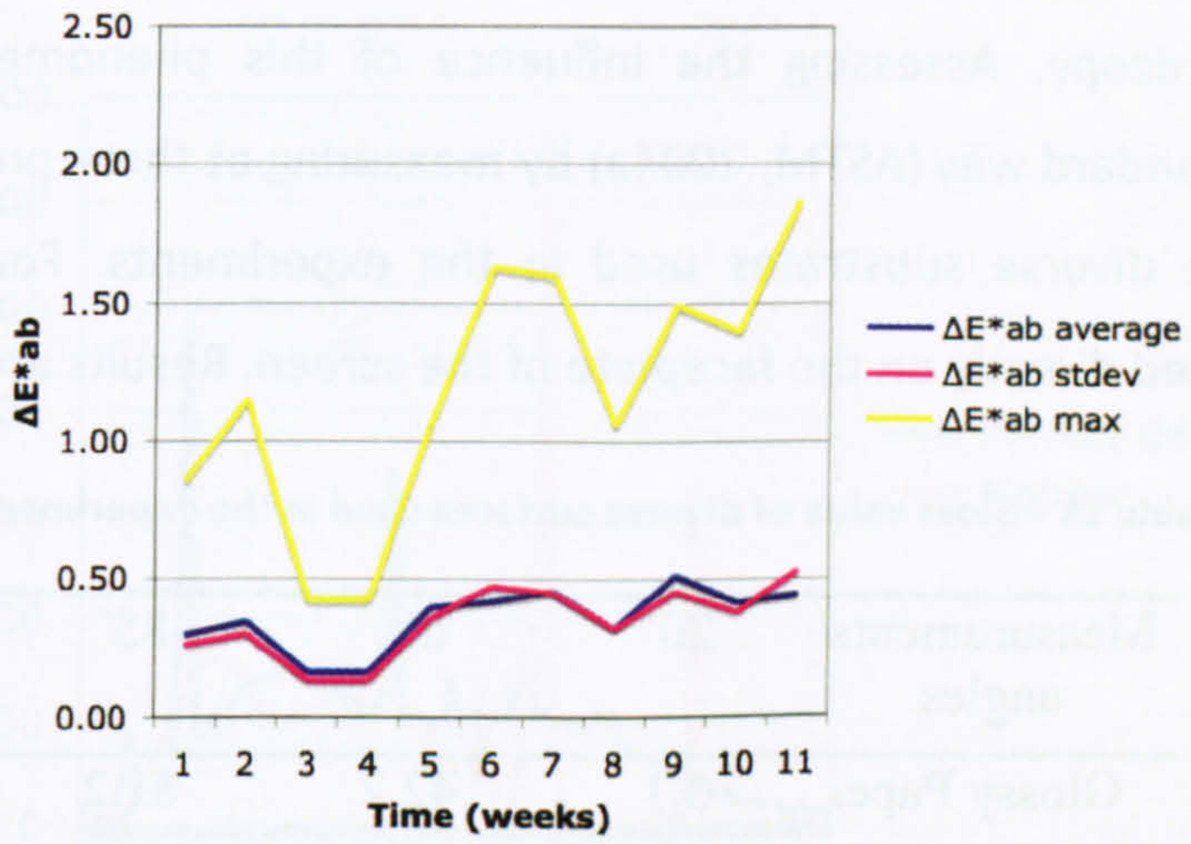


Figure 30 - Lacie 214: repeatability

g) Gamut

The gamut of the *212* monitor is shown in the next figure. A similar gamut was obtained for the other monitor.

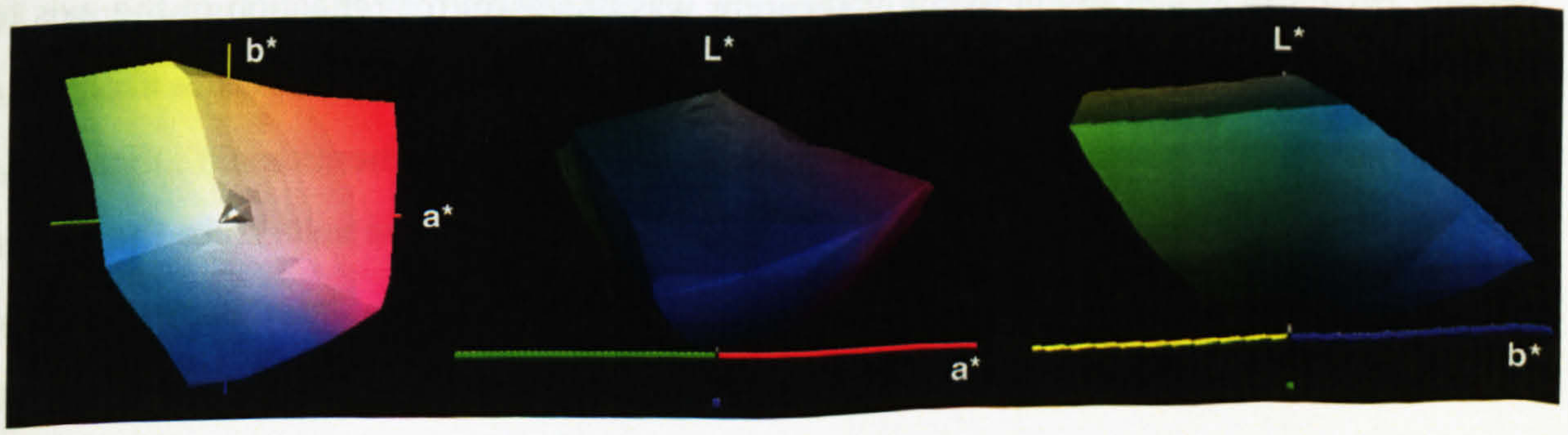


Figure 31 - Lacie 212: Gamut

4.7 Common Characteristics

All the important characteristics specific to the diverse components of the colour reproduction framework have been introduced. This section will focus on properties that affect both systems.

4.7.1 Specular Gloss

"Gloss is associated with the capacity of a surface to reflect more light in some directions than in others" (ASTM, 2004a). The geometry of illumination and viewing determines the influence of gloss, and the reasons behind the configuration adopted were to minimise gloss as much as possible. The lack of any illumination shining on the faceplate of monitors implies that they will not be concerned with this effect, but it will certainly occur in the case of hardcopy. Assessing the influence of this phenomenon was therefore performed in the standard way (ASTM, 2004a) by measuring at three pre-defined angles the gloss values of the diverse substrates used in the experiments. For the monitors, the glossmeter was placed directly on the faceplate of the screen. Results are shown in Table 15.

Table 15 - Gloss value of diverse surfaces used in the experiments

Measurements angles	20°	60°	85°
Glossy Paper	5.1	42.2	81.2
Matt Paper	1.5	4.4	9.9
Lacie Monitor	6.3	21.5	43.8

It is important to note that the measurements geometries mentioned are different from the CIE Standard geometry of illumination and viewing *45/0* used in this study. In the present case, the axis of the glossmeter receptor was at the mirror reflection of the axis of the incident beam for each measurement configuration (ASTM, 2004a). The more likely type of gloss that might occur would thus correspond to the 20° angle, rather than 60°. It could be thought rightly the matt paper would be a better choice, as its gloss values are the lowest ones. The adopted viewing geometry however ensures that the interference on the appearance of stimuli due to gloss was minimized. Furthermore, since glossy types of paper significantly increase the reproducible colour gamut in comparison to matt ones, the choice

was appropriate. The data recorded to derive the characterisation model were gathered in the real experimental conditions, entailing that gloss, if it happens, was automatically integrated in the implemented model.

4.7.2 Fluorescence

Fluorescence can be employed to described specimens "*in which radiant power in a range of excitation wavelengths is absorbed, and corresponding radiant power is emitted in a range of longer emission wavelengths*" (ASTM, 1990a). Since monitors are strictly emitting kind of devices, they are not concerned by this phenomenon. On the other hand, it could well happen for hardcopy. Figure 32 illustrates the spectral power distribution of the daylight simulators measured by employing a white tile approximating a perfect diffuser.

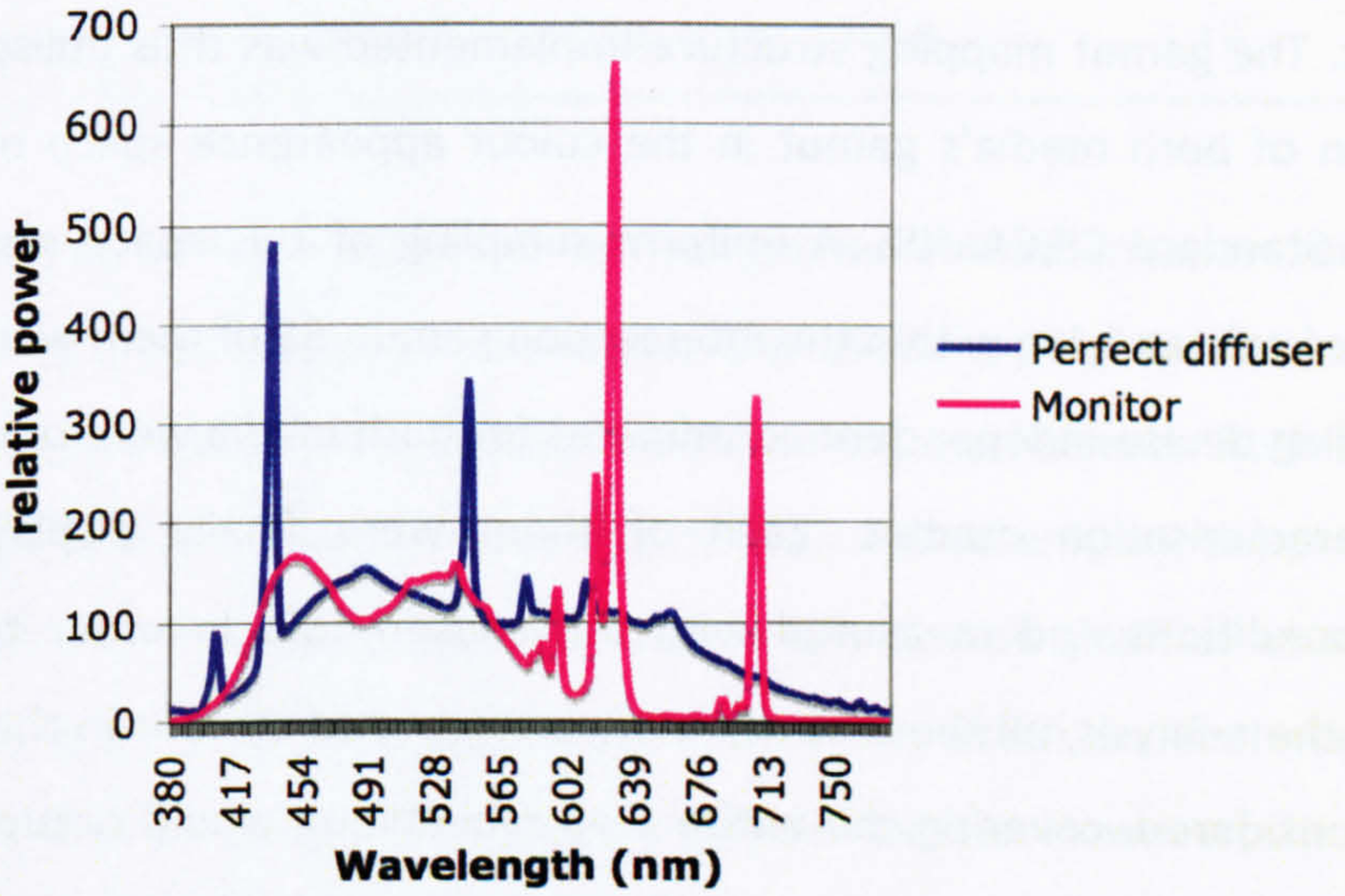


Figure 32 - Spectral power distribution of different whites

It can be seen that the amount of power emitted by the daylight simulators below 380 nanometers seem to be virtually nil, indicating that fluorescence caused by lower wavelength absorbance has little chance of occurring. If any fluorescence occurs with longer wavelength, the manner with which the data used to derive the colour reproduction framework were collected ensured that this phenomenon, may it happen, will be included in the overall model.

4.8 Framework

Evaluating the general performance of each media on their own is of little use in this case. What is crucial for this study is to evaluate the overall degree of accuracy achieved by the whole reproduction framework under the real experimental conditions. The average reproduction error, although usually providing the most meaningful criterion to judge the accuracy of a reproduction has its importance lessened in this case, since the most noticeable appearance differences will correspond to maximum reproduction errors. Both criteria therefore have to be scrutinised with equal care to provide a pertinent overview of the colour precision attained.

The overall accuracy of the reproduction framework can only be meaningfully estimated for the colours whose appearance attributes lay within the intersection of both devices' gamut. The gamut mapping structure implemented was thus utilised to determine the intersection of both media's gamut in the colour appearance space employed in this study, the CIE Standard CIECAM02. A uniform sampling of this space was performed to establish a set of colours lying within the intersection gamut. 59 of them were recorded, and the corresponding device-independent coordinates for both media were obtained with their respective characterisation models. Each of them were finally displayed under real experimental conditions and measured with the chosen TSR. In order to maximise the significance of the analysis, all the test colours were grouped by four in the central area of the medium considered, covering the entire area real stimuli would occupy. Furthermore, each colour was measured three times at different locations within the area they covered. This strategy was designed to take into account the spatial uniformity parameter into the investigation, as the previous analysis showed that each medium exhibited different levels of accuracy. Moreover, the intrinsic variability of the hardcopy illumination was also taken into account, as the tubes were switched on and off for periods of two minutes, exactly like during the experiments, and the three measurements were made at random times during the *on* intervals.

Two different types of data were recorded to estimate the overall precision. The tristimulus values were first computed from the spectral data using the method previously described. In order to check the validity of the calculation, the tristimulus values given by the

TSR were also gathered. A comparison between the two types of tristimulus values revealed that they were linearly related. The only differences between them were a relatively small difference in magnitude of less than 5%. The following table present the differences obtained from the colorimetric data directly read from the single SP used, as this represent the most meaningful testing method.

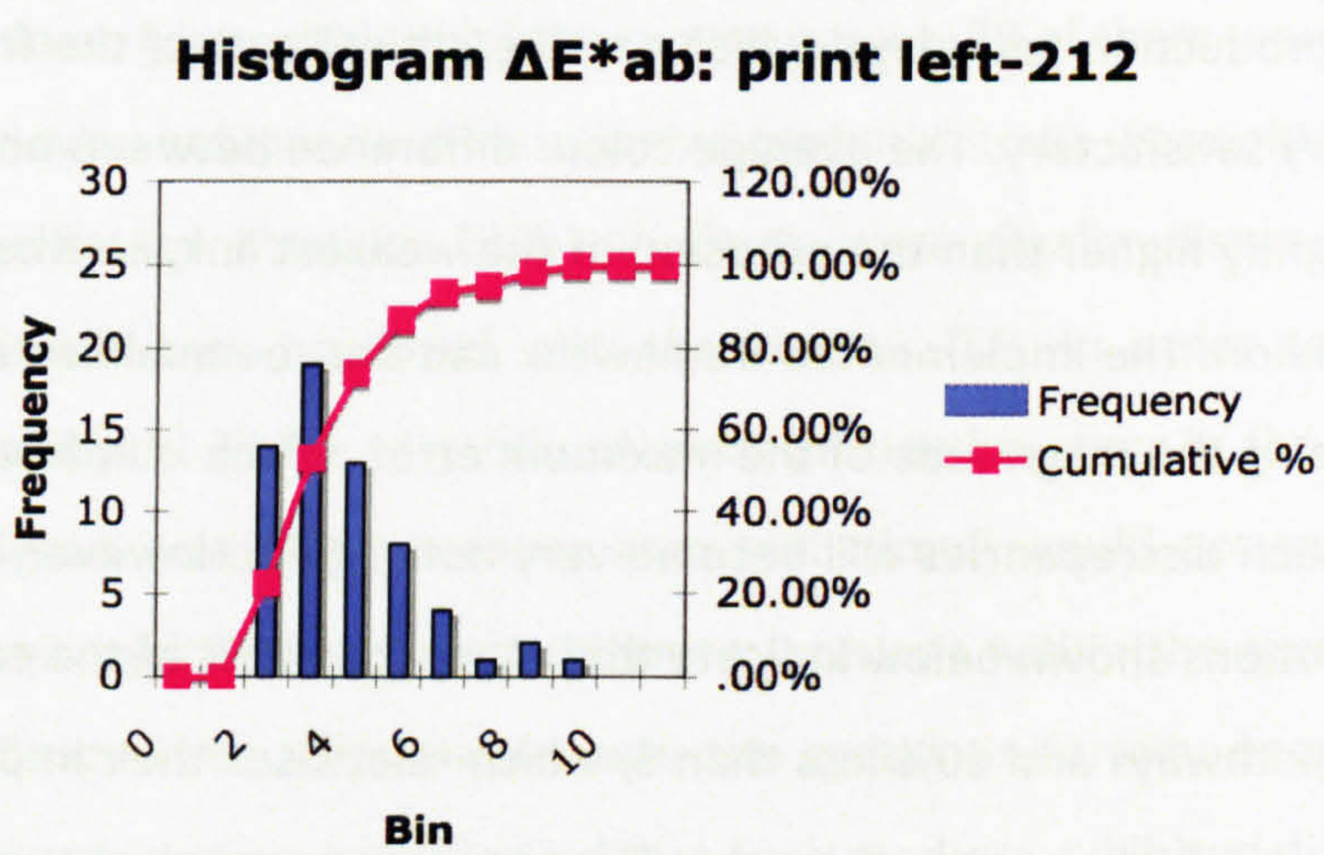
Table 16 - Whole framework - ΔE*ab error statistics of the diverse reproduction pathways

ΔE^*_{ab}	average	variability	maximum
Print left – 212	3.28	1.61	8.95
Print right – 214	2.78	1.68	9.48
212 - 214	0.79	0.19	1.40
Print right – Print left	1.61	0.69	3.16

Given the performance obtained by the printer characterisation model, the two print-monitor reproduction pathways, which are the critical part of the framework, can be considered as very satisfactory. The average colour difference between both sets of devices is indeed only slightly higher than the precision of the weakest link, i.e. the printer, and so is there their dispersion. The implemented framework can thus be qualified as acceptable, the only drawback being the magnitude of the maximum error, which is unfortunately increased to levels where such discrepancies will become very noticeable. However, the histograms of the errors distributions shown below indicate that more than 95% of the errors are less than 6 ΔE^*_{ab} for both pathways and 60% less than 3, which relativises their impact. The precision of the colour reproduction achieved by the monitors is unsurprisingly excellent. It is even lower than the accuracy achieved by their respective characterisation models, which is excellent and indicate that these models are actually partly responsible for the magnitude of the colour differences errors. A more unexpected result emerges from the print-to-print pathway, as the quality of the match between both sides of the apparatus outclass very significantly the performance of the printer. As previously mentioned, the analysis of the printer repeatability showed similar results, but that was certainly due to the automatic method of measurements used, as the same exact area were repeatedly assessed. However, the investigation was designed this time to take into account every possible source of

variability. Given the lack of uniformity of the printer, and also to a lesser extent of the illumination, but especially because of its temporal instability, it could be very reasonable to expect worse results. Obtaining a maximum error lying exactly at the discriminability threshold for complex images is thus an excellent achievement.

Overall, the precision achieved can be considered satisfactory. The results are indeed not significantly worse than the intrinsic accuracy of the biggest limiting factor, namely the printer. The maximum reproduction error may be a source of worry, because its magnitude implies that observers will definitely notice such an appearance discrepancy. However, most errors are located around the 3 ΔE^*_{ab} boundary, usually considered as the just noticeable difference observable for pictorial images. However, given the current reproduction abilities of typical printers, it would be illusory to hope implementing a reproduction framework where all errors would lie within the 3 ΔE^*_{ab} limit. The quality of the built framework can therefore be considered as adequate.



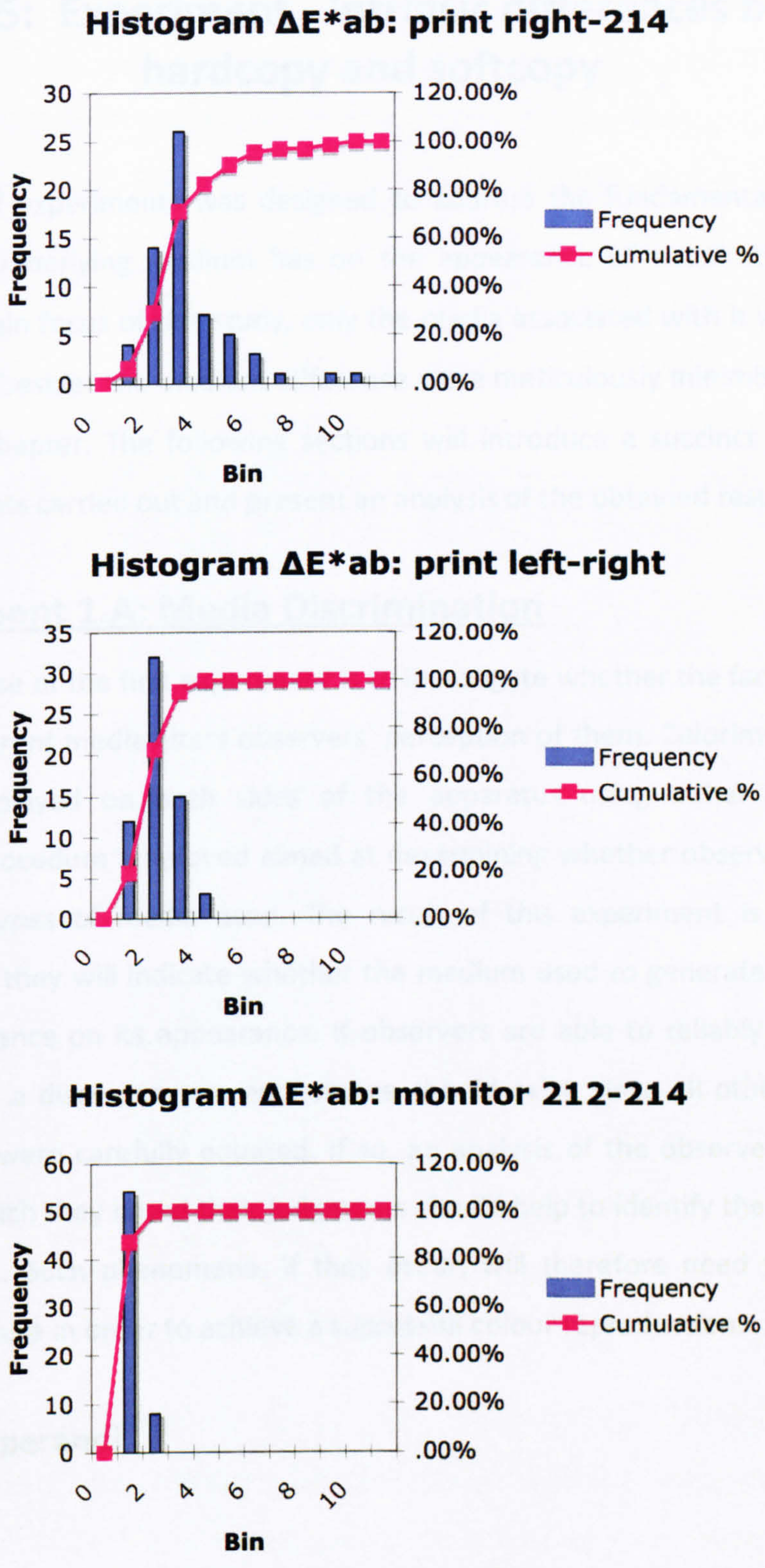


Figure 33 - Whole framework: histogram of the ΔE*ab errors distribution of the diverse reproduction pathways

Chapter 5: Experiment - Intrinsic differences between hardcopy and softcopy

A series of experiments was designed to address the fundamental question as to what effect the underlying medium has on the appearance of visual stimuli. Since soft proofing is the main focus of this study, only the media associated with it were considered. Any other factors besides the medium difference were meticulously minimised, as described in the previous chapter. The following sections will introduce a succinct overview of the diverse experiments carried out and present an analysis of the obtained results.

5.1 Experiment 1.A: Media Discrimination

The purpose of the first experiment is to investigate whether the fact that stimuli are displayed on different media alters observers' perception of them. Colorimetrically identical stimuli were displayed on both sides of the apparatus using either media, and the psychophysical procedure employed aimed at determining whether observers were able to distinguish the types of media used. The result of this experiment is of fundamental importance since they will indicate whether the medium used to generate a stimulus has a quantifiable influence on its appearance. If observers are able to reliably discriminate the media employed, a discrepancy in appearance should exist, since all other aspects of the external context were carefully equated. If so, an analysis of the observers' comments on the factors on which they based their judgments should help to identify the possible roots of these differences. Such phenomena, if they occur, will therefore need to be taken into account in the future in order to achieve a successful colour reproduction.

5.1.1 Modus Operandi

5.1.1.1 Stimuli

Twelve images covering a wide range of image content and colorimetric characteristics were selected for this study (Morovic & Wang, 2003). The two main classes of stimuli, i.e. pictorial images and business graphics, were almost equally represented, as this set of experiments aims at encompassing the widest range of applications possible. The sampling was performed so as to incorporate as many image types as those identified by the

experimental guideline published by the CIE TC 8-03 (CIE, 2004). The colour gamut of both media being very different, the gamuts of all experimental images had to be modified in order to make the comparison meaningful. The characterisation models of both media were used to generate their gamuts, from which the intersection gamut was derived by interpolation. The gamut mapping algorithms implemented for the other experiments were then employed to restrict the gamuts of all experimental images to the intersection of both media's gamuts.

Figure 34 - Experimental stimuli

5.1.1.2 Viewing conditions

The viewing apparatus previously introduced was used to control very accurately the viewing conditions.

5.1.1.3 Observers

Seven women and six men aged between 21 and 34 volunteered to participate in the experiment. They all completed the full Ishihara Test for Color Blindness successfully (Ishihara, 1980). They were asked to use the provided chin rest throughout the experiment,

and their positions were first calibrated to ensure that it complies with the overall geometrical design. Written instructions were finally given to them, and the experiment started once they signified they understand them. Since the process of changing the stimuli could indicate the type of medium, small opaque black curtains were drawn in front of each aperture before and after any trial and observers were also kindly asked to listen to some music. Changing both stimuli between trials could take up to 30 seconds. Once the stimuli were ready, the curtains were opened and the headphones removed, and observers were prompted to look at the stimuli displayed simultaneously through both apertures. On any trial, each of the following four possible combinations had an equal probability of occurring: (1) both stimuli were generated by two monitors; (2) both stimuli were generated by two prints; (3) the left stimuli was generated by a monitor and the right by a print; and (4) the left stimuli was generated by a print and the right by a monitor. The order of presentation was randomised between each session, and the four possible combinations for each image were never presented in a row. Overall, each observer participated in three sessions, each session lasting for about 40 minutes.

Their first task consisted of *discriminating* the types of media employed, i.e. reporting whether both stimuli were generated using the same medium or different ones. After a response has been made, observers were prompted to rate their confidence in their answer on a five-point rating scale. Category *-2* corresponded to "not very confident" and Category *+2* corresponded to "very confident". Their second task was to *identify* the type of medium used to generate both stimuli, and to state if possible the reasons of their choice. This experiment was repeated three times in order to investigate observers' repeatability. The instructions given to observers were the following:

"In front of you stands a screen containing two apertures. Pairs of images will be presented at the same time through those apertures, one through each. Each image can be generated by two different media: a monitor (softcopy), or an illuminated print (hardcopy). Thus, at one time, you can be shown either two softcopies, or two hardcopies, or one hardcopy and one softcopy. Your first task is to report whether the images were generated by the same medium or not. To do so, please rate the confidence you have in your judgement on a scale where:

2 same : you are sure the media are the same.

1 same : you are fairly sure the media are the same.

0 : you are not sure at all.

1 different : you are fairly sure the media are different.

2 different : you are sure the media are different.

Please also state the type of medium (monitor, print) used to generate the left and the right hand side stimuli. You might also be asked the reasons of your choices.

This is not a test. We are seeking your opinion."

5.1.2 Psychophysical Methods

"Detection Theory is a general psychophysical approach to measuring performance" (Gescheider, 1997). Central to this theory are two values derived from experimental data, denoted *hit rate (H)* and *false alarm rate (F)*, whose exact definitions are dependent on the design of the experiment at hand. Classically, the measurement of performance, such as *discriminability*, is most often assessed using the proportion of correct responses, denoted *p(c)*, obtained from linearly combining *H* and *F*. However, it is actually rarely adequate for this kind of tasks, as it is very prone to observers' biases, i.e. their willingness to use one answer rather than the other. Furthermore, a given proportion correct attained, even by an unbiased observer, has a different meaning, in the sense that it represents a different level of performance, depending on the task assigned to observers.

A more appropriate performance index is the detection-theory metric *d'* (Macmillan & Creelman, 2004). It has the desirable property of representing a uniform level of performance for any task, and is independent of bias. In fact, the theory also provides an index of bias, denoted *c*. Several methods are available for carrying out the analysis depending on the tasks performed by observers, and the natural paradigm for studying observers' first task, i.e. discrimination, is the *same-different* design. This technique is particularly well adapted for this kind of tasks, as it does not require neither from the

observer nor the experimenter to be able to articulate or even be aware in which way the stimuli actually differ. Observers can use whatever cues are useful in discriminating between the two stimuli, and the objective is not to determine which ones were used but simply to establish the differentiability of the stimuli.

The second part of this experiment also involves discrimination, but in a slightly different manner. The first task require observers to make binary decisions – *same* or *different* – from the presentation of four possible types of stimuli combinations, which can be designated as a *classification* task. On the other hand, the paradigm of the second task is much more basic, as observers are asked to identify the type of medium, which corresponds to a *recognition* design. The procedure involved belongs to the *yes-no* type of design, although the actual responses are different from yes or no. *d'* can however be used in both cases to represent the level of performance attained by observers. Other interesting metrics are also provided by the detection theory and will be introduced in due course.

5.1.3 Results

5.1.3.1 Media discrimination

In order to carry out the detection theory analysis, we defined *hits* as those trials on which an observer judged the two stimuli as being generated by two different media when in fact they were, and *false alarms* as those trials on which an observer judged the two stimuli as being generated by two different media when in fact they were *not*. The sensitivity metric *d'* can be simply obtained from the difference between the z-scores of the *hits* and *false alarms* rates. Theoretically, the range of *d'* is infinite, but a [*H*, *F*] pair of [0.99, 0.01] corresponds to a *d'* of 4.65, which is usually considered as an effective ceiling. The resulting value, displayed in Table 17, suggests that, on average, observers were able to discriminate reliably whether the two stimuli were generated by the same type of medium or not. This conclusion might seem obscure for people with no experience concerning the detection theory, and so proportion correct is also commonly employed to clarify an analysis performed from a detection theory point of view. *p(c)* can be more intuitively interpreted, as its range is defined in a much clearer way than *d'*, with 0.5 represents chance performance and 1.0 perfect discrimination. If derived from the *d'* metric, *p(c)* does not suffer from the disadvantages previously mentioned, which can be easily observable in Table 17 where very

different values are obtained depending on how the analysis is performed. The influence of observers' bias clearly affects the results. The detection theory bias metric *c* reveals the preponderance of "same" answers among observers' responses, indicating that they show some bias towards saying "same". The proportion correct $p(c)_{S\text{-}D}$ metric, if computed directly from the raw data, ignore this factor and thus indicate a diminished sensitivity compared to when proportion correct is derived from the *d'* metric, corresponding to the $p(c)_{S\text{-}D\ unbiased}$. The level of accuracy represented by $p(c)_{S\text{-}D\ unbiased}$ confirm the conclusion that observers could reliably distinguish whether the two stimuli were generated by the same medium or not.

Table 17 - Experiment 1 - Media discrimination: sensitivity and bias metrics

H	*F*	*d'*	$p(c)_{S\text{-}D}$	$p(c)_{S\text{-}D\ unbiased}$	*c*
0.69	0.22	2.43 ± 0.10	0.61 ± 0.02	0.74 ± 0.02	0.14±0.02

However, a more efficient way to eliminate observer bias is to obtain a full receiver-operator characteristics (ROC) curve, where different hit rates are plotted versus their corresponding false alarm rates. This type of curves summarises very well the prediction of detection theory. All the [*H*, *F*] pairs lying on such a curve represent a unique level of sensitivity, but each point corresponds to a different bias and/or criterion. When performance is at chance (*d'*=0), the ROC is the major diagonal, where the hit and false alarms rates are equal. As the sensitivity increases, the locus of the curves shifts toward the upper left corner. Examples of such curves are given in Figure 35. The advantage of producing these curves is that the characteristics of the underlying process can be estimated unambiguously from its path, whereas a single [*H, F*] rate pair is not sufficient. The most efficient way to collect enough data is the confidence rating procedure. It is assumed that each ratings correspond to a different criteria adopted by observers to judge the stimuli. A complete ROC curve can thus be inferred from the [*H, F*] pairs pooled over subjects and corresponding to different criteria.

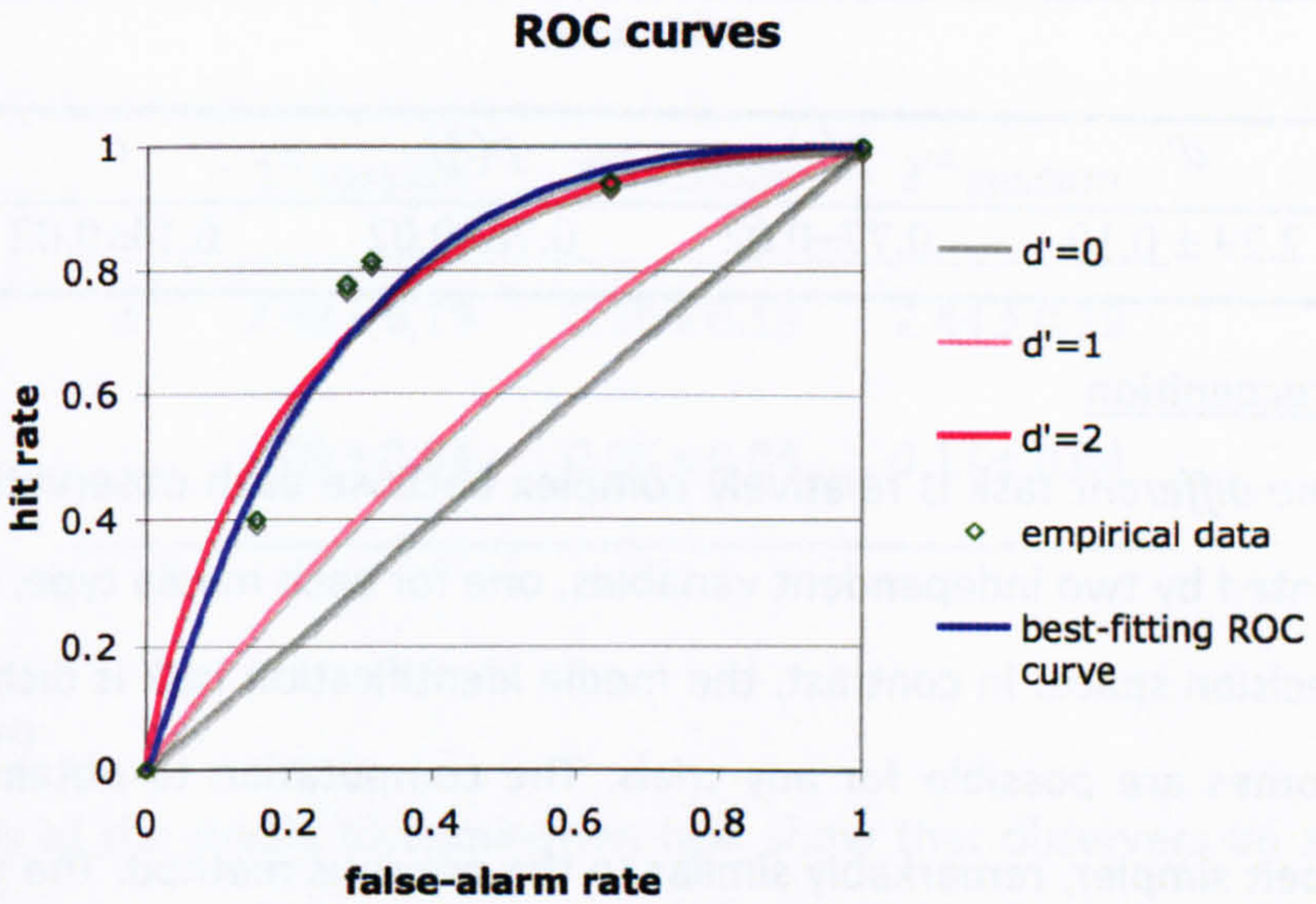


Figure 35: Experiment 1 - Media discrimination: ROC curves

The parameters of the best-fitting ROC curve were estimated from the collected [*H*, *F*] pairs using a jackknifing technique involving maximum-likelihood estimation, since simply averaging the data for each observer can yield misleading results. The ROC curve obtained for the discrimination task is shown alongside the examples in Figure 35. Since there is no goodness-of-fit statistics available for jackknifed *same-different* ROC curves (Francis & Irwin, 1995), the evaluation has to be made by eye, and it can be concluded that the obtained curve provide a reasonable fit to the data. However, it can observed that the shape of the ROC curve is not symmetrical with regards to the negative diagonal on the hits vs. false alarms diagram. The reasons for this behaviour are related to the strategy adopted by observers to perform the task, which will be analysed in more details later. The sensitivity estimates have to be revised to take into account this factor, and the best-fitting parameters obtained from the jackknifed ROC curve are shown in Table 18, where the only differences with the previous table are the slight increase in magnitude of *p(c)*. Although *d'* is not affected by this phenomenon, proportion correct cannot depict such behaviour, and its magnitude might not represent accurately the true performance achieved by observers. Another measure of discriminability, *p(A)*, which represents proportion of the area of the entire [*H*, *F*] graph that lies beneath the ROC curve, is more appropriate in such cases. *p(A)* has furthermore the double advantage of being a non-parametric measure and to be equivalent to *p(c)*. It therefore corresponds to the most reliable and intuitive way to represent the level of performance achieved by observers in this case.

Table 18: Experiment 1 - Media discrimination: sensitivity and bias metrics estimated from the best-fitting ROC curve

d'	*p(c)*	*p(A)*	*c*
2.29 ± 0.10	0.77±0.02	0.78±0.02	0.14±0.02

5.1.3.2 Media recognition

The *same-different* task is relatively complex because each observation of any trials can be represented by two independent variables, one for each media type, entailing a two-dimensional decision space. In contrast, the media identification task is dichotomous, since only two outcomes are possible for any trials. The computation to obtain the sensitivity metric *d'* is, albeit simpler, remarkably similar to the previous method. The *yes/no* paradign was employed to analyse the data. The performance measure in this case is much lower than for the discrimination task. The corresponding *p(c)* indicates that observers experienced difficulties to perform this task. Furthermore, one of the test images exhibited some spatial artifacts that clearly gave away the type of medium used. The data collected from this stimulus were therefore removed from the subsequent analysis. The results obtained from the ROC analysis are not presented here, as no additional conclusions emerge from them.

Table 19: Experiment 1 - Media recognition: sensitivity and bias metrics

d'	$p(c)_{Y\text{-}N\ unbiased}$	*c*
1.01 ± 0.10	0.57±0.02	-0.14±0.02

5.1.3.3 Observers' Repeatability

Table 20 displays the different performance metric *d'* obtained for the three different sessions. Although the accuracy metric obtained for the last session is significantly different from the two others, this result should not be interpreted as a lack of repeatability. Rather, it suggests that observers were in fact *learning* the specific task they were asked to perform, a phenomenon widely observed in the Signal Detection Theory field.

Table 20: Experiment 1 - Observers' repeatability: sensitivity and bias metrics for the three experimental sessions

	1st session	2nd session	3rd session
d'	2.08 ± 0.19	2.25 ± 0.19	2.44 ± 0.19
c	0.09 ± 0.04	0.05 ± 0.04	0.17 ± 0.04

5.1.4 Discussion

The results of the media discrimination task show that observers on average could perceive quite reliably whether the two stimuli were generated by either the same or different media. This was to be expected since the average precision of the colour reproduction framework built corresponded to the commonly accepted value for the just-noticeable-difference observable in pictorial images (Fairchild & Alfvin, 1995). Furthermore, the equivalent proportion correct value in this case was almost exactly halfway between chance performance and perfect accuracy, which correspond to the most common location where the threshold is set according to the classical threshold theory. It could therefore be argued that this provides an indirect way to verify the location of the just noticeable difference for pictorial images.

Despite the relative good accuracy achieved by observers, they often stressed the difficulties they experienced to perform this task. The differences between the two types of stimuli, although noticeable, were close enough to create confusion. The amount of bias they showed towards "same" answers, as quantified by the *c* metric, reflect well this tendency. Additionally, the obtained ROC curve, besides providing an accurate measure of sensitivity, also allows an inference to be made about the decision strategy that observers adopted. Its shape, more precisely its symmetry with regards to the negative diagonal on the hits vs. false alarms diagram, indicate that the strategy observers chose was not optimal, emphasizing the trouble they were experiencing. This is consistent with previously obtained results, where it was concluded that the multi-dimensionality of colour perception forbids the use of an optimal strategy (Logvinenko & Maloney, 2006).

Observers thus experienced difficulties to produce judgements for the first task, but they manage on average to discriminate quite reliably between the two classes of stimuli. After completing the three sessions, observers were prompted to explain in more details how they perform both tasks. The majority of them reported that a slight difference in chroma between the two media was the main cue helping them to discriminate between both classes of stimuli. Their answers corroborate the analysis performed so far in the sense that the overall precision of colour reproduction was not high enough to go unnoticed by observers. However, their performance in the second task was much worse, although the theoretical level of difficulty of this part is least. So, although they could perceive differences between the information generated by both types of medium, they could not utilise these differences to infer reliable information about the exact type of medium employed. The visual signal they received from both types of media had slightly different colorimetric characteristics, but was not significantly different depending on the medium used to generate it. It can be therefore concluded that, under experimental conditions similar to the present ones and subject to the device's precision, the type of medium used to generate a stimulus does not influence its appearance. In other words, a colorimetric match is an appearance match.

5.2 Experiment 1.B and 1.C: Accuracy and Pleasantness

Another important aspect determining the suitability of soft proofing is whether judgments made on the basis of hardcopy simulation are transferable to prints, and two experiments were designed and carried out to assess this issue. Judging the quality of a reproduction cannot be performed in absolute terms, but is always dependent on a specific application. Accuracy is a criterion that plays a major role in typical colour reproduction framework, and formed the basis for the first experiment. Observers were asked to judge how precisely a hardcopy reproduction (or its simulation) approximates, in terms of colour appearance, a softcopy original. However, final outputs should not only provide a faithful rendering of the original scene, but should also be pleasing since they will most probably be used as stand-alone. The second experiment investigated the overall pleasantness of the reproductions produced by each GMAs that were tested. Observers were presented with two reproductions generated by two different GMAs *on the same medium* and asked to choose their preferred one in terms of colour appearance. From a more conceptual point of

view, this preference experiment is fundamentally different from the previous one, although the set-up and procedures employed are very similar. It exclusively targeted "within-media" judgments, or more specifically how such decisions made on the basis of prints exclusively are transferable to soft proofs only, whereas the previous experiment investigated how "cross-media" judgments transfer to "within-media" ones. The combination of both should provide a clear overview of the suitability of soft proofs as a surrogate to hardcopies.

Each experiment had to be duplicated in order to collect decisions based both on hardcopy stimuli (termed 1st reproduction workflow) and on softcopy simulations of hardcopy stimuli (2nd reproduction workflow). Since GMAs are by nature destined for cross-media colour reproduction applications workflow, an assessment of their performance would provide an ideal way to compare the two studied reproduction workflows. However, it is important to understand that their intrinsic absolute performance is of no value in this specific case. Rather, the analysis should focus on the correlation of the results obtained by each workflow, instead of considering each case individually.

5.2.1 Modus Operandi

5.2.1.1 Stimuli

The twelve images chosen for the previous task were also employed for these two experiments. In the first one, the original stimuli, i.e. not gamut-mapped, were displayed on a monitor, and the reproductions obtained with the three different GMAs were shown to observers first on prints and then on displays. Given the colorimetric stability of the print, the same reproduction stimuli were used for the second experiment.

5.2.1.2 Gamut mapping algorithms

Three standard gamut-mapping algorithms were employed in this study. First of all, minΔE is the simplest of all, as it consists in replacing out-of-gamut colours with the closest in-gamut colours. The standard colour difference formula ΔE^*_{ab} is used to determine the closest in-gamut neighbourg. On the other hand, LLIN is an algorithm that preserves hue angles, but performs a linear compression for both lightness and chroma. Finally, GCUSP is the most elaborate of all and is best described by his author *"First the lightness of a colour is compressed linearly in a way which applies full compression only to achromatic colours and*

which alters the lightness of high-chroma colours to a lesser extent. [...] Next, colours are compressed using a spherical mapping towards the point on the lightness axis which has the same lightness as the cusp at a given angle" (Morovic & Luo, 1997).

5.2.1.3 Viewing conditions

Once again, the viewing conditions needed to be configured with extreme care in order to ensure the quality of the results. The physical set-up built for the previous experiment satisfied the present requirements adequately since it has the double advantage of hiding all cognitive cues and also to equate the surround conditions very accurately.

5.2.1.4 Observers

Sixteen and fourteen colour-normal observers participated in the two experiments respectively. As previously, they were all aged between 21 and 34 years old. The complicated process destined to disguise the type of medium employed to generate the diverse stimuli was of no use in this case, although observers were still asked to use the provided chin rest throughout the experiments. Written instructions were given to them, and the experiment started once they signified they understood them.

The observers' task in the first experiment consisted of judging the overall accuracy, in terms of colour appearance, of the reproduction stimuli using a scale of numbers ranging from category -3, which corresponded to "very inaccurate", to category +3, "very accurate". Their task in the second experiment simply consisted in indicating which of the two presented reproductions they preferred, once again focusing on the colour appearance attributes of the stimuli. Some preliminary sessions were performed with some observers in order to determine two stimuli that were systematically categorised respectively among the best and worst. Each observer was subsequently shown those two stimuli before each session in order to calibrate their answers. The order of presentation was randomised between each session, and the four possible combinations for each image were never presented in a row. Overall, each observer participated in three sessions, each session lasting for about 40 minutes. The instructions given to observers were the following:

"You will be shown pairs of images, one of them being a reproduction of the other. The location of the original and of its reproduction will be alterned between right and left.

We would like to know how accurate you think the reproduction is. Please express your opinion on a scale of numbers from -3 to 3 where:

-3 : very inaccurate

-2 : moderately inaccurate

-1 : slightly inaccurate

0 : average

1 : slightly accurate

2 : moderately accurate

3 : very accurate

Use numbers between -3 and 3 to represent equal intervals of accuracy. You may not use fractions or decimals, only integers. You will also be asked to judge specific parts of the images.

This is not a test. We are seeking your opinion."

5.2.2 Psychophysical Methods

The preferred method to analyse experimental data arising from a category-judgement procedure is to convert observers' responses into interval scale according to Torgerson's law of categorical judgement (Engeldrum, 2000). Due to the spread of this technique, no more details will be given, but the reader is invited to read the cited reference for more details.

For the second experiment, the obtained pair-comparison responses can also be converted into interval scales if additional assumptions are made about the distribution of the data. Many models are available to describe the probability density function that characterise the scale value differences, among which the Thurstone's case V model of the Law of Comparative Judgement is without doubt the most widely used in the imaging community. However, another model, based on work from Bradley and Terry (1984), is more appropriate to analyse such data, as it *"provides tractable maximum-likelihood estimators*

for scales, simultaneous confidence intervals and hypothesis tests for model fit, uniformity, and differences among populations of judges" (Handley, 2001), which are not readily available from the Thurstone model. Furthermore, the Bradley-Terry model is also able to handle directly incomplete data matrix. In practice, both models yield nearly identical scale estimates for complete data.

5.2.3 Results and Discussions

5.2.3.1 Overall results

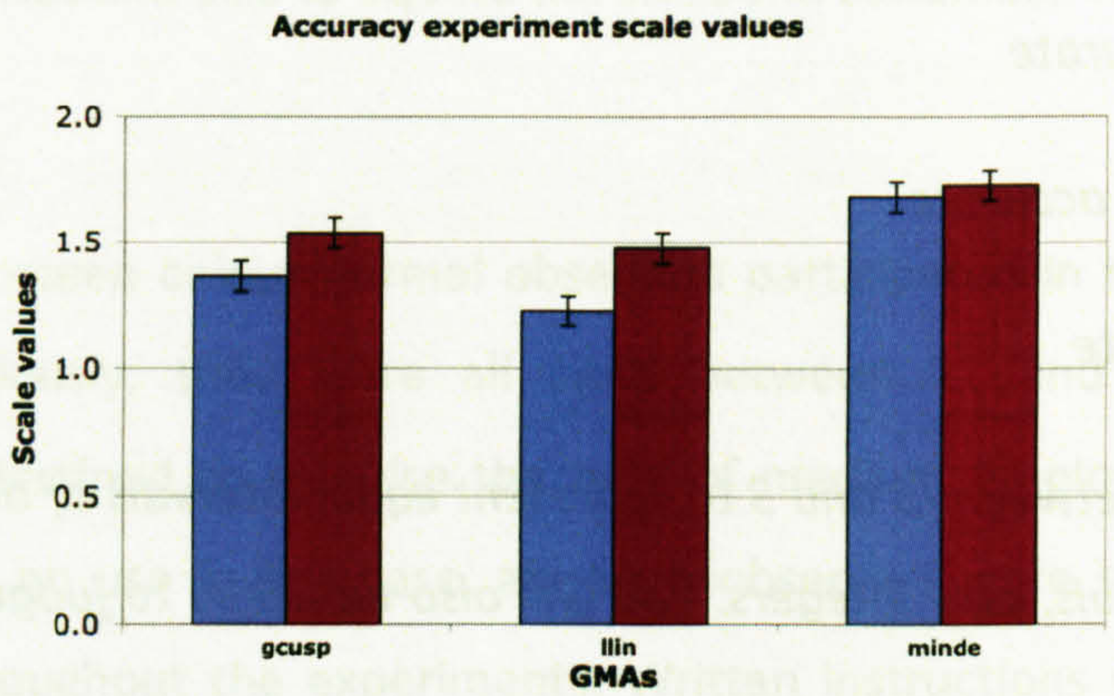


Figure 36: Experiment 2 - Accuracy: scale values (red bars: CRT-CRT workflow, blue bars: CRT-print workflow)

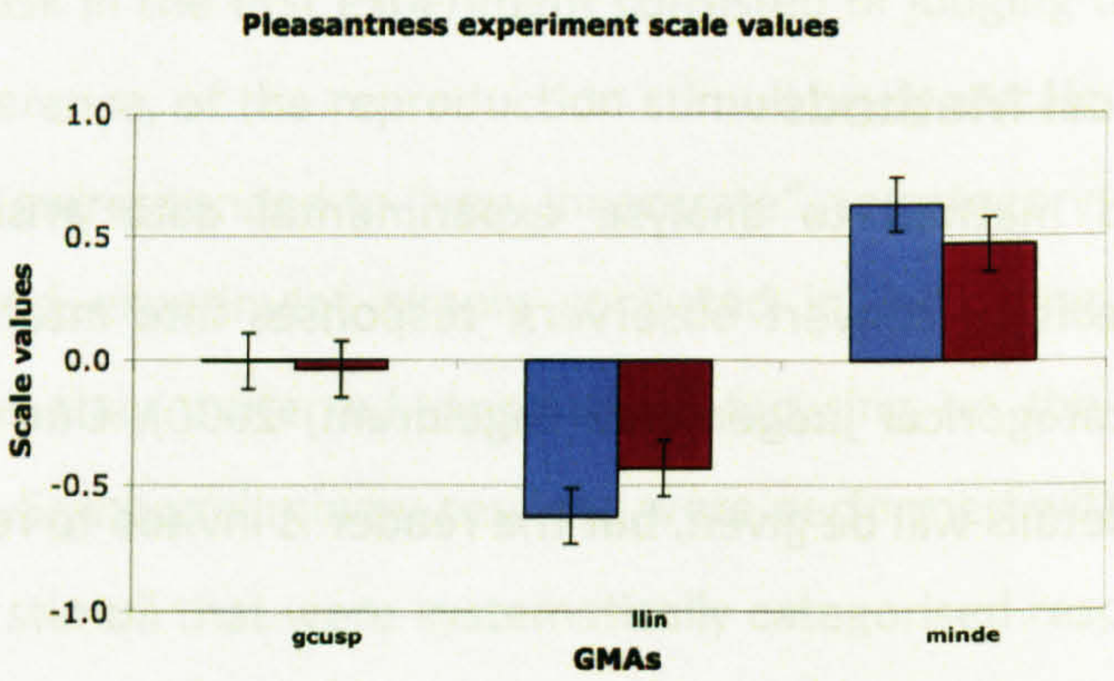


Figure 37: Experiment 3 - Pleasantness: scale values (red bars: CRT-CRT workflow, blue bars: CRT-print workflow)

The overall scale values obtained by each GMA for the second and third experiments are plotted in Figure 36 and Figure 37 respectively. Since the scale values obtained according to the two psychophysical techniques employed are not on an absolute scale, it is not reasonable to compare directly the absolute scale values from reproduction workflow 1

(blue bars) with those from workflow 2 (red bars). However, it is the correlation of the relative performances of the GMAs *between* the two workflows *within* each experiment that should be assessed. From these results, it can be seen clearly that the same overall conclusion can be made from both workflows. That is, in each experiment, the relative difference between each GMA's scale values is approximately the same regardless of the target medium used, as confirmed by the very high coefficient of determination R^2 between the scale values obtained for both workflows in each experiment (Table 21 col. 1). As previously mentioned, a hypothesis test can be performed on the pleasantness results to determine whether the data obtained from each workflow are significantly different or not. The test statistic for uniformity of these two groups is T=1.10 with 2 degrees of freedom, which is not significant, since the 95% chi-square cut-off is 5.99. There is thus no statistical difference in how the pleasantness of GMAs outputs depending on the medium used to generate the stimuli. Furthermore, for each experiment, the rank order of the algorithms is almost the same for both workflows. The only minor discrepancy comes from the accuracy experiment results, where the GCUSP and LLIN GMAs are not significantly different for the CRT-CRT case, whereas they are just different for the second workflow. Nonetheless, this difference has the merit to highlight the only noticeable discrepancy between the two workflows, i.e. that the differences between judgements made on the basis of a hardcopy simulation (Figure 36) tend to be less spread than those obtained with real hardcopy stimuli (Figure 37). The relatively poor precision of printers, in terms of colour reproduction, may explain the higher dispersion present in the hardcopy workflow.

Table 21 : Experiment 2 and 3: correlation of the scale values between the two workflows

	Overall	Pictorial Images	Business graphics	Business - outliers
Accuracy	0.998	0.996	0.492	0.984
Preference	0.996	0.989	0.980	0.987

5.2.3.2 Image Types Results

Another major goal of this study consisted in investigating whether the relative performance of the GMAs for both workflows was influenced by the type of images used. Figure 38-A and Figure 39-A shows the resulting scale values for business graphics for the accuracy and the preference experiment respectively, while those for pictorial images are illustrated in Figure 38-B and Figure 39-B.

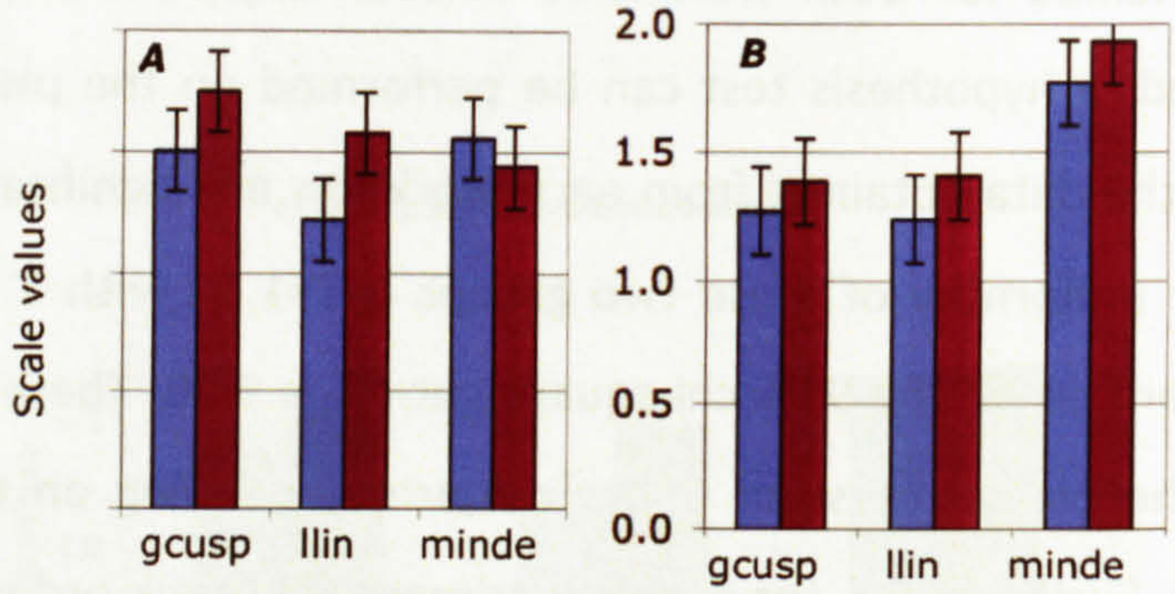


Figure 38: Experiment 2 – Accuracy: scale values for business graphics (A) and pictorial images (B) (red bars: CRT-CRT workflow, blue bars: CRT-print workflow)

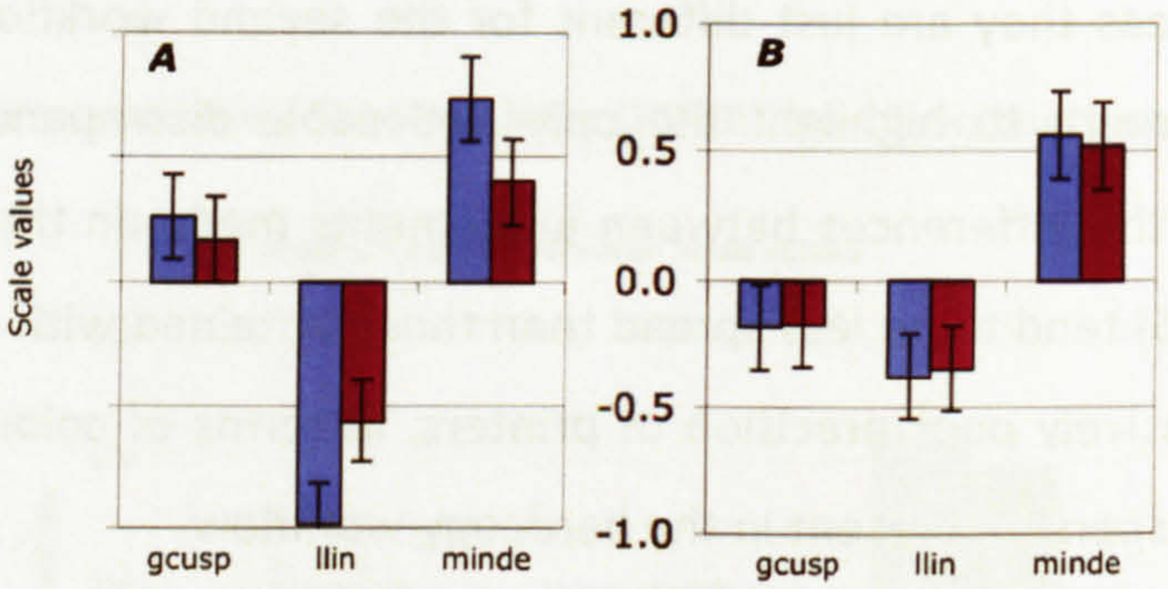


Figure 39: Experiment 3 – Pleasantness: scale values for business graphics (A) and pictorial images (B) (red bars: CRT-CRT workflow, blue bars: CRT-print workflow)

From these results, the high correlation between the two workflows for the pictorial images is easily observable (Table 21 col. 2), which strongly confirm the conclusions previously drawn. However, business graphics images do not show this tendency. Their results for the third experiment clearly illustrate the origin of the higher dispersion of scale values previously observed in the print workflow. Their results for the second experiment not only confirm this fact, but also suggest that the overall GMAs ranking differs depending on the workflow used, as the rank of LLIN algorithm is significantly different between the two workflows for the accuracy experiment. This is also confirmed by the weak correlation

observed (Table 21 col. 3). If that was the case, judgements based on softcopy simulations of a print would not be in agreement with those obtained with real hardcopies.

The analysis was further pursued by computing the scale values obtained by each GMAs for every test image, rather than image types. By comparing the ranking order and also the coefficient of determination of the GMAs between the two workflows for every individual image, it was found that a few did not concur with the overall results. For the accuracy experiment, two business graphics images, *pollution* and *air*, and one pictorial image, *cou*, generated markedly different results between the two workflows. Similarly, for the preference experiment, only a single image, *pollution*, was found to behave in a different way. Since most of them belong to the business graphics type, they might be responsible for the odd behaviour exhibited by this type of images. Indeed, removing those outliers from the business graphics dataset brought its results back in agreement with the overall conclusion, as Figure 40 illustrates. The correlation between the two workflows also increased dramatically (Table 21 col. 4).

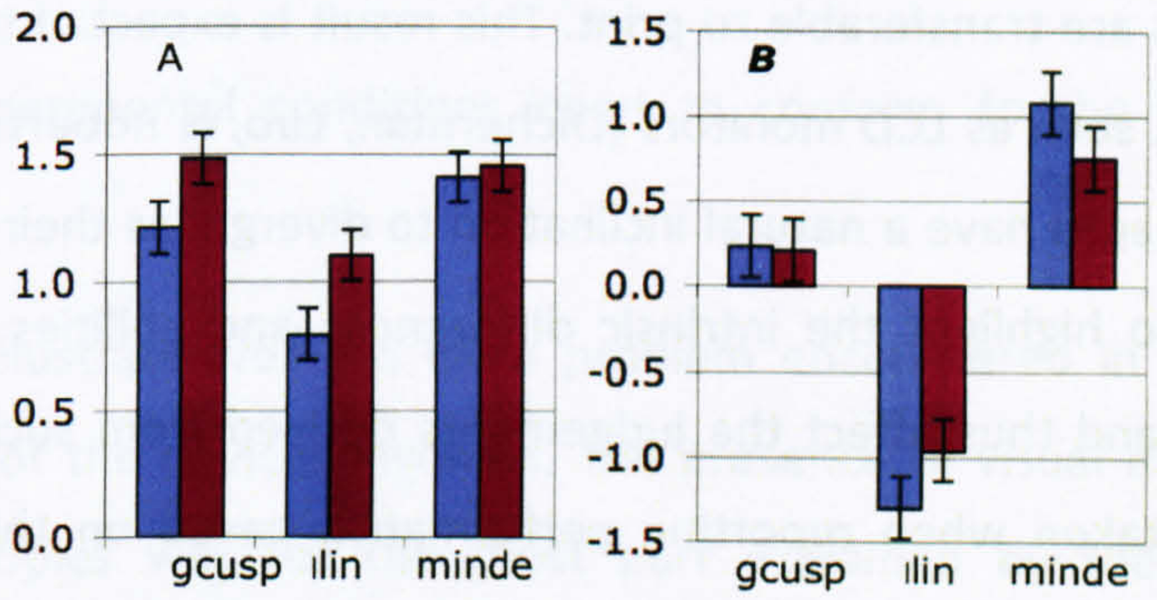


Figure 40: Business graphics images: average accuracy(A) and preference (B) scale values of individual GMAs for both experiments

5.2.4 Discussion

Overall, there was a 75% agreement in terms of ranking order between the two workflows for the accuracy experiment (60% for the business graphics stimuli, and 86% for pictorial images) and 92% for the preference experiment (80% for the business graphics stimuli, and 100% for pictorial images). Those results confirm once again that judgements made on the basis of pictorial images are very robust, and do not seem to suffer from any loss when simulations are used instead of hardcopies. On the other hand, business graphics

images are more prone to discrepancies, although the majority of stimuli seem to concur overall.

Several observers whose results concur with the observed discrepancies between the two workflows were asked to describe more precisely the reasons of their choices regarding the incriminated images. *cou*, an African sunset, and the only pictorial image exhibiting a different behaviour, really put the colour reproduction abilities of both media under serious stress, as it contains very dark areas alongside highly saturated sky colours. *pollution,* the only image to exhibit problems in both experiments, mainly consists of a single and saturated yellowish-red cast, which both media did not manage to reproduce in the same way, the printer giving emphasis to the yellow part while the CRT reproduced red better. The problem encountered in the *air* image resulted from the large and smooth red-to-green gradient it contains, which the lack of uniformity of the printer did not allow to render properly.

Despite the observed discrepancies, the overall conclusion still holds that judgements based on soft proofs are transferable to print. This result is expected to also apply to other display technologies, such as LCD monitors (Oicherman, Luo, & Robertson, 2006). However, business graphics images have a natural inclination to diverge, as their characteristics make them more prone to highlight the intrinsic differences and abilities of the reproduction devices of interest, and thus affect the judgements derived from such stimuli. Great care must therefore be taken when reporting performance based on this type of stimuli. A consequent stimuli set is not superfluous in this case. Nevertheless, the general conclusion that can be unambiguously drawn from these results is that, providing that the viewing conditions are very carefully equated and that a significant number of test images is used, soft proofs are suitable as surrogates for the final print in judging the quality of colour reproduction.

Chapter 6: Transition

The results obtained from this series of experiments are unambiguous. If the restricted experimental conditions specified in this study are met, judgements are transferable between prints and monitors. However, the constraints imposed here represent a significant departure from actual industrial settings. It is therefore essential to assess the pertinence of the obtained conclusions in a more realistic context.

Let's examine the conditions that have to be met for the conclusions to hold:

- Spatial information needs to be reproduced without defects or artefacts.
- The characterisation models of the devices involved obviously need to be of a respectable standard.
- The differences between two stimuli being judged have to be greater than the accuracy of the reproduction framework to be assessed meaningfully.
- The experimental conditions need to conform to the settings previously described.

The first two illustrate well the main problem encountered in this study, i.e. the reproduction abilities of the devices involved. The presence of visual differences between softcopies and hardcopies was for the most part explained by the printer's level of performance. While the monitors were representative of the ones commonly used in prepress houses, the printer employed, without being inadequate, was below the standard of typical state-of-the-art proofers used in the Graphics Arts industry. Such devices can achieve a colorimetric accuracy of less than 1 ΔE^*_{ab}, and evidently do not generate any spatial artefacts, thus producing outputs of the highest quality. The probability of achieving a true colorimetric match is therefore much higher with such devices. Obtaining a visual match is also within reach, as colour reproduction inaccuracies will be well below the visual threshold for discerning differences between pictorial images. As a consequence of the third issue mentioned, it will be possible to critically appraise finer and subtler tonal differences. The validity of the conclusion is therefore determined to a large extent by the fourth condition. Given the numerous restrictions placed on the experimental design, it is trivial to

conclude that typical industrial settings clearly do not comply with the conditions imposed here. The critical issue therefore consists in evaluating the impact of the discrepancies on the transferability of judgements.

Not all of the differences involved the geometric configuration or the viewing conditions. Few concerned the implementation details of the colour reproduction workflow. For instance, the gamut of available colours was limited to the intersection of both devices' gamut in order to ensure the pertinence of the comparison. This does not correspond to what is being done in the Graphics Arts industry, as all devices are used to their full potential. As previously mentioned, most print colours can be reproduced on monitors, as their gamut is much larger to that of a print (if the same white point is assumed). However, there still exist few discrepancies, as small colour regions "adjacent" to the printer's primaries, i.e. cyan, magenta and yellow, are outside of the gamut of standard monitors. This difficulty can easily be remedied to by the adoption of newer monitors that utilise *Adobe RGB* primaries. This colour space covers the entire colour reproduction area of the *ISO-Coated* colour space, which is one of the standard colour sets of the prepress industry. Even the colour reproduction area of prints with six or more primaries is almost entirely encompassed by the *Adobe RGB* colour space. Obtaining a colorimetric match, both in terms of accuracy and gamut, is therefore feasible, providing that the correct equipment is used.

However, the accurate reproduction of the spectral and spatial information of a stimulus is a necessary but not sufficient condition to guarantee an appearance match. The external conditions in which stimuli are appraised are also critical. The viewing distance is especially important, as the discrete manner in which the colour data signal is encoded is vastly different between hardcopies and softcopies. Halftone patterns could easily become visible if the viewing distance was not sufficient. However, since viewing booths are routinely utilised to appraise prints, hardcopies are most commonly viewed from a distance of at least a meter, which conform to the settings specified in the experimental design. On a larger dimensional scale, the perception of texture is similarly influenced by the viewing distance. However, most paper types do not exhibit any significant level of texture. Newspaper is a common exception, but its inherent poor quality significantly limits the ability to generate an accurate proof anyway. Overall, neither halftoning nor texture is likely

to affect the quality of the proof, as hardcopies are almost always viewed from sufficiently far enough to disguise such discrepancies.

The viewing distance is however not the only geometric factor influencing an object's appearance. As previously stated, the geometry of illumination and viewing plays a crucial part in determining the appearance of a print. For instance, the amount of viewing flare will significantly affect the appearance of dark stimuli. Even though its effect was nullified by the elaborate experimental setup, its presence cannot be avoided under real conditions. Similarly, the geometry of illumination and viewing cannot realistically be made to precisely match the measurement geometry, as was done in the experimental setup. Viewing booths usually generate a largely diffuse illumination, in stark contrast to the 45/0 geometry routinely used to provide colorimetric measurement. Such discrepancy may have some drastic impact on the resulting appearance, as geometric metamerism may come into effect. Geometric metamerism is defined as two colours which appear identical when viewed at one angle, but which no longer match if the viewing angle is changed. Geometric metamers are actually very common, as they may be caused either by differences in the physical characteristics of materials, such as their refractive indices, or differences between the surface structures of the two materials. Typically, one is shiny and smooth, while the other is dull and rough. By extension, we can also conclude that a material's surface does play a role in determining an object's appearance, especially if the geometric layout is altered, the geometry of lighting for instance. Therefore, although instruments may report a colorimetric match between two patches made of different materials, it is very unlikely that human observers will share this judgement. Since it is impossible not to view a stimulus at varying angles under non-experimental conditions, it is highly likely that the varying geometry will give rise to perceptual differences. As a matter of fact, the influence of surface reflections and varying geometries, although still largely disregarded by most of the research community, has been studied a handful of times.

Although colour measuring instruments were designed *"to be insensitive to small variations"* (Rich, 1988), the complexity of the interactions between light and matter cannot be brushed aside so trivially. Johnson and Stephenson (1983) showed that a 10° variation in the illumination direction produced a 0.1 variation in the Y tristimulus value of a green ceramic tile. They also reported that a change in aperture half angle from 2° to 8° at 0/45

produced a 1.7 unit CIELAB colour difference or a variation of 0.28 in the Y tristimulus value of that same tile. Similarly, the differences between the extremes of two sets of CIE recommended tolerances on 45/0 colour measurements vary between 0.2 and 0.7 ΔE*ab. Rich also reported on disparities originating from geometrical differences. *"The absolute tristimulus values do not agree. This is directly attributable to the difference in geometry. The National Bureau of Standards has shown that there are systematic differences between the absolute scale of reflectance factor for 6/d and 45/0 geometries"* (Rich, 1988). As can be intuitively envisaged, discrepancies due to diverging geometries are maximised when materials with disparate surface characteristics are considered. *"For the measurement of plane-surface intermediate gloss specimens and of textured-surface specimens, including textiles, where the first-surface reflection component may be distributed over a wide range of angles, measurement may be made with the specular component included, but the resulting color coordinates may not correlate best with visual judgments of the color. For the measurement of plane-surface, low-gloss (matte) specimens, the specular component may either be excluded or included, as no significant difference in the results should be apparent"* (ASTM, 2003b). Bidirectional measurements may therefore take more than simply colour into account, depending on the surface's roughness. *"The documentation issued* [with a spectrophotometer utilising the bidirectional 45/0 geometry] *listed the median illumination angle as 0° and the median viewing angle as 45° while the illumination half angle is 6° and the viewing half angle is 6.5°. An instrument with such half angles, while within the CIE specification, is indeed capable of measuring both specularly and diffusely reflected light simultaneously"* (Rich, 1988). Since the chromaticity of surface reflections essentially corresponds to the illuminant's chromaticity, the net effect will consist in a saturation decrease. Few other sources report the influence of gloss on colour measurements. *"The predicted color was compared to the measured color for cyan, magenta, yellow, red, green, and blue images that spanned a wide range of gloss. The RMS color error between the predicted and measured color of all of these images was found to be about 3 CIELAB ΔE*ab units"* (Dalal & Natale–Hoffman, 1999). Similarly, Leekley et al. (1978)investigated the changes in colour when printing onto various substrates with differing gloss and proposed a measure of degradation as a function of gloss. The influence of gloss is especially important in our context because, as indicated in the previous citation from the ASTM E1331 standard,

surface reflections affect the appearance of very glossy samples the most, and such samples are the most commonly used in the Graphics Arts industry.

Although surface reflections affect the perception and measurement of colour very significantly, geometrical disparities have a considerably more fundamental impact than anything that was previously mentioned. Emmel offers an eloquent account on the subject. *"In the graphic arts, most measuring instruments use a 45°/0° measuring geometry wherein the incident light beam is collimated with an incidence of 45°, and the detector is placed at an angle of 0°.* [...] *The entering collimated light beam follows within the medium a path of length X/cos(α), which is shorter than the average path of length 2X followed by diffuse light. Because the detector is at an angle of 0°, only the light emerging with an angle of 0° is detected. This emerging light beam follows in the medium a path of length X, which is also shorter than the average path length of 2X followed by diffuse light. Because the total path length of the light beam within the medium is shorter, the absorption of the light beam within the medium is not the same as for diffuse light"* (Emmel, 2003). The lighting setup therefore determines the level of absorption within the medium, hence also the reflected colour. Estimating the appearance of an object from a single bidirectional measurement is therefore a gross approximation. More attributes need to be taken account to replicate correctly an object's appearance.

In this context, the purpose of the second part of this thesis is to present means of improving the accuracy of soft proofing by increasing the amount of information rendered on screen. Ultimately, the aim is to provide an interactive computer-based simulation faithfully rendering the whole appearance of printed paper illuminated by the viewing booths typically employed in prepress houses. The task is however daunting. Although the imaging community traditionally focuses on geometric and chromatic (and its counterpart, achromatic) information, appearance is a complex phenomenon that has a much greater dimensionality.

A review of the complete sets of appearance attributes will assist us in adopting the right strategy to tackle this intricate problem. We will follow the classification established by a leading expert, John Hutchings, in one of the most challenging field in terms of appearance specification, the food industry (MacDougall, 2002). Although other ontologies (Caivano,

1997; Lozano, 2006) may differ, their elements are essentially the same. Any object viewed by a human observer generates two distinct types of perceptions in his/her psyche. *"Basic perceptions are of size, shape, surface texture, colour, translucency, gloss and their patterns and uniformities. Derived perceptions, formed through repeated* [...] *experience, comprise visual expectation* [such as, for food, creaminess and value]"(MacDougall, 2002). Psychology and semiotics are beyond the realm of this study, and the second type of perception will thus be omitted from the rest of the discussion. We will instead concentrate on basic perceptions, which originate from the physics of the constituent materials of the perceived object.

"There are three types of physically definable material properties – optical properties, physical form, and temporal properties.

- *Optical properties include light distribution occurring over the surface and within the depth of the material as well as reflectance, transmission, spectral and goniophotometric factors. These properties can be used in specifications of visually perceived object colour, translucency, transmission and gloss.*
- *Physical form includes statements of size, shape, pattern, surface texture, and those strength or viscosity properties that become visible by virtue of the temporal aspects of the object.*
- *Temporal properties include attributes that change with time, such as ageing and the way a material deforms under pressure or the jelly that wobbles when shaken"*(MacDougall, 2002).

Temporal properties are obviously irrelevant in the context of this study. Similarly, physical form is of little interest. Generating an accurate computer simulation of a piece of paper in terms of its size and shape is trivial. On the other hand, replicating the flexibility of paper is a lot more challenging than it first appears. It is however more of an engineering task, and we will focus on modeling rigid pieces of paper as a starting point. Texture is also not so much of a concern, as most types of paper commonly used in the Graphic Arts industry have a smooth surface with no visible texture at a macroscopic scale. Exceptions to this rule are newsprint, with its large grain, and specialist media for printing banners,

brochures, greeting and business cards, iron-on transfers and other creative projects. Again, we will restrict our scope to proofing scenarios by focusing on paper types presenting no visible texture effects.

Remain optical properties. Transmission, and its perceptual equivalent translucency, are often ignored in the printing industry, as only the reflected light presents interest. Anyhow, the magnitude of this phenomenon is at most diminutive, and is further eliminated by our use of a white backing in all situations, as mentioned in section 4.3. We will thus make the supposition that all our specimens are opaque. On the other hand, reflectance and spectral properties are inherent to the whole industry in general, and to colour management systems in particular. The reproduction of spectral information and, equivalently, chromatic information has been extensively studied (Fairchild, 2004; Hunt, 1998), and, if we omit gamut discrepancies, is operating well enough in our context. Best-of-breed soft proofing systems indeed achieve precision levels (IPA, 2005) that are below the visual threshold of human observers. From a theoretical point of view, it should therefore not be possible for a normal observer to distinguish *colour* differences between an original piece and a soft proof. However, as previously mentioned, this is not the case, and this is caused by divergences on how the term *colour* is defined.

From a human observer point of view, "[c]*olor is perceived away from the specular angle where* [the reflectance] *has only insignificant variation with respect to incident and perceiving angle*" (Kuo, Ng, & Wang, 2002a). The terms of this definition are broad and encompass a large range of illumination and viewing geometries. On the other hand, colour measurements are performed according to the CIE recommendations, which specify a very strict and limited set of illumination and viewing geometries. Furthermore, most instruments employed in the Graphic Arts industry only use a pair of these geometries, *45/0* and *0/45*, which are essentially equivalent. State-of-the-art colour management systems therefore perform their operation from a single sample taken out of the billion that may be obtained from the whole hemisphere of reflection. Much consideration have obviously been undertaken to select the most appropriate configuration, in a part of the hemisphere where variation due to geometry were deemed to be "*insignificant*" (Kuo, Ng, & Wang, 2002a). This is however not strictly speaking true. Metallic samples are a perfect example of that point (Klinker, 1993 p.30), but even printed paper may also exhibit some non Lambertian

behaviour in that region (see section 8.5.2). Furthermore, this type of illumination is highly directional, and do not represent well lighting environment used in the Graphic Arts industry, such as viewing booths. More complex interactions between light and matter are not taken into account, so what instruments measure is different from what observers perceive.

Optical properties, i.e. all the aspects involved in the reflection and transmission of light, constitute the cornerstone of the colour-imaging field. The measurement of reflectance and transmittance has greatly evolved over the last few decades, resulting in substantial improvements both in terms of accuracy and speed. These technological achievements, however great they may be, cannot however mask the fact that the underlying assumptions have severe shortcomings. The number of samples collected to estimate the reflective and transmissive properties of diverse medium is not sufficient to represent correctly the higher dimensionality of light and matter interactions. In fact, given the complexity inherent in the perception of colour, it was only natural that the problem was initially simplified to the study of simple stimuli, assumed to be near-perfect Lambertian diffusers, viewed and illuminated under restricted geometries specially designed to avoid more complex phenomena. However, this simplified approach continues to form the basis of current state-of-the-art colour management systems, and it should thus come as no surprise that such methods break down, at least partially, when employed in industrial applications, where the higher dimensionality of light and matter interactions become inevitable.

In that context, the remainder of this study will focus on modeling the optical properties of printed paper in a more exhaustive manner. This implies enhancing materials simulation beyond simple colour appearance prediction to more complete representations. The objective will be to produce a full goniometric simulation of the light scattering that occurs in the hemisphere above the surface of a print. Obtaining such a representation will allow the precise quantification of how light incident on a surface is reflected into a continuum of directions. The aim is not to depreciate the importance of colour, which will still remain a fundamental element to the success of the reproduction, but to incorporate additional elements to enhance the realism of the simulation. Perhaps the most salient feature of this new model will be the inclusion of gloss, the visual attribute resulting from specular reflections. But the intention is actually broader than simply reproducing colour and

gloss. They are not two distinct dimensions of appearance, but rather two instances of the same phenomenon. Each of them actually corresponds to specific geometries of illumination and viewing, CIE *45/0*, *d/0* or their reciprocal for colour, and near the specular angle for gloss. Rather than concentrating on these specific geometries, what is proposed here is a full goniometric simulation of the spatial distribution of light reflected by a print. On a more conceptual level, the aim is to make colour management frameworks evolve past simple pixel-to-pixel colorimetric reproduction towards a higher dimensional appearance simulation. The objective consists of supplanting simple flat colour reproduction with an interactive three-dimensional simulation of an object's appearance viewed under lifelike lighting environments. Ultimately, the intent is to more precisely replicate the interactions between light and matter.

Chapter 7: Physics of light and matter interactions

"Building a reflection model that is close [...] *to reality means using an established theory of reflection"*(Watt & Watt, 1992 p.48). In order to better simulate the optical properties of printed paper, the interactions between light and matter will first be considered from a physics perspective. We will start by considering the interactions' characteristics for a wide range of material in order to ensure the universality of the discussion. Although the proposed framework is aimed at the Graphics Arts industry, the problem it addresses are not unique to this domain, and adopting a similar approach could prove to be beneficial to many others. The following considerations will however promptly return to the materials at the heart of this study, paper and colorants.

The amount of literature devoted to the physics of optics may easily be overwhelming. A meticulous reader interested in a more mathematics and physics approach will probably find Born and Wolf (1980) treatment of the problem most compelling, while the more didactic approach adopted by Weisskopf (1968) may suit others better. Alternatively, Nassau (1983) provides an eloquent systematization of *all* the physical phenomena at the origin of colour.

7.1 Duality of optics

Optical phenomena can be broadly examined from two different perspectives. Traditionally, the interactions between light and matter can be accounted for in a very simple geometric manner, and the physics branch is aptly termed *geometric optics. "Light is treated as a set of rays, emanating from a source, which propagate through transparent media according to a set of three simple laws. The first law is the law of rectilinear propagation, which states that light rays propagating through a homogeneous transparent medium propagate in straight lines. The second law is the law of reflection, which governs the interaction of light rays with conducting surfaces (e.g. metallic mirrors). The third law is the law of refraction* [Snell's law]*, which governs the behaviour of light rays as they traverse a sharp boundary between two different transparent media (e.g. air and glass)"*(Fitzpatrick, 1999). The incident beam, the reflected beam, and the refracted beam lie in the same plane,

denoted the incidence plane. This is however a significant simplification of optics, and fails to account for many important optical effects such as diffraction and polarization.

Quantum optics issues omitted, the most complete description of optical phenomena is obtained from considerations of the electromagnetic nature of light (Born & Wolf, 1980). Classical electromagnetic wave theory is derived from Maxwell's prominent four equations (Maxwell, 1865), and is concerned with the propagation of electromagnetic waves through a medium or in a vacuum, as well as through different media. The latter case is evidently the most interesting from our point of view, as we are interested in what happens to an electromagnetic wave at an air/object interface. A full electromagnetic wave theory model of the propagation of light would however be too computationally demanding at present. High frequency approximations to wave propagation fortunately exist and reduce the complexity of the problem to a more manageable degree. The previously introduced *geometric optics* is the simplest of all. By considering the relationships that exist at the boundary between two media that are assumed to be perfect dielectrics, it is possible to derive the law of reflection and refraction from the equations describing waves propagation. As previously suggested, *geometric optics* cannot explain many phenomema such as diffraction, interference, polarization, aberrations and other complex effects. *Physical optics* (also known as *wave optics*) offers an intermediary solution between the precise theoretical framework established in wave electromagnetism and the approximate approach of *geometric optics*, which ignores wave effects altogether. By modeling the propagation of complex wavefronts through optical systems, *physical optics* accounts for all the previously mentioned phenomena in a correct manner.

At first glance, the dichotomy existing between the models describing the propagation of light might appear peculiar. They are based either on elementary geometrical considerations, or require extensive usage of some of the most fundamental physics theory.

"It turns out, however, that wave effects are only crucially important when the wavelength of the wave is either comparable to, or much larger than, the size of the objects with which the wave is interacting. When the wavelength of the wave becomes much smaller than the size of the objects with which the wave is interacting, then the interactions can be accounted for in a very simple geometric manner [...]. *Since the wavelength of visible light is only of order a micron, it is very easy to find situations in which the wavelength of light is very much smaller*

than the size of the objects with which the light is interacting. Thus, 'wave-less' optics, which is usually called geometric optics, has a very wide range of applications"(Fitzpatrick, 1999).

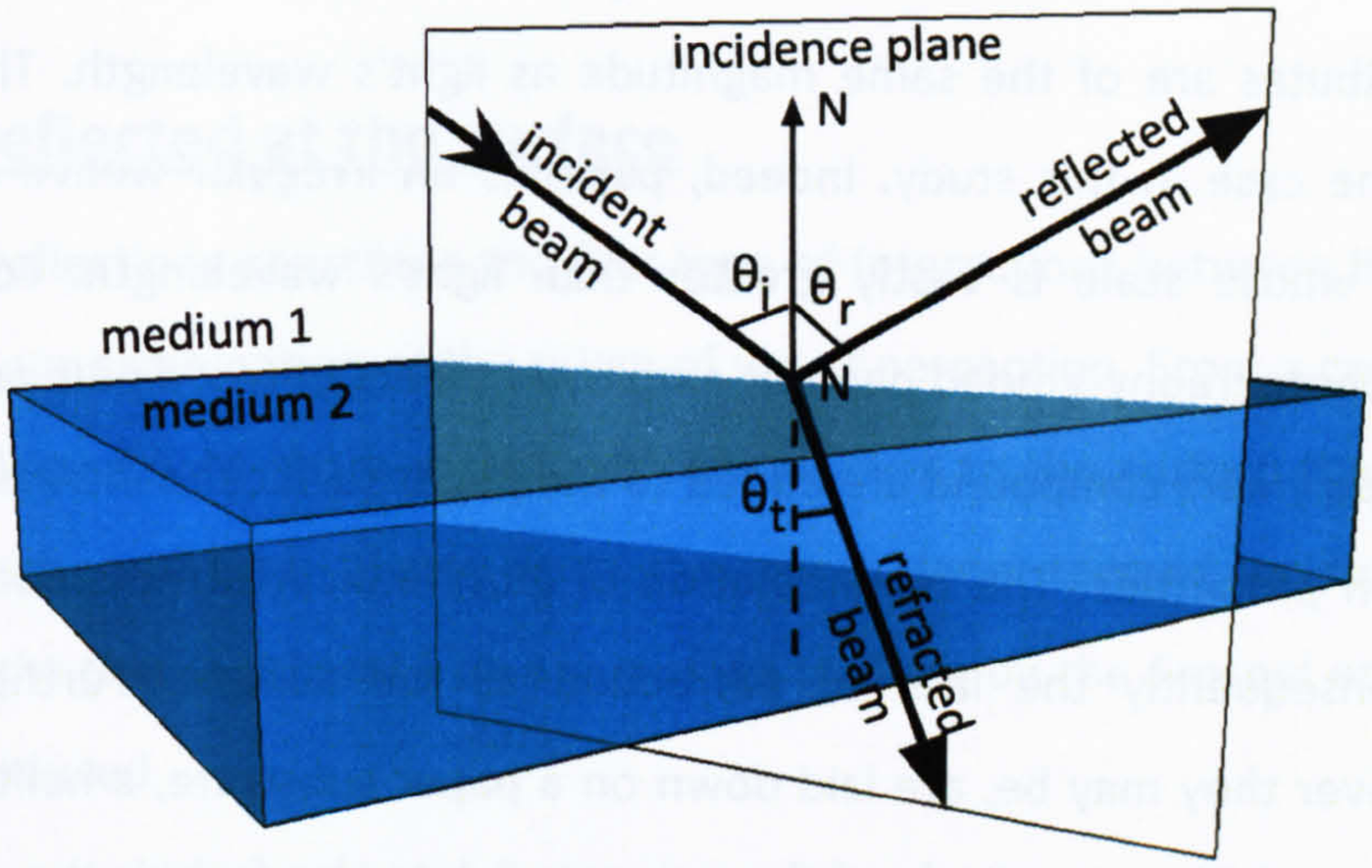


Figure 41 - Refraction and reflection of light at the interface between two media of different refractive indices.

It is yet unclear which approach will be more suitable to fulfill the needs of this study. An important consideration however emerges from this succinct analysis. The behaviour of light at an air/object interface is twofold: part of it is reflected back in the air, while the other part is refracted through the medium (Figure 41). We will now successively study the characteristics of the reflection and refraction process in more details.

7.2 Assumptions

Due to the complexity of the interactions between light and matter, certain assumptions need to be made to reduce the problem to more tractable proportions. First of all, we will assume most of the time that the first medium from which light originates is air. The following discussion concerning the physics of reflection and refraction will attempt to provide a comprehensive treatment on the subject, hence a more general approach may be adopted. Otherwise, please assume that when an interface is considered, it will be between air and the material under consideration.

7.2.1 Diffraction

The type of optical phenomena emanating from the interactions between light and matter depend on the scale of the entities forming the matter. Diffraction only occurs when a medium's attributes are of the same magnitude as light's wavelength. This is however unlikely to be the case in this study. Indeed, paper is an irregular weave of translucent cellulose fibers, whose scale is vastly greater than light's wavelength. For high-quality printing, where photography-graded paper is employed, fibers of wood pulp are not the only constituents, as polymers compound are added to the paper surface to increase its gloss and its ink absorption properties. The manifestation of gloss indicates the smoothness of the surface, and consequently the lack of diffraction at the surface. Furthermore, when colorants, whatever they may be, are laid down on a paper substrate, a halftone pattern is most commonly used. The magnitude of the generated dots also forbids the materialization of diffraction at the surface of the substrate. However, under the surface, i.e. within the colorants or the coating, this is another matter. Diffraction may well occur, especially for pigments-based inks. However, the vast scattering of light under the surface will most likely diffuse diffracted photons in an isotropic way (section 7.6.3). Diffraction phenomenon are therefore unlikely to appear in the light reflected from a hardcopy, and will therefore be mostly omitted from the remainder of the discussion. Geometric optics model should therefore suffice to satisfy the needs of this study.

7.2.2 Polarization

"Natural sunlight and lamp light are by and large completely unpolarized"(Wolff & Kurlander, 1990 p.45). However, light *"can become partially polarized or even completely polarized when specularly reflected off material surfaces.* [...] *After the first surface reflection, a light wave almost always has a polarization state different from it originally had after emanating from a source.* [...] *After two reflections, even initially unpolarized light can have an intensity different from the intensity predicted by lighting models that do not account for polarization"*(Wolff & Kurlander, 1990 p.44). In this study, we will only consider prints that are directly illuminated by light, as would be the case in pre-press houses. Furthermore, only surface reflections are affected by polarization, which confines this potential problem to a rather small proportion of the reflected light. Last, the differences between the two extremes polarization states are small under commonly employed angles

of incidence, as can be envisaged from Figure 42. In conclusion, polarization issues can be safely omitted from this study. It is nevertheless important to keep them in mind for future evolutions of the developed model.

7.3 Light reflected at the surface

Surface reflections constitute the first type of interactions between light and matter encountered in our examination of the origin of visual perception. From a geometrical point of view, the reflected light beam makes with the normal to the surface the same angle as the incident beam, and both beams lie in the same plane, termed incidence plane. The proportion of light that is reflected at the surface is given by the Fresnel equations, which will now be introduced.

7.3.1 Fresnel equations

From a physics perspective, surface reflections are embodied by the Fresnel equations, which describe the behaviour of light when moving between media of differing optical properties. Derived from the Maxwell equations, they provide a basis for determining the fractions of the incident energy that is reflected and transmitted at an interface. The reflectance and the transmittance of an optically smooth surface are expressed in terms of the:

- angle of illumination θ.
- polarization of light.
- index of refraction, which in its most general form is a complex number $\tilde{n} = n + jk$, whereby its real part is the simple index of refraction n and its imaginary part the extinction coefficient k.

Since both n and k vary with wavelength, as demonstrated by the ubiquity of chromatic aberrations, the Fresnel equations are also dependent on the light's wavelength. They can account for surface reflections generated by any types of material. The treatment is however different depending on the type of material, and the extinction coefficient k is accountable for this discrepancy. k is a measure of how well a particular substance absorbs electromagnetic radiation. If the radiation hardly penetrates the material, such as for metals,

the extinction coefficient is high. The derivation of the Fresnel equations in this case is tedious and beyond the scope of this study. The final set of equations are provided by Glassner (1995), and curious readers are invited to read Born and Wolf (1980) treatment on inhomogeneous plane waves for further information. Conversely, the derivation is relatively straightforward for other types of materials, such as non-metals (dielectrics). Their extinction coefficient is nil, indicating that they do not conduct electricity, which simplifies their refractive index by reducing it to its real part *n* only. Since we further assumed that the incident light is not polarized, the Fresnel coefficient for dielectrics can simply be expressed as (Cook & Torrance, 1981; Glassner, 1995):

$$F = \frac{1}{2}\frac{(g-c)^2}{(g+c)^2}\left\{1+\frac{[c(g+c)-1]^2}{[c(g+c)+1]^2}\right\} \tag{7}$$

where

$$\begin{aligned} c &= \cos\theta \\ g &= n^2 + c^2 - 1 \end{aligned} \tag{8}$$

Figure 42 and Figure 43 illustrate the behaviour of the Fresnel reflections for different types of materials. The first figure exemplifies the difference of reflected intensity between a metal and a dielectric. The surface reflections from the former are in general much stronger than for the latter. However, as the incident angle approaches grazing angle, all materials become 100% reflective. The second figure illustrates the variations in surface reflections as a function of wavelength and incidence angle.

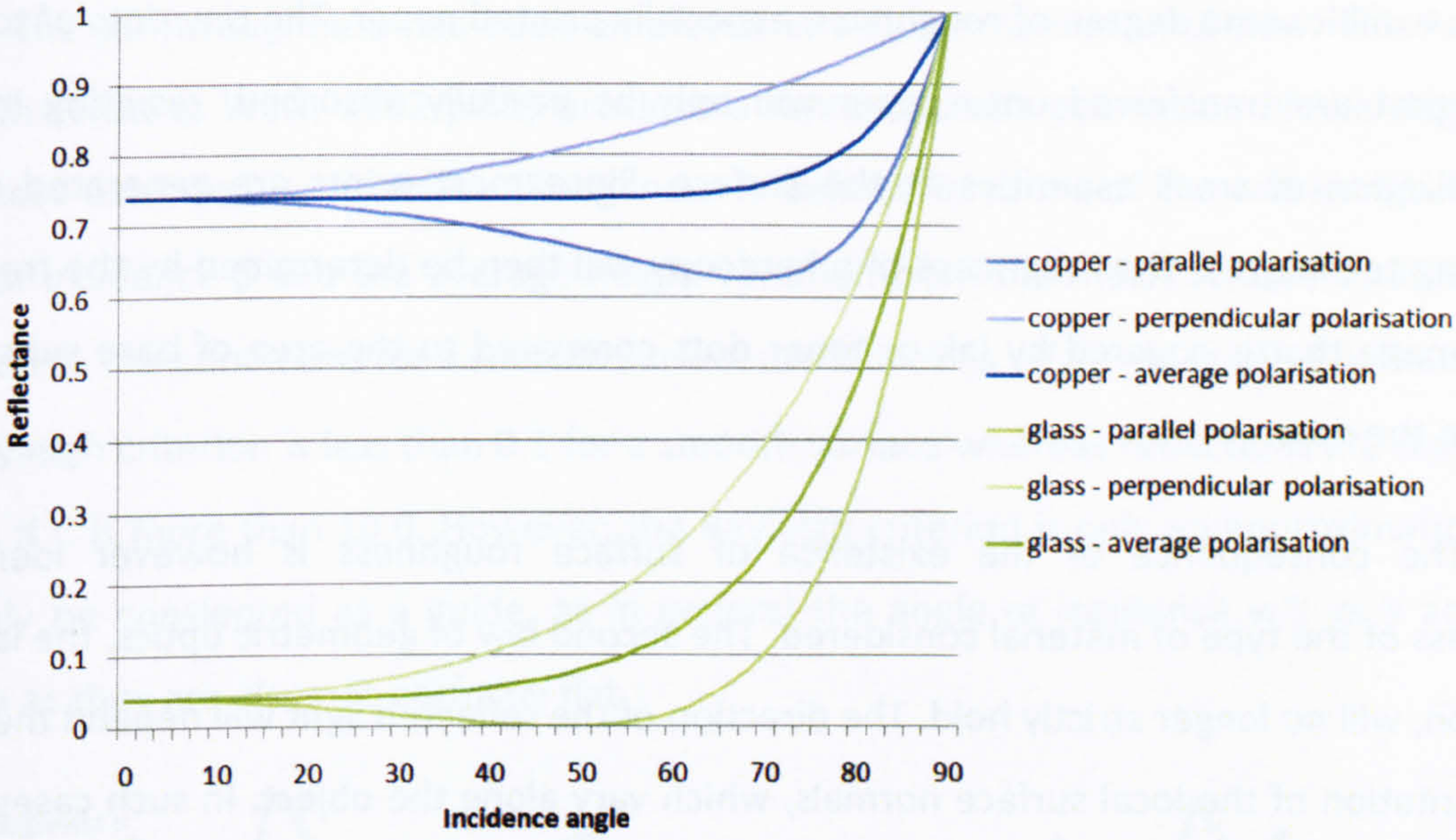


Figure 42 - Reflectance of glass and copper as functions of angle of incidence and polarization plane for a wavelength of 700 nm.

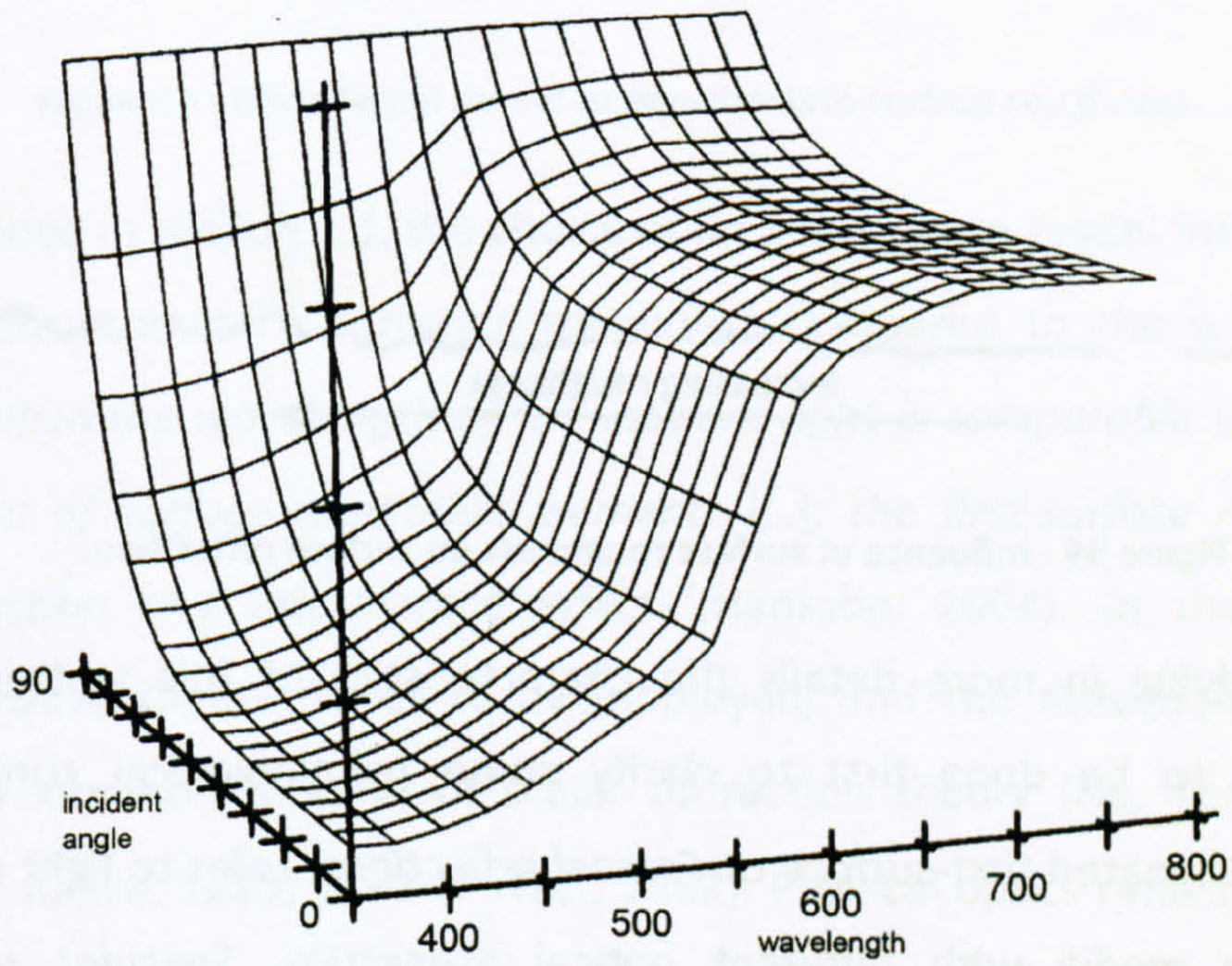


Figure 43 – Average (in terms of polarisation) reflectance of copper as functions of angle of incidence and wavelength. Reproduced from Klinker (Klinker, 1993) .

7.3.2 Rough surfaces treatment

The equations given in the previous section only apply to ideal surfaces that are considered to be clean and optically smooth. *"Practical surfaces depart from this ideal case, because they are rough and possibly oxidized and contain surface contaminants"*(Watt & Watt, 1992 p.50). Unless they were specifically designed to be optically smooth, most

surfaces exhibit some degree of roughness, especially printed paper. The tiny dots of toners or inks that are transferred onto paper will only be partially absorbed, resulting in the materialization of small asperities at the surface. Since most prints are generated using halftoning techniques, the roughness of a hardcopy will then be determined by the fraction of the image that is covered by ink or toner dots compared to the area of bare substrate between the dots.

The consequence of the existence of surface roughness is however identical regardless of the type of material considered. The second law of geometric optics, the law of reflection, will no longer strictly hold. The direction of the reflected light will depend then on the orientation of the local surface normals, which vary along the object. In such cases, the light is scattered to some degree around the global angle of perfect mirror reflection. As can be seen from Figure 44, the degree of surface roughness will determine the geometrical distribution of surface reflections.

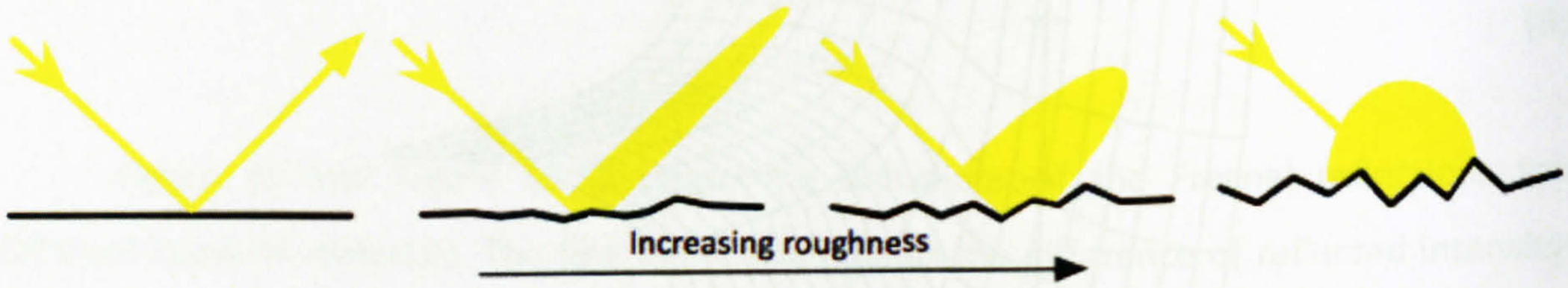


Figure 44 - Influence of surface roughness on surface reflections.

Before studying in more details the characteristics of this scattering process, a digressions needs to be done first to clarify some terminological confusion. *Surface reflections* (also designated *first-surface* or *Fresnel reflections*) refer to light reflected at the interface between media with different optical properties. *Specular reflections* (also designated *regular reflections*) refer to light reflected in the direction of ideal mirror-like reflection. This is a conceptual distinction, rather than a purely geometric one. For an optically smooth surface, the surface reflections will be specular, whereas for a rough surface, the surface reflections will be diffused, or scattered, to a certain extent around the specular direction.

$$C = \frac{4\pi\sigma \sin\beta}{\lambda} \tag{9}$$

The term roughness has been employed in a fuzzy manner until now. The Rayleigh criterion (Watt & Watt, 1992 p.51) provides a more rigorous estimation of the degree of roughness exhibited by a surface. Roughness is defined in terms of the wavelength λ, the angle of incidence β and the average height variation from the mean surface σ, as seen in Equation 9. The conditions for a smooth surface are stated as $\sigma/\lambda \to 0$ or $\beta \to 0$. Practically, the Rayleigh criterion is less than 0.1 for a smooth surface whereas reflections are uniformly diffuse if *C* is more than 10.0. However, the Rayleigh criterion is only an approximation and can only be considered as a guide, as in general the angle of incidence will vary across a surface as they are also not perfectly flat.

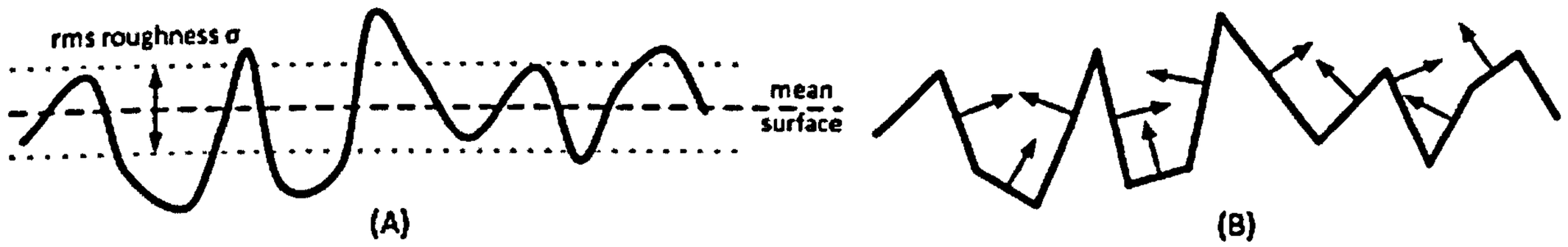


Figure 45 - Microfacets model to approximate surface roughness.

As mentioned in section 7.1, the choice of an appropriate model for light reflection is determined by the scale of a surface's roughness compared to the wavelength of the incident light. *"When the wavelength of the incident light is comparable to or smaller than the projected size of surface roughness elements* [...]*, the first-surface reflection process introduces diffraction and interference effects"*(Hansson, 2004). In that case, classical electromagnetic wave theory needs to be employed, and the reflection process will be governed by the Kirchhoff integral of scalar diffraction theory (He, Torrance, Sillion, & Greenberg, 1991; Kajiya, 1985; Watt & Watt, 1992). Physical optics reflection models, such as the ones proposed by He *et al.* (1991) and Kajiya (1985), provide an approximation to this complex equation. However, we made the reasonable assumption that the features of the surfaces presently studied were much larger than light's wavelength. The problem can thus be tackled using the simpler approach provided by geometric optics. In that case, "[a] *common approach to modelling surface roughness is to use a so-called microfacet model, where the surface is still assumed to be clean and to consist of very small perfect microfacets oriented in various directions around the mean surface.* [...] *Because any small area on the surface is assumed to consist of a collection of microfacets of orientation distributed about the mean or average surface over that area, specular reflection will occur as for an optically*

smooth surface but will be spread about the mirror direction"(Watt & Watt, 1992 p.52). The concept is illustrated in Figure 45 (B). Essentially, the aim is to broaden classical geometric optics by using a statistical approach. *"The complete geometrical specification of a reflecting surface is rarely known* [...]. *However, small scale variations of the electromagnetic field on the surface are averaged out when viewed from a distance. This averaging over points on a surface is statistically equivalent to averaging over an entire class of surfaces with the same statistical description"*(He, Torrance, Sillion, & Greenberg, 1991 p.177). The roughness of a surface is usually represented by a distribution function of the slope or orientation of the reflecting planes of the microfacets. Diverse facet slope distribution models have been published over the years. Gaussian distributions frequently appear to represent the perturbation of the heights relatively to the mean surface, as in Davies (1954) and Torrance and Sparrow (1967). However, *"Beckmann provided a comprehensive theory that encompasses all* [kind of] *materials and is applicable to a wide range of surface conditions ranging from smooth to very rough"*(Watt & Watt, 1992 p.53). The Beckmann distribution *D* is widely used in most modern reflection models and is given by:

$$D = \frac{1}{m^2 \cos^4 \alpha} e^{\left(-\frac{\tan^2 \alpha}{m^2}\right)} \tag{10}$$

where *m* is the RMS slope of the microfacets, and *D* returns the proportionate area of microfacets orientated at angle α to the average normal of the surface.

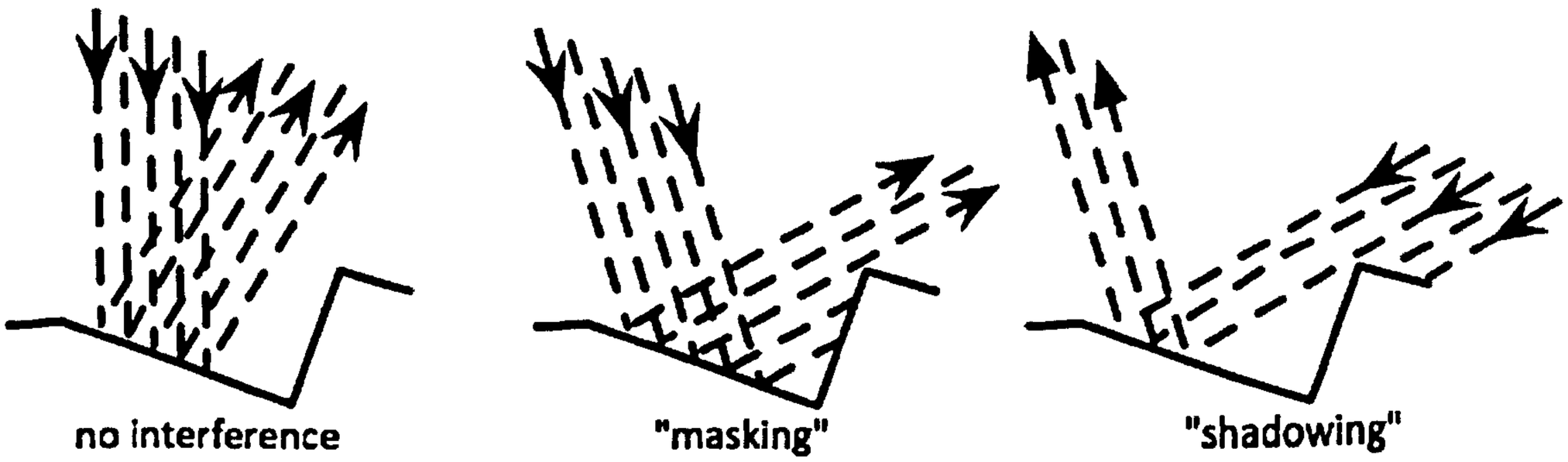


Figure 46 - Geometric attenuation factors.

The last important notion relative to the treatment of rough surfaces is the geometric attenuation factor. *"The surface geometry for rough surfaces also introduces conditions*

where parts of the surface shade or mask other parts of the surface"(Hall, 1989 p.25), as can be seen from Figure 46. "*This effect manifests itself at large angles of incidence or reflection, where parts of the surface are shadowed and/or masked by other parts, reducing the amount of reflection*"(He, Torrance, Sillion, & Greenberg, 1991). The geometric attenuation factor *G* describes the fraction of the microfacets oriented to reflect light from the source to a viewer, which are visible to both the light and the viewer, and not self-shaded by the surface. The factor is defined as:

$$G = \min\{1, G_s, G_m\} \tag{11}$$

where G_m corresponds to the masking term and G_s to the shadowing term, which are respectively obtained in their vector forms by the following equations:

$$\begin{aligned} G_m &= \frac{2(N.H)(N.V)}{V.H} \\ G_s &= \frac{2(N.H)(N.L)}{V.H} \end{aligned} \tag{12}$$

where N, V, L, and H are all unit vectors, N representing the local surface normal, V the incident light beam, L the reflected light beam and H is the half-vector between L and V.

7.4 Light refracted

A light beam that hits an interface separating two media of different optical properties is partially reflected into the first medium and partially refracted into the second medium (Figure 41). The former process had been treated in the previous section. This section will be devoted to the latter.

As previously mentioned, the proportion of light that is transmitted through the interface is also determined by the Fresnel equations. From a geometrical point of view, the refracted beam makes with the normal of the surface an angle that is related to the angle of the incident beam by Snell's law. Once again, all three beams lie in the same incidence plane. This change in direction is however not the only propagation alteration endured by refracted light. While the potential interactions between light and air are fairly limited, the possibilities are exponentially increased when other materials are considered. Nassau (1983) actually devoted an entire book to survey the entire range of possible interactions. However,

we will mostly focus here on the needs of the traditional industrial partners of colour science, the textile, painting and also evidently the printing industries. Furthermore, we will not delve in to the complexity and the richness of each individual interaction phenomenon. Instead, we will only consider their consequences on a macroscopic scale. The optical phenomena affecting the propagation of light through matter can then be narrowed down to three main categories, which will be introduced successively.

7.4.1 Absorption

"The term light absorption refers to all processes that reduce the intensity of a light beam when interacting with matter"(Emmel, 2003). Once again, diverse mechanisms are responsible for such phenomenon, but the outcome is essentially always the same. Electromagnetic waves in the visible band are dissipated into another kind of energy, most commonly through thermal agitation. The classic method to determine the absorption of light by a certain material is represented by the *Beer-Lambert-Bouguer* (Lambert, 1760), which is an empirical relationship relating the absorption of light to the properties of the material through which the light is travelling. The variation due to absorption of the flux of a collimated beam travelling through a material is usually given as:

$$d\phi = -\varepsilon(\lambda)\, c \ln(10)\, \phi\, dx \tag{13}$$

where the entering intensity flux ϕ is decreased after passing through a infinitely small thickness dx of concentration c, $\varepsilon(\lambda)$ being the absorption coefficient of the medium. In essence, the law states that there is a logarithmic dependence between the transmission of light through a substance and the concentration of the substance, and also between the transmission and the length of material that the light travels through. Thus if d and α are known, the concentration of a substance can be deduced from the amount of light transmitted by it:

$$T(\lambda) = \frac{\phi(X)}{\phi(0)} = 10^{-[Xc\varepsilon(\lambda)]} \tag{14}$$

This leads us to considering a third macroscopic process by which light interacts with matter, transmission. If the thickness of the material is not sufficient to absorb all the light, part of the light that entered the material will be refracted into another medium at the

interface located at the other physical end of the material. The absorption properties of a material regulate its transmissive abilities, which in turn *"determine the opacity, translucency or transparency of materials. A material is opaque if no part of the refracted light beam is able to penetrate through the entire material. Otherwise, the material is transmittant, appearing either transparent or translucent, depending on how much of the light is scattered in the material body"*(Klinker, 1993 p.31). The transmissive properties of a certain material can be quantified by the transmittance factor, given in equation 14, which is defined as the ratio between the incoming flux $\phi(0)$ and the flux $\phi(X)$ emerging from a layer of thickness X. However, we assumed that printed paper was an opaque material. Transmission phenomenon and its consequences will therefore be omitted from the rest of this study.

7.4.2 Fluorescence / Emission

The second type of macroscopic interactions affecting the propagation of light through matter is fluorescence. It shares certain similarities with absorption, as it *"occurs when light is absorbed at one energy, and only part of the energy is reemitted* [...]. *This necessarily produces light at a lower energy, that is, at a larger wavelength"*(Nassau, 1983 p.59). A fluorescent material absorbs electromagnetic energy, typically in the ultraviolet band of the electromagnetic spectrum but it may also occur in the visible band, and re-emits it at a longer wavelength. In the context of this study, we are obviously mostly interested in the re-emissions that occur in the visible band of the electromagnetic spectrum. *Fluorescence* is actually an instance of *luminescence*, which encapsulates the diverse process by which energy is emitted from a material at a different wavelength from that at which it is absorbed. Another occurrence of such processes is *phosphorescence*, which is essentially delayed *fluorescence*, whereby the re-emission of energy is not immediate but occurs during a timeframe that is dependent on the material. Once again, phenomenon such as fluorescence will be omitted from this study.

7.4.3 Scattering

The scattering of light is the last of the major physical processes that contribute to the visible appearance of most objects. It is defined as a general physical process whereby propagating waves or moving particles, such as light, are forced to deviate from a straight

trajectory when travelling through a certain medium. The cause of this deviation is usually the presence of one or more localized non-uniformities in the medium through which it passes, resulting in local variations of the refractive index. The types of non-uniformities that may cause scattering are too numerous to list, but a small sample includes particles, bubbles, droplets, density fluctuations, or even textile fibers in clothing. The usage of the term is however commonly broadened to include other interactions beyond those occurring within the medium. For instance, the diffusion of the reflected light by a rough interface is also often referred as scattering of light. By analogy with the absorption formulation, the variation due to scattering of a collimated light beam travelling through a material is usually given as (Koschmieder, 1924):

$$d\phi = -\sigma(\lambda)\, c\, \ln(10)\, \phi\, dx \tag{15}$$

where the entering intensity flux ϕ is decreased after passing through a infinitely small thickness dx of concentration c, $\sigma(\lambda)$ being the scattering coefficient of the medium. Figure 47 provides an illustration of the phenomenon. Part of the incident radiation does not cross the whole thickness of the medium, but is instead dispersed in multiple random directions.

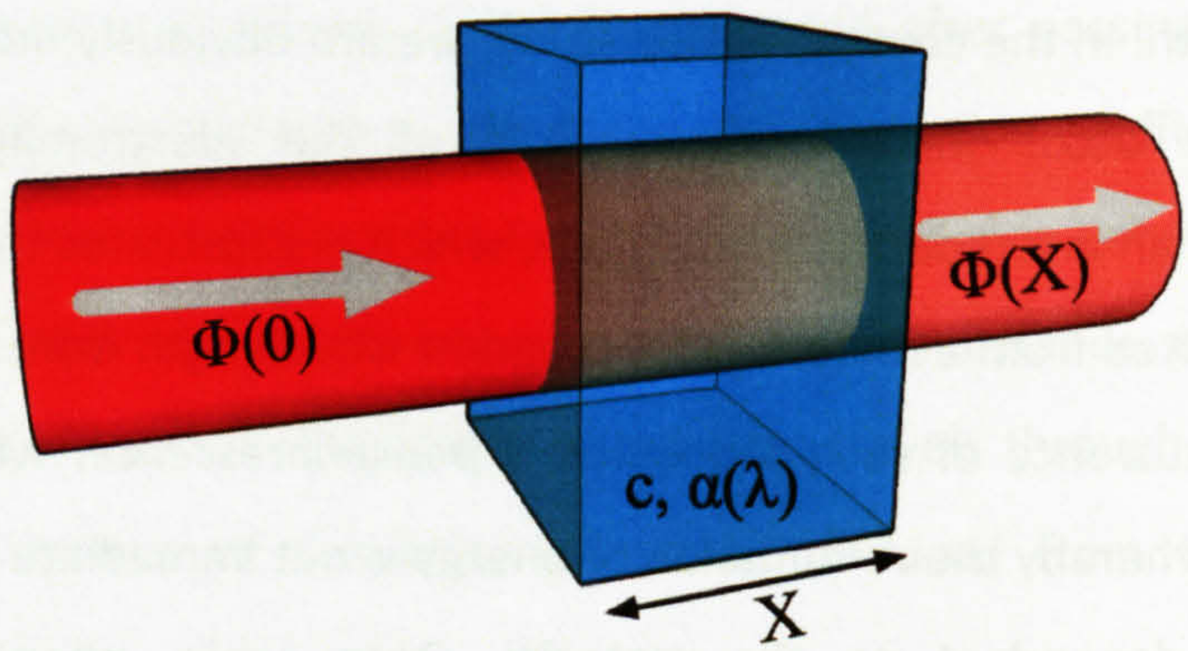


Figure 47 – Illustration of the outcome of absorption and scattering phenomena.

The complexity of scattering is tightly coupled with the density of the non-uniformities present in the medium. In its simplest incarnation, scattering is assumed to be single or independent, i.e. the light scattered by a non-uniformity does not interact with others. Since it is virtually impossible to locate with enough precision the location of the

non-uniformities relatively to the path of the radiation, the outcome of single scattering, which tends to depend strongly on the exact incoming trajectory, appears random to an observer. Single scattering is therefore often described by probability distributions. The Rayleigh scattering theory and its generalization, the Mie theory (Glassner, 1995), are the two established theories employed to describe independent scattering. They can for instance be employed to explain some of the most striking natural colours. "[B]*lue light is more strongly scattered than red light. Sunlight scattered in the atmosphere is mostly blue, which explains why the sky is blue. At sunrise and at sunset, the light from the sun has to traverse a thicker atmospheric layer than at noon, so most of the blue light is scattered, and the remaining unscattered light is mostly red*"(Emmel, 2003).

However, as the density of the non-uniformities or the thickness of the medium increase, the independence assumption no longer holds, and the single scattering gives way to multiple scattering, i.e. a radiation may be scattered many times before it escapes the medium. The principal distinction between the effects of single and multiple scattering is that single scattering can usually be treated as a stochastic phenomenon while multiple scattering tends to be more deterministic. Indeed, the randomness of multiple scattering interactions is approximately averaged out by the large number of scattering events, so that the final path of the radiation appears to be an isotropic distribution of intensity as the radiation is spread out. This is exemplified by a light beam passing through thick fog. This is especially pertinent for the study of pigment inks, where high densities of pigments have to be reached to obtain saturated colours. Multiple scattering is analogous to diffusion, and the terms multiple scattering and diffusion are interchangeable in many contexts. Unlike independent scattering, there is no comprehensive theory modelling multiple scattering of large non-uniformities densely packed. However, by taking advantage of the deterministic nature of multiple scattering, this problem can be tackled satisfactorily by considering the macroscopic consequences of multiple scattering rather than for each individual non-uniformities. Such solutions will be introduced in the next section.

7.4.4 Phenomenological models

A phenomenological approach to physical modelling consists in describing the outcome of a process rather than its intrinsic mechanisms. Due to the complexity of multiple

scattering for unconstrained conditions, i.e. large and randomly shaped non-uniformities that are densely packed, no exact quantitative solutions exist at the moment. Phenomenological models provide a mean to describe the electromagnetic distribution of radiation resulting from multiple scattering without having to consider the multitude of individual events.

Given the similitude of the previously introduced interactions between light and matter, the scope of the problem addressed by phenomenological problem is not usually limited to multiple scattering only. Rather, all other optical phenomena, i.e. absorption and emission, are also included for the sake of completeness. The most comprehensive theory describing the flux variation of a collimated beam travelling through a certain medium is the *radiative transfer equation* first proposed by Chandrasekhar (2003). In essence, the equation provides a single framework to account for all the interactions between light and matter described above. The alterations endured by a beam of radiation as it propagates through a medium are expressed in terms of the energy loss ensued from absorption, the energy gain stemmed from emission and energy redistribution arising from scattering. The radiative transfer equation is undeniably extremely powerful, but its ubiquity comes at the price of an intimidating complexity that makes it difficult to deploy.

In practice, approximations of the radiative transfer equations are commonly employed. The theory established by Kubelka and Munk (1931) is without a doubt the most widespread one. It was originally developed for paint films placed on top of a reflecting surface, but was successfully implemented in a larger range of applications especially for the painting and printing industries. The theory aims at estimating the reflectance R of a medium as a function of its absorption and scattering properties. The scattering is assumed to be isotropic as a result of multiple scattering, hence the global reflectance is assumed to be uniformly diffuse, as is the incoming illumination. Emission phenomena, such as fluorescence, are also omitted from the scope of this theory. Despite its popularity, the Kubelka-Munk theory does not provide more than an approximative solution to the radiative transfer equation, and the failures of the model have been extensively reported over the years. Numerous attempts at improving its accuracy have been published, and the *multiple fluxes* theory proposed by Mudget and Richards (1971) provides a more solid framework. While the Kubelka-Munk model only considers two fluxes, the approach taken by Mudget

and Richards consists in dramatically increasing the number of fluxes involved in the computation. The principle still consists in considering the balance of energy in an infinitesimal layer of thickness dx, hence pairs of upward and downward fluxes still form the core of this method. However, the angular dependency of each interaction, especially scattering, is better taken into account because each fluxes pair represents the variation of flux in a conical subdivision of the space. The problem can still be represented by a system of linear differential equations, but of much larger proportions. The peculiarities of this model are beyond the scope of this study, hence the curious reader is invited to read original article from Mudget and Richards. Alternatively, Völz (2001) devotes a section of his book introduces an adaptation of this model used in the painting industry, where a four-fluxes version of this theory is routinely employed.

7.5 Synopsis

The diverse processes describing the major types of interaction between light and matter have been introduced individually in the previous sections. This section will on the other hand consider them all together in order to assess the global consequences of their successive impact on light.

Let us follow the propagation of a light beam as it interacts with a layer of translucent and coloured substrate. The substrate is assumed to be located in a surrounding environment constituted of air. The whole process is illustrated in Figure 41. As the beam reaches the air-substrate interface, part of it is reflected back into the air, while the rest is refracted in the substrate. The reflection follows the second law of geometric optics, I.e. the light is reflected in the specular direction. However, depending on the roughness of the surface, the reflected beam may be scattered around the specular direction, to an extent governed by the roughness of the surface. The light that has not been reflected at the interface penetrates the layer, and changes its course according to Snell's law. As the medium contains colorant particles, part of the light is absorbed at certain wavelengths. Depending on the properties of the absorbing entities, the energy may be entirely transformed into another type of energy, typically thermal energy, or only partially transformed, with the remaining energy being re-emitted. The re-emitted radiation as well as the refracted light that has not been absorbed may be scattered in random directions if

they encounters colorant particles. More precisely, while the path alteration due a single scattering event is random, the macroscopic distribution of scattered light may be considered isotropic if multiple scattering is assumed, i.e. if the colorant concentration is sufficiently high enough. On average, the refracted light that is not absorbed is therefore scattered diffusely within the medium. Scattered beams continue their course within the medium, and may be absorbed or scattered again depending on where their trajectory lead them. The length of time they spend in the medium is obviously variable, but it is safe to say that they eventually reach one of the medium's interfaces with the surrounding environment. As for the initial air-substrate interface, part of the light is refracted in the surrounding environment while the rest is reflected inside the substrate, a process known as *internal reflection*. However, the characteristics of the refraction process are different in that case. Indeed, the substrate typically has a higher refractive index than the surrounding environment. Hence, by virtue of the Fresnel equations, the proportion of internal reflection at the substrate-air interface will be much greater than the proportion of light that has been reflected at the air-substrate interface. A significant fraction of light therefore remains in the medium, and is increasingly scattered through the same absorption and scattering mechanisms. Eventually, the light that is not absorbed will finally leave the substrate. Depending on which interface the light is exiting the medium, the light refracted back into the surrounding environment will be considered as either reflected or transmitted. Overall, the balance of energy can be expressed as follow:

$$\begin{pmatrix} entering \\ light \end{pmatrix} = \begin{pmatrix} reflected \\ light \end{pmatrix} + \begin{pmatrix} absorbed \\ light \end{pmatrix} + \begin{pmatrix} emitted \\ light \end{pmatrix} + \begin{pmatrix} scattered \\ light \end{pmatrix} + \begin{pmatrix} transmitted \\ light \end{pmatrix} \quad (16)$$

Several interesting observations can be inferred from this succinct analysis. Let us start by recalling some of the assumptions made previously. Transmission is not only irrelevant in the context of this study, but it was anyway deemed to be insignificant as the usage of white backing effectively renders hardcopies sufficiently opaque for transmission phenomenon to be safely ignored. On the other hand, the light that is reflected back towards observers' eyes is at the heart of this investigation. Two different sources for reflection were introduced in the previous paragraph. *Surface reflections*, previously introduced in section 7.3, constitutes the first form of light reflection, while the second corresponds to the light that has penetrated the medium but changed its trajectory due to

multiple scattering and finally exited. To pursue the structural inclination of the terminology, this type of reflection will be termed *body reflections*. The optical properties of an object, which account for a significant proportion of its appearance, are determined by the combination of the two reflections.

The intrinsic characteristics of these two reflections are vastly different. First and foremost, they vary in their geometrical properties. The geometrical distributions of the *surface reflections* vary greatly, from rectilinear if the surface is optically smooth to approximately hemispherical if the surface is very rough. On the other hand, the distribution corresponding to *body reflections* is mostly invariant. Indeed, *body reflections* originate from the light scattered and absorbed within the medium. Multiple scattering representing the most commonly encountered set of conditions, the macroscopic distribution resulting from the scattering process will be isotropic on average. The light bouncing back to the substrate-air interface will thus reach the surface under any possible degrees of inclination with equal chances. Hence the exiting light will also have an isotropic distribution. Overall, the geometric distribution of all the light reflected back from the medium could therefore be considered as the sum of two components, one invariant and isotropic, originating from *body reflections*, while the geometry of the second is fluctuating. Again, a terminological clarification is necessary. There is a conceptual different between *body reflections* and *diffuse reflections*. *Body reflections* correspond to light that has exited from the medium. They can often be characterised as being diffuse, but *diffuse reflections* cannot reciprocally be assumed to originate from *body reflections*. For instance, *surface reflections* may be diffuse depending on the characteristics of the surface.

Secondly, another source of dissimilarity between *surface* and *body reflections* is their disparate spectral composition. For the former, the light has little chance of interacting with the medium as it does not penetrate it. The spectral composition of *surface reflections* is therefore often considered to be the same as the incident light. This is however not strictly speaking true for all kind of medium, and this issue will be broached later on. On the other hand, this is not the case at all for *body reflections*. The energy of the incoming radiation will be altered as it interacts with matter. In the case of fluorescence, the absorption of a photon of a certain energy, typically in the ultraviolet band of the electromagnetic spectrum, triggers the emission of another photon with a lower energy level, hence a longer

wavelength, usually within the visible band. Fluorescence simply manifests itself as an increase in the spectral reflectance at certain wavelengths, an increase that may make the reflectance reach levels higher than 100%. It is only critical to take into account this phenomenon if illuminants with diverse ultraviolet contents are employed, in which cases different reflectance will be obtained for different lighting environments. Otherwise, it will implicitly be present in the reflectance of an object, but will not have any important impact on its appearance. As mentioned in section 7.4.2, this case is assumed in this study, and fluorescence will be omitted for the remainder of the discussion. On the other hand, the impact of the absorption phenomenon is usually much stronger and can be summarised in a simpler way. While the light travels through the medium and bounces back and forth from scattering, it is increasingly absorbed at wavelengths that are characteristic for the material. The physics of the phenomenon are on the contrary rather complex. *"Wave theory does not deal with colour due to selective absorption. This is usually considered in a quantum mechanical framework"*(Watt & Watt, 1992 p.48). Overall, the outcome of these two processes will greatly transform the spectral composition of the incoming radiation, hence *body reflections* will have a very different spectral composition from *surface reflections.*

To summarise, the light reflected back to observers' eyes therefore originates from two different processes, *surface reflections* and *body reflections.* Their characteristics differ in two fundamental aspects, their geometric distribution and their spectral composition. *"In describing appearance, wavelength (or spectral) variability is primarily responsible for color, while geometric (or directional) selectivity is primarily responsible for gloss, luster, translucency, and like attributes. However, geometric conditions not only affect geometric variables such as gloss and transparency, but also affect color, diffuse reflectance, and transmittance. Likewise spectral conditions can affect the measurement of geometric attributes of appearance. Therefore both the spectral and geometric conditions of measurement must be identified in specifying an appearance attribute of a specimen"* (ASTM, 2004c). However, the theoretical aspect of the proposed representation should not escape the vigilance of an inquisitive reader. What is presented above does not offer more than an abstract depiction of the functioning of all the processes involved. Actual materials differ greatly in their appearance, which points to significant discrepancies in their optical behaviour. The substrate sample considered for the purpose of this analysis was after all

assumed to be translucent and coloured, which, as can be intuitively envisaged, is representative of only a small percentage of all existing materials. The following section will present in more details the principal types of materials as well as the characteristics of their interactions with light.

7.6 Material classes

"An important criterion for classifying the appearance of materials is the electric conductivity of the material. This distinction separates dielectric materials from non-dielectric (conducting) materials, such as metals"(Klinker, 1993 p.30). Although this classification scheme does not include all existing materials, it will nevertheless provide an adequate differentiation between most common materials. More comprehensive schemes based on diverse criteria also exist (Askeland, 2005).

7.6.1 Electrical conductors

Conductors are materials that contain movable charges of electricity. While most conductors are metallic, there are many non-metallic conductors as well, including all plasmas and semiconductors. The light reflection from conductors depends mainly on a single reflection process. *"An impinging electromagnetic wave can stimulate the motion of electrons near the surface, which in turns lead to re-emission (reflection) of a wave. There is little depth penetration* [...]. *As a result, reflection from a metal occurs essentially at the surface"*(Cook & Torrance, 1981). The result is that, for metallic objects, only reflections originating at their interface with a surrounding medium can be seen. *"It turns out that electromagnetic waves cannot propagate very far through a conducting medium before they are either absorbed or reflected"*(Fitzpatrick, 1999). *"Thus internal reflections are not present to contribute to a diffuse component"*(Cook & Torrance, 1982). This strong absorption is also coupled with strong surface reflectance. *"Metallic objects have a very high surface reflectance. At normal incidence, for example, silver and aluminium reflect over 90% of all visible light* [...]. *The high reflectance power of metals is related to their ability to conduct electricity"*(Klinker, 1993 p.30). However, this strong reflectance is not always wavelength independent. *"For many metals, there are sufficient free electrons to result in nearly equal reflection of all wavelengths. This results in the reflective grayish colors typical of steel, aluminium or nickel. Reflective characteristics of other metals are affected by electrons that*

are bound to the atoms of the metal. This lack of freedom results in selective absorption of light at frequencies where the oscillation of the electron is damped. This results in the color of metals such as gold, copper or bronze"(Hall, 1989 p.159). As an illustration, the spectral reflectance of diverse metals with polished surfaces is given in Figure 48.

7.6.2 Electrical insulators

Electrical insulators, also known as dielectrics, are substances that are highly resistant to electric current. Dielectric materials can be solids, liquids, or gases, as well as vacuums. The following discussion will obviously focus on solid dielectrics, but it is worth remembering that air is also an electrical insulator. Unlike metals, the light reflected by dielectrics does not emanate only from surface reflections, because their optical properties possess characteristics resembling closely those of the abstract medium presented in section 7.5. Most of the light impinging on a dielectric actually penetrates the medium, as *"dielectrics generally reflect only about 4% directly from the material surface"*(Klinker, 1993 p.30). The veracity of the figure should be subject to caution, as the exact proportion is obviously dependent on the precise nature of the dielectric and also the incidence of the incoming light, as can be envisaged from Figure 49 and Figure 42. Light refraction is therefore a prominent phenomenon in dielectrics, but it will not necessarily give rise to body reflections. Indeed, dielectrics exhibit a wide range of optical properties, which requires refining the classification scheme in order to provide a comprehensive account of their diversity. We will adopt the categorisation strategy established by Klinker, whereby the uniformity of a dielectric constitutes the classification criterion. Uniformity is defined relatively to the wavelength of visible light. *"Dielectrics are considered to be non-uniform if the size of the particles is large compared to the wavelengths of light"*(Klinker, 1993 p.31). This distinction is similar in essence to the one presented in section 7.1, whereby the theoretical framework employed to describe the behaviour of surface reflections is determined by the scale of the surface roughness compared to light wavelengths. Material homogeneity is another criterion that is often employed to classify dielectrics. It can however be misleading, as an inhomogeneous material with particles smaller than the wavelengths of light behaves in a uniform manner, although it still has to be considered inhomogeneous. In absolute terms, the definition of uniformity is also not entirely immune of criticisms, but it was nevertheless deemed appropriate for the needs of this analysis.

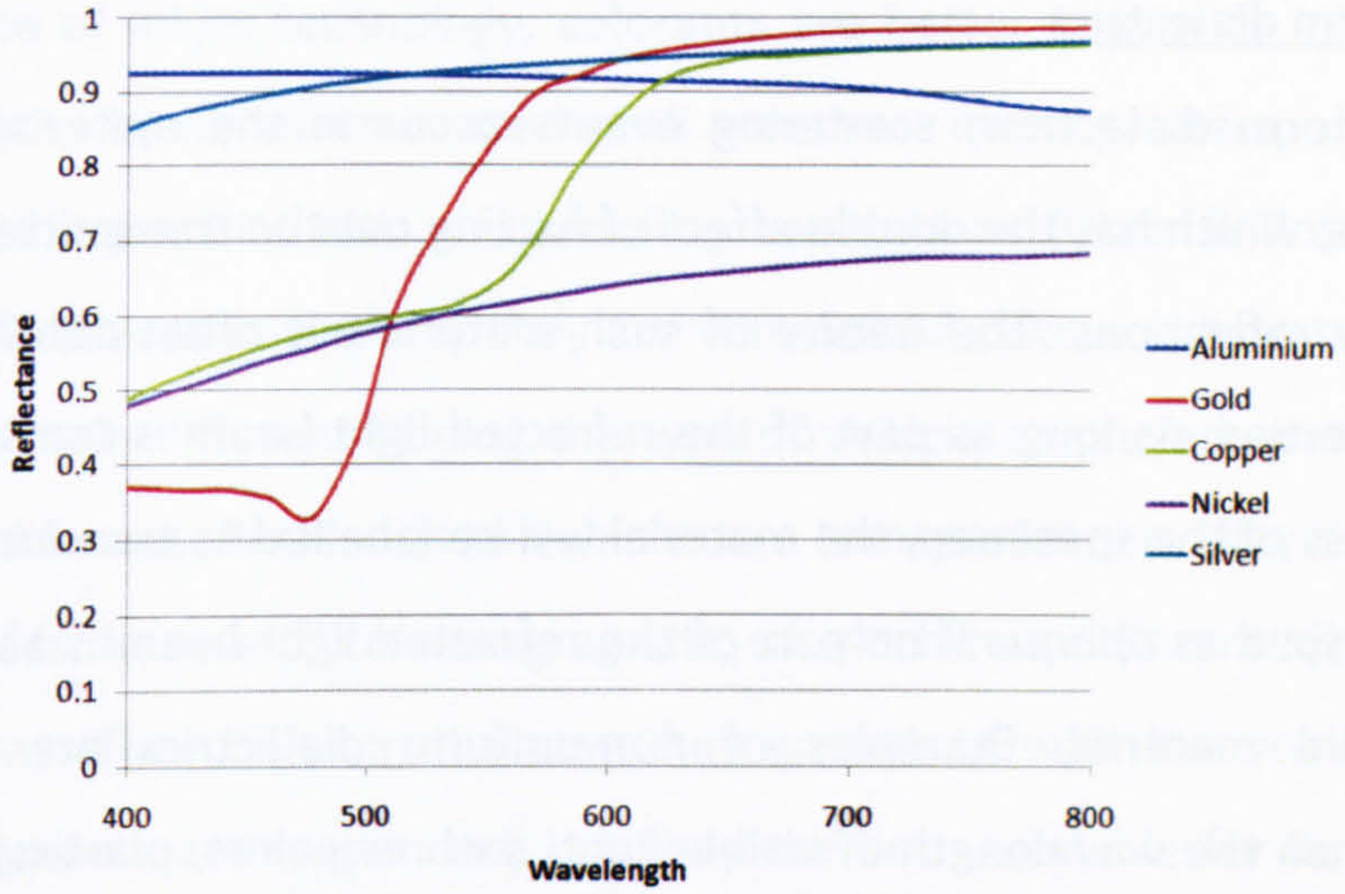


Figure 48 - Spectral reflectance of diverse polished metals under normal incidence at 30˚C.

7.6.2.1 Uniform dielectrics

"Since uniform materials do not contain any particles large enough to scatter light, no light beam is reflected back to the material surface from where the light entered. As a result, no body reflection exists for such materials. Any light entering the material body is either absorbed or exits at the opposite side"(Klinker, 1993 p.31). The correct idiom to describe such materials is *transparent* (ASTM, 2003a). Examples of uniform dielectrics include glass, clear water, perfect gemstones and some crystals such as quartz.

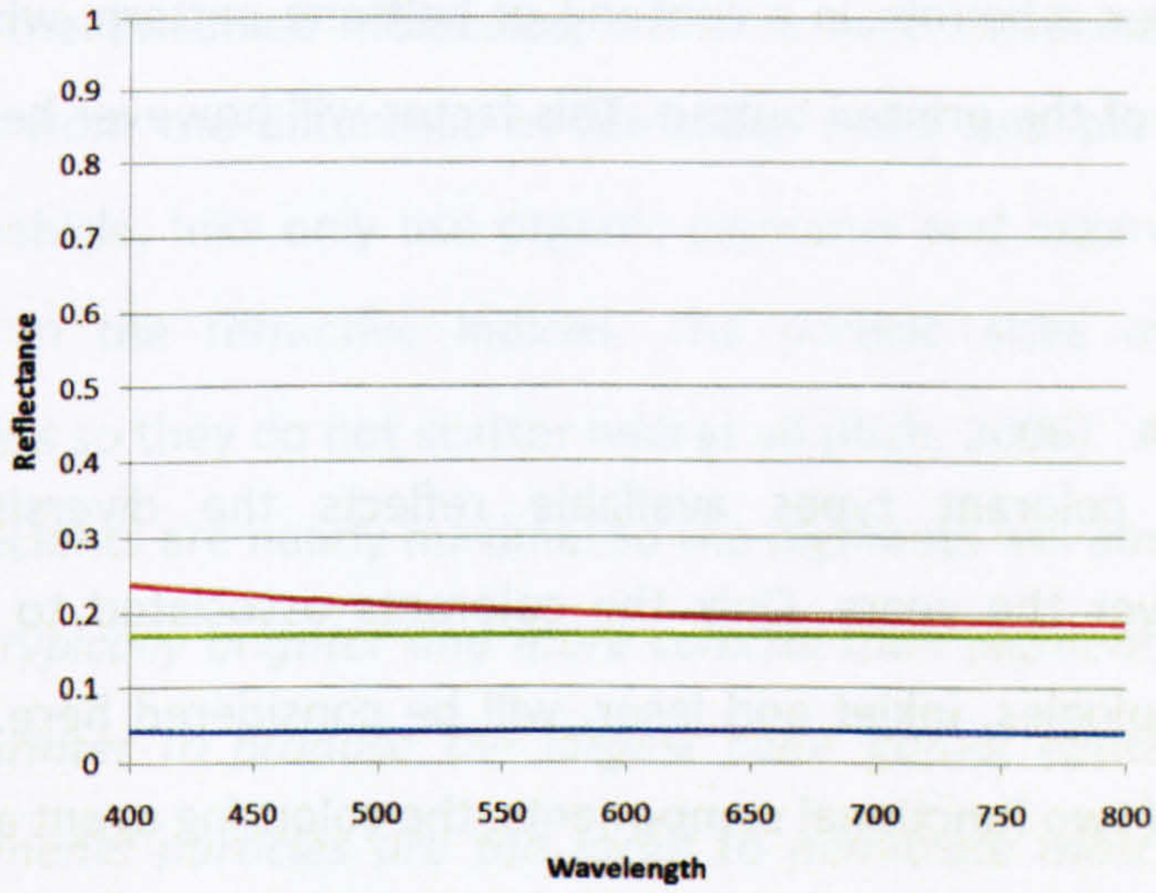


Figure 49 - Spectral reflectance of diverse polished dielectrics under normal incidence at 30˚C.

7.6.2.2 Non-uniform dielectrics

Unlike uniform dielectrics, scattering events occur in the material body of non-uniform dielectrics, which has the double effect of fuzzing out the transmitted light and also engendering body reflections. The nature of such materials is often determined by their transmissive properties. As long as part of the refracted light beam is transmitted through the whole thickness of the specimen, the material will be labelled as *translucent*. Otherwise, a material is described as *opaque* if no part of the refracted light beam is able to penetrate through the entire material. Examples of non-uniform dielectrics are materials with pigments larger than the wavelength of visible light, such as paints, plastics, ceramics dyed textiles and paper.

7.6.3 Paper and colorants

The discussion so far has been conducted at a very general level. It will be re-focused in this section on the optical properties of the materials at the heart of this study. The act of printing consists, in the modern sense, of applying a colorant onto a substrate, most commonly paper, in order to reproduce documents or photos. Prints are therefore composed of heterogeneous materials, each of them is a non-uniform dielectric, and their optical properties will be introduced further. Another important component taking part in the determination of optical properties is the colorants' geometrical layout. Colorants are typically laid down on paper substrate in a contone or halftone pattern, which has critical impact on the appearance of the printed output. This factor will however be left out of the conversation for now.

7.6.3.1 Colorants

The multitude of colorant types available reflects the diversity of printing technologies developed over the years. Only the colorants associated to the two most widespread printing technologies, inkjet and laser, will be considered here. Colorants are almost always composed of two functional components: the colouring agent and the vehicle. The role of the vehicle consists in delivering the colouring agent to the paper, while the colouring agent is responsible for selective absorption at certain wavelengths.

In the case of inkjet technology, colorants are better known as inks. They come in two different forms, liquid, which is used in all consumer inkjet printers, and solid, which is usually reserved for graphics arts and proofing applications (Ohta & Rosen, 2006 p.133). Despite major differences in the way they are transferred onto a substrate, both types are fairly similar, the main difference being the type of vehicle employed, water for liquid inks versus wax for solid ones. Both sorts of vehicles are transparent, i.e. they do not absorb nor scatter light. That is the role of colouring agents, of which two different types are available, dyes and pigments. They differ in the way they are mixed with the vehicle, i.e. dyes are dissolved and pigments are suspended, but they are both available for liquid or solid inks. *"Dyes are molecules designed to absorb specific frequencies of light, and they are chemically dissolved in the ink vehicle. With dimensions on the order of a nanometer, dye molecules are small enough to penetrate into any absorbent medium"*(Ohta & Rosen, 2006 p.135). On the other hand, "[p]*igments are particles about 50 to 150 nanometers in diameter composed of tens of thousands of dye molecules bound together. Unlike dyes, pigments are not dissolved in the ink vehicle: they are formulated with dispersing agents to create a stable suspension"*(Ohta & Rosen, 2006 p.135). It would be reasonable to presume their optical properties to differ, especially their scattering abilities, because of the differences in dimension. Unexpectingly, this is not the case, as scattering is very low, if not absent, in both instances. In the cases of dyes, the only possibility for scattering would be Rayleigh scattering from the dissolved molecules, which is very small indeed. For pigments, the scattering arises from the difference in refractive index and particle size of the pigment relative to the vehicle. Inks only use organic pigments and organic resins so there is very little difference in the refractive indices. The particle sizes of ink pigments are 150 nanometers or less so they do not scatter well at all (Rich, 2006). At these particle sizes, the absorption coefficients are nearly maximal so the pigments will absorb almost as strongly as dyes. *"Dyes are typically brighter and more colorful than pigments, and this usually allows the dye-based printer to produce the largest color gamut"*(Ohta & Rosen, 2006 p.135). However, "[p]*igments particles are too large to penetrate most inkjet printers or photo media. Instead, they form a thin film on the surface* "(Ohta & Rosen, 2006 p.136), which is advantageous when printing on paper because the pigment stays on the surface of the paper. This is desirable because more ink on the surface of the paper means less ink needs to be used to create the same intensity of colour. However, this also tends to increase the

roughness of prints' surfaces, hence generating diffuse surface reflections that may decrease the overall colourfulness.

In the case of laser printing, inks are replaced by toners, which are simply charge-responsive dye or pigment colorants. In essence, most toners can be thought as powder inks, although some come in liquid form (Ohta & Rosen, 2006 p.180). However, unlike inks, extra components are also present to provide the magnetic or electrostatic properties required by the electrophotographic process. *"In general, toner colorants are pigments or dyestuffs* [...] *that are coated with usually a plastic or resinous binder and dispersed in a suitable mineral spirit with a very minor proportion* [...] *of charge control or directing agents which tend to direct and accurately control good deposition and lay-down of the colorant during latent image development and fixing"*(Bradley, 1977). Since toner particles are typically few microns in diameter (Ohta & Rosen, 2006 p.181), they do not generate diffraction phenomenon but they do scatter light. However, when printed and fused, the toner forms a coloured polymer film that does not scatter much. Regarding their colour-rendering abilities, the previous discussion on the merits of dyes versus pigments also applies here, apart from the differences due to the depth of colorant penetration. Indeed, the nature of the electrophotographic process consists in fusing the toner particles at the surface of the substrate, hence dyes and pigments will be located at the same position. The vehicle, termed carrier in this case to denote its electric charge, is designed to be transparent, as for inks.

The characteristics of all types of colorant are therefore very similar. Be it inks or toners, pigments- or dyes-based, their intrinsic optical properties can fundamentally be considered as equivalent. Roughly speaking, the role of the vehicle can simply be reduced to delivering colorant to the paper, and is hence formulated as a transparent medium. The absorption properties of the colorant are regulated by the colouring agents, i.e. dissolved dyes of suspended pigments. The sum of the diverse discrepancies obviously has a practical impact on colorants' colour-rendering abilities, but they can generally be described in the same manner. Colorants are transparent medium that absorb some parts of the incoming light without scattering any of it.

7.6.3.2 Paper

The substrate plays a key role in the printing process. Many types of substrate are available, such as plastics, textiles and ceramics, but paper is most commonly employed. Paper is a thin material produced by the amalgamation of wood pulp fibers whose natural tint is yellowish (Kirk & Othmer, 1978 p.768-802). Its pleasing whiteness is obtained by chemical processing during the manufacturing process, where the pulp first undergo a bleaching process often followed by the embedment of fluorescent optical brighteners within the fiber weave. By nature, pulp fibers are transparent, which may seem peculiar at first since paper as a whole can be considered as fairly opaque. However, a morphological analysis of paper reveals the presence of multiple air gaps between fibers (Weisskopf, 1968 p.68). This can be observed from the microscopic cross-section (Figure 50 (E)) of one of the plain paper used in this study. These air interstices provide numerous interfaces where Fresnel reflections may occur. *"The scattering of light is thus accounted for by applying the Fresnel laws to the cell wall/air interface"*(Schramm & Meyer, 1998). Therefore, because of fibers' random shape and orientation, an incident light beam will be reflected and refracted in many directions. Absorption occurs when light travels through fiber cell walls, and can be accounted for by the *Beer-Lambert-Bouguer* law. The light that penetrates the fibers is again partially reflected at the inside surfaces, and after several reflections and refractions, it comes back to the eye of the observer from various directions. The fibers act as macroscopic non-uniformities, and are therefore responsible for paper's scattering abilities. The veracity of this analysis can be verified in a simple and entertaining manner. By pouring water on a sheet of raw paper, the air gaps will be filled with water, thereby reducing the differences in refractive index between paper and interstices, and the opacity of the paper will be greatly reduced.

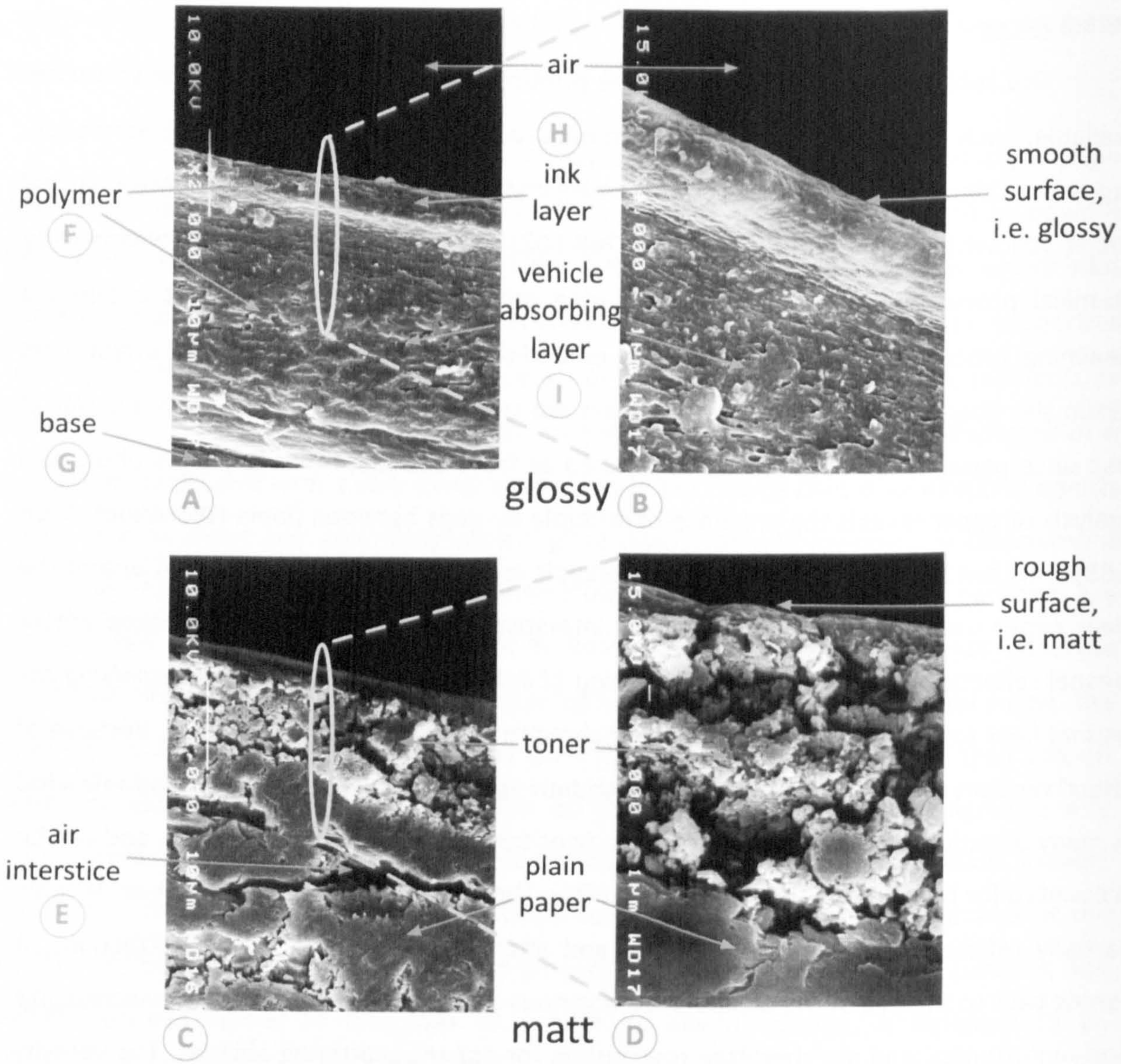


Figure 50 - Microscopic cross-sections of two different types of printed paper used in the simulation presented in the next chapters.

Plain paper is by far the most commonly employed paper-based substrate, but for higher quality printing, ink-receptive coatings (Figure 50 (F)) are added to the paper surface in order to improve its imaging properties. *"Because untreated paper can allow excessive penetration of ink, coatings are applied to hold the colorant near the surface to achieve high color saturation"*(Ohta & Rosen, 2006 p.140). In such cases, wood pulp (Figure 50 (G)) is only present to fulfill the mechanical requirements and for providing a white and scattering background. For the highest printing quality, it may even be replaced by a bright-white photobase or white polymer film. A cross-section of a coated substrate is shown in Figure 50 (A). The base is made of either plain paper or a white photobase (Figure 50 (G)) depending

on the substrate quality. The top two layers represent the ink-receptive coatings which *"control the movement of the ink vehicle on the surface of the print medium affecting color bleed, area-fill uniformity, and dot gain"*(Ohta & Rosen, 2006 p.140). They are made of transparent polymer compounds, which are qualified as porous or encapsulating depending on the way they handle inks. The role of the first (Figure 50 (H)) consists in absorbing and anchoring the colorant near the surface, which is advantageous because more ink on the surface of the paper means less ink needs to be used to create the same intensity of colour. The second polymer layer (Figure 50 (I)) aims at facilitating the quick absorption of the binder to avoid ink spreading. For similar reasons, the polymer layers are also isolated from the paper bulk by an impermeable barrier in order to keep the ink from migrating further along the fibers (Emmel, 1998 p.12). In the case of a non-paper based base, such as for white photobase, this is not required, as the polymers forming the base are non-absorbent.

It is important to note that the mechanical and chemical formation presented here only offers an approximate representation of actual printing media. There exist great variations between different manufacturers. The arrangement previously introduced is obviously optimised for inkjet technology. *"Each major inkjet printer manufacturer develops its own ink formulations, and because inkjet print quality depends on the print medium, inks and ink-receptive coatings are generally designed together for best performance"*(Ohta & Rosen, 2006 p.138). The printing technology employed also greatly influences the requirements placed on media. While inkjet printing put a lot of constraints on the chemical properties of the medium, laser printing offer a more flexible approach, as a toner layer is fused onto a medium by thermal energy and mechanical pressure. Once the medium is cooled to room temperature, the toner image will be fixed onto the medium. Therefore, *"*[l]*aser color printers should be able to handle any kind of medium"*(Ohta & Rosen, 2006 p.189). Regardless of the technology used or manufacturers' specificities, the underlying principles remain however the same, and so do substrate optical properties. The polymers forming the ink-receptive coatings are transparent and non-absorbent, while the base strongly scatters light in all directions.

7.6.3.3 Print output

Printing on paper can essentially be summarized as adding another transparent layer on top of a white substrate. This is obviously not the case for inkjet printing, as inks will penetrate the substrate and blend with the polymer layers or paper bulk, but this is a reasonable approximation for laser printing. The end result is however equivalent. From an optical point of view, any printed sheet of paper is the sum of two components: a transparent layer absorbing pre-determined parts of the visible spectrum sitting on top of a highly reflecting and scattering white base. The optical properties of a hardcopy therefore closely resembles that of the abstract substrate presented in the synopsis of the physics of the interactions between light and matter. However, significant differences arise from the heterogeneous nature of prints. As for the abstract substrate, *surface reflections* account for only a small fraction of the reflected light, since hardcopies' constituents are dielectrics. The characteristics of these *surface reflections* will also not be uniform across the whole surface. It is indeed trivial to appreciate that they will depend on both the halftone pattern used as well as the precise location of impact. On the other hand, the properties of *body reflections* will be essentially the same. Despite evident inhomogeneity and structural differences, the scattering process, under the conditions exhibited by prints, is probabilistic in essence and gives rise to an isotropic distribution that will nullify these differences. Printed paper therefore exhibit all the characteristics of a non-uniform translucent dielectric, i.e. isotropic *body reflections* with spectral composition varying across the surface and *surface reflections* whose goniometric distributions also vary across the surface.

7.7 Goniometric representation of light-matter interactions

Since issues related to transmission and texture are excluded from this study, *surface* and *body reflections* are the optical properties that determine a print's appearance. More precisely, appearance is governed by their respective goniometric distribution and spectral composition, as was previously mentioned. Precisely replicating this appearance would imply designing an accurate model of the interactions between light and matter. However, the previous sections clearly highlighted the fact that all these interactions occur at a nanoscopic scale. Attempting to model the behaviour of every individual photon and molecules would be absolutely impractical. The problem can be reduced to more tractable proportions by adopting a larger scale approach. Indeed, we have seen that scattering

processes can be described in a probabilistic framework, and the resulting distribution is isotropic when considered at a macroscopic scale. Similarly, obtaining a faithful representation of surface's relief is extremely difficult, but its reflective properties may be very well reproduced at a macroscopic level by considering a theoretical surface that has the same statistical description. It is therefore perfectly possible to accurately reproduce an object's appearance by providing a mathematical description of these optical characteristics, rather than carry out a detailed simulation of the physics involved. The appearance of an object can thus be approximated by only considering the global outcome of these interactions rather than each of them individually. Such a mathematical representation is provided by the bidirectional reflectance-transmittance distribution function (BRTDF), also known as the bidirectional scattering distribution function (BSDF) (Glassner, 1995 p.664). The BRTDF describes the way light is transmitted and reflected at an interface between materials with disparate optical properties. Since most materials are opaque, a special case of this function, the bidirectional reflectance distribution function (BRDF), is more widely known. The BRDF represents the scattering of light occurring in the hemisphere above the surface of a material, whereas in the case of a BRTDF or BSDF, the full sphere of scatter directions is of interest. Transmission being omitted in the context of this study, only the BRDF will be considered in the remainder of this thesis, and its principal characteristics will now be introduced.

7.7.1 Prerequisite

7.7.1.1 Geometrical notations

The geometric setting of a BRDF requires the specification of three-dimensional angles relative to a differential patch of the surface instigating the scattering event. Since the patch is differential in size, it can be reasonably approximated as a planar surface possessing a well-defined surface normal. Together, the normal and the patch determine the orientation of the hemisphere of directions the BRDF describes. A light source subtending a solid angle $d\omega_I$ casts luminance on this patch. The reflected light propagating in the volume subtending a solid angle $d\omega_R$ is under consideration. The four unit vectors required to represent the principal elements involved in the definition of a BRDF are illustrated in the following Figure 51:

- N: normal to the surface.
- I: incident light direction. θ_I and φ_I correspond respectively to the elevation and azimuth angles of the incident direction.
- R: scattering direction. θ_R and φ_R correspond respectively to the elevation and azimuth angles of the scattering direction.
- F: bisecting angle of I and R. The angle between F and N is denoted α. The angle between I and F is denoted ψ. The azimuth angle of F is denoted φ_F.

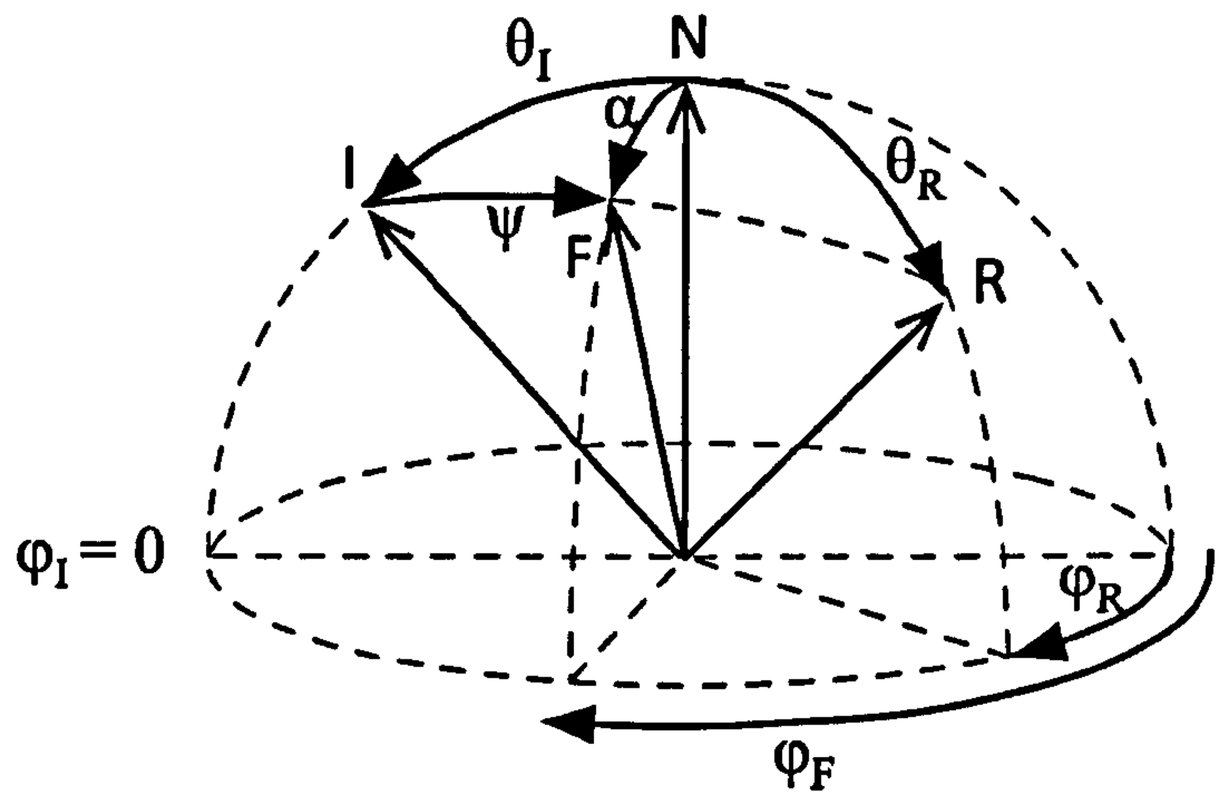


Figure 51 – Geometric notations.

7.7.1.2 Radiometric terms

The definition of the BRDF makes use of several radiometric quantities, such as radiance and irradiance, which will be assumed to be known and understood in the remainder of this study. Further information may be found in Glassner's book (1995 p.647).

7.7.2 Bidirectional reflectance distribution function

"The bidirectional reflectance distribution function describes how light incident on a surface is reflected into a continuum of directions. It is defined as the ratio of the differential reflected radiance leaving [the point of impact] *in direction* [R] *and the differential irradiance arriving from* [I]"(Kautz, 2003) at a given frequency of light. In simpler terms, it models what ratio of light for some incoming direction I leaves in some outgoing direction R, and is most often defined as:

$$\rho(\vartheta_I,\varphi_I,\vartheta_R,\varphi_R,\lambda) = \frac{L(\vartheta_I,\varphi_I,\vartheta_R,\varphi_R,\lambda)}{E(\vartheta_I,\varphi_I,\lambda)\cos\vartheta_I d\omega_I} \quad (17)$$

where L is the reflected radiance (also called radiant exitance), E is the incoming irradiance, λ represents the wavelength of light, and $\cos\vartheta_I d\omega_I$ indicates the angle-dependent nature of the BRDF. The cosine-term introduces the reduction in flux intercepted by the differential surface patch as the angle between the normal and the polar component of the light direction increases. The solid angle indicates the amount of area on the surface of the hemisphere through which the inbound flux is intercepted.

The BRDF is therefore a five-dimensional function of four angles (or two vectors) and wavelength. This definition can also be extended for materials exhibiting heterogeneous BRDFs, i.e. where the BRDF varies with the position of impact on the surface. The BRDF of such materials is qualified as *spatially varying* or *shift-variant*. The polarization of the incident light may also be considered, which brings the tally of parameters to eight for the most general expression of the BRDF. Despite this high-dimensionality, it is based on few assumptions besides the ones already mentioned in this study. *"First, it assumes that light exits the surface at the same location where it hit the surface"* (Kautz, 2003). This is obviously not the case for papers, but it is nevertheless a reasonable approximation to do. While some light absorbed from a point A may exist at another point B, the scattering process is isotropic, which results in an energy balance across the whole surface. In other words, the lack of reflection in point A will on average be compensated by light absorbed at point B and exited at point A. Secondly, "[t]*he definition of the BRDF further implies that light is reflected immediately, i.e. light is not stored for some time and then re-emitted later (phosphorescence). Finally, the BRDF cannot model materials that change the frequency of the incident light (fluorescence)*" (Kautz, 2003). This assumption has already been dealt with in great details in section 7.2. Despite the generality of the definition, not all functions are acceptable as BRDF because they have to comply with certain rules. *"The BRDF has to fulfil two important properties to be physically plausible. First it needs to be energy conserving, i.e. no more energy must be reflected than is received"* (Kautz, 2003). As noted by Ward, this is a recurrent problem in this field of study. *"The fact that normalization was not adequately treated in He's otherwise impeccable derivation shows just how much normalization is overlooked and undervalued in reflectance modelling"* (Ward, 1992). Another important

property that BRDFs must respect is the reciprocity principle commonly encountered in electromagnetism in general and optics in particular. *"Furthermore, it must obey the Helmholtz reciprocity, meaning that the BRDF must be symmetric in* $[R]$ *and* $[I]$" (Kautz, 2003). More simply stated, *"results* [should] *not* [be] *changed if the geometries of incident and viewing beams are interchanged"* (ASTM, 2004c). Finally, as for the energy conservation issue, there are many widespread misconception concerning BRDFs' properties. *"First, a BRDF is not bounded to the range [0,1] – a common misconception about BRDFs. Although the ratio* $[L]$ *to* $[E]$ *must be in [0,1], the division by the cosine term in the denominator implies that a BRDF may have values larger than 1. Secondly, a BRDF is not a unit-less function. Since the BRDF definition above includes a division by the solid angle (which has units steradians (sr)), the units of a BRDF are inverse steradians (*sr^{-1}*)"* (Wynn, 2004).

7.7.3 Dichromatic reflection models

The BRDF provides a very precise and complete description of light-matter dynamics. However, its high-dimensionality generates many difficulties for the acquisition of data as well as their processing. Even if advances in technology simplified the capturing process, the sheer amount of raw data forbids their direct use. Instead, researchers exploited the high redundancy present to reduce the BRDF equation to simpler mathematical expressions. As mentioned numerous times, the reflected light originates from two different processes, *surface reflections* and *body reflections*, and so the BRDF may also be divided into two elements. *"The bidirectional reflectance may be split into two components, specular and diffuse. The specular component represents light that is reflected from the surface of the material. The diffuse component originates from internal scattering (in which the incident light penetrates beneath the surface of the material)"* (Cook & Torrance, 1982). All modern reflection models, qualified as *dichromatic* to reflect their strategy, adopted this approach to provide fast and accurate approximations to BRDFs. The following discussion will provide a concise introduction to the most popular ones. As an indication of their dichotomous nature, the discussion will successively introduce the diverse strategies adopted for the *surface* and *body reflections* rather than describing each model's characteristics successively.

7.7.3.1 Surface reflections

Three different approaches may essentially be found in literature for the simulation of *surface reflections*. Of very little research interest is the simulated approach. In such models, the characteristics of the specular lobe are not obtained from physics laws, but derived from generic mathematical functions chosen according to certain criteria. *"These models begin with the objective of reproducing a certain appearance phenomenon such as plastic or rolled metal. They are typically developed in the context of realistic image generation, and they often trade physical accuracy for computational efficiency"* (Walker, 1999). The best-known model in that category was derived by Phong, who proposed to model the specular highlight with a cosine lobe. Although it is physically implausible, its simplicity and the satisfactory images it produces have led to its widespread use in modern computer graphics software. However, it can only approximate the reflectance of a plastic-like surface, which explains the lack of realism often encountered in the early days of computer graphics (Cook & Torrance, 1981).

Another category consists of the empirical approach. *"Starting from a set of measured data, these models attempt to find a representation that best fits the values"* (Walker, 1999). These models focus essentially on finding a suitable mathematical representation of the specular lobe, or alternatively on generating an efficient data structure to store the large amount of data present (Cabral, Max, & Springmeyer, 1987; Westin, Arvo, & Torrance, 1992). Among the many published models, Ward (1992) proposed a model based on Gaussian lobes, Lafortune *et al.* (1997) on cosinusoidal functions, Schroeder and Sweldens (1995) on spherical wavelets while Koenderik *et al.* (1996) attempted to reconstruct BRDFs by adopting a decomposition approach based on Zernike polynomials. Despite being popular thanks to their implementation ease, the universality of these models is not guaranteed since they are not based on the laws of physics.

On the other hand, analytical models attempts to be as true to the surface physics as possible by starting from first principles, most often derived from geometric or physical optics. Their structures follow exactly the premises established in section 7.3. To summarize, the proportion of light reflected at the surface is determined by a Fresnel term, and the spread generated by the surface roughness is simulated by assuming that the surface is made of randomly orientated microfacets that behave like mirrors. Another term is also

included to represent the shadowing and masking occurring between adjacent microfacets. The model originally derived by Cook and Torrance (1981) is a prime example of such an approach. It is by far the most widely known and used, but diverse closely related models also exist. They only diverge in the probability distribution representing the orientation of the microfacets population. While Cook and Torrance chose the distribution originally derived by Beckmann, Blinn (1977) employed a standard normal distribution, and Obein (2000) a Lorentz distribution. While these methods are all based on a geometric optics approach, there were also few attempts to generate simulations for a wider range of materials and surface finishes by taking into account physical optics phenomena. Most notably, the model presented by He *et al.* (1991) provides means to simulate diffraction scattering by the surface roughness. Similarly, Oren and Nayar (1994) proposed a model for approximating diffraction scattering from the shadowing and masking term and inter-reflections. A more general model, which accounts for polarization, Fresnel, and shadowing/masking effects, has been described by Bahar (1987). Stogryn (1967) applied a more general form of the Kirchhoff theory, thus taking polarization effects and the correct dependency of the Fresnel reflectivity into account. More complete descriptions of these models may be found in Walker (1999), Kautz (2003), Dempski (2004), and Glassner (1995).

7.7.3.2 Body reflections

While there is an abundance of strategies for surface reflections, only a single approach is generally used for body reflections. All dichromatic reflection models follow the principles established in the first ever reflectance model proposed by Lambert (Walker, 1999), which assumes the scattering to be constant in all exitant directions. In other words, light falling on such a surface is scattered so that the apparent brightness of the surface to an observer is the same regardless of the observer's angle of view. A material exhibiting such a behaviour is referred as an ideal diffuser (or a Lambertian surface). Only a few isolated materials come close to approximating an ideal diffuser, such as unfinished wood, concrete, or even skin, but no natural material is truly Lambertian. Indeed, to be qualified as Lambertian, *all* the reflections from a surface should be equal in any direction, surface reflections included. But to isotropic surface reflections correspond very specific surface relief, which are unlikely to occur or remain due to natural wear and tear. Advanced dichromatic reflection models do not make such an assumption. Instead, they presume that

only body reflections are Lambertian, which seems at first to be a reasonable conjecture since we have seen that the scattering emanating from body reflections is isotropic for densely packed non-uniformities, which represents the most common case.

While this looks to be an undisputable assumption due to the hundreds of articles broaching the issue (Dempski & Viale, 2004; Glassner, 1995), there are few evidence disseminated in the literature that this is not so. Firstly, Cook and Torrance remarked that "*the reflectance varies only slightly for incidence angles within about 70° of the surface normal*" (Cook & Torrance, 1981), which clearly implies a departure from an ideal diffuser after that threshold. While they remain vague on the exact nature of the departure, a clearer representation is provided by Orchard. "*Even with isotropic scattering, the true emergent flux per unit solid angle is not uniform*" (Orchard, 1968). Orchard even supplies a figure of the complete scattering distribution, which is reproduced here (Figure 52 (A)). While his article focuses on the optics of pigments in the context of the painting industry, the remarks made remain entirely valid in our context, as the underlying principles are actually remarkably similar. Indeed, the scattering distribution we measured (Figure 52 (B)) correspond almost exactly to what Orchard reported, which confirms both facts, first that paint and printed paper behave similarly from an optical point of view, and second that body reflections are not Lambertian. However, the precise reasons at the origin of this behaviour are even harder to pinpoint. Despite our best efforts, no conclusive references discussing the topic were brought to light.

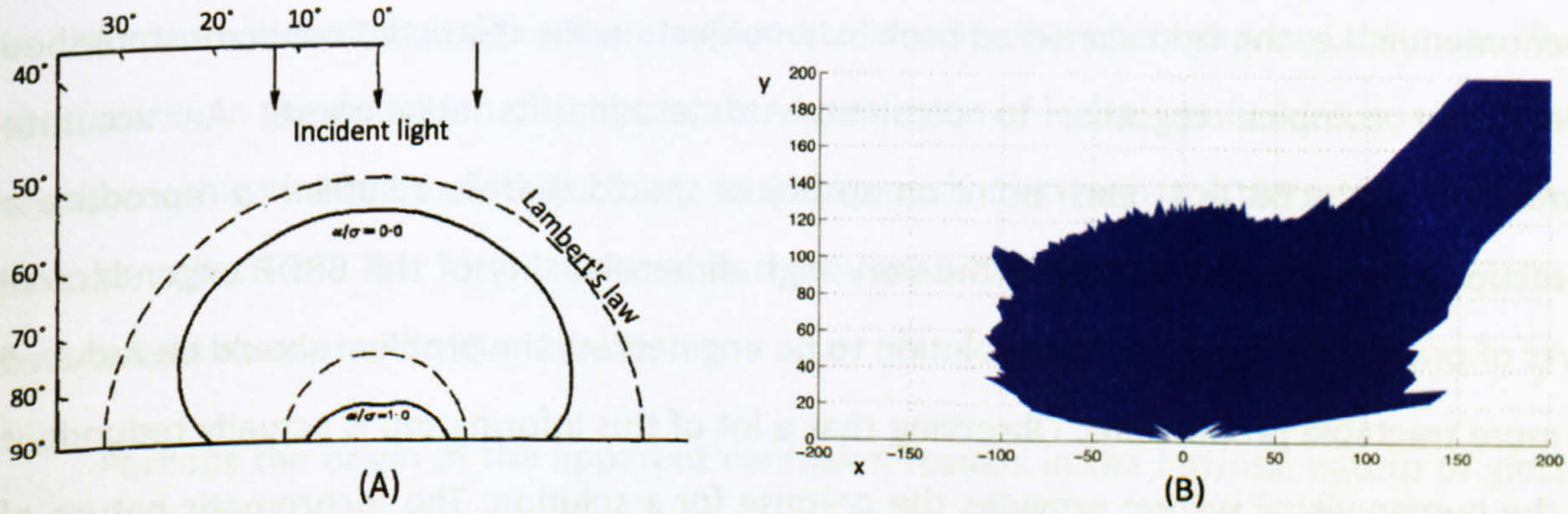


Figure 52 - Directional distribution of light reflected obtained by Orchard (A) and measured on glossy specimens (B).

However, some obvious explanations immediately come to mind. First of all is the total internal reflection angle. *"If light travels from a medium with high index of refraction into one with smaller index of refraction, Snell's law shows that for an angle of incidence [*superior*] to a certain critical angle, all the light is reflected internally"* (Glassner, 1995). Since reflected light emerges from paper and colorants into the air, this will clearly happen in our case. The critical angle is entirely determined by the refractive indices of the two media under consideration. The refractive indices of the diverse papers and colorants used in this study were found to be comparable to the refractive index of most materials, i.e. 1.5 (Dempski & Viale, 2004). The corresponding critical angle is about 70°, which concur extremely well with the observations made in the previous paragraph. Indeed, if surface roughness is taken into account, it is trivial to imagine the effect on the distribution of reflections, i.e. a progressive reduction past the critical angle. A second possible explanation also involves the surface relief. As can be envisioned from Figure 44, the path of light is perturbed the most under grazing angles, where the influence of masking and shadowing becomes maximal, which entails the same consequences as in the previous case, i.e. a progressive reduction as the angle of the reflected light approaches grazing angles.

7.8 From physics to perception

The light reflected back to observers' eyes originates from two different processes, *surface reflections* and *body reflections*. From the detailed descriptions previously given, any normally educated human being would intuitively identify these phenomena to the notions of colour and gloss. Colour and gloss are therefore two manifestations of the same phenomenon, i.e. the light scattered back by an object. In the restricted context established here, they combine together to completely determine its appearance. An accurate simulation of the BRDF at each point on an object should therefore suffice to reproduce a hardcopy's appearance. However, the very high dimensionality of the BRDF engenders all sorts of practical difficulties. For a solution to be engineered, the problem should be reduced to more tractable proportions. Observing that a lot of this information is actually redundant to the human visual system provides the premise for a solution. The trichromatic nature of the human visual system signifies that only three quantities need to be considered to replicate an object's colour. While this is not strictly speaking true, as was discussed in great length in the first part of this thesis, the approximation is nevertheless sufficient for the

purpose of the current argument. Similarly, body reflections, from which the perception of colour originates, exhibit an approximately isotropic goniometric distribution. Therefore, if only colour was to be considered, the corresponding BRDF can be simplified to the extreme. However, for an object's appearance to be accurately simulated, the whole BRDF needs to be appropriately reproduced, hence body *and* surface reflections need to be precisely simulated. While the characteristics of colour have been previously introduced in great details, the brief description of gloss that was given in the first part cannot suffice to generate a conclusive opinion on the requirements needed to generate an accurate simulation of gloss. The remainder of this chapter will thus be devoted to provide an in-depth introduction to the perceptual phenomenon referred to as gloss.

7.8.1 Gloss

Gloss is not an optical property of materials, but rather a perceptual attribute of objects, as indicated by the definition changes published in the latest version of the vocabulary approved by the CIE (1987). Gloss is the psychophysical counterpart of surface reflections, and can be approximately correlated to the amount of incident light that is reflected at the specular reflectance angle of the mean of a surface. Unlike colour, it is more elusive in nature, as proven by the cryptic definition given by the ASTM: "*angular selectivity of reflectance, involving surface-reflected light, responsible for the degree to which reflected highlights or images of objects may be seen as superimposed on a surface*" (ASTM, 2003a) or also "[g]*loss is associated with the capacity of a surface to reflect more light in some directions than in others*" (ASTM, 2004a). Perhaps a more intuitive explanation would be to give the common verbal qualification employed to identify this phenomenon, *shininess*. The lack of clarity in the definition often generates a profusion of lexical confusions, to which we remedy in this thesis by distinguishing unambiguously between the terms "*gloss*" and "*specular reflection*". The former is used here to mean the visual attribute, and the latter refers to light measured instrumentally.

Perhaps the origin of the apparent confusion resides in the intrinsic nature of gloss. Unlike colour, the perception of gloss is not unique, but rather gloss only reveals its multiple facets when the conditions of observation are altered. As Hunter (Hunter, 1937, 1975)

initially observed in the seminal work he performed in the 1930's, there are at least six different visual phenomena related to apparent gloss:

- *Specular gloss: perceived brightness associated with the specular reflection from a surface.* This essentially corresponds to the perceived brightness of a highlight. Since the human visual system is not very good at estimating absolute levels of luminance, the specular gloss is often determined by comparing a specimen to a reference sample.

- *Sheen: perceived shininess at grazing angles.* The concept of sheen is essentially equivalent to specular gloss, the only difference being the geometrical configuration of measurement. Sheen is often used to assess the gloss of surfaces that appears matt under other standard viewing angles.

- *Contrast gloss: perceived relative brightness of specularly and diffusely reflecting areas for the same specimen.* It applies when the relation between the specular and the diffuse component of reflection is of the same order of magnitude, hence for surfaces with a very low level of gloss such as textiles.

- *Distinctness-of-image: perceived sharpness of images reflected in a surface.* For specimens exhibiting a high level of gloss, the gloss can be evaluated by estimating the sharpness of the reflected image of the illuminant. The light source should obviously present high spatial contrast, such as sharp borders for instance.

- *Haze: perceived cloudiness in reflections near the specular direction.* As for distinctness-of-image gloss, haze is only observed with very glossy samples. Haze usually appears in an area right next to the specular direction, and is identified by the presence of a cloudy impression next to the illuminant reflection

- *Absence-of-texture gloss: perceived surface smoothness and uniformity.* This is usually caused by surface non-uniformities that result in reflection variation in a fine scale. Strictly speaking, this is not really gloss, but an effect due to texture.

7.8.2 Gloss dimensionality

The diversity of gloss manifestations is symptomatic of the complexity of the phenomenon. As previously mentioned, the BRDF is a high-dimensional function, and Hunter's observations suggest that gloss exploits all of these dimensions. Most of the published articles on the subject indeed largely concur with this intuition, apart from one paper written by Billmeyer and O'Donnell (1987). While investigating the perceptual dimensions of gloss, a multidimensional analysis of the visual judgements collected on a limited series of samples lead them to conclude that a single dimension was sufficient to explain observers' judgements. Again, this conclusion goes against intuition as well as the bulk of the literature published on the topic. *"It usually requires more than one measurement to identify properly the glossy appearance of any finish"* (ASTM, 2004d). "[I]*t is well known that the gloss phenomenon involves more than can be represented by a single number. Additional significant information about gloss is contained in the bi-directional reflectance distribution function (BRDF) measured by detecting light reflected at angles beyond the equal/opposite specular angle*" (Arney, Ye, & Wible, 2006). A more precise quantification of the gloss phenomenon was more recently performed by Pellacini *et al.* (2000). Using the power of computer graphics to their advantage, they concluded that gloss has two independent dimensions. The first one correlates well with the apparent contrast of the reflected image, which is qualitatively similar to Hunter's *contrast gloss*, while the second correlates with the apparent sharpness or distinctness of the reflected, which again is similar to Hunter's *distinctness-of-image* attribute. Earlier studies by Hunter and Judd (1939) and Harrison and Poulter (1951) corroborated this fact and demonstrated that gloss perception depends not only on the quantity of light reflected in the specular direction, but also on the width of the specular peak. These results however need to be mitigated by scrutinizing their range of validity. "[W]*e want to make clear that strictly speaking, the model we've developed only accurately predicts appearance within the range of glossy paints we studied, under the viewing conditions we used.* [...] *we fully expect that for other materials and under other conditions different gloss attributes such as sheen and haze may play a greater role*"(Pellacini, Ferwerda, & Greenberg, 2000). Similarly, the results obtained by Hunter and Judd are relevant to paints only, while printed papers were the focus of Harrison and Poulter. Therefore, while a general description of gloss may require more than two dimensions, the perception of gloss resulting from paints or printed papers can, according to

these studies, be successfully represented by only two dimensions. They should therefore be perfectly applicable to the materials currently studied, as previously explained.

Chapter 8: Empirical analysis

The goal sought in this second part is to better replicate the overall appearance of hardcopies, but the discussion is kept on general terms to maximise the universality of the findings. The perceptual attributes of interest in the context of this study were found to originate from an object's optical properties, and more precisely from the scattering of light in the hemisphere above an object's surface. A complete and convenient mathematical description of this scattering is given by the BRDF, which may be divided into two components, as was previously introduced. To summarise, the light reflected back to observers' eyes therefore originates from two different processes, *surface reflections* and *body reflections*. Their characteristics differ in two fundamental aspects, their geometric distribution and their spectral composition. While body reflections are approximately isotropic but generally vary in their spectral composition across an object's surface, surface reflections exhibit essentially the spectral composition as the incoming light's, but their goniometric distribution is greatly influenced by the roughness of the object's surface. As a matter of fact, previous research concluded the perceptual attributes engendered by surface reflections were two-dimensional, with the two dimensions representing the geometrical configuration of the specular lobe. Furthermore, colour is also widely assumed by the major part of the research community to be three-dimensional, whereby tristimulus values are the determinant factor influencing the perception induced by body reflections.

On the other hand, the BRDF possesses eight dimensions in its most general form. On more realistic terms, six of these dimensions are important, the sixth being its spatial variation. While the ultimate goal is to reproduce the perceptual phenomenon induced by reflection phenomena, our primary target is to replicate their physical characteristics. If we omit colour for an instant, limiting ourselves to two dimensions may not be sufficient to successfully replicate the specular lobe. After all, the previously cited articles did not try to model it nor the BRDF, but simply to establish the dimensionality of the visual attribute termed gloss. We will therefore have to perform our own analysis to establish how many and which physical characteristics of the BRDF are pertinent to our study. A large set of print samples will be generated and evaluated, but first, the diverse recommended methods to

measure physical quantities, BRDFs, surface and body reflections, as well as their psychophysical counterparts, colour and gloss, will be presented first.

8.1 Metrology

The instrument employed to measure a material's BRDF is termed *spectrogoniometer. "Goniophotometry is a general procedure for evaluating the manner in which an object or a material geometrically distributes the unabsorbed portion of incident light"* (ASTM, 2004b). The traditional design consists of two rotating arms, one for the light source and the other for the photometric sensor. A whole BRDF may be obtained by rotating each arm in turns so that all possible angular combinations are covered. While this provides a complete picture of a material's BRDF, the process is extremely laborious. Fortunately, cleverer designs exist, such as the one developed in the spectrogoniometer used in this study, which will be introduced later.

Another way to reduce the workload involved in measuring BRDFs is to take advantage of the high redundancy present in them. Indeed, as previously introduced, the body reflections of the materials studied here are essentially isotropic for most part of the hemisphere of reflectance. This is not true at grazing angles, and may not be true for other materials. But even in such cases, the rate of changes as a function of incidence angle is generally low in frequency. The extent of body reflections obviously depends on the width of the specular lobe. The glossier the material, the more redundancy will be present. Generating an estimate of a large portion of the BRDF can thus be inferred from a small number of samples. This method is frequently used in the industry for the measurement of metallic, pearlescent and special effect paint finishes. *"The variation in color, as a function of geometry, is usually measured by spectrophotometry with several specified sets of geometric conditions. Measurement of this kind, at a few selected angles, is called "multiangle spectrophotometry," as distinguished from measurement over a broad range of angles, which is called goniospectrophotometry"* (ASTM, 2004e). Two different 3-angle illumination/1-angle receiving geometric conditions are commonly used: the 25°/45°/72.5 illumination, 45° receiving geometry defined by DIN (2001), and the 15°/45°/110° illumination, 45° receiving geometry advocated by Dupont (E. I. du Pont de Nemours & Co.). Estimates of the most important attributes, i.e. colour and gloss, of simple BRDFs are thus

easily captured thanks to multi-angle spectrophotometers. They might however not be sufficient to recover precisely the whole BRDF though, as they perform a single measurement of the specular lobe.

Other instruments may also provide estimates of the BRDF. However, they were designed with a more specific intention in mind, and do not consider the whole hemisphere of reflectance, but rather focus on specific characteristics of the BRDF. For instance, colour-measuring instruments such as spectrophotomers concentrate only on body reflections at a very specific set of angles. The gamut of instruments dedicated to the measurements of colour, i.e. body reflections, has been introduced at length in section 4.2, and no further details will be given here. The other important trait of BRDFs, specular reflections, is handled by a class of instruments designated under the generic umbrella term of glossmeters. A traditional glossmeter provides gloss ratings that are obtained by comparing the specular reflectance from a specimen to that from a black glass standard. Three different geometries are recommended (ISO, 2004) depending on the shininess of the specimen:

- *"The 60° geometry is used for intercomparing most specimens and for determining when the 20° geometry may be more applicable.*
- *The 20° geometry is advantageous for comparing specimens having 60° gloss values higher than 70.*
- *The 85° geometry is used for comparing specimens for sheen or near-grazing shininess. It is most frequently applied when specimens have 60° gloss values lower than 10" (ASTM, 2004a).*

The resulting gloss rating is generated thanks to a non-linear relationship that maps the photometric readings to visual gloss ranking obtained through psychophysics experiments. "*Measurements* [made by glossmeters] *correlate with visual observations of surface shininess made at roughly the corresponding angles*" (ASTM, 2004a). "*Based on the experiment done by Hunter, 60-degree specular gloss measurement ranging from 15 to 80 correlates well with human beings observation*" (Kuo, Ng, & Wang, 2002b). This method therefore provides an easy-to use and affordable solution to estimate the shininess of a material. However, a single measurement is generated, which is not sufficient to recover the

multi-dimensional aspect of the specular lobe. *"The measure of gloss is a simplification of the BRDF down to a single appearance related quantity"* (Westlund & Meyer, 2001). "[T]*here is no quantification of the angular distribution of the scattered light*" (Hansson, 2004).

The width of the specular lobe can however be captured thanks to hazemeters. Although Hunter's qualitative description has already been given, the more technical approach taken by ASTM provides a clearer definition of haze: *"for a specified specular angle, ratio of flux reflected at a specified angle (or angles) from the specular direction to the flux similarly reflected at the specular angle by a specified gloss standard"* (ASTM, 2004a). Haze is produced by irregularities in the reflecting surface that affect the distribution of flux reflected around the specular angle. The common verbal qualification employed to identify this phenomenon is *clarity*. An important attribute of haze is that it only is a phenomenon of high gloss surfaces. Indeed, the specular reflections originating from matt surfaces are widely spread across the whole hemisphere of reflectance. "[The measurement of haze] *is applicable to nonmetallic specimens having a 60° specular gloss value greater than 70° in accordance with Test Method D 523*" (ASTM, 1999). There are essentially two different methods for measuring haze. The most accurate one consists in comparing gloss reflectance factor measurements made at certain angles to similar measurements made at angles immediately adjacent to the original one. *"In Test Method A, gloss reflectance factor is measured at 30° to the specimen normal using narrow illuminator and receiver aperture angles (0.5° wide maximum).* [...] *Narrow-angle (2°) reflection haze is measured at 28 or 32° or both, and wide-angle (5°) reflection haze at 25 or 35° or both. The ratio of reflectance factors for 28 or 32° or both, perpendicular and parallel to the machine direction of the specimen, is computed as a measure of directionality. In Method B, specular gloss is measured at 20° according to Test Method D 523, and narrow-angle reflection haze is measured at 18.1 and 21.9°"* (ASTM, 2004d). A cheaper alternative consists in simply comparing two standard gloss measurements at different angles to obtain a haze rating. *"Measurements of 60° and 20° specular gloss are made on a specimen. The haze index is computed as the difference between the two measurements"* (ASTM, 1999).

8.1.1 Practical considerations

Several instruments are therefore available to obtain diverse types of information about the BRDF, from the entire hemisphere of reflectance to more specific aspects. The choice of instruments employed in this study needs to be dictated not only by the requirements imposed, but also by the practicality of the solution implemented. It would indeed be an immense waste of time if the proposed solution was completely unrealistic. For instance, while the use of spectrogoniometers makes perfect sense from a scientific point of view, envisaging their deployment in Graphic Arts studio is simply nonviable, due to their prohibitive cost as well as the productivity decrease yielded by the intense labour needed to operate them. *"Goniophotometer instruments available commercially and reported in the literature are expensive and complex devices compared to gloss meters and are not commonly used for routine paper characterization"* (Arney, Ye, & Wible, 2006). On the other hand, certain instruments, such as spectrophotometers, are already widespread, and are able to give information about a certain aspect of the BRDF, namely body reflections. Similarly, glossmeters are hazemeters, though not widely used yet, are inexpensive enough for their inclusion in Graphic Arts studios' worflow to be considered. They also provide information about another important aspect of the BRDF, surface reflections. However, whether their readings are sufficient to accurately recover the geometry of specular lobes remains to be established with certainty. Though the information presented previously suggested that two measurements were sufficient to explain the perception of gloss, there is no conclusive evidence that they would be sufficient, nor that these measurements would be the appropriate ones. Therefore, an empirical analysis will first be conducted to determine the true dimensionality of the BRDFs of a typical set of print samples employed in the Graphics Arts industry.

8.1.2 Spectrogoniometer

The BRDF measurements were obtained with *Eldim's EZcontrast* spectrogoniometer, a remarkable instrument that allows capturing a BRDF in a matter of seconds thanks to the combination of a Fourier lens and a CCD sensor. The working principle is actually rather simple, and is illustrated in Figure 53. The Fourier lens provides a Fourier transform image of the sample, i.e. *"every light beam reflected from the test area with a θ incident angle will be focused on the focal plane at the same azimuth and at a position $x=F(\theta)$"* (ELDIM). In

layman's term, this means that to each angle of incidence will correspond a certain area on the CCD. A two-dimensional map of the angular characteristics of the hemisphere of reflectance can thus be captured. The acquisition process is also effortless as there is no mechanical movements, and last but not least fast, since a whole hemisphere is captured in a single shot. The manufacturer specifications are given in Table 22. Of interest, it can measure the distribution of reflectance in almost the entire hemisphere, i.e. 360 degrees in azimuth and 80 degrees in elevation, with an angular resolution of 0.4 degrees. The light source is a filtered Xenon arc lamp generating a collimated light beam that approximates closely the spectral power distribution of the standard CIE D65 illuminant. By using five coloured filters, BRDFs in each of the CIE tristimulus values X, Y and Z colour bands can be captured. Unfortunately, very little information concerning this set of filters is available, apart from the fact that two of them target the green region of the visible spectrum, two others the blue part, and the last one the red part. Strictly speaking, this instrument does not measure BRDFs, but tristimulus values, hence reflected radiance indirectly. However, a simple normalization by the incoming irradiance can transform the output data into a full-fledged BRDF.

The measurement protocol strictly followed the guidelines established in section 4.2. There was unfortunately no time to fully evaluate the abilities of this apparatus, such as its absolute accuracy. However, routine maintenance had been performed shortly before the measurements phase, and its warm-up routine involved a black level calibration. Given the plausibility of the obtained results, it was therefore safe to assume that it was functioning correctly. Since all the measurements occurred within a three days period, its short-term repeatability was established by measuring multiple times each day a yellow patch printed on *DP-Euro offset gloss* paper. The presence of the specular lobe implied that the radiance distribution could not be assumed to be normal, hence using standard statistical methods was out of question. Furthermore, the strength of the noise present in body reflections also forbade the use of non-parametric test such as the Mann-Whitney U-test. A simple visual analysis unambiguously revealed the short-term repeatability to be more than sufficient, as informally indicated by the high correlation coefficient (> 0.99) obtained between different measurement sessions.

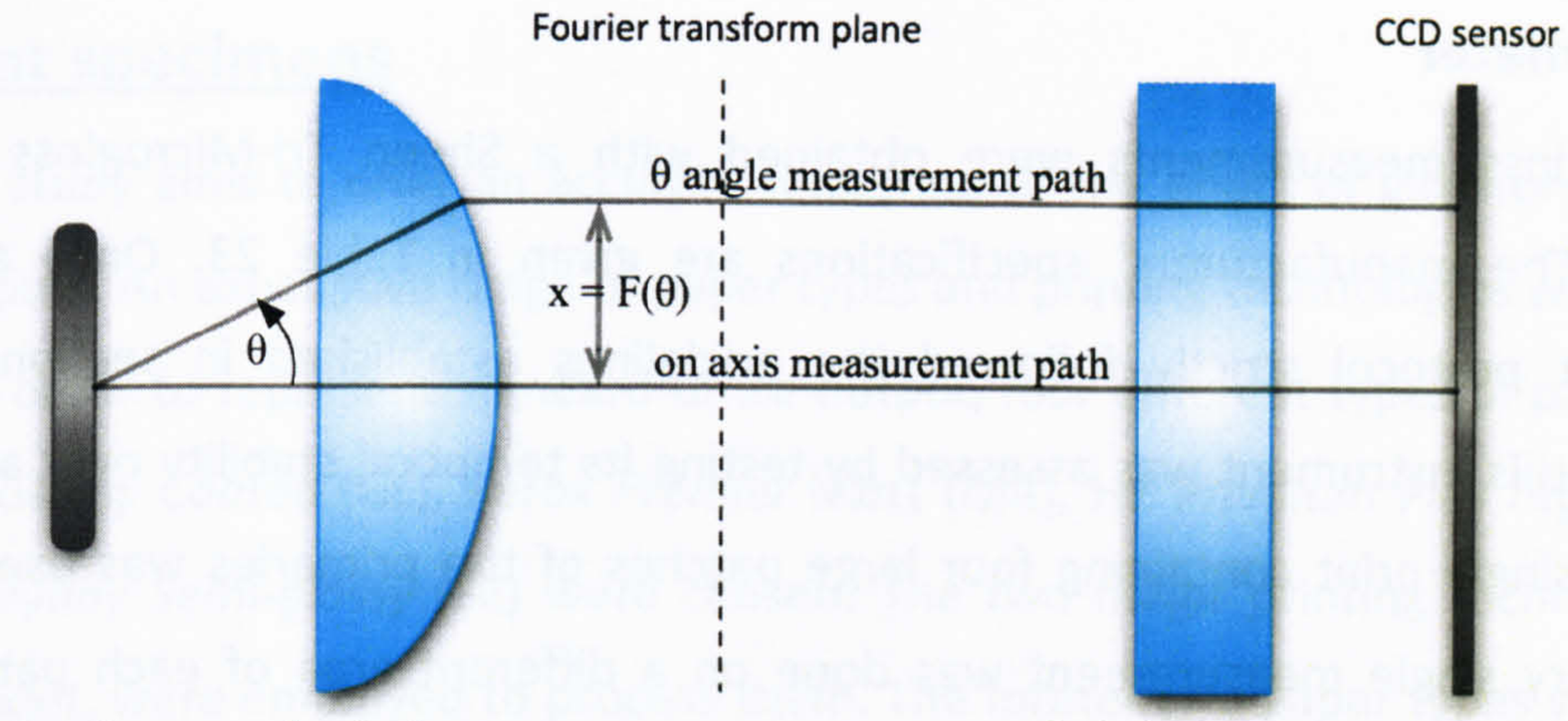


Figure 53 – Schematic of Eldim EZ-Contrast spectrogoniometer

Table 22 - Manufacturer specifications for Eldim's EZContrast spectrogoniometer.

Sensor configuration	Peltier cooled CCD Photopic response
Sensor resolution	6M
Field coverage - Incident angle - Azimuth angle	 +/- 80° 0/360°
Angular resolution	+/- 0.4°
Measuring area - Diameter min - Diameter max	 100µm 2mm
Luminance range - Min - Max	 0.001 $cd.m^{-2}$ 1 000 000 $cd.m^{-2}$
Accuracy - Angle - Luminance (cd/m2) - Chromaticity (xy,u'v') RMS	 < 0.15° +/- 3% 0.005 for any colour stimulus

8.1.3 Glossmeter

The gloss measurements were obtained with a Sheen Tri-Microgloss Plus 160 glossmeter. The manufacturers' specifications are given in Table 23. Once again, the measurement protocol strictly followed the guidelines established in section 4.2. The suitability of this instrument was assessed by testing its temporal stability over a period of five days. A single print containing four large patches of the primaries was used for that purpose. Every single measurement was done on a different area of each patch, hence spatial uniformity does also have an influence on the results. The standard deviations of the diverse measurements series are given in Table 24. The large majority of observed standard deviations are about or less than 1 gloss unit, which is consistent with the manufacturer's specifications, and the largest recorded difference, which is clearly an outlier, was four gloss units. These results were deemed perfectly acceptable.

Table 23 - Manufacturer specifications for Sheen Tri-Microgloss Plus 160 glossmeter.

Measurement geometry	20°, 60°, 85°
Range	0.1 – 1000 gloss units
Resolution	0.1 gloss unit
Accuracy	± 1gloss unit
Repeatability	< 0.5 gloss unit

Table 24 - Temporal repeatability of the Sheen Sheen Tri-Microgloss Plus 160 glossmeter. The results are expressed in gloss units.

	DP	HI	IS	SE	GL	MA	IM	NA
Cyan	0.12	0.06	0.14	0.25	0.06	0.00	0.04	0.09
Magenta	0.19	0.12	0.15	0.12	2.23	0.15	0.04	0.12
Yellow	0.09	0.29	0.16	0.29	0.10	0.06	0.07	0.03
Black	0.85	0.95	0.98	1.08	1.35	0.58	0.19	0.34

8.2 Print specimens

This study aims to offer an accurate rendering of the effect of gloss on all types of printed outputs. An exhaustive range of paper types and printing technologies was therefore selected. In order to represent standard office output, four different types of papers, Xerox *Colortech+ Glossy Coated* (GL), Xerox *Premier Matt* (MA), HP *Premium Plus High-gloss* (HI) and HP *Everyday Semi-glossy* (SE) were chosen. The two major printing technologies, i.e. laser and inkjet, were employed to process them. The former two paper types were printed on a Xerox *Phaser 7300*, while the latter two on a HP *Designjet 20PS*. The remaining paper/printer combinations were chosen to represent professional applications. In order to represent typical newspaper output, the *National Press* (NA) and *Improved Newsprint* (IM) papers types were printed on a standard press. Lastly, *Coated Satin* (IS) and *DP-Euro Offset gloss* (DP) were printed on best-of-class inkjet printers to represent state-of-the-art printer-based proofing systems, and by extension high-quality press.

For each of these paper types, BRDF measurements of the 16 combinations of the four primaries at full coverage were performed with an incident light at an elevation of 45 degrees, i.e. replicating the CIE 45:0 geometry. Additionally, measurements of the four primaries at 33% and 66% coverage were obtained. Cyan was also measured with an incident light at an elevation of 45 degrees, as for the first measurement, but with an azimuth of 90 degrees, in order to test for anisotropy. Finally, the cyan colorant was also measured when illuminated by incident lights at an elevation of 20 and 70 degrees, in order to sample approximately the available range of incident light. Measuring at angles larger than 90 degrees, such as the angle of 110 degrees commonly employed in the industry, do not make much sense here, as the isotropy of the samples entails that such measurements would simply mirror those obtained at 70 degrees. All of these print samples were also subjected to gloss measurements. Additionally, the gloss ratings for colour ramps of each primary were estimated. Lastly, colour characterization was derived for each paper from measurements of a standard IT8.7/3 chart. The measurement protocol strictly followed the guidelines established in section 4.2.

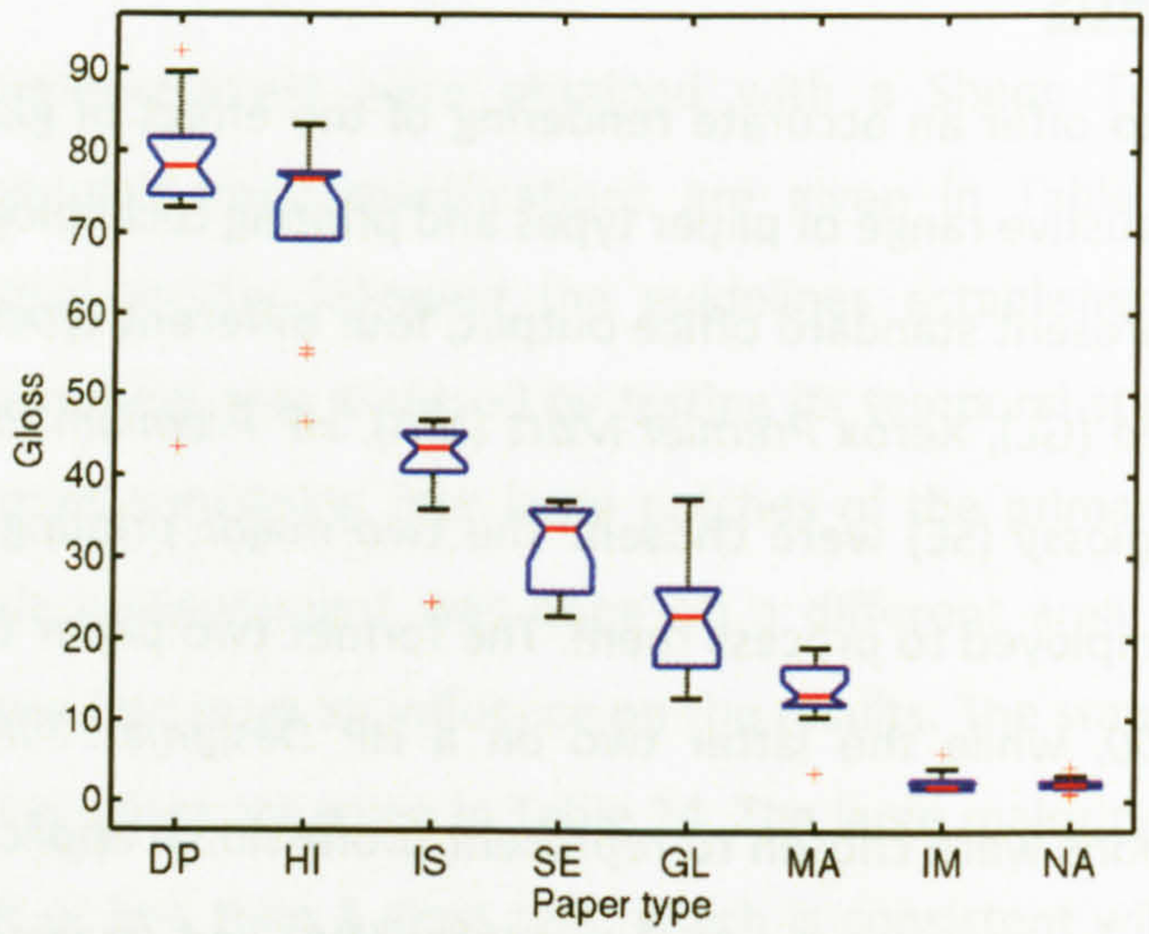


Figure 54 - Box-and-whisker plot representing the gloss ratings of all paper types for the 60° measurement geometry.

A summary of the gloss ratings recorded for each paper/colorant combinations is shown in the box-and-whisker plot in Figure 54. The choice of paper types samples the gloss range approximately evenly, from totally matt papers to very gloss prints. However, it is important to note that, despite what the figure shows, the gloss range does not have 100 as its upper limit. Indeed, a gloss rating is obtained by comparing the specular reflectance from a specimen to that from a black glass standard. Many materials generate a gloss rating higher than 100, electrical conductors for instance. While glossier materials exist, personal experience proved that it would be rather difficult to find a paper exhibiting as much gloss as *DP-Euro Offset gloss*, let alone more. Hence the choice of paper types was deemed to be suitable to effectively sample the entire gloss range of conventional printed papers.

8.3 Anisotropy

An important practical concept implied by the BRDF equation, but not covered until now, is anisotropy. In the context of this study, anisotropy corresponds to the lack of radial symmetry in a material's scattering of light. "*Materials that exhibit anisotropic scattering behavior often possess some type of surface microgeometry that interferes with the symmetric scattering of light. The prototypical examples of anisotropic materials are rolled metal and varnished wood, although sand, textiles, and even paint exhibit anisotropic scattering. Each of these materials might be characterized by directional ridges and valleys*

that prevent light from scattering independent of the azimuth angle" (Walker, 1999 p.6). Since paper is an irregular weave of cellulose fibers, there is a high chance that some of the chosen paper specimens exhibit anisotropic behaviour to a certain degree. For all paper types, measurements of the 60° gloss for the 16 primaries combinations were performed at four different azimuth angles: along the length of a standard sheet (which will be referred to as vertical), perpendicular to the length (horizontal), and also at 45° and 135° to the length. A t-test was performed on each combination of directions to test whether their means were statistically equal or not at the 0.01 significance level. The results are shown in Table 25. The boolean value false and true indicates whether the null hypothesis, i.e. equal means, can be respectively rejected or accepted.

Table 25 - Results of a t-test analysis performed on different measurement directions to reveal the eventual presence of anisotropy. The boolean value false and true indicates whether the null hypothesis can be respectively rejected or accepted.

	DP	HI	IS	SE	GL	MA	IM	NA
h-v	*T*	*F*	*T*	*F*	*F*	*F*	*F*	*F*
45-135	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*
h-45	*T*	*T*	*T*	*F*	*T*	*F*	*T*	*T*
v-135	*T*	*T*	*T*	*T*	*F*	*T*	*T*	*T*

Interestingly, the papers where a difference between horizontal and vertical measurements appears use cellulose fibers as their mechanical support, while those where no difference exists use white polymers. The presence of fibers thus does generate an anisotropic behaviour, as was expected. However, the magnitude of these differences is relatively minor, which makes sense since most of the affected papers are the least glossy. For most papers, the differences between horizontal and vertical measurements are between two and three gloss units, i.e. always less than 10% of their absolute gloss rating. On the other hand, there was no statistical difference between measurements made at 45° or 135°. In addition, the measurements performed at an angle generated gloss readings that were intermediary between the horizontal and vertical values. Moreover, these measurements exhibited in most cases no statistical differences with either the horizontal or the vertical measurements. Therefore, since the levels of gloss anisotropy were benign at most, especially compared to the differences between paper types, it was decided to ignore

these small discrepancies by performing every single gloss measurements at an angle of 45 degrees to the length in order to get the better of an already minor anisotropic behaviour. Given the task ahead, this is a reasonable approximation that could undoubtedly be lifted with further work.

8.4 BRDFs visualisation

Before delving into an in-depth analysis of the characteristics of the measured BRDFs, a quick primer will be given in this section to provide a general overview of the actual BRDFs obtained. Figure 55 represents the BRDFs of a yellow colour patch printed at full coverage on diverse types of paper. From left to right are shown specimens exhibiting decreasing levels of gloss. Each CIE X, Y and Z channels are displayed successively on the ordinate axis. For each paper types, the best-fitting data range was derived to optimise the amount of information displayed. Each column, i.e. paper type, thus has its own specific range, whose extrema are indicated on the colour bar next to each column. All the ranges used are linear.

The most striking feature of these series of plots concerns the magnitude of the specular lobe. By examining the data range's maximum value, it is obvious to conclude that the specular lobe varies non-linearly with the level of gloss. The specular lobe of the glossiest paper (*DP*) is indeed so large that the hemisphere of body reflections almost disappears from the plot. The existence of the body reflections for that paper is however confirmed by Figure 56, where the data range is set to the same values for all plots, in order to focus on body reflections alone. The differences between the specular lobes are also emphasized in these series of plots. In contrary to its magnitude, the width of the specular lobe increases as gloss ratings decrease, until it become so large and short that it blends in with body reflections. It also becomes clear to see that, unlike what Figure 55 suggested, body reflections are actually very similar for all papers. There are indeed very little variations between the different types of paper, which all exhibit large CIE X and Y radiance hemispheres coupled with almost non-existent CIE Z body reflections, which is typical of the reflectance pattern of a pure yellow print.

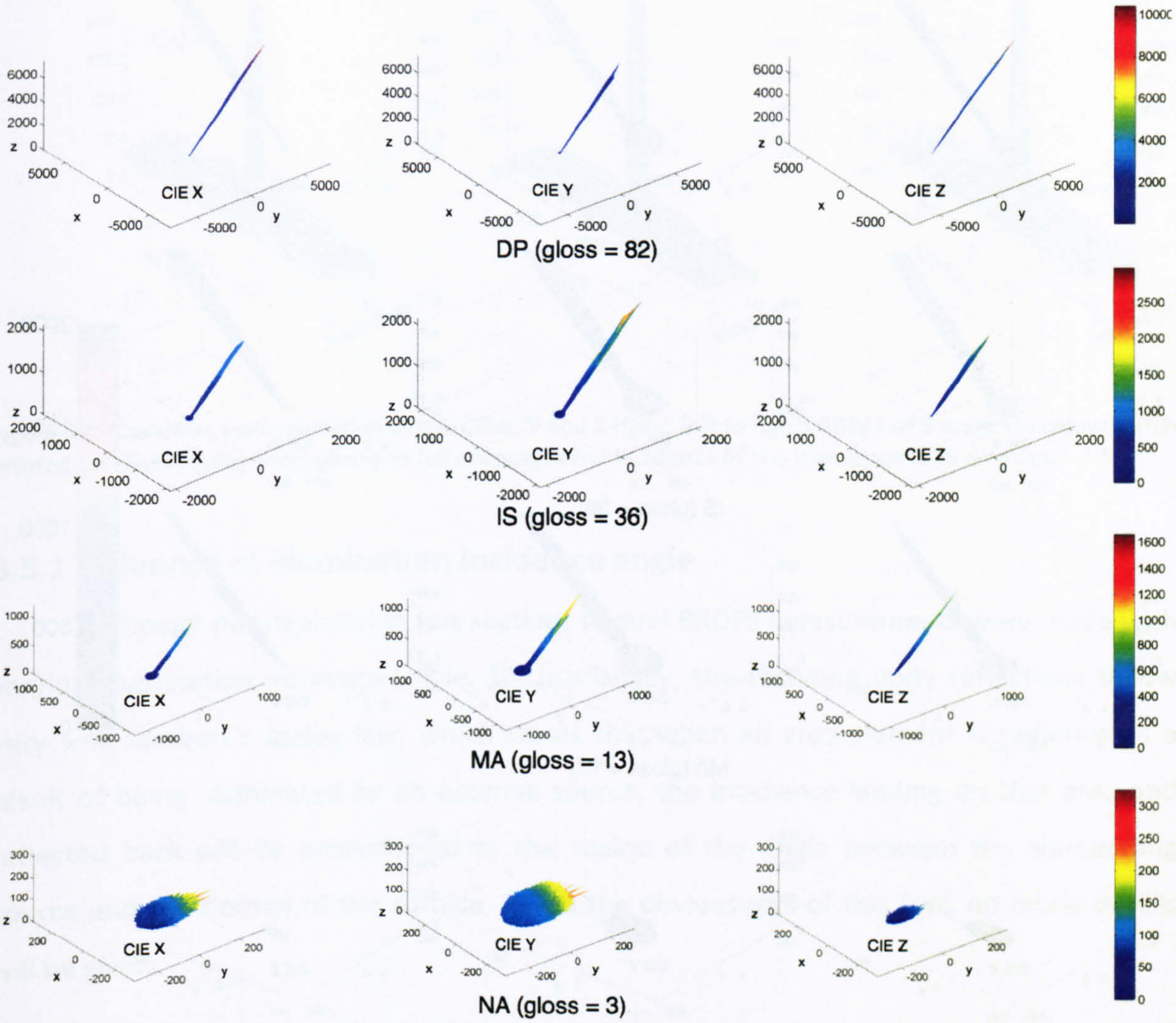


Figure 55 - Representation of the BRDFs of papers with gloss levels varying from very glossy (top) to totally mat (bottom). In all cases, the depicted radiance distributions correspond to a yellow colour patch (100 % coverage). In each column are represented (from left to right) the BRDFs of CIE X, Y and Z. The incidence of the illuminant is 45 degrees. The magnitude of the tristimulus values is indicated on the colour bar. The axes are denoted *x*, *y* and *z* to illustrate the orthonormality of the space. A more intuitive representation consists of considering the corresponding polar coordinates, whereby the polar angles represent the azimuth and the elevation, and the radius represent the magnitude of the tristimulus values.

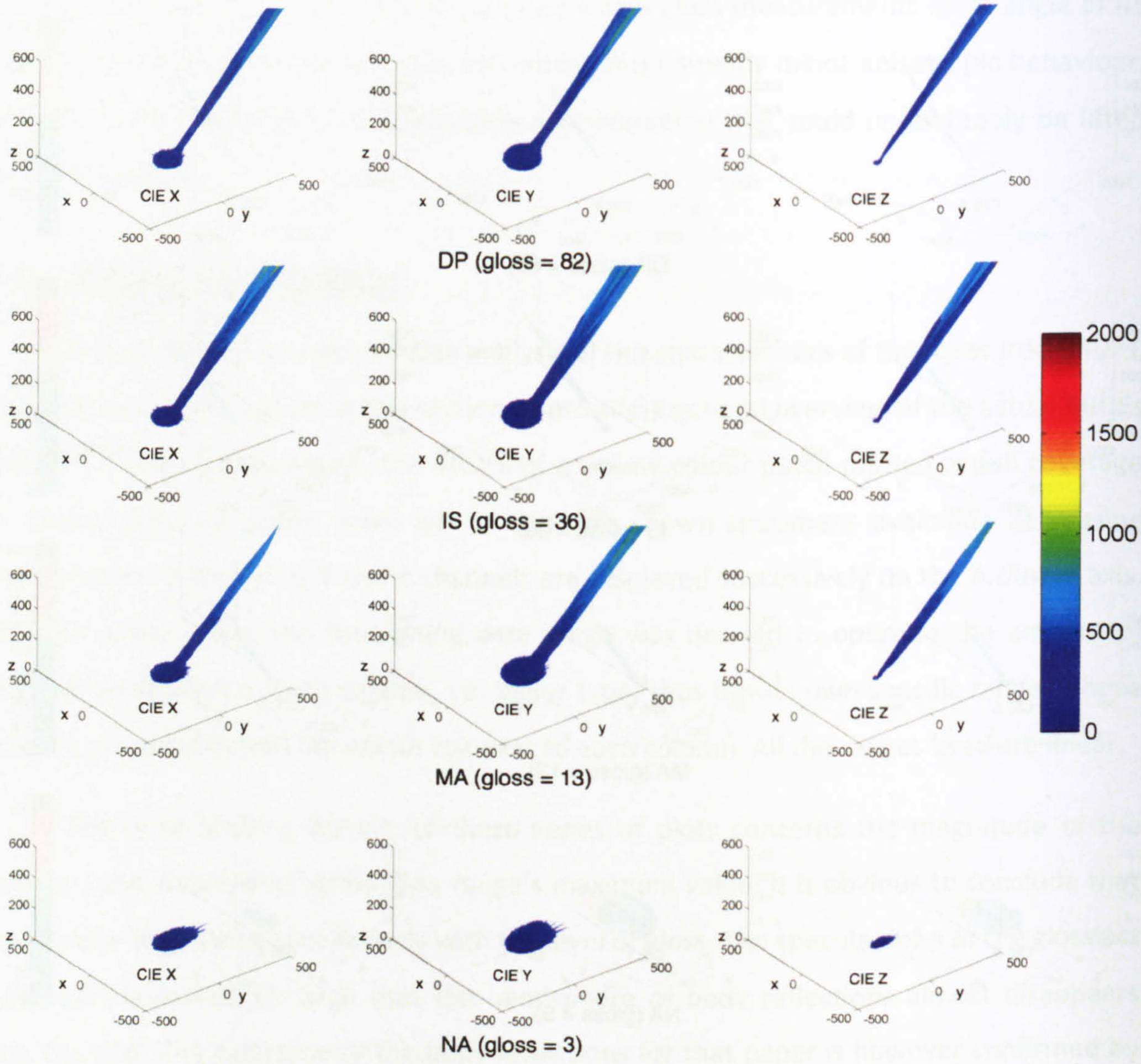


Figure 56 - Same as Figure 55, apart from the tristimulus range which has been set to the same quantities for all plots.

8.5 Characteristics of body reflections

We will now focus on the characteristics of body reflections. Rather than using the spherical plots introduced in the previous section, another data visualisation technique will be employed here, whereby the spherical coordinates are plotted as such in a Cartesian orthonormal coordinates system. This essentially "flattens" the hemisphere of reflections. As seen from the following figure, a much better assessment of body reflections can be obtained from such plots. While the plots only show the BRDFs of a magenta colour patch printed on HP *Everyday Semi-glossy* at full coverage, similar patterns can be observed for other papers and colours.

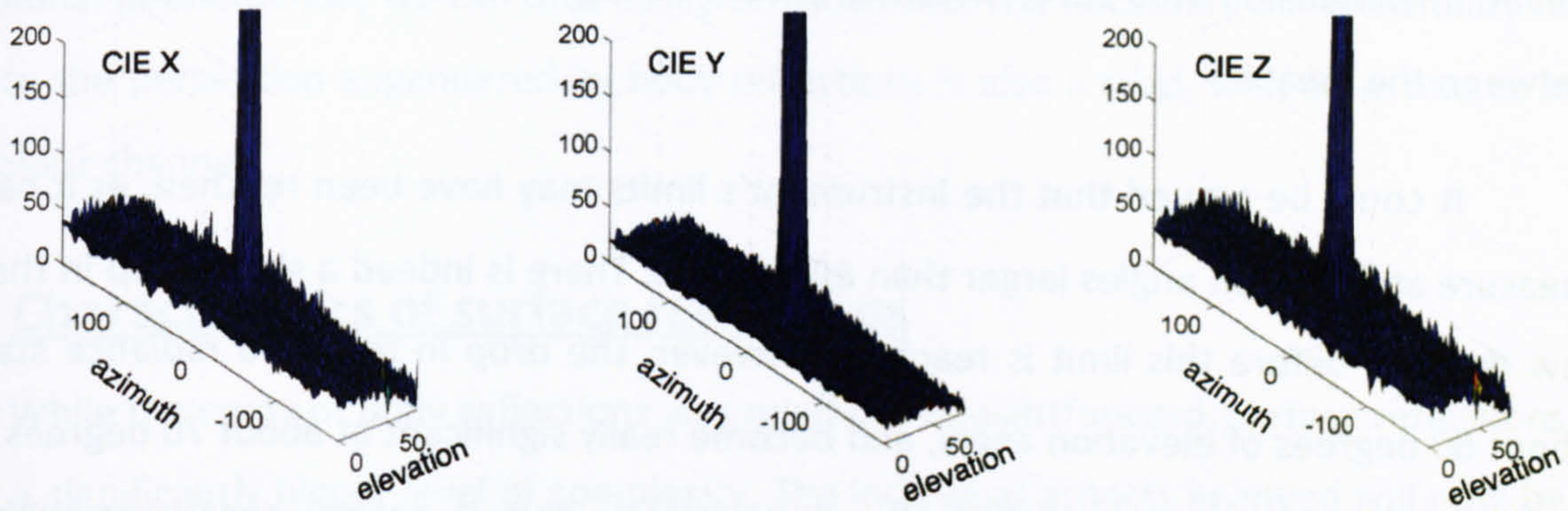


Figure 57 - Cartesian representation of the CIE X, Y and Z (from left to right) BRDFs of a magenta colour patch printed on HP Everyday Semi-glossy at full coverage. The incidence of the illuminant is 45 degrees.

8.5.1 Influence of illumination incidence angle

Although not depicted in this section, several BRDFS measurements were made with varying illumination incidence angle. Unsurprisingly, the resulting body reflections follow very well Lambert's cosine law, which states that when an area element is radiating as a result of being illuminated by an external source, the irradiance landing on that area and reflected back will be proportional to the cosine of the angle between the illuminating source and the normal of the surface. Given the obviousness of this fact, no mode details will be given.

8.5.2 Lambertian reflectance

At a first glance, body reflections seem to correspond well to those of a Lambertian diffuser, i.e. isotropic scattering across the hemisphere of reflectance. However, as previously mentioned, further examination revealed that there is a clear departure from this ideal behaviour at grazing angles. If the variations along the azimuthal dimension are not significant, those along the elevation dimension are, as can be seen from Figure 58, which provides a clearer impression of the phenomenon. The measured radiances were averaged along the azimuthal dimension, resulting in a projection of the whole BRDF onto the elevation angle-measured radiance plane. As the viewing direction approaches grazing angles, the amount of radiance reflected is quickly diminishing. This phenomenon can be observed for all types of paper and all printed colours. A two-way ANOVA was realized on the radiance means along azimuthal and elevation dimensions. The obtained p-value for the

elevation dimension was zero, which is a very strong indication that a difference exists between the means.

It could be argued that the instrument's limits may have been reached, as it cannot measure at elevation angles larger than 80 degrees. There is indeed a sharp drop in the last few degrees before this limit is reached. However, the drop in reflected radiance start at about 60 degrees of elevation angle, and become really significant at about 70 degrees. This is well before the instrument's limits. Furthermore, this is exactly consistent with the observations made by Cook and Torrance and Orchard (section 7.7.3.2). While circumstantial evidence does not constitute an absolute proof, this undeniably corroborates the existence of the reflectance drop. This departure from uniformity is however confined to grazing angles, which are not critical to the quality assessment of prints. Moreover, this discrepancy is of little significance compared to the magnitude and the shape of the specular lobes. As for anisotropy, we will therefore simplify our task by ignoring this issue and assume that body reflections are Lambertian as a starting point. Once again, further work could easily overcome this issue.

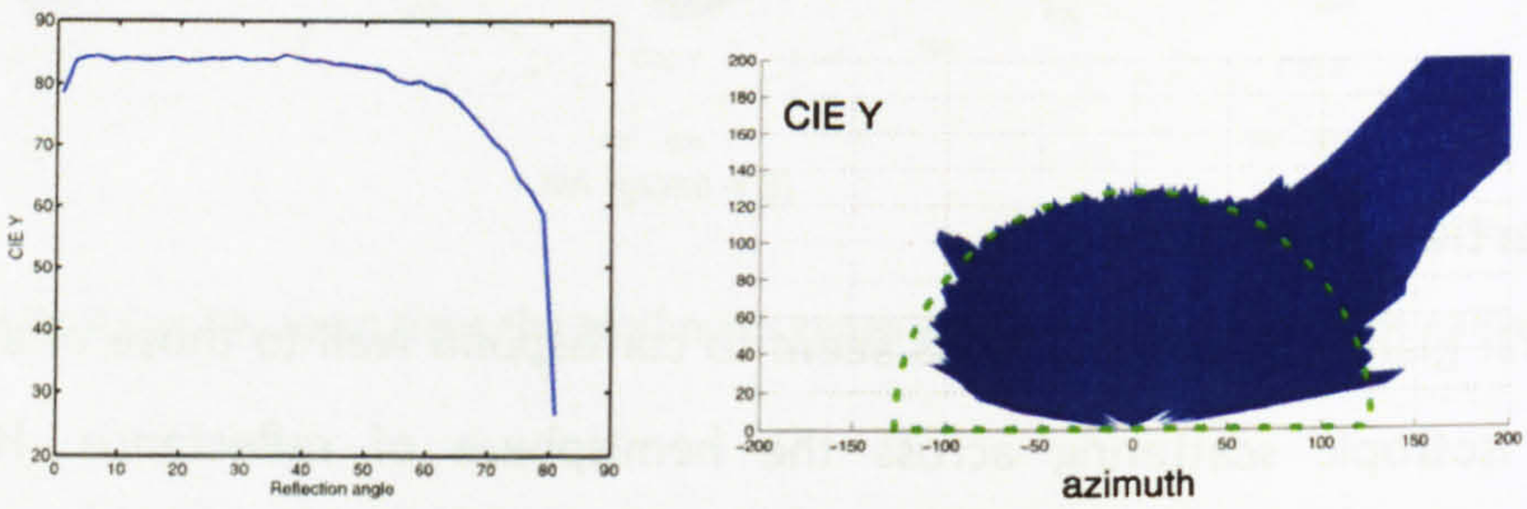


Figure 58 - Diverse representations of the changes of body reflections as a function of reflection angle. On the left is shown a cartesian plot, while a spherical plot is given on the right hand side.

8.5.3 Dimensionality of body reflections

As a concluding remark, body reflections measured in a single colour channel, whatever they may be, are undoubtedly uni-dimensional. A single estimate is sufficient to successfully replicate them with precision. Taking into account the drop in reflectance at grazing angles would not affect this state of affairs, as the pattern of drops observed is the same regardless of a print's colour or paper type. Human colour perception being a three-

dimensional phenomenon, we can thus easily conclude that the necessary dimensionality to replicate the perception engendered by body reflections is also a triad, with an estimate in each colour channel.

8.6 Characteristics of surface reflections

While the study of body reflections was relatively straightforward, surface reflections exhibit a significantly higher level of complexity. The individual aspects involved will now be presented in turns.

8.6.1 Magnitude of specular lobe

As previously observed in Figure 55, the height of the specular lobe appears to be related to surface's gloss ratings in a non-linear manner. The box-and-whisker plot displayed in Figure 59 provides another outlook on that issue. The similarity with the other box-and-whisker plot (Figure 54) showing the diverse ranges of gloss measured on the different types of paper is undeniable. They differ only in the relative differences between the paper types, whereby the plot of the specular peaks' magnitudes appears to be an exponentiated version of the gloss ratings plot. To an increase in gloss thus corresponds an exponential increase in the height of the specular lobe. The two data sets are fairly correlated (Table 26), and considering the logarithm of the specular lobes' heights rather than their absolute magnitudes improves the correlation further, which is unsurprising given the manner in which the human visual system handles the perception of brightness (Hunt, 1998).

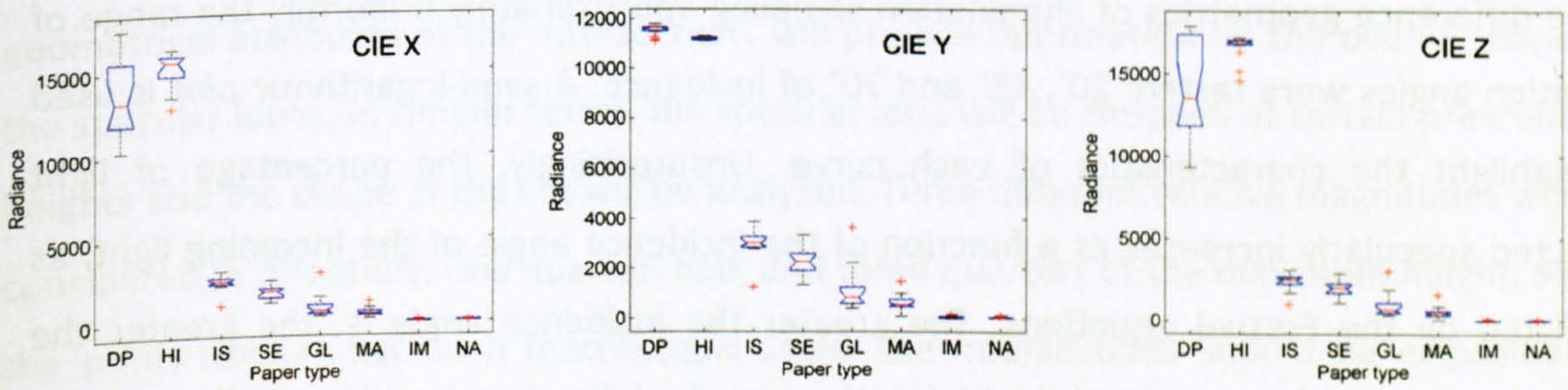


Figure 59 - Box-and-whisker plot representing the magnitude of the specular lobes of all colours measured for all paper types.

Table 26 – Channel-by-channel correlation between gloss ratings (60° measurements) and the magnitude of the specular lobe. All colours and paper types were considered.

	CIE X	CIE Y	CIE Z
Lobe height – gloss ratings	0.911	0.941	0.896
Log of Lobe height – gloss ratings	0.951	0.931	0.946

The magnitude of the specular lobe therefore clearly constitutes one of the dimensions of surface reflections. While it provides a good indication of a print's gloss ratings, the relative weakness of the correlation suggests that the relationship is not that simple. The discrepancies could be partially attributed to noise, which certainly plays a role, but the variances in heights exhibited by the two glossiest papers cannot be explained by noise only. Indeed, while their variances in the CIE Y channel are modest, their variances in the CIE X and Z channels are huge in comparison. Since the results are the average values of three successive measurements, noise and other environmental factors have to be excluded. The height variations observed are of another nature, as will be explained later.

8.6.2 Influence of illumination incidence angle

The following figure (Figure 60) illustrates the evolution of the specular lobe's magnitude as a function of the illumination incidence angle. The recorded heights correspond to a cyan colour patch printed at full coverage on all the selected paper types. Three difference geometries of illumination sampling approximately uniformly the range of elevation angles were tested: 20°, 45° and 70° of incidence. A semi-logarithmic plot is used to highlight the characteristics of each curve. Unsurprisingly, the percentage of light reflected specularly increases as a function of the incidence angle of the incoming light, as predicted by the Fresnel equations. The greater the incidence angle is, the greater the specular reflection becomes and the higher the specular lobe stretches. In a similar manner as previously, this increase is also non-linear.

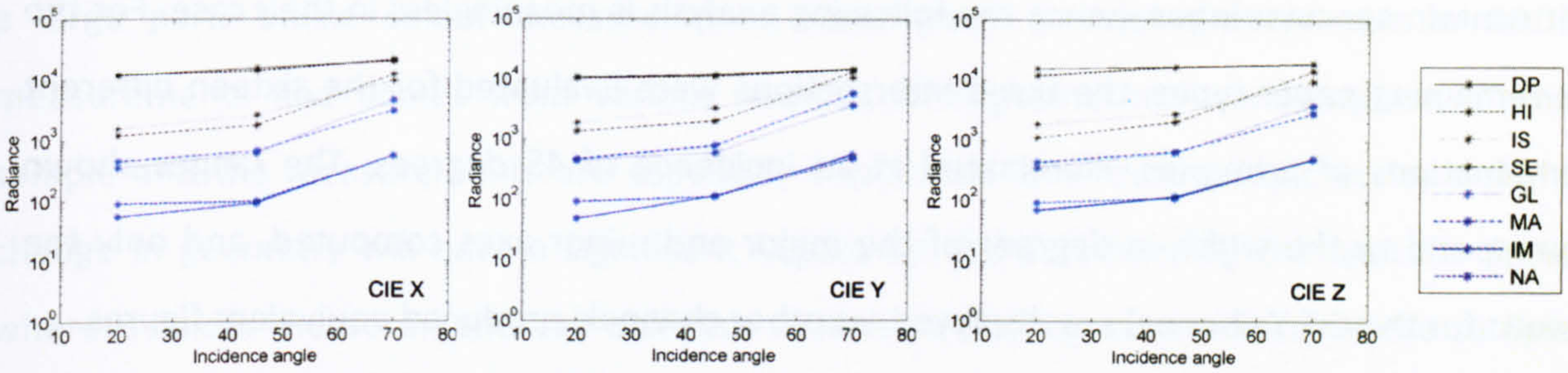


Figure 60 - Semi-logarithmic plot of the maximum radiance of the specular lobe as a function of the illumination incidence angle. In all cases, the depicted radiance distributions correspond to a cyan colour patch (100 % coverage) printed on all the selected paper types. Each plot represents (from left to right) the maximum radiance in the CIE X, Y and Z colour channels.

8.6.3 Extent of the specular lobe

As mentioned during the literature survey, the width and shape of the specular lobe is often considered as the second dimension, after its magnitude, determining the gloss perception engendered by a certain material. "[G]*loss depends not only on the specular luminous flux but also on the particular shape of the specular peak of the reflected light*" (Obein, Knoblauch, & Viénot, 2004). It is indeed trivial to imagine two BRDFs whose specular lobes have the same height but differ in width, or where the transition back to body reflections do not take place at the same pace. Surely the resulting gloss perception should be different. It is therefore of crucial importance to faithfully simulate not only the magnitude of the BRDF, but also its shape.

In order to investigate the shape of the recorded specular lobes, we will consider their intersections with pre-defined iso-luminance planes. An examination of the geometrical attributes of the intersections will provide information on the overall shape of the specular lobes. In simpler terms, the specular lobe will be chopped at certain pre-defined heights and the shape of the cut will be analysed. Three different relative magnitudes will be considered in this study: one quarter, half, and three quarters of the maximum height. Since the prints appear flat on a macroscopic scale, the intersections should be ellipsoidal in nature. The best-fitting ellipses were obtained by a regression method in order to minimise the influence of the omnipresent noise. Information about the overall shape of these intersections can thus be summarised by stating the major and minor axes of these ellipses. Figure 61 contains a box-and-whisker plot representing the ranges of these axes per type of paper. The two matt paper types are not represented as, strictly speaking, their BRDFs do

not contain specular lobes, hence the following analysis is meaningless in their case. For the six remaining paper types, the three intersections were evaluated for the sixteen different combinations of primaries, illuminated at an incidence of 45 degrees. The ranges shown correspond to the width in degrees of the major and minor axes computed, and only the results for the CIE Y channel are displayed, as other channels produced equivalent figures.

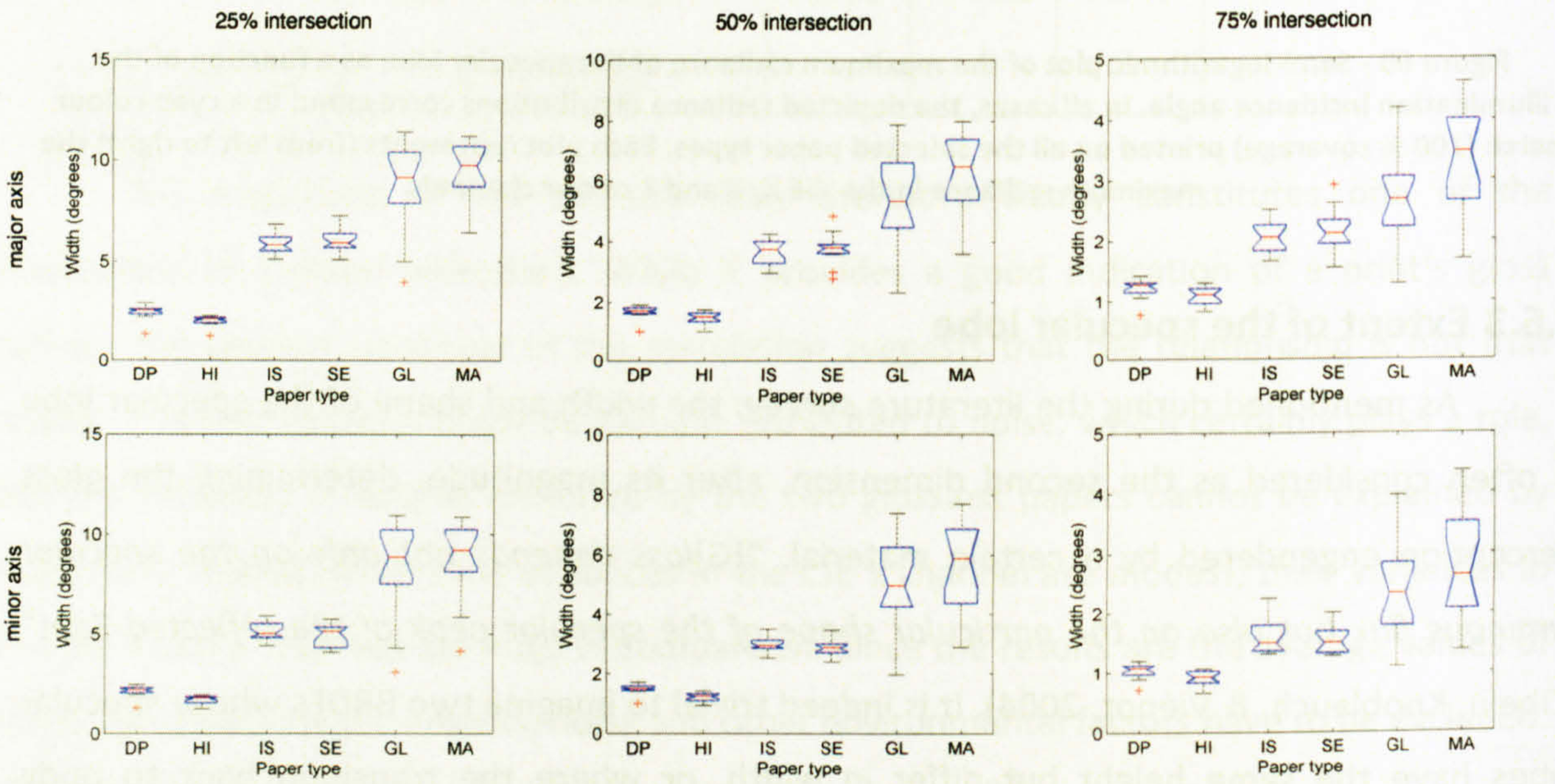


Figure 61 - Box-and-whisker plot representing the observed ranges for the width of the specular lobes of all colours measured for all glossy papers. Along the columns are represented three different cross-sections: one quarter, half, and three quarters of the maximum height. The top and bottom rows respectively represent the major and minor axes obtained.

As for the magnitude of specular lobes, the overall trend is also very clear here. The glossier the paper, the narrower the specular lobe. The specular lobe of the glossiest paper is actually so narrow that it is never more than two degrees of width on average, and is often less than one degree large at the three-quarter intersection. Similarly, even the extent of the specular lobe of the least glossy paper is rather limited, as the calculated width varies from ten degrees to below four at the three-quarter intersection. A visual corroboration can be obtained from the BRDFs displayed in Figure 56. The visual differences between the most and least glossy paper types cannot be qualified as large, as least in terms of their width. On the other hand, the variances recorded are vastly different. While glossier papers exhibit approximately the same width regardless of the colours, this is not the case for less glossy papers. This is not necessarily one of their features, as noise, or perhaps uncertainty of measurements would be a better term in this case, can certainly explain this discrepancy for

a large part. Unlike colour measurements, noise plays a much greater part in BRDF measurements and comes from various sources. The geometrical configuration of the sample and the measurements are especially important. On a macroscopic scale, a slight change in geometry will have a significant impact on the data forming the specular lobe, whereas colour measurements are oblivious to it. Even the slightest rotation of the sample being measured will modify the lobe's spatial location. Its width will also be significantly affected by any bend present in the specimen. While great care has been taken while measuring the diverse BRDFs, it is practically impossible to obtain a perfect geometrical configuration for the measurements. This is especially pertinent for less glossy papers, as their base was made of cellulose fibers, which is significantly less rigid than their glossier photographic counterparts.

Table 27 - Channel-by-channel correlation between gloss ratings (60° measurements) and the width of the specular lobe at three different height. For the paper types considered (all but matt specimens), all colours were investigated.

	CIE X	CIE Y	CIE Z
Lobe width – gloss ratings at 25%	-0.904	-0.894	-0.898
Lobe width – gloss ratings at 50%	-0.875	-0.858	-0.871
Lobe width – gloss ratings at 75%	-0.743	-0.777	-0.792

The magnitude of the variance also impacts on the quality of the correlation between the recorded width and the corresponding gloss ratings (Table 27). While the overall trend is similar to that of the magnitudes', apart from the obvious fact that the correlation is here negative, the correlation is also notably lower, no doubt due to the presence of the high variance for the least glossy papers. The high variance observed on less glossy papers is not the only interesting issue, as the consistency observed for glossier paper is also of interest. Remember that while analysing the specular lobes' magnitudes, it was observed that their variances were very large. Noise was suggested as an explanation, but given the robustness

observed here, it should actually be ruled out. This provides further indications that another factor plays a role in the variance observed for the specular lobes' heights.

8.6.4 Goniophotometer vs. glossmeter

Measuring gloss consists of estimating the amount of reflected light in a specified angular range defined by international standards (ASTM, 2004a; ISO, 2004). It should therefore be possible to approximate glossmeter readings from BRDF measurements by performing a numerical cubature on the appropriate angular range. The corresponding analysis shows a strong correlation between the two sets of results.

Table 28 - Channel-by-channel correlation between measured gloss ratings (60° measurements) and gloss estimated from BRDFs measurements. All colours and paper types were considered.

	CIE X	CIE Y	CIE Z
Correlation	0.969	0.960	0.973

8.6.5 Chromaticity of specular lobe

There appears to be a widespread assumption among the colour research community, even from very respectable sources, that the colour of surface reflections equates the illumination's. "*These mirror-like (or specular* [...]*) reflections are produced by the top surface of objects, and, unless the objects are metals, the light reflected is not coloured by the surface and is therefore the same colour as the light source*" (Hunt, 1998). "[...] *the common assumption that the surface reflectance from colour prints is spectrally non-selective*" (Granberg, 2002). If the chromaticity of the surface reflections was that of the illuminant, a single point would be seen in Figure 62, which displays the chromaticities of the specular lobes for all recorded BRDFs. A large spread can be observed instead. Contrary to common assumption, the chromaticities of surface reflections are therefore different from the chromaticity of the illuminant.

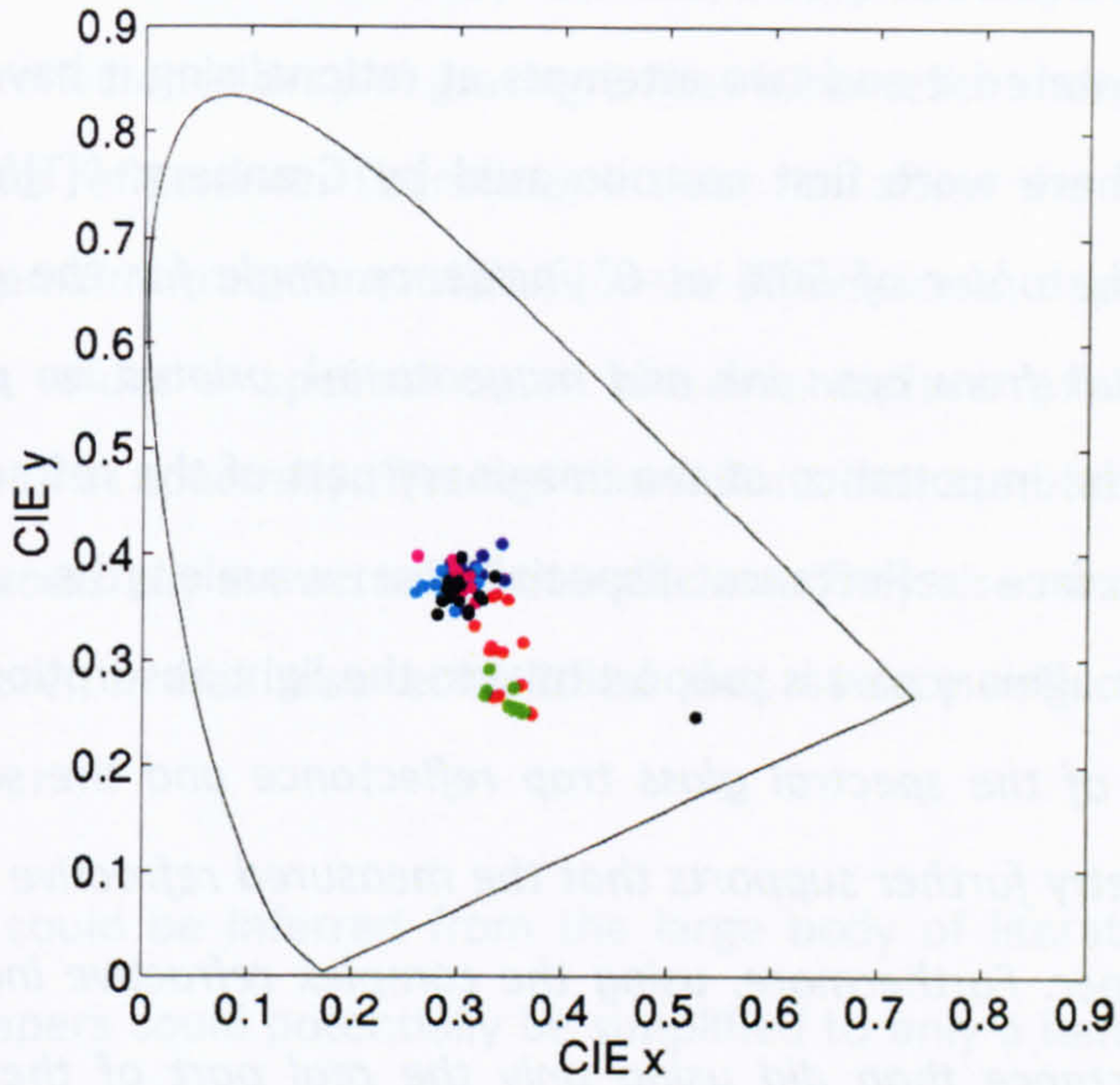


Figure 62 - Chromaticity plot of the measured specular lobes for all colours and paper types. To each colour correspond a single paper type. Each point of the same colour represents different CMYK combinations.

The standard deviations between the chromaticities of the sixteen combinations of primaries range between 0.0110 and 0.0581 chromaticity units for all paper types. These are large differences, corresponding to up to 26 ΔE*ab if the same lightness is assumed. This phenomenon can account for the large variance observed while analysing the specular lobe's heights. A large variance was only detected in the CIE X and Z channels, whereas data for the CIE Y channel remained very consistent. This can easily be explained by the fact that glossmeters actually perform a luminance measurement, i.e. in the CIE Y channel. The large variance observed was therefore simply due to the fact that specular lobes have varying chromaticities.

Body reflections could provide an obvious explanation for the occurrence of this phenomenon. Indeed, they do exhibit very significant spectral variations. It could thus be reasonably envisaged for surface reflections to be affected by them, for instance simply as a combination of their spectral selectivity and the illuminant's spectral power distribution. However, this is very unrealistic due to the sheer magnitude of specular lobes. For high-gloss papers, body reflections are insignificant in comparison. Furthermore, this possibility has been extensively analysed, and no apparent link between the chromaticity of the lobe and the chromaticity of the diffuse component was uncovered. The spectral selectivity of surface

reflections is however undeniable. Despite the lack of awareness of this issue, few researchers have encountered it and two attempts at rationalising it have been published. The results presented here were first corroborated by Granberg. "[T]*he results showed spectral variations in the order of 50% at 0° incidence angle for the gloss and surface reflectance from cyan ink from cyan ink and magenta ink printed on paper*" (Granberg, 2002). He emphasised the importance of the imaginary part of the refractive index, which contributes to the surface reflectance especially at wavelengths with strong light absorptions, since the imaginary part is proportional to the light absorption coefficient. "*The almost identical shape of the spectral gloss trap reflectance and the surface reflectance predicted from elipsometry further supports that the measured refractive index is physically relevant for printed paper. Furthermore, using the complex refractive index gave a better prediction of the reflectance than did using only the real part of the refractive index*" (Granberg, 2002). On the other hand, Arney *et al.*, while still noticing strongly chromatic surface reflections, observed an exactly opposite behaviour. "*A highly transparent ink (e.g., yellow ink measured with red light) reflected approximately twice as much specular light than a highly absorbing ink (e.g., yellow with blue or cyan with red)*" (Arney, Anderson, Franz, & Pfeister, 2006). They establish that a significant amount of specular reflections actually originates from below the surface of a print. Their study reveals the presence of air gaps between colorant and paper, increasing the number of interfaces where Fresnel reflections occur. Surface reflections therefore do not originate only from air-ink/toner interfaces, but also ink/toner-air and air-paper interfaces. This diversifies the spectral composition of reflected light, as each material possess a different index of refraction.

8.6.6 Synopsis

The empirical analysis has uncovered few additional issues on top of what was already inferred from the literature survey. While the mathematical definition of the BRDF is general enough to actually take all of them into account, there is a widespread presumption that most BRDFs, especially for materials like printed papers, can be simplified down to the sum of two distinct components that vary either in the spatial or spectral domain. "*In describing appearance, wavelength (or spectral) variability is primarily responsible for color, while geometric (or directional) selectivity is primarily responsible for gloss, luster, translucency, and like attributes*" (ASTM, 2004c). This is however far from the truth.

First of all, we have seen that body reflections vary not only colorimetrically, but also geometrically. The reflectance drop at grazing angles is however always the same, regardless of the material or its reflectance. Therefore, it does not affect the dimensionality of body reflections, which remain three-dimensional (or uni-dimensional, depending on whether we consider tristimulus values or spectral data). Likewise, surface reflections also vary in the spatial domain as well as spectrally. The two dimensions that are often cited as necessary to reproduce them successfully are actually not sufficient. They also carry a strong colorimetric content. Furthermore, it was also established that surface reflections exhibit a minor anisotropic behaviour.

Although it could be inferred from the large body of literature on the topic that BRDFs of printed papers could potentially be simplified to only a handful of terms, say one dimension for body reflections (from a spectral perspective) and two for describing the geometry of surface reflections, this is however an erroneous approach if one tries to precisely simulate, from a mathematical perspective, such BRDFs. On the other hand, there are few mitigating factors. First of all, the anisotropy, although clearly present, can be considered as rather inconsequential, especially since simply choosing an appropriate geometric configuration for measurements can attenuate its importance to the point that it is almost not noticeable. Secondly, the feasibility of the proposed solution, i.e. whether it could realistically be implemented in actual industrial settings, is paramount. This indubitably removes spectrogoniometers from the equation, and since glossmeters are only able to provide readings in the luminance channel, replicating the chromaticity of surface reflections is out of the question. Lastly, the goniometric non-uniformity of body reflections will also be omitted from the rest of this study in order to minimise the complexity of the task ahead. After all, the changes of body reflections only occur at grazing angles, which are not of primary importance in the assessment of prints' quality.

While these simplifications are guaranteed to diminish the precision achieved from a mathematical perspective, their implications are far less for psychophysical evaluations. They indeed concern extreme cases (grazing angles), or virtually imperceptible phenomenon (anisotropy). The major characteristics of BRDFs, such as the relative magnitude of body reflections, and the width and height of the specular lobe should still be reproduced, providing that the simulation perform approximately correctly. As for the chromaticity of

gloss, monitors are not able to reach levels of luminance sufficient to replicate the magnitude of surface reflections, hence this is a moot point. Furthermore, perhaps accurately reproducing the colour of gloss may not be so important anyway. "[I]*t is somewhat surprising that chromatic effects are not generally noticed when examining the gloss characteristic of printed materials. Perhaps the extreme glint of the gloss coupled with the highly chromatic character of the underlying, diffuse image masks the much lesser chromatic character of the gloss*" (Arney, Anderson, Franz, & Pfeister, 2006).

Chapter 9: BRDF Modeling

The purpose of this study is to improve the accuracy of soft proofing by increasing the amount of information rendered on screen. Ultimately, the aim is to provide an interactive computer-based simulation faithfully rendering the total appearance of printed paper illuminated by viewing booths typically employed in prepress houses. This first implies enhancing material simulation beyond simple colour appearance prediction to more complete representations. As the title suggests, what is proposed here is a full goniometric simulation of the spatial distribution of light reflected by a print, which is usually represented by the bi-directional reflectance distribution function. The objective has however to be tempered by practical considerations. As previously mentioned, the viability of the proposed solution in actual industrial settings is of the utmost importance. Spectrogoniometers were thus deemed unsuitable for the task. Instead, the intent consists of approaching BRDFs from the readings of simpler and cheaper instruments, glossmeters and spectrophotometers, that are readily available in prepress houses. This decision has evidently some unavoidable consequences on the accuracy achievable. Furthermore, the task was simplified further by omitting few BRDFs characteristics from the simulation. While the precision of the simulation may be affected on absolute terms, it was however rationalised that these discrepancies are unlikely to have a major impact on the visual quality of the simulation.

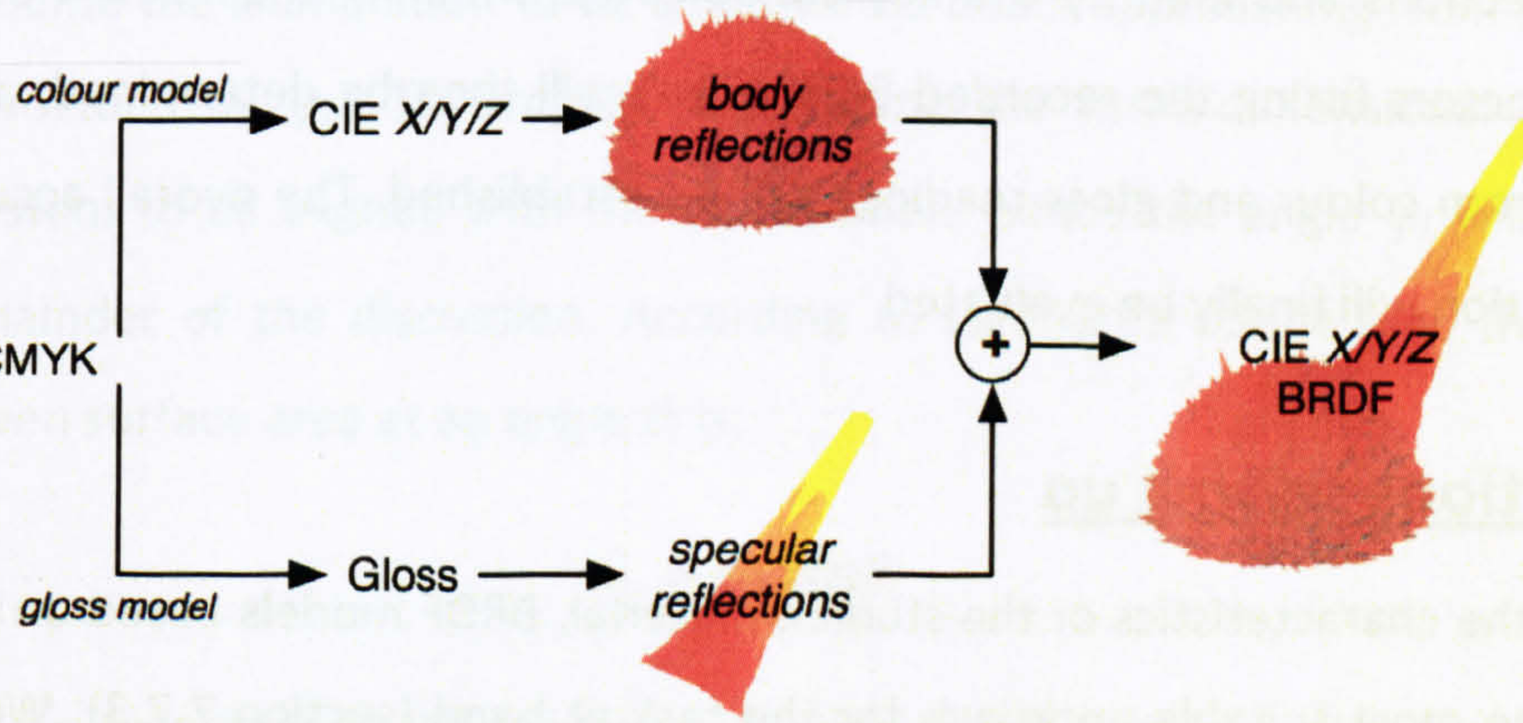


Figure 63 - Conceptual representation of a workflow simulating the goniometric distribution of reflected light.

The proposed solution should ideally integrate into existing workflows rather than displace them. Existing colour workflows with proven track record are already in place and mastered by prepress houses staff. Discarding this expertise would be in every way detrimental, especially since an accurate reproduction of colour is a core aspect of their business. In that context, the proposed workflow (Figure 63) is essentially an extended version of standard colour characterisation models. Instead of providing a single output (CIE XYZ), the entire distribution of reflected colours under all possible angles is estimated (BRDFs of CIE X, Y and Z). Three different versions of the workflow will be derived, one for each colour channel. The workflow itself comprises two pathways, each having two stages. The two pathways arise from the fact that BRDFs can generally be divided into two independent components, i.e. body and surface reflections. Each component will be modelled individually, and the final results will simply be obtained by summing them. The two stages are very similar in both pathways. The first consists of mapping CMYK data to estimates of body and surfaces reflections, i.e. CIE XYZ and gloss respectively. The second stage will then transform these estimates into the parameters of an actual BRDF model, from which simulation of entire BRDFs can be generated. Once all the diverse parameters are derived, a pixel's BRDF may thus be simply obtained from its CMYK coverage data.

The second stage of the process provides the most challenging issues. Indeed, robust models already exist for the first stage, so only few details will be given on that topic. This section will thus focus in turn on providing an explicit mathematical description of surface and body reflections and identify the key variables determining their behaviour. A set of optimal parameters fitting the recorded BRDFs best will then be determined, and ways to derive them from colour and gloss readings will be established. The overall accuracy of the proposed solution will finally be evaluated.

9.1 Equations setting up

Given the characteristics of the studied material, BRDF models based on geometrical optics offer the most suitable approach for the task at hand (section 7.7.3). We will adopt the model developed by Cook and Torrance (1981), as it handles rough isotropic surfaces and is widely used. Alternatively, the equivalent Ward (1992) model may also be of interest, as it is computationally faster and can deal with anisotropic BRDFs.

To recap, the model developed by Cook and Torrance is a dichromatic reflection model, whereby sub-surface reflections are assumed to be Lambertian, and surface reflections are described by the Fresnel equations modulated by a roughness term. By using the geometrical notations defined in section 7.7.1.1 and reproduced here (Figure 64), the radiant exitance estimated by the Cook-Torrance model can be expressed symbolically as:

$$L(\theta_I,\varphi_I,\theta_R,\varphi_R,\lambda) = c_{body}L_{body}(\theta_I,\varphi_I,\lambda) + c_{surf}L_{surf}(\theta_I,\varphi_I,\theta_R,\varphi_R,\lambda) \quad (18)$$

where c_{body} and c_{surf} represent the proportion of light reflected from body and surface reflections respectively. Each term of this equation will now be expanded in turn. We will omit the wavelength term for the sake of clarity.

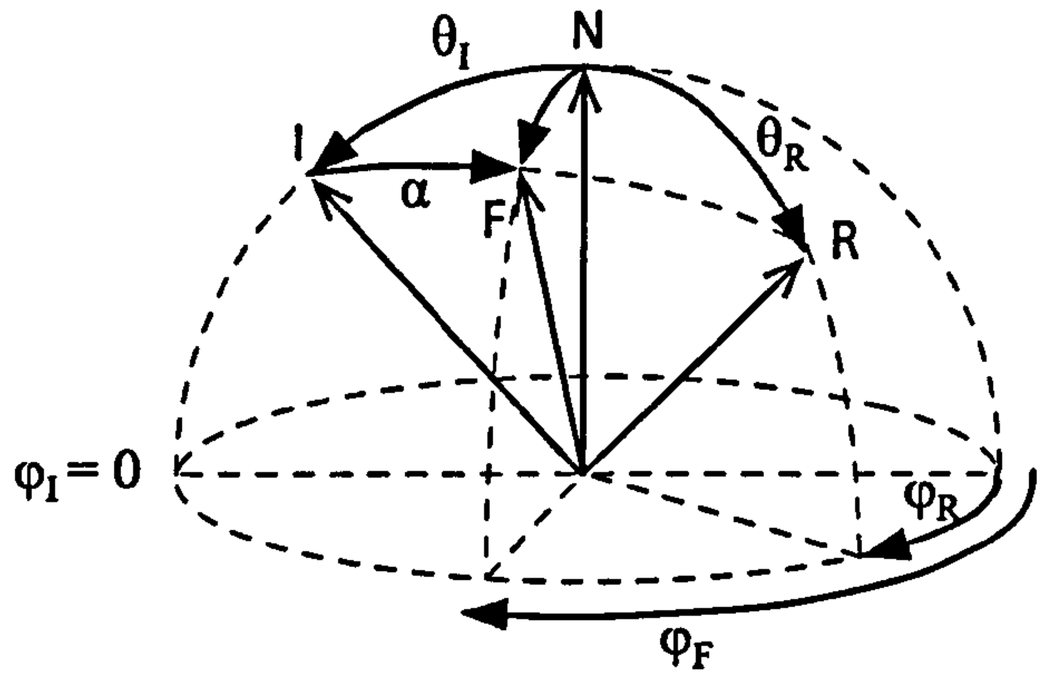


Figure 64 - Geometric notations.

a) Irradiance

We assume the illumination to be a collimated beam approaching from the direction specified by vector $I(\theta_I, \varphi_I)$. To simplify, we assume the three-dimensional orthonormal coordinate system to be aligned with the illumination, hence the angle φ_I will be omitted from the remainder of the discussion. According to Lambert's cosine law, the irradiance falling on a given surface area at an angle θ_I is:

$$E_I = E_O \cos\theta_I \quad (19)$$

where E_O is defined as the irradiance at normal incidence, i.e. the raw irradiance of the spectrogoniometer.

b) Measurement surface area

The measurement surface area S_R depends on the reflection direction as follow:

$$S_R = \frac{S_{RO}}{\cos\theta_R} \tag{20}$$

where S_{RO} is a disc whose dimensions are given in the spectrogoniometer's specifications.

c) Luminous flux

The luminous flux reaching an element of surface area S_R from which the reflection towards $R(\theta_R, \varphi_R)$ will be initiated is given by:

$$\phi_{I,R} = S_R \cdot E_I \tag{21}$$

d) Surface reflections

In the Cook-Torrance model, a surface is assumed to consist of a collection of micro-facets of different orientation. Because each of these micro-facets behaves like a perfectly reflecting or mirror faces, specular reflections will occur as for an optically smooth surface. Therefore, for an incident light coming from a direction $I(\theta_I)$, all the light reflected towards a direction $R(\theta_R, \varphi_R)$ must originate from micro-facets whose normal is aligned with the vector H, the bisector of I and R. The perturbation generated by the differences in orientation will spread the specular reflections and generate a lobe. These can be described by a probability distribution function $P_H(\alpha)$ of the orientation of the micro-facets. $P_H(\alpha)$ represent the probability of finding a micro-facet whose normal is orientated in the direction of $H(\alpha, \varphi_H)$. The number of micro-facets orientated in the direction of $H(\alpha, \varphi_H)$ can be expressed as:

$$n_H = P_H(\alpha) n_{tot} \tag{22}$$

where n_{tot} represent the total number of micro-facets present in the element of surface area S_R. Likewise, the total surface area occupied by micro-facets orientated in the direction of $H(\alpha, \varphi_H)$ is given by:

$$S_H = f \cdot \cos(\alpha) \cdot P_H(\alpha) \cdot n_{tot} \tag{23}$$

where f represent the surface area of a single micro-facet. According to the laws of geometric optics, only S_H contribute to specular reflections in the direction of $R(\theta_R, \varphi_R)$. Therefore, only a fraction of the incoming luminous flux will be reflected towards $R(\theta_R, \varphi_R)$:

$$\phi_{R,H} = \phi_{I,R} \cdot \frac{S_H}{S_R}$$
$$\phi_{R,H} = \phi_{I,R} \cdot \frac{f \cdot \cos(\alpha) \cdot P_H(\alpha) \cdot n_{tot}}{S_R} \quad (24)$$

This expression can be simplified by a few judicious observations. Firstly, the measurement area S_R is equal to the sum of all surface elements S_H for $H(\alpha, \varphi_H)$ describing all possible directions in the hemisphere:

$$S_R = \int_0^{2\pi} d\varphi_F \int_0^{\pi/2} S_H d\alpha$$
$$S_R = \int_0^{2\pi} d\varphi_F \int_0^{\pi/2} f \cdot \cos(\alpha) \cdot P_H(\alpha) \cdot n_{tot} d\alpha \quad (25)$$
$$S_R = 2\pi f n_{tot} \int_0^{\pi/2} \cos(\alpha) \cdot P_H(\alpha) d\alpha$$

From that expression, we can isolate n_{tot}:

$$n_{tot} = \frac{S_R}{2\pi f \int_0^{\pi/2} \cos(\alpha) \cdot P_H(\alpha) d\alpha} \quad (26)$$

Let's anticipate on the results by stating that $P_H(\alpha)$ very quickly decreases to nil for all specimens tested in this study (section 0). In such cases, it is reasonable to make the following approximation:

$$\cos(\alpha) \sim 1 \quad (27)$$

which yields:

$$\int_0^{\pi/2} \cos(\alpha) \cdot P_H(\alpha) d\alpha \sim \int_0^{\pi/2} P_H(\alpha) d\alpha \quad (28)$$

Finally, let's reflect on the fact that the sum of probabilities $P_H(\alpha)$ over the hemisphere of integration is necessarily equal to one, which yields:

$$\int_0^{2\pi} d\varphi_F \int_0^{\pi/2} P_H(\alpha) d\alpha = 1$$
$$\int_0^{\pi/2} P_H(\alpha) d\alpha = \frac{1}{2\pi} \quad (29)$$

Equation 26 can thus be simplified thanks to equations 28 and 29:

$$n_{tot} \approx \frac{S_R}{f} \quad (30)$$

The reflected luminous flux can thus be simply expressed as:

$$\phi_{R,H}(\alpha) = \phi_{I,R} \cdot \cos\alpha \cdot P_H(\alpha) \quad (31)$$

The last expression would however only be valid if micro-facets were reflecting 100% of incoming light, which is not the case. Firstly, the amount of light reflected and refracted as a ray of light strikes a boundary is determined by the Fresnel equations, which are functions of the incoming light's incidence angle and the complex refractive index. Since the materials at the heart of this study are dielectrics, the imaginary part of the refractive index may be omitted. While there are suggestions the imaginary plays an important role in determining the chromaticity of surface reflections, approaching this quantity from measurements is beyond the realm of this study. Furthermore, the constraints imposed also forbid the simulation of this characteristic of surface reflections. The real part of the refractive index n will thus suffice in our context. The Fresnel factor is denoted $F(\theta_I, n)$. Secondly, another factor modulating the amount of surface reflections is the roughness of the surface. Shadowing and masking phenomena are taken into account by the geometric attenuation factor, which is defined as follow:

$$G(\theta_I, \theta_R, \alpha, \psi) = \min\left\{1, \frac{2\cos\alpha\cos\theta_r}{\cos\psi}, \frac{2\cos\alpha\cos\theta_i}{\cos\psi}\right\} \quad (32)$$

The luminous flux $\phi_{R,H}$ reflected towards $R(\theta_R, \varphi_R)$ is attenuated by these two factors and becomes:

$$\phi'_{R,H} = \phi_{R,H} \cdot F(\theta_I, n) \cdot G(\theta_I, \theta_R, \alpha, \psi) \quad (33)$$

Since the model provide estimates of exitant radiances, the radiance reflected towards $R(\theta_R, \varphi_R)$ is:

$$L_R(\theta_I,\theta_R,\varphi_R) = \frac{\phi'_{R,H}}{\omega S_R \cos\theta} = \frac{\phi'_{R,H}}{\omega \frac{S_{RO}}{\cos\theta}\cos\theta}$$
$$L_R(\theta_I,\theta_R,\varphi_R) = \frac{\phi'_{R,H}}{\omega S_{RO}} \quad (34)$$

where ω is the solid angle of the measurement cone. To summarise, the radiance reflected from surface reflections, for an illumination originating from $I(\theta_I)$ and measured in the direction $R(\theta_R, \varphi_R)$, is:

$$L_{surf}(\theta_I,\theta_R,\varphi_R) = \frac{1}{\omega S_{RO}} F(\theta_I,n)\cdot G(\theta_I,\theta_R,\alpha,\psi)\cdot P(\alpha)\cdot\cos\alpha\cdot\frac{S_{RO}}{\cos\theta_R}\cdot\cos\theta_I\cdot E_O$$
$$L_{surf}(\theta_I,\theta_R,\varphi_R) = \frac{E_O}{\omega}\frac{\cos\alpha\cos\theta_I}{\cos\theta_R}F(\theta_I,n)\cdot G(\theta_I,\theta_R,\alpha,\psi)\cdot P(\alpha) \quad (35)$$

e) body reflections

Part of the incoming luminous flux $\phi_{I,R}$ penetrates the surface and is reflected back towards the direction $R(\theta_R, \varphi_R)$. If ρ_d represents the proportion of light reflected below the surface, the behaviour of such a Lambertian diffuser is expressed as:

$$L_{body}(\theta_I) = \rho_d \frac{E_I}{\pi}$$
$$L_{body}(\theta_I) = \rho_d \frac{E_O}{\pi}\cos\theta_I \quad (36)$$

f) total reflections

For dichromatic models, the total radiance reflected is the weighted sum of surface and body reflections. Therefore, the luminance reflected by printed paper in a given direction can be expressed as:

$$L(\theta_I,\theta_R,\varphi_R) = c_{body}\rho_d\frac{E_O}{\pi}\cos\theta_I + c_{surf}\frac{E_O}{\omega}\frac{\cos\alpha\cos\theta_I}{\cos\theta_R}F(\theta_I,n)\cdot G(\theta_I,\theta_R,\alpha,\psi)\cdot P(\alpha) \quad (37)$$

which can be simplified to:

$$L(\theta_I,\theta_R,\varphi_R) = C_{body}\cos\theta_I + C_{surf}\frac{\cos\alpha\cos\theta_I}{\cos\theta_R}F(\theta_I,n)\cdot G(\theta_I,\theta_R,\alpha,\psi)\cdot P(\alpha) \quad (38)$$

if C_{body} ans C_{surf} are two constants defined as:

$$C_{body} = c_{body}\rho_d\frac{E_O}{\pi}$$
$$C_{surf} = c_{surf}\frac{E_O}{\omega} \quad (39)$$

In order to obtain the BRDF from this expression, we simply need to normalise by the incoming irradiance:

$$\rho(\theta_I,\theta_R,\varphi_I) = \frac{L(\theta_I,\theta_R,\varphi_I)}{E_O(\theta_I)\cos\theta_I} \quad (40)$$

g) Equation unknowns

The model as expressed by equation 38 contains five degrees of freedom:

- c_{body} and c_{surf}, which account for one as their sum is necessarily equal to 1.
- E_O, the incoming irradiance.
- ρ_d, which expresses body reflections.
- n, the real part of the refractive index.
- $P(\alpha)$, the probability distribution of the micro-facets orientation.

Unfortunately, it was not possible to estimate the irradiance generated by the spectrogoniometer. It was however kept constant throughout the measurement phase. Since E_O was necessary for the deduction of the actual values of ρ_d, c_{body} and c_{surf}, the other set of constants C_{body} and C_{surf} were employed instead. A specimen's characteristics are thus represented by four parameters, C_{body}, C_{surf}, n and $P(\alpha)$, which will now be quantified in turn.

9.2 Optimal parameters

9.2.1 Body reflections constant C_{body}

In the case of body reflections, a single parameter (per colour channel) is sufficient to describe them, as they are assumed to behave in a Lambertian way. It thus seems obvious to derive the value of the constant C_{body} from simple tristimulus values, and indeed, a strong correlation between the two sets of data can be observed from Figure 65. The diverse values of C_{body} for all possible paper/ink combinations were determined by averaging the aspecular data for the recorded BRDFs. The aspecular data were defined as all radiances not participating in the specular lobe, which obviously varied depending on the paper. Also, the drop in reflectance at grazing angles could have interfered with the analysis and was thus ignored by simply disregarding incidence angles larger than 70°. The corresponding tristimulus values were measured under the CIE 45/0 geometry of illumination and viewing. A strong linear correlation can be observed for all three channels, indicating that the C_{body} constants can simply be derived from tristimulus values.

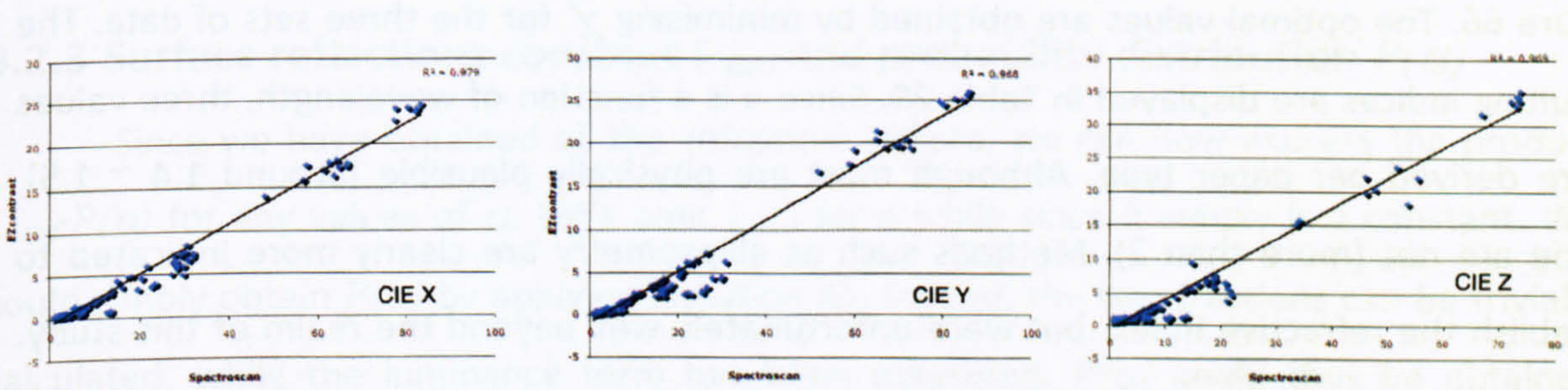


Figure 65 - Correlation between the average values of body reflections for all paper/ink combinations and their corresponding tristimulus values for all three colour channels. The illumination incidence angle for the BRDF measurement was 45°, while the geometry of measurements for tristimulus values was 45/0.

9.2.2 Refractive index *n*

In order to derive the remaining parameters, the constant C_{body} was subtracted from each BRDF measurement. Indeed, since the last three parameters control surface reflections, all the analysis will be carried on the specular lobe data alone. Let's restrict the investigation field even further by only considering data belonging to the illumination's incidence plane. This will not affect the results since the measured BRDFs are radially isotropic. The refractive index n will be obtained first. Unlike the C_{body} constants, which vary

depending on the colour printed, a unique n will be derived for each paper type. Indeed, refractive indices are a property of materials, and there are only five materials present at most, a substrate and four inks. While every material have different refractive indices, most of their constituents are organic in this case, i.e. their refractive indices should be approximately equivalent, hence a single estimate should suffice for our needs. Accessing such fine details was anyway impossible in the context of this study. A single refractive index will thus be estimated.

Let's isolate n from the other two parameters by rewriting equation 35:

$$C_{surf} \cdot P(\alpha) = L_{surf}(\theta_I, \theta_R, \varphi_R) \cdot \frac{\cos\theta_R}{\cos\alpha\cos\theta_I} \frac{1}{F(\theta_I, n) \cdot G(\theta_I, \theta_R, \alpha, \psi)} \quad (41)$$

The product $C_{surf}P(\alpha)$ is evidently independent of the illumination's incidence angle. Since measurements at three different incidence angles of illumination were performed, the goal consists in adjusting the value of n so that the three curves obtained from the combinations of expression 41 and the three measurements are superposed, as illustrated in Figure 66. The optimal values are obtained by minimising χ^2 for the three sets of data. The resulting indices are displayed in Table 29. Since n is a function of wavelength, three values were derived per paper type. Although most are physically plausible (around 1.4 ~ 1.5), some are not (more than 2). Methods such as ellipsometry are clearly more indicated to establish the refractive index, but were unfortunately well beyond the realm of this study. The results presented here should be more considered as numerically optimized parameters rather than physical quantities. The changes that would be entailed by obtaining the real indices could be compensated by opposite modifications of the other two unknowns variables, C_{surf} and $P(\alpha)$.

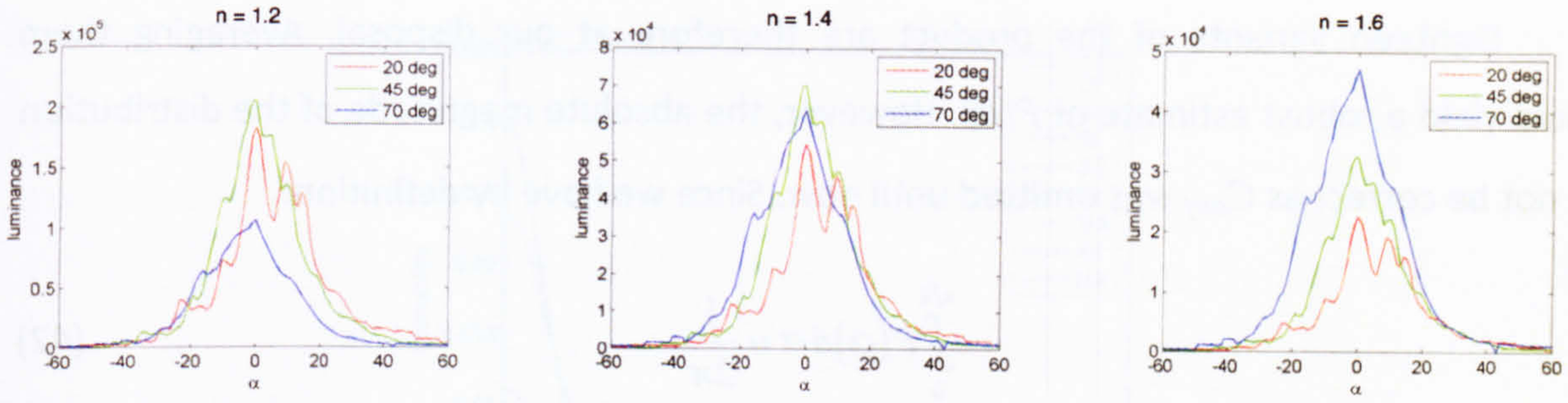


Figure 66 - Representation of the $C_{surf}P(\alpha)$ product computed from three measurements with varying incidence angles and for diverse refractive indices. The depicted curves correspond to the CIE X channel of a cyan colour patch (100 % coverage).

Table 29 – Refractive indices for all paper types.

	DP	HI	IS	SE	GL	MA	IM	NA
CIE X	2.5	2.5	1.4	1.2	1.3	1.4	1.3	1.5
CIE Y	2.7	2.6	1.4	1.2	1.3	1.4	1.4	1.6
CIE Z	2.6	2.7	1.4	1.2	1.4	1.5	1.6	1.7

9.2.3 Surface reflections constant C_{surf} and probability distribution $P(\alpha)$

Since we have obtained all the refractive indices, we can now express the product $C_{surf}P(\alpha)$ for any values of α. Let's omit C_{surf} for a while since it simply is a constant. We could simply obtain $P(\alpha)$ by applying Equation 41. Indeed, the two fractions can be trivially calculated, while the luminance term has been measured. $P(\alpha)$ could thus be obtained iteratively, by simply computing for all values of α the two fractions, and extracting the luminance corresponding to that α from the BRDF measurements. Actually, we have at our disposal multiple estimates of this luminance term since:

- measurements at three different incidence angles were performed.
- measurements were performed in each colour channel. Since $P(\alpha)$ represents a geometrical attribute, a unique definition should prevail.
- measurements for negative α should be equal to those of positive α since the BRDF is isotropic.

Eighteen variants of the product are therefore at our disposal. Averaging them should yield a robust estimate of $P(\alpha)$. However, the absolute magnitude of the distribution will not be correct as C_{surf} was omitted until now. Since we have by definition:

$$\int_0^{\pi/2} P(\alpha)d\alpha = \frac{1}{2\pi} \tag{42}$$

the exact value of C_{surf} can be derived so that the integral is equal to $1/2\pi$, and the correct $P(\alpha)$ will naturally ensue. The diverse C_{surf} and $P(\alpha)$ are respectively given in Table 30 and illustrated in Figure 67. In the same way as for the refractive index, three different values of C_{surf} are obtained per paper type, whereas since $P(\alpha)$ represents the microscopic geometry of surfaces, a single instance should be derived per paper type. The probability distribution of matt is approximately flat, indicating that micro-facets are orientated approximately randomly. The surface can thus be qualified as rough, hence its matt character. On the other hand, one is more likely to find micro-facets whose orientation is close to the normal on glossy papers, which means that such surfaces are mostly flat, and thus glossy.

Table 30 - C_{surf} values for all paper types, in $(lm.m^{-2}.sr^{-1})x10^6$.

	DP	HI	IS	SE	GL	MA	IM	NA
CIE X	3.02	2.19	5.67	13.2	3.61	2.05	0.77	0.49
CIE Y	2.49	1.96	6.99	15.9	4.61	2.65	0.73	0.50
CIE Z	2.96	2.14	5.85	14.6	2.66	1.63	0.43	0.36

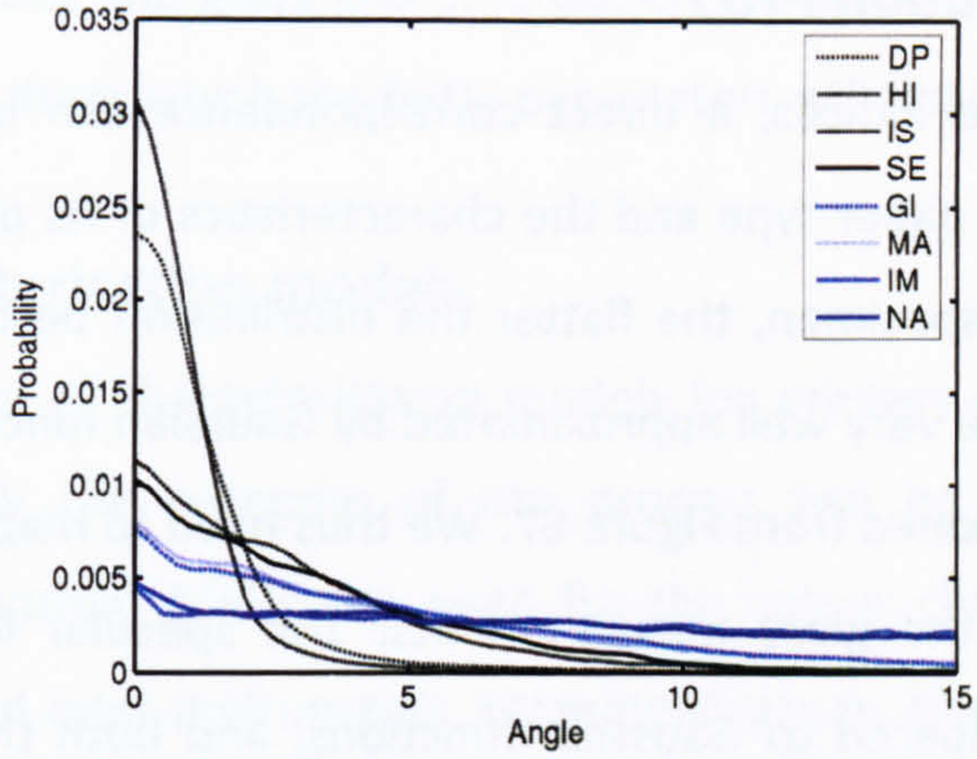


Figure 67 - Probability distribution functions of the micro-facets orientation for all papers.

9.3 Relating colour and gloss measurements to BRDF parameters

We have obtained the best fitting parameters representing the measured BRDFs. In the case of body reflections, the mapping between tristimulus values and C_{body} constants is trivial, as previously explained. No further details will thus be given on that topic. Instead, this section will focus on presenting methods to map BRDF parameters and gloss readings.

9.3.1 Refractive index *n*

As previously mentioned, the real refractive indices of all the diverse materials used during the course of this study are actually not too dissimilar from each other. Indeed, all colorants and substrates are constituted mostly of organic materials, indicating that the real part of their refractive index is approximately 1.5 (section 7.6.3.2), no matter what inks or paper types are used. While greater differences were obtained for the mathematically derived values presented here, it was argued that the lack of proper instrumentation was at fault. Nonetheless, the reported variations in refractive indices are minimal, and clearly cannot taken into account the differences in perceived or measured gloss. There is absolutely no connection between the gloss measurements (Figure 54) and the refractive index (Table 29). We therefore cannot predict the latter from the former. Also, contemplating the possibility for pre-press houses to measure this term is absurd, since it could not even been done in this academic study. It is however conceivable for the printer / paper / colorant manufacturer to provide such data. We will suppose this is the case and consider the refractive indices as given quantities.

9.3.2 Probability distribution $P(\alpha)$

Unlike the refractive indices, a direct correspondence can be observed between average gloss readings of a paper type and the characteristics of its probability distribution $P(\alpha)$. The less glossy the specimen, the flatter the distribution becomes. The probability distribution functions can be very well approximated by Gaussian functions centered in zero (r^2 = 0.996), as can be envisioned from Figure 67. We thus need to map the characteristics of these Gaussian curves to the gloss measurements. The specular lobes themselves can actually be reasonably compared to Gaussian functions, and both their width and height were fairly well correlated to gloss readings. Similarly, the width parameter of the Gaussian functions fitted to the diverse $P(\alpha)$ is well correlated to gloss (r^2 = 0.947), but the height is not. However, the height is very highly correlated to the width (r^2 = 0.999). The probability distribution functions can therefore be approximated by Gaussian curves whose parameters are derived from the gloss measurements. The correlation between the estimated and the best-fitting parameters is high (r^2 = 0.981).

9.3.3 Surface reflections constant C_{surf}

At a first glance, the magnitudes of C_{surf} do not appear to be related to gloss readings at all. The largest values are obtained for HP *Everyday Semi-glossy* (SE) paper, which only exhibits a mid-range gloss, while the two glossiest papers, *DP-Euro Offset gloss* (DP) and HP *Premium Plus High-gloss* (HI), share similar surface reflection constants as the least glossy non-matt paper, Xerox *Premier Matt* (MA). It may seem odd at first that C_{surf}, which contributes to the magnitude of the specular lobe, does not appear to be correlated to gloss, because the magnitude of the specular lobe is definitely correlated to gloss, albeit moderately. However, the height of the specular lobe is determined by the product between the specular constant, the probability distribution function and the Fresnel term. The latter two quantities are now known, and the log of the lobe is well correlated to gloss measurements. The specular component is easily obtained from these data. The correlation between the estimated and the best-fitting parameters is excellent (r^2 = 0.991).

9.4 Characterisation models

The second stage of the workflow, whereby colour and gloss are mapped to BRDFs parameters, has now been treated. However, the workflow models BRDFs from CMYK data.

We therefore need colour and gloss characterisation models to transform CMYK data into colour and gloss values, from which the BRDF parameters will finally be generated.

9.4.1 Colour characterisation models

Since the subject of characterisation models for printers was extensively treated in section 4.5.2.2f) , only the outcome of the process will be stated here. A standard Neugebauer model (Sharma, 2002) was used for the colour characterization of all paper types. In order to avoid inter-instruments agreement issues, the average body reflections measured using the spectrogoniometer were used rather than a spectrophotometer's output. This will not affect the results as a strong linear correlation exists between the two sets of data. The results can be found in Table 31. The accuracy achieved was judged as satisfactory.

Table 31 - Accuracy of colour and gloss characterisation models. The colour errors are expressed in ΔE^{*}_{00} units and the gloss errors in gloss units.

	DP	HI	IS	SE	GL	MA	IM	NA
Gloss	5.70	5.85	4.10	2.20	5.08	3.74	1.33	1.38
Colour	3.09	3.90	3.14	3.83	3.26	3.58	4.63	4.82

9.4.2 Gloss characterisation models

Unlike colour, modelling gloss from CMYK input is a topic that has hardly been broached by the research community. Only a handful of models were published (Kuo, Ng, & Wang, 2002b; Ng, 2005) on that subject, and they are all based on the same premise. Gloss is a manifestation of surface reflections, and its characteristics are dictated by the geometrical features of a material's surface. Therefore the parameters of a gloss model should be related to the elements that may perturb the topography of such a surface. In the case of printing, the tiny dots of toners or inks that are transferred onto paper will only be partially absorbed, resulting in the materialization of small asperities at the surface. Since most prints are generated using halftoning techniques, the roughness of a hardcopy will then be determined by the fraction of the image that is covered by ink or toner dots compared to the area of bare substrate between the dots. *"Except for the uncoated paper, we can assume that the roughness on the surface of the paper is negligible. Hence, when toner particles begin to adhere on a paper surface, the profile of the surface becomes*

rougher. This results in the reduction of measured gloss for paper with medium and high gloss. However, when a significant amount of toner particles are laid down on the paper surface, gaps between toner particles begin to disappear, which results in smoother surface. This, in turns, increases gloss" (Kuo, Ng, & Wang, 2002b). This is especially true for xerographic printers, whereby toner particles are fused onto a medium by thermal energy and mechanical pressure. This is less relevant in the case of inkjet technology, because inks are for the most part absorbed by the substrate onto which they are laid. The variations in gloss as a function of colorant coverage could thus be expected to be less for inkjet printers, but surprisingly this is not the case. Four of the studied paper types were investigated, two printed on an inkjet printer, and the remaining two on a xerographic printer. Colour ramps for each of the primaries were printed and their gloss levels evaluated. The resulting curves are shown in Figure 68.

The observed changes in gloss do not correspond to what Kuo *et al.* reported, as the greatest variations involve the papers printed on an inkjet technology. Moreover, only one of the papers present the type of fluctuations detailed by Kuo *et al.*, as the other three exhibit mostly monotonic behaviours. More precisely, the variations of gloss in each colour ramp, even for the four paper types not represented here, can be reasonably well described by simple polynomial functions. Nevertheless, all the observations presented here are in agreement with the basic principle of the model outlined by Kuo *et al.*, i.e. gloss is a function of the surface geometry, the CMYK coverage controls the quantity of asperities generated by the printing process, and this change in roughness can be related to the gloss level perceived. The model devised by Kuo *et al.* clearly evokes the ideas utilised in the Neugebauer model. The predicted gloss level is assumed to be the weighted average of the gloss generated by each individual colorant. As for the Negebauer model, the weights are also obtained from the Demichel equations, i.e. from each colorant's relative coverage. However, there are two significant departures. Firstly, only the four primaries are considered, and not all their combinations. Secondly, the gloss generated by a colorant is not fixed, but is obtained from a polynomial function of its coverage. Given the soundness of its foundations, this model was adopted to provide the link between CMYK input data and gloss estimates. The performance for each paper type is summarised in Table 31. The testing

sample was composed of 64 randomly chosen colours. Overall, the accuracy was deemed to be suitable for the task at hand.

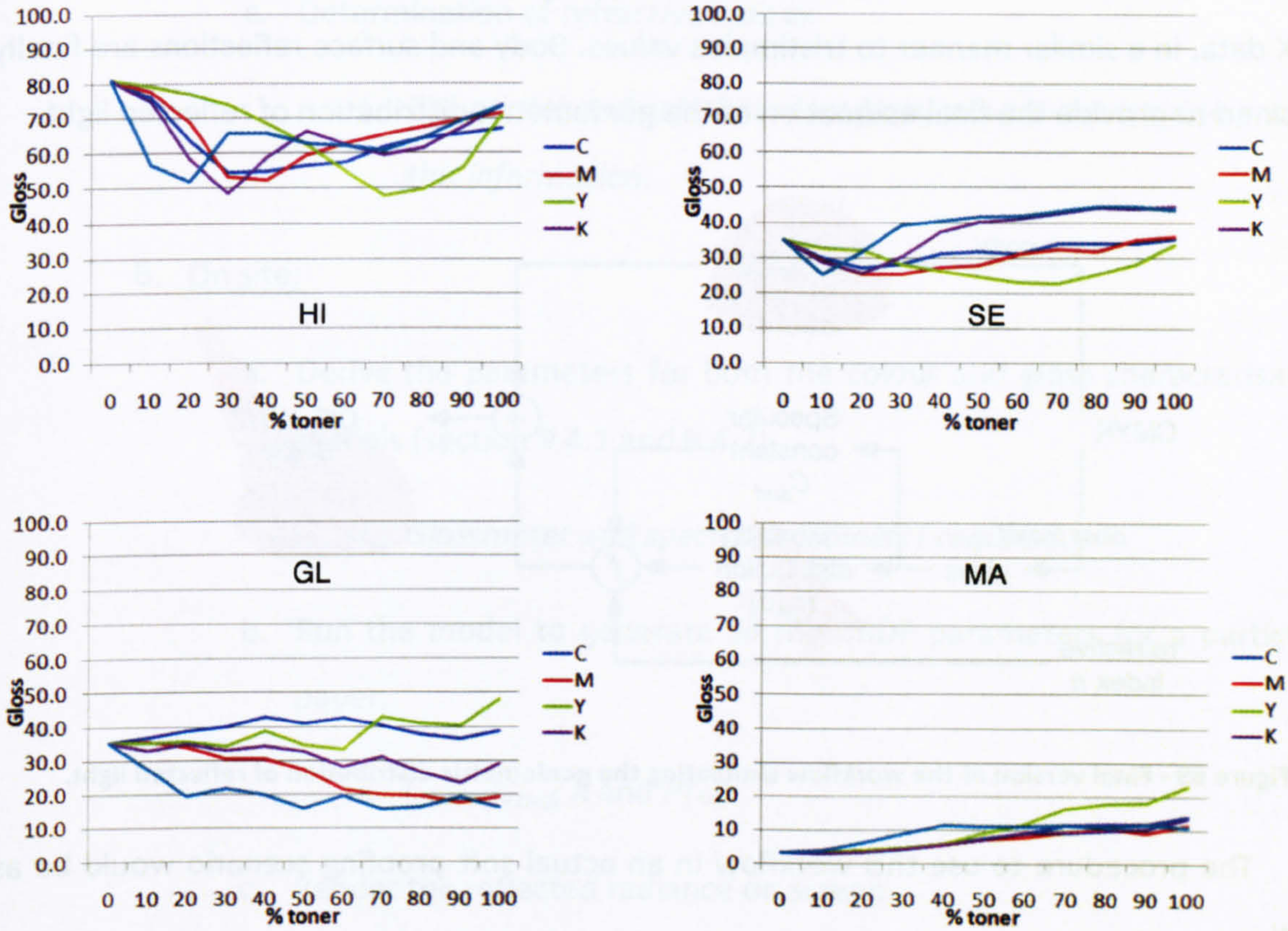


Figure 68 - Variations of the gloss of the four primaries as a function of colorant coverage. Four paper types were tested.

9.5 Synopsis of workflow

The final version of the workflow can now be unveiled (Figure 69), since all components have been treated in full. A hemispherical distribution of reflected colours is generated in all three tristimulus channels from simple CMYK colorant coverage values. The Cook-Torrance model, a well-known model used in the computer graphics field, provides the BRDF simulation. Since it is a dichromatic model, it assumes that reflections can be divided into two independent components. Body reflections are estimated from simple colorimetric data. On the other hand, surface reflections require three parameters, refractive indices, probability distributions and a constant. The refractive index cannot be linked with gloss. However, it is very reasonable to imagine that the manufacturer of the printer, the paper or the colorant could provide it. It is therefore used as a fixed input to the workflow. Since only

the real part of the refractive index is employed, we also make the reasonable assumption that a single value per channel is sufficient, since all colorants are made of similar organic components. The remaining two parameters, the probability distributions and the constants, can be accurately derived from gloss. The gloss ratings can themselves be obtained from CMYK data, in a similar manner to tristimulus values. Body and surface reflections are finally combined to provide the final estimation of the goniometric distribution of reflected light.

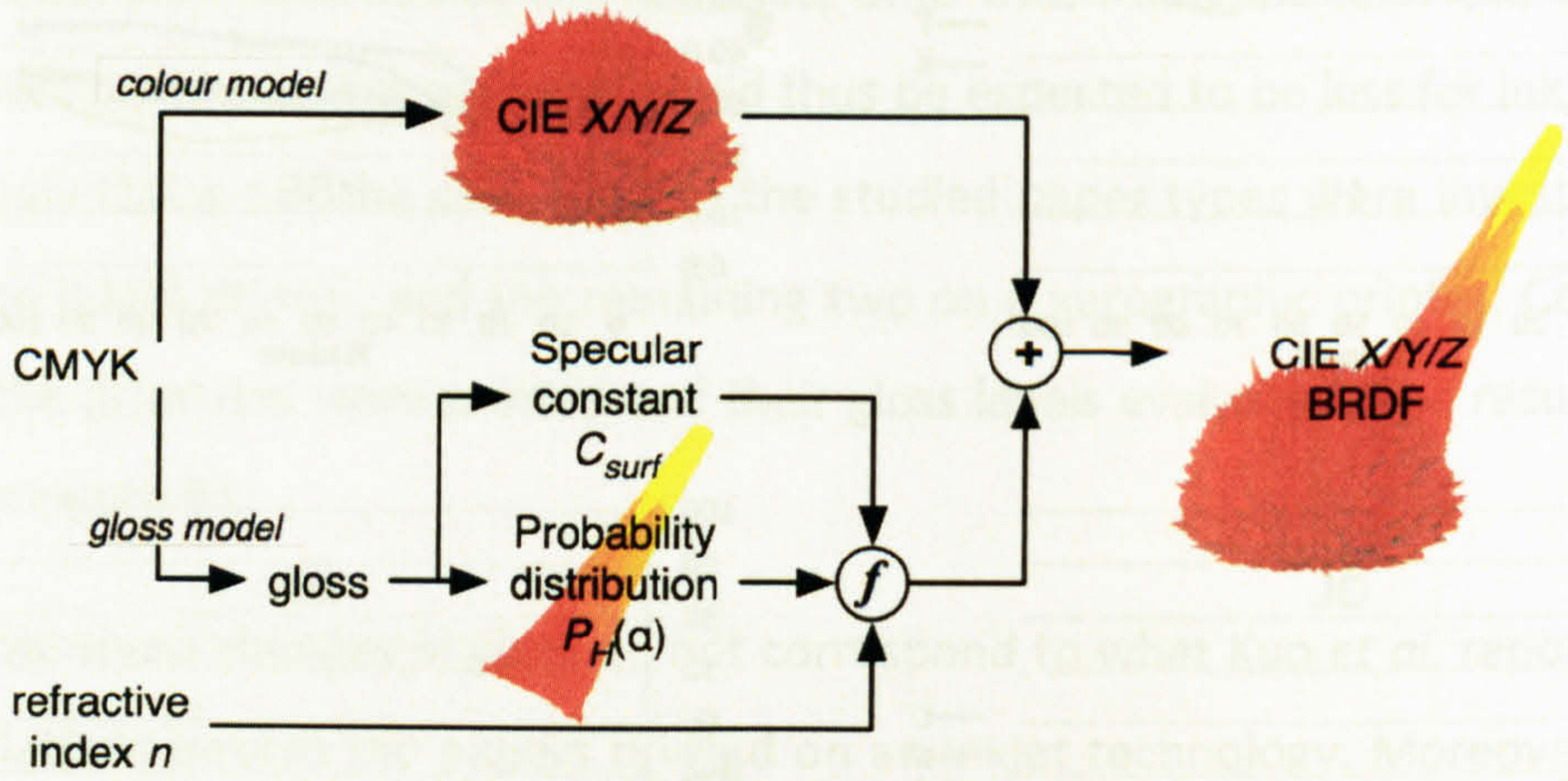


Figure 69 - Final version of the workflow simulating the goniometric distribution of reflected light.

The procedure to use this workflow in an actual soft proofing scenario would be as follow.

A. <u>In laboratories:</u>

   a. Determination of the characteristics of surface reflections, $P(\alpha)$ and C_{surf} (section 9.2.3)

      i. *Measurements of the BRDFs of the sixteen combinations of the four primaries are required to determine the parameters describing the characteristics of surface reflections.*

   b. Derive a link between , $P(\alpha)$ and C_{surf} and gloss (section 9.3.2 and 9.3.3).

      i. *Strictly speaking, this should be performed for every combinations of paper types and colorants to be employed, but experience shows that paper exhibiting similar gloss levels have*

remarkably similar BRDFs Hence a more practical approach would be to derive that function only once from a large set of paper types with different gloss levels.

c. Determination of refractive indices

i. *The printer and colorant manufacturers are expected to provide this information.*

B. <u>On site:</u>

a. Derive the parameters for both the colour and gloss characterisation models (section 9.4.1 and 9.4.2).

i. Glossmeter and spectrophotometer required.

b. Run the model to generate all the BRDF parameters for a particular paper.

i. C_{body}, C_{surf}, n and $P(\alpha)$.

c. Render the reflected radiance on screen.

i. *This implies having a lighting model, which is not described here. Please refer to the conclusion for a discussion on the topic.*

9.6 Goodness of fit

The accuracy of the whole workflow will now be analysed. A preliminary assessment of the diverse metrics available needs however to be made first.

9.6.1 Performance metrics

The most appropriate performance metric to assess the quality of this workflow would be one that gives information about the differences in perceptual terms. Colour difference formulae, such as ΔE^*_{00}, would thus seem to be an obvious choice. However, the exponential nature of the specular lobe seriously complicates matters. Indeed, if body reflections exhibit radiances that are typical of relative colorimetry, i.e. in a range between 0

and 100, the specular lobe may reach radiances of more than 20000 for the glossiest paper. The definition of the white point therefore becomes crucial in such cases, and that is where all the problems originate. Adopting the conventional definition for white points, i.e. the tristimulus values for the bare substrate, do not make any sense, as absolute differences as high as few thousands ΔE^*_{00} may be generated in such cases. Moreover, this kind of errors may occur even if the specular lobe is perfectly well modelled. A slight bend of the specimen may for instance shift the location of the recorded specular lobe by few degrees, which will lead to enormous errors.

One could argue that the colorimetric attributes of the specular lobe should be used instead as the white point. However, the chromaticities of surface reflections vary greatly, which leads to an ambiguous decision as to which set of attributes should be employed. But even so, it is not clear whether such formulae are valid or not in this case. Indeed, colour appearance models, and by extension colour difference formulae, are predicated upon the fact that the state of adaptation is *single and steady*, which is highly doubtful in the case of surface reflections. How could the human visual system be simultaneously adapted to colour *and* gloss, when there are four orders of magnitude of differences between them? There is therefore ample evidence that colour difference formulae are not suited for the task at hand. They are probably not entirely irrelevant, and may under certain circumstances provide pertinent information. However, there is no guarantee that they still relate to human judgements, nor that they provide the same kind of accuracy.

Purely mathematical performance metrics will be used instead. Even if this type of quality metrics do not equate with human subjective perception, they will nevertheless provide a rational judgement of whether the outcome of the physical interactions between light and matter interactions are being simulated correctly. After all, even though the final aim consists of replicating the appearance of prints, what is really being reproduced here is a purely physical entity, the goniometric distribution of reflected light. Such metrics are thus perfectly relevant in our context. Three different metrics are particularly useful for our purposes. First of all, the coefficient of determination r^2 was chosen because it presents the great advantage of being immune to some of the errors originating from geometrical noise. The second metric selected is the peak signal-to-noise ratio (PSNR), which is the ratio between the maximum power of a signal and the power of residuals that affects the fidelity

of its reproduction. It is usually expressed in a decibel scale, hence it is well suited for handling the exponential nature of BRDFs. This metric is often used in the image compression field to measure the quality of a reconstruction, and values above 30 dB are considered as good and above 40 excellent, while scores below 20 are poor. The final metric used in this study is the root mean square error of approximation (RMSEA), which is a statistical measure of goodness of fit. It compares the residual errors to a model's degrees of freedom (15 in our case), and models with an RMSEA of 0.5 or less are considered as good, whereas models whose RMSEA is 0.10 or more are a poor fit to the data. The exact expression of these three metrics may be found in any engineering and/or statistics textbook.

9.6.2 Sample sets

Given the difficulty of obtaining BRDF measurements, it was not possible to record enough readings to generate both a training and a testing data sets. The performance of the workflow will thus be evaluated with data that were used to derive its parameters, which may hinder the validity of the findings. However, such cases are usually problematic for simple and unique processes, whereby outputs are obtained directly from inputs via a single mathematical expression with a few parameters. In our case, the whole workflow is much more complex, as it includes multiple stages and pathways. The BRDF parameters that are being used for the simulation are not the best-fitting ones that were derived mathematically, but rather are numerically computed from estimates of gloss, which are themselves derived from CMYK colorant coverage. We therefore consider such an evaluation to be valid. The performance of the workflow was thus assessed by taking into account all available BRDFs readings.

9.6.3 Visual analysis

Let's start the evaluation process by performing visually a qualitative analysis of the simulation results. It is important to note that the following results are based on actual simulations, and not simple fits. All the simulations obtained were generated from raw CMYK data. The combined three-dimensional and exponential natures of BRDFs make non-interactive visualisations of such entities rather difficult. We will therefore focus our efforts on two specific planes that highlight the most important characteristics of BRDFs (Figure 70):

- The incidence plane.
- The azimuthal plane, which is a conical plane whose incidence angle is constant and equal to the illumination's, and whose azimuthal angle varies between $-\pi/2$ and $\pi/2$.

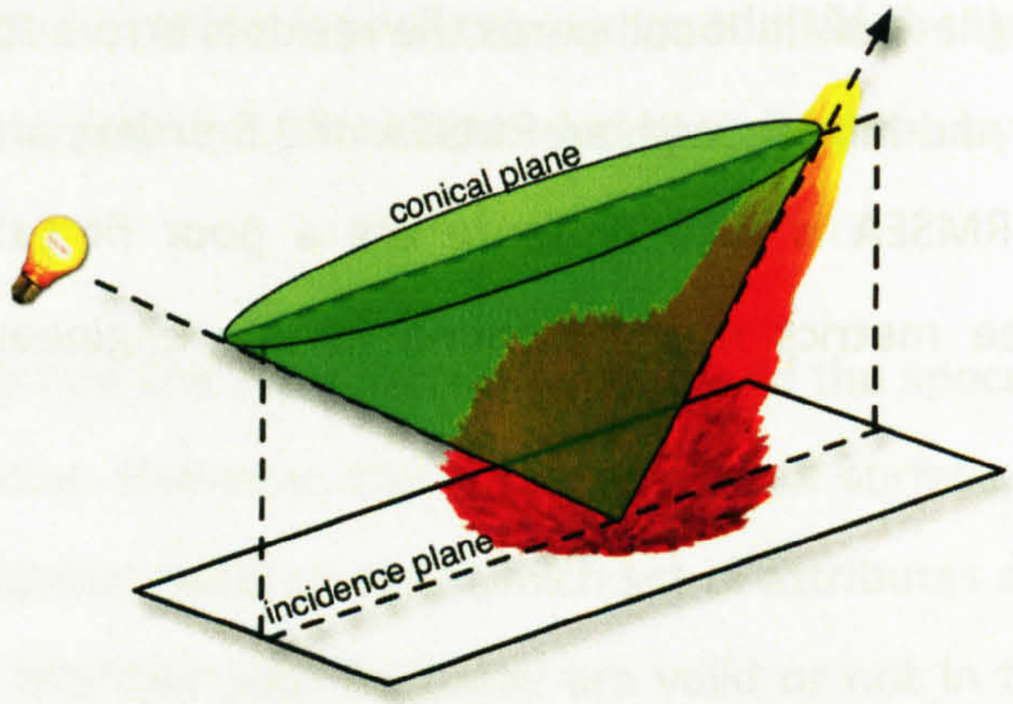


Figure 70 - Representation of the incidence and conical planes. Both planes contain the incident illumination direction and the specular direction.

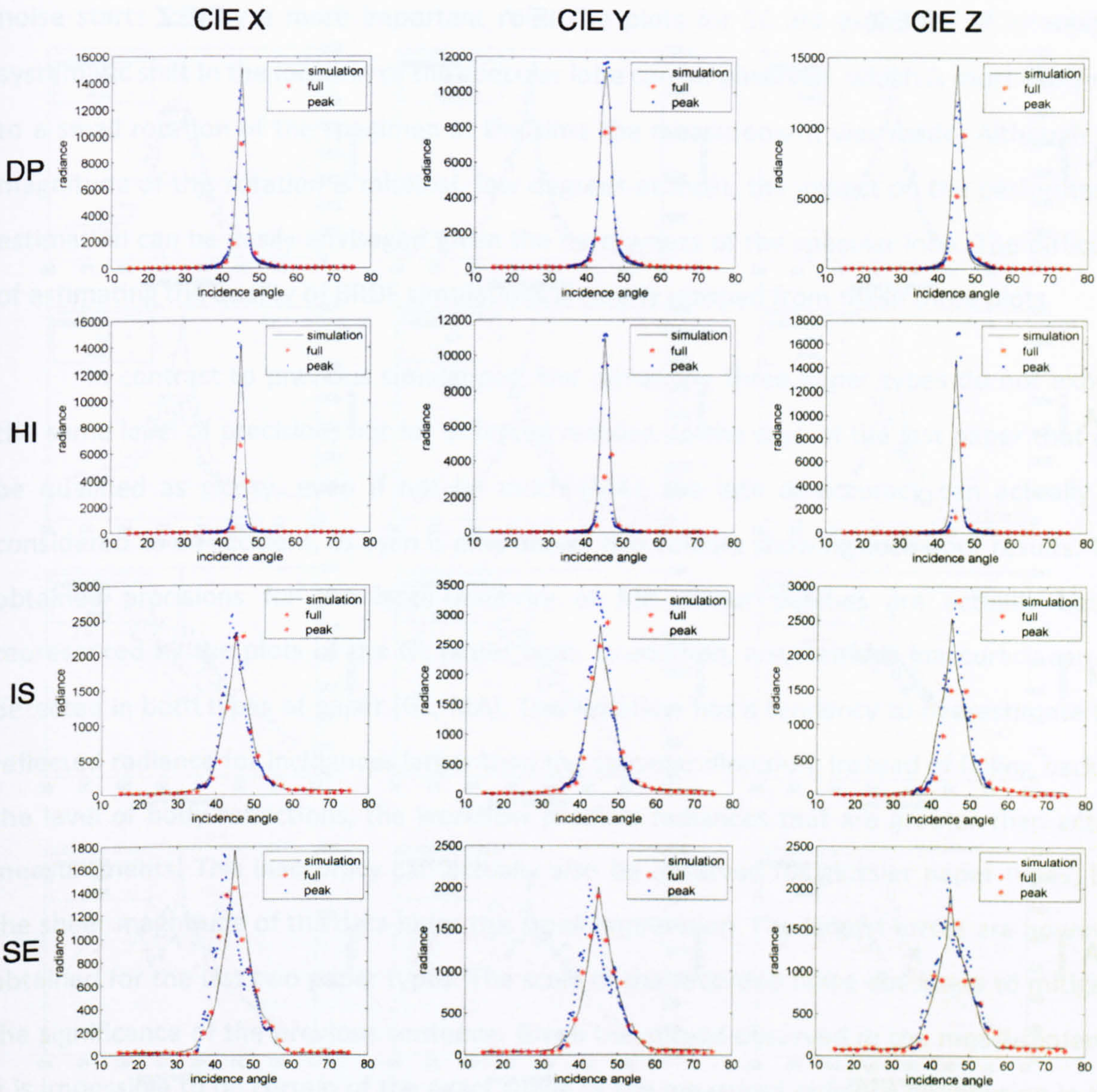


Figure 71 - Representation of the radiances measured and simulated on the incidence plane. In all cases, the depicted radiance plots correspond to a uniform cyan colour patch (100% coverage) illuminated at 45°. To each row correspond a different paper types, with decreasing gloss levels from top to bottom. In each column are represented (from left to right) the radiances of the CIE X, Y and Z channels. The red and blue dots correspond to measurements at different angular resolutions, while the black line correspond to the simulation.

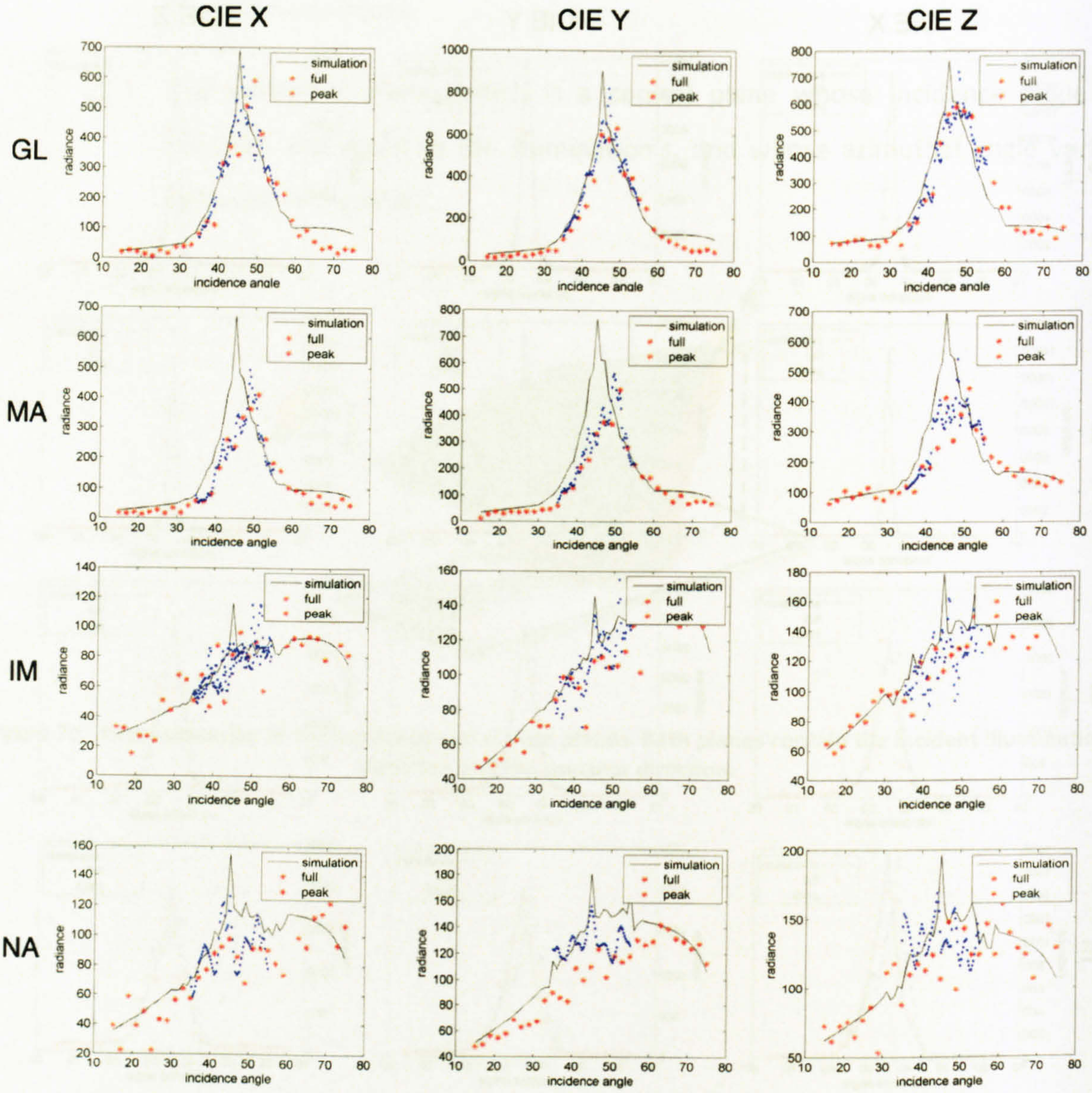


Figure 72 - see previous figure.

Overall, most of the simulations in the incidence plane (Figure 71 and Figure 72) follow the measurements very well. Although the plots all represent the radiance distribution of cyan, the results obtained for this colour are actually very representative of the general trend. The most noticeable feature of this series of plots is the accuracy diminishing as gloss ratings drop. For highly glossy specimens (DP, HI), the simulation can be qualified as excellent. The angular extent of the specular lobe is very well reproduced. On the other hand, there is a significant discrepancy between simulated and measured heights of the specular lobe. As was thoroughly explained previously, this is due to the incapacity of the workflow to estimate the chromaticity of the specular lobe. The same comments apply for the next three paper types (IS, SE, GL), which are also very well simulated, even though

noise starts to play a more important role. The plots for SE are especially of interest. A systematic shift in the location of the specular lobe can be observed, which is most likely due to a small rotation of the specimen at the time the measurement was made. Although the magnitude of this rotation is minimal, few degrees at most, the impact on the performance estimation can be easily envisaged given the narrowness of the specular lobe. The difficulty of estimating the quality of BRDF simulations is clearly grasped from these three plots.

In contrast to previous simulations, the remaining three paper types do not exhibit the same level of precision, but for different reasons. In the case of the last paper that can be qualified as glossy, even if not by much (MA), the lack of accuracy can actually be considered as an accident, as cyan is only one of two colours showing such poor results. The obtained precisions for the large majority of MA colour patches are actually better represented by the plots of the GL paper type. In addition, a systematic inaccuracy can be detected in both types of paper (GL, MA). The workflow has a tendency to overestimate the reflected radiance for incidences larger than the specular direction. Instead of falling back to the level of body reflections, the workflow predicts radiances that are greater than actual measurements. This inaccuracy can actually also be observed for glossier paper types, but the sheer magnitude of the data hides this small imprecision. The largest errors are however obtained for the last two paper types. The scale of the recorded noise does help to mitigate the significance of the previous sentence. Given the spread observed in the measurements, it is impossible to be certain of the exact shape of the measured radiance distribution in the incidence plane. Conversely, this also implies that the radiance distribution simulated should be considered as one of the possible distributions suggested by the data, and thus cannot theoretically be rejected for lack of precision. The general trend of changes is indeed approximately respected. On the other hand, the biggest flaw of the simulation is perhaps its lack of smoothness. While the surface of these matt papers is rough, and so a fair amount of fluctuations can be expected from the measurements, their reflectance distribution should be statistically isotropic, and so should the simulation. More details on the reasons of this failure will be provided later. Let's now assess the quality of the simulation on the azimuthal plane.

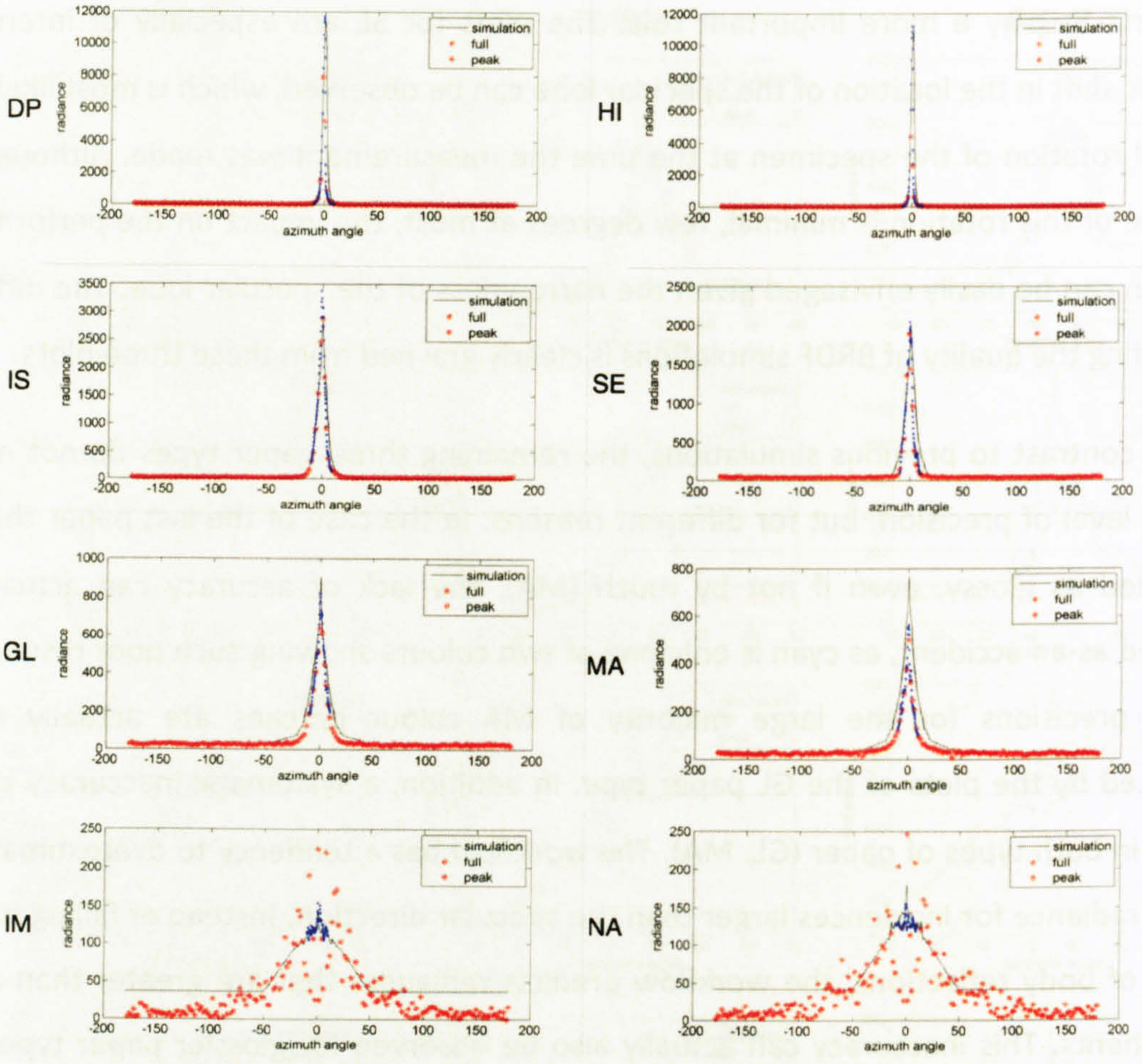


Figure 73 - Representation of the radiances measured and simulated on the azimuthal plane plane. In all cases, the depicted radiance plots correspond to a uniform cyan colour patch (100% coverage) illuminated at 45°. Only simulations of the CIE Y channel are shown, as similar behaviour were observed in the other two channels. All paper types are represented. Once again, the red and blue dots correspond to measurements at different angular resolutions, while the black line correspond to the simulation.

Very similar behaviours can be observed for the simulations on the azimuthal plane (Figure 73). The accuracy for the five glossiest paper types varies from excellent to good, and the MA paper type can also be included if the colour displayed, cyan, is discounted as an outlier. On the other hand, the precision of the matt papers simulations is even poorer than previously. While the general trend was approximately preserved in the incidence plane, none of the features of the measured radiances is properly reproduced here. The magnitude and extent of surface reflections is poor, and the magnitude and changes of body reflections is significantly overestimated. It may be thought that estimating and reproducing body reflections is a trivial task, but precisely determining the boundary between body and surface reflections is virtually impossible for such papers. The strength of the noise complicates matters further. Overall, as for the incidence plane, the precision of the

simulation on the azimuthal plane decreases with diminishing gloss ratings, and can be considered as satisfying for all paper types but the matt ones.

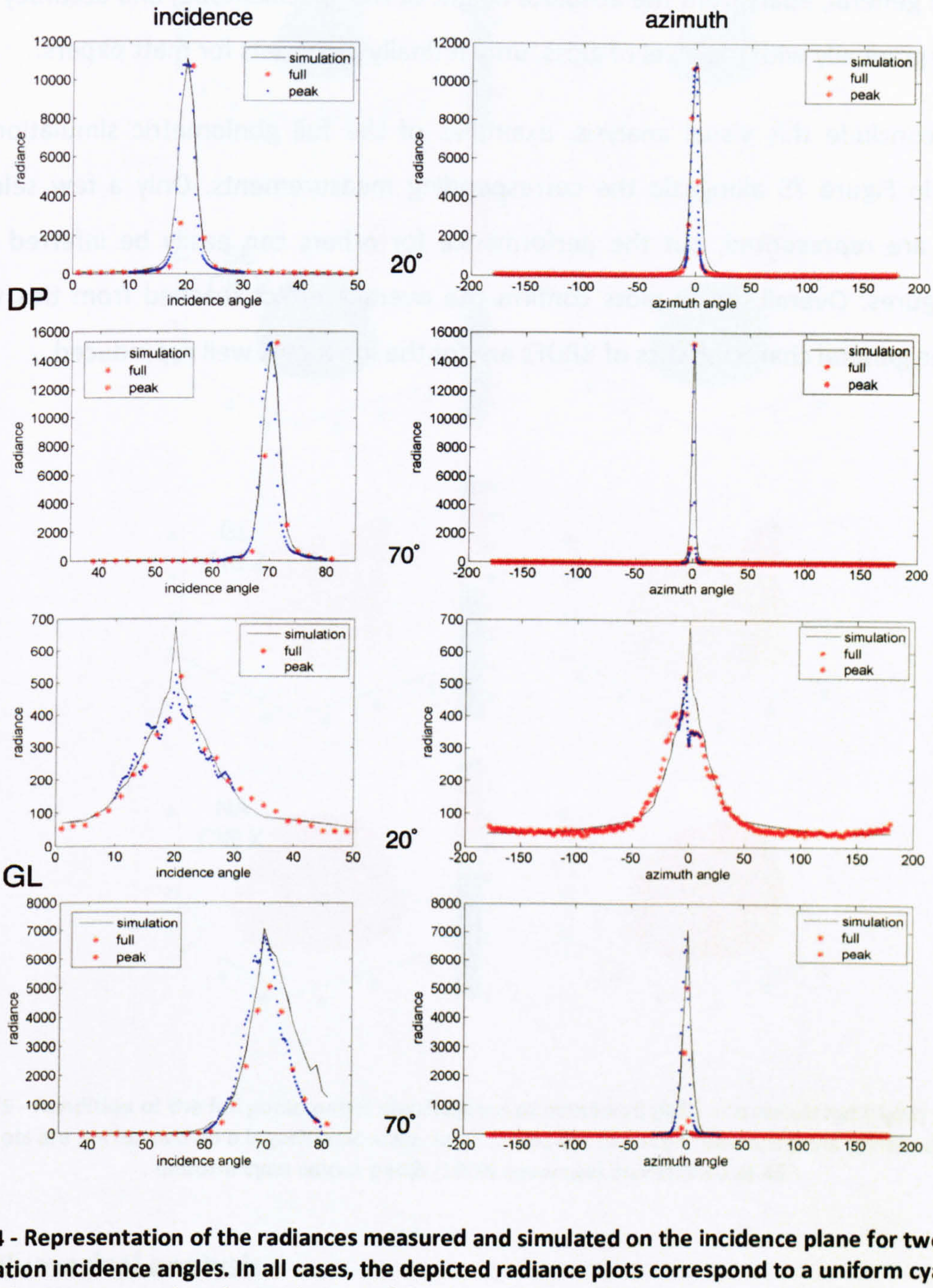


Figure 74 - Representation of the radiances measured and simulated on the incidence plane for two different illumination incidence angles. In all cases, the depicted radiance plots correspond to a uniform cyan colour patch (100% coverage). Only simulations of the CIE Y channel of two paper types are shown, as similar behaviour were observed for other channels and paper types. Once again, the red and blue dots correspond to measurements at different angular resolutions, while the black line correspond to the simulation.

A single illumination incidence angle has been considered until now. Since BRDFs predict the reflectance distribution for any incidence illumination, let's analyse the behaviour of the workflow for two other incidence angles, 20° and 70°. Space considerations

only permit to present a small subset of the thousands of the simulations generated (Figure 74). Nevertheless, the observations were once again similar. Simulations for glossy papers are good in general, apart from the absolute height of the specular lobe, and accuracy tend to diminish gradually with the level of gloss, until it finally plummets for matt papers.

To conclude this visual analysis, examples of the full goniometric simulation are presented in Figure 75 alongside the corresponding measurements. Only a few selected specimens are represented, but the performance for others can easily be inferred from previous figures. Overall, these plots confirm the overall verdict reached from this visual analysis: the general characteristics of BRDFs are for the most part well reproduced.

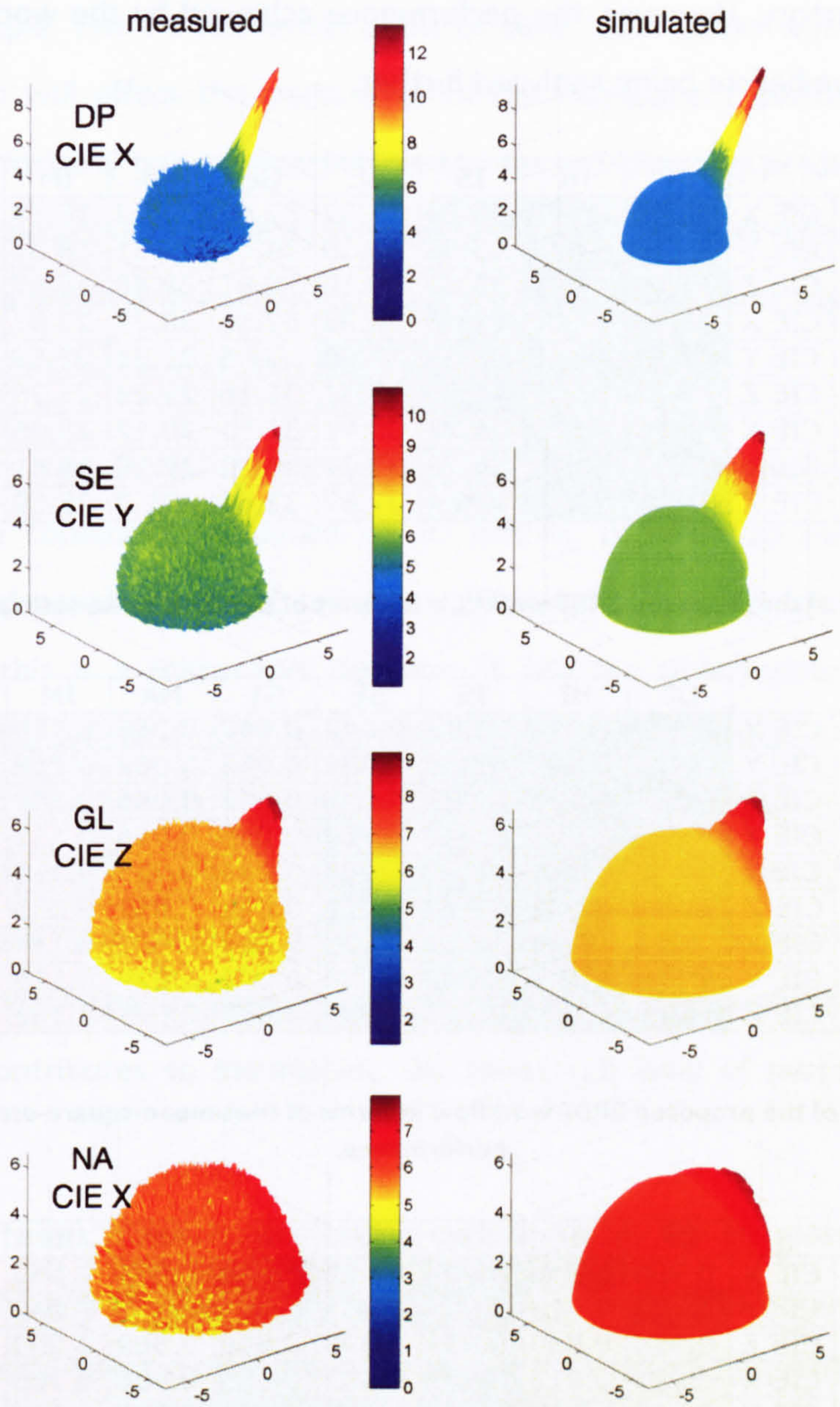


Figure 75 - Rendition of the full goniometric distributions of measured (left) and simulated (right) radiances. All plots are performed on a logarithmic scale. In all cases, the depicted radiance plots correspond to a uniform cyan colour patch (100% coverage) illuminated at 45°.

9.6.4 Numerical analysis

The numerical accuracy of the whole workflow was tested using the three performance metrics previously introduced. Figure 76, Figure 77 and Figure 78 present the results obtained for all paper types, all three colour channels and three different incidence angles. Overall, the findings obtained from the visual analysis are confirmed. The general accuracy achieved by the proposed model is good. The results for the six glossy papers are

particularly satisfactory. However, the performance achieved by the workflow needs to be put into perspective before being analyzed further.

		DP	HI	IS	SE	GL	MA	IM	NA
20°	CIE X	38.44	38.73	36.28	32.32	32.62	32.42	18.78	20.33
	CIE Y	43.50	41.60	35.58	33.39	34.01	32.51	18.87	20.59
	CIE Z	41.90	41.82	35.84	33.10	29.61	29.48	18.97	20.50
45°	CIE X	38.83	37.86	39.4	34.33	32.35	30.79	23.67	22.2
	CIE Y	42.22	40.45	40.13	36.68	32.6	31.21	22.59	22.2
	CIE Z	38.73	37.71	41.26	35.31	31.16	29.28	23.76	21.75
70°	CIE X	39.83	39.19	36.27	31.91	31.51	30.17	25.66	26.89
	CIE Y	42.02	41.69	36.76	32.27	31.22	28.16	28.41	30.65
	CIE Z	42.97	42.3	36.4	32.67	33.05	31.7	28.59	28.93

Figure 76 - Accuracy of the proposed BRDF workflow in terms of signal-to-noise-ratio performance (in dB).

		DP	HI	IS	SE	GL	MA	IM	NA
20°	CIE X	0.022	0.027	0.030	0.042	0.047	0.066	0.159	0.138
	CIE Y	0.026	0.027	0.028	0.038	0.053	0.062	0.158	0.126
	CIE Z	0.020	0.024	0.029	0.036	0.042	0.065	0.149	0.121
45°	CIE X	0.027	0.024	0.025	0.032	0.056	0.055	0.117	0.140
	CIE Y	0.023	0.023	0.026	0.033	0.043	0.065	0.116	0.125
	CIE Z	0.021	0.025	0.023	0.031	0.042	0.052	0.133	0.132
70°	CIE X	0.024	0.023	0.029	0.035	0.048	0.072	0.114	0.115
	CIE Y	0.025	0.022	0.027	0.036	0.044	0.079	0.116	0.128
	CIE Z	0.023	0.026	0.027	0.034	0.037	0.075	0.102	0.131

Figure 77 - Accuracy of the proposed BRDF workflow in terms of root-mean-square-error of approximation performance.

		DP	HI	IS	SE	GL	MA	IM	NA
20°	CIE X	0.908	0.956	0.952	0.925	0.949	0.941	0.560	0.735
	CIE Y	0.965	0.913	0.963	0.937	0.959	0.958	0.642	0.760
	CIE Z	0.924	0.909	0.931	0.898	0.866	0.869	0.553	0.614
45°	CIE X	0.927	0.915	0.971	0.902	0.925	0.906	0.708	0.669
	CIE Y	0.908	0.935	0.981	0.936	0.935	0.923	0.755	0.711
	CIE Z	0.913	0.892	0.966	0.924	0.871	0.802	0.692	0.664
70°	CIE X	0.946	0.964	0.896	0.807	0.896	0.824	0.902	0.896
	CIE Y	0.933	0.957	0.909	0.829	0.892	0.831	0.899	0.899
	CIE Z	0.953	0.961	0.920	0.849	0.905	0.851	0.878	0.868

Figure 78 - Accuracy of the proposed BRDF workflow in terms of r^2 performance.

Unlike colour measurements, noise plays a much greater part in BRDF measurements and comes from various sources. The geometrical configuration of the measurement procedure is especially important. On a macroscopic scale, a slight change in geometry will have a significant impact on the data forming the specular lobe, whereas colour measurements are oblivious to it. Even the slightest rotation of the sample being measured will modify the lobe's spatial location. Its width will also be significantly affected by any bend

present in the sample. From a numerical point of view, such issues will be considered as inaccuracies, which will affect the pertinence of the subsequent performance estimate. Although this is intrinsically not an issue inherent to the workflow, its predicted accuracy will be lowered as a consequence. For reference, the two glossiest papers measured here have a specular lobe whose width is less than two degrees at mid-height. The angular resolution of the spectrogoniometer is also of crucial importance in such cases.

Other factors that contribute towards reducing the overall accuracy of the workflow originate from the constraints imposed on it. Indeed, in order to ensure its viability, glossmeters were deemed the most suitable apparatus to provide correlates of surface reflections. While this is a reasonable decision, it has the unfortunate disadvantage of prohibiting the measurement of the specular lobe's chromaticity, and thus also of reproducing correctly its heights in all colour channels. Secondly, body reflections are also affected, which is especially crucial because, for any of the glossy papers, it represents more than 99% of the total hemispherical distribution of reflectance. As previously mentioned, body reflections do not behave strictly speaking as Lambertian diffusers. The combination of all these factors contributes to diminishing the measured level of performance from its optimal value.

In absolute terms, the accuracy of the simulation for the six glossy papers can be considered as good, and even better if all the issues mentioned above are taken into account. Their signal-to-noise ratio is almost always higher than 30dB, and sometimes even higher than 40 dB, which is excellent. Although the correlation between simulated and measured data is not always greater than 0.9, the visual inspection of the radiance distributions showed the overall width and height of the lobe is generally very well reproduced by the model. The model scales down from very high to intermediate to very low gloss in a very graceful manner, but the figures confirmed the previous conclusions for totally matt samples. The precision of the simulation obtained for them is much lower than any other specimens. If the combination of the noise and the low radiance magnitudes characteristic of matt samples amplifies this lack of precision, there is no denying that the proposed model offers no more than an approximate simulation of matt reflections. The reasons for this state of affairs will now be unveiled.

9.6.5 Discussion

Several issues that emerged from the development of the workflow and the analysis of its performance will now be discussed in turn.

9.6.5.1 Matt specimens failure

Every evaluation, be it visual or numerical, of the quality of the simulation for matt specimens lead to the same unanimous conclusion, i.e. their goniometric distribution of reflectance is poorly reproduced. Noise has been cited as a contributing factor, and its observed magnitude indicated that it certainly is playing an important part in explaining the mediocrity of the results for matt samples. While looking at Figure 72 and Figure 73, it is indeed trivial to appreciate how its strength may hinder the mathematical derivation of optimal parameters. However, we believe there is also a more fundamental origin to the problem. The BRDF model chosen in this study, the Cook-Torrance model, is based on premises that simply do not apply to the measurements recorded for matt specimens. Figure 79 provides a first clue corroborating this supposition. What is shown is a linear projection on the incidence plane of the hemispherical reflectance of a matt sample.

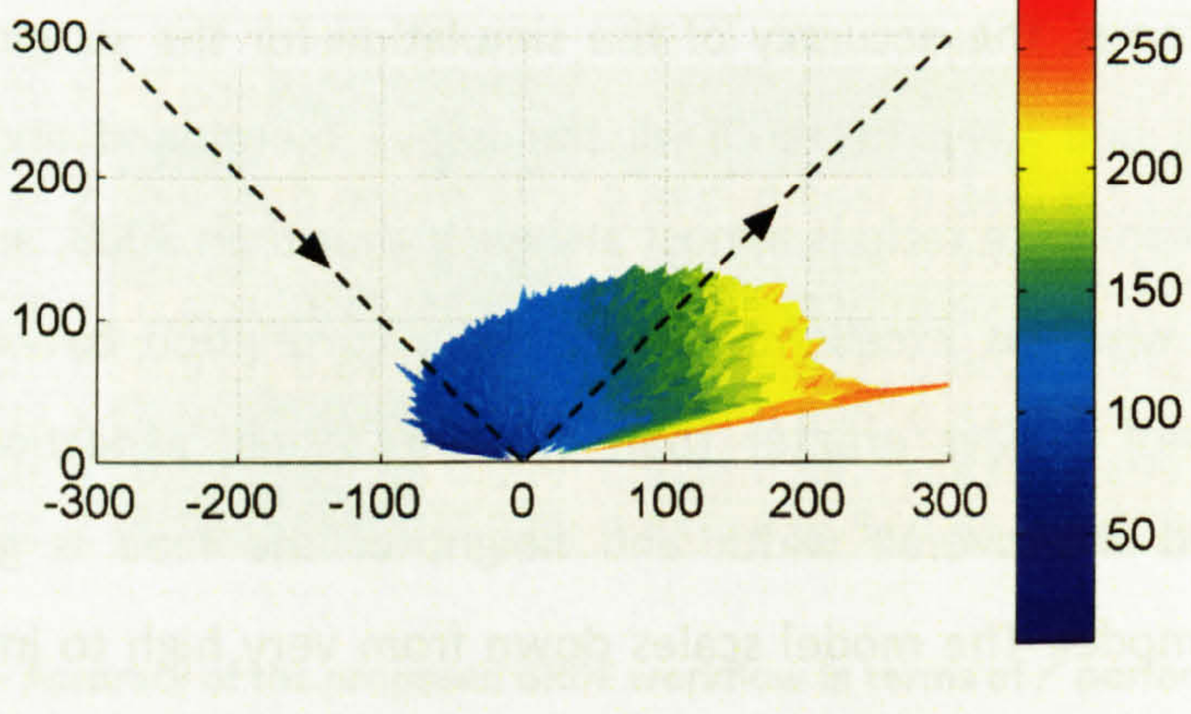


Figure 79 - Rectlinear projection on the incidence plane of the hemispherical reflectance of a matt sample. The depicted CIE Y radiance plots correspond to a uniform cyan colour patch (100% coverage).

Let's quickly recap the properties of the Cook-Torrance model: it is a dichromatic model, whereby reflectance is assumed to originate from two distinct components with specific colorimetric and goniometric characteristics. More precisely, body reflections are assumed to be Lambertian, while specular reflections supposedly generate a specular lobe whose spatial extent is limited to a certain degree around the specular direction. This clearly

does not correspond to what is being observed here. First of all, the maximum reflectance is not in the specular direction, but for larger reflection angles. Secondly, neither a specular lobe nor a transition between body and surface reflections can be detected. Instead, the changes in reflectance magnitude can be simply modelled as a linear increase of reflection angles. This can also be established from the plots corresponding to matt specimens in Figure 72. Not only this contradicts the dichromacy assumption, but it also invalidates the significance of the shadowing-masking term. Conceptually, this expression makes the reasonable assumption that the reflection is attenuated at large angles of incidence or reflection, because parts of the surface are shadowed and/or masked by other parts. On the contrary, recorded measurements of matt samples exhibit an exactly inverse behaviour, whereby the greatest reflections are observed around the largest angle of reflection accessible to the spectrogoniometer, 80°. Given this fact, one could expect to witness a similar behaviour at the other end of the range of reflection angles, but this is not the case at all. Reflectance under such angles, and also for azimuthal angles far away from the incidence plane (Figure 73), is at its lowest, to such a point that it is virtually non-existent. There is therefore ample evidence to conclude that dichromatic models such as the one proposed by Cook and Torrance are not well suited to reproduce the reflectance distribution of such specimens. Their behaviour is completely different from the predictions of such conventional models, which rationalise the failure of the workflow to accurately reproduce the goniometric reflectance distribution of matt specimens. Despite the best efforts of this study's author, no published BRDF models that could reproduce the characteristics of the observed reflectance distribution were uncovered. This could form the basis for future work to improve the proposed workflow.

9.6.5.2 Refractive indices magnitudes

Refractive indices were one of the parameters that had to be derived during the development phase of the workflow, as they are required to operate the Cook-Torrance model. It was unfortunately beyond the realm of this study to provide accurate physical measurements for the materials in question, and the parameters were derived from a mathematical regression on the BRDF measurements. While the obtained estimates did perform well in terms of overall accuracy, their absolute magnitude do not correspond to what can be expected from actual printing materials. Almost all colorants only use organic

pigments and organic resins so their refractive indices share a similar value of about 1.5. On the contrary, the ones obtained is this study deviated from this value in two different ways, which will now be justified in turn.

First of all, the refractive indices of the two glossiest paper types were above 2.0, which only occur in the natural world for rare materials such as diamonds and moissanite. Despite the high quality of these prints, this clearly indicates a failure of the regression method employed to estimate the refractive indices for these two paper types. However, the same procedure was used to evaluate the refractive indices of the other types of paper, and the results were perfectly reasonable in these cases, which naturally bring to questioning what is the peculiarity of these two paper types. A simple analysis of the regression method used is sufficient to help clarifying the issue at hand. The value of the refractive indices were obtained by maximising the overlap between the curves describing the expression $C_{surf}P(\alpha)$ for the three illumination incidence angles measured. Since the changes of body reflections magnitudes is insignificant compared to the specular lobe, this procedure essentially consists of adjusting n in order to superpose the specular lobes obtained for all incident angles. Moreover, the two affected paper types exhibit the highest gloss ratings, hence their specular lobes are the narrowest. The specular lobe of the glossiest paper is actually so narrow that it is never more than two degrees of width on average, and is often less than one degree large. Similarly, the specular lobe of the second glossiest paper is even narrower (Figure 71). Since the angular resolution of the spectrogoniometer is 0.4°, only a handful of measurements are available at most to perform the optimisation. A visual analysis of the variation of the three curves obtained for each paper as a function of refractive index demonstrates that it is very difficult to decide which value offer the best overlap. Unlike other papers, there is no optimal solution, and the differences engendered from using the optimal parameters or the true physical values are actually minor. The erroneous magnitudes of the refractive indices obtained for the two glossiest papers are therefore due to an insufficiency of data caused by the fact that the spectrogoniometer's physical limits have been reached.

The second issue related to refractive indices concern all other paper types. Unlike previously, their magnitudes are this time below the expected value, although not by much. A simple explanation for this phenomenon stems from the examination of the canonical

conditions specified to establish the reference refractive index of a certain material. No matter what method is used, ellipsometry, goniometry or spectrophotometry, a certain number of constraints are imposed on the characteristics of the specimen used. More specifically, its surface often needs to be flat and optically polished. On the other hand, the prints used in this study clearly did not and could not adhere to these recommendations. Accurately estimating and reproducing their inherent roughness was actually an intrinsic part of this study. This difference in roughness is the key to understanding why the estimated refractive indices are lower than the reference values. The influence of surface roughness on materials' optical properties has already been discussed at length during the introduction to this second part. In order to describe the behaviour of light at an interface, different physics frameworks are used depending on the scale of the surface features compared to the wavelength of light. In a similar manner, the refractive index is also influenced by surface roughness. This can conceptually be envisaged by considering Figure 45. A reference refractive index is obtained for a flat and optically polished surface, which does essentially correspond to the theoretical mean surface represented. Actual surfaces always depart from this ideal, and their topography present irregularities at a microscopic scale. The method used here to obtain refractive indices considers the surface on a macroscopic scale, i.e. measurements are integrated over a large area, which is essentially equivalent to considering the theoretical mean surface. But these two surfaces are vastly different, as parts of the real surface are at a higher altitude than the mean surface, whereas others are below it. In other words, the method employed here takes into consideration an hybrid surface, whose *virtual* refractive index varies with spatial location. In some places, it is equal to the actual refractive index of the material tested, whereas in others it is simply equal to the refractive index of air. The estimation procedure therefore generates a value that is bounded by these two fundamental indices. As a matter of fact, few studies (Nieto-Vesperinas, Sanchez-Gil, Sant, & Dainty, 1990) have been published on the topic. They reported and modelled the measured decrease in refractive index as a function of surface roughness. The concept hybrid surfaces therefore rationalises the discrepancies between estimated and standard refractive indices. To conclude, it is unlikely that obtaining the correct refractive indices would modify the overall performance of the proposed workflow. The changes induced could be compensated by other parameters, as both the probability

distribution of the microfacets orientation and the specular constant are directly derived from the values of refractive indices.

9.6.5.3 Micro-gloss

Despite the meticulous attention that has been put into the design of this workflow, one of the most salient features of gloss has been unintentionally overlooked. That trait is absolutely fundamental to the notion of gloss, and largely explains why gloss is so difficult not only to reproduce but also simply to assess. As with everything else concerning gloss, it is intimately linked with the geometry of a material's surface. "*Despite the fact that the gross level is a function of a number of different parameters, most of the gloss variations depend directly or indirectly on the topography variations at various dimensional scales. This holds for even the flattest and highest quality surfaces*" (Bui, 2003). So far, only the macroscopic and microscopic scales have been taken into consideration in this study. On a macroscopic level, every specimen was assumed to be perfectly flat, and its microscopic topography was represented by a probability distribution of microfacets' orientations. However, topographic fluctuations are not restricted to extreme scales, but small altitude variations exist at the surface of a paper at a mesoscopic scale. Unless the surface of an object is extremely smooth, its fine structure will inevitably be perceived by observers. This phenomenon is usually referred to as microgloss, and started only very recently to get attention from the research community (Hansson, 2004). As a consequence, the instruments currently available to estimate surface reflections were not developed to take microgloss into consideration. "*Commonly available gloss measurement instruments are not designed to capture these effects, but only to provide an average gloss measurement of the image region - typically on the order of several square millimetres*" (Ng et al., 2003). Small-scale gloss variations are however very important in determining the appearance of gloss. "*When the* [microfacets' orientations] *are spatially correlated, certain fine structures will become apparent when the viewing angle is very close to the specular angle of those facets. For example, the normal vectors along the toner/ink blistering exhibit strong spatial correlation, which, in turns, contributes to a significant microgloss appearance*" (Ng et al., 2003). This is perfectly exemplified by Figure 80. On the left side is represented the density scan of a colour patch that is half blue and half green, while the corresponding microgloss scan is shown on the right. The microgloss measurement was performed precisely at the specular

direction, whereas the colour scan was done significantly away from it. Although the patch appears perfectly smooth at aspecular angles, the small-scale gloss variations become very apparent in the microgloss scan. Since glossmeters perform their measurements over a relatively large area, they are oblivious to such fluctuations. Ng *et al.* repeated the procedure with diverse specimens and observed that "[a]*lthough the average of microgloss reading is found to linearly correlate with the glossmeter reading with* r^2 *statistics being over 93%, each sample possesses different microgloss structure. This indicates that it is impossible to quantify the amount of microgloss existing on a print sample using a current glossmeter*" (Ng et al., 2003).

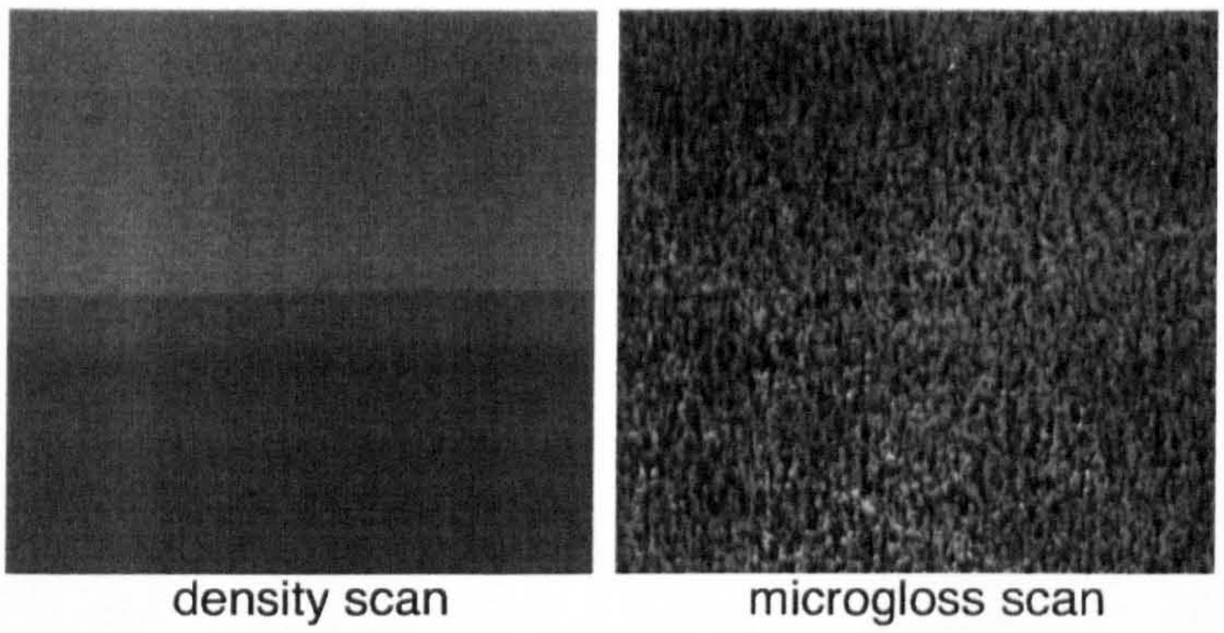


Figure 80 – Density (i.e. colour) scan (left) and microgloss scan (right) of the same hardcopy. The print is half blue and half green. (reproduced from Ng *et al.* (Ng et al., 2003)).

As can be envisaged from Figure 80, the existence of microgloss alters considerably the overall perception of gloss, as well as its appraisal. "[W]e *found that the sample which exhibits less gloss nonuniformity is perceived more frequently as being glossier than the other sample even though it has lower specular gloss reading. As a result, it is highly likely that multiple gloss attributes will contribute simultaneously to perceived gloss*" (Kuo, Ng, & Wang, 2002a). "*Studies have indicated that perceived differential gloss and even average gloss preference can be significantly altered by the presence of microgloss artifacts in the image regions*" (Ng et al., 2003). It is therefore of critical importance for the applicability of a BRDF simulation to reproduce microgloss meaningfully. At the time of this writing, no mathematical model of microgloss existed to our knowledge. Only empirical approaches have been published, whereby a surface's topography was captured using some three-dimensional imagery systems, such as photometric stereo methods (Hansson, 2004), and

simply mapped onto a geometric model of a hardcopy. Such simulations are obviously only valid for materials possessing the same optical properties, and also lack randomness, as the same topography will be reproduced for all simulations. Nevertheless, the results look very convincing, as can be envisaged from Figure 81. Moreover, the arbitrary nature of bona fide mathematical models forbid the generation of perfect matches of the topographic profile of real specimens. The use of empirical methods as a starting point could therefore be justified, as long as the simulation is performed in an appropriate manner. Furthermore, the simplicity of such approaches, both in terms of equipment and computational requirement, makes them viable in industrial settings.

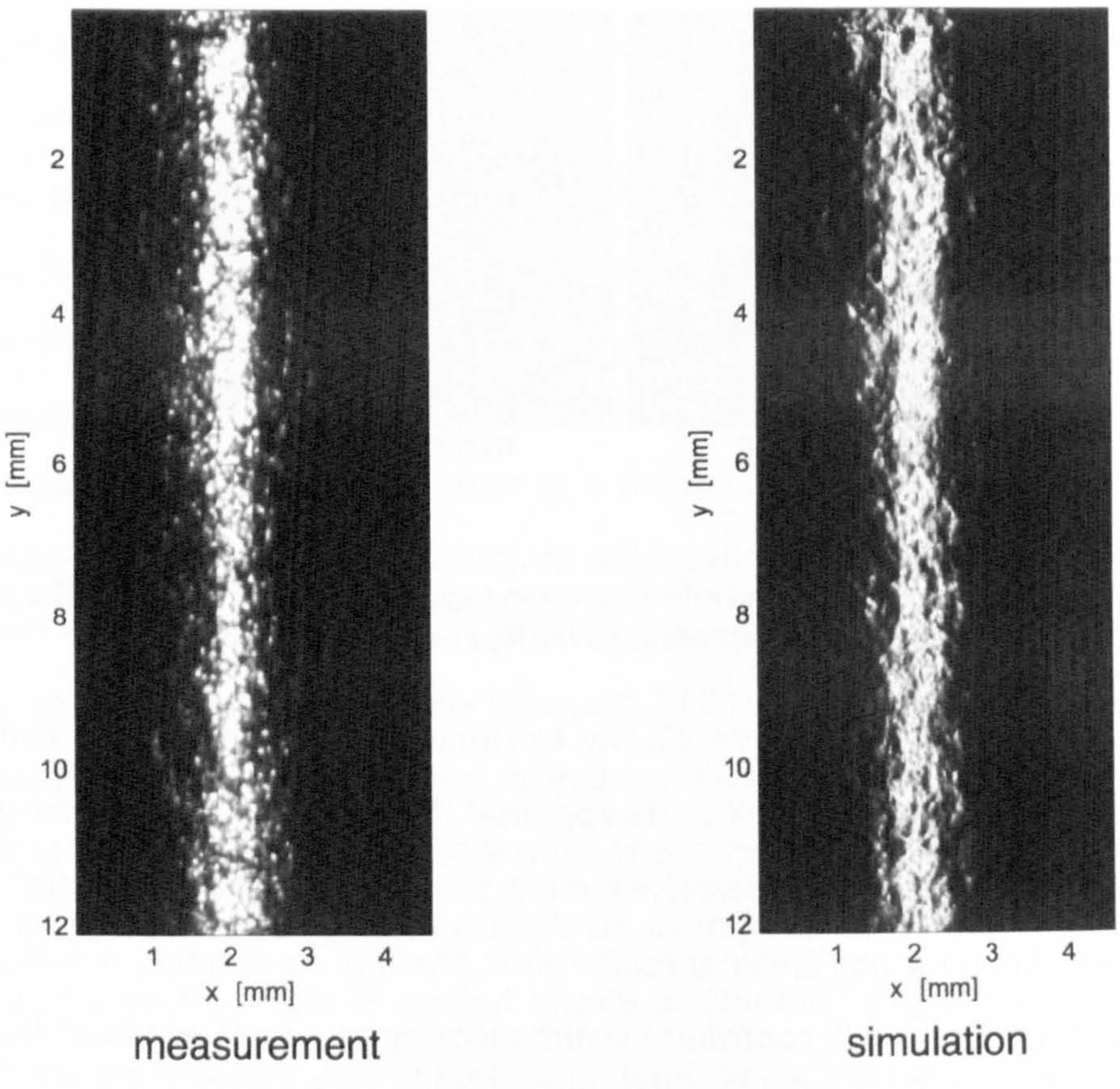


Figure 81 – Actual microgloss scan (left) versus Hansson's simulation. (reproduced from (Hansson, 2004)).

Chapter 10: Conclusion

Colour management has greatly matured over the last decade to the point that proofing, i.e. assessing the acceptability of the finished piece, can now be performed effectively on monitors, a process termed soft proofing. Soft proofing systems are now claimed to give a closer colorimetric approximation to the actual final output than conventional hard copy proofing systems (IPA, 2005). Nevertheless, it is a common experience that colour experts in industries such as Graphic Arts do not always agree that an instrumentally predicted match is a good predictor of an appearance match.

In that context, a series of experiments was first conducted to assess the suitability of soft proofs as surrogates for the final print in judging colour-reproduction quality. Incidentally, a secondary goal consisted in uncovering what might be the causes explaining the reluctance to trust soft proofs. A complex viewing apparatus was specifically designed in order to evaluate whether soft proofing can be successful under optimal conditions. The first general conclusion that was drawn is that while observers could reliably observe the media difference because of the intrinsic inaccuracies of the devices involved, they could not identify the exact type of medium used to generate the stimuli. Although obtaining an exact colorimetric match is probably impossible for complex images, the type of medium used to generate stimuli does not influence their appearance. Secondly, judgments made on the basis of soft proofs were found to be transferable to prints, subject to the fact that the tonal variations evaluated are greater than the accuracy of the colour reproduction workflow. Soft proofing was thus deemed viable under such restricted experimental conditions.

However, typical industrial settings clearly do not comply with the conditions that were imposed here. Actually, the viewing conditions under which hard and soft proofs are usually assessed are completely arbitrary. On the other hand, the experimental setup was designed to replicate the CIE-recommended measurement geometries, which specify a very strict and limited set of illumination and viewing geometries. Furthermore, most instruments employed in the Graphic Arts industry only use a pair of these geometries, 45/0 and 0/45, which are essentially equivalent. State-of-the-art colour management systems therefore perform their operation from a single sample taken out of the multitude that may be

obtained from the whole hemisphere of reflection. It was argued that the amount of samples collected to estimate the reflective properties of diverse medium is not sufficient to represent correctly the higher dimensionality of light and matter interactions. More complex interactions between light and matter are not taken into account, so what instruments measure is different from what observers perceive. This simplified approach continues to form the basis of current state-of-the-art colour management systems, and it should thus come as no surprise that such methods break down, at least partially, when employed in industrial applications, where the higher dimensionality of light and matter interactions become inevitable. Current international standards on soft proofing and viewing conditions therefore need to be reviewed and improved in order to make soft proofing successful in industrial applications.

In that context, the second part of this thesis presented means of improving the accuracy of soft proofing by increasing the amount of information rendered on screen. We presented a workflow that extends traditional colorimetry beyond simple colour appearance prediction to a higher dimensional representation, which includes gloss. The intention is to provide a full goniometric simulation of the spatial distribution of light reflected by a print (BRDF) based upon data gathered with simple and inexpensive instruments, glossmeters and spectrophotometers. The proposed workflow is essentially an extended version of standard colour characterisation models. Instead of providing a single output (CIE XYZ), the workflow estimates the entire distribution of reflected colours under all possible angles (BRDFs of CIE X, Y and Z). Although entire BRDFs were derived from only two readings per channel, the accuracy of the simulation was found to be excellent for most paper types. The model scales down from very high to intermediate to very low gloss in a very graceful manner, but not quite all the way to totally matt samples. It was established that such specimens exhibit behaviour significantly different from the assumptions on which the underlying mathematical model is based. This issue could provide the starting point for further improvement to the proposed model. Other interesting topics with evident research potential also emerged from the performed analysis. Firstly, from a mathematical perspective, deriving and integrating an accurate model for the chromaticity of the specular lobe would certainly improve the overall precision of the simulation. Similarly, replicating precisely the goniometric properties of body reflections, especially their drop at grazing

angles, should also increase the performance. Secondly, the simulation also needs to be further refined by incorporating micro-gloss effects into the equation. As emphasized in this article, gloss is all about geometry, unlike colour. Any gloss or BRDFs simulation will be incomplete without modelling the small variations present on a mesoscopic scale. Thirdly, performing a psychophysical evaluation of the workflow performance is a *sine qua non* condition to determine its fitness for the intended purpose. While it was argued the level of performance was good in most respects, the evaluation was based on purely mathematical metrics, and not on psychometric data. That implies the underlying physics are well reproduced, but because of the inherent non-linearity of the human visual system, it does not guarantee that this will translate into an appearance match. A first approach to resolving this issue would be the inception of a gloss difference metric, one that is equivalent and directly comparable to colour difference formulae. Another method, perhaps the most appropriate, would be to perform a full-fledged psychophysical experiment.

While this proposal sounds perfectly sensible, its development is however far more complex than it first appears. Pellacini presents the deceptive simpleness of the issue very well "[A]*ll previous gloss studies have looked exclusively at locally illuminated surfaces in uniform surrounds. This practice is understandable given the difficulty of controlling complex environments, but it's strange considering that one of the most salient things about glossy surfaces is their ability to reflect their surroundings*" (Pellacini, Ferwerda, & Greenberg, 2000). Indeed, as was repeatedly stated during this thesis, the origins of visual perception are the interactions between light and matter. These two entities epitomise perfectly the old adage that goes "*They are like two sides of the same coin*". One cannot be without the other, but there was only time to study one of them during the course of this thesis. An accurate modelling of prints' reflective properties was derived, but the subject of lighting has barely been broached. However, setting up such a computer simulation implies taking into account more than simple chromatic properties, the lighting's only attributes considered so far. As mentioned by Pellacini, reproducing the geometrical layout is absolutely paramount to the success of the experiment, as the gloss sensation is essentially a mirror image of the diverse illuminants. Furthermore, rendering in real-time the interactions between these two entities require very significant engineering efforts, as will now be briefly introduced.

As previously stated, the aim sought is to provide an interactive computer-based simulation faithfully rendering the appearance of printed paper illuminated by viewing booths typically employed in prepress houses. A critical aspect therefore concerns the generation of highly efficient representation of complex lighting environments. By far the most common used lighting models used in Computer Graphics are based on the *local illumination* paradigm, which are limited by nature to handle simple configurations, a finite set of simple light sources illuminating only few objects while discarding any eventual inter-reflections. Examples of this type of light sources include ideal point lights, directional lights and spotlights. Screenshots of a prototype application using such an illumination model are given in Figure 82. The BRDF of a medium-glossy paper exhibiting differential gloss (the black colorant is glossier than the others) was simulated using the workflow presented in the second part of this work. A fluorescent tube horizontally placed was simulated by a series of sixteen point light sources. In order to keep computational cost to a minimum, none of the complex inter-reflections between objects present in a scene are considered. Furthermore, although their abilities seem to actually suit well the characteristics of the problem at hand, i.e. a viewing booth in a dark environment, light sources covering a large area can actually only be approximated with point light sources. While this is not a problem for materials exhibiting low to medium gloss levels, a custom-built prototype application clearly established that the existence of such a stratagem became obvious with highly glossy objects. The plausibility of the simulation is

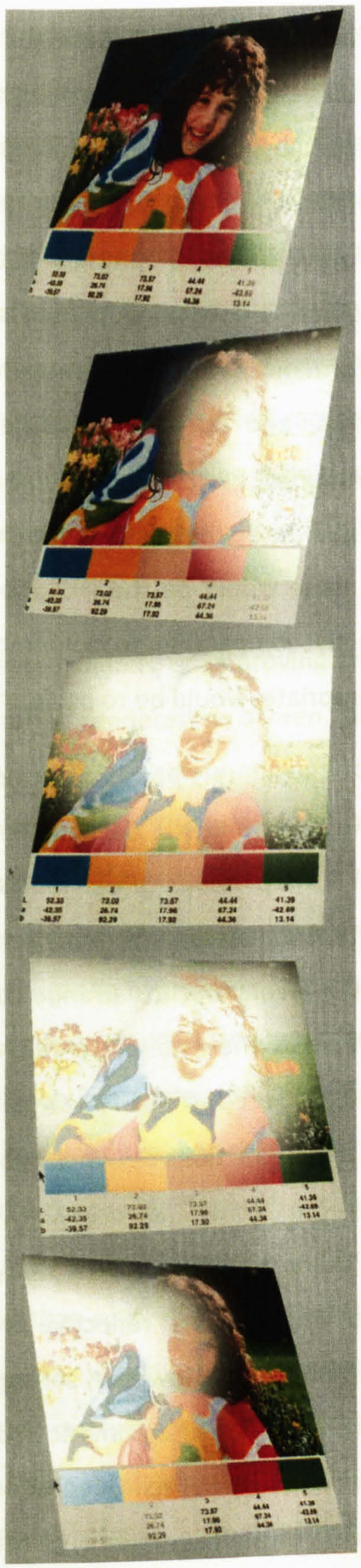


Figure 82 - Screenshots of a prototype application using a local illumination model. The BRDF of a medium-glossy paper was simulated using the workflow presented previously.

thus seriously compromised by all these factors.

Global illumination models, on the other hand, were designed to handle lifelike situations. *"The task* [...] *is to model the interplay of light among large-scale objects of an environment in order to approximate the quantity and quality of light reaching the eye of an observer. The problem is global in that all objects can interact with one another by means of scattered light"* (Arvo, 1993). This problem can be stated in the terms of the transport theory, a general technique for analysing all macroscopic phenomena resulting from the interaction of infinitesimal particles with a medium. From a radiometric point of view, the transport equation – also called radiance equation in that case - *"completely describes* [, for every point in the environment,] *the intensity of light at a point in every direction, wavelength and polarisation"* (Glassner, 1995). Despite its aesthetic, the equation is actually much too arduous to solve analytically for scenes of arbitrary complexity. Only approximate solutions can be had, such as ray tracing, radiosity, or hybrid solutions of the two. They offer the most accurate answer at the expense of complexity, i.e. their applicability is limited to offline rendering.

Another approximation to the radiance equation that satisfies the real-time constraint has emerged as one of the most exciting new techniques presented over of the last few years at the SIGGRAPH conference. Pre-computed Radiance Transfer (PRT) *"attacks a difficult problem in the area of interactive global illumination: how to shade an object lit with arbitrarily complex lighting, properly accounting for shadows and interreflections among surfaces, with the object and lighting changing over time, and all updating at interactive rates"* (Haines, 2006). From the point of view of this study, the great advantage of PRT is to offer a representation of the whole viewing environment, as if a camera had captured the whole sphere surrounding an object, rather than a discrete version, limited to a finite number of sources. Ramamoorthi and Hanrahan (2004) demonstrated that it is possible to generate a heavily compressed yet accurate representation of all the incident illumination at a point by using spherical harmonics. *"Spherical harmonics are something of a spherical analogue of a Fourier series. Just as a series of scaled and shifted sine waves of different frequencies can reconstruct a 1D signal, a series of spherical harmonics can approximate an image on a sphere"* (Haines, 2006). Similarly, a BRDF representing the reflective properties of a material is also a spherical function, hence the radiance transfer at

any location on an object can also be represented by another set of spherical harmonics. Obtaining the exiting radiance at a specific location only requires combining the two sets of spherical harmonics, which simply consists of a series of dot products (Kautz, 2003). Such a representation therefore presents tremendous advantages in our quest to generate a goniochromatic simulation of the appearance of prints. It is a completely general approach, and is the only one that is able to cope with complex interactions in real-time. The illumination conditions provided by conventional viewing booths used in the Graphics Arts industry can be exactly reproduced. As a matter of fact, not only viewing booths can be reproduced, but a simulation of an actual Graphics Arts studio can be simulated if the room is photographed in an appropriate manner. Perhaps even more startling, simulating the effects of standard CIE illuminants such as D65 can be done not only by considering their chromaticities alone, but by actually recreating the sky as it was at the time of the initial measurement. The exact position of the sun can be estimated from simple astronomical calculations, and the resulting lighting environment generates simulations that are *"very realistic"* (Green, 2003). Generating an interactive computer-based simulation faithfully rendering the appearance of printed paper, or any other materials in fact, is definitely within grasp, but the amount of work involved is sufficient to overwhelm another PhD student.

Glossary

Body reflection: reflection produced by absorption and scattering of light by colorants within a colored material.

BRDF: bi-directional reflectance distribution function.

Conductor: materials that contain movable charges of electricity, often by complex refractive index.

Dielectric: see insulator.

Diffuse reflection: reflection in which flux is scattered in many directions by diffusion at or below the surface.

Diffusion: change of the angular distribution of a beam of radiant flux by a transmitting material or a reflecting surface such that flux incident in one direction is continuously distributed in many directions. May originate from surface roughness or scattering below the surface.

Flare: light falling on an image, in an imaging system, which does not emanate from the subject point.

Geometric optics: high frequency approximation (i.e. macroscopic) to light, describing light propagation in terms of rays.

GMA: gamut-mapping algorithms.

Interface: boundary between two media with dissimilar optical properties.

Insulator: substances that are highly resistant to electric current, their extinction coefficient is often nil.

Light: electromagnetic radiant power that is visually detectable by the normal human observer, having wavelengths from about 380 nm to about 780 nm.

Medium: (optics) material through which electromagnetic waves propagate.

Rectilinear propagation: 1st law of geometric optics, states that light rays propagating through a homogeneous transparent medium propagate in straight lines.

Reflection: 2nd law of geometric optics, governs the interaction of light rays with conducting surfaces.

Refraction: 3rd law of geometric optics, governs the behaviour of light rays as they traverse a sharp boundary between two different transparent media.

Regular reflection: see specular reflection.

SLD: self-luminous display, i.e. a CRT or LCD monitor.

SP: spectrophotometer.

Specular: pertaining to flux reflected from the surface of an object, without diffusion, at the specular angle.

Specular reflection: reflection without diffusion, in accordance with the laws of optical reflection, as in a mirror.

Surface reflection: reflection perceived as belonging to the surface of a specimen, without the specimen appearing to be self-luminous.

TSR: teleradiospectrometer.

Veiling reflections: reflection of light by an image-bearing surface, that reduces the apparent contrast of the image.

References

Alessi, P. J. (1994). CIE guidelines for coordinated research on evaluation of colour appearance models for reflection print and self-luminous display image comparisons. *Color Reseach and Application, 19*(1).

Alessi, P. J. (1996). *CIE TC 1-27 - An update on colour appearance model evaluation for hardcopy/softcopy image comparison.* Paper presented at the CIE Expert Symposium '96 on colour standards for image technology.

Alessi, P. J., & Madden, T. E. (1992). *State of the art hardcopy/softcopy image-matching techniques.* Paper presented at the Comparison of color images presented in different media.

ANSI. (1999). *IT8.7/3-1993: Graphic Technology - Input Data for Characterization of 4-Color Process Printing*.

Arney, J. S., Anderson, P. G., Franz, G., & Pfeister, W. (2006). Color Properties of Specular Reflections. *Journal of Imaging Science and Technology, 50*(3).

Arney, J. S., Ye, L., & Wible, J. (2006). Analysis of Paper Gloss *J. Pulp & Paper Sci.*

Arvo, J. (1993). *Transfer Equations in Global Illumination.* Paper presented at the ACM SIGGRAPH 1993.

Askeland, D. R. (2005). *Science and Engineering of Materials*: Brooks Cole.

ASTM. (1990a). D 991 - 90: Standard practice for color measurement of fluorescent specimens. *ASTM International Standards on Color and Appearance Management.*

ASTM. (1990b). E 1345 - 90: Standard practice for reducing the effect of variability of color measurement by use of multiple measurements. *ASTM International Standards on Color and Appearance Management.*

ASTM. (1991). E 1341 - 91: Standard practice for obtaining spectroradiometic data from radiant sources for colorimetry. *ASTM International Standards on Color and Appearance Management.*

ASTM. (1999). D 4039 - 93: Standard Test Method for Reflection Haze of High-Gloss Surfaces. *ASTM International Standards on Color and Appearance Management.*

ASTM. (2003a). E 284 - 03: Standard Terminology of Appearance. *ASTM International Standards on Color and Appearance Management.*

ASTM. (2003b). E 1331 - 96: Standard Test Method for Reflectance Factor and Color by Spectrophotometry Using Hemispherical Geometry. *ASTM International Standards on Color and Appearance Management.*

ASTM. (2004a). D 523 - 89: Standard Test Method for Specular Gloss. *ASTM International Standards on Color and Appearance Management.*

ASTM. (2004b). E 167 - 96: Standard Practice for Goniophotometry of Objects and Materials. *ASTM International Standards on Color and Appearance Management.*

ASTM. (2004c). E 179 - 96: Standard Guide for Selection of Geometric Conditions for Measurement of Reflection and Transmission Properties of Materials. *ASTM International Standards on Color and Appearance Management*.

ASTM. (2004d). E 430 - 99: Standard Test Methods for Measurement of Gloss of High-Gloss Surfaces by

Goniophotometry. *ASTM International Standards on Color and Appearance Management*.

ASTM. (2004e). E 2175 – 01: Standard Practice for Specifying the Geometry of Multiangle Spectrophotometers. *ASTM International Standards on Color and Appearance Management*.

Bahar, E., & Chakrabarti, S. (1987). Full wave theory applied to computer-aided graphics for 3D objects. *IEEE Computer Graphics and Applications, 7*(7).

Bala, R. (2003). Device characterization. In G. Sharma (Ed.), *Digital color imaging handbook*: CRC Press.

Bartleson, C. J., & Breneman, E. J. (1967). Brightness perception in complex fields. *Journal of the Optical Society of America, 57*.

Berns, R. S. (1991). Color tolerance feasability study comparing crt-generated stimuli with an acrylic-lacquer coating. *Color Reseach and Application, 16*(4).

Berns, R. S. (1996). Methods for characterizing CRT displays. *Display, 16*.

Berns, R. S., & Choh, H. K. (1995). Cathode-ray-tube to reflection-print matching under mixed chromatic adaptation using RLAB. *Journal of Electronic Imaging, 4*(4).

Berns, R. S., & Gorzynski, M. E. (1991). *Simulating surface colors on crt displays: the importance of cognitive clues.* Paper presented at the AIC Colour and Light '91.

Billmeyer, F. W., & O'Donnell, F. X. D. (1987). Visual gloss scaling and multidimensional scaling analysis of painted specimens. *Color Research and Applications, 12*(6).

Blinn, J. F. (1977). *Models of light reflection for computer synthesized pictures.* Paper presented at the ACM SIGGRAPH 1977.

Born, M., & Wolf, E. (1980). *Principles of Optics* (6th ed.): Pergamon.

Bradley, E. T. (1977). USA Patent No.: U. S. P. a. T. Office.

Bradley, R. A. (1984). Paired comparisons: Some basic procedures and examples. In P. R. Krishnaiah & P. K. Sen (Eds.), *Handbook of Statistics,* (Vol. 4): Elsevier Science Publishers.

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs, I. the method of paired comparison. *Biometrika, 39*.

Brainard, D. H. (1998). Color constancy in the nearly natural image, 2: achromatic loci. *Journal of the Optical Society of America, 15*(2).

Brainard, D. H., & Ishigami, K. (1995). *Factors influencing the appearance of crt colors.* Paper presented at the IS&T/SID Color Imaging Conference.

Braun, K. M., & Fairchild, M. D. (1996). *Psychophysical generation of matching images for cross-media color reproduction.* Paper presented at the IS&T/SID Color Imaging Conference.

Braun, K. M., & Fairchild, M. D. (1997). Testing five color-appearance models for changes in viewing conditions. *Color Reseach and Application, 22*(3).

Braun, K. M., Fairchild, M. D., & Alessi, P. J. (1996). Viewing techniques for cross-media image comparisons. *Color Reseach and Application, 21*(1).

Brenner, E., Cornelissen, F. W., & Nuboer, J. F. W. (1989). Some spatial aspects of simultaneous colour contrast. In J. J. Kulikowski, C. M. Dickinson & I. J. Murray (Eds.), *Seeing contour and colour*: Pergamon, Northern Eye Institute, Oxford.

Bui, T. H. (2003). *Paper Optics: Surface roughness and non-uniformity of paper.* Linkoping University.

Cabral, B., Max, N., & Springmeyer, R. (1987). *Bidirectional Reflection Functions from Surface Bump Maps.* Paper presented at the ACM SIGGRAPH 1987.

Caivano, J. (1997). Semiotics and cesia: Meanings of the spatial distribution of light. *Colour and Psychology. From AIC Interim Meeting 96 in Gothenburg. Stockholm: Scandinavian Colour Institute, Colour Report F50.*

Chandrasekhar, S. (2003). *Radiative Transfer*: Dover Publications Inc.

Choh, H. K., Park, D. S., Kim, C. Y., & Seo, Y. S. (1996). *Effects of ambient illumination on the appearance of crt colors.* Paper presented at the IS&T/SID Color Imaging Conference.

Choi, K. F. (1994). *Comparison of color difference perception on soft-display versus hardcopy.* Paper presented at the IS&T/SID Color Imaging Conference.

CIE. (1987). *Publication CIE 17.4: International Lighting Vocabulary.*

CIE. (1994). Technical Committee 1-27: CIE guidelines for coordinated research on evaluation of colour appearance models for reflection print and self-luminous display image comparisons. *Color Reseach and Application, 19*(1).

CIE. (1995). Technical Committee 1-34: Testing colour-appearance models: guidelines for coordinated research. *Color Reseach and Application, 20*(4).

CIE. (1996). Technical Report 122: The relationship between digital and colorimetric data for computer-controlled CRT displays.

CIE. (1998). Technical Report 13.2: Method of specifying and measuring colour rendering properties of light sources.

CIE. (2000). *Technical Committee 8-04: Experimental guideline.*

CIE. (2003a). *Technical Committee 8-01: A Colour Appearance Model for Colour Management Systems: CIECAM02.*

CIE. (2003b). *Technical Report 15.3: Colorimetry.*

CIE. (2004). *Technical Report 156: Guidelines for the evaluation of gamut mapping algorithms.*

Cook, R. L., & Torrance, K. E. (1981). A reflectance model for computer graphics. *Computer Graphics, 15*(3).

Cook, R. L., & Torrance, K. E. (1982). A reflectance model for computer graphics. *ACM Transactions on Graphics, 1*(1).

Dalal, E. N., & Natale–Hoffman, K. M. (1999). The Effect of Gloss on Color *Color Reseach and Application, 24*(5).

Davies, H. (1954). The reflection of electromagnetic waves from a rough surface. *Inst. of Electr. Eng.*(101).

Dempski, K., & Viale, E. (2004). *Advanced Lighting and Materials with Shaders* Wordware Publishing Inc.

DIN. (2001). 6175-2:2001-03: Tolerances for automotive paints - Part 2: Goniochromatic paints. *Deutsches Institut für Normung e.V. (DIN).*

ELDIM. www.eldim.fr.

Emmel, P. (1998). *Modèles de prédiction couleur appliqués à l'impression jet d'encre.* École polytechnique fédérale de Lausanne, Lausanne.

Emmel, P. (2003). Physical models for color prediction. In G. Sharma (Ed.), *Digital color imaging handbook*: CRC Press.

Engeldrum, P. G. (2000). *Psychometric scaling: a toolkit for imaging systems development*: Imcotek Press.

Fairchild, M. D. (1993). *Chromatic adaptation in hard-copy/soft-copy comparisons.* Paper presented at the Color Hard Copy and Graphic Arts II.

Fairchild, M. D. (1995). Considering the surround in device-independent color imaging. *Color Reseach and Application, 20*(6).

Fairchild, M. D. (2004). *Color Appearance Models* (2nd ed.): John Wiley & Sons.

Fairchild, M. D., & Alfvin, R. L. (1995). *Precision of color matches and accuracy of color-matching functions in cross-media color reproduction.* Paper presented at the IS&T/SID Color Imaging Conference.

Fairchild, M. D., Berns, R. S., & Lester, A. A. (1996). Accurate color reproduction of crt-displayed images as projected 35-mm slides. *Journal of Electronic Imaging, 5*(87-96).

Fairchild, M. D., & Johnson, G. M. (1999). Color appearance reproduction: visual data and predictive modelling. *24, 2.*

Fairchild, M. D., & Reniff, L. (1995). Time-course of chromatic adaptation for color-appearance judgements. *Journal of the Optical Society of America A, 12.*

Fitzpatrick, R. (1999). *Electromagnetism and Optics, a non-calculus survey course*. from http://farside.ph.utexas.edu/teaching/302l/lectures/lectures.html.

Francis, M. A., & Irwin, R. J. (1995). Decision strategies and visual-field asymmetries in same–different judgments of word meaning. *Memory & Cognition 23*(3).

Gescheider, G. A. (1997). *Psychophysics: The Fundamental*: Lawrence Erlbaum Associates.

Glassner, A. S. (1995). *Principles of digital image synthesis*: Morgan Kaufmann.

Granberg, H. (2002). *Paper optics and perception: Spectral light absorption and gloss of cyan & magenta inks*: Acreo Tryckteknisk Forskning, Report No. acr009805, Norrkoping.

Green, P. (1999). *Understanding Digital Color*: Graphic Arts Technical Foundation.

Green, R. (2003). *Spherical Harmonic Lighting: The Gritty Details* Sony Computer Entertainment America

Greenberg, D. P. (1996). *The symbiotic relationship between computer graphics and colour imaging.* Paper presented at the IS&T/SID Color Imaging Conference.

Haines, E. (2006). An Introductory Tour of Interactive Rendering. *IEEE Computer Graphics and Applications 26*(1).

Hall, R. (1989). *Illumination and Color in Computer Generated Imagery*: Springer-Verlag.

Handley, J. C. (2001). *Comparative Analysis of Bradley-Terry and Thurstone-Mosteller Paired Comparison Models for Image Quality Assessment.* Paper presented at the IS&T PICS.

Hansson, P. (2004). Simulation of small-scale print gloss variations based on topographic data. *NORDIC PULP AND PAPER RESEARCH JOURNAL, 19*(1).

Harrison, V., & Poulter, S. (1951). Gloss measurement of papers – the effect of luminance factor. *British Journal of Applied Physics, 2.*

He, X. D., Torrance, K. E., Sillion, F. X., & Greenberg, D. P. (1991). A Comprehensive Physical Model for Light Reflection. *Computer Graphics, 25*(4).

Henley, S. A., & Fairchild, M. D. (2000). *Quantifying mixed adaptation in cross-media color reproduction.* Paper presented at the IS&T/SID Color Imaging Conference.

Herbert, F. H., Kirkenaer, J. S., & Ladson, J. A. (2003). *Absolute and relative colorimetric evaluation for precise color on screen.* Paper presented at the Color Imaging VIII: Processing, Hardcopy, and Applications.

Hunt, R. W. G. (1995). *Why is black and white so important in colour?* Paper presented at the IS&T/SID Color Imaging Conference.

Hunt, R. W. G. (1998). *Measuring Colour* (3rd ed.): Foutain Press.

Hunt, R. W. G. (2004). *The Reproduction of colour* (6th ed.): John Wiley & Sons.

Hunter, R., & Judd, D. (1939). Development of a method of classifying paints according to gloss

. *ASTM Bulletin*(March 1939).

Hunter, R. S. (1937). Methods of Determining Gloss. *Proceedings ASTM, 36.*

Hunter, R. S. (1975). Chapter 6, "Scales for Gloss and Other Geometric Attributes," and Chapter 13," Instruments for the Geometric Attributes of Object Appearance". In *The Measurement of Appearance*: Wiley-Interscience.

Hunter, R. S., & Harold, R. W. (1987). *The measurement of appearance* (2nd ed.): John Wiley & Sons.

IEC. (1999). *61966-2-1: Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space – sRGB.*

IPA. (2005). *IPA Color Proofing RoundUP Special Report*: www.ipa.org.

Ishihara, S. (1980). *Ishihara's Test for Color Blindness*: Kanehara.

ISO. (1999). 12646: Graphic technology—Displays for colour proofing—Characteristics and viewing conditions.

ISO. (2000a). 3664: Viewing conditions – Graphic technology and photography.

ISO. (2000b). 12233: Photography -- Electronic still-picture cameras -- Resolution measurements.

ISO. (2004). 2813: Paints and varnishes -- Determination of specular gloss of non-metallic paint films at 20 degrees, 60 degrees and 85 degrees.

Jennes, J. W., & Shevell, S. K. (1995). Color appearance with sparse chromatic context. *Vision Research, 35*.

Johnson, N. L., & Stephenson, H. F. (1983). *The Influence of Geometric Tolerances on 45°/0° and 0°/45° Colorimetric Measurements.* Paper presented at the CIE 20th Session, Amsterdam.

Johnson, T., and Brammer, J. (2002). *The effects of gloss and viewing flare on colour appearance.* Paper presented at the TAGA Symposium.

Jones, L. A. (1963). Psychological concepts: perceptual and affective aspects of color. In O. C. o. Colorimetry (Ed.), *The science of color*: Optical Society of America.

Kajiya, J. T. (1985). *Anisotropic reflection models.* Paper presented at the SIGGRAPH'85, San Francisco.

Katoh, N. (1995). *Appearance match between soft copy and hard copy under mixed chromatic adaptation.* Paper presented at the IS&T/SID Color Imaging Conference.

Katoh, N., & Nakabayashi, K. (2001). *Applying mixed adaptation to various chromatic adaptation transformation (CAT) models.* Paper presented at the IS&T PICS.

Katoh, N., Nakabayashi, K., Ito, M., & Ohno, S. (1998). Effect of ambient light on color appearance of softcopy images - mixed chromatic adaptation for self-luminous displays. *Journal of Electronic Imaging, 7*.

Kautz, J. (2003). Hardware Lighting and Shading: a Survey. *Computer Graphics forum 1*.

Kirk, R., & Othmer, D. (Eds.). (1978). *Encyclopedia of Chemical Technology* (3rd ed. Vol. 16): John Wiley & Sons.

Klinker, G. J. (1993). *A Physical Approach to Color Image Understanding*: A K Peters.

Koenderink, J. J., van Doorn, A. J., & Stavridi, M. (1996). *Bidirectional Reflection Distribution Functions Expressed in Terms of Surface Scattering Modes.* Paper presented at the European Conference on Computer Vision.

Koschmieder, H. (1924). Theorie der horizontalen sichtweite. *Beitr. Phys. Atmos., 12*(33-53).

Kubelka, P., & Munk, F. (1931). Ein Beitrag zur Optik der Farbanstriche. *Zeitschrift für technische Physik, 12*.

Kuo, C., Ng, Y., & Wang, J. C. (2002a). *Differential Gloss Visual Threshold Under Normal Viewing Conditions.* Paper presented at the IS&T's NIP18: 2002 International Conference on Digital Printing Technologies.

Kuo, C., Ng, Y., & Wang, J. C. (2002b). *Gloss Patch Selection based on Support Vector Regression.* Paper presented at the IS&T's 2002 PICS Conference

Kwak, Y., MacDonald, L. W., & Luo, M. R. (2001). *Colour appearance comparison between lcd projector and lcd monitor colours.* Paper presented at the AIC'01.

Lafortune, E., Foo, S.-C., Torrance, K. E., & Greenberg, D. P. (1997). *Non-linear Approximation of Reflectance Functions.* Paper presented at the ACM SIGGRAPH 1997.

Laihanen, P. (1994). Exact soft proofing. *Journal of Imaging Science and Technology, 38*(5).

Laine, J. (2002). *Experiments on adaptive soft-copy color reproductio.* Paper presented at the IS&T/SID Color Imaging Conference.

Lambert, J. H. (1760). *Photometria*.

Leekley, R. M., Tyler, R. F., & Hultman, J. D. (1978). *The Effect Of Surface Reflection On Color.* Paper presented at the TAGA.

Lo, M., Luo, M. R., & Rhodes, P. A. (1996). Evaluating colour models' performance between monitor and print images. *Color Reseach and Application, 21*(4).

Logvinenko, A. D., & Maloney, L. T. (2006). The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Perception & Psychophysics, 68*.

Lozano, R. D. (2006). A New Approach to Appearance Characterization. *Color Reseach and Application, 31*(3).

MacDonald, L. W., Morovic, J., & Xiao, K. (2002). A Topographic Gamut Mapping Algorithm. In *Colour Imaging Science: Exploiting Digital Media*: John Wiley & Sons.

MacDougall, D. B. (2002). *Colour in food: Improving quality* (1 ed.): Woodhead Publishing Limited and CRC Press LLC

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: a user's guide*: Lawrence Erlbaum Associates.

Marin, J., & Bassinger, G. (2003). *Gatf 2003 Soft Proofing Study*: Graphic Arts Technical Foundation.

Mason, R. P. (1992). The influence of surface properties on image interpretation. In *Comparison of color images presented in different media* (Vol. 2, pp. 825-845): TAGA and ISCC.

Masters, T. (1993). *Practical Neural Network Recipes in C++*.

Maxwell, J. C. (1865). A Dynamical Theory of the Electromagnetic Field. *Philosophical Transactions of the Royal Society of London*(155).

Moroney, N., Fairchild, M. D., Hunt, R. W. G., Li, C., Luo, M. R., & Newman, T. (2002). *The CIECAM02 Color Appearance Model.* Paper presented at the IS&T/SID Color Imaging Conference.

Morovic, J., & Luo, M. R. (1997). *Gamut Mapping Algorithms Based on Psychophysical Experiment.* Paper presented at the IS&T/SID Color Imaging Conference.

Morovic, J., & Wang, Y. (2003). *Influence of Test Image Choice on Experimental Results.* Paper presented at the IS&T/SID Color Imaging Conference.

Morovic, P., Xu, H., & Luo, M. R. (1999). *Inter-comparison of colour measuring instruments.* Paper presented at the International Colour Management Forum.

Mudget, P. S., & Richards, L. W. (1971). Multiple scattering calculations for technology. *Applied Optics, 10*(7).

Nassau, K. (1983). *The Physics and Chemistry of Color: The Fifteen Causes of Color* (2nd ed.): John Wiley & Sons.

Ng, Y. (2005). *INCITS W1.1 2005-011: Image Quality of Printers*.

Ng, Y., Zeise, E., Mashtare, D., Kessler, J., Wang, J., Kuo, C., et al. (2003). *Standardization of Perceptual based Gloss and Gloss Uniformity for Printing Systems (INCITS W1.1).* Paper presented at the IS&T's PICS.

Nieto-Vesperinas, M., Sanchez-Gil, J. A., Sant, A. J., & Dainty, J. C. (1990). Light transmission from a randomly rough dielectric diffuser: theoretical and experimental results. *Optics Letters, 15*(22).

Nokia. (2005). http://www.construnet.hu/nokia/Monitors/TEST/monitor_test.html.

Obein, G., Knoblauch, K., & Viénot, F. o. (2004). Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision, 4*.

Obein, G., Leroux, T., & Viénot, F. o. (2000). Bi-directional reflectance distribution factor and gloss scales. *Proceedings of SPIE, 4299*.

Ohta, N., & Rosen, M. (Eds.). (2006). *Color desktop printer technology*: CRC Taylor & Francis Group.

Oicherman, B., Luo, R. M., & Robertson, A. R. (2006). *Observer metamerism and colorimetric additivity failures in soft-proofing.* Paper presented at the IS&T/SID Color Imaging Conference.

Orchard, S. E. (1968). A new look at ligment optics. *J. Oil. Col Chem. Assoc.*

Oren, M., & Nayar, S. K. (1994). *Generalization of Lambert's Reflectance Model.* Paper presented at the ACM SIGGRAPH 1994.

Oskoui, P., & Pirrota, E. (1998). *Influence of background characteristics on adapted white points of CRTs.* Paper presented at the IS&T/SID Color Imaging Conference.

Owens, J. C., & VanSant, B. W. (1999). *Image quality verification in the development of hardware and media for the Kodak digital lab system.* Paper presented at the IS&T's PICS.

Pellacini, F., Ferwerda, J. A., & Greenberg, D. P. (2000). *Toward a Psychophysically-Based Light Reflection Model for Image Synthesis.* Paper presented at the SIGGRAPH.

Ramamoorthi, R., & Hanrahan, P. (2004). A Signal-Processing Framework for Reflection. *ACM Transactions on Graphics, 23*(4).

Rich, D. (2006).

Rich, D. C. (1988). The Effect of Measuring Geometry on Computer Color Matching. *Color Reseach and Application, 13*(2).

Rich, D. C., Alston, D. L., & Allen, L. H. (1992). Psychophysical verification of the accuracy of color and color-difference simulations of surfaces samples on a crt display. *Color Reseach and Application, 17*(1).

Schramm, M. T., & Meyer, G. W. (1998). *Computer Graphic Simulation of Light Reflection from Paper* Paper presented at the IS&T's 1998 PICS Conference.

Schroder, P., & Sweldens, W. (1995). *Spherical Wavelets: Efficiently Representing the Sphere.* Paper presented at the ACM SIGGRAPH 1995.

Sharma, G. (2002). *Digital Color Imaging Handbook*: CRC Press Inc.

Singer, B., & D'Zmura, M. (1994). Color contrast induction. *Vision Research, 34*.

Stogryn, A. (1967). Electromagnetic scattering from rough, finitely conducting surfaces. *Radio Science, 2*(4).

Stokes, M., Fairchild, M. D., & Berns, R. S. (1992). *Colorimetrically quantified visual tolerances for pictorial images.* Paper presented at the Comparison of color images presented in different media.

Sueeprasan, S., Luo, M. R., & Rhodes, P. A. (2001). Investigation of Colour Appearance Models for Illumination Changes Across Media. *Color Reseach and Application, 26*(6).

Sueeprasan, S., & Luo, R. (2001). *Incomplete chromatic adaptation under mixed illuminations.* Paper presented at the IS&T/SID Color Imaging Conference.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review, 34.*

Tiplitz-Blackwell, K., & Buchsbaum, G. (1988). Quantitative studies of color constancy. *Journal of the Optical Society of America, A5.*

Togerson, W. S. (1954). A law of categorical judgement. In L. S. Clark (Ed.), *Consumer behavior* (pp. 92-94): New-York University Press.

Torrance, K. E., & Sparrow, E. (1967). Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am., 57*(9).

Usui, N., & Imamura, A. (1995). *Printing simulator.* Paper presented at the TAGA Symposium.

Völz, H. G. (2001). *Industrial Color Testing* (2nd ed.): Wiley VCH.

von Helmholtz, H. L. F. (1866). *Treatise on Physiological Optics*: Leopold Voss.

Wachtler, T., Albright, T. D., & Sejnowski, T. J. (2001). Nonlocal interactions in color perception: nonlinear processing of chromatic signals from remote inducers. *Vision Research, 41.*

Walker, P. A. (1999). *A Visualization System for Bidirectional Reflectance Distribution Functions.* University of Oregon.

Ward, G. J. (1992). *Measuring and Modeling Anisotropic Reflection.* Paper presented at the ACM SIGGRAPH 1992.

Watt, A., & Watt, M. (1992). *Advanced Animation and Rendering techniques: Theory and Practice*: ACM Press and Addison-Wesley.

Weisskopf, V. F. (1968). How Light Interacts with Matter. *Scientific American, 219*(3).

Wesner, M. F., & Shevell, S. K. (1992). Color perception within a chromatic context: Changes in red/green equilibria caused by noncontiguous light. *Vision Research, 32.*

Westin, S. H., Arvo, J. R., & Torrance, K. E. (1992). *Predicting Reflectance Functions from Complex Surfaces.* Paper presented at the ACM SIGGRAPH 1992.

Westlund, H. B., & Meyer, G. W. (2001). *Applying Appearance Standards to Light Reflection Models.* Paper presented at the ACM SIGGRAPH 2001.

Wolff, L. B., & Kurlander, D. J. (1990). Ray Tracing with Polarization Parameters. *IEEE Computer Graphics and Applications, 10*(6).

Wynn, C. (2004). *An Introduction to BRDF-Based Lighting*: NVIDIA Corporation.

Yule, J. A. C. (1967). *Principles of Color Reproduction: Applied to Photomechanical Reproduction, Color Photography, and the Ink, Paper, and Other Related Industries*: John Wiley & Sons.

Zaidi, Q., Yoshimi, B., Flanigan, N., & Canova, A. (1992). Lateral interactions within color mechanisms in simultaneous induced contrast. *Vision Research, 32*.