UNIVERSITY OF
NEWCASTLE

Colour Constancy in Simple and Complex Scenes

David Bramwell[1]
Department of Physiological Sciences
Medical School
Framlington Place
Newcastle upon Tyne
NE2 4HH

Submitted to the University of Newcastle upon Tyne
for the degree of Doctor of Philosophy

January 1997

1 Supported by the Newcastle University Research Committee

Abstract

Colour constancy is defined as *the ability to perceive the surface colours of objects within scenes as approximately constant through changes in scene illumination.* Colour constancy in real life functions so seamlessly that most people do not realise that the colour of the light emanating from an object can change markedly throughout the day. Constancy measurements made in simple scenes constructed from flat coloured patches do not produce constancy of this high degree. The question that must be asked is: *what are the features of everyday scenes that improve constancy?*

A novel technique is presented for testing colour constancy. Results are presented showing measurements of constancy in simple and complex scenes. More specifically, matching experiments are performed for patches against uniform and multi-patch backgrounds, the latter of which provide colour contrast. Objects created by the addition of shape and 3-D shading information are also matched against backgrounds consisting of matte reflecting patches. In the final set of experiments observers match detailed depictions of objects – rich in chromatic contrast, shading, mutual illumination and other real life features – within depictions of real life scenes.

The results show similar performance across the conditions that contain chromatic contrast, although some uncertainty still remains as to whether the results are indicative of human colour constancy performance or to sensory match capabilities. An interesting division exists between patch matches performed against uniform and multi-patch backgrounds that is manifested as a shift in CIE xy space.

A simple model of early chromatic processes is proposed and examined in the context of the results.

Acknowledgements

There are many people who have provided support and assistance throughout the preparation of this thesis. Unfortunately, space limits the list of influences and friendships that can be formally acknowledged here. This journey has been an interesting and satisfying experience. There are those whose contribution has been so great that I feel special mention is required.

To my parents I owe a large debt. Their gentle encouragement and useful advice has been a huge factor, not only throughout my thesis, but throughout my life. They have my greatest respect and thanks.

My supervisor, Anya Hurlbert, deserves a great deal of credit for the work presented in this thesis. I was not the easiest student to manage and appreciate the freedom I was afforded in examining avenues of research that I found interesting. Her management style encouraged the development of my own ideas and capabilities and I am pleased to say that throughout my thesis our relationship has been more akin to friends and equals than supervisor and student.

I also feel my examiners should be mentioned, Paul Whittle and Gary Green. Their careful examination of my thesis and thoughtful questioning changed my viva from an ordeal to a useful exchange of ideas.

Barry Ledeatte also deserves a mention. We both started our PhDs in the vision lab at the same time and not only has the exchange of ideas been enlightening, but also, the mutual support provided by someone going through exactly the same experiences proved invaluable for retaining focus (and sanity!!).

Finally, everyone in the Neural Systems Research Group deserves a mention for providing a relaxed and lively atmosphere in which to work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Intelligence is difficult to define. Most people associate intelligence with tasks we find difficult. Playing competitive chess, for example, is considered to require some sort of high intelligence, whereas visual tasks, which every human finds easy, whatever his or her IQ, are not thought to require intelligence.

The artificial intelligence community's goal of creating a machine that can play chess exemplifies the high level of intelligence we accord to that task, but that goal may be far easier achieved than building a vision machine with anything like the effortless visual capabilities of a human being.

It is when we attempt to analyse a task into basic operations, for example instructions for a computer, that the staggering complexity of some of these processes becomes apparent. A task that can be performed by a child in seconds could keep the world's fastest computer busy for a very long time or prove to be near insoluble within the limitations of the computer's command set and resources.

The task performed by the visual system is easy to define in a simple input–output structure.

- Input: The arrays of light photons arriving on the two arrays of photoreceptors on the retinae of the eyes. A more complex system might add some knowledge of approximate location, expectation of the objects it may find at that location, memories of previous visual experiences, some *a priori* knowledge of the physics of the world (factors that spurred the evolution of our transducers)

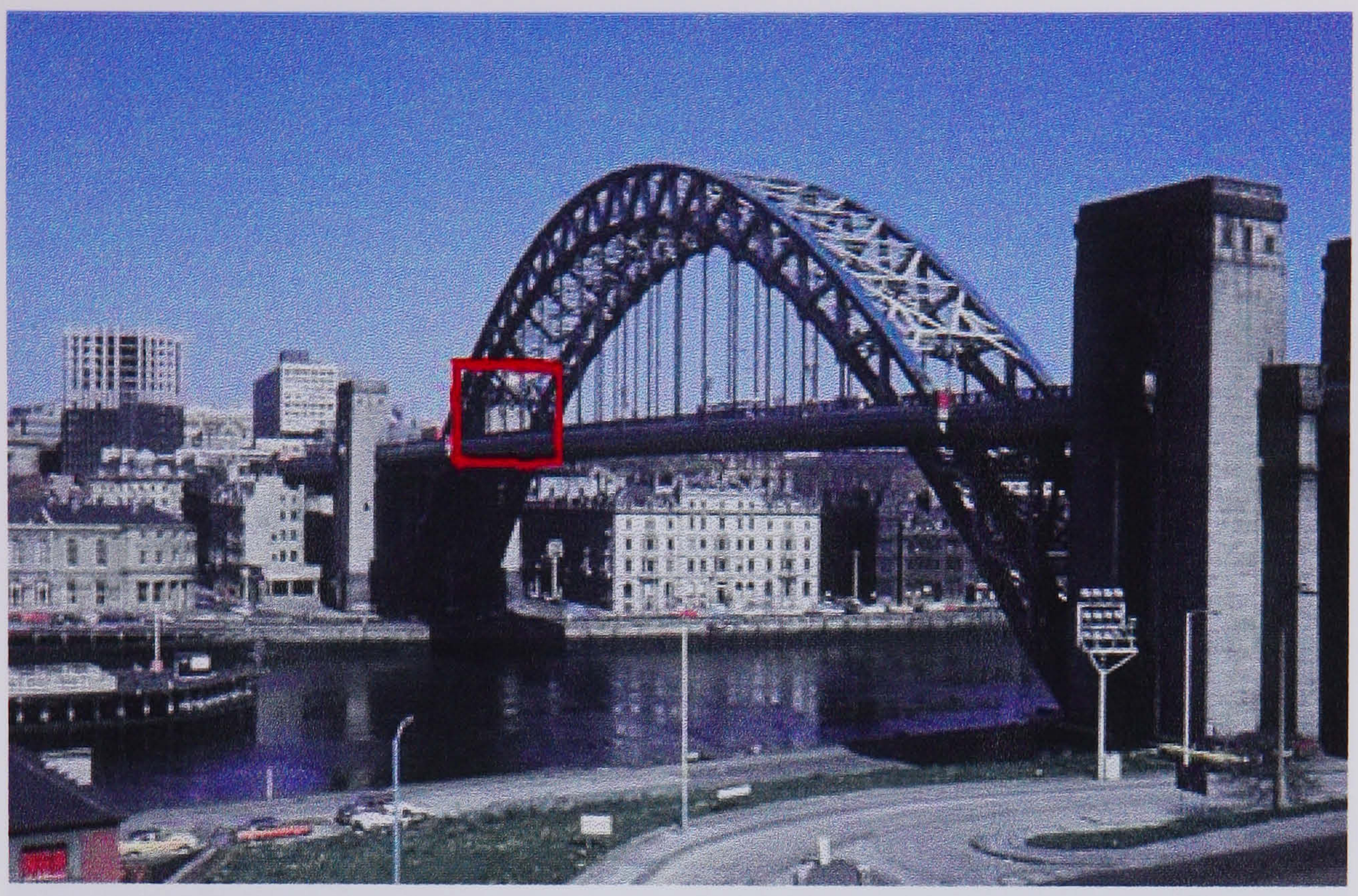

Figure 1.1: An example of an outdoor scene (©Audio Visual Centre, University of Newcastle Upon Tyne)

and input from other senses along with the desire to satisfy some instinct or drive (for example, finding food).

- Output: Recognition of objects, people, surroundings, prevailing weather conditions (illumination) and imminent danger; localisation of self in spatial relation to other objects; estimation of self velocity and of object velocities.

To us this task seems easy. It is not. The problem can be highlighted by comparing Figure 1.1, which depicts an outdoor scene, with Figure 1.2. We immediately see and recognise the forms and objects within the picture of Figure 1.1. Figure 1.2 shows a small section of the scene (the part in the red square) and also shows excitations for red, green and blue channels (here, arbitrarily chosen wavebands of the visible spectrum; the excitations in the L, M and S cone channels would be similar). The images represent the inputs to the three receptor classes where each square signifies the input to a single receptor. The receptor inputs are related to each other, but exactly how and which picture elements should be classified together is the whole crux of the problem of visual perception.

Visual perception can be subdivided into many aspects that can, to an extent, be researched independently. For example, techniques for detecting object outlines

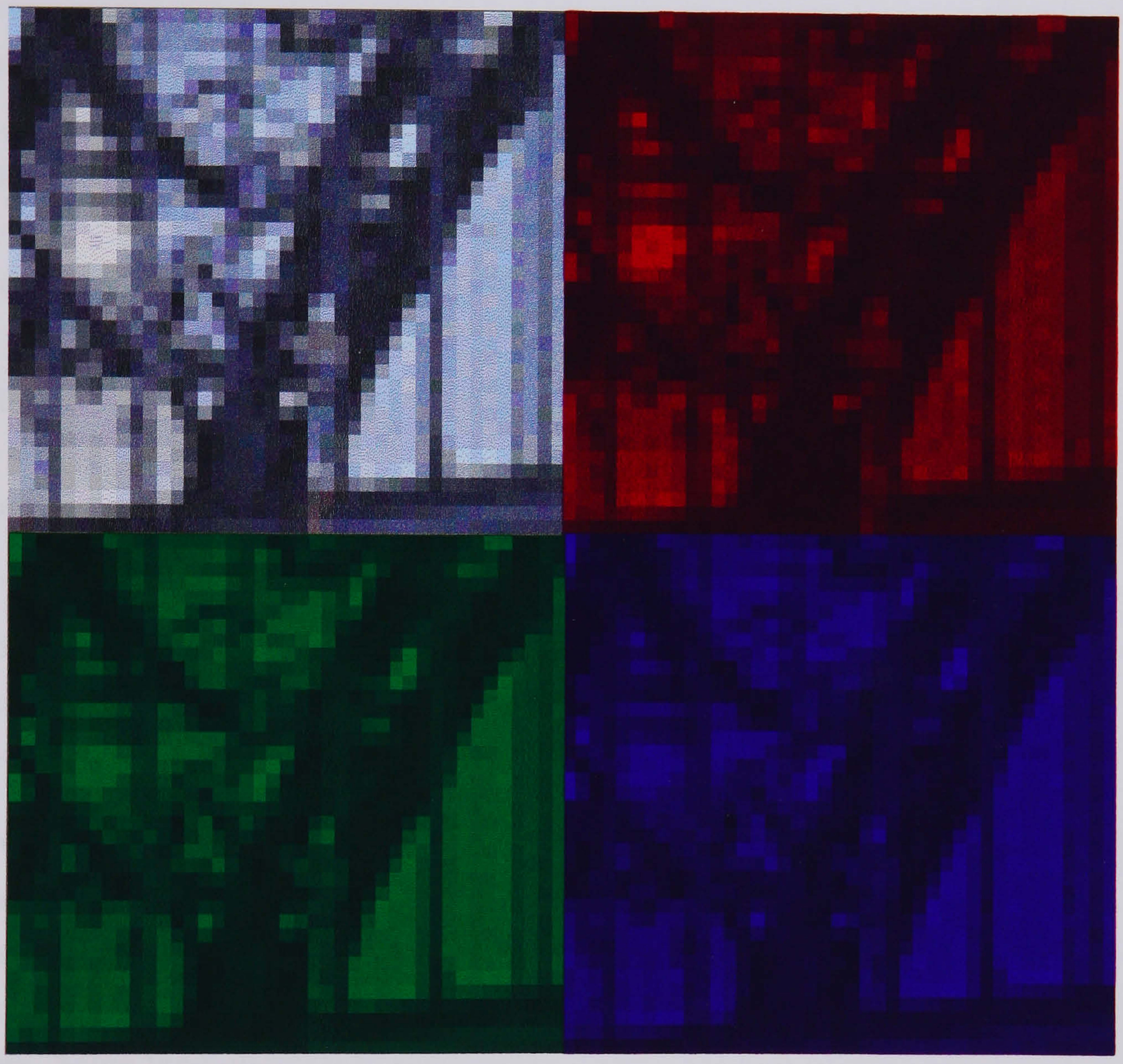

Figure 1.2: Small section of outdoor scene showing: full colour, red channel, green channel and the blue channel

and image segmentation may exist independently of routines for computing motion fields (for example). Whether we will ever get a working visual system from 'bolting together' many such systems is an open question, but the problem is simply too large and complex to handle in any other way at present.

This thesis is primarily concerned with a task performed by the visual system known as colour constancy. Colour constancy is the tendency for object surface colours within a scene to appear the same through changes in scene illumination. Again, our visual system performs the task underlying colour constancy so well that we just do not realise the task is required. The best way to show the need for colour constancy is to show the output of a system without it. Figure 1.3 shows the output from a colour camera for the same scene under different illumination conditions (these scenes can be considered to be the input for the visual system, i.e. the signals impinging on the retina). If we were viewing the scene naturally it would look much the same as the first image for all of the images. We can detect the change in illumination but generally do not experience the significant shifts in colour undergone by the objects within the scene.

The work in this thesis aims to categorise human colour constancy performance in simple (arrays of flat matte patches) and complex (full colour images) scenes and to use this information to assess the level of constancy that can be explained by known processes within the lower levels of the visual system.

Chapter 2 covers the basic theory underlying studies of low level human vision along with basic experimental procedural factors, and some experimental results from the literature. Work of other studies on low level vision is reviewed and presented where the findings are useful for comparison purposes to the results of this thesis. Some fundamentals of colorimetric theory are also covered along with other representation spaces for stimuli. Some psychophysical techniques relevant to the procedures used in this thesis are discussed. All experimentation is performed with computer presentations, and the capabilities and limitations of such systems are considered along with their calibration.

Chapter 3 reviews the results of several studies on adaptation in the human visual system and on colour and lightness constancy. Possible methodologies for solving the 'colour constancy problem' are discussed. Transformations of the colour

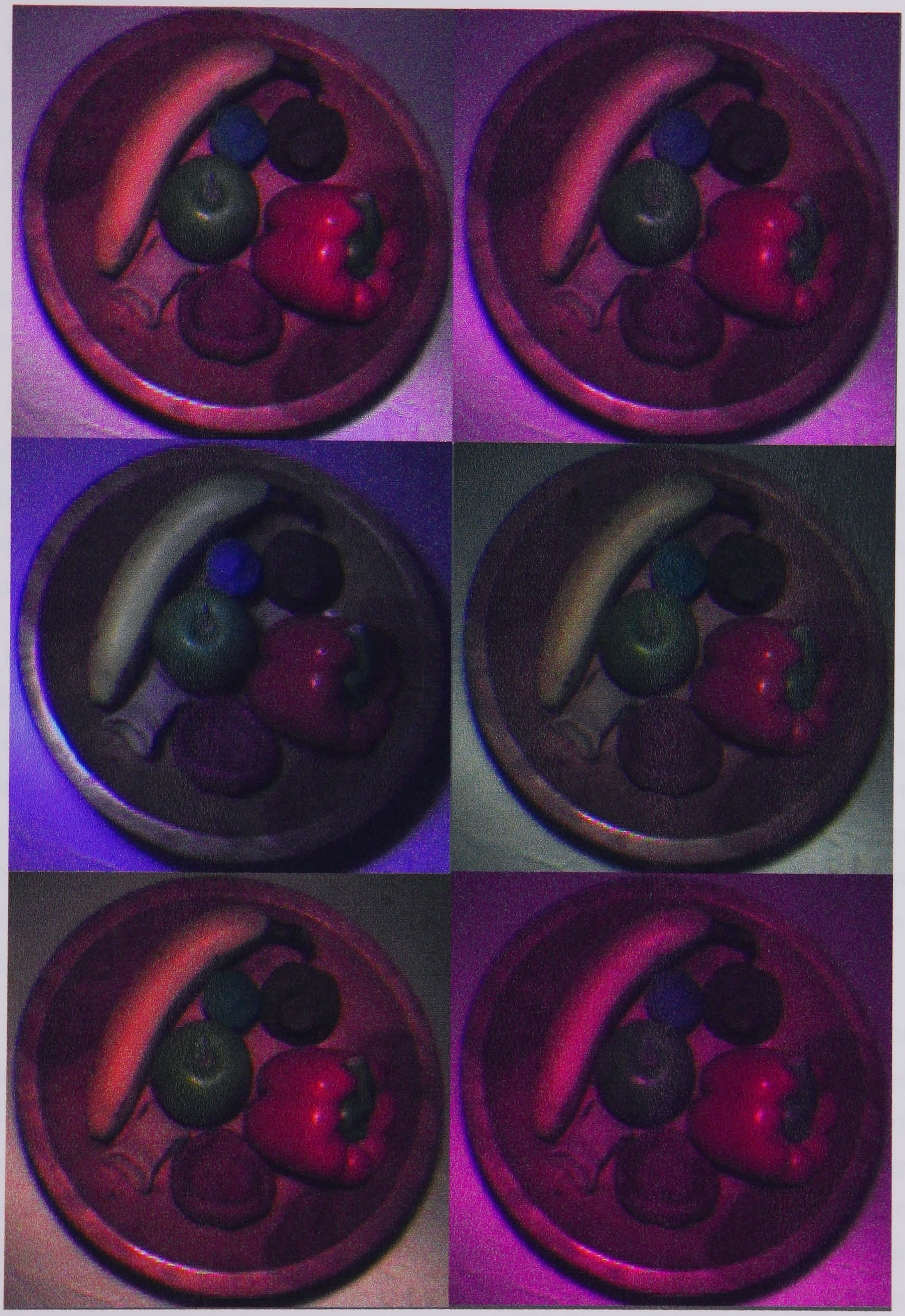

Figure 1.3: A scene under different illuminations viewed without colour constancy

signals that are known to occur are also assessed and possible implications for colour constancy discussed.

Chapter 4 reviews some interesting results on the possible contribution of high level factors to the perception of surface properties. Special emphasis is placed on studies in which perceived structural features of a scene affect the surface lightness or colour judgements.

Then follows four results sections. Each is self contained and describes full stimulus specifications, motivation, the techniques used and the results obtained from the procedure. Particular points of interest are discussed with the results.

Chapter 5 describes experiments used to verify general procedural aspects and control experimentation. The topics covered include the effects of: the viewing techniques, stimulus configuration and matching under varying illuminations. The results are presented and assessed and provide some base level performance results against which subsequent results can be measured.

Chapter 6 examines colour constancy for flat matte patches against uniform and multi-coloured backgrounds. The motivation behind this section is two-fold. First a base level of performance is measured with our specific equipment and paradigms. This allows for the comparison of our results with similar past studies and with the new data to follow from our subsequent experiments. The second probes the role of colour contrast in the matching of patch colours. It is known that the visual system adapts to factors in its visual environment and changes its performance to suit. Exactly which factors in the scene produce a given state is unknown.

Chapter 7 increases the scene complexity slightly by adding 3-D shading to form recognisable objects. This chapter probes the effects of increasing information and also tests other effects such as object recognition and colour memory. In everyday life scenes may be differentially illuminated; for example, a desk may be lit by a tungsten bulb, while the rest of the room is bathed in daylight. This differential illumination does not cause a problem for our perception of constant object colours, but would for many colour constancy algorithms. Is the extra information of shading and object recognition enough for us to compensate for differential illumination?

Chapter 8 increases the information further by testing constancy performance in depictions of real scenes. As many 'real life' features are provided as possible (recognisable objects, 3D surface shading, colour contrast, mutual illumination. shadows. highlights etc.). In some respects, the colour constancy performance shown in typical laboratory experiments is less than would be expected from the performance levels we seem to experience in everyday life. We ask here whether these extra features will improve constancy in experiments.

Chapter 9 gives a discussion of all of the results obtained in context with each other and other studies and examines some properties of the visual system that they suggest. Chapter 10 concludes the thesis with a summary of the motivations, experimentation, results and implications of the work presented.

Chapter 2

Fundamentals of Colour Vision and Colorimetry

2.1 Introduction

The extent to which the human visual system possesses colour constancy will be limited by the basic performance characteristics of the eye's imaging system. A basic understanding of some of these processes is paramount in the design and execution of colour constancy experiments. This chapter covers the basic theory required to quantify and assess the performance levels of observers throughout the experimentation.

2.2 The human visual system

2.2.1 Properties of the eye

The human visual system has four types of photoreceptor, the rods and three cone types.

The rods have a broadband response – their spectral sensitivity function forms a bell-shaped curve non-zero between 380nm – 620nm, with a peak response wavelength, $\lambda_{max} \approx 502$ nm. They are used in low light intensity situations (< 10 Td (Wyszecki and Stiles 1982 page 559)). In retinal topography terms, they are believed to be absent in the foveola (central 1.7–2.0^o of visual field) and extremely

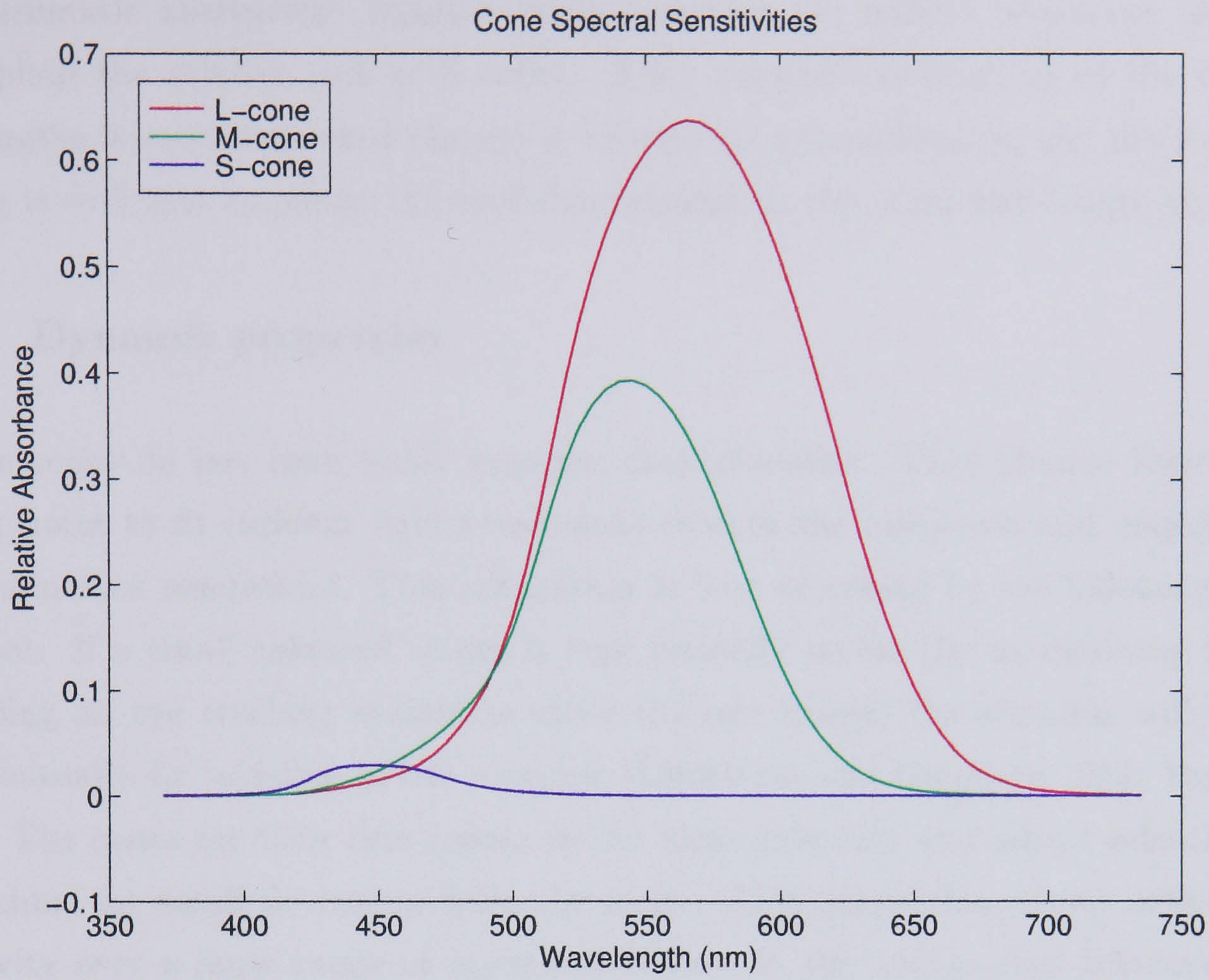

Figure 2.1: Cone Spectral Sensitivities (Smith and Pokorny)

rare in the fovea (central 5.2^o of visual field). They are assumed to provide a very small response at the light levels used in this thesis' experiments and hence not to contribute to the visual tasks tested. Specifically, if we take worst case conditions of a pupil diameter of 2mm (it is likely to be larger than this given the luminances involved) then at $5cd/m^2$ the retina will be irradiated with approximately 15.7 Trolands. Given the above light intensity and topographic constraints the rods will be omitted from any further consideration in this thesis.

The three cone types have broadband responses with peaks nominally designated at 555 nm, 545 nm and 444 nm, for the long (L), medium (M) and short (S) wavelength cones respectively. Figure 2.1 shows the cone responses as derived by Smith and Pokorny (Smith and Pokorny 1975).

Topographically the L and M cones appear to be arranged in a regular hexagonal structure throughout the fovea with many fewer S cones interspersed. A ring of S cones exists just outside the foveal area. The ratio of L+M cones to S cones is approximately 100:1. Marimont and Wandell (Marimont and Wandell 1994) suggest that chromatic aberration[1] considerations based on the optical properties of the eye explain the relative lack of S cones. They suggest the blurring of the short wavelengths is such that when the eye is focused on a broadband target the S cone spacing is sufficient to report the available contrast in the short wavelength signal.

2.2.2 Dynamic properties

The cones do not have static response characteristics. They change their operating point to fit incident light constraints (within the minimum and maximum light intensities resolvable). This adaptation is best described by the following experiment: If a small coloured target is kept retinally stable (by maintaining fixation using an eye tracking system to move the test image) the stimulus will fade and eventually be invisible to the observer (Ditchburn and Ginsborg 1952; Yarbus 1967). The cones set their zero points to the mean intensity and adjust sensitivity to maximal for small deviations from the mean. This adaptation allows increased sensitivity over a large range of input intensities. It also means that information about absolute light levels is lost and only relative differences can be detected and signalled to higher processes. Exactly how the adaptation level is controlled is not fully known but it is known that local and global spatial interactions have effects as well as temporal modulations.

2.2.3 Trichromacy

It follows that any stimulus entering the eye will result in a triplet of excitations from the cones. These three values can be said to specify the actual colour we see. This process also represents a massive reduction in information, as the cones convert the infinite spectrum of incident light into three excitations. This process can be represented mathematically in the following manner:

$$L = \int_{-\infty}^{\infty} l(\lambda) f(\lambda) d\lambda \tag{2.1}$$

$$M = \int_{-\infty}^{\infty} m(\lambda) f(\lambda) d\lambda \tag{2.2}$$

[1]The distortion in a broadband signal that occurs when the signal passes through media of different refractive indices resulting in some wavelengths being refracted more than others.

$$S = \int_{-\infty}^{\infty} s(\lambda) f(\lambda) d\lambda \qquad (2.3)$$

where $l(\lambda), m(\lambda)$ and $s(\lambda)$ are the cone spectral sensitivity functions for the long, medium and short wavelength cones respectively (as shown in figure 2.1) and $f(\lambda)$ is the spectrum of the incident light.

This reduction of information brings about a phenomenon called *metamerism.* Two input spectra are termed metameric if they produce the same responses (i.e. $L_1 = L_2$, $M_1 = M_2$ and $S_1 = S_2$). It follows that two metamers are indistinguishable by a human observer. This principle underlies the operation of the television set. The light spectra emitted from the phosphor triads is vastly different to that emanating from the original object, but since the light emitted is a metamer to the original we find the two indistinguishable.

All light spectra entering the eye will be converted to a triplet of values; hence any stimulus we apply can be described by a point in a three–dimensional space. LMS space is a physiologically based system and is not always convenient for the representation of stimuli for psychophysical experimentation. The reason for this is that the cone excitations do not clearly relate to sensations as reported by observers.

A form of $Brightness + Colour$ specification would be more convenient as perceptually we experience these properties as distinct. A linear transformation can be applied that re–orientates the LMS axes in such a way as to separate the sensations of brightness from 'colourfulness'. Such a space was defined by the CIE in 1931 (CIE 1932). The system is based on measurements of the visual system's V_λ function, which describes the luminous efficiency of the eye. This function describes the magnitude of a monochromatic stimulus required to match another monochromatic stimulus in terms of brightness. Formally, two stimuli $P_1(\lambda)$ and $P_2(\lambda)$ are perceived as equally bright when

$$\int_\lambda P_1(\lambda) V_\lambda(\lambda) d\lambda = \int_\lambda P_2(\lambda) V_\lambda(\lambda) d\lambda \qquad (2.4)$$

The best way to understand the V_λ function is to consider two monochromatic lights. If we define monochromatic sources P_1 at 600nm and P_2 at 640nm then equation 2.4 reduces to:

$$\begin{aligned} P_1 V_\lambda(600) &= P_2 V_\lambda(640) \\ P_1 &= \frac{V_\lambda(600)}{V_\lambda(640)} P_2 \end{aligned} \quad (2.5)$$

The V_λ function may be derived using hetero–chromatic flicker photometry. This technique relies on the fact that the human flicker fusion frequencies (when a rapidly flickering target is seen as solid) for luminance and chromanance differ. Two coloured stimuli are alternately displayed at above the chromatic, but below the luminance fusion frequencies. The observer adjusts the luminance of one of the stimuli until he obtains minimal or zero flicker. At this point the luminances of the two coloured patches will be perceptually equal.

If we define $\overline{y}(\lambda)$ as V_λ and define two further linearly independent (in that neither can be constructed from a combination of the other two) functions, $\overline{x}(\lambda)$ and $\overline{z}(\lambda)$, then we can define a trichromatic space XYZ where the XZ plane contains all the *colourfulness* and the Y axis denotes increases in *luminous intensity.* Figure 2.2 shows the $\overline{x}(\lambda)$, $\overline{y}(\lambda)$ and $\overline{z}(\lambda)$ function as defined in the CIE 1931 standard. One further adjustment is required to make a trichromatic space that is useful in representing colour stimuli. The XYZ space representation does not have a consistent colour distribution in the XZ plane with respect to changes in Y. This problem can be reduced by normalising the chromaticity coordinates. So we have now defined xyY space where:

$$X = \int_{-\infty}^{\infty} \overline{x}(\lambda) f(\lambda) d\lambda \quad (2.6)$$

$$Y = \int_{-\infty}^{\infty} \overline{y}(\lambda) f(\lambda) d\lambda \quad (2.7)$$

$$Z = \int_{-\infty}^{\infty} \overline{z}(\lambda) f(\lambda) d\lambda \quad (2.8)$$

$$x = \frac{X}{X+Y+Z} \quad (2.9)$$

$$y = \frac{Y}{X+Y+Z} \quad (2.10)$$

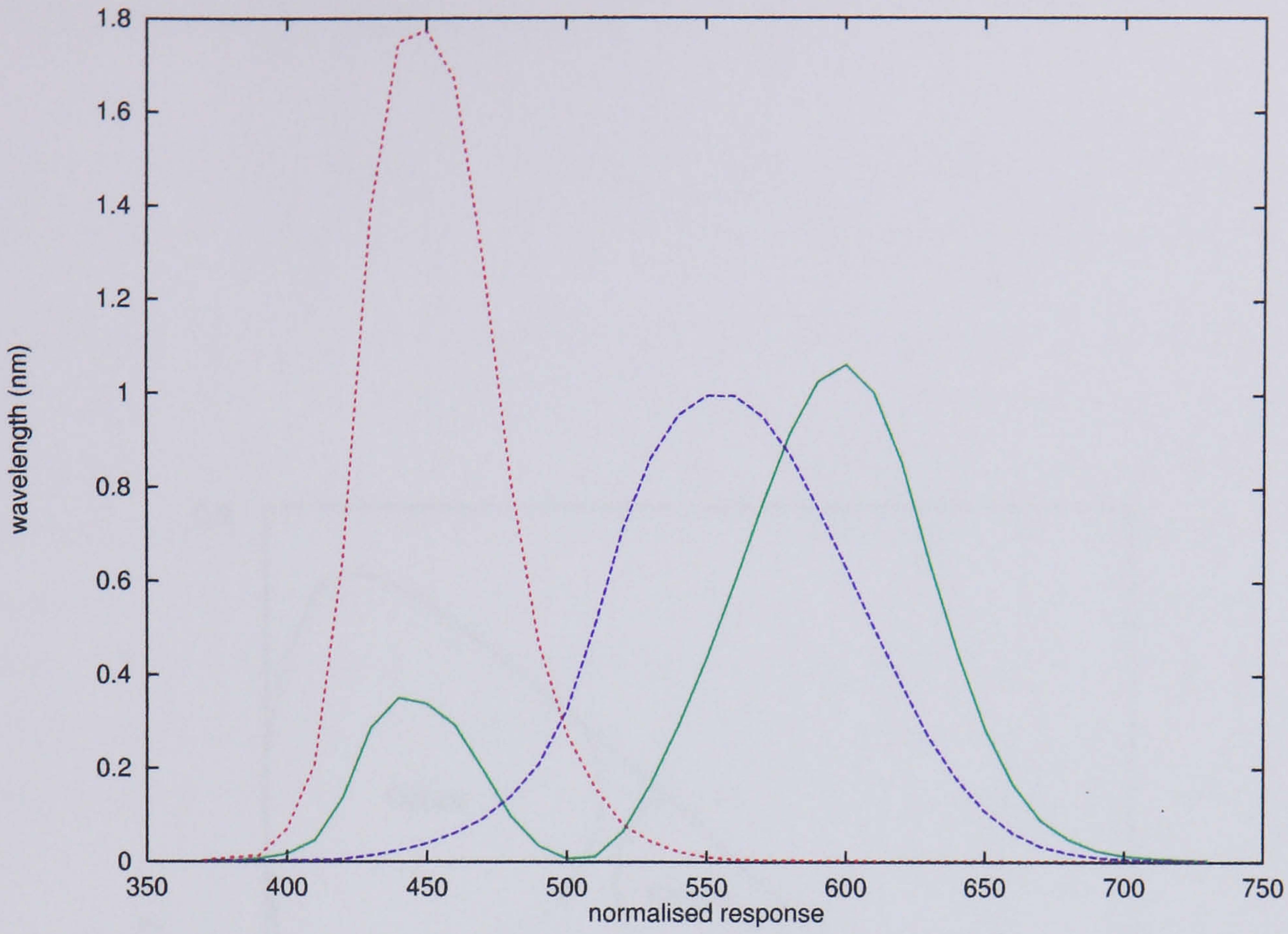

Figure 2.2: $\overline{x}(\lambda)$, $\overline{y}(\lambda)$ and $\overline{z}(\lambda)$ functions

xyY space is a mathematical construct that can be used to specify all possible unique inputs to the human visual system. All systems can be specified by their input to output mapping. Formally $O(x) = T.I(x)$ where $I(x)$ specifies system inputs, $O(x)$ specifies the outputs of the system resulting from $I(x)$ (possibly also related to previous inputs to the system) and T is the system *transfer function.* It is one of the goals of vision research to find the visual system's transfer function. The technique usually used to find a transfer function is to specify a set of inputs and observe the resultant outputs.

A simple experiment of this form would be to show a human observer a monochromatic light against an iso-luminant neutral background. The observer reports whether a target is visible. If this procedure is repeated for a range of monochromatic sources then the range of visible colours can be defined. Figure 2.3 shows the results of such an experiment displayed in xyY space. This experiment shows immediately that the mapping is not linear since the transfer function is band–limited with respect to the possible input wavelengths.

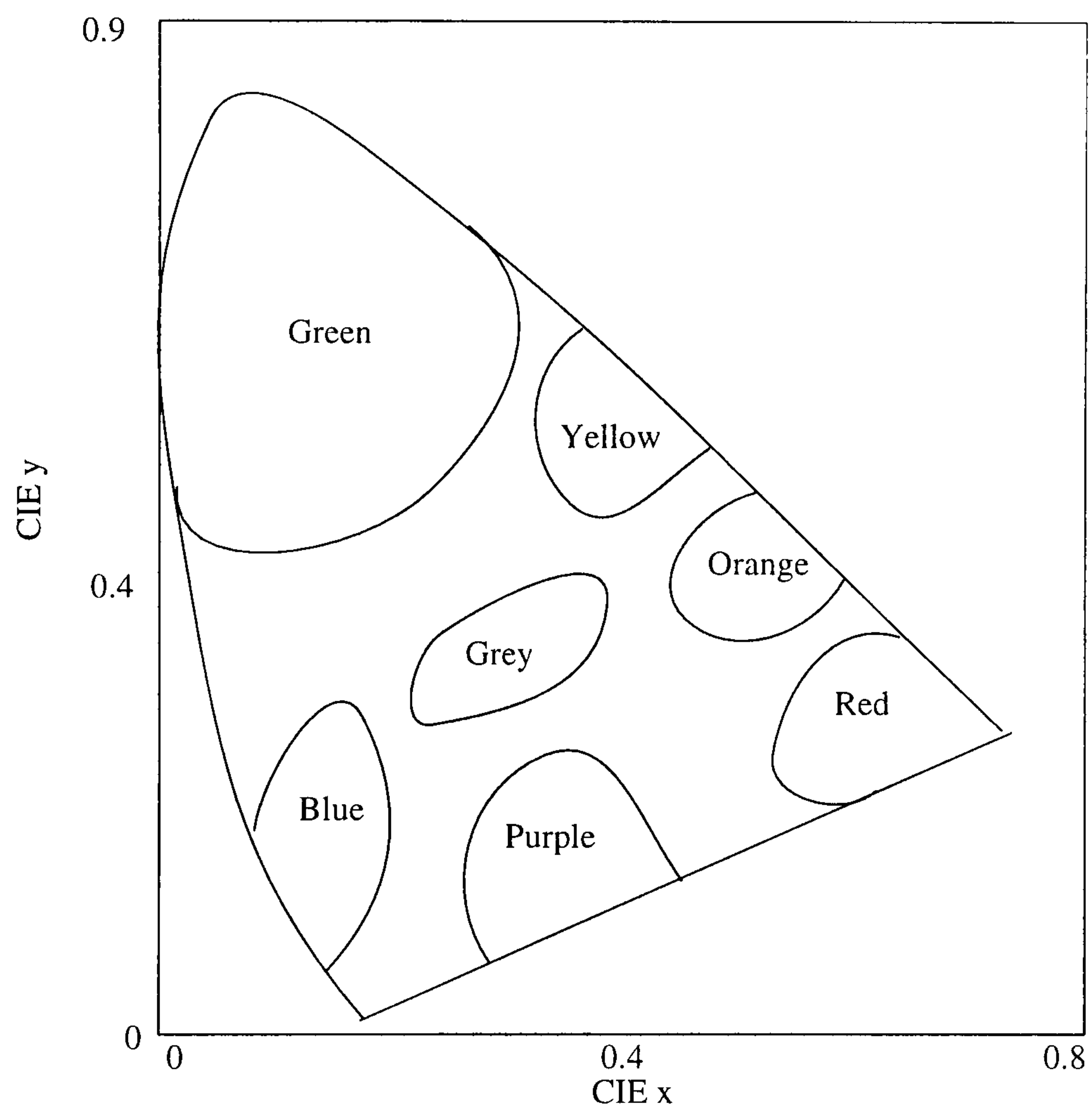

Figure 2.3: Limits of human colour perception in CIE xyY space: the triangular locus represents the boundary between visualisable and non visualisable colours. Approximate colour classification is also shown.

2.3 Colour discrimination

One of the first questions to be asked is: *How large a difference in chromaticity is required for a human observer to perceive two coloured patches as distinct?*

MacAdam (MacAdam 1942) performed the first major study into human observers' colour discrimination abilities. He used an ingenious technique to achieve very fine colour resolution. In brief, the polarised light from two coloured filters (one horizontally and the other vertically polarised) is combined by a Rochon prism. The orientation of this prism determines how much light of a given polarisation it lets through. At zero degrees we obtain the light from only one of the filters, 45^o provides an equal mixture of the two and so on. This system provides a smooth variation of chromaticity along the lines joining the two filter's chromaticity points in xyY space.

The procedure involves the choice of a test patch chromaticity through which several chromaticity lines joining realisable filters pass. The observer is shown the test chromaticity and asked to perform a match by adjusting an adjacent patch (by rotating the Rochon prism via a geared wheel mechanism). When a large number of matches have been performed the distribution of the matches with respect to their frequency of occurrence approximates a normal distribution centred on the test chromaticity. The results from several of the test lines can be pooled to produce a 2–dimensional map of discrimination performance for a given test chromaticity in a given iso-luminant plane. The standard deviations of these distributions can be used as a measure of discrimination performance around a given test chromaticity. Brown and MacAdam (Brown and MacAdam 1949) found the standard deviation of the distributions to be more stable than just noticeable differences when comparing results across regions in colour space. Figure 2.4 shows some of the results obtained by MacAdam in the described experiment. Points to note are the significant variations in size of the ellipses. MacAdam could find no projective transformation that would create a *perceptually uniform* colour space in which a given distance in one area of the space would correspond to a perceptually similar colour response in another area of the space.

Follow up studies performed by Wyszecki and Fielder (Wyszecki and Fielder 1971) showed that the ellipses obtained from MacAdam's procedure can vary be-

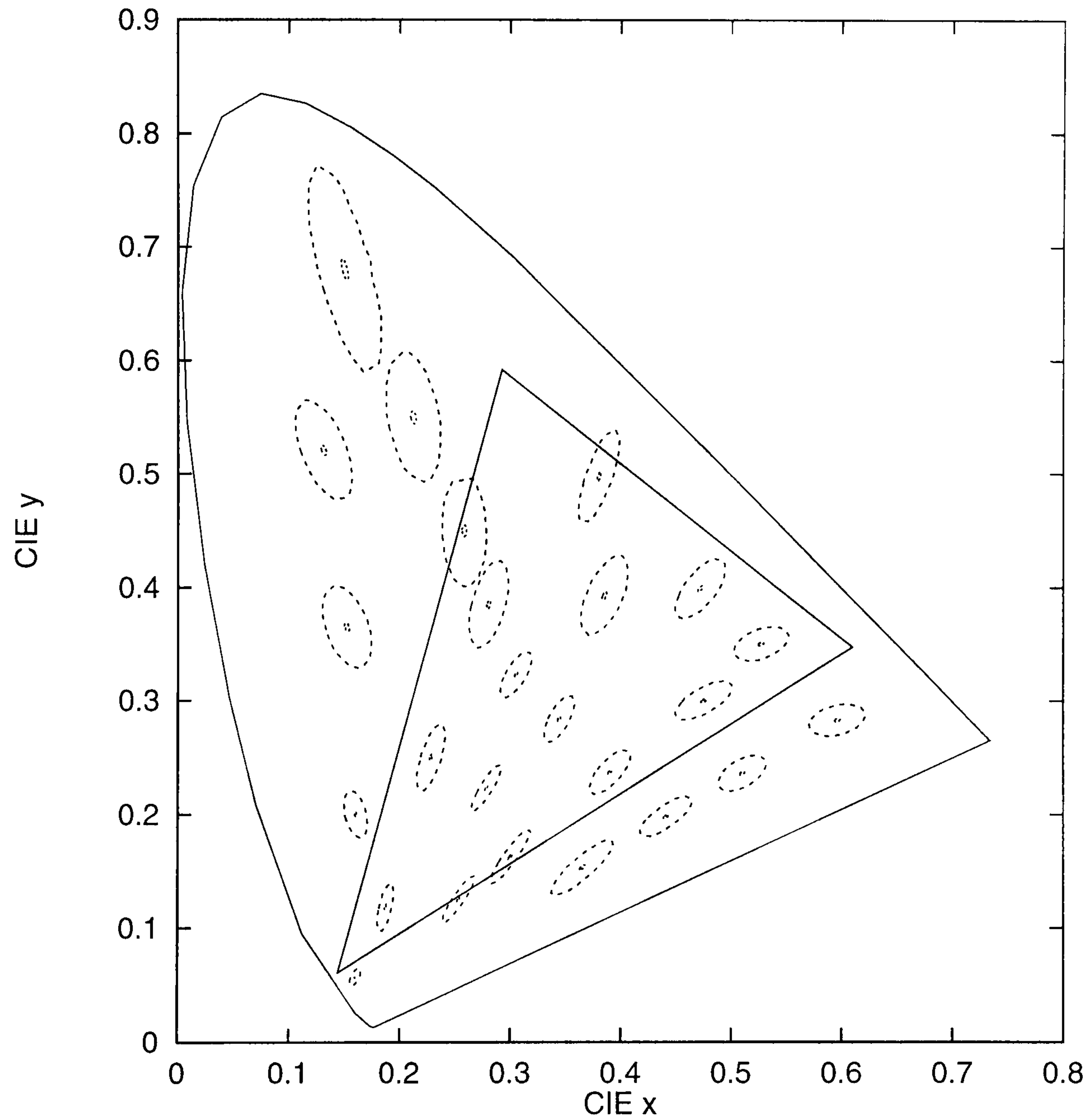

Figure 2.4: MacAdam one standard deviation ellipses for colour discrimination. Triangle shows EIZO T560i phosphor coordinates. Large ellipses are 10 times actual size of MacAdam ellipses.

tween observers and even between subsequent repeats of the experiment by the same observer. The general trend of results shown by MacAdam remains valid, but small variations from these results should be expected when performing any similar psychophysical study.

2.3.1 Cone based spaces

CIE xyY space is convenient for specifying input stimuli to the visual system but may not be the best space to view the psychophysical output. Cone based spaces can be more informative in showing the underlying relationships of the matched data. For example, if we wished to test the von Kries proposal of adaptation based on cone ratios, then plotting the results of a matching experiment in a L_{test} versus L_{match} plot should produce a straight line through the origin if the hypothesis is supported.

It can be difficult to examine and compare stimuli plotted as points in three dimensional spaces. MacLeod and Boynton (MacLeod and Boynton 1979) proposed a space that is convenient for examining the effects of chromatic stimuli on cone classes. In most colour discrimination studies the possibly confounding effects of variation in luminance are avoided. This property is utilised in the new space which normalises for luminance variation by dividing the cone signals by L + M.

If we define

$$\begin{aligned} l &= \frac{L}{L+M} \\ m &= \frac{M}{L+M} \\ s &= \frac{S}{L+M} \end{aligned}$$

then we can plot points in a three dimensional space: lms, which has luminance normalised cone responses on each of its axes. Since $m = 1 - l$ it can been seen that an iso–luminant set of stimuli will be situated in a single plane in lms space. This feature allows the space to be collapsed to produce the MacLeod–Boynton space plot in which s is plotted on the abscissa and l plotted on the ordinate. This space is particularly useful for examining adaptational mechanisms in the cones.

It can also be illuminating to attempt to show signals that would be expected from the opponent mechanisms known to exist in the visual system. Again the luminance component is usually kept constant which allows the space to be collapsed to a two dimensional plot in which $L - M$ is plotted on the ordinate and $S - L + M$ along the abscissa. This space is usually used to examine points relative to a mean, i.e. the cones have been assumed to have performed adaptation to features of the stimulus and all signals are shown relative to this point. This space is particularly useful for examining possible cortical gain phenomena.

Cone based spaces are examined in more detail in chapter 3.

2.4 Computer controlled displays

The equipment used by MacAdam was excellent for the task which it was designed, but very limited. The versatility and convenience of computer controlled displays causes them to be popular in vision research. Unfortunately, because high resolution displays and video generators are not designed to be used in vision experimentation, the validity of their use must be considered and their limitations addressed.

The cathode ray tube uses the electron stream generated from a hot cathode which is then accelerated by a high anode voltage and collimated to excite phosphor dots on a screen. This stream can be diverted by magnetic fields and hence made to scan across the screen in a controlled manner. The excitation of the phosphor coating causes light in the visible region of the electro-magnetic spectrum to be emitted, the intensity of which is proportional to the energy contained in the electron beam.

2.4.1 Gamma correction

Most modern phosphors produce exponential ($emission = current^{\alpha}$) light emission in response to a linear increase of electron current. Figure 2.5 shows an example of such a response. A linear response is preferable for experimentation; the linearisation procedure is usually termed gamma correction. In its simplest form gamma correction involves deriving the exponent of the CRT current to emission characteristic and producing an inverse. For example to correct the CRT example given in

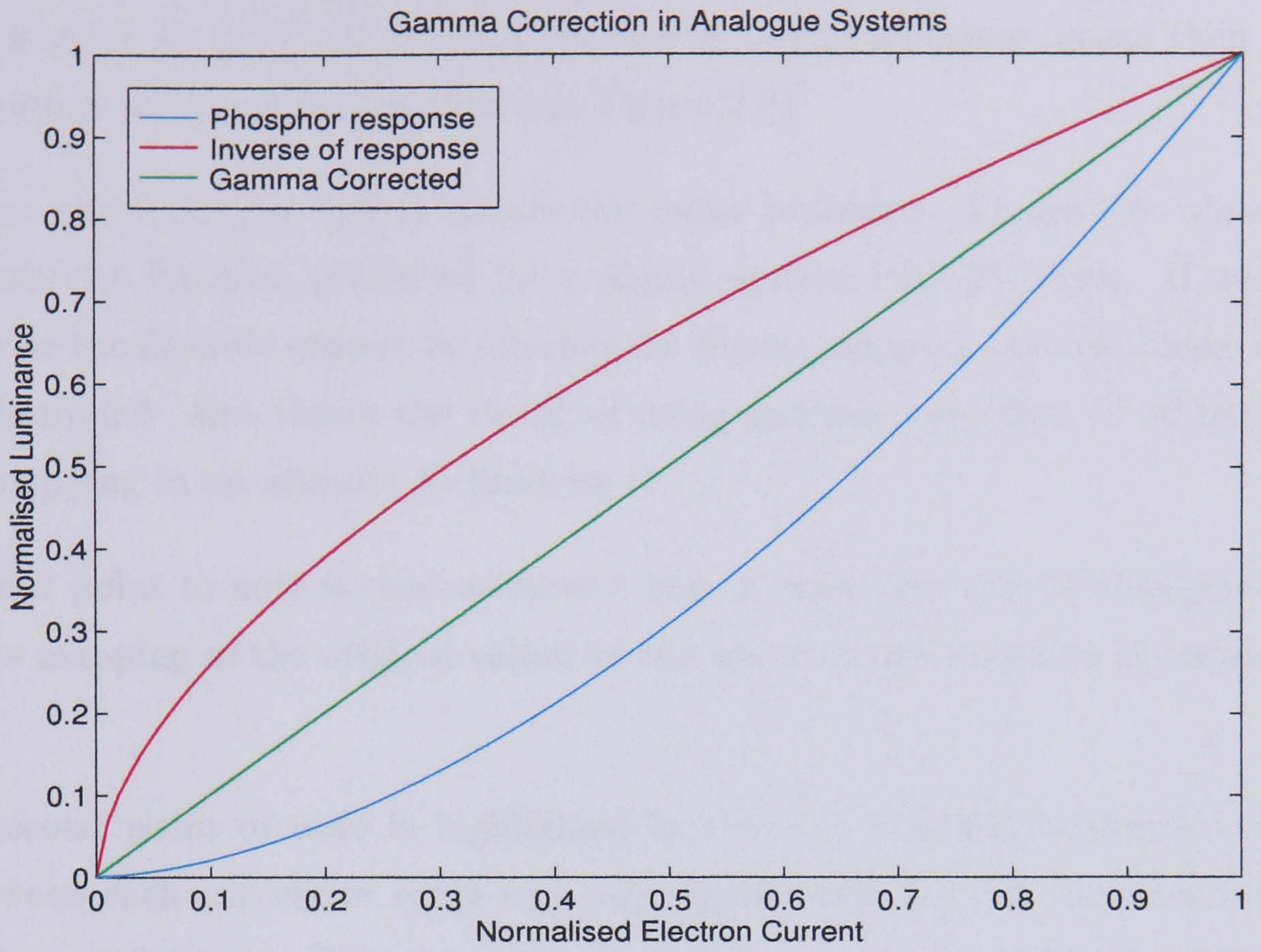

Figure 2.5: Analogue monitor response curves (Luminance = $current^{\alpha}$ where $\alpha =$ 1.7)

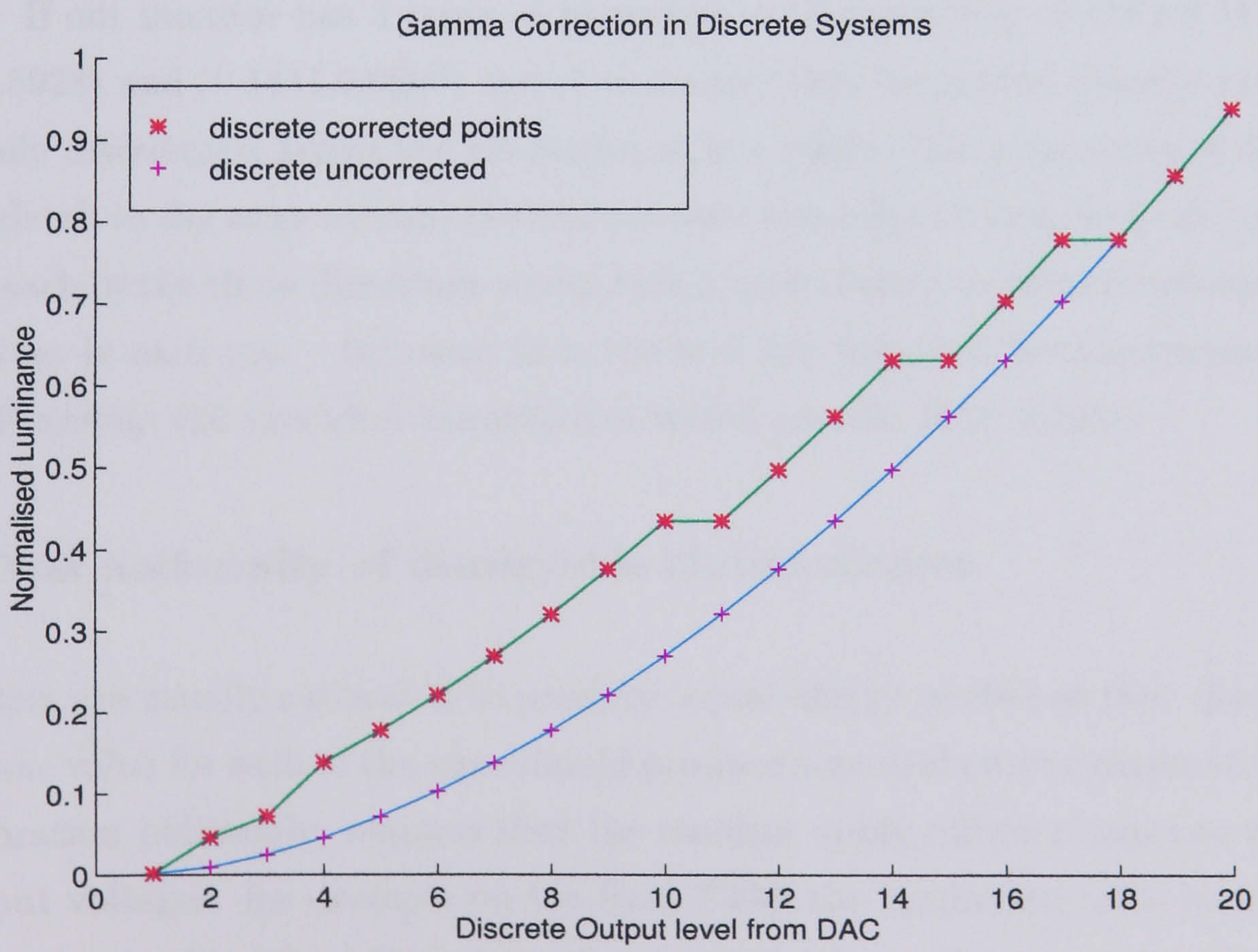

Figure 2.6: Discrete monitor response curves ($\alpha = 1.7$, levels = 20)

Figure 2.5 we should raise the input current to the power I^{γ} ($\gamma = \frac{1}{\alpha}$). i.e. Linear Emission is given by $(I^{\alpha})^{\gamma}$. If we could do this by purely analogue means then this transformation would suffice (as shown in Figure 2.5).

Gamma correction by digital means can cause problems. Figure 2.6 shows a sample response function produced for a digital system with 20 levels. If we try and linearise the discrete output by altering the digital mapping then problems may occur. Figure 2.6 also shows the result of using gamma correction to adjust the discrete mapping in an attempt to linearise it.

The first point to note is that we have a loss in resolution due to changing the one to one mapping of the original values to the many to one mapping of corrected values.

The second point of note is highlighted in this low resolution example. The steps between each successive value can vary significantly (as can been seen from the gradient variation). This may or may not be a problem depending on the variable being tested and the technique used, but should always be considered. As an example let us consider the digital resolution that would be required to perform a MacAdam style discrimination task for a patch centred on (0.305,0.323) in CIE xy space. If our monitor has 3 types of phosphors of chromaticity (0.6099,0.3473), (0.2922,0.5923) and (0.1441,0.0607) and if we assume that the grid of discrete values is uniformly distributed across the iso-luminant test plane (this assumption will be tested in detail in the next section) then to measure two sub-one-standard-deviation points in each of the three directions would take approximately a discrete resolution of 888 values in each gun – far more than the ≈ 6 bits obtained from linearisation of an 8 bit system (10 bits after linearisation would provide 1024 values).

2.4.2 Non uniformity of displayable chromaticities

Monitors are usually calibrated to produce 'equal energy white'; so that specifying the same value for each of the guns should produce a neutral output chromaticity. This calibration philosophy requires that the monitor apply differing gains to each of the input voltages: for example on the Eizo T560i the luminance ratio required for white is approximately 4.16:10.58:1.87 (red:green:blue). If we consider this in the context of discretisation of the input voltages then it is obvious that an increase

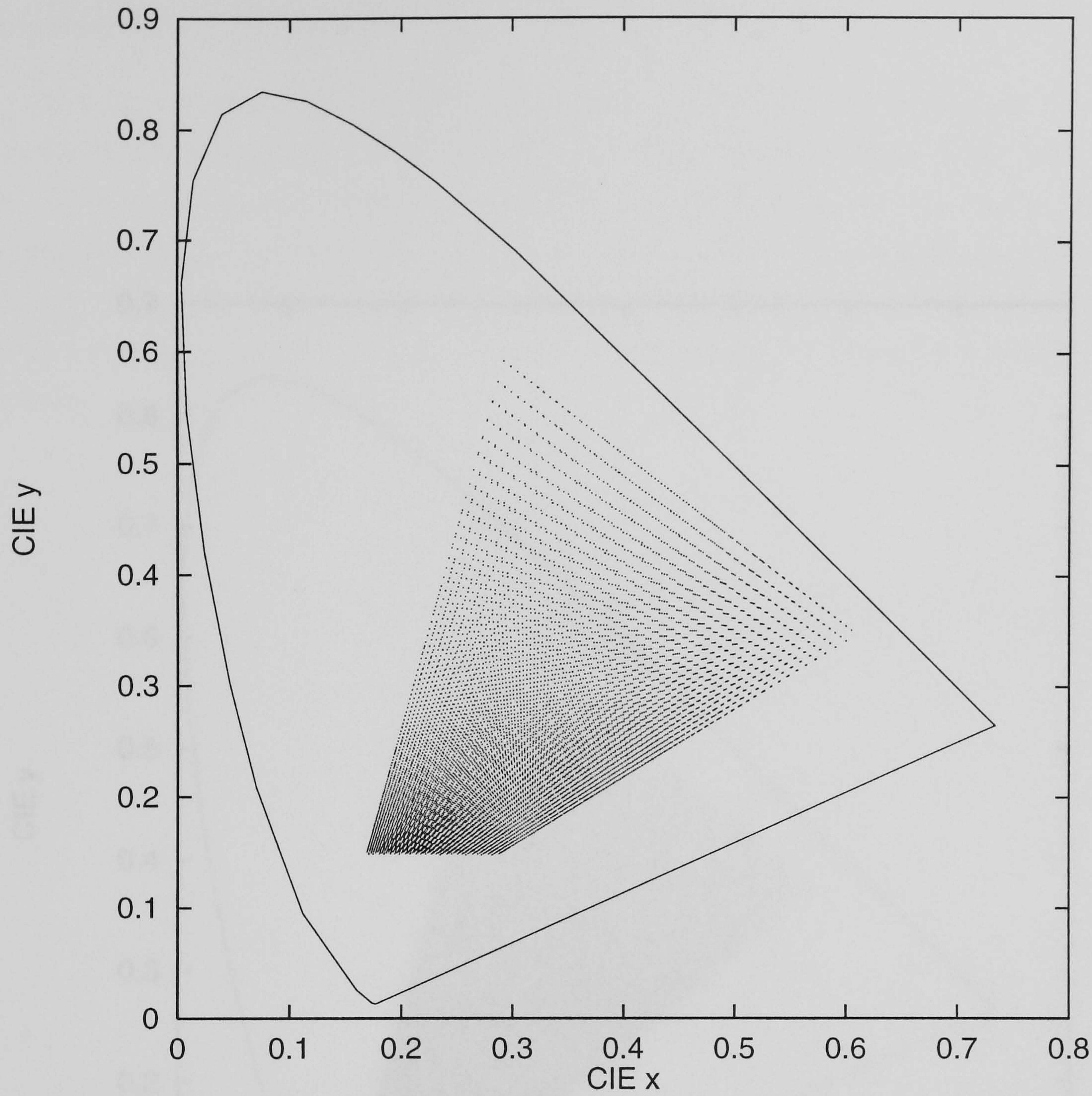

Figure 2.7: Unique displayable chromaticities in the iso-luminant plane Y = 6 cd/m^2

in the green gun DAC value will produce a larger change in luminance than a similar change in blue. If we attempt to produce the displayable chromaticities in a given iso-luminant plane (say 6.0 $\pm$ 0.1 cd/m^2) then it is clear that we have greater resolution in the blue gun than in the green. This means that in any iso-luminant plane the distribution of chromaticities will be non uniform. Figure 2.7 shows the distribution for 6.0 cd/m^2.

If we now consider what happens when we alter the iso-luminant plane two more factors emerge: firstly the distribution of colours alters with luminance – a chromaticity that is displayable at one luminance may not be available at another (see Figure 2.8); and secondly the gamut of available colours diminishes as luminance

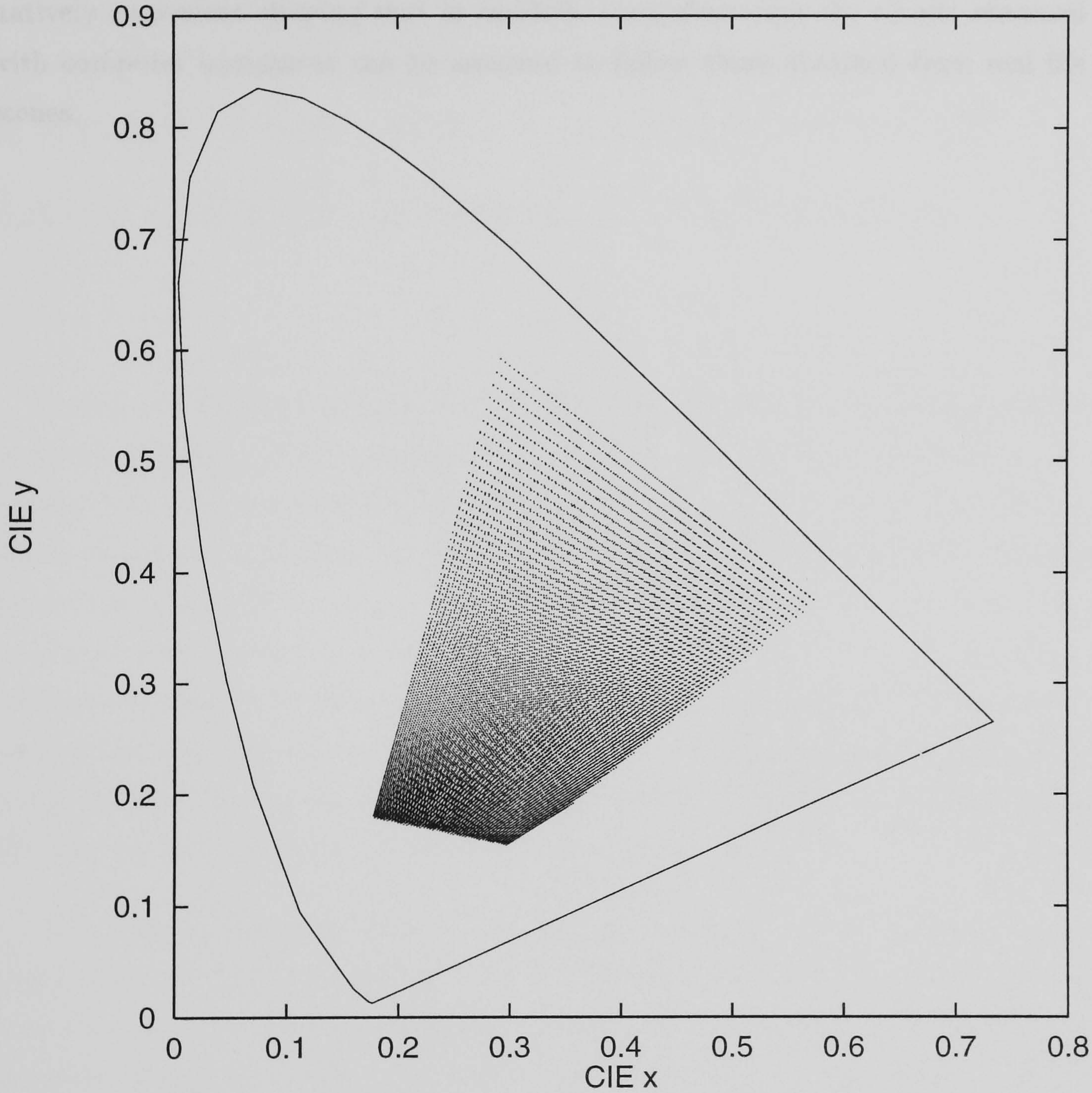

Figure 2.8: Unique displayable chromaticities in the iso-luminant plane Y = 9 cd/m^2

is increased.

In summary computer controlled displays provide repeatability, ease of use and versatility but have drawbacks that must be carefully considered whenever any study is performed. Savoy and O'Shea (Savoy and O'Shea 1993) compared the results for a computer display with those for the same 'real life' scene. The results were qualitatively equivalent showing that in carefully controlled cases the results obtained with computer equipment can be assumed to follow those obtained from real life scenes.

Chapter 3

Properties and Constraints of Low Level Visual Processing

Colour and lightness constancy are known to be affected by high level cognitive processes (Gilchrist 1977; Arend and Reeves 1986). Throughout this thesis an operational division is made between *sensory* (low level) and *perceptual* (high level) colour constancy. The term sensory colour constancy is used when purely sensory processing is referred to – i.e. the initial receptor signal transforms performed by the visual system towards the goal of colour constancy. The term perceptual colour constancy is used for processes used in discounting the illuminant that involve higher level factors such as scene interpretation, memory or instructionally induced cognitive processes. The two classes of mechanism are likely to operate in tandem in real life, but can be considered separately in controlled studies.

This chapter discusses some of the relevant features of the basic sensory processes of the human visual system. The first section reviews the possible structure of the system and examines information transmission constraints that predict the proposed morphology. The next section outlines the methodology and results of studies of colour matching that highlight factors that must be considered in assessing contrast and adaptation studies. The final section outlines colour contrast and colour adaptation experiments.

Chapter 4 examines evidence for high level effects on lightness and colour constancy and especially any implications relevant to the results of this thesis.

3.1 Low level sensory processes

Regardless of the exact processes that achieve colour constancy in humans, the phenomenon must be mediated through the signal processing pathways within the visual system.[1] Study of the basic operation of the visual system can provide useful information on cone signal transforms that may be key processes in human colour constancy. This chapter is concerned with studies of the low level visual system's response to simple stimuli (usually a single surface against a single background).

A demonstration developed by Gelb (Gelb 1929) in the late 1920's clearly showed that the visual system does not respond solely to absolute cone inputs. The experiment involves two patches of reflecting paper and an illuminant. In the first part of the demonstration the light is shone on a dark paper (the remaining field of view is entirely dark). Observers report the patch's appearance as white. Gelb then introduces a second surface into the scene that reflects more of the illuminant's light, as compared to the first. Observers now report the new patch to be white and the original patch as dark grey. Exactly the same signals are received by the cones for the original dark paper in the two cases (ignoring the possibility of light scatter within the eye), yet the perception of surface lightness significantly changes.

Land and McCann (Land and McCann 1971) show the same effect in a more quantitative manner. The experiment uses an array of black and white overlapping patches. The energy emanating from a perceptually white patch and a perceptually dark patch is measured with a photometer. The illumination across the whole collection of papers is now increased so that the dark patch is radiometrically identical to the original white patch (which is now brighter). The dark patch is still perceived as dark and the white patch as white. Exactly the same retinal illuminance (for the test patch) can be shown to produce the whole gamut of sensations from black through all the greys to white.

These two experiments illustrate two modes of viewing, the first being aperture mode (sometimes called void space, as an object or surface is presented in the void) in which the observer behaves in a photometric manner and reports surface perceptions that are a combination of illumination and surface reflectance. The second is surface

[1] This does not necessarily have to be the case. We do not know the 'locus' of colour constancy, so we cannot predict which signals are used, but judging by the efficiency of implementation of most neural systems, it is the most likely hypothesis.

mode in which a surface is placed within the context of a scene. In this case the observer reports colour perceptions that depend more on brightness and/or colour values of the surface relative to its surround and that are more analogous to surface reflectance properties.

This transform of absolute to relative values (or contrast) makes sense from considerations of our natural environment, which no doubt played a role in the development of our visual processes. The range of reflectances of natural surfaces varies from brightest to darkest in the ratio of approximately ninety to one[2]. Variations in illumination can be in the range of tens of thousands to one. Given limited resolution[3], if the visual system used an absolute scale of encoding, it is highly likely that reflectance changes in scenes would be indistinguishable from noise. Alternatively, since information on illumination intensity is of limited interest[4]; it is useful to concentrate as much of the available bandwidth on the coding of the surface reflectances of objects with which we wish to interact.

It is logical that the visual system adjusts the operating range of the receptors to match the surface properties observed within the visual field. In engineering terms this kind of system could be described as a variable offset and gain amplifier. There is little doubt that the early visual system performs such a process. What is not clearly understood is the exact interaction between the signals impinging on the retina and the offset and gain that results. Early in the 19th century, von Kries (von Kries 1905) suggested that the cone responses were scaled independently relative to the proportion of light received by each cone class. In later expositions of von Kries adaptation, a time constant is sometimes also included that scales the receptors not only relative to spatial cone excitation but also relative to the recent excitation within the cone class. Whereas in simple stimuli, von Kries adaptation can predict early vision responses it does not predict lightness and colour perception more generally.

At this point it is worth jumping forward and considering slightly later stages of

[2]Krinov (Krinov 1947) examined the surface properties in natural scenes and found variation in the range of 30:1. More recent analysis (Delbrook 1996) which includes synthetic materials suggests the higher ratio of 90:1.

[3]All neural systems are inherently noisy and as such have limited signalling bandwidth regardless of the signal encoding.

[4]From a biological survival standpoint it is useful to be able to recognise the approach of the night or storms etc., but small fluctuations such as those due to cloud cover are unimportant.

processing. The reasoning behind this is that it soon becomes clear that there are at least two variable offset and gain amplification stages present in sensory visual processing which can to an extent be examined independently. Each stage has different criteria which set the offset and gain and the second stage parameters are likely to depend on the output from the first. The first stage is also believed to be monocular (retinal) whereas the second stage probably operates after image fusion (cortical)[5]. If we assume such processing it becomes clear that we must be careful to examine the results of studies in the context of both stages because, in general, our measured output is in terms of the system as a whole.

There are several reasons for assuming the existence of two major variable offset and gain stages. The first line of reasoning follows from the standpoint of a maximisation of information transmission. The first stage offset and gain amplifier operates on cone signals (in part mediated by bleaching of cone photopigment and in part by fast neural adaptation via the horizontal and bipolar cell network (De Valois and De Valois 1988)). Re-examination of the cone spectra shown in figure 2.1 shows that the L and M cones have spectra that overlap significantly. It follows that the signals produced will be highly *correlated.*

Three transmission channels, two of which essentially contain the same information are not efficient from an information transmission standpoint. A better transmission strategy would be to decorrelate maximally the three cone signals and send that information instead. This decorrelation would improve signal to noise ratios and hence increase effective bandwidth. The effect of the decorrelation is essentially to redefine the axes of the LMS 3-space. Consider plotting all of the cone excitations from a stimulus: a cluster of points would result in the LMS 3-space surrounded by areas devoid of points (the highly correlated nature of the L and M axes ensure this). What decorrelation does is to define another set of axes that can represent all points in the filled sub-volume and little of the redundant space. This results in a shifting and magnification of the sub-volume maximising available signal bandwidth.

[5]This does not preclude the existence or formation of opponent signals at retinal levels. The assertion is that a gain and offset stage based on opponent channels operates after image fusion. The cortical locus of the second mechanism also is suggested by interocular induction effects which re-scale the sensations from both eyes in a manner dependent on the stimuli in both (Shevell and Humanski 1984).

Buchsbaum and Gottschalk (Buchsbaum and Gottschalk 1983) performed such a decorrelation analysis of the cone spectral sensitivities. They used a technique analogous to principal component analysis in which basis vectors are created that systematically try to account for as much available power as possible. For example the first principal component in their analysis accounts for 97.2 % of the signal energy that could be represented by the three cone types. This first component turns out to be analogous to a luminance channel. The second basis vector accounts for a further 2.78 % of the signal energy and signals red–green chromatic information. The third vector accounts for 0.015 % of the signal energy and signals blue–yellow chromatic information. The decorrelation transform they produced is:

$$\begin{bmatrix} lum(basis1) \\ r-g(basis2) \\ b-y(basis3) \end{bmatrix} = \begin{bmatrix} 0.887 & 0.461 & 0.0009 \\ -0.46 & 0.88 & 0.01 \\ 0.004 & -0.01 & 0.99 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{3.1}$$

which is shown graphically in figure 3.1. Recordings made from cat and monkey (Derrington, Krauskopf and Lennie 1984; Hubel and Wiesel 1968; De Valois, Abramov and Jacobs 1966) have shown that the early visual system has cells specifically designed to produce the differential signals that would be required in the decorrelation of the cone signal.

More evidence for the existence of such a transform comes from psychophysical studies. Hue cancellation experiments show that red-green and blue-yellow colour perceptions are relatively independent phenomena. A red patch can be returned to an achromatic appearance by the addition of green; the same applies to blue and yellow (Jameson and Hurvich 1961). This highlights the special position of the opponent channels in our visual perception, i.e. a red patch cannot be made to appear grey by the addition of blue.

Barbur, Harlow and Plant (Barbur, Harlow and Plant 1994) found that matching of chromatic test patches against multi coloured backgrounds (consisting of small squares) is virtually unaffected by continuously present, but randomly varying luminance contrast which suggests that the gain and offset control of the luminance and chromatic channels are, to a large degree, independent.

Currently we are assuming the early visual system to have two stages of variable offset and gain amplification. The first stage is a cone signal based operation that

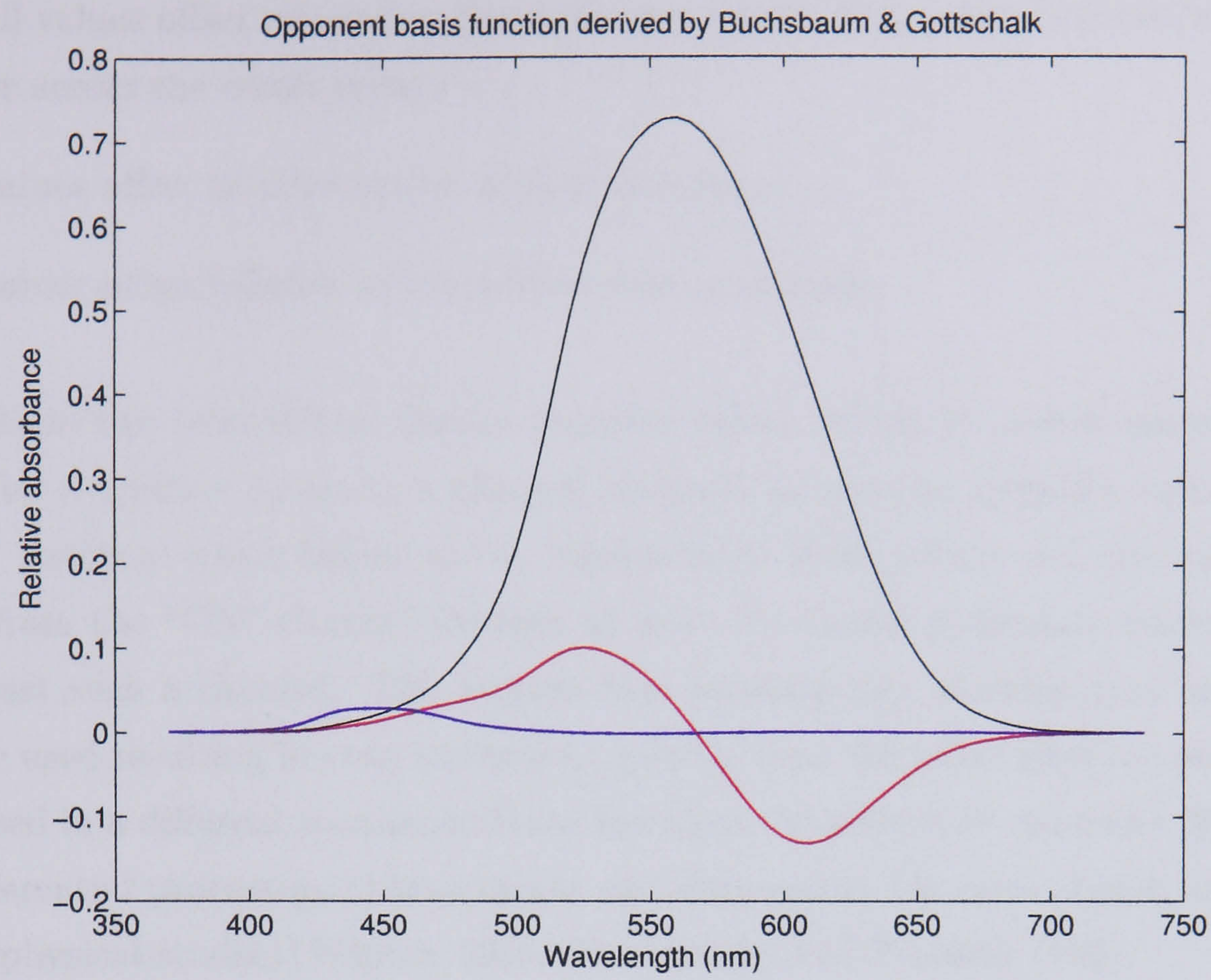

Figure 3.1: The colour opponent channels produced by Buchsbaum and Gottschalk

optimises the usable signal in each cone class based on the stimulus present, and possibly also upon stimuli in recent history. The output of this first stage then undergoes a linear series of transformations that decorrelate the cone outputs to produce three channels (i.e. luminance, r-g, b-y) as input to the second stage for variable offset and gain amplification.

Now, the decorrelation of the cone signals optimises the information transmission of all possible cone inputs – a given stimulus might not contain anywhere near this range of input. This raises the need for the second stage of variable offset and gain amplification which can be used to process further the cone opponent signals so as to optimise information transmission. The process involves shifting and expanding the axes of the cone opponent space so as to represent signal only (much the same as auto scaling a graph). The gain and offset of this second stage are likely related to the current input values and recent history of values but could be affected by a number of stimulus features.

There is another point that must be addressed in the simple two stage variable offset and gain amplifier model, that of offsetting. There are three major possibilities:

- all values offset relative to the minimum within-class value (calculated locally or across the whole scene)
- values offset relative to the within-class mean,
- values offset relative to the within-class maximum.

The second two possibilities require negative values, which the brain can only represent by a positive signal in a channel assigned as carrying negative values. The "OFF" pathway which begins in the bipolar layer of the retina and remains segregated from the "ON" channel through at least the lateral geniculate nucleus, provides just such a channel. This implies that possibly four variable gain amplifiers may be used resulting in cone excitations greater than the offset (increments) being processed in a different manner to those less than the offset (decrements). Evidence for differential processing of increments and decrements has been shown in several psychophysical studies (Whittle 1986; Chichilnisky and Wandell 1996).

We are now back at the stage where we can begin to evaluate studies that probe the adaptational properties of the visual system. We can assess the features of the stimuli and decide which stage of amplification was maximally affected and hence assess the parameters that determine the offset and gain of the amplification. We can also assess to what extent the model is valid and identify cases that necessitate higher level aspects of visual processing. It is also worth noting that so far we have not actually lost any information about the image: if all the gains and offsets are known, then apart from any noise losses (which should be small due to the encoding used) we can exactly reconstruct the array of cone excitations. All processing thus far has been about maximising the desired signal (surface reflectances) to noise ratio in the limited bandwidth available.

How could the cone offsets be calculated?

The answer to the question posed in the section heading is by no means clear. Several studies have shown strong offset effects based on stimulus features. This section examines some of these studies and continues the efficiency of information transmission arguments in an attempt to highlight the factors that may be used in offset calculation.

First, let us consider two major divisions amongst the possibilities for offset calculation. The offset may be calculated from:

1. features in the whole field of view – possibly a scene average of everything in view, the highest cone excitations in each class, etc.
2. a weighted calculation based on spatially local excitations.

Temporal effects may equally apply to both classes, so will be considered as a separate issue. It is also possible that a weighted combination of 1) and 2) is used.

In all of the following arguments it will be assumed that the visual system has a discrete number of levels that it can represent. This is not unreasonable as we know the visual system to have limited transmission bandwidth and to suffer from noisy transmission. These two factors enforce the physical constraint of a finite number of levels that can be distinguished from noise regardless of the transmission technique.

A calculation based purely on global field aspects can be very undesirable. Consider a simple indoor scene of a shiny jug on a table illuminated by sunlight through a window, against the backdrop of the rest of the room that is illuminated only by reflected light. The jug has a specular highlight which provides the brightest point within the scene.

If we examine an offset adjustment based on the global mean then a luminance histogram would reveal a large usage of grey levels above the mean (the jug and table) and a large usage below the mean (all of the background) with very little usage around the mean (a double peaked histogram). This distribution does not use efficiently the available resolution. A simple improvement would be to split the luminance levels into two and centre the first offset in the middle of the 'dark' histogram and the second in the middle of the 'light' histogram. To reconstruct the original would require knowledge of the two offsets and the areas to which they applied, but would allow almost maximal use of signal bandwidth. This transformation essentially encodes the image in a manner closer to the surface reflectance of the objects – the dual peaked histogram is a product of the illumination variation.

A similar argument can be made against basing offset calculation on a global scene maximum. If we consider the sample scene again, the brightest point is the specular highlight – which is much brighter than any other point in the scene. If this point were used as a normalisation factor we would have a scale from the darkest black to the white of the specular highlight. This would result in a large number of unused resolution levels between the specular highlight and the next brightest object within the scene.

The inadequacies of purely global offsets can be strikingly demonstrated through the use of digital cameras which tend to use global measures to control gain. It is easy to see the problems that features like specular highlights can cause simply by panning a digital camera across a scene like the one previously described. As soon as the camera has the specular highlight in view the rest of the scene darkens. A second problem is more subtle but can easily be seen in comparison with a camera designed to behave like the visual system (this camera uses near neighbour interactions to control the gain of each cell (Andreou 1996)). This problem is shadows: with a normal camera they tend to be too dark and many image processing techniques find them a serious problem. In the local-offset-camera the shadowed areas are dimmer but still have enough contrast so as to allow features to be discerned that are simply lost to normal cameras. An interesting addition to this is that the images still look remarkably 'natural' even though information about local absolute intensity has been discarded.

So from an information transmission standpoint, cone response offsetting and gain based on local features seems to be the best solution. This still leaves the questions: How large a local region? Which features of the local region should be used? and How are they combined? An extra dimension should be added at this point – time. The cone responses could also be related to their own previous excitations.

The eye very rarely looks in the same place for any length of time. Even when we fixate the difficulty the eye muscles have in holding the eye in place causes very small jitter in the image we see. So in everyday existence our cones rarely 'see' exactly the same point in an image continuously. Several studies (Ditchburn and Ginsborg 1952; Yarbus 1967) provide us with a rather striking example of cone adaptation. In their experiment they track the movements of the eye and shift the stimuli to mirror those movements. This procedure creates a *retinally stabilised image.* The results show that the image fades and eventually becomes invisible to the observer.

At first this result seems to point to an extremely local adaptation process, i.e. a single cone responding only to temporal changes. This could be attributed to bleaching of the photopigment within each cone which after a time stabilises and allows very little response to the stimulus (so little that it could not reliably be differentiated from noise). In fact, image fading is consistent with local and global schemes – as long as there exists a level in the hierarchy that responds only to changes in signals. For example if a cone response remains stable, and so does the offset level by which it is adjusted, then the same signal will result. If this signal is differenced to similar signals over time then the image will fade.

A process that could mediate the local offsetting could be *lateral inhibition* in which excitation of neighbouring cells acts as to inhibit the excitation of a centre cell. Several studies (Barlow 1953; Hartline 1940; Kuffler 1953) have shown the existence of lateral inhibitory processes within the retina, but unfortunately enough ambiguity exists in understanding of the exact function of the retinal cell types to forestall the prediction of the extent and exact functional characteristics of the reported lateral inhibitory processes.

To return to the questions posed earlier on the nature of local interaction: some support for the local interactions hypothesis, and estimates of its extent, come from the work of Walraven (Walraven 1973) and Valberg (Valberg 1974). The studies involved examining colour induction effects in simple centre-surround stimuli (a small test patch against a larger background). They found that enlarging the surround progressively affected perception of the test up to around $1.5^o - 2^o$ beyond the test. In an extension to this procedure Fairchild and Lennie (Fairchild and Lennie 1992) showed that extending the background to completely fill the field of view also had no further effect. From further adaptational studies with centre-surround stimuli of various durations,Fairchild and Lennie conclude that chromatic adaptation is spatially localised[6] with a time course of the order of 10 seconds.

Fairchild and Lennie used a task that required the observer to adjust the centre to appear achromatic. It can be argued that this is a very special task and as such may produce slightly different results from other procedures. The nature of the concern can be voiced as follows: given that it is known that the visual system uses

[6]They report that the mechanism of adaptation controlling colour appearance is not sharply localised, although at distances $> 1^o$ it has little effect.

second-stage encoding in which one channel is specifically luminance, an achromatic setting task requires the observer simply to minimise the output of the other two channels. This possibility is not a great concern and may even be useful in that it is likely to isolate the proposed 'first-stage' adjustments, but the results may not be entirely indicative of the function of the system as a whole when a different technique or stimulus is used.

From the evidence presented we could assume that 'first-stage' adaptation is a function that operates in terms of local spatial and temporal excitations and that the spatial interaction is limited to a small area around each cone. Cornelissen and Brenner (Cornelissen and Brenner 1992) performed several experiments which add further support. The first experiment placed an achromatic border between the test patch and the inducing surround and the results show the induction to be affected by the achromatic border width (up to 2^o). The second study shows that increasing the size of the patch has no effect on the induction, i.e. the extent of border which affects the test patch does not scale with object size.

The use of a weighted centre-surround operator (with temporal memory) makes sense from an information optimisation standpoint. The memory provides some *hysteresis* to the system which may serve to reduce the effect of eye movements and other rapid mean level shifts that are likely not to be indicative of the prevailing scene illumination. The weighted nature causes very local interactions to be highly significant and progressively more distant interactions less significant. If we consider the system as a whole, every cell to some extent affects every other cell (with psychophysically measurable effects up to 2^o) so the system can be thought to be strongly locally coupled with weak (but possibly significant) global interactions.

This analysis is likely to be a massive over-simplification, but may be illustrative for global behaviour. First, each stage in the model system is essentially linear, which is unlikely to be true on a small scale, but for many experiments the results appear to be at least piece-wise linear (or log linear). So we can use the linear model to assess at what point non-linear processing within each stage should be considered. The two stage model itself is non-linear, although it is likely to behave in a linear manner when overall scene contrast and mean remain constant (i.e. the second stage gain and offset are fixed).

The visual system does behave in a surprisingly linear (or log linear) manner for a range of stimuli. Photoreceptor response, for example, can show highly linear characteristics in response to simple flash stimuli. In addition, responses to brightness contrast are also remarkably linear (provided a basic discrimination threshold is exceeded) (Whittle and Challands 1969).

Major losses in linearity are likely to occur when special function systems begin to take effect. For example, in a brightness study performed by Whittle (Whittle 1992), when a test patch of incremental brightness exceeds certain levels relative to the background the response function begins to shift away from what would be expected by a ratio hypothesis. As suggested by Whittle this may be a compensatory mechanism built into the system that attempts to counter increased light scatter caused by physical properties of the eye's imaging system (reducing contrast). Marimont and Wandell (Marimont and Wandell 1994) come to a similar conclusion concerning the relative rarity of blue cones – that the chromatic aberration of the eye's physical optical system essentially dictates that more would be of little use.

It is worth noting that using two stages of variable offset and gain amplification in which the first stage is controlled by a known local spatio-temporal weighting function still allows full reconstruction of the original excitation pattern (as long as some offset and gain information is also retained and noise is not a significant factor). All we have developed so far is a highly efficient means for representing as much useful (object not illuminant based) image content as possible in a restricted bandwidth scenario. (Note: all information reduction due to spatial frequency filtering etc. has not been considered in this model). This is significant because it is known that high level scene interpretation can affect surface colour and lightness perception (these studies are considered in more detail in the next chapter). This re-adjustment of perceived surface properties suggests that all of the lightness and colour assignment of surfaces is not performed by purely static linear processes and also that it is important that early processes do not discard too much information.

It is also worthy of note that the two stage encoding suggested should go a very long way towards lightness and colour constancy. It is by no means sufficient to describe human behaviour in such circumstances, but should to a large degree remove significant portions of the illumination artifacts within images. This does not mean that we can create an image that is the same as the original with the illuminant artifacts removed, but rather that the *output* from the processing will be similar for the same scene under differing illumination conditions. This distinction is important and is not highlighted enough in studies of colour constancy. Most studies (including this thesis) measure colour constancy performance by the level of compensation that would remove the effects of the illuminant and transform the scene we view to a new scene under a standard illuminant. All a colour constant system really implies is that the external representation of surfaces within two scenes is transformed to approximately the same internal perceptual surface representation (Forsyth 1990).

3.2 Effects of low level processes on lightness and colour constancy

As mentioned in the previous chapter, MacAdam produced the first major study of human colour discrimination as defined by observers' adjustment of a test field colour. The experiment employed a test stimulus consisting of a circle (2^o diameter) split into two halves against a 21^o radius uniform coloured field. The results clearly show the variation in human discrimination capabilities for colours displayed against CIE illuminant C.

Particular, and sometimes seemingly insignificant, details of stimulus configurations could have strong effects on the results of an experiment. We can use MacAdam's discrimination study as an example of some of the possible factors that contribute to his results. His procedure used simultaneously viewed abutting patches. This configuration can be thought to show the properties of low level cone interactions more than observer colour discrimination abilities. The task can be performed by an observer adjusting the dial until the two patches become homogeneous, which entails the use of differential signals at the patch borders and not necessarily any assessment of perceptible colour differences. The results of MacAdam are therefore valid only for this stimulus configuration, and are likely to set an upper limit on colour discrimination performance.

Additionally, Boynton, Hayhoe and MacLeod (Boynton, Hayhoe and MacLeod 1977) show a 'gap effect' in discrimination of achromatic and chromatic tests. Whereas luminance discrimination was impaired by the addition of a small gap (luminance step) between the two samples, chromatic discriminations were either unaffected or improved (depending on the colours used). The improvement in this case could be due to an increase in the definition of a border between two isoluminant coloured regions.

The colour discrimination abilities of human observers are also affected by the state of adaptation of the system (Fairchild and Lennie 1992; Burnham, Evans and Newhall 1957) and as such the MacAdam ellipses are really only a snapshot of discrimination capabilities for the colours tested under the state of adaptation induced in the observer.

Another significant feature of MacAdam's experiment is that of the match technique. In MacAdam's experiment the observer adjusted a single dial. This one-dimensional parametric matching was facilitated by the choice of two filters to act as end points of a vector in the selected CIE xy plane. This filter choice, made by the experimenter, means that only fine adjustment needs to be made to make a match – since the general region of colour space has already been selected.

In a similar study Newhall, Burnham and Clark (Newhall, Burnham and Clark 1957) performed colour matches with almost identical stimuli, but used a colorimeter (consisting of three illuminants that the observer could increase independently to create a mixture that matched the test presentation). They reported discrimination ellipses on average 27.6 *times* the area of those obtained by MacAdam. The difference between the results could arise from two main sources. The difficulty of a three parameter match compared to a single parameter match could increase errors but should improve with practice. The second and more likely source is a form of tracking response in which the visual system alters parameters continuously as the matching patch changes. As mentioned previously, Fairchild and Lennie assert that adaptation can take up to 10 seconds to complete (the time to stabilise is related to the magnitude of the change). So in MacAdam's experiment the small adjustments in colour can be quickly compensated for and a steady state reached, but in experiments in which large shifts occur the time to stabilisation of the colour signal may be up to 10 seconds. The observer may therefore be matching to a transient percept. The match the observer sees and signals at one instant may therefore vary widely and the matching error will be correspondingly large. The error could be worse if large areas are changed. DePriest, Krauskopf and Lennie (DePriest, Krauskopf and

Lennie 1991) made recordings from cells that took minutes to stabilise after a step change in chromatic background (luminance steps take less time).

Another possibly significant feature of MacAdam's stimulus (and of centre surround stimuli as a whole) can be highlighted by considering the two stage variable offset and gain amplifier model again. In the impoverished environment of the centre surround configuration there is very little information. In fact, only four signals exist:

1. the contrast of the test patch with the background;
2. the contrast of the test patch to the match patch;
3. the contrast of the match patch to the background;
4. the contrast of the rest of the visual field (usually black) with the uniform background.

These may reduce to three values when a match is made[7]. There is no problem for the first stage amplification which will offset the levels and signal edge contrasts to the opponent mechanisms. Now – remembering the assumption that the second stage amplifier adjusts offset and gain to just contain all of the contrast information within the image – we can see that this form of stimulus would cause the second stage amplifier to essentially adjust its gain values to track the observer adjustments (within noise and maximal amplification constraints) and allow colour discrimination that is unlikely to occur when larger arrays of contrast are present in the image. This is a completely hypothetical assertion, but the possibility should not be discounted especially in the light of the findings of Webster and Mollon (Webster and Mollon 1991) which showed contrast adaptation effects in post receptoral mechanisms (which are analogous to the second level amplification hypothesised). Although the effects were only shown for temporal modulations a similar pattern of excitation could be induced through eye movements.

Colour induction is a phenomenon whereby the surroundings of a test patch can cause a change in the perceived colour of the patch. It is believed to be, at

[7]If the observer balances the cone contrast ratios of the test and match with respect to their backgrounds.

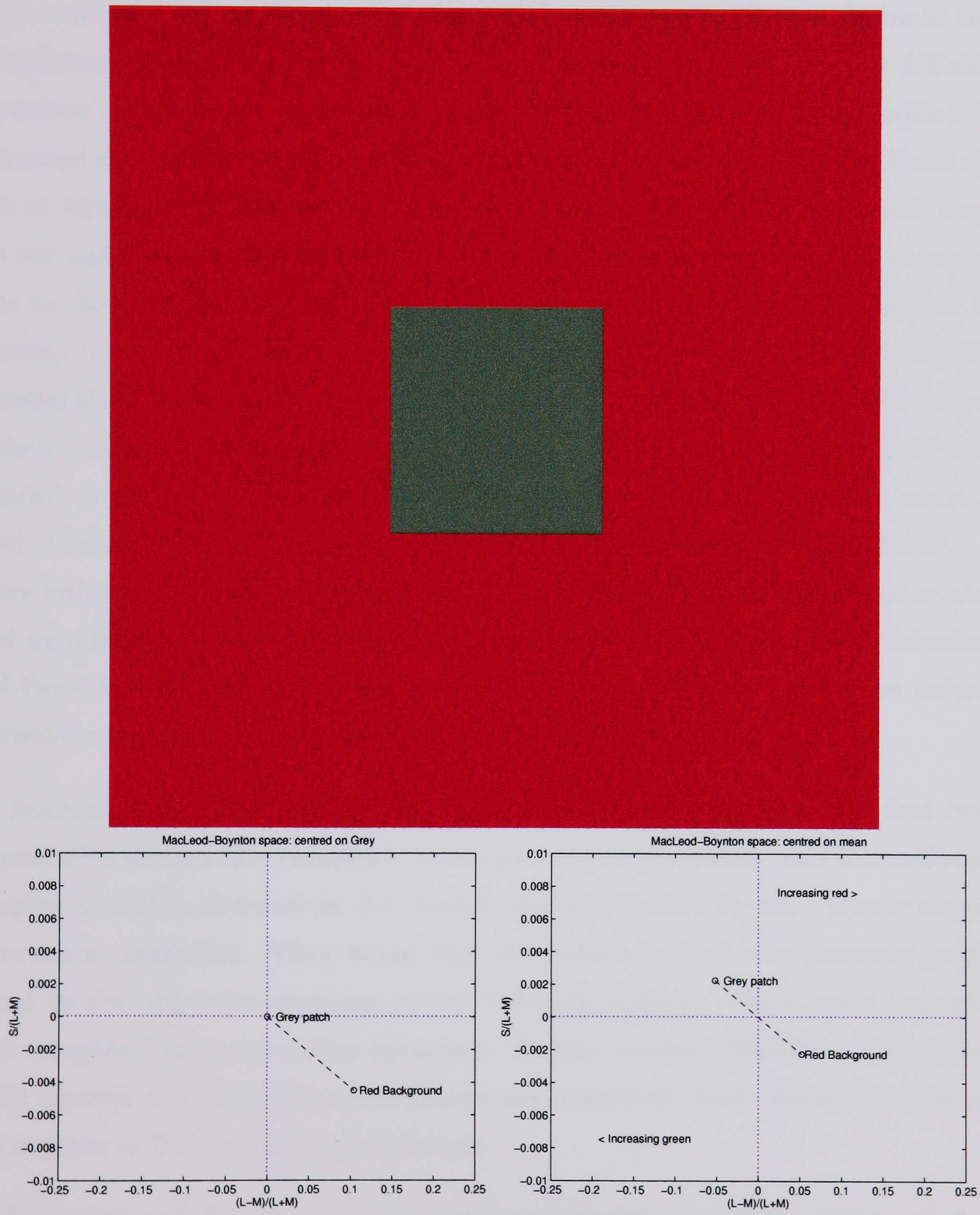

Figure 3.2: The shift explanation of colour induction

least in part, mediated by cortical function (Shevell and Humanski 1984; Boynton 1956; Krauskopf and Zaidi 1986) and hence could be a product of the offset and gain mechanism suggested for the second stage of amplification. Consider Figure 3.2 which shows a simple centre-surround stimulus with an achromatic centre and a red inducing surround. The effect of chromatic induction would be to make the achromatic patch appear greenish grey. The lower left plot in figure 3.2 shows the colour opponent space plot for this simple stimulus. If we apply a simple transformation that shifts the mean of the points to the origin (shown in the lower right plot of figure 3.2 we can see that this would make the inducing surround appear less red and the achromatic patch less red – which is equivalent to being more green (due to the opponent channel coding). If we also perform a gain normalisation that expands the stimuli to fill the space we can see that the effect will be increased for minimal stimuli (especially if the gain operation occurs first) and will be a function of the surround and test patch colours. Following this argument, if a range of colour contrast is present in the stimulus the second stage gain and offset are essentially fixed at smaller values (as long as the extent and mean are set by the contrast) and hence induction effects are reduced (this form of adaptive gain adjustment is examined quantitatively in section 6.4). Studies (Shevell and Wesner 1990; Cornelissen and Brenner 1989) do show a reduction in chromatic induction when the inducing surround consists of many patches of various colours.

Jameson and Hurvich (Jameson and Hurvich 1961) performed the first major quantitative studies into chromatic induction. The experiments used Munsell colour samples as match alternatives (i.e. match choice is limited by the discrete range of reflectances available). They found that the induced colour was inversely proportional to the opponent response of the inducing squares, which again suggests a post receptoral locus for colour induction. Further studies (Walraven 1976, Shevell 1978) showed that induction phenomena are consistent with colour selective gain attenuation in the opponent mechanisms.

Blackwell and Buchsbaum (Blackwell and Buchsbaum 1988) showed that the effects of an inducing surround on the centre increases exponentially as a function of surround size (in the range 0 to 35 minutes of arc – at which point they show no further results). They also reported that the effect of a surround on the centre reduces exponentially as a function of separation of the centre and the inducing surround, in the range of 0 to 14 minutes of arc. (Although there is still a contribution at $14'$ they do not report the exact value at which no effect is seen). Blackwell and Buchsbaum also studied the relationship between colour of the centre and colour of the inducing surround. For the colours tested, induction was maximal when the angle between the hues (in an r-g, b-y plane) was in the range 35 to 65 degrees. Only three centre colours are reported and the results only show the angle dependence up to 90^o, so the results cannot be conclusively assessed in terms of compliance with opponent channel gain explanations. The authors do claim that colour induction seeks to remove the common elements within the centre and the surround, which is equivalent to the shift operation used earlier.

The shift technique is very similar to that used by Webster and Mollon (Webster and Mollon 1991) to explain their results. Additionally, their results suggest that a simple gain for each opponent channel is *not* sufficient. For their stimuli two sets of gain were required but they were not aligned with the opponent axes, but rather with the axes of temporal modulation and its orthogonal axis. This raises some interesting possibilities.

Firstly, this may suggest that multiple chromatic channels exist in cortex (as proposed by Webster and Mollon). This scheme would imply a third amplification stage to the model in which n chromatic channels split the hue circle into n segments each with variable gain, which would relegate the first two stages to the role of efficient signal transmission processing. Note that Webster and Mollon do show a slight dominance of the opponent axes, which suggests some gain adjustment may occur in the opponent channels also.

A second possibility has been explored in detail by Atick, Li and Redlich (Atick, Li and Redlich 1993). They suggest that a further stage of axis manipulation occurs that decorrelates the opponent channels (this processing would replace the offset and gain of the proposed second stage). The decorrelation is dependent upon the signals present in the channels and in graphical terms is analogous to a rotation by θ (the angle made by the principal axis of temporal modulation with the positive r-g axis) followed by scaling of the new x and y axes (to fit excitation) followed by a rotation of $-\theta$.

An interesting point raised by Atick, Li and Redlich is that the stimuli used by Webster and Mollon were too minimalist to allow direct calculation of the decorrelation vectors. If the visual system does decorrelate the signals at this point it will experience the same ambiguity and have to fix variables to values that may be related to the processing capabilities of the system as a whole or stimulus features that are not used for the computation in other scenes. (Note: the rotation- scale-rotation transform that Atick, Li and Redlich assumed did fit the data very well, but was not directly calculated from a decorrelation of the opponent channels as they had hoped. The form of the solution they obtained does suggest such a transform but assumptions had to be made as to the role of certain parameters in their solution).

Unfortunately, the design of an experiment that could distinguish between the two suggestions is difficult as depending on the number of chromatic channels assumed the two schemes would produce qualitatively the same results. What may prise the two apart is if the number of chromatic channels were relatively small which would lead to slight inhomogeneity in adaptational response to certain angles in the hue circle. The adaptational effects defined by decorrelation would be homogeneous for all test angles. If the number of channels is reasonably high then the variance in observer matches may also make the results of such an experiment inconclusive.

3.2.1 Viewing techniques for colour matching

We can now move on to more procedural aspects involved in colour matching. We have seen that the particular stimulus configuration can greatly affect results. One of the first considerations must be whether or not it is valid to match something in one retinal location to that in another and extrapolate the results to other retinal locations. One answer lies in a study by Burnham, Clark and Newhall (Burnham, Clark and Newhall 1957) which found no significant differences for matches made for four different stimulus configurations. The stimulus consisted of a $2^o \times 2^o$ bipartite field. This field can be rotated in steps of 90^o to produce the four stimulus configurations. They found no significant difference between the four cases. Cone variation throughout the retina (i.e. as retinal eccentricity increases the frequency and type of cones alter) dictates that there should be variation in colour perception, but the variations appear to be symmetrical. Thus provided test and match stimuli are presented in a 'balanced' manner (i.e. symmetric through the centre of the retina) the results should mirror visual system limits and not cone mosaic variation.

In colour constancy experimentation we wish to compare the perceived colour of a surface under one set of illumination conditions with that of the same surface under a different set of illumination conditions. Consideration of the foregoing arguments with respect to retinal and cortical adaptation show that this task will require careful design of the match technique.

There are five main possibilities:

1. simultaneous binocular – both eyes may see both illumination conditions simultaneously. The state of monocular adaptation in this case is uncontrolled. The scenario can be said to exist in real life where parts of scenes are subject to very different lighting conditions. If purely local rapid spatial mechanisms only underlie colour constancy capability, then this viewing configuration should produce a high degree of performance.

2. simultaneous monocular – a septum separates the two eyes which view spatially separated stimuli (not fused into a single image). One eye views one illumination condition and the other eye another. Monocular adaptation should be highly controlled but cortical interactions are unpredictable.

3. fused haploscopic display – a septum separates the eyes which view different illumination conditions. The display is fused through the use of a stereoscope. Monocular adaptation is highly controlled but cortical interactions are still unpredictable. This form of display appears to enhance the adaptational effects: the adapting backgrounds may become so homogeneous that when adaptation is complete a chromatic illumination difference can be difficult to detect. This is not usually the case with non-fused displays. This display also holds monocular adaptation more stably – in the non-fused case the observer usually moves his eyes between the two stimuli which means that one eye will be directed towards the dark septum. In the fused case both eyes continuously view the stimulus or adapting backgrounds.

4. successive binocular – the observer views one illumination condition with both eyes and then (usually after a forced delay) views another illumination condition. Unless there is a delay of over 2 minutes (DePriest, Krauskopf and Lennie 1991) chromatic monocular adaptation may not be completely controlled. Cortical adaptation – although not affected by differential excitation – may also be unpredictable. This case also requires the use of memory which in itself can affect matches.

5. successive monocular – the observer views one illumination condition with one eye and then after a delay the other condition with the other eye. Several implementational issues exist in this technique. For example, what should the other eye view when not viewing the stimulus? If it views a black field then monocular adaptation is not entirely predictable. Alternatively, if a diffuse mean field is displayed then the same cortical interaction problems may result.

Newhall, Burnham and Clark (Newhall, Burnham and Clark 1957) examined the properties of matching under simultaneous and successive conditions (both binocular). They used the same stimulus configuration[8] so that any errors could be attributed to the viewing technique used. In summary they found that successive matching yielded: higher variability of replicative matching, shorter match times, systematically higher purities (colour saturation) and somewhat higher luminances. The found that the memory match condition (successive) produced discrimination ellipses on average 6.2 times the areas of those for simultaneous matching. Another feature of the study that differentiates it from some other successive forms is that the test was shown only once at the start of a trial.

Eastman and Brecher (Eastman and Brecher 1972) performed a similar study in which they compared viewing techniques. The major difference in this study was the use of a shutter mechanism that controlled the amount of time each eye could see each illumination field. The sequence consisted of: left eye views field for 4 seconds (right occluded), both eyes occluded for 4 seconds, right eye views other field for 4 seconds (left occluded), then both eyes occluded for 4 seconds. The sequence was repeated until a match was complete. The 4 second duration was chosen because Hubel and Wiesel (Hubel and Wiesel 1968) found that cortical interactions to binocular stimuli died down after around a second – an extra 3 seconds was added as a safety margin. This technique does, to a certain degree, solve the cortical interaction problem but may do so to the detriment of monocular adaptation. The timings ensure a steady state of monocular adaptation but not necessarily a complete one. The study concluded that the successive monocular viewing technique produced the best results (in terms of compensation for an illuminant shift), closely

[8]There was a minor stimulus difference in that in the simultaneous case the test and match fields were created by dividing the $2^\circ \times 2^\circ$ match field into two abutting fields. The complete match field was used in the successive case.

followed by simultaneous monocular and the worst being simultaneous binocular (fused haploscopic was not tested).

More recently Fairchild and Munsell (Fairchild and Munsell 1994) performed a study examining variants of all of the techniques. They performed correlation analysis on the results showing that there is a high correlation between successive and simultaneous (fused) monocular displays. The ranking of the results in terms of illuminant compensation follow those of Eastman and Brecher i.e. successive monocular produced the highest degree of adaptation closely followed by fused haploscopic.

Several studies (Brainard and Wandell 1992; Brainard and Wandell 1991; Bäuml 1994; Newhall, Burnham and Clark 1957; Troost and deWeert 1991) use extended variants of the successive binocular technique. Observers are either trained to memorise the colour of a test patch, match to a memory colour or match to a colour category (i.e. grey). This technique does have the advantage that very natural viewing can be used and retinal and cortical adaptation allowed to complete. The disadvantages are that the long term memory required could mean results are more related to colour classification issues than pure colour constancy performance. Boynton et al (Boynton, Fargo, Olson and Smallman 1989) report that discrimination for coloured patches, that clearly differ when presented simultaneously, can change when a delay of 10 seconds is introduced between presentations. Colours that, when presented simultaneously, are clearly different can be reported as the same by observers after the delay is introduced.

3.3 Effects of low level processes on lightness and colour constancy

The results from several chromatic adaptation studies have been discussed in the previous sections with respect to the hypothetical two (or possibly three) stage model. There have been many more studies with results that are particularly relevant to the experimental work covered in this thesis. This section highlights some illustrative examples.

Burnham, Evans and Newhall (Burnham, Evans and Newhall 1957) performed colour matches under differential illumination conditions using simultaneous monocular displays (non-fused). The stimuli were of the centre surround type. They found they could predict each shift with an affine transformation of the chromaticity co-ordinates i.e.

$$X = a_{11}X^{'} + a_{12}Y^{'} + a_{13}Z^{'} + a_{14}$$
$$Y = a_{21}X^{'} + a_{22}Y^{'} + a_{23}Z^{'} + a_{24}$$
$$Z = a_{31}X^{'} + a_{32}Y^{'} + a_{33}Z^{'} + a_{34}$$

The a_{ij} coefficients were derived from empirical data for *each* illumination condition. The coefficients differ significantly and no analysis was performed to try to predict them from the illumination shifts used. In the context of the multi stage model proposed, this experiment most likely shows the adaptation capabilities of the first stage of processing i.e. retinally based compensation.

More recently Brainard and Wandell (Brainard and Wandell 1991) performed a similar experiment in which observers adjusted a match patch to the same colour as a test patch, the colour of which they had been trained to remember. The stimulus used is classed as a colour constancy stimulus as it has several different reflecting surfaces from which the observer may be able to infer properties of the illumination (the stimulus consisted of a square array of Munsell patches against a uniform background, which could also be seen between the patches where a small gap existed). They found they could explain their data with a bilinear model related to the prototype patch actual and perceived reflectance and the illuminant changes in terms of the basis vectors that can describe both illuminants. The degree of compensation for illuminant shifts they obtained was estimated as only half the amount required for perfect colour constancy. Again, in terms of the proposed simple model, this experiment mainly shows the response of the first compensation stage. Although the stimuli contained colour contrast, the patches were drawn from the set of Munsell patches at random. This means the effects on a second stage mechanism are inconsistent and most likely to be manifested as per trial matching errors. Interestingly some of the subjects tested simply could not reliably memorise the test patches - though those who did apparently did so quite repeatably.

In a later study Brainard and Wandell (Brainard and Wandell 1992) used much the same stimulus configuration and matching procedure as described in the previous section. An interesting point they raise is how local interactions between the match patch and surrounding patches could affect the results. They found in pilot studies that moving the match patch was critical, as using a fixed location introduced strong contrast artifacts in the match results. They do not, however, tell us if moving the test patch was strictly necessary or whether changing the surrounding patches was sufficient to disrupt the match dependence on the local spatial neighbourhood.

Brainard and Wandell also mention that, in the colour training phase, feedback as to the accuracy of a match was useful in reducing the variance of observer matches. This seems an interesting distinction in that active participation in the adjustment assists in the memory process. The observers only trained under the reference illumination condition and no feedback was given for the main experiment, although this does raise the question as to the effects of the specific reference illuminant chosen and what feedback training to a different illuminant would produce.

The experimental procedures employed by Walraven, Benzschawel, Rogowitz and Lucassen (Walraven, Benzschawel, Rogowitz and Lucassen 1991) and Lucassen and Walraven (Lucassen and Walraven 1993) are rare in that they used a stimulus with a large amount of chromatic contrast which was kept constant throughout the experiments[9]. The test and match arrays were exactly the same and consisted of 35 Munsell patches on a 100 % reflecting background. The patches were separated so that the background could clearly be seen between them. The test stimuli were chosen so as to make a circle in CIE xyY space centred on the achromatic point. This choice is interesting as it allows easy assessment of the effects of the illuminant shift by examination of the distortion of the circle. Another interesting feature is that it is easy to tell at a glance if the matches are merely offset or whether gain changes occurred which distort the shape of the circle.

In the first study (Walraven, Benzschawel, Rogowitz and Lucassen 1991) the contrast explanation of colour constancy was explored. As in other studies, several matches were made under different illumination conditions (test illuminant against a standard illuminant; viewing was of the non fused haploscopic variety). They examined the match results in terms of the cone excitations of the test and match patches with respect to the background (the 100 % reflecting surface). They concluded that, in general, the observer matches could be explained by cone ratios of

[9] Strictly speaking this is not the case as the 35 patch reflectances used in the stimulus remained the same but illumination did vary throughout the experiments resulting in patch contrast changes with respect to the background and an accompanying interocular contrast change.

the test and match patches with respect to the backgrounds. The correlation found was high for the L and M cones but a deviation was noted for the S cones that could not be carefully examined due to the relatively small illuminant shifts employed.

The second study (Lucassen and Walraven 1992) used larger illuminant shifts so that the deviations of the S cone could be further examined. In this experiment both test and match presentations were displayed under chromatic illumination (i.e. the match area was not under a standard illuminant). They found that a compressive non-linearity was required to allow the prediction of the S cone matches. This extra factor can be interpreted directly into the simple model proposed. The first stage provides the offset and gain that produces the von Kries style cone ratio predictions and the second stage the compressive non linearity (in the form of another gain adjustment that is related to the in-channel contrast within the scene). They found that the compressive non-linearity was only required for the S cone. This is not a problem for the model for two reasons: first the model predicts that the second stage gain should be applied to the opponent channels and not the L and M signals. Second, as in most colour constancy experimentation, the stimuli were designed to be near isoluminant – a property that restricts the range of values that the L and M cones can assume (especially the ratio of test and match patch compared to the background). This may restrict the variation so much that the non-linearity predicted for the L and M cones is so small that it is indistinguishable from matching errors.

Chapter 4

High Level Influences on Lightness and Colour Constancy

The division was made in chapter 3 between *sensory* and *perceptual* effects in colour and lightness constancy. This chapter examines studies in which high level scene features (i.e. segmentation into objects, perceived depth, knowledge, instructions etc.) have an influence on the surface perception of objects within a scene.

An early example that image segmentation can have effects on perceived surface lightness was provided by Koffka (Koffka 1935). In his illusion (Figure 4.1) we can see that the addition of the dividing line changes the perceived lightness of the two circular halves from the nearly uniform appearance of the complete circle. This example raises interesting questions as to the calculation of lightness based on edge ratios. It is suggested that since absolute image intensity information is lost due to photoreceptor habituation, then colour and lightness must be re-attributed by a fill-in process based on edge ratios. Consider the complete circle: in this case the very different edge ratios of the two halves of the circle are attributed to differing backgrounds and hence are consistent with a circle of uniform reflectance. When the circle is divided the edge ratios at the dividing line allow an alternative interpretation of the scene – now we can see a dark annulus on a white sheet of paper beside which is placed a dark sheet of paper with a half annulus on top of it.

Gilchrist (Gilchrist 1977) managed to quantify the change in lightness that a surface may undergo with perceived configuration in a cleverly designed experiment.

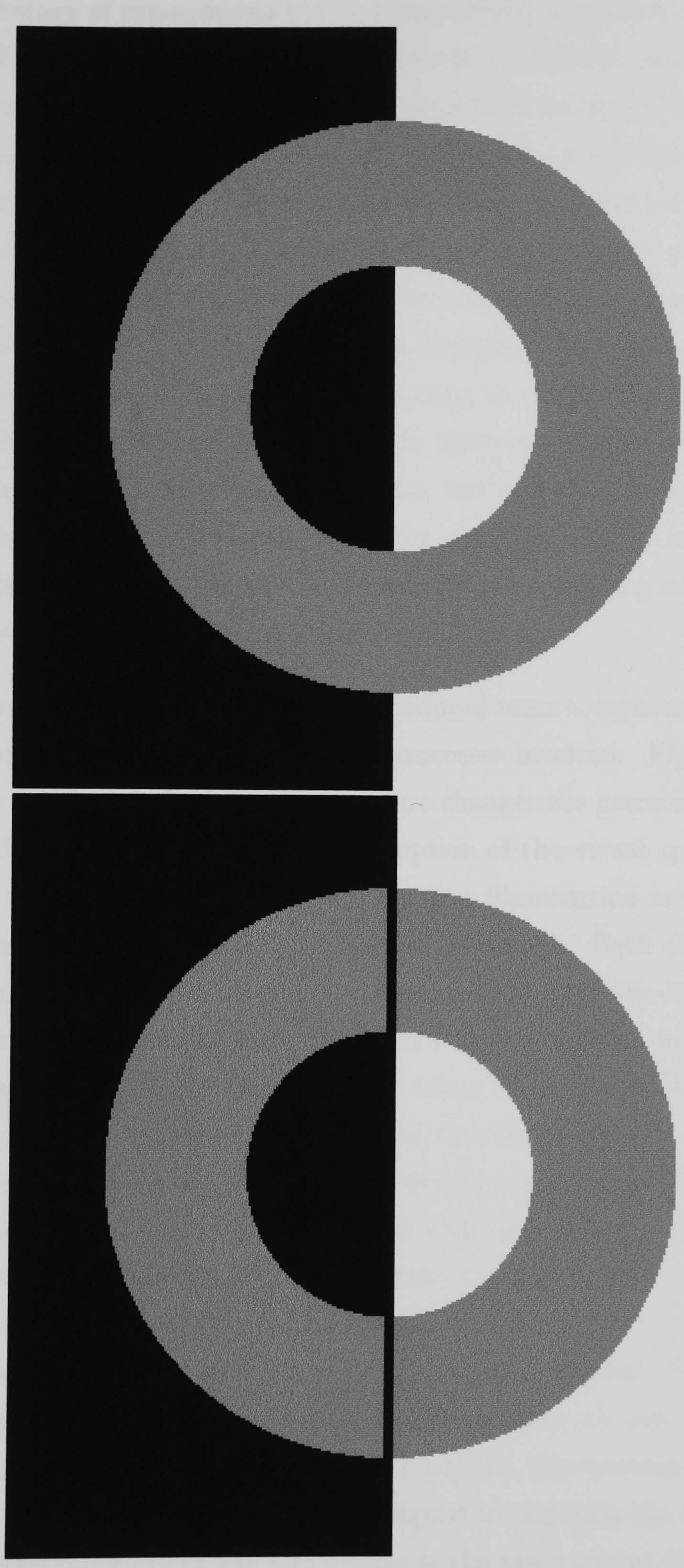

Figure 4.1: The Koffka ring illusion

In the first stage of experimentation an observer was asked to make lightness judgements for two trapezoidal surfaces (monocularly). Once this stage was complete the observer was allowed to use both eyes which showed the actual scene characteristics – one of the trapezoidal surfaces was a trapezoidal paper perpendicular to the observer's line of sight, the second was a square paper that was laid flat and appeared trapezoidal due to perspective. Once the observer was made aware of the spatial arrangements more lightness judgements were made and the perceived lightness of the surfaces changed. The matches for the upright surface (assumed to be in direct illumination) were made darker (corresponding to darker perceived lightness) and the matches for the flat patch increased in lightness. This experiment is striking in the changes that it induces in perception, but some ambiguity exists in the cognitive processes that the observer used. For example: did the observer cognitively change lightness judgements to compensate for the pervading scene features or did the patches markedly change in appearance?

Adelson (Adelson 1993) has recently produced some compelling illusions in which no ambiguity exists as to the cognitive processes involved. Figure 4.2 shows an example in which perceived spatial structure changes the perceived lightness of surfaces. It also highlights a 'built in' assumption of the visual system – in times of ambiguity we appear to prefer to assume that illumination is from above. Consider patches B1 and B2 in the top right of the figure. Both of these patches are printed with the same grey level i.e. reflect exactly the same amount of light, yet B1 appears to have a much lower reflectance than B2. It is usually assumed that this division occurs due to B1 apparently being under direct illumination and B2 being in a region of shadow. The argument that simultaneous contrast factors are causing the illusion can be shown to be false by considering the top left of figure 4.2. Surfaces A1 and A2, again, are printed with exactly the same grey level, in fact all of the surfaces have the same grey levels, the only difference between the top left and top right is that the shapes of some of the patches have been changed so as to produce a different three dimensional scene perception. We can see that the local spatial neighbourhood does indeed make patches A1 and A2 appear slightly different but by no means as different as B1 and B2. The bottom half of figure 4.2 is a pseudo colour version of the top half designed to highlight the construction of the illusion. Everything about the two figures is the same except the grey levels have been re-mapped to random colours.

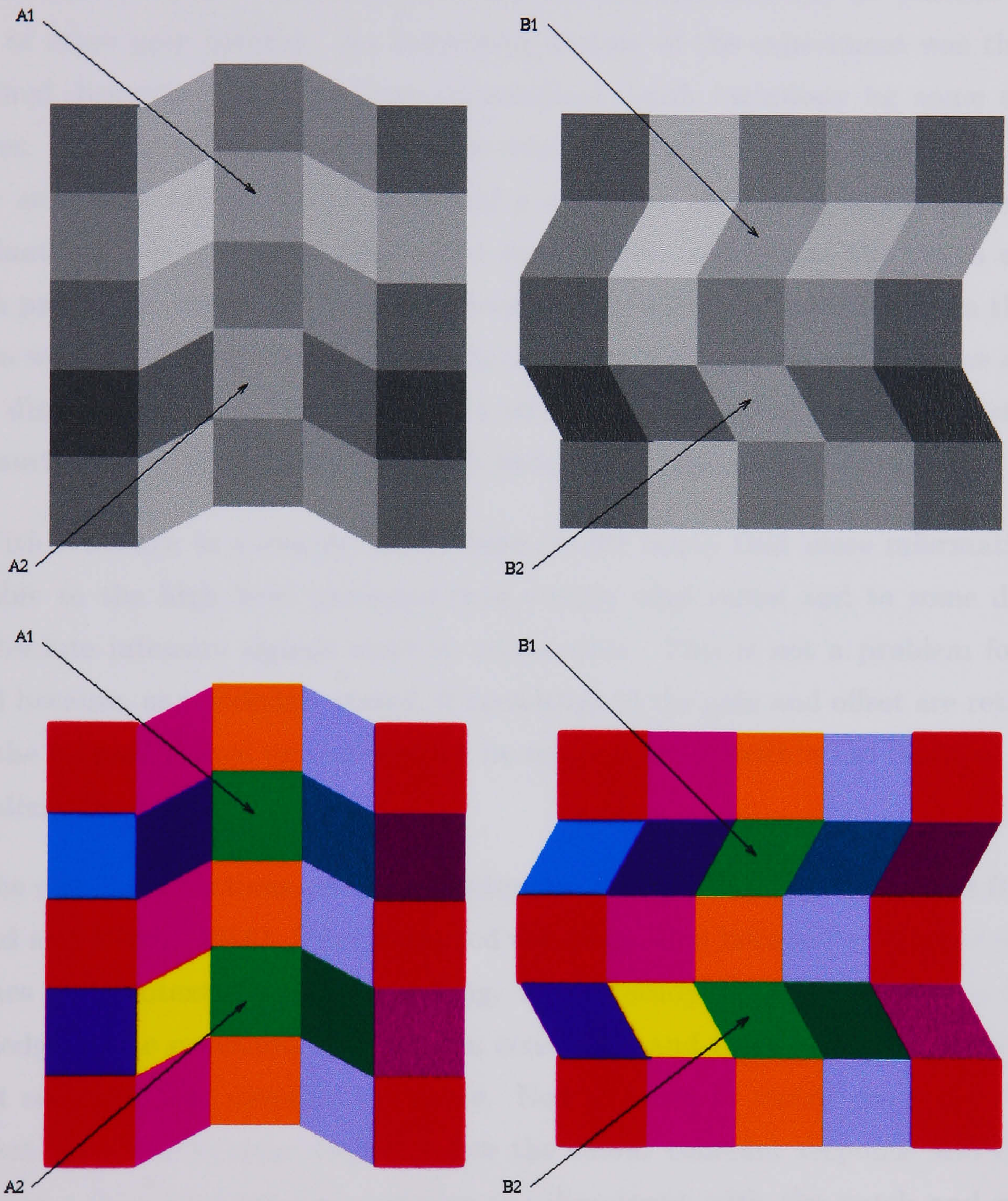

Figure 4.2: Adelson's illusion (pseudo colour version shows grey patch equivalences)

Logvinenko (Logvinenko 1996) has shown similar effects for coloured surfaces. He suggests that the rules that produce such effects are subtly different and the effect is much weaker.

Another perceived illumination effect was shown by Schirillo and Shevell (Schirillo and Shevell 1993) in which they performed lightness matches for patches in an array of other grey patches. An interesting feature of the experiment was the use of retinal disparity to induce three dimensional depth variations for some of the patches. The collection of patches that were perceptually in the foreground were under simulated bright illumination and a collection farther in depth in dimmer illumination. They found a small effect on lightness matches as the depth of the match patch was varied. What was interesting in their study was that when the far papers were brightly illuminated and the near papers dimly illuminated the depth effect disappeared. This may highlight a second assumption of the visual system – close surfaces are usually more intensely illuminated than distant surfaces.

With reference to a simple model these results imply that more information is available to the high level processes than simply edge ratios and to some degree the absolute intensity signals must be recoverable. This is not a problem for the model because, as previously stated, if knowledge of the gain and offset are retained then the original retinal excitations can be reconstructed (within the limits of noise degrading the signal).

The role of instructions given to the observer was examined by Arend and Reeves (Arend and Reeves 1986). They extended the distinction between void space colour matches and contextual colour matching. In this study the observer has access to knowledge of the prevailing illumination conditions and tends to see it as a 'transparent coloured film' covering the scene. Now if asked to match the exact colour of a test they can to some degree ignore the colour constant response and report the exact colour they see i.e. combine the illuminant with the patch and give a colorimetric match. Alternatively, if the observer is asked about the surface colour of the test and the illuminant colour then they can separate the two properties and report on both. This distinction is a major problem in attempting to quantify human colour constancy capabilities due to the fact that the matching procedure tends to produce a colorimetric match.

Arend and Reeves attempted to bridge the gap by emphasising the perceptual difference by varying the instructions given to the observer. The first set of instructions required the observer to adjust the match patch to appear exactly the same, in terms of hue, saturation and luminance, as a test patch (displayed in an array of patches under a test illuminant). The second set of instructions highlighted the fact that the display consisted of two simulations of coloured reflected surfaces under two different illuminants. The observer was required to adjust the match patch to 'appear to be cut from the same piece of paper' as the test patch. The results showed a major improvement in constancy for the paper match condition – in fact some of the highest constancy yet produced in experimentation.

There are, however, several objections to the study. First, the viewing technique was of the simultaneous binocular type (this was intentional) and as such retinal adaptation to both illuminants caused a reduction in constancy for the hue-saturation match condition (for the metric they chose to use). Second, the 'paper match' requires a high degree of conscious thought and may require specific knowledge of illuminant/surface interactions. The three observers in the study had science backgrounds and were familiar with the sort of simulation used and how the colour of the patches were produced. These shortcomings were actually pointed out by the authors themselves. In a similar unpublished study we repeated the experiment using Fine Art students, whose colour mixing knowledge is based on subtractive mixing. They produced terrible constancy, but produced highly predictable results when the subtractive colour model was taken into account. The motivation behind the experiment was both timely and very informative, but unfortunately did not provide us with a procedure to circumvent the 'multiple viewing modes' problem.

Cornelissen and Brenner (Cornelissen and Brenner 1990) examined behavioural changes in observers under instructional variation. An eye tracking system was used to record which patches the observers spent time looking at. In both cases (hue-saturation and paper match) the observers spent a great deal of time simply looking at the test and match patches, but in the paper match case observers did spend more time examining the whole display.

Craven and Foster (Craven and Foster 1992) provide us with an alternative methodology for the solution of the 'colour constancy problem'. The idea is that the visual system can differentiate between changes in surface colour and changes in illumination. The technique they used was to show a collection of coloured surfaces under a test illuminant and then in a second interval show a scene in which the illuminant had been changed or several of the patches had been changed (in a manner that kept the space average for the two conditions constant). They found that observers could reliably differentiate the two cases even in remarkably short time intervals (Foster, Craven and Sale 1992). A second variant of the experiment was used to examine the time dependence of the phenomena. In that experiment, three collections of papers were shown at the same time and the observer was required to define which of the outside two collections was due to an illuminant change.

It is interesting to consider this experiment within the confines of the proposed two stage model – the second stage of which was assumed to respond to the edge contrast within a region and adjust gain and offset to compensate. If we consider the stimulus used in this study then in the illuminant shift case we have a clear change of offset but should not need gain adjustments. In the material shift case we may require a shift in offset and of gain. So this task may produce distinctly different effects in the opponent colour channels – which may be a basis for performing the task.

McCann (McCann 1993) provides another example of small stimulus features that can produce strikingly differing percepts. The stimulus used contains two centre surround displays. The displays have different reflectances – one shifted towards blue and the other towards yellow. The blue patch is then illuminated with a yellow illuminant and the yellow patch with a blue illuminant in such a manner that the two displays appear to match. McCann then added 'colour constancy test patches' (thin bands of reflectance) of white, black, light red, dark red, etc. His results showed that any 'constancy test patch' that provides a new maximum cone excitation within a given class destroys the observed match. He also found that the introduction of a patch that did not provide a new maximum for any cone class did not destroy the match. The results showed that the introduction of a new maximum affected all of the viewed image. The amount of influence was dependent on the extent and proximity of the new patch to an area of interest. In terms of the proposed simple model this phenomenon corresponds to a gain change in the second stage mechanism to compensate for a contrast change that is out of range for the current gain settings.

Specular highlights could provide a method for calculating the illumination and hence to allow for its compensation. The colour of a specular highlight is a mixture of the illuminant colour and the underlying surface colour. Specular highlights are also scene features that are reasonably easy to find within an image. A small problem with the technique is that it can not provide an exclusive method for colour constancy because many natural scenes contain few specular highlights. Hurlbert, Lee and Bülthoff (Hurlbert, Lee and Bülthoff 1989) performed an experiment in which conflicting information from scene highlights and the object body colour was presented to the observer (i.e. the specular highlight suggested an illuminant of one colour and the mean colour of objects within the scene suggested a different illuminant). It was found that the observer, to a large degree, discounted the information from the specular highlight.

Chapter 5

Experimental Results: General Procedures and Control Experiments

5.1 Introduction

This thesis is primarily concerned with examining human colour constancy performance in simple and complex scenes. There are four experimental results chapters covering different experimental stimulus complexities. The chapters are arranged in terms of increasing scene complexity and may examine several stimulus aspects at each detail level. The experiments range from matching patches against uniform backgrounds to matching full colour objects in complex scenes.

Each experiment takes the same basic form in that observers are requested to match a patch or object in a test scene with a match patch or object in a reference scene.

In this chapter, technical aspects common to all the experiments are described along with the basic procedural verification that has to be performed when beginning to experiment with any new paradigm. The results from this chapter also provides a base level of observer performance to which we can compare subsequent experimental results.

More specifically this chapter examines aspects of the viewing techniques and stimulus configuration used, chromatic adaptation, dial match and method of constant stimuli procedures for data collection and colour memory.

In chapter 6, the results of measurements of colour constancy in flat scenes are presented. In these experiments, some previous work (Arend and Reeves 1986; Valberg and Lange-Malecki 1990; Shepherd and Whittle 1995; Wandell 1993) is repeated. The specifics of the experiments are different in terms of exact viewing technique used and stimulus design. This repeated work is useful to place the presented results in context with the other studies.

A two stage model was outlined in the theory section that predicts different match results for centre surround stimulus configurations and matches performed in a colour contrast rich environment. Chapter 6 presents results of matches performed in both conditions and assesses the implications with respect to the simple model.

Chapter 7 covers colour constancy of shaded objects in flat scenes, extending work in the previous section to probe the effects of increasing the information in the scene by adding luminance variation to the test and reference presentations in the form of an object.

The question of 'self constancy' is also addressed in this section. The procedure involves differential illumination of a test object and the background (this differential lighting can occur frequently in real life and does not appear to cause great difficulty for everyday colour constancy). The motivation can be stated as: Is the extra information present in an object representation enough to allow colour constancy to be performed regardless of conflicting information from the surround?

The final results chapter covers colour constancy in real life scenes. This section uses full colour images and allows the observer to alter the colour of an object within such a scene. In this way cross illuminant colour matching in real life scenes can be performed. The motivation behind this section is to try to 'bridge the gap' between the high level of constancy achieved in real life and the relatively low levels produced in flat scene experiments. The number of variables present in the scenes (shading, specular highlights, mutual illumination etc.) are hard to quantify so insights about exactly how colour constancy is performed are unlikely, but an improvement in constancy under these conditions would raise questions that need to be addressed about the use of simplistic scenes in testing colour constancy.

5.2 General procedural aspects

5.2.1 Stimulus display

To simplify explanations the following conventions will be used:

- A test is defined as a patch, object or part of a bitmap image that we are matching to.
- A match is defined as a patch, object or part of a bitmap image whose colour is adjusted under observer control to produce a match. The term will usually be qualified i.e. match patch, match object.
- A background is that part of the stimulus that is used to control adaptation in the observer, it may be uniform, composed of many coloured patches, or consist of the non-test part of a bitmap image.
- The field presented to the left eye is designated as test and is used to show presentations of differing colours under differing illumination conditions.
- The field presented to the right eye is designated as match and contains presentations that are under a fixed illuminant (D_{65} or equivalent). The colour of the match is either adjusted by the observer or is changed by the program between trials in the method of constant stimuli scenario.

Each field contains a simulation of a scene under daylight or artificial illumination (the bitmap images are directly photographed and not simulated). The scenes are constructed from simulations of the interaction of the full visible spectra (interpolated to 1nm from 5nm samples) of Munsell patches with daylight illuminants calculated from the basis functions derived by Judd, MacAdam and Wyszecki (Judd, MacAdam and Wyszecki 1964).

A two minute adaptation period is enforced at the start of each experiment and the adapting background remains constant throughout an experiment[1].

[1]The patches used in the background are spatially shuffled between trials, thus reducing the effects of simultaneous contrast while maintaining the same overall adaptational effects. Control experiments in which the background did not change produced qualitatively the same results, but the possibility still remains that there may be a slight fluctuation in observer adaptational levels.

All stimuli and experimental displays are produced on a Silicon Graphics Crimson workstation with VGX graphics, and displayed on an Eizo T560i 17" Trinitron monitor. The monitor is calibrated using an automatic procedure outlined in chapter 2 of this thesis (for more detail the reader is referred to Bramwell (Bramwell 1993)). The procedure produces approximately 115 linear values per gun. The gradient of these linearised scales is controlled in the calibration process so as to produce good additivity between the guns. A second program is used at regular intervals that checks for drift and also additivity. The program randomly chooses a legal set of RGB values and takes 10 readings for each using the Minolta CS100 Chroma Meter. The readings are averaged and compared to the values calculated from the RGB to XYZ conversion matrix produced in the calibration. If any error is greater than the variance in the averaged readings then the system is re-calibrated.

Spatial inhomogeneity has been found to be similar to that reported by Brainard (Brainard 1989) with a consistent luminance fall-off from centre to edges of the screen. To counter this all stimuli are presented in a 'balanced' manner. Relative to the centre of the screen they share similar positions and hence experience similar spatial inhomogeneity. Gun mis-alignment is not significant across the screen. The luminances are kept low to avoid any non linearity that can be introduced from power supply loading. The low luminances also avoid slew rate[2] problems that may occur in high contrast and high temporal and spatial frequency signals (i.e. there is a finite time required to heat the cathode to produce an increase in electron current and similarly for it to cool; large changes in very short time intervals may not produce the required results. For a 76Hz display at 1280x1024 resolution each pixel is produced in less than 10 nanoseconds).

For reasons of consistency the stimuli are shown in the same configuration regardless of viewing conditions. For example, the experiment that compares binocular to haploscopic viewing has exactly the same stimulus on screen but employs differing viewing conditions.

All displays have similar general properties. Part a) of Figure 5.1 shows a schematic representation of the full display. The display is split exactly in half ver-

[2]'Real' amplifiers can require significant amounts of time for their output response to reach the level required by the current input. This loss in linearity is usually termed the slew rate and is a function of the magnitude of the change required at the output and the time available for the change.

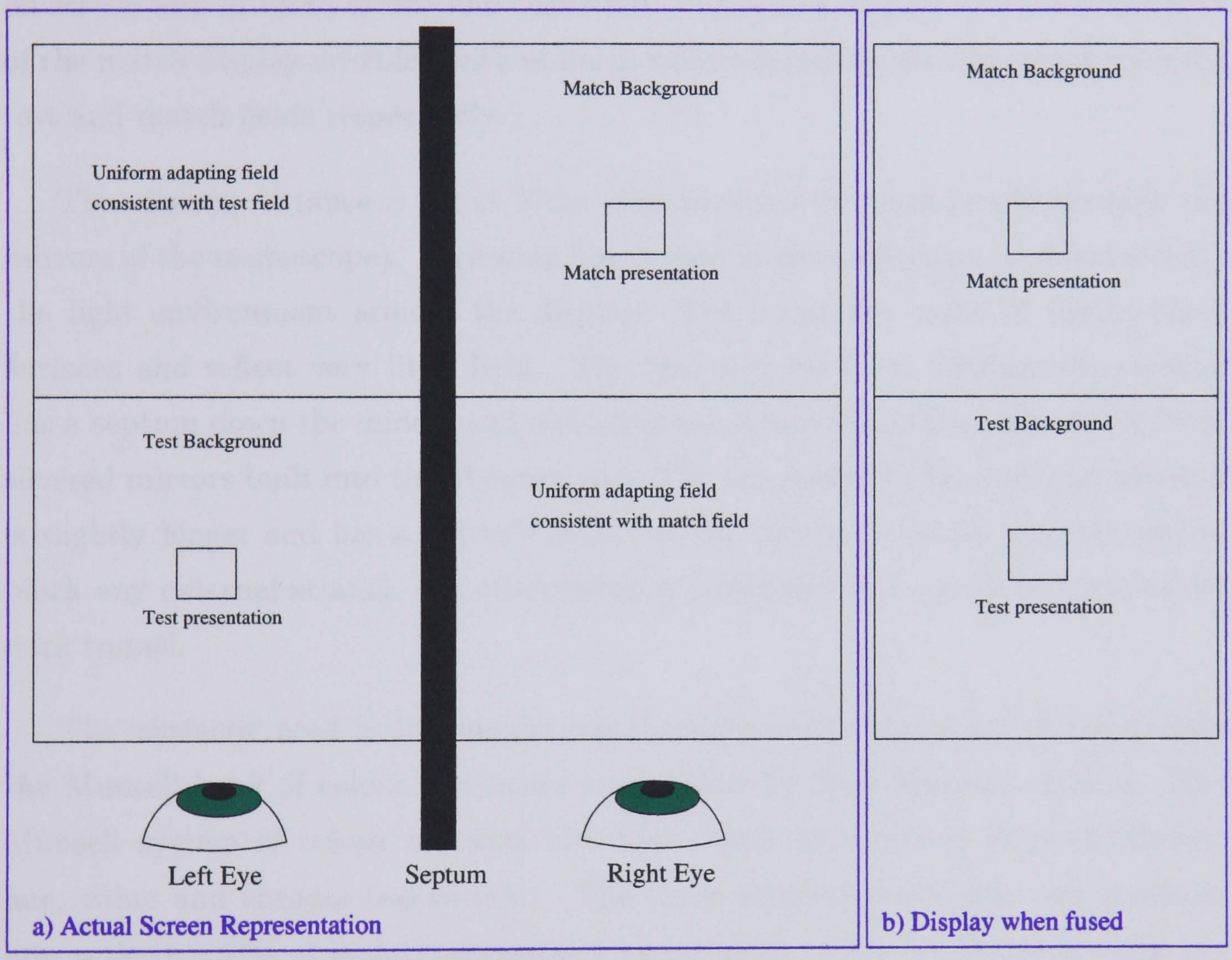

Figure 5.1: A schematic representation of the stimulus display

tically producing two fields of $16^o \times 25^o$, at the viewing distance of 57 cm. The test presentation is shown in the lower left half of the screen and the match presentation shown in the upper right part of the screen. When viewed in the fused haploscopic case the two presentations appear in the same vertical plane as shown in Figure 5.1 part b). Presentations are centred with respect to their particular field, i.e. test presentations are centred in the lower left quadrant of the display and match presentations centred in the upper right. To clarify this point further, examine figures 5.9, 6.8, 6.15, 7.6 and 8.22 which show several stimulus presentations as they appear on screen and in fused form. The upper part of the test display and the lower part of the match display are filled with uniform fields adaptationally consistent with the test and match fields respectively.

The viewing distance is set at 57cm (this includes the path length through the mirrors of the stereoscope). A viewing box is used in all conditions to control strictly the light environment around the display. The boxes are made of matte black surfaces and reflect very little light. The box used for fused haploscopic viewing has a septum down the middle and a stereoscope created from four adjustable front silvered mirrors built into the observer end. The box designed for binocular viewing is slightly longer and has a "diver's mask" at the observer end for comfort and to block any external stimuli. All observation is performed in a specially constructed dark tunnel.

The 'surfaces' used in the simulations throughout this thesis are all taken from the Munsell book of colour and hence are labelled by their Munsell notation. The Munsell system of colour notation identifies colour in terms of three attributes: hue, value and chroma (saturation). The three attributes of colour are arranged into orderly scales of equal (perceptual) visual steps (when the Munsell chips are viewed under D_{65}). A Munsell colour is specified in the form: hue value/chroma. For example, the specification of a green reflecting patch with a 'value' of 5 and a chroma of 4 would be G 5/4. The Munsell notation specifies 10 hues; red (R), yellow-red (YR), yellow (Y), green-yellow (GY), green (G), blue-green (BG), blue (B), purple-blue (PB), purple (P), red-purple (RP). The final hue designator is N which represents a neutral (achromatic) reflecting surface. The simulations used in generating the stimuli for the experiments presented in this thesis use the measured reflectance spectra of Munsell chips from the matte reflecting surface collection.

In all the simulations, Munsell chips of value 5/ are used except in the backgrounds, where values of 4/ and 5/ are allowed. As mentioned previously, screen RGB values are obtained from a simulation of full Munsell spectra with full illuminant spectra. A conversion factor was selected that converts the values from the simulation to those displayable on screen (most daylight illuminants produce light intensity in the thousands of candela per metre squared range whereas standard monitors can produce a maximum only of around 120 cd/m^2). This factor reduces the intensity of the illuminant spectrum by a multiplicative constant. (This is a widely used technique and is expected not to affect results significantly as the human visual system is driven chiefly by scene contrast and to a large degree can discount absolute intensity levels). The conversion factors chosen produce approximately 6 cd/m^2 on screen for 5/ patches (this value of course varies slightly depending on the illuminant and patch). Except for those experiments in which luminance variation is studied explicitly, the stimuli are constructed to be of similar luminance. The test patches are selected to have the same Munsell value, but the exact luminance required to represent the patch is dependent on the interaction of its reflectance spectrum with the illuminant spectrum and as such exact isoluminance cannot be ensured. Slight differences with the background may, in fact, be desirable as luminance edges may assist in the colour discrimination process (Boynton, Hayhoe and MacLeod 1977). Unfortunately the nature of the discretisation process means that the scale of luminances displayable in certain parts of xyY space are larger than others. These factors can mean a loss in veridicality of the simulation, but the variation will always be less than 0.15 cd/m^2(max total error 2.5 % : max red 0.04 cd/m^2(0.6 %) , max green 0.09 cd/m^2(1.5 0.02 cd/m^2(0.33 %)).

In any one experimental run, only one illumination condition is tested (i.e. one test illuminant and one reference illuminant, not necessarily the same), so as to keep observer adaptation as constant as possible. Observers naive to the purposes of the experiment are never told of the illumination difference or allowed to see the stimuli except through the stereoscope.

5.2.2 Dial matches

At the start of an experiment a set of instructions is displayed to the observer on the monitor screen. Figure 5.2 shows the bitmap presented to the observers. The

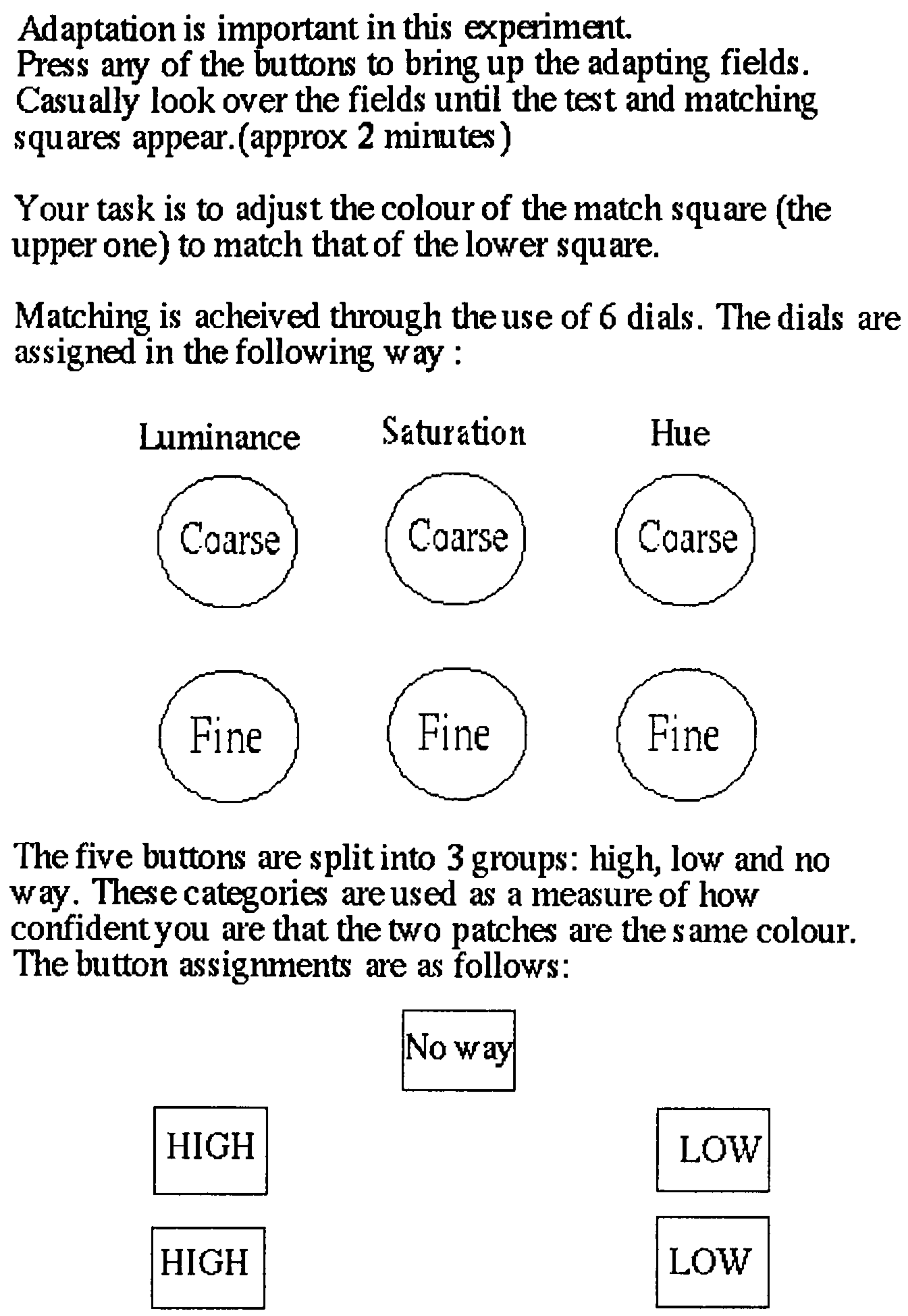

Figure 5.2: The instructions displayed on screen prior to dial matching

observer is required to press any button to signal that he/she has read and understood the instructions. The two minute adaptation period then begins. During this time adapting backgrounds are displayed that are consistent with the experiment to follow. The test presentation then appears alongside the match presentation. (The two fields are super-imposed by the stereoscope to fuse into a single image in which the test presentation appears below the match presentation). The observer is allowed free viewing and unrestricted time to make the match. When a match is completed the observer presses a button and the next trial is shown. The mapping of the dials relative to the colour space is slightly changed between trials to avoid the possibility of an observer memorising dial positions.(Figure 5.6 shows a single trial in schematic form.)

A custom built unit is used for the dial matching, a photograph of which is

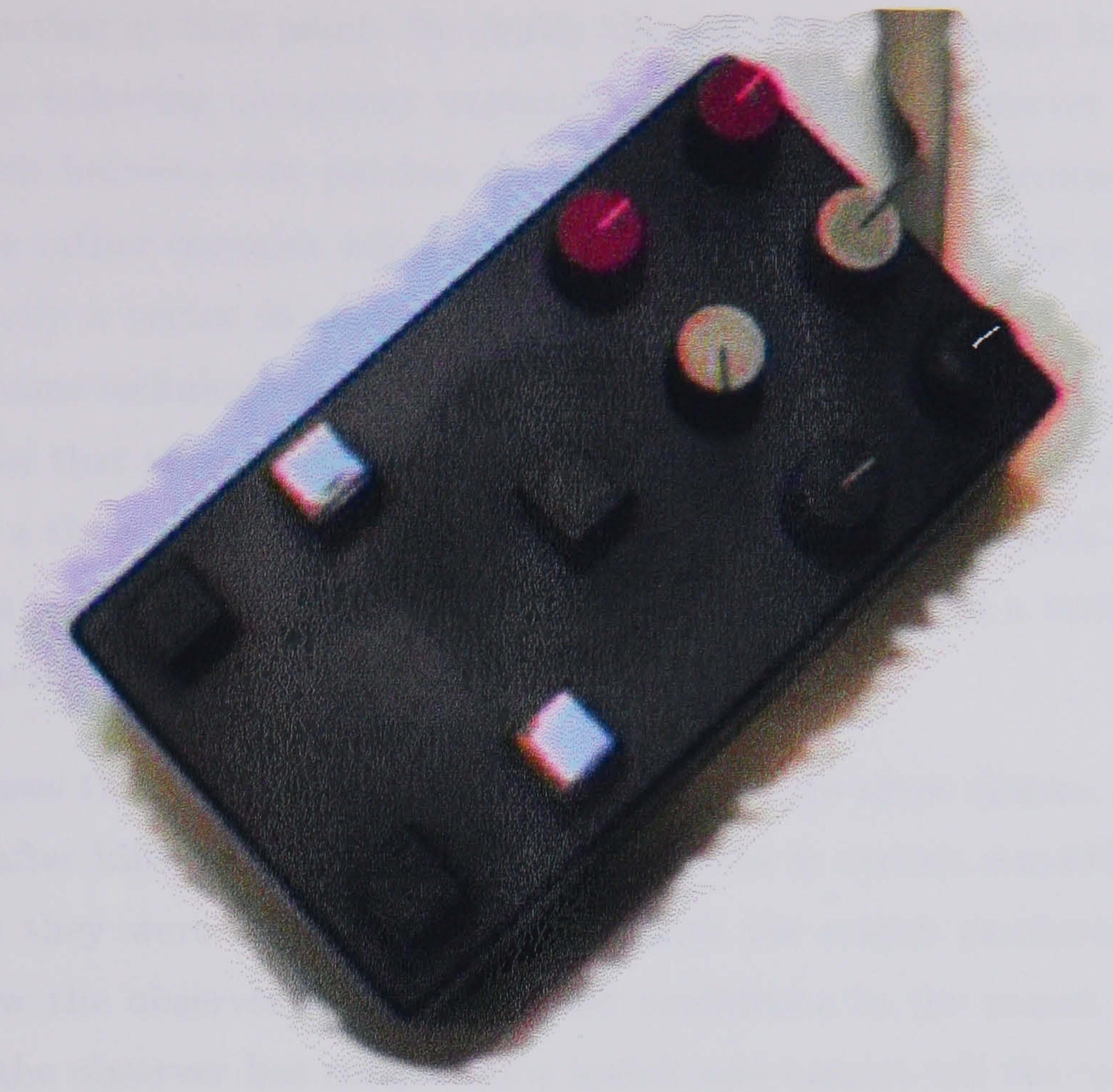

Figure 5.3: The box used for dial matching

shown in Figure 5.3. In experiments designated as 'dial matches' input is obtained from the observer in one of two ways.

In the first scenario, the six dials allow coarse and fine adjustment of three parameters: relative luminance, saturation and hue angle. The resolution available from the unit exceeds the displayable colour resolution. The term 'relative luminance' is used because in all dial matches the mathematically 'correct' (in terms of the perfect colour constant match of the test presentation) luminance is initially displayed and the dials allow offsetting of the luminance relative to this value (this value may be calculated by the program or specified by the experimenter). In our experience with this technique the luminance dial is rarely used, and if it is, only small adjustments are made.

In the second scenario only two or four of the dials are used. Again the dials are split into coarse and fine scales. What the dials adjust is experiment specific and will be described separately for each experiment. In overview, the dials adjust one or two parameters that traverse a solution space specified by the experimenter. This form of adjustment is mainly used in the real scene experiments and will be discussed further at that point. To clarify the use of the technique for the moment consider the following imaginary experiment. We wish an observer to perform a colour match between two patches on the same uniform background. Instead of allowing the rather complex adjustment in terms of luminance, hue and saturation we can specify a vector in colour space along which we know the match to reside (much the same technique used by MacAdam). We can assign a parametric equation to this vector that allows it to be traversed by variation of a single variable. Thus we can reduce a three dimensional match to a single dimensional match by restricting the solution space. The experimental code supports the use of a vector or a plane by specification of its end points.

In all cases the five buttons of the unit are split into three classes. This division was made after initial experimentation showed that in certain conditions observers stated that they were not entirely satisfied with the match produced. The three classes allow the observer to signal his/her confidence in the match. 'HIGH' signals when the observer has completed a match and cannot tell the test and match presentations apart, 'LOW' when the observer can just tell the presentations to be different but cannot get a better match, and 'NO WAY' when a satisfactory match cannot be obtained. This may also assist in raising observer morale, in that they are allowed the concept that a suitable match may be extremely difficult or even not possible.

5.2.3 The method of constant stimuli

This technique is a two dimensional extension of the one dimensional procedure often used in psychophysics to produce a psychometric function. The technique is described in detail in Bramwell and Hurlbert (Bramwell and Hurlbert 1996). The procedure involves the building of a statistical representation of the probability of an observer giving a particular response to the given task. In the one dimensional scenario several values of the parameter under test are chosen, say, contrast of a grating. The observer is presented with each of the values in a random order and asked to make a response based on a criterion, for example: Is there a grating present in the display or is the display field simply uniform? The response is recorded and the procedure repeated a number of times. This results in a plot of probability versus contrast. A function can be fitted to these data points that allows us to approximate

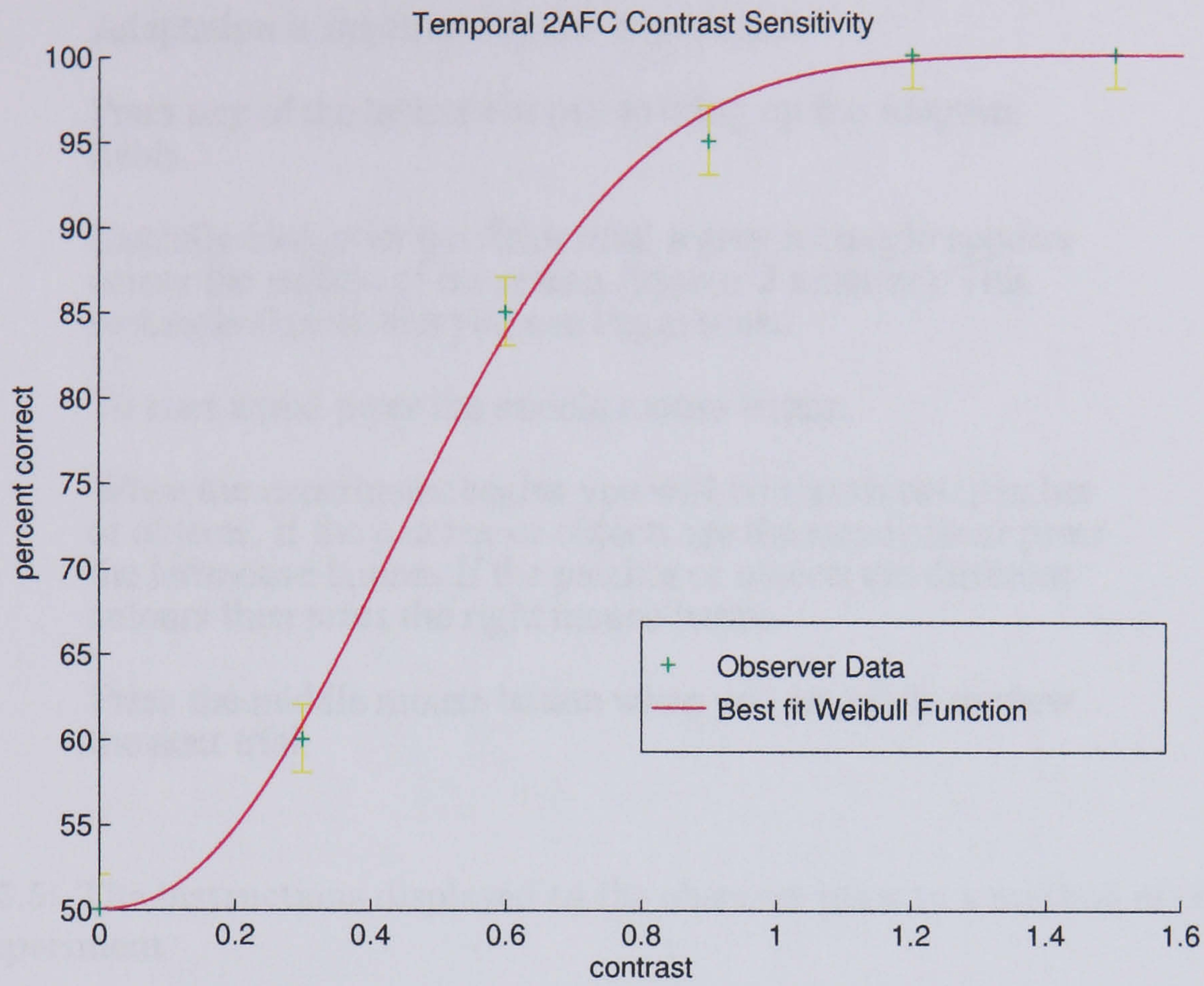

Figure 5.4: Sample results for a contrast sensitivity experiment

observer response at other stimulus levels. Figure 5.4 shows an example of the possible results of such an experiment and the fitted function (in this case a Weibull function).

The procedure has been extended into two dimensions for the purposes of our experimentation. Several 'candidate' patches are chosen in a region of interest (this procedure will be outlined in detail later) and the experiment runs in the same way as the one dimensional case i.e. the observer is presented with a test and match patch and asked to respond whether the patches are the same or different in colour. The results produce a contour map of performance. When analytic functions are fitted, features of the observer response can be measured and analysed, such as the centre location, dimensions and orientation of the discrimination ellipse.

The time course of events for the method of constant stimuli case is as follows: first, an image outlining instructions is presented to the observer (Figure 5.5). The observer presses a mouse button to initiate a two minute adaptation period in which backgrounds consistent with the rest of the experiment are shown. After this time a grey bar appears across the middle of the screen to signal that the observer may

Adaptation is important in this experiment.

Press any of the mouse buttons to bring up the adapting fields.

Casually look over the fields until a grey rectangle appears across the middle of the screen.(approx 2 minutes). This rectangle signals that you can begin trials.

To start a trial press the middle mouse button.

When the experiment begins you will be shown two patches or objects. If the patches or objects are the same colour press the leftmouse button. If the patches or objects are different colours then press the right mouse button.

Press the middle mouse button when you are ready to view the next trial.

Figure 5.5: The instructions displayed to the observer prior to a method of constants style experiment

start trials when ready. The mouse is used as the input device: the middle button initiates a trial, the left button signals a 'yes' response, and the right button signals a 'no' response. The program will only accept buttons at certain times; for example, the observer cannot request a new trial without giving a response for the previous trial. The exact criterion for a 'yes' response is task specific and will be discussed with each experiment.

Several timing factors of each trial can be controlled. Figure 5.6 shows the time course of a single trial with respect to the available parameters. In almost all cases the trial timings are as shown in parentheses. Any exceptions will be pointed out on a per experiment basis. All method of constant stimuli procedures in this thesis are performed using a one alternative binary choice procedure (1ABC). The motivation behind this choice is discussed in Bramwell and Hurlbert (Bramwell and Hurlbert 1996). The main reasons for not using 2AFC are: there is a possibility that the observer could 'learn' the appearance of a test patch during an experimental run in which some of the alternatives are obviously incorrect. The stimulus that should be shown in the 'correct' interval is also an issue: since we know human colour constancy is not perfect, using the perfect colour constancy point in the correct interval will, at best, provide a standard by which to judge subsequent trials.

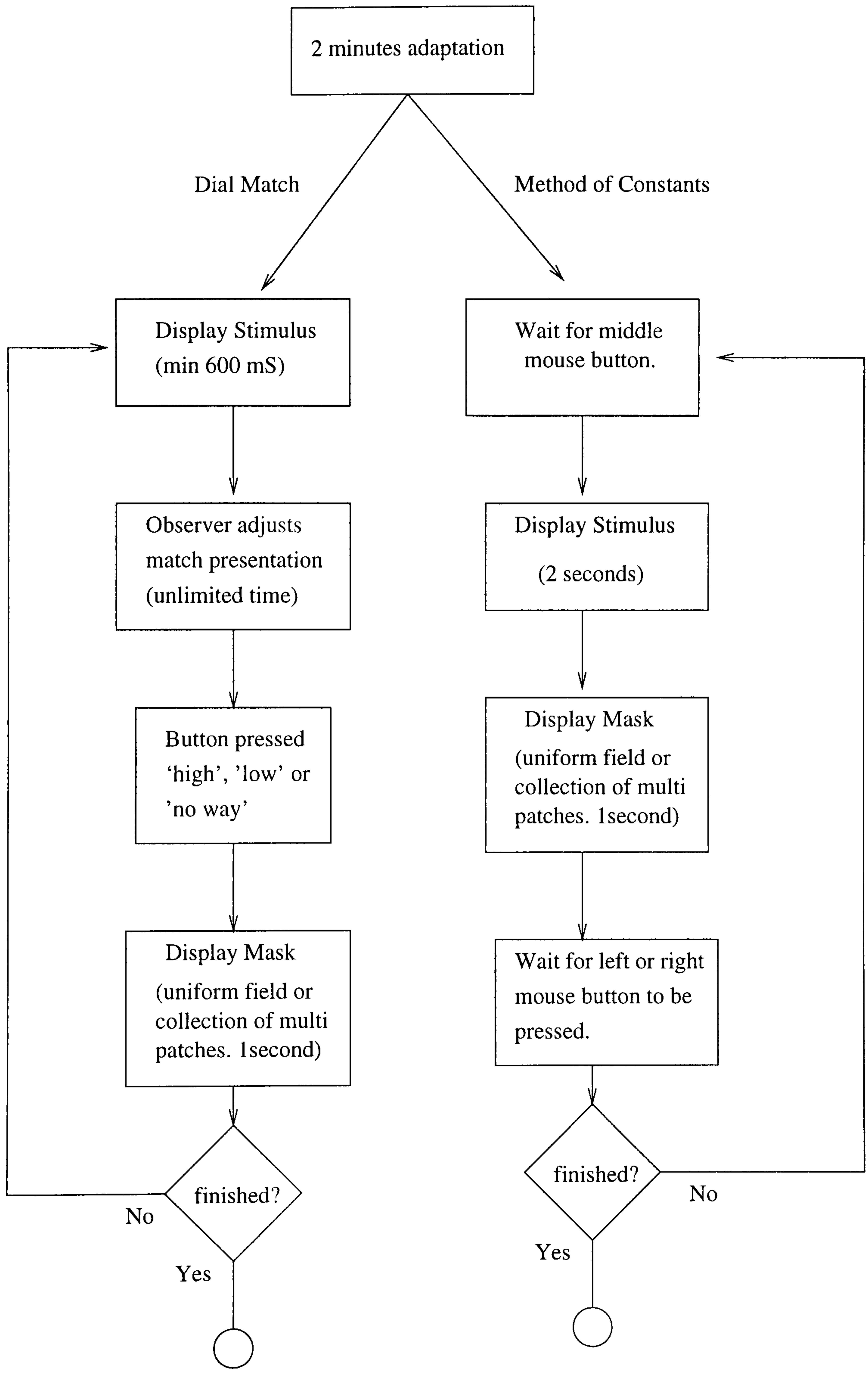

Figure 5.6: Experimental flow and timing parameters

Observer that produced the presented data

Number of trials actually performed

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.368	0.371	0.00746	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.421	0.428	0.00787	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	—	—	—	—	0.314	0.43	0.00893	100_3
B 5/4	D_{40}	D_{65}	—	—	—	—	0.271	0.344	0.00476	100_3

Data presented to 3 s.f.

No data available or field not applicable.

Figure 5.7: Tabulation conventions.

5.3 Tabulation conventions

The form of tabulated results have been held as consistent as possible. Figure 5.7 shows an example table and highlights some of the conventions used.

5.4 Plotting conventions

The results are all presented in one of two forms. The first form shows the CIE xy plane with points of all luminances printed on the same spot (i.e. viewing CIE xyY space from directly above). For reasons of consistency symbols have been chosen to represent certain features which will be used in all the CIE xy space plots. Figure 5.8 shows the chosen symbols. Colour is used to illustrate differing data sets on the same plot. Where a number is placed next to a hue–saturation point the value gives the daylight temperature in hundreds of degrees Kelvin. Munsell patch designations may also be given.

Symbol Conventions

⊕	Chromaticity of test presentation. Usually termed the hue–saturation point (HS point).
×	Chromaticity the test presentation would be if displayed under the reference illuminant. (CC point).
+	Chromaticity of mean. May be mean of match point or of object. Colour may vary to allow association with a data set.
o	HIGH confidence match. (Colour may vary).
✱	LOW confidence match. (Colour may vary).
——	Monitor gamut available at 6 cd.
······	Parametric match space.

Figure 5.8: Symbol conventions for CIE xy space plots.

The second form of plot shows cone ratios for the three cone classes along with a vector ratio that combines the three cone values onto a single plot. If the matches produced conform to a cone ratio explanation of the match points then all points plotted should appear along the line $x = y$. That is, if matching is cone ratio conformant:

$$\frac{L_{test}}{L_{testbackground}} = \frac{L_{match}}{L_{matchbackground}} \tag{5.1}$$

$$\frac{M_{test}}{M_{testbackground}} = \frac{M_{match}}{M_{matchbackground}} \tag{5.2}$$

$$\frac{S_{test}}{S_{testbackground}} = \frac{S_{match}}{S_{matchbackground}} \tag{5.3}$$

The vector ratio is calculated in the following manner:

Let L_t = magnitude of the test ratio vector (LHS of equation 5.1) and L_m = magnitude of the match ratio vector for the L cone class (RHS of equation 5.1 and similarly for equations 5.2 and 5.3. Then

$$V_t = \sqrt{L_t^2 + M_t^2 + S_t^2} \tag{5.4}$$

$$V_m = \sqrt{L_m^2 + M_m^2 + S_m^2} \tag{5.5}$$

The intention behind the vector representation is to highlight cases in which variation in a single cone class is co-dependent upon another cone class. For example, if one wishes to maintain isoluminance then the L and M cone values become highly dependent which may force deviation from the $x = y$ line. If this variation is correlated with a similar, compensatory variation in another cone then the vector representation will show very little deviation from the $x = y$ line. The plot is equivalent to comparing the lengths of vectors representing the test and match presentations in scaled versions of LMS space.

5.5 Control experiments

5.5.1 Viewing technique: binocular versus fused haploscopic

Aims:

This study tests the role of monocular adaptation. The observer is asked to match the same patches in the same surrounds in the same stimulus display for three different viewing conditions. In the binocular case the observer can freely view the whole screen with both eyes. In the fused haploscopic case the display is split into two halves, each eye viewing a different, simultaneous stimulus presentation which is fused into one by the stereoscope. In the fused binocular case, viewing is the same as the fused haploscopic case except the septum is removed to allow light from each side to mix.

Stimulus, set-up and methods:

Figure 5.9 shows the stimulus configuration as displayed on the screen. The stimulus consists of two separate simulations. The test field is a simulation of a Munsell test patch (N 5/ , Y 5/4, G 5/4 or B 5/4) against a Munsell N 4/ background under a test illuminant (D_{40}) and the right field has a simulated background of N 4/ under D_{65} against which the match patch is presented. The patches are squares subtending $3^o \times 3^o$. The patches are displayed in a 'close' form which places them almost abutting vertically when the images are fused. This display is preferred so as to remove any possible distortion of results that may occur due to the observer saccading from test to match presentation (which is examined experimentally in

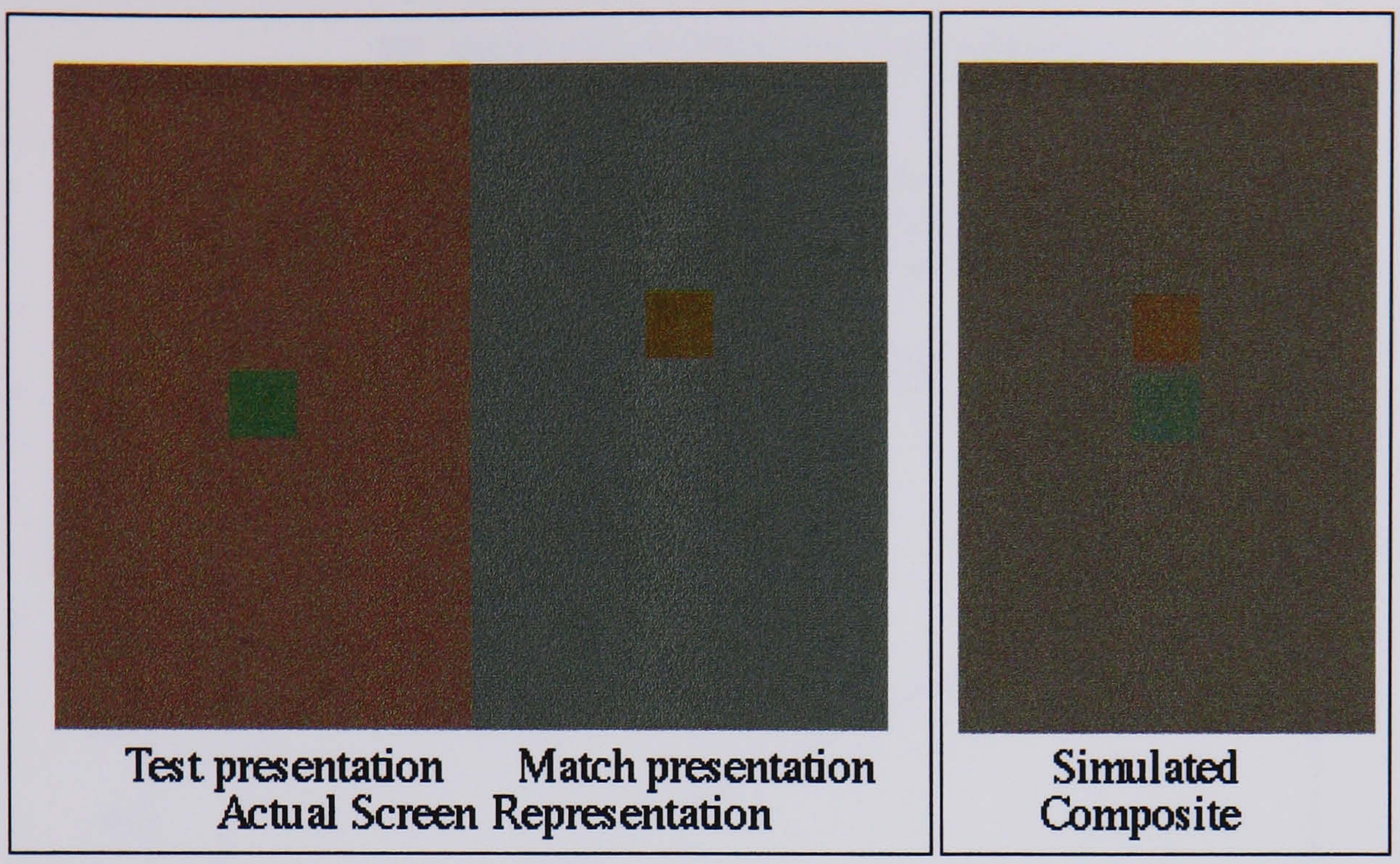

Figure 5.9: Stimulus example for the viewing technique experiment

section 5.5.6). The observer adjusts luminance, hue and saturation of the match patch so as to match the test patch. The observer is instructed to make the patches appear the same colour regardless of surround, i.e. to perform a hue–saturation match.

Results:

Two observers performed matches. Results for DB (male, experienced in colour matching, knowledgeable as to the purpose of the experiment and of the stimulus construction) are shown in Figure 5.10 and for TE (female, experienced in colour matching, naive as to the purposes of the experiment) in Figure 5.11.

Analysis:

Table 5.1 shows the results of some simple statistical analyses of the non-fused binocular match results. Tables 5.2 and 5.3 show the same analyses for the fused binocular and haploscopic conditions respectively.

The degree of colour constancy can be seen qualitatively by assessing the distance of the matches from the colour constancy point (marked 'x'). The binocular case

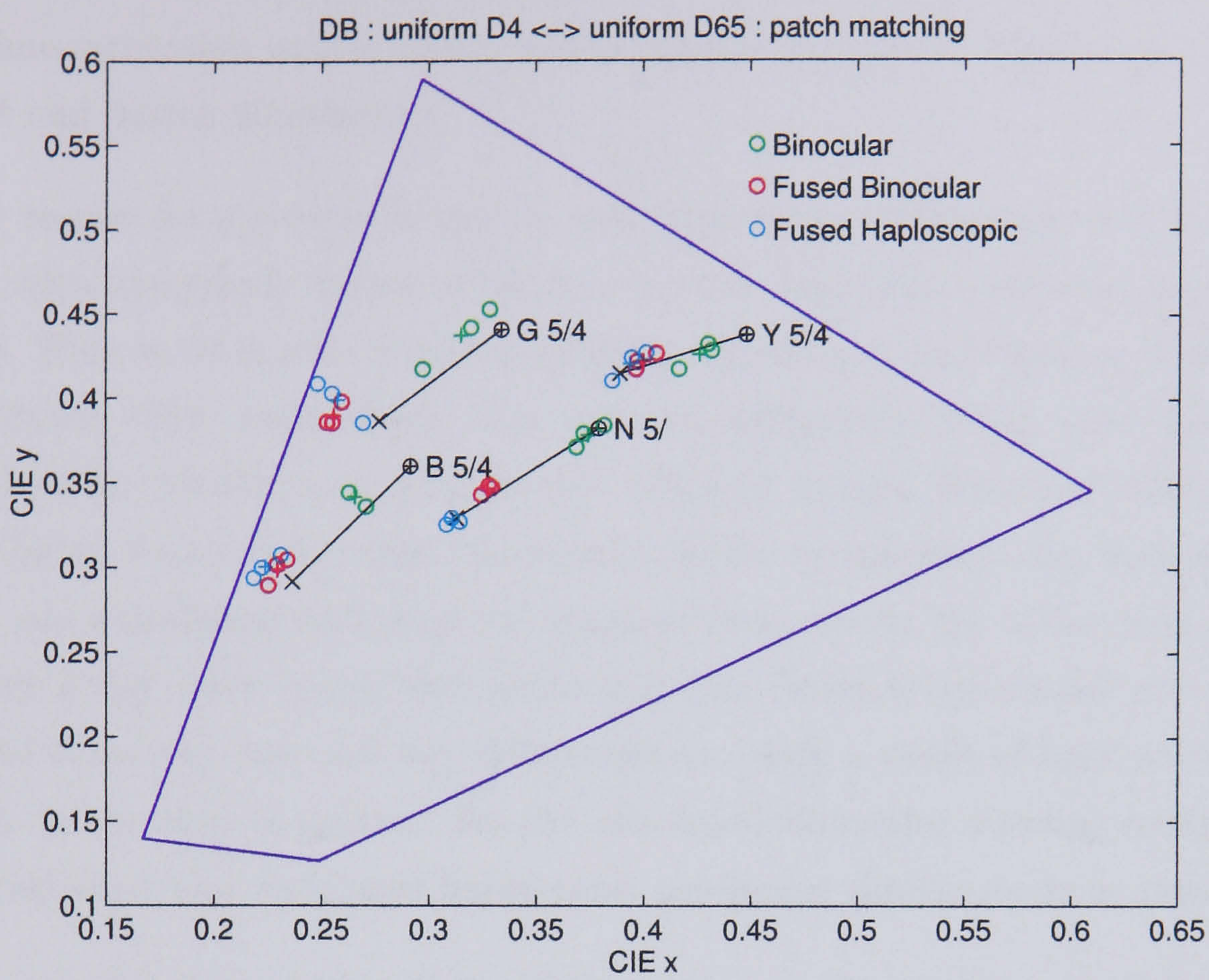

Figure 5.10: Comparison of viewing conditions for observer DB.

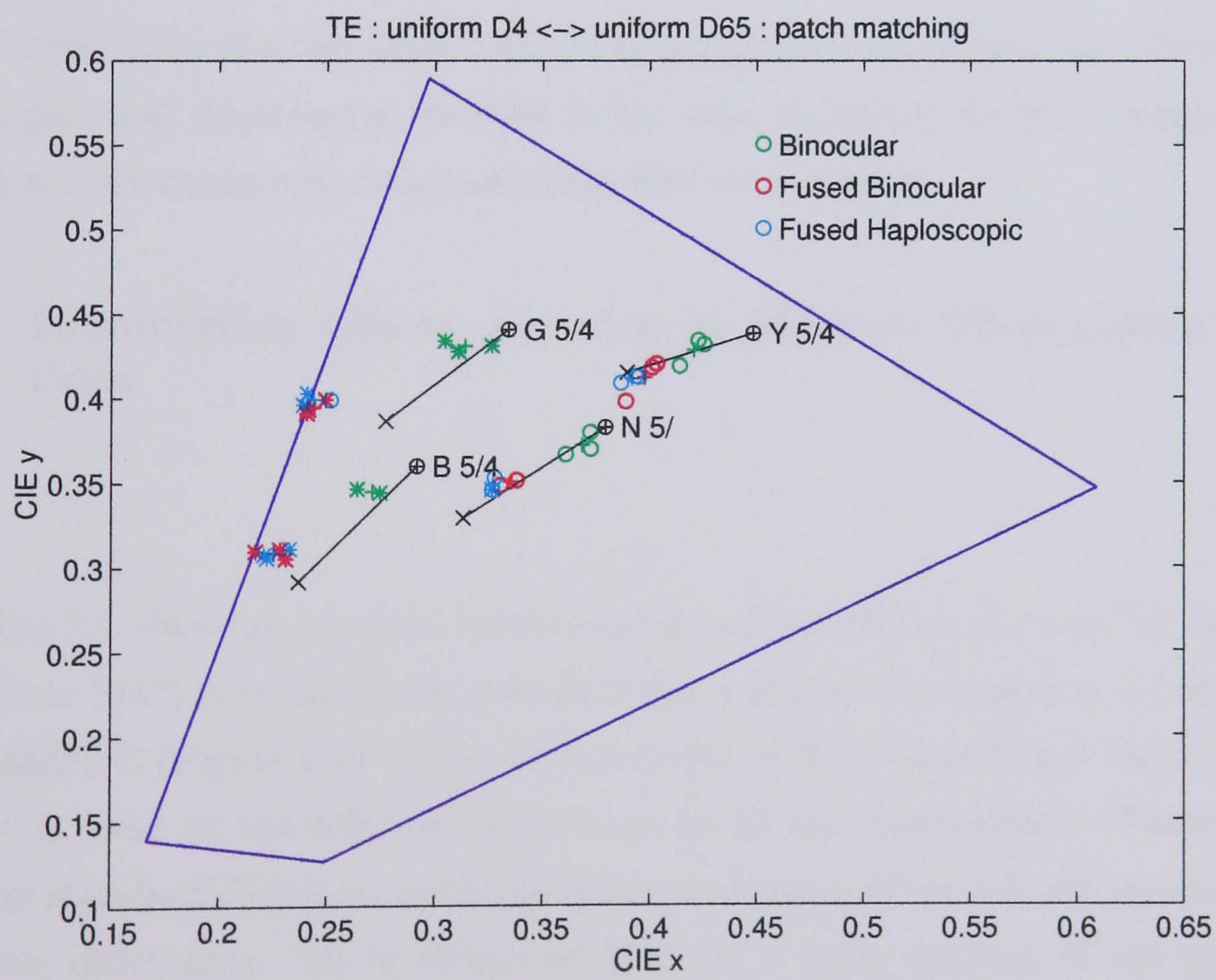

Figure 5.11: Comparison of viewing conditions for observer TE.

produces the worst constancy. Interestingly the matches do not correspond to a direct hue-saturation match either, which suggests a form of hybrid adaptation to the test and match illuminants.

The results for the fused binocular and haploscopic viewing conditions are very similar with marginally better constancy for the haploscopic case for some of the patches. This result is not entirely surprising. Cornelissen and Brenner (Cornelissen and Brenner 1989) have shown that matches performed under strict fixation as opposed to free viewing can produce very different results. The fused nature of the stimuli linked with the fact that the observer looks up and down (i.e. each eye looks only at one simulation and does not traverse from one to the other as is required in the binocular case) means that monocular adaptation is reasonably controlled in the fused binocular case and any differences are likely a result of light scatter. The variance in matches is greatest for the non-fused binocular viewing configuration with fused binocular and fused haploscopic producing similar variance results.

The significant deviations from colour constancy obtained for the G 5/4 and B 5/4 patches are believed to be linked to the centre surround stimulus configuration and are examined in detail in chapter 6.

The results for the two observers are in qualitative agreement for all conditions and suggest that haploscopic viewing is the best technique to use to probe colour constancy performance in simultaneously presented scenes.

5.5.2 Illumination effects: matches in different illumination conditions

Aims:

Figure 2.4 shows some of the colour discrimination ellipses obtained by MacAdam (MacAdam 1942) from matching two abutting coloured fields against a background of illuminant C (Figure 5.12 shows the spectrum of illuminant C and D_{65}, the latter of which is used as the reference illuminant in all the simulations). These ellipses show one standard deviation in the obtained matches and do not correspond to just noticeable differences. (It is estimated that for a large portion of the space one just noticeable difference corresponds to approximately three times the MacAdam standard deviation ellipse) . It can be clearly seen that discrimination abilities vary

DB	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.374	0.378	0.00698	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.426	0.426	0.00881	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	0.315	0.437	0.0179	100_3	—	—	—	—
B 5/4	D_{40}	D_{65}	0.266	0.341	0.00506	100_3	—	—	—	—

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.368	0.371	0.00746	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.421	0.428	0.00787	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	—	—	—	—	0.314	0.43	0.00893	100_3
B 5/4	D_{40}	D_{65}	—	—	—	—	0.271	0.344	0.00476	100_3

Table 5.1: Observer DB and TE match results for non fused binocular viewing

DB	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.328	0.346	0.003	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.4	0.422	0.00544	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	0.256	0.389	0.00606	100_3	—	—	—	—
B 5/4	D_{40}	D_{65}	0.23	0.298	0.00692	100_3	—	—	—	—

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.336	0.35	0.00357	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.398	0.412	0.0111	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	—	—	—	—	0.244	0.394	0.0049	100_3
B 5/4	D_{40}	D_{65}	—	—	—	—	0.225	0.308	0.00625	100_3

Table 5.2: Observer DB and TE match results for fused binocular viewing

DB	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.311	0.327	0.00292	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.394	0.42	0.00895	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	0.258	0.399	0.0119	100_3	—	—	—	—
B 5/4	D_{40}	D_{65}	0.224	0.3	0.00683	100_3	—	—	—	—

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.327	0.349	0.00362	66.67_2	0.326	0.345	—	33.33_1
Y 5/4	D_{40}	D_{65}	0.392	0.411	0.0038	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	0.252	0.398	—	33.33_1	0.24	0.399	0.00345	66.67_2
B 5/4	D_{40}	D_{65}	—	—	—	—	0.225	0.308	0.0055	100_3

Table 5.3: Observer DB and TE match results for haploscopic viewing

with the colour tested. The illumination acting on the patch changes the colour and hence the discrimination abilities of the observer. It is not known exactly where the information reduction that causes the ellipses occurs. Two main possibilities concern us with respect to haploscopic matching. Firstly, the restriction could occur at the retina (monocularly). Alternatively, the transformation could be cortical and occur after image fusion (binocularly) and possibly after colour constancy mechanisms.

Which of these two possibilities occurs is important to us in assessing the errors in haploscopic matching. Consider Figure 5.13 which shows a worst case prediction for an experiment in which an observer performs a match between two patches against differing backgrounds producing different states of monocular adaptation. If the discrimination ellipses are retinal in origin then we can see that the two monocular presentations will reside in different regions of colour space and hence differ in size and possibly orientation. The combination of match errors for each illumination condition could result in larger overall match errors. In the reduced case of the two backgrounds being exactly the same the two ellipses would be the same and combine to produce the same discrimination ellipse. Hence, haploscopic matching as a tool for examining colour constancy could produce errors that are related to

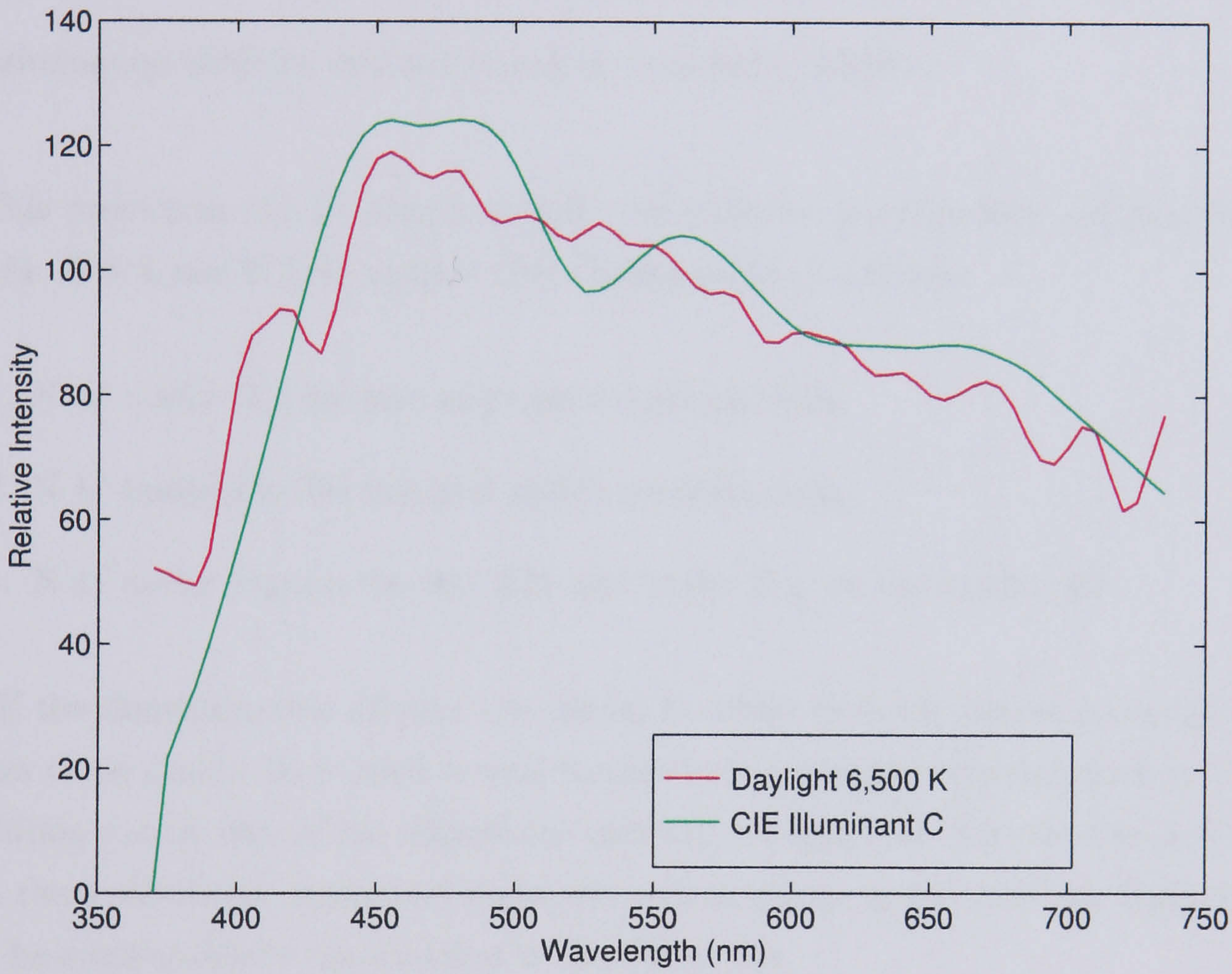

Figure 5.12: The energy spectrum of illuminant C

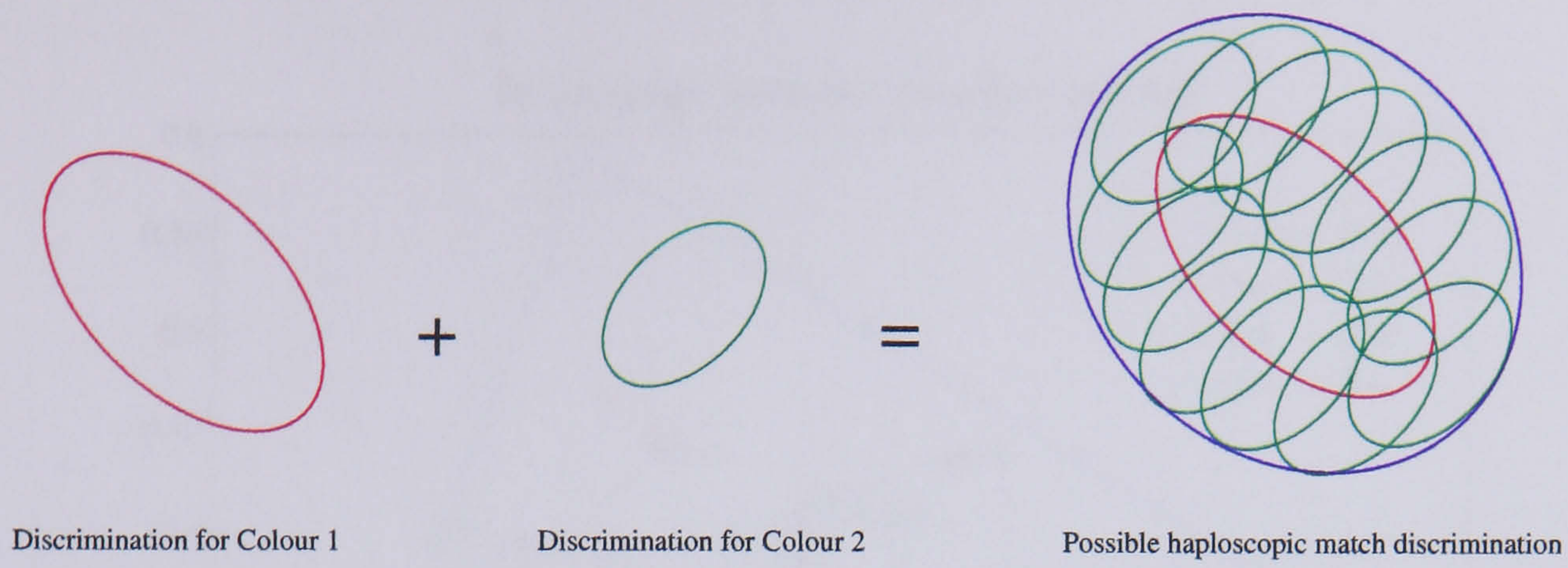

Figure 5.13: Worst case resultant discrimination for a haploscopic match experiment

discrimination abilities and not to colour constancy abilities.

This prediction can be simply tested. An observer matches four patches (N 5/, Y 5/4, G 5/4 and B 5/4) against three background conditions:

1. N 4/ under D_{40} for test and match presentations,
2. N 4/ under D_{65} for test and match presentations,
3. N 4/ under D_{40} on the test side and under D_{65} on the match side.

If the discrimination ellipses are retinal in origin then the match errors in condition three should be related to and larger than the largest match errors in either condition one or two. If the ellipses are cortical in origin then the matches in condition three should be dependent upon the area of colour space in which they reside and be comparable to the matches in condition two.

Results:

Two observers performed the full set of matches for this experiment. The results for TE, JH (female, experienced in colour matching, naive to the purposes of the experiment) are shown in Figure 5.14 and Figure 5.15 respectively. Other observers performed the experiment and produced qualitatively the same results.

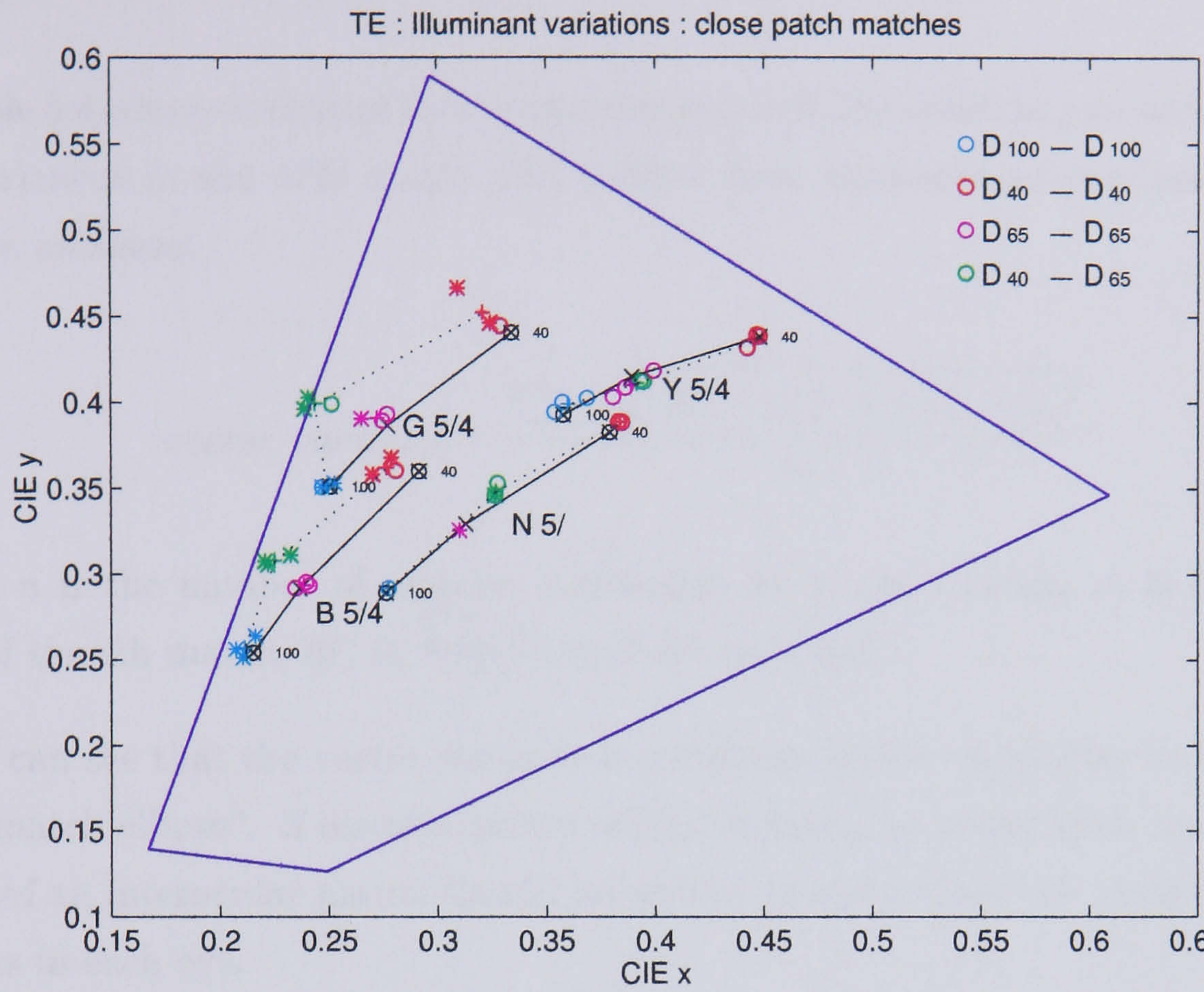

Figure 5.14: Matches under different illumination conditions for observer TE.

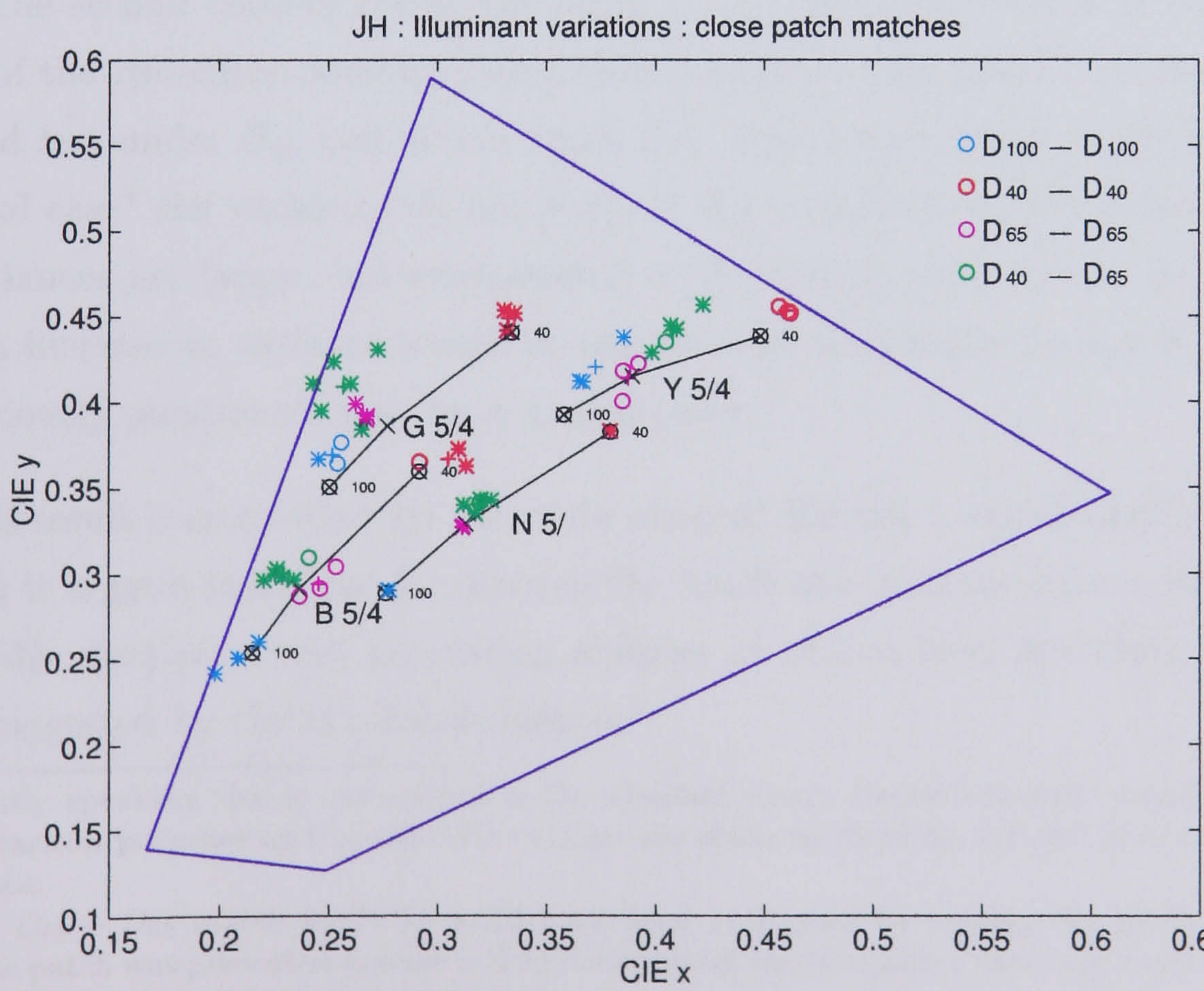

Figure 5.15: Matches under different illumination conditions for observer JH.

Analysis:

Table 5.4 shows statistics for the matches for each illumination pair and observer. The variances in the CIE x and CIE y have been combined to produce a vector variance measure:

$$vector\ variance = \frac{\sum_{i=1}^{n} \sqrt{(\overline{m_x} - x_i)^2 + (\overline{m_y} - y_i)^2}}{n} \tag{5.6}$$

where: n is the number of matches performed for a given patch, x_i is the CIE x value of the ith match, $\overline{m_x}$ is $\frac{\sum_{i=1}^{n} x_i}{n}$, and similarly for y.

We can see that the vector variance is a measure of the variability in the *radius* of the match ellipse[3]. If discrimination ability is retinal in origin then the discrimination of an interocular match should be related to and exceed the variances of the matches in each eye.

Table 5.5 tests the hypothesis for observer TE. The first column shows the match results for the illumination conditions of test under D_{40} and match under D_{65}. The second column shows the result of the maximum variances obtained in either of the two corresponding match experiments i.e test under D_{40} match under D_{40} and test under D_{65} and match under D_{65}. Apart from the N 5/ entry which is a special case[4] the variances do not support the retinal origin hypothesis (some of the variances are larger, but examination of the ellipses shown in figure 2.4 shows that an increase in variance would be expected for all patches except N 5/, which as previously mentioned, may be a special case).

This result is interesting for two main reasons. Firstly, it shows that haploscopic viewing is a good technique for probing the limits of colour constancy. Second, the colour discrimination and processing abilities at retinal level are likely to exceed those suggested by the MacAdam ellipses[5].

[3]Strictly speaking this is not correct in the absolute sense. Nevertheless the measure is useful for comparison purposes and is valid if no claims are made to absolute variability in terms of CIE xyY space.

[4]The $D_{65} - D_{65}$ match for N 5/ could have been performed by making the patch 'disappear' (since the patch was presented against a N 4/ background and luminance variation is not considered) resulting in the zero variance reported

[5]If information were lost at the retina then we would not expect to see the comparable resolution in the different illumination conditions. This reasoning assumes that, in the retinal origin scenario,

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{40}	0.384	0.388	0.000706	100_3	—	—	—	—
Y 5/4	D_{40}	D_{40}	0.446	0.436	0.00392	100_3	—	—	—	—
G 5/4	D_{40}	D_{40}	0.329	0.444	—	33.33_1	0.317	0.456	0.0125	66.67_2
B 5/4	D_{40}	D_{40}	0.281	0.36	—	33.33_1	0.275	0.362	0.00402	66.67_2
N 5/	D_{65}	D_{65}	—	—	—	—	0.31	0.325	$0^\dagger$	100_3
Y 5/4	D_{65}	D_{65}	0.389	0.41	0.00898	100_3	—	—	—	—
G 5/4	D_{65}	D_{65}	0.276	0.391	0.00219	66.67_2	0.265	0.39	—	33.33_1
B 5/4	D_{65}	D_{65}	0.239	0.293	0.00299	100_3	—	—	—	—
N 5/	D_{100}	D_{100}	0.278	0.291	—	100_3	—	—	—	—
Y 5/4	D_{100}	D_{100}	0.36	0.399	0.00658	100_3	—	—	—	—
G 5/4	D_{100}	D_{100}	0.247	0.35	—	33.33_1	0.25	0.352	0.00304	66.67_2
B 5/4	D_{100}	D_{100}	—	—	—	—	0.211	0.257	0.00619	100_3

† A problem in the experimental parameters reduced this match to 'make the patch disappear' i.e. the perfect match occurred when the patch was the same as the background.

JH	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{40}	—	—	—	—	0.379	0.383	—	100_3
Y 5/4	D_{40}	D_{40}	0.46	0.453	0.00262	100_3	—	—	—	—
G 5/4	D_{40}	D_{40}	—	—	—	—	0.333	0.449	0.00533	100_3
B 5/4	D_{40}	D_{40}	0.292	0.365	—	33.33_1	0.312	0.367	0.00528	66.67_2
N 5/	D_{65}	D_{65}	—	—	—	—	0.312	0.327	0.000323	100_3
Y 5/4	D_{65}	D_{65}	0.387	0.413	0.00948	100_3	—	—	—	—
G 5/4	D_{65}	D_{65}	—	—	—	—	0.266	0.394	0.00436	100_3
B 5/4	D_{65}	D_{65}	0.246	0.294	0.00899	100_3	—	—	—	—
N 5/	D_{100}	D_{100}	—	—	—	—	0.278	0.29	0.000222	100_3
Y 5/4	D_{100}	D_{100}	—	—	—	—	0.372	0.42	0.0143	100_3
G 5/4	D_{100}	D_{100}	0.255	0.37	0.00621	66.67_2	0.246	0.367	—	33.33_1
B 5/4	D_{100}	D_{100}	—	—	—	—	0.208	0.251	0.00929	100_3

Table 5.4: Observer match statistics for various illumination conditions

TE	variance $D_{40}^{test} : D_{65}^{match}$	max variance*	difference
N 5/	0.00362	0.000706	0.00291
Y 5/4	0.0038	0.00898	-0.00518
G 5/4	0.00345	0.00219	0.00126
B 5/4	0.0055	0.00402	0.00148

*maximum variance from either matches made under $D_{40}^{test} : D_{40}^{match}$ or $D_{65}^{test} : D_{65}^{match}$

Table 5.5: Comparison of variances for observer TE

5.5.3 Memory effects: how accurately can the patch colours be remembered?

Aims:

In order to allow comparisons between conditions and to keep the number of matches down to a reasonable level it is inevitable that an observer will be shown the same test patch several times, and thereby have multiple opportunities to adjust its matching patch under D_{65}. The motivation behind this experiment is to test whether memory effects play any role in the matching procedure.

Stimulus, set-up and methods:

A highly trained observer (DB), who has performed a large number of matches over several years, and who is very familiar with the Munsell patches tested, performed dial matches in response to colour names. (The names are given in written form in the centre of the test field; see Figure 5.16). The matches performed were with the understanding that they should be as close as possible to the 'remembered' patch corresponding to that name. Test and match patches are presented against a simulated background of Munsell patch N 4/ under illuminant D_{65}. The five remembered patches were: N 5/ (named as GREY), Y 5/4 (YELLOW), G 5/4 (GREEN) and B 5/4 (BLUE). The names ORANGE and RED were included for comparison purposes, i.e. to determine the accuracy of matching to a colour category not specifically representative of a remembered patch.

---

the signal is encoded after discrimination and that the original signals are not compared at cortex.

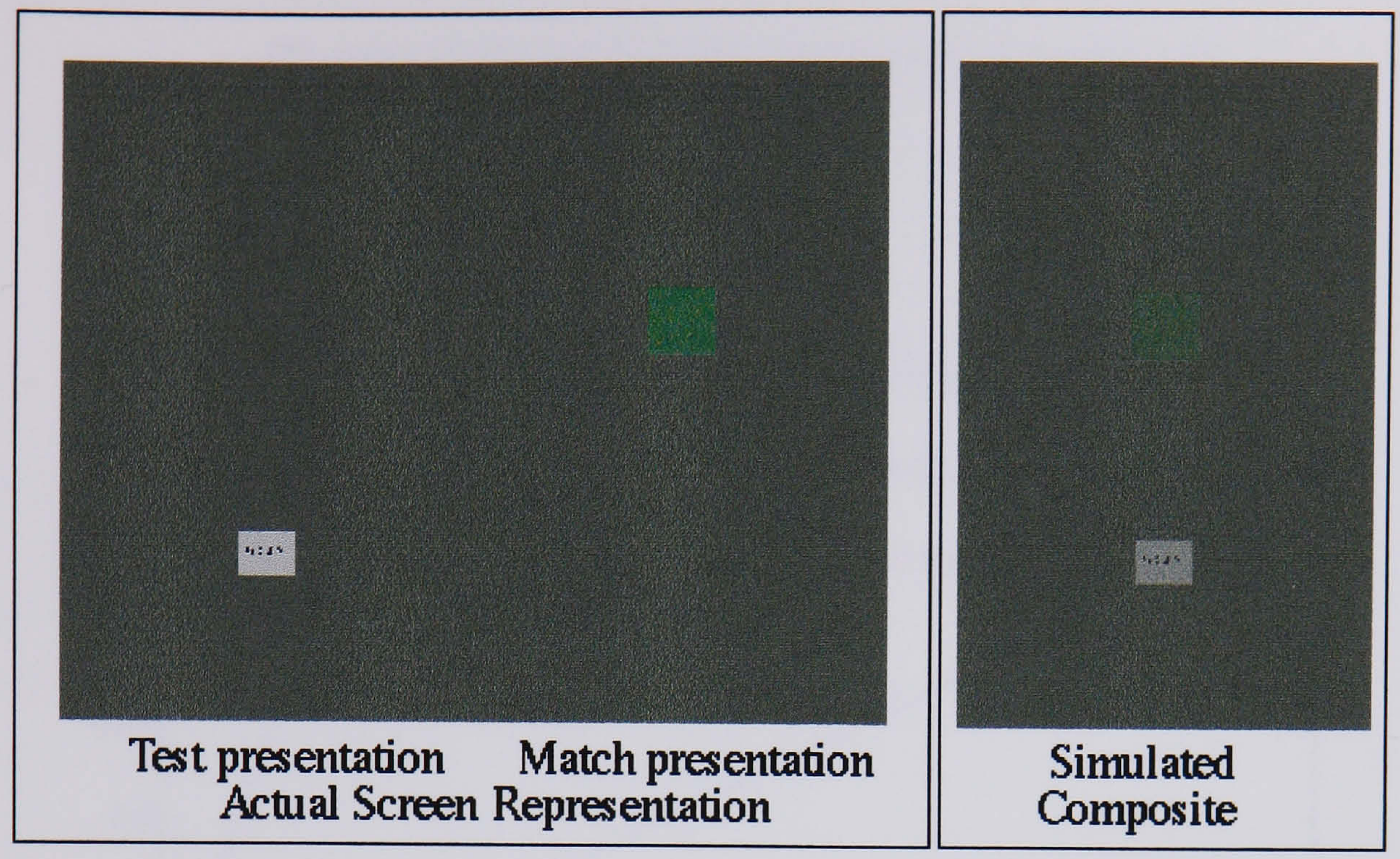

Figure 5.16: An example of a stimulus used in the memory match experiment. The colour name appears in the square on the left.

Results:

The results for DB are shown in Figure 5.17.

Analysis:

Table 5.6 shows statistics for memory matches to the colour names. Statistics for matches to actual patches for the same observer are also given where applicable. These matches used exactly the same viewing conditions and setup.

N 5/ is again a special case in this experiment producing a variance that should be considered a function of the paradigm and not representative of observer capabilities. The observer (DB) has a high degree of knowledge about the stimulus and matching setup which may have significantly affected his N 5/ memory match.

The rest of the patches show significantly higher variances for the memory matches over the direct matches. The means of the memory matches do not consistently correlate with those obtained for the Munsell patches either. These results suggest that it is safe to assume that memory of patch colours is not a major factor in the colour constancy performance reported in this thesis. It also suggests that

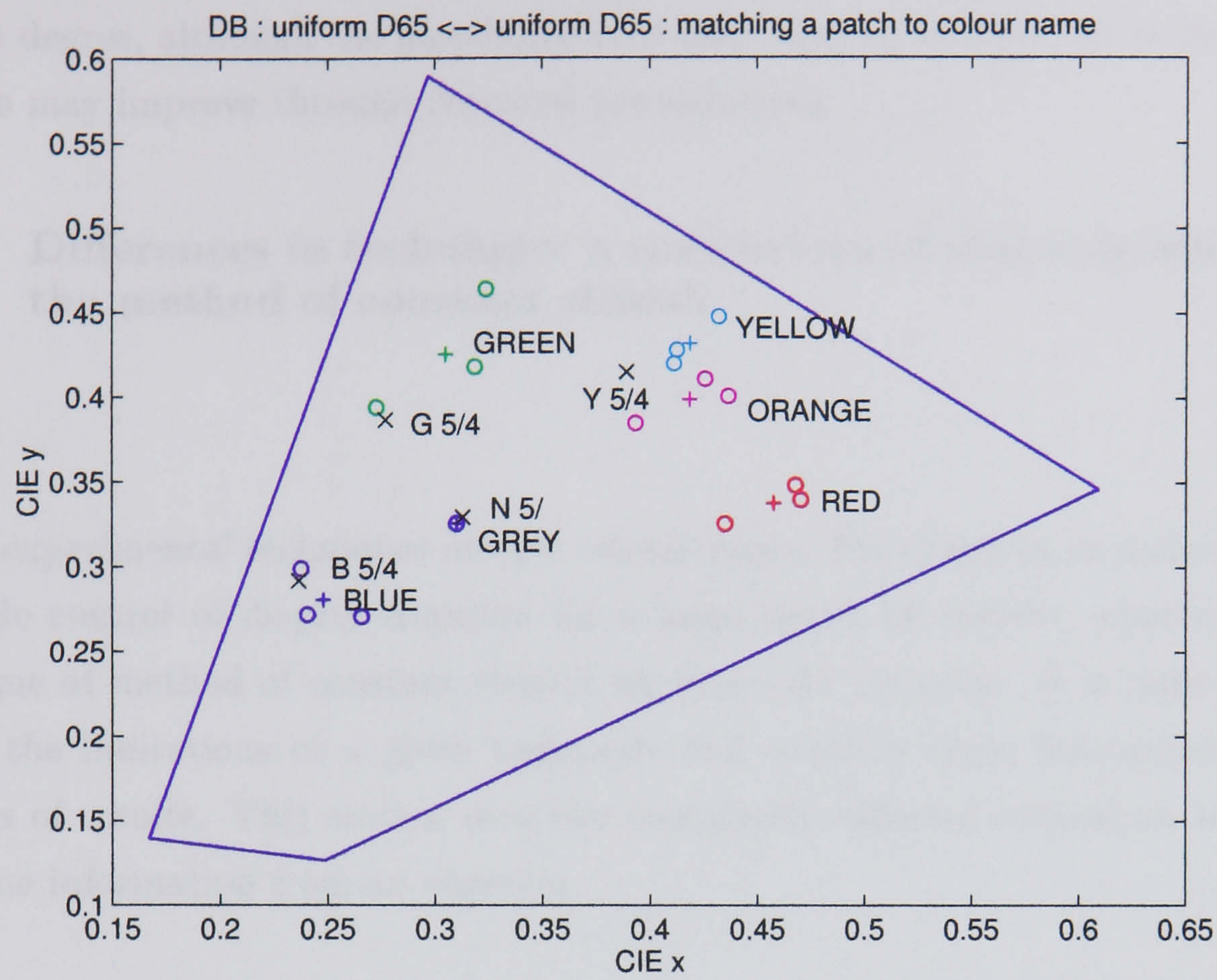

Figure 5.17: Memory matches to colour names performed by observer DB.

DB	Illuminants		Observer confidence high		
Patch	test	match	mean x	mean y	variance
GREY	D_{65}	D_{65}	0.310	0.325	0
RED	D_{65}	D_{65}	0.458	0.338	0.0177
YELLOW	D_{65}	D_{65}	0.419	0.432	0.014
GREEN	D_{65}	D_{65}	0.305	0.425	0.0343
BLUE	D_{65}	D_{65}	0.248	0.28	0.0176
ORANGE	D_{65}	D_{65}	0.419	0.399	0.0204
N 5/ $_{measured}$	D_{65}	D_{65}	0.310	0.316	0.00791
Y $5/4_{measured}$	D_{65}	D_{65}	0.399	0.416	0.00561
G $5/4_{measured}$	D_{65}	D_{65}	0.275	0.394	0.00475
B $5/4_{measured}$	D_{65}	D_{65}	0.230	0.276	0.00695

Table 5.6: Observer DB match statistics for memory matches and $D_{65} - D_{65}$

repeated presentations of the same patch do not affect the match results to a significant degree, although the possibility does exist that an observer's discrimination abilities may improve through repeated presentations.

5.5.4 Differences in technique: a comparison of dial matching with the method of constant stimuli

Aims:

All experimental techniques involve compromises. For example, in dial matching we trade control of display duration for a large choice of colours, whereas in the technique of method of constant stimuli we trade the opposite. It is important to realise the limitations of a given technique and consider these limitations in the analysis of results. This section uses two completely different techniques to derive the same information from an observer.

Stimulus, set-up and methods:

This experiment is performed in two stages. First the observer performs ten matches for patches N 5/, Y 5/4, G 5/4 and B 5/4 against a uniform N 4/ background. Simulated illumination of D_{40} is used on the test presentation and D_{65} for the match presentation. Of these ten matches five high confidence matches are chosen for each patch and made into a method of constants run. The trials for the patches are interleaved and randomly presented. Each match option is displayed several times allowing a 'likelihood of reporting the patches as different' to be measured. These results can be linearly interpolated to create a 'percentage reported as different' contour map.

If the contour is closed at this point then a two dimensional Gaussian surface can be fitted and values for the centroid of the distribution, its orientation and dimensions can be produced. In most cases more points will be needed, which can be selected from consideration of the contour map already produced. Once the points have been selected a new experimental run is set up and another contour plot produced.

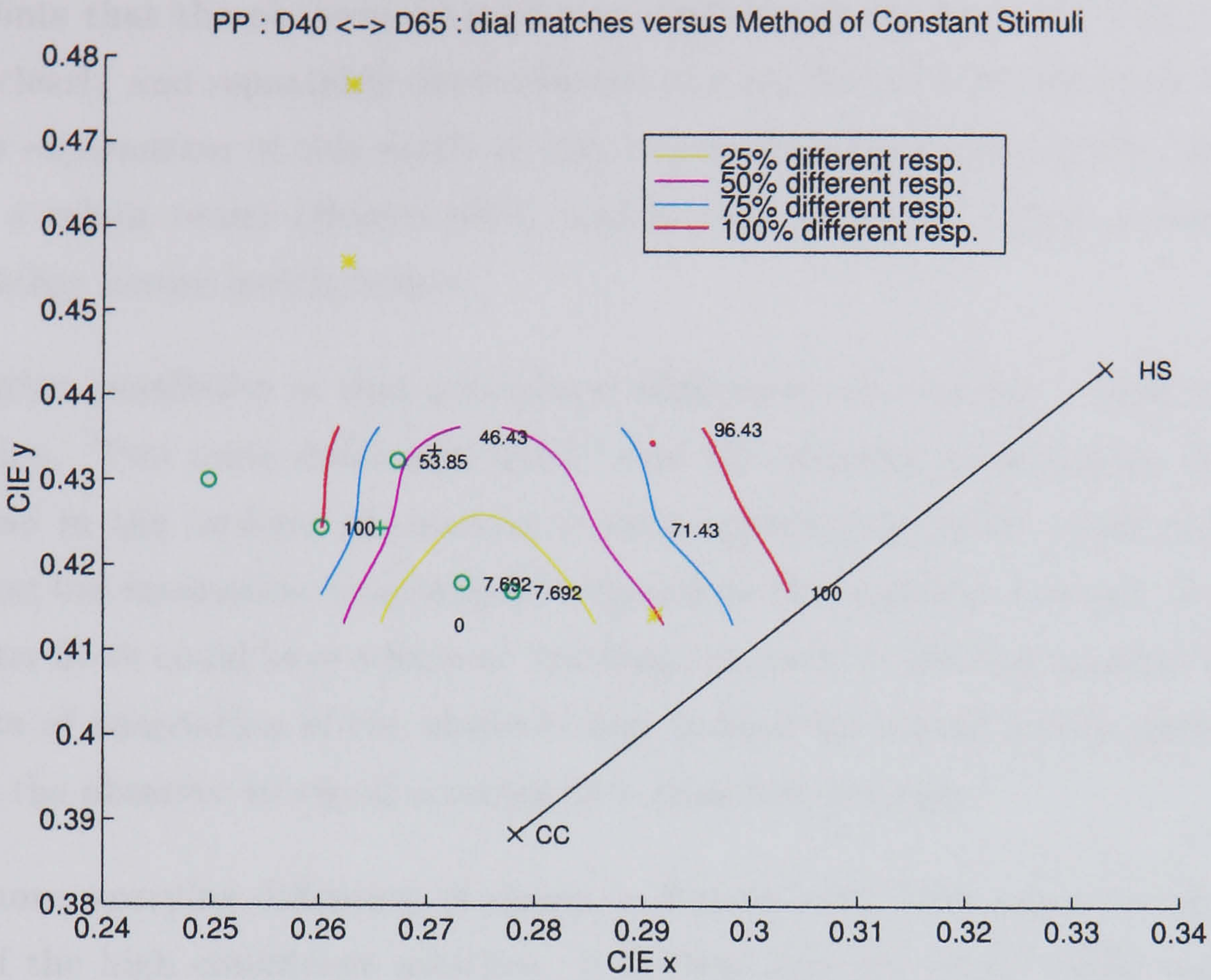

Figure 5.18: Comparison of dial matching with 2D method of constant stimuli.

Results:

Figure 5.18 shows the results obtained for G 5/4 by this procedure for observer PP (male, experienced observer, naive as to the purposes of the experiment). A small number on the plot shows the percentage different responses obtained for a match point just to the left of the number.

Analysis:

(Note: The results presented in this and the following section are mainly designed to be illustrative of the general trend. A far more thorough analysis of the method of constants technique with more rigorous statistical analysis is provided in Bramwell and Hurlbert (Bramwell and Hurlbert 1996), from which some of the conclusions made here are derived).

The first point of interest is shown in Table 5.7. All of the match points given in this table were chosen from HIGH observer confidence dial matches. Interestingly when these same points are used in a method of constants procedure we can see that points that the observer claimed were indistinguishable during dial matching can be clearly and repeatably discriminated in a method of constants procedure. A possible explanation of this result is that colour matches are normally distributed around a colour centre (Brown 1952), and as such we would expect a variation of detectability across match points.

Another possibility is that procedural differences are causing a shift in colour perception. Two main differences exist: first the stimulus presentation is strictly controlled in the method of constant stimuli experiment, which could result in a consistent but incomplete adaptational response to the stimulus. Second, as outlined in chapter 3, we could have a form of 'tracking response' in the dial matches in which the state of adaptation of the observer lags behind the actual screen presentation causing the observer to signal a match to a transient percept.

A more worrying difference is shown in Figure 5.18. The green '+' shows the mean of the high confidence matches. It is clear that the mean of the method of constants results does not coincide with the mean of the dial matches. This could signal that we do not have enough dial matches for an accurate representation of observer abilities or that a colour shift is inherent in one of the techniques.

In general[6], the method of constants procedure suggests better constancy performance than that suggested by dial matching. It also suggests a lower variance in the matches, i.e. better discrimination abilities. In summary, we can say that the two techniques produce 'qualitatively' the same results and, if anything, the performance of the visual system is likely to exceed that suggested by dial matching. Since dial matching is a great deal quicker and all the results are used comparatively (i.e. a dial match under one condition is compared with a dial match from another condition) the rest of the experimentation is conducted using dial matches.

5.5.5 Bias considerations in the method of constant stimuli

Aims:

One-alternative-binary-choice procedures can be susceptible to observer bias. The multiple experimental runs procedure outlined in the previous section could

[6]These conclusions are made with reference to data not presented in this thesis.

match x	match y	% 'different'
0.322	0.340	26.92_{26}
0.333	0.357	100_{26}
0.322	**0.331**	$\mathbf{0.00}_{26}$
0.314	0.338	100.0_{26}
0.316	0.340	96.15_{26}

Table 5.7: Method of constants results for HIGH confidence matches (patch N 5/ against multi patch, test under D_{40}, match under D_{65})

produce erroneous results if bias is a problem. The criterion given to the observer is strict and well defined, i.e. if the observer can detect any difference in the patches he/she must say that they are different. This should reduce the risk of observer shifts in criterion. The following experiment is designed to determine whether the particular sets of stimuli used in the method-of-constant-stimuli runs may induce a bias in observer response.

Stimulus, set-up and methods:

All stimulus features are as reported in the previous section.

Two experimental runs are constructed that provide different match alternatives to the same test patches. The average of the collection of match patches is shifted between the experiments and, for comparison purposes, one of the match options is the same in both experiments. If a bias is induced through the collection of match alternatives then it is expected that the two distributions will not combine into a coherent elliptical surface and that the percentage of 'different' responses given for the common match point will be significantly different.

Results:

Figures 5.19, 5.20 and 5.21 show the contour plots produced for two sets of match alternatives for N 5/, Y 5/4 and B 5/4 respectively. A small number on the plot shows the percentage 'different' responses obtained for a match point just to the left of the number.

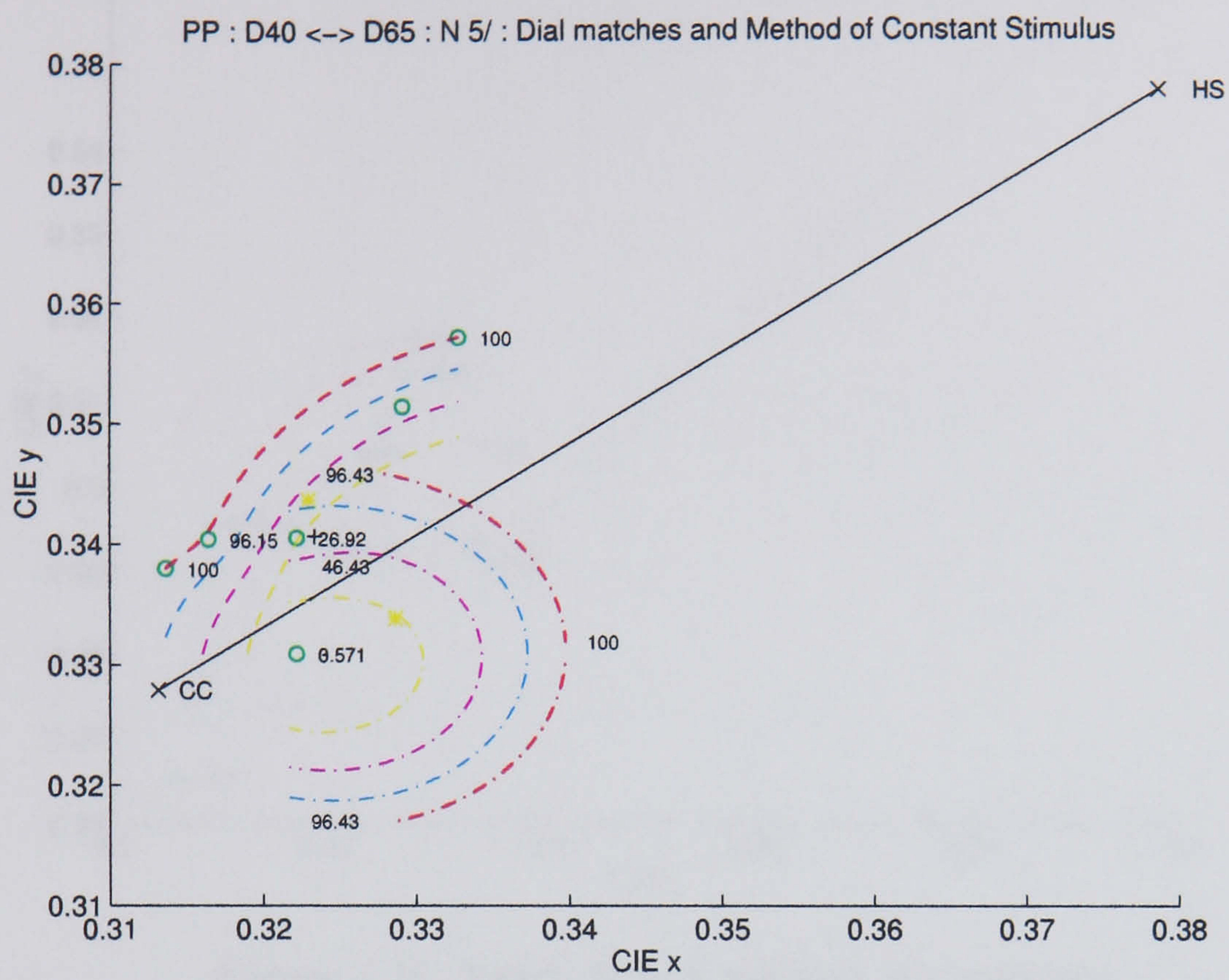

Figure 5.19: Patch N 5/ method of constants.

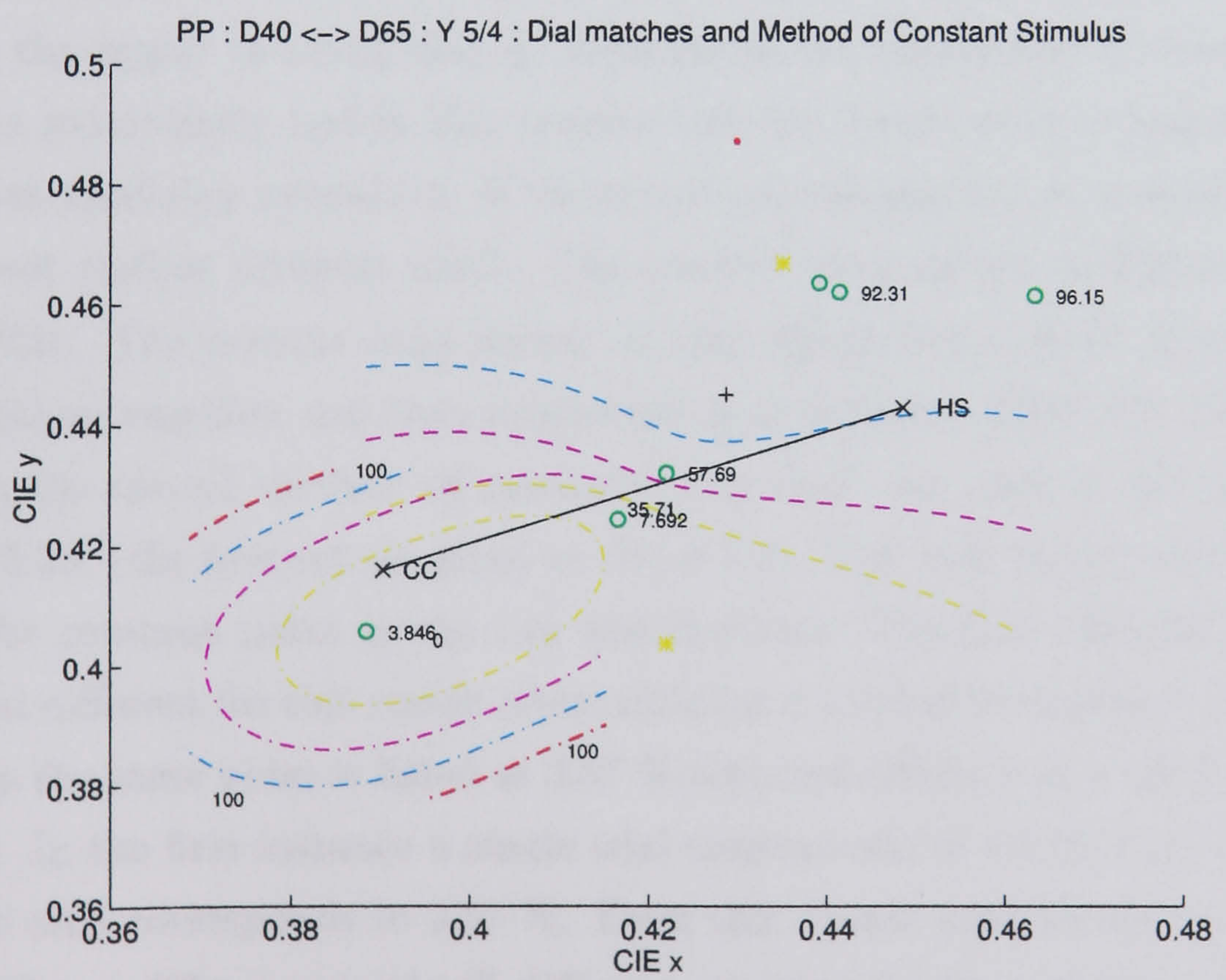

Figure 5.20: Patch Y 5/4 method of constants.

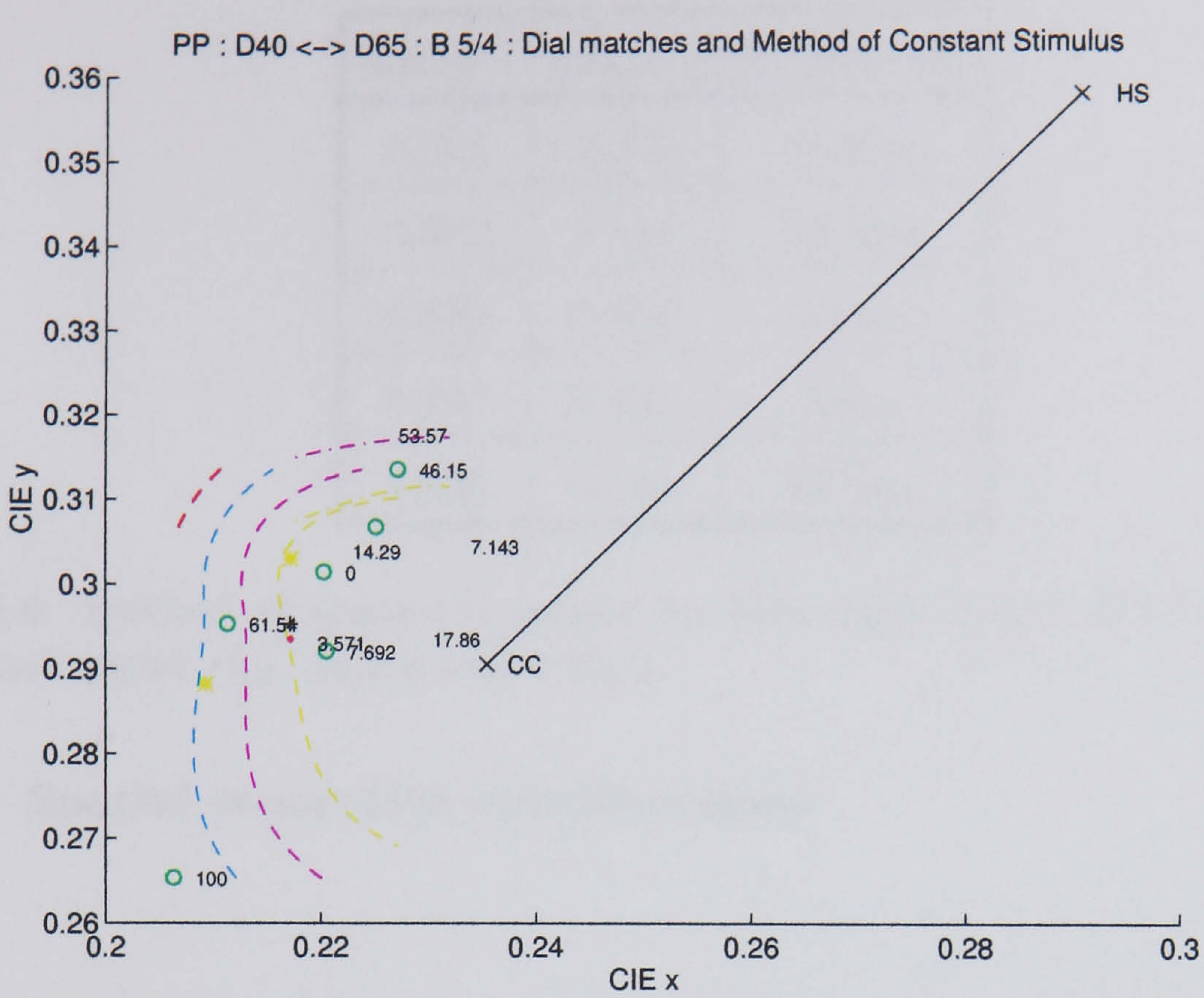

Figure 5.21: Patch B 5/4 method of constants.

Analysis:

Examination of the figures shows that in general the distributions concur, although the degree to which they do depends on the distribution of match points. N 5/ looks particularly bad in this respect but this discrepancy is largely a function of the interpolation procedure. If the actual percentages are examined then a more consistent surface presents itself. The contour map shown in Figure 5.18 highlights this. The contour map shown in that figure was created from joining two distributions together and then interpolating as a whole. Table 5.8 lists the points used in the second method of constants run that was used in the production of Figure 5.18 (the first set are given in Table 5.7). The bold entries in the two tables show the common point in the two distributions. The first distribution lists 0 % reported different for this match alternative for a total of 26 repeated presentations, whereas the same point is listed as 3.57 % reported different in a total of 28 presentations. In the first instance a single trial corresponds to 3.8 %, and in the second, a single trial corresponds to 3.57 %. Even this simple analysis shows that there is no significant difference in the % different reported for the common point.

match x	match y	% 'different'
0.322	0.338	46.43_{28}
0.322	0.317	96.43_{28}
0.322	**0.331**	$\mathbf{3.57_{28}}$
0.340	0.332	100_{28}
0.323	0.346	96.43_{28}

Table 5.8: Method of constants results for bias check (patch N 5/ against multi patch, test under D_{40}, match under D_{65})

5.5.6 Spatial separation considerations

Aims:

Many of the match studies examined in chapter 4 require observers to match patches that are abutting or have a small separation. In order to test aspects of complex surrounds on matching it becomes necessary to move the patches apart. This experiment compares the case of matches of abutting patches to patches separated by 5^o.

Stimulus, set-up and methods:

The stimuli are simulations of patches N 5/, Y 5/4, G 5/4 and B 5/4 against uniform backgrounds of N 4/. The test field is under simulated D_{40} illumination and the match side is under simulated D_{65}. In the first experiment the patches are presented almost abutting vertically (when the displays are fused). In the second experiment the patches are moved to the location used for the patch match experiments in the next chapter (i.e. patches centred within the test and match fields resulting in a vertical separation of 5^o, when fused).

Results:

The results for observers DB and TE are shown in Figure 5.22.

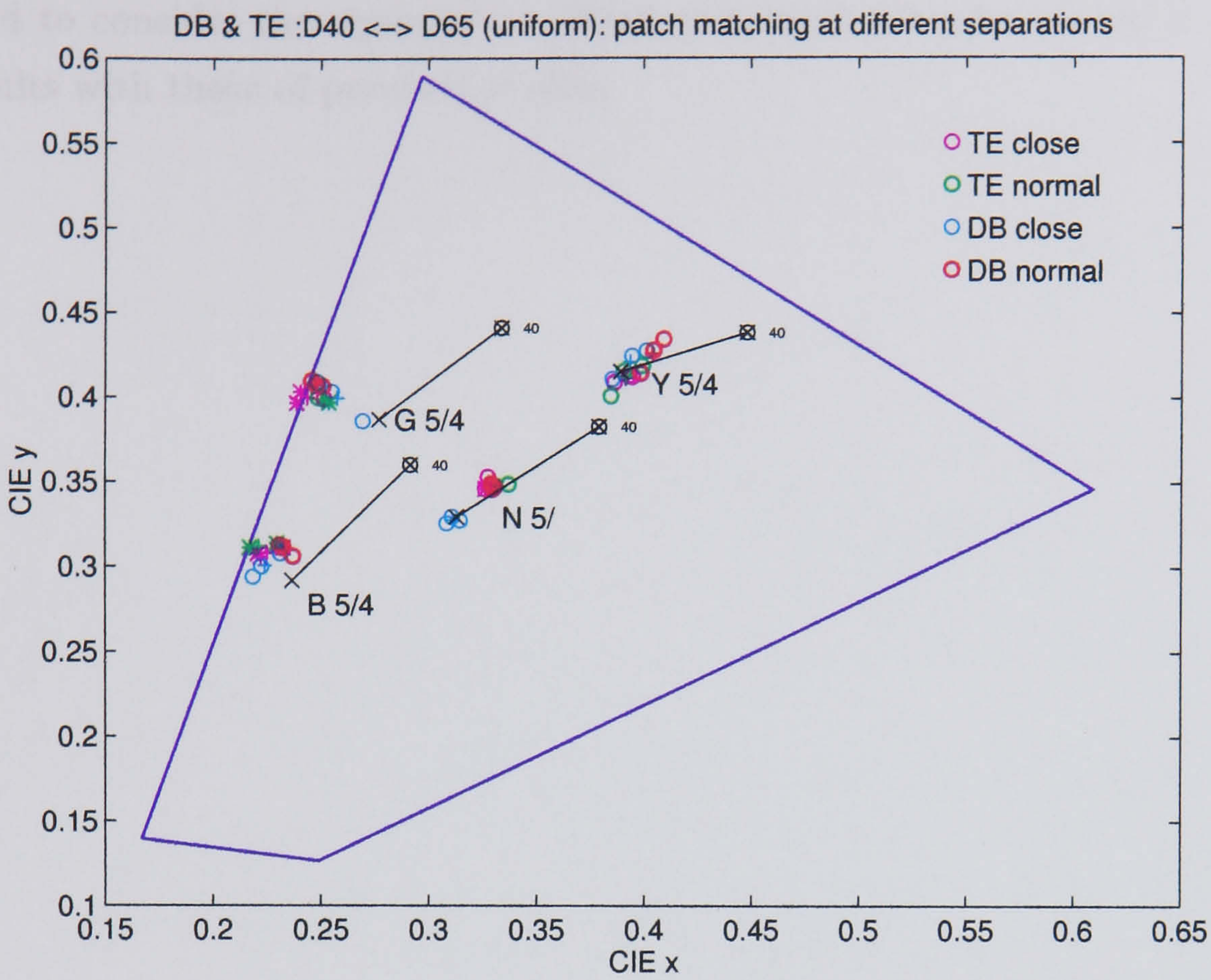

Figure 5.22: Comparison of patch display separations for observers DB and TE.

Analysis:

The main reason for expecting a difference in 'close' and separated patches are related to adaptational differences and the possible influences of local contrast. The observers are not requested to fixate between the match and test patches and as such we expect variation in local adaptation due to eye movements. In both cases one eye looks at the adapting background of the other field while the other eye views one of the patches. The exact state of adaptation will depend on the amount of time spent looking at each patch and the amount of influence spatially local areas have on colour appearance (in the close case we would expect the patches to have greater influence on each other due to their small spatial separation).

Table 5.9 presents match statistics for the two observers. The difference entries show no consistent shift in mean or variance for either observer. The differences are just as likely due to the small number of samples as to patch separation. These results suggest it is unlikely that the results presented in this thesis are significantly affected by the location of the test and match patches.

All experimental results are examined in a comparative manner which obviates the need to consider the separations. What this result allows us to do is compare our results with those of previous studies.

DB	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/ $_{close}$	D_{40}	D_{65}	0.311	0.327	0.00292†	100_3	—	—	—	—
Y 5/4_{close}	D_{40}	D_{65}	0.394	0.42	0.00895	100_3	—	—	—	—
G 5/4_{close}	D_{40}	D_{65}	0.258	0.399	0.0119	100_3	—	—	—	—
B 5/4_{close}	D_{40}	D_{65}	0.224	0.3	0.00683	100_3	—	—	—	—
N 5/	D_{40}	D_{65}	0.33	0.347	0.00122†	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.404	0.425	0.00836	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	0.248	0.407	0.00229	100_3	—	—	—	—
B 5/4	D_{40}	D_{65}	0.234	0.309	0.0035	100_3	—	—	—	—
N 5/ $_{diff}$	—	—	-0.019	-0.02	0.0017	100_3	—	—	—	—
Y 5/4_{diff}	—	—	-0.01	-0.005	0.00059	100_3	—	—	—	—
G 5/4_{diff}	—	—	0.01	-0.008	0.009	100_3	—	—	—	—
B 5/4_{diff}	—	—	-0.01	-0.009	0.00333	100_3	—	—	—	—

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/ $_{close}$	D_{40}	D_{65}	0.327	0.349	0.00362†	66.67_2	0.326	0.345	—	33.33_1
Y 5/4_{close}	D_{40}	D_{65}	0.392	0.411	0.0038	100_3	—	—	—	—
G 5/4_{close}	D_{40}	D_{65}	0.252	0.398	—	33.33_1	0.24	0.399	0.00345	66.67_2
B 5/4_{close}	D_{40}	D_{65}	—	—	—	—	0.225	0.308	0.0055	100_3
N 5/	D_{40}	D_{65}	0.333	0.347	0.00294†	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.392	0.411	0.00904	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	0.249	0.399	—	66.67_2	0.254	0.396	—	33.33_1
B 5/4	D_{40}	D_{65}	—	—	—	—	0.221	0.311	0.00563	100_3
N 5/ $_{diff}$	—	—	-0.006	0.002	0.00068	100_3	—	—	—	—
Y 5/4_{diff}	—	—	—	—	0.00524	100_3	—	—	—	—
G 5/4_{diff}	—	—	-0.109	-0.001	—	66.67_2	-0.014	0.003	—	33.33_1
B 5/4_{diff}	—	—	—	—	—	—	-0.003	-0.003	-0.00013	100_3

† In this experiment the task could not be performed by 'making the patch disappear' and as such is likely to be more representative of the variance in matches for patch N 5/.

Table 5.9: Match statistics and differences for DB and TE for two patch separations

Chapter 6

Experimental Results: Patch Matching against Uniform and Multi-Patch Backgrounds

6.1 Introduction

It was suggested in chapter 3 that the contrast present in the scene as a whole could have a large effect on match performance. The experiments in this chapter are designed to examine matches against uniform and multi-patch backgrounds. Section 6.4 at the end of this chapter examines the proposed model in the context of the results obtained in this chapter.

The new feature added in this chapter is the multi-patch backgrounds. Before describing them in detail a few general features will be discussed. The backgrounds are designed to produce a large amount of chromatic contrast without biasing the scene average chromaticity. The patches have been made small in size to reduce local colour contrast effects. The backgrounds are created from Munsell reflecting surfaces so that an accurate simulation can be produced of the same collection of papers under differing illumination conditions.

A selection of 24 Munsell patches were chosen for the multi-patch background. Table 6.1 lists the Munsell designations of the patches along with their calculated chromaticities and cone excitations under D_{65}, D_{100} and D_{40}. Figure 6.1 shows the multi-patch background positions in CIE xy space under D_{65} and D_{40}. Fig-

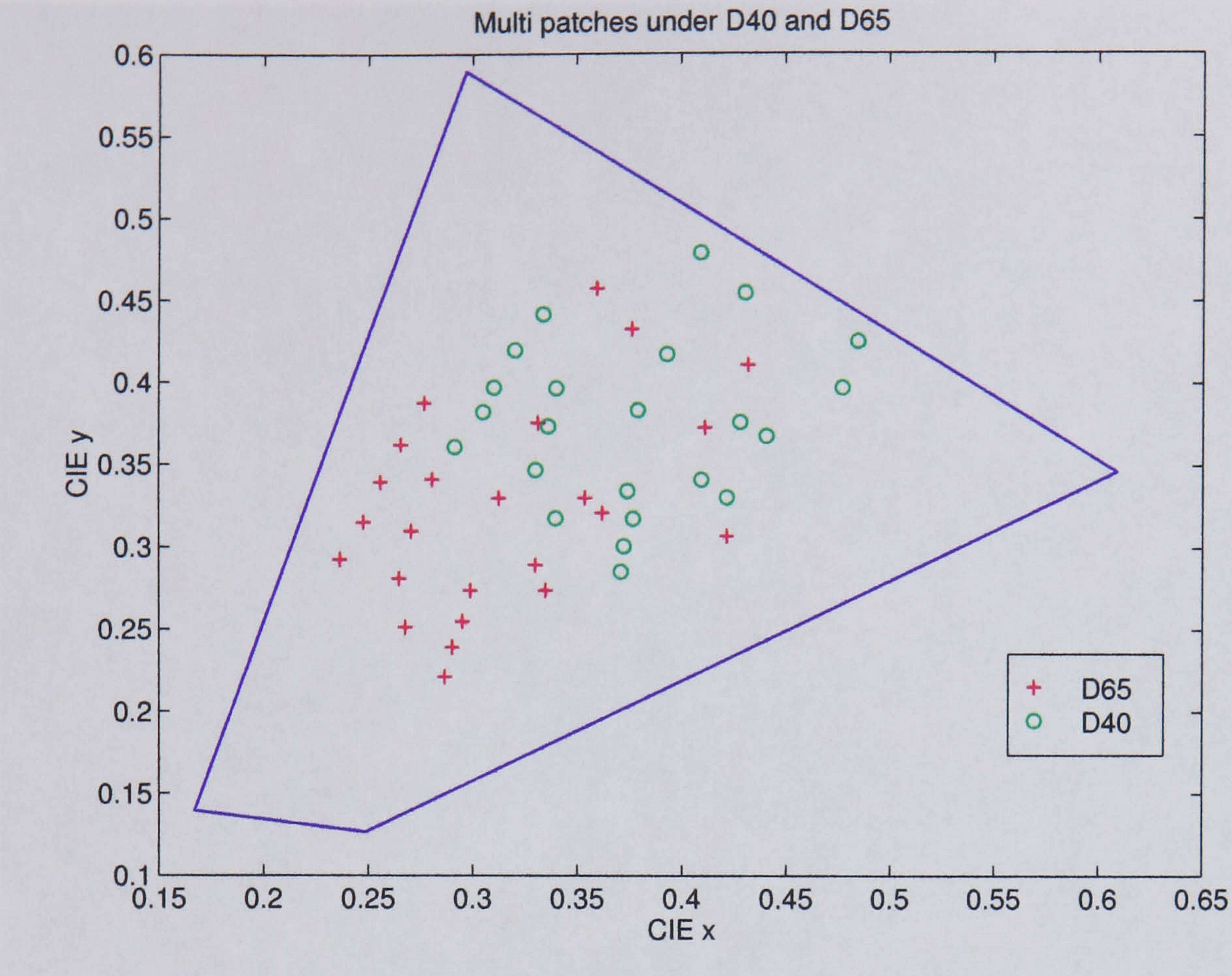

Figure 6.1: Multi patch coordinates under D_{40} and D_{65}

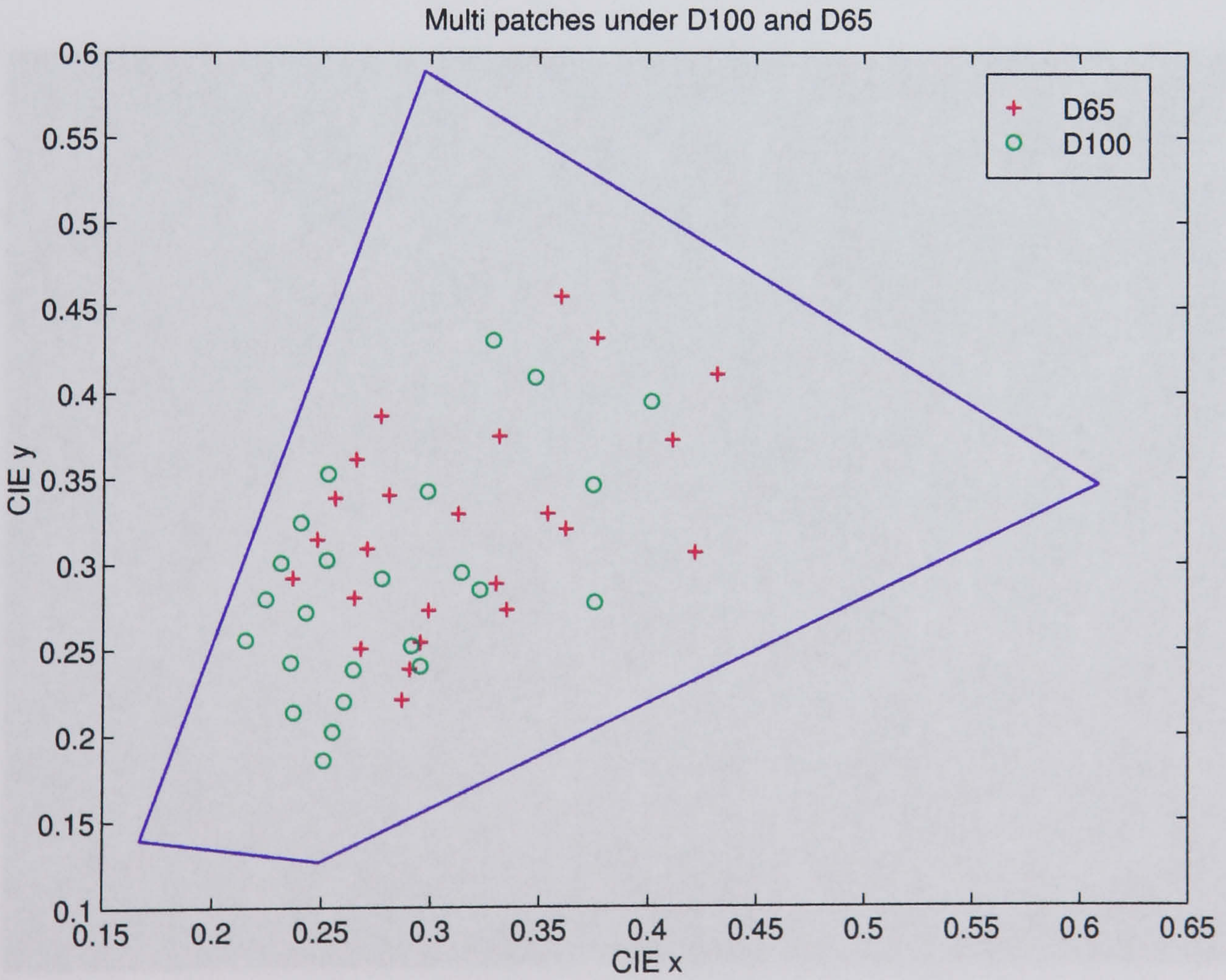

Figure 6.2: Multi patch coordinates under D_{100} and D_{65}

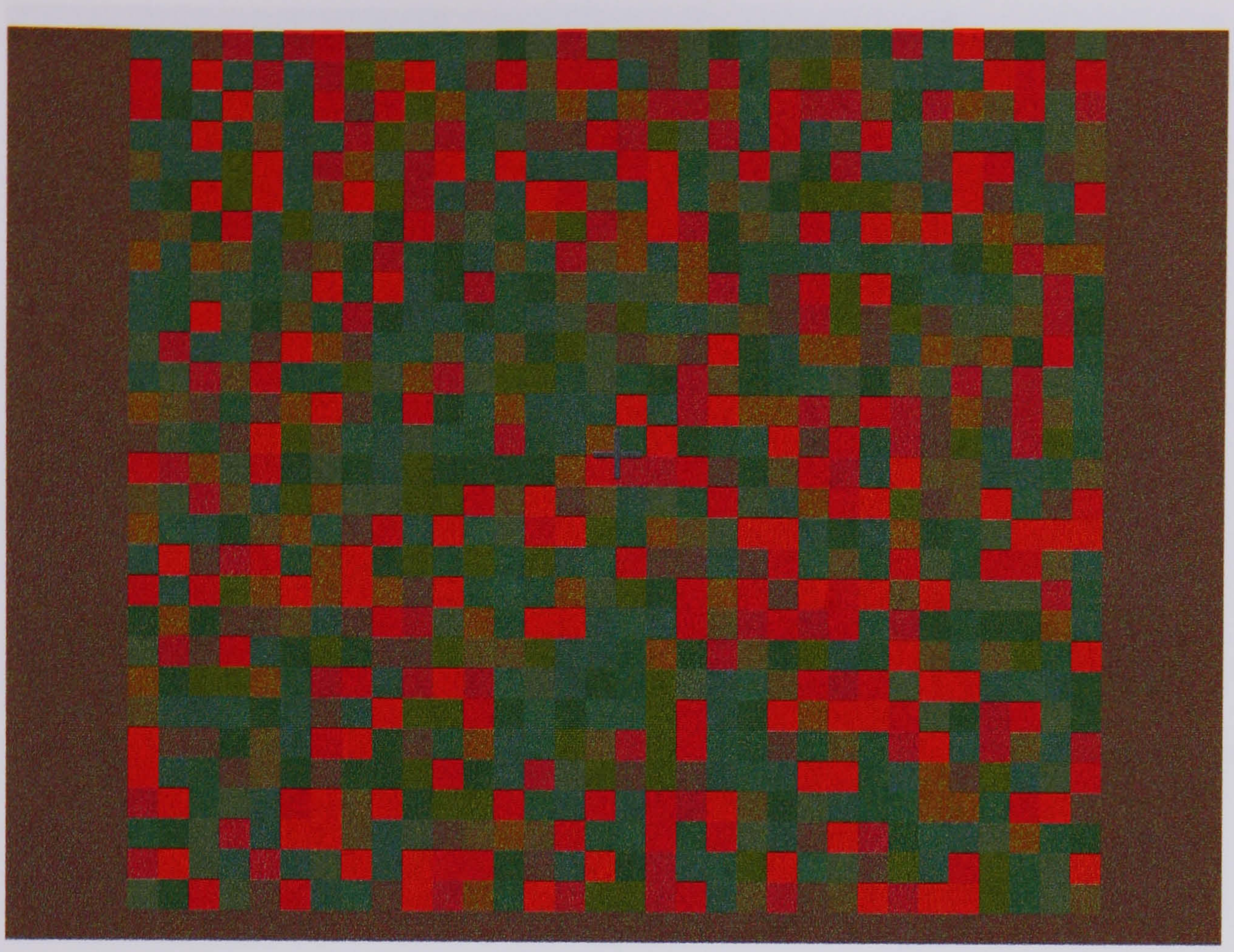

Figure 6.3: Example of multi patch background under D_{40}

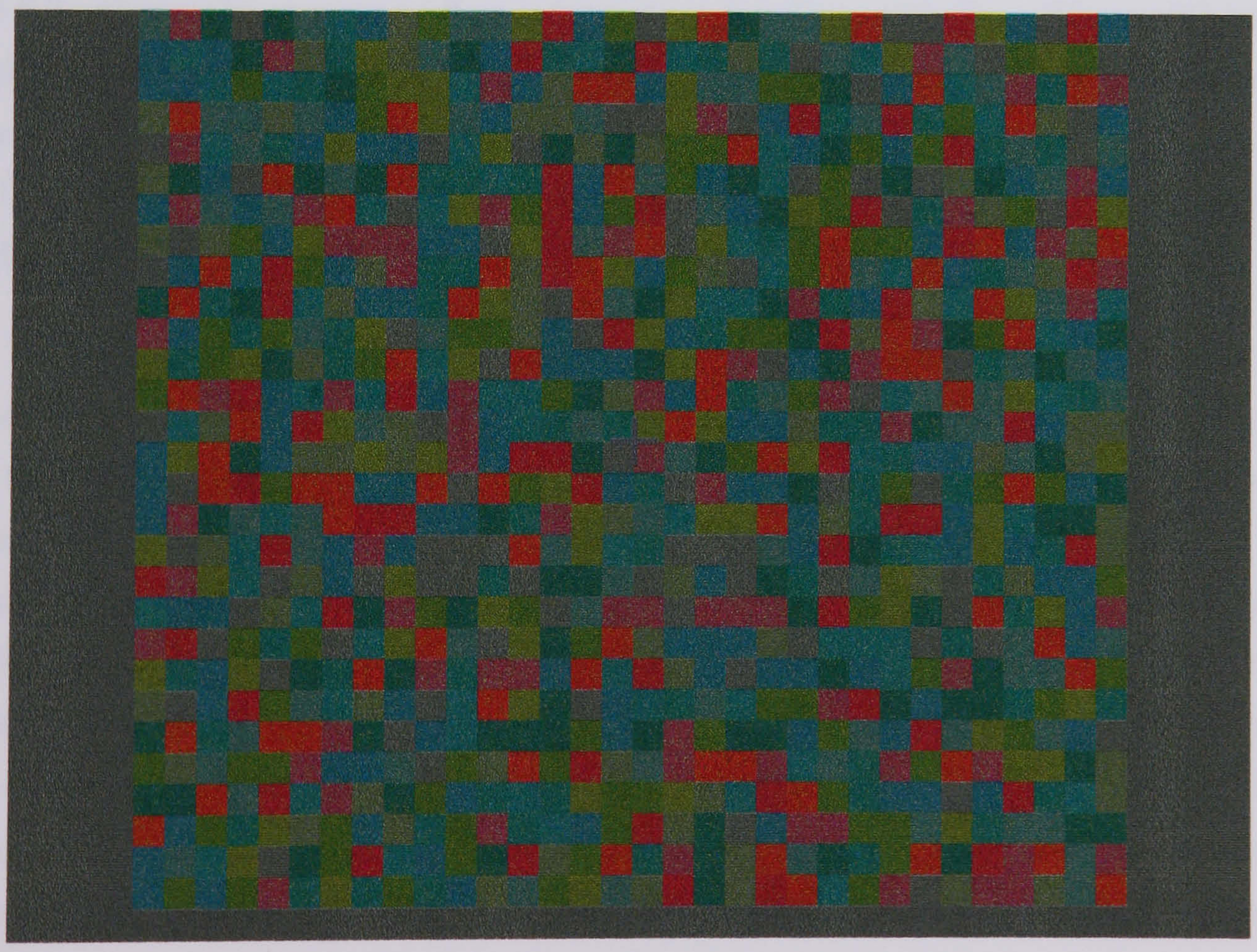

Figure 6.4: Example of multi patch background under D_{65}

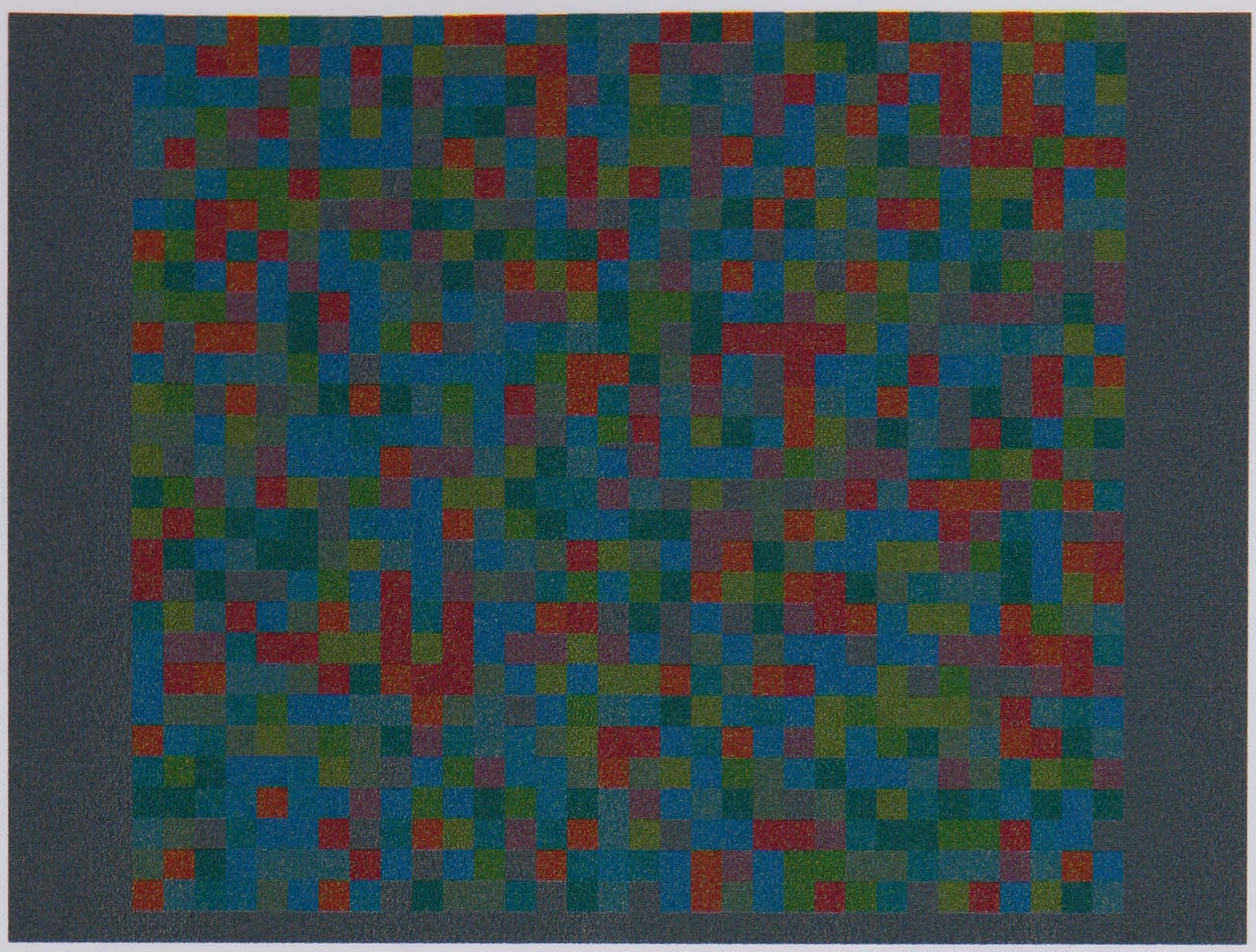

Figure 6.5: Example of multi patch background under D_{100}

ure 6.2 shows the corresponding information for D_{100} and D_{65}. The multi-patch background is created from a square array of 65x65 of these patches (i.e. each patch is used more than once) displayed in random spatial order, which changes from trial to trial. Figures 6.3, 6.4 and 6.5 show examples of the multi patch background under three illuminants.

Valberg and Lange-Malecki (Valberg and Lange-Malecki 1990) claim they can produce an equivalent uniform background for chromatic fields but do not specify the weighted function they used to calculate it. A uniform field of the same mean chromaticity may not be equivalent to the multi-patch background, but the equivalence is tested by experimentation. The LMS excitations of the multi patch background, the test patches and equivalent backgrounds for the illuminants tested along with the illuminant spectra are shown in figures 6.6 and 6.7. The chromaticity coordinates of the uniform and equivalent backgrounds are listed in Table 6.2. Two different methods were used to calculate the equivalent uniform backgrounds. The method used for some experiments in the following chapter calculates the background from a simple mean of the Munsell patches along with simple frequency of use information; the background chromaticities produced by this method are la-

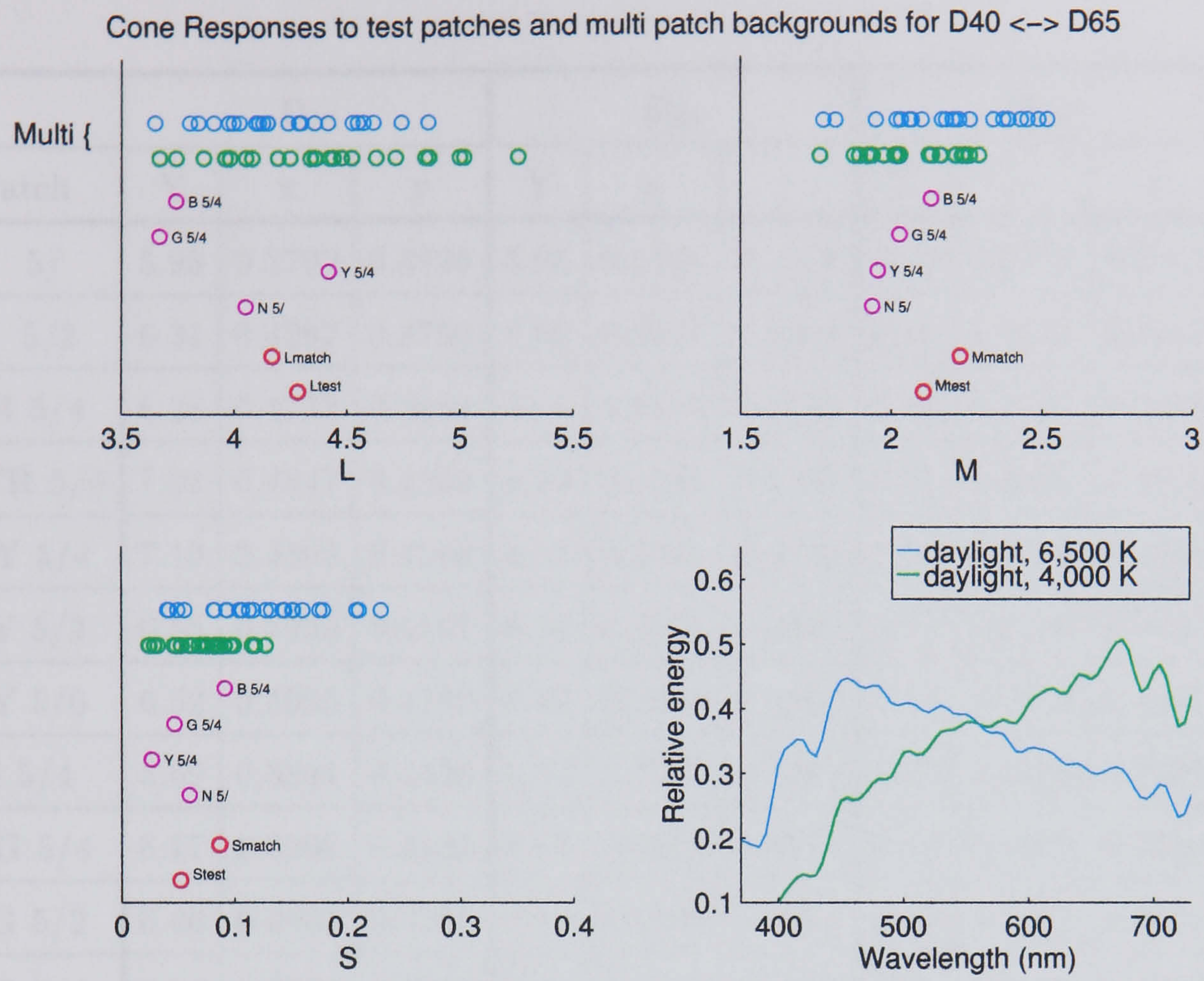

Figure 6.6: LMS plots of the test patches, equivalent backgrounds (match for E_{65}; test for E_{40}) and multi patch backgrounds under D_{40} and D_{65}

belled E'. Observer dial match results with E' backgrounds show a consistent shift suggesting that either perceptually the equivalent uniform background for a multi-patch simulation is not simply the mean of the patches, or the method of calculation is inaccurate. The method of equivalent background calculation was shown to be the problem when a second method removed the observed offsets. The second method is more advanced and yields better estimates of an equivalent uniform background, here labelled E. This method involves the generation of many of the backgrounds by the experimental program and averaging over a whole experimental run, taking the random number generator into account. The issue of matching differences caused by differences in equivalent backgrounds is examined in more detail in the next chapter.

	D_{40}			D_{65}			D_{100}		
Patch	Y	x	y	Y	x	y	Y	x	y
N 5/	5.98	0.3792	0.3826	5.96	0.3126	0.3290	6.13	0.2775	0.2913
R 5/2	6.31	0.4282	0.3756	6.05	0.3537	0.3292	6.16	0.3141	0.2947
YR 5/4	6.28	0.4773	0.3968	5.96	0.4113	0.3722	5.86	0.3747	0.3458
10YR 5/6	7.03	0.4847	0.4249	6.59	0.4321	0.4106	6.54	0.4016	0.3946
10Y 5/4	7.10	0.4308	0.4540	6.95	0.3765	0.4320	7.00	0.3479	0.4089
GY 5/2	6.55	0.3933	0.4167	6.54	0.3315	0.3746	6.75	0.2986	0.3423
GY 5/6	6.52	0.4095	0.4782	6.49	0.3598	0.4565	6.64	0.3285	0.4307
G 5/4	5.69	0.3341	0.4406	5.82	0.2771	0.3865	6.09	0.2528	0.3526
10G 5/4	6.17	0.3205	0.4188	6.45	0.2658	0.3611	6.72	0.2403	0.3240
BG 5/2	6.40	0.3403	0.3958	6.53	0.2808	0.3401	6.84	0.2522	0.3021
BG 5/4	6.37	0.3104	0.3961	6.64	0.2561	0.3383	7.06	0.2312	0.3005
10BG 5/4	6.21	0.3052	0.3812	6.44	0.2478	0.3143	6.88	0.2242	0.2793
B 5/2	6.21	0.3052	0.3812	6.05	0.2708	0.3088	6.28	0.2427	0.2715
B 5/4	5.87	0.3363	0.3726	6.17	0.2366	0.2916	6.58	0.2151	0.2554
PB 5/4	5.87	0.2916	0.3599	6.84	0.2649	0.2802	7.14	0.2357	0.2422
10BP 5/6	6.60	0.3301	0.3461	6.74	0.2679	0.2506	7.01	0.2371	0.2132
P 5/4	6.64	0.3397	0.3168	6.00	0.2990	0.2728	6.25	0.2644	0.2380
P 5/6	6.16	0.3742	0.3335	6.15	0.2953	0.2542	6.30	0.2602	0.2193
P 5/8	6.23	0.3769	0.3167	6.61	0.2903	0.2383	6.75	0.2550	0.2018
P 5/10	6.73	0.3724	0.3000	6.39	0.2868	0.2205	6.58	0.2508	0.1853
10P 5/4	6.56	0.3709	0.2845	6.71	0.3300	0.2885	6.77	0.2912	0.2520
10P 5/6	6.89	0.4096	0.3405	6.59	0.3349	0.2730	6.76	0.2951	0.2400
10RP 5/4	6.97	0.4218	0.3298	7.08	0.3620	0.3202	7.09	0.3225	0.2847
10RP 5/8	7.38	0.4408	0.3673	6.51	0.4217	0.3064	6.48	0.3754	0.2772

Table 6.1: Munsell patches used in the multi-patch background under the three illuminants

Illuminant	Patch	Y	x	y
D_{40}	N 5/	5.98	0.3792	0.3826
D_{65}	N 5/	5.96	0.3126	0.3290
D_{100}	N 5/	6.13	0.2775	0.2913
E_{40}	—	6.38	0.3811	0.4110
E_{65}	—	6.39	0.3229	0.3678
E_{100}	—	6.56	0.2924	0.3366
E'_{40}	—	6.0	0.39	0.39
E'_{65}	—	6.0	0.312	0.33

(Notes: D designates a daylight illuminant; E designates a calculated uniform background created from averaging all the chromaticities in several multi-patch displays; E' designates a calculated uniform background created from averaging the chromaticities of the Munsell patches in a single multi-patch display.)

Table 6.2: Coordinates of backgrounds and illuminant designations

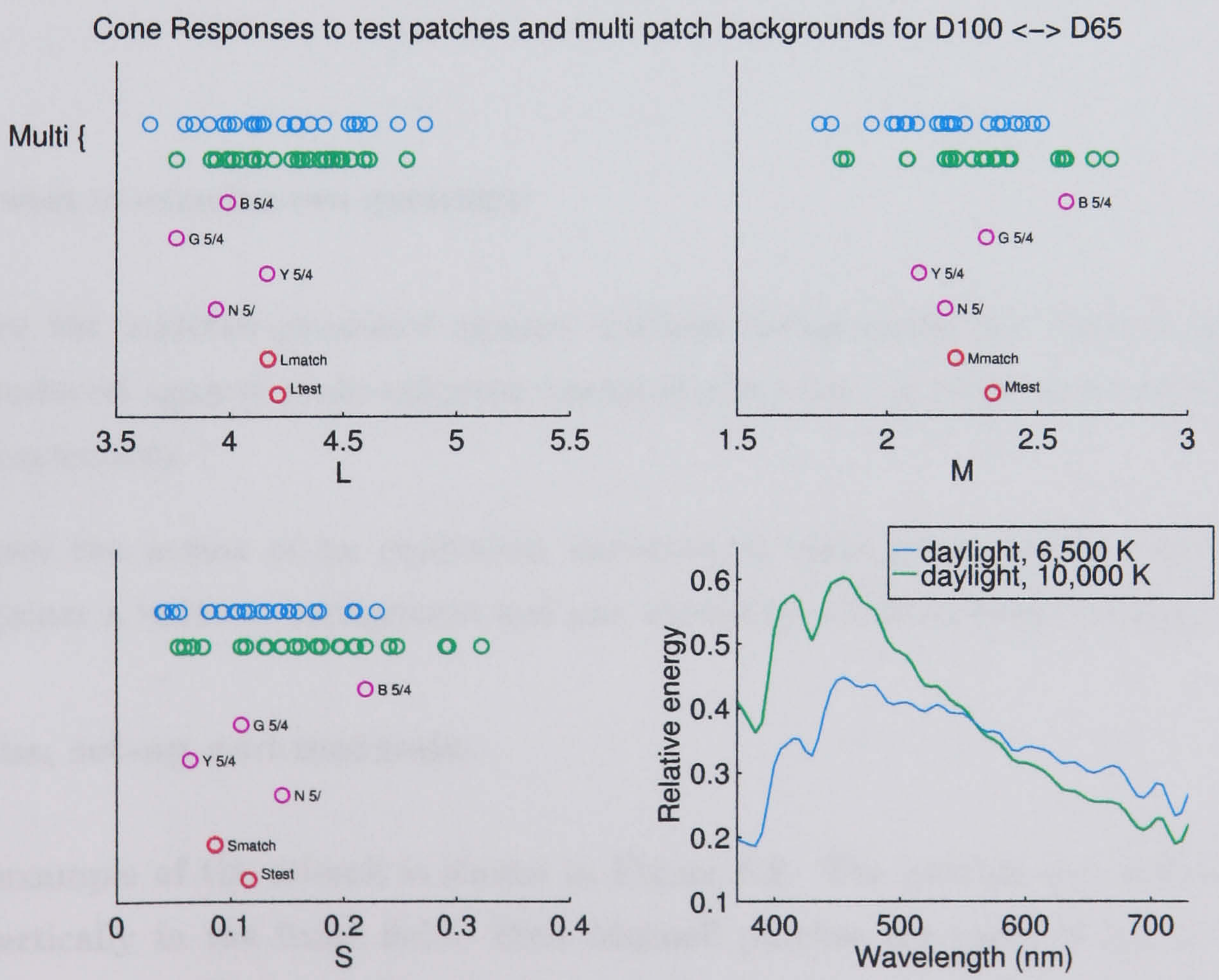

Figure 6.7: LMS plots of the test patches, equivalent backgrounds (match for E_{65}; test for E_{100}) and multi-patch backgrounds under D_{100} and D_{65}

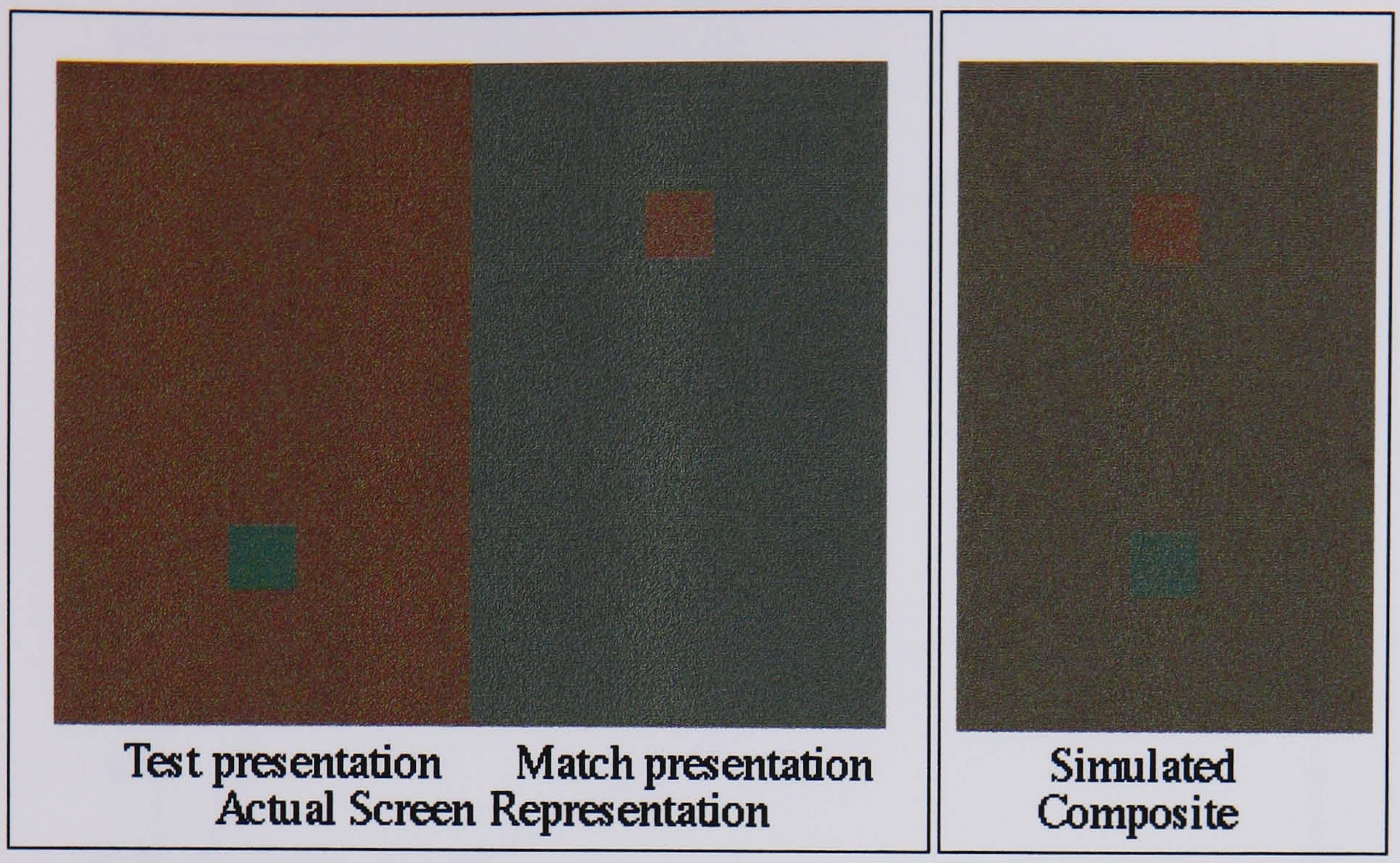

Figure 6.8: Example of a stimulus used in the patch against uniform equivalent background experiment

6.2 Patches on uniform backgrounds

Aims:

We wish to examine two questions:

1. Are the matches produced against uniform backgrounds the same as those produced against multi-coloured backgrounds with the same space-averaged chromaticity ?
2. Does the notion of an equivalent background make sense between matches against a uniform background and one consisting of many small patches?

Stimulus, set-up and methods:

An example of the stimuli is shown in Figure 6.8. The patches are centred 5^o apart vertically in the fused field. Four Munsell patches are used; N 5/, Y 5/4, G 5/4 and B 5/4. Each patch is presented under the simulated illuminant for the condition tested against the equivalent background for that illuminant i.e. for a test illuminant of D_{65} the test patches are calculated as being under D_{65} against

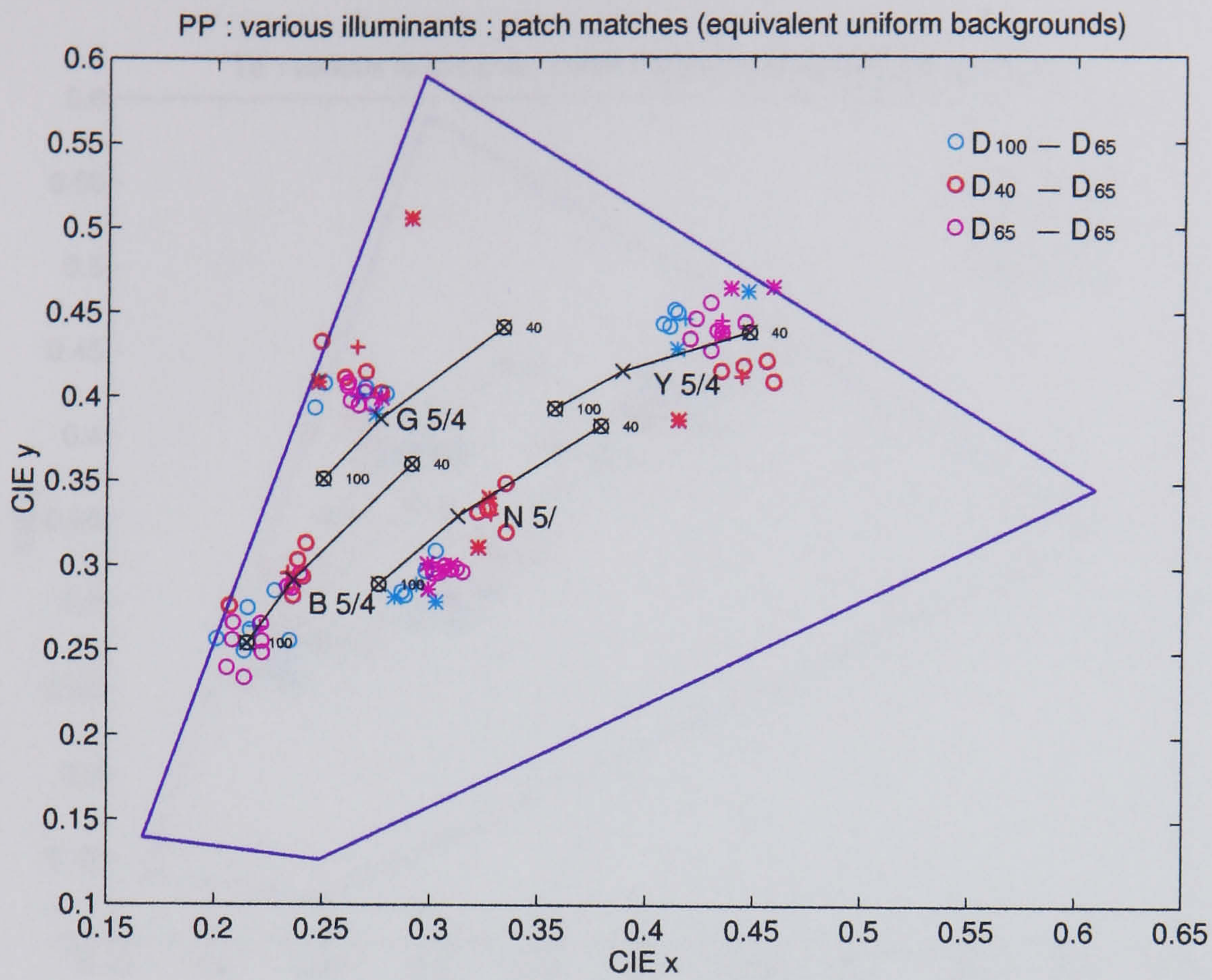

Figure 6.9: Patch matches on uniform equivalent backgrounds for observer PP

the equivalent of the multi-patch background under D_{65} (i.e. E_{65}). The observer performs a match by the adjustment of luminance, hue and saturation of the match patch.

Results:

Figure 6.9 shows the results for PP for three illuminant conditions. The corresponding plots for TE, JH and DB are shown in figures 6.10 – 6.12, respectively.

Analysis:

Table 6.3 lists match statistics for two of the observers, DB and TE (naive as to the purposes of the experiment). The results are indicative of the general trend of all observers. We can see that the variances in the matches remain approximately constant for each patch in each of the three conditions. Colour constancy performance can be assessed by comparing the matches made in the E_{65}^{test} : E_{65}^{match} with

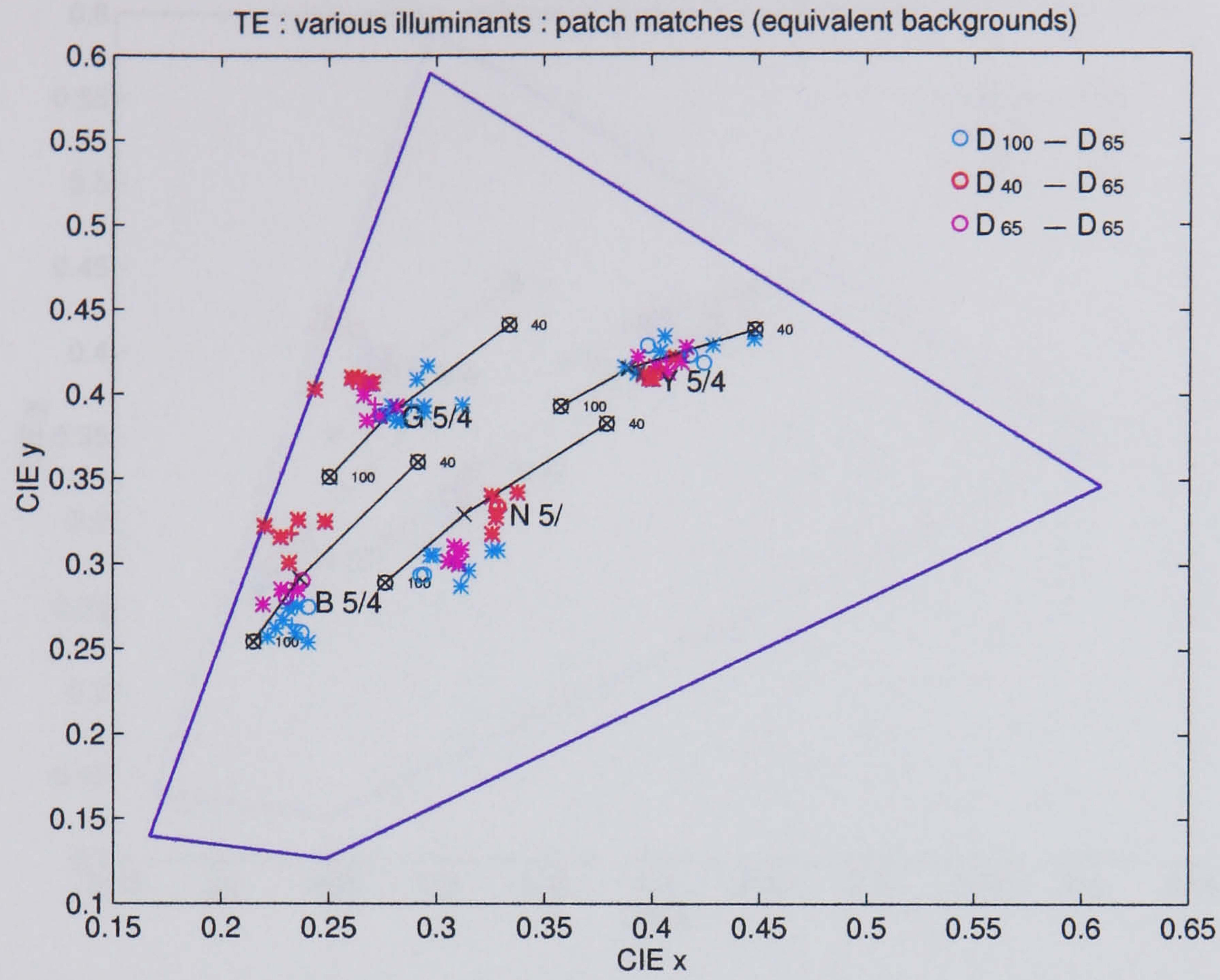

Figure 6.10: Patch matches on uniform equivalent backgrounds for observer TE.

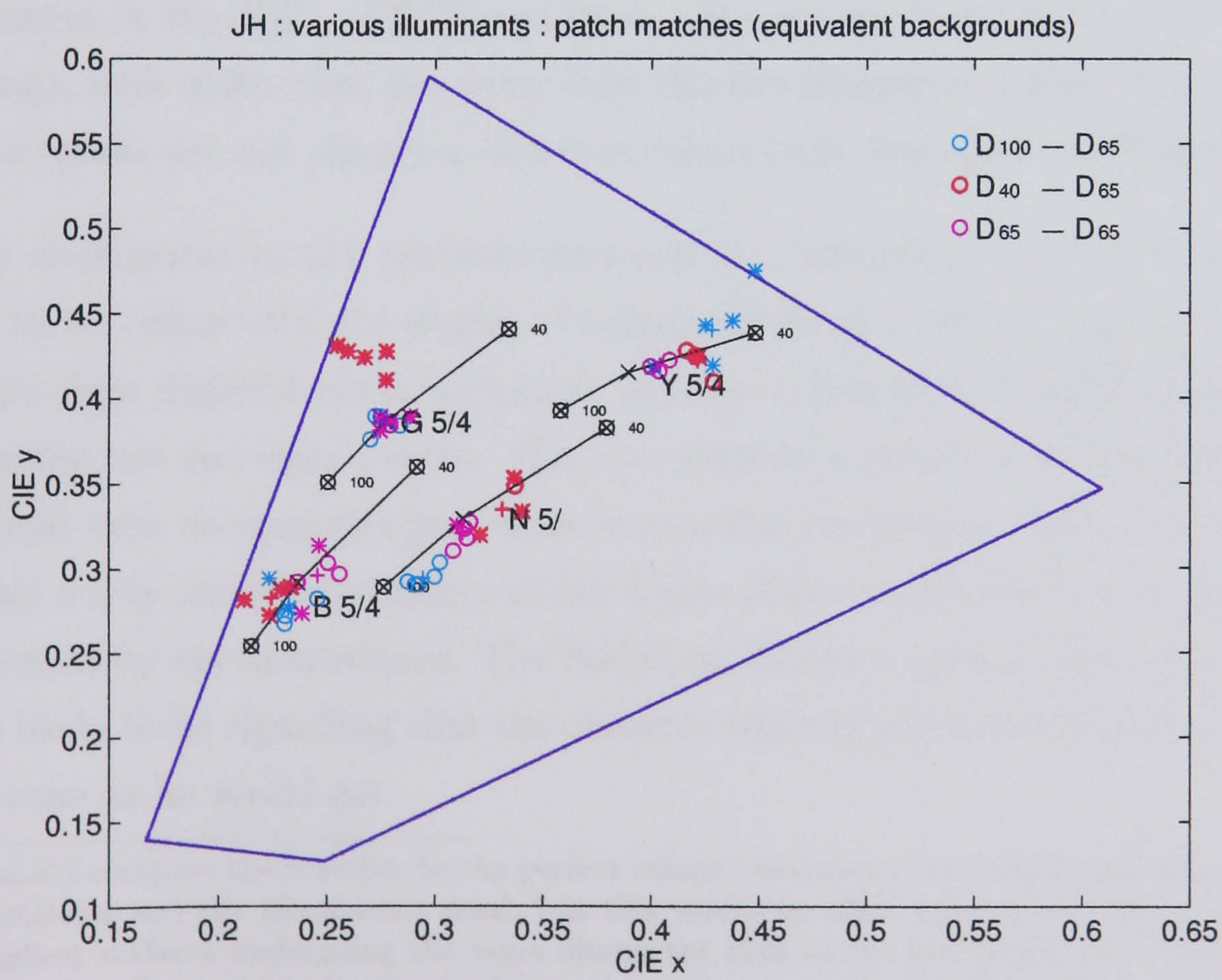

Figure 6.11: Patch matches on uniform equivalent backgrounds for observer JH.

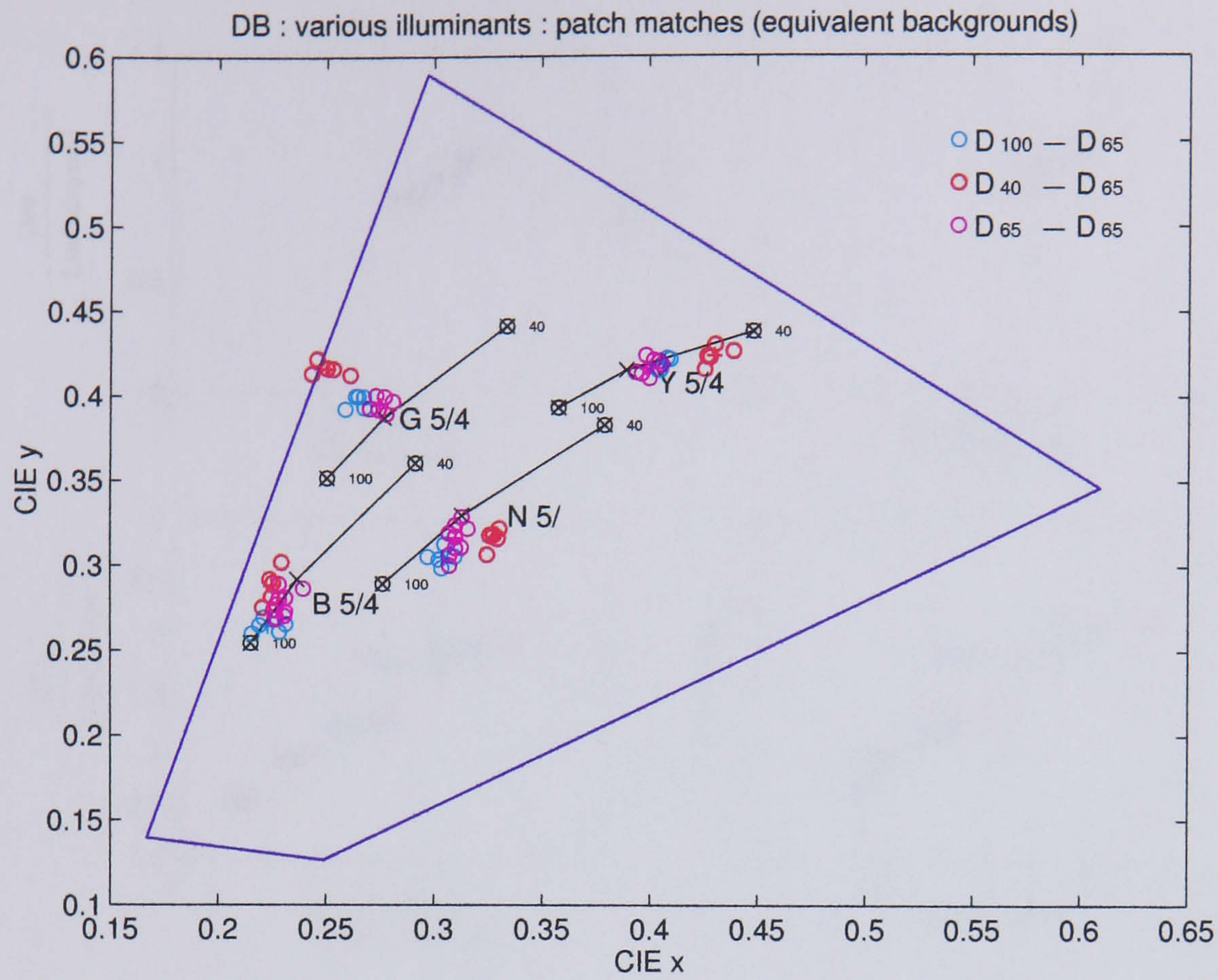

Figure 6.12: Patch matches on uniform equivalent backgrounds for observer DB.

the other two conditions[1]. Both observers show imperfect colour constancy with performance in the $E_{40}^{test} : E_{65}^{match}$ condition not only consistently low, but also, interestingly, with shifts that are away from the hue-saturation point. This suggests that the results are not simply a failure to adequately discount the illuminant.

The confidences in the matches reported are interesting in that they do not appear to correlate with the degree of colour constancy shown. More surprisingly observers show dissatisfaction with some matches where there is no illuminant shift between the test and match fields. This case requires a simple hue-saturation match and should have no complications from interocular interaction. DB's confidence results may not be entirely indicative of the degree of discriminability in the patches as was required by the instructions. The HIGH confidence response from this observer is more likely to be signalling that the observer was entirely satisfied that the match was as close as he would get.

[1]We could compare the matches to the perfect colour constancy point calculated from the Munsell reflectances and the illuminants used, but this would in effect assume that the backgrounds were constant surfaces undergoing the same illuminant shift as the test patch, which they are not necessarily. In other words, the E_{40} background is not necessarily exactly equivalent to the same surface as the E_{65} background, viewed under D_{40}.

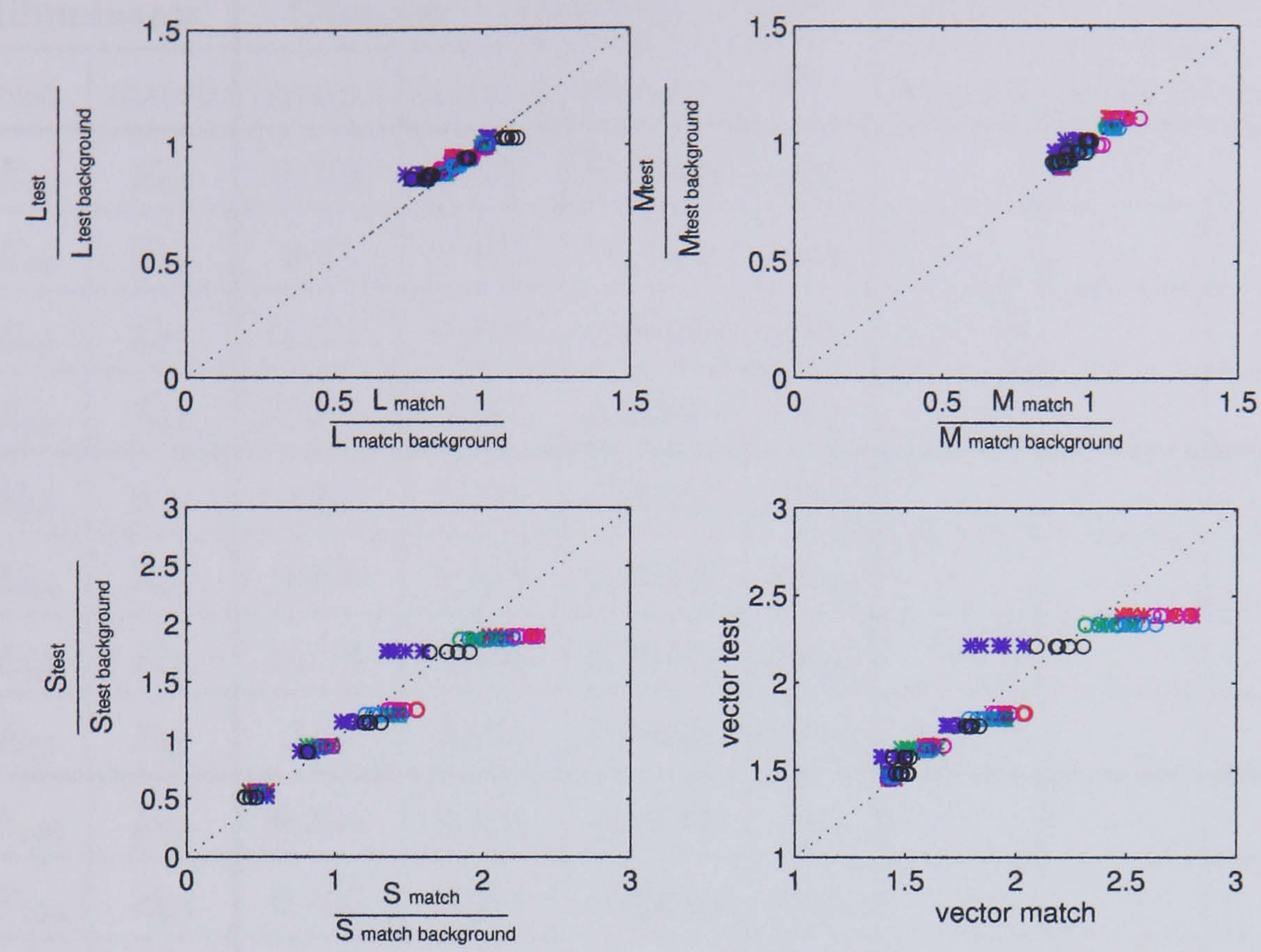

Figure 6.13: Cone ratio plot for patch matches on uniform equivalent backgrounds for observers DB and TE.

Shepherd and Whittle (Shepherd and Whittle 1995) have reported that the matches produced in their study (using a very similar experimental setup) are consistent with a cone ratio match hypothesis, i.e. the observer simply equates the cone ratios between patch and background in the test and match presentations. Figure 6.13 shows the cone ratio plots for observer DB and TE for all the matches listed in Table 6.3 (a key to the colours is presented in figure 6.2). If the results agreed with the cone ratio hypothesis then all of the match points would fall on the $x = y$ line of the three plots. In general the results presented here are in good agreement with those reported by Shepherd and Whittle. Although, the deviations from the $x = y$ line here seem quite large, this effect is due to the relatively small range of values tested. If the scales are expanded to match those of other studies then the agreement with the cone ratio hypothesis is clearer.

DB	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	E_{40}	E_{65}	0.328	0.315	0.00454	100_5	—	—	—	—
Y 5/4	E_{40}	E_{65}	0.43	0.423	0.0061	100_5	—	—	—	—
G 5/4	E_{40}	E_{65}	0.251	0.415	0.00599	100_5	—	—	—	—
B 5/4	E_{40}	E_{65}	0.225	0.287	0.00804	100_5	—	—	—	—
N 5/	E_{65}	E_{65}	0.31	0.316	0.00791	100_{10}	—	—	—	—
Y 5/4	E_{65}	E_{65}	0.399	0.416	0.00561	100_{10}	—	—	—	—
G 5/4	E_{65}	E_{65}	0.275	0.394	0.00475	100_{10}	—	—	—	—
B 5/4	E_{65}	E_{65}	0.23	0.276	0.00695	100_{10}	—	—	—	—
N 5/	E_{100}	E_{65}	0.304	0.304	0.00568	100_5	—	—	—	—
Y 5/4	E_{100}	E_{65}	0.407	0.419	0.00396	100_5	—	—	—	—
G 5/4	E_{100}	E_{65}	0.265	0.396	0.00444	100_5	—	—	—	—
B 5/4	E_{100}	E_{65}	0.223	0.263	0.00652	100_5	—	—	—	—

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	E_{40}	E_{65}	0.329	0.333	—	20_1	0.329	0.331	0.0103	80_4
Y 5/4	E_{40}	E_{65}	—	—	—	—	0.404	0.414	0.00842	100_5
G 5/4	E_{40}	E_{65}	—	—	—	—	0.262	0.406	0.00794	100_5
B 5/4	E_{40}	E_{65}	—	—	—	—	0.233	0.317	0.0124	100_5
N 5/	E_{65}	E_{65}	—	—	—	—	0.309	0.304	0.00451	100_5
Y 5/4	E_{65}	E_{65}	—	—	—	—	0.407	0.419	0.00907	100_5
G 5/4	E_{65}	E_{65}	—	—	—	—	0.272	0.393	0.00952	100_5
B 5/4	E_{65}	E_{65}	0.234	0.285	0.00612	40_2	0.228	0.281	0.00719	60_3
N 5/	E_{100}	E_{65}	0.294	0.293	0.000844	20_2	0.313	0.302	0.014	80_8
Y 5/4	E_{100}	E_{65}	0.414	0.423	0.0108	30_3	0.417	0.426	0.0228	70_7
G 5/4	E_{100}	E_{65}	0.275	0.387	—	10_1	0.29	0.394	0.0135	90_9
B 5/4	E_{100}	E_{65}	0.239	0.267	0.00766	20_2	0.232	0.263	0.00864	80_8

Table 6.3: Match statistics for matching against uniform backgrounds

Key for cone ratio plots

o TE – test under D100; match under D65.

o TE – test under D65; match under D65.

o TE – test under D40; match under D65.

o DB – test under D100; match under D65.

o DB – test under D65; match under D65.

o DB – test under D40; match under D65.

Figure 6.14: Colour key for figures 6.13 and 6.20.

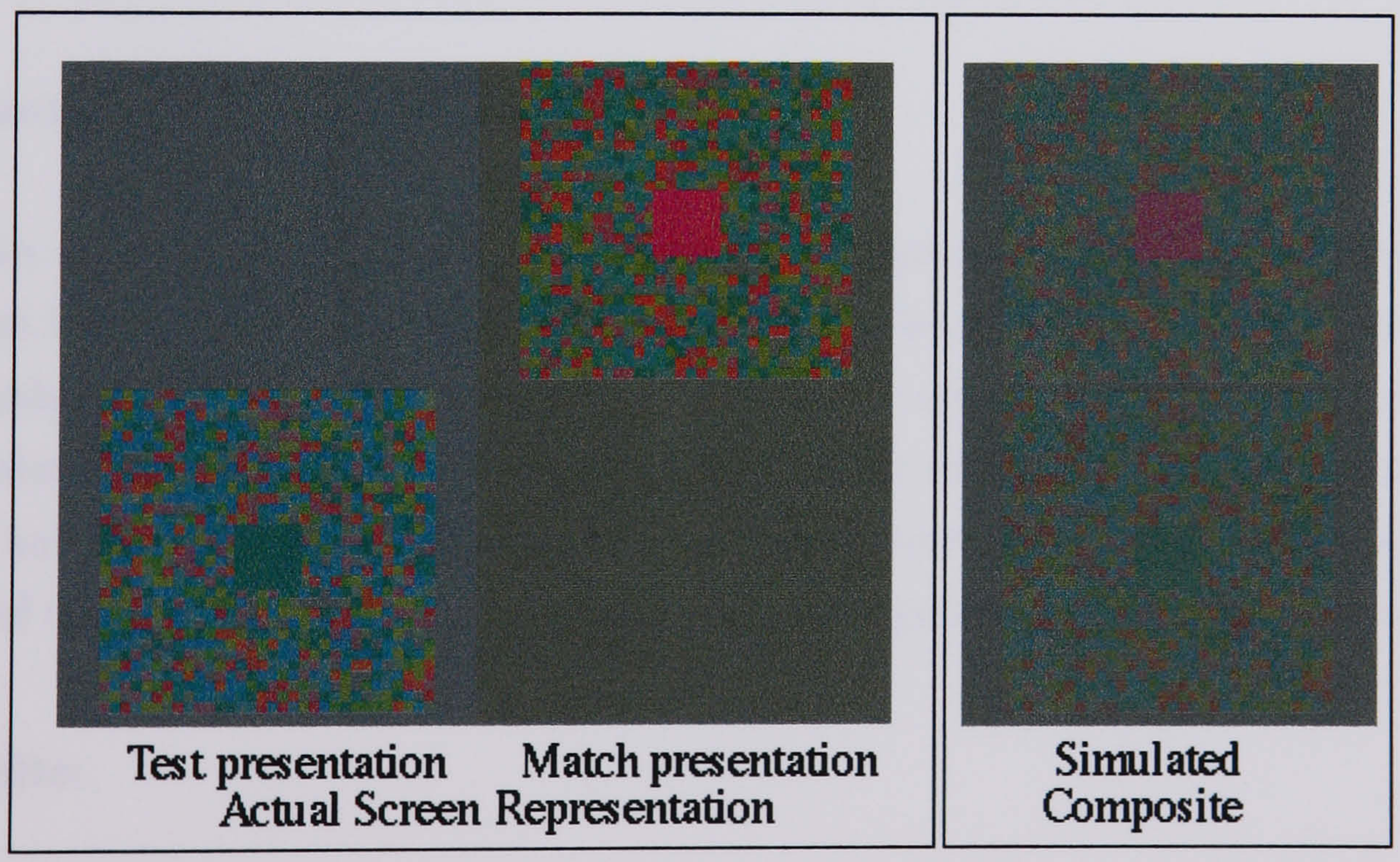

Figure 6.15: Example of a stimulus used in the patch against multi patch background experiment

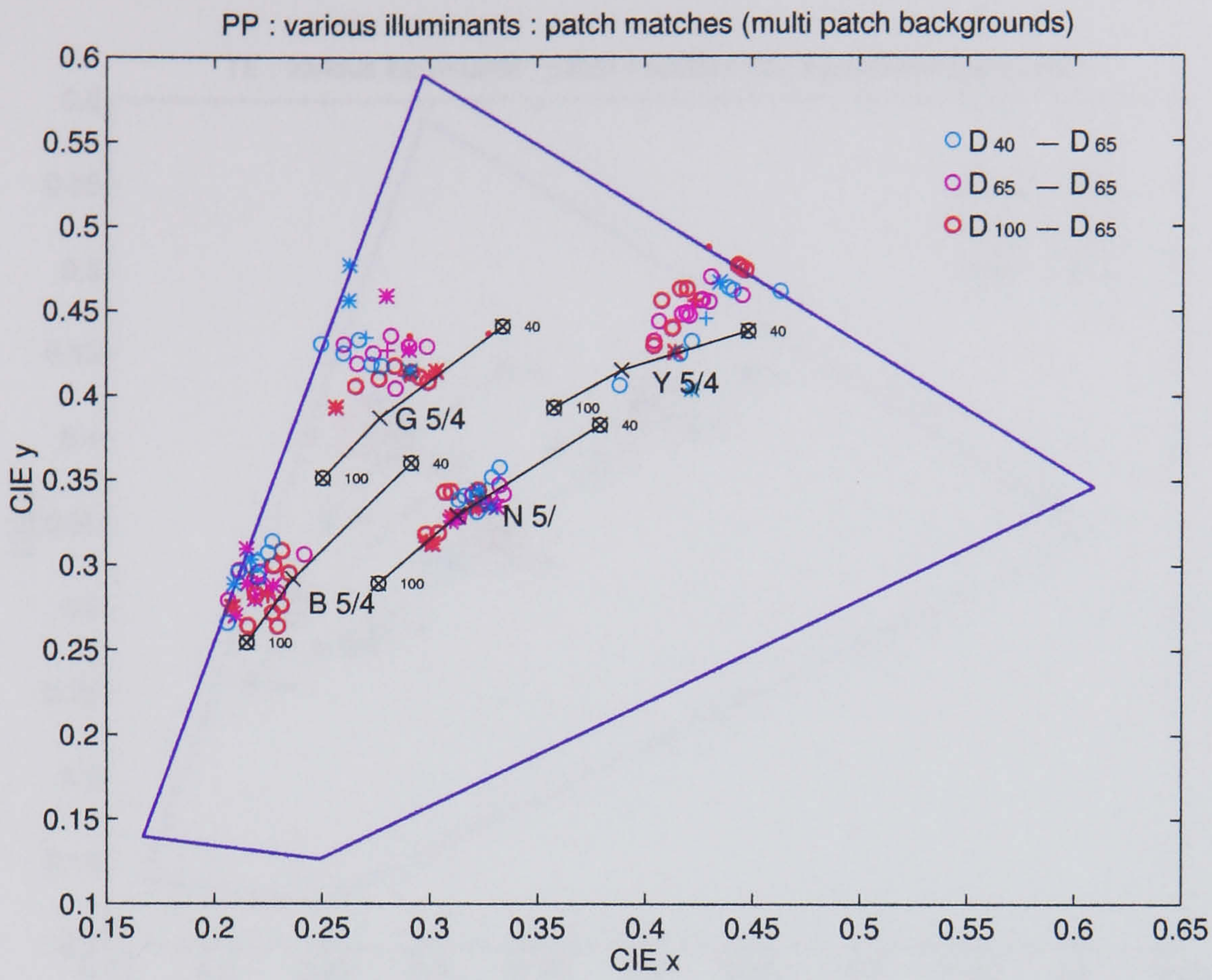

Figure 6.16: Patch matches on multi-patch backgrounds for observer PP.

6.3 Patches on multi-coloured backgrounds: effects of chromatic contrast in the background

Stimulus, set-up and methods:

An example of the stimuli used is shown in Figure 6.15. All stimulus features except for the multi-patch background are exactly the same as in the uniform backgrounds experiment. The multi-patch backgrounds are generated for each trial by repeated random use of the 24 Munsell patches. The multi-patch backgrounds cover only half of each vertical half-field to avoid interference with binocular fusion. The rest of the field is filled with a simulation of N 4/ under the particular illuminant.

Results:

Figure 6.16 shows the results for PP for three illuminant conditions. The corresponding plots for TE, JH and DB are shown in figures 6.17, 6.18 and 6.19 respectively.

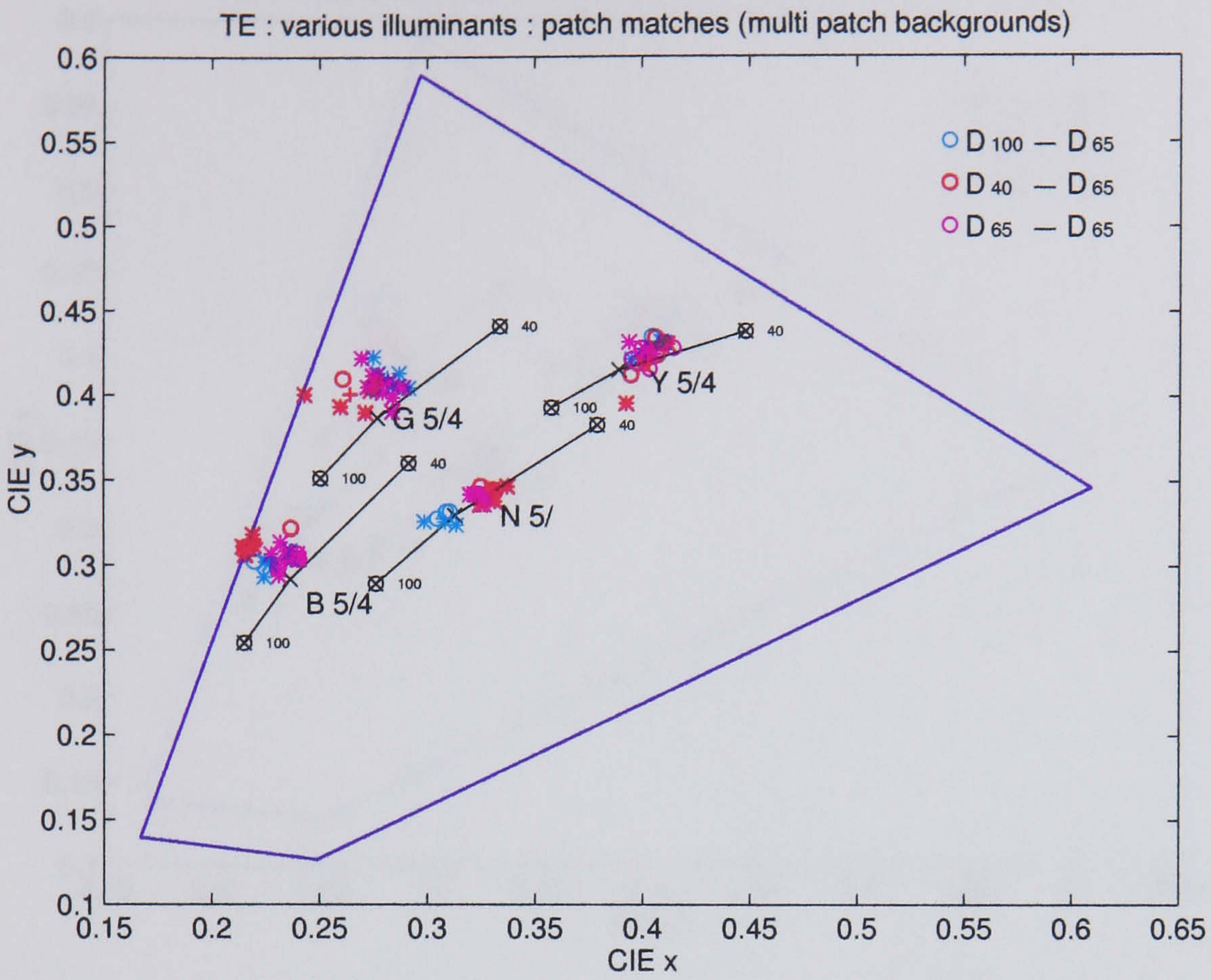

Figure 6.17: Patch matches on multi-patch backgrounds for observer TE.

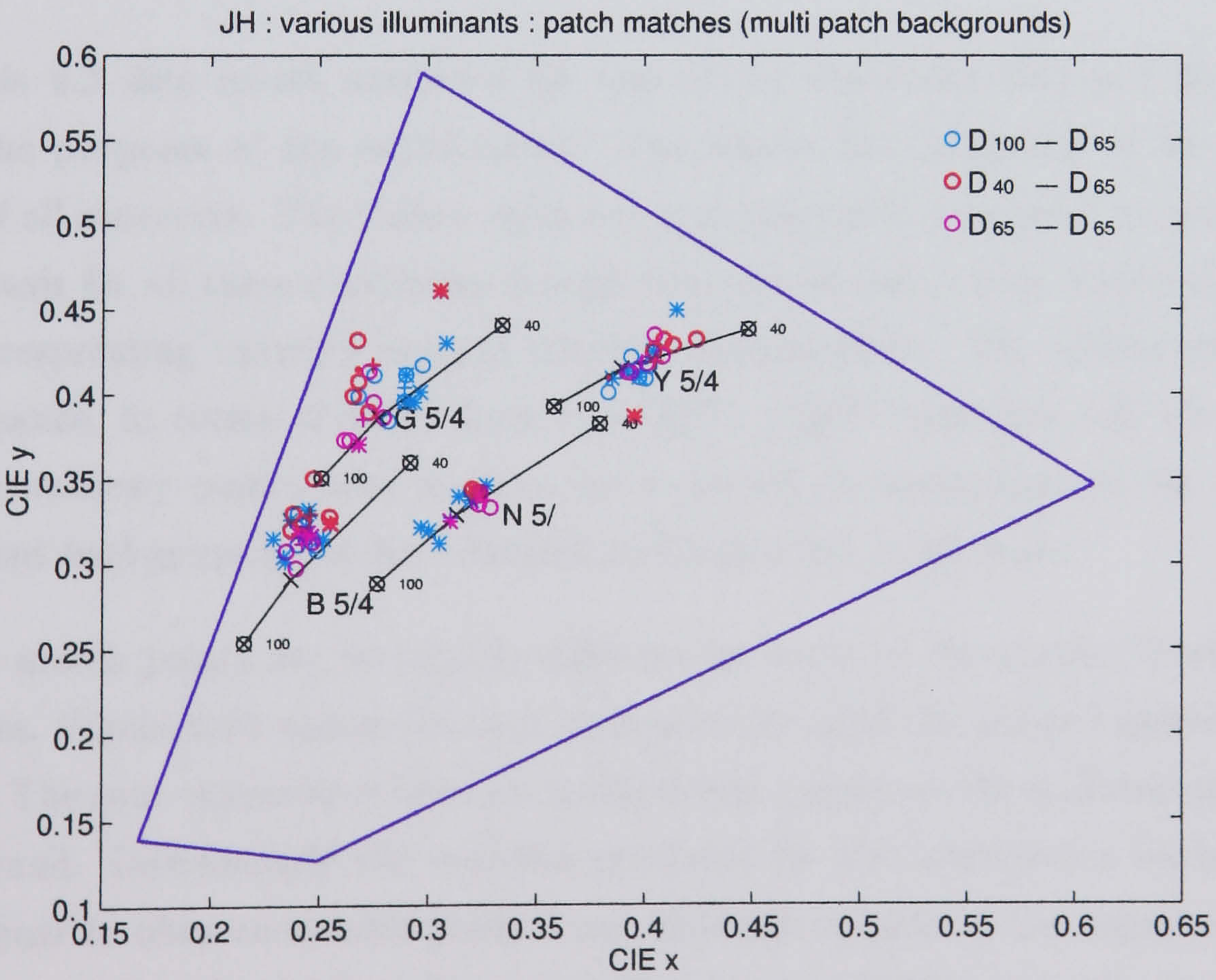

Figure 6.18: Patch matches on multi-patch backgrounds for observer JH.

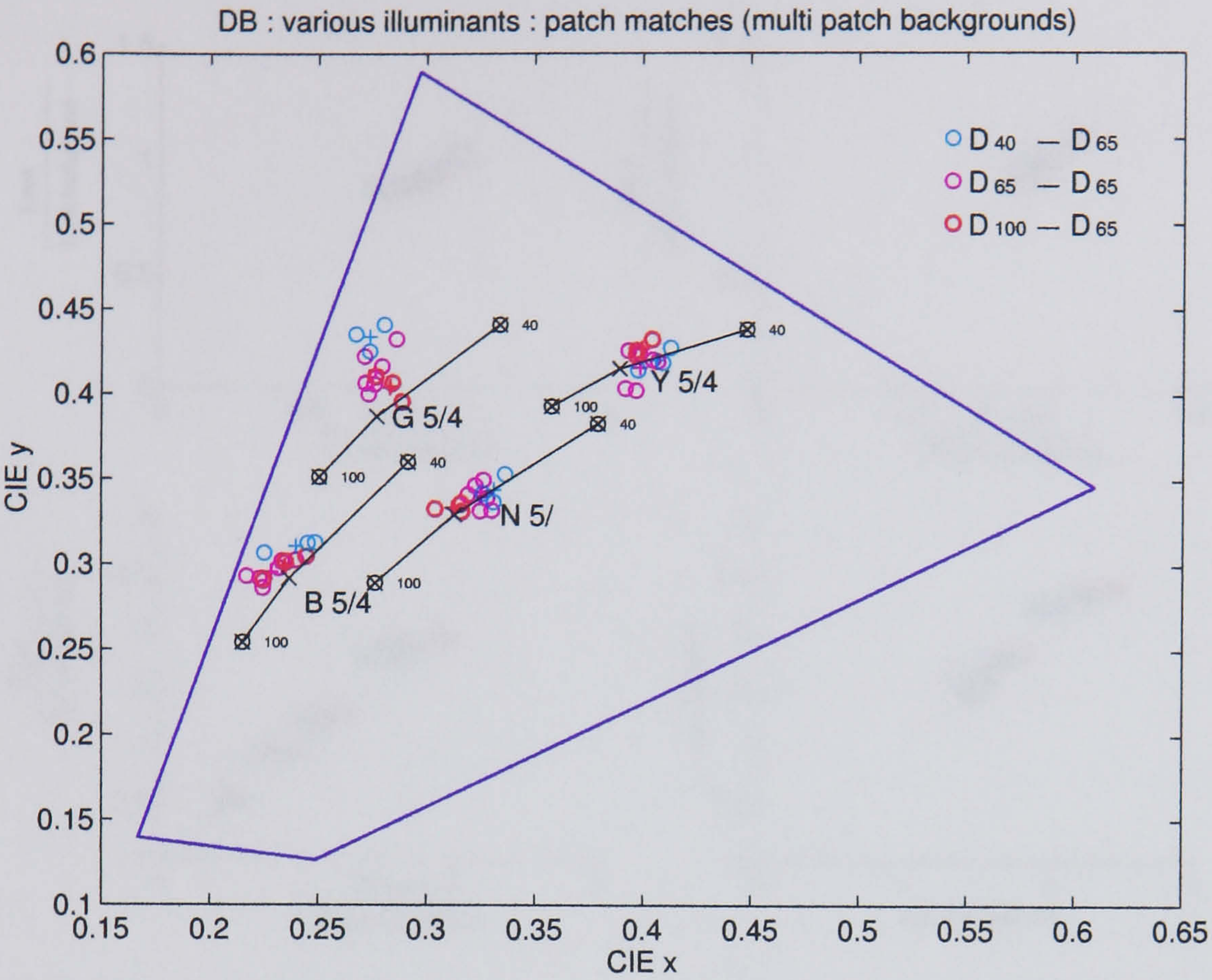

Figure 6.19: Patch matches on multi-patch backgrounds for observer DB.

Analysis:

Table 6.3 lists match statistics for two of the observers, DB and TE (naive as to the purposes of the experiment). The results are indicative of the general trend of all observers. The match variances are reasonably consistent on a patch by patch basis for all three conditions though they do appear to be a little larger than the corresponding matches against uniform backgrounds. The colour constancy performance, in terms of shifts from the D_{65}^{test} : D_{65}^{match} matches and the perfect colour constancy points does seem to be improved in comparison to the uniform equivalent backgrounds results – though still imperfect in all cases.

The match points are noticeably different for some of the patches between the two cases. Figure 6.20 shows the cone ratio plots for all of the points reported in Table 6.4. The cone values have been normalised with respect to the uniform equivalent background. Interestingly the matches produced for the multi-patch backgrounds also appear to obey cone ratio predictions, although in general the degree of colour constancy produced is higher. There is still some uncertainty as to the equivalence of the uniform background to the multi-patch background. The cone ratio results

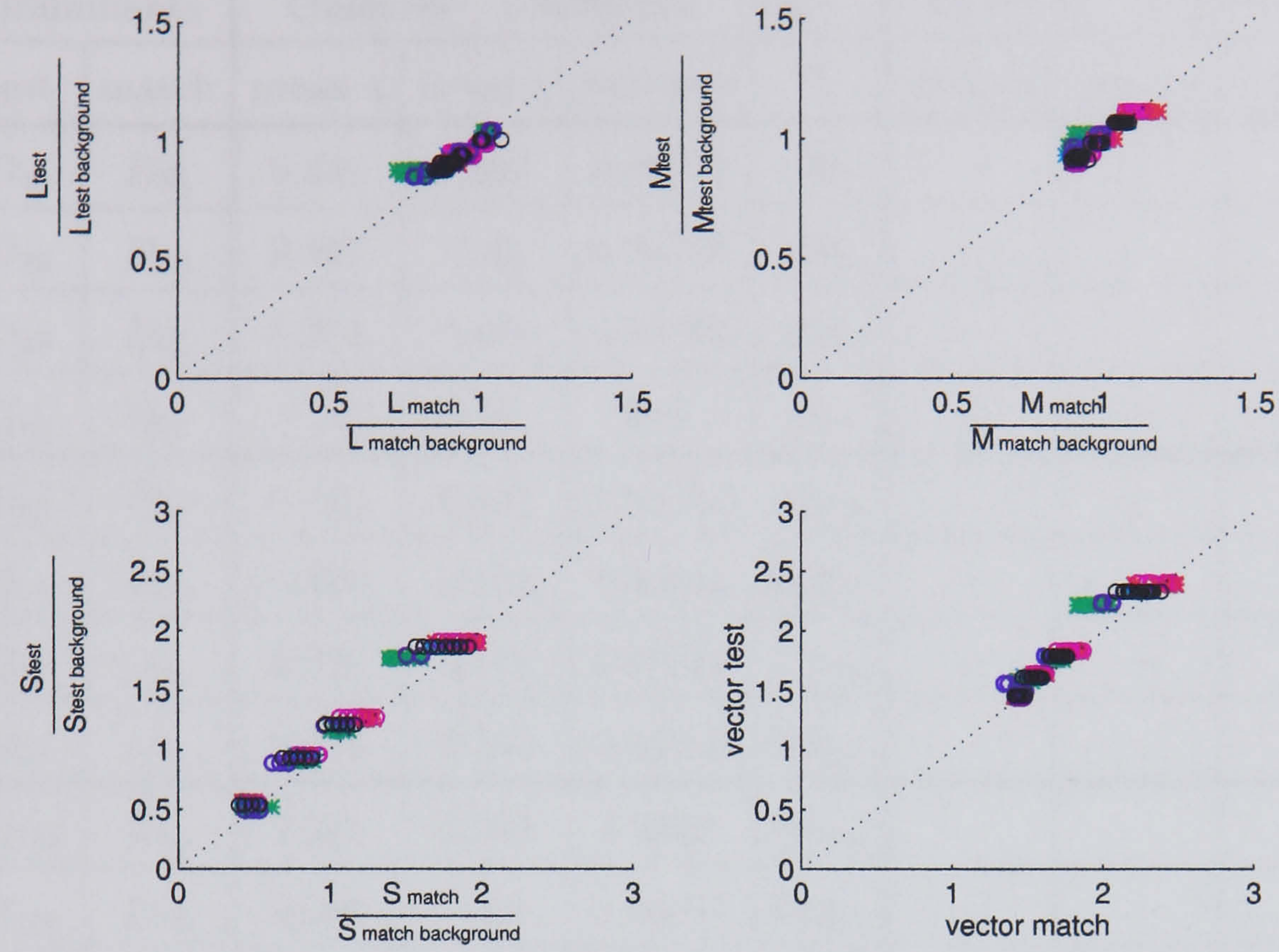

Figure 6.20: Cone ratio plot for patch matches on multi-patch backgrounds for observers TE and DB.

suggest that either the uniform background is a suitable equivalent or that the normalisation factor need only be consistently calculated for each illuminant case and that the exact calculation is unimportant as long as it captures a good estimation of the overall excitation in each cone class.

6.4 Analysis of matching against uniform and multi-patch backgrounds

This section compares the results obtained in this chapter and examines them in more detail in the context of the proposed two stage model.

The main aim of experiments in this chapter is to measure colour constancy performance for flat colour patches in uniform and multi-patch backgrounds. It was proposed in chapter 3 that early visual processing may take the form of a two stage amplifier system.

The first stage can be formally described with:

DB	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.331	0.343	0.00771	100_3	—	—	—	—
Y 5/4	D_{40}	D_{65}	0.407	0.42	0.00778	100_3	—	—	—	—
G 5/4	D_{40}	D_{65}	0.274	0.433	0.00832	100_3	—	—	—	—
B 5/4	D_{40}	D_{65}	0.24	0.31	0.01	100_3	—	—	—	—
N 5/	D_{65}	D_{65}	0.322	0.342	0.00634	100_{18}	—	—	—	—
Y 5/4	D_{65}	D_{65}	0.399	0.424	0.00844	100_{18}	—	—	—	—
G 5/4	D_{65}	D_{65}	0.282	0.411	0.00931	100_{18}	—	—	—	—
B 5/4	D_{65}	D_{65}	0.235	0.299	0.00843	100_{18}	—	—	—	—
N 5/	D_{100}	D_{65}	0.315	0.333	0.0058	100_{15}	—	—	—	—
Y 5/4	D_{100}	D_{65}	0.396	0.42	0.00827	100_{15}	—	—	—	—
G 5/4	D_{100}	D_{65}	0.281	0.403	0.0112	100_{15}	—	—	—	—
B 5/4	D_{100}	D_{65}	0.232	0.296	0.00711	100_{15}	—	—	—	—

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N 5/	D_{40}	D_{65}	0.328	0.345	0.00298	33.33_2	0.331	0.341	0.00572	66.67_4
Y 5/4	D_{40}	D_{65}	0.403	0.423	0.00983	50_3	0.402	0.414	0.0146	50_3
G 5/4	D_{40}	D_{65}	0.261	0.409	—	16.67_1	0.265	0.398	0.013	83.33_5
B 5/4	D_{40}	D_{65}	0.237	0.321	—	16.67_1	0.217	0.31	0.00375	83.33_5
N 5/	D_{65}	D_{65}	0.323	0.342	0.000876	16.67_2	0.324	0.34	0.00246	83.33_{10}
Y 5/4	D_{65}	D_{65}	0.404	0.425	0.00794	66.67_8	0.401	0.427	0.0041	33.33_4
G 5/4	D_{65}	D_{65}	—	—	—	—	0.279	0.406	0.00841	100_{12}
B 5/4	D_{65}	D_{65}	0.237	0.302	0.00522	25_3	0.234	0.303	0.00662	75_9
N 5/	D_{100}	D_{65}	0.308	0.33	0.00295	50_3	0.307	0.324	0.00555	50_3
Y 5/4	D_{100}	D_{65}	0.4	0.428	0.00825	33.33_2	0.403	0.426	0.00887	66.67_4
G 5/4	D_{100}	D_{65}	0.275	0.404	—	16.67_1	0.283	0.411	0.00818	83.33_5
B 5/4	D_{100}	D_{65}	0.22	0.301	—	16.67_1	0.23	0.301	0.00713	83.33_5

Table 6.4: Match statistics for matching against multi-patch backgrounds

$$F_1 = G_1.I + O_1 \tag{6.1}$$

and in matrix form as

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{12} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} + \begin{bmatrix} O_L \\ O_M \\ O_S \end{bmatrix} \tag{6.2}$$

the matrix form of G_1 caters for the possible interaction between cone classes – if no interaction is required then the matrix will take a diagonal form. It should be made clear at this point that we are not suggesting a single amplifier, but several covering regions of the visual field, the extent of which we are leaving as an open question. The a_{ij} and O_k entries are assumed to be functions of local and global cone excitations.

There are two ways in which we can construct a cone opponent representation at this point. We could convert the raw cone signals (obviating the need for the first stage) or we could convert the processed signals from the first stage and feed them into a second. The distinction is not critical at this point as the two scenarios produce qualitatively the same results (i.e. a linear transformation can convert one to the other). Where the distinction may become very important is when gain limits are imposed or noise is introduced into the system. We will consider the model version in which the output of one stage feeds directly into another as this allows examination of the effects of gain and noise on the system as a whole.

The second stage is similar to the first but operates on a transformed version of the output of the first stage. The transform we will use is the maximal decorrelation opponent representation derived by Buchsbaum and Gottschalk (Buchsbaum and Gottschalk 1983).

Formally

$$I_2 = T.F_1 \tag{6.3}$$

which in matrix form is

$$\begin{bmatrix} lum \\ r-g \\ b-y \end{bmatrix} = \begin{bmatrix} 0.887 & 0.461 & 0.0009 \\ -0.46 & 0.88 & 0.01 \\ 0.004 & -0.01 & 0.99 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \tag{6.4}$$

The second stage gain and offset transform may be expressed as

$$F_2 = G_2.I_2 + O_2 \tag{6.5}$$

by substitution we obtain

$$F_2 = G_2.T.(G_1.I + O_1) + O_2 \tag{6.6}$$

where G_2 and O_2 are functions of I_2 and/or I locally and globally.

In summary, the gain and offset of the first stage are functions of local and global cone excitations. (Each cone class is processed independently but may have its offset or gain adjusted by signals produced from the other two cone types). The output from the first stage is assumed to be representative of the surface contrast within the viewed scene. The second stage gain and offsets are functions of the local and global contrast in each of the opponent channels. (Again each channel is processed independently but the gain and offsets used may be affected by the other two channels). The output from the second stage is assumed to produce a normalised representation of *opponent channel* contrast. The transformations used are consistent with a processing system that endeavours to devote as much of the available signal bandwidth to the representation of scene contrast.

The experiments in this chapter compare matches made against uniform backgrounds and patches made against multi-patch backgrounds. There are still too many unknown parameters to allow the model to be applied to the stimuli. We can solve this problem by making some assumptions and assessing the results. The following assumptions are used:

- Due to the simple nature of the stimuli we assume that all of the gain and offset parameters are set by global measures.
- The first stage gain parameters are set independently for the test and match presentations.
- The second stage gain and offsets are calculated from the combined contrast results for the test and match fields.
- There is no restriction on gain or offset or any noise in the system.

6.4.1 Uniform backgrounds

We can now consider the state of the system for each of the stimuli. Figure 6.21 presents six plots that show various stages within the model. The data in the plots corresponds with a haploscopic match experiment of G 5/4 against N 5/ under D_{40} (test side) and D_{65} (match side) i.e. patches against uniform backgrounds. The perfect colour constancy match has been used for the match side. The first plot shows the CIE xy space representation of the points (1 = G 5/4 under D_{40}, 2 = N 5/ under D_{40}, 3 = G 5/4 under D_{65} and 4 = N 5/ under D_{65}). The second plot shows the MacLeod-Boynton space representation of the same points, highlighting the absolute excitations elicited in the cones. The third plot shows the output of the first amplifier stage, where we have assumed a Von Kries style transformation in which each cone class is normalised by the uniform field excitation. The cone values are divided by the cone excitations of the corresponding uniform field, i.e. the test side cone excitations are divided by the cone excitations produced by the uniform test background, and similarly for the match side. This normalisation strategy causes the points for the two uniform backgrounds to super-impose at x ($\frac{L-M}{L+M}$)= 0, y ($\frac{S}{L+M}$) = 0.5. We now have a contrast based representation of the input signals. The opponent transformation is now applied producing the fourth plot. (The achromatic channel is not considered at this point). The fifth plot shows the result of shifting the opponent signals to produce a mean of zero in each channel. This transform is analogous to the offset operation and may produce chromatic induction type changes in the perceived colours. The final plot shows the global output of the proposed two stage model. The final plot is produced by scaling each of the axes so that all of the signals present fit in the interval -1 to 1 (channel based gain adjustment).

This particular implementation is one of many possibilities that would fit the morphology proposed. In an alternative realisation the shift of the opponent space could be applied after the expansion. The example given is mathematically equivalent to the described implementation. The gain and offset values obtained would differ, but equivalent processing could be performed. This example highlights the fact that implementational details will change the quantitative results obtained but should not alter qualitative predictions significantly.

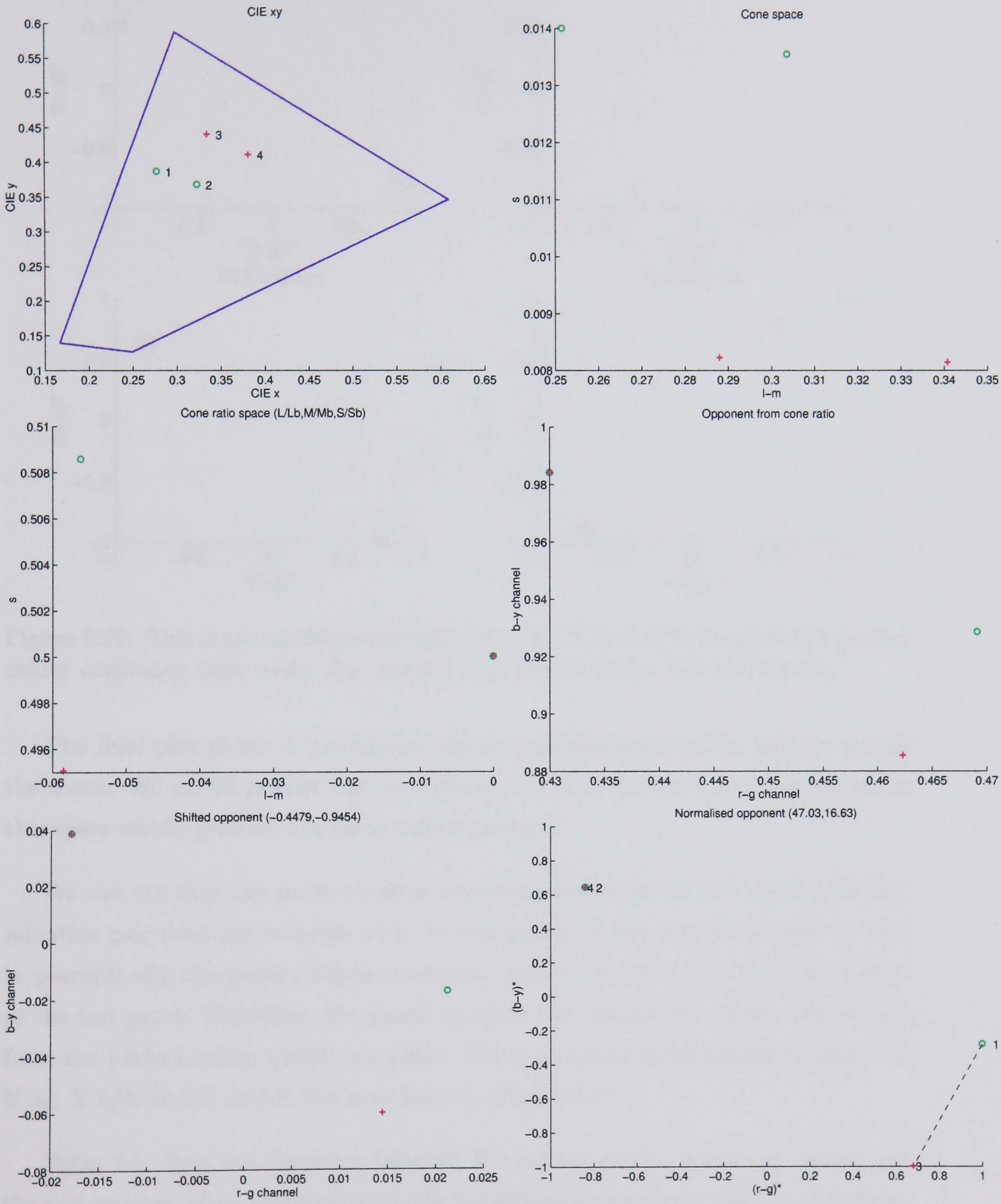

Figure 6.21: Details of two stage model prediction for G 5/4 perfect colour constancy (test under D_{40} against E_{40}, match under D_{65} against E_{40})

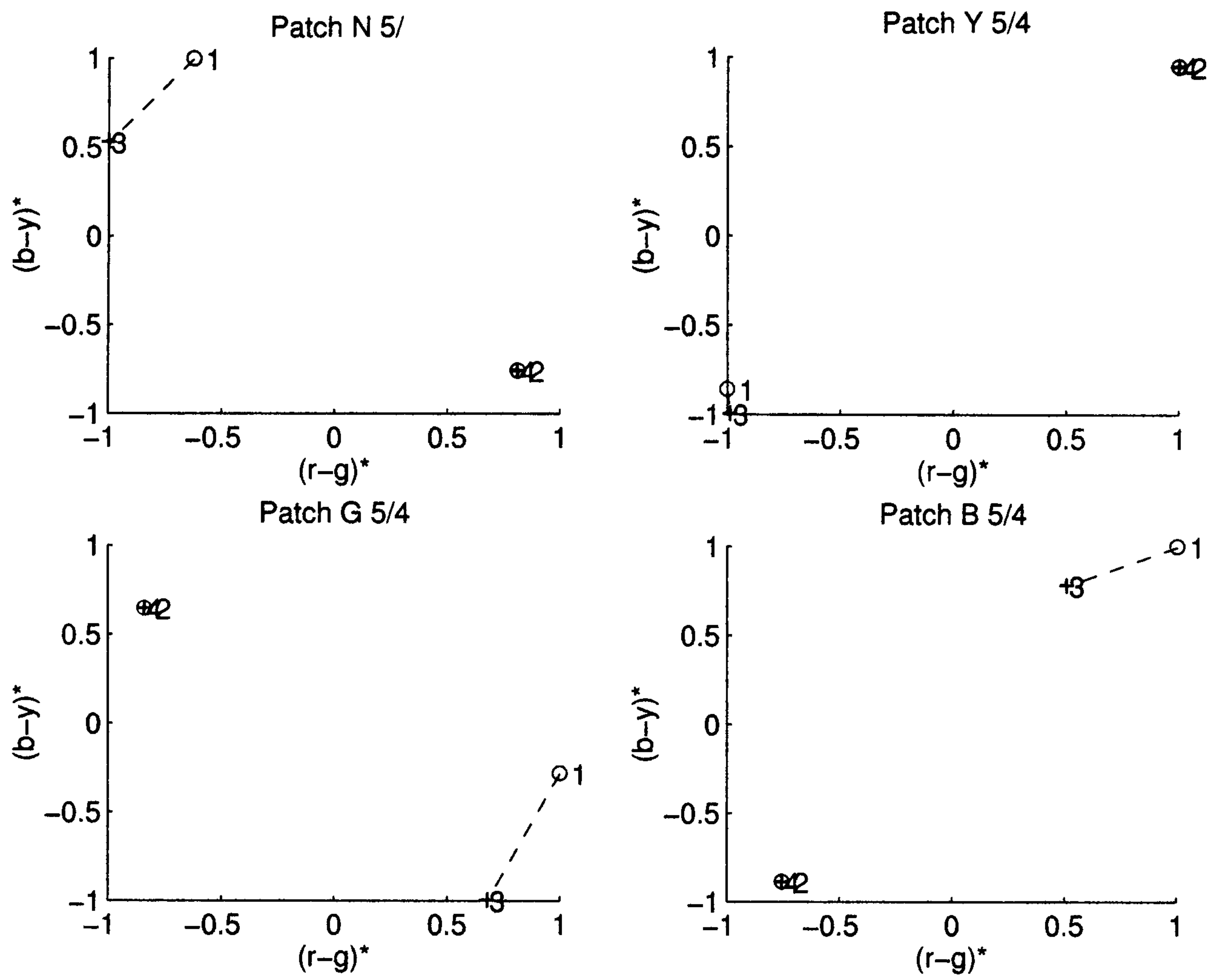

Figure 6.22: Two stage model predictions for N 5/, Y 5/4, G 5/4 and B 5/4 perfect colour constancy (test under D_{40} against E_{40}, match under D_{65} against E_{40})

The final plot shows a 'perceptual space' representation of the colours within the scene. We would predict that two points that map to the same coordinates in this space would produce the same colour percept.

We can see that the perfect colour constancy match for G 5/4 under this illumination pair does not coincide with the test match in the perceptual space – that is, perceptually, the perfect colour constancy point does not appear the same colour as the test patch. Therefore, the model predicts that the chosen match will deviate from the perfect colour constancy point. The perceptual space representations for N 5/, Y 5/4, G 5/4 and B 5/4 are shown in Figure 6.22.

Table 6.5 lists the distances between the perfect colour constancy points and the test patches when represented in the perceptual space along with the Euclidean distance of the mean of the matches found in the experimentation for observer TE. If the general transformation procedure is valid we can make the following prediction: all of the test points will vary from the perfect colour constancy points, with the

TE†	error in perceptual space			error in CIE space			error
patch	x	y	$\sqrt{x^2+y^2}$	x	y	$\sqrt{x^2+y^2}$	ratio
N 5/	0.3732	0.4749	0.604	0.0143	0.0199	0.0245	24.6
Y 5/4	0.01369	0.1384	0.1391	0.00273	0.00403	0.00487	28.6
G 5/4	0.32	0.716	0.7843	0.0251	0.0115	0.0251	31.3
B 5/4	0.4917	0.2183	0.5379	0.0116	0.0163	0.0200	27.0

† the experimental results presented here correspond to those shown in Figure 5.14. This set was chosen due to the low variance in the matches.

Table 6.5: predicted and measured errors for observer TE matching patches against uniform backgrounds (test under D_{40}, match under D_{65})

largest shift for G 5/4, next largest for N 5/, then B 5/4 followed by good colour constancy for Y 5/4. The errors in matching in CIE xy space do correlate with the errors predicted from the perceptual space. The error ratio column shows that the Euclidean errors are in reasonable agreement with, on average, the error in the perceptual space being 27.9 times that in the CIE xy plane. Other data sets show the same general agreement for a variety of illumination conditions but there are some exceptions which may be due to matching error.

This raises an interesting question: Are the observers minimising the error in the proposed perceptual space? We can run the model simulation using the observers' match points instead of the perfect colour constancy points. Table 6.6 lists the perceptual space errors resulting from using the mean of the observer matches for the data set used in table 6.5.

We can see that in most cases the colour constancy prediction represents more of a minimisation of error in the perceptual space. There are two possibilities: the model could be completely wrong or some of the assumptions are incorrect. Since qualitatively the errors in the perceptual space predict the errors in the matches it is likely that the general set of transformations used in the model are valid – the assumptions used appear to produce erroneous quantitative results.

One possibility is that 'perceptually' the errors in Table 6.5 are indiscriminable i.e. the noise within the system coupled with the high gains means that some of the values are effectively the same. If we reduce the maximum gain of the system then

TE match data, patches against uniform backgrounds (D_{40} – D_{65})							
Patch	r-g gain	b-y gain	r-g offset	b-y offset	$\sqrt{\delta x^2 + \delta y^2}$	b-y error	r-g error
N 5/	34.49	11.64	-0.415	-1.05	1.072	0.9651	0.4662
Y 5/4	21.01	3.968	-0.3867	-0.7529	0.2439	0.1784	0.1662
G 5/4	34.18	20.16	-0.4505	-0.9346	0.5971	0.597	0.006593
B 5/4	19.85	2.433	-0.4723	-1.324	0.4725	0.3201	0.3476

perfect colour constancy, patches against uniform backgrounds (D_{40} – D_{65})							
Patch	r-g gain	b-y gain	r-g offset	b-y offset	$\sqrt{\delta x^2 + \delta y^2}$	b-y error	r-g error
N 5/	41.22	8.479	-0.4103	-1.074	0.604	0.3732	0.4749
Y 5/4	24.01	3.997	-0.3886	-0.7511	0.1391	0.01369	0.1384
G 5/4	47.03	16.63	-0.4479	-0.9454	0.7843	0.32	0.716
B 5/4	16.49	2.229	-0.4757	-1.384	0.5379	0.4917	0.2183

Table 6.6: Two stage model predictions of errors for perfect colour constancy and observer data

the differences in the errors also reduce. Unfortunately signal noise at this level of the system is very difficult to predict and hence so is maximum gain.

6.4.2 Multi-patch backgrounds

We can perform the same analysis for the multi-patch backgrounds. The only difference between the processing for multi-patch and uniform backgrounds is in the cone space normalisation factor. In the multi-patch case, the cone inputs are normalised by the calculated effective uniform background (see Table 6.2).

The normalised space representations of the perfect colour constancy matches against multi-patch backgrounds are shown in Figure 6.23. The red crosses show the locations of stimuli presented to the left eye (i.e. test and background under D_{40}). The green circles show the same information for the right eye (i.e. match under D_{65}). The test and match patches have been highlighted with purple circles and are joined together with a dotted line that shows the Euclidean error in the perceptual space between the test patch and the perfect colour constancy point.

perfect colour constancy, patches against uniform backgrounds ($D_{40} - D_{65}$)							
Patch	r-g gain	b-y gain	r-g offset	b-y offset	$\sqrt{\delta x^2 + \delta y^2}$	b-y error	r-g error
N 5/	4.412	0.798	-0.4159	-1.375	0.05995	0.03994	0.0447
Y 5/4	4.444	0.7828	-0.4142	-1.35	0.02724	0.002534	0.02712
G 5/4	4.357	0.7919	-0.4187	-1.365	0.04518	0.02964	0.03409
B 5/4	4.318	0.8132	-0.4208	-1.398	0.1514	0.1287	0.07964

Table 6.7: Two stage model predictions of errors for perfect colour constancy matches on multi-patch backgrounds

The gains and offsets of the opponent channels resulting from the simulation are listed in Table 6.7. The errors between test and match in the perceptual space are also listed. It can easily be seen that the patches in the background essentially 'lock' the gain and offset of the proposed second stage mechanism. The match errors predicted are small compared to those against uniform fields. Colour constancy performance did increase in our experiments but the inter observer variability is too high to allow an error comparison of the type used for the uniform backgrounds.

It is possible that the multi-patch background simulation is more affected by the assumptions made in this very simple simulation. In the uniform field case the spatial integration and weighting assumption (i.e., all points in the visual field equally weighted) is of little consequence. In the multi-patch stimuli the local neighbourhood of the patch can vary from trial to trial. It is easy to see that if the local weighting were strong then the 'locking' of the gain and offset may not be as strong as predicted and could vary between trials.

In summary, the simple two stage model presented here is a useful exploratory tool for examining the types of signal manipulation that may occur within the visual system. The model is by no means complete, but even in this form can predict features of matching data with some success.

An interesting feature of the model can be seen from examining the perceptual space co-ordinates of the patches in the multi-patch backgrounds. We can see that for the lower values of (b-y)* the coordinates of the patches almost overlay and as the value increases so do the errors between points. We can see that to a large extent the transform removes the effects of the illumination difference.

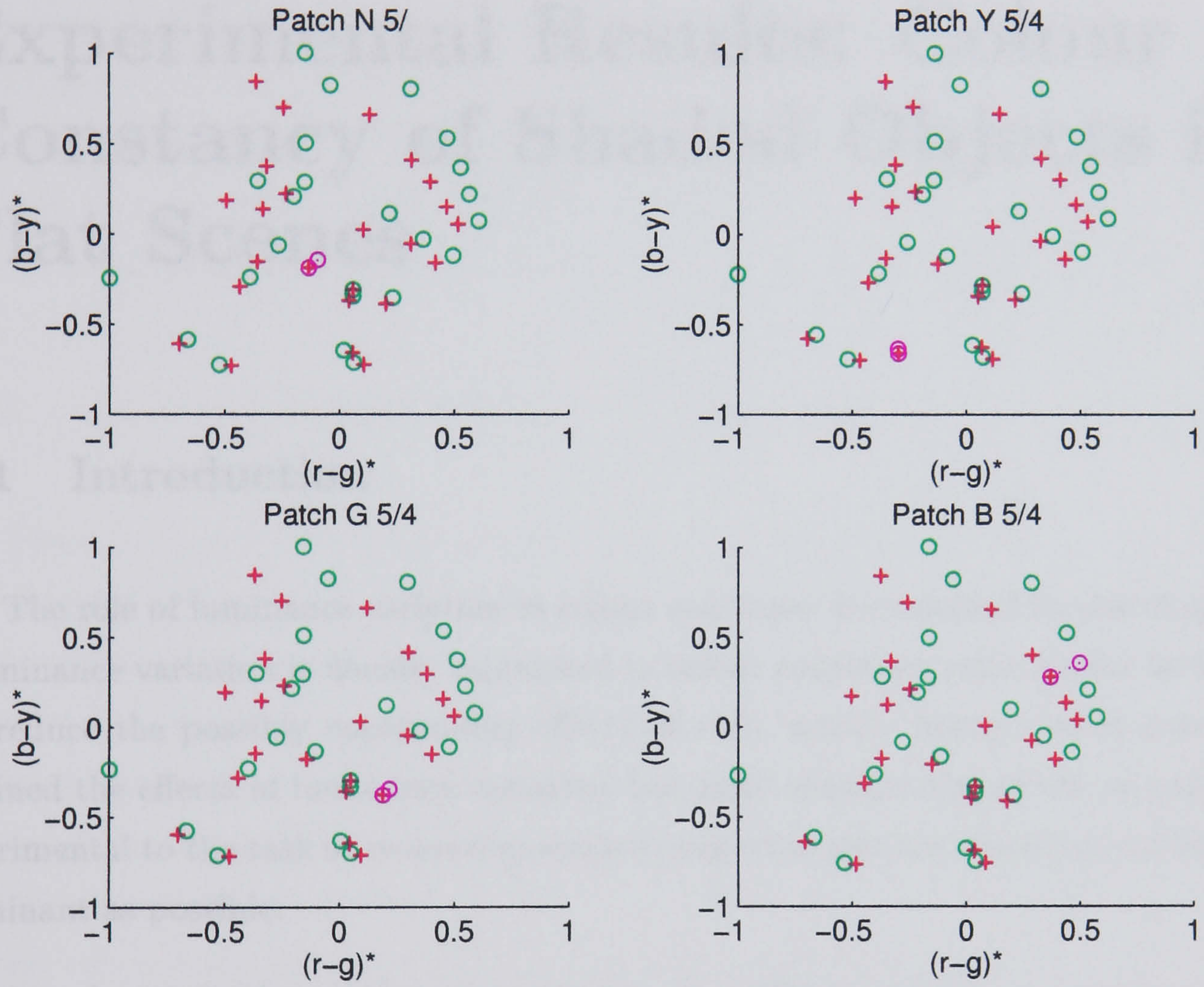

Figure 6.23: Two stage model predictions for N 5/, Y 5/4, G 5/4 and B 5/4 perfect colour constancy against multi-patch backgrounds (test under D_{40}, match under D_{65}). Note: the error in the perceptual space is so small in some cases that the dotted line is not visible.

Chapter 7

Experimental Results: Colour Constancy of Shaded Objects in Flat Scenes

7.1 Introduction

The role of luminance variation in colour constancy is examined in this chapter. Luminance variation is usually minimised in colour constancy experiments in order to reduce the possibly confounding effects of such 'noise'. Some studies have examined the effects of luminance variation, but most consider the effects as possibly detrimental to the task of recovering surface properties and use stimuli as uniformly luminant as possible.

Arend and Goldstein (Arend and Goldstein 1987) conducted an experiment in which a 'realistic luminance profile' due to the scene illumination was simulated across the test scene. They were primarily concerned with probing the extent to which the observers could discount the illumination profile.

Lucassen and Walraven (Lucassen and Walraven 1992) examined the role of absolute illumination level in a (non fused) simultaneous haploscopic match display. They found that the luminance offset between the two fields affected the matches only minimally.

The experiments in this chapter ask a different question: does the extra information present in an object's luminance profile possibly *improve* the degree of colour constancy?

The experiments are split into two main sections. The first examines object matching against different uniform and multi-patch backgrounds. The general procedural aspects are kept very similar to those used in the patch matches in order to facilitate comparisons. The second set examines the colour constancy performance of observers when matching shaded objects. Two main scenarios are explored. The first is essentially a standard colour constancy simulation with the exception that the observer adjusts an object in terms of luminance, hue and saturation until it matches a test object. The second explores 'self constancy', posing the following question. Since the object may be easily classed as distinct from the background and as such may be subject to different illumination, does the visual system recognise this fact and compensate accordingly for any apparent colour shifts?

7.2 Object simulation

There are many differences between a flat reflecting surface and a real life object, for example luminance variation, surface textures, chromatic variation, shape etc. In this chapter we wish to add the minimal amount of features in order to represent an object. This can be done by providing the correct 2D contour and luminance profile for an object. The resulting depiction is not the same thing as a full colour picture of an object; for this reason, the objects simulated by the addition of shape and luminance information will always be qualified with the word 'mono'. The word 'mono' refers to the *monochromatic* nature of the stimuli; i.e. the whole object is, in colorimetric terms, the same colour at every point (the perceived colour may vary).

The shading is achieved by the following procedure. First a monochrome image is taken of the test object. This image is then edited to remove all non-object information. The levels in the image are processed to present a given contrast (set by the experimenter but consistent across test objects, and designed to make the objects look as natural as possible under the illumination conditions tested). When the object is loaded the monochrome levels are linearly transformed into CIE Y values. A Munsell patch is chosen for the surface colour and all illuminant interactions performed on this chosen spectrum. The spectrum calculations produce the xyY space co-ordinates that would be used to draw a test patch in the experiments of

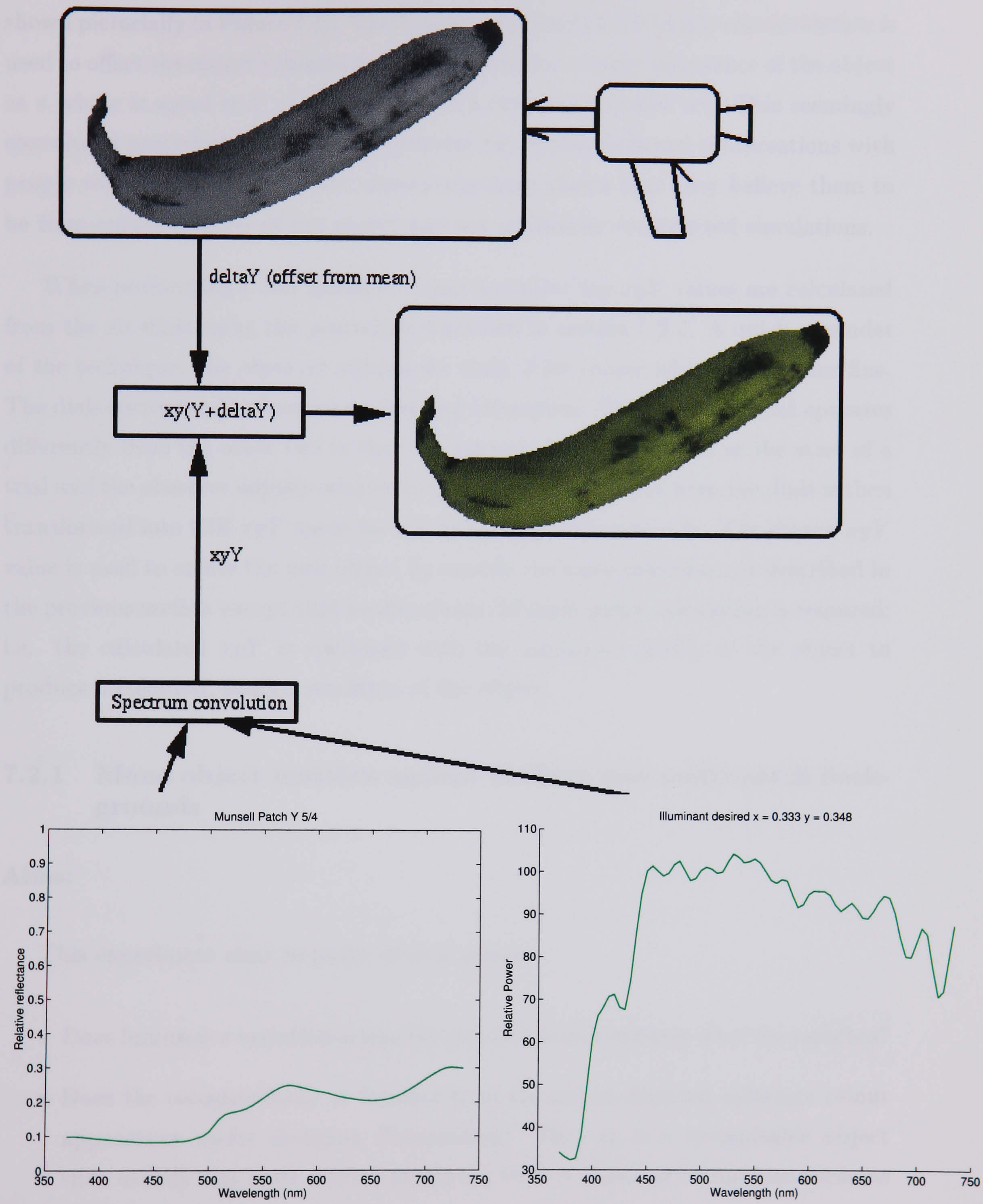

Figure 7.1: Details of mono object creation

the previous chapter. The xy values are used with the CIE Y values from the object monochrome image to produce the RGB values sent to the screen (this sequence is shown pictorially in Figure 7.1). The Y value produced in the spectrum calculation is used to offset the object's luminance scale so that the average luminance of the object as a whole is equal to Y (i.e. the same as a test patch would be). This seemingly convoluted procedure produces quite convincing results. Informal conversations with people who have seen the object simulations have shown that they believe them to be from colour pictures of the object and not artificially constructed simulations.

When performing a dial match of object to object the xyY values are calculated from the six dials using the procedure described in section 5.2.2. A quick reminder of the technique: the observer adjusts six dials, 3 for coarse adjustment, 3 for fine. The dials correspond to luminance, hue and saturation. The luminance dial operates differently from the other two in that the 'correct' luminance is set at the start of a trial and the observer adjusts relative to this value. The output from the dials is then transformed into CIE xyY space for use by the experimental code. The chosen xyY value is used to create the new object by exactly the same procedure as described in the previous section except that no illuminant/Munsell patch calculation is required; i.e. the calculated xyY is combined with the luminance profile of the object to produce a coloured, shaded rendition of the object.

7.2.1 Mono object matches against uniform and multi-patch backgrounds

Aims:

This experiment aims to probe several points:

- Does luminance variation across the test and match surfaces alter the matches?
- Does the recognisability or familiarity of the object depicted influence colour appearance under changing illumination? That is, is a recognisable object that usually has a set colour associated with it matched in the same way as one that does not?

- Does image segmentation affect the matches (the objects can be considered not to belong to the background and hence may have reduced influence from the background)?

The experimentation in this chapter employs novel stimuli. With this in mind this first block of experimentation is mainly geared towards measuring baseline performance for matching mono shaded objects to mono shaded objects. It is yet to be proven that observers can reliably perform this kind of match at all, let alone across large portions of colour space and in the presence of differential adaptation. This section uses a selection of chromaticity points covering as much of the available gamut as possible. Matches to these points are made against various backgrounds so as to allow the assessment of matching abilities for several regions of colour space under different states of observer adaptation. Unfortunately the perfect colour constancy points that correspond to the test chromaticities are not known and as such conclusions as to colour constancy performance in objects would be limited to approximations. Section 7.2.2 uses stimuli constructed using the full simulation of surface reflectance and illuminant interactions. Combining the results from the two sections will allow assessment of overall observer performance for mono object matches across colour space and in terms of colour constancy.

Stimulus, set-up and methods:

Two objects are used, a banana and a clay blob. Figure 7.2 shows the banana monochrome image. Cross-sectional slices are also shown to illustrate the luminance profile of the object. The same information for the blob object is shown in Figure 7.3.

The objects are centred within their respective fields (centres displaced by 5^o vertically when fused). Figures 7.4 and 7.6 show examples of the banana object against uniform and multi-patch backgrounds. Figures 7.5 and 7.7 show the corresponding figures for the blob object.

The current experiment tests the effects of matching mono objects. In order to get an idea of the observer's general object matching capabilities several chromaticities are used which broadly cover a CIE xy plane. Figure 7.8 shows the candidate chromaticity points. The gamut for 10 cd/m^2 is shown on this plot in order to highlight a problem with the data set: due to the gamut restriction of the monitor

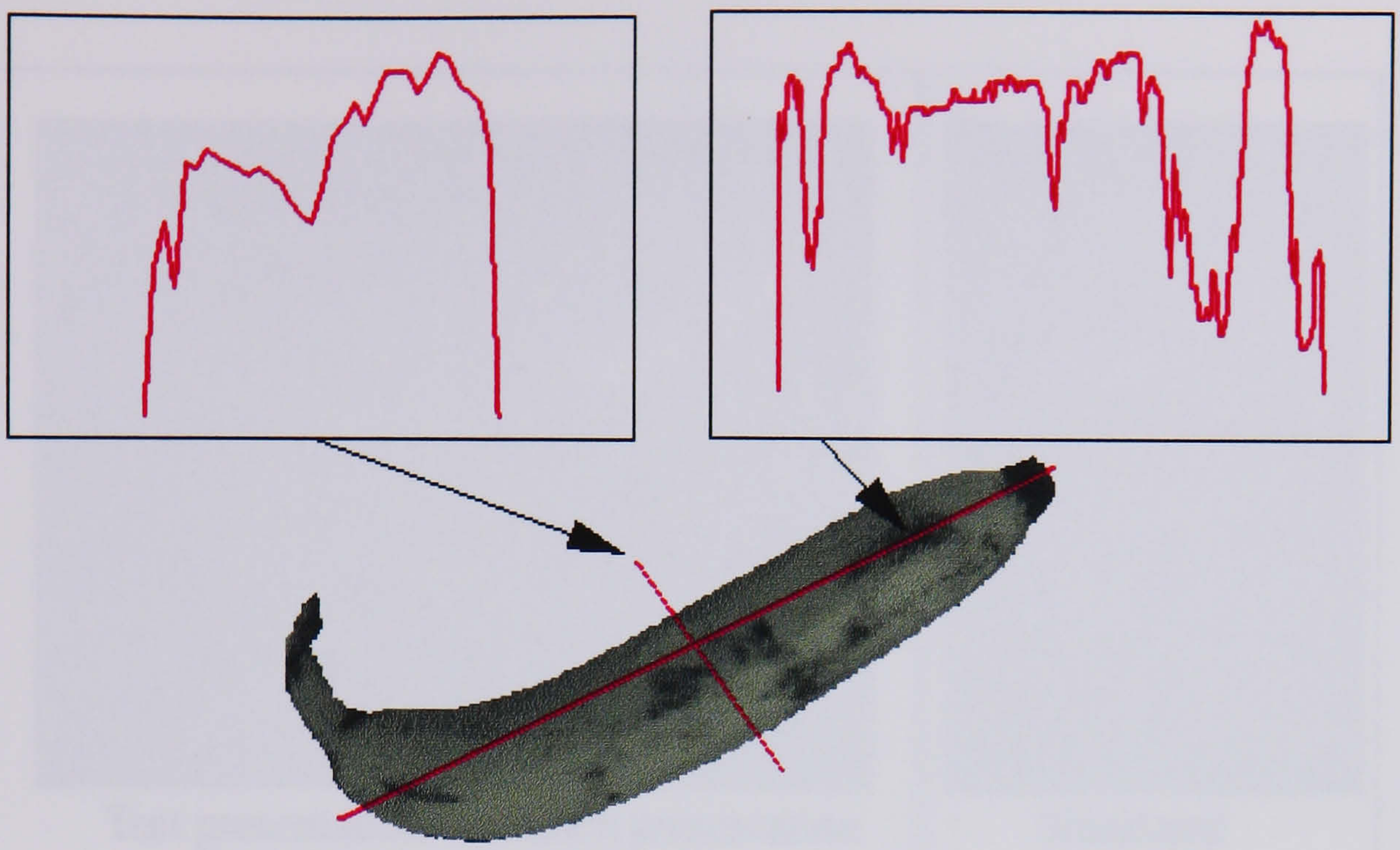

Figure 7.2: The banana, monochrome image specification

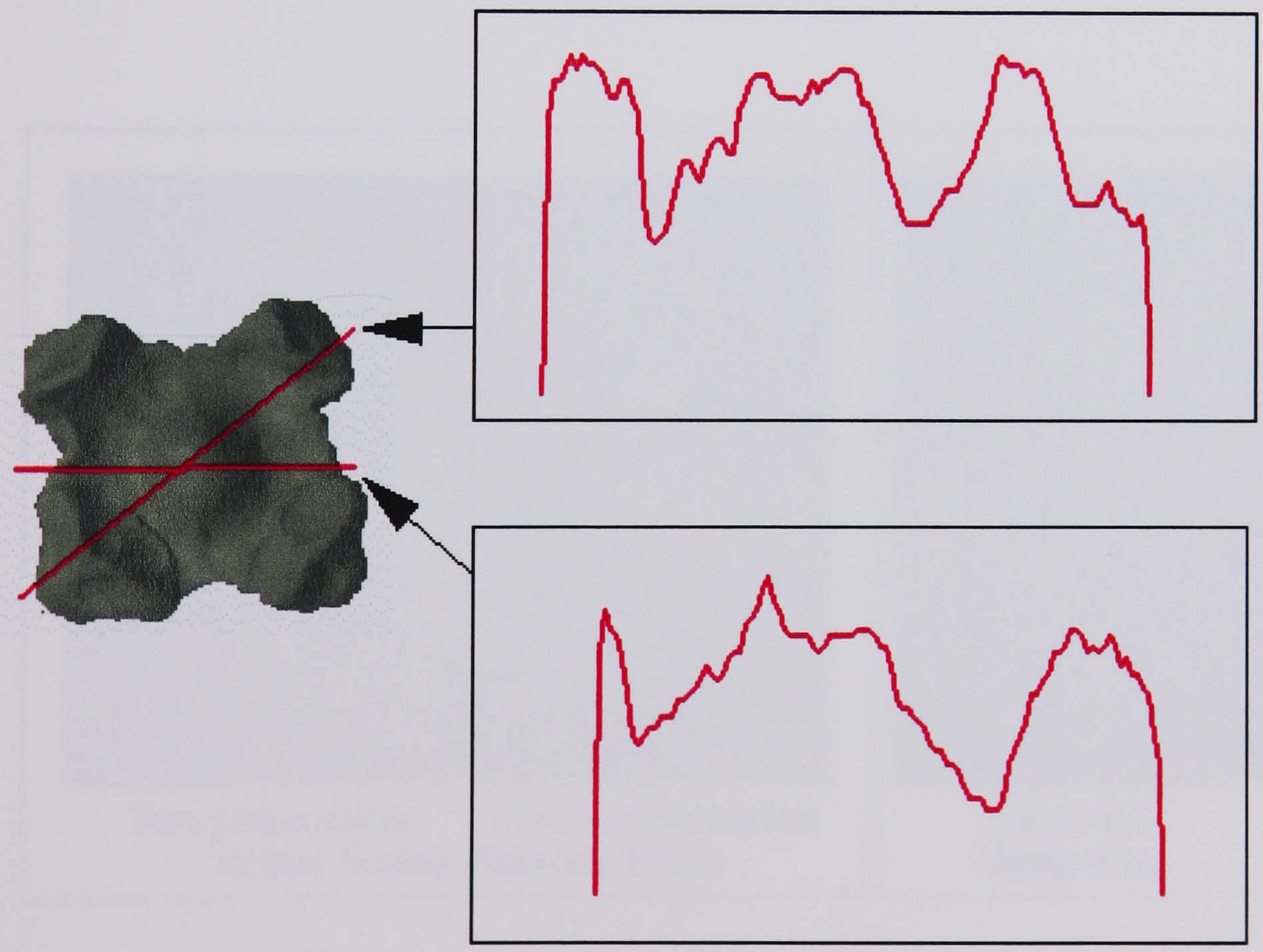

Figure 7.3: The blob, monochrome image specification

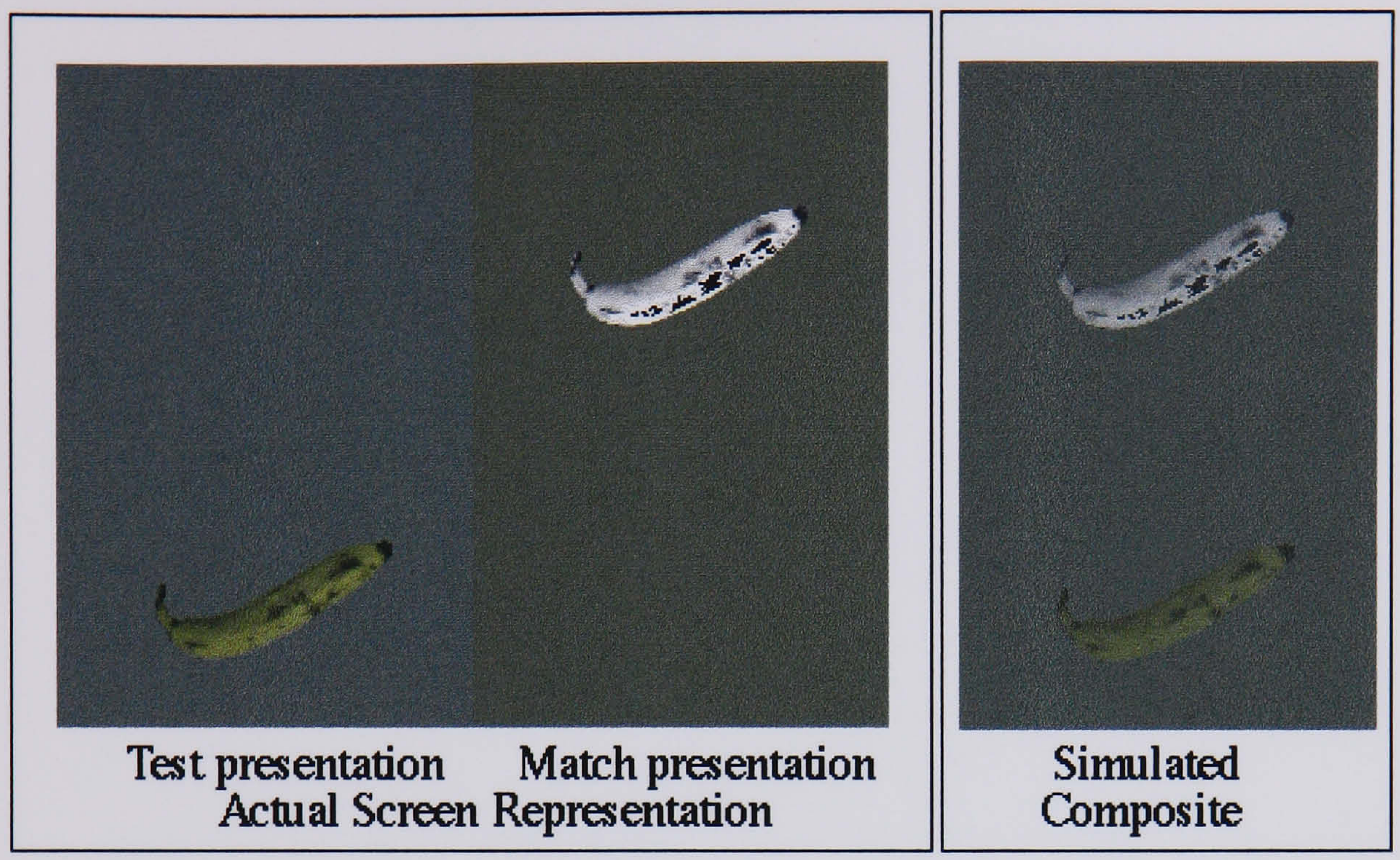

Figure 7.4: Example of the banana against a uniform background

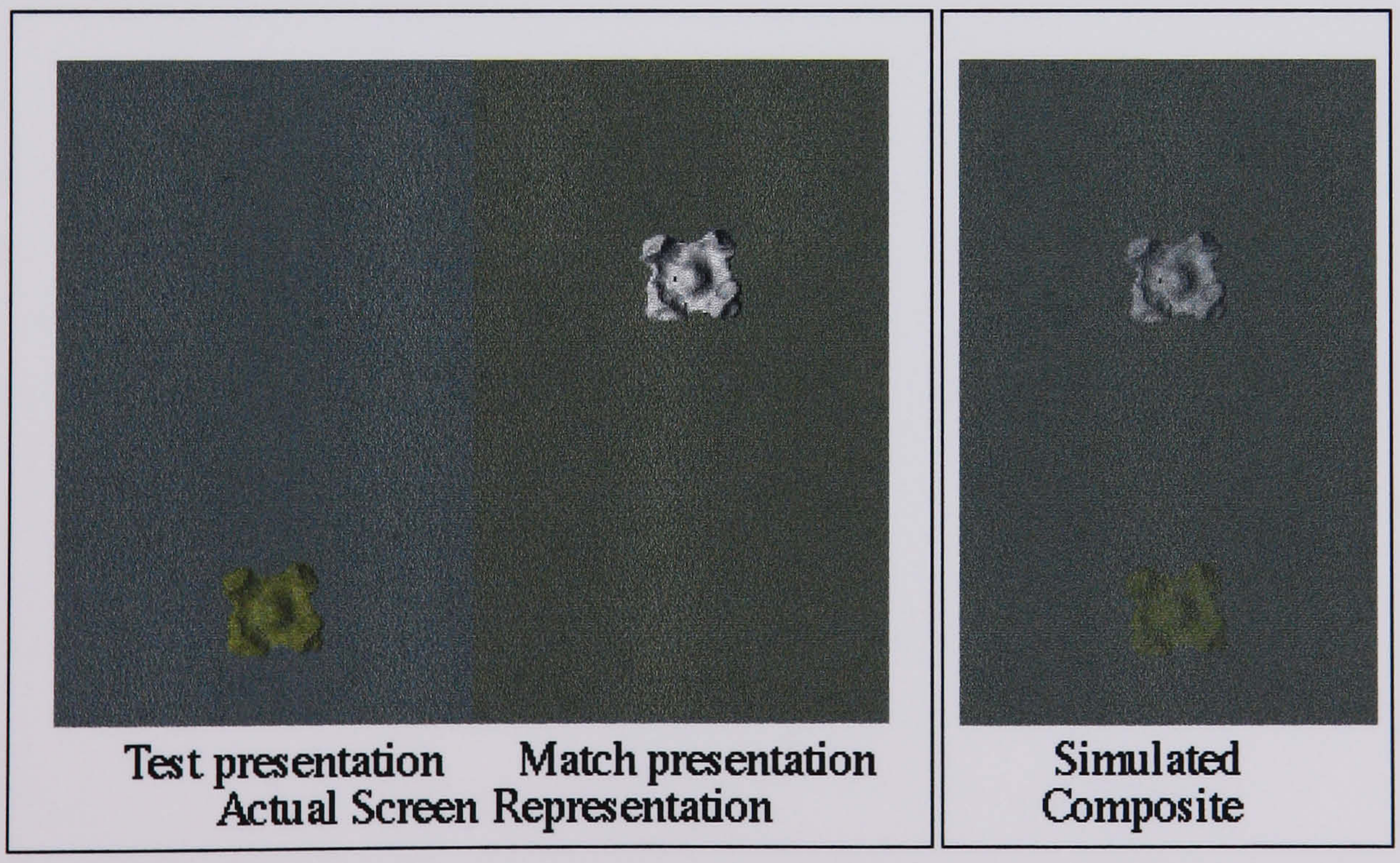

Figure 7.5: Example of the blob against a uniform background

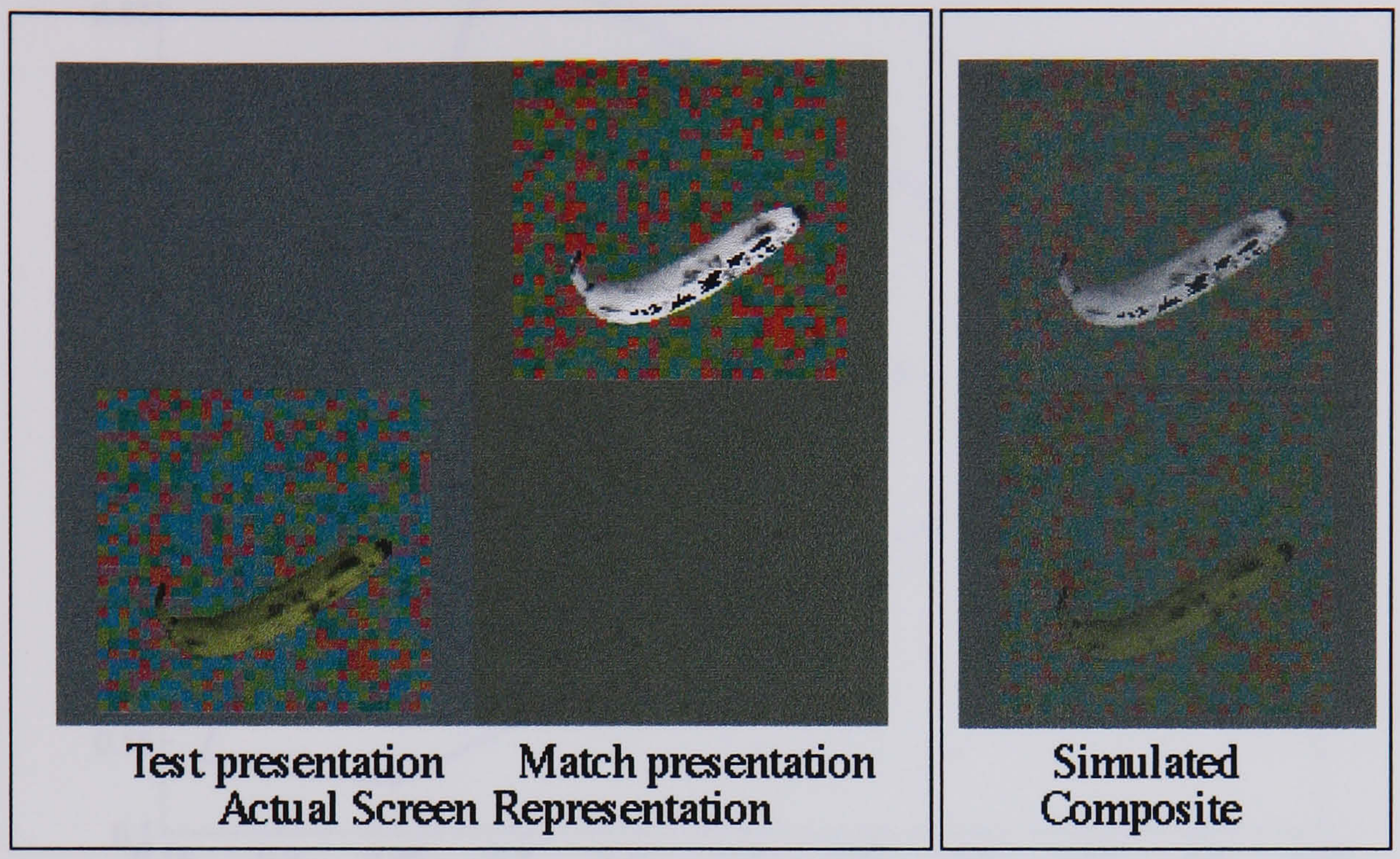

Figure 7.6: Example of the banana against a multi coloured background

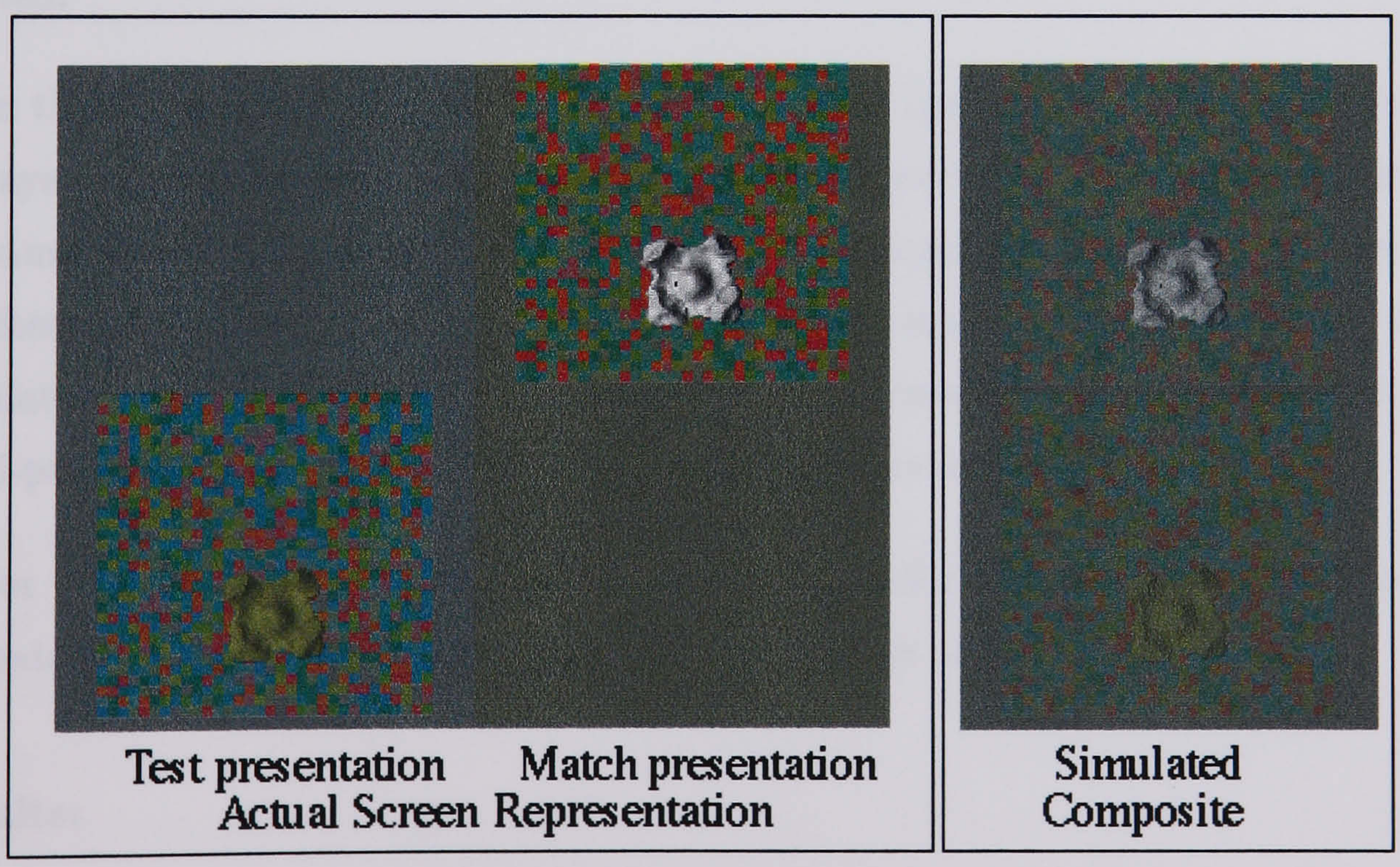

Figure 7.7: Example of the blob against a multi coloured background

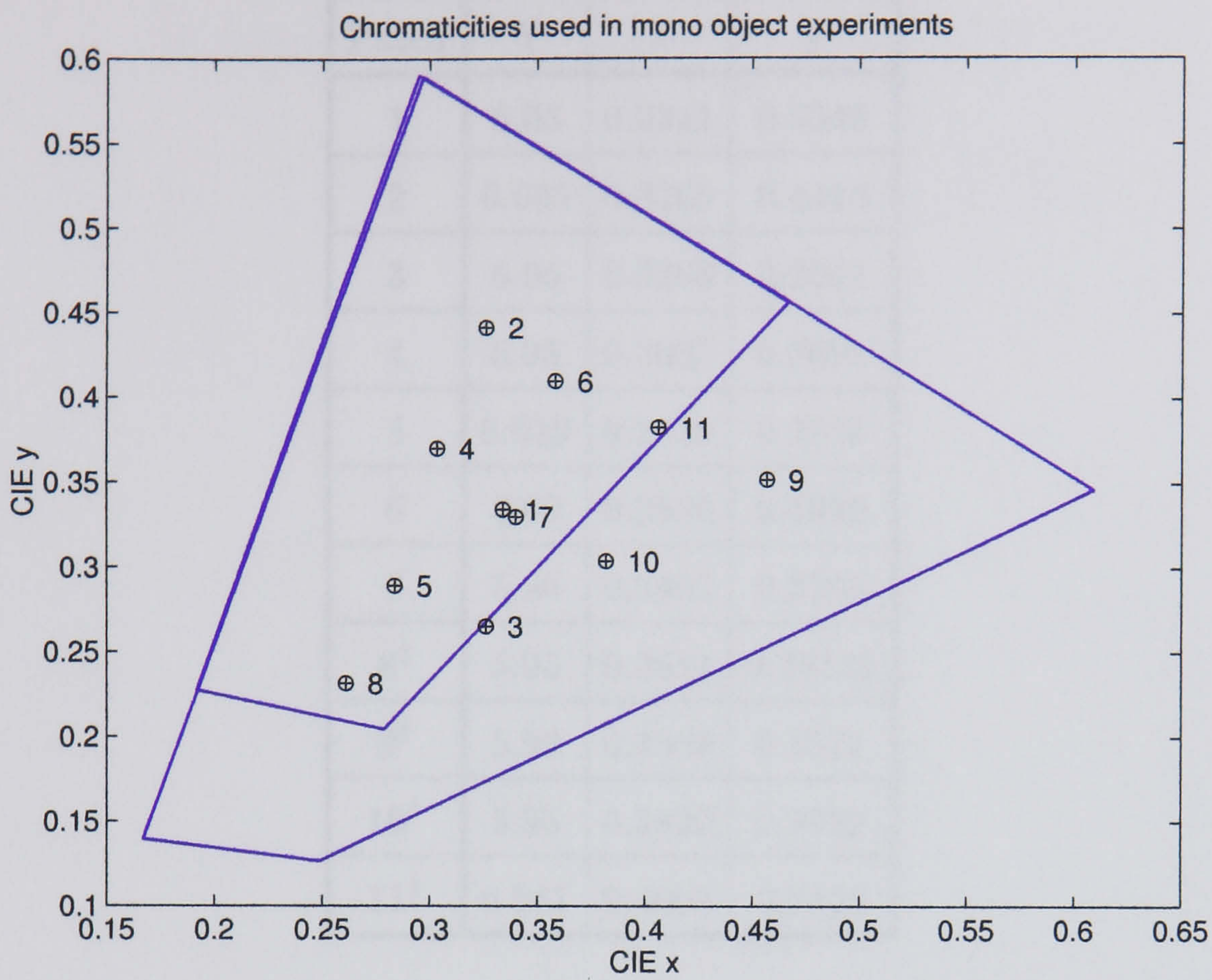

Figure 7.8: Test chromaticity points used in the object experiments

insufficient luminance variation exists at all the candidate points to allow the object to be represented. Where gamut restrictions caused problems test points are simply omitted.

In this experiment the same test chromaticities (patches and mono objects) are displayed against different backgrounds. The backgrounds are generated in exactly the same way as for the uniform and multi-patch experiments for flat patches. The matches are always performed on a background of multi-patches under D_{65} or the calculated uniform equivalent. Test objects are presented against a background of multi-patches under D_{40} or calculated uniform equivalent.

For comparison purposes the experiment is performed with the corresponding flat patches with all other parameters remaining the same.

Results:

The results for TE are shown in figures 7.9, 7.10 and 7.11 for banana, blob and patch matches respectively. The results for matches against uniform and multi-patch backgrounds have been combined and are shown on the same plots.

Patch	Y	x	y
1	6.03	0.3341	0.3340
2	6.045	0.3265	0.4410
3	6.06	0.3263	0.2651
4	5.95	0.3037	0.3698
5	6.019	0.2838	0.2890
6	6.00	0.3586	0.4096
7	5.96	0.3402	0.3296
8†	5.95	0.2611	0.23140
9†	5.99	0.4569	0.3521
10†	5.95	0.3820	0.3039
11†	6.047	0.4063	0.3825

† the luminance range present in the shaded objects was not displayable for these test patch chromaticities.

Table 7.1: Coordinates for patches used in object experiments

Similarly, the results for JH are shown in figures 7.12, 7.13 and 7.14 for banana, blob and patch matches respectively.

Analysis:

Before considering the results of the object match it is worth examining the patch match results shown in figures 7.11 and 7.14. The results presented in these plots show the first matches performed by TE and JH. The data sets shown on these plots are segregated to highlight learning effects: set 1 was performed on one day and set 2 on the next. TE takes to the task quickly and produces matches of reasonable repeatability almost immediately though improvement can be seen in the most recent matches (figure 5.11). JH's initial matches are very noisy and, interestingly, a large number are signalled as high confidence matches. Little improvement is in evidence between the two days but more recent data (Figure 5.15) show marked improvements in certain areas of the space. It is difficult to tell whether the improvement is due to the observer becoming more capable at the task of performing a match or if

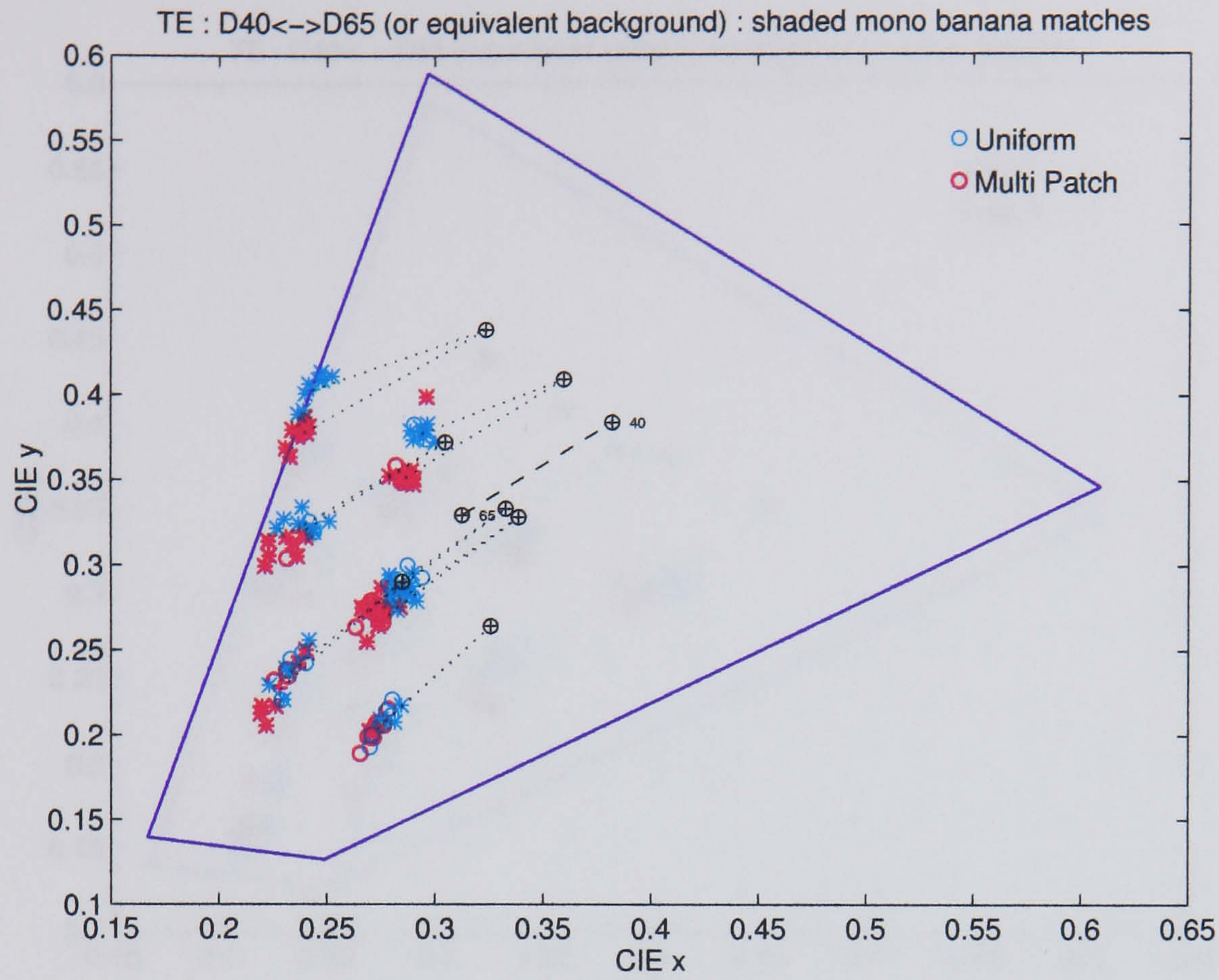

Figure 7.9: Matches for mono banana object for observer TE. (Note: the dashed line shows the extent of the inter-ocular illuminant shift. The points show the locations of N 5/ under D_{40} and D_{65})

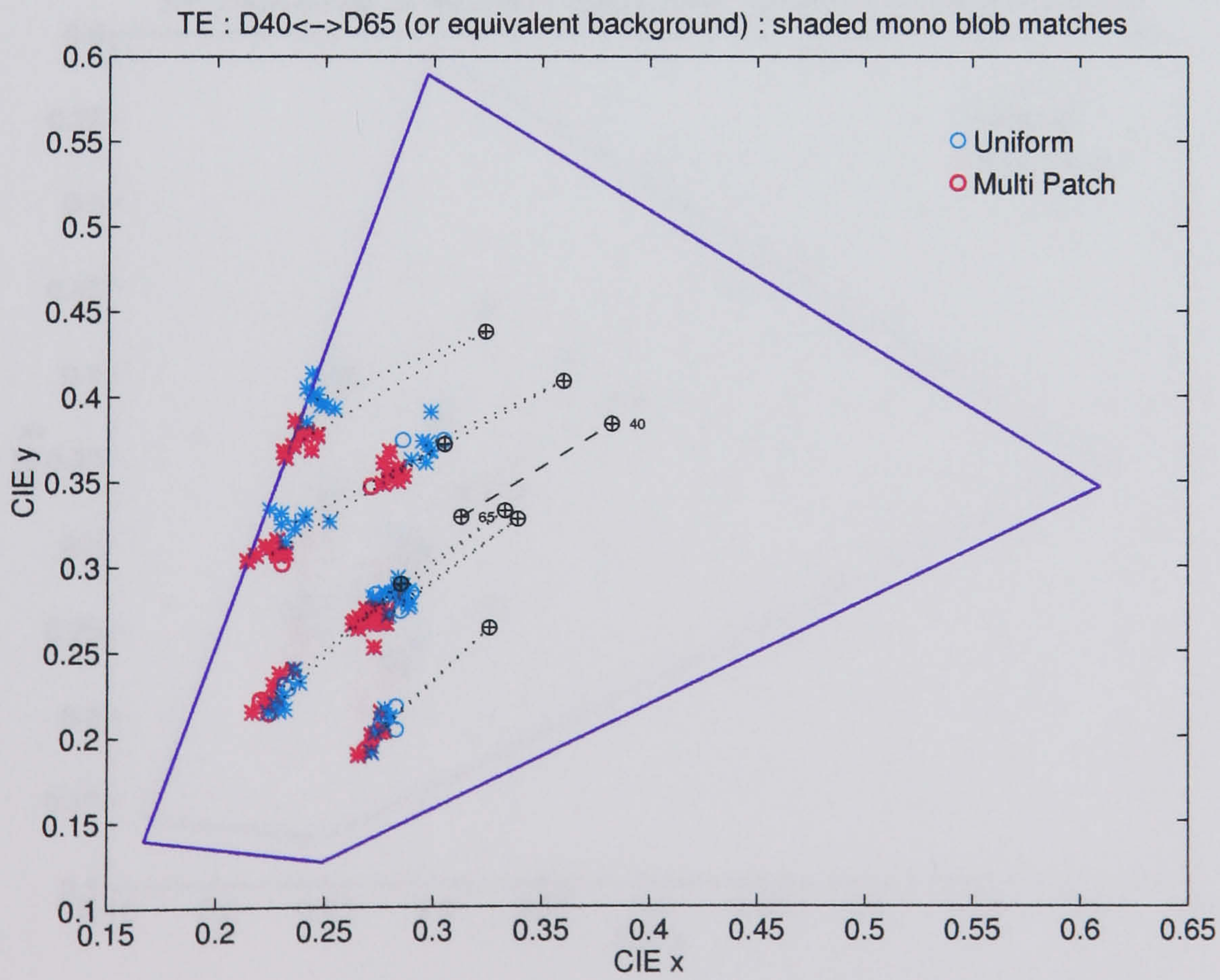

Figure 7.10: Matches for mono clay blob object for observer TE.

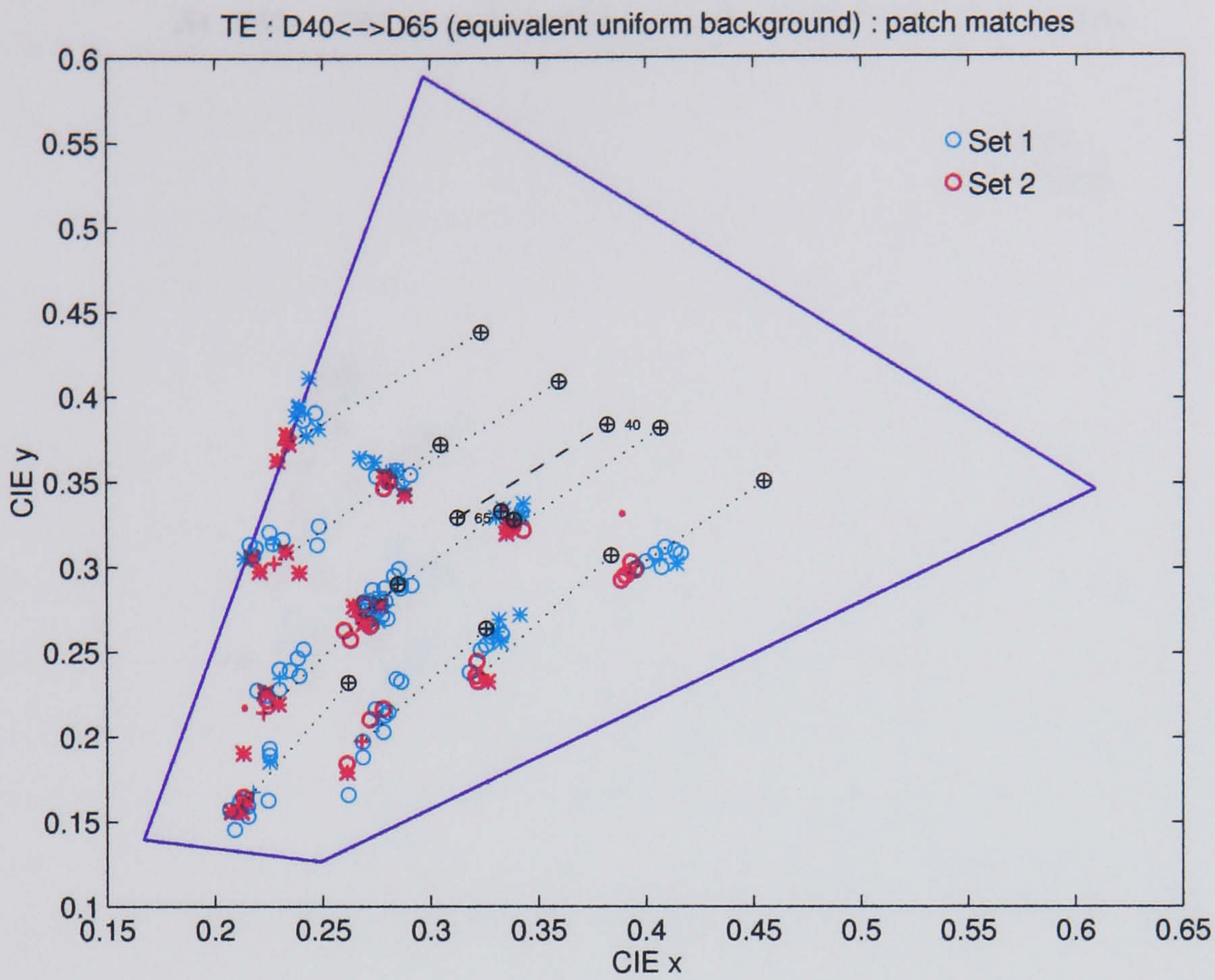

Figure 7.11: Patch matches on uniform equivalent backgrounds for observer TE.

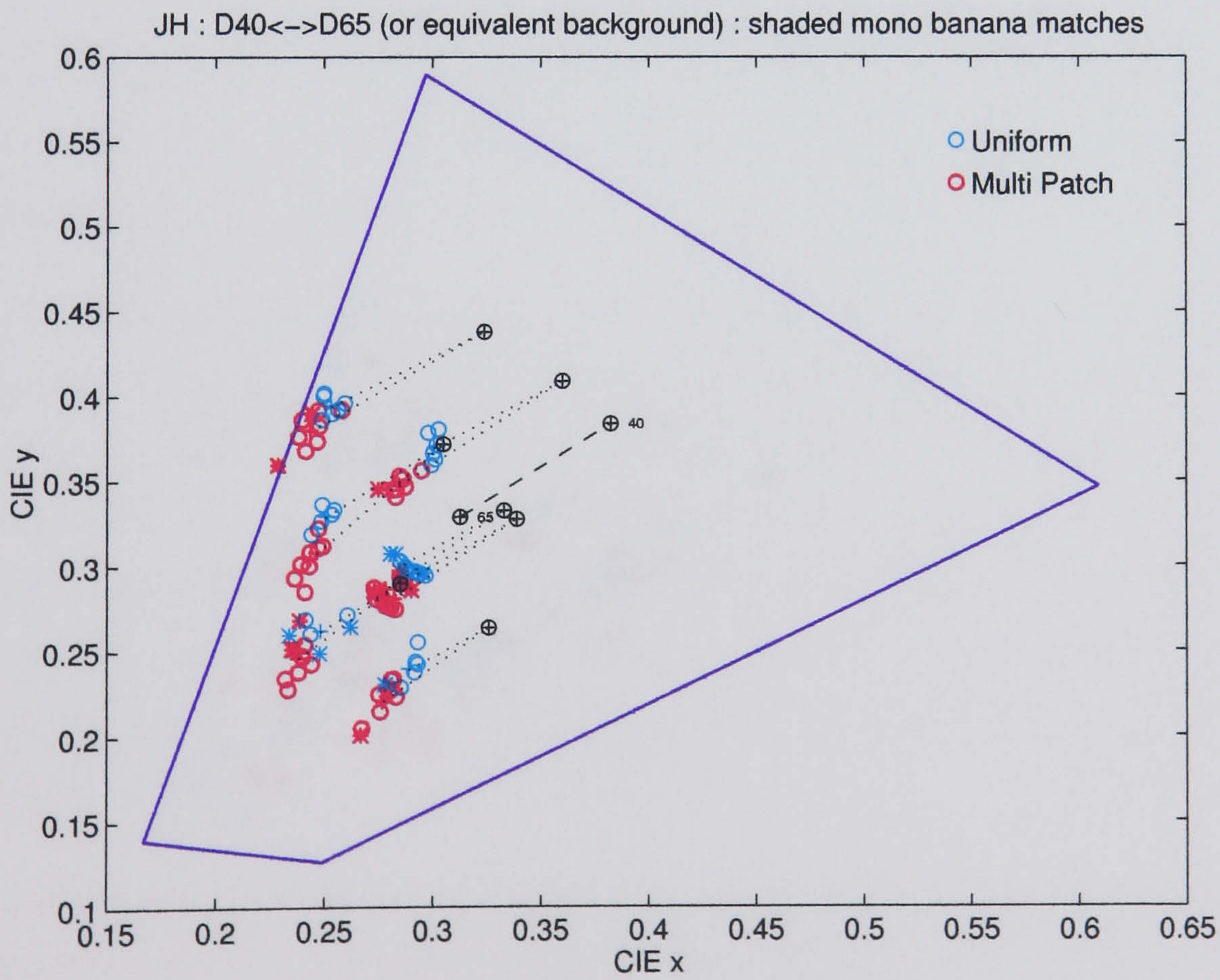

Figure 7.12: Matches for mono banana object for observer JH.

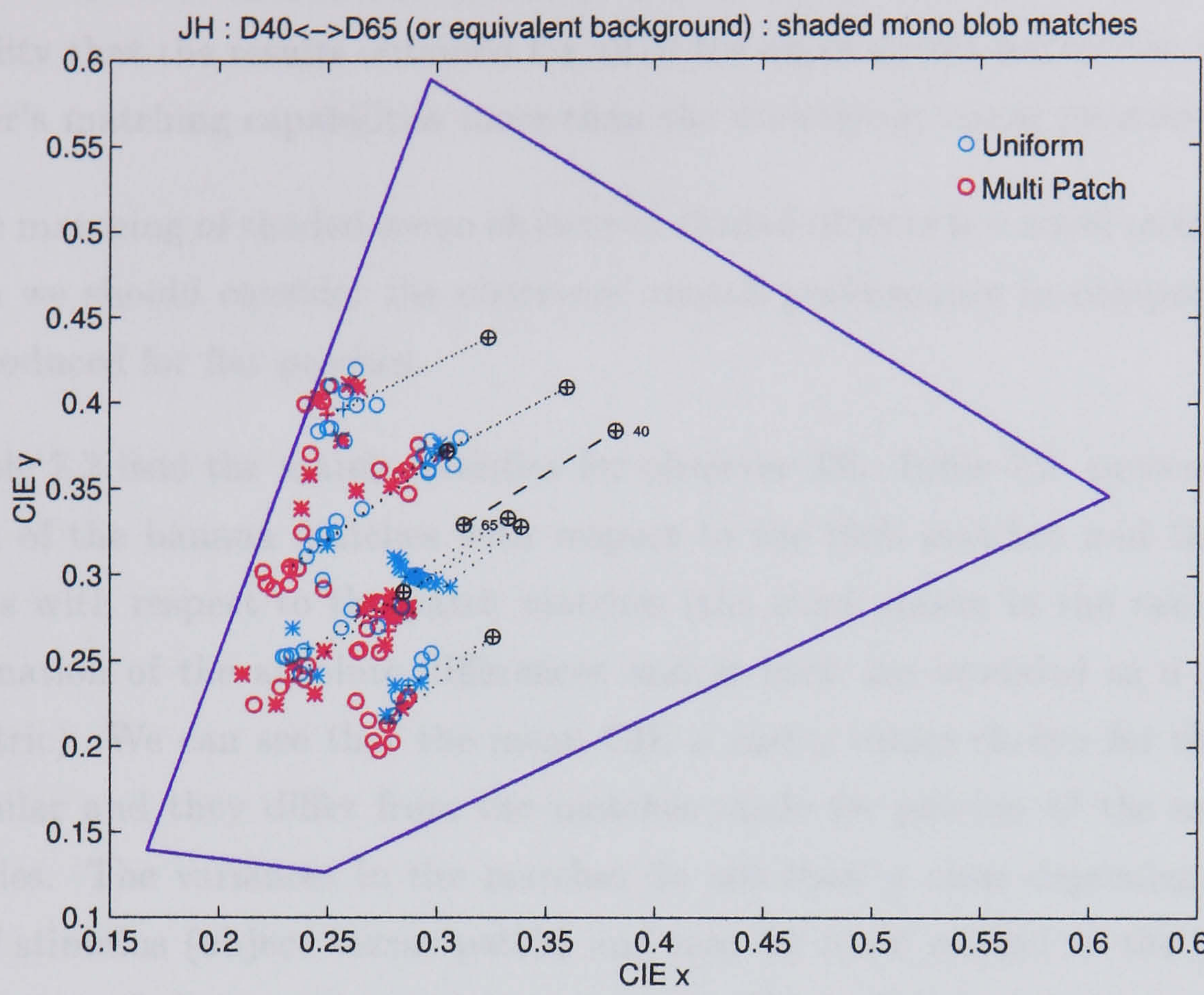

Figure 7.13: Matches for mono clay blob object for observer JH.

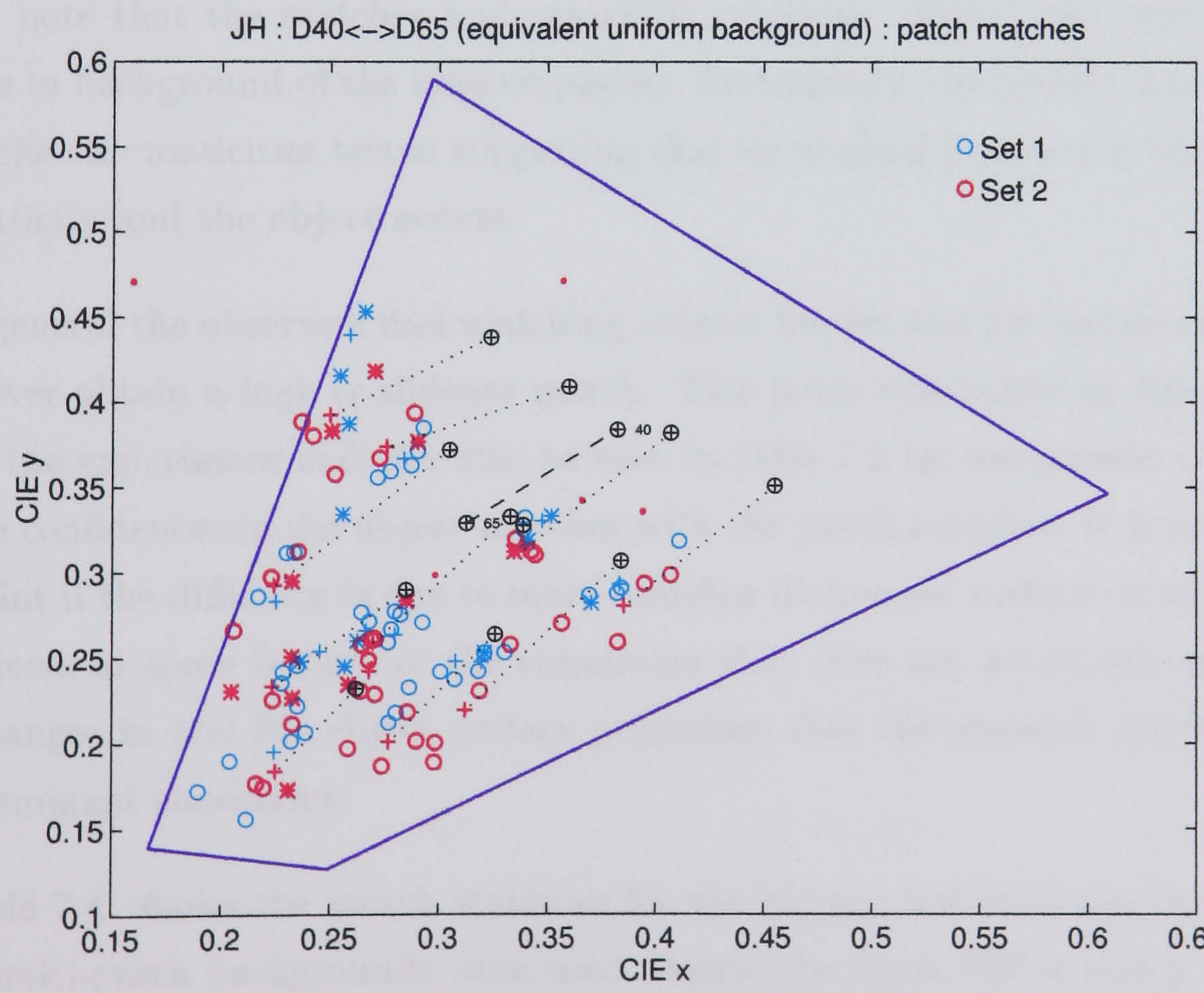

Figure 7.14: Patch matches on uniform equivalent backgrounds for observer JH.

an improvement in discrimination ability is occurring. There will always remain the possibility that the results obtained for all of the experiments mirror the individual observer's matching capabilities more than the underlying visual processes.

The matching of shaded mono objects to shaded objects is a novel technique and as such we should consider the observers' match performance in comparison with that produced for flat patches.

Table 7.2 lists the match statistics for observer TE. Table 7.3 presents a comparison of the banana matches with respect to the blob matches and the banana matches with respect to the patch matches (the sums shown in the table refer to a summation of the absolute differences and as such are intended as a global error metric). We can see that the mean CIE x and y values chosen for the objects are similar and they differ from the matches made for patches of the same chromaticities. The variances in the matches do not show a clear dependence on the type of stimulus (object versus patch) and may be more related to the particular combination of chromaticity and luminance profile used rather than the use of an object or a patch. As mentioned at the start of this section colour constancy performance cannot be assessed with this particular set of experimental conditions, but we can note that the matches undergo shifts consistent with those expected from changes in background of the kind employed. Fortunately, the results are consistent across the chromaticities tested suggesting that no unusual interaction between test chromaticity and the object occurs.

In general the observers find matching objects harder and for certain chromaticities never obtain a high confidence match. This point was raised by the observers during the experiment and can also be seen in table 7.2 by comparison of the percentage confidences in the object matches with the patch matches. It is not clear at this point if the difficulty is due to more complex illuminant surface perceptions for the objects or some feature of the simulation that does not accurately mirror the true changes in real life object surface properties that the observer expects in the test illuminant conditions.

Table 7.4 shows the match statistics for the banana and blob produced by TE when multi-patch backgrounds were used. Again the mean CIE x and y values for the objects correlate well with far less agreement in the variances.

For the patch matches it was found that the degree of colour constancy exhibited by the observers increased when the multi-patch background was used. Table 7.5 shows the differences in mean CIE x and y matches for the banana against uniform equivalent (E'_{40}) and the multi-patch backgrounds. It was mentioned in chapter 6 that the initial calculation of the equivalent backgrounds, E', used in this chapter has subsequently been found to be somewhat inaccurate, hence the background averages were re-estimated producing the E illuminants. An estimate of the shift induced by the differences in calculation technique can be calculated from E'_{40}, E_{40}, E'_{65} and E_{65}. The estimate turns out to be $\delta x = -0.002$, $\delta y = -0.0588$ (i.e. the results under the E' illuminants should be shifted by these amounts from those obtained under E). This correction has been applied to the points and is shown in brackets in the table itself[1]. The difference between the uniform and multi-patch backgrounds appears to be reduced when object matches are performed. For some test chromaticities there is evidence of shifts away from perfect colour constancy, but where they occur they appear to affect matches in both background conditions.

In summary we find that observers can match simulations of monochromatic objects with accuracy comparable to patches. The task is more difficult and obtaining a perfect perceptual match can be extremely difficult and frustrating for observers. Colour induction effects appear to be markedly reduced, i.e. the background does not appear to induce a large hue shift in the objects, as indicated by the reduction in difference between the uniform and multi-patch backgrounds that was evident with patch matching.

7.2.2 Mono object matches under consistent and inconsistent illumination conditions

Aims:

In everyday life examples of differential scene illumination occur frequently. For example, whenever we look out of a window from an artificially illuminated office, our experience is that we can see the change in illumination but we can also still accurately relate scene colours in the office and through the window. This scenario

[1] A control experiment was performed using the E illuminant equivalent backgrounds and results showed that the shift apparent in the results presented here is mostly an artifact of the chromaticity of the uniform background used.

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
1_{banana}	E'_{40}	E'_{65}	0.27	0.273	0.00553	55.56_5	0.27	0.276	0.00581	44.44_4
2_{banana}	E'_{40}	E'_{65}	0.239	0.377	—	11.11_1	0.237	0.376	0.00692	88.89_8
3_{banana}	E'_{40}	E'_{65}	0.272	0.201	0.00666	88.89_8	0.27	0.202	—	11.11_1
4_{banana}	E'_{40}	E'_{65}	0.232	0.303	—	11.11_1	0.238	0.32	0.0252	88.89_8
5_{banana}	E'_{40}	E'_{65}	0.23	0.234	0.00367	33.33_3	0.228	0.223	0.0167	66.67_6
6_{banana}	E'_{40}	E'_{65}	0.284	0.353	0.00493	22.22_2	0.287	0.35	0.00381	77.78_7
7_{banana}	E'_{40}	E'_{65}	0.277	0.271	0.00481	44.44_4	0.276	0.267	0.00757	55.56_5
1_{blob}	E'_{40}	E'_{65}	—	—	—	—	0.268	0.269	0.00542	100_9
2_{blob}	E'_{40}	E'_{65}	—	—	—	—	0.238	0.374	0.00773	100_9
3_{blob}	E'_{40}	E'_{65}	0.271	0.197	0.00507	22.22_2	0.273	0.199	0.00798	77.78_7
4_{blob}	E'_{40}	E'_{65}	0.23	0.301	—	11.11_1	0.225	0.309	0.00588	88.89_8
5_{blob}	E'_{40}	E'_{65}	0.222	0.217	0.00442	22.22_2	0.226	0.225	0.0098	77.78_7
6_{blob}	E'_{40}	E'_{65}	0.271	0.347	—	11.11_1	0.28	0.355	0.0065	88.89_8
7_{blob}	E'_{40}	E'_{65}	0.272	0.275	—	11.11_1	0.278	0.268	0.00574	88.89_8
1_{patch}	E'_{40}	E'_{65}	0.277	0.285	0.00946	100_9	—	—	—	—
2_{patch}	E'_{40}	E'_{65}	0.244	0.388	0.00362	22.22_2	0.242	0.39	0.00894	77.78_7
3_{patch}	E'_{40}	E'_{65}	0.276	0.207	0.018	100_9	—	—	—	—
4_{patch}	E'_{40}	E'_{65}	0.229	0.314	0.0112	88.89_8	0.214	0.304	—	11.11_1
5_{patch}	E'_{40}	E'_{65}	0.232	0.236	0.0104	88.89_8	—	—	—	—
6_{patch}	E'_{40}	E'_{65}	0.281	0.355	0.00732	55.56_5	0.279	0.357	0.00985	44.44_4
7_{patch}	E'_{40}	E'_{65}	0.281	0.278	0.00709	88.89_8	0.276	0.268	—	11.11_1
8_{patch}	E'_{40}	E'_{65}	0.217	0.165	0.0151	88.89_8	0.226	0.185	—	11.11_1
9_{patch}	E'_{40}	E'_{65}	0.405	0.305	0.00809	88.89_8	0.415	0.302	—	11.11_1
10_{patch}	E'_{40}	E'_{65}	0.326	0.252	0.00768	55.56_5	0.334	0.264	0.00715	44.44_4
11_{patch}	E'_{40}	E'_{65}	0.342	0.332	—	11.11_1	0.337	0.333	0.00662	66.67_6

Table 7.2: Match statistics for object and patch matches against uniform backgrounds for TE

TE match differences† relative to the banana			
Patch	mean δx	mean δy	mean $\delta variance$
1_{blob}	0.00174	0.00506	0.00738
2_{blob}	-0.00104	0.00185	0.0103
3_{blob}	-0.000372	0.00314	-0.000261
4_{blob}	0.0121	0.0106	0.00762
5_{blob}	0.00291	0.0034	0.00616
6_{blob}	0.00675	-0.00309	-0.00189
7_{blob}	-0.000832	-0.000392	-0.00902
sum_{blob}	0.0258	0.0275	0.0426
1_{patch}	-0.00649	-0.0107	-0.00208
2_{patch}	-0.00476	-0.0138	-0.00191
3_{patch}	-0.004	-0.0053	-0.0113
4_{patch}	0.00997	0.00514	0.00305
5_{patch}	-0.00393	-0.00928	0.00152
6_{patch}	0.00627	-0.00516	-0.00645
7_{patch}	-0.004	-0.00873	-0.00257
sum_{patch}	0.0394	0.0581	0.0289

† the HIGH and LOW confidence matches were combined to produce the statistics presented in this table.

Table 7.3: Differences in match statistics for the blob object and patches relative to the matches for the banana

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
1_{banana}	D_{40}	D_{65}	0.282	0.29	0.00587	44.44_4	0.28	0.286	0.0054	55.56_5
2_{banana}	D_{40}	D_{65}	—	—	—	—	0.244	0.403	0.00861	100_9
3_{banana}	D_{40}	D_{65}	0.276	0.207	0.00886	66.67_6	0.28	0.211	0.00579	33.33_3
4_{banana}	D_{40}	D_{65}	0.243	0.322	0.0034	22.22_2	0.239	0.324	0.00834	77.78_7
5_{banana}	D_{40}	D_{65}	0.235	0.238	0.0104	44.44_4	0.232	0.236	0.0104	55.56_5
6_{banana}	D_{40}	D_{65}	0.293	0.378	0.00521	22.22_2	0.295	0.377	0.00485	77.78_7
7_{banana}	D_{40}	D_{65}	0.291	0.285	0.00543	33.33_3	0.288	0.283	0.00627	66.67_6
1_{blob}	D_{40}	D_{65}	0.276	0.283	0.00274	22.22_2	0.28	0.286	0.00615	77.78_7
2_{blob}	D_{40}	D_{65}	—	—	—	—	0.246	0.398	0.00734	100_9
3_{blob}	D_{40}	D_{65}	0.279	0.21	0.00502	55.56_5	0.276	0.205	0.00957	44.44_4
4_{blob}	D_{40}	D_{65}	—	—	—	—	0.236	0.326	0.00874	100_9
5_{blob}	D_{40}	D_{65}	0.232	0.229	0.00144	22.22_2	0.231	0.222	0.00873	77.78_7
6_{blob}	D_{40}	D_{65}	0.295	0.374	0.00941	22.22_2	0.296	0.371	0.00765	77.78_7
7_{blob}	D_{40}	D_{65}	0.287	0.281	0.00576	33.33_3	0.286	0.279	0.0059	66.67_6

Table 7.4: Match statistics for object matches against multi-patch backgrounds for TE

Patch	x difference (corrected)	y difference (corrected)	variance difference
1	0.012 (0.01)	0.017 (-0.0418)	0.00034
$2^{\dagger}$	0.007 (0.005)	0.027 (-0.0318)	0.00169
3	0.004 (0.002)	0.006 (-0.0528)	0.0022
4	0.011 (0.009)	0.019 (-0.0398)	—
5	0.005 (0.003)	0.004 (-0.0548)	0.00673
6	0.009 (0.007)	0.025 (-0.0338)	0.00028
7	0.014 (0.012)	0.014 (-0.0448)	0.00062

† calculated from low confidence match (all other entries calculated from high confidence matches only).

Table 7.5: Match statistics differences for the banana object against uniform and multi-patch backgrounds

implies that image segmentation (thought to be a 'low level' process) and classification (thought to be a 'high level' process) may have an effect on colour constancy performance.

We know that flat patch scenes behave in a highly cone ratio conformant manner and hence any differential illumination of patch and background will likely be attributed to a surface colour change of the test patch. In fact, in such 'low information' scenes a surface colour change is just as valid a solution (and probably statistically more likely) as the change in illuminant case. At what levels of information can we reliably tell a surface colour change from an illuminant change and compensate accordingly?

The following experiment probes this question by adding specifically three forms of information to the simple test surface/background configuration:

1. a luminance profile representative of the surface of an object;
2. a clear segmentation criterion (the object can easily be segmented from the background); and
3. object recognition – a banana is an easily recognisable object with an expected surface colour.

The object can be differentially illuminated compared to the surrounding background and the same experiment performed as in the consistent illumination case.

Stimulus, set-up and methods:

Two objects, one recognisable and of a known colour (a banana) and the other not (a clay blob) are matched against multi-patch backgrounds. The multi-patch background is simulated as being under the same (consistent) or a different illuminant (inconsistent). Figure 7.15 shows examples of the consistent and inconsistent illumination stimuli (the cases use the same patch for the object colour, Y5/4 so as to highlight the effects of the illumination variation). The objects take on four possible surface reflectances provided by Munsell chips N 5/ , Y 5/4, G 5/4 and B 5/4. The stimulus configuration is the same as that shown in Figure 7.6 and Figure 7.7.

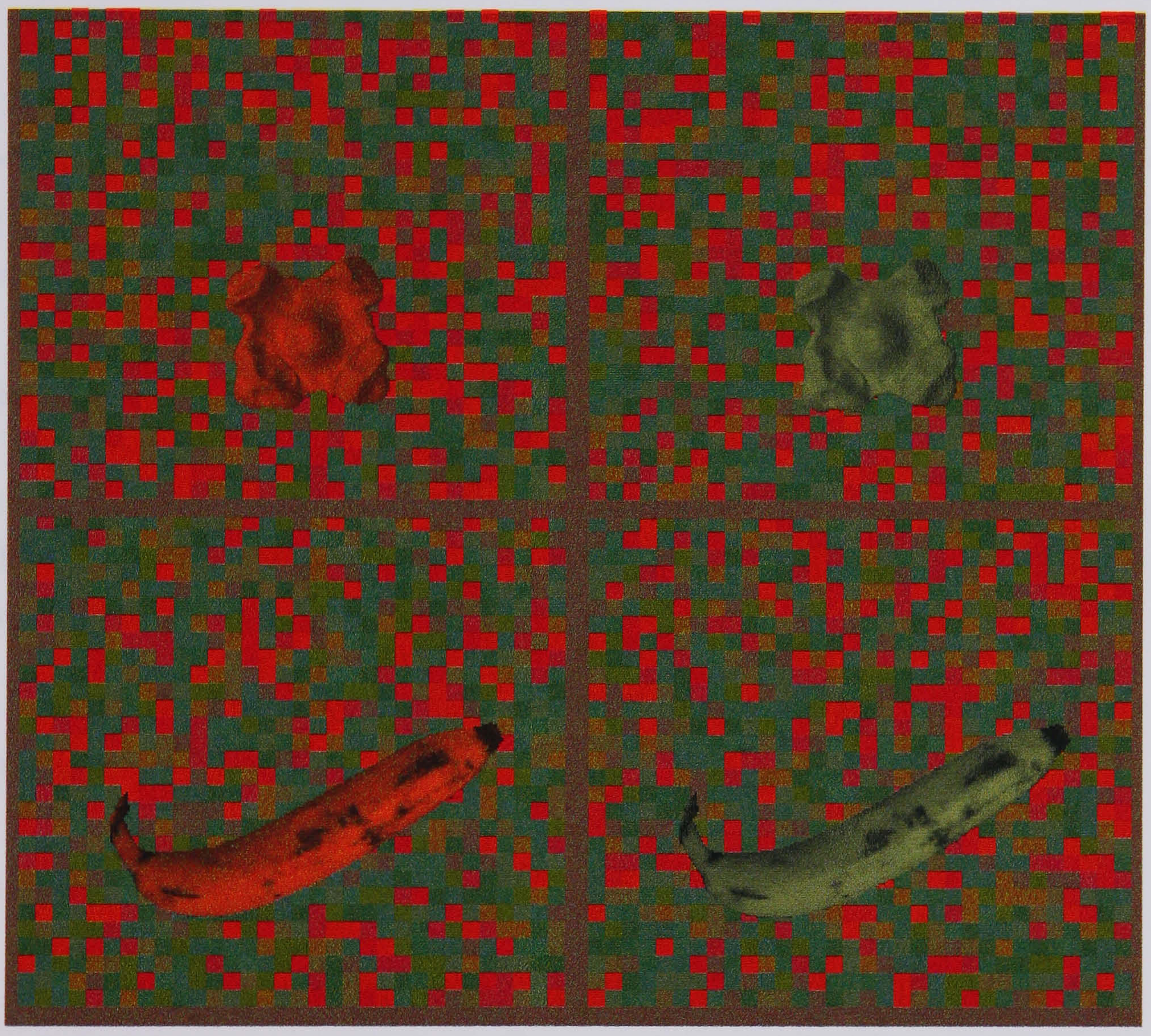

Figure 7.15: Examples of the banana and blob under consistent (left) and inconsistent (right) illumination

To give the observer a clear impression of the background illumination all matches are performed against multi-patch backgrounds (as in the previous section). The matches are made against a multi-patch background under D_{65}. The following illumination conditions are used:

- consistent – test multi-patch background under D_{40}, test object under D_{40}.
- inconsistent – test multi-patch background under D_{100}, test object under D_{40}.

If the observer can correctly assess the illumination conditions on the object and produce good colour constancy then the matches for both conditions should fall in the same region. If they do not, and use properties of the multi-patch backgrounds, then a large shift in the matches is expected between the two conditions.

Results:

Figures 7.16, 7.17 show the match results for observer TE matching the banana and blob respectively. The consistent and inconsistent illumination conditions are shown on the same plots.

Similarly, figures 7.18, 7.19 show the match results for observer JH matching the banana and blob respectively.

Analysis:

The matches of both observers support the hypothesis that the observers are assuming the existence of a single illuminant and attributing variation in object colour to a change in surface properties. Tables 7.6 and 7.7 show the differences very clearly.

The differences between matches for the objects is small implying that object recognition and attribution of a known surface colour to the banana are not major factors in these results. The matches produced are very similar to those produced for the corresponding patches against uniform equivalent backgrounds. This is interesting as a multi-patch background was used in these experiments. This may suggest

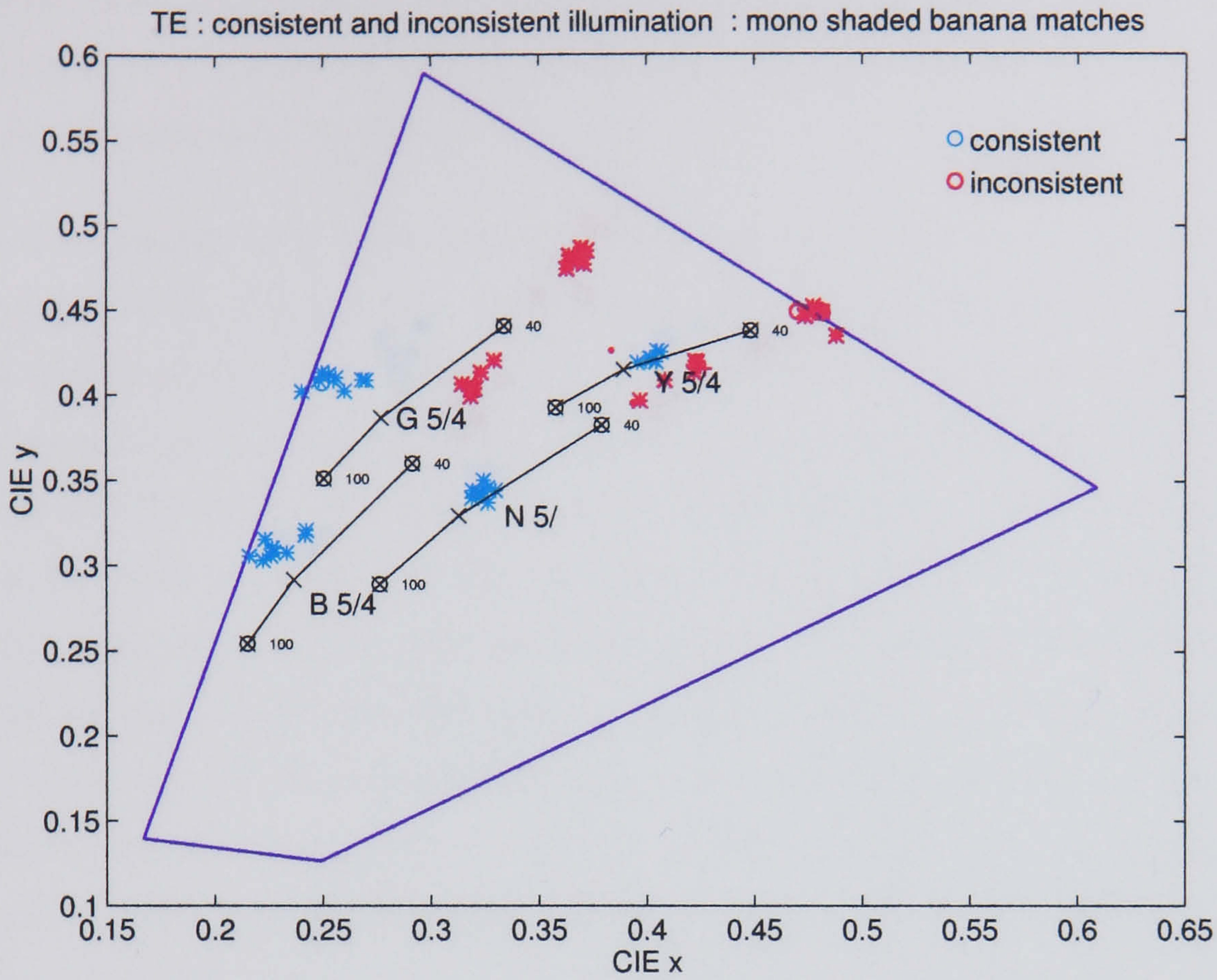

Figure 7.16: Matches for mono banana object under consistent and inconsistent illumination conditions for observer TE.

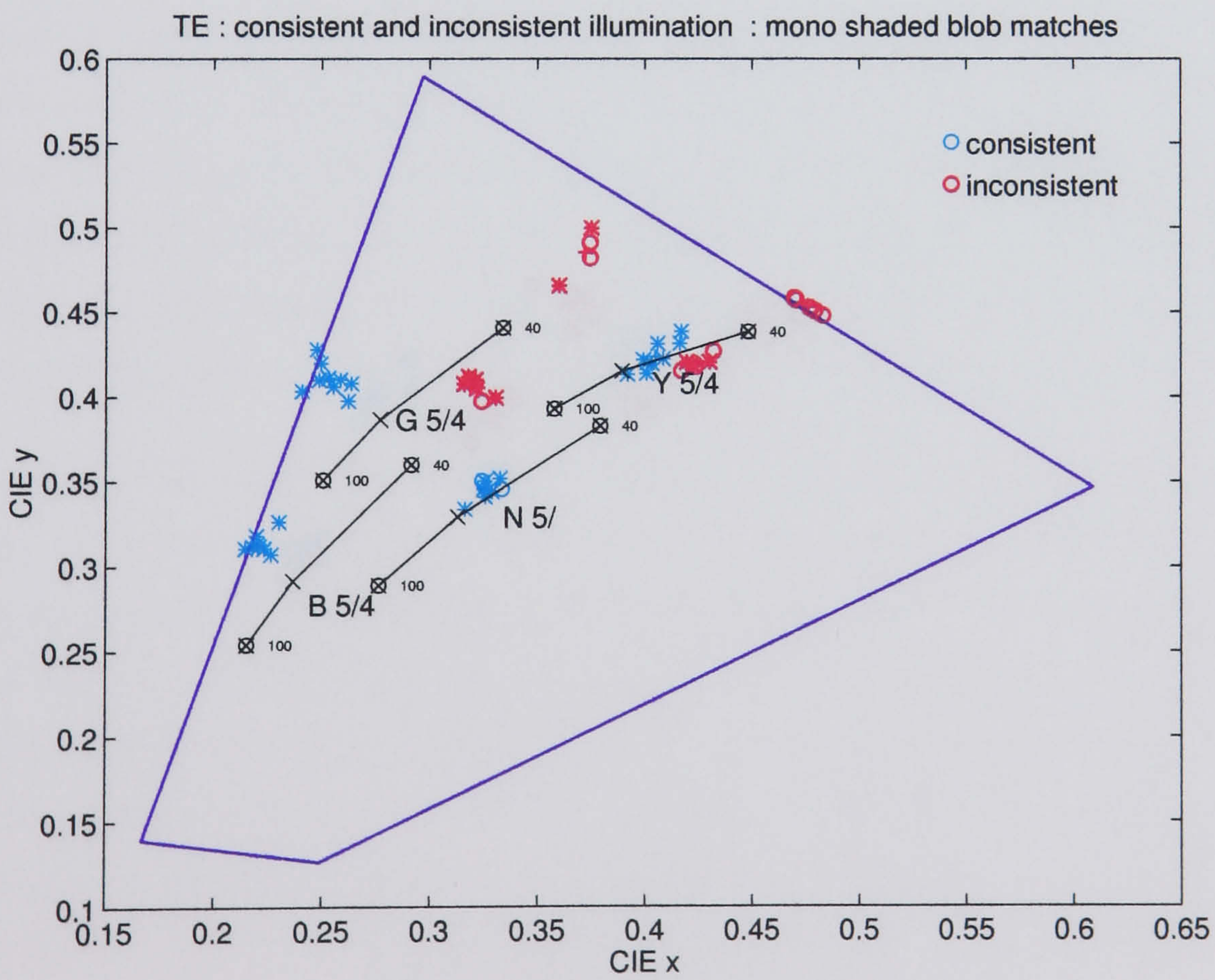

Figure 7.17: Matches for mono clay blob object under consistent and inconsistent illumination conditions for observer TE.

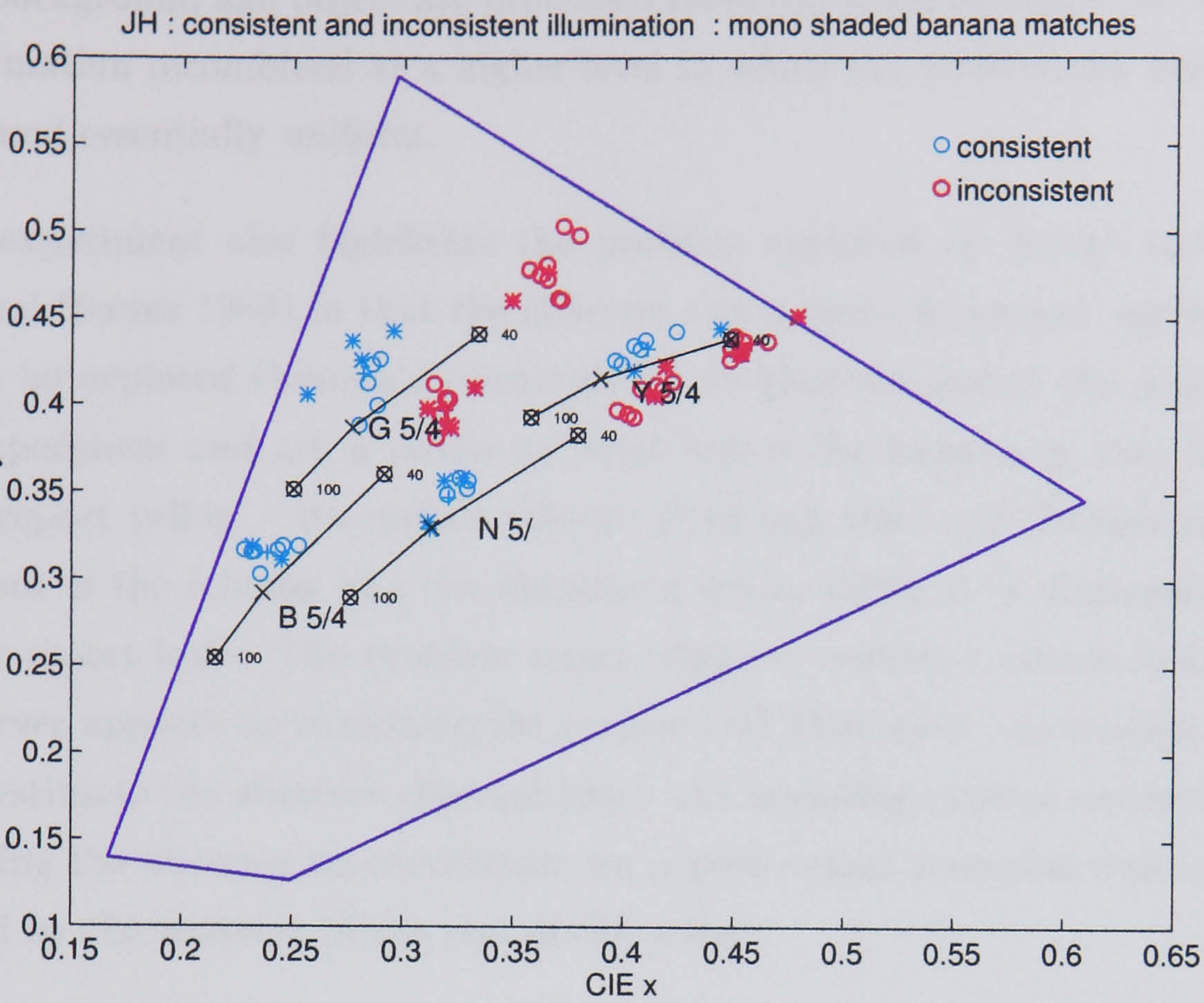

Figure 7.18: Matches for mono banana object under consistent and inconsistent illumination conditions for observer JH.

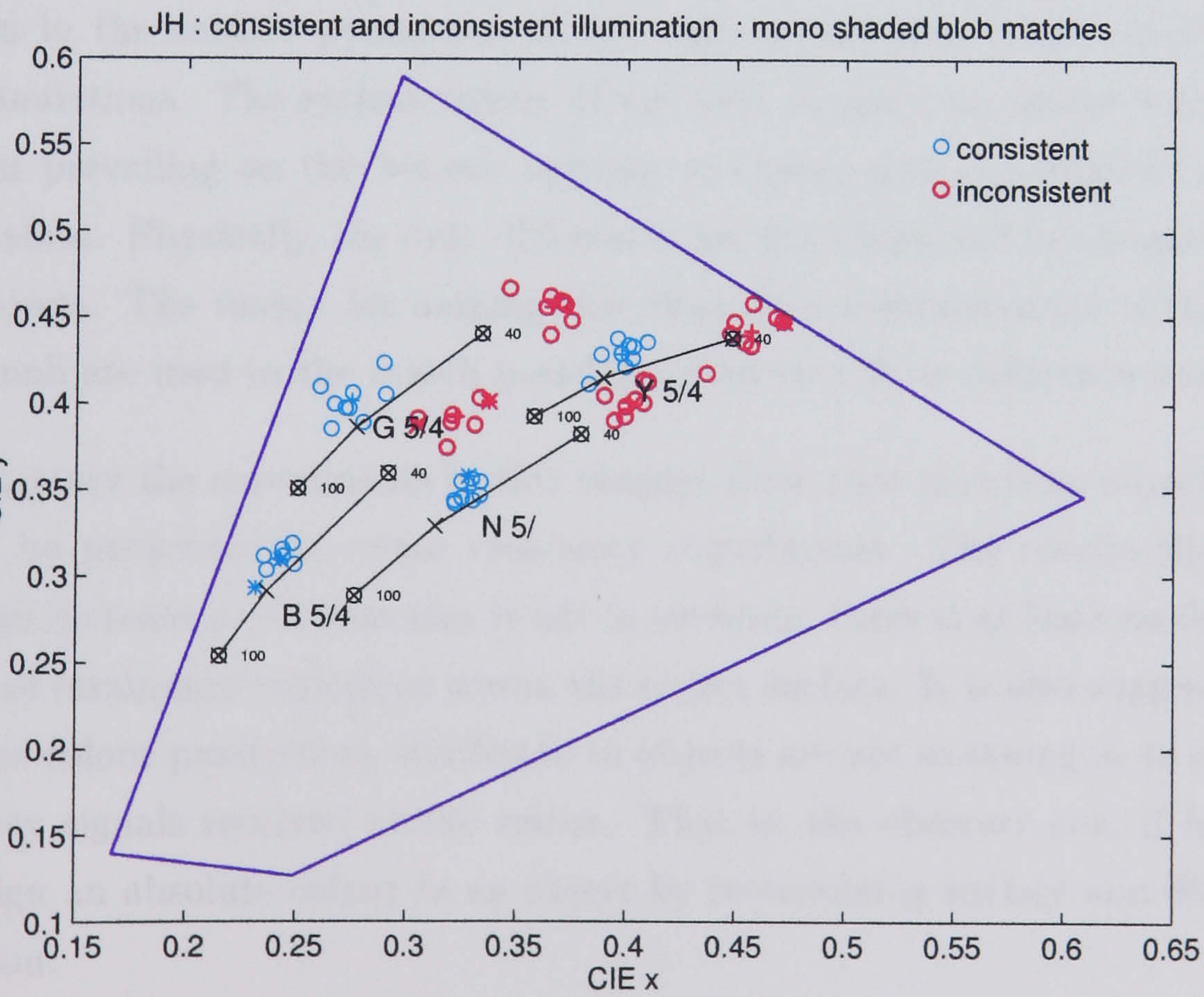

Figure 7.19: Matches for mono clay blob object under consistent and inconsistent illumination conditions for observer JH.

that the background and object are processed relatively independently and illumination information recombined at a higher level in which the multi-patch background is considered essentially uniform.

This experiment also highlights the problem explored by Arend and Reeves (Arend and Reeves 1986) in that the observer can operate in several 'modes'. This point can be explored through an example: if we generate one of the stimuli used in this experiment and ask a passer-by what colour the banana is, they are most likely to report yellow – its surface colour. If we ask the more detailed question, what colour is the banana and the illuminant under which it is displayed, people can easily report both. The problem arises when we request a match to be made. The observer appears to recombine the surface and illuminant information and objectively estimate the absolute chromaticity – the matching process encourages this by requiring the observer to concentrate on a pure colour sensation and not to be influenced by the contents of the rest of the scene.

An interesting 'demonstration' can be produced by manipulation of the illuminants and surface colours for the banana and blob objects. Four scenes are created, two with the banana, and two with the blob. For certain illuminant/surface pairs a division in the surface perception of the banana and blob results in the mono object simulations. The surface colour of the blob appears to change whereas the illuminant prevailing on the banana appears to change and the surface colour remain constant. Physically, the only differences are the shape and luminance profiles of the objects. The reason for naming this example a demonstration is that if the same stimuli are used in the match paradigm then very little difference results.

In summary the experiments in this chapter show that matching objects to objects can be performed in colour constancy experiments. The results show that, whereas an increase in performance is not in evidence, there is at least no detrimental effect of luminance variations across the object surface. It is also suggested that the surface colour perceptions attributed to objects are not so strong as to outweigh the sensory signals received at the retina. That is, the observer can, if he so desires, assign an absolute colour to an object by recombining surface and illuminant information.

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N $5/_{ban}$	D_{40}	D_{65}	0.322	0.342	—	11.11_1	0.324	0.343	0.00457	88.89_8
Y $5/4_{ban}$	D_{40}	D_{65}	—	—	—	—	0.403	0.422	0.00392	77.78_7
G $5/4_{ban}$	D_{40}	D_{65}	0.25	0.406	—	11.11_1	0.256	0.408	0.00915	88.89_8
B $5/4_{ban}$	D_{40}	D_{65}	—	—	—	—	0.229	0.31	0.00879	100_9
N $5/_{blob}$	D_{40}	D_{65}	0.329	0.347	0.00515	22.22_2	0.326	0.344	0.00582	77.78_7
Y $5/4_{blob}$	D_{40}	D_{65}	—	—	—	—	0.405	0.423	0.0097	100_9
G $5/4_{blob}$	D_{40}	D_{65}	—	—	—	—	0.253	0.41	0.00941	100_9
B $5/4_{blob}$	D_{40}	D_{65}	—	—	—	—	0.221	0.313	0.00579	100_9
N $5/_{ban}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	—	—	—	—	0.427	0.415	0.0226	100_6
Y $5/4_{ban}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	0.477	0.449	0.00495	50_3	0.477	0.449	0.00391	50_3
G $5/4_{ban}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	—	—	—	—	0.368	0.48	0.00539	100_6
B $5/4_{ban}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	—	—	—	—	0.321	0.407	0.00749	100_6
N $5/_{blob}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	0.424	0.419	0.00708	50_3	0.424	0.419	0.00426	50_3
Y $5/4_{blob}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	0.476	0.453	0.00505	100_6	—	—	—	—
G $5/4_{blob}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	0.375	0.486	0.00458	66.67_4	0.368	0.482	0.0187	33.33_2
B $5/4_{blob}$	$D_{100}^{obj} : D_{40}^{back}$	D_{65}	0.324	0.397	—	16.67_1	0.322	0.406	0.00571	83.33_5

Table 7.6: Match statistics for object matches under consistent and inconsistent illumination conditions for observer TE

JH	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
N $5/_{ban}$	D_{40}	D_{65}	0.326	0.353	0.00503	44.44_4	0.316	0.34	0.0142	55.56_5
Y $5/4_{ban}$	D_{40}	D_{65}	0.407	0.432	0.00775	77.78_7	0.424	0.432	0.0231	22.22_2
G $5/4_{ban}$	D_{40}	D_{65}	0.286	0.409	0.0132	55.56_5	0.278	0.427	0.0167	44.44_4
B $5/4_{ban}$	D_{40}	D_{65}	0.239	0.314	0.00919	77.78_7	0.239	0.314	0.00774	22.22_2
N $5/_{blob}$	D_{40}	D_{65}	0.328	0.348	0.00686	88.89_8	0.328	0.358	—	11.11_1
Y $5/4_{blob}$	D_{40}	D_{65}	0.398	0.428	0.00908	100_9	—	—	—	—
G $5/4_{blob}$	D_{40}	D_{65}	0.275	0.401	0.0125	100_9	—	—	—	—
B $5/4_{blob}$	D_{40}	D_{65}	0.244	0.312	0.00683	77.78_7	0.237	0.301	0.00986	22.22_2
N $5/_{ban}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.409	0.404	0.0112	77.78_7	0.416	0.413	0.00891	22.22_2
Y $5/4_{ban}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.455	0.434	0.00753	77.78_7	0.466	0.44	0.0167	22.22_2
G $5/4_{ban}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.368	0.478	0.014	88.89_8	0.349	0.459	—	11.11_1
B $5/4_{ban}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.317	0.398	0.00944	55.56_5	0.321	0.395	0.0113	44.44_4
N $5/_{blob}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.401	0.401	0.00921	100_9	—	—	—	—
Y $5/4_{blob}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.455	0.441	0.0138	88.89_8	0.472	0.447	—	11.11_1
G $5/4_{blob}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.366	0.456	0.0088	100_9	—	—	—	—
B $5/4_{blob}$	$D^{obj}_{100} : D^{back}_{40}$	D_{65}	0.323	0.392	0.0117	77.78_7	0.321	0.394	0.0179	22.22_2

Table 7.7: Match statistics for object matches under consistent and inconsistent illumination conditions for observer JH

Chapter 8

Experimental Results: Colour Constancy in Real Life Scenes

8.1 Introduction

In some respects, the colour constancy performance shown in experiments is less than would be expected from the performance levels we seem to experience in everyday life. The effects of high level scene interpretation on surface perception were examined in chapter 4. In those experiments, seemingly minor changes to the scene could cause radically different surface percepts.

The experiments in this chapter address the possibility that aspects of real life scenes which are usually omitted from standard experiments have a large effect on the colour constancy performance of observers. With this goal in mind, the stimuli used in the experiments in this chapter are depictions of real life scenes. The effects of illumination are not simulated – they are recorded from real scenes under real illuminants. This avoids the assumptions present in scene simulation and adds many features which are usually avoided, such as: mutual illumination, shadows, specular highlights, recognisable objects against recognisable backgrounds etc. Several computational studies (Funt, Drew and Ho 1991; Healey 1991; Lee 1986) suggest that the information present in these scene 'features' could assist in the calculation and compensation of the illumination.

8.2 Object matching in real scenes

The procedure used to draw the mono objects in the previous chapter was developed further to allow the adjustment of objects within a scene. A 24 bit camera (with all automatic gain and offset adjustment disabled after white balancing to daylight) is used to take pictures of the real life scenes under different illumination conditions. This image is then processed[1] to allow it to be used in the calibrated system, in which resolution is reduced to around 115 unique linear levels per gun. Stencils are created for each object for which we require to adjust the colour. The stencils and images can be manipulated to change the properties of all or part of the image. The manipulations that can be performed include:

- xyY shifting – an area of the image has the xyY value of each pixel calculated and a constant added to each variable, after which the new value is converted back to calibrated RGB.
- LMS manipulation – the LMS values of each pixel are calculated and multiplied by a value specified for each cone class by the experimenter.
- mono image – every pixel is converted into xyY space, and the Y value is used with a user specified xy to create new xyY triples that are converted to calibrated RGB.

The experimental program has the capability to perform quad–linear interpolation from four input images, which may be created by one or several of the above manipulations. These images are used to specify a parametric solution space for the observer to traverse. The program uses the fact that XYZ and calibrated RGB spaces are linearly related. This allows the traversal of a vector in XYZ space (and

[1] In radiometric terms the images produced are not exact depictions of the real scenes. The camera produces 3 'images' corresponding to 'red', 'green' and 'blue'. These images are created from integration of the incident spectra (from a point in the scene) after filtering much in the same way as the cones collect light. The three images produced should be a linear transformation of the corresponding CIE XYZ images. This 'calibrating' transformation is not performed and hence each point in the images on screen may not be a metamer of the same point in the real scene. This was not considered a major problem as the results of all the experiments are only considered with respect to each other and not in absolute terms. The reason for not calibrating the images is that transforming an image that is quantised in one 3 dimensional space can cause a large loss in resolution when transformed into another. The small offsets due to differing primaries was considered better than unpredictable fluctuations in contrast in parts of the image.

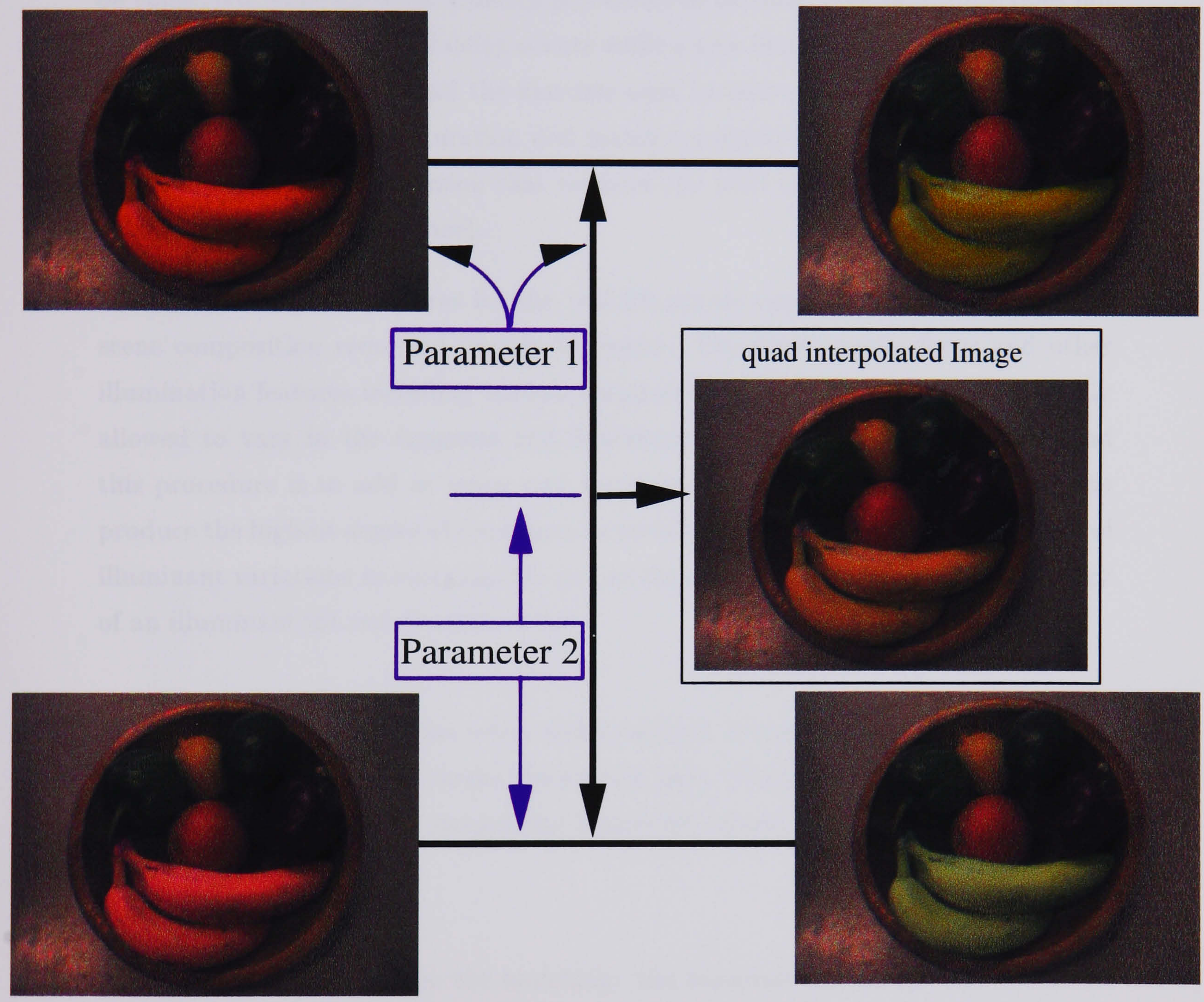

Figure 8.1: The procedure used to create an interpolated image from four examples

hence xyY space) through traversal of an equivalent vector in RGB space. From careful choice of border images we can generate a continuum of images in which each pixel in a given area can be adjusted in terms of hue and saturation, for example. Figure 8.1 shows schematically the procedure used to create a new image from the four samples. The advantage of this technique is that colour contrast can be conserved while allowing a match procedure to be used. The procedure does not operate in real time; a small delay occurs while a new image is generated. Observers soon adjust to this delay and the matches tend to take around the same time as a full luminance, hue and saturation dial match (probably due to the reduced search space in the parametric version that removes the need to examine what should be redundant areas of xyY space).

When taking the pictures for the real life scenes care was taken to ensure the scene composition remained exactly the same. Illuminant angle, focus and other illumination features including mutual illumination and shadows were intentionally allowed to vary in the tungsten and fluorescent images. The motivation behind this procedure is to add as many real world features as possible in an attempt to produce the highest degree of constancy as possible. We easily deal with these sort of illuminant variations in everyday life and so the exact luminance spread and energy of an illuminant are not considered here.

Three images are used: one taken under daylight (Figure 8.2), one under tungsten (Figure 8.3) and the third under fluorescent light (Figure 8.4). Figure 8.5 shows the xy space gamut of the images and Figure 8.6 shows the luminance histograms of the images.

Three objects are used in the matching: the bananas, the orange and the green pepper. Figure 8.7 shows the CIE xy gamuts and luminance profiles of the bananas under the three illumination conditions. Figures 8.8 and 8.9 show the same information for the orange and green pepper.

It has been suggested (Results in Chapters 6 and 7; Shepherd and Whittle 1995; Wandell 1993) that observer matches can be predicted from considering the cone ratios of the stimulus against its background. Foster and Nascimento (Foster and Nascimento 1994) performed a large number of simulations, using realistic illuminant

Figure 8.2: Real life scene under daylight illumination

Figure 8.3: Real life scene under tungsten illumination

Figure 8.4: Real life scene under fluorescent illumination

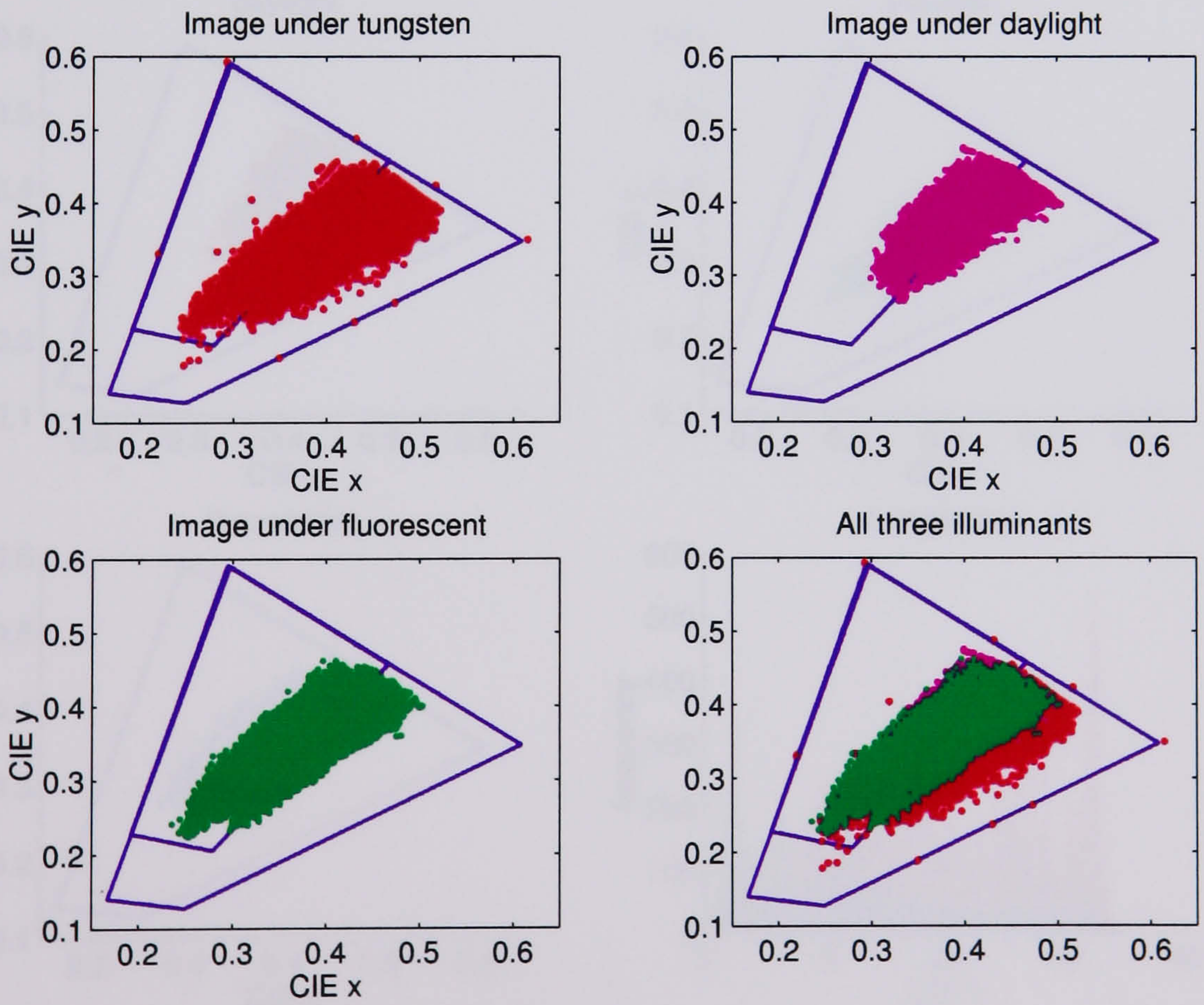

Figure 8.5: CIE xy plane information for the real life scenes

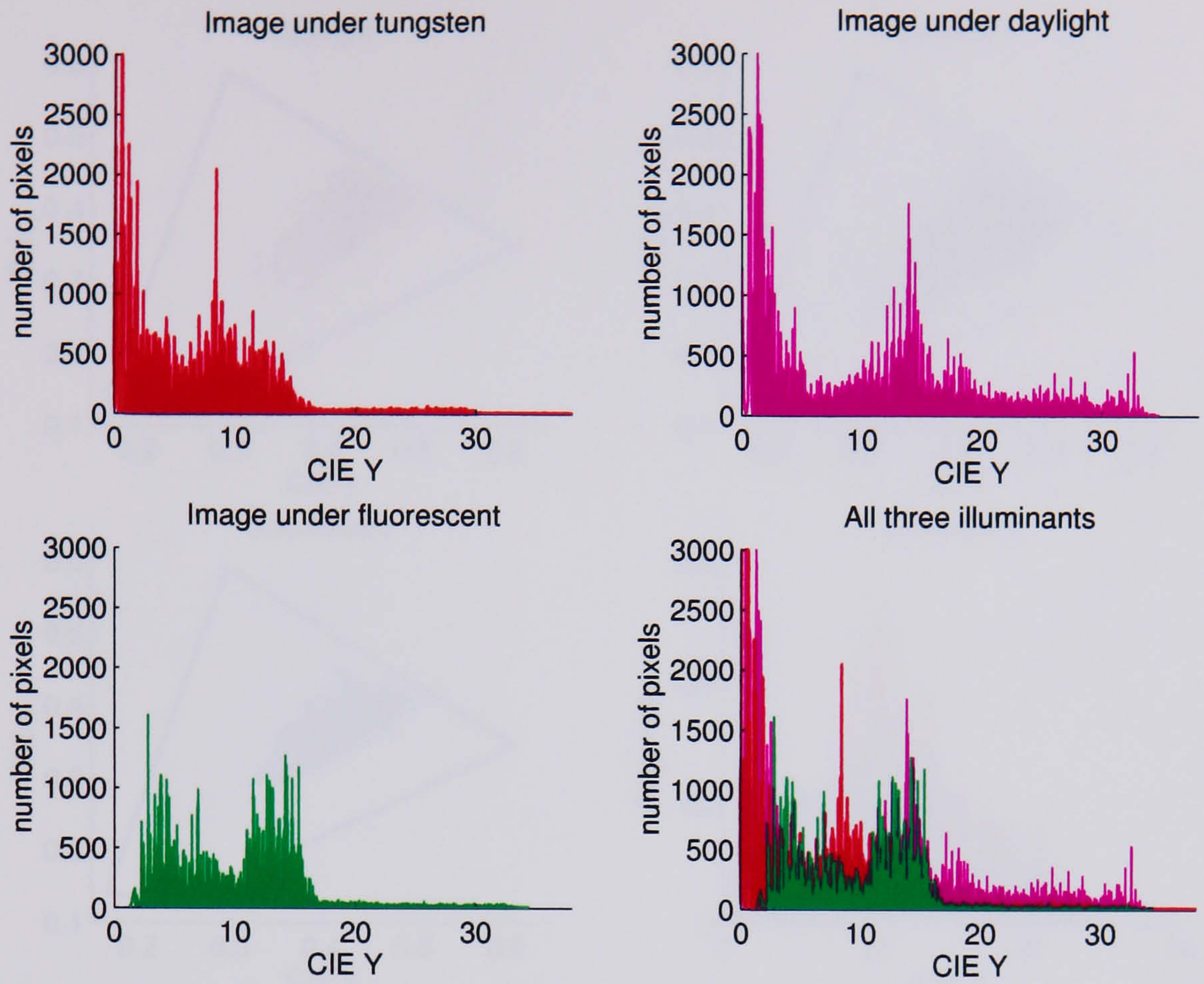

Figure 8.6: Luminance histograms for the real life scenes

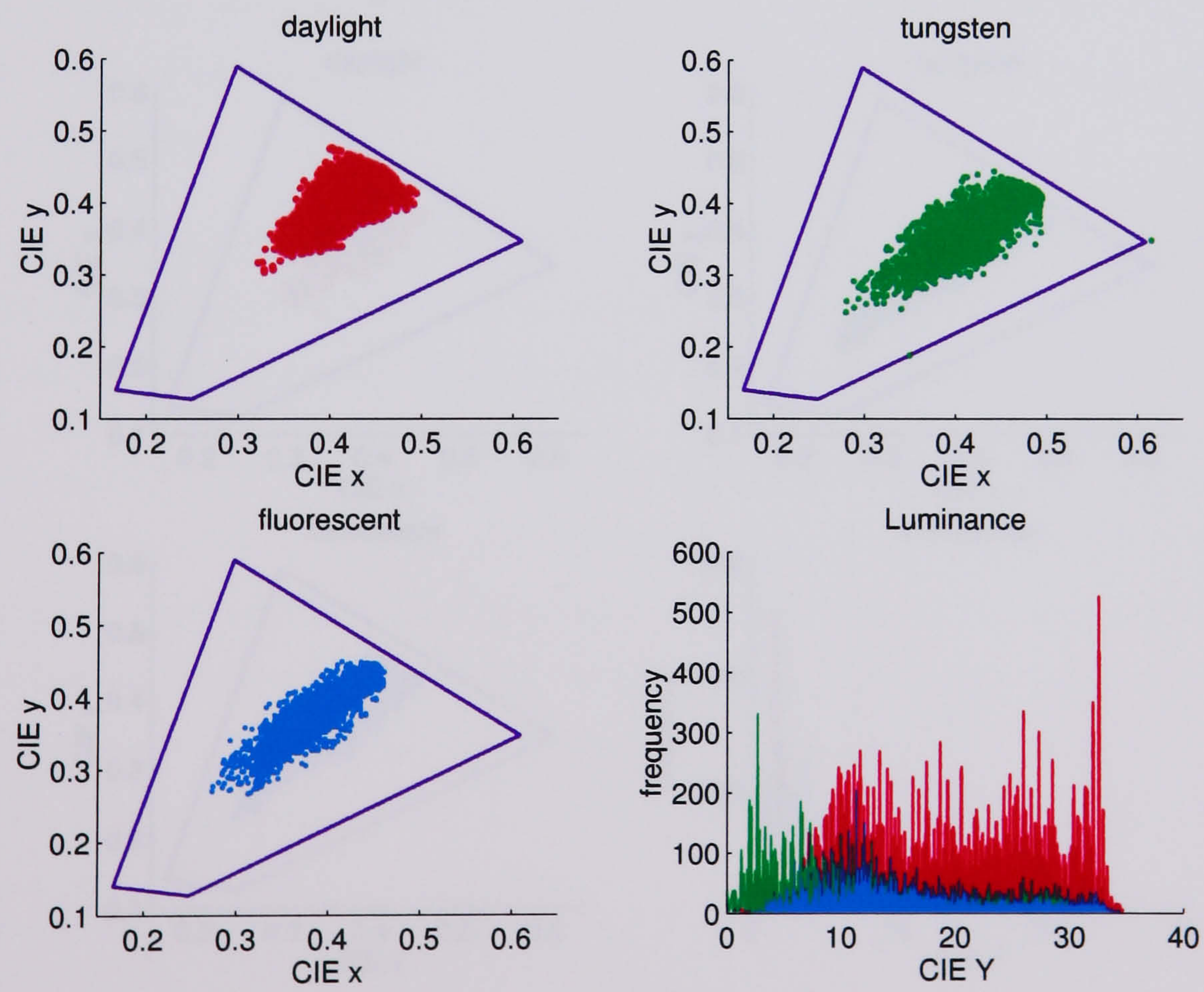

Figure 8.7: Colorimetric information for the bananas under various illumination

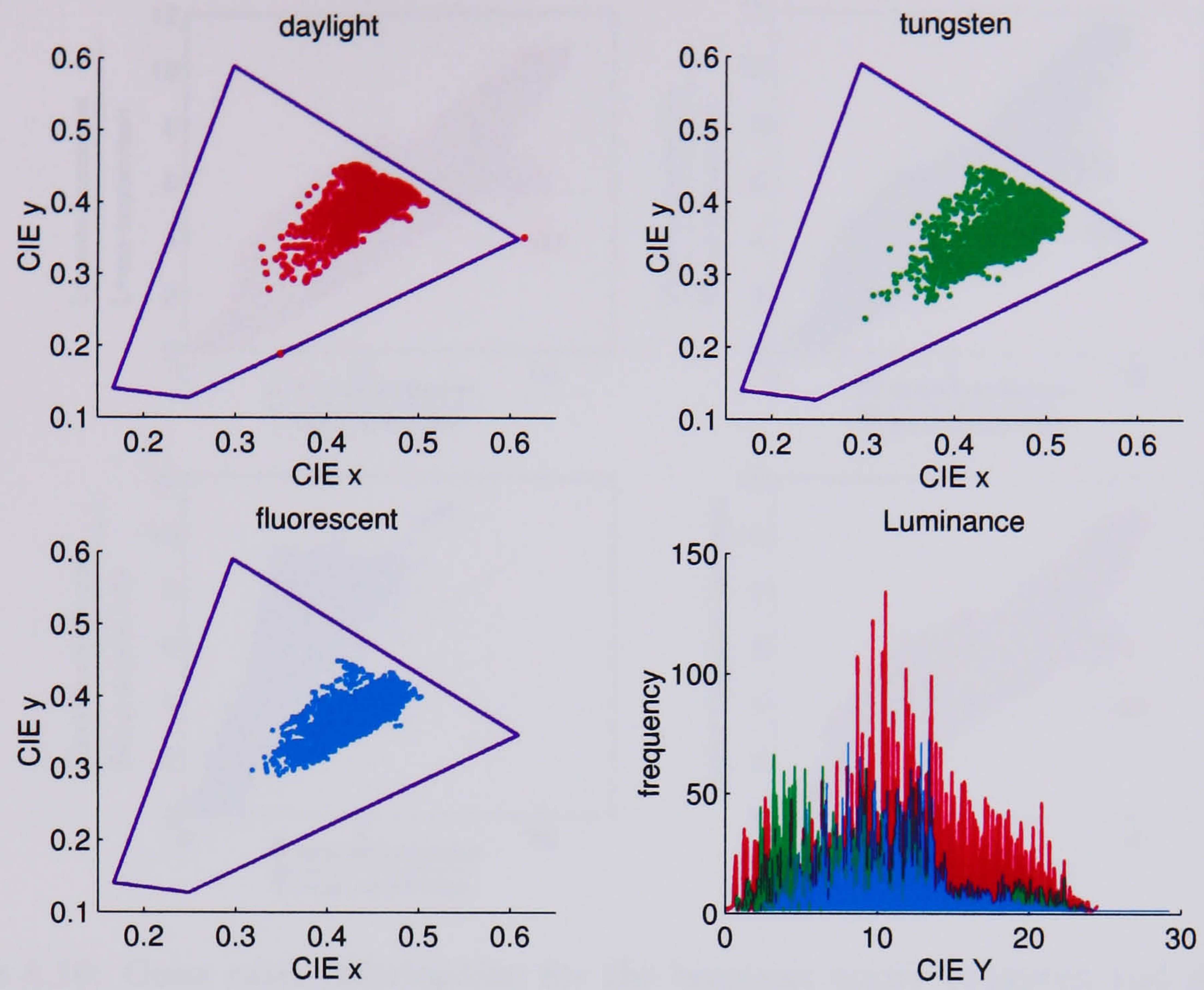

Figure 8.8: Colorimetric information for the orange under various illumination

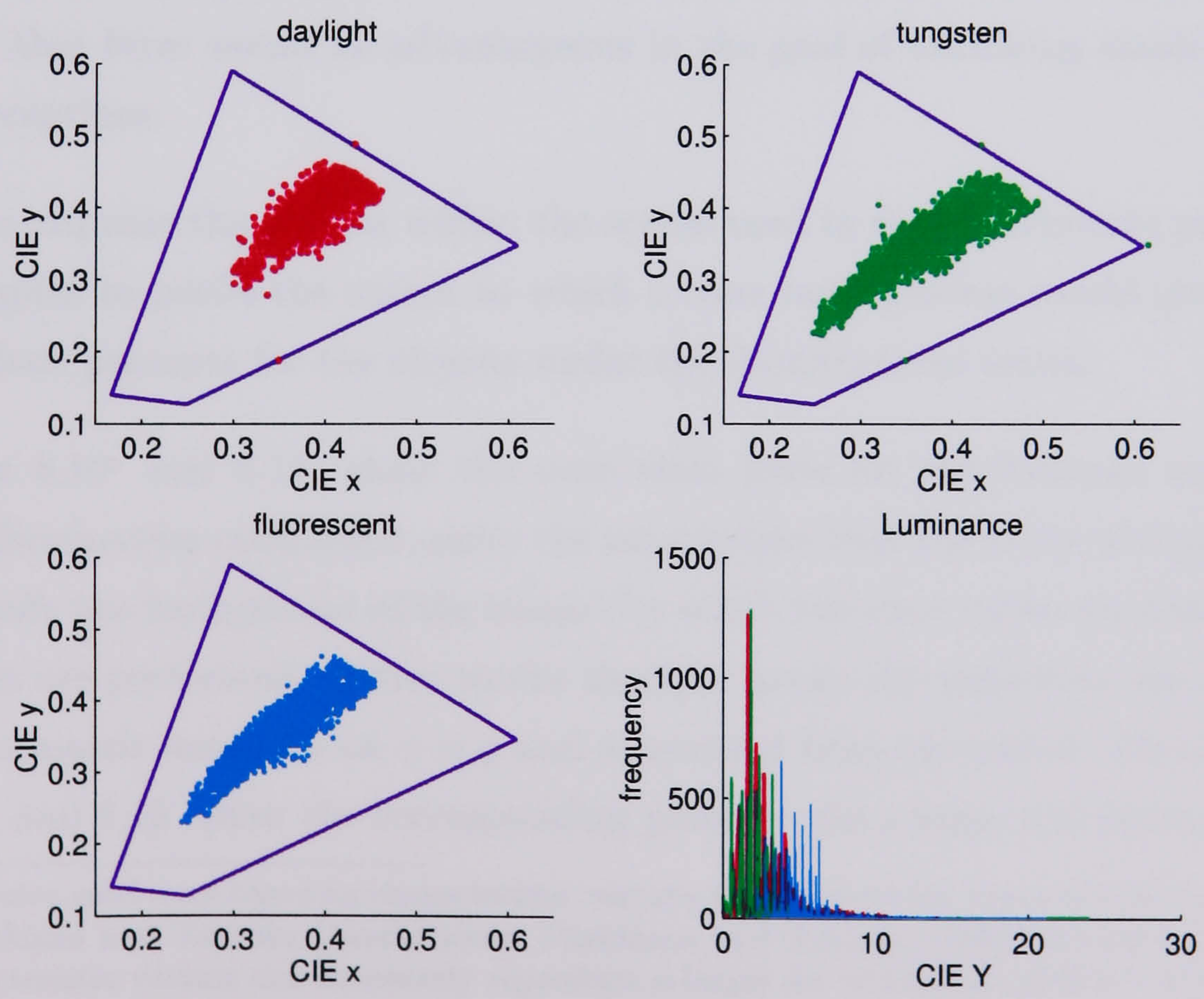

Figure 8.9: Colorimetric information for the green pepper under various illumination

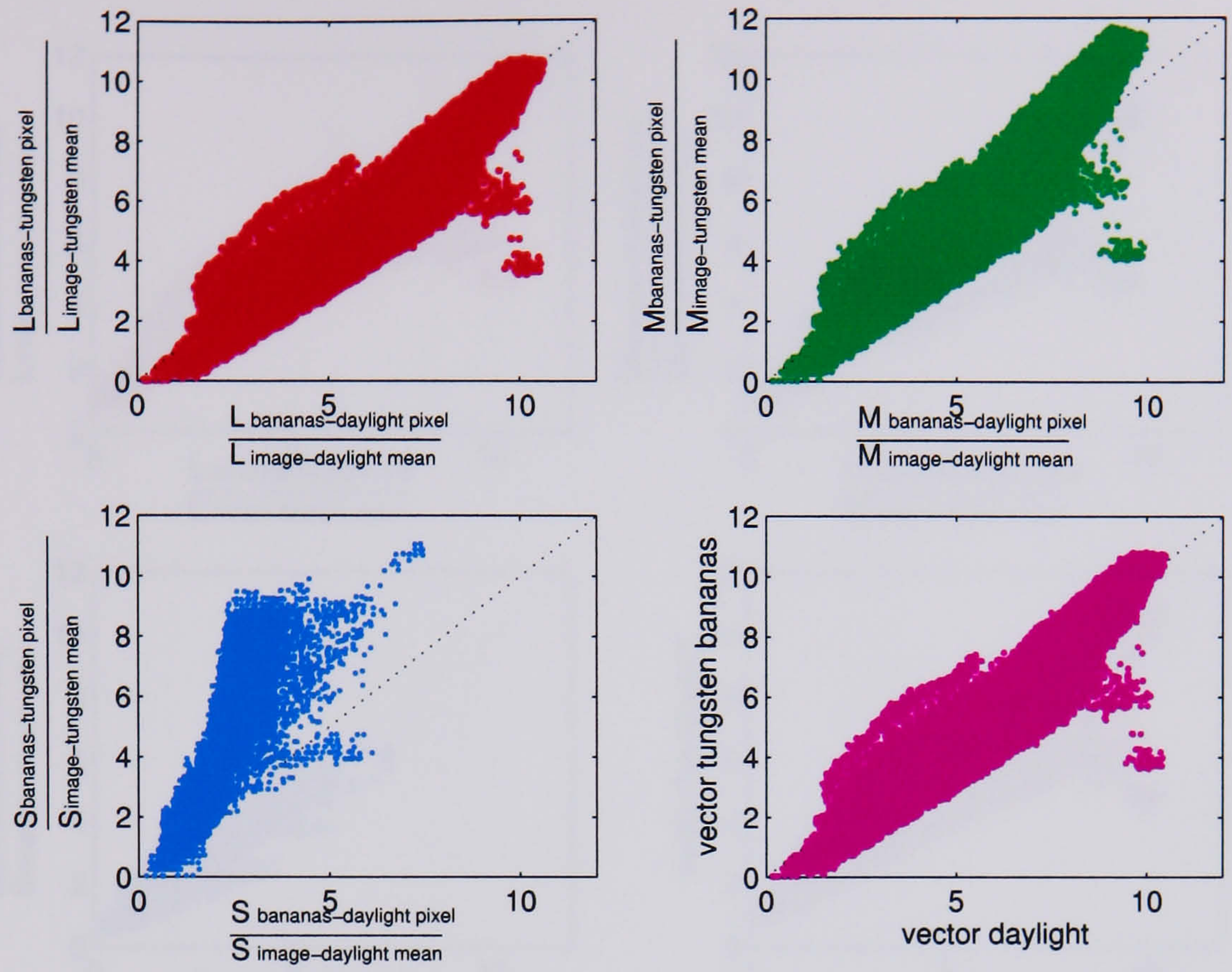

Figure 8.10: Cone ratio information for the bananas under tungsten and daylight

and surface spectra[2], into the stability of the cone ratio prediction as a measure of obtaining surface properties from a scene. They concluded that a normalisation process of that form would be advantageous in the goal of obtaining stable surface colour perceptions.

We can examine the objects within the scenes used in the experiments presented in this chapter to probe the extent to which a cone ratio process would give stable surface colour percepts for the objects under the illuminations tested.

Figures 8.10 and 8.11 show the cone ratio plots for the bananas under the different illumination conditions under the assumption that the scene average LMS values specify the background of the image (by which the cone ratios are scaled) and all matches are performed relative to the daylight image (by definition the daylight to daylight match would lie on $x = y$ and is omitted from the plots). Figures 8.12, 8.13, 8.14, and 8.15 show the corresponding plots for the orange and green pepper.

[2]The spectra used were based on characteristic vectors derived from the Munsell spectra. Jaaskelainen, Parkkinen and Toyooka (Jaaskelainen, Parkkinen and Toyooka 1990) have shown that this set of characteristic vectors can accurately reproduce a larger set of natural coloured objects, such as flowers, leaves and berries.

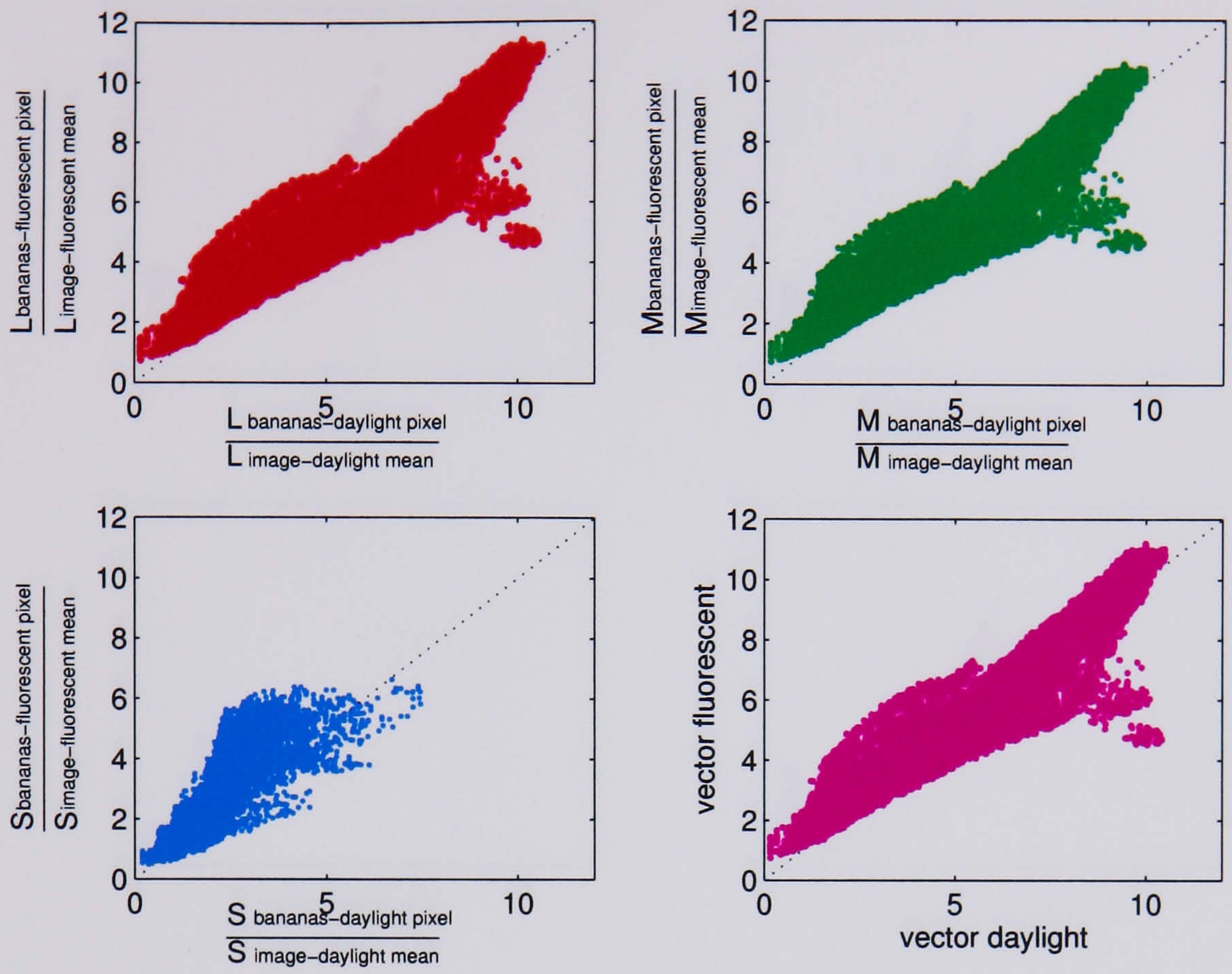

Figure 8.11: Cone ratio information for the bananas under fluorescent and daylight

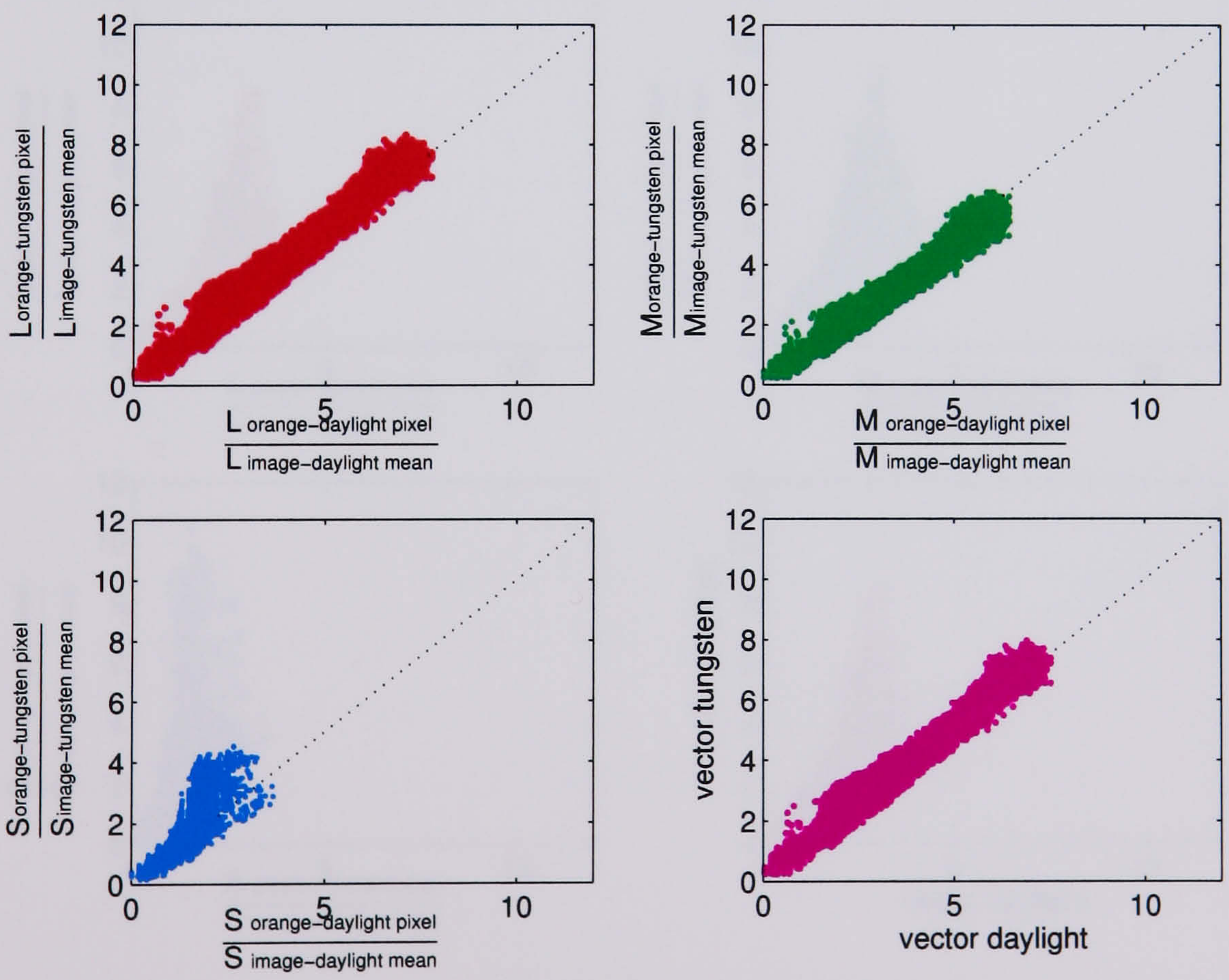

Figure 8.12: Cone ratio information for the orange under tungsten and daylight

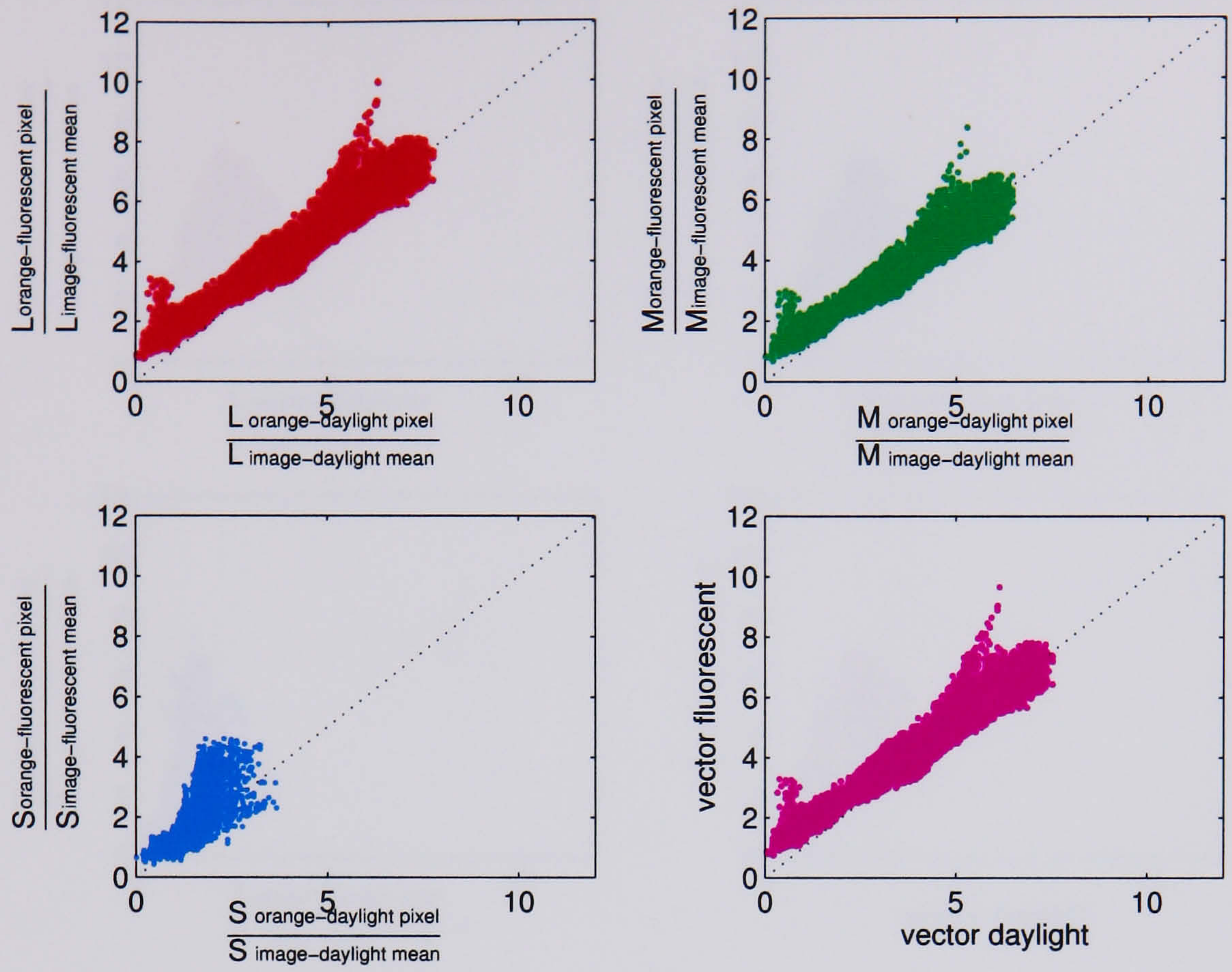

Figure 8.13: Cone ratio information for the orange under fluorescent and daylight

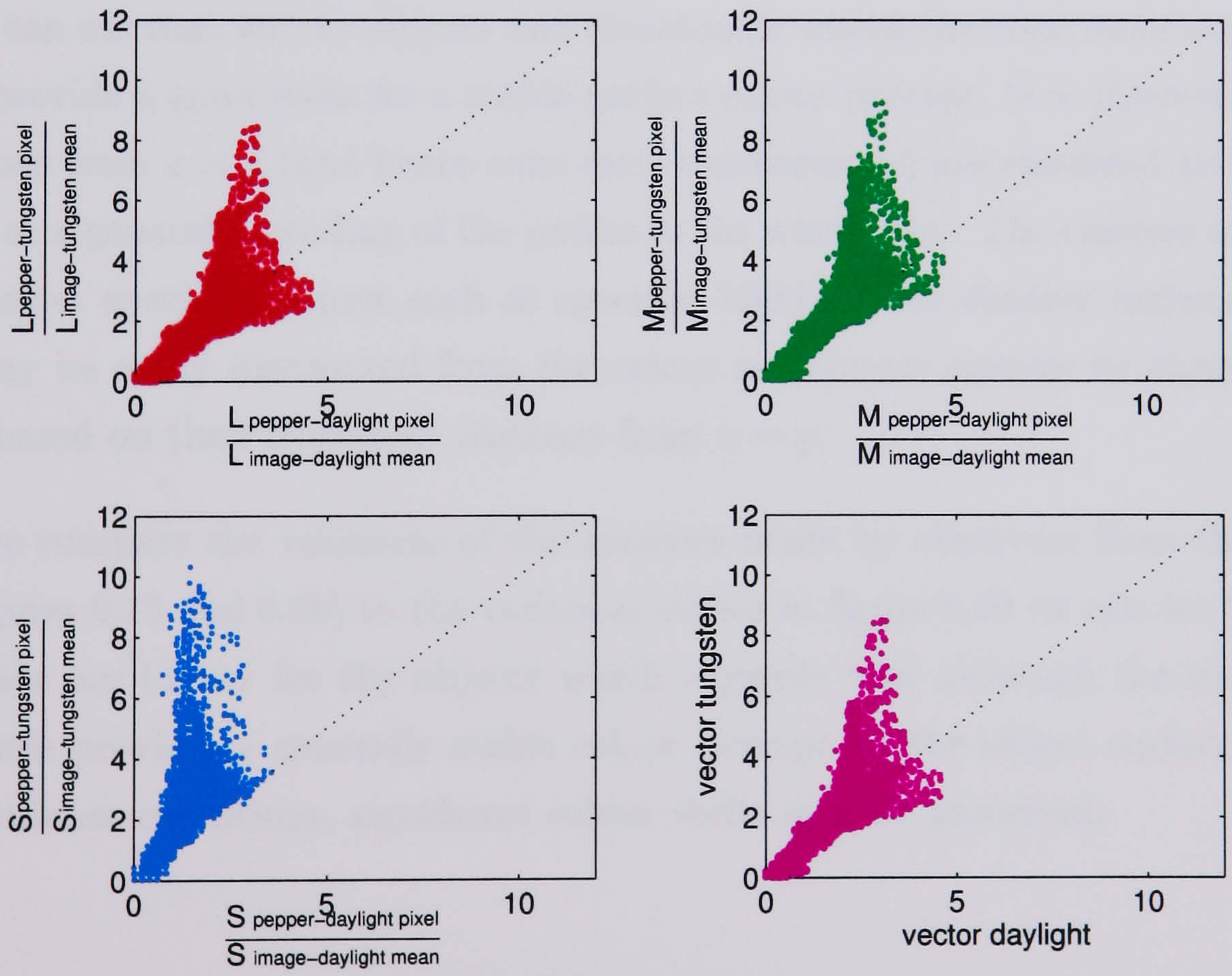

Figure 8.14: Cone ratio information for the green pepper under tungsten and daylight

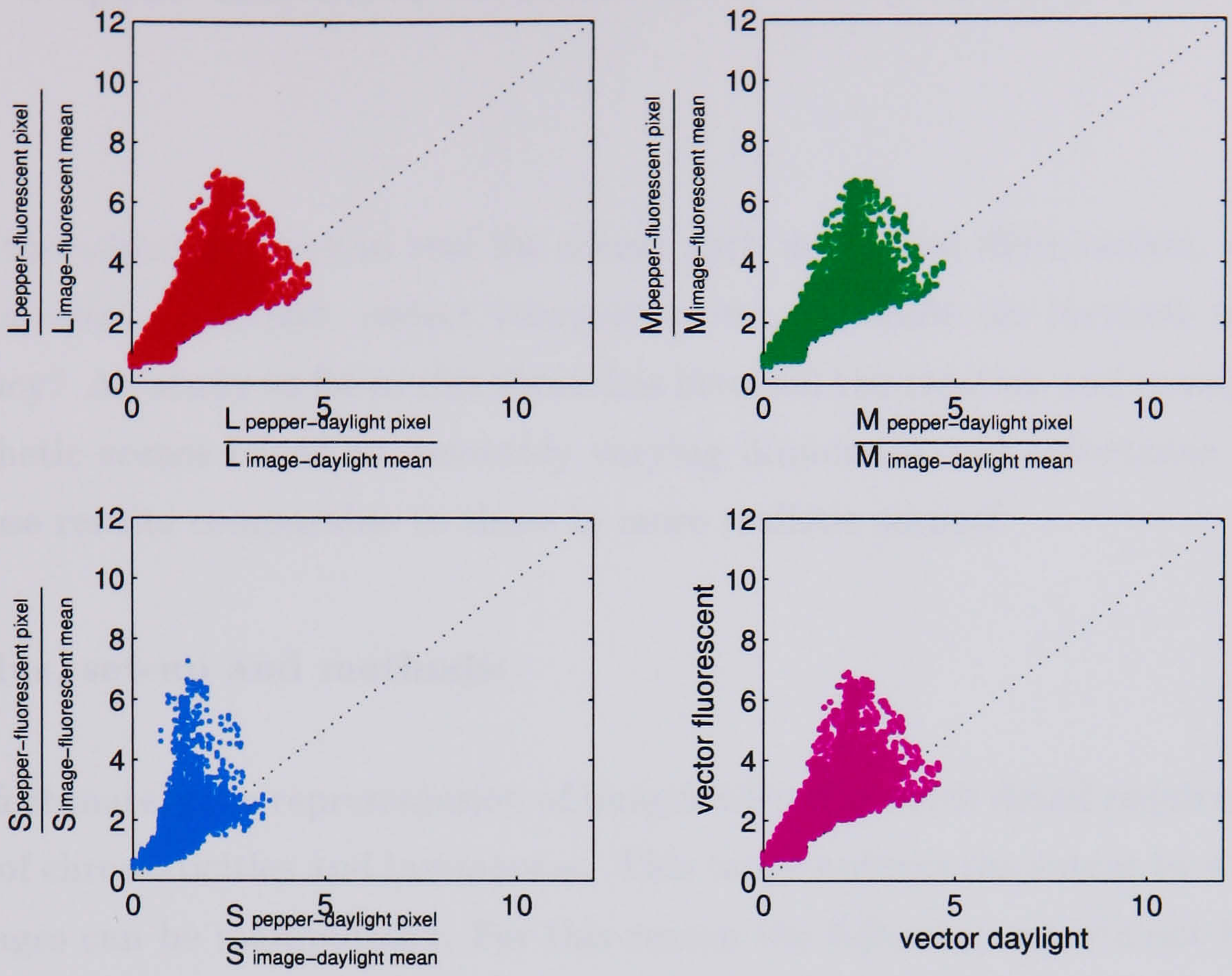

Figure 8.15: Cone ratio information for the green pepper under fluorescent and daylight

We can see that for the objects and illuminants tested the cone ratio hypothesis would provide a good basis for a stable surface colour percept. It is interesting that deviations from $x = y$ (and hence cone ratio conformance) are clustered and do not appear as a general spreading of the points in the whole plot. The clusters are likely to represent specific features such as specular highlights or shadow regions and as such may be easily discounted from the colour assessment process by thresholding points based on their Euclidean distance from $x = y$.

If we compare the variances of the matches made by observers from the $x = y$ line (figures 6.13 and 6.20) to the variances shown in figure 8.10 we can see that the deviations are higher for the objects which suggests that although the cone ratio hypothesis provides a generally stable colour percept for the object surfaces across the illuminant conditions, significant colour shifts may be perceived.

8.3 Object matches of fruit in a real life scene

Aims:

Do the added features in real life scenes such as mutual illumination, coloured shadows, colour contrast, object recognition etc., promote an increase in colour constancy? All study so far in this thesis has involved the creation and manipulation of synthetic scenes based on smoothly varying illuminant and reflectance spectra. Are these results comparable to those in more realistic scenes?

Stimulus, set-up and methods:

Unfortunately the representation of images with this much detail requires a large gamut of chromaticities and luminances. This range reduces the extent to which the full images can be manipulated. For this reason the following experiment has been set up to 'disprove a hypothesis'. The hypothesis is that all of the extra features in the scene should produce a very high degree of colour constancy. If this is the case then the matches for each of the cases should cluster in the same area of the solution space. If the theory does not hold then the matches will be shifted and may be unobtainable with the solution spaces supplied. The experiment in its current form, might not yield the exact match points for the objects under the different illumination conditions.

Four images were created from the daylight image that allowed easily noticeable hue and saturation changes in the test object. Figure 8.16 shows these images for the bananas, Figure 8.17 the images for the orange and Figure 8.18 the images for the green pepper. The corresponding plots in xy space are shown in figures 8.19, 8.20 and 8.21. The observer uses two dials controlling the parametric traversal of the solution space to match an object in the match presentation to that shown in the test presentation.

The observers made matches to the three objects under differing test image illumination conditions; an image under daylight to an image under daylight (Figure 8.22), an image under tungsten light to an image under daylight (Figure 8.23) and an image under fluorescent light to an image under daylight (Figure 8.24). The solution spaces are set up for the daylight match and hence, if a high degree of

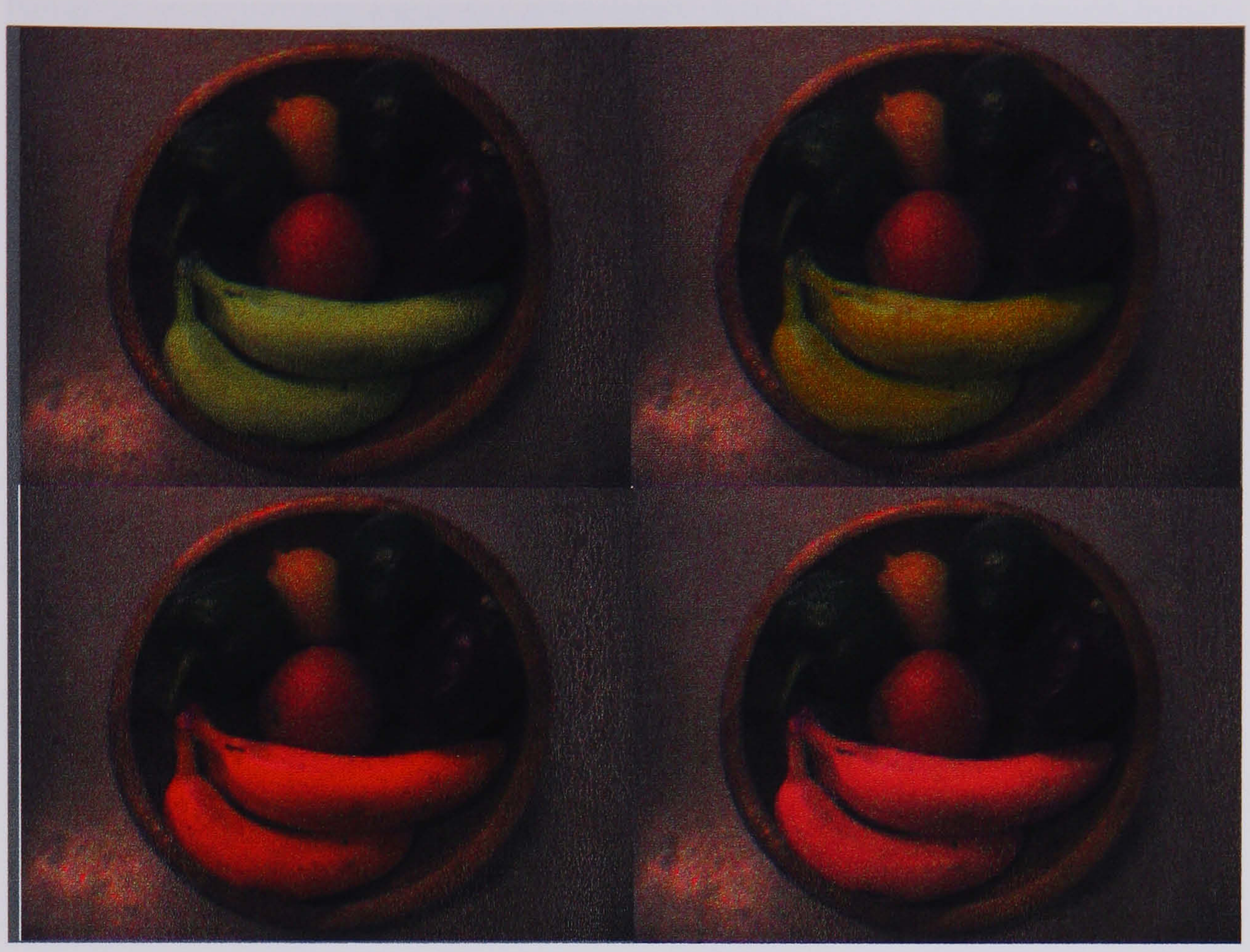

Figure 8.16: Four images used in the quad interpolation of the bananas

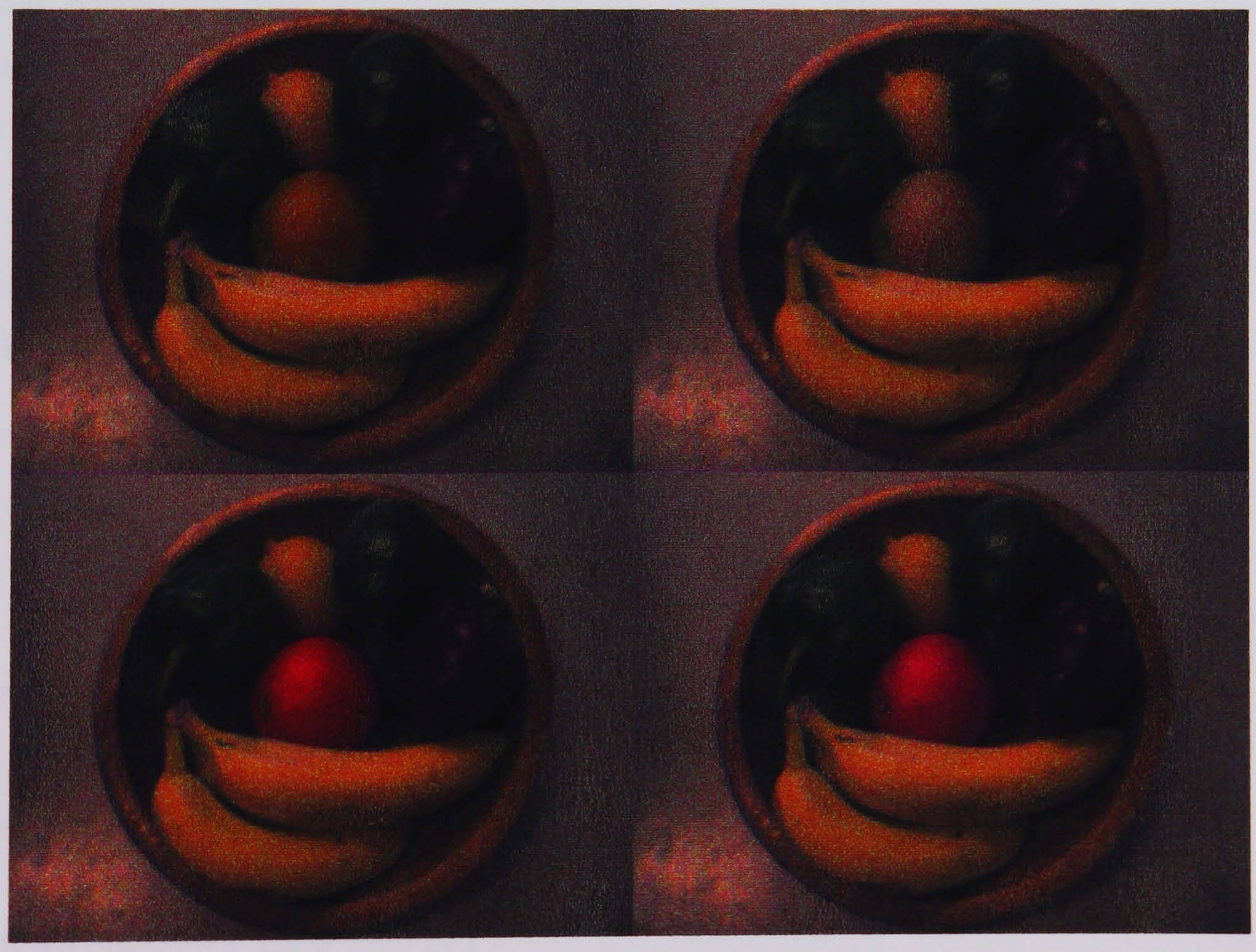

Figure 8.17: Four images used in the quad interpolation of the orange

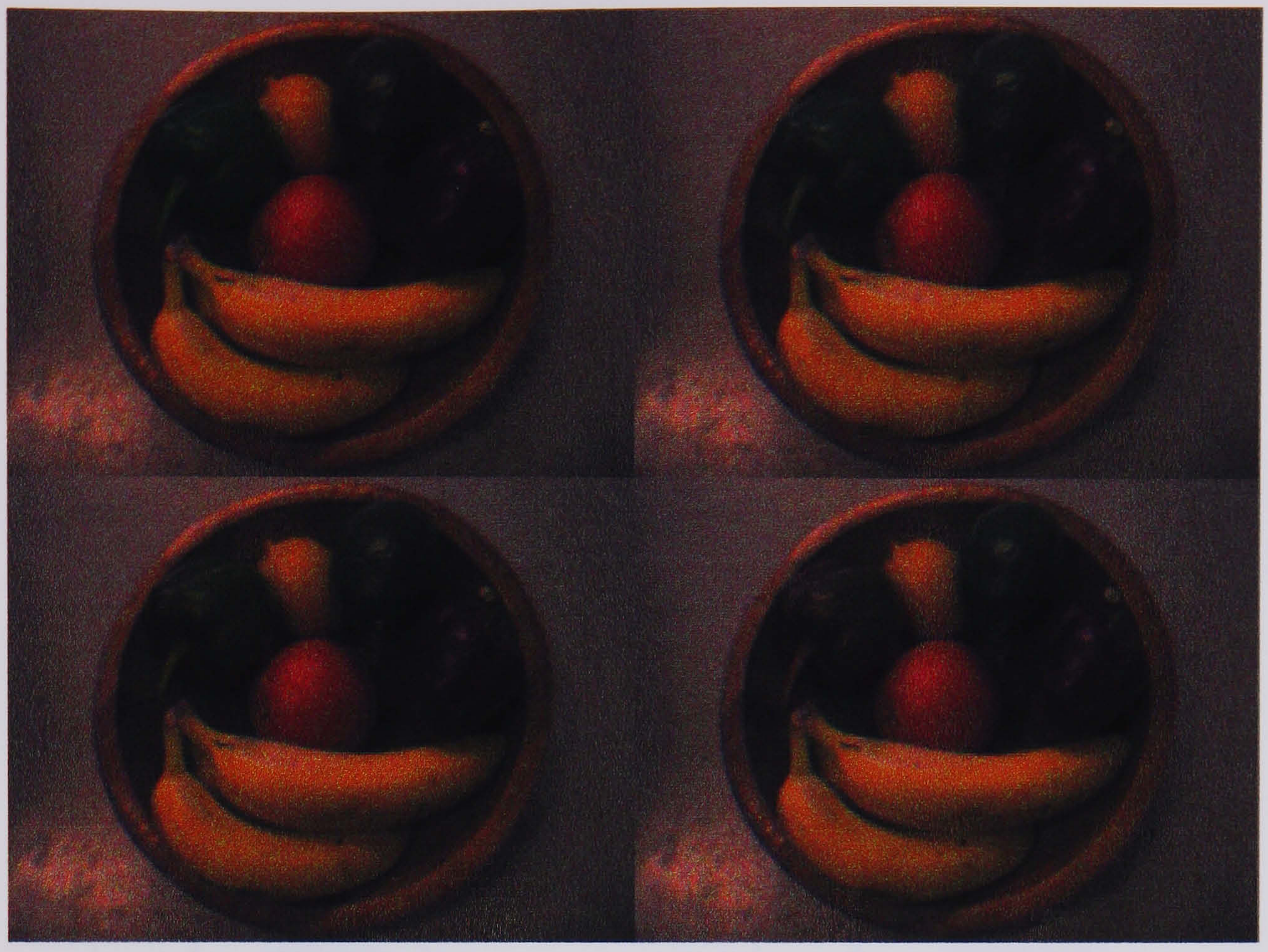

Figure 8.18: Four images used in the quad interpolation of the green pepper

colour constancy does not result then the observer may be unable to complete a match. The observer is instructed that, if a match does not seem possible then set the dials as close as can be obtained and press the 'LOW' confidence button.

Only a single illuminant condition is tested in a given run and the observer is forced to adapt for two minutes prior to any matching. The adapting field is a uniform field with the same xyY co-ordinates as the mean of the real scene image.

To derive the contribution of the new technique to the variance in the matches the observers performed a patch equivalent to the daylight – daylight match. Figure 8.25 shows an example of the stimulus. The patches displayed correspond to the mean chromaticities of the object images tested (all normalised to 6 cd/m^2). The procedure and instructions remain exactly the same.

Results:

Two observers made matches in this set of experiments, TE and JH (Both female, experienced in colour matching and naive to the purposes of the experiment).

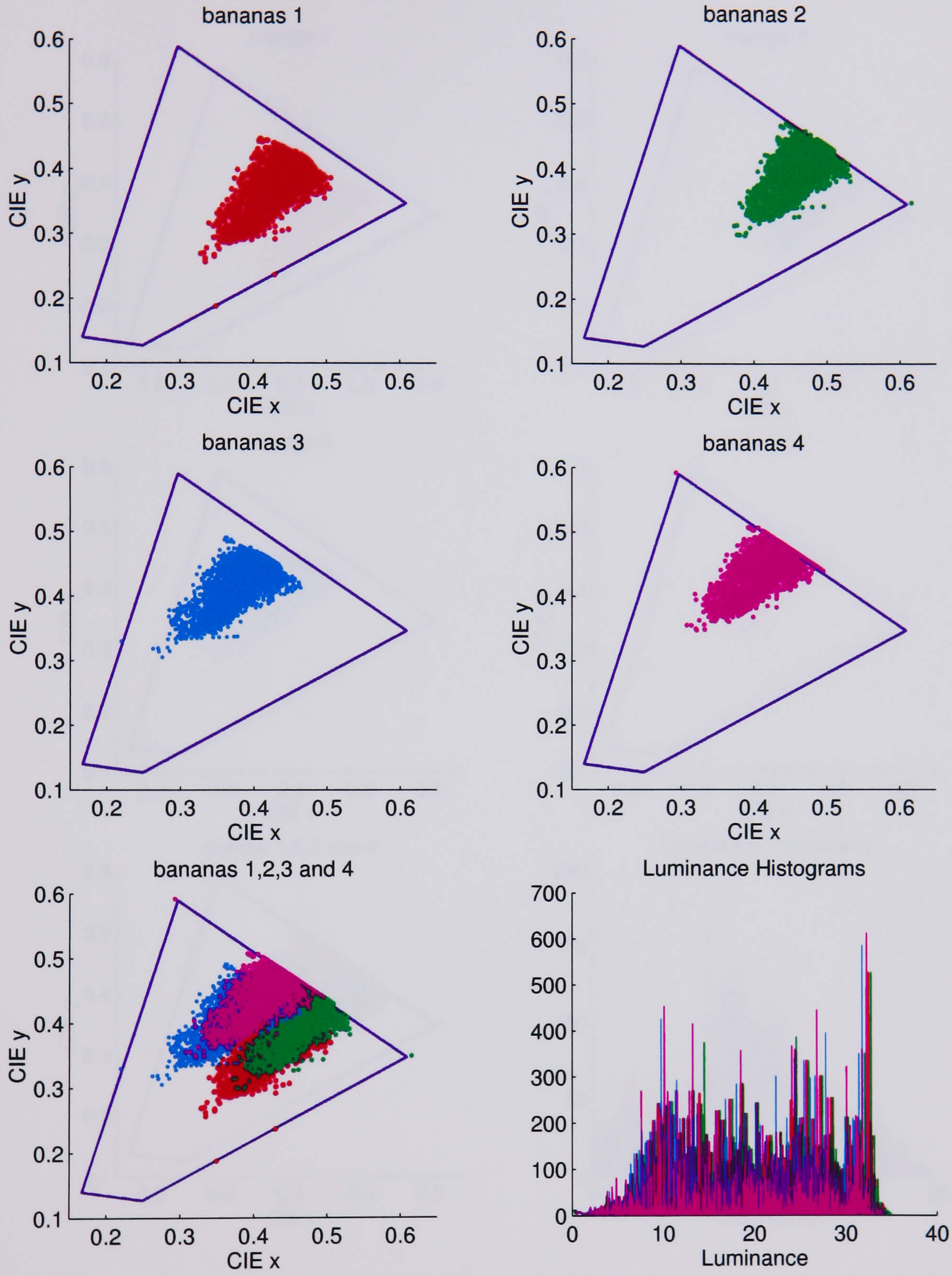

Figure 8.19: Colorimetric information for the four images used in the quad interpolation of the bananas

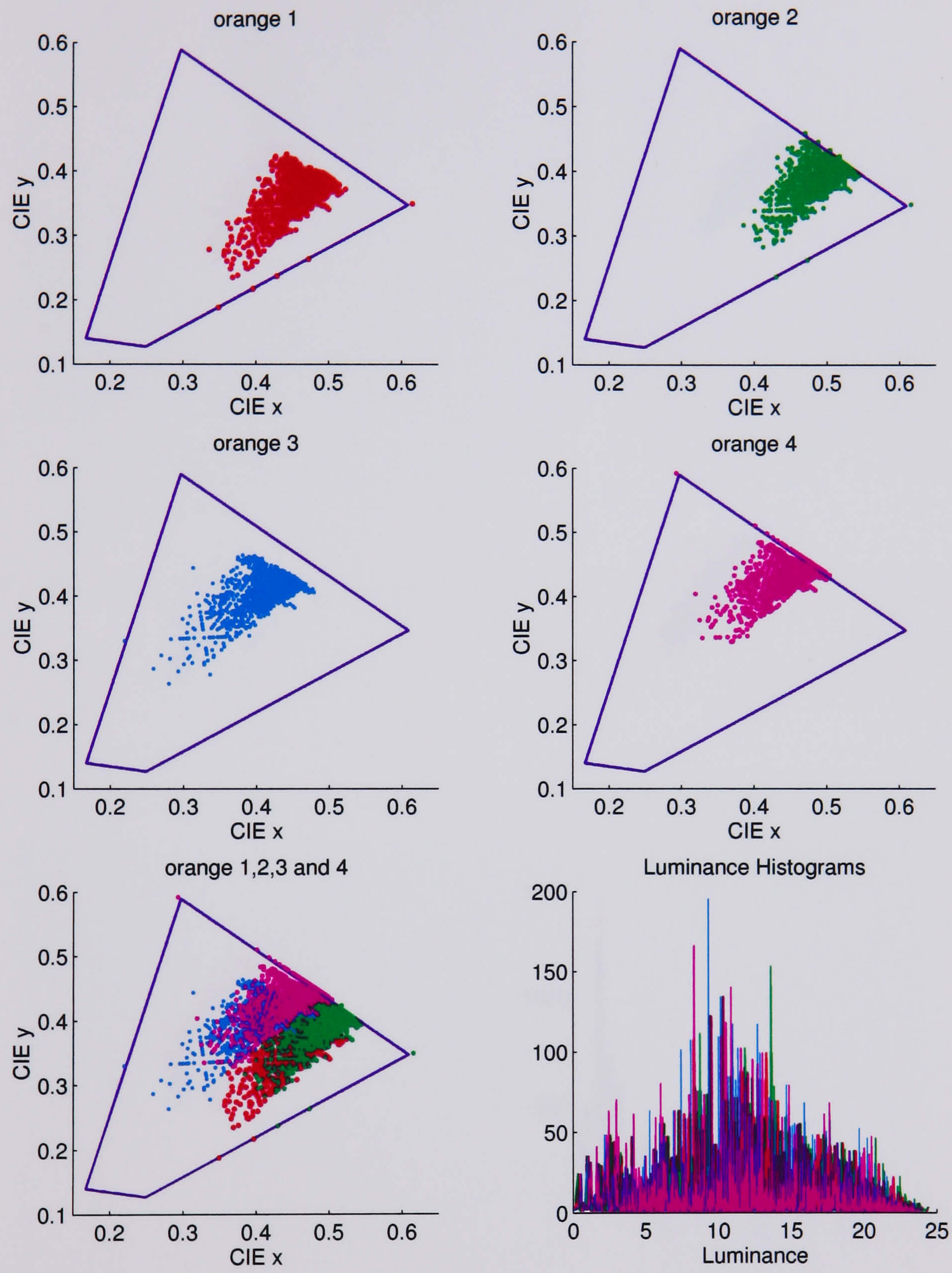

Figure 8.20: Colorimetric information for the four images used in the quad interpolation of the orange

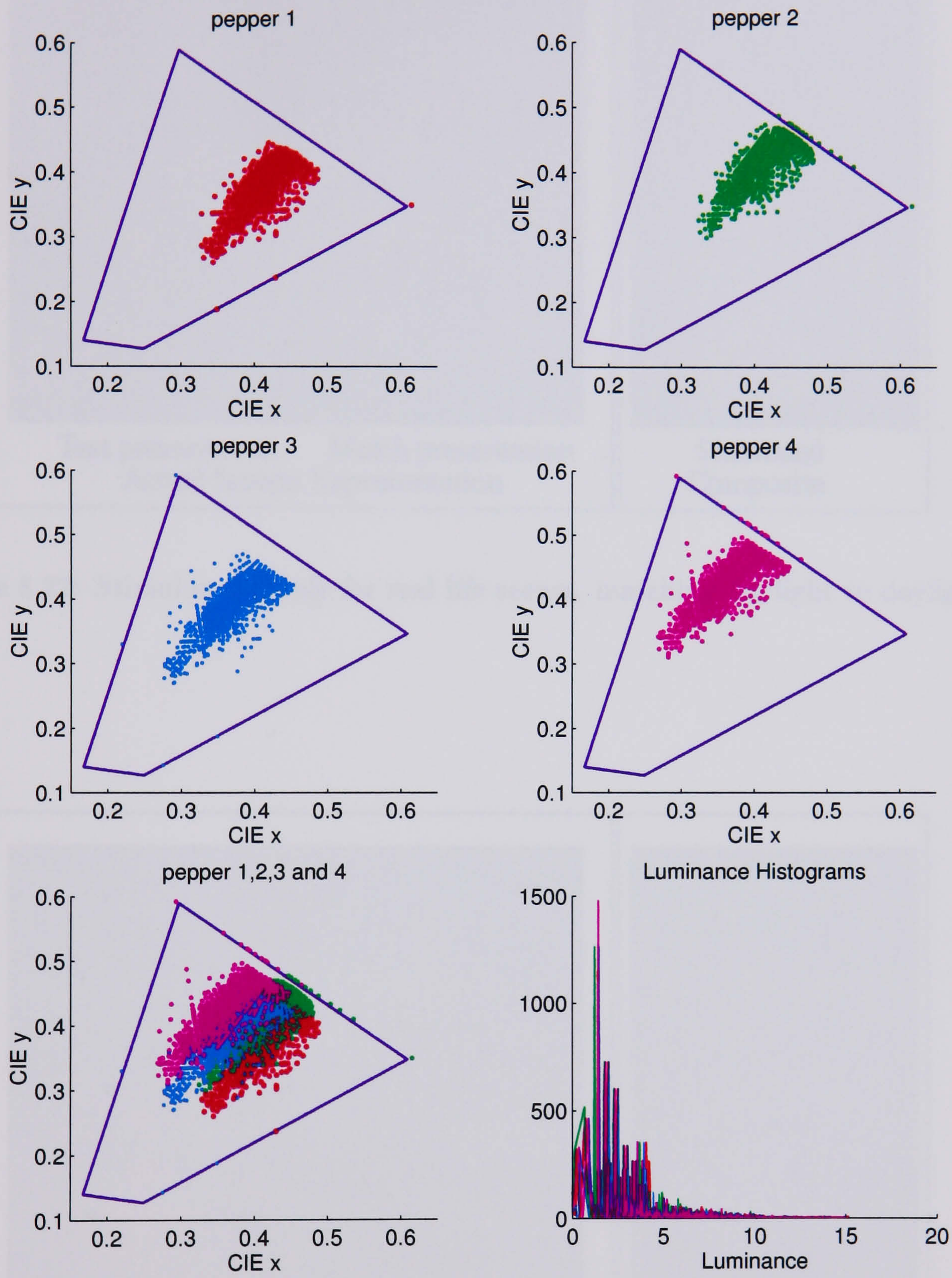

Figure 8.21: Colorimetric information for the four images used in the quad interpolation of the green pepper

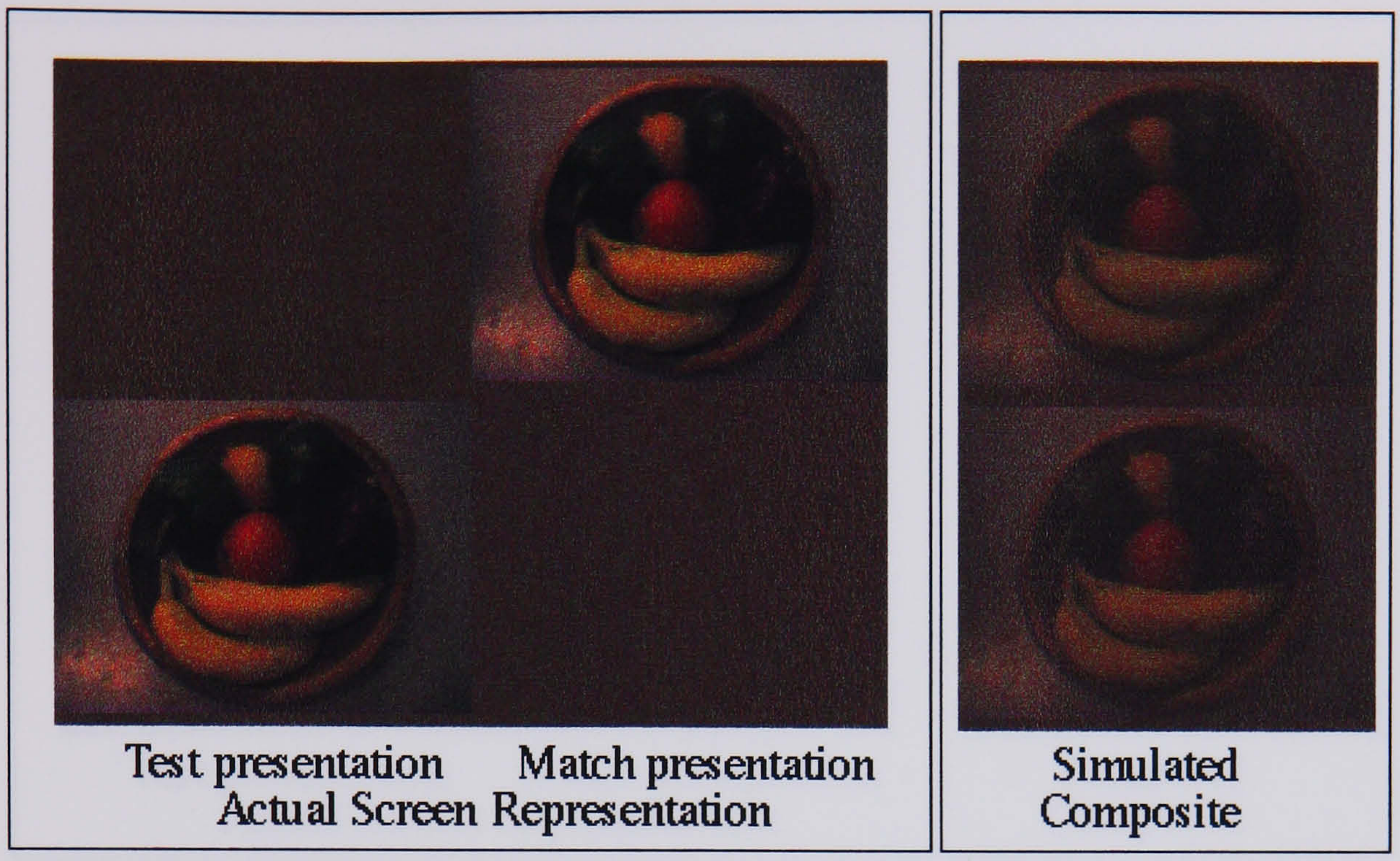

Figure 8.22: Stimulus example for real life scenes, matching daylight to daylight

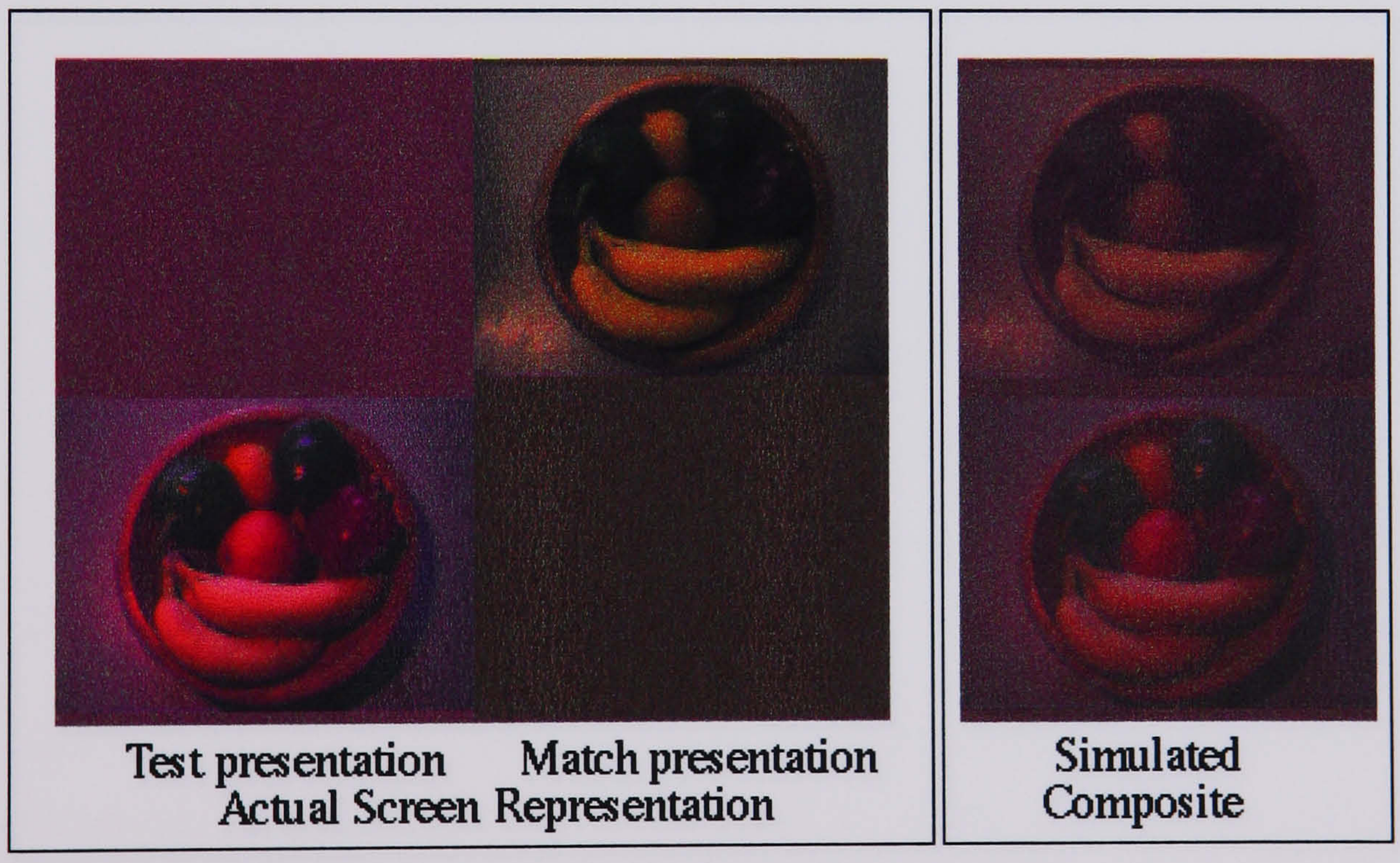

Figure 8.23: Stimulus example for real life scenes, matching tungsten to daylight

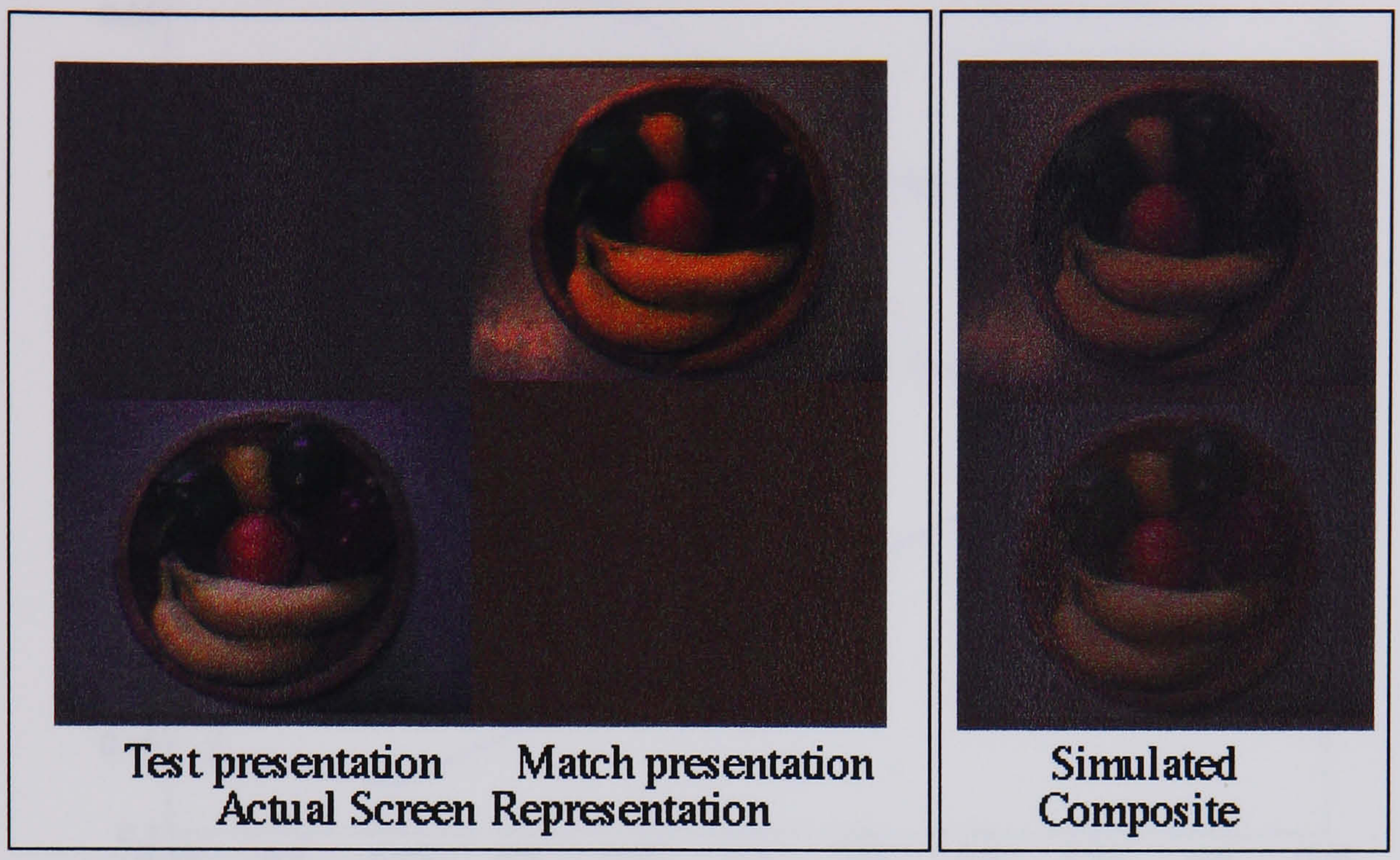

Figure 8.24: Stimulus example for real life scenes, matching fluorescent to daylight

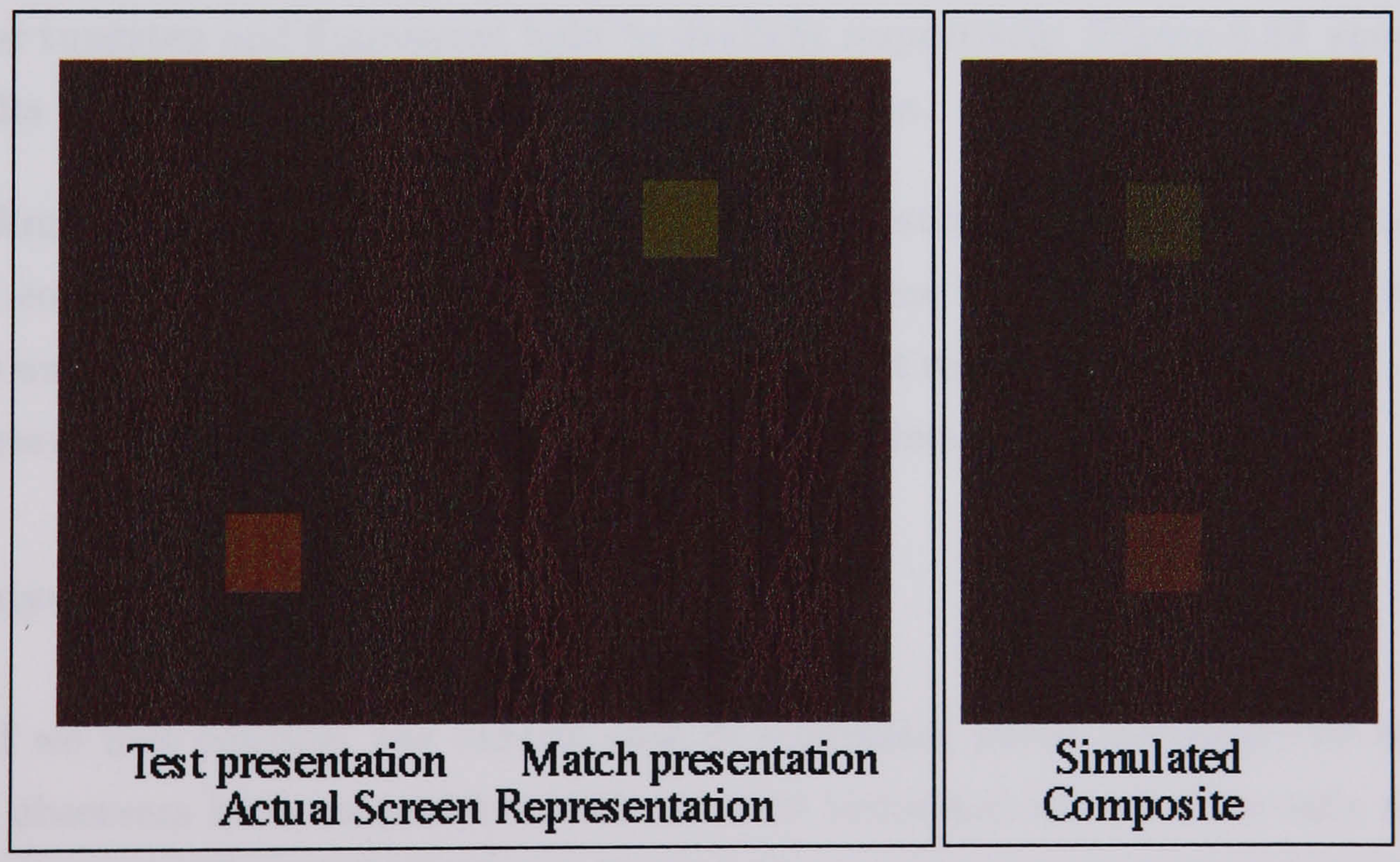

Figure 8.25: Stimulus example for matching daylight scene - patch equivalents

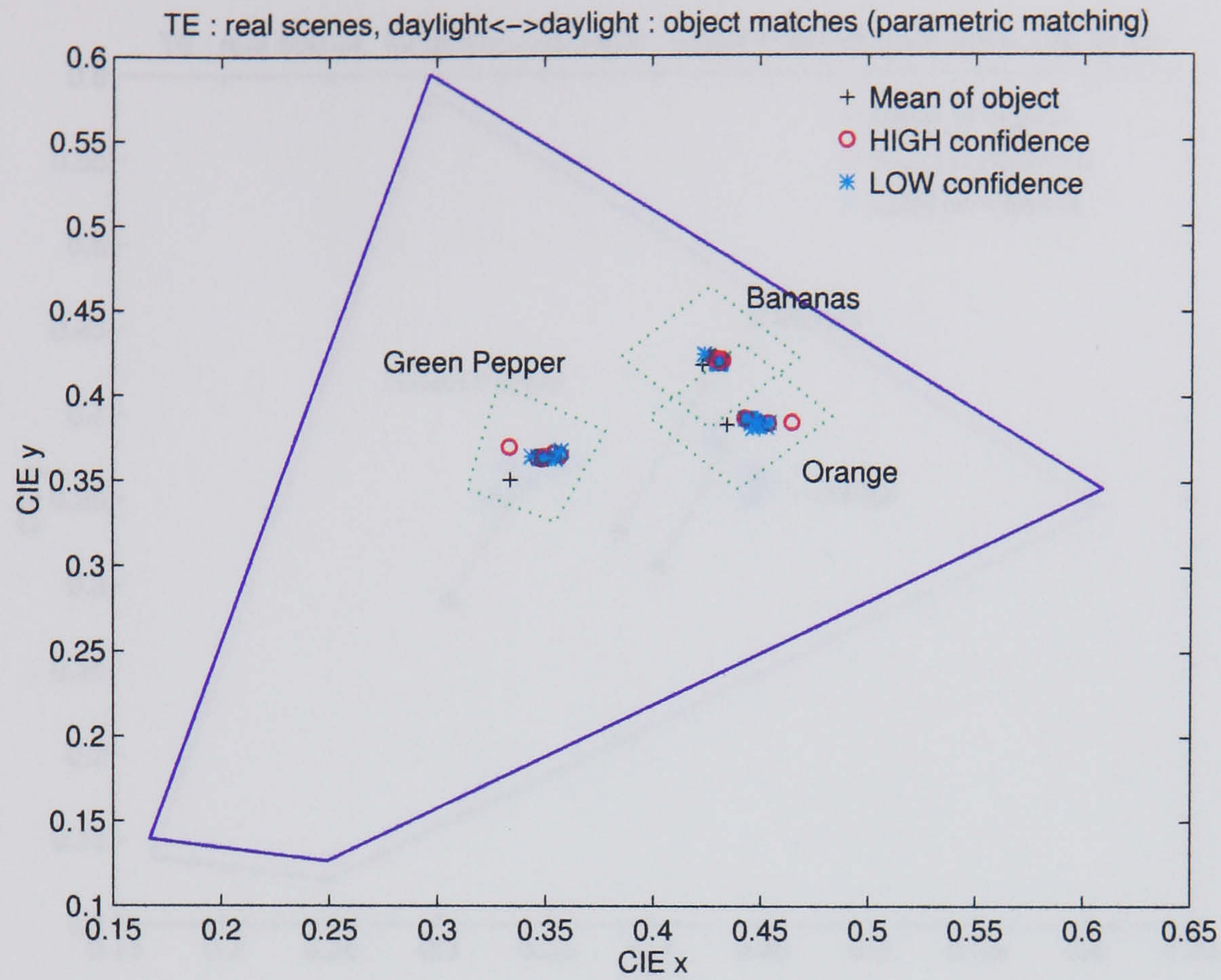

Figure 8.26: Real scene object matches for test and match images under daylight illumination for observer TE.

Figure 8.26 shows the object match results for TE in the no illuminant shift condition. Figures 8.27 and 8.28 show the same information for matching objects under tungsten and fluorescent light to daylight respectively. Figure 8.29 shows the results of the patch equivalent matching experiment.

Similarly, figure 8.30 shows the object match results for JH in the no illuminant shift condition. Figures 8.31 and 8.32 show the same information for matching objects under tungsten and fluorescent light to daylight respectively. Figure 8.33 shows the results of the patch equivalent matching experiment.

Analysis:

If we first consider the control case of equivalent patch matching, we can see that observers have no problem with the new technique and produce very repeatable matches. This is interesting in itself as the match variance using the parametric technique seems to produce matches with far lower variance than matches that require luminance, hue and saturation adjustment. This is highlighted in Table 8.1 in

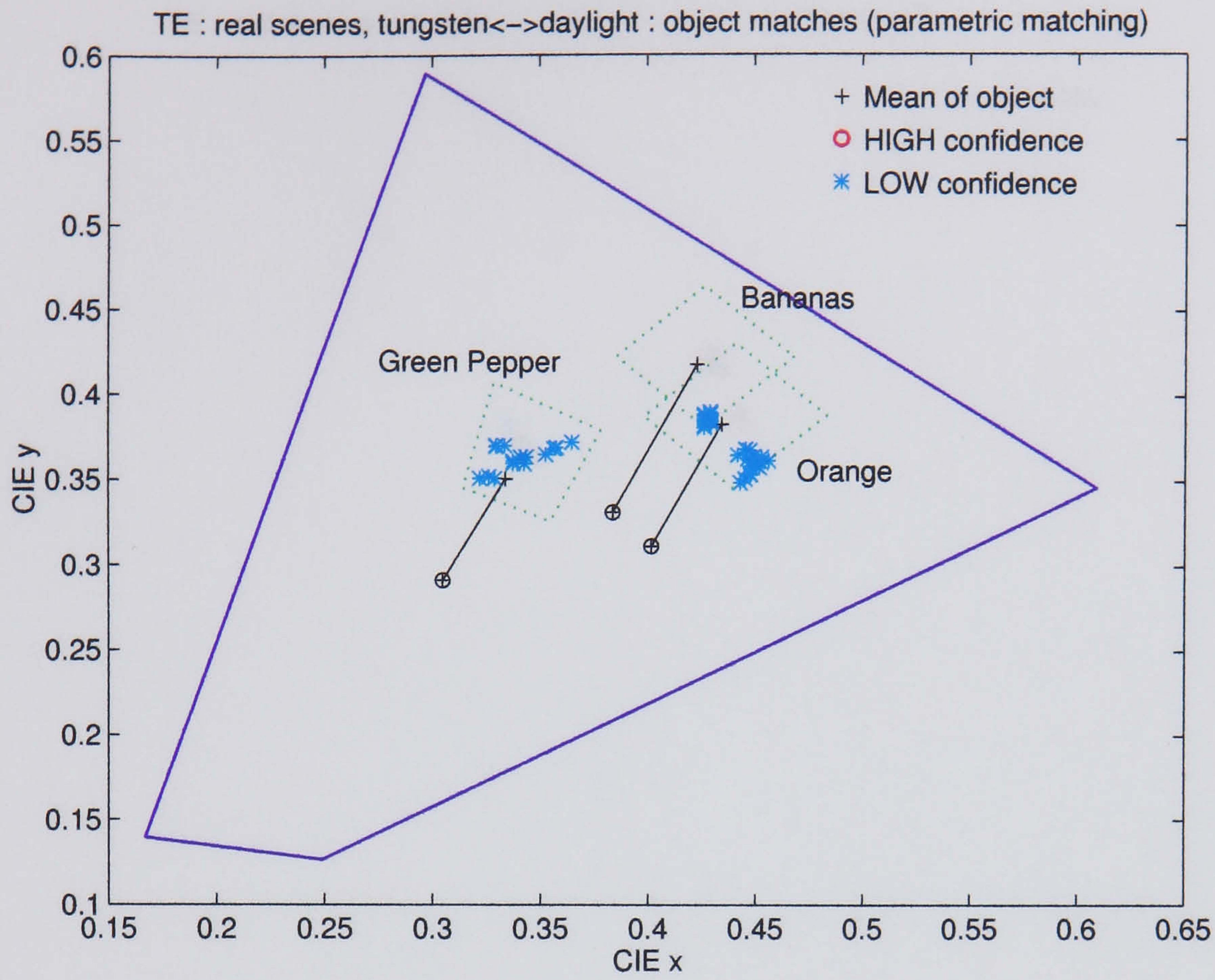

Figure 8.27: Real scene object matches for a test scene under tungsten illumination and match scene under daylight illumination for observer TE.

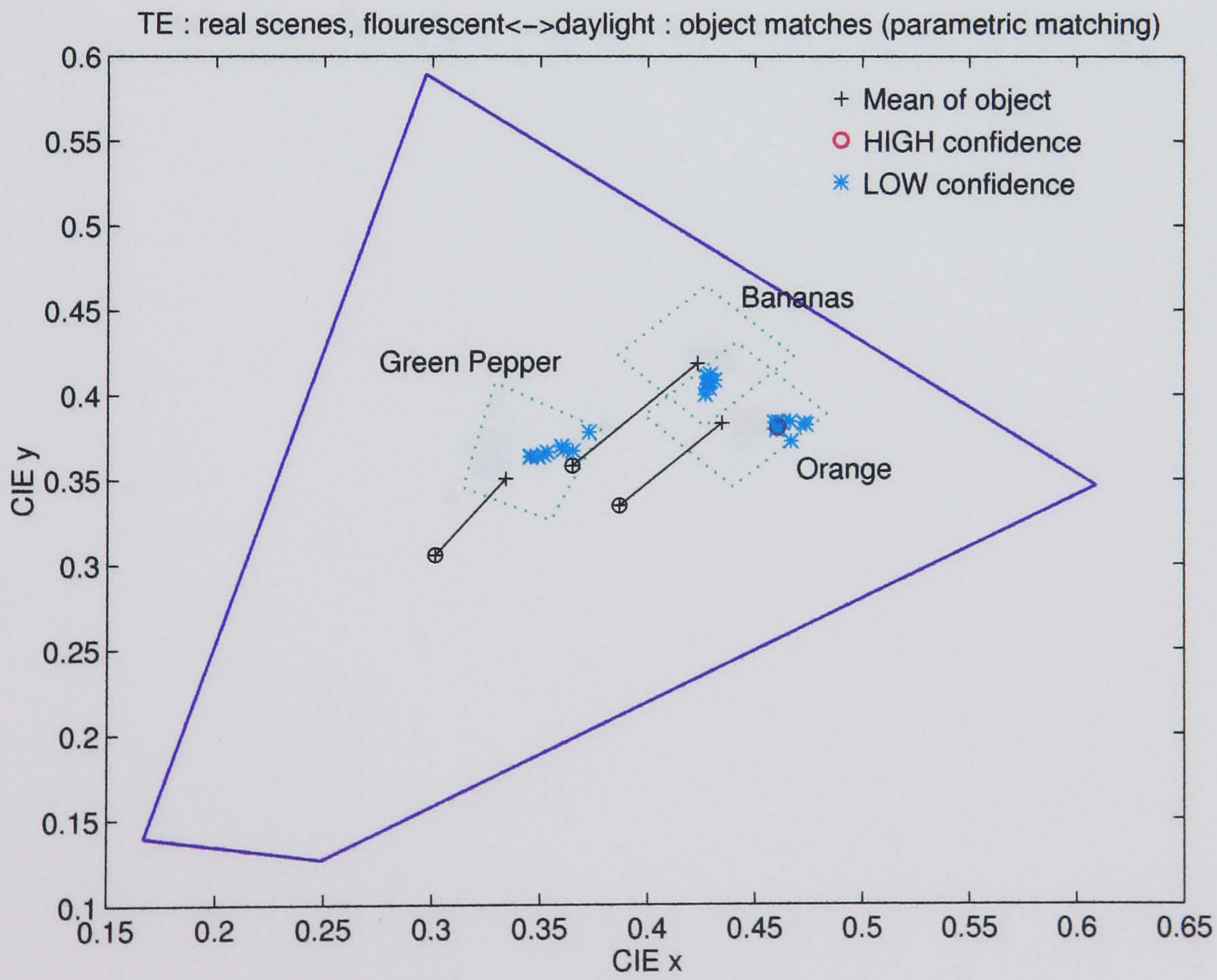

Figure 8.28: Real scene object matches for a test scene under fluorescent illumination and match scene under daylight illumination for observer TE.

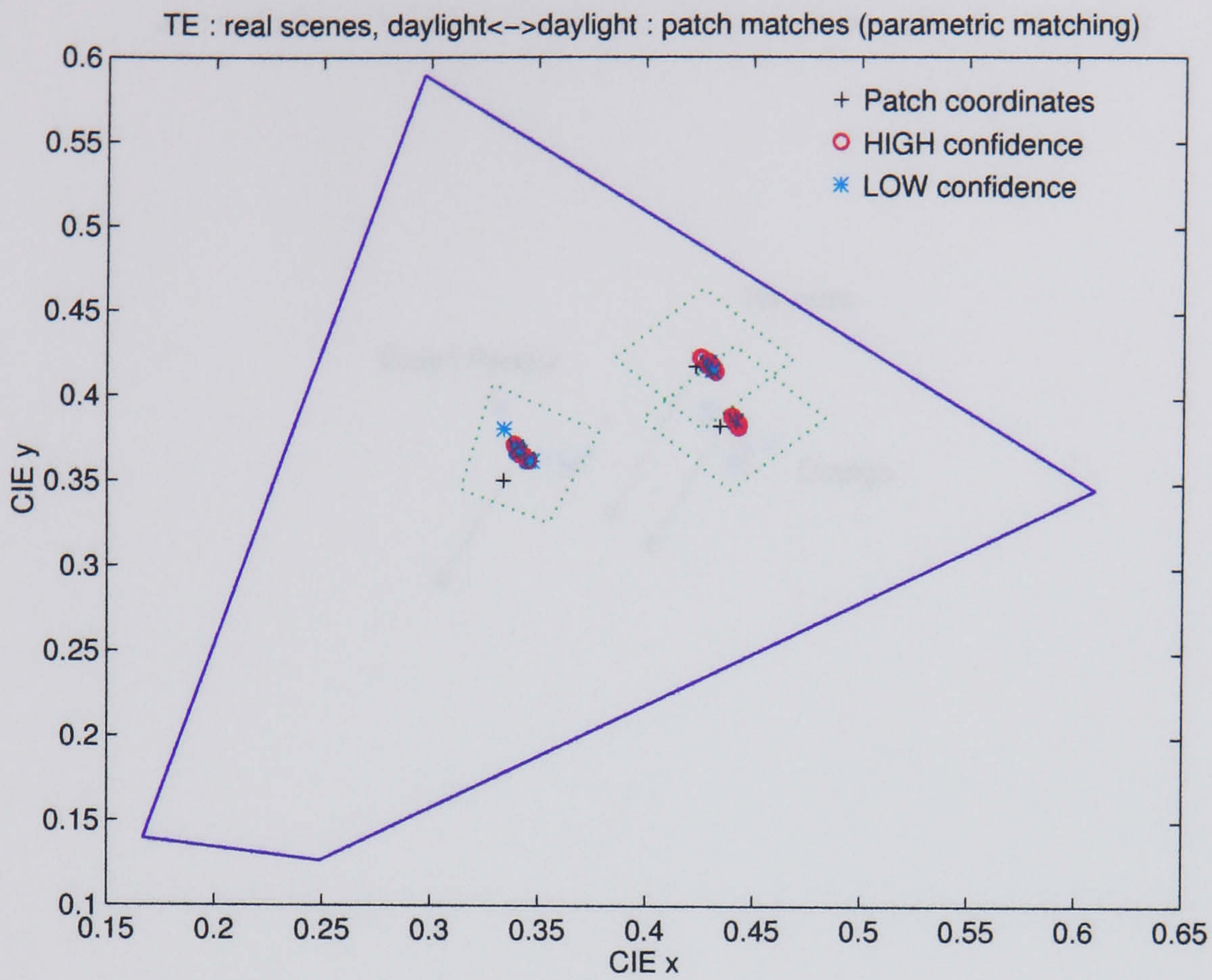

Figure 8.29: Flat patch match results for equivalent patches using the quad-linear interpolation technique for observer TE.

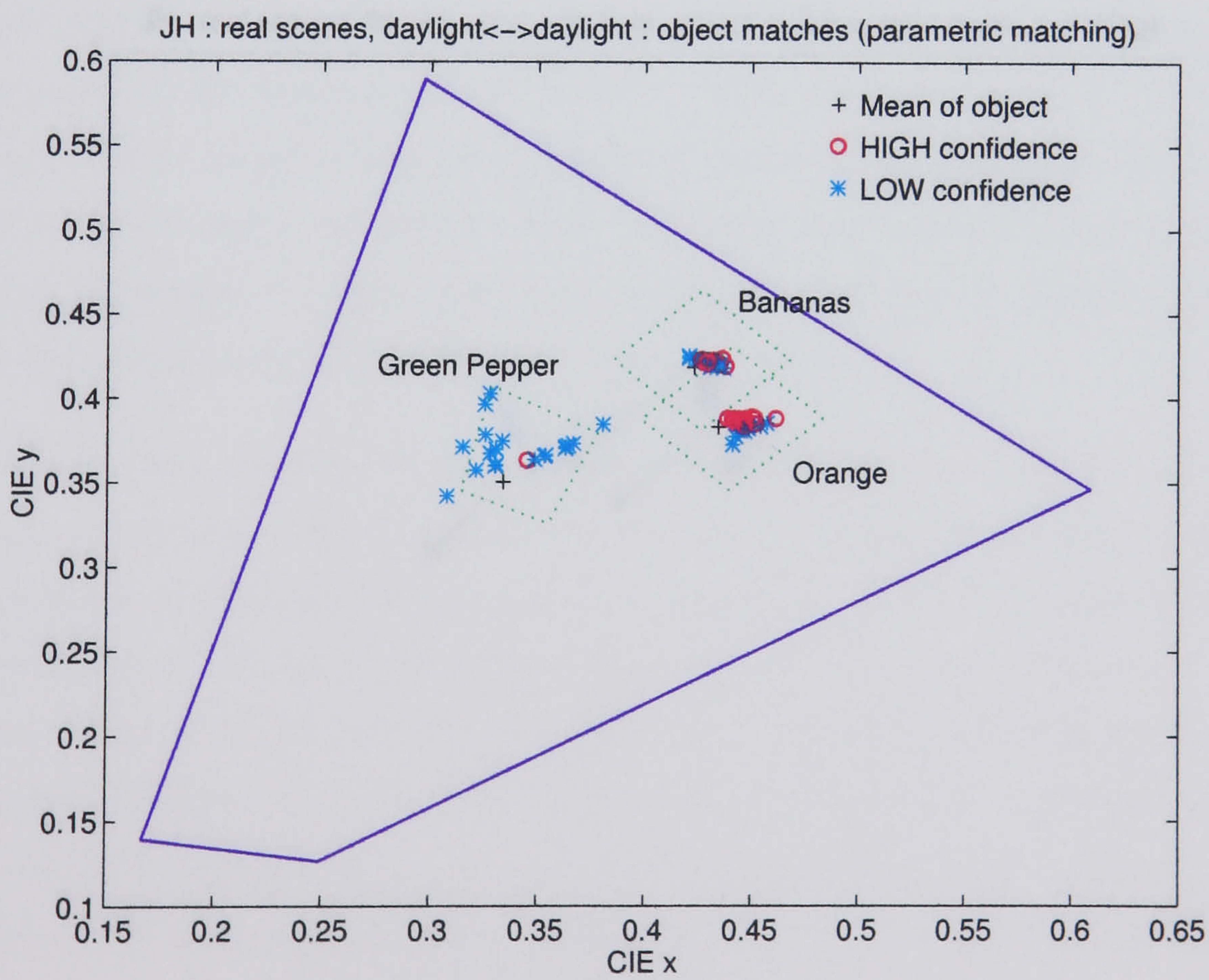

Figure 8.30: Real scene object matches for test and match images under daylight illumination for observer JH.

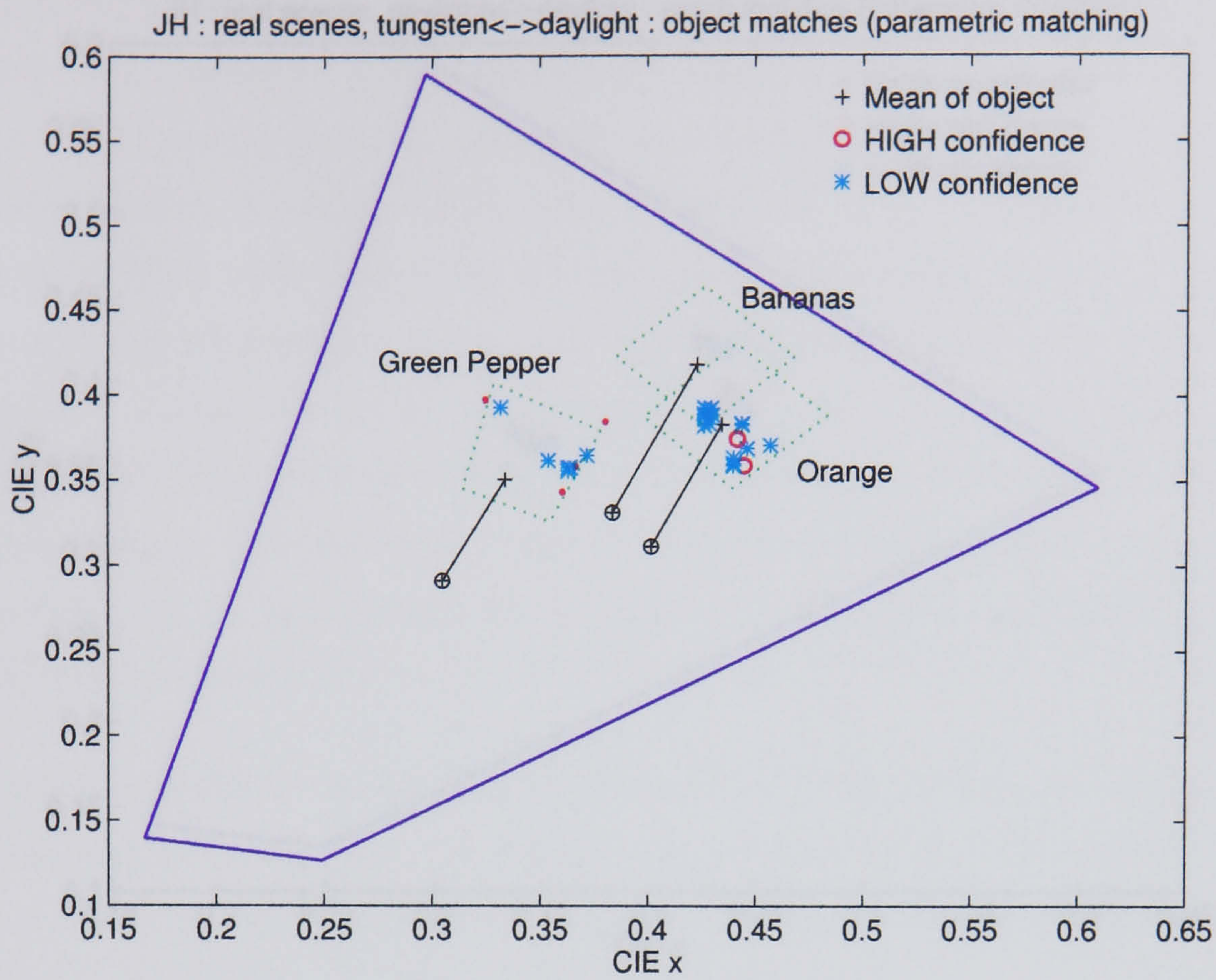

Figure 8.31: Real scene object matches for a test scene under tungsten illumination and match scene under daylight illumination for observer JH.

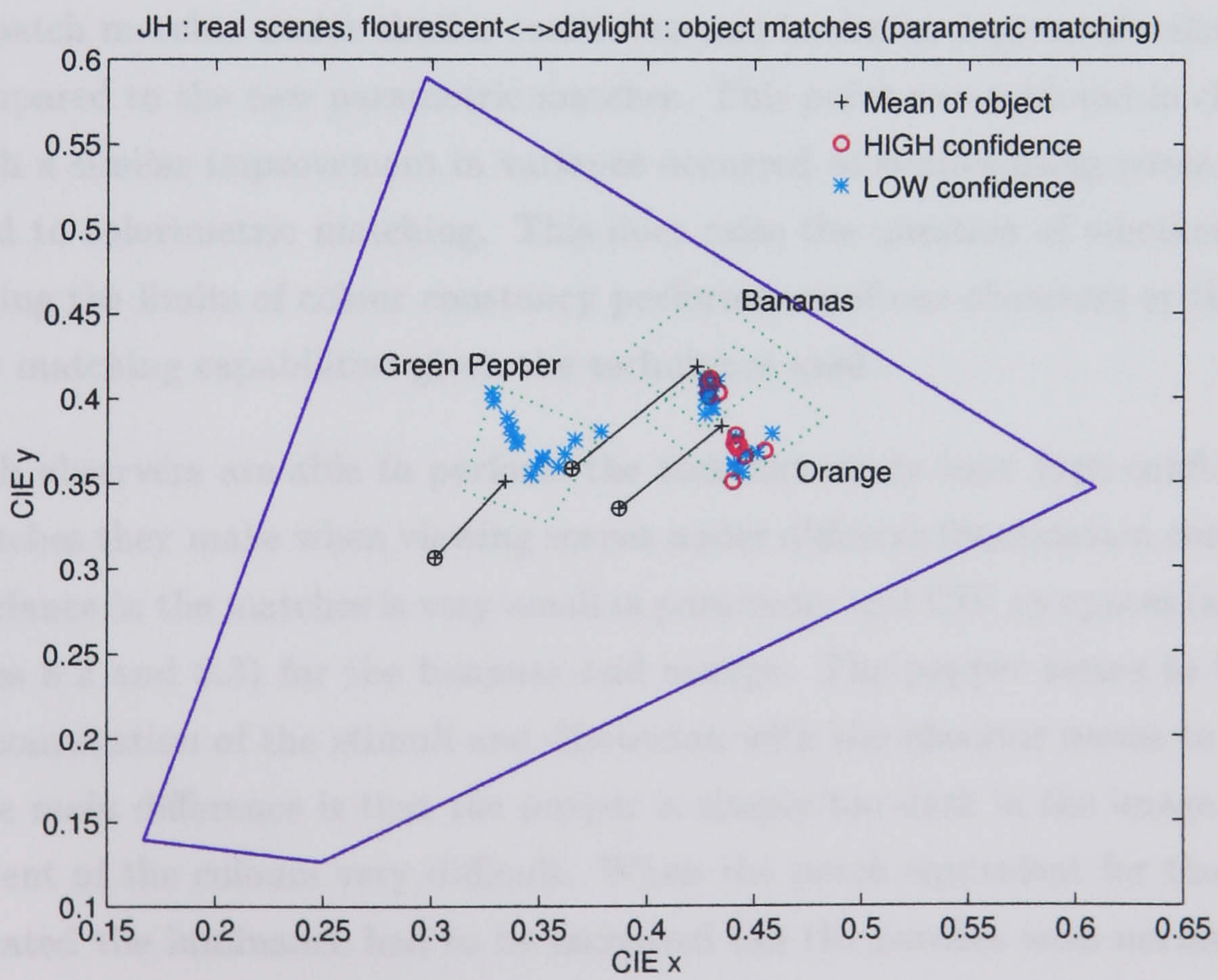

Figure 8.32: Real scene object matches for a test scene under fluorescent illumination and match scene under daylight illumination for observer JH.

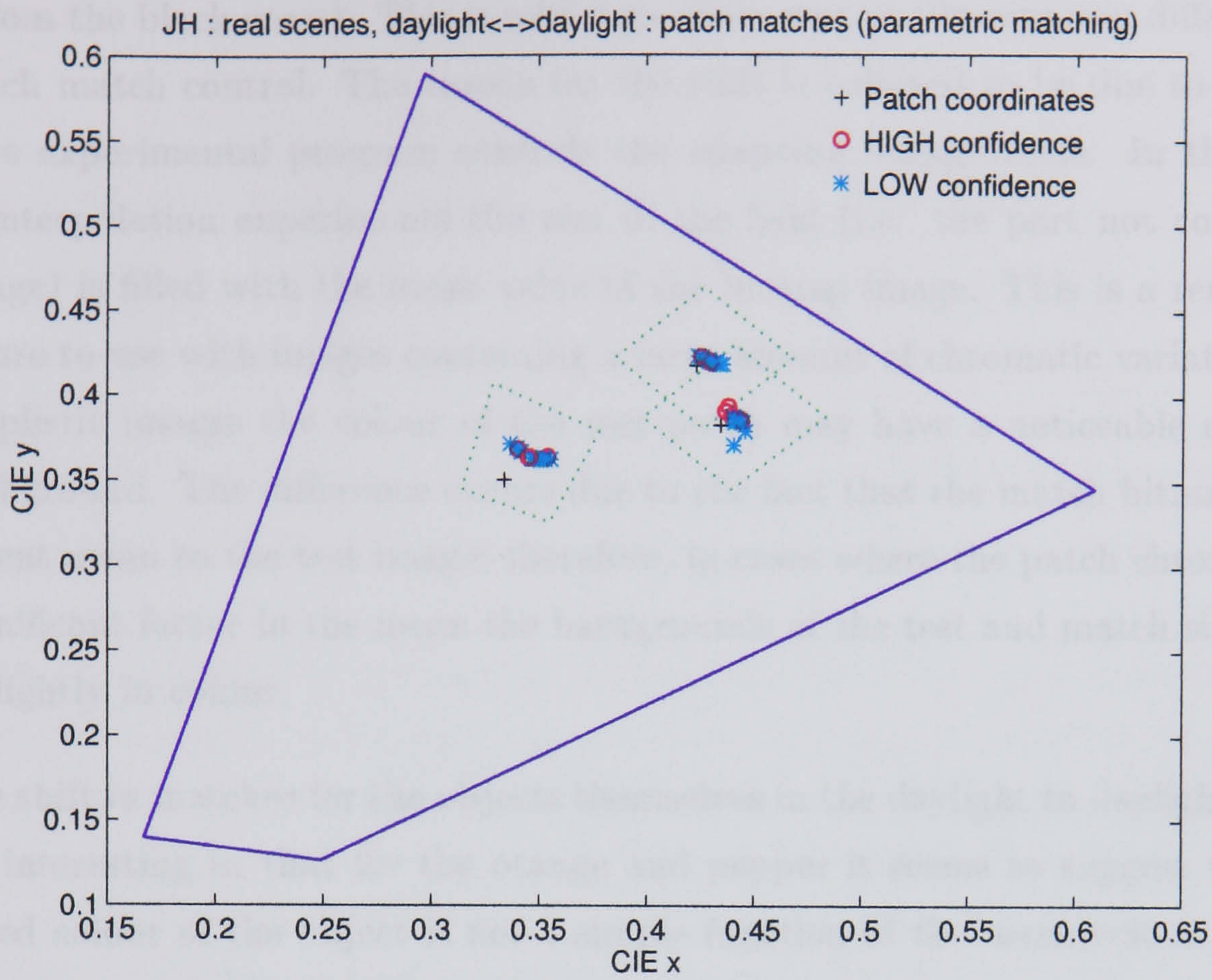

Figure 8.33: Flat patch match results for equivalent patches using the quad-linear interpolation technique for observer JH.

which patch matches under similar conditions and in similar regions of colour space are compared to the new parametric matches. This point was explored in chapter 3 in which a similar improvement in variance occurred in studies using parametric as opposed to colorimetric matching. This does raise the question of whether we are measuring the limits of colour constancy performance of our observers or the limits of their matching capabilities given the techniques used.

Both observers are able to perform the task but rarely have high confidence in the matches they make when viewing scenes under different illumination conditions. The variance in the matches is very small in parameter and CIE xy spaces (as shown in tables 8.2 and 8.3) for the bananas and orange. The pepper seems to fare less well. Examination of the stimuli and discussion with the observer seems to suggest that the main difference is that the pepper is simply too dark in the image making judgement of the colours very difficult. When the patch equivalent for the pepper was created the luminance had to be increased (all the patches were normalised to 6 cd/m^2 in that experiment) for any colour to be seen in it at all. A clear shift in the matches can be seen for the green patch (as shown by shift of the match points away from the black cross). This is odd since there was no illumination difference in the patch match control. The reason for the shift is believed to be due to the way that the experimental program controls the adapting backgrounds. In the quad-linear interpolation experiments the rest of the field (i.e. the part not containing the image) is filled with the mean value of the bitmap image. This is a reasonable procedure to use with images containing a large amount of chromatic variation, but for simplistic images the colour of the test patch may have a noticeable effect on the background. The difference occurs due to the fact that the match bitmaps have a different mean to the test image; therefore, in cases where the patch chromaticity is a significant factor in the mean the backgrounds of the test and match sides may differ slightly in colour.

The shift in matches for the objects themselves in the daylight to daylight match case is interesting in that for the orange and pepper it seems to suggest that the perceived colour of the object is not a simple function of the mean colour of all of the points on its surface (as depicted by the black cross).

The experiments in this chapter were set to 'disprove an hypothesis'. The hypothesis was that all of the extra information in the scene should improve colour constancy. Examination of the figures and tables shows clearly that not only was the degree of colour constancy less than we obtained for flat patches in similar illumination conditions but in some cases the observer could not get to the match due to the restrictions of the parameter space. Re-examination of figures 8.16, 8.17 and 8.18 shows that the extent of the parametric space represents relatively large shifts in colour. Not only is the degree of colour constancy exhibited low but in several cases it should be even worse.

This result is very interesting since it appears at odds with the perception of the scene that an observer obtains during casual viewing of the stimuli. Such casual viewing does not give an impression of large shifts in colour. The scenes look remarkably similar when viewed in the experimental context. This could be another manifestation of the problem outlined in the last chapter in which the observers' matches suggest a 'colorimetric' mode of viewing rather than the surface mode we are interested in. The low degree of constancy exhibited could paradoxically signal higher colour constancy performance in the scenes. If the observer has a fully

TE	Illuminants		Observer confidence high				Observer confidence low			
Patch	test	match	mean x	mean y	variance	%	mean x	mean y	variance	%
bananas_{mean}	day mean	day mean	0.429	0.419	0.0017	80_{16}	0.429	0.419	0.00169	20_4
pepper_{mean}	day mean	day mean	0.342	0.367	0.00309	55_{11}	0.342	0.367	0.0048	45_9
Y 5/4	E_{65}	E_{65}	—	—	—	—	0.407	0.419	0.00907	100_5
G 5/4	E_{65}	E_{65}	—	—	—	—	0.272	0.393	0.00952	100_5

Table 8.1: Comparison of parametric and hue saturation matches

accurate impression of the illumination colour then when the visual system 'recombines' the illumination and surface colour information during the matching process the colour shift will be greater than if the observer only partially recovered the illuminant colour.

Another possibility is that 'something' is lost in the depictions of the real life scenes that reduces our colour constancy capabilities. It was noted in section 3.1 that modern cameras use a global gain control technique that coupled with the loss in resolution due to the digitisation process can alter scene features such as shadows to the point that information is simply lost. If the reduction in constancy performance in depictions of coloured scenes is, in fact, due to the distortions in the image and not the factors discussed in the rest of this section then these results may provide interesting information. The distortions in the digitised scenes can be accurately quantified and as such we can accurately define the 'missing' information. If addition of this information does improve performance in these experiments then we know that the colour constancy system must, to some extent, rely upon these features.

In summary, matching CRT-displayed depictions of real objects in depictions of real scenes viewed in a fused haploscopic display shows no evidence for improved constancy. The results suggest that there might even be a reduction in colour constancy as compared with similar conditions in flat scenes. It is suggested that these results may not be indicative of the colour constancy performance of observers and may depend on the particular experimental techniques used in its measurement.

TE	Illumination†	Observer confidence high				Observer confidence low			
object	test	param 1	param 2	variance	%	param 1	param 2	variance	%
bananas	daylight	0.674	0.436	0.0632	33.33_7	0.566	0.428	0.061	66.67_{14}
orange	daylight	0.725	0.357	0.102	28.57_6	0.661	0.379	0.0856	71.43_{15}
pepper	daylight	0.7	0.438	0.15	38.1_8	0.755	0.396	0.114	61.9_{13}
bananas	tungsten	—	—	—	—	0.0475	0.0303	0.0474	100_{15}
orange	tungsten	—	—	—	—	0.229	0.0677	0.106	100_{15}
pepper	tungsten	—	—	—	—	0.86	0.6	0.261	100_{15}
bananas	fluorescent	—	—	—	—	0.11	0.284	0.0603	100_{10}
orange	fluorescent	0.719	0.176	—	10_1	0.778	0.134	0.102	90_9
pepper	fluorescent	—	—	—	—	0.83	0.344	0.193	100_{10}

† the match object was always under daylight.

Table 8.2: Real object matches for observer TE (in parameter space)

TE	Illumination	Observer confidence high				Observer confidence low			
object	test	mean x	mean y	variance	%	mean x	mean y	variance	%
bananas	daylight	0.43	0.421	0.00127	33.33_7	0.43	0.42	0.002	66.67_{14}
orange	daylight	0.45	0.385	0.00617	28.57_6	0.447	0.384	0.00272	71.43_{15}
pepper	daylight	0.349	0.364	0.0049	38.1_8	0.351	0.364	0.00366	61.9_{13}
bananas*	tungsten	—	—	—	—	0.428	0.385	0.00256	100_{15}
orange*	tungsten	—	—	—	—	0.449	0.359	0.00632	100_{15}
pepper	tungsten	—	—	—	—	0.34	0.363	0.0121	100_{15}
bananas	fluorescent	—	—	—	—	0.428	0.406	0.00272	100_{10}
orange*	fluorescent	0.46	0.38	—	10_1	0.465	0.381	0.00586	90_9
pepper	fluorescent	—	—	—	—	0.355	0.366	0.00848	100_{10}

* some matches restricted by parameter space.

Table 8.3: Real object matches for observer TE (in CIE xy space)

Chapter 9

Discussion

9.1 Introduction

This chapter examines all of the experimental results presented in this thesis. The implications of the results with respect to the model proposed and prediction of colour matches are then examined followed by consideration of details of the experimental procedures that may have affected the results.

9.2 Summary of experimental results

In general all of the results indicate that colour matches made across illuminant shifts, as measured here, are dictated by sensory processing. The matches made against uniform fields produce consistent colour shifts in CIE xy space away from colour constancy which are not in evidence for matches made against complex backgrounds. There are no major qualitative differences found for matches performed against contrast-rich backgrounds with different degrees of complexity.

All of the results show imperfect colour constancy, although some surface-illuminant combinations show better colour constancy than others. The degree of failure is consistent with many other similar studies in which a hue-saturation match is requested (Lucassen and Walraven 1992; Lucassen and Walraven 1993; Brainard and Wandell 1986; Bäuml 1994; Walraven, Benzschawel, Rogowitz and Lucassen 1991; Arend and Reeves 1986). The degree of constancy appears to be more related to the specific patch and background chromaticities than to the degree of complexity present in the stimulus. For example, Munsell patch Y 5/4 produces good constancy for most daylights, whereas the degree of constancy evident for patch G 5/4 varies significantly with the illumination. A specifically interesting point is shown in the match results for patch G 5/4. The shift in the match point is not indicative of a lack of illuminant compensation as the chromaticity the observers select is not situated on the line joining the perfect colour constancy point to the hue-saturation point, as we would expect if adaptation were not complete. This suggests that the dominant mechanism responsible for the matches is either unable to operate correctly in these forms of stimuli or that its major goal is not to achieve colour constancy.

The experiments introduced the novel feature of requiring observers to report their confidence in the matches they produce. This produced interesting results in several of the experiments. First, in the control and patch experiments, it can be seen that the low confidence points tend to be further from the centroid of the match distribution. This suggests that the errors reported in previous experiments may be under-estimating the capabilities of the visual system by assigning discriminable points to threshold or sub-threshold positions. The second interesting result is the reduction in observer confidence as the complexity of the scenes increases. This could indicate two things: first, the simulations might be too unsophisticated resulting in objects undergoing transformations that seem unnatural. Second, the effect may indicate that the observer has a stronger surface colour perception and becomes increasingly dissatisfied with the results produced when he performs in 'colorimetric mode'. This suggests that the results might not be indicative of colour constancy abilities, *per se*, but might be more indicative of the techniques used.

Chapter 5 examined mainly technical issues of match experiments. Some interesting points did arise from these essentially control experiments. The experiment in which matches are made against different illuminant simulations suggests a cortically based encoding of the retinal signals. This follows from the fact that matches appear to be related to the portion of colour space they reside, and not to the portions in which the stimuli of each eye reside. This suggests that discrimination capabilities are restricted by cortical and not retinal mechanisms[1].

[1]If resolution were lost at the retina then discrimination ellipses related to the stimuli present in both eyes would be expected, not an ellipse related to the portion of colour space the patch resides in.

The experiment in which the colour memory of an observer was examined raises the question: how good does colour constancy have to be? The system only needs to discount the illuminant to a degree to which an object surface colour matches some internal representation. The difficulty in matching and the simplistic stimuli may not make these results entirely indicative of human colour memory, but the point is still relevant. Does our colour constancy performance correlate with the tolerances for surface object colours?

The method of constant stimuli experiment starts to raise questions of the effects of procedural aspects on the results. The results are qualitatively the same as those produced by dial match techniques, but, if anything, suggest that performance may exceed that produced in dial matching.

Chapter 6 examined colour contrast in the backgrounds of simple match stimuli. The results follow cone ratio predictions (Shepherd and Whittle 1995; Wandell 1993), but the contrast results also suggest that the cone ratio model may not be sufficient to quantify behaviour accurately. That is, both matching against uniform and multi-patch background experiments produce results that follow cone ratio predictions, but a consistent chromaticity shift between the data sets means that the cone ratio hypothesis of colour matching is not sufficient to predict the results in CIE xy space. The results also suggest that colour constancy performance is improved (in terms of absolute chromaticity shifts from the perfect colour constant solution) by using the multi-patch backgrounds.

Chapter 7 examined the effects of adding 3-D shading to the test and match. The experiment employed two objects: a banana – an example of a recognisable object with a known surface colour, and a clay blob – an object that should promote no preconceptions in the observer. The results show that not only can observers reliably match shaded object to shaded object, but they can do so in a reasonably colour constant manner. The colour constancy obtained is imperfect and similar to that obtained for flat patches against multi-patch backgrounds. Confidence in the matches is reduced, possibly signalling increased surface colour awareness within the observer. An experiment was presented that tested the possibility of self constancy – does a shaded object, which segments easily from the background, retain its perceived surface colour when differentially illuminated to the multi-patch background upon which it is super-imposed? This situation can occur frequently in real life and seems to present no great difficulty for colour constancy mechanisms [2]. The results showed no evidence of self constancy and produced predictable matches from a sensory processing point of view.

The experimental stimuli used in chapter 8 take the level of detail almost as far as we can whilst using computer presentation[3]. Renditions of real scenes under real illuminants were digitised. The scenes were intentionally rich in complex features such as: shadows, colour contrast, mutual illumination, highlights etc. Several objects were matched between scenes under different illuminations. The results show imperfect colour constancy, and may even represent a reduction in performance as compared to the flat patches – although firm conclusions cannot be made as the exact illumination conditions were not comparable.

9.3 Implications for modelling and predicting colour constancy

A simple two stage model was explored throughout this thesis. The assumptions and specifics of the model were not formalised or explored in detail and as such firm conclusions as to its veridicality with respect to the visual system cannot be made. The general methodology is by no means novel and specifics of the model differ from others (Wandell 1993; Atick, Li and Redlich 1993) chiefly in implementational issues. The models are essentially equivalent[4]. Wandell suggests a series of transforms very similar to those suggested, but he does not restrict his transformations to the cone and opponent encoding suggested in this thesis. Atick, Li and Redlich perform a maximal decorrelation of input signals based on the inputs to the system with respect to space and time. Their procedure essentially concatenates the transformation matrices into a single transform that must be recalculated continuously.

Interesting predictions from the model were shown in section 6.4. The model qualitatively predicted that the matches against uniform backgrounds should be

[2] Admittedly a single object being the only part of the whole visual scene under a different illuminant is rare, but may occur.

[3] The limit is imposed by current technology and is not a conceptual or procedural problem. Cameras that capture local contrast and the ability to record and display more discrete levels would improve the stimuli.

[4] The unpredictability of noise at each stage in the system may cause slight differences to occur in the output produced.

shifted in CIE xy space in comparison with the same patches against a multi patch background. Not only were the predictions borne out but the size of the shift correlates well with the predicted size of match error. The model is by no means complete and uses several unwarranted and unexplored assumptions, but the general mode of operation does seem to predict the data remarkably well.

The model predicts these errors by examining differences in the final representation (the perceptual space). This same procedure highlights an inadequacy in the cone ratio hypothesis of match results. Both matches against uniform and multi-patch backgrounds conform well to the cone ratio hypothesis. Unfortunately, the hypothesis does not appear to be complete since it cannot accurately predict where the matches will be in the CIE xy plane. If cone ratios were the only factor in the match response then there is no reason for the consistent chromaticity shifts in CIE xy space between the uniform and multi-patch conditions.

An implication of the model has hitherto remained unexplored. The methodology essentially splits the information input to the system into two streams. The first stream is the surface-based information with which we have been mostly concerned. The other stream consists of the gains and offsets used by each amplifier at each of the stages. The gains and offsets are related to the illumination characteristics within the field of view of each amplifier pair, i.e. there are a number of two stage amplification stages operating upon signals from differing spatial locations within the visual field. This second stream would contain information related to the absolute intensity and chromatic properties of the illumination prevalent in each region. These signals could be hierarchically 'grown' into larger regions dependent on a similarity criterion. This would allow the construction of probable illumination for large regions of the visual field. Even simple similarity criteria would allow the image to be segmented into areas of consistent illumination. The results of the 'region' growing could be fed back to the amplifier stages to lock similar illumination regions to the same gains and offsets, thus reducing local offset effects due to areas of uniform colour.

This encoding scheme has another interesting property – it treats as distinct information about illumination and surface properties while retaining the capability to reconstruct fully the original excitation pattern[5]. This suggests observers would have access to information about not only surface colour and illumination, but also the absolute excitation at a given point on the object surface.

9.4 Consideration of experimental procedures

As the complexity of the stimuli increases, the division between what is perceived when casually viewing the scene and the matches produced for the same stimuli increases. In casual viewing, the perception of object surface colour does seem to improve as does the impression of illumination. Casual viewing suggests a higher degree of colour constancy in the complex scenes, whereas matching seems to suggest a reduction.

This contradiction may be due to the nature of the information we seem to be able to require our visual system to produce. For a single object we can be simultaneously aware of its surface colour, the illumination it is under and its apparent absolute colour. This can be seen quite easily by considering a single object – for example, a banana on an overcast day. If you ask yourself what colour is the banana, the answer will be yellow. Now ask, what is the illuminant and you will feel that bluish seems appropriate. Now you can objectively look at the banana and say what colour it actually is and you will see it as a blue tinged yellow, or desaturated yellow. We appear to be able to either access the transformed sensory signals (and possibly gain and offset factors) and reconstruct the original retinal signal[6] or somehow reconstruct the original chromaticity from the surface reflectance and illuminant perceptions. This is a similar suggestion to that of Arend and Reeves (Arend and Reeves 1986) which prompted the experiment in which observers were actually asked to perform a surface colour match. They showed a large increase in constancy – the exact reasons for which may not entirely instructional (this issue is examined in chapter 4).

This raises the possibility that the observer is responding in 'objective analysis mode' whenever requested to perform a dial match of hue, saturation and bright-

[5]The reconstruction capability will be effected by the noise in the system as a whole and assumes that the offset and gain is not only measurable but retained for use by later stages.

[6]Of course, some information will be completely lost to later processes.

ness. This essentially reduces all colour constancy experiments that use dial match techniques to categorisation of sensory processing of multi-coloured scenes.

The original motivation behind the method of constants technique presented in section 5.5.4 was an attempt to shift the mode of the observer. The technique is still in development and at the moment appears to produce results comparable to dial matches. This may be due to the instructions: the observers are asked to respond 'same' if the patches or objects are the same *colour*. A question that may be more suitable would be: are the two presented *objects* the same? The answers to whether the method of constant stimuli technique will improve matters are inconclusive as observers experienced in the match task appear to treat the task in the same way as a colour match – they have been trained to compare the test and match objectively. We intend to repeat the experiment using subjects that are naive not only to the purposes of the experiment, but also to colour matching in a laboratory environment. It is hoped that presenting objects to such observers coupled with instructions requiring an object-based response may be more fruitful in examining human colour constancy performance.

Chapter 10

Conclusions

This thesis set out to examine human colour constancy performance in simple and complex scenes. The simple interpretation of the results is that colour constancy performance in simple and complex scenes is qualitatively very similar, as measured in asymmetric matching experiments. The results suggest that low level sensory processes are dominant factors in the constancy obtained. The addition of object 3-D shading, object recognisability, mutual illumination, highlights and other 'real life' scene features appears to have very little effect on the constancy produced – in the matching experiments performed here.

The qualifying remark: 'in the matching experiments here' is used to highlight the possibility that the sensory influences may be dominant due to procedural aspects in the experiments. The results for complex scenes seem to be somewhat at odds with casual viewing of the test scenes, from which good impressions of surface colour are obtained. This suggests that the complex scene results represent the degree to which absolute chromatic information is available to the observer at higher levels, i.e. the object representations within the scenes do not provide features so strong as to mask the sensory information. It appears that sensory information of absolute stimulus characteristics is available at the highest levels and can be consciously 'selected' as an attribute of interest. The cognitive component appears to tie all of the seemingly contradictory results and observations together. The paradox can be expressed in the strong and accurate impression of object surface and scene illuminant colour we obtain from casual viewing of the real scenes. We are able to differentiate, and access, the distinct properties of absolute chromaticity.

surface colour and illuminant of an object. It appears that matching tasks require the observer to access the absolute form of information and as such perform in a highly predictable manner based on low level sensory encoding.

The results may also suggest an alternative: does the visual system strive for perfect colour constancy or does it strive for efficiency of useful information encoding? The two alternatives can produce very similar results but are subtly different concepts. Firstly, perfect colour constancy is unnecessary if we can successfully operate in our environment with a close approximation to it, i.e. if imperfect constancy provides all of the information we require about the ripeness of a piece of fruit then the exact chromaticity is irrelevant. There are survival situations in which perfect colour constancy could be detrimental – knowledge of the approach of night or of a thunderstorm can be critical.

Second, several computational studies (Funt, Drew and Ho 1991; Healey 1991; Lee 1986) suggest that complex scene features allow 'colour constancy' to be solved to a high degree of accuracy. This raises the question: why do we never get perfect colour constancy even in scenes when more than sufficient information exists for us to assess accurately, and hence discount the illuminant? This does assume that the visual system uses these high level features, which is still an open question. No matter how complex the scene, or how many 'advantageous' features are present we can still assess the absolute chromaticity of an object and detect chromaticity variation with changes in illumination. This tends to suggest that the visual system strives for efficient representation of useful information within a scene, which happens to coincide under most conditions with a surface based representation of the world. The two scenarios are so close as to be almost indistinguishable for many scenes. An efficient encoding of a scene happens to coincide with a colour constant view of the world. Usui, Nakauchi and Miyamoto (Usui, Nakauchi and Miyamoto 1992) have shown, using computer simulations of scenes consisting of Munsell patches under various illuminants, that the maximal decorrelation of the signals produced from the whole data set produces perfect colour constancy. In light of these results, the fact that we do not have perfect colour constancy could suggest that either we do not use decorrelation as a basis for colour constancy, or our decorrelation transform is not completely plastic i.e. cannot completely adjust to match the given stimulus, or the visual system is subject to more internal noise than the study assumed or is strongly dependent on features of the scene in view (i.e. we may not have access to the 'whole data set' that allows the maximal decorrelation to converge to perfect colour constancy).

Thirdly, colour constancy has been demonstrated in honeybees (Werner, Menzel and Wehrhahn 1988) and fish (Ingle 1985), and, in the studies from which comparisons can be made, appears to be qualitatively similar to that in humans. This suggests that a large amount of detailed scene analysis is not required for the solution of colour constancy (to the levels exhibited by humans) and that sensory processing of signals may be the major system underlying human colour constancy.

In vision research, it often happens that highly detailed models of a section of the visual system are presented. The models can be gloriously complex and predict some aspect of visual function very accurately. But, ironically, it often happens that another functional unit should be placed at the end of the block diagram labelled 'then perception occurs'. Unfortunately, I am now at the point where I am forced to do the same thing. A basic model was presented in this thesis that, in qualitative terms, seems a reasonable interpretation of the functional stages that occur in the early visual system. Qualitatively the model is not a new suggestion (Wandell 1993; Whittle 1994; Webster and Mollon 1991; Lucassen and Walraven 1993). Although most authors do not express it in the form presented here, all predict that one linear stage is not sufficient to explain the results obtained, just as does the model presented in this thesis. The model essentially transforms the input image to maximise the signal to noise ratio – the signal being information about the surface properties of objects. This transformation does go towards the goal of colour constancy, but is only a transform and as such does not tell us anything of how scene perception may be mediated. How the transformed signal is interpreted to produce our view of the world is still a very open question.

In conclusion, this thesis has examined colour constancy in simple and complex scenes and found the results consistent with mostly low level signal transformations of the retinal input. Whether high level factors have a larger role in our everyday viewing of scenes is still a strong possibility and as such, it is suggested that colour constancy in humans is a combination of low signal transforms adjusted by high level processing in order to return the information we require for the task in hand.

Bibliography

Adelson, E. H. (1993). Perceptual organization and the judgement of brightness. *Science 262*, 2042–2044.

Andreou, A. (1996). Personal communication and experimentation. Technical report, Telluride Workshop in Neuromorphic Engineeriing.

Arend, L. E. and R. Goldstein (1987). Lightness models, gradient illusions, and curl. *Perception and Psychophysics 42*(1), 65–80.

Arend, L. E. and A. Reeves (1986). Simultaneous color constancy. *Journal of the Optical Society of America 3*(10), 1743–1751.

Atick, J. J., Z. Li and A. N. Redlich (1993). What does post–adaptational color appearance reveal about cortical color representation? *Vision Research 33*(1), 123–129.

Barbur, J., A. Harlow and G. Plant (1994). Insights into the different exploits of colour in the visual cortex. *Proceeding of the Royal Society of London series B-Biological Sciences 258*(1353), 327–334.

Barlow, H. B. (1953). Summation and inhibition in the frog's retina. *Journal of Physiology 119*, 69–88.

Bäuml, K. (1994, February). Color appearance: Effects of illuminant changes under different surface collections. *Journal of the Optical Society of America 11*(2), 531–542.

Blackwell, K. T. and G. Buchsbaum (1988). The effect of spatial and chromatic parameters on chromatic induction. *Color Research and Applications 13*, 166–173.

Boynton, R., L. Fargo, C. X. Olson and H. Smallman (1989, Oct). Category effects in color memory. *COLOR research and application 14*(5). 308–313.

Boynton, R. M. (1956). Rapid chromatic adaptation and the sensitivity functions of human color vision. *Journal of the Optical Society of America 46.*

Boynton, R. M., M. M. Hayhoe and MacLeod (1977). The gap effect: chromatic and achromatic visual discrimination as affected by field separation. *optica acta 24*(2), 159–177.

Brainard, D. H. (1989, February). Calibration of a computer controlled color monitor. *Color Research and Applications 14*(1), 23–34.

Brainard, D. H. and B. A. Wandell (1986). Analysis of the retinex theory of color vision. *Journal of the Optical Society of America 3*, 1651–1661.

Brainard, D. H. and B. A. Wandell (1991). A bilinear model of the illuminant's effect on color appearance. In M. S. Landy and A. J. Movshon (Eds.), *Computational models of visual processing*, pp. 171–186. MIT Press.

Brainard, D. H. and B. A. Wandell (1992). Asymmetric color matching: how color appearance depends on the illuminant. *Journal of the Optical Society of America 9*(9), 1433–1448.

Bramwell, D. (1993, July). Monitor calibration. Technical Report 2, Neural Computation and Sensory Systems Research Group.

Bramwell, D. and A. Hurlbert (1996). Measurements of colour constancy using a forced-choice matching technique. *Perception 25*, 229–241.

Brown, W. R. J. (1952, April). Statistics of color-matching data. *Journal of the Optical Society of America 42*(4), 252–256.

Brown, W. R. J. and D. L. MacAdam (1949). Visual sensitivities to combined chromaticity and luminance differences. *Journal of the Optical Society of America 39*(10), 808–834.

Buchsbaum, G. and A. Gottschalk (1983). Trichromacy, opponent colours coding and optimum colour information transmission in the retina. *Proceedings of the Royal Society of London B 220*, 89–113.

Burnham, R., R. Evans and S. M. Newhall (1957, January). Prediction of color appearance with different adaptation illuminations. *Journal of the Optical Society of America 47*(1), 35–42.

Burnham, R. W., J. R. Clark and S. M. Newhall (1957, october). Space error in color matching. *Journal of the Optical Society of America 47*(10). 959–966.

Chichilnisky, E.-J. and B. Wandell (1996). Seing gray through the ON and OFF pathways. *Visual Neuroscience 13*, 591–596.

CIE (1932). *CIE proceedings 1931.* Cambridge University Press.

Cornelissen, F. W. and E. Brenner (1989). Simultaneous colour contrast depends on the experimental conditions. *Perception 18*, 529.

Cornelissen, F. W. and E. Brenner (1990). The role of eye movements in color vision. *Perception 19*, 333.

Cornelissen, F. W. and E. Brenner (1992). Influence of separation on chromatic induction scales with eccentricity. *Perception 21*, 75.

Craven, B. J. and D. H. Foster (1992). An operational approach to color constancy. *Vision Research 32*, 1359–1367.

De Valois, R. L., I. Abramov and G. H. Jacobs (1966). Analysis of response patterns of LGN cells. *Journal of the Optical Society of America 56*, 966–977.

De Valois, R. L. and K. K. De Valois (1988). *Spatial Vision.* Oxford University Press.

Delbrook, T. (1996). Personal communication. Technical report, Telluride Workshop in Neuromorphic Engineeriing.

DePriest, D., J. Krauskopf and P. Lennie (1991). Slow mechanisms of chromatic adaptation in macaque lgn. *Investigative Opthalmology & Visual Science 32*, 1252.

Derrington, A. M., J. Krauskopf and P. Lennie (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology 357*, 241–265.

Ditchburn, R. W. and B. L. Ginsborg (1952). Vision with stabilized images. *Nature 170*, 36–37.

Eastman, A. A. and G. A. Brecher (1972, April). The subjective measurement of color shifts with and without chromatic adaptation. *Journal of IES*, 239–246.

Fairchild, M. and P. Lennie (1992). Chromatic adaptation to natural and incandescent illuminants. *Vision Research 32*, 2077–2085.

Fairchild, M. D. and D. B. Munsell (1994). Testing colour appearance models in cross-media image reproduction. *Journal of Photographic Sciences 42*(3). 87–88.

Forsyth, D. A. (1990). Colour constancy. In A. Blake and T. Troscianko (Eds.). *AI and the Eye*, pp. 201–227. John Wiley and Sons Ltd.

Foster, D. H., B. J. Craven and E. R. H. Sale (1992). Immediate color constancy. *Ophthal. Physiol. Opt. 12*, 157–160.

Foster, D. H. and S. M. C. Nascimento (1994). Relational colour constancy from invariant cone-exitations ratios. *Proceedings of the Royal Society of London B 257*, 115–121.

Funt, B., M. Drew and J. Ho (1991). Color constancy from mutual reflection. *International Journal of Computer Vision 6*, 5–24.

Gelb, A. (1929). *Handbuck norm. und pathol. Psychologie*, Chapter Die "Farbenkonstanz" der Sehdinge, pp. 594–678.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science 195*, 185–187.

Hartline, H. K. (1940). The receptive fields of optic nerve fibers. *American Journal of Physiology 130*, 690–699.

Healey, G. (1991). Estimating spectral reflectance using highlights. *Image and Vision Computing 9*(5), 333–337.

Hubel, D. H. and T. N. Wiesel (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London) 195*, 215–243.

Hurlbert, A., H.-C. Lee and H. Bülthoff (1989, May). Cues to the color of the illuminant. *Investigative Ophthalmology and Visual Science Supplement 30*(3).

Ingle, D. J. (1985, february). The goldfish as a retinex animal. *Science*, 651–653.

Jaaskelainen, T., J. Parkkinen and S. Toyooka (1990). Vector-subspace model for color representation. *Journal of the Optical Society of America A 7*, 725–730.

Jameson, D. and L. M. Hurvich (1961). Opponent chromatic induction: experimental evaluation and theoretical account. *Journal of the Optical Society of America 51*(46-53).

Judd, D. B., D. L. MacAdam and G. Wyszecki (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America 54*, 1031–1040.

Koffka, K. (1935). *Principles of Gestalt Psychology*. Harcourt, Brace and Co.

Krauskopf, J. and Q. Zaidi (1986). Spatial factors in color induction. *Investigative Ophthalmology and Visual Science 27*(suppl), 291.

Krinov, E. (1947). Spectral'naye otrazhatel'naya sposobnost'prirodnykh obrazovanii. *Izd. Akad. Nauk USSR (Proc. Acad. Sci. USSR).* translation by G. Belkov, "Spectral reflectance properties of natural formations," Technical Translation TT-439 (National Research Council of Canada, Ottawa, Canada, 1953).

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology 16*, 37–68.

Land, E. H. and J. J. McCann (1971). Lightness and retinex theory. *Journal of the Optical Society of America 61*, 1–11.

Lee, H.-C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America 3*, 1694–1699.

Logvinenko, A. (1996). The Adelson illusion in colour. presentation at Colour in Context, meeting, Durham University.

Lucassen, M. and J. Walraven (1992). The luminance variable in color constancy - evidence for luminance normalised cone signals. *Investigative Opthalmology & Visual Science 33*(4), 705.

Lucassen, M. and J. Walraven (1993). Quantifying color constancy - evidence for nonlinear processing of cone-specific contrast. *Vision Research 33*(5-6), 739–757.

MacAdam, D. L. (1942, May). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America 32*(5), 247–273.

MacLeod, D. I. and R. M. Boynton (1979, August). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America 69*(8), 1183–1186.

Marimont, D. H. and B. Wandell (1994, december). Matching color images: the effects of axial chromatic aberration. *Journal of the Optical Society of America 11*(12), 3113 – 3122.

McCann (1993). Color sensations in complex images. In *IS&T and SID's Color Imaging Conference: Transforms & Transportability of Color.*

Newhall, S., R. Burnham and J. R. Clark (1957, January). Comparison of successive with simultaneous color matching. *Journal of the Optical Society of America 47*(1), 43–56.

Savoy, R. and O'Shea (1993). Color constancy with reflected and emitted light. *Perception 22*(Supplement), 61.

Schirillo, J. A. and S. K. Shevell (1993, December). Lightness and brightness judgements of coplanar retinally noncontiguous surfaces. *Journal of the Optical Society of America 10*(12), 24422452.

Shepherd, A. and P. Whittle (1995). Cone ratio predictions of patch matches against uniform adapting fields. draft version supplied to A. C. Hurlbert.

Shevell, S. K. (1978). The dual role chromatic backgrounds in colour perception. *Vision Research 18*, 1649–1661.

Shevell, S. K. and R. A. Humanski (1984). Color perception under contralateral and binocularly fused chromatic adaptation. *Vision Research 24*(9), 1011–1019.

Shevell, S. K. and M. F. Wesner (1990). chromatic adapting effect on an achromatic light. *OSA Technical digest series 15*, 149.

Smith, V. and J. Pokorny (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research 15*, 161–171.

Troost, J. and C. deWeert (1991). Naming versus matching in color constancy. *Perception & Psychophysics 50*(6), 591–602.

Usui, S., S. Nakauchi and Y. Miyamoto (1992). A decorrelating neural network for colour constancy. *IEEE*, 1030–1031.

Valberg, A. (1974). Color induction: dependence on luminance, purity and dominant or complementary wavelength of inducing stimuli. *Journal of the Optical Society of America 64*, 1531–1540.

Valberg, A. and B. Lange-Malecki (1990). "colour constancy" in mondrian patterns: a partial cancellation of physical chromaticity shifts by simultaneous contrast. *Vision Research 30*(3), 371–380.

von Kries, J. (1905). Influence of adaptation on the effects produced by luminous stimuli. *Handb. Physiol. Menschen 3*, 109–282.

Walraven, J. (1973). Spatial characteristics of chromatic induction: the segregation of lateral effects from stray light artifacts. *Vision Research 13*, 1739–1753.

Walraven, J. (1976). Discounting the background – the missing link in the explanation of chromatic induction. *Vision Research 16*, 289–295.

Walraven, J., T. L. Benzschawel, B. E. Rogowitz and M. P. Lucassen (1991). *Testing the contrast explanation of color constancy*, pp. 369–377. New York: Plenum Press.

Wandell, B. (1993, november). Color appearance: The effects of illumination and spatial pattern. *Proceedings of the National Academy of Science 90*, 9778 – 9784.

Webster, M. A. and J. Mollon (1991). Changes in colour appearance following post-receptoral adaptation. *Nature 349*, 235–238.

Werner, A., R. Menzel and C. Wehrhahn (1988, January). Color constancy in the honeybee. *Journal of Neuroscience 8*(1), 156–159.

Whittle, P. (1986). Increments and decrements: luminance discrimination. *Vision Research 26*, 1677–1691.

Whittle, P. (1992). Brightness, discriminability and the "crispening effect". *Vision Research 32*(8), 1493–1507.

Whittle, P. (1994). *Lightness, brightness, and transparency*, Chapter The psychophysics of contrast brightness, pp. 35–110. Erlbaum.

Whittle, P. and P. D. C. Challands (1969). The effect of background luminance on the brightness of flashes. *Vision Research 9*, 1095–1110.

Wyszecki, G. and G. Fielder (1971). New color-matching ellipses. *Journal of the Optical Society of America 61*, 1135–1152.

Wyszecki, G. and W. Stiles (1982). *Color Science.* John Wiley and Sons, New York.

Yarbus, A. L. (1967). *Eye movements and vision.* Plenum Press, New York.