Available online at www.sciencedirect.com
SCIENCE DIRECT®
Computer Vision and Image Understanding 94 (2004) 92–114
Computer Vision and Image Understanding
www.elsevier.com/locate/cviu

Induction operators for a computational colour–texture representation

M. Vanrell,* R. Baldrich, A. Salvatella, R. Benavente, and F. Tous

Computer Vision Center, Dept. d'Informàtica, Edifici O, Campus Universitat Autònoma de Barcelona, Bellaterra 08193, Barcelona, Spain

Received 1 December 2002; accepted 29 October 2003

Abstract

The aim of this paper is to outline a perceptual approach to a computational colour–texture representation based on some colour induction phenomena. The extension of classical grey level methods for texture processing to the RGB channels of the corresponding colour texture is not the best solution to simulate human perception. Chromatic induction mechanisms of the human visual system, that has been widely studied in psychophysics, play an important role when looking at scenes where the spatial frequency is high as it occurs on texture images. Besides others, chromatic induction includes two complementary effects: chromatic assimilation and chromatic contrast. While the former has been measured by Wandell and Zhang [A spatial extension of CIELAB for digital colour image reproduction, in: SID, 1996] and extended to computer vision by Petrou et al. [Perceptual smoothing and segmentation of colour textures, in: 5th European Conference on Computer Vision, Freiburg, Germany, 1998, pp. 623] as a perceptual blurring, some aspects on the last one still remain to be measured, but it has to be a computational operator that simulates the contrast induction phenomenon performing a perceptual sharpening that preserves the structural properties of the texture. Applying both, the perceptual sharpening and the perceptual blurring, we propose to build a tower of images as an induction front-end that can be the basis of a perceptual representation of colour–textures.

Keywords: Colour; Texture; Colour–texture model; Perceptual sharpening; Perceptual blurring; Colour induction

* Corresponding author. Fax: +34-3-581-16-70.
E-mail addresses: maria@cvc.uab.es (M. Vanrell), ramon@cvc.uab.es (R. Baldrich), annasg@cvc.uab.es (A. Salvatella), robert@cvc.uab.es (R. Benavente), ftous@cvc.uab.es (F. Tous).

1077-3142/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2003.10.017

1. Introduction

Any scene of the world is projected on our retina as a map of different regions that are the projections of 3D surfaces. The properties of these projections are derived from the position and orientation of the surfaces in the scene, the observer location, and the light that provoke the neuronal excitation of the visual system. In computer vision, people usually deal with a set of surface properties, shape, orientation, colour, and texture. In this work we will only deal with the last two and their dependency on the final perception.

Although both are inherent properties of surfaces, these two visual cues have usually been studied separately [3,4]. This is probably due to their usual representations, while colour is a point feature given by the value of a pixel in several bands or channels, texture has to be modelled as a spatial relationship of the point with its neighbours. In Fig. 1 we see the RGB channels of a colour image, we can observe that the spatial information perceived from the colour image does not appear as it is in any of the channels. Therefore specific representations have to be built in order to deal with both cues at the same time.

The study of colour–texture representations has received increasing attention. The objective of many researchers has been to find co-joint representations of spatial and chromatic information which capture the spatial dependencies (in particular, correlation) *within and among* spectral bands. One of the most frequent approaches is to define a feature vector joining grey level texture features and colour features [5,6]. Another one is to extend classical texture models, such as Markov Random fields and the autocorrelation function, in order to deal with multichannel images [7,8], or wavelet analysis extended to colour images by combining the results in colour channels [9]. Other works, like [10], convert RGB values into a single code from which texture measurements are computed as if it were a grey scale image. Spatio-chromatic representations are computed in [11,12] over the smoothed Laplacian of the image, and the structural tensor that is usually used to represent local texture properties is extended to colour images in [13].

However, we want to highlight the approach presented in [2,14] that is based on known perceptual mechanisms of the human visual system. Colour–texture

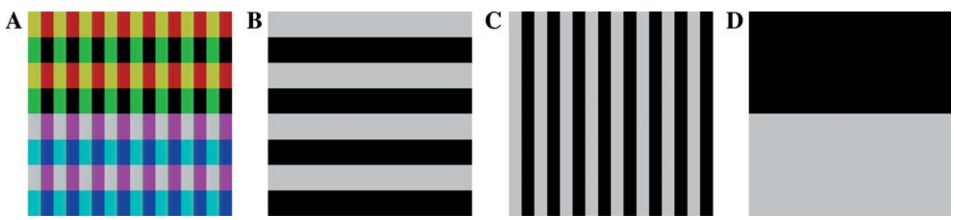

Fig. 1. (A) A colour image; (B) red channel of (A); (C) green channel of (A); and (D) blue channel of (A). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

interaction is represented as a perceptual blurring which depends on the spatial frequency of the coloured patterns and the observer position. This approach is based on important conclusions from psychophysical works on colour–texture interaction [1,15–18]. We will review the details of this perceptual blurring in Section 5.

Following on from this prior work, in this paper we propose a global perceptual approach of colour–texture representation that combines the introduced perceptual blurring with a first approach to a perceptual sharpening. The combination of these two operators can simulate the visual process of colour–texture perception with the colour induction mechanisms produced by different spatial frequencies.

To this end, in Sections 2 and 3 we give, separately, the basis of computational representations of colour–texture. In Section 4 we briefly introduce the colour induction mechanisms and their relationship to spatial frequency. Section 5 explains the essentials of perceptual blurring that simulates chromatic assimilation, and in Section 6 we introduce a perceptual sharpening that simulates the colour contrast phenomenon. In Section 7 we consider colour–texture perception as a visual process that integrates both induction mechanisms, this will be the basis for a global representation of colour–texture that is introduced in Section 8. Finally, in Section 9 we discuss some properties of the inductors operators and the advantages of using them in computing textural properties and in Section 10 we summarise the open problems that still remain in order to complete the proposed computational model.

2. Colour

Colour is the visual cue derived from the human visual processing of the electromagnetic radiation that reaches the retina. This process can be seen as a change in representation, which, in general, implies a dimensionality reduction. Although colour was not given much importance in the first decades of computer vision, since most previous work has been undertaken on grey-level images, the situation has changed and colour has become a very important visual cue for most of the vision tasks, such as object recognition [19], image indexing [20], tracking [21], shape extraction from colour variations [22], etc.

A computational representation can be easily built by a 3D representation. A wide range of different spaces can be used: *device-dependent* spaces are the most common in computer vision, this is the case of the RGB provided by any digitalisation process; other common representations are those *perceptually correlated* with colour properties, e.g. HLS; recently, different versions of *physiologically based* spaces try to provide the colour-opponency first described by Hering and finally demonstrated by Hurvich and Jameson [23]; from colour science and for calibrated conditions there are *uniform* spaces such as CIELAB or CIELUV [24] where Euclidean distances agrees with similarity judgements; finally, *invariant* spaces try to deal with colour representations that present invariance properties to the light and geometry of the scene conditions [8,19,25].

3. Texture

Texture is the visual cue derived from non-homogeneous surfaces in scenes. Depending on the surface reflectance, positioning of the observer, and lighting conditions, we can obtain different texture images from the same surface. Although there are some recent works dealing with the recovery of the physical reflectance properties of a texture [26,27] and some other works that have recovered 3D shape information from texture [28,29], the most traditional approach in computer vision has been the analysis of texture images without taking any consideration on the image formation process. Extensive reviews can be found in [4,30–32], where it is shown that texture has been studied for different purposes such as segmentation, classification or synthesis. Despite the large number of works, there is still a lack of a standard texture definition and does not exist a widely accepted texture representation space, as it exists for colour. Interesting work intended to define a standard texture space based on perceptual considerations should be considered [33–35], since such work could establish the basis of a standard computational representation. Before we go deeply into computational representations we will do a short inside on psychophysics theories on texture perception, that have been the basis for some of the works in computer vision.

In psychophysics, the aim has been to understand how the human visual system represents textures and which are the mechanisms used for texture segregation. Texture is one of the most complex visual cues and for the moment there is not a uniquely accepted theory. Two basic approaches are confronted as being the basis for a visual internal representation of texture. On one hand, a local feature extraction process has been supported by the Julesz's texton theory [36], and on the other hand a global spatial frequency analysis seems to be indispensable as it has been demonstrated by Beck et al. [37]. Let us go deeply in these two approaches.

The first approach, the Julesz's texton theory is based on the fact that differences between two textures, are due to differences in the first order statistics, or densities, of the texton attributes, it ignores the positional relationships between adjacent textons. Texton attributes are defined as the blob properties, these are, size and contrast for general blobs, and orientation for elongated blobs. Other textons can be line endings or terminators, but a more exhaustive list of texton has not been developed yet. Although all the texton theory conclusions are based on psychophysical experiments, Julesz associates the feature extractors with simple or complex cortical receptive fields described by Hubel and Wiesel [38].

The second approach, led by Beck et al. [37] and supported by other researchers [39,40] advocates that, differences in first order statistics of local properties independently of the blob arrangement is not enough to be able to capture the segregation of textures, since in a wide range of cases, differences are due to patterns emerging from the different arrangements of image blobs. In these cases a global spatial-frequency analysis is needed in order to represent different textures.

In Fig. 2 we demonstrate the complementary character of these two approaches. While the textures (A) and (B) can be easily differentiated in the frame of the Julesz's texton theory due to differences in blob contrast; textures (B) and (C) are equal

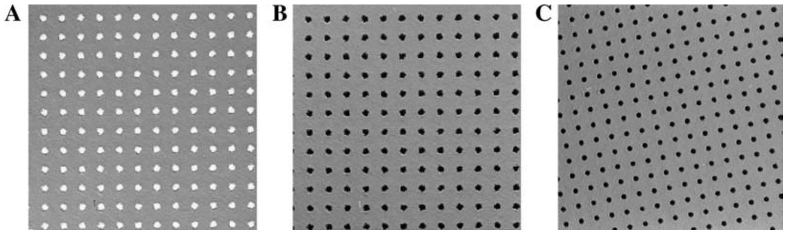

Fig. 2. Examples of textures formed by simple blobs and their emergent patterns.

according to this theory, since there is no difference in terms of texton attributes. Differences between textures (B) and (C) can be easily derived in the frame of a global frequency analysis, for which a difference in emergent orientations can be considered. Therefore, any texture representation will have to combine both, global and local properties.

4. Colour induction

Colour induction are the colour phenomena that changes the colour appearance of a stimulus due to the influence of the scene contents in the field of view. There are different types of induction phenomena such as colour adaptation, colour assimilation or colour contrast amongst others.

Colour adaptation is involved in any scene interpretation and occurs when scene colours are perceived without being affected by the illuminant influence, this ability presented by the HVS has been modelled in computer vision by different colour constancy methods [41–43] that usually imply global image transformations.

Colour assimilation does not depend on global colour illuminant effects but on the direct surrounding colour of a certain stimulus [44], this surrounding colour is usually called the inducing stimuli or inductor. As colour assimilation, colour contrast does depend on the direct surrounding colour acting as an inductor, but it can also depend on remote inductors acting as a global colour illuminant [45–47]. In this work, and because we are on defining a colour–texture model, we will deal with those induction mechanisms that only depend on the direct surround. Thus we assume that inductor stimuli are only based on contrast edges of the surround and therefore we will regard to dependencies on local spatial frequencies [48].

Without regarding to causes provoking induction, the essential difference between the effects of contrast and assimilation mechanisms is the direction of the chromatic change provoked by the inductor, in this case, the surrounding colour [49].

Colour contrast occurs when the chromaticity of the test stimulus changes away from the chromaticity of the inducing stimulus, an example of this effect can be seen in Fig. 3A, where we can see how a given test stimulus, T, is perceived as P1 when surrounded by S1 and as P2 when surrounded by S2, that is, P1 is closer to S2 and P2 is closer to S1.

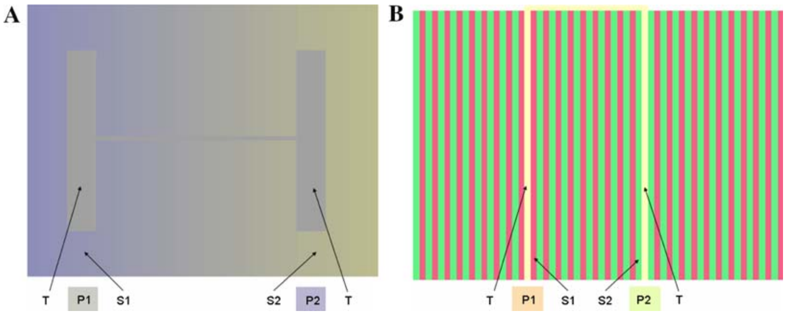

Fig. 3. Colour induction. (A) Chromatic contrast; (B) chromatic assimilation.

Colour assimilation occurs when the chromaticity of the test stimulus changes towards the chromaticity of the inducing stimulus, assimilation effects are shown in Fig. 3B, [1] where we can see how a given test stimulus, T, is perceived as P1 when surrounded by S1 and as P2 when surrounded by S2, in this case P1 is closer to S1 and P2 is closer to S2. Chromatic coordinates of these perceived samples are shown in Fig. 4. For colour contrast, perceived stimuli moves away from the corresponding surround, and for colour assimilation chromaticity moves towards the surround.

Chromatic coordinates of these perceived samples are shown in Fig. 4. For colour contrast, perceived stimuli moves away from the corresponding surround, and for colour assimilation chromaticity moves towards the surround.

Considering the given definitions and examples, it is obvious that any perceptual approach towards a colour–texture representation should take into account the colour induction effects we have introduced above. In psychophysics we find a wide range of works dealing with the induction phenomena or the influence of surrounding chromaticity on the appearance of colour [15–17,47,48,50–53]. In all these works, the authors present different aspects of colour human induction measurements. The influence from direct surrounds or remote inducers, the asymmetry of the measurements due to changes from luminance or the dependency on spatial frequency of patterns are some of the aspects that are measured and analysed. Conclusions from all these measurements help to provide answers about how the perceptual mechanisms are organised in the human visual system. They help in building a more precise model of how the human visual system acts from the retinal representation of colour to the final judgements of colour appearance. Considerations are done in terms of different physiological aspects as cone absorption rates and their retinal distribution, optical chromatic aberrations or the existence of opponent-colour signals in the visual pathways.

[1] Appearance of colours can vary depending on the printing device and therefore the induction effects may vary substantially if images are not exactly reproduced.

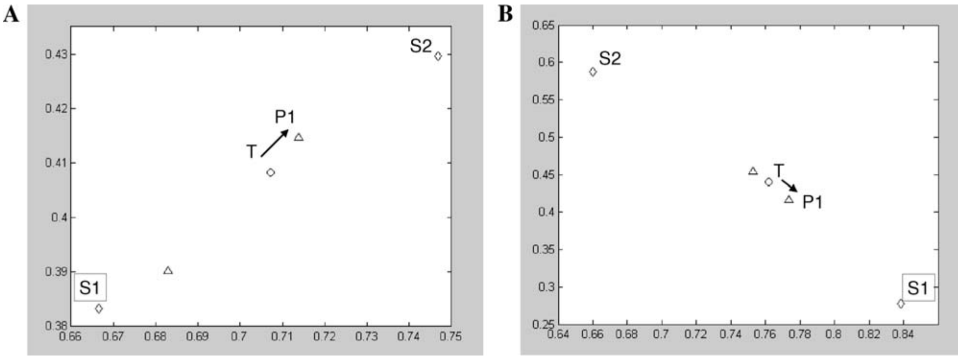

Fig. 4. Chromatic coordinates. (A) Chromatic contrast; (B) chromatic assimilation.

The most interesting conclusions derived from all these works and from a computer vision point of view can be summarised in the two following points:

1. Changes in colour appearance due to the spatial frequency of patterns can be described by a two-step pattern–colour separable model [16,17]:

   First step. A colour transformation to a new coordinate space that is independent of the image content. The best correspondence of the derived data is achieved by the opponent-colour transformation that occurs in the first steps of the visual pathways [23].

   Second step. In the new coordinate frame, colour representation is transformed by a gain factor that is dependent on the image content, and it is different in each colour channel.

2. The relationship between spatial frequency and the two types of colour induction can be summarised as follows [48,54]:

   Transition frequency. A spatial frequency of 4 cpd is a transition frequency from assimilation induction to contrast induction.

   Extreme frequencies. Spatial frequencies at 9 and 0.7 cpd assures assimilation and contrast induction, respectively, for any inductor.

Frequency measures are given in cpd units (cycles per degree) that represents the number of cycles for 1° of visual angle. Visual angle is a common way to express a spatial measure which can include the effect of viewing distance and the size of the stimulus. In Fig. 5, we can see coloured square-wave patterns at different spatial frequencies. These plots correspond to an image subtending 6° of visual angle when observed at a distance of 30 cm. [2] From 0.5 to 1 cpd we can perceive images with two-coloured bars, blue and yellow, their colour perception is stressed by the colour contrast phenomenon due to a low spatial frequency. When the frequency increases we tend not to perceive separate blobs but a global colour that is the result of the two basic colours plus the frequency effect.

[2] We have considered 30 cm as a usual distance to be located while reading a paper.

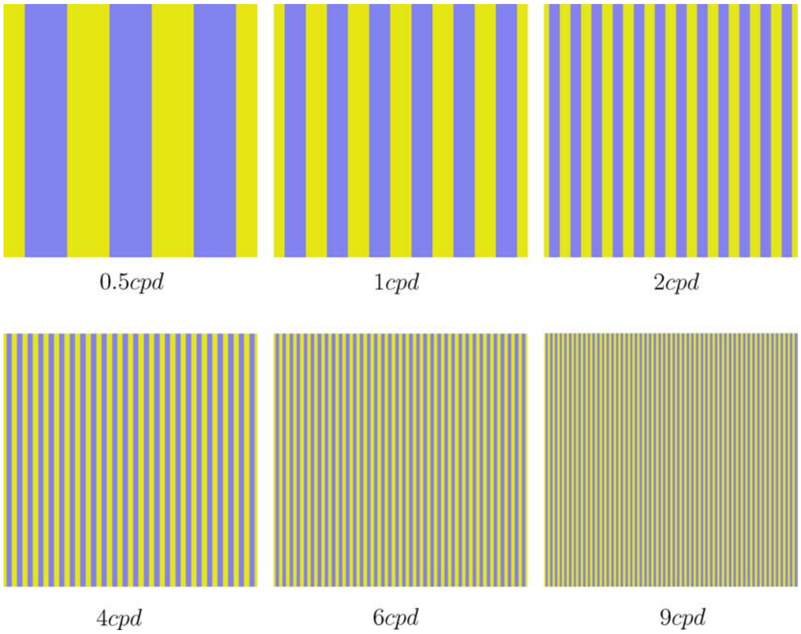

Fig. 5. Colour Induction at different spatial frequencies. Frequencies are computed by considering observer position at 30 cm from the image. Images are displayed on 6° of visual angle.

5. Colour assimilation as a perceptual blurring

A computational model of colour assimilation has already been given by an isotropic blurring of an image on the opponent-colour space, this has been proposed by Petrou et al. [2] as a perceptual blurring. Colour assimilation effects were first measured by Wandell et al. [1], where they take perceptual measurements of quality on printed patterns. To achieve this goal, they propose the Spatial-CIELAB space that is given by a two-steps process: an opponent colour channel transformation and a spatial blurring of the image channels. The blurring process convolves the image channels with kernels formed by weighted sum of exponential functions. In this way, they build a perceptual representation of a colour image presenting the assimilation effect of the HVS. This assimilation operator applied to an image, I in XYZ coordinates, is given by

$$\mathrm{Ass}(I)_{\vec{\omega}}^{\vec{\sigma}} = \mathrm{CIELAB}(\mathcal{B}(\mathrm{Opp}(I), \vec{\omega}, \vec{\sigma})) \tag{1}$$

that is finally transformed to the CIELAB space to deal with colour reproduction error measurement. The blurring operator is defined as

$$\mathcal{B}(I, \vec{\omega}, \vec{\sigma}) = (\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3) \text{ where } \mathcal{B}_c = I_c * f_c, \tag{2}$$

where I_c is the cth channel of a colour image I, and the function

$$f_c = m_c \sum_i \omega_i E_i \ : \ E_i = k_i \exp\left\{\frac{x^2 + y^2}{\sigma_i^2}\right\}, \tag{3}$$

where c represent each of the three opponent channels, m_c is a scale factor chosen to make the kernel f_c sum to one and k_i, is again a factor scale selected to make E_i sum to 1. The values of ω_i and σ_i have been determined from psychophysical measurements [1], ω_i are fixed values and σ_i are given in degrees of visual angles, thus, depending on the observer position we have

$$\sigma_{\text{pixels}} = d \cdot \frac{R}{S} \cdot \tan(\sigma_{\text{degrees}}) \tag{4}$$

where d is the observer distance to the texture, S is the length of one edge of the area where the image is displayed, and R is the number of pixels of image I along S, in Fig. 6, we can see an scheme of this unit conversion process, the size of the kernels is always built to cover the area of 1° of visual angle. An approximation of the profiles of these isotropic filters is shown in Fig. 7, where the filters have been built for an image of 550 pixels, displayed on a visual field of 20 cm per edge and observed from 40 cm.

Taking into account the previous conversion expressions we can redefine the assimilation operator as a function of the observer distance:

$$\mathrm{Ass}(I, d) = \mathrm{Ass}(I)^{\vec{\sigma}}_{\vec{\omega}}. \tag{5}$$

To illustrate how this transformation behaves on a given image we show in Fig. 8 the results of applying the Spatial-CIELAB transformation on two images (A) and (B), presenting an important colour assimilation effect. We can see on the profiles below, how the Spatial-CIELAB transformation makes the green-blue band to become bluish when it is surrounded by blue and to become reddish when surrounded by red. Then, images (C) and (D) are the images that presents the assimilation effect. It is important to note that when looking at these images an assimilation effect is also

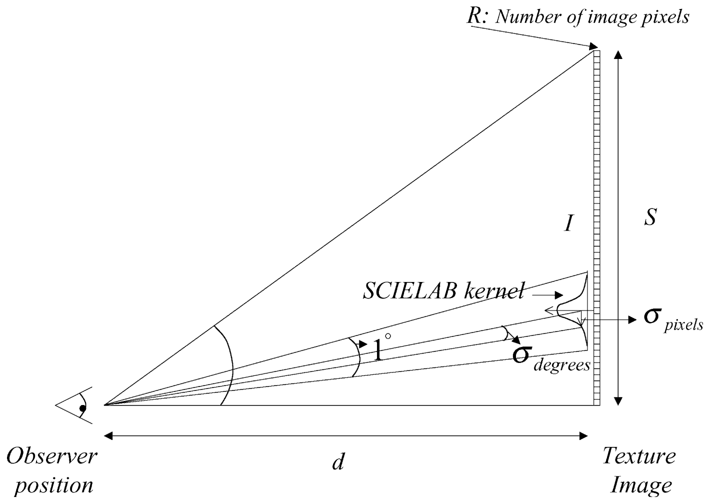

Fig. 6. General scheme for unit conversion from visual angle to image pixels.

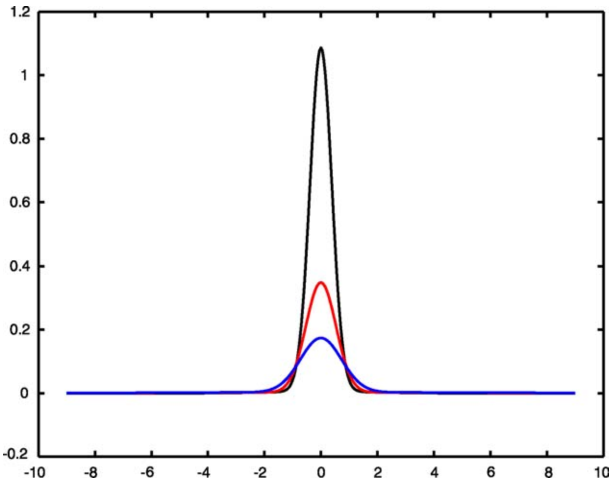

Fig. 7. Example of profiles of the 2D symmetric filters for the Spatial-CIELAB. Black, red, and blue colour lines represent the kernel for the intensity, red-green, and yellow-blue channel, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

present. To interpret these images we should point out that the isolated colours of the centre band are the colours that we perceive in the central bands of the original images.

Perceptually blurred images will be the basis for global measurements of colour–texture images, due to the fact that they represent the perceived colour appearance of textured images containing high spatial frequencies and viewed from a long distance.

6. Colour contrast as a perceptual sharpening

As we have already explained in Section 4, colour contrast is a complementary mechanism to colour assimilation. Colour contrast arises on regions of low spatial frequency and shifts the chromaticity of the stimulus in a direction away from the chromaticity of the surround. In this section we define a computational operator that simulates the colour contrast phenomenon.

This operator enhances differences in the transitions among colours of regions presenting low frequencies. While the assimilation effect has been solved by a blurring operator, it seems quite natural that the contrast effect will have to be solved by a sharpening operator. A detailed discussion of different colour sharpening operators can be found in [55].

To this purpose we define a sharpening operator, taking as a basis one of the most common sharpening operators [56], that is given by

$$S(I, \gamma, \sigma) = (S_R, S_G, S_B) \text{ where } S_c = I_c - \gamma LoG_\sigma(I_c), \tag{6}$$

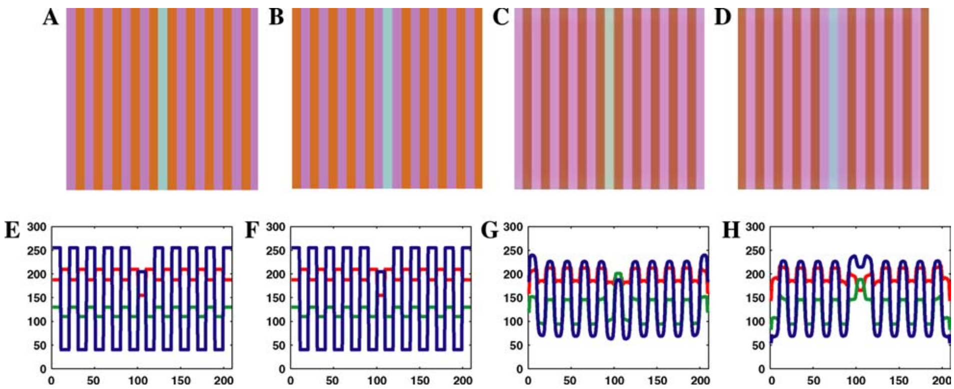

Fig. 8. (A,B) Examples of two images presenting important assimilation effects. (C,D) Previous images transformed by Spatial-CIELAB. (E–H) are the RGB profiles of images (A–D), respectively.

where I_c is the cth channel of a colour image I, usually given by an RGB representation, γ is a constant controlling the amount of enhancement and $LoG_\sigma(I_c)$ is the Laplacian of the image I_c blurred with a Gaussian, that is, convolved with the function

$$\nabla^2 G_\sigma = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right]\exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\}. \tag{7}$$

One of the problems with this operator is that colour artifacts might appear. A possible solution is to change to a new space where one of the channels is the intensity information and, then, to apply the sharpening operator to this channel and transform the obtained image to the RGB space. In this case, chromatic contrast is not involved, only brightness contrast is achieved. However, when the sharpening is applied to the RGB channels of the image it is a common practice to maintain the same parameters for all the channels. It is known that filtering the intensity channel by a certain blurring kernel has a more important effect than doing the same process on the chromatic channels [57].

Considering the conclusions on pattern–colour separability presented in Section 4 an inductor operator should act on an opponent-colour space. Then, an inductor operator that implements a chromatic contrast will be denoted by Con, and defined as

$$\mathrm{Con}(I)^{\vec{\sigma}}_{\vec{\gamma}} = \mathrm{RGB}(\mathcal{S}(\mathrm{Opp}(I), \vec{\gamma}, \vec{\sigma})), \tag{8}$$

where $\mathcal{S}$ is a sharpening operator that is applied to each channel of the image and which has parameters γ_i and σ_i where i corresponds to the ith channel of the image. These parameters act as independent gain factors for each channel.

Based on the classical sharpening defined by S, and taking into account the considerations about simultaneous contrast done in the work of Grossberg and Todorovic [58], we propose a sharpening that will allow to simulate chromatic contrast induction. This is a psychophysical work on brightness contrast based on on–off lateral geniculate cells, modelling responses in the boundary contour system by a sum of exponential functions that is nearly equivalent to the Laplacian of Gaussian. They describe a *Diffusive Filling-In* process to perform the brightness percept. It is done by a set of contiguous connected biological cells that pass signals between each other. The cell activity can spread to neighbouring cells, then to neighbours of the neighbours, and so on. This process is inhibited by *Boundary Contour Units*, that is, they decrease the diffusion constant by reducing the electrical activity between cell membranes.

Inspired on these ideas, this process can be computationally performed by an interpolation process of some kind of boundary detector signal. To clarify the process we will apply it to the 1D signal depicted in Fig. 9A. In the next paragraphs we will explain how to simulate this process with a computational procedure.

In order to simulate the described process on a I_c image, we firstly need to identify its inhibited and activated areas, whichever the colour channel is being analysed. When applied, for example, to the red–green dimension, the active areas will be the reddish ones and the inhibited areas the greenish ones. Computationally it is equivalent to the intensity of the stimulus. In the preceding example a red area is positive and a green area negative. However, following [58], the final response at a given pixel depends on the colours around it. There will be positive and negative responses when comparing against its surround. A yellow area is a negative area when its surround is red but positive when green.

Filtering the image with a Laplacian operator is well suited to such situations, since its response is positive in the transition between dark and light, and negative in the transition from light to dark. The result of this step on a 1D example is shown in Fig. 9B.

We will exploit the fact that the result of convolving I_c with a Laplacian operator represents its edge locations by zero crossings, i.e., a change between positive an negative response or vice versa. These are usually represented by the maximum and minimum of the border responses having different sign. Because they might not be adjacent, the zero-crossing detection must be done in a small neighbourhood (i.e., 3×3) around the actual zero-crossing. In this way we assure to obtain the zero-crossing in their maximum and minimum values of the $LoG(I_c)$.

Let us call $Z(I_c)$ the binary image having 1 at those points where there is a zero-crossing in the image I_c. Fig. 9C shows the location of the zero-crossings (i.e., $Z(I_c)$) in red dots.

For subsequent processes, here we need to add to $Z(I_c)$ the points on the image frame. This will allow to extent the next operator effect all over the image.

Once we have built the inhibition/activation areas of I_c, represented by $Z(I_c)$, now we can easily obtain the energy of this points by computing an image product $Z(I_c) \cdot LoG(I_c)$. In the example these are the red dots in Fig. 9D.

The following step is to build a surface where its value in a certain point will indicate the level of activation of this point and is the result of a spreading process on the inhibition/activation areas. It must have some properties:

1. The points on the boundaries must preserve its energy, i.e., the relationship between adjacent regions must be maintained.
2. The number of zero-crossings between points of the boundaries must be preserved, i.e., there will not be more regions than in the input energy image.

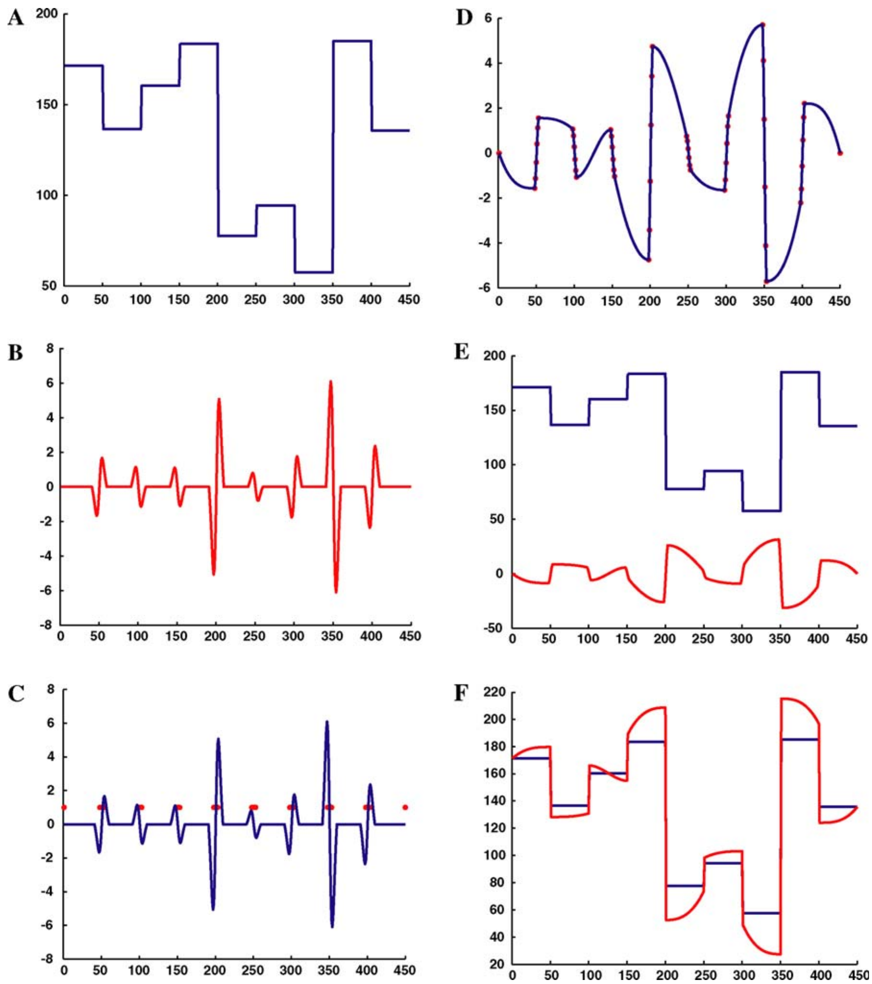

Fig. 9. Example of the process of perceptual sharpening operator on a 1D signal. (A) The original signal. (B) The Laplacian of Gaussian of (A). (C) Marks in red dots the position of the zero-crossing points of (B) represented in the $Z(I_c)$ image in text, these values are set to 1. (D) The result of interpolating the signal of the Laplacian of Gaussian on the previous points over the whole signal. It correspond to the $\mathcal{I}((\mathrm{C}),(\mathrm{B}))$ operator. (E) Plots, in blue line, the original signal and, in red line, the correction factor multiplied by the parameter γ i.e., $\gamma_c SLoG(I_c, \sigma_c)$ in Eq. (10). (F) Plots the final result $\mathcal{S}_c$ (in red) versus the original signal (in blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

Let us call $\mathcal{I}(X, Y)$ the operator that constructs this surface from the energies of a set of boundary points which are those having 1 value in image X and whose activation energy is given by its corresponding position in Y.

To build this surface an immediate solution is to use some kind of surface interpolation. For simplicity and because it copes with the above mentioned restrictions we will use linear interpolation. The problem of interpolation is that it needs to have equally spaced signal samples. This is not the case of the zero-crossings. To solve this point we use a Delaunay triangularisation as a previous step to achieve a uniform spaced set of points. For an intuitive and non-rigorous idea on how it works we can say that we need to generate the best triangles among the input data points, and to interpolate their values across the triangle plane.

Then, for a given channel I_c

$$SLoG(I_c, \sigma) = \mathcal{I}(Z(I_c), LoG(I_c, \sigma)) \tag{9}$$

is the spread of $LoG(I_c)$. In Fig. 9D we can see a profile in blue of how it behaves. Following the same schema as in the classical sharpening operator of Eq. (6) we define the operator $\mathcal{S}$ of a colour image, I, as

$$\mathcal{S}(I, \vec{\gamma}, \vec{\sigma}) = (\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3) \text{ where } \mathcal{S}_c = I_c - \gamma_c SLoG(I_c, \sigma_c) \tag{10}$$

and it runs the following steps:

Step 1. Transform the image to the opponent colour representation.

Step 2. For each channel c:

1. compute the Laplacian of Gaussian of sigma σ_c,
2. interpolate the responses at the edges of regions inside them,
3. subtract 2.2 from channel c of step 1 given a weight γ_c.

Step 3. Return to the original colour representation.

The parameters $\vec{\gamma}$ and $\vec{\sigma}$ play an important role for an inductor operator, we can roughly explain that $\vec{\sigma}$ parameter selects the boundaries or the scale of the regions that will be preserved by the chromatic contrast effect, and $\vec{\gamma}$ parameter fix the gain of the induction effect.

To fix their values to define a perceptual operator further psychophysical measurements are required from those given in [48]. However there are strong relationships between them. Values for $\vec{\gamma}$ depend on the spatial frequency of the image. The

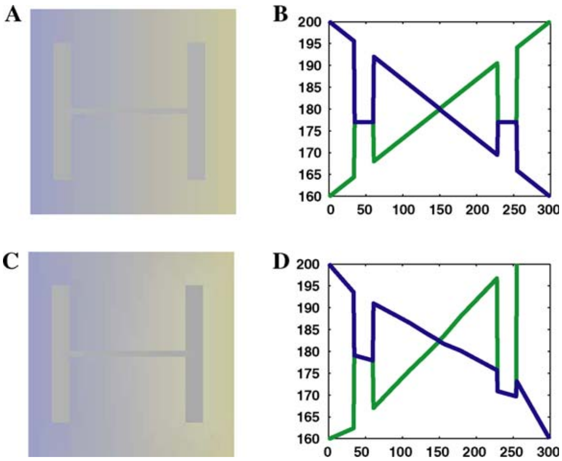

Fig. 10. (A) Original image, (C) perceptual sharpening of (A). (B,D) present a profile of an horizontal line of images (A,C), respectively.

spatial frequency depends on the observer distance to the image. The observer distance implies to select more or less image details that is provided by the $\vec{\sigma}$ parameter. Above considerations allow to redefine the contrast induction operator as a function of the observer distance, that is

$$\mathrm{Con}(I, d) = \mathrm{Con}(I)^{\vec{\sigma}}_{\vec{\gamma}}. \tag{11}$$

Fig. 10 shows the effects of the perceptual sharpening on a synthetic image and the corresponding profiles of an image line. On the other hand, Figs. 11 and 12 show the result of this sharpening on four natural images.

7. Colour–texture perception as a visual process

In Sections 2 and 3 we have explained, respectively, how colour and texture can be computationally represented as separate cues. However, we need to

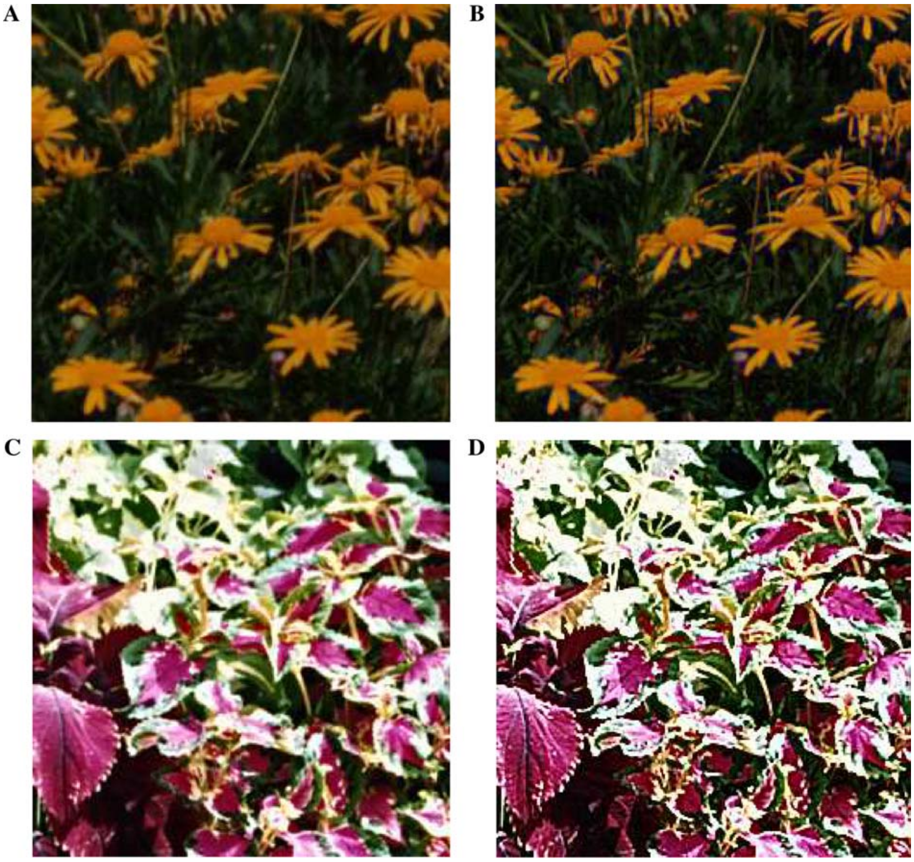

Fig. 11. (A,C) Original images (Vistex:`Flowers.0001`, Vistex:`Leaves.0005`). (B,D) Results of the perceptual sharpening on (A,C), respectively.

Fig. 12. (A,C) Original images (from a Canon Ixus digital camera, Vistex scene:`Corridor.0000`). (B,D) Results of the perceptual sharpening on (A,C), respectively.

work on images from real world, which is neither a grey world, nor a Mondrian world.

In Section 4 we have briefly introduced the colour induction phenomenon that explains how colour changes depend on spatial frequencies, and how opponency can represent these interactions. These perceptual considerations can be simulated with two types of computational operators, first we have explained in Section 5 the perceptual blurring which enables us to model the colour of a surface corresponding to a texture with high spatial frequency properties, second, we have explained in Section 6 a perceptual sharpening which enables us to model the colour of a surface corresponding to a texture with low spatial frequencies.

The spatial frequency is not an inherent property of a texture, it is the result of a vision process, that can be achieved by a successive change on the observing conditions. It can be achieved by changing from a peripheral to foveal vision, or simply by moving the observer position. Consequently, a perceptual representation of a colour–texture will have to be represented by the tower of images that simulates this vision process. This approach has already been proposed by Petrou et al. [2] as perceptual tower.

In this work we propose to complete this perceptual tower by adding the colour representation of textures when spatial frequencies are lower than 4 cpd, which is the threshold measured by Smith et al in [48] as the threshold frequency from colour assimilation to colour contrast in normal colour observers.

8. A computational colour–texture model

Considering the assumptions done in the previous section, we now define the complete perceptual tower which we intend as a general perceptual representation of coloured textures. Therefore, for a given image we can build a perceptual tower representing colour information as it was observed from different observer distances. It can be considered as a colour–texture front-end representing colour–texture interaction and the basic step for further processing, a parallel approach to a scale-space representation for grey-level images [59].

For a given texture image, I, we can estimate a predominant spatial frequency from its Fourier spectrum, that is denoted as v_{cpp}, and can be obtained by computing

$$v_{\mathrm{cpp}} = v \; : \; \mathrm{BPH}(v) = \max(\mathrm{BPH}(\mathrm{FS}(I))), \tag{12}$$

where FS is the Fourier spectrum of an image, and BPH is its frequency energy histogram. Frequency is given in cycles per pixel (cpp) units, and then we can get this spatial frequency in cpd, v_{cpd}, by using this conversion expression:

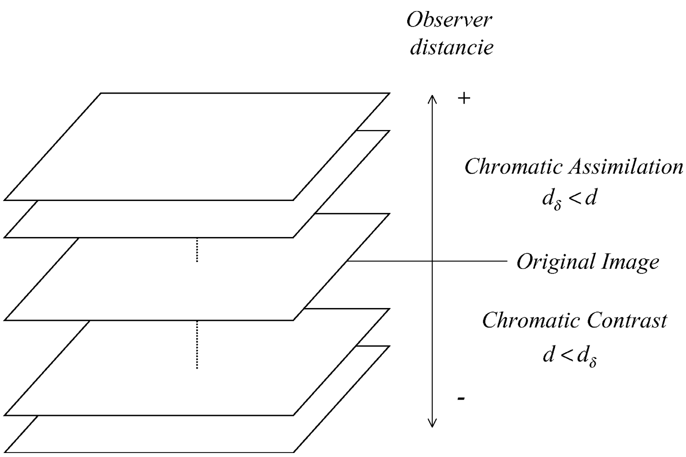

Fig. 13. General scheme of the colour–texture front-end.

$$v_{\mathrm{cpd}} = v_{\mathrm{cpp}} \cdot \frac{R}{a\tan(\frac{S}{d})} \tag{13}$$

where d is the observer distance to the texture, S is the length of one edge of the area where the image is displayed, and R is the number of pixels of image I along S. By considering this expression we can build a perceptual tower by computing:

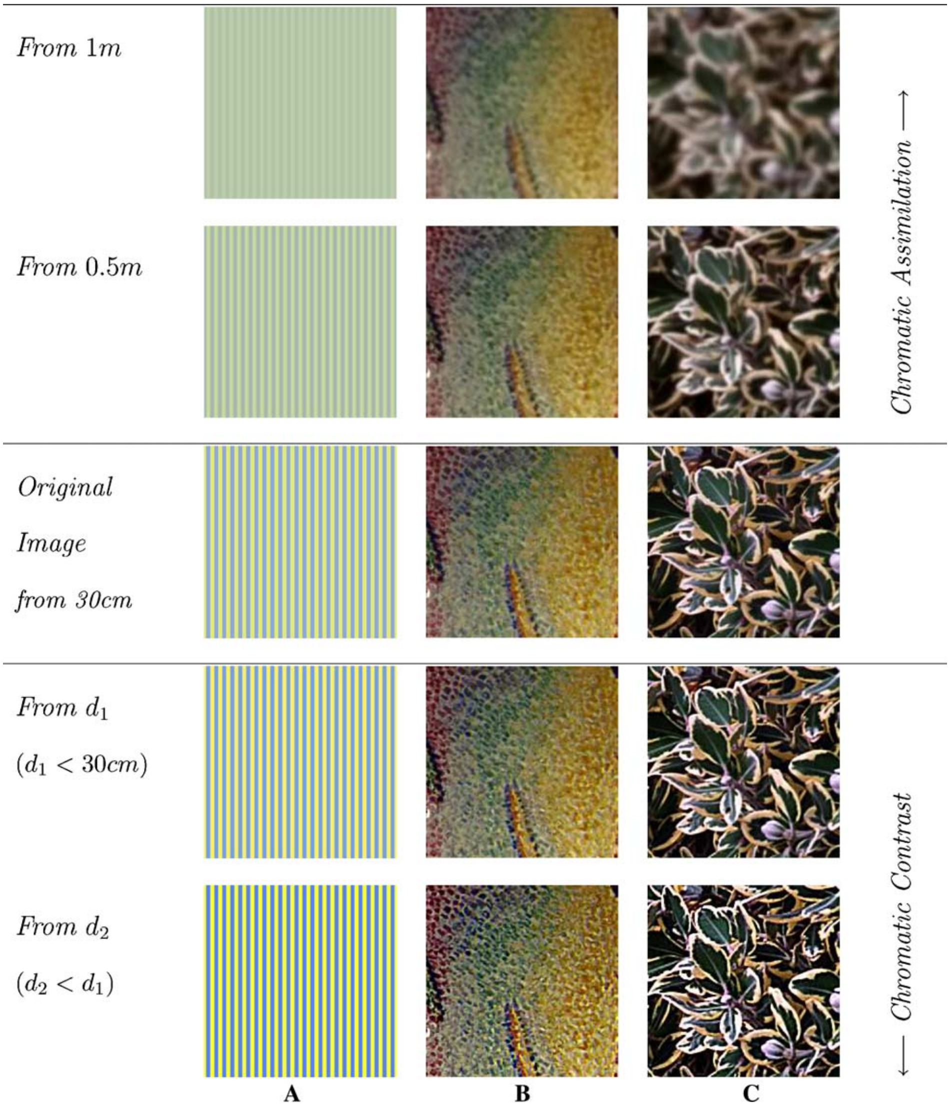

Fig. 14. Examples of perceptual towers. Columns present the perceptual towers for three different images. (A) Synthetic image, the original image present a 4 cpd frequency when observed at 30 cm. (B) Image from VisTex dataset presenting spatial frequency similar to (A). (C) Image from VisTex dataset presenting higher spatial frequency than (A,B).

$$T(I)_{d_i} = \begin{cases} \mathrm{Con}(I, d_i) & \text{for } d_i \in (0, d_\delta), \\ \mathrm{Ass}(I, d_i) & \text{for } d_i \in (d_\delta, +\infty), \end{cases} \tag{14}$$

where d_δ is the observer distance that makes $v_{\mathrm{cpd}} = 4\,\mathrm{cpd}$, that is the threshold frequency that changes from chromatic contrast to assimilation, as has been measured in [48], it is given by

$$d_\delta = \frac{S}{\tan(R \cdot \frac{v_{\mathrm{cpp}}}{4})}. \tag{15}$$

The perceptual tower will provide the colour appearance according to the spatial frequency of the image content and related to the observer position. Fig. 13 shows a general scheme of the proposed colour–texture front-end that acts as a scale-space representation of a colour image, but representing the interaction between colour and spatial frequencies in a perceptual sense.

In Fig. 14 we can see the perceptual tower for three different images. For images in which assimilation occurs we can consider the representation to be a true perceptual representation since it is based on psychophysical studies of assimilation. However, no similar studies of chromatic contrast have been conducted and so for images in which contrast occurs, the representation is not truly perceptual. We can see how the assimilation process blur the images when distance increase and how colour regions are enhanced when distance decrease.

9. Discussion

In this paper we outlined a framework to deal with the interaction between texture and colour in a perceptual sense. Our contribution extends previous working [2] by incorporating into the perceptual tower an inductor operator, perceptual sharpening, which models the effect of chromatic contrast (a complementary effect to assimilation which is modelled by perceptual blurring).

Combining these two operators we can build a global framework that acts as a primal sketch for colour representation. This perceptual tower can provide a computational representation of colour appearance. It requires the psychophysical measurements of the chromatic contrast effects.

Both operators present an essential property, due to they are applied on an opponent-colour space, it allows to act more selectively on the brightness or chromatic information simulating the HVS pathways. This cannot be achieved when a multiscale filtering is applied on the RGB channels of images.

On the other hand, the way in which the perceptual sharpening is defined implies that it does not break structural properties of the image blobs. This will be an essential property for texture analysis, due to the need of finding global density of the blob attributes, as it is stated by the Julesz texton theory presented in Section 3. In this sense, perceptual sharpening provides a real contribution in the blob segmentation step, avoiding usual refinements to the zero-crossings of the

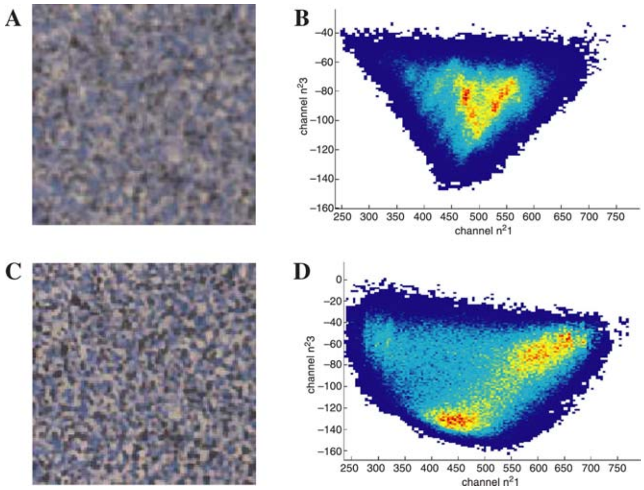

Fig. 15. (A) Original image; (B) projected histogram of image (A); (C) perceptual sharpening of image (A); and (D) projected histogram of image (C).

Laplacian [60]. In Fig. 15 we can see how the perceptual sharpening changes the colour histogram providing a clear definition of colour clusters that will contribute to get a better segmentation of blobs as the first step to compute their attributes. An objective validation of this property is given in [55].

As we have introduced in Section 3, perception of textures is based on two complementary processes, first, by computing a global density of blob attributes and second, by computing attributes of patterns emerging globally from the texture. Because the first step is based on results of a blob segmentation, the images of the perceptual tower corresponding to short-distance appearance will be the basis, that is, those images for which the perceptual sharpening has been applied can provide an excellent representation to get blob properties. On the other hand, those images presenting the assimilation effect can provide a global view or a long-distance view where the emergent patterns can be better extracted.

The main advantage of this approach for computer vision applications is given by the fact that for any colour image we can build different appearance images representing different observations of the original one. The usefulness of this colour primal sketch is straightforward for subsequent visual tasks like segmentation. In more high level visual tasks, such as, content based retrieval the perceptual tower can add some interesting refinement or subtlety to queries. For instance, the image (A) of Fig. 12, could be retrieved from two different queries in two different contexts. Thus, we could ask for someone wearing a greenish T-shirt, or we could ask for someone wearing a stripped yellow and blue T-shirt. Namely, we can add some nuances to any query, we can ask for a global view of the image or we can ask for a detailed view of the image, depending on the accuracy we require to the retrieval question.

10. Future work

The first step to provide a full perceptual approach will require a psychophysical measurement of chromatic contrast, this will imply a tedious task of psychophysical experimentation. Once it is done, we will have to match the measurements with the sharpening parameters.

In the work presented we have taken all the attention on getting induction operators presenting good properties to represent colour appearance, however we have not taken attention to the computational cost of these operators. In this sense the sharpening operator algorithm should be improved in future versions since the interpolation step is computationally expensive.

Finally, we have generally assumed a predominant spatial frequency in texture images, however it is not always true, and therefore it can imply some changes on computing the distance parameters of the perceptual towers.

Acknowledgments

This work has been funded by project TIC2000-0382 of the MCYT of the Spanish Goverment.

References

[1] X. Zhang, B.A. Wandell, A spatial extension of CIELAB for digital color image reproduction, in: SID, 1996.
[2] M. Petrou, M. Mirmehdi, M. Coors, Perceptual smoothing and segmentation of colour textures, in: 5th European Conference on Computer Vision, Freiburg, Germany, 1998, pp. 623–639.
[3] B.A. Wandell, The synthesis and analysis of color images, IEEE Trans. Pattern Anal. Machine Intell. PAMI-9 (1) (1987) 2–13.
[4] T.R. Reed, J.H.D. Buf, A review of recent texture segmentation and feature extraction techniques, CVGIP: Image Understanding 57 (3) (1993) 359–372.
[5] C. Carson, S. Belongie, H. Greenspan, J. Malik, Region-based image querying, in: CPR Workshop on Content-Based Access of Image and Video Libraries, 1997.
[6] S. Tan, J. Kittler, Colour texture classification using features from color histogram, in: Proc. 8th Scandinavian Conf. on Image Processing, 1993.
[7] D. Panjwani, G. Healey, Markov random field models for unsupervised segmentation of textured color images, IEEE Trans. Pattern Anal. Mach. Intell. 17 (10) (1995) 939–954.
[8] G. Healey, L. Wang, Illumination-invariant recognition of texture in color images, J. Opt. Soc. Am. 12 (9) (1995) 1877–1883.
[9] F. Lumbreras, J. Serrat, R. Baldrich, M. Vanrell, J.J. Villanueva, Colour texture recognition through multiresolution features, in: Quality Control by Artificial Vision, 2001, pp. 114–121.
[10] A. Gagalowicz, S.D. Ma, C. Tournier-Lasserve, Efficient models for color textures, in: 8th ICPR, 1986, pp. 412–414.
[11] T. Caelli, D. Reye, On the classification of image regions by colour, texture and shape, Pattern Recognition 26 (4) (1993) 461–470.
[12] B.V. Funt, G.D. Finlayson, Color constant color indexing, IEEE Trans. Pattern Anal. Mach. Intell. 17 (5) (1995) 522–529.
[13] J. Weickert, Coherence-enhancing diffusion of colour images, Image Vision Comput. 17 (1999) 201–212.

[14] M. Mirmehdi, M. Petrou, Segmentation of color textures, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2) (2000) 142–159.
[15] K.-H. Baeuml, B. Wandell, Color appearance of mixture gratings, Vision Res. 36 (18) (1996) 2849–2864.
[16] A. Poirson, B. Wandell, The appearance of colored patterns: pattern–color separability, J. Opt. Soc. Am. 10 (12) (1993) 2458–2470.
[17] B.A. Wandell, Color appearance: the effects of illumination and spatial pattern, Tech. rep., Psychology Department, Standford University, November 8, 1995.
[18] X. Zhang, J. Farrell, B.A. Wandell, Applications of a spatial extension to CIELAB, in: SPIE, 1997.
[19] D. Slater, G. Healey, The illumination-invariant recognition of 3D objects using local color invariants, IEEE Trans. Image Processing 18 (2) (1996) 206–210.
[20] M. Swain, D. Ballard, Color indexing, Internat. J. Comput. Vision 7 (1) (1991) 11–32.
[21] J. Crowley, F. Berard, Multi-modal tracking of faces for video communications, in: IEEE Proc. CVPR, 1997.
[22] M.D.B.V. Funt, M. Brockington, Recovering shading from color images, in: Proc. ECCV, 1992, pp. 124–132.
[23] L. Hurvich, D. Jameson, An opponent-process theory of color vision, Psychol. Rev. 64 (6) (1957) 384–404.
[24] G. Wyszecki, W. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, second ed., John Wiley and Sons, 1982.
[25] T. Gevers, A.W.M. Smeulders, Color based object recognition, Pattern Recognition 32 (1999) 453–464.
[26] K. Dana, S. Nayar, B. Ginneken, K. J, Reflectance and texture of real-world surfaces, in: Proc. CVPR97, 1997, pp. 151–157.
[27] B. Ginneken, J. Koenderink, Texture histograms as a function of irradiation and viewing direction, Internat. J. Comput. Vision 31 (2/3) (1999) 169–184.
[28] A. Witkin, Recovering surface shape and orientation from texture, Artificial Intell. 17 (1981) 17–45.
[29] J. Garding, Shape from texture and contour by weak isotropy, Artificial Intell. 64 (1993) 243–297.
[30] L. Gool, P. Dewaele, A. Oosterlinck, Survey: Texture analysis anno 1983, Computer Vision Graphics Image Process. 29 (1985) 336–357.
[31] H. Wechsler, Texture analysis—a survey, Signal Process. 2 (1980) 271–282.
[32] M. Tuceryan, A. Jain, Handbook of Pattern Recognition and Computer Vision, World Scientific, 1993, Ch. Texture Analysis, pp. 235–276.
[33] A. Rao, G.L. Lohse, Towards a texture naming system: identifying relevant dimensions of texture, Vision Res. 36 (1996) 1649–1669.
[34] A. Rao, G.L. Lohse, Identifying high level features of texture perception, CVGIP: Graphical Models Image Process. 55 (1993) 218–233.
[35] H. Tamura, S. Mori, T. Yamawaki, Textural features corresponding to visual perception, IEEE Trans. System Man Cybern. SMC-8 (1978) 460–473.
[36] B. Julesz, J. Bergen, Textons, the fundamental elements in preattentive vision and perception of textures, Bell Syst. Technol. J. 62 (1983) 1619–1645.
[37] J. Beck, A. Sutter, R. Ivry, Spatial frequency channels and perceptual grouping in texture segregation, Computer Vision Graphics Image Process. 37 (1987) 299–325.
[38] D. Hubel, T. wiesel, Receptive fields and functional architecture of monkey striate cortex, J. Physiol. 195 (1968) 215–243.
[39] L. Harvey, M. Gervais, Internal representation of visual texture as the basis for the judgment of similarity, J. Exp. Psychol.: Human Perception Performance 7 (4) (1981) 741–753.
[40] L. Harvey, M. Gervais, Visual texture perception and Fourier analysis, Perception Psychophys. 24 (6) (1978) 534–542.
[41] L.T. Maloney, B.A. Wandell, Color constancy: a method for recovering surface spectral relfectance, J. Opt. Soc. Am. 3 (1) (1986) 29–33.
[42] D. Forsyth, A novel algorithm for color constancy, Internat. J. Comput. Vision 5 (1) (1990) 5–36.
[43] G. Finlayson, Color in perspective, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-18 (10) (1996) 1034–1038.

[44] J. McCann, Assimilation and contrast, in: IS$T-SID 9th CIC, 2002, pp. 91–96.
[45] S. Shevell, J. Wei, Chromatic induction: border contrast or adaptation to surrounding light?, Vision Res. 38 (1998) 1561–1566.
[46] C.S. Barnes, J. Wei, S. Shevell, Chromatic induction with remote chromatic contrast varied in magnitude, spatial frequency, and chromaticity, Vision Res. 39 (1999) 3561–3574.
[47] S.K. Shevell, J. Wei, A central mechanism of chromatic contrast, Vision Res. 40 (2000) 3173–3180.
[48] V.C. Smith, P.Q. Jin, J. Pokorny, The role of spatial frequency in color induction, Vision Res. 41 (2001) 1007–1021.
[49] G. Wyszecki, Handbook of Perception and Human Performance, Wiley, 1986, Ch. Color Appearance.
[50] E. Miyahara, V. Smith, J. Pokorny, The consequences of opponent rectification: the effect of surround size and luminance on color appearance, Vision Res. 41 (2001) 859–871.
[51] V. Smith, W. Jin, J. Pokorny, Color appearance: neutral surrounds and spatial contrast, Vision Res. 38 (1998) 3265–3269.
[52] A. Poirson, B. Wandell, Pattern–color separable pathways predict sensitivity to simple colored patterns, Vision Res. 36 (4) (1996) 515–526.
[53] E. Chichilnisky, B. Wandell, Photoreceptor sensitivity changes explain color appearance shifts induced by large uniform backgrounds in dichoptic matching, Vision Res. 35 (1995) 239–254.
[54] C. Fach, L. Sharpe, Assimilative hue shifts in color gratings depend on bar width, Perception Psychophys. 40 (1986) 412–418.
[55] R. Baldrich, Perceptual approach to a computational colour texture representation for surface inspection, Ph.D. thesis, Computer Vision Center, Dept. d'Informàtica, Universitat Autònoma de Barcelona, 2001.
[56] R. Gonzalez, P. Wintz, Digital Image Processing, Addison-Wesley, Reading, MA, 1987.
[57] B. Wandell, Foundations of Vision, Sinauer Associates, Inc., ISBN 0-87893-853-2, 1995.
[58] S. Grossberg, D. Todorovic, Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena, Perception Psychophys. 43 (1988) 241–277.
[59] T. Lindeberg, J.-O. Eklundh, On the computation of a scale-space primal sketch, J. Visual Commun. Image Representation 2 (1) (1991) 55–78.
[60] H. Voorhees, Finding texture boundaries in images, Tech. Rep. 968, MIT Artificial Intelligence Laboratory, 1987.