ELSEVIER

Displays 23 (2002) 213–222

DISPLAYS

www.elsevier.com/locate/displa

Spectral-reflectance function recovery for improved colour-constancy experiments

Juan L. Nieves*, F. Pérez-Ocón, J. Hernández-Andrés, J. Romero

Departamento de Óptica, Facultad de Ciencias, Universidad de Granada, E-18071 Granada, Spain

Received 21 June 2001; revised 4 March 2002; accepted 4 March 2002

Abstract

A set of symmetric memory-matching data is presented to analyse some implications of long-term memory factors within classical colour-constancy paradigms and separation algorithms. Using simulated Mondrian-type colour surrounds on a CRT monitor, subjects make a series of colour matches between a test and a matching surface; the surfaces are rendered under the same standard illuminant (equal-energy illuminant). The 16 test surfaces used were categorised into four apparent-hue collections. The analysis of the colour differences show that subjects maintained good mental representations of the surfaces, although a shift in luminance was found. With these results, we investigated how errors in remembering surface colours might be translated into errors in reconstructing surface reflectances. Thus, a description of the remembered surfaces is provided, and the spectral differences are analysed via a goodness-of-fit coefficient (GFC). As it is derived from colour-differential thresholds and GFC values, the analysis of the recalled spectral-reflectance functions shows little loss of information in the observer's task, despite imperfect mathematical recovery of the surfaces. The similarities between test and matching surfaces suggest that colour-constancy algorithms could benefit of memory matches when an illuminant change takes place, and use spectral-tolerance bands defined over the surfaces comprising a scene to improve their implementation.

Keywords: Colour vision; Colour-constancy; Reflectances; Separation algorithms; Colour-differential thresholds

1. Introduction

Colour memory is usually defined in terms of the observer's ability to recognise a colour stimulus or remember it for a period of time [1]. Extensive research has been conducted on colour memory and its influence on experiences requiring detection and reproduction of *past-colour* stimuli [2–4]. However, the influence of memory can be extended to the study of other colour-vision phenomena (e.g. colour-constancy), in which memory is usually disregarded. Colour-constancy is defined as a stable perceived colour under different illuminant conditions [5]. For good colour-constancy, the visual system needs both a stable representation of the reflectance function of objects and a description of the spectral changes in the illuminant. When this classical problem is reproduced in the laboratory, the intrusion of memory is minimised to avoid confounding the observer's task, although it is clearly important in everyday life.

From this viewpoint several studies have tried to quantify colour-constancy in humans [6–9]. Colour-constancy has been found to be imperfect and affected by many factors, including compensation for illuminant changes and temporal factors. The influence on the observer's task is clearly evident when the subjects make matches under different instructions: paper- or surface-match conditions (in which perceived colour is associated with a surface) show good colour-constancy, whereas hue-match conditions (in which perceived colour is associated with a light rather than a surface) show poor colour-constancy [10]. In addition, temporal factors can intervene. Although the visual system uses rapid processes to attain good colour-constancy [11], long-term factors influence stable colour appearance and improve the colour-constancy achieved [12]. Jin and Shevell [13] have found that colour memory alters the perceived colour of surfaces and changes the mental standard used in experimental colour-constancy trials. The results, being consistent with the use of complex chromatic backgrounds reflect the importance of memory in these kinds of experiments [6,14].

To date, colour-constancy has been quantified using several indexes [7,12] but perfect colour-constancy is usually based on a full compensation for the illuminant.

* Corresponding author. Tel.: +34-958-246367; fax: +34-958-248533.
E-mail address: jnieves@ugr.es (J.L. Nieves).

0141-9382/02/$ - see front matter
PII: S0141-9382(02)00040-9

This point should be taken into account because many colour-constancy algorithms are designed to recover reflectance functions (mathematically as exact as possible) to characterise object colours [15]. The visual system reportedly uses a similar procedure to discount the variable illuminant conditions, a hypothesis which has been used by the colour-constancy algorithms based on a linear representation of surface reflectances and illuminants [6,16–20]. Recent colour-constancy models have used of affine invariant properties of colour signals at opponent stages [21]. This constant colour-appearance approach avoids the problem of the surface-reflectance recovery, given that all that is required is to evaluate the red–green and yellow–blue signals to identify illuminant changes [22].

Nevertheless, studies demonstrate that perfect colour-constancy is difficult to achieve unless both the subject's task and the visual performance are well specified (e.g. colour-constancy tends to be good when surface-match is required with several chromaticities in view). Bramwell and Hurlbert [23] have developed a forced-choice constant-stimuli procedure for measuring colour-constancy, which avoids several of the earlier drawbacks. Observers view the test and the reference surfaces haploscopically with each eye adapted to different illuminants. The observers' task is to respond 'yes' or 'no' in order to measure the locus of subjective colour-constancy. The results support the idea that colour-constancy is as good as the underlying colour-discrimination ability of each observer across illuminant changes. Another caveat is that the use of computer displays to simulate scenes (flat colour surfaces seen under simulated sources of light which are perceptually identical to real surfaces) can confound the observers' task. This is particularly pronounced for colour-constancy and leads to significant changes in the colour-constancy index. Kraft and Brainard [24] have shown that the best colour-constancy the observers achieve dramatically changed depending upon the cues considered in the scene (i.e. glossy highlights, mutual reflections, flat surfaces, three-dimensional objects, etc.). The study shows that constancy worsens when part of these cues are removed from the scene, although observer always achieved some residual colour-constancy.

With these considerations, we analysed some implications derived from the intrusion of memory in a typical separation algorithm for achieving colour-constancy. For this, a series of symmetric colour matches were made with 16 surfaces while keeping the illuminant conditions constant (equal-energy illuminant). Test and matching phases were separated by 1 min and no feedback was given; both a stable colour memory and a stable adaptation state were achieved with this procedure. If the colour-vision task were to recognise objects seen at different times and under different illuminations, we would expect the performance of memory and constancy to be matched. The distances between the test and matching surfaces were quantified according to colour-difference formulas, and compared with the recalled surface reflectance derived from linear models. To establish the mathematical differences between the recovered reflectances and the matches, we analysed the spectral representations via a goodness-of-fit coefficient (GFC) based on Schwartz's inequality.

In the second part of the experiment, the memory matches were also evaluated with a set of asymmetric colour matches made with variable illuminant conditions—the classical colour-constancy experimental paradigms. Our premise in this experiment was that the performance of colour-constancy algorithms should improve if the colorimetric differences were small enough (in comparison with pre-defined colour-differential thresholds) to ensure a good mathematical recovery of the spectral content. This is an important factor, which is not taken into account in the variety of colour-constancy separation schemes and algorithms proposed in the literature. Our approach resembles the spectral sharpening introduced by Finlayson et al. [25], who examined sensor transformations which led to optimal colour-constancy models. Although our aim is more modest and is not to introduce a separation scheme or algorithm, the results suggest that the traditional interpretation of perfect colour-constancy is a highly restrictive criterion. It does not reflect the observers' everyday experience of this phenomenon and thus one would not necessarily expect a satisfactory or perfect recovery.

2. Methods

2.1. Visual display and stimuli

Stimuli were presented on a Samsung CSD5577 colour monitor controlled by an 8-bit Tigastar graphic card. The monitor was calibrated using a Spectrascan PR-704 from Photo Research at 1 m from the screen. The values of spectral radiance for each gun were measured, maintaining the other two at zero and for the permitted DAC values (0–255). The calibration procedure generated stimuli according to their CIE XYZ tristimulus values, and was based on acceptance of the hypothesis of spatial independence with a simple scale factor [26,27]. Due to deterioration of phosphors with time, periodic calibrations were made to ensure stable chromaticity and luminance of the stimuli.

The surfaces were 16 square matte surfaces rendered under a standard illumination, and an equal-energy illuminant—characterised by a flat spectral power distribution and (0.333,0.333) chromaticity coordinates—was used as reference. The standard surface was presented at a central square area on the screen (1.7° of visual angle) and was surrounded by a Mondrian-type background ($11.3 \times 14.3°$ of visual angle). All the standard surfaces had a fixed luminance of 22.0 cd/m^2 and were distributed in different regions of the CIE-*xy* diagram. The chromaticity coordinates of the surfaces rendered under the equal-energy illuminant are shown in Fig. 1. The Mondrian background simulated chromatically complex scenes, and was made up

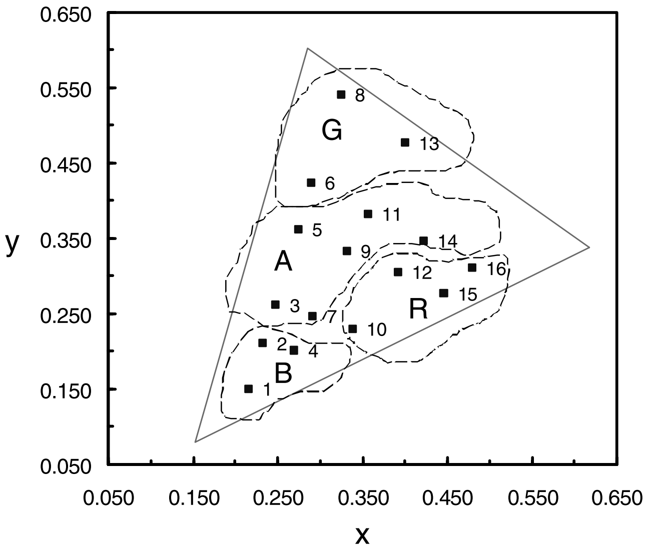

Fig. 1. Test surface collection. The plot shows the CIE-*xy* coordinates of the 16 standard surfaces when the surfaces were rendered under the reference illuminant (equal-energy illuminant). The surfaces were categorised into four classes according to their apparent hue: achromatic (A), blue (B), green (G), and red (R). The permitted gamut of colours is also shown (solid lines).

34 flat square matte colour surfaces selected from our own collection of colour samples [28]. The spatial-averaged chromaticity and luminance of the collection roughly equal the equal-energy chromaticity coordinates of the adapting background.

2.2. Procedure

A series of symmetric matches were made under the same test and matching illuminant conditions. An equal-energy illuminant was used as an adaptation standard. Before each experimental session the observers were adapted for 3 min to darkness and then for another 3 min to a uniform achromatic background with 18.8 cd/m^2 and the equal-energy chromaticity coordinates (Fig. 2). A beep indicated the end of this period and the standard surface was presented for 10 s surrounded by a Mondrian background. Observers were instructed to memorise hue, saturation, and brightness of the surface. To minimise colour-contrast phenomena over the central area, we randomly changed the shape and position of the surrounding surfaces during the standard phase. After the learning phase, the session continued with a new adaptation to the achromatic background during which the observers had to remember the standard surface. After 1 min, the matching phase began; the observers adjusted the central area of the screen, which was initially black, varying its chromaticity and luminance with the keyboard. Six keys permitted the observer to modify the red, green, and blue gun values, with two additional keys acting as a brightness control. No time limit was imposed for this task although the recommendation to the observer was less than 1 min. The experimental block lasted approximately 35 min and it was composed of eight trials with different standard surfaces. Each surface was matched five times, the total number of observations being 240 (16 surfaces × 5 times × 3 observers).

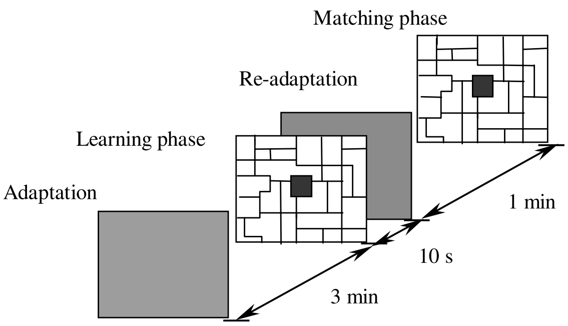

Fig. 2. Time course for the experimental sessions. In the learning and the matching phase, the display consisted of an array of randomly shaped and positioned colour patches (Mondrian surround) under the equienergy illuminant.

Table 1
Mean chromaticity coordinates (*x*,*y*) and standard deviations (SD) for each of matching surfaces. Data are grouped according to four apparent-hue collections into which surfaces were categorised

Surfaces		*x*			*y*		
		Mean	SD	Δx[a]	Mean	SD	Δy[b]
Blue	1	0.214	0.015	0.003	0.155	0.033	0.005
	2	0.230	0.020	0.003	0.206	0.032	0.004
	4	0.272	0.023	0.002	0.204	0.031	0.004
Green	6	0.290	0.024	0.001	0.429	0.045	0.006
	8	0.320	0.005	0.005	0.546	0.027	0.006
	13	0.394	0.022	0.008	0.483	0.031	0.006
Red	10	0.345	0.055	0.006	0.236	0.036	0.006
	12	0.397	0.031	0.004	0.305	0.020	0.001
	15	0.447	0.042	0.000	0.273	0.027	0.003
	16	0.479	0.030	0.001	0.305	0.015	0.005
Achrom	3	0.241	0.016	0.007	0.258	0.030	0.004
	5	0.271	0.018	0.005	0.368	0.019	0.006
	7	0.287	0.034	0.005	0.241	0.019	0.006
	9	0.330	0.017	0.003	0.333	0.009	0.000
	11	0.353	0.028	0.005	0.388	0.053	0.006
	14	0.422	0.034	0.001	0.350	0.032	0.004

[a] $\Delta x = |x_{\mathrm{Match}} - x_{\mathrm{Test}}|$.
[b] $\Delta y = |y_{\mathrm{Match}} - y_{\mathrm{Test}}|$.

2.3. Observers

The three subjects who took part in the experiment all had normal colour vision (Ishihara and Farnsworth D-15) and were corrected to normal acuity according to standard colour tests.

Table 2
Relative luminance ($\Delta Y/Y$) found for each observer. The luminance increment was derived as the difference between the observer's matches and the standard surfaces. The right column also shows the CIE-*xy* colour-difference (d*s*) between standard and recalled surfaces

Surfaces		$\Delta Y/Y$			d*s*
		JR	JH	FJ	
Blue	1	0.02	−0.02	0.05	8
	2	0.04	−0.06	0.11	2
	4	0.04	0.00	0.16	2
Green	6	0.09	0.15	0.19	4
	8	0.09	0.11	0.15	5
	13	0.14	0.05	0.17	8
Red	10	0.17	0.11	0.20	3
	12	0.06	0.06	0.16	3
	15	0.09	0.09	0.14	2
	16	0.08	0.09	0.14	3
Achrom	3	0.08	0.02	0.13	7
	5	0.06	−0.09	0.13	8
	7	0.00	0.04	0.15	3
	9	0.04	0.01	−0.04	3
	11	0.08	−0.04	0.11	7
	14	0.09	0.13	0.16	3

3. Results

3.1. Quantifying chromaticity and luminance differences

The results are analysed as a function of four apparent-hue collections into which standard surfaces have been arbitrary categorised (see Fig. 1). In the first stage, we quantify the chromaticity (Δx and Δy) and luminance (ΔY) differences between the match and the standard surfaces, and a repeated-measures analysis of variance (MANOVA) was made to test for possible interactions among matches of the surfaces by different observers. With a significance level of 0.05, the analysis showed no significant interactions between chromatic differences and surfaces ($F = 3.363$, $p = 0.208$). Also, the surfaces matched by the observers does not influence the chromatic differences ($F = 3.362$, $p = 0.208$), although chromatic differences Δx, Δy and ΔY for recalled surfaces are slightly different ($F = 10.44$, $p = 0.064$). Table 1 presents the average-chromaticity coordinates of all the matches; these average values will be use in the following calculations of the colour differences. The greatest standard deviations were 0.055 for *x*-coordinate (surface 10) and 0.053 for *y*-coordinate (surface 11), with Δx- and Δy-error being smaller than 0.008 between test and matching surfaces. The observers' errors tended towards the corresponding canonical colour, except for the achromatic group of surfaces that presented an almost equally distributed errors around the chromaticity coordinates of the standard colours.

With respect to the luminance variations, Table 2 indicates a general trend towards a relative increase of the recalled luminance. This trend is consistent with earlier studies of colour memory [29], and does not depend on the surface matched by the observers. The differences tend to be greater for the green and red surfaces (i.e. surfaces 13 and 10) and smaller for the blue surfaces.

In the second stage, we fixed a discrimination criterion to evaluate the earlier results. The assumption underlying colour-discrimination studies is that threshold discrimination contours form an ellipse in the chromaticity diagram (an ellipsoid in the three-dimensional colour space). From these contours [30–32], a set of colour-difference formulas can be derived in the form of values of the metric-tensor coefficients g_{ij} as follows,

$$(\mathrm{d}s)^2 = g_{11}(\mathrm{d}x)^2 + g_{22}(\mathrm{d}y)^2 + 2g_{12}(\mathrm{d}x)(\mathrm{d}y) \quad (1)$$

where the distance d*s* is widely acknowledged to be representative of minimum perceptible colour difference.

To evaluate the discrimination ability for symmetric-matching task in our experiment, we measured the colour differences revealed in the matches by means of a set of five colour-difference formulas, one for each of five regions in the CIE1931 chromaticity diagram [33]. The chromaticity tolerance for colorimetric purposes has been considered to be d*s* values below three colour-difference units. However, taking into account the inter-observer variability in colour-discrimination tasks [34], we also used colour tolerances of five colour-difference units. Nevertheless, in colour-discrimination experiments the observer's task is much more critical and restrictive than the ones used here (because of the imperfection of colour-constancy) as deduced elsewhere [8]. Thus we did not consider such a restrictive colour-difference unit in our colour-constancy experiments. Given the observers' everyday experience, higher colour-difference values could be applied as an error measure in these kinds of experiments.

The right column in Table 2 shows the colour differences associated with each of the surfaces matched. There are several surfaces with greater colour-difference units: surface 1 (blue), surface 13 (green), and surfaces 3 and 5 (achromatic), all with d*s* values around 5–8 units. These largest differences could be considered normal for technical applications. All observers concurred that it was extremely difficult to match the hue and saturation of these surfaces due to their proximity to the permitted gamut of colours in the CRT. By contrast, the other surfaces were well reproduced during matching, given the aforementioned restrictive chromaticity-discrimination criterion. These results and those from Table 1 suggest that matching and test surfaces would be seen as equal surfaces for chromaticity differences Δx and Δy around 0.008 and 0.006, respectively. These values can be extrapolated as error levels or estimators for colour-constancy experiments where illuminant changes occur. Below these values, colour-induced displacements derived from illuminant changes and the adaptational responses derived from the

Table 3
The CIELuv colour-differences separately for each of three observers. Data correspond to the colour-, chroma-, and hue-differences (labelled as ΔE^*_{uv}, ΔC^*_{uv}, and ΔH^*_{uv}, respectively)

Surfaces		ΔE^*_{uv}			ΔC^*_{uv}			ΔH^*_{uv}		
		JR	JH	FJ	JR	JH	FJ	JR	JH	FJ
Blue	1	9.4	11.3	8.6	−4.9	−8.0	−2.2	8.0	8.0	8.1
	2	7.7	3.2	12.0	7.5	2.1	11.3	0.3	0.3	0.3
	4	3.6	5.2	6.3	−3.3	−5.2	2.4	0.1	0.1	0.1
Green	6	7.6	10.1	12.1	6.8	8.5	9.9	1.0	1.0	1.0
	8	9.6	10.9	12.7	8.6	9.6	11.2	2.8	2.8	2.8
	13	11.7	9.1	12.9	6.9	4.0	7.9	8.0	7.9	8.1
Red	10	9.6	7.9	10.4	2.7	0.4	3.6	6.7	6.6	6.7
	12	5.9	5.8	9.9	5.0	4.9	7.8	1.9	1.9	2.0
	15	9.5	9.5	12.5	8.6	8.6	11.3	2.2	2.2	2.2
	16	9.7	10.3	13.5	8.2	8.8	11.7	4.2	4.2	4.2
Achrom	3	10.4	7.9	12.4	9.9	7.7	11.4	1.6	1.6	1.6
	5	8.8	5.8	11.3	8.3	4.2	10.0	1.9	1.8	1.9
	7	7.1	8.5	13.0	6.8	8.1	11.5	2.3	2.3	2.3
	9	3.2	2.8	3.2	2.7	2.7	2.7	0.7	0.7	0.7
	11	8.0	6.6	8.8	4.1	2.3	4.5	6.2	6.0	6.2
	14	4.8	6.1	7.1	0.1	1.2	2.0	3.5	3.5	3.5

observer's matches will be considered fairly equal, thereby achieving good colour-constancy.

The drawback of this analysis is that we have no information concerning the influence of the luminance in the colour-difference values evaluated from Eq. (1). Although the chromaticity differences between the test and matched surface were minor, we do not know to what extent the relative increase in the luminance derived from observers' matches changes the perceived colour. To solve this problem, we converted our data to the CIELUV uniform colour space [35] for evaluating the size of colour differences. In calculating the CIELUV colour-difference ΔE^*_{uv} we used the usual expression:

$$\Delta E^*_{uv} = [(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2]^{1/2} \quad (2)$$

where ΔL^*, Δu^* and Δv^* are the differences in the corresponding rectangular coordinates between the two colours. It is also useful to calculate the hue-difference ΔH^*_{uv} to provide additional information of the colorimetric differences:

$$\Delta H^*_{uv} = \left[(\Delta E^*_{uv})^2 - (\Delta L^*)^2 - (\Delta C^*)^2\right]^{1/2} \quad (3)$$

where ΔC^* is the difference in chroma between the two colours. The white point for the transformation to $L^*u^*v^*$ space was the tristimulus values of a perfect reflecting diffuser under the equal-energy illuminant.

Table 3 summarises the colour, chroma, and hue differences calculated for the three observers. As might be expected from the earlier results, the luminance differences derived from the matches induced similar changes in the computed colour-difference. The mean distance between the test and the match was 8.6 ΔE^*_{uv} units, with greater values corresponding to observer FJ. Three or five ΔE^*_{uv} units are typically taken as a measure of a just-noticeable difference in technical and industrial applications. Thus, for most observers the colours could be clearly distinguished (at least without any intrusion of memory in the experiment). On the contrary, the chroma differences (average value of 5.3) and the hue differences (average value of 3.2) were minor. This explains why the main significant effect of the memory in the perceived colour was the luminance variations instead of the chromaticity variations between the test and the matched surfaces. Furthermore, the relatively small values calculated for ΔC^*_{uv} and ΔH^*_{uv} suggest that memory had not completely degraded the perceived colour of the surfaces. However, we should not forget that the restrictive character of the criterion used was based upon colour-discrimination judgements. Hence, the question arises concerning the possibility of retaining other colorimetric characteristics of the stimuli, such as reflectance spectra for surface processing. Next, we evaluate the spectral differences derived from memory in comparison with the size of these colour differences.

3.2. Analysis of the recalled spectral-reflectance functions

The spectral-reflectance function is expected to be a significant factor for colour-constancy algorithms, since it is the colour signal defined from surface reflectance that specifies the colour of objects under different illuminants [15]. Regarding the earlier results, the question arises as to how errors in remembered surface colours may be translated into errors in reconstructed surface reflectance. We analysed the reflectance function revealed in the matches under equal-energy illumination to quantify the threshold for acceptability over the set of recalled surfaces. The method

Table 4
GFC for all the surfaces matched by each of three observers. The average values and the standard deviations are also shown for each individual collection (right columns), and for each individual observer (lower rows)

		JR	JH	FJ	Mean	SD
Blue	1	0.99683	0.99929	0.99951	0.99854	0.00098
	2	0.99739	0.99952	0.99831		
	4	0.99918	0.99796	0.99889		
Green	6	0.99706	0.99557	0.99591	0.99506	0.00553
	8	0.99646	0.99918	0.99947		
	13	0.98101[a]	0.99416	0.99671		
Red	10	0.99887	0.96224[a]	0.99969	0.99632	0.01074
	12	0.99984	0.99995	0.99876		
	15	0.99965	0.99951	0.99945		
	16	0.99886	0.99995	0.99911		
Achrom	3	0.99573	0.99498	0.99346	0.99395	0.00631
	5	0.99928	0.99648	0.98616[a]		
	7	0.99530	0.99951	0.99418		
	9	0.99968	0.99159	0.99843		
	11	0.99592	0.98042[a]	0.97893[a]		
	14	0.99966	0.99156	0.99976		
Mean		0.99692	0.99387	0.99604	0.99561[b]	0.00705[b]
SD		0.00453	0.00982	0.00577		

[a] Poorest recovery; GFC values under 0.99.
[b] Overall averaged GFC and standard deviation.

used to reconstruct the recalled reflectance function is based on the assumption that the reflectance can be closely approximated by the sum of a small number of basis functions [19,36,37].

Given the chromaticity coordinates and luminance of the recalled surface, the spectrum of its reflectance $S_r(\lambda)$ can be recovered from a linear combination of spectral basis functions S_i as:

$$S_r(\lambda) = \sum_i^m \langle S_r(\lambda) | S_i(\lambda) \rangle S_i(\lambda) = \sum_i^m \sigma_i S_i(\lambda) \tag{4}$$

where the symbol $\langle \cdot | \cdot \rangle$ represents the usual inner product, and σ_i are the basis coefficients. In the application of the earlier equation, the first three vectors of the Parkkinen's basis were used [38]. From Eq. (4), a simplified expression can be found for the coefficients σ_i (see Ref. [37] on pp. 317–318 for full details about the intermediate steps),

$$\vec{\sigma} = \Lambda^{-1} \vec{\pi} \tag{5}$$

in which $\vec{\pi}$ is the tristimuli vector which is known from the matches and Λ^{-1} the matrix which describes the colour signal. Note that the method requires that we know the spectral power distribution of the illuminant, which appeared in the definition of Λ^{-1}.

To compare the recalled reflectance revealed in the observers' matches with the colour-difference data, we evaluated the differences between the standard reflectance $S(\lambda)$ and the recalled reflectance $S_r(\lambda)$, using a GFC based on the Schwartz's inequality:

$$\mathrm{GFC} = \frac{\left| \sum_j S(\lambda_j) S_r(\lambda_j) \right|}{\left| \sum_j [S(\lambda_j)]^2 \right|^{1/2} \left| \sum_j [S_r(\lambda_j)]^2 \right|^{1/2}} \tag{6}$$

The spectrum was sampled from 400 to 700 nm ($\Delta\lambda = 5$ nm). The GFC values runs from 0 to 1, so that the mathematical reconstruction of the function $S_r(\lambda)$ would be better as the GFC values approach unity. Usually, Eq. (6) defines the square root of the variance-account-for coefficient (VAF). As a goodness criterion, we considered the following: GFC ≥ 0.99 represents good recovering for colorimetric purposes; GFC ≥ 0.999 indicates quite good recovering; and finally, values GFC ≥ 0.9999 signifies almost an exact mathematical recovering [16,39].

The results of this analysis are summarised in Table 4, which shows the GFC values for each of standard surfaces and observers. In addition, we determined the average values for each of four collections, for each of three observers, and for all the matches as a whole. The GFC was thus used as a measure of the accuracy of colour memory. The results reveal that exact mathematical recovery was found for only two of the surfaces (surfaces 12 and 16 for observer JH). Since there are evident spectral differences, this would suggest that the mental standard was degraded by memory, but only 10% of the recalled surfaces registered less than 0.99, indicating a good recovery. Furthermore, the recovery of the other surfaces is quite good from a colorimetric viewpoint; notably, more than 35% surpassed the 0.999 value considered to signify quite good recovery. In terms of inter-observer variability, we find small differences among observers, with GFC values exceeding 0.99 (good recovery).

With respect to the influence of hue in the matches, the average GFC within each hue collection shows low values for blue surfaces (close to 0.999). The achromatic surfaces appear to be more influenced by memory although the GFC values exceed the colorimetric recovery criterion. These results are clearly shown in Fig. 3. The worst result corresponds to surface 11, for which two of observers registered a poor GFC value (close to 0.9). Generally, however, the spectral characteristics of the reflectance recalled from memory were well retained. Thus, if our experimental goal were to reproduce surface spectral-reflectance, which is very common in colour-constancy algorithms and experiments, the images would be visually acceptable. As might be expected, the main differences are derived from vertical shifts along y-axis, indicating an upward trend in the recalled luminance.

It should be noted that the earlier results and discussion cannot be extrapolated to the human visual system. It is implausible that the visual system would look at the recover reflectance because it is inherently limited by its ability to

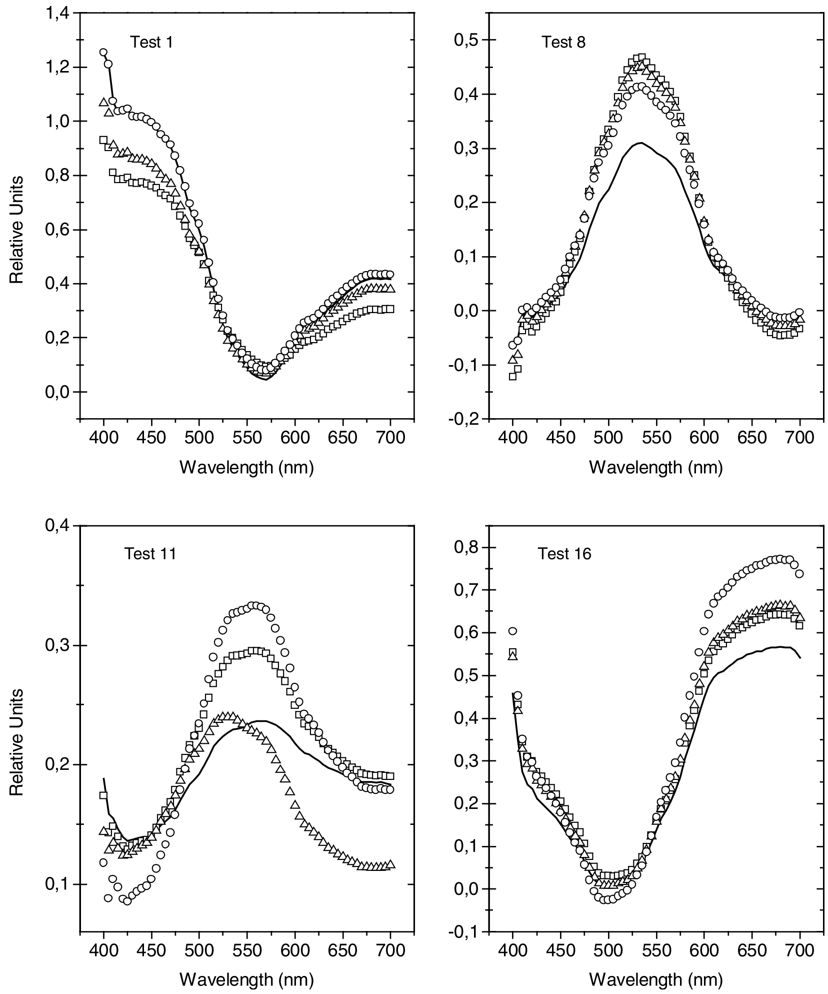

Fig. 3. Spectral-reflectance functions recalled from memory. The plot shows the matching surfaces set by each of three observers: JR (□), JH (△), and FJ (○); the reference surface (—) is also shown for comparison. Data are for examples of spectral recovery corresponding to blue (test-1), green (test-8), achromatic (test-11), and red (test-16) hue groups selected.

recover the full spectrum of the light signals it receives. However, colour-constancy separation algorithms can perform that operation and we would expect the performance of memory and constancy to be matched in this context.

3.3. Remarks on asymmetric matches derived from colour-constancy experiments

The earlier results indicate that discrimination criteria can be applied to colour-constancy algorithms and experiments. In these kinds of experiments asymmetric matches are usually made in which observers try to compensate for different illuminant changes. Although different indexes have been proposed to quantify colour-constancy [7,12], there are still problems to be solved, such as the reference illuminant. It has been suggested that the visual system could use either a canonical illumination to achieve colour constant appearance [40] or a similar equivalent-illuminant model that provides a good mental representation of object colours [14]. In addition, colour-discrimination can also influence colour-constancy [23]. Based on these ideas, we can use the earlier memory matches to estimate asymmetric matches when an illuminant change takes place. We did not assume an error in estimating the match illuminant, as Brainard et al. [14] have done, but a colour-constancy representation with respect to a confidence band in estimating the colour signal under different illuminants.

An example of this situation is shown in Fig. 4, in which surface-colour signals recovered from a classical colour-constancy experiment [41] are compared with the original standard surface-colour signals. These data were obtained by changing the illuminant conditions through the matching phase. Five test illuminants (an A-, 10000K-, D_{65}-, and F_{11}-illuminant) were used for comparison. Once the surfaces were matched, the spectra of the recalled surfaces were calculated from Eqs. (4) and (5). The illuminant component $E(\lambda)$ involved in Eq. (5) corresponds now to each of the new five test illuminants employed.

The results presented in Section 3.1 indicate that, in general, test surfaces were well retained in memory under equal-energy illumination. Nevertheless, poor recovery resulted for surface 11 (colour-difference approaching 10

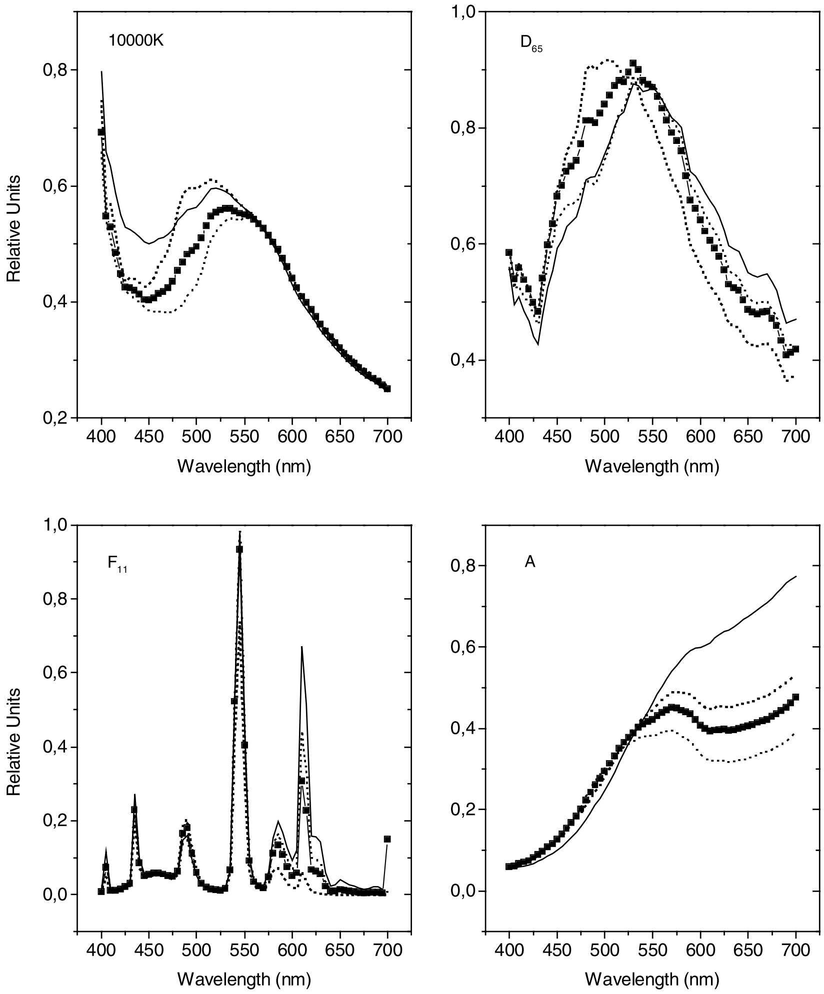

Fig. 4. Estimated colour-signal spectra (■) computed from observer's matches under different test illuminant conditions. Data are for surface 11 and observer JR. Each panel also shows the original or reference colour-signal function (—) and the confidence bands derived from colour-memory results (···).

units) and spectral differences appeared between the matched and test surface (see Table 4). Thus, with this surface, we tested the idea of using recovered reflectances instead of xyY coordinates as an additional way of supporting colour-constancy separation algorithms. For better visualisation of the influence by changes in the illuminant, we present the results as a function of the colour signal (product of the spectral-reflectance function and the power spectral distribution of the illuminant). Fig. 4 shows the recovered colour signals (solid symbols) derived from observer's matches and the original colour signals (solid lines) which were in turn derived from the standard surface reflectance for each of the test illuminants. If the test illuminant were completely discounted, there would have to be a perfect coincidence between the spectral-reflectance recalled and the surface test.

When we evaluate colour-constancy from colorimetric criteria—through a colour-constancy index similar to that proposed by Arend et al. [7]—we find a weak constancy for observers' matches under the 10000K- and the A-illuminant (constancy-index values of 0.41 and 0.35, respectively). On the other hand, when applying memory factors to this analysis a better constancy is suggested than is predicted by colorimetric considerations alone. This can be achieved by adding confidence bands to the reference surface (based upon GFC analysis in Section 3.2) as indicated in Fig. 4 with dotted lines. These confidence bands correspond to ± 1 standard deviation as derived from memory results by each of three observers. When this is done, spectral differences between predicted and obtained constancy for surface 11 diminish. The GFC values for all the illuminant conditions exceeded 0.99, with the A-illuminant being the only exception (GFC of only 0.8932). We may expect the results associated with this illuminant since the data derived from it also showed poor constancy. Consequently, it is clear that the colour-memory factors alone cannot account for colour-constancy. The results not only confirm the imperfection of colour-constancy but also reveal that colour memory could facilitate this phenomenon. It is not self-evident how the visual system would compute these confidence bands to generate the predictions found. However, in a qualitative or a quantitative way, one may expect colour-constancy algorithms to do so.

4. Discussion and conclusions

We have analysed the implications derived from memory in asymmetric matches made with and without illuminant changes. Several studies have reported the degradation of colour stimuli held in memory with very long adapting periods. Apart from rapid illuminant changes, where relational colour judgements can be made [42], colour-constancy is also linked to greater periods of time with slow illuminant changes. In these situations, observers must be able to retain a colour stimulus in memory, so that possible mental-standard deterioration influences colour-constancy [13,43]. To avoid this deterioration, observers can use familiar objects to ensure a reliable colour appearance; this possibility has not been considered here, as our aim was to isolate the observers' task from their familiar environment (familiar regarding particular shapes, sizes, etc.).

However, with this simple experiment, we have found that observers retain the surfaces memorised quite well. The adapting time used here was intermediary between the shortest and the longest in similar studies [13], but the results agree with those associated with complex backgrounds—a standard illuminant used during learning phase is more easily discounted when there are multiple surfaces in view. These results indicate, if we take into account the task involved, that colour differences between the standard and matching surface were small. For most of the surfaces colour differences are below 5 units, which can be considered normal in industrial or technical applications. Observers showed relatively good recall for hue and chroma of the standard stimuli. For luminance, we detected a relative increase that coincides with other findings concerning classical memory matches. In addition, the differences among observers were minor, so that the results can be extrapolated as thresholds or estimators for surface processing. This suggests new possibilities of retaining other characteristics of the stimuli, which can improve the evaluation of the degree of colour-constancy under illuminant changes.

Since surface colour perception is linked to the spectral properties of illuminants and surfaces in a scene, the earlier results were reinforced by the analysis of reflectance recovered from the matches. This was done to quantify the threshold for acceptability over the set of recalled surfaces. Although differences arose between standard and recalled reflectance, they proved negligible and thus we cannot consider the mental standard to have been completely degraded. Since vision is developed and adapted to the real world [24], the analysis and measurements of colour-constancy should also include the observer's visual capacities and abilities. From colorimetric considerations, the GFC, which evaluates spectral similarity, is proposed as a standard to evaluate the differences between the reference and the recovered reflectance. The GFC values derived from the experiments exceeded 0.99 which represented good recovering for colorimetric purposes [16,39].

However, separation algorithms for colour-constancy are intended to compensate for illuminant changes by using exact mathematical recovery of reflectance from a scene. This can be efficient but of little usefulness in the real world where other cues should be taken into account [24]. The results obtained here signal the importance of considering both the mathematical constraints on colour-constant separation algorithms and the constraints on visual performance for particular visual tasks. The goal of computational theories is to develop algorithms that achieve colour-constancy for a given set of surfaces. Occasionally, changes in colour appearance help the visual system discriminate aspects of its environment, i.e. sudden changes in weather (sunny-cloudy), walks through dense leafy sites, etc. Colour-memory data and colour-discrimination thresholds can be used in these situations as standards to enhance colour-constancy algorithms. Thus, tolerance bands or units could be defined over each surface in view, when colour-constancy must be evaluated in a scene under different illuminant conditions. It is a matter of future studies to analyse to what extent the accuracy of the computational theories can be improved by introducing these tolerance bands, which are derived from the limited capacities of human visual performance.

Acknowledgments

This work was supported by the Comisión Interministerial de Ciencia y Tecnología (CYCYT), Ministerio de Educación y Ciencia (Spain), BMF2000-1473.

References

[1] C.J. Bartleson, Colour in memory in relation to photographic reproduction, Photographic Science Engineering 5 (1961) 327–331.
[2] S.E. Clark, Retrieval of colour information from preperceptual memory, Journal of Experimental Psychology 82 (1969) 263–266.
[3] C.J. Bartleson, Memory colors of familiar objects, Journal of the Optical Society of America 50 (1960) 73–77.
[4] R.W. Burnham, J.R. Clark, A test of hue memory, Journal of Applied Psychology 39 (1955) 164–172.
[5] J. Pokorny, S.K. Shevell, V.C. Smith, Colour appearance and colour constancy, in: P. Gouras (Ed.), The Perception of Colour, Macmillan, London, 1991, pp. 43–61.
[6] D.H. Brainard, B.A. Wandell, Asymmetric colour matching: how colour appearance depends on the illuminant, Journal of the Optical Society of America A 9 (1992) 1433–1448.
[7] L.E. Arend, A. Reeves, J. Schirillo, R. Goldstein, Simultaneous colour constancy: papers with diverse Munsell values, Journal of the Optical Society of America A 8 (1991) 661–672.
[8] A. Valberg, B. Lange-Malecki, Colour constancy in Mondrian patterns: a partial cancellation of physical chromaticity shifts by simultaneous contrast, Vision Research 30 (1990) 371–380.
[9] K.T. Blackwell, G. Buchsbaum, Quantitative studies of colour constancy, Journal of the Optical Society of America A 5 (1988) 1772–1780.
[10] L.E. Arend, A. Reeves, Simultaneous colour constancy, Journal of the Optical Society of America A 3 (1986) 1743–11751.

[11] D.H. Foster, B.J. Craven, E.R.H. Sale, Immediate colour constancy, Ophthalmic and Physiological Optics 12 (1992) 157–160.

[12] I. Kuriki, K. Uchikawa, Limitations of surface-colour and apparent-colour constancy, Journal of the Optical Society of America A 13 (1996) 1622–1636.

[13] E.W. Jin, S.K. Shevell, Colour memory and colour constancy, Journal of the Optical Society of America A 13 (1996) 1981–1991.

[14] D.H. Brainard, W.A. Brunt, J.M. Speigle, Colour constancy in the nearly natural image. I. Asymmetric matches, Journal of the Optical Society of America A 9 (1997) 2091–2110.

[15] B.A. Wandell, The synthesis and analysis of colour images, IEE Transactions on Pattern Analysis Machine Intelligence 9 (1987) 2–13.

[16] J. Romero, A. García-Beltrán, J. Hernández-Andrés, Linear basis for representation of natural and artificial illuminants, Journal of the Optical Society of America A 14 (1997) 1007–1014.

[17] J.L. Dannemiller, Spectral reflectance of natural objects: how many basis functions are necessary? Journal of the Optical Society of America A 9 (1992) 507–515.

[18] M. D'Zmura, P. Lennie, Mechanisms of colour constancy, Journal of the Optical Society of America A 3 (1986) 1662–1672.

[19] L.T. Maloney, B.A. Wandell, Colour constancy: a method for recovering surface spectral reflectance, Journal of the Optical Society of America A 3 (1986) 23–33.

[20] G. Buchsbaum, A spatial processor model for object colour perception, Journal of the Franklin Institute 310 (1980) 1–26.

[21] Q. Zaidi, Identification of illuminant and object colors: heuristic-based algorithms, Journal of the Optical Society of America A 7 (1998) 1767–1776.

[22] J.L. Nieves, J. Romero, J.A. García, E. Hita, Visual system's adjustments to illuminant changes: heuristic-based model revisited, Vision Research 40 (2000) 391–399.

[23] D.I. Bramwell, A.C. Hurlbert, Measurements of colour constancy by using a forced-choice matching technique, Perception 25 (1996) 229–241.

[24] J.M. Kraft, D.H. Brainard, Mechanisms of color constancy under nearly natural viewing, Proceedings of the Natural Academy of Sciences USA 96 (1999) 307–312.

[25] G.D. Finlayson, M.S. Drew, B.V. Funt, Spectral sharpening: sensor transformations for improved color constancy, Journal of the Optical Society of America A 5 (1994) 1553–1563.

[26] J.A. Díaz, J.R. Jiménez, E. Hita, L. Jiménez del Barco, Optimizing the constant-channel chromaticity and colour gamut of CRT colour displays by control of brightness and contrast level, Applied Optics 35 (1996) 1711–1718.

[27] L. Jiménez del Barco, J.A. Díaz, J.R. Jiménez, M. Rubiño, Considerations on the calibration of CRT colour displays assuming constant channel chromaticity, Color Research and Application 20 (1995) 377–387.

[28] A. García-Beltrán, J.L. Nieves, J. Hernández-Andrés, J. Romero, Linear bases for spectral reflectance functions of acrylic paints, Color Research and Application 23 (1998) 39–45.

[29] S.M. Newhall, R.W. Burnham, J.R. Clark, Comparison of successive with simultaneous colour matching, Journal of the Optical Society of America 47 (1957) 43–56.

[30] G. Wyszecki, G.M. Fielder, New matching–matching ellipses, Journal of the Optical Society of America 61 (1971) 1135–1152.

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

[32] D.L. MacAdam, Visual sensitivities to colour differences in daylight, Journal of the Optical Society of America 32 (1942) 247–274.

[33] J. Romero, J.A. García, L. Jiménez del Barco del Barco, E. Hita, Colour difference formulas tested by use with chromatic discrimination data and new formulas, Applied Optics 28 (1989) 4697–4702.

[34] J.A. García, J. Romero, A. García-Beltrán, J.R. Jiménez, Interobserver variability of chromaticity discrimination and colour representation, Journal of Optics 24 (1993) 65–69.

[35] ASTM Committee, E-12, Standard practice for computing the colors of objects by using the CIE system (E. 308-95), in: ASTM standards on colour and appearance measurement, American Society for Testing and Materials, Philadelphia, PA, 1996, pp. 262–263.

[36] M. D'Zmura, G. Iverson, Colour constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces, Journal of the Optical Society of America A 10 (1993) 2148–2165.

[37] J.M. Troost, C.M.M. Weert, Techniques for simulating object colors under changing illuminant conditions on electronic displays, Color Research and Application 17 (1992) 316–327.

[38] J.P.S. Parkkinen, J. Hallikainen, T. Jaaskelainen, Characteristic spectra of Munsell colors, Journal of the Optical Society of America A 6 (1989) 318–322.

[39] J. Hernández-Andrés, J. Romero, A. García-Beltrán, J.L. Nieves, Testing linear models on spectral daylight measurements, Applied Optics 37 (1998) 971–977.

[40] L.E. Arend, How much does illuminant colour affect unattributed colors? Journal of the Optical Society of America A 10 (1993) 2134–2147.

[41] J.L. Nieves, A. García-Beltrán, J. Romero, Response of the human visual system to variable illuminant conditions: an analysis of opponent-colour mechanisms in colour constancy, Ophthalmic and Physiological Optics 20 (2000) 44–58.

[42] B.J. Craven, D.H. Foster, An operational approach to colour constancy, Vision Research 32 (1992) 1359–1366.

[43] E.W. Jin, S.K. Shevell, Colour memory: discounting of the illuminant during learning depends of the number of different illuminants experienced during recall, Investigative Ophthalmic and Visual Science 38 (Suppl.) (1997) 4214.