Chapter 4. Color image analysis by Three dimensional Fourier transform and correlation

In this chapter we propose the description of color and multi-spectral images by three dimensional functions. We face the analysis of color, and multi-channel systems from the point of view of signal theory We also study the three dimensional Fourier transform and three dimensional correlation for these functions and give an interpretation of their three dimensional spectrum, taking into account the special nature of the color variable.

4.1. Images as three dimensional light distributions

One conceives an image as a light intensity distribution on a surface that is quantified by a function that depends on two real variables (η_x, η_y), which map the points of the surface. Because of the wave nature of light, one can consider that the intensity at each point results from the contribution of an infinite number of monochromatic (harmonic) waves. The contribution to the intensity of each one of these monochromatic waves is characterized by the spectral density of the light.

Usually, the spectral distribution of the light is not uniform over the set of points where the image is defined. This way, one considers that the intensity of light of an image is quantified by a function of three variables $i(\eta_x,\eta_y,\lambda)$ in which the third axis can be interpreted as the spectral distribution of the light at the point (η_x,η_y).

4.2. Multi-channel acquisition of images

Because of the wavelength dependency of the light distribution of an image, the same signal can generate different responses when acquired by detectors with different spectral sensitivity. The response, $f(x,y)$, of a photo-detector array that we consider to be linear and shift invariant, when detecting a signal $i(\eta_x,\eta_y,\lambda)$ is given by

$$f(x,y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{0}^{+\infty} W(\eta_x - x\Delta_x,\eta_y - y\Delta_y,\lambda)i(\eta_x,\eta_y,\lambda)d\lambda d\eta_x d\eta_y \,, \qquad (4.1)$$

Here $W(\eta_x,\eta_y,\lambda)$ is the impulse response function of the detector array. x and y are zero or positive integer numbers smaller than D_x-1 and D_y-1, respectively, that index the pixels of the array. Δ_x, Δ_y are the spacing between pixels.

In the most usual case, the spectral response of the detector is the same for all the pixels, then we can write the impulse response function as a function of separable variables:

$$W(\eta_x,\eta_y,\lambda) = W_S(\eta_x,\eta_y)S(\lambda)\,. \tag{4.2}$$

That is, the response $f(x,y)$ describes a sample of the three dimensional signal, integrated over a range of the spectrum determined by the spectral sensitivity of the detector, that acts as a weight function.

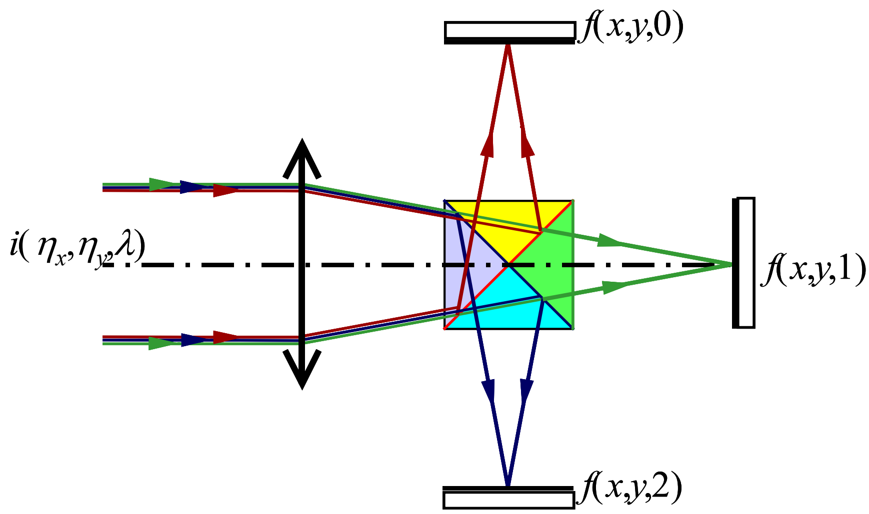

Figure 4.1. Scheme of a multi-channel detector. The same beam is split in three, and each one of the three beams passes through a dichroic filter before being acquired by a detector array.

Several elements are involved in the spectral sensitivity of the acquisition system. Between them one counts the transmission of the glasses and the reflectance of the mirrors of the optical system, and also the efficiency of the detector array. Additionally, one can use dichroic filters to modify the spectral sensitivity of the detector. Moreover, several detectors can be combined in the same device, and the same image can be sampled by detectors with different spectral responses (see Figure 4.1). This way, a collection of N different samples are generated from the same signal. The samples of the same image obtained by the different detectors are called the *channels* of the image. And an imaging system with more than one detector, having these different spectral sensitivity is called a *multi-channel* system, but it is also called color system, *multi-spectral* system or *hyper-spectral* system.

We illustrate this in Figure 4.2, the image of the Earth, that the human visual system would perceive as represented in Figure 4.2a, is acquired with three different detectors. Their spectral response is shown in Figure 4.2b. The response of the three detectors is different from zero at different ranges of the electromagnetic spectrum. This way Figure 4.2c corresponds to the detector sensitive to the wavelengths on the visible range. Figure 4.2d corresponds to a range of the infrared radiation with wavelengths about 6.4 μm, called the water vapor band. And Figure 4.2e corresponds to a range of the infrared about 13.6 μm, known as the thermal band.

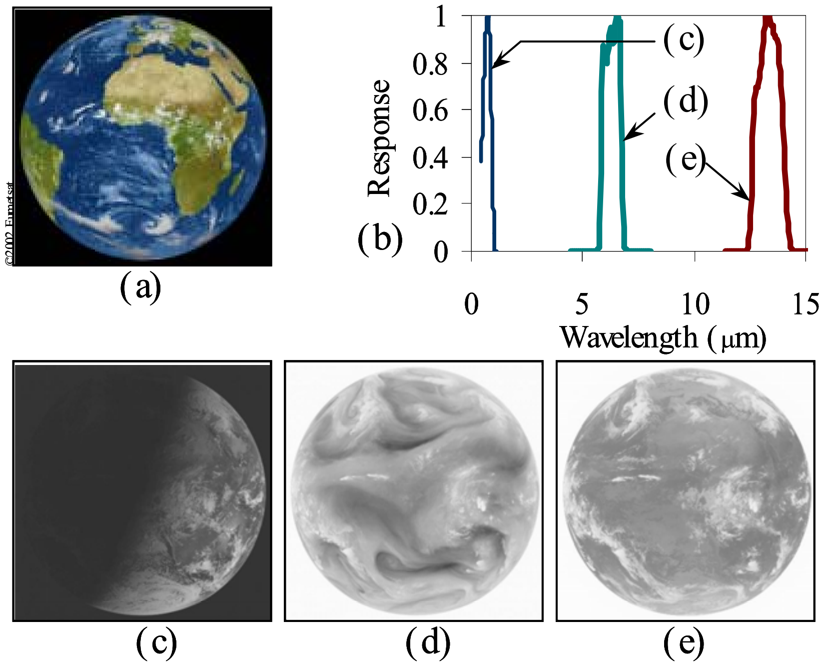

Figure 4.2. Image acquisition with a multi-channel camera. (a) Color reproduction of the earth, as seen by the human visual system. (b) Spectral response of the detectors. (c) visible band channel, (d) water vapor band channel, (e) thermal band.

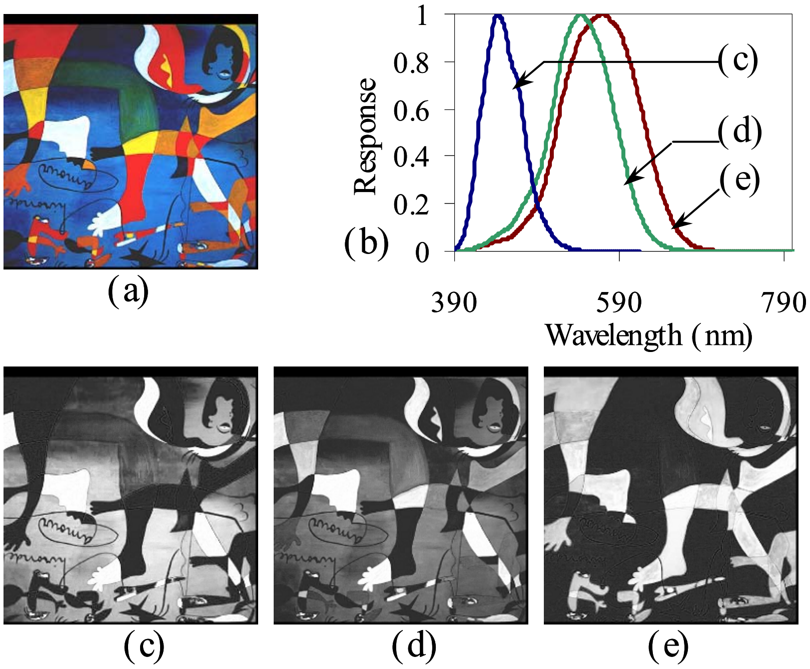

Figure 4.3. Image acquisition by the human visual system. (a) Color reproduction of the scene. (b) Normalized spectral absorption of the iodopsin pigments. (c) Blue channel of the color image. (d) Green channel. (e) Red channel.

Also the human vision system is a multi-channel system. Three varieties of a pigment called iodopsin, that have different spectral absorption response are present in the cone-cells of the eye retina, They constitute the three types of detectors, and allow the eye to behave as a multi-channel system. We have illustrated this in Figure 4.3. The image in Figure 4.3a is perceived as a color image because it is acquired by the three type of cone-cells in the retina. Their normalized spectral response is represented in Figure 4.3b. If one represents separately the response of each one of the three detectors, the images in Figure 4.3c, d and f are obtained One defines the *colors* as the different sensations (the response) that the human visual system perceives when stimulated by different spectral distributions. In analogy to the human visual system, in this work we also refer to the response of a multi-channel system as color. The non-uniformity on the perception of the color distribution of Figure 4.3a comes from the fact that the three acquired channels are different each other.

4.3. Spectrum sampling

Two different interpretations raise from the multi-channel detection of images. One of them consists of considering the multi-channel acquisition as a sampling of the spectral distribution, and the other consists of considering that the multi-channel systems have vector response in which the spectra are mapped onto a *N*-dimensional vector space. In this chapter we establish the relation between these two interpretations.

We consider the multi-channel image acquisition as a sample of the spectral distribution of the points of the image. Therefore, the response of the system is considered to be a three dimensional function, because the signal is three dimensional too. The multi-channel detector is now characterized by a three dimensional impulse response function. However, one must take into account that the spectral response of the multi-channel device is not, in general, shift invariant (See Figure 4.3b). So we write the impulse response function of the device as a function of four variables, as follows (three are real: η_x, η_y, and λ, and one is integer: n).

$$W_{3D}(\eta_x, \eta_y, \lambda, n) = W_S(\eta_x, \eta_y) S_n(\lambda), \tag{4.3}$$

Here n is the index for the detector whose spectral response is characterized by $S_n(\lambda)$ This way the three dimensional response of a multi-channel system is given by the next expression:

$$f(x, y, n) = \int_0^{D_y \Delta_y} \int_0^{D_x \Delta_x} \int_0^{+\infty} W_{3D}(\eta_x - x\Delta_x, \eta_y - y\Delta_y, \lambda, n) i(\eta_x, \eta_y, \lambda) d\lambda d\eta_x d\eta_y, \tag{4.4}$$

In addition, it is easy to find that the n-th channel acquired by the system from the image is given by

$$f_n(x, y) = f(x, y, n). \tag{4.5}$$

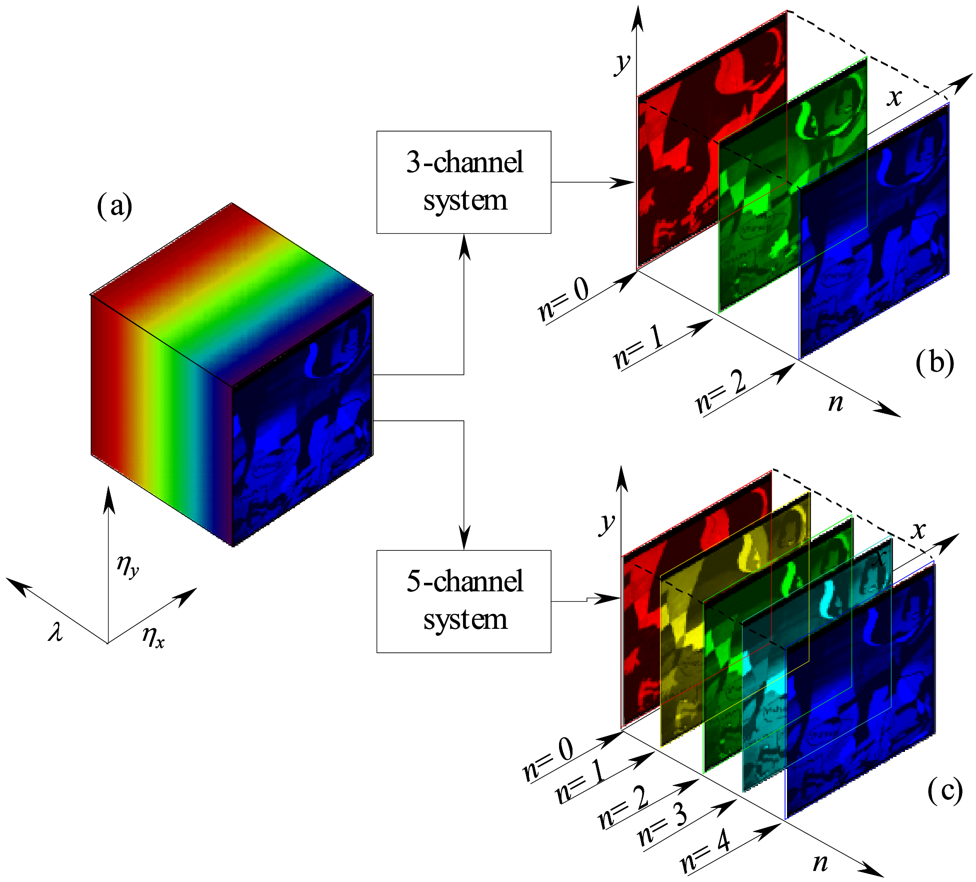

Figure 4.4. Sampling of three dimensional light distribution. (a) Original distribution. (b) Sampling by a system with 3 detectors.(c) Sampling by a system with 5 detectors.

Though multi-channel systems are not generally shift invariant along the wavelength axis, the spectral response of their detectors are similar each other, but centered at different wavelengths (See Figure 4.3b as an example). This way, to consider that the three dimensional response of the multi-channel system is a sample of the three dimensional signal is a realistic approximation.

The sampling of a continuous three dimensional image by multi-channel systems is illustrated in Figure 4.4. The continuous function in Figure 4.4a is sampled by a 3-channel system, and its response is the discrete function in Figure 4.4b. There the wavelength axis is sampled at three points. Each one of the represented planes is called a channel of the three dimensional function. If the system is a 5-channel system the response to the same signal is that of Figure 4.4c. There, the color axis is sampled at five points, that is, the three dimensional function has five channels.

Many artificial systems try to emulate the human visual system response, so they are composed of three detectors (N=3) with spectral sensitivities centered on the red, green and blue ranges of the visible range of the electromagnetic spectrum, like the iodopsin pigments responses are. One refers to the responses of each one of these three detectors as the red (R), green (G) and blue (B) channels of the color image respectively. We establish the convention that the red channel is the n=0 channel, and the green and blue channels are the n=1 and n=2 channels respectively. In general, for a color **C** with components r, g and b, the response of a RGB color system is $c(0)=r$, $c(1)=g$, and $c(2)=b$.

Also the electromagnetic spectrum of a signal is a distribution, that in this case depends on the wavelength. Therefore the Fourier transform can be applied to it. That leads to the distribution of chromatic frequencies (see for example **[Romero95]**) for a given signal. Because the number of detectors of any real system is limited, the sampling of the spectral distribution does not contemplate the higher chromatic frequencies of the spectral distributions of the images, and therefore many different spectral distributions lead to the same response of the system. For the human visual system, the identification of different spectral distributions by the same color is usually known as *metamerism*.

That means that the response of a multi-channel system to any signal can be emulated by a number of different signals, or by a linear combination of signals (because the system is linear). In addition, the minimum number of signals that complete the set of all the possible responses of a multi-channel system with N detectors is N signals per pixel. However, depending on the spectral responses of the system it is necessary to use negative coefficients in the linear combinations. That is, the set of all the possible responses of a multi-channel system has the structure of a N-dimensional vector space. Furthermore, the result of applying any linear operation to a color is completely determined by the result of applying the operation to the colors of the vector space basis.

4.4. The color space of the Human Visual System

In the case of the human visual system, the vector space of all the possible responses is a three dimensional space, as corresponds to the three varieties of iodopsin in the retina. This vector space is often called the *color space*. We extend the use of this term to the vector space of the possible responses of any multi-channel system. The colors of the basis of the color space are often called the *primary colors*. One uses to choose the red, green and blue as the primary colors, because they maximize the quantity of colors that can be generated with positive components. This is expressed by the Grassman's color mixture law, as follows:

$$c\mathbf{C} = r\mathbf{R} + g\mathbf{G} + b\mathbf{B}, \tag{4.6}$$

where r, g and b are the components of the color $c\mathbf{C}$ in the basis defined by **R**,**G** and **B**. To illustrate the color mixture law, we represent in Figure 4.5 the r, g, b components (known as *tristimulus values*) for the responses for all the monochromatic spectral distributions in the basis formed of three monochromatic signals (**R** at 645 nm, **G** at 525 nm and **B** at 445nm).

Note that the tristimulus value for the primary colors is one for one of the detectors and zero for the other two detectors

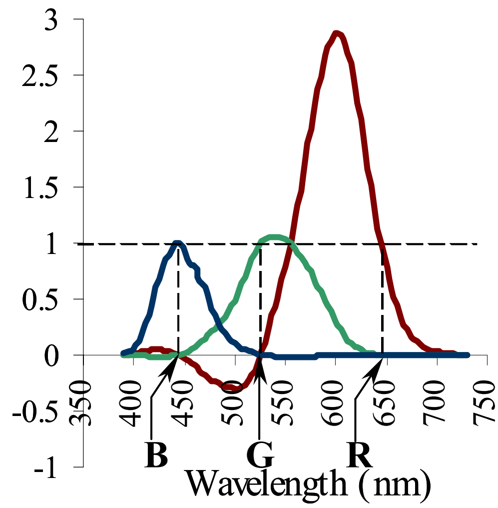

Figure 4.5. Tristimulus values for monochromatic impulses of equal power.

.The tristimulus values of the monochromatic spectra define the coordinates of a curve parameterized by the wavelength on the color space. Usually this curve is orthogonally projected onto the plane r+g+b=1, which is the plane that contains all the colors with *1/3-intensity*†, defined as the average of the response to a signal of the detectors. The r, g, b coordinates of this projection are called the *color matching functions*. Any spectrum can be conceived as the addition of monochromatic stimuli, as follows:

$$i(\lambda) = \int_0^{+\infty} i(\lambda')\delta(\lambda - \lambda')d\lambda' , \qquad (4.7)$$

† The intensity is defined as the average of r, g and b. Therefore the primary colors have intensity 1/3

where $\delta(\lambda-\lambda')$ is the spectrum for a unit power monochromatic stimulus of wavelength λ'. Therefore, the response of the system to any spectrum is a vector that intersects the 1/3-intensity plane ($r+g+b=1$) in the region bounded by the color matching functions. However, only the response of the system to stimuli that have positive components in the basis defined by the primary colors can be emulated by mixing these primary colors.

Figure 4.6a is a representation of the color vector space for the human visual system. The primary colors are indicated by the vectors **R**, **G** and **B**. The 1/3-intensity plane is also represented (see the triangle defined by its intersection with the $r=0$, $g=0$, and $b=0$ planes), and the curve defined by the color matching functions is represented on it. Note that the curve is out of the first octant and that it intersects the coordinate axes on the primary colors. That indicates that they have at least one negative component and therefore they cannot be reproduced by additive mixtures of **R**, **G** and **B.** We have also represented in Figure 4.6a the secondary colors, cyan(**C**), magenta (**M**) and Yellow (**Y**), that are the addition of equal parts of two primary colors (**C**=**B**+**G**, **M**=**B**+**R**, **Y**=**R**+**G**). The primary colors, the secondary colors, the white (**W**=**R**+**G**+**B**) and the black are the vertices of a cube, that contains the color gamut that can be generated with values of r, g and b between 0 and 1 (or between 0 and 255). Figure 4.6b represents the colors in the intersection of the 1/3-intensity plane with this cube, projected onto the **RG** plane. This triangle is usually called the *Maxwell's triangle*. All the colors with intensity equal to 1/3 that can be represented by mixture of **R**, **G** and **B** have been represented. The dashed line represent the points corresponding to the response to the pure spectral colors. Because they have at least one negative component, they can not be reproduced by positive combinations of **R**, **G**, and **B** (Except the primary colors). Therefore they are out of the Maxwell's triangle.

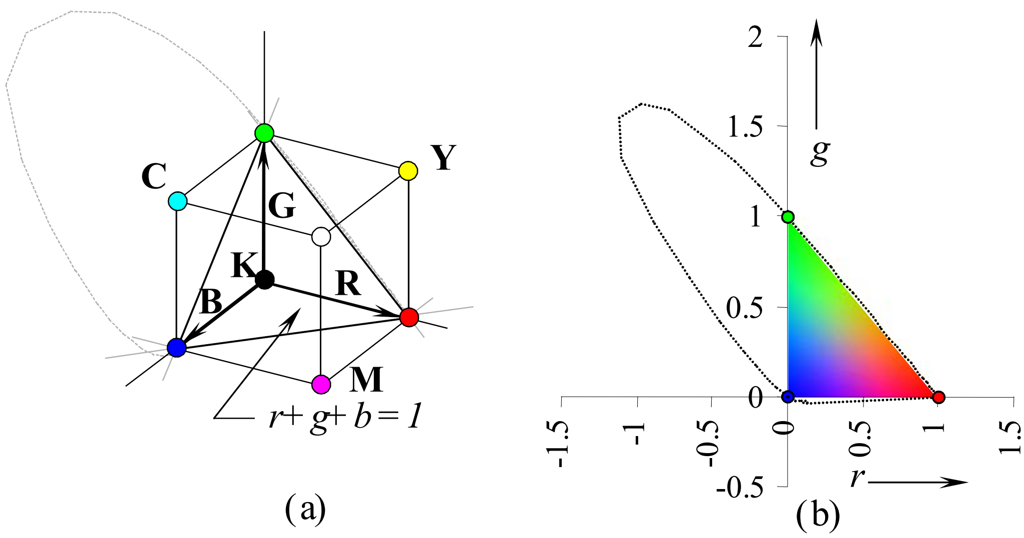

Figure 4.6. Representation of the color space. (a) three dimensional representation of the primary colors (**R**, **G** and **B**). (b) Projection of the colors on the unit intensity plane onto the **RG.** The dashed line represents the monochromatic colors.

Aside from the {**R**,**G**,**B**} basis, the human visual system color space can be generated by different vector bases. Some usual bases are the {**X**,**Y**,**Z**}, in which the components of any color are positive, being the component on **Y** the luminance of the color and also the {**A**,**T**,**D**} basis, built from psycho-physical models of human perception.

It is also interesting to map the color space using a cylindrical coordinate system (ζ,ρ,θ), in which the privileged axis is perpendicular to the equal-intensity planes. It is represented in Figure 4.7a. This coordinate system is directly related to the *HSI* (hue, saturation, intensity) representation of color, given by the next expressions **[Gonzalez92]**:

$$I = \frac{1}{3}(r + g + b), \tag{4.8a}$$

$$S_H = 1 - \frac{3}{r + g + b}\min(r, g b), \tag{4.8b}$$

$$H = \tan^{-1}\left[\frac{\sqrt{3}(g - b)}{2r - g - b}\right]. \tag{4.8c}$$

The relation between (ζ,ρ,θ) and *HSI* is given by:

$$I = \sqrt{3}\zeta\ , \tag{4.9a}$$

$$S_H = \frac{\rho}{\rho_{MAX}(\theta)}, \tag{4.9b}$$

$$H = \theta\ , \tag{4.9b}$$

One can identify the intensity (I) to the axial coordinate of the cylindrical coordinate system (ζ). The hue (H) is the angular coordinate (θ), represented in Figure 4.7b for an equal-intensity section of the cylinder; and the saturation (S_H) is the radial coordinate normalized to take its maximum value (one) on the coordinate planes. This representation system is interesting because the coordinates have a direct spectral interpretation. The intensity is correlated to the total energy of the signal, the saturation gives information about the proximity of the colors to the pure spectral colors, and the hue gives information about the main wavelength range of the signal that generates the color.

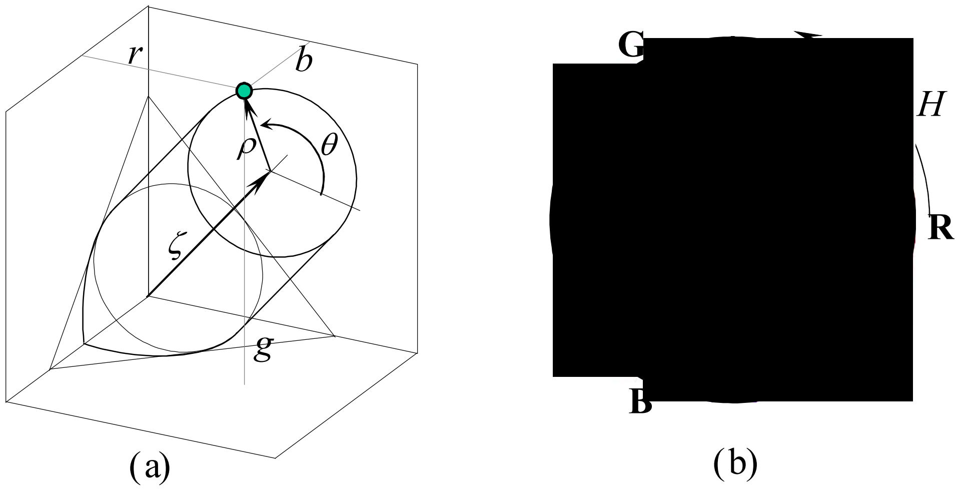

Figure 4.7. (a) Cylindrical coordinate system of the color space. (b) colors in the *HS* representation of the 1/3-intensity plane. The shadowed colors represent colors that are not reproducible by mixture of **R**, **G** and **B**.

Therefore, the *HSI* representation establishes a relation between the geometrical distribution of the colors in the color space and some properties of the spectral distribution that generate the colors.

4.5. Color Fourier spectrum

Linear transformations of the color space can be considered as linear transformations of the signals that generate the colors. These operations can also be performed in the frequency domain. So, it results interesting to define the Fourier transform of the color distribution, we refer to this operation as the *color Fourier transform*, and it is defined by the following expression:

$$F_C(m) = \frac{1}{N}\sum_{n=0}^{N-1} c(n)\exp\left(-i2\pi\frac{mn}{N}\right). \tag{4.10}$$

$F_C(m)$ is the *color Fourier spectrum* of the color with components $c(0),\ldots, c(N-1)$. Analogously, one says that $F_C(x,y,m)$ is the color Fourier spectrum of a color image with components if

$$F_C(x,y,m) = \frac{1}{N}\sum_{n=0}^{N-1} f(x,y,n)\exp\left(-i2\pi\frac{mn}{N}\right). \tag{4.11}$$

The color Fourier transform is a linear operation on the color space, therefore, once the basis has been established, it can be expressed as a $N\times N$ matrix ($\mathbf{F}$), whose matrix elements F_{mn} are complex numbers given by

$$F_{mn} = \frac{1}{N}\exp\left(-i2\pi\frac{mn}{N}\right). \tag{4.12}$$

As an example we can consider the matrix for N=3:

$$\mathbf{F} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & \exp\left(-\frac{2\pi}{3}\right) & \exp\left(+\frac{2\pi}{3}\right) \\ 1 & \exp\left(+\frac{2\pi}{3}\right) & \exp\left(-\frac{2\pi}{3}\right) \end{bmatrix}. \tag{4.13}$$

The inverse color Fourier transform $\mathbf{F}^{-1}$ operator is given by

$$\mathbf{F}^{-1} = N\mathbf{F}^{\dagger}, \tag{4.14}$$

that means that except by a normalization factor, the color Fourier transform is a unitary[†] operator: i.e., orthogonal vectors in the color space are transformed in orthogonal vectors in the color Fourier spectrum vector space (defined to be orthogonal if their hermitian product is zero). Note that while the color space is a N-dimensional vector space of vectors with real components, the color Fourier transform is a linear operation onto a N-dimensional vector of complex components, that therefore has $2N$ degrees of freedom. Nevertheless, the spectra of the colors have symmetrical real part and anti-symmetrical imaginary part, what involves to have N boundaries on the vector space of the spectra. Therefore, the total number of degrees of freedom for the manifold of the color Fourier spectra of the color space is N.

The m=0 channel of the color Fourier spectrum of an image corresponds to the DC term of its Fourier series. It is formed by the un-weighted addition of all the channels of the image, therefore, aside from the total intensity of the signal, it does not contain any information about the spectral distribution of the image. Because physical color distributions are real-valued, the m=0 channel of the color-frequency spectrum is real-valued too. Therefore, the values at the DC channel of the color-

[†] One can normalize the Fourier transform operation as $F_{mn} = \frac{1}{\sqrt{N}} \exp\left(-i2\pi \frac{mn}{N}\right)$ so as to have a unitary operator $\mathbf{F}^{\dagger}\mathbf{F}=\mathbf{I}$, that preserves both orthogonality and the norm.

frequency spectrum of all the elements of the color space map a one-dimension manifold. Every point of this one-dimensional manifold corresponds to all the colors whose spectra have the same DC value. These colors determine a $N-1$ dimensional manifold of the color space, which is determined by the equation

$$\sum_{n=0}^{N-1} c(n) = k, \tag{4.15}$$

and is represented by the hyper-plane Ω_k in Figure 4.8. The m=0 channel of the color Fourier spectrum is the same for all the colors whose representation in the color space has the same orthogonal projection onto the straight line ω in Figure 4.8, which is orthogonal to the planes Ω_k

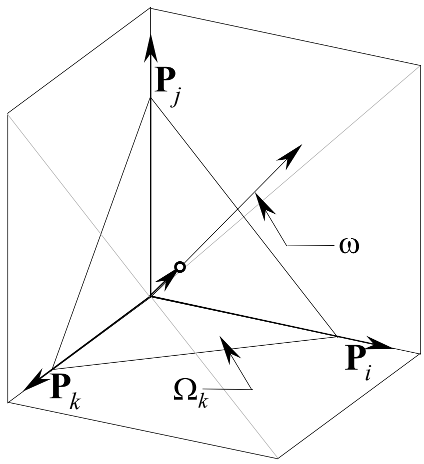

Figure 4.8. Representation of the hyper-plane of all the colors that have the same DC value in the color-frequency spectrum (Ω). $\mathbf{P}_i$, $\mathbf{P}_j$, $\mathbf{P}_k$ are three of the N primary impulses of the N-dimensional color space.

This is demonstrated next. The projection of a color described by a vector **c** with components $c(n)$ onto the line ω is given by its scalar product with the director vector of ω, $\mathbf{u}_\omega$

$$\mathbf{c}\cdot\mathbf{u}_w = \sum_{n=0}^{N-1} c(n)u_w(n). \tag{4.16}$$

All the components of $\mathbf{u}_\omega$ are equal, therefore we can factorize them, as follows:

$$\begin{aligned}\mathbf{c}\cdot\mathbf{u}_w &= u_w\sum_{n=0}^{N-1} c(n) \\ &= Nu_w F_C(0)\end{aligned}. \tag{4.17}$$

This way, the DC channel is proportional to the orthogonal projection of the color onto the line ω.

In the case of the *RGB* color space, the DC channel of the color-frequency spectrum is identified as the intensity of the transformed color. That is, the same value in the m=0 channel of the color-frequency spectrum is shared by all the colors that have the same intensity, independently of their chromaticity. These colors compose the planes Ω_k.

The other channels ($m\neq 0$) of the color-frequency spectrum are, in general complex-valued, even when considering real-valued color compositions. They establish a mapping of the color space onto the complex plane, which is a two-dimensional vector space. Because the color Fourier transform is an hermitian operator,. the projections of the color space given by the different channels are orthogonal each other and orthogonal to

ω in particular. That is, the manifolds constituted by colors whose spectra have only one nonzero component are orthogonal each other if the nonzero component is different.

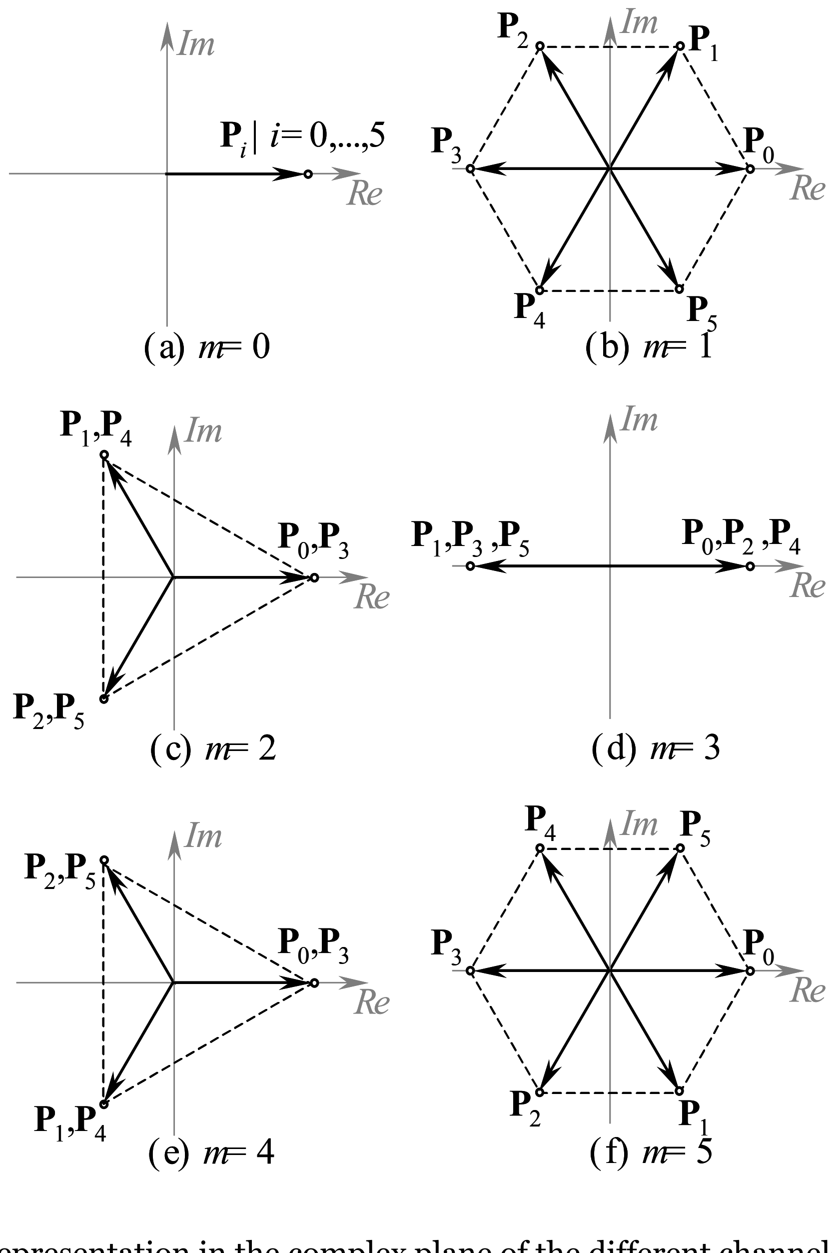

Figure 4.9. Representation in the complex plane of the different channels (a to f) of the color Fourier spectrum of the primary colors for a color space of dimension N=6.

To illustrate the mapping of the color space onto the complex plane, we consider as an example the case of a 6-dimensional color space (N=6). Taking into account that the color Fourier transform is a linear operation, the spectra of the whole color space can be obtained by linear combinations of the spectra of the primary colors. The color-frequency spectrum of each primary color has been performed and the resulting values have been represented in the complex plane for the different channels (m=0,...,5) in Figure 4.9. For each channel, the values of the color-frequency spectra of the different primary colors are distributed on the vertices of an M-angles polygon, where M is any of the integer factors of the number of channels N. In the case of N=6 the figure can be an hexagon (Figure 4.9b and f), a triangle (Figure 4.9c, e) and also the two ends of a segment of the real axis (Figure 4.9a and d). One can observe that each primary color takes a different position at each channel of the color spectrum. Therefore, the values of the m≠0 channels can be thought as different two-dimensional projections of the colors on the complex plane.

In the case of N=3, the planes parallel to Ω_k are the only planes orthogonal to ω, therefore the channel m=1, and also m=2, that is the complex conjugate of m=1, establish an orthogonal projection of the color space onto Ω_k. The mentioned projection can be derived from the color Fourier transform expression in eq.4.13. We consider the color Fourier transform of a color with components r,g and b:

$$\begin{bmatrix} F_C(0) \\ F_C(1) \\ F_C(2) \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & \exp\left(-\frac{2\pi}{3}\right) & \exp\left(+\frac{2\pi}{3}\right) \\ 1 & \exp\left(+\frac{2\pi}{3}\right) & \exp\left(-\frac{2\pi}{3}\right) \end{bmatrix}\begin{bmatrix} r \\ g \\ b \end{bmatrix}. \tag{4.18}$$

$F_C(0)$ maps the line ω, and the real and imaginary part of $F_C(2)$ can be taken as the parameters of Ω. Then, we can write:

$$\begin{bmatrix} F_C(0) \\ \operatorname{Re}F_C(2) \\ \operatorname{Im}F_C(2) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -i & i \end{bmatrix}\begin{bmatrix} F_C(0) \\ F_C(1) \\ F_C(2) \end{bmatrix}. \tag{4.19}$$

Therefore

$$\begin{bmatrix} F_C(0) \\ \operatorname{Re}F_C(2) \\ \operatorname{Im}F_C(2) \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & \cos\left(-\frac{2\pi}{3}\right) & \cos\left(+\frac{2\pi}{3}\right) \\ 0 & \sin\left(-\frac{2\pi}{3}\right) & \sin\left(+\frac{2\pi}{3}\right) \end{bmatrix}\begin{bmatrix} r \\ g \\ b \end{bmatrix}. \tag{4.20}$$

And by replacing the cosine and sine functions by their values, we obtain,

$$\begin{bmatrix} F_C(0) \\ \operatorname{Re}F_C(2) \\ \operatorname{Im}F_C(2) \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix}\begin{bmatrix} r \\ g \\ b \end{bmatrix}. \tag{4.21}$$

Note that the row-vectors in the matrix are orthogonal each other, but they are not unit vectors. They can be normalized by dividing them by their norm. This leads to the following operator:

$$\begin{bmatrix} \zeta \\ \eta \\ \xi \end{bmatrix} = \begin{bmatrix} \sqrt{3}F_C(0) \\ \sqrt{6}\operatorname{Re}[F_C(2)] \\ \sqrt{6}\operatorname{Im}[F_C(2)] \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{6} & 0 \\ 0 & 0 & \sqrt{6} \end{bmatrix}\begin{bmatrix} F_C(0) \\ \operatorname{Re}[F_C(2)] \\ \operatorname{Im}[F_C(2)] \end{bmatrix}. \tag{4.22}$$

This way one can establish a liner relation, **A,** between the (r,g,b) components of a color and a new set of components (ζ,η,ξ) determined by the values of the color-frequency spectrum.

$$\mathbf{A}=\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \tag{4.23}$$

and

$$\begin{bmatrix} \zeta \\ \eta \\ \xi \end{bmatrix} = \mathbf{A}\begin{bmatrix} r \\ g \\ b \end{bmatrix}. \tag{4.24}$$

A is the transformation operator between two orthonormal vector bases of the color space, therefore it is an orthogonal operator, i.e. it verifies $\mathbf{A}^T\mathbf{A}=\mathbf{I}$.

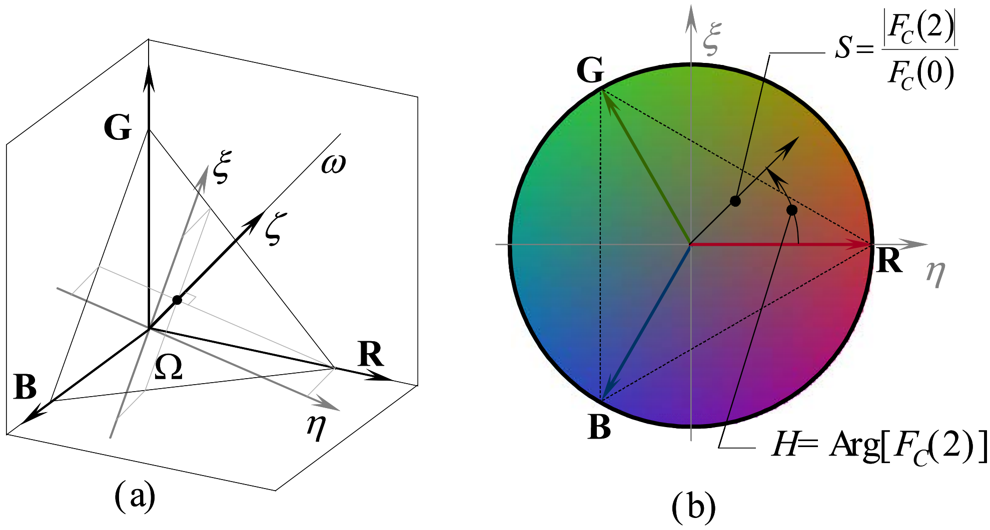

Figure 4.10. (a) Representation of the vectors of the base defined by the DC value of the color-frequency spectrum, and the real and imaginary parts of the m=2 channel. (b) Distribution of the colors with intensity equal to 1/3 in the complex plane according to the value at the m=2 channel of their color-frequency spectrum.

We represent in Figure 4.10a the ortho-normal basis of the color space defined by the color Fourier transform. Note that the m=0 channel of the color-spectrum is proportional to the projection ζ of the colors onto the line ω. Additionally, the channels m=1 and m=2 establish a projection of the colors onto the plane Ω. We give in Figure 4.10b a representation of this mapping. We have placed at each point (η,ξ) the color with 1/3-intensity whose spectrum, valued at the m=2 channel is $\eta+i\xi$.

Furthermore, we can retrieve the relations between the color spectrum and the hue, saturation and intensity representation of a color:

$$H = Arg[F_C(2)], \tag{4.25a}$$

$$S = \frac{|F_C(2)|}{F_C(0)}, \tag{4.25b}$$

$$I = F_C(0). \tag{4.25c}$$

Here, *S* corresponds to the saturation of the color. However, in this case the normalization is independent of the hue of the colors. It takes its maximum value (*S*=1) in the primary colors.

This way, the relation of the *HSI* representation of the colors with the characteristics of the spectral distribution of the signals that generate the colors can be justified by the usual properties of the Fourier transform.

The intensity (*I*) is proportional to the DC channel of the color-frequency spectrum of the spectral distribution. Because the DC term is conceived as the average value of the spectral distribution of the signal, it results obvious that the intensity represents a measure of the total power of the signal.

The saturation (*S*) is defined as the ratio between the high color-frequencies and the DC term. It constitutes a measure of the proximity of the spectrum of the signal to a uniform spectrum signal (white), or to a monochromatic signal. If the spectral distribution of the signal is uniform in the wavelength axis, only the DC term of its color-frequency spectrum is nonzero, therefore its saturation is zero. On the other hand, for a pure spectral signal, the color-spectrum has a uniform distribution.

Therefore the ratio between the high-color frequencies and the DC term is one.

Also the meaning of the hue (H) can be derived from the properties of the Fourier transform. It is defined as the phase (or argument) of the $F_C(2)$ channel. The translation theorem establishes the relation between the phase distribution of the color-frequency spectrum and the shift of the signal from the origin. Therefore, one can consider that the hue contains the information about the location of the signal in the wavelength axis.

Let us define the energy of a spectral distribution as

$$E_C = \sum_{n=0}^{N-1} c(n)^2 \, . \tag{4.26}$$

We refer to E_C as the *color energy*[†]. In the case of the *RGB* color space it can be written as:

$$E_C\{r, g, b\} = r^2 + g^2 + b^2 \, . \tag{4.27}$$

E_C is the measure of the distance of the signal to the origin of the color space (the black color), independently of the hue or the saturation of the color. Its relation to the intensity and the saturation is derived next.

[†] r, g and b are usually based on intensity measures, and they represent a measure of the physical energy. Nevertheless, E_C must beunderstood from the point of view of signal theory, and is defined so as to be preserved by Parseval's theorem.

By applying the Parseval's theorem we can write the color energy in terms of the color Fourier spectrum, as follows:

$$\begin{aligned} E_C\{r,g,b\} &= 3|F_C(0)|^2 + 3|F_C(1)|^2 + 3|F_C(2)|^2 \\ &= 3|F_C(0)|^2 + 6|F_C(2)|^2 \end{aligned}. \tag{4.28}$$

And by dividing the two sides of the equation by $|F_C(0)|^2$ we have

$$\frac{E_C\{r,g,b\}}{|F_C(0)|^2} = 3 + 6\frac{|F_C(2)|^2}{|F_C(0)|^2}. \tag{4.29}$$

We replace the expressions for the intensity and the saturation. Then we obtain

$$\frac{E_C}{I^2} = 3 + 6S_I^2. \tag{4.30}$$

That means that colors with equal intensity do not have the same color energy because the ratio of the color energy to the intensity has a quadratic dependence on the saturation of the colors.

4.6. Three dimensional Fourier transform of color images

The combination of the color Fourier transform with the usual two dimensional Fourier transform of images leads to the definition of the three dimensional Fourier transform of three dimensional functions that describe color images. Let us consider a color image described by a three-dimensional function $f(x,y,n)$, defined for integer values of x, y and n between zero and D_x, D_y and N respectively, which are the size in the two spatial directions and the number of channels of the image. The frequency spectrum of $f(x,y,n)$ is also a three dimensional function. The three dimensional discrete Fourier transform of $f(x,y,n)$ leads to a sample of its three dimensional spectrum noted as $F(u,v,m)$:

$$F(u,v,m) = \frac{1}{D_x D_y N} \sum_{n=0}^{N-1} \sum_{y=0}^{D_y-1} \sum_{x=0}^{D_x-1} f(x,y,n) \exp\left[-i2\pi \left(\frac{ux}{D_x} + \frac{vy}{D_y} + \frac{mn}{N} \right) \right], \qquad (4.31)$$

where u and v are the spatial frequencies in the x and y directions respectively, and m indicates the color frequency.

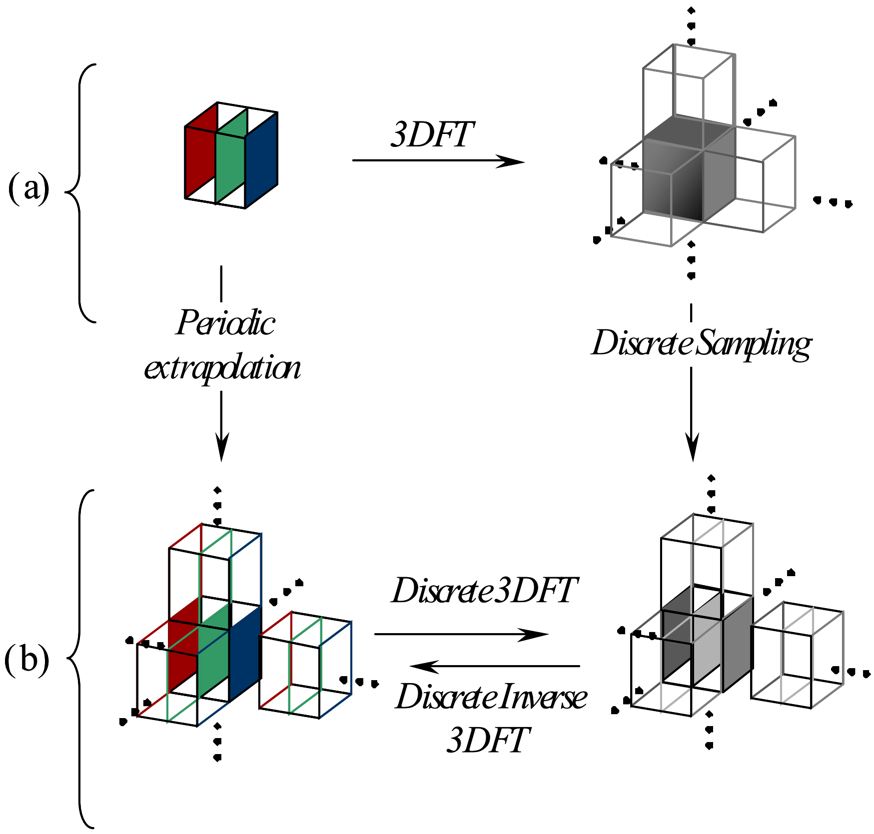

Figure 4.11. Illustration of the three dimensional Fourier transform of color images. (a) The continuous Fourier transform of a discrete signal leads to a continuous periodic spectrum. (b) The discrete Fourier transform links the periodic extrapolation of the signal with a discrete sample of its spectrum.

Because $f(x,y,n)$ is a discrete function, its spectrum is periodic (see Figure 4.11a), and the sampling theorem indicates that the period of $F(u,v,m)$ is D_x, D_y and N in the u, v and m directions respectively. A single period of the three dimensional spectrum contains all the information carried by the sample of the image. Therefore, one can consider that the size and the number of channels of the three dimensional spectrum is the same as the size and number of channels for the sample of the image.

The inverse three dimensional Fourier transform gives the expression of three dimensional function in the direct domain in terms of its three dimensional spectrum,

$$f(x,y,n)=\sum_{n=0}^{N-1}\sum_{y=0}^{D_y-1}\sum_{x=0}^{D_x-1}F(u,v,m)\exp\left[+i2\pi\left(\frac{ux}{D_x}+\frac{vy}{D_y}+\frac{mn}{N}\right)\right]. \qquad (4.32)$$

In general, this expression is valid only for x, y and n between zero and D_x, D_y and N respectively, because $F(u,v,m)$ is a discrete sample of the three dimensional spectrum of $f(x,y,n)$. However, if we consider $f(x,y,n)$ to be periodic, then the expression is exact for any set of arguments because in this case the sample of the spectrum corresponds to the spectrum, which is discrete for periodic functions. One considers that the three dimensional discrete Fourier transform and the inverse three dimensional discrete Fourier transform link the periodic extrapolation of a sample of the three dimensional signal to the discrete sample of its three dimensional spectrum (see Figure 4.11b). This way, one can assign the index $n+kN$ to the n-th channel of the three dimensional spectrum, where k is any integer number. In the case of *RGB* color images, in which we have established the correspondence of the red channel to $n=0$, the green channel to $n=1$ and the blue channel to $n=2$, it is sometimes useful to consider that the blue channel is assigned to $n=-1$, and that it is the opposite channel to the green.

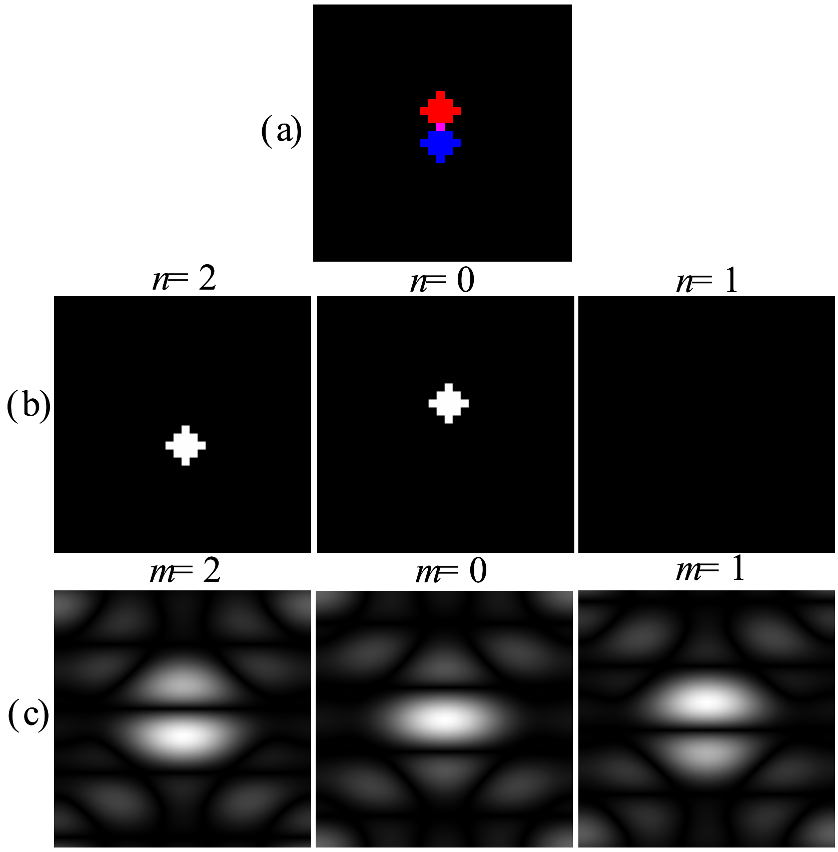

Figure 4.12. Three dimensional Fourier transform of a color image.(a) Original color image. (b) RGB channels of the color image. (c) channels of its three dimensional spectrum.

Because the three-dimensional functions that describe color images are real-valued functions, their spectrum have symmetric real part and anti-symmetric imaginary part. Therefore the m-th channel and the $(N–m)$-th channel of the three dimensional spectrum are complex conjugate each other.

The three dimensional description of color images involves the coupling between the spatial information and the color distribution of the images. This feature differentiates it from other color imaging techniques, including those that involve parallel processing of the color channels.

To illustrate this, we just consider the color image in Figure 4.12a. It consists of two objects, with the same shape and different color, that overlap in one pixel. Each object can be considered as a square with a 5-pixel long diagonal, which is oriented along the vertical. One square is red and the other is blue, and the pixel where they overlap is colored in magenta. The *RGB* channels of the figure are represented in Figure 4.12b. Each square appears in a different channel. This way, there is no signal in the n=1 channel, that corresponds to the green channel. The magnitude of the channels of the spectrum obtained by performing the three dimensional Fourier transform are represented in Figure 4.12c. One can observe that all the channels of the image are involved in the generation of each one of the channels of its three dimensional spectrum. This way, while the n=1 channel (green) is null in the direct domain image, the m=1 channel of the three dimensional spectrum takes nonzero values.

Therefore, unlike the multi-channel techniques based on the parallel processing of the channels, there is a coupling of the information carried by the different channels that is useful for some color image processing applications. An example to illustrate this is given in Figure 4.13. We have considered the two sample scenes in Figure 4.13a and b, and we have performed the two dimensional Fourier transform of the channels, processed in parallel, and the three dimensional Fourier transform of the color image. In the first case, noted as multi-channel, we represent the addition of the magnitude of the two dimensional spectra of the three channels. For the second case, noted as the three dimensional case, we consider the m=0 channel of the three dimensional spectrum of the color image.

One can observe that for the multi-channel case, both scenes lead to the same result, even when the scene in Figure 4.13a has two squares and Figure 4.13b has only one. This is because the squares in Figure 4.13a are on different channels of the color image. Therefore, for both cases the R and B channels have a single square, and there is no interference between the two objects of the scene in Figure 4.13a.

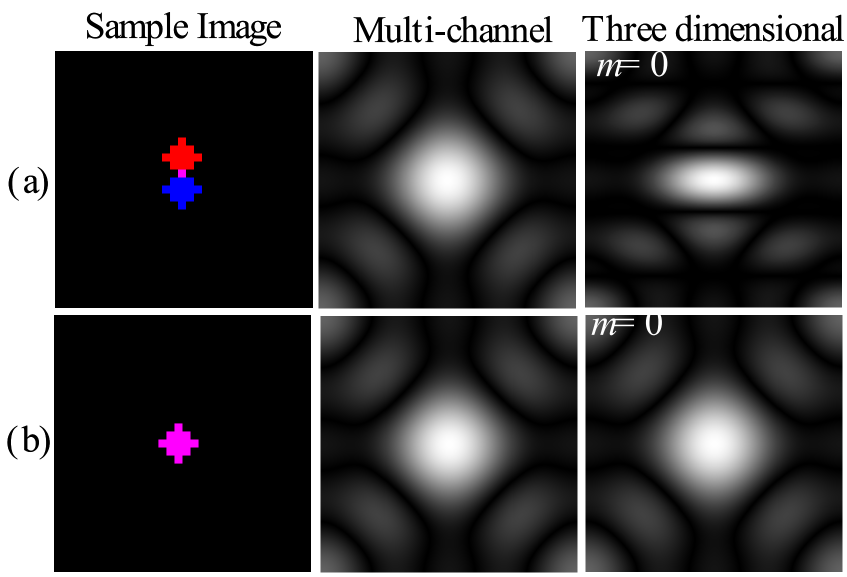

Figure 4.13. The three columns represent respectively the sample scene, the addition of the magnitude of the two dimensional spectra of the channels, and the m=0 channel of the three dimensional spectrum of the signal for a two object scene (a) and for a single object scene (b).

Contrasting this, the three dimensional Fourier transform leads to different results for the two scenes. In this case there is interference between the channels. Therefore, the two objects in Figure 4.13a produce an interference pattern that leads to a result different from the obtained for Figure 4.13b, in which the scene is composed by a single square. They are shown in the third column of Figure 4.13.

4.7. Properties of the three dimensional Color Fourier transform

Usual properties of the Fourier transform have a special interpretation in the case of color images, given the particular nature of the color variable. The interpretation for some of these properties are given next.

4.7.1. *Three dimensional impulse function and uniform function.*

An impulse function is a distribution that is null everywhere except one point (one pixel for discrete distributions). The unit impulse distribution is described by a delta function, the Dirac's delta function for continuous functions and the Kroneker's delta function for discrete distributions. When considering the three dimensional description of color images, one has to take into account that each point of the three dimensional function corresponds to a point of a single channel of the color image. This way, a three dimensional impulse function corresponds to a function that is null everywhere except one point where it takes a color distribution corresponding to one of the primary colors.

The magnitude of the three dimensional spectrum of a three dimensional impulse function is uniform along the three dimensions. That is, the value of the magnitude is the same for all the spatial and color frequencies.

Another interesting case is the three dimensional uniform function, that takes the same value at all the pixels and all the channels. It describes an image composed by a uniform field, colored in white, that is the color that results from adding equal amounts of all the primary colors. For this function, the spectrum is null-valued at all the pixels of all the channels, except at the origin pixel of the m=0 channel, that corresponds to the DC value of the function. This way, the three dimensional spectrum of a uniform white image is a centered three dimensional impulse function.

Let us also consider the case of the monochromatic images, that is, images with only one nonzero channel. The color distribution of each pixel of the image can be considered as an impulse function along the color axis. Therefore, the distribution along the color-frequency axis of the corresponding three dimensional spectrum is a constant (see Figure 4.14a).

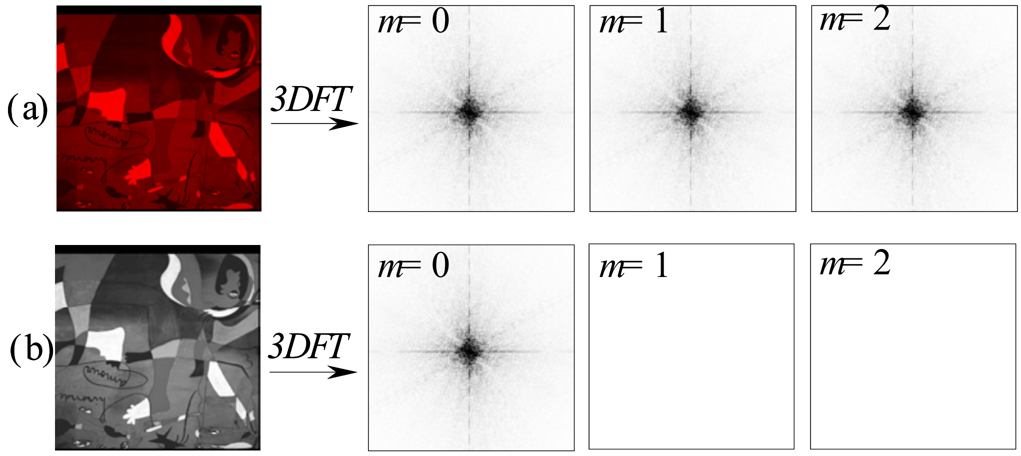

Figure 4.14. Three dimensional spectrum of a monochromatic image (a), and of a grayscale image

The opposite case are the grayscale images, in which the color distribution of every pixel is uniform (Figure 4.14b). Therefore the distribution along the color frequency axis of the three dimensional spectrum of a grayscale image is an impulse function. In other words, the three dimensional spectrum of grayscale images has only one nonzero channel.

4.7.2. ***Translation theorem***

The translation theorem states that the spectra of translated functions differ only in a linear phase factor. For three dimensional functions it is expressed as follows:

$$\begin{aligned} FT_{3D}\left[f(x-x_0, y-y_0, n-n_0)\right](u,v,m) &= FT_{3D}\left[f(x,y,n)\right](u,v,m) \\ &\times \exp\left[-i2\pi\left(\frac{ux_0}{D_x}+\frac{vy_0}{D_x}+\frac{mn_0}{N}\right)\right]. \end{aligned} \quad (4.33)$$

Let us note that we consider cyclic translations because we deal with discrete functions and discrete spectra.

From the translation theorem it follows that the magnitude of the three dimensional spectrum of a color image is not affected by cyclic translations either spatial or along the color axis. We illustrate this in Figure 4.15. We have consider the three color images represented in Figure 4.15a. Each one contains the same object shifted from the origin by three pixels in the vertical direction. The objects have also been shifted along the color axis. That is, the red square is at the origin of the color axis, but the green and blue ones are shifted by one and two channels respectively.

The magnitude of the spectrum of the considered images is uniform along the color axis for the three scenes, because they are monochromatic images. We represent in Figure 4.15b the magnitude of one of the channels of the three dimensional spectrum of each one of the sample images. One can observe that the magnitude is equal in the three cases. We have also represented in Figure 4.15c the phase distribution of the three channels of the three dimensional spectrum. Because the square is shifted from the origin along the vertical direction the phase increases linearly along the y-coordinate in the three cases. In addition, for the red object, the phase of the three channels is the same, but in the case of the green and the blue squares there is a phase delay between the channels that grows linearly with the channel index, m. This is because these scenes are translations of the first one along the color axis. Because the shift for the green square is in the opposite direction to the shift for the blue square, the sign of the phase delay between channels have opposite sign and equal absolute value.

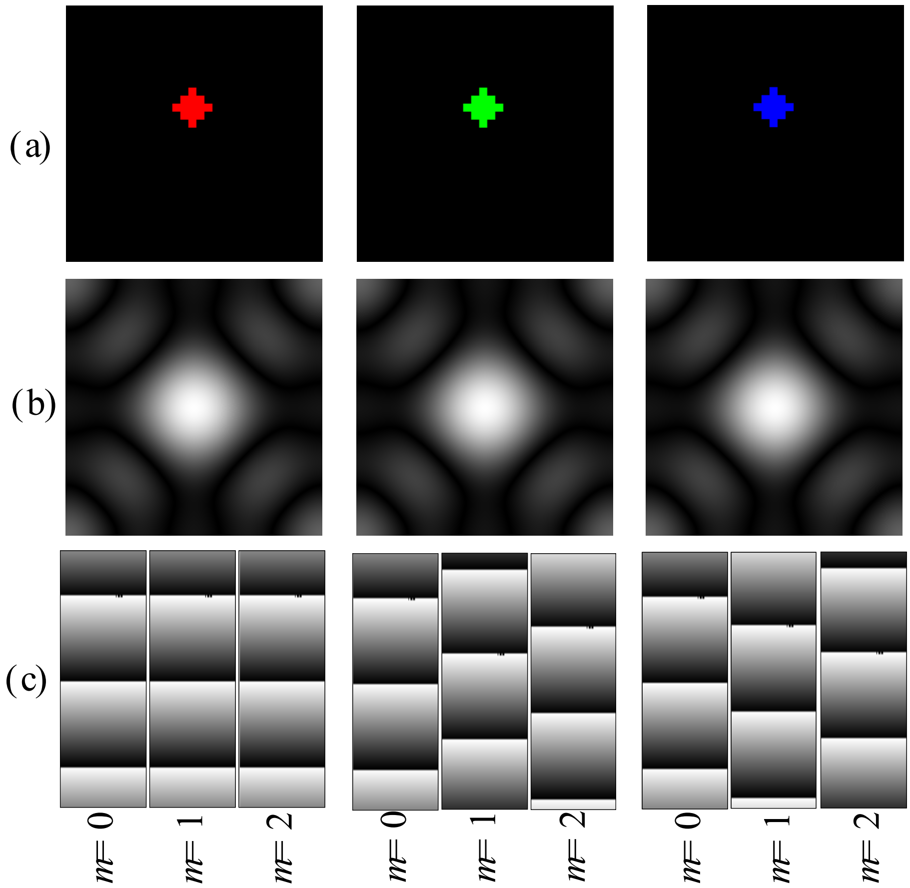

Figure 4.15. Illustration of the translation theorem. (a) Sample images (b) Magnitude of their three dimensional spectra. (c) Phase of their three dimensional spectra

4.7.3. *Convolution theorem*

The usual definition of convolution can be extended to color images described by three dimensional functions, as follows:

$$[f * g](x, y, n) = \sum_{n'=0}^{N-1} \sum_{y'=0}^{D_y-1} \sum_{x'=0}^{D_x-1} f(x', y', n')\, g(x - x', y - y', n - n'), \tag{4.34}$$

Because we use discrete functions, it results convenient to assume that the translations are cyclic along the two spatial axes and also along the color axis. Under these circumstances, the convolution theorem states that the three dimensional spectrum of the convolution of two functions is given by the product of their spectra, as follows:

$$FT_{3D}[f * g](u,v,m) = \{FT_{3D}[f] \cdot FT_{3D}[g]\}(u,v,m), \qquad (4.35)$$

here $FT_{3D}[\ ]$ indicates the three dimensional Fourier transform operation. Similarly, we have that the spectrum of the product of two functions is given by the convolution of their spectra:

$$FT_{3D}[f \cdot g](u,v,m) = \{FT_{3D}[f] * FT_{3D}[g]\}(u,v,m) \qquad (4.36)$$

Note that for the three dimensional description of color images, the different channels of the image contribute to the convolution result. This is illustrated in Figure 4.16. We have considered the same sample image as in Figure 4.12. The scene is composed of two squares (one in red and the other one in blue), one on top of the other. The three dimensional function that describes the scene (Figure 4.16a) is composed of two identical squares centered at points with different spatial and color coordinates. This way one can consider that the image is the convolution of a single centered monochromatic square, (Figure 4.16c) with a function with two punctual impulses located at the channels and positions of the squares in the color image (Figure 4.16b).

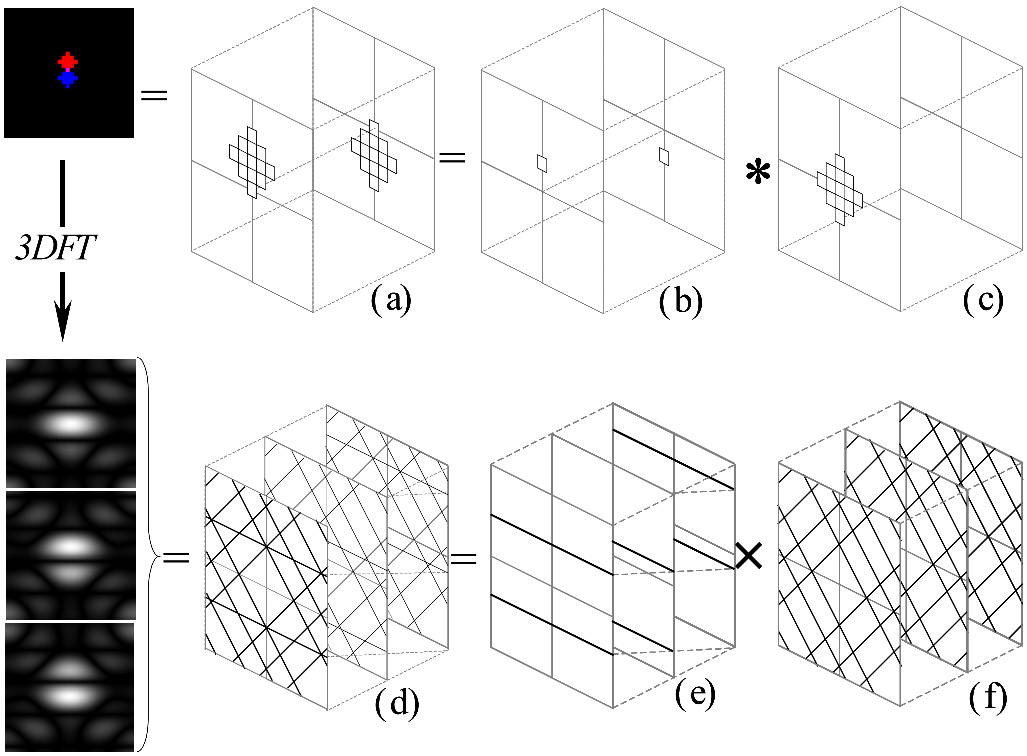

Figure 4.16. Convolution theorem (a) Sample image as a three dimensional function, expressed as the convolution of two impulses (b) and a monochromatic square (c). The three dimensional spectrum (d) is expressed as the product of the spectra for the two impulses(e), and for the square (f).

From the convolution theorem, it follows that the three dimensional spectrum of the color image, represented in Figure 4.16d, is the product of the three dimensional spectra of the two convolved functions. These spectra are represented in Figure 4.16e and f by the points where their magnitude values are zero. The three dimensional spectrum of the monochromatic square, is given by the same sinc(x+y)sinc(x–y) function in all the channels And the spectrum of the two impulses consists of the interference pattern of the two impulses, that is given by a cosine function whose argument is constant in the planes normal to the line that links the two impulses. Because the impulses are located at different points and channels, the zeros of the cosine function are located along horizontal lines, with vertical positions that depend on the channels. This interference pattern multiplies the spectrum of the square, therefore, the three dimensional spectrum (Figure 4.16d) for the color image will take null values in the zeros of the spectrum of the square, and also on the zeros of the spectrum of the two impulses, leading to the interference between the squares.

4.7.4. ***Babinet's principle***

Babinet's principle states that the magnitude of the spectrum of an image and its complementary is the same everywhere except at the origin. This can be also extended to the case of color images. In this case the complementary image comes from replacing each primary color by its complementary to the white, that is by the addition of all the other colors. This is represented in Figure 4.17. There the complementary to the image in Figure 4.17a is the image in Figure 4.17b. Note that the three overlapped squares in red, green and blue become cyan, magenta and yellow, respectively. The addition of the images in Figure 4.17a and b is a white field. Therefore, the addition of their spectra leads to the spectrum of the white field, that is zero everywhere except at the origin of the m=0 channel. That is, out of the origin of the m=0 channel, the spectra of an image is equal with the opposite sign to the spectra of its complementary image, and therefore their magnitudes are equal, as can be seen in Figure 4.17c and d, where the magnitude of the three dimensional spectra for both images is represented.

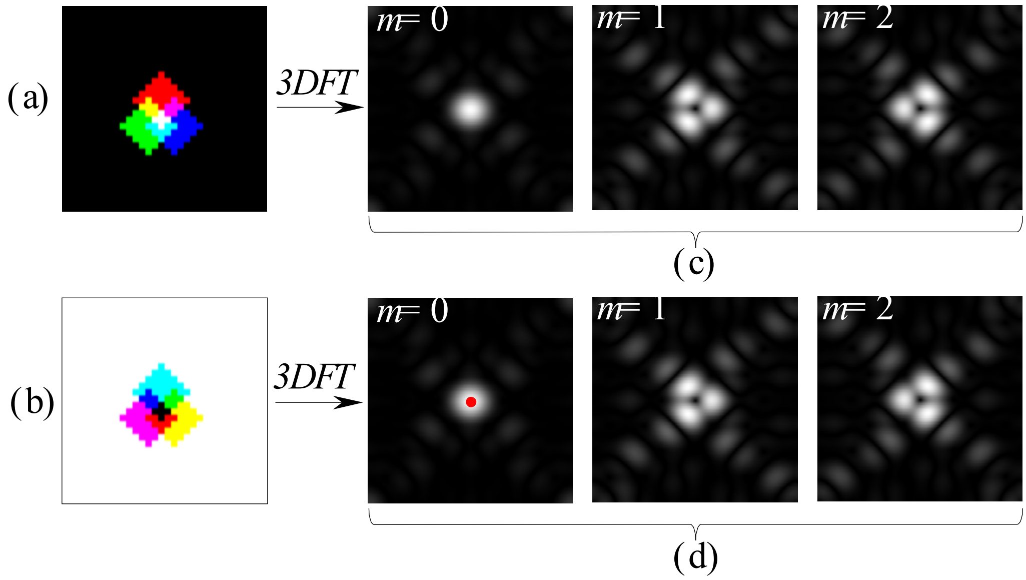

Figure 4.17. Illustration of the Babinet's principle.(a) Sample image (b) Complementary image (c) Magnitude of the three dimensional spectrum the sample image, and of the complementary image (d). The origin of the m=0 channel has been removed in (d).

4.7.5. *Inversion theorem*

The inversion theorem states that the application of two direct Fourier transform to an image, leads to the same image with the axes inverted. In the case of the three dimensional color Fourier transform we have to consider the inversion also along the color axis. We illustrate this in Figure 4.18. We have considered the sample image in Figure 4.18a composed of four objects in different colors (the four initial letters for the words 'red', 'green', 'blue' and 'cyan'). We show in Figure 4.18b, c and d the images resulting from the application of two direct Fourier transforms of the spatial distribution, of the color distribution and of the three dimensional distribution of the original image, respectively.

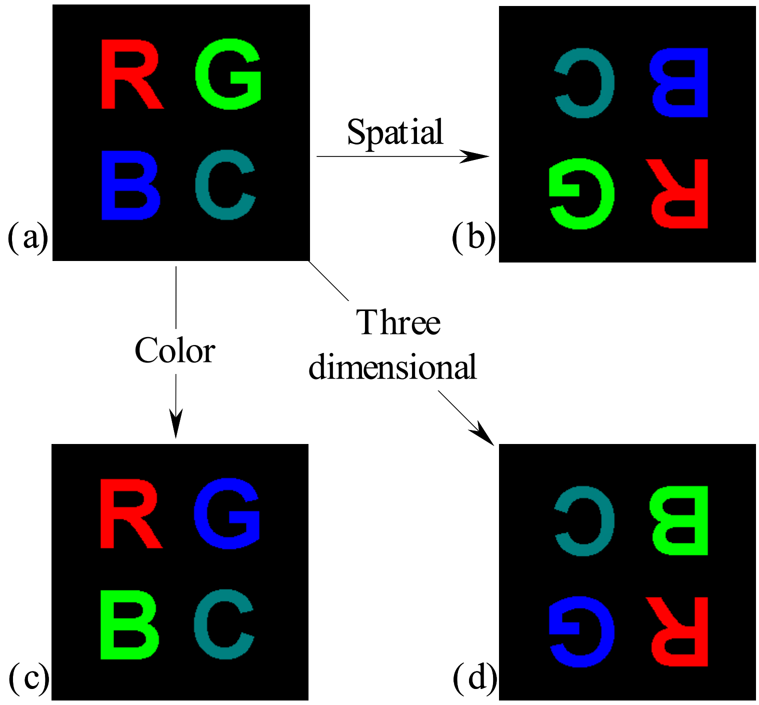

Figure 4.18. Illustration of the inversion theorem. (a) Sample image. (b) Result of applying twice the Fourier transform along the spatial variables. (c) Result of applying twice the Fourier transform along the color axis. (d) Result of applying twice the three dimensional Fourier transform

One can observe in Figure 4.18b that the spatial distribution is inverted with respect to the original image, but the colors of the objects are not altered. In Figure 4.18c the spatial distribution has not changed but the color distribution of the objects is changed. This way, the green object in Figure 4.18a has turned blue in Figure 4.18c, and vice versa. Neither the red nor the cyan objects have changed, because they have color distributions that are symmetrical along the color axis: the red is nonzero only at the origin of the color axis, and the cyan has equal amounts of green and blue.

Figure 4.18d, shows the image resulting from the application of two direct three dimensional Fourier transforms to the original image. In this case the three variables of the image have been transformed. One can observe the inversion of both the color and the spatial distribution of the objects. This way, the letters have been inverted, and the 'G' is colored in blue, while the 'B' is colored in green.

4.8. Three dimensional correlation for color pattern recognition

The description of color images by three dimensional functions and the generalization of the Fourier transform operation for these functions, allows us to generalize the three dimensional correlation as a tool for color pattern recognition. Once again, the particular character of the color dimension requires a special interpretation of the correlation properties to ensure its applicability to color pattern recognition tasks.

The correlation for discrete three dimensional functions is defined as follows:

$$[f \otimes g](x,y,n) = \sum_{n'=0}^{N-1} \sum_{y'=0}^{D_y-1} \sum_{x'=0}^{D_x-1} f(x',y',n')\, g^*(x'-x, y'-y, n'-n). \qquad (4.37)$$

Again, we consider cyclic translations along the three dimensions of the image. From the convolution theorem one can express the correlation of two functions $f(x,y,n)$ and

$g(x,y,n)$ in terms of their three dimensional spectra, $F(u,v,m)$ and $G(u,v,m)$ respectively.

$$[f \otimes g](x,y,n) = FT_{3D}^{-1}\{[F \cdot G^*](u,v,m)\}(x,y,n), \qquad (4.38)$$

where FT3D-1{ } indicates inverse three dimensional Fourier transform.

Because the correlation is a function of the spatial and the color variables (it stands on the direct domain), each channel of the three dimensional function can be associated to a primary color, in the same way as each couple (x,y) can be associated to the coordinates of a pixel. However, in general the correlation using nonlinear filters is a complex-valued function, or it takes positive and negative values, therefore it may lead to color compositions that have no physical significance. This problem can be avoided by representing a positive real-valued function of the complex number, such as the magnitude, or the argument. Nevertheless one has to refer to this assignment as a pseudo-coloration.

The utility of the correlation for pattern recognition usually comes from three major properties: the autocorrelation takes its maximum value at the origin, it is shift invariant and it is larger than the cross-correlation with any function of equal energy. These three mathematical properties of the correlation have a special interpretation when applied for pattern recognition of color images. We analyze them next.

4.8.1. *Autocorrelation at the origin*

The autocorrelation is the correlation of a function with itself. It can be demonstrated that the value of the autocorrelation is maximum at the origin, independently of the distribution of the image considered.

We consider the Cauchy-Schwarz inequality for the series in the definition of the correlation.

$$
\begin{aligned}
[f\otimes f](x,y,n) &= \sum_{n'=0}^{N-1}\sum_{y'=0}^{D_y-1}\sum_{x'=0}^{D_x-1} f(x',y',n')\,f^*(x'-x,y'-y,n'-n) \le \\
&\left[\sum_{n'=0}^{N-1}\sum_{y'=0}^{D_y-1}\sum_{x'=0}^{D_x-1} f(x',y',n')\,f^*(x',y',n')\right]^{\frac{1}{2}} \times \\
&\left[\sum_{n'=0}^{N-1}\sum_{y'=0}^{D_y-1}\sum_{x'=0}^{D_x-1} f(x'-x,y'-y,n'-n)\,f^*(x'-x,y'-y,n'-n)\right]^{\frac{1}{2}}
\end{aligned}
\quad . \qquad (4.39)
$$

Because we consider periodic functions, the two factors at the right side of this equation are equal, each other, and equal to the auto correlation at the origin. Therefore we can write.

$$
[f\otimes f](x,y,n) \le [f\otimes f](0,0,0)\,. \qquad (4.40)
$$

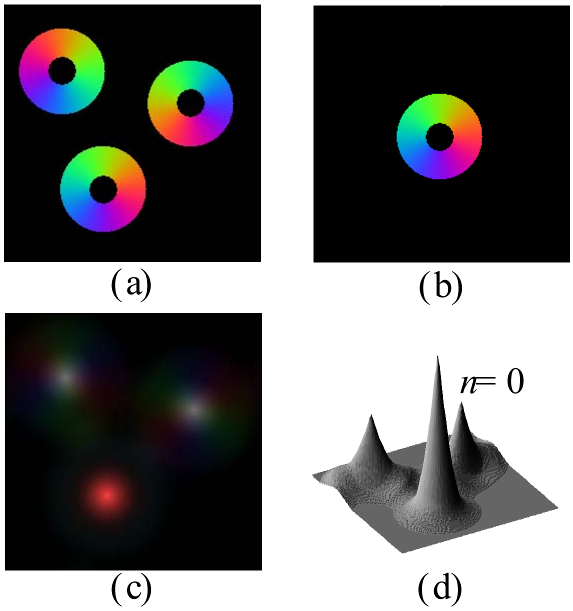

Figure 4.19. Recognition of targets with non-uniform color distribution (a) Scene, (b) target. (c) Pseudo-colored representation of the squared magnitude of the three dimensional correlation. (d) Profile of squared magnitude of the n=0 channel of the three dimensional correlation.

In the case of color images, the autocorrelation is a three dimensional function that has its maximum channel n=0, independently of the color distribution of the objects of the image. This ensures that the peak of the three dimensional correlation corresponding to the target object presents its maximum in the n=0 channel, independently of the color distribution of the target.

That is, the distribution along the color axis of the correlation is not the same as the color distribution of the objects of the scene. Therefore, even for objects with non-uniform color distributions the autocorrelation will present its maximum peak in the n=0 channel. This is illustrated in Figure 4.19. We have considered a scene with three rings with identical shape and different color distributions (Figure 4.19a). All the rings contain the same set of colors, but they are differently distributed on the object. In addition, the average[†] color of each object is the white. The target object is presented in Figure 4.19b, it matches ring placed nearest to the bottom of the scene. The squared magnitude of the three dimensional correlation has been represented in Figure 4.19c. To represent all the channels at once, we have used a pseudo-coloration in which the n=0 channel is encoded in red, and the n=1,2 are encoded in green and blue respectively . A profile of the squared magnitude at the n=0 channel of the three dimensional correlation has been represented also in Figure 4.19d. One can observe that the target object is clearly recognized by a red peak, what means that its maximum is in the n=0 channel, even when the average color of the target is the white.

Furthermore, the other objects in the scene, present very low peaks, even when they have the same colors as the target ring (but differently distributed). This is because in the three dimensional description of color images the color information is coupled with the spatial distribution. That is, it takes into account both the colors of the objects, and the way these colors are distributed in the objects.

[†] Here the average color of an object is understood as the color whose components are the average of the components of all the colors of the object.

4.8.2. *Shift Invariance along the color axis*

The three dimensional correlation is invariant to shifts of the images. That is, its values do not change, but are shifted by the same amount as the image is shifted. The correlation is shifted in the same direction if the shifted function is the first operand (scene) and in the opposite direction if the shifted function is the second operand (reference). In this sense, the three dimensional correlation is also invariant to shifts along the color axis.

Nevertheless, a translation applied to a three dimensional function has a different effect on the color image that it describes, depending on the component of the shift along the color axis. If the function is shifted along the spatial axes (or along any direction contained in the plane defined by them), the objects present in the image change their position, but they do not change their shape, and therefore one whishes to recognize them as the same object. However, if the function is shifted along the color axis, the color composition of the objects is altered, and therefore the colors change. Then, one does not recognize them as the same object anymore.

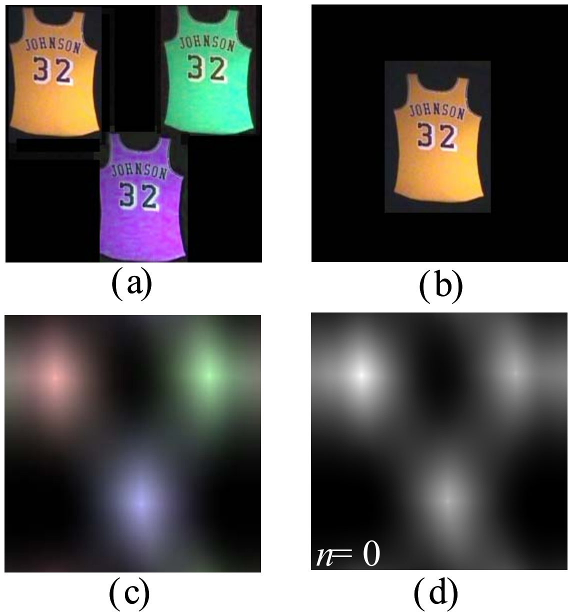

Figure 4.20. Example of color pattern recognition by three dimensional correlation. (a) Scene, (b) reference. (c) Pseudo-coloration of the squared magnitude of the three dimensional correlation. (d) Squared magnitude of the three dimensional correlation at the channel n=0

Therefore, for color pattern recognition purposes, the shift invariance of the three dimensional correlation has to be restricted to pure spatial shifts so as to discriminate objects with equal shape, and color distributions related by cyclic translations. To do this, only the n=0 channel of the correlation has to be valued.

We illustrate this in Figure 4.20. We have considered the three dimensional correlation between the images in Figure 4.20a (the scene) and b (the reference). In the scene there are three shirts with identical shape and different colors. The colors of the green and violet objects come from shifting the yellow shirt along the color axis by one channel, and two channels respectively. This scene has been correlated with a reference image that contains the target object alone (Figure 4.20b). We have represented in Figure 4.20c the squared magnitude of the obtained correlation. In the image, the n=0, 1 and 2 channels have been pseudo-colored in red, green and blue respectively. One can observe that there are three peaks, corresponding to the three objects in the scene. However, the distribution of the peaks along the channels is different for the three objects. This way, the maximum corresponding to the yellow object is a reddish peak, what is indicating that its maximum is at the n=0 channel. Because the green and violet objects are translations of the target object along the color axis, they are recognized by peaks at the n=1 and n=2 channels. If we just consider the n=0 channel of the correlation (Figure 4.20d), the target object can be discriminated from the other objects because it presents a maximum peak.

The relation between the vector space interpretation of color and the signal theory approach of color detection given in this chapter provide a new framework for the study of transformations of color images and for the design of correlation filters for color pattern recognition that take into account both the colors and the spatial distributions of images. This constitutes the subject of Chapter 6.