where (p_x, p_y) are the coordinates of $\mathbf{p}$. From these definitions we construct two sets of connected components. The inner points of the connected components are areas where there is no sudden changes, and will be considered homogeneous areas. We will say that a connected component is white when all of its surround is darker than itself. In the same way, black connected components are lighter than all of their surround. From this point, $\neg Z_w(I)$ defines the set of plausible white connected components, $C^w(I) = \{C_i^w(I)\}$, whereas $\neg Z_b(I)$ defines a set, $C^b(I) = \{C_i^b(I)\}$, of plausible black connected components. From now on, we will use the term *region* instead of connected component. The following step is to distinguish those $C_i^w(I)$ and $C_i^b(I)$ that are, actually, white or black regions. We define the white regions as

$$W(I) = \{C_i^w(I) \in C^w(I) \mid \sum_{\mathbf{p} \in C_i^w(I)} -sgn(LoG(I,\sigma)(\mathbf{p})) = |C_i^w(I)|\}, \tag{4.16}$$

and in a similar way it is defined the set of black regions

$$B(I) = \{C_i^b(I) \in C^b(I) \mid \sum_{\mathbf{p} \in C_i^b(I)} sgn(LoG(I,\sigma)(\mathbf{p})) = |C_i^b(I)|\} \tag{4.17}$$

Up to this point, not all the pixels are classified as belonging to a black or white region. Those unclassified pixels will be merged in a neutral class, $N(I)$. The pixels in $N(I)$ belong to regions that are surrounded by lighter and darker regions at the same time, and so, can not be classified as black or white regions.

$$N(I) = (C^b(I) \cup C^b(I)) - W(I) - B(I) \tag{4.18}$$

Thus far, all regions of the image are classified in one of the three types of regions. Moreover, we need to specify how much black or white these regions are. The final image $ELoG$ (*Expanded Laplacian of Gaussian*) will measure how different is a region from its surround assigning at each pixel of the region the maximum difference of all the pixels in this region with its surround (i.e: the laplacian of gaussian).

$$ELoG(I_{\mathbf{p}}, \sigma) = \begin{cases} min_{\mathbf{p_k} \in W_i(I)}(LoG(I,\sigma)_{\mathbf{p_k}}) & : \quad \mathbf{p} \in W_i(I) \\ max_{\mathbf{p_k} \in B_i(I)}(LoG(I,\sigma)_{\mathbf{p_k}}) & : \quad \mathbf{p} \in B_i(I) \\ 0 & : \quad \mathbf{p} \in N_i(I) \end{cases} \tag{4.19}$$

We use $ELoG(I_{\mathbf{p}}, \sigma)$ when applying the process to the pixel $\mathbf{p}$ of I, and $ELoG(I, \sigma)$ when it is calculated all over the pixels of the image I. Two examples of the $ELoG$ operator applied on monochromatic stimulus are shown in figure 4.12. The upper graphics show the original stimuli in blue lines. Graphics (c) and (d) show the result of laplacian in blue lines and the output of $ELoG$ in red lines. Black regions have positive response whereas white regions are negative. Regions between darker and lighter ones have 0 response. We want to remark how the maximum and minimum inside each region is expanded all around it.

What remains to conclude is to apply the sharpening formula using $ELoG(I)$ instead of $LoG(I)$. We will call the new operator *Expanded Sharpening* $(ES(I, \vec{\gamma})_{\vec{\sigma}})$,

$$ES(I, \vec{\gamma})_{\vec{\sigma}} = I - \vec{\gamma} ELoG(I, \vec{\sigma}) \tag{4.20}$$

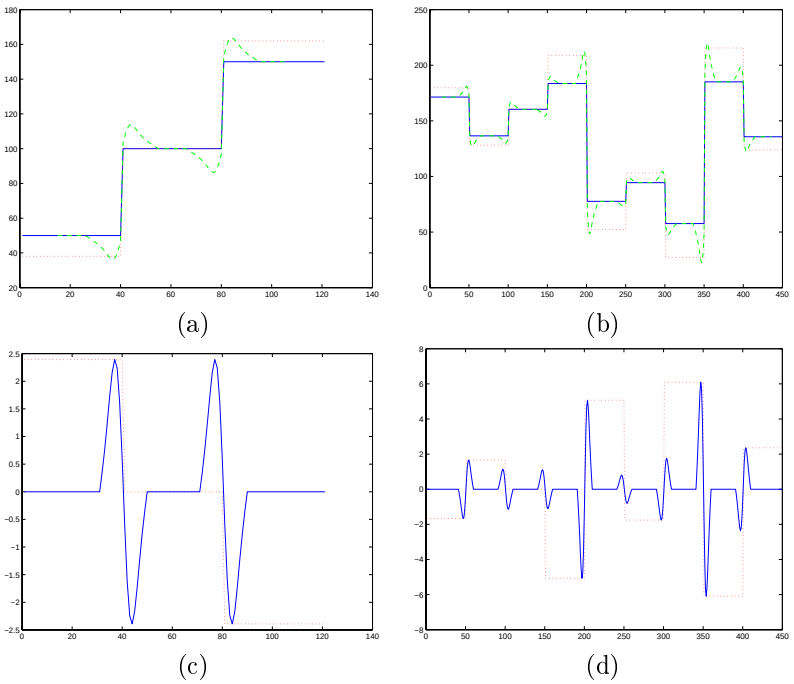


Figure 4.12: Graphic explanation of the effect of operator $ES(I)$, on (a) and (b): In solid lines the original stimulus, in dotted lines the output from the $ES(I)$ operator, compared with the T operator in dashed lines. On (c) and (d): In solid lines the Laplacian of Gaussian response, and in dotted lines the output form $ELoG(I)$.

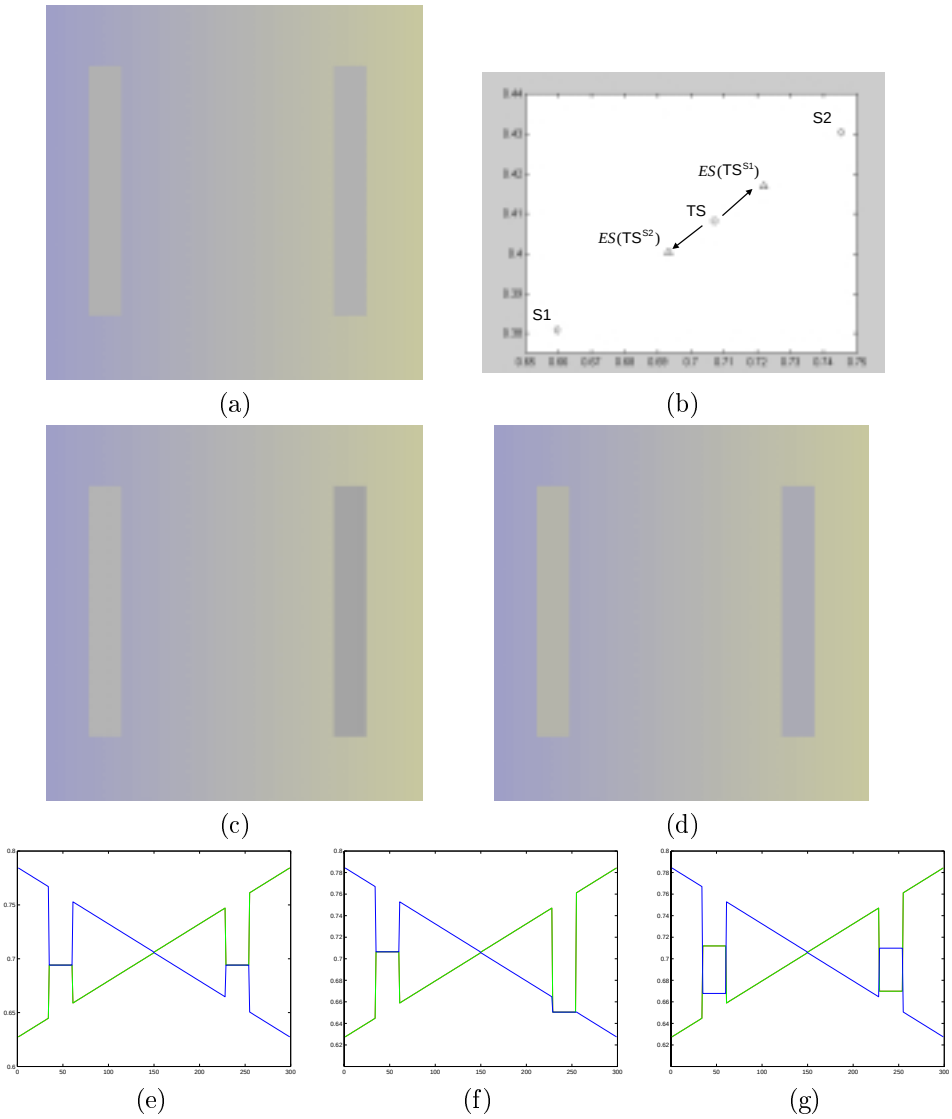


Figure 4.13: Region perceptual sharpening: (a) the input image, (c) and (d) the results with two different parameter configurations. (b) shows the displacement of the chromatic values of the test stimulus in (d) with regard to the original stimulus. And (e),(f),(g) comparison of the profiles of an horizontal line in the RGB space, from left to right: original, (c) and (d).

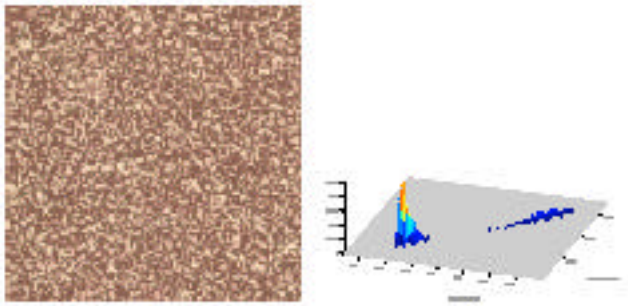
Returning to figure 4.12 in (a) and (b) the red lines are the response of the new region perceptual operator ES compared to the responses of the local perceptual operator T, green lines. Whereas $T(I)$ only have effect in a short neighbourhood, $ELoG(I)$ works on the whole region. The question is, will it work? and the answer is not so simple. When dealing with perceptual vision the way to validate a model is by means of psychophysical experiments. Some times they are done with a very reduced group of individuals and a short set of test, it is because of the intrinsic complexity of this type of experiments. The kind of tests are a uniform background scene with regular polygons, there are some that are more complex than others. They could be gratings or two simple squares [95, 98, 110, 81]. In any case they should be done by scientists of this field.

In this thesis this problem will not be broached as it is a computational approach to perceptual vision and we are not trying to imitate the human vision but to approximate the images to what humans see. As a matter of fact, this operator has been inspired in the experiments before mentioned. These experiments analyse the reaction to certain isolated stimulus and lead us to look for the different stimuli in the image.

We have introduced the example in figure 4.11 to see the leaks of operator T, in this case the result is what is expected. It is depicted in figure 4.13, where the original colours of the example are used. In (a) there is the input image and in (c) and (d) two examples of the operator changing the parameter $\vec{\gamma}$. (e), (f) and (g) are the horizontal profiles of the central line of the image for the input image and both examples (c) and (d). The red response is occluded by the green one as they are the same. It can be appreciated a shift of the test stimulus against the surround. The chromatic coordinates of the test stimulus for the first example (c) are the same as in the input image as the parameter $\vec{\gamma}$ has been adjusted to work only in the black–white pathway. In the second example $\vec{\gamma}$ is a constant vector, and thus, all pathways are equally weighted. In this case there is a change in the chromatic coordinates of the final stimuli. This situation is plotted in the graphic Fig.4.13(b), where the two surround of the stimuli have chromatic coordinates S1 and S2 (left to right) and the original test stimulus is TS. TS1 and TS2 are the chromatic values of the test in the output image. It is clear that they behave in the same direction that the HVS does.

One such examples of the operator on images of small isotropic texture is illustrated on figure 4.14.

This operator is based on the fact that a region is conceived as inhibited or activated in intensity, red–green or blue–yellow channel. If there is a problem it will be in the definition of the regions and the assumption that a region can be only inhibited or activated for a certain channel. But the reality shows that under certain circumstances it can be inhibited and at the same time activated. This is the case of the example of the bars. What happens when both bars are joined together with a slim bar of the same stimulus? Taking as example the blue–yellow channel, the left grey bar will be an activated region whereas the right bar will inhibit. But as they are connected, they are the same region. When a region is inhibited and activated simultaneously then it belongs to the set of neutral regions that show no reaction and then the result is the same input image. The modified experiment is shown in figure 4.15, where the stimuli are the same colour as before and the results of the operator



(a) (b)

Figure 4.11b The result of applying JSEG on the image in fig. 4.11(a), and a view of its histogram over the red and green bands.

Figure 4.11c Example of the region perceptual clustering operator in a case where its performance is not good enough. The text stimuli are connected and they constitute two perceptual regions being inhibited and activated at the same time.

are not shown because they are exactly the same image.

4.3.3 Spread perceptual sharpening

In short, the local perceptual sharpening operator fails because it does not extend to the centre of the stimulus and the region perceptual sharpening fails because,although it comes from psychophysical ideas, it does not consider one region to have two different behaviours at the same time. To solve this conflict we have designed an alternative operator that combines the good properties of the previous operators. The idea is to use the LoG edge enhancing to locate the boundary of regions and to use some of the definitions of the region perceptual sharpening to reduce the number of points used in the operator. The intensity of inhibition/activation in this points will be scattered to the centre of the region no matter which kind of region it is.

Starting form equations 4.13 and 4.14, which define the points that form the borders of the regions of the image, we can take the local inhibition or activation of a region taking the LoG in these points. The following step is to construct a surface where its height in a certain point indicates the level of activation of this point, taking into account the intensity on the points of edges that define the region to which it belongs. This surface must have some properties:

1. The points on the boundaries must preserve its energy, i.e: the relationship between adjacent regions must be maintained.

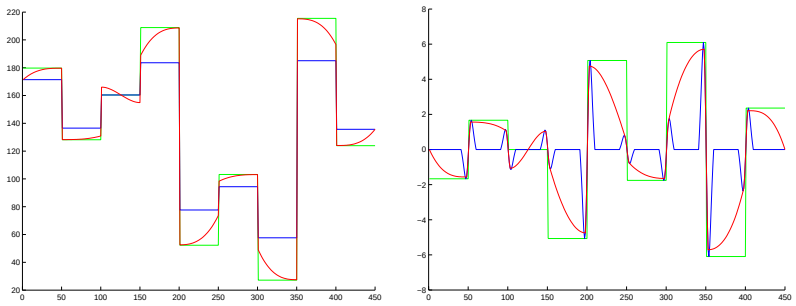
2. The zero crossings between points of the boundaries must remain equal, i.e: there will not be more regions than in the input energy image.

3. Zero crossing can only be added inside a neutral region (defined in Eq. 4.18).

Let us call $\mathcal{S}(\mathcal{X},\mathcal{Y})$ the operator that constructs this surface from the energies of a set of boundary points, X, giving the activation energy on points Y. An immediate solution is to use some kind of surface interpolation, but not all possible. Some of the possibilities are: nearest neighbourhood interpolation, linear interpolation and cubic Hermite interpolation. Some that are not possible are those based on spline interpolation. The choice of the interpolation method will affect the smoothness of the resultant image. The smoothness is achieved constraining interpolation to certain conditions on the continuity of the first and second derivatives. The complexity of these methods is considerable and it has to be kept in mind when working with large images. In this case, it is reasonable to use linear interpolation, instead. Now we can define the new operator. Since our definition of the operator spreads the energy of the region borders into its inside, we will call it *Spread Sharpening* (SS), and similarly the resulting energy surface is a spread modification of the LoG surface. Then,

$$SLog(I,\sigma) = \mathcal{S}(LoG(I,\sigma)_{Z_w(i) \bigcup Z_b(I)}, I), \tag{4.21}$$

is the spread Log taking as a control points the energy of the points where there is a change on the inhibition/activation, and evaluated all over the points of image I. Following the same schema than in the local and region perceptual sharpening operators (Eqs. 4.12, 4.20) the final operator will be defined as:



(a) (b)

Figure 4.16: Graphic explanation of the functioning of region perceptual sharpening (SS). In (a) the original signal (blue line), the ES output (green) and the SS (red) are plotted. (b) is the response of $SLoG$ in red in front of the LoG in blue.

$$SS(I,\vec{\gamma})_{\vec{\sigma}} = I - \vec{\gamma} ILoG(I,\vec{\sigma}) \tag{4.22}$$

Taking the 1D signal of figure 4.12(b) we will illustrate the effects of the operator. Figure 4.16(a) plots in blue line the original input, in (b) the blue line is the LoG response of the signal, green line is the $ELoG$ response of the previous operator and red line is the $SLoG$ response. It is evident the spread effect of the function $\mathcal{S}$. The function solves de problems of neutral regions and makes the edges influence the inner part of the region. The final output signal is shown in red in (a) compared to the output of operator ES in green.

We noted that the region perceptual sharpening failed when applied to image in figure 4.15. Let us test the performance of this last operator. Figure 4.17(a) depicts the resultant image, whereas in (b) we have shown, as an example, the output from the inhibition/activation function $SLoG$ of the blue–yellow pathway. The profiles shown are from the original image and the output image. The profile from the output of ES applied on the same image is not shown because it is exactly the same as the profile from the input image.

4.3.4 Examples

These operators should be tuned to the contents of the image to adjust the frequencies at which they work better taking into account the distance from which the images are seen. Other parameter to adjust are the ratios of each opponent channel in the contrast response. These adjustments should come from psychophysical measurements. Whereas it seems that there begins to be a consensus on the first set of parameters, contrast begins at least at 1.7 cpd, it is not clear the influence of each channel in the response. Psychophysics agree that the intensity is the most sensitive channel, in second term there is the red–green channel and finally the blue–yellow channel. However we did not find any literature on which their ratios are.

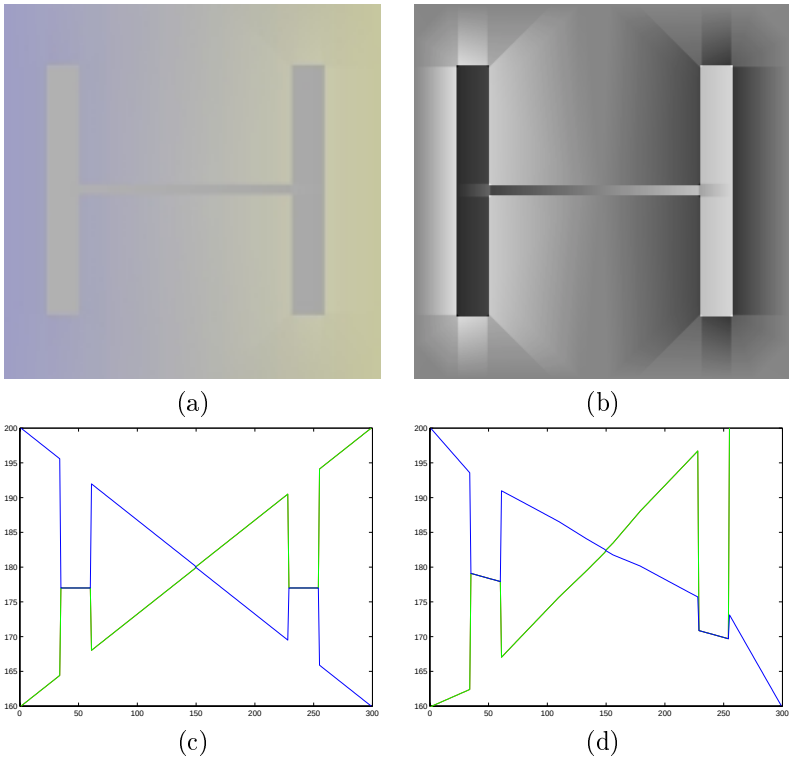


Figure 4.17: Graphic explanation of the spread perceptual sharpening operator: (a) is the spread perceptual operator output applied to the image in figure 4.15, (b) is the $SLoG$ response using linear interpolation of the blue–yellow pathway, (c) is the profile of an horizontal line from the original image in the RGB space and (d) in the case of (a).

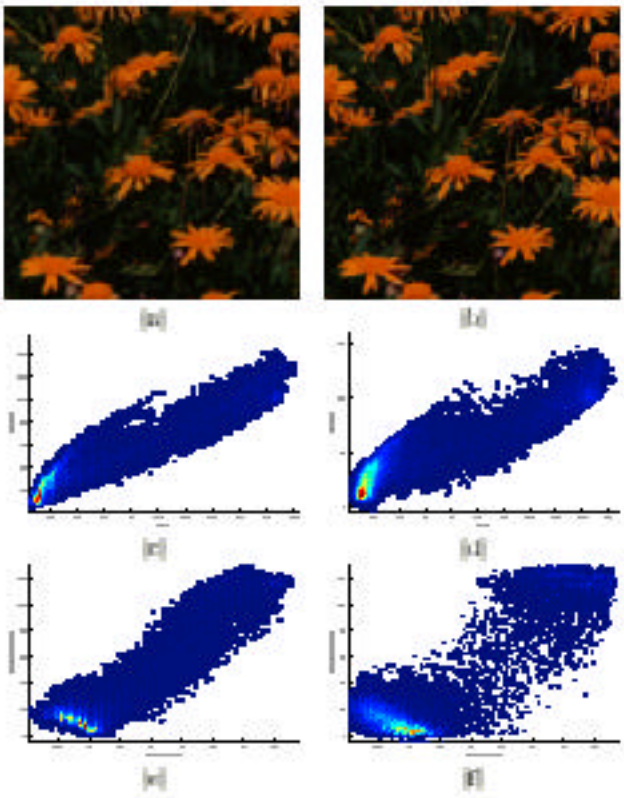


Figure 4.18: Example of the spread perceptual sharpening (Watson,). (a) and (b) original and spread perceptual sharpened result. (c) and (d) red-green histograms of images (a) and (b). (e) and (f) red/green-blue/yellow histograms of images (a) and (b).

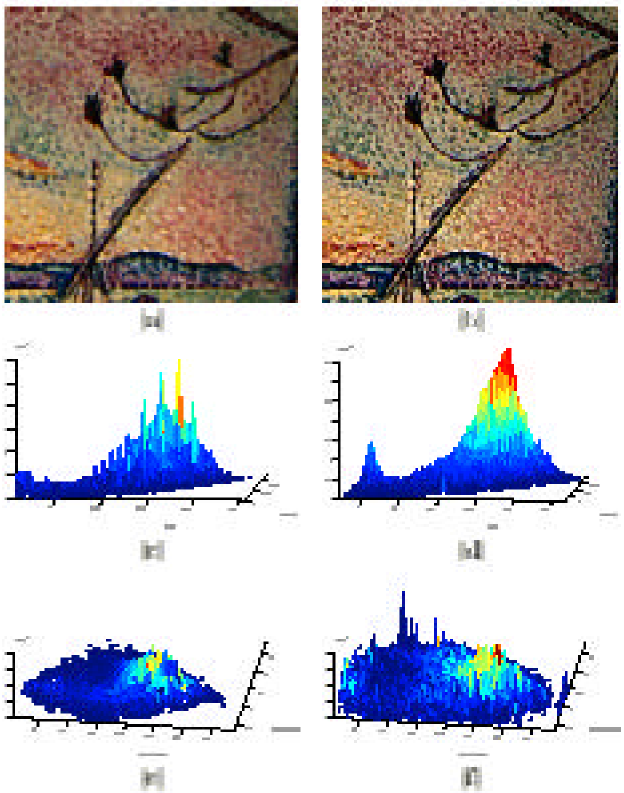


Figure 4.20: Example of the spread perceptual sharpening (Ch-tree________). (a) and (b) original and spread perceptual sharpened result. (c) and (d) red-green histogram of images (a) and (b). (e) and (f) intensity–red/green histogram of images (a) and (b).

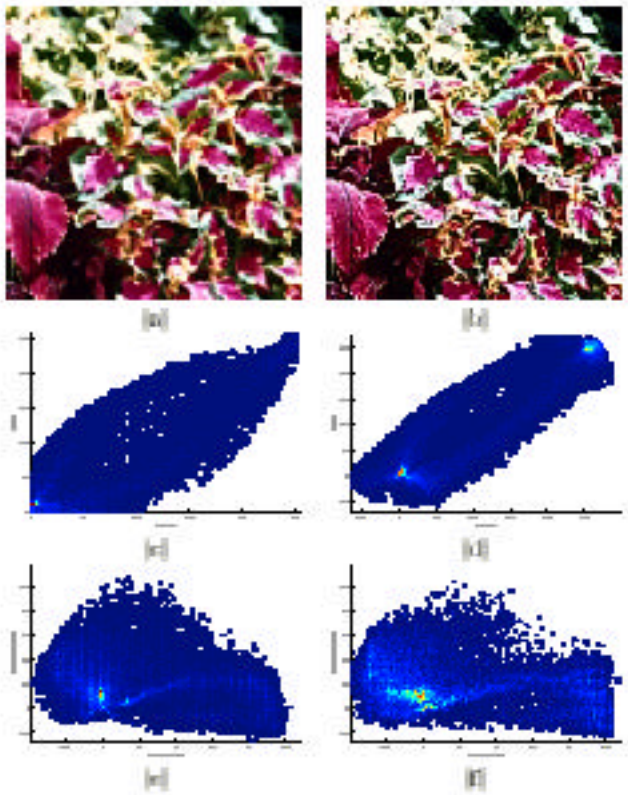


Figure 4.33: Example of the spatial perceptual sharpening [Weiss_____]. (a) and (b) original and spatial perceptual sharpened result. (c) and (d) green-blue histograms of images (a) and (b). (e) and (f) red/green-blue/yellow histograms of images (a) and (b).

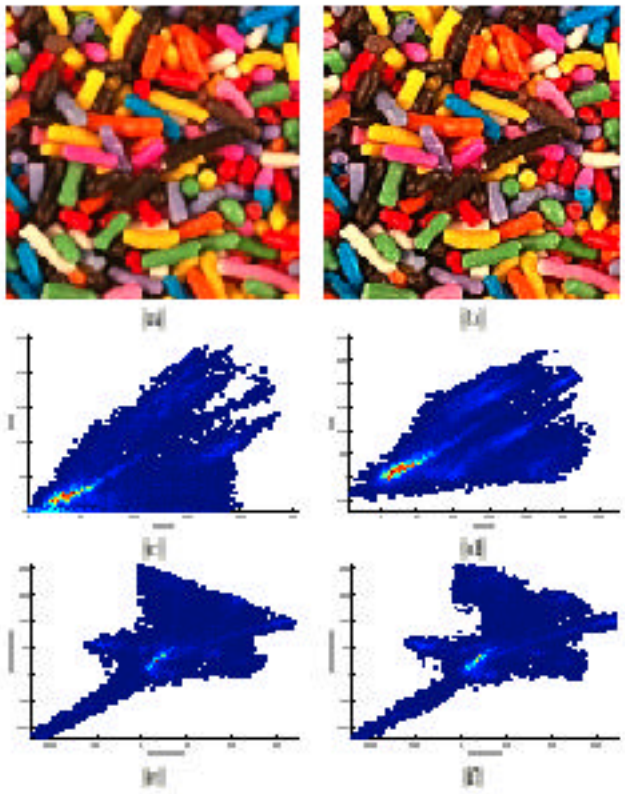


Figure 4.22 Example of the spatial perceptual sharpening (Vistex). (a) and (b) original and spatial perceptual sharpened result. (c) and (d) rich-prim histograms of images (a) and (b). (e) and (f) intensity-weighted histograms of images (a) and (b). (g) and (h) the 2D-histograms reflecting one of the dimensions in the HCH space. (e) and (f) 2D-histograms in the opponent space.

scale ($\vec{\sigma}$)	$M_1^{ps} > M_1^o$	$M_2^{ps} > M_2^o$	$M_1^{ps} > M_1^o \wedge$ $M_2^{ps} > M_2^o$	M_1^{ps} imp.	M_2^{ps} imp.
high	73.78	67.07	48.17		
medium	99.39	42.68	42.07		
low	63.41	58.54	38.41		
all	99.39	85.98	85.37	159.4	97.1

Table 4.2: Spread percetual sharpening on VisTex image database. Values are in %.

To study the operators we have analysed a set of images from the texture image database VisTex from MIT MediaLab. It is a large database from which we have selected some of them to illustrate the effects of the spread perceptual operator. The figures 4.18, 4.19, 4.20 and 4.21 show four of these images. In all cases: (a) is the original image, (b) is the output image for which the parameters have been chosen empirically, (c) and (d) are the projected histograms of (a) and (b) respectively rejecting one of the dimensions, in each case the dimension rejected was the one that enables to show a better view, and finally, (f) and (g) are the projected histograms on the opponent space, the rejected dimensions have also been chosen to best display the effects of the operator.

The effects are more visible when analysing the opponent space. In the first example the division between green and orange is greater in the perceptual sharpened image than in the original one. Although the printed images do not show a large difference it exists and it is very useful in segmenting colours. The second example show one case where there are some colours but they can not be intuited from neither the RGB nor the opponent RGB 2D–histogram. When the image is perceptually sharpened the opponent histogram show peaks belonging to the colours on the image, that do not appear in the original. The effects on the third example can be seen even in the RGB histogram. Two narrow peaks (those in red and yellow) show the localisation of the red and green leaves. In the last example the effects of the operator are weaker than in the previous images. However this is natural if we consider that the number of colours is large and their spatial location does not produce a colour contrast effect. In this case the operator does not spread colours but concentrates the distributions of colours.

4.4 Validation

While psychophysical test are not done we have validated the operators looking for indexes that show a better discrimination between colours. If the operators perform well the resultant images should be easier to segment in the predominant colours. Following the scheme in [46], we segment the image in two clusters and get a measure of how good this clustering is. A new segmentation is done with three clusters and the measure is calculated. If the previous measure is better than the new one we stop, if not the number of cluster is incremented and the comparison is done again until

scale ($\vec{\sigma}$)	$M_1^{ps} > M_1^o$	$M_2^{ps} > M_2^o$	$M_1^{ps} > M_1^o \wedge$ $M_2^{ps} > M_2^o$	M_1^{ps} imp.	M_2^{ps} imp.
high	82.32	67.68	57.93		
medium	73.78	66.46	52.44		
low	68.29	70.73	48.78		
all	87.80	86.59	76.22	102	87.2

Table 4.3: Region perceptual sharpening on VisTex image database. Values are in %.

scale ($\vec{\sigma}$)	$M_1^{ps} > M_1^o$	$M_2^{ps} > M_2^o$	$M_1^{ps} > M_1^o \wedge$ $M_2^{ps} > M_2^o$	M_1^{ps} imp.	M_2^{ps} imp.
high	73.78	78.66	58.54		
medium	75.61	75.61	59.15		
low	71.34	79.88	59.76		
all	86.59	90.24	79.27	38.06	46.94

Table 4.4: Local perceptual sharpening on VisTex image database. Values are in %.

process	$M_1^{ps} > M_1^o$	$M_2^{ps} > M_2^o$	$M_1^{ps} > M_1^o \wedge$ $M_2^{ps} > M_2^o$	M_1^{ps} imp.	M_2^{ps} imp.
SS vs None	99.39	85.98	85.37	159.4	97.1
ES vs None	87.80	86.59	76.22	102	87.2
T vs None	86.59	90.24	79.27	38.06	46.94

Table 4.5: Summary on perceptual sharpening on VisTex image database. Values are in %.

a maximum on the measure is reached. Instead of using a k-means algorithm as in [46] we used an *Expectation–Maximisation* mixture of gaussians which is more general and fits better the data. Both methods are briefly explained in section 5.2. There are a number of ways of measuring how good a clustering is, Coleman and Andrews enumerate some of them in [21]. In this validation experiment we have selected the following two:

$$M_1 = tr(\mathbf{S}_b)tr(\mathbf{S}_w) \tag{4.23}$$

$$M_2 = \frac{\mathbf{S}_b}{\mathbf{S}_w} \tag{4.24}$$

where $tr(\cdot)$ indicates "trace" or sum of the diagonal elements of a matrix, $\mathbf{S}_w$ is the within groups scatter matrix, a measure of how condensed the cluster is, and $\mathbf{S}_b$ is the between scatter matrix, a measure of the distance between clusters. The scatter matrices are defined in a better context in equations A.2 and A.3 in section A. We will use them here just as a tool.

The measure used should have a maximum when the best clustering is reached. Although M_2 is better when evaluating the dispersion of clusters, it is not upper bounded whereas M_1 is. We have used M_1 to iterate the clustering process and both M_1 and M_2 to measure the behaviour of the operators.

Another problem is that the parameter $\vec{\sigma}$ involved in the operators should be settled specially for each image, however to automatically find the best scale for each image is still an open issue that will derive form this thesis. What we will do is to try three different scales: high, medium and low, keeping the best clustering. The original image is also clustered using the same criterion. When M_i is applied on the original clustered image we will denote it as M_i^o and when done with the sharpened images M_i^{ps}, whichever it is the used operator. The sharpening is done on the two chromatic channels, the intensity is left as it is to show the computational chromatic contrast behaviour. The experiment is done on 164 images of the VisTex image database. Tables 4.2, 4.3 and 4.4 show the results for the SS, ES and T operator. The last two columns of the tables are the percentage of improvement of the measures wit respect to the measure on the original image.

From these results we can conclude that the operators are performing a good separation of colours on the image. And as we suspected, the more complex is the operator the better is its performance. The last row of the tables takes the best clustering in the three scales. Table 4.5 summarises the three operator to show their evolution.

4.5 Discussion

After an introduction to the colour induction phenomena it is concluded that computer vision lacks of an approach to the chromatic contrast effect. While there exists a computational model of colour assimilation, this is not the case for colour contrast.

Our contribution in this subject materialises in three new operators. The first one takes the traditional RGB sharpening operators to a space where colour appearance

is best modeled, adding spatial constraints to the generated responses. It has been illustrated it can work in some circumstances where the stimuli are small, and it can be adequate to very high frequency textures, but there are many situations where it does not fit well. This drives us to psychophysical literature that gives us the trail to search for inhibited and activated regions on the different visual pathways. Once more, there are many cases where the operator is useful but in some others it can not simulate the human visual perception. It was an inflection point to the search of a more general chromatic contrast operator. The result was the *Spread perceptual sharpening* operator that gather the experience in the preceding operators. The core of the operator is the idea of spreading the inhibition and activation of the cells on the transition between regions.

The capacity of the operators to differentiate colours has been tested on a texture image database, performing a segmentation and measuring how good it was compared to the image itself. The results obtained shows a good progression. Although the third operator is the more complete, the knowledge of the scene can advise to choose one of the other operators. This is a matter of complexity. Each operator can cope with more circumstances than the previous one but at the expense of computer resources.

There are open issues that have to be addressed in a near future and they are outside the scope of this thesis. The first one is to find the way to combine both spatial blurring and contrast induction in the same scene. One approach could be to look for different frequency regions in an image and applying the most suitable colour induction. On the other hand, we have presented a method that can be adjusted to viewing distance ($\vec{\sigma}$) and to the weight of each channel in the chromatic contrast effect ($\vec{\gamma}$). Both set of parameters have to be analysed from a psychophysical point of view, and then transferred to the computer vision field.