Chapter 4

Computational operators for colour texture perception

Most of the previous works dealing with computational representations for colour texture have been directed to extend gray level representations to every one of the RGB channels. As we have already seen in the introduction chapter, to deal with colour texture we need operators that combine co–jointly the spatial and the colour information in a way that simulates the especial behaviour demonstrated by the human visual system. In this chapter we will analyse colour induction as the most important phenomena that acts on the colour texture perception, and we will propose a computational operator for a perceptual sharpening that allows to complement previous results on perceptual blurring, providing both a general model for colour induction, the first one in chromatic contrast and the last one in chromatic assimilation.

4.1 Colour Induction

Colour induction is a colour phenomena that changes the colour appearance of a stimulus due to the influence of the scene contents in the field of view. In this category we have to include the colour adaptation phenomena introduced in chapter 3, which is always involved in any scene interpretation. Adaptation models or colour constancy methods usually are global visual mechanisms.

In this section we will deal with other induction phenomena that depend on the surrounding colour of a certain stimulus. The surrounding colour is called the inducing stimuli or inductor [116]. Depending on the direction of the chromatic change provoked by the inductor, we will distinguish two types of colour induction:

Chromatic Assimilation occurs when the chromaticity of the test stimulus changes towards the chromaticity of the inducing stimulus. An example of assimilation phenomena is shown in figure 4.1.(a).

Chromatic 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 figure 4.1.(b).

In figure 4.1.(c) and (e), we can see a plot of the chromaticity coordinates of the stimuli presented in images (a) and (b). We denote the test stimulus as TS, that is, the image region that is affected by an inducing surround. These inductors are denoted as S1 and S2.

In the first column of figure 4.1 we see the effects of the assimilation, the test stimulus moves its appearance towards the appearance of its own surround. The TS is yellow, and it appears pink when surrounded by S1, that is red, i.e. yellow moves toward red and becomes pinkish. The same TS becomes greenish when it is surrounded by S2 that is green.

In the second column of figure 4.1, we see the effects of the colour contrast, the test stimulus moves its appearance away from the appearance of its own surround. The TS is grey and it appears yellowish when it is surrounded by the S1 bluish surround. Complementary, the same TS appears bluish when it is surrounded by a S2 yellowish surround. In this case, the induction phenomena is behaving inversely as it behaves in assimilation. Chromaticities of the perceived stimuli are going far from the surround chromaticity. This phenomena is called simultaneous contrast when it is given on achromatic images. A typical example of simultaneous contrast or brightness contrast is shown in figure 4.2, where the same stimulus seems darker when surround is lighter and lighter when the surround is darker.

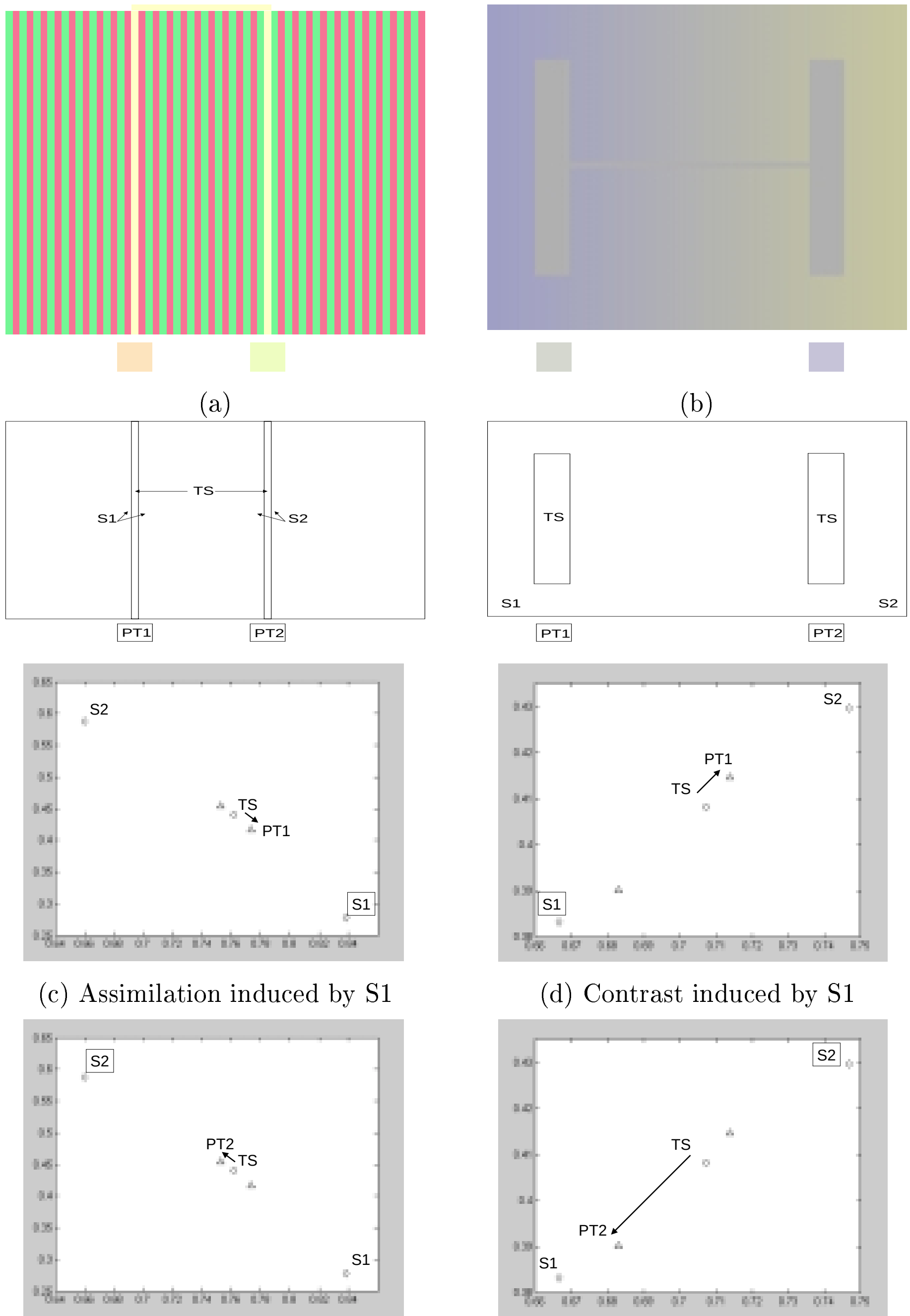
Considering the given definitions and examples, it is obvious that any perceptual approach towards a colour texture representation should take into account the colour induction effects we have introduced above.

In psychophysics we find a wide range of works dealing with the induction phenomena or the influence of surrounding chromaticities on the appearance of colour [98, 81, 96, 97, 95, 24, 94, 2, 86, 109, 20, 85, 112, 110]. In all these works, 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 pursue to give answers about how this perceptual mechanisms are organised in the human visual system. They help in building a more precise model on how human visual system acts from the retinal representation of colour to the final judgements on 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.

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

1. Changes on colour appearance due to the spatial frequency of patterns can be described by a two-step pattern-colour separable model [85, 109]:

    - 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 given by the opponent-colour transformation.



(a) (b)

(c) Assimilation induced by S1 (d) Contrast induced by S1

(e) Assimilation induced by S2 (f) Contrast induced by S2

Figure 4.1: Colour Induction. (a) Colour Assimilation. (b) Colour Contrast. (c), (d), (e) and (f) plot chromaticity coordinates of the RGB values of the images (a) and (b) denoted as given in the below graphics. (c) and (e) Chromaticity moves towards the inducing surround.(d) and (f) Chromaticity moves away from the inducing surround.

Figure 4.2: Simultaneous Contrast

  - Second step, in the previous coordinate frame, colour representation is transformed by a gain factor that is dependent of the image content.

2. The relationship between spatial frequency and the two types of colour induction can be summed up as follows [98, 29]:
    - A spatial frequency of 4 cpd. is a transition frequency between assimilation induction to contrast induction.
    - Spatial frequencies at 9 cpd. 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 degree of visual angle. The visual angle is a common way to express a spatial measure that allows to adjust the observer distance and the displayed window size to different possibilities. In figure 4.3, we can see coloured square-wave patterns at different spatial frequencies. These plots are given on image size corresponding to the diameter of 6 degrees of visual angle when observed at 30cm. From 0.5 cpd tp 2 cpd we can perceive images with two coloured types of blobs, blue and yellow. As 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.

Considering the above conclusions, we can derive a computational model for colour texture image representation based on the pattern-colour separable model shown at figure 4.4. Where Opp represents the opponent-colour transformation and A_ϕ and C_γ are respectively the assimilation and contrast operators, that represent the induction effects on each colour channel and for the corresponding range of spatial frequencies in the image. We have also indicated the possibility to insert other special phenomena that has been referred in the bibliography. A combination step of the resulting signals is represented by a P transformation.

This model has to allow to derive colour texture properties from the set of perceptually defined images. While the perceptual blurring has to allow defining global colour properties, the perceptual sharpening has to allow a better segmentation of different coloured blobs and the computation of their attributes.

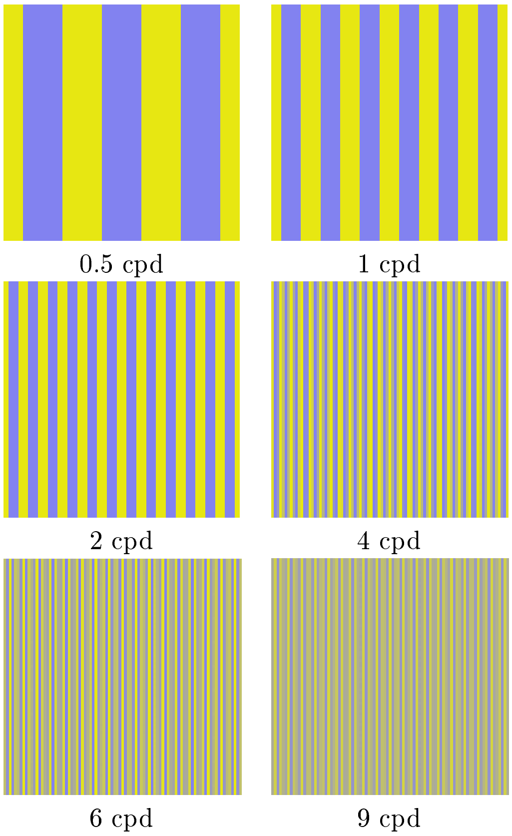


Figure 4.3: Colour Induction at different spatial frequencies. Frequencies are computed by considering observer position at 30cm from the image. Images are displayed on 6 degrees of visual angle.

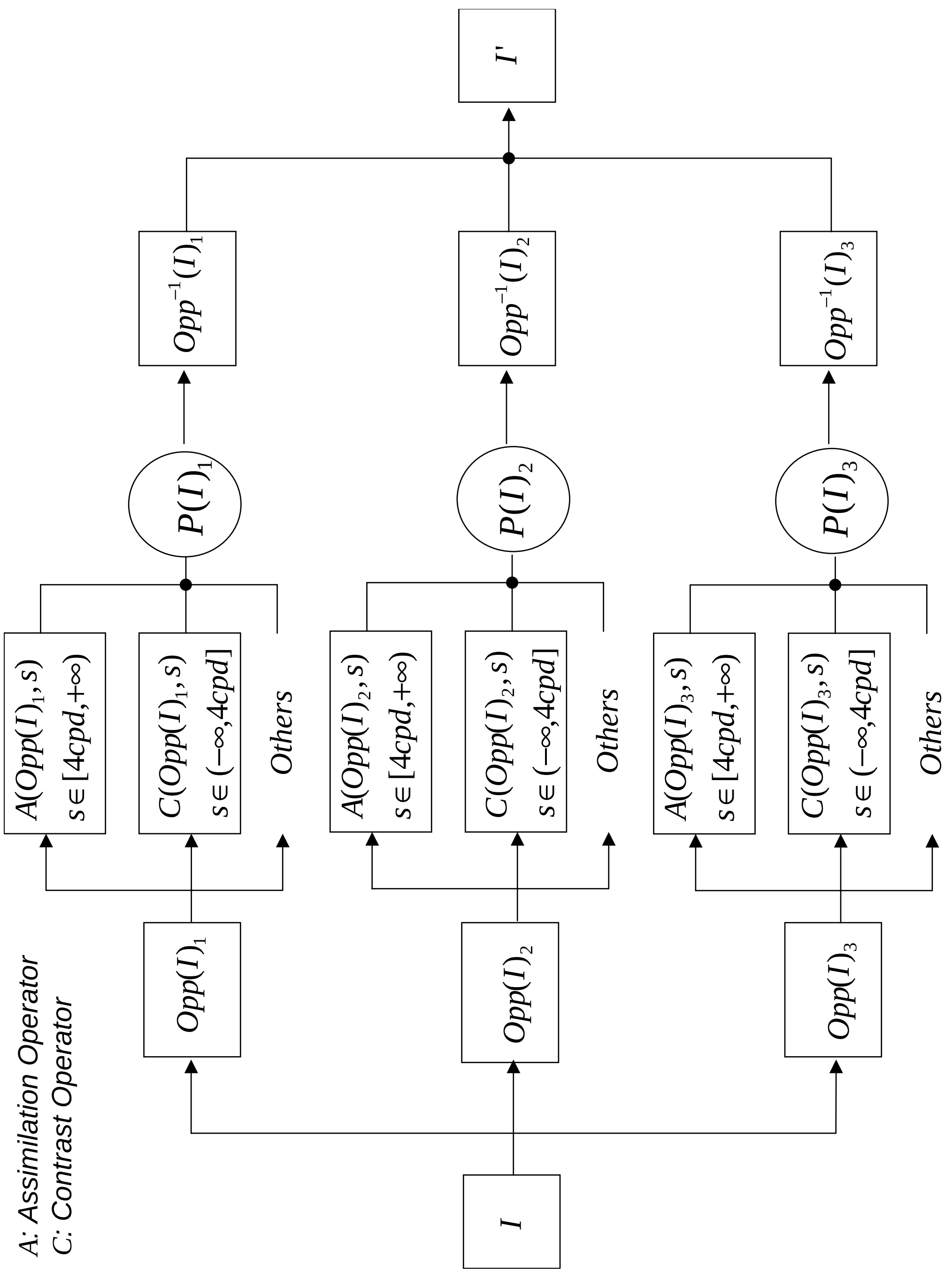


Figure 4.4: A pattern-colour separable model for colour induction.

In the next sections we will go deeply on how to define computational operators implementing the induction operators, but before to do it, we will introduce the opponent-colour space.

4.1.1 Opponent-Colour Space

The concept of opponent colours was first described by Hering in 1878, he made some interesting observations about some pairs of colours one never sees together at the same place and at the same time. While we are able to see a reddish or a yellowish orange, and a bluish or a greenish cyan, we never can observe a greenish red or a bluish yellow neither the opposite. These two hue pairs, red-green and blue-yellow are called opponent colours.

From this observation Hering hypothesised the existence of a unique visual pathway to encode red and green, and a unique visual pathway to encode blue an yellow. The same hypothesis was done for a visual pathway encoding achromatic black and white signals. It takes to formulate a neural representation of colours.

This opponent process model was left behind while the trichromatic theory of colour was stabilising the basis of the modern colorimetry based on the colour-matching experiments and all the derived standard spaces. It was resurrected when a hue-cancellation method was defined by Hurvich and Jameson [58] to quantitatively measure colour-opponency.

Due to the efforts of Hurvich and Jameson with the hue-cancellation experiment and plus the quantitative data provided by direct neurophysiological responses obtained from some measurements in the retinal neurons of a fish and in the lateral geniculate nucleus of non human primates, the opponent processing has been no longer questioned.

From the Hurvich and Jameson measurements a general opponent model schema can been derived, we show a computational approach of it in figure 4.5. There are some variations of the transform to this space from a trichromatic Young-Helmholz space, all of them follow the same schema of colour incompatibility the difference lies in the coefficients, α_i, β_j and γ_k, that combine the input signals.

In computer vision we usually only have a colour image representation in a RGB space of an unknown camera and under undefined conditions. Among others, a common representation of the opponent colour space is the one used in [100] that is defined as:

$$Opp(p) = p \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}, \tag{4.1}$$

$$RGB(p) = Opp^{-1}(p) = p \cdot \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & -1/2 & 0 \\ 1/6 & 1/6 & -1/3 \end{pmatrix} \tag{4.2}$$

where p is a 3D-vector of the RGB coordinates of the given space.

To be able to better establish the parameters for the spatial operators we will define in the next sections, we will use an orthonormal basis, given by:

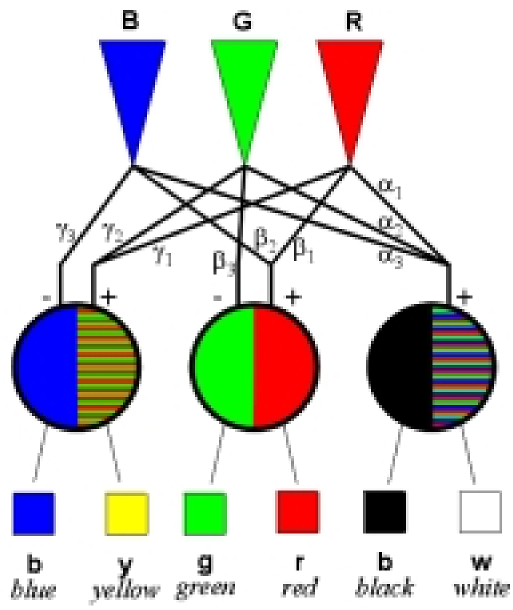


Figure 4.5: An Opponent colour vision model for a computational approach.

$$Opp(p) = p \cdot \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix}, \tag{4.3}$$

$$RGB(p) = Opp^{-1}(p) = p \cdot \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \end{pmatrix} \tag{4.4}$$

In both cases the first dimension represents the intensity or dark-white channel, the second dimension represents the red-green chromaticity channel and the third dimension represents the yellow-blue chromaticity channel.

4.2 Colour Assimilation as a perceptual blurring

As has been previously introduced, colour assimilation is the perceptual mechanism that takes chromaticities of regions with very high frequencies towards the chromaticities of the neighbouring regions. This effect is the result of a spatial blurring, that is usually implemented in computer vision with the convolution of the image with a gaussian spatial filter [75, 68]. However, in this case the spatial filter will not be applied on the RGB space as it is usually done, it will be applied to the opponent-colour space.

The idea of building a perceptual tower for a multiscale representation simulating different views of the same scene from different observer distances and considering the human colour perception has been taken to a computer vision model for the first time in [11, 79, 84, 80]. All these works have been based on pshychophysical measurements of colour appearance on human subjects given by the Spatial–CIELAB space defined in [119], this measurements have given the parameters of the spatial filters needed to simulate human assimilation on a CIELAB colour space. We will go deeply on this space in the next section.

In order to correctly apply Spatial–CIELAB blurring in images, the sensor blurring should be removed and be substituted by the perceptual one. In the thesis of Boukoubalas [12] there is an interesting explanation on how to do this.

4.2.1 S–CIELAB: Spatial CIELAB

S–CIELAB is a spatial extension to the CIELAB[1] colour metric that is used for measuring the quality of colour reproduction in digital images. It has been defined to improve the error computation on non-uniform spatial regions.

The Spatial–CIELAB representation is based in the two-step model defined by Wandell et al in [85, 109]. Firstly, a step to an opponent-colour space from the CIELAB representation is done, and secondly a convolution with a kernel whose shape has been psychophysically determined for each colour dimension. Finally, the filtered channels are transformed again to the standard CIELAB [117], that is actually representing the Spatial–CIELAB.

In this case, the opponent representation is built from the standard XYZ colour space, and is given by:

$$Opp(p) = p \cdot \begin{pmatrix} 0.279 & -0.449 & 0.086 \\ 0.72 & 0.29 & -0.59 \\ -0.107 & -0.077 & 0.501 \end{pmatrix} \tag{4.5}$$

where p is given by (X, Y, Z) following the standard CIE 1931.

The spatial filters for each opponent channel are built as a sum of gaussian functions, that is:

$$f_k = m_i \sum_i \omega_i E_i \tag{4.6}$$

where k represent every one of the three opponent channels, m_i is a scale factor chosen to make that the kernel sums to one, and

$$E_i = k_i \exp^{\frac{x^2+y^2}{\sigma_i^2}} \tag{4.7}$$

again the k_i factor scale is selected to make that E_i sums to 1. The measured values to substitute the parameters ω_i and σ_i are given for each opponent channel in table 4.1. Where the spreads are given in degrees of visual angle. Depending on

[1]The CIELAB space is an important international standard for colour measurement. The main property of CIELAB space is the uniformity with respect to human colour judgements.

Opponent Channel	Weights ω_i	Spreads σ_i
1 (I)	0.921	0.0283
	0.105	0.133
	-0.108	4.336
2 (R-G)	0.531	0.0392
	0.330	0.494
3 (Y-B)	0.488	0.0536
	0.371	0.386

Table 4.1: Parameters of the Spatial–CIELAB spatial kernels.

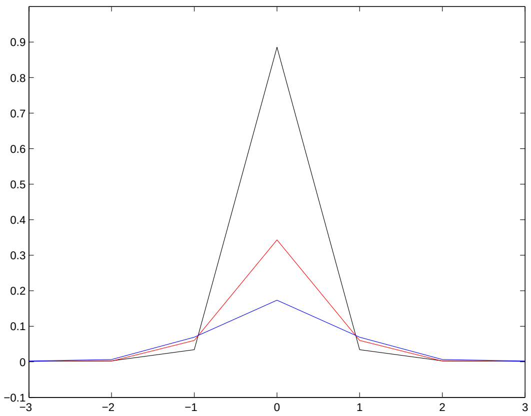


Figure 4.6: Profiles of the two-dimensional symmetric kernels for the Spatial–CIELAB. Black, Red and Blue colour lines represent the kernel for the Intensity, Red-Green and Yellow-Blue channel respectively.

the observing conditions the equivalence in pixels is easily computed by the following expression:

$$\sigma_{pixels} = d \cdot \tan(\sigma_{degrees}) \cdot R \tag{4.8}$$

where, σ_{pixels} and $\sigma_{degrees}$ represent the spreads in pixels and in degrees of visual angle, respectively; d is the distance in cm between the stimulus and the observer (or the camera in computer vision), and R is the display resolution that is given in pixels/cm.

The profiles of these symmetric filters are shown in figure 4.6, where the filters have been built to simulate the human colour perception of an image of 550 pixels, displayed on a visual field of 20cm and observed from 40cm.

To illustrate how this transformation behaves on a given image we shown in figure 4.7 the results of applying the Spatial–CIELAB transformation on two images presenting an important colour assimilation effect. We can see on the profiles below,

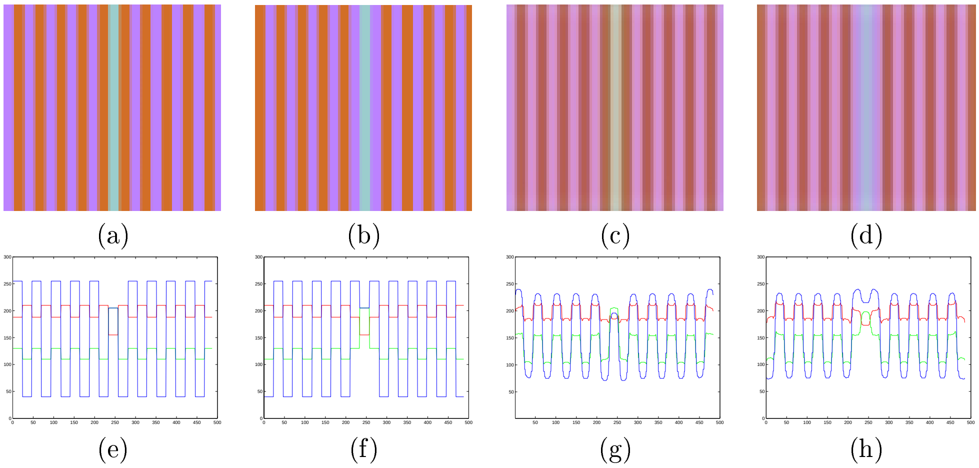


Figure 4.7: (a), (b) Examples of two images presenting important assimilation effects. (c) and (d) Previous images transformed by Spatial–CIELAB. (e), (f), (g) and (h) are the RGB profiles of images (a), (b), (c) and (d), respectively.

how the Spatial–CIELAB transformation makes that the green-blue band is becoming bluish when is surrounded by blue and it becomes reddish when surrounded by red.

4.3 Colour Contrast as a perceptual sharpening

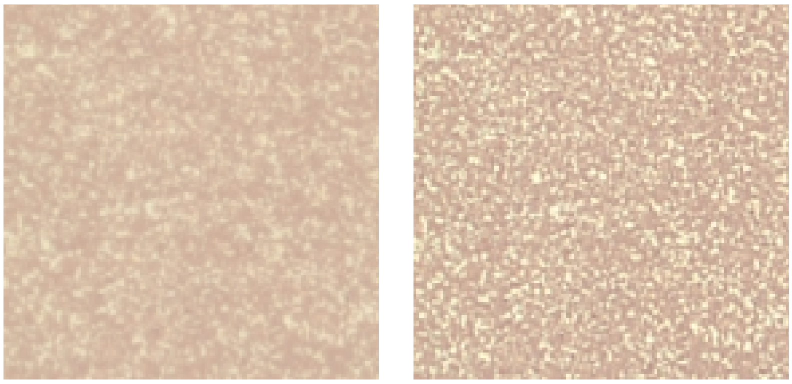
Colour contrast is the complementary mechanism to the assimilation that takes chromaticities of regions with spatial low frequency. Whereas a computational model for colour assimilation has been proposed in computer vision, a computational operator that simulates colour contrast phenomena has not been proposed in the computer vision literature.

In the following sections we will present the main contribution of this work that is devoted to this end, that is, to define an operator that enhances differences in the transitions among colours of regions presenting lower frequencies. While the assimilation effect has been solved by a blurring operator, it seems quite natural that the contrast effect will have to be implemented by a sharpening operator.

The final foal of this operator is to produce a sharpened image that allows a better segmentation of texture blobs in order to be able to compute their local attributes, following human perceptual considerations.

4.3.1 Local perceptual sharpening

In this section and in the subsequents we will progressively define sharpening operators presenting good properties to represent colour contrast. The first and the most common sharpening filter is defined as:



(a) (b)

Figure 4.8: Effect of the traditional sharpening operator:(a) original image, (b) sharpening of the image using the usual transform on the RGB space

$$S_c(I, \gamma) = I_c - \gamma \nabla(I_c) \tag{4.9}$$

where I_c is the c–th channel of a colour image I of dimensions $N \times M$, $\nabla(I_c)$ is the laplacian of the image channel c ($\nabla(I) = \partial^2 I/\partial x^2 + \partial^2 I/\partial y^2$) and γ is a constant that controls the amount of the enhancement. This process is done for each channel separately. Nonetheless, the laplacian operator is very noise sensitive. To avoid this problem, the laplacian of a gaussian (LoG) is used, that is, to smooth the image before the enhancement in order to reduce noise effects. The resulting operator has the following expression:

$$S_c(I, \gamma) = I_c - \gamma LoG(I_c), \tag{4.10}$$

$$LoG(I) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right] e^{\frac{x^2+y^2}{2\sigma^2}} \tag{4.11}$$

where the $LoG(I)$ expression is centered on zero and with gaussian standard deviation σ. Whichever it is the method used, there is a post–process to clip the output of the responses outside the range of the image (usually $[0 \ldots 255]$ in the rgb–space). We will use the notation $S(I, \vec{\gamma})$ to indicate that the operator $S_c(I, \vec{\gamma}[c])$ is applied for each channel of the image and merged together to form a new n–spectral band image.

The first attempt to chromatic contrast perception enhancement is the usual brightness sharpening, but applied to all the bands of the image, that is: $S(I, \vec{\gamma})$. This operator has been applied to the colour texture image of figure 4.8. Apparently, there is a clear enhancement of the texture that form the image. In the original image the transition from one blob to another blob of different colour is very smooth. Even that, some texture is appreciated. Enhancing the image makes the colour blobs more

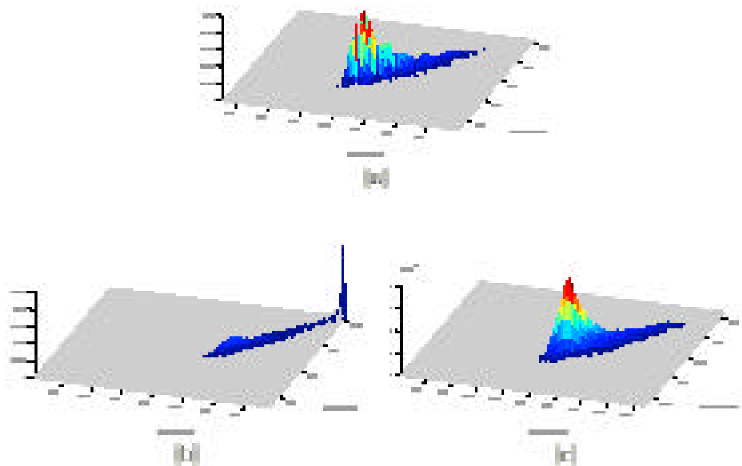
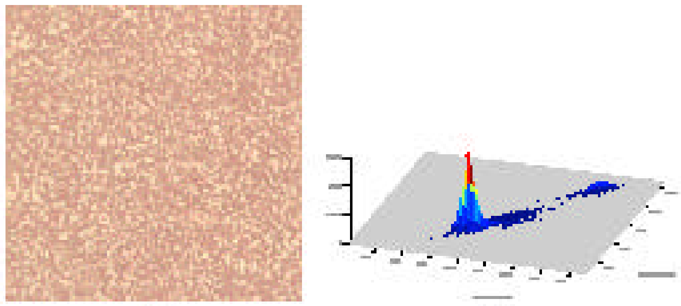


Figure 4.9: (a) Red-green histogram of figure 4.8(a). (b) Red-green histogram of figure 4.8(b). (c) Red-green histogram of image 4.8(a) after a sharpening operator without forcing the dynamic range to be the same than in the original image.

distinguishable, let us see how this effect is projected on the colour image distribution. In this example, the blue channel is the one with less colour information (the colours that appear are mainly formed by red and green). In figure 4.9 we show the histogram of both images using only the red and green channels. In figure 4.9(a) we see the colour distribution of the original image, although we can appreciate two kinds colours in the image, the distribution presents only one important peak. In figure 4.9(b) we see how the sharpen operator makes to appear the important peaks, one corresponding to pink kinds on the other corresponding to white kinds.

What we will use is a modification of the mask unsharp sharpening operator. Instead of operating on the RGB space we will operate on the opponent space that give us a more perceptual approach as we have previously introduced.

This change will not be enough to get good contrasted images. The fact that the range of the image is forced to the maximum and minimum of the dynamic range of the representation cause colour distortions to appear. This is the case of the peak belonging to white in figure 4.9(b), where a maximum is formed. When applying the operator without restricting the range of the result (fig. 4.9(c)) one of the two main colours disappears. And finally, in both cases the range of the result image is quite different from the range of the input image. In short, neither the limit on



(a) (b)

Figure 4.20: Effect of the perceptual space sharpening operator. (a) Sharpening of the image using the local sharpening operator T' on the figure 4.8(a). (b) the red-blue histogram of (a)

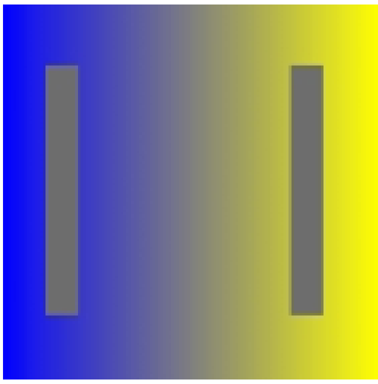
dynamic range nor the freedom to surpass this range are a good choice to improve the underlying texture. What we propose is to mix both points by resizing the output into the range of the neighbourhood of the input pixel. The neighbourhood limits the output result in the range of the minimum and maximum in a certain neighbourhood ω defining the spatial extension to look around. The expression of this new operator, T', is as follows:

$$T'(I)_{x,y} = RESIZE[S(C)\gamma(I), \gamma]_{min[I,\omega]}^{max[I,\omega]} \quad (4.22)$$

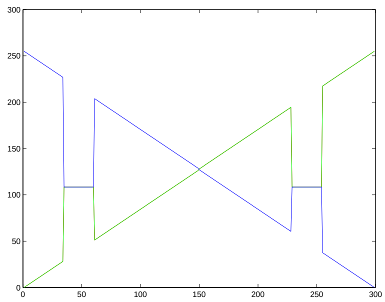
where the superindex and subindex $max[I,\omega]$ and $min[I,\omega]$ are the maximum and minimum range for each pixel inside a neighbourhood ω. The maximum and minimum threshold can be carried out by means of morphological operators. In this case $max[I,\omega] = I \oplus \omega$ which is the morphological dilation and $min[I,\omega] = I \ominus \omega$ being the morphological erosion, in both cases using ω as the structuring element.

The parameter γ plays an important role for the purposes. It allows to control which component of the perceptual stimulus we want to enhance. As an example, figure 4.20 shows the result of the defined operator with a constant γ. In this case the visual appearance of the image is closer to the original image than in the image in figure 4.8(b). Although the texture seems not to be as sharpen as it is using the traditional sharpening, the histogram shows the opposite. Using T' makes the blobs to be very well formed, without side effects, and the range of the resulting image is in keeping with the original image.

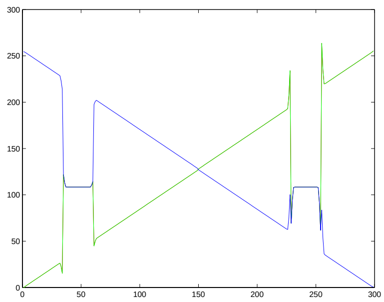
The problem of using equation 4.22, whichever the colour space or the range limit that is applied, is that it reinforces the differences in the transitions between different colour areas, but not the area itself. In fact, the behaviour of γ is similar to an edge detector. That makes the pattern to look like being dramatically enhanced when dealing with small texture images, but not when working with more homogeneous



(a)



(b)



(c)

Figure 4.11: Example of the Local perceptual operator for large stimuli: (a) original image, (b) RGB profiles of image (a), (c) RGB profiles of image (a) applying operator T.

ones. Some extra colours are created nearby the edges of the image, and this is what makes the texture emerge substantially. Although the edges are enhanced, and so is the texture, the main area of the colour stimulus is not changed as it should be for a chromatic contrast effect. An example of this situation is figure 4.11. (a) is the image used in explaining colour contrast but enhancing its colours. It can be observed that the effect is on the edges of the stimuli and not on the stimuli themselves. The centres of them are the same when there should be chromatically different. In fact, the operator should spread the edge response all over the areas that form the edge.

4.3.2 Region perceptual sharpening

Based on the work of Grossberg[2] [50] and the operator T, we want to construct a new operator to improve chromatic contrast simulation. The main idea is to recognise inhibited and activated areas, whichever the colour dimension is analysed. When applied, for example, to the red-green opponent colour dimension, the active areas will be the reddish ones and the inhibited areas the greenish ones. Computationally its is equivalent to the intensity of the stimulus. In the preceding example a red area is positive and a green area negative. However, whichever is the sign of the area it can not be considered neither positive nor negative unless it is compared with another area. 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 negative when green. The laplacian operator performs well in such definition because its response is positive in the transition between dark and light, and negative on the contrary. We will use the fact that the laplacian indicates the edge location by a zero cross, i.e: a change between positive an negative response or vice versa. We define an homogeneous area as the points that lie inside the regions surrounded by zero cross points. But as we are working in a discrete domain, the zero–crossings are not well locate. Then, a zero cross are those points where there is a change of sign of the laplacian between it and one of its neighbours. That makes us define two sets of zero–crossings:

$$Z_w(I) = \{\mathbf{p} \in I \mid \exists \mathbf{p_i} \in H(\mathbf{p}) : sgn(LoG(I,\sigma)_{\mathbf{p_i}}) = -1 \wedge sgn(LoG(I,\sigma)_{\mathbf{p}}) = 1\} \quad (4.13)$$

$$Z_b(I) = \{\mathbf{p} \in I \mid \exists \mathbf{p_i} \in H(\mathbf{p}) : sgn(LoG(I,\sigma)_{\mathbf{p_i}}) = 1 \wedge sgn(Log(I,\sigma)_{\mathbf{p}}) = -1\} \quad (4.14)$$

where $\mathbf{p}$ stands for the pixels of the image I, $H(p)$ for the pixels belonging to the neighbourhood of $\mathbf{p}$, and $sgn(\mathbf{p})$ is the sign of the intensity value in $\mathbf{p}$, which equals 1 when positive and -1 when negative. $Z_w(I)$ are the zero-crossings taken at the falling edge, and $Z_b(I)$ at the raising edge. $Z_w(I)$ coincide with the limits of the light (or white) areas and $Z_b(I)$ are the limits of dark (or black) areas. I is a one–channel image, being it the responses to one of the opponent channels. Usually $H(\mathbf{p})$ is defined as

$$H(\mathbf{p}) = \{(p_x, p_y - 1), (p_x + 1, p_y), (p_x, p_y + 1), (p_x - 1, p_y)\} \quad (4.15)$$

[2]This a psychophysical work on brightness contrast based on on–off lateral geniculate cells, modeling responses in the boundary contour system by a sum of exponential functions that is nearly equivalent to the laplacian of gaussian.