Vision Research 208 (2023) 108234

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Understanding the effect of correlated colour temperatures on spatio-chromatic properties of natural images

Juan Ojeda, Javier Romero, Juan Luis Nieves *

Department of Optics, University of Granada, 18071 Granada, Spain

ARTICLE INFO

Keywords:
Colour
Colour vision
Natural images
Correlated colour temperature

ABSTRACT

Despite the natural occurrence of global and local daylight changes in natural scenes, the human visual system typically adapts well to these changes and develops stable colour perception. In a previous study, the influence of daylight characterized by its Correlated Colour Temperatures (CCT) on different chromatic descriptors was analysed (Ojeda et al., 2017). The results showed that chromatic information is almost constant for CCT values above 14,000 K, with local extremes occurring in the range of low CCTs. The aim of this work is to extend the analysis of the CCT dependence of the illuminant to those that consider the spatio-chromatic structure, including second order descriptors (gradients, spectral slope, spectral signature, and PCA) and higher order descriptors (kurtosis, skewness, and number of relevant colours). Our results show that most of the descriptors exhibit horizontal asymptotic behaviour for CCTs above 15,000 K and local extremes in the range of 3,900 K-9,600 K. For those descriptors that could be analysed in CIELAB space, sufficient statistical evidence was obtained to consider skewness, kurtosis, and the independent spectral slopes of the L* channel as equal in the range of CCTs used. However, the slight variations in spectral signatures and the directions of the principal components when applying PCA to image patches are not statistically significant and cannot be considered equal under different illuminants. The number of relevant colours (NRC) exhibits sensitivity to temperature variations and behaves similarly to the other descriptors, due to its small number.

1. Introduction

The statistical characterization of natural images has not only allowed for the consideration of the probabilistic relationship between these images and the tasks being performed, as well as the quantum fluctuations inherent to light absorption and the noisy nature of neural chemical processes (Geisler, 2008), but also the confirmation that the human visual system has evolved to efficiently process this type of stimuli (Geisler, 2008; Simoncelli and Olshausen, 2001).

In adapting to the environment, the human visual system provides stable colour perception that is independent of local and global changes in daylight occurring in natural images. This phenomenon, known as colour constancy, is not perfect (Foster, 2011) and the effects of different illuminations on the colour and visibility of objects can be significant, particularly in terms of chromatic reproduction (Schanda, 2007; Masuda and Nascimento, 2013). As the colour signal characterizing the light entering the eye depends on both the spectral reflectance of objects and the spectral power distribution of the light illuminating them, characterized by their correlated colour temperature (CCT), the dependence of aspects related to the efficient encoding of the spectral distribution of the colour signal and chromatic diversity on different illumination conditions has been studied in previous work.

When applying principal component analysis (PCA) to the colour signal of hyperspectral images of terrestrial environments under different illuminants (Chiao et al., 2000), the principal components obtained provide similar amounts of variance and spectra, demonstrating the low influence of the first principal components with natural illumination variations and supporting colour constancy in these environments (Chiao et al., 2000).

However, other descriptors related to the chromatic richness of natural scenes do show some dependence on the illuminant. Martínez-Verdu et al. (2007) found that both the content and shape of colour solids in uniform colour spaces under different illuminants depend on the illuminant used. The colour gamuts corresponding to illuminants with CCTs in the range of 5,500–10,000 K (unless they have a very narrow bandwidth) are larger than those for illuminants outside of this range, with different colours appearing discernible under different illumination conditions.

* Corresponding author.
E-mail address: jnieves@ugr.es (J. Luis Nieves).

https://doi.org/10.1016/j.visres.2023.108234
Received 18 January 2023; Received in revised form 27 March 2023; Accepted 30 March 2023
Available online 25 April 2023

Masaoka et al. (2013) proposed that the number of discernible colours depends on the colour appearance model (hue, lightness, and chroma), the colour space, and the threshold of colour differences used. Their results showed that the number of discernible colours, without any Von Kries transformation of chromatic adaptation to the data, has a maximum around 4,000 K in the CIELAB and CIE94 colour spaces.

Masuda and Nascimento (2013) analysed the fidelity, or degree of naturalness, of the reproduced colours and the attractiveness, or preference, of 12 hyperspectral images of food counters as a function of the CCT of physically existing daylight and metameric illuminants (illuminant D65). For both types of illuminants, they found a clear dependence of these indices on the CCT of the illuminant source, with the most natural colours obtained at a CCT of 6,040 K for daylight and the most preferred colours at a temperature of 4,410 K for daylight, slightly higher at 6,200 K and 4,550 K, respectively, for the metamers of the D65 illuminant. When these descriptors were related to the chromatic range corresponding to each illuminant, it was found that the illuminants chosen for chromatic preference were related to those with the largest colour gamut, while those selected to optimize naturalness had more symmetrical ranges (with an aspect ratio closer to unity).

Foster (2021) used the spectral properties of the light reflected by each surface in hyperspectral image samples and the principles of information theory to computationally determine the minimum upper bound of the number of surfaces that can be considered chromatically the same over a given time period. While the average number of distinguishable surfaces was 10,000 in CIECAM02 space, the number of surfaces that could be considered the same after a 2-minute interval was 600, after 10 min it was 200, and after one hour it was 70. This significant reduction in the recognition of surfaces by colour, even at short time intervals, is much smaller when only pure global changes of the spectrum are considered. In this case, Foster (2021) estimated the number of chromatically equal surfaces under 4,000 K and 6,500 K illuminants to be between 4,400 and 5,500, depending on the internal noise model used.

In a previous study (Ojeda et al., 2017), the influence of daylight on different first-order statistical descriptors of colour (such as colour gamut characterized by the best-fitting ellipse parameters, colour volume, and number of discernible colours (NDC)) of 350 colour images simulated under 108 natural illuminants with CCTs in the range of 2,735 K to 25,889 K was analysed. The results suggested a low dependence of all descriptors for illuminants with CCTs above 10,000 K, as well as local extremes in a similar temperature range (2,950 K-6,300 K) to that in which the most natural and preferred colours or the highest NDC were obtained in previous studies (Masuda and Nascimento, 2013; Masaoka et al., 2013).

However, a comparison of means test applied to the distributions of colour volume and number of discernible colours for each category and the different illuminants determined that, at a 95% confidence interval, these descriptors could be considered statistically significantly equal. This result indicates that the palette of all colours in the scene at all available intensities, as well as the number of unit cubes into which we can segment the CIELAB colour space containing at least one pixel, are invariant, which could be due to the process of chromatic adaptation in the first stages involved in the phenomenon of colour constancy.

Our main hypothesis was that changes in CCT would have a significant impact on spatio-chromatic descriptors, and that understanding these changes would provide a more complete understanding of the impact of lighting on visual perception. For instance, visual segmentation requires to extract and to classify image statistics by pixel correlations (Haralick and Shapiro, 1992). Implicitly or explicitly, computer vision algorithms usually compare neighbouring image areas to get a region-based segmentation. In such approaches, understanding if CCT would change the chromatic properties of image features would help to implement those algorithms. We must consider that natural vision can easily segment image regions (and without effort can disambiguate surface changes from illuminant changes). Based on the hypothesis that changes in CCT can have a significant impact on spatio-chromatic descriptors, the knowledge of the influence of the illuminant on these descriptors is an important step towards understanding the ability of the visual system to adapt to different lighting conditions and may have practical implications such as lighting design or colour management. In order to further explore this dependence, the aim of this work is to extend the analysis of illuminant dependence from a previous study focused on chromatic descriptors (Ojeda et al., 2017) to second-order statistics such as gradients, spectral slope, spectral signature, and PCA, as well as higher-order ones like kurtosis, skewness, and number of relevant colours (NRC), which jointly take into account the spatio-chromatic structure of natural images.

2. Methods

2.1. Images

In the present work we have used 600 colour images from the MIT Scene UNderstanding (SUN) database (Xiao et al., 2010). Because some semantic categories did not contain enough examples to be representative, we decided not to use the whole set of images and categories and classified them into twelve semantic categories (of 50 images per category) depending on the content of the images: Forests and parks, fields, coasts, flowers and fruits, mountains, beaches, rivers and waterfalls, highways, cities, buildings, interiors, and streets. These in turn are grouped into two global categories (Rosch et al., 1976) called rural images (the first seven classes) and human environments (the last five).

To study the influence of the illuminant on the statistics of the natural images, each RGB image was normalized to the range (0,1) and its CIE tristimulus values calculated and simulated under a set of 108 spectral power distributions (SPD) characterized by their respective correlated colour temperatures (CCTs) in the wavelength range 400–700 nm (Hernández-Andrés et al., 2001). The simulation was done using Bradford's chromatic matching algorithm (Süsstrunk et al., 2000).

For the spectral power distribution (SPD) of the illumination source the data from Hernández-Andrés et al. (2001) obtained for hemispherical daylight were used. This data set is based on global spectral irradiances over a horizontal surface from direct sunlight, if present, and from the whole sky (Peyvandi et al., 2016), under different atmospheric conditions, over two years in Granada, Spain (a non-industrial urban area, of geographical coordinates 37°11′N 3°37′W and 680 m altitude) and covering a wide range of CCTs from 3,766 K to the 25,889 K (Hernández-Andrés et al., 2001). SPDs of daylights covering the range of CCTs below 3,766 K were simulated with SBDART (Peyvandi et al., 2016; Ricchiazzi et al., 1998) (since hemispheric daylight with CCTs below 3,600 K can be observed in rare conditions), a computer tool that allows to calculate radiative energy transfer within the atmosphere through parallel planes on the Earth's surface under different climatic conditions.

To take advantage of the property of uniform colour spaces in which the chromatic properties of the environment correspond to those of its visual representation (McDermott and Webster, 2012), finally the simulated images were transformed to CIELAB colour space (Ruzon, 2020).

2.2. Data analysis

The pixels of an image, in addition to the L*, a*, and b* values that define their colour, also contain information about their relative position in the scene. To account for this, it is necessary to use statistical descriptors of second-order (between pairs of pixels) or higher order (groups of more than two pixels).

2.2.1. Second order descriptors between adjacent pixels: Gradients

The second-order colour space information between adjacent pixels, which have been previously defocused using a Gaussian filter to reduce

noise and the effect of local edges, is characterized by the intensity gradients of the different signal planes. A large absolute value in the gradient of the L*, a*, or b* channel reflects an abrupt change in luminance or colour between adjacent pixels, allowing for the localization of edges of objects or surfaces.

To characterize the simultaneity of edges in the three planes L*, a*, and b*, and as an approximation of the joint probability density, the three joint histograms corresponding to the gradients in the different planes of each scene were calculated. Fig. 1 shows the edges and respective gradients joint histograms in a logarithmic scale, corresponding to an image of the flowers and fruits category, in which a greater grey darkness is associated with a higher frequency in the combination of values.

2.2.2. B.2 s order descriptors between non-adjacent pixels: Spectral slope and signature and PCA

The approximation of the second-order spatio-chromatic information for pixels that are not necessarily adjacent can be done both in the frequency domain and in the spatial domain. In the former, the spectral formulation of the autocorrelation based on the Wiener-Khinchin Theorem (Khintchine, 1934) is used to determine power and spectral signature descriptors. In the spatial domain, the covariance of each plane of an opponent colour space is used in the calculation of the PCA.

To focus on variations of pixel values and discard the mean value of each channel, we set the mean value of each plane to zero by subtracting its corresponding mean from the original values of each channel. To equalize the values of the edges of our images and reduce spectral dispersion that can generate unwanted effects when calculating the power spectrum (square of the modulus of the Fourier transform), we used the Kaiser-Bessel window function, which maximizes energy in frequency space for a limited energy and spatial extent (Harris, 1978).

When working with individual images, the dependence of the power spectrum on frequency can vary significantly, even between images of the same semantic category. However, when averaged across images of each semantic class on a logarithmic scale, results for spectral slopes close to 2 are obtained (Field, 1987; Burton and Moorehead, 1987; Tolhurst et al., 1992; Ruderman and Bialek, 1994). Table 1 shows the spectral slope values obtained for the two general categories in the L*, a*, and b* planes and three different orientations 0°, 45°, and 90°.

Although the spectral signature of a set of images is defined as the mean of the amplitude spectrum for all the images in the set (Torralba and Oliva, 2002), the term soon became associated with the isoline plot obtained by cutting the amplitude or power spectrum with different horizontal planes, so that the points enclosed by each isoline contain a certain percentage of the total energy of the amplitude or spectrum (Torralba and Oliva, 2003). Fig. 2 shows us in the first column, the neperian logarithm of the power spectrum for the three channels L*, a*, and b* of an image of the "beach" category. In the second column, each

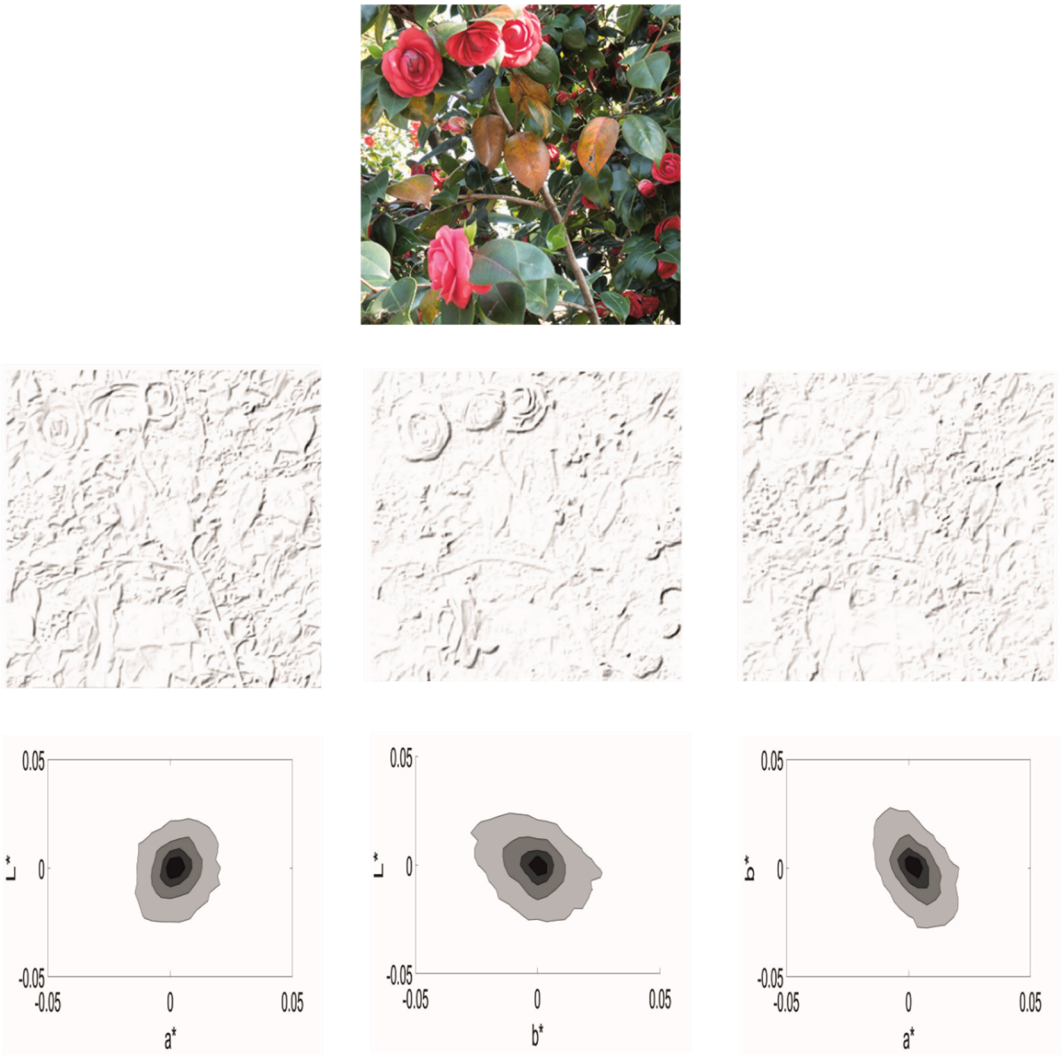

Fig. 1. In the top row, original image. The second row represents the edges corresponding to the three planes L*, a* and b*. Bottom row, joint histograms of the different channels.

Table 1
Mean slopes of the fit lines between the logarithm of the power spectrum and the logarithm of the spatial frequency for the three channels L*, a* and b* in three directions (horizontal θ = 0º, diagonal θ = 45º and vertical θ = 90º) for rural and human environment images.

	L*			a*			b*		
	Horizontal	Vertical	Diagonal	Horizontal	Vertical	Diagonal	Horizontal	Vertical	Diagonal
Rural images	−2,33	−2,48	−2,67	−2,16	−2,31	−2,58	−2,21	−2,37	−2,55
Images of human environments	−2,28	−2,46	−2,82	−2,18	−2,28	−2,60	−2,19	−2,31	−2,63

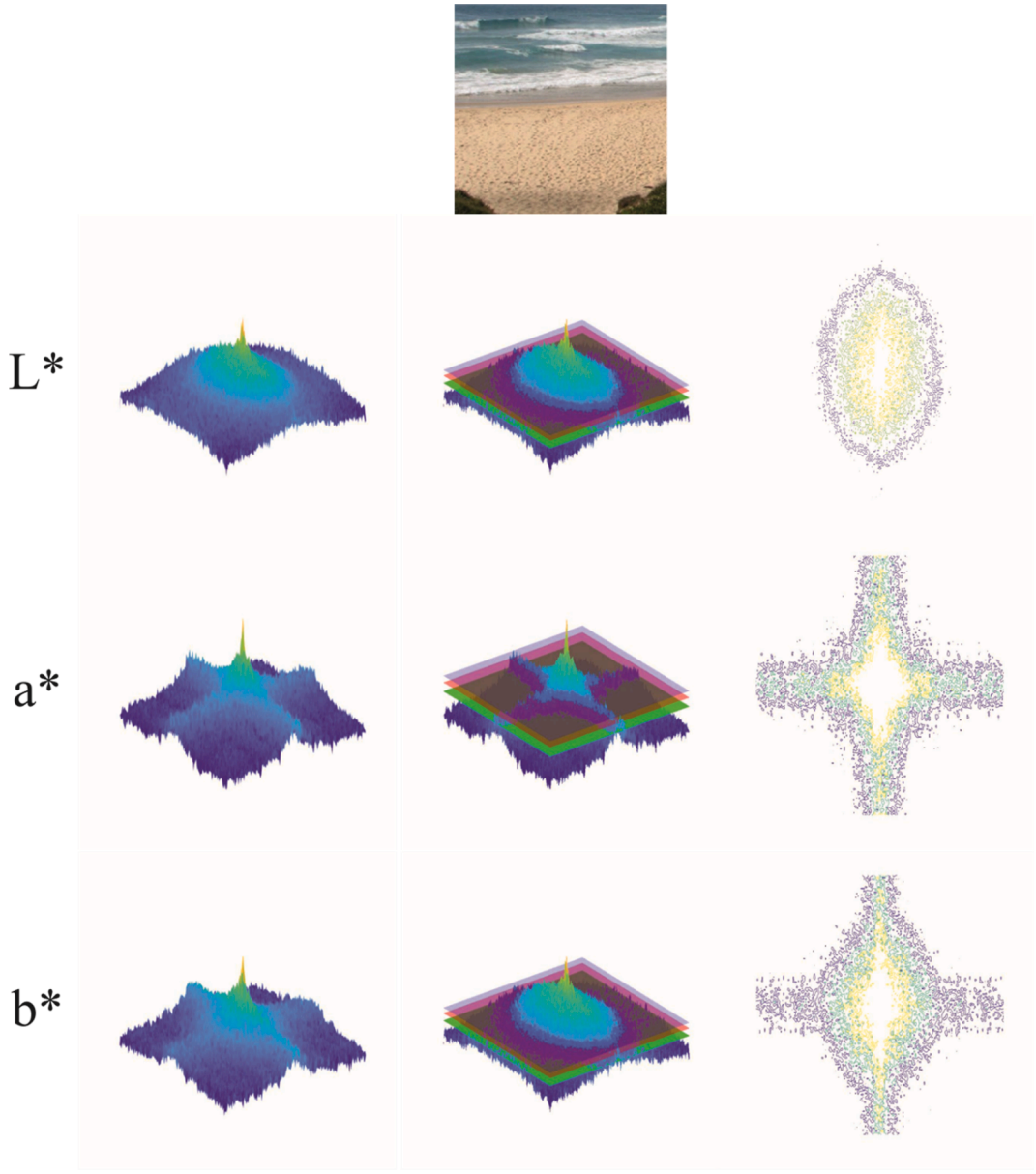

Fig. 2. Power spectra of the top image for the three channels L*, a* and b* (first column); intersection of horizontal planes with the power spectrum giving us the regions enclosing 60% (blue plane), 80% (red plane) and 90% (green plane) of the total energy (second column) and the corresponding spectral signatures (60% light green, 80% dark green and 90% blue) (third column).

of the spectra is intersected by three horizontal planes so that the part of the signal above them encompasses 60, 80, and 90% of the total energy of the signal (blue, red, and green planes, respectively). Finally, the last column represents the spectral signature corresponding to the three channels for the described energy levels (60% light green, 80% dark green, and 90% blue). The geometric structures obtained by maintaining certain amounts of energy provide us with global, non-localized information about the dominant orientations and the different spatial scales that compose the image. The components of the spectral signature at high frequencies indicate significant changes and details of the image texture; on the other hand, the components at low frequencies represent the main structure of the image (Fang et al., 2013).

To improve the representation of the global characteristics of the power spectrum as well as to include the variability that the individual images of each class would introduce in it, instead of a spectral signature, some authors (Fang et al., 2013) have proposed a distribution map

that consists in representing the contours of all the images of the class with a certain degree of transparency in a shared domain. This new representation provides more information about the less predominant orientations (other than horizontal and vertical) and at high frequencies so its use can improve the semantic classification of images. Fig. 3 shows the spectral signatures and their corresponding distribution maps along with the 80% energy contour (red line) of the rural and human environment images.

Principal component analysis (PCA) is a statistical technique that consists of a linear transformation of the coordinate system, so that the new orthogonal axes are oriented in the directions that capture, in a decorrelated way, the greatest amount of variance in the data. Principal components can be found by ordering the eigenvectors of the covariance matrix of our data, so that the corresponding eigenvalues are ordered in decreasing order.

To study the spatio-chromatic structure of our images, we applied PCA to a selection of 17,500 square patches of 8x8 pixels from both rural and human environment images (50 patches in each image from the first and 70 patches from the second). The pixel data of each patch were ordered as a column vector, where the values corresponding to the first dimension of the colour space were followed by those of the second and finally the values of the third (Tailor et al., 2000). The first column of Fig. 4 shows the principal components obtained from our rural images in the RGB colour space arranged from left to right and from top to bottom, according to the decreasing order of their eigenvalues (variance). From a spatial point of view, the principal components are neither localized nor oriented, and their variance decreases with increasing spatial frequency. By showing (Fig. 4 (b)) the chromaticities of the 192 pixel values of some of their components in an isoilluminant plane, where the horizontal axis corresponds to variations in R-G and the vertical axis to changes in B-Y and where the luminance is deduced from the luminosity of the points, they can be classified into three types (Provenzi et al., 2016):

- Homogeneous chromatic components: These have the greatest variance, no spatial structure, and are highly selective to one colour (Kellner and Wachtler, 2013).
- Achromatic components: They encode the luminance variations and have a greater variance than the opposite chromatic colour components.
- Non-homogeneous chromatic components: They are colour opponents, so their representation corresponds to a line that crosses opposite quadrants (44 and 190 of RGB) not necessarily coinciding with the direction of the axes. Within the latter, the variances decrease from the BY opponent characteristics to the RG opponents.

To quantify the pixel alignment directions of the principal component pixels of the individual patches, we again ran PCA on the coordinates of all pixels of each principal component and used the eigenvectors to estimate the same, (Kellner and Wachtler, 2013)

$$\alpha = atan\frac{PC1_2}{PC1_1} \quad (1)$$

where α is the angle formed by the first principal component with the horizontal positive semi-axis, and PC11 and PC12 are the coordinates of the first principal component obtained by applying PCA to the chromatic values of the different patches, considering them in an isoilluminant plane. The third column of Fig. 4 shows the histogram of the angles obtained for the 192 principal components. We see that most of them are close to a horizontal 0º or vertical 90º alignment, these corresponding to the colour opposition mechanisms in the different colour spaces used. However, a small number of components show different angles, as well as colours of their pixels that are different from the main ones (e.g., 81 and 108 of the RGB space), indicating that the opposing colour directions are still correlated to some degree in the data (Wachtler et al., 2001).

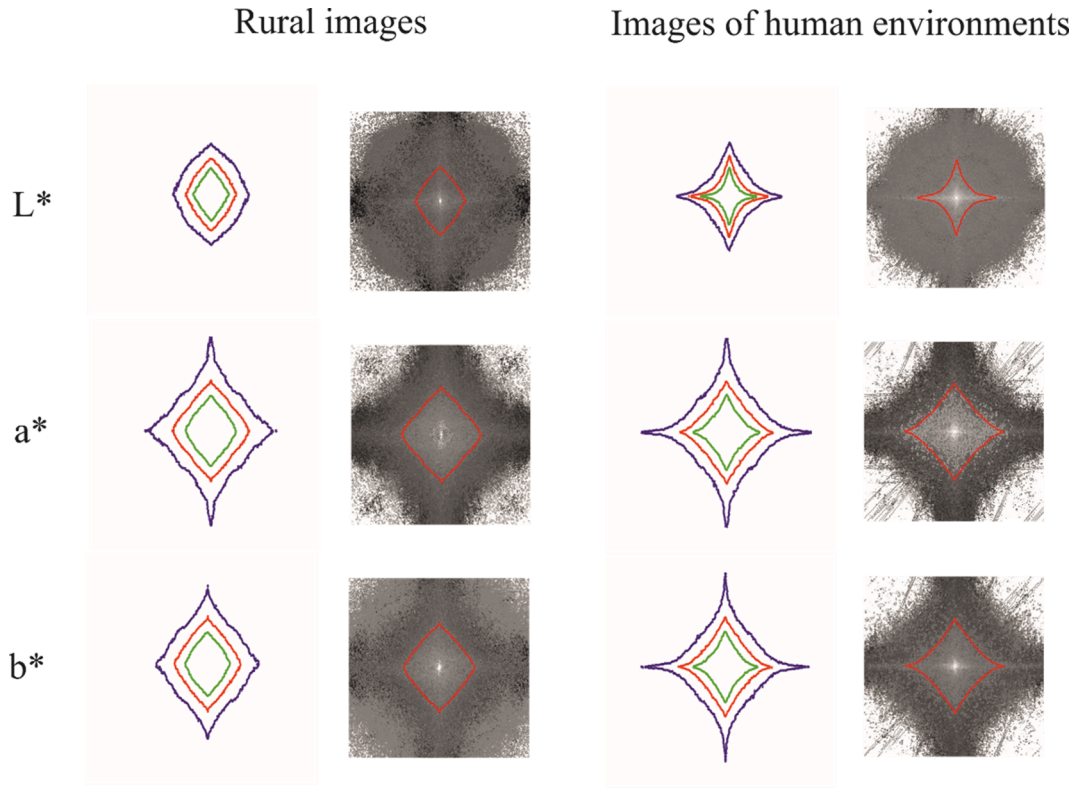

Fig. 3. Spectral signatures of L*, a*and b* planes of rural and human environment images (on columns 1 and 3) corresponding to 60% (green), 80% (red) and 90% (blue), and power spectrum neperian logarithm distribution maps (on columns 2 and 4) with the 80% energy contour (red line).

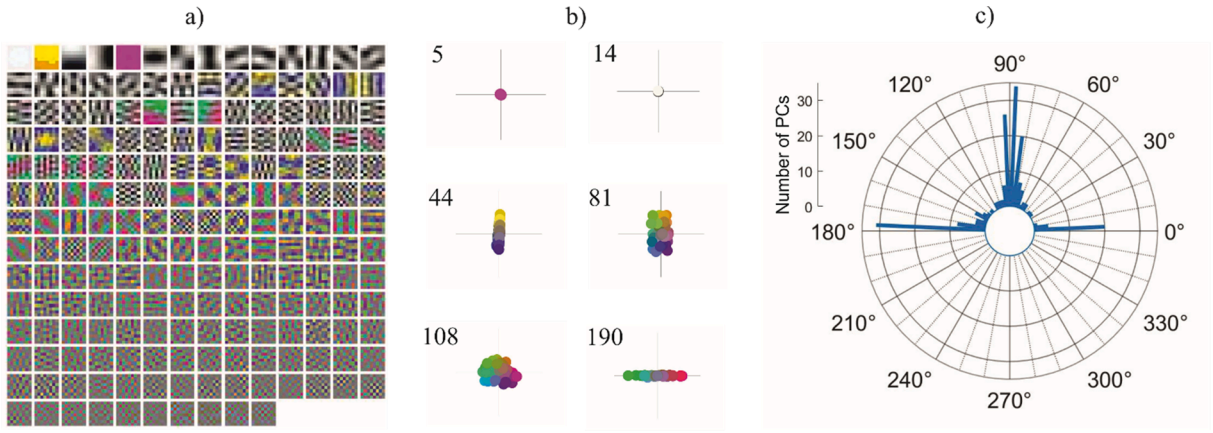

Fig. 4. (a) 192 principal components obtained from our rural images in RGB colour space arranged from left to right and from top to bottom according to the decreasing order of their eigenvalues. (b) Chromaticities of the 192 pixel values of some components in an isoilluminant plane. The horizontal axis corresponds to variations in the RG channel while the vertical axis corresponds to changes in the BY channel. The luminance can be deduced from the luminance of the points. (c) Circular histogram showing the alignment angles (between 0º and 180º) of the pixels of the principal components obtained.

2.2.3. *Higher order descriptors*

Since second-order statistics only uses amplitude information in the frequency domain, skewness and kurtosis are the simplest global statistics that can capture the phase structure in the spectral domain of the distribution of our data. (Reinhard et al., 2001; Thomson et al., 2000). These dimensionless measures which are worth by definition 0 and 3 respectively for Gaussian distributions of the data are defined as respectively,

$$S = \frac{E\{(x-c)^3\}}{E\{(x-c)^2\}^{3/2}} \quad K = \frac{E\{(x-c)^4\}}{E\{(x-c)^2\}^2} \tag{2}$$

Skewness characterizes the degree of skewness of a distribution around its mean being an indicator of the difference between the mean and median of a data set. Kurtosis is a measure of the lack of "normality" of the data distribution, focusing on the extent (size) of the tails of the distribution relative to a Gaussian. Prior to the study of these descriptors, the data were whitened using PCA, which not only removes second-order dependencies, but provides us with a distribution of zero mean and unit variance.

In a proposal in which it is not necessary to introduce predefined colour categories or to determine which areas are visually salient in the scene, Nieves et al. (2020) establish, based on the colorimetric content of each image, the novel concept of relevant colours, defined as the categorical discernible colours that describe the chromatic diversity of a pictorial work. The computational algorithm they applied allowed them to estimate a reliable colour palette for each of them based on 4 parameters: size of the cube into which the CIELAB space is divided, a minimum percentage (or general threshold) of the total pixels in a cube and to take into account that Milojevic et al. (2018) found that the most saturated colours can act as predictors of how an observer would categorize the colour distribution of natural objects, they considered as relevant colours also those that corresponded to those cubes that showed values of L* or C*=$\sqrt{a^{*2}+b^{*2}}$ greater than a given value. The colours considered as relevant are determined as the mean values of the pixel colours in each selected cube. The mean number of relevant colours (NRC) for all images analysed in that work was 18 (with a SD of 6).

In this work, we apply this algorithm to our set of natural images by setting the parameter value to 20 CIELAB units for the cube edge size,

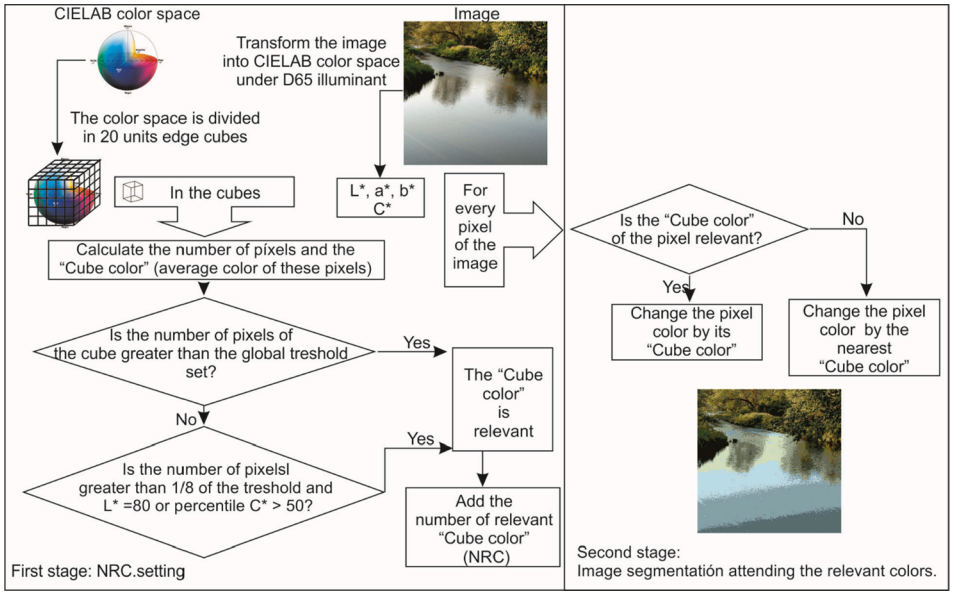

Fig. 5. Schematic of the algorithm for the calculation of the number of relevant colours (adapted from Nieves et al. [Nie20]) that we have applied to determine the NRC of each of our natural images.

4% for the threshold, a value of L* >80, and a percentile value of C* > 50. Fig. 5 shows a schematic of the algorithm used. Once the CIELAB colour space has been divided into edge cubes of the size of 20 CIELAB units ($\Delta E^* = 20$) and the RGB components of each pixel of an image have been transformed to that colour space (L*, a*, and b* components), we calculate the number of pixels that would be inside each of the cubes into which the colour space has been divided, and the mean value of the L*, a*, and b* components that we associate with the colour of the cube. From which we also determine the chroma of the cube.

If the number of pixels in the cube is greater than the overall threshold, then the colour of the cube will be relevant. If the number of pixels in the cube is between the threshold value and one-eighth of the threshold, we will also consider it relevant in the case where we can consider the colour of the cube as a saturated colour, i.e., when it is satisfied that $L^* \geq 80$ and the percentile $(C^*) \geq 50$. The number of cubes that, for each image, fulfil these conditions is the number of relevant colours, and the colour of each of the cubes, the colour of the "palette" of chromatically relevant colours in the corresponding image (Fig. 5 left).

The algorithm also allows for easy segmentation of the original image by simply checking for each pixel of the original image if it corresponds to a relevant colour. If this is the case, the pixel colour is replaced by its corresponding relevant colour; otherwise, the pixel colour is replaced by the colour of the nearest cube (from a Euclidean view) in the CIELAB colour space (Fig. 5 right).

Our results shown in Table 2 for the different semantic categories of natural images is like the one already mentioned in Nieves et al. (2020).

3. Results

3.1. Second order descriptors between adjacent pixels: Gradients.

Fig. 6 illustrates the joint histograms in logarithmic scale of the mean gradients in various planes for each scene for the categories of rural and human environment images under three different illuminants characterized by colour temperatures of 2,735 K, 6,478 K, and 25,889 K. All of them display a maximum near zero (since most pixels in both rural and human environment images do not correspond to the edges of the images, regardless of the daylight under which they are exposed), an approximately ellipsoidal distribution, and some degree of symmetry with similar widths in different regions.

However, when examining rural images, we see that in the L*-a* and L*-b* joint histograms, the luminance channel (L*) has a height that remains practically constant for the three temperatures, while for illuminants around 6,478 K, the chromatic channels (a* and b*) have higher frequencies for higher values in absolute value of their gradients (the histogram widens in that direction). In the a*-b* joint histogram, it is the yellow-blue opponent channel that has practically the same height, with the red-green channel having a histogram with higher frequencies for higher gradients around CCTs of the 6,478 K illuminant. This increase in edges is not in the direction of b* but in the direction of the bisector of the second and fourth quadrants of the joint histogram, resulting in a greater presence of light green edges. Images of human environments have joint histograms covering larger areas, indicating a greater presence of edges in the three channels than in rural images. With respect to the influence of these with illuminant temperature, we can observe that the highest frequency of higher gradients in the red-green plane is reached at 25,889 K in both the L*-a* and a*-b* histograms, while the highest number of edges in the yellow-blue plane occurs at low illuminant temperatures (2,735 K) in both the L*-b* histogram and the a*-b* histogram. When we extend the analysis to all values of the temperature range, the variations of the heights of the different channels with respect to their maximum values in the joint histograms decrease from the L* channel (8.3% in both categories), to the a* channel (33.3% for human environments and 50% for rural images) to the b* channel, which is the one with the highest dependence on the illuminant (42.8% for human environments and 57.2% for rural images).

It is observed that most edges combine information from two channels. The level of independence of these channels was determined by calculating the average mutual information between the gradients of the different channels, measured in bits (since 64 elements were used, the range should be between 0 and 6 bits). When analysing the mutual information in relation to the illuminant temperature, it varies between 0.01 and 0.594 bits, indicating that a gradient in one dimension cannot predict the most probable gradient in the other dimension. The mutual information between the red-green and yellow-blue channels has the largest range of variation with illuminant temperature, ranging from 0.123 to 0.306 bits for rural images and 0.158 to 0.260 bits for images of human environments. This decrease is observed across all semantic categories at a colour temperature of 5,858 K. This small mutual information suggests that the edges corresponding to the yellow-blue and red-green channels are not entirely independent and exhibit greater dependence between them, particularly at lower illuminant temperatures, than between the luminance channel and one of the chromatic channels.

3.2. Second order descriptors between non-adjacent pixels: Spectral slope.

When studying the behaviour of the spectral slope under different illuminants characterized by their colour temperature, and for three different orientations (0 degrees (horizontal), 45 degrees (diagonal) and 90 degrees (vertical)) and different semantic categories, we observed that the chromatic channels appear to be dependent on the illuminant, with their ranges varying between 0.027 and 0.25 for channel a* and between 0.087 and 0.664 for channel b*.

A statistical analysis of the mean slope distributions at different orientations in the L* plane was made using the Kruskall-Wallis test. A p value of 1.000 was obtained for all the semantic categories studied and the three orientations (horizontal 0°, diagonal 45° and vertical 90°) with a degree of freedom 107 and the values of the chi-squared of the table shown in the Appendix. These results revealed the mean slope of each category at each orientation can be considered significantly equal for different illuminants characterized by the CCT, at a 95% confidence interval.

3.3. Second order descriptors between non-adjacent pixels: Spectral signature.

Fig. 7 displays the spectral signatures and their distribution maps for all semantic categories and three illuminants corresponding to the colour temperature of 2,735 K, 6,498 K, and 25,889 K. The geometric structures obtained indicate that they are practically independent of the illuminant for the luminance channel. However, for the two chromatic channels, there are changes in the details and texture of the image, with variations in the high frequencies corresponding to directions other than

Table 2
NRC obtained for the different sets of natural images.

Forests and parks	Fields	Coasts	Flowers and fruits	Mountains	Beaches	Rivers and waterfalls	Rural images	Highways	Cities	Buildings	Interiors	Streets	Images of human environments
15 ± 4	10 ± 4	15 ± 6	22 ± 6	15 ± 5	10 ± 4	15 ± 5	15 ± 6	14 ± 5	20 ± 4	18 ± 5	19 ± 5	22 ± 5	19 ± 6

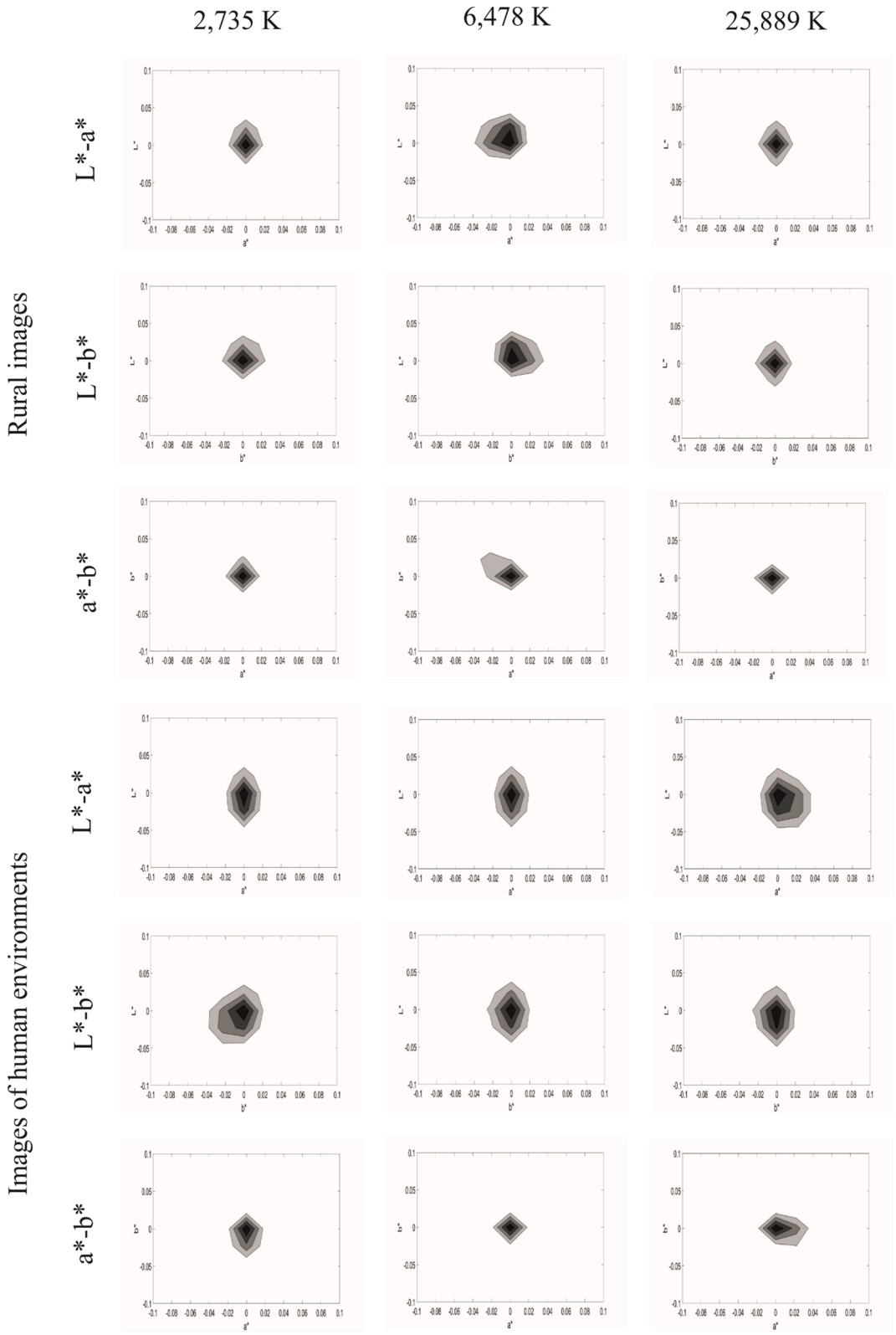

Fig. 6. Joint histograms in logarithmic scale for the three channels L*, a* and b* of the rural images and the images of human environments under three illuminants characterized by temperatures of 2,735 K, 6,478 K and 25,889 K.

0° and 90° being clearly observed in the distribution maps. These variations are more pronounced for illuminants corresponding to the more extreme temperatures than for 6,478 K, which are more concentrated in the 45° orientation. Meanwhile, the main structure of the image (associated with the low frequencies) remains generally stable, as well as the higher frequencies in the horizontal and vertical directions.

3.4. Second order descriptors between non-adjacent pixels: PCA.

In Fig. 8, the upper part shows the 192 principal components

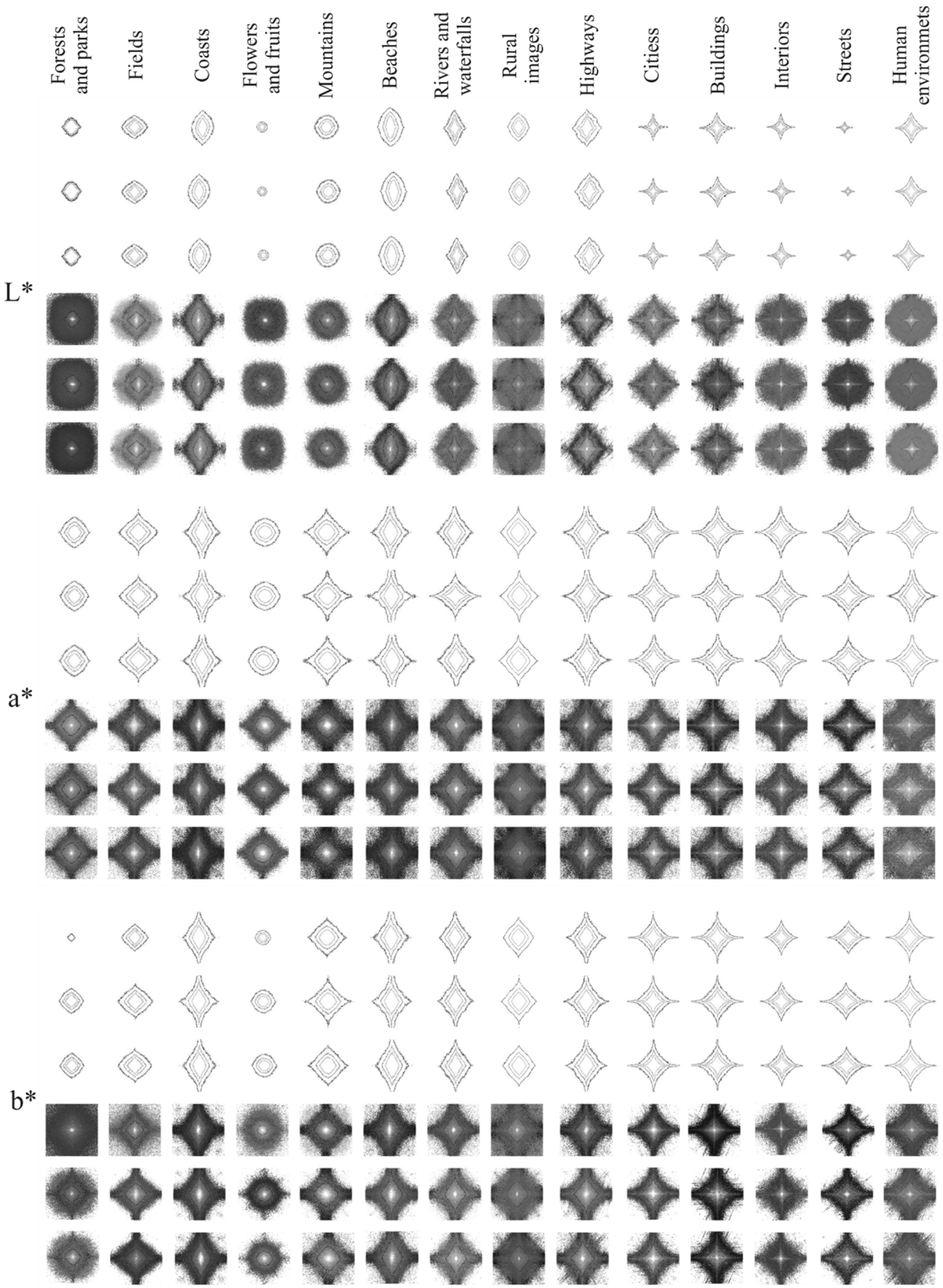

Fig. 7. Spectral signature (the first three rows of each figure) and distribution maps (the last three rows of each figure) for the three channels L*, a* and b* and the different semantic categories for three illuminants characterized by colour temperatures 2,735 K (the first row and the fourth row), 6478 K (the second row and the fifth row) and 25,889 K (the third row and the sixth row).

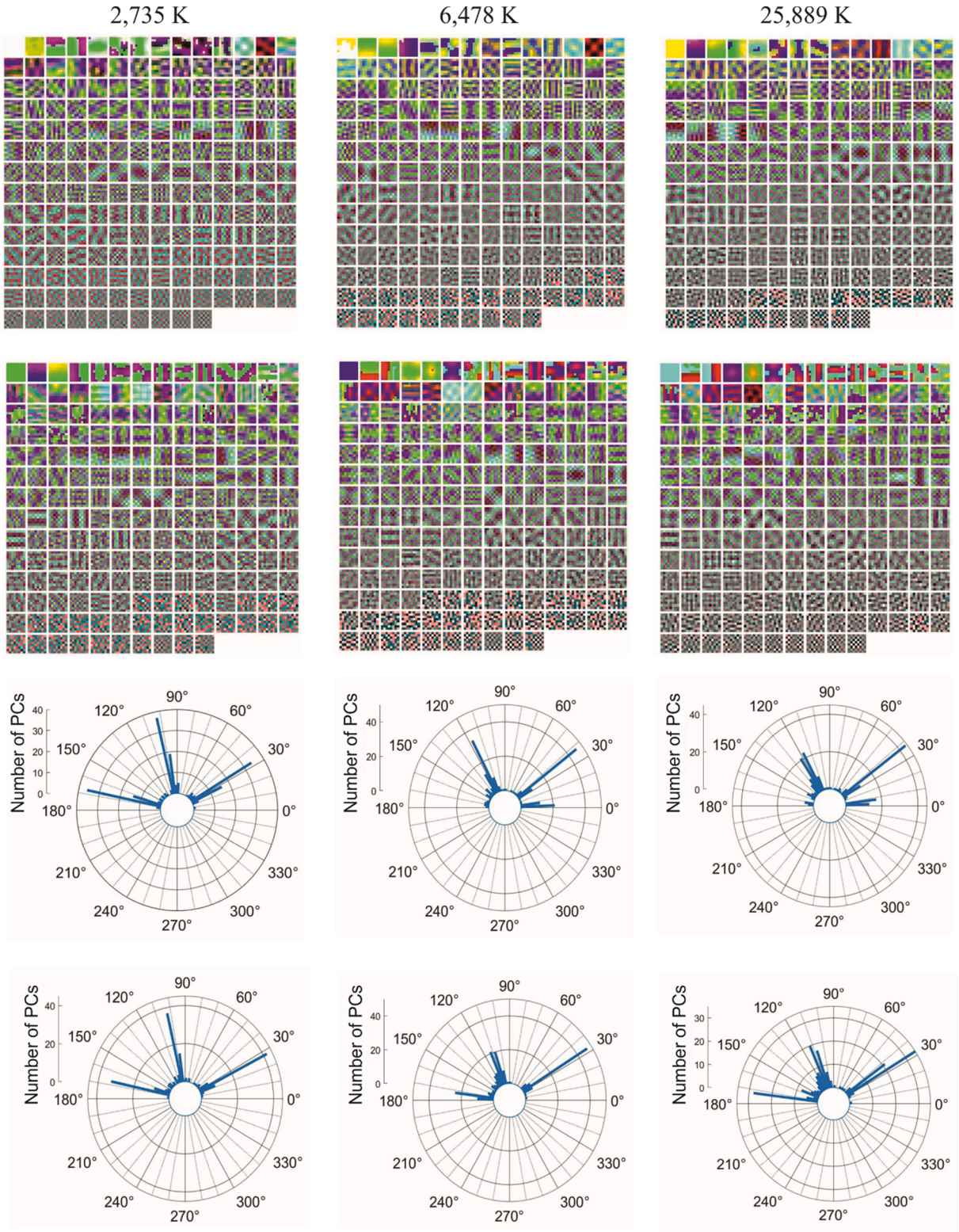

Fig. 8. The 192 principal components corresponding to 17,500 patches of rural (first row) and human environment (second row) images simulated in CIELAB colour space under three illuminants: 2,735 K, 6,478 K and 25,889 K, circular histograms corresponding to the angle (between 0° and 180°) formed by these components with the positive horizontal semi-axis for rural(third row) and human environment(fourth row) images.

obtained from the selection of patches of rural images (first row) and images of human environments (second row) in the CIELAB colour space under three illuminants (CCTs corresponding to 2,735 K, 6,478 K, and 25,889 K). The lower part displays the circular histograms of the angles (between 0° and 180°) of the directions obtained for the same. It is observed that the orientations, for both rural and human environment images, present three predominant directions: a first direction corresponding to values close to a horizontal alignment (10° for rural images under an illuminant with medium or high CCTs or 170° in other cases) and two directions approaching perpendicular to each other as the

illuminant CCT increases, varying between 30° − 40° and 100° − 120°, corresponding to colour opponent mechanisms. These directions coincide with those presented by the principal components of the original images. Also, for medium or high CCTs of the illuminant, more principal components appear with angles different from the most frequent ones noted, indicating a higher correlation between the colour opponent directions for illuminants with higher temperatures. Since the eigenvalue associated with each eigenvector of the eigenvector decomposition of the covariance matrix of the data explains the percentage of variance associated with that principal component, Fig. 9 shows the histogram corresponding to the eigenvalues of the first 25 principal components of the image patches under the three illuminants already discussed. As expected, for each illuminant, it shows a rapid decrease in the first components explaining >90% of the variance of the data with the 13 first components for rural images and 9 for those of human environments. Moreover, the variance explained by the first component is higher as the illuminant temperature increases (67.3%, 70.9%, and 71.0% for rural images and 69%, 71.5%, and 72.4% for those of human environments, respectively) and the differences from the fifth component onwards are less than half a percentage point.

3.5. Higher order descriptors: Skewness and kurtosis.

In Fig. 10, the top two plots show the skewness of the luminance channel for simulated images under 108 different illuminants. The skewness for the luminance channel is consistently negative, indicating that images have more light pixels than dark pixels, and is relatively independent of illuminant temperature in all semantic categories, except for a few outliers for rural images. The skewness of the chromatic channels, on the other hand, is relatively constant with illuminant temperature for temperatures above 15,000 K. In particular, the red-green channel, which is the most symmetric of the channels, shows local maxima in the range 4,600 K − 6,173 K for rural image categories with less asymmetric distributions and local minima at higher temperatures in the remaining categories. For the skewness of the a* channel in images of human environments, there is a local maximum around an illuminant temperature of 4,600 K. Depending on the category, this relative extreme can translate into an asymmetric distribution. Lastly, the skewness of the yellow-blue channel, which is always positive, shows the greatest dependence on the illuminant and a rapid growth for low temperatures up to the range of 4,600 K − 6,339 K, where all categories reach a local maximum, more pronounced for those categories that are more asymmetric, before decreasing more slowly to a horizontal asymptotic trend that indicates a lack of dependence of the descriptor with the colour temperature of the illuminant for high temperatures (>15,000 K).

The behavior of the kurtosis descriptor was analysed in a similar manner to the skewness descriptor. The results indicate that all semantic categories in all channels display a positive mean kurtosis and higher values in human environment categories. The luminance channel exhibits practical independence of its value over the entire temperature range, while the chromatic channels display little dependence on the CCT of the illuminant for high temperatures (>15,000 K). As seen in Fig. 11, a marked relative maximum in both chromatic channels is found in the range 5,345 K −6,340 K. The relative variation of kurtosis with respect to the maximum kurtosis value was on average 25.1% and 33.75% for the red-green and yellow-blue opponent channels respectively in rural images, and slightly lower at 24.1% and 26.1% in human environment images.

3.6. Relevant colours under different illuminant conditions

Lastly, Fig. 12 illustrates the relationship between the number of significant colours and the illuminant temperature within the range studied. Like other measures of colour richness, such as colour volume and NDC, it displays an asymptotic behaviour at temperatures above 15,000 K, with greater variations at lower temperatures and relative extremes and local maxima in the range of 3,900 K to 6,000 K.

When comparing the distributions of colour volume and number of discernible colours for each category and different illuminants using a mean comparison test, it was determined that for a 95% confidence interval, these descriptors can be considered statistically equal. Nevertheless, there is not enough statistical evidence to make the same conclusion when categorizing them using the NRC.

4. Discussion and conclusions

The analysis of the relationship between different descriptors used in this work and the variation of the spectral power distribution of daylight, as characterized by its correlated colour temperature (CCT), reveals three notable aspects. First, an asymptotic behaviour is observed for CCTs above 15,000 K, indicating that the statistics are independent of the CCT in that range. Second, the largest variations in the descriptors with CCT are found for illuminants with the lowest colour temperatures. This dependence is not caused by the use of two data sets to characterize the spectral power distribution of the illuminant source. Finally, in cases where a strictly monotonic relationship is not shown, the descriptors exhibit a relative extreme in the CCT range of 3,900 K-9,600 K. This temperature range coincides with those in which previous studies (Masuda and Nascimento, 2013) have found the most natural and preferred images by observers and the largest colour ranges in uniform colour spaces (4,410 K and 6,200 K respectively). Masaoka et al. (2013) found the maximum NDC without any chromatic adaptation of the data in CIELAB and CIE94 colour spaces around a CCT of 4,000 K, and Martínez-Verdu et al. (2007) found the largest colour ranges of colour solids in uniform colour spaces in the CCT interval 5,500 K-10,000 K. The location of these relative extremes around 6,500 K, which is associated with daylight, suggests that other aspects of colour vision may be optimized for these illuminations (Ojeda et al., 2017).

For those descriptors where it was possible to analyse the degree of their dependence on illuminant temperature for each of the different viewing mechanisms in CIELAB space (L*, a*, and b* channels), the L*

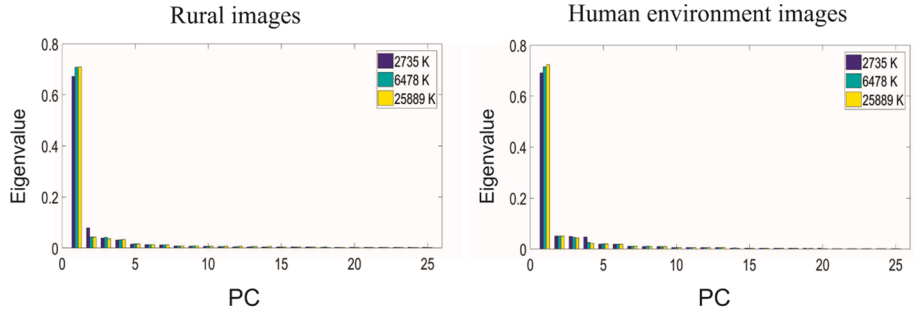

Fig. 9. Normalized eigenvalues of the first 25 principal components corresponding to the PCA of rural and human environment image patches under three illuminants: 2,735 K, 6,478 K and 25,889 K.

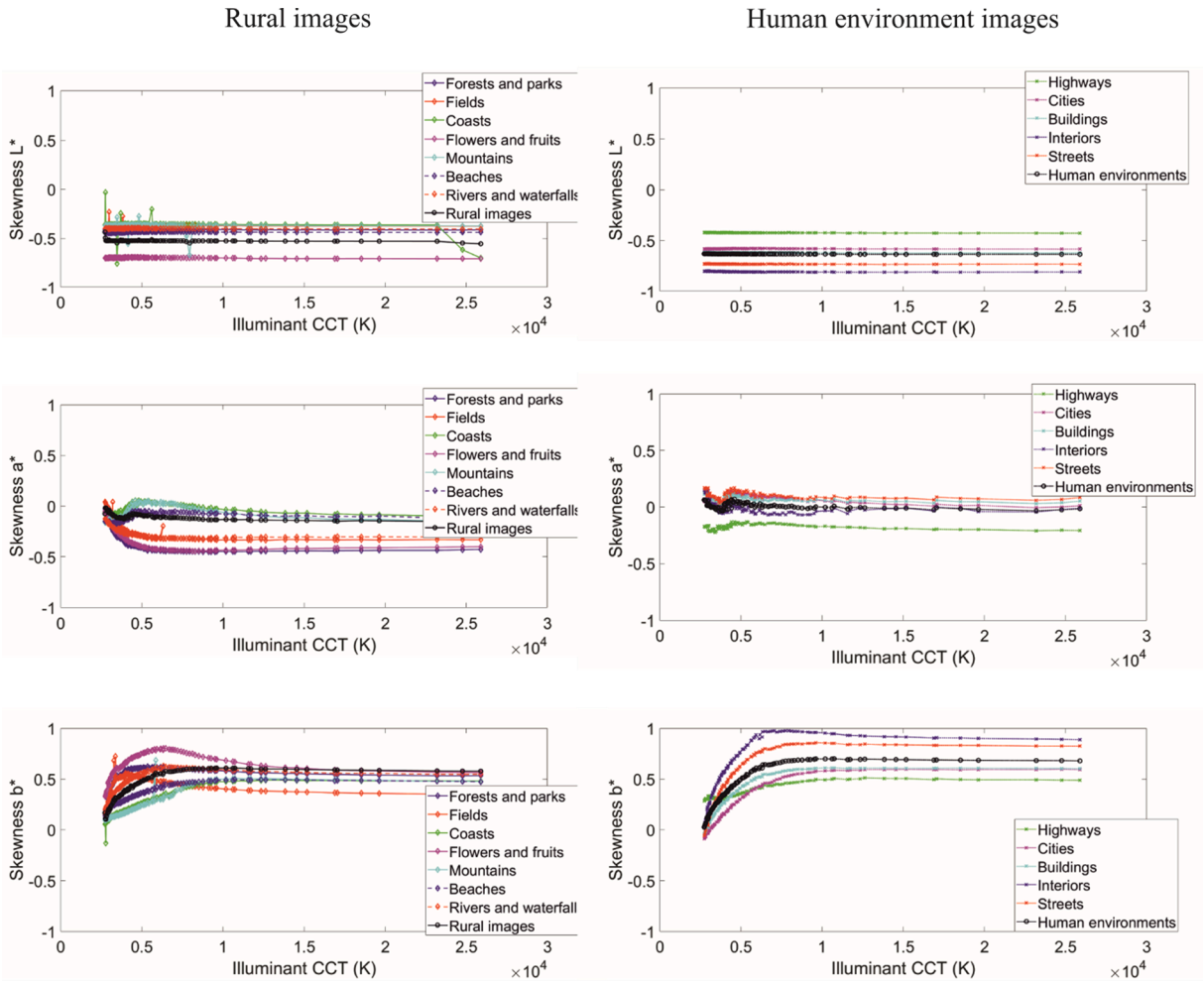

Fig. 10. Mean value of the skewness of the L*, a* and b* components for rural (first column) and human environment (second column) images as a function of illuminant CCT.

channel shows the least dependence on CCT. There is sufficient statistical evidence to consider the descriptors standard deviation, skewness, kurtosis, and spectral slopes independent of it. Of the two chromatic channels, a* and b*, it is always the latter that shows the greatest sensitivity to illuminant variations. This behaviour can be explained by the proximity of the spectral absorption peaks of the L and M cones (Lovell et al., 2005). From an evolutionary perspective, the lower sensitivity of the a* channel to changes in the illuminant allows trichromatic visual systems to optimally detect edible fruits or leaves throughout the day, except at dawn and sunset when the b* channel is more effective. However, other studies (Panorgias et al., 2012) have suggested that the vision mechanisms in higher primates related to blue-yellow variations have been greatly affected by the physical characteristics of daylight.

When all three viewing mechanisms in CIELAB space are considered simultaneously, there is not enough statistical evidence to consider them all independent of the CCT of the illuminant at a 95% confidence interval. While descriptors of chromatic diversity, such as colour volume and the number of discernible colours (i.e. first order descriptors), can be considered statistically significant, others, such as the patchy index (i.e. high order descriptors), are not, showing greater sensitivity to variations in the illuminant temperature, which could be due to the process of chromatic adaptation involved in the phenomenon of colour constancy (Ojeda et al., 2017).

The slight variations obtained in the orientations of the principal components and in their eigenvalues when applying PCA to image patches to study their spatio-chromatic structure, suggest a low dependence on the CCT of the illuminant, as indicated by the results of Chiao et al. (2000) and Nascimento et al. (2016) when applying this analysis to the colour signal spectra of natural hyperspectral ground images under different illuminants or local illumination conditions. However, these small differences, from a statistical significance point of view (Kruskall-Wallis test) for a 95% confidence interval, are not sufficient to consider the orientations of the principal components of the image slices equal under different illuminants, although they are sufficient to consider the eigenvalues equal.

CRediT authorship contribution statement

Juan Ojeda: Methodology, Investigation, Software. **Javier Romero:** Conceptualization, Writing – original draft, Supervision. **Juan Luis Nieves:** Visualization, Writing – review & editing, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

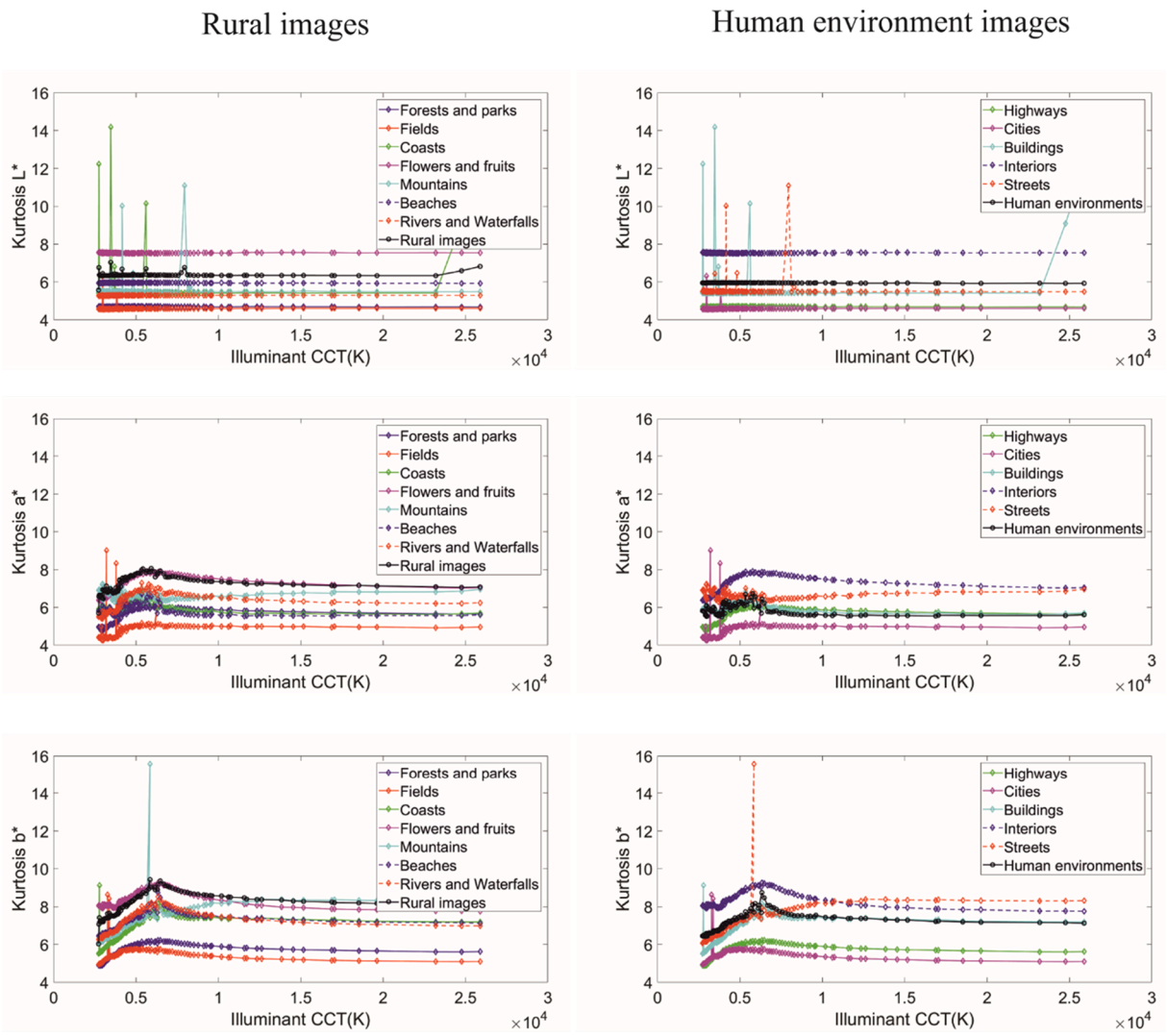

Fig. 11. Mean value of kurtosis of L*, a* and b* components for rural (first column) and human environment (second column) images as a function of illuminant CCT.

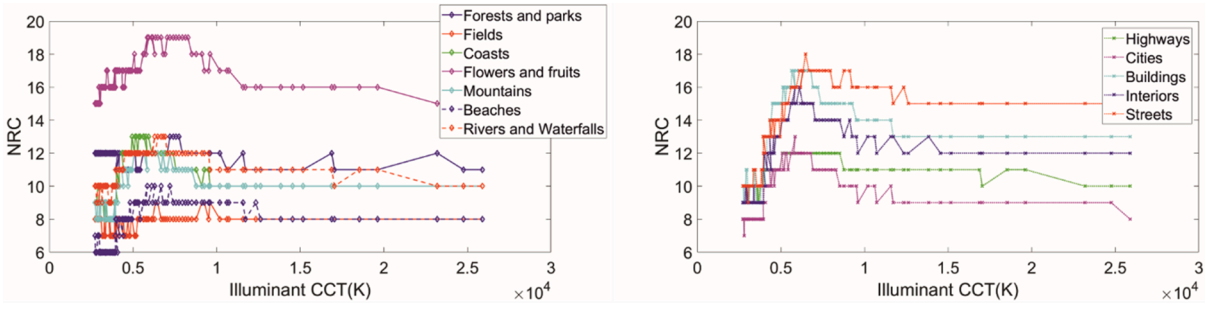

Fig. 12. NRC for rural (left plot) and human environment (right plot) images as a function of illuminant CCT.

Data availability

Data will be made available on request.

Acknowledgements

We thank Computational Colour and Spectral Imaging Erasmus+ master programme (610605-EPP-1-2019-1-NO-EPPKA1-JMD-MOB) for supporting the development of this work.

Appendix

A statistical analysis of the mean slope distributions at different orientations in the L* plane was made using the Kruskall-Wallis test. A p value of 1.000 was obtained for all the semantic categories studied and the three orientations (horizontal 0°, diagonal 45° and vertical 90°) with a degree of freedom 107 and the values of the chi-squared shown here:

	Forests and parks	Fields	Coasts	Flowers and fruits	Mountains	Beaches	Rivers and waterfalls	Highways	Cities	Buildings	Interiors	Streets
0°	0.545	0.226	0.337	0.889	0.501	0.321	0.223	0.191	0.267	0.097	0.108	0.296
45°	0.511	0.570	0.739	0.690	0.722	1.478	0.524	1.143	0.694	0.327	0.406	0.402
90°	0.564	0.139	0.217	0.906	0.374	0.239	0.071	0.380	0.150	0.179	0.279	0.523

References

Burton, G. J., & Moorehead, I. R. (1987). Colour and spatial structure in natural scenes. *Applied Optics, 26*, 157–170.

Chiao, C. C., Cronin, T. W., & Osorio, D. (2000). Colour signals in natural scenes: Characteristics of reflectance spectra and effects of natural illuminants. *Journal of the Optical Society of America A, 17*, 218–224.

Fang, H., Tam, G.-K.-L., Aubrey, A. J., Rosin, P. L., Wallraven, C., Cunningham, D., … Chen, M. (2013). Visualizing Natural Image Statistics. *IEEE Transactions on Visualization and Computer Graphics, 19*, 1228–1241.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4*, 2379–2394.

Foster, D. H. (2011). Colour constancy. *Vision Research, 51*, 674–700.

Foster, D. H. (2021). Fluctuating environmental light limits number of surfaces visually recognizable by colour. *Scientific Reports, 11*, 1–10.

Geisler, W. S. (2008). Visual Perception and the Statistical Properties of Natural Scenes. *Annual Review of Psychology, 59*, 167–192.

Harris, F. J. (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE, 66*, 51–84.

R.M. Haralick and L.G. Shapiro, Computer and Robot Vision vol. 1. Reading, M.A.: Addison-Wesley, 1992.

Hernández-Andrés, J., Romero, J., Nieves, J. L., & Lee, R. L., Jr. (2001). Colour and spectral analysis of daylight in southern Europe. *Journal of the Optical Society of America A, 18*, 1325–1335.

Kellner, C. J., & Wachtler, T. (2013). A distributed code for colour in natural scenes derived from center-surround filtered cone signals. *Frontiers in Psychology, 4*, 1–11.

Khintchine, A. (1934). Korrelationstheorie der stationären stochastischen Prozesse. *Mathematische Annalen, 109*, 604–615.

Lovell, P. G., Tolhurst, D. J., Párraga, C. A., Baddeley, R., Leonards, U., Troscianko, J., & Troscianko, T. (2005). Stability of the colour-opponent signals under changes of illuminant in natural scenes. *Journal of the Optical Society of America A, 22*, 2060–2071.

Martínez-Verdu, F., Perales, E., Chorro, E., de Fez, D., Viqueira, V., & Gilabert, E. (2007). Computation and visualization of the MacAdam limits for any lightness, hue angle and light source. *Journal of the Optical Society of America A, 24*, 1501–1515.

Masaoka, K., Berns, R. S., Fairchild, M. D., & Abed, F. M. (2013). Number of discernible object colours is a conundrum. *Journal of the Optical Society of America A, 30*, 264–277.

O. Masuda and S. M. C. Nascimento, "Best lighting for naturalness and preference", Journal of Vision 13 (7) : 4, 1–14 (2013).

McDermott, K. C., & Webster, M. A. (2012). Uniform colour spaces and natural image statistics. *Journal of the Optical Society of America A, 29*, 182–187.

Milojevic, Z., Ennis, R., Toscani, M., & Gegenfurtner, K. R. (2018). Categorizing natural colour distributions. *Vision Research, 151*, 18–30.

Nascimento, S. M. C., Amano, K., & Foster, D. H. (2016). Spatial distributions of local illumination colour in natural scenes. *Vision Research, 120*, 39–44.

Nieves, J. L., Gomez-Robledo, L., Chen, Y.-J., & Romero, J. (2020). Computing the relevant colours that describe the colour palette of paintings. *Applied Optics, 59*, 1732–1740.

Ojeda, J., Nieves, J. L., & Romero, J. (2017). How daylight influences high-order chromatic descriptors in natural images. *Applied Optics, 56*, 120–127.

A. Panorgias, J.J. Kulikowski, N. R. Parry, D.J. McKeefry and I.J. Murray, "Phases of daylight and the stability of colour perception in the near peripheral human retina", Journal of Vision Mar 1;12(3):1. (2012).

Peyvandi, S., Hernández-Andrés, J., Olmo, F. J., Nieves, J. L., & Romero, J. (2016). Colourimetric analysis of outdoor illumination across varieties of atmospheric conditions. *Journal of the Optical Society of America A, 33*, 1049–1059.

E. Provenzi, J. Delon, Y Gousseau and B. Mazin, "On the second order spatiochromatic structure of natural images", Vision Research, Mar;120:22-38 (2016).

Reinhard, E., Shirley, P., & Troscianko, T. (2001). *Natural image statistics for computer graphics* (pp. 1–9). University of Utah School of Computing. Technical report UUCS-01-002,.

Ricchiazzi, P., Yang, S., Gautier, C., & Sowle, D. (1998). SBDART: A research and teaching software tool for plane-parallel radiative transfer in the earth's atmosphere. *Bulletin of the American Meteorological Society, 79*, 2101–2114.

Rosch, E., Mervis, C. B., Gray, W. D., Johnsosn, W. D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.

Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters, 73*, 814–817.

M. Ruzon (2020). https://es.mathworks.com/matlabcentral/fileexchange/24009-rgb2lab.

Schanda, J. (2007). *Colourimetry: Understanding the CIE system*. John Wiley & Sons.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience., 24*, 1193–1216.

Süsstrunk, S., Holm, J., & Finlayson, G. D. (2000). Chromatic adaptation performance of different RGB sensors. *Proceedings of SPIE Colour Imaging, 4300*, 172–183.

Tailor, D. R., Finkel, L. H., & Buchsbaum, G. (2000). Colour-opponent receptive fields derived from independent component analysis of natural images. *Vision Research, 40*, 2671–2676.

Thomson, M. G. A., Foster, D. H., & Summers, R. J. (2000). Human sensitivity to phase perturbations in natural images: A statistical framework. *Perception, 29*, 1057–1069.

Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics, 12*, 229–232.

Torralba, A., & Oliva, A. (2002). Depth estimation from image structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*, 1–13.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation Neural Systems, 14*, 391–412.

Wachtler, T., Lee, T.-W., & Sejnowski, T. J. (2001). Chromatic structure of natural scenes. *Journal of the Optical Society of America A, 18*, 65–77.

J. Xiao, J. Hays, K. Ehinger, A. Oliva and A. Torralba, "SUN database: large-scale scene recognition from Abbey to Zoo", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3485-3492 (2010).