

Contrast and Color

Jean-Michel Morel, Ana Belen Petro and Catalina Sbert

November 3, 2010

1 Color: description and representation

1.1 Introduction

The study of Color involves several branches of knowledge: physics, psychology, mathematics, art, biology, physiology,... Each one of these disciplines contributes with different informations, but the way towards the absolute understanding of color has just begun.

According to the *Collins* dictionary (2000): “Color is:

- a. an attribute of things that results from the light they reflect, transmit, or emit in so far as this light causes a visual sensation that depends on its wavelength,
- b. the aspect of visual perception by which an observer recognizes this attribute,
- c. the quality of the light producing this aspect of visual perception.”

This definition illustrates the complexity of the notion of color and roughly sketches the three factors on which color depends: light, physical objects and our visual system.

Color does not exist by itself; only colored objects exist ([15]). Three elements are necessary for the existence of color:

- A light source, to light the scene.
- The objects, which reflect, spread, absorb or diffract the light.
- A receptor, which captures the spectrum reflected by the object.

Color basically depends on these three elements, in such a way that, if one of them is not present, then we are not able to perceive color. On one hand, color is a physical attribute, due to its dependence on a light source and on physical characteristics of the objects; on the other hand, it is a psychophysical and physiological attribute, since it depends on our visual perception.

Light consists of a flux of particles called photons, which can be regarded as tiny electromagnetic waves ([17]). These waves must have a length between 380nm and 780nm to stimulate our visual system. The wavelength content of a light beam can be assessed by measuring how much light energy is contained in a series of small frequency intervals. The light can then be described by its spectral distribution of energy.

The color of an object is defined and measured by its reflection spectrum ([17]). When light hits an object, the following three phenomena can happen: the light can be absorbed and the energy converted to heat, as when the sun warms something; it can pass through the object, as when the sun's rays hit water or glass; or it

can be reflected, as in the case of a mirror or any light-colored object. Often two or all three of these phenomena occur simultaneously. Moreover, the same object can have different “colors”, depending on the light source, on the geometry of the light devices, or on the change of some of its physical characteristics.

Finally, the receptor can be of different nature, for example, a photographic camera, a video camera, the eyes,... These last ones are the receptors in the human vision, they are the “gateway” for external information and this is the beginning of a large and complicated process which we explain in detail in the next subsections.

In the next sections, we will take a brief look at the human visual system and at the principal characteristics of color. In Section 1.4 different systems for the representation of color information are presented.

1.2 How do we see color? Human vision

Briefly, vision is an active process depending both on the operations of the brain, performed thanks to eye information, and on the external, physical environment. The human eye and brain together translate light into color. But, how is this conversion done?

In the human vision system, the information related with color perception is encoded at two fundamental levels. The first level occurs in the receptors which are located on the retina. It is a level related with the existence of three types of receptors called cones. Their responses are the starting point for the second level, which is based on the activity of three opponent mechanisms. We start by describing the first level, also called the *trichromatic level*.

In human vision, cones and rods in the eyes are the receptors of information from the external world. In particular, the rods are responsible for our ability to see in dim light and the cones are the color receptors. The number of rods and cones is different, and it differs depending on the part of the retina where they are located; rods generally outnumber cones by more than 10 to 1, except in the center of the retina, the fovea. In Figure 1 we can see a graphical representation of the distribution of rods and cones in the eye. The cones are concentrated in the so-called *fovea centralis*. Rods are absent there but dense elsewhere. We can note the absence of both rods and cones between -15 and -19 degrees of visual angle. This is the region of the blind spot where there are no cones or rods and the collected ganglion axons exit the retina as the optic nerve.

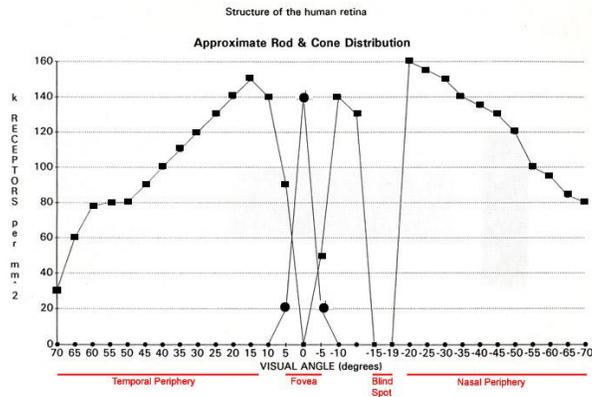


Figure 1: Approximate distribution of rods and cones across the retina. Cone distribution is signified by the line with the circle markers and the rod distribution is signified by the square markers.

Each rod or cone contains a pigment that absorbs some wavelengths better than others. Depending on the

absorption peak of the pigment, we can distinguish three different types of cones: L, for the long wavelengths; M, for the medium wavelengths; and S, for the short ones (see Fig. 2). Due to the overlap between the curves, the wavelengths which affect one particular cone affect the other ones, too.

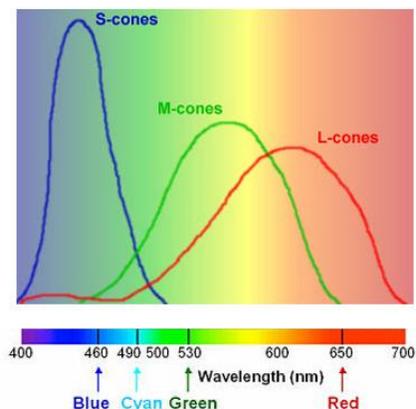


Figure 2: *Absorption spectra for the three classes of cones in the retina.*

These cones are wrongly called blue, green and red cones, respectively, due to their absorption peaks: at 440 nm for cone S, at 545 nm for cone M and at 580 for cone L (see Fig. 2). These names are wrong because monochromatic lights, whose wavelengths are 440, 545 and 580 are not blue, green and red, but violet, blue-green and yellow-green. For this reason, if we were to stimulate cones of just one type, we would not see blue, green and red but, probably violet, green and yellowish-red instead.

We talk of *tristimulus values* to refer to the joint response of the three types of cones to an input light. It turns out that the human visual system cannot distinguish between two physically different color lights as long as they produce the same tristimulus values. Colors with different spectral power distribution but the same tristimulus values are known as *metameric color stimuli* and they match in color for a given observer. Metamerism is what makes color encoding possible: there is no need to reproduce the exact spectrum of a stimulus. It is sufficient to produce a stimulus that is visually equivalent to the original one. This yields a principle for color reproduction, for instance, in televisions or printers (see Section 1.3).

We have to consider that the number of each type of cones in our retina is different. As a consequence, the eye response to blue light is much weaker than its response to red or green, because almost $2/3$ of the cones process the longer light wavelengths (reds, oranges and yellows).

This theory about the three types of receptors in our retina does not explain some psychophysical observations (such as post-image or simultaneous contrast), which can be understood by considering the existence of a second level of information processing provided by the cones.

All these observations have led to a theory called the *opponent process theory* which proposes that trichromatic signals from the cones feed into a subsequent neural stage and exhibit three major classes of information processing in the superior cells of the retina. Three different biochemical mechanisms occur at this level, which respond in an opponent way to different wavelengths. The first mechanism is red-green, which responds positively to red and negatively to green; the second one is yellow-blue, which responds positively to yellow and negatively to blue; and, the third one is white-black, which responds positively to white and negatively to black.

Positive response means a biochemical process in the receptors that produces a chemical substance; whereas, negative response means the breakdown of this chemical molecule.

In Figure 3 we can see the connection between the trichromatic phase and the opponent process phase. There, we can observe how the S, M and L cones are connected to the bipolar cells to produce the opponent responses. The blue-yellow bipolar cells are excited by the medium and long cones and inhibited by the short cones. The red-green cells are excited by the long and short cones and inhibited by the medium cones. The red-green cells are excited by the long and short cones and inhibited by the medium cones.

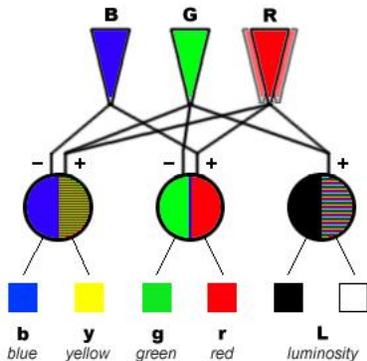


Figure 3: *Diagram of the connection between the three cones and the opponent cells.*

It is worth noting that the human visual system does not discriminate all the colors with the same efficiency and if two colors are very close perceptually, they can be confused. In the perceptual system, the sensitivity to the distance between colors is not related to the photoreceptors behavior in an uniform way ([15]).

Finally, we must to talk about the chromatic adaptation. The visual system has the capacity to adapt to the average luminosity of the environment. Color constancy can be described as the fact that if a scene is lighted up from a given source and this light source is changed, then the object colors are not changed for the observer. On the contrary, if the illumination change only affects one object, then the object color is changed. In other words, an observer adapts his colors perception in reference to the whole visual field.

The color constancy phenomenon can also be observed by examining a white piece of paper under various types of illumination, for example, daylight and tungsten light. The spectral distribution of these two illuminants is very different. Thus, the paper color should look different under the different illuminants. It turns out that the paper retains its white appearance under both light sources. This is due to the fact that the cones adapt their response to the reflectance level in the environment. We can conclude that the visual mechanism to perceive the contrast of the colors depends on the relative proportions between the different illuminations in the image.

This phenomenon does not affect the display devices, since they emit their own light, unless there is a colored light source illuminating them.

1.3 Creating colors

There exist two different methods of mixing colors: the first one is additive mixture, which is the addition of different wavelengths and is produced by light mixture. The second one is subtractive mixture, which is the subtraction or cancelation of bands of wavelengths by the combination of light absorbing materials. It is produced by mixing paints.

In additive mixing, we start the mixture with the absence of color stimuli, and we add three lights (“colors”) of different wavelengths, that cover different parts of the spectrum, to stimulate the S, M and L cones receptors in the eye. Any three colors that are linearly independent (i.e., none of them can be obtained as a mixture of the other two) can be considered in order to obtain an additive mixture. These three colors are called primary colors. Red, green and blue are the most commonly used as additive primary colors (see Fig. 4). In this type of mixture, the black color is the result of no colors mixed at all. On the contrary, the addition of these three primary colors yields the white color.

By means of this kind of mixture, color is produced on a color display such as a television or computer monitor. The surface of a color display is made up of hundreds of tiny dots of phosphor. Phosphors are chemical substances that emit light when they are bombarded with electrons and the amount of light given off depends on the strength of the electron beam. The phosphors on the screen are in groups of three, one bluish (for the short wavelengths), one greenish (for the medium wavelengths) and one reddish (for the long wavelengths). By varying the intensity levels of the phosphors, different levels of lightness are achieved. The human eye does not have enough resolution to distinguish each isolated element. For this reason the perception is a light beam with the three primary colors blending at each point, creating any hue.

Subtractive mixing is the process of filtering parts of the spectrum of the reflected light. It is based on the capacity of the surface to reflect some wavelengths and absorb others. When a surface is painted with a pigment or dye, a new reflectance characteristic is developed based on the capacity of the pigment or dye to reflect and absorb the different wavelengths of light. An example shall help to better understand this kind of mixture. Consider a surface painted with a yellow pigment which reflects wavelengths 570-580nm and another surface painted with cyan pigment which reflects 440-540nm. The color resulting from the mixture of both pigments will be green. This is because the yellow pigment absorbs the shorter wavelengths and the cyan pigment absorbs the entire longer wavelengths. As a consequence the only reflected wavelengths are some medium wavelengths, which create the sensation of green. Yellow, cyan and magenta are the most commonly used subtractive primary colors (see Fig. 4). As mentioned above, the yellow ink absorbs the short wavelengths and the cyan ink absorbs the long wavelengths, while the magenta ink absorbs the medium wavelengths. They are called the secondary colors and they are obtained by adding two primary colors. Yellow is obtained from red and green, cyan from green and blue, and magenta from red and blue. Here white is the result of no mixing (the entire spectrum is reflected) whereas mixing the three subtractive primary colors yields no reflection of light, i.e., black.

Some mechanical devices, such as printers, use the secondary colors to obtain the rest of colors. In printers, the original image is separated into its cyan, yellow and magenta components. A film is made for each separation and then a plate is produced from the film. White paper (which reflects all wavelengths) is run through the stations of a color press to accept layers of ink from each plate. The different quantity of ink mixed produces a higher or lower level of absorption in the short, medium and long wavelengths. In theory, it is possible to create any reflective color by mixing a combination of cyan, magenta and yellow. In practice, however, the inks that printers use are not perfect. This becomes more obvious when all three colors are mixed to obtain black color. The color that results is muddy brown, due to the impurities in the inks. That is why printers use black ink to get the best results.

1.4 Color spaces

In order to use color as a visual cue in image processing, an appropriate method for representing the color signals is needed. The different color specification systems or color models (color spaces or solids) address this need. Color spaces provide a rational method to specify, order, manipulate and effectively display the object colors taken into consideration. The color space choice depends on the previous knowledge that we have about “color” and on the application that we want to give to this information (for example, defining colors, discriminating



Figure 4: *Additive and subtractive mixture.*

between colors, judging similarity between colors,...).

The color models are normally three-dimensional spaces, due to the fact that our perception of color is trichromatic. The different spaces differ in the choice of the coordinate system, which defines the space. In the classical literature, four basic color model families can be distinguished ([13]):

1. **Physiologically inspired color models**, which are based on the three primary colors which stimulate the three types of cones in the human retina. The *RGB* color space is the best-known example of a physiologically inspired color model.
2. **Colorimetric color models**, which are based on physical measurements of spectral reflectance. Three primary color filters and a photometer are usually needed to obtain these measurements. The *CIE* chromacity diagram is an example of these models.
3. **Psychophysical color models**, which are based on the human perception of color. Such models are either based on subjective observation criteria or are built by taking into account the human perception of color.
4. **Opponent color models**, which are based on perception experiments, utilizing mainly pairwise opponent primary colors.

1.4.1 *RGB* Color Space

The red, green and blue receptors in the retina define a trichromatic space whose basis is composed by pure colors in the short, medium and high portions of the visible spectrum. As we have mentioned in Section 1.3, it is possible to reproduce a large number of colors by additive mixture using the three primary colors.

The *RGB* model is the most natural space and the most commonly used in image processing, computer graphics and multimedia systems. This is due to the fact that display devices use the addition mixture and the three primary colors to reproduce and “encode” the different hues (see Section 1.3). This model is based on the cartesian coordinate system. The pixel’s red, green and blue values in a digital color image are the three coordinates of the model and they are represented by a cube, since the three values are non-negative and they are considered as independent variables. Pure red, green and blue are situated in three vertices of the cube, while the other three vertices correspond to pure yellow, cyan and magenta (see Fig. 5). Black has coordinates $(0,0,0)$ and, at the opposite vertex, stands the white color. We call the line that joins the black vertex to the white vertex the *grey axis*, in which the three coordinate values are equal. In mathematical terms, the *RGB*

cube is made up of all points of three coordinates, (r, g, b) , with $0 \leq r, g, b \leq M$. The range of the three coordinates is usually from 0 to 255. The space is sometimes normalized so that $M = 1$.

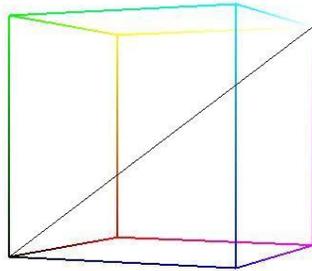


Figure 5: *The RGB color model.*

The main disadvantage of this space is the perceptual non-uniformity, *i.e.* the low correlation between the perceived difference of two colours and the Euclidean distance in this space. Its psychological non-uniformity is another problem. In this space, the information about the chromaticity (hue and saturation) and intensity components of color is not independent. Finally, considering the representation of a natural image in the *RGB* space, we observe a powerful correlation between the different coordinates. It is due to the special distributions of the sensors in the camera and to the post-processing operations. Therefore independent operations over each coordinate are not possible.

In Figure 6 and Figure 7 we can observe two images with three different points of views of their representation inside the RGB cube.

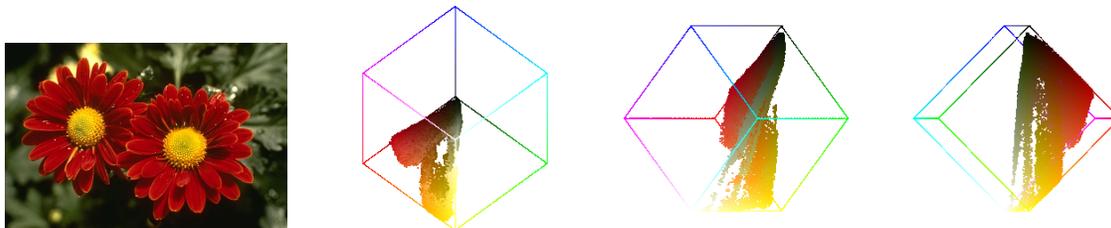


Figure 6: Original image with three different points of view of their RGB cube. The colors of the image are situated inside the RGB cube in their corresponding site.

1.4.2 *HSI Color space*

The *RGB* space is not easy for user specification and recognition of colors. The user cannot easily specify a desired color in the *RGB* model. Models based on lightness, hue and saturation are considered to be better suited for human interaction. These three color features are defined as:

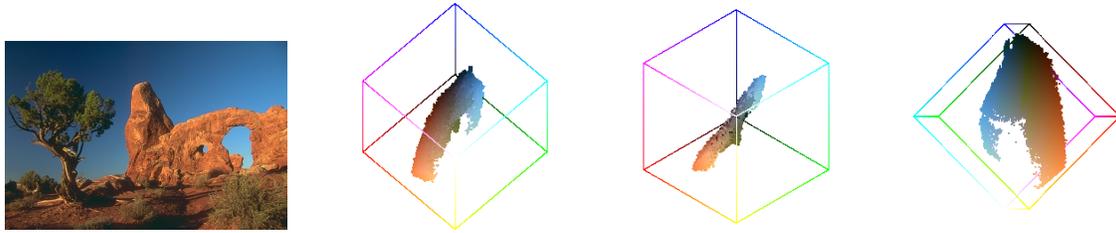


Figure 7: Original image with three different points of view of their RGB cube. The colors of the image are situated inside the RGB cube in their corresponding site.

- **Intensity or lightness** is the visual sensation through which a surface that is illuminated by a given luminous source seems to project more or less light. It corresponds to light, dark or faint terms. In some sense, lightness may be referred to as relative brightness.
- **Hue** is the visual sensation that corresponds to the color purity. The hue is defined by the dominant wavelength in the spectral distribution.
- **Saturation** measures the proportion on which the pure color is diluted with white light. It corresponds to pale, faded or brilliant terms.

The *HSI* color model owes its usefulness to two main facts. First, the intensity component is decoupled from the chrominance information represented as hue and saturation. Second, hue is invariant with respect to shadows and reflections. This is due to the fact that hue is the dominant wavelength in the spectral distribution and is independent on the intensity of the white light. We would wish the three variables to be independent. However, saturation depends on the intensity. If the intensity level is very high or very low, then, the saturation variable can only take very low values.

The *HSI* color space can be described geometrically from the *RGB* cube (Fig. 8). It is constructed by placing an axis between the black point and the white point in the *RGB* space, that is the diagonal of the *RGB* cube. As we have mentioned before, this axis is often referred to as the *grey axis*. Any color point on the *HSI* space is defined by its three components with respect to this axis: hue, which is the angle between a reference line and the line passing through the color point and orthogonal to the grey axis; saturation, which is the radial distance from the point to the grey axis; and intensity, which is the magnitude of the orthogonal projection of the color point onto the grey axis.

The main drawback of this representation in the continuous case is that the hue has a nonremovable singularity on the grey axis. This singularity occurs whenever $R = G = B$.

The white color is the result of an approximately equal stimulation of the S, M and L cones. We know that the grey axis is composed of the colors which have the three variables identical, i.e. $R = G = B$. Then, we can conclude that the greys are not colors, or that they are colors without hue, since the white and the grey are just differentiated by their intensity level.

2 Histogram processing for image enhancement

In this section we will expose some techniques for color image enhancement. We will consider the color images in the RGB space. In this space, the color image can be considered as three gray level images, the three color

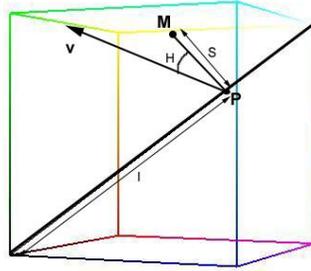


Figure 8: Representation of the HSI space.

channels, red, green and blue. All the techniques we will explain in this section are applied to each color channel separately. Then we will define the concepts for a gray level image.

One of the simplest and most useful tools in digital image processing is the histogram. The histogram of a digital image of dimensions $n_x \times n_y$, with L total possible levels in the range $[min, max]$ is defined as the discrete function

$$h(l_k) = n_k,$$

where l_k is the k th intensity level in the interval $[min, max]$ and n_k is the number of pixels in the image whose intensity level is l_k . Generally the values of min and max are 0 and 255 respectively.

Often, it is useful to work with normalized histograms, obtained simply by dividing all elements $h(l_k)$ by the total number of pixels in the image which we denote by $N = n_x \times n_y$:

$$p(l_k) = \frac{h(l_k)}{N} = \frac{n_k}{N}, \quad k = 0, 1, \dots, L - 1.$$

From basic probability, $p(l_k)$ is an estimate of the probability of occurrence of intensity level l_k .

From the histogram we can obtain the following statistical properties:

Mean : is the average of all the intensity levels and informs us of the general image brightness, is defined by

$$\bar{l} = \sum_{k=0}^{L-1} l_k p(l_k)$$

Variance: measure the dispersion from the mean, is defined by

$$\sigma^2 = \sum_{k=0}^{L-1} (l_k - \bar{l})^2 p(l_k)$$

A large variance corresponds to high contrast image and vice versa.

Histograms are the basis for numerous spatial domain processing techniques, since the histogram provides a global description of the appearance of an image. For example, Figures 9-12, show the histogram of four basic types of images. Figure 9 shows a histogram where the intensity levels are concentrated at the left,

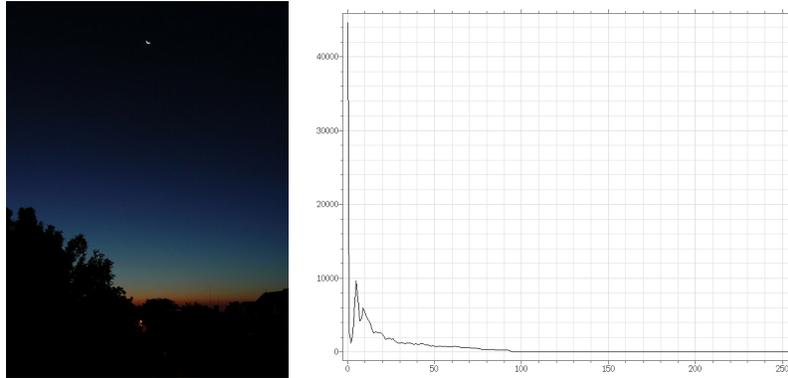


Figure 9: An example of the histogram of a dark image.

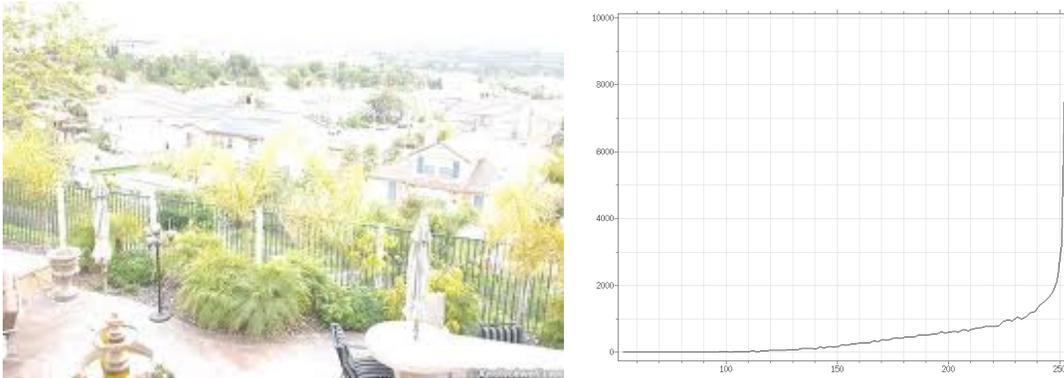


Figure 10: An example of the histogram of a bright image.

corresponding to a dark image. In Figure 10 is just the opposite, the histogram is concentrated at the right, corresponding to a bright image. The histogram shown in Figure 11 has a narrow shape which indicates little dynamic range and corresponding to an image having low contrast. All the intensity levels occur toward the middle of the gray scale, corresponding to a greyish image. Finally, Figure 12 shows a histogram with significant spread, corresponding to an image with high contrast.

The shape of the histogram of an image gives us useful information about the possibility for contrast enhancement. In the next subsections we present methods for manipulating histograms obtaining a more contrasted image.

The histogram can be modified by means of some functions which expand the intensity gray levels, compress or translate them [6]. We begin the study of image enhancement techniques operating directly on the pixels of an image. Image processing functions may be expressed as

$$g(x, y) = T[f(x, y)]$$

where $f(x, y)$ is the input image, $g(x, y)$ is the processed image, and T is a monotone function. The transformation T can be applied to the gray level intensity or in each color channel. Here we present one example of

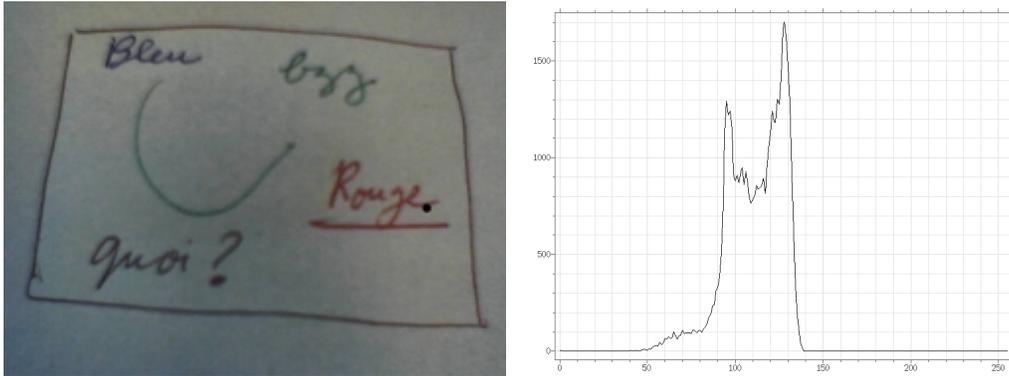


Figure 11: An example of the histogram of an image with a little dynamic range.

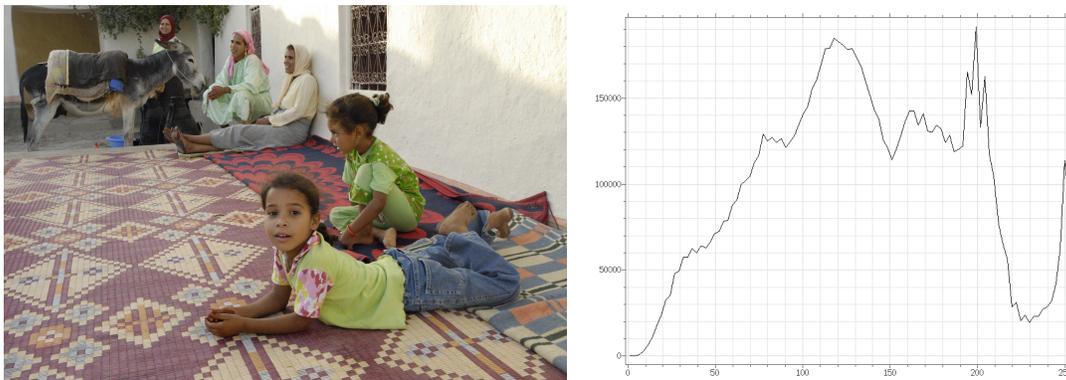


Figure 12: An example of the histogram of a well contrasted image.

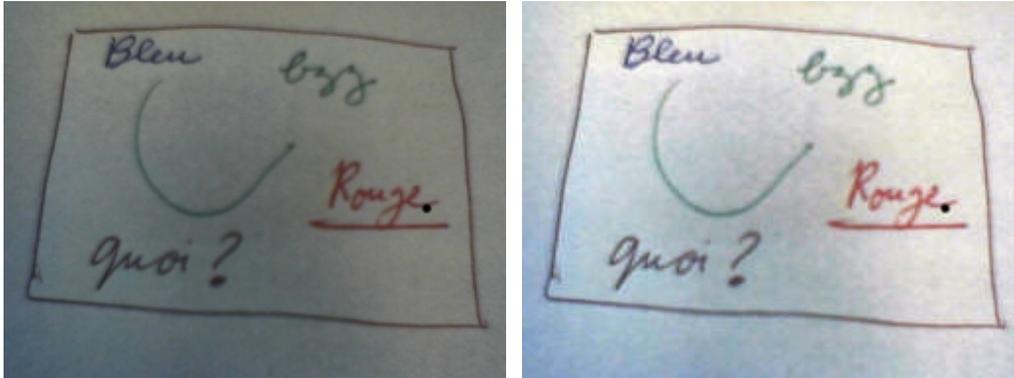


Figure 13: An example of stretching at the maximum the histogram of the three color channels.

transformation, which is the more simplest since it is a linear transformation applied to each color channel.

2.1 Linear transformation. Simplest color balance

It is the simplest way to stretch or to contract the histogram of an image and it is defined by

$$T(l) = \frac{s_{max} - s_{min}}{l_{max} - l_{min}}(l - l_{min}) + s_{min} \quad (1)$$

where l_{max} and l_{min} are the maximum and minimum values of the intensity levels in the image, s_{max} and s_{min} are the desired maximum and minimum values in the histogram. If the ratio

$$\frac{s_{max} - s_{min}}{l_{max} - l_{min}}$$

is smaller than 1 this transformation contract the histogram, if the ratio is larger than 1 the transformation stretch the histogram and if $s_{max} = 255$ and $s_{min} = 0$ the transformation stretches, as much as it can, the histogram.

An special case of this kind of linear transformation is what we call **Simplest color balance**. This algorithm attempt to correct underexposed images, or images taken in artificial lights or special natural lights, such as sunset.

There are many sophisticated algorithms in the literature performing color balance or other color contrast adjustments. The performance of these many color correction algorithms can be evaluated by comparing their result to the simplest possible color balance algorithm proposed here. The assumption underlying this algorithm is that the highest values of R, G, B observed in the image must correspond to white, and the lowest values to obscurity. If the photograph is taken in darkness, the highest values can be significantly smaller than 255. By stretching the color scales, the image becomes brighter. Figure 13 shows the original image, whose highest value is and the result of stretching the three color channels to $[0, 255]$ using Equation (1).

If there was a colored ambient light, for example electric light where R and G dominate, the color balance will enhance the B channel. Thus the ambient light will lose its yellowish hue. Although it does not necessarily improves the image, the simplest color balance always increases its readability. Figure 14 shows an example of a sunset image and the histogram of the three color channels. Observe that the blue channel is less contrasted

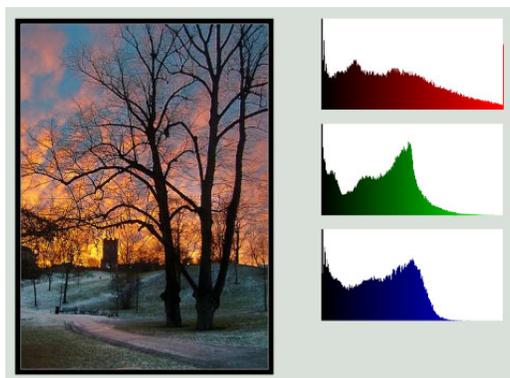


Figure 14: An example of a sunset image and the histogram of each channel.

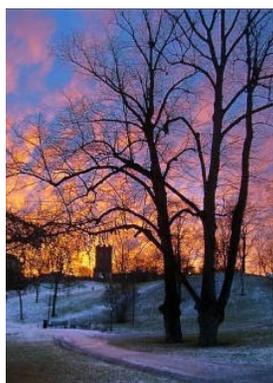


Figure 15: The result of stretching the values to $[0,255]$.

than the others. Figure 15 shows the result of stretching, as much as it can, the values of the three channels so that they occupy the maximal possible range $[0, 255]$.

However, many images contain a few aberrant pixels that already occupy the 0 and 255 values. Thus, an often spectacular image color improvement is obtained by “clipping” a small percentage of the pixels with the highest values to 255 and a small percentage of the pixels with the lowest values to 0, before applying the affine transform (1). Notice that this saturation can create flat white regions or flat black regions that may look unnatural. Thus, the percentage of saturated pixels must be as small as possible.

The proposed algorithm computes the minimum and maximum, l_{min} and l_{max} in Equation (1), such that the number of pixels with values less than l_{min} plus the number of pixels with values greater than l_{max} are some percentage of the total number of pixels. Then, using this minimum and maximum, stretches the intensity values at the maximal possible, $s_{min} = 0$ and $s_{max} = 255$.

Given an image and given the saturation level $s \in [0, 1)$, for example, $s = 0$ means 0 “per cent” saturation, $s = 0.03$ means 3 “per cent” saturation. The algorithm will saturate at most $N \times s$ pixels, half at the beginning and half at the end of the histogram. Now we explain how we compute the values l_{min} and l_{max} .

A efficient implementation is achieved by an histogram-based variant, the steps of the algorithm are:

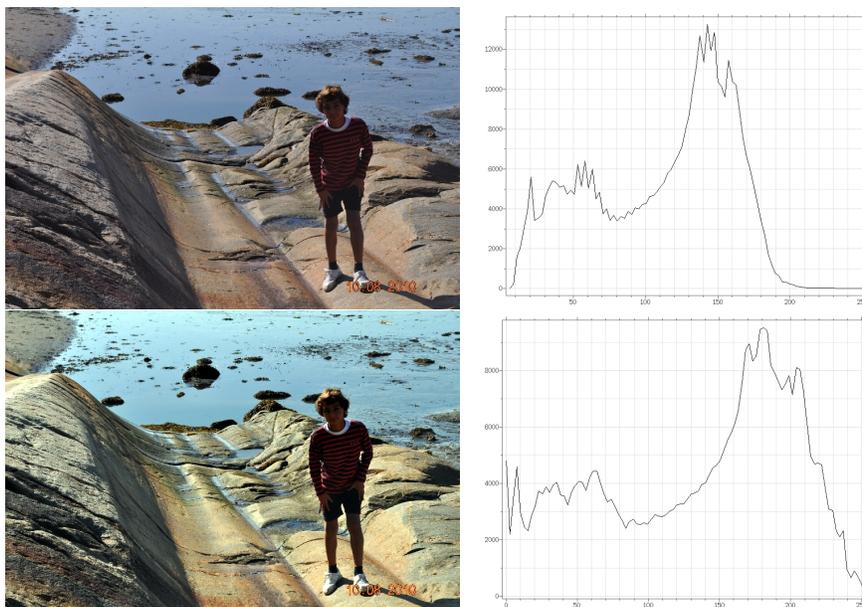


Figure 16: Top Left: Original image. Top right: Histogram of the original image. Bottom Left: Result with 2 of saturation. Bottom right: Histogram of the result.

1. Build a cumulative histogram of the pixel values.

$$H(l_k) = \sum_{j=0}^k h(l_j).$$

The cumulative histogram bucket labeled l_k contains the number of pixels with value lower or equal to l_k .

2. l_{min} is the lowest value such that $H(l_{min}) \geq N \times s/2$, the number of pixels with values lower than l_{min} is at most $N \times s/2$. If $s = 0$ then l_{min} is the lowest histogram label, i.e. the minimum pixel value of the input image. l_{max} is the highest value such that $H(l_{max}) \leq N \times (1 - s/2)$, the number of pixels with values higher than l_{max} is at most $N \times s/2$. If $s = 0$ then l_{max} is the highest histogram label, i.e. the maximum pixel value of the input image.
3. Saturate the pixels.
4. Affine transform (1) with $s_{min} = 0$ and $s_{max} = 255$.

The proposed algorithm provides both a white balance and a contrast increasing. However, note that this algorithm is not a real physical white balance: It won't correct the color distortions of the capture device or restore the colors or the real-world scene captured as a photography. Such corrections would require a captured sample of known real-world colors or a model of the lighting conditions.

Now we present some examples of the **Simplest color balance algorithm**. Figure 16 shows an image with its histogram, the result after saturated a 2 “per cent” and the result histogram.



Figure 17: Left: Original image. Center: Result with a 0 of saturation. Right: Result with 2 of saturation.



Figure 18: Left: Original image. Center: Result with a 0 of saturation. Right: Result with 3 of saturation.

Figure 17 shows an image in completely unnatural blue light, this kind of images are often used to illustrate color balance, or color contrast adjustment algorithms. A trivial affine transform corrects it adequately by removing the bluish effect. A still more contrasted result is obtained by saturating only 2 “per cent”. Another example with a little improvement with a 0 of saturation is shown in Figure 18. This little improvement is due to the little lights in the bottom of the image. Using a 3 of saturation, we obtain a brighter and contrasted image.

2.2 Insufficiency of histogram processing

Here we show examples where the Simplest color balance is insufficient or not adequate. For example for images with a back-lighting the algorithm proposed produce little improvements. Figure 19 is an example. For images like a sunset the algorithm can produce very unnatural colors. Figure 20 shows the sunset image and the result after 3 per cent of saturation, the orange pixels decrease and a completely unnatural blue color is created.



Figure 19: Left: Original image. Right: Result with 3 of saturation.



Figure 20: Left: Original image. Right: Result with 3 of saturation.

3 Modifications in the gradient domain

3.1 Poisson image editing

The concept of image editing encompasses the operations by which the local content of one or several images is selected and manipulated to create new syncretic images. The simplest such operation is a copy-paste of a part of an image into another. Image stitching, by which several images are fused into a panorama is another example. Local contrast or color adjustments after selecting manually a part of the image (for example the shadows) is another variant. The inpainting operation, by which an object is removed in the image and replaced by a texture is also a classic editing operation. For many examples of such image editing problems and surveys of the techniques we refer to [12, 16, 2].

Most image editing operations modify an input image I by manually selecting a region Ω in it. The task is to fill in this region by picking information from the rest of the image $I \setminus \Omega$, or from another image or, in the case of a selective contrast change, from the image in Ω itself. The most recent algorithms to solve this problem are based on partial differential equations. Two types of PDE, parabolic equations and elliptic equations, are the most widely used. Image inpainting [1] is based on parabolic equations. Matting operations [14] are based on the Poisson equation. The main challenge in image editing is to avoid suspicious color or texture alterations which would reveal the silhouette $\partial\Omega$ of the image region which has been modified. As amply demonstrated by [12], the Poisson equation (2) is an extremely efficient response to this problem.

$$\begin{cases} \Delta u(x) = f(x) & \text{if } x \in \Omega \\ u(x) = g(x) & \text{if } x \in \partial\Omega \end{cases}, \quad (2)$$

where $\Delta u(x, y) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$ is the Laplacian of u .

In 2003 the authors of [12] proposed to use (2) for “seamless editing”. The use of this mathematic tool is motivated by the fact that psychophysics observations suggest that our vision perceives the Laplacian of the images instead of the image itself. Thus the image u is satisfactorily specified by its Laplacian. On the other hand knowing the Laplacian inside the domain and the image value on the boundary is enough to reconstruct the image, and this is done by the Poisson equation. This work have been widely cited and various improvements have been presented [10, 7, 4, 14].

We detail the image Poisson reconstruction from “a guidance vector field”, in the terminology of Perez et al. [12]. Let R , a closed subset of \mathbb{R}^2 , be the image domain, and let $\Omega \subset R$ be a closed subset with boundary $\partial\Omega$. The problem is to find the image whose gradient field is the closest, in L_2 -norm, to the prescribed “guidance vector field” \mathbf{v} defined in Ω , under given boundary conditions on $\partial\Omega$. The image is known in the rest of the image, $R \setminus \Omega$ (Figure 21). This problem writes

$$\min_u \int_{\Omega} |\nabla u - \mathbf{v}|^2, \quad \text{with } u|_{\partial\Omega} = f|_{\partial\Omega}, \quad (3)$$

and its solution is the unique solution of the Poisson equation with Dirichlet boundary conditions

$$\Delta u = \text{div } \mathbf{v}, \quad \text{with } u|_{\partial\Omega} = f|_{\partial\Omega}. \quad (4)$$

The variational problem (3) can be discretized using a finite difference approach. Using the notation of Perez et al. [12], it amounts to solve the following linear equations

$$|N_p|u_p - \sum_{q \in N_p \cap \Omega} u_q = \sum_{q \in N_p \cap \partial\Omega} f_q + \sum_{q \in N_p} v_{pq}, \quad \text{for all } p \in \Omega, \quad (5)$$

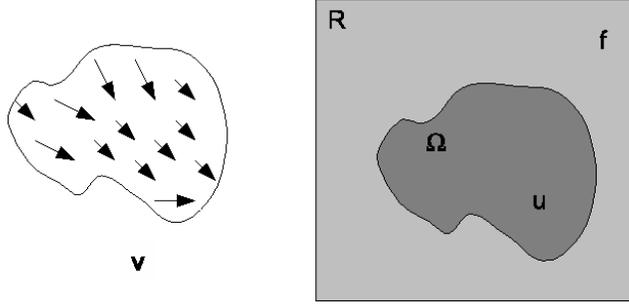


Figure 21: The unknown function u interpolates f in domain Ω under a guidance vector field \mathbf{v} .

where N_p is the 4-connected neighbors of p and $v_{pq} = \mathbf{v} \left(\frac{p+q}{2} \right) \cdot \overrightarrow{pq}$. Since the region Ω can have an arbitrary shape, the authors in [12] must solve the linear system (5) using iterative solvers, for example the Gauss-Seidel iteration with overrelaxation or a multigrid method.

The basic choice for the guidance vector field \mathbf{v} is a gradient field taken directly from a source image. This type of implementation is as the copy-paste operation. The seamless cloning tool thus obtained ensures the compliance of source and destination boundaries. It can be used to conceal undesirable image features or to insert new elements in an image, but with much more flexibility and ease than with conventional cloning, as illustrated in Figure 22. Denoting by g the source image, the interpolation is performed under the guidance of

$$v = \nabla g;$$

and (4) now reads

$$\begin{cases} -\Delta u(x) = -\Delta g(x) & x \in \Omega \\ u(x) = f(x) & x \in \partial\Omega \end{cases}$$

In this case, the discrete equation (5) is

$$|N_p|u_p - \sum_{q \in N_p} u_q = \sum_{q \in N_p} v_{pq}, \quad \text{for all } p \in \Omega, \text{ where}$$

$$v_{pq} = g_p - g_q, \text{ then } \sum_{q \in N_p} v_{pq} = |N_p|g_p - \sum_{q \in N_p} g_q$$

With the tool described previously, no trace of the destination image f is kept. However, there are situations where it is desirable to combine properties of f with those of source image g , for example to add objects with holes, or partially transparent ones, on top of a textured or cluttered background. An example is shown in Figure 23, in which a text layer is to be peeled off the source image and applied to the destination image, without the need for complex selection operations. One possible approach is to define the guidance vector field \mathbf{v} as a linear combination of source and destination gradient fields but this has the effect of washing out the textures. However, the Poisson methodology allows non-conservative guidance fields to be used, which gives scope to more compelling effect. In this case, the selection of the vector guidance field is:

$$\mathbf{v} = \begin{cases} \nabla f & \text{if } |\nabla f| > |\nabla g| \\ \nabla g & \text{otherwise} \end{cases}$$



Figure 22: We want to copy the region of the source image into the destination image. With a simple copy-paste, the boundary is visible, with the Poisson image editing a seamless cloning is obtained.

And, in this case, (5) is equivalent to

$$|N_p|u_p - \sum_{q \in N_p} u_q = \sum_{q \in N_p} v_{pq}, \quad \text{for all } p \in \Omega, \text{ where}$$

$$v_{pq} = \begin{cases} f_p - f_q & \text{if } |f_p - f_q| > |g_p - g_q| \\ g_p - g_q & \text{otherwise} \end{cases}$$

The second member of the discrete equation is written

$$\sum_{q \in N_p} F(f_p - f_q, g_p - g_q), \quad p \in \Omega, \quad F \text{ is an odd function defined by}$$

$$F(s, t) = \begin{cases} s & \text{if } |s| > |t| \\ t & \text{otherwise} \end{cases}$$

3.2 Local contrast adjustment

Perez et al. use the method of Fattal et al. [5] to modify smoothly the image dynamic range. The idea is to transform the gradient field of the logarithm of the image to reduce the large gradients and to increase the small ones. The transformed gradient is used to reconstruct the logarithm of the image, by solving the Poisson equation. In fact in [12] the authors select a region Ω , for example an under-exposed region, and transform the gradient in the log-domain by

$$\mathbf{v} = \alpha^\beta |\nabla f|^{-\beta} \nabla f, \quad (6)$$

where $\alpha = 0.2$ times the average gradient norm of f over Ω and $\beta = 0.2$, that is, the guidance vector field is a concave function of the gradient in the log-domain.



Figure 23: Left: Text image. Middle: Texture. Right: Result of the mix.

The Perez et al. method is relatively complex. Indeed, it depends on two parameters, α , β . The selection of the dark region is manual and only permits to select a few simple image regions. Thus, it is not adapted to treat the dark regions of images with back light or excessively dark shadows. Indeed, in such images the dark regions can be many and have a complicated topology. Thus, they can hardly be selected manually. The alternative proposed here is to select the dark region using a threshold T . $T = 50$ seems to be the right threshold and was the default value in all treated examples. Over the dark region the gradient vector is amplified by a factor α . Since the a stronger factor would amplify too much the noise, again in most examples the default value was 2.5. With the default values for T , α the proposed algorithm is fully automatic:

- Select the dark region Ω by the threshold T .
- Define the guidance vector field by

$$\mathbf{V} = \begin{cases} \nabla f & \text{in } R \setminus \Omega \\ \alpha \nabla f & \text{in } \Omega \end{cases} \quad (7)$$

- Solve the Poisson equation with Neumann boundary conditions using the Fourier transform as explained in a posterior section.

Figures 24 and 25 show results obtained using the proposed algorithm. Note that the complexity of the dark regions.



Figure 24: Top Left: original image. Top Right: Zones with intensity below $T = 50$. Bottom: Result using guidance vector field (7) with $\alpha = 2.5$.



Figure 25: Top Left: original image. Top Right: Zones with intensity below $T = 50$. Bottom: Result using guidance vector field (7) with $\alpha = 2.5$.

3.3 Retinex

In 1964 Edwin H. Land formulated the Retinex theory, the first attempt to simulate and explain how the human visual system perceives color. His theory and an extension, the “reset Retinex” were further formalized by Land and McCann in 1971. Several Retinex algorithms have been developed ever since. These color constancy algorithms modify the RGB values at each pixel to give an estimate of the physical color independent of the shading.

The Retinex original algorithm was both complex and not fully specified. Indeed, this algorithm computes at each pixel an average of a very large and unspecified set of paths on the image. For this reason, Retinex has received several interpretations and implementations which, among other aims, attempt to tune down its excessive complexity.

But the original Retinex algorithm can be formalized as a (discrete) partial differential equation. More precisely, it can be shown (see [11]) that if the Retinex paths are interpreted as symmetric random walks, then Retinex is equivalent to a Neumann problem for a Poisson equation. This result gives a fast algorithm involving just one parameter, also present in the original theory.

The Retinex Poisson equation (given below) is very similar to Horn’s and Blake’s equations, which were proposed as alternatives to Retinex. It also is one of the “Poisson editing” equations proposed in Perez et al. The final principle of the algorithm is extremely simple. Given a color image I , its small gradients (those with magnitude lower than a threshold t) in each channel are replaced by zero. The resulting vector field is no more the gradient of a function, but the Poisson equation reconstructs an image whose gradient is closest to this vector field. Thus, a new image is obtained, where small details and shades of the original have been

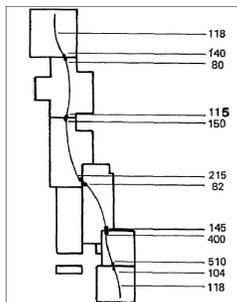


Figure 26: The original Land-McCann scheme without the reset mechanism: $\frac{140}{80} \times \frac{115}{150} \times \frac{215}{82} \times \frac{145}{400} \times \frac{510}{104} = \frac{6.25}{1}$. “Luminance of Mondrian (illuminated from below) at particular points along the path from top to bottom. The numbers at the bottom indicate the ratios of luminance at adjacent edges along the path.”

eliminated. The elimination of the shades creates more homogeneous colors. This fact, according to Land and McCann, models the property of our perception to perceive constant colors regardless of their shading.

3.3.1 Original Retinex algorithm

The basic Retinex model is based on the assumption that the HVS operates with three retinal-cortical systems, each one processing independently the low, middle and high frequencies of the visible electromagnetic spectrum. Each system produces one lightness value which determines, by superposition, the perception of color in the HVS. On digital RGB images, the lightness is represented by the triplet (L_R, L_G, L_B) of lightness values in the three chromatic channels.

Inspired by several experiments, Land and McCann observed that edges are the main image feature invariant to illumination, and therefore the main source of information to achieve color constancy ([9]). They also realized that a luminance ratio threshold between two adjacent points maintains the edge if there is one between those points, but eliminates the gentle slopes caused by nonuniform illumination. Thus, obtaining the lightness values boils down to processing the entire image in terms of luminance ratios. In the case of two widely separated areas in the image, they therefore consider the sequential product of ratios across edges on a path joining both areas (see Fig. 26). Since this generalized lightness ratio would then depend upon the chosen path, the Retinex algorithm considers all possible paths starting at random points and ending at the pixel at which the lightness value is computed. This lightness is then defined as the average of the products of ratios between the intensity values of subsequent edge points in the path. In order to remove the effects of nonuniform illumination over the scene, the ratio is considered unitary if it does not differ from 1 by more than a fixed threshold value.

The formula giving the lightness value L of a pixel $x = (i, j)$ computed by Retinex in a given chromatic channel was proposed in [8], but the process itself is described in [9]. The image data $I(x)$ is the intensity value for each chromatic channel at x . Land and McCann consider a collection of N paths $\gamma_1, \dots, \gamma_k, \dots, \gamma_N$ starting at x and ending at an arbitrary image pixel y_k . Let n_k be the number of pixels of the path γ_k , and denote by $x_{t_k} = \gamma_k(t_k)$ for $t_k = 1, \dots, n_k$ and by $x_{t_k+1} = \gamma_k(t_k + 1)$ the subsequent pixel of the path so that $\gamma_k(1) = x$ and $\gamma_k(n_k) = y_k$.

Definition 1 The **lightness** $L(x)$ of a pixel x in a given chromatic channel is the average of the relative lightness at x over all paths, that is

$$L(x) = \frac{\sum_{k=1}^N L(x; y_k)}{N}, \quad (8)$$

where $L(x; y_k)$ denotes the **relative lightness** of a pixel x with respect to y_k on the path γ_k defined by

$$L(x; y_k) = \sum_{t_k=1}^{n_k} \delta \left[\log \frac{I(x_{t_k})}{I(x_{t_k+1})} \right], \quad (9)$$

and, for a fixed contrast threshold t ,

$$\delta(s) = \begin{cases} s & \text{if } |s| > t \\ 0 & \text{if } |s| < t. \end{cases} \quad (10)$$

The reset mechanism proposes an adaptation of the above definition to ensure that all paths start from regions where the maximal luminance value is attained. We quote from [9]:

One that seems simple, but is not, is to scan the entire scene to find the area or areas with the highest reflectance. (...) Although this technique is mathematically valid we feel that it is not readily transposed into biological mechanisms. We therefore sought a technique that can automatically establish the highest reflectance without a separate first scanning step. We adopted the convention that the initial ratio is the ratio of the signal of the second receptor to that of the first (followed by the third to the second, etc.). Then, regardless of the true reflectance of an area, our technique supposes that the first receptor in any path is reporting from an area that reflects 100% of the light. (...) Attainment of a sequential product greater than 1.0 indicates that the sequence should be started afresh and that this new, higher reflectance should be next supposed to be 100%. (...) As the path proceeds, the sequential product always starts over at unity when the path encounters an area with a reflectance higher than the highest previously encountered.

Notice that the above quotation uses products of ratios (see Fig. 26). Taking the logarithm of this product yields a sum of logarithms of ratios. In the reset formulation of Retinex, the average giving $L(x)$ is taken over paths on which all partial sums leading to the complete sum $L(x; y_k)$ must be non-positive:

$$\forall j = 1, \dots, n_k - 1, \sum_{t_k=j}^{n_k} \delta \left[\log \frac{I(x_{t_k})}{I(x_{t_k+1})} \right] \leq 0. \quad (11)$$

The value of the reset Retinex solution at a pixel depends on the memory of each single path. To the best of our knowledge, this fact rules out any PDE formalization for reset Retinex. As shown by the above Land and McCann quotation, the main goal of the reset mechanism was to ensure that all paths starting from x end at points y_k that are image extrema. This goal is not achieved by reset Retinex. The reset mechanism only selects paths along which *there is no value larger than the initial value*. This observation justifies defining an “*Extrema Retinex*”, namely a variant where all paths only start from image extrema. Extrema Retinex is an easy adaptation of Def. 1:

Definition 2 (*Extrema Retinex*) The **lightness** $L(x)$ of a pixel x in a given chromatic channel is the average of the relative lightness at x over all paths linking x to an arbitrary image extremum y_k , the path meeting no other extremum before reaching y_k . We therefore have $L(x) = \frac{\sum_{k=1}^N L(x; y_k)}{N}$ where $L(x; y_k)$ is given by (9).

3.3.2 A PDE Formalization of Retinex Theory

It can be proven that the output of the Retinex algorithm proposed by Land and McCann is the only solution of a discrete partial differential equation with Neumann boundary conditions.

Theorem 1 *The lightness value in a chromatic channel L defined in Def. 1 is the unique solution of the discrete Poisson equation*

$$\begin{cases} -\Delta_d L(x) = F(x) & x \in R \\ \frac{\partial L}{\partial n} = 0 & x \in \partial R \end{cases} \quad (12)$$

where $\Delta_d L$ denotes the discrete Laplace operator $\Delta_d L(i, j) =: L(i + 1, j) + L(i - 1, j) + L(i, j + 1) + L(i, j - 1) - 4L(i, j)$ and

$$\begin{aligned} F(x) = & f\left(\frac{I(x)}{I(x_{-0})}\right) + f\left(\frac{I(x)}{I(x_{+0})}\right) + \\ & + f\left(\frac{I(x)}{I(x_{0-})}\right) + f\left(\frac{I(x)}{I(x_{0+})}\right), \end{aligned} \quad (13)$$

where $f(x) = \delta(\log(x))$, δ has been defined in (10), $x_{-0} = (i - 1, j)$, $x_{0-} = (i, j - 1)$, $x_{+0} = (i + 1, j)$, and $x_{0+} = (i, j + 1)$.

A straightforward adaptation of the proof of Theorem 1 gives, with the same notation, the Retinex equation in the *Extrema* case:

Corollary 1 *Let \mathbf{Y} be the set of image maxima. The Extrema Retinex lightness value in a chromatic channel L defined in Def. 1 is the unique (M, N) symmetric and periodic solution of the discrete Poisson equation*

$$\begin{cases} -\Delta_d L(x) = F(x) & x \notin \mathbf{Y} \\ L(x) = 0 & x \in \mathbf{Y} \end{cases} \quad (14)$$

For the proof of the Theorem see [11].

Remark 1 *The equation obtained in Theorem 1 is very similar, to the Poisson editing equation proposed in Perez et al. [12]. These authors propose a texture flattening application, whose goal it is to wash out the texture and keep only the edges. Using the same notation as in the previous section they consider a guidance vector field*

$$\mathbf{v} = \begin{cases} \nabla I & \text{if there is an edge} \\ 0 & \text{in other case} \end{cases},$$

but they do not specify which kind of edge detector they use. Using as edge detector a threshold functions of the gradient like the function δ (10) applied to the gradient, then the texture flattening of Perez et al is equal to the Retinex equation and to Blake's equation [3].

Land's Retinex theory was postulated as a perception model and attempted to explain the HVS and in particular classic color perception illusions. In optical illusions the information gathered by the eye is processed in the brain to give a perception that does not tally with a physical measurement of the stimulus source. Applying the Retinex algorithm to an illusory image, it is expected that the result will be an image showing the same tendencies in the alteration of colors as the HVS.

As a first classic example, Fig. 27 shows Adelson's checker shadow illusion. In the left image a green cylinder standing on a black and white checker-board casts a diagonal shadow across the board. The image has been

so constructed that the white squares in the shadow, one of which is labeled "B," have actually the very same gray value as the black squares outside the shadow, one of which is labeled "A." The classic illusion is that the squares A and B appear respectively black and white, despite the fact that they share the same gray value, (88). If Retinex is faithful to human perception, it should make B much brighter than A. The image on the right shows the result of applying the PDE Retinex algorithm. The gray value in square A is now 75 and in square B it is 100, making the square A effectively darker than the square B, in agreement with our perception. The Retinex primary goal was to simulate our perceptual color illusions. Thus the criterion here is whether the perceptual tendencies are adequately simulated and not by any image improvement.

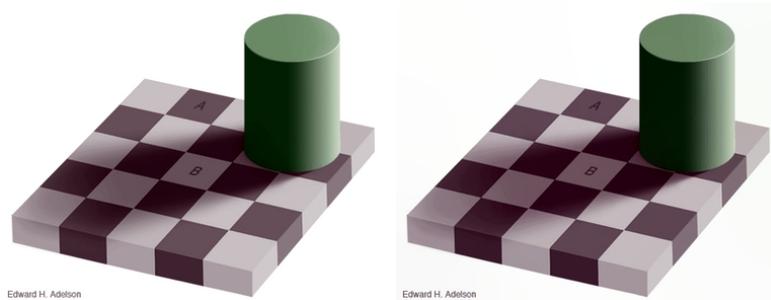


Figure 27: The Adelson’s checker shadow illusion and the same image applying the Retinex algorithm with $t = 3$.

Our next illusion is about simultaneous contrast, namely the fact that the appearance of a color depends on the colors surrounding it. Fig. 28 shows a background with a smooth variation and two circles with the same gray value (170). One of them is placed in the darker part of the image, and the other one in the brighter part. The usual perception is that the circle in the darker part looks conspicuously brighter than the other. Again, this illusion is so strong that it needs masking the whole image by a white sheet, excepting only both disks, to check that they indeed share the same luminance. If we use a threshold t larger than the background variation, in this case we take $t = 3$, the result is an image with constant background (130). The left circle gets a 105 gray value and the right circle a 235 gray value, which is coherent with our perception.

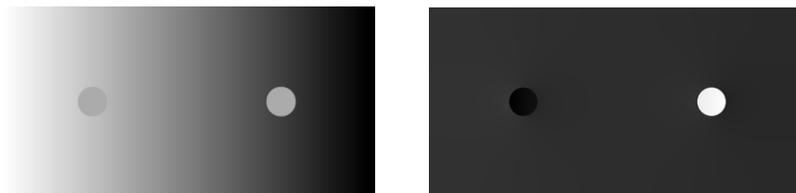


Figure 28: Simultaneous contrast illusion. Left: the original image. Center: Retinex result with $t = 3$. By Retinex the background slope is eliminated and therefore the shape-background contrast enhanced. In the left image the two disks have exactly the same grey level. Retinex is an attempt to formalize the process by which the right disk appears conspicuously brighter than the left one. After applying Retinex, this disk indeed *becomes* brighter.

To understand the effect of the threshold in Retinex Fig. 29 shows a noisy original and the result of Retinex with increasing threshold values $t = 1, 3,$ and 5 . The background clutter and the shades are progressively filtered out when t increases, but the main edges are kept. At $t = 5$, however, edges start losing contrast and low contrasted details could disappear.

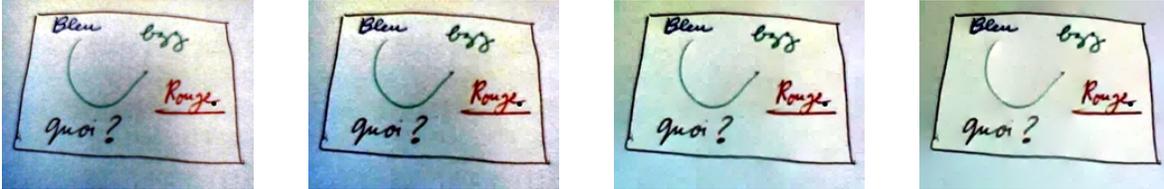


Figure 29: Left: Original image. Center left: Retinex with $t = 1$. Center right: Retinex with $t = 3$. Right: Retinex with $t = 5$. Observe how colors are slightly enhanced and the background clutter progressively eliminated when the threshold t grows.

A final example, Figure 30, demonstrates how Retinex can be used for removing shadows. It is not always effective, but here we can observe a good example. The shadow removal works only if the boundary of the shadow is blurry, and therefore has a small gradient.



Figure 30: Left: Original image. Center: Balanced image. Right: Retinex result. Observe how shadows are eliminated.

4 Implementation of Poisson Equation using Fourier Transform

The aim of this section is to solve the above variational problem (3), and, consequently the associated Dirichlet problem (4), using the Fourier method. To do that the boundary of the domain Ω must coincide with the coordinate lines, which in general is not the case. This can be done by extending the definition domain of the guidance vector field to the whole image domain R , and solving the variational problem over all the whole image domain. The guidance vector field \mathbf{v} defined on Ω is extended to R by

$$\mathbf{V} = \begin{cases} \mathbf{v} & \text{over } \Omega \\ \nabla f & \text{otherwise} \end{cases}, \tag{15}$$

where f be the image, which is known over $R \setminus \Omega$. Then the problem is to minimize

$$\min_u \int_R |\nabla u - \mathbf{V}|^2. \tag{16}$$

The minimizer is uniquely determined by the Euler-Lagrange equation

$$\Delta u = \operatorname{div} \mathbf{V}, \quad \text{over } R, \quad (17)$$

with homogeneous Neumann boundary conditions

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \quad \text{over } \partial R, \quad (18)$$

where \mathbf{n} is the direction orthogonal to the boundary.

Since equation (17) has constant coefficients and the boundaries of R coincide with the coordinate lines, we can apply the Fourier transform method to solve the partial differential equation. The Neumann boundary condition is implicitly imposed by extending the original image symmetrically across its sides, so that the extended image, which is four times bigger, becomes symmetric and periodic. Once this is done, we can apply the Fourier transform. The discrete Fourier transform (DFT) permits to compute directly the Fourier coefficients of a band limited and periodic function u from its samples u_{jl} on a $J \times L$ grid by

$$\widehat{u}_{mn} = \sum_{j=0}^{J-1} \sum_{l=0}^{L-1} u_{jl} e^{-i \frac{2\pi jm}{J}} e^{-i \frac{2\pi ln}{L}} \quad (19)$$

for $m = 0, \dots, J-1$ and $n = 0, \dots, L-1$. The Fourier series formula recovers u from its Fourier coefficients by

$$u_{jl} = \frac{1}{JL} \sum_{m=0}^{J-1} \sum_{n=0}^{L-1} \widehat{u}_{mn} e^{i \frac{2\pi jm}{J}} e^{i \frac{2\pi ln}{L}}. \quad (20)$$

Thus, by a simple differentiation, the Poisson equation $\Delta u = \operatorname{div}(\mathbf{V})$ translates into a relationship between the Fourier coefficients of u and \mathbf{V} :

$$\left(\left(\frac{2\pi m}{J} \right)^2 + \left(\frac{2\pi n}{L} \right)^2 \right) \widehat{u}_{mn} = \frac{2\pi im}{J} \widehat{V}_{1mn} + \frac{2\pi in}{L} \widehat{V}_{2mn}, \quad (21)$$

where $\mathbf{V} = (V_1, V_2)$.

Thus the strategy for solving (17-18) by FFT techniques is:

- Quadruplicate by symmetry the discrete domain and \mathbf{V} .
- Compute the discrete Fourier transforms of V_1 and V_2 .
- Compute the discrete Fourier transform of the solution \widehat{u}_{mn} as

$$\widehat{u}_{mn} = \frac{\frac{2\pi im}{J} \widehat{V}_{1mn} + \frac{2\pi in}{L} \widehat{V}_{2mn}}{\left(\frac{2\pi m}{J} \right)^2 + \left(\frac{2\pi n}{L} \right)^2}. \quad (22)$$

- Obtain the samples u_{jl} of the solution by the inverse discrete Fourier transform.
- Restrict them to the initial domain.

4.1 Computation of the vector field \mathbf{V}

The guidance vector field \mathbf{V} defined in the problem of Poisson Image Editing depends on the gradient of the original image or on the gradient of some source image. To compute the guidance vector field, multigrid methods use finite differences approximating the gradient of the image. Actually the Fourier method permits exact expressions for all derivatives. If that the guidance vector field is the gradient of some function f , $\mathbf{V} = \nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$, then

$$\widehat{V}_{1mn} = -\frac{2\pi im}{J} \widehat{f}_{mn} \quad \widehat{V}_{2mn} = -\frac{2\pi in}{L} \widehat{f}_{mn} \quad (23)$$

To solve (21) for each choice of the guidance vector field it is therefore necessary to compute the Fourier transform of the images whose gradients compose the vector field.

For example:

- In the classic copy-paste problem, the guidance vector field \mathbf{v} is a gradient field from another source image g . The guidance field therefore is

$$\mathbf{V} = \begin{cases} \nabla g & \text{over } \Omega \\ \nabla f & \text{otherwise} \end{cases}, \quad (24)$$

where f be the image which will be kept on $R \setminus \Omega$.

- In the local contrast adjustment the guidance vector field is defines in (7).
- In the Retinex the guidance vector field is

$$\mathbf{V} = \begin{cases} \nabla f & \text{if } |\nabla f| > T \\ 0 & \text{in other case} \end{cases} \quad (25)$$

5 Projects

The explained methods in this course are global methods. The aim of these projects is propose two kind of local methods for image enhancement. The first one must be implemented in for the IPOL web page. The second one is a modelization problem.

5.1 Local contrast enhancement by “center-surround” filters

The aim of this project is to stretch locally the histogram from three measures, the mid point of the local dynamic range, the mean intensity value or the median intensity value.

Given an image I defined on a rectangle R , for each point $x \in R$, consider a x -neighborhood of radius r , $B_r(x)$. The algorithm consists on:

- Compute $m(x) = \min_{B_r(x)} I(y)$ and $M(x) = \max_{B_r(x)} I(y)$.
- Compute the values:

$$mp(x) = \frac{m(x) + M(x)}{2} \quad \text{the mid point of the local dynamic range} \quad (26)$$

$$mean(x) = (I * \chi_{B_r(x)}) \frac{1}{\pi r^2} \quad \text{the mean intensity value on } B_r(x) \quad (27)$$

$$med(x) = median_{B_r(x)} I(y) \quad \text{the median intensity value on } B_r(x) \quad (28)$$

- Stretch the local histogram as

$$I'(x) = I(x) + k(I(x) - V(x)) \quad (29)$$

where k is some constant corresponding to the factor range of stretching, and $V(x)$ is one of the three values $mp(x)$, $mean(x)$ or $med(x)$.

The model can be applied to the intensity gray level or to the three channels independently.

When the model is applied to the gray level intensity it must adapt the three color channels proportionally, that is, if I' is the new gray level intensity then the new color channels are $C' = (I'/I)C$ where C represents R , G or B .

The conclusions of the project must be compare the possible results using the gray level intensity or the three color channels independently concluding if it's better to apply the model to the gray level intensity or to the three color channels.

On the other hand, it must compare the three different models using the mid point, the mean or the median values, concluding which gives the better results.

5.2 Design an algorithm for the local histogram equalization preserving the level sets

The project consists on the lecture and compression of the work of Vicent Caselles, Jose Luis Lisani, Jean-Michel Morel and Guillermo Sapiro, **Shape Preserving Local Histogram Modification** to design an optimal strategy to implement the model . This work consists on a local histogram equalization algorithm preserving the level sets of the image.

The equalization of the histogram is a transformation T such that the new image has a uniform distribution function. Let be H the distribution function of the image u defined in a domain R and with l_{min} and l_{max} its minimum and maximum values:

$$H(l) = \frac{\text{Area}(x \in R : u(x) \leq l)}{\text{Area}(R)}$$

Then the change of variable

$$v(x) = (l_{max} - l_{min})H(u(x)) + l_{min}$$

gives a new image whose distribution function is uniform in the interval $[l_{min}, l_{max}]$.

References

- [1] M. Bertalmio, G. Sapiro, C. Ballester, and V. Caselles. Image inpainting. In *SIGGRAPH '00*, pages 417–424, 2000.
- [2] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture image inpainting. *IEEE Transactions on Image Processing*, 12, August 2003.
- [3] A. Blake. Boundary conditions for lightness computation in mondrian world. *Comput. Vis. Graph. Image Process.*, 32:314–327, 1985.
- [4] Q. Chuan, W. Shuozhong, and Z. Xinpeng. Image editing without color inconsistency using modified poisson equation. In *Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, pages 397–401, 2008.

- [5] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. *ACM Transactions on graphics*, 21(3):242–256, 2002.
- [6] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley Publishing Company, 1992.
- [7] J. Jia, J. Sun, C.-K. Tang, and H.-Y. Shum. Drag-and-drop pasting. *ACM Transactions on graphics*, 3:631–636, 2006.
- [8] E. Land. Recent advances in retinex theory and some implications for cortical computations: Color vision and the natural image. *Proc. Nat. Acad. Sci.*, 80:5163–5169, 1983.
- [9] E. Land and J. McCann. Lightness and retinex theory. *J. Opt. Soc. Amer.*, 61(1):1–11, Jan 1971.
- [10] Anat Levin, Assaf Zomet, Shmuel Peleg, and Yair Weiss. Seamless image stitching in the gradient domain. In *Proc. of the Eighth European Conference on Computer Visison*, volume 4, pages 377–389, 2004.
- [11] J. M. Morel, A. B. Petro, and C. Sbert. A pde formalization of retinex theory. *Image Processing, IEEE Transactions on*, 19(11):2825–2837, 2010.
- [12] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. *ACM Transactions on graphics(SIGGRAPH '03)*, 22(3):313–318, 2003.
- [13] K.N. Plataniotis and A.N. Venetsanopoulos. *Color Image Processing and Applications*. Springer, 2000.
- [14] J. Sun, J. Jia, C. K. Tang, and H. Y. Shum. Poisson matting. *ACM Transactions on graphics*, 23(3):315–321, 2004.
- [15] A. Trémeau, C. Fernandez-Maloigne, and P. Bonton. *Image numérique couleur, de l'acquisition au traitement*. Dunod, 2004.
- [16] Jue Wang and Michael F Cohen. *Image and Video Matting: A survey*. Now Publishers, May 2008.
- [17] G. Wyszecki and W.S.Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley, 1984.