Università degli Studi di Parma

Dottorato di Ricerca in
Scienza e Tecnologia dei Materiali Innovativi

Tesi di Dottorato di Ricerca

Hyperspectral colour imaging and spectrophotometric instrumentation

Coordinatore:
Chiar.mo Prof. Enrico Dalcanale

Dottorando:
Andrea Della Patria

Tutor:
Chiar.mo Prof. Claudio Oleari

XXIV ciclo

Abstract

The trichromatic nature of commercial photography is strictly connected with the nature of human colour vision, although the characteristics of usual colour imaging devices are quite different from the human visual system.

The increase in the number of colour channels for spectral (either multispectral or hyperspectral) imaging is an active field of research with many potential applications in different fields.

Each element of the captured scene is specified in the spectral image by the spectral reflectance factor. This measurement is independent of the particular illumination of the scene and allows the colorimetric computation in a device-independent colour space for any chosen illuminant and any observer.

This thesis describes the project and construction of a compact spectrophotometric camera, which can be used in both portable and in-situ applications. The compactness is made possible by a suitable image spectral scanning based on an Induced Transmission Filter (ITF).

This filter is made by a set of thin-film coatings of dielectric materials with high and low refraction index, whose shape like a wedge induces a wavelength selective transmittance, continuously variable along one direction and uniform in the perpendicular direction. Such a filter, classified as Linearly Variable Filter (LVF), operates continuously from 430 nm to 940 nm and allows hyperspectral imaging.

In traditional scanners the whole apparatus is moved along a path as long as the

scene, whereas in this instrument the camera body is still and the LVF is the only moving part.

The sequence of operations for wavelength and radiometric calibrations are discussed. The expected acquisition times and number of images as a function of the spectral sampling step are considered. The resulting properties make the instrument easy to use and with short acquisition times.

Moreover, overviews of the historic evolution of colour vision fundamentals, colour spaces and spectral imaging technology are given for introducing the reader to the essential concepts useful for the understanding of the text.

Dedicated to:

Stefania and *Fabrizio*,

Claudio,

my parents,

Franco and *Marisa*,

Ivana and *Giulio*,

Massimiliano, *Ilenia* and *Benedetta*

Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my gratitude to my tutor, Prof. Claudio Oleari, whose understanding, and patience added considerably to this experience. He provided me with direction, technical support and became more of a mentor and friend, than a professor. Finally, I would like to thank Prof. Fernando Fermi for the time out from his busy schedule for fruitful discussions.

A very special thanks goes out to Dr. Angela Piegari and Dr. Anna Sytchkova, from the Enea Optical Coatings Lab in Rome, for the design and manufacturing of the linearly variable filter.

I must also acknowledge Dr. Luca Pezzati, my boss, from the National Institute of Optics (INO) of the National Research Council (CNR) in Florence for the opportunities he offered me to enter the beautiful world of spectral imaging.

Last but not least, I say "thank you very much!" to my family (Stefania and Fabrizio) for their invaluable patience and encouragement, with my whole soul best!

Introduction

Currently, *digital-era* is at the height of its fortune and colour in imaging applications has fully turned into digital, without much regret for conventional film-based photography, also thanks to the constantly improving quality and rapidly decreasing prices of digital cameras.

This is true, of course, just for trichromatic cameras, namely those which rely on three-filters technology to represent colours by means of three channels (i.e. Red, Green and Blue, hence (RGB)).

Human colour vision is trichromatic too and this fact, besides being in strict connection with the great success of colour imaging, modern television, conventional photography, computer displays, etc, forms the basis for colorimetry. If on the one hand this simplifies image capture and colour computation, on the other hand it may introduce differences from the human eye that sometimes are evident.

There is one question which excited my curiosity the first time I have been asked:

> *"Is a photograph some kind of document?"*

The answer is not a trivial matter as other persons, including people of academic rank, showed a puzzled expression and their thoughts were not of the same opinion. Often it is not enough to be a simple digital-consumer to state that the correct answer is *"no"*, since a little confidence with colour science and imaging provide the awareness of the limits of trichromatic photography.

A photograph is not a document in that it freezes the scene as it was at the instant

it was shoot, without information about the illumination/light-collection geometry and without any kind of knowledge of the spectral composition of the illumination. Most of the times, what's worse is the inability of camera's colour sensitivity curves to match the sensitivity curves of the human eye or to be represented by a linear combination.

A digital photograph is therefore a collection of numbers in groups of three, i.e. *RGB*, which return an image of a scene of course, but that are evidently meaningful only in the particular reference frame of the camera that captured it (device-dependent colour).

Photography for commercial purposes aims at returning *photorealistic* images of a scene. Many of the operations are carried out either automatically or with the active contribution by the user, for the sake of matching the result to the idea of the scene in the mind of the user himself. There are cultural reasons, psychophysical and physiologic phenomena of human colour vision that participate to form this photorealistic ideal. In particular, psychophysical and physiologic phenomena of human colour vision are classified as colour appearance and have been receiving great attention from the photographic industry for many years. For this reason, a photograph departs indeed from the purpose of documentation and, at the same time, the need to overcome this limit encourages the research for a photographic technology that considers documentation as a result of the comparison between the photograph and the captured reality.

For a digital photograph of a scene to be archived for future exact reproduction (as it was at the moment of capture) a totally different strategy must be adopted, which imposes to drift away from the trichromatic approach in favour of the spectral one (multispectral or hyperspectral).

Increasing the number of colour channels is a practice that is slowly gaining ground, although it is mainly limited to research applications in various fields. Spectral

imaging, by its nature, is not affected by the limitations of colorimetric imaging that is based on metameric matching and that suffers from the particular choice of the illuminating source.

This new and more elaborate approach represents colours by their spectral power distributions and allows for the separation of the spectral properties of the objects from the illumination through the spectral reflectance. Since spectral reflectance is independent of the characteristic of the image acquisition systems, multispectral and hyperspectral images can be transformed to any colour space and be rendered under any illumination of choice.

This thesis is about the design and creation of a spectrophotometric camera for high-quality acquisition of colorimetric and hyperspectral images. It is of fundamental importance the knowledge of all the components that participate to the imaging process and of how they affect the image quality. Moreover, these considerations become more urgent when compactness is required.

The opportunity to devise such a compact instrument came from the almost fortuitous availability of a linearly variable filter (LVF), that was deposited at the Optical Coatings Laboratory at ENEA-Casaccia in Rome. The instrument underwent several changes, every time trying to minimize the number of optical elements and of moving parts. Even though it took a great deal of time, the final concept is an elegant, compact and simple instrument with one moving part only (the LVF).

After the electronic modules for acquisition will be available (commissioned to an external factory which also performed the opto-mechanical alignment), the camera will be calibrated according to the procedures detailed in the dissertation, and its performances will be evaluated.

The first two are background chapters, providing a short theoretical background of concepts about colour and instrumentation for spectral imaging.

Chapter 1 provides an overview of colour science, encompassing major historical

premises, colour vision, colour spaces and colour computation.

Chapter 2 is a review of the most important and successful instrumentations for multispectral and hyperspectral imaging. Their basic characteristics and specifications are described, with particular focus on their operating principle. The list of such instruments is surely not comprehensive, but efforts were made to include only the instruments that introduced elements of innovation at the moment of their release, or that were particularly appreciated for the specific applications they were used.

Chapter 3 illustrates the details of the devised hyperspectral camera, from its opto-mechanical layout, through the working principle, up to the calibration procedures. Besides bibliographic cross-references scattered throughout the text, further readings are suggested in the appendixes for getting into details with important arguments.

Appendix 1 is a short review about the compression of spectral data, of the utmost importance when dealing with large amount of data, as is the case with hyperspectral imaging.

Appendix 2 explores the details of the Principal Component Analysis (PCA), a multivariate statistical technique often used in spectral imaging to reduce the number of images without sacrificing the variance of their spectral content.

Appendix 3 illustrates the physical principle underlying the Induced-Transmission filters, on which the LVF is based.

Appendix 4 deals with the calibration of trichromatic imaging devices that employ lookup tables (LUT) and how these tables can support colorimetric computation.

Appendix 5 gives some details about the main architectures of Charge-Coupled Devices (CCD), like the matrix-sensor used in the project of this camera.

Finally, the conclusions contain a discussion of possible future work and potential development for the compact hyperspectral instrumentation.

1

The limits of trichromatic photography

The advent of digital era gave a relevant boost to the availability of technologies and devices that exploit colour at different levels and with different strategies. It has become a common practice to capture scenes by means of compact digital cameras and photo-cameras, as well as making hard copies by compact digital printers, thus avoiding costly equipments and the usage of turning to the photographer. This widespread fan of these brand new possibilities was not for free, as it introduced severe limitations that are inherent in the technology itself. Furthermore, the *for-everyone* attribute made things even more complicated. In fact, for example, prior to digital era the art of photography was confined to a limited number of professionals or amateurs who were quite skilled in the management of the tools they were equipped with. Currently, it is not rare to hear non-professional users (including teenagers and kids) discussing about colour, mostly about their perception of colours as reproduced by their user-friendly equipment. If from a general point of view this is an undeniable improvement, nonetheless it opens several problems concerning the correct awareness and interpretation of the output by the end users. One of such problems is met when one faces the reproduction of colors, either on prints or viewing media (i.e. monitors,

display, projection, etc.).

1.1 The reproduction of colour

To this purpose, it is worthwhile to point out that different levels of classification exist in the practice of colour reproduction, according both to the fields of application and to the manifold techniques used to achieve it. In particular, two sets of reproduction modes can be identified: in the first, the measuring instrument plays a fundamental role and some original quantities remain unaltered with reproduction; thus, depending on the particular case, the reproduction can be:

- *spectral*
- *colorimetric*
- *exact*

In the second set, none of the quantities are reproduced exactly, as colour appearance itself is the measurement; as a consequence, the reproduction is said to be *photo-realistic*, so can be:

- *corresponding*
- *equivalent*

A *spectral* reproduction, which is typical of the paint industry, is achieved when reflectance is reproduced in such a way that a coloured object is not distinguishable from its original under all the illuminant and for every viewing position.
In *colorimetric* reproduction, which is the main feature of print industry, the chromaticity and the relative luminance of coloured objects constitute the task of the reproduction. In this case, when viewed in reflection, the reproductions may suffer

from *metamerism*[1], as the illuminant varies.

A *colorimetric* reproduction turns to *exact* when also the absolute luminance is included, provided that equal viewing conditions are ensured for both the original and its reproduction (Oleari, 2008).

It depends upon whether the levels of luminance in the original and its copy are different or not that photo-realistic reproductions are named *equivalent* or *corresponding*, respectively.

Great colour differences may occur under these circumstances; in particular, it is a common practice to prefer reproductions at higher contrast and more colourful than the originals.

In every field of colour reproduction typical criteria were defined and accepted; as an example, those criteria that are valid for the television must not be confused with those accepted in the paint and plastic industry.

1.2 Colour

A seemingly easy question to answer, not so much indeed, is the following:

"What is *colour?*"

As far as colour vision is concerned, it is of great relevance to point out that colour is an attribute of visual sensation; as such, it is an extremely subjective entity and, above all, is incommunicable. However, the properties of the mechanisms ruling colour vision and the relationships between physical stimuli and the responses of these mechanisms can be described yet. Furthermore, while colour itself is incommunicable this does not prevent different people agreeing about "colour-matching", namely agreeing that two physically different stimuli look chromatically the same. This constitutes the basis of colorimetry.

[1] METAMERISM is a property of two matching colour stimuli with different spectral radiances

Speaking of colour, it is not unusual to run into misunderstandings because the word *colour* is given so many meanings, i.e. paint, *RGB*- or CIE-values, perceptual sensations, spectral radiation, colour system notation, etc. Here is given maybe the most widely accepted technical definition of colour, issued from the Committee on Colorimetry of the Optical Society of America in 1940:

> Colour consists of the characteristics of light other than spatial and temporal inhomogeneities; light being that aspect of radiant energy of which a human observer is aware through the visual sensations which arise from the stimulation of the retina of the eye

At least three fine different aspects of reality can be indicated by the word *colour*, as it can describe the property of an object, the characteristics of light rays and a set of sensations. It is up to the speaker's ability to employ the right words if he is emphasizing one of the three aforementioned particular meanings. Except for extraordinary cases, when the difference is critical, commonly the distinction among these three meanings is disregarded in actual practice. It seems that the identification of the intended meaning relies upon the mere knowledge of the existence of these different aspects. This can save plenty of meaningless words, mostly in the case of philosophic discussions about colour that result in raising the confusion about these three aspects. The efforts towards simplification of the definition of colour inspired Dean B. Judd (a great contributor to colour science) to formulate the following sentence with words having less than four syllables:

> Colour is that aspect of the appearance of objects and lights which depends upon the spectral composition of the radiant energy reaching the retina of the eye and upon its temporal and spatial distribution thereon

One of humans' most important senses is vision, which is stimulated by light and colour. This fascinating interaction arose the thinking of many philosophers like Plato

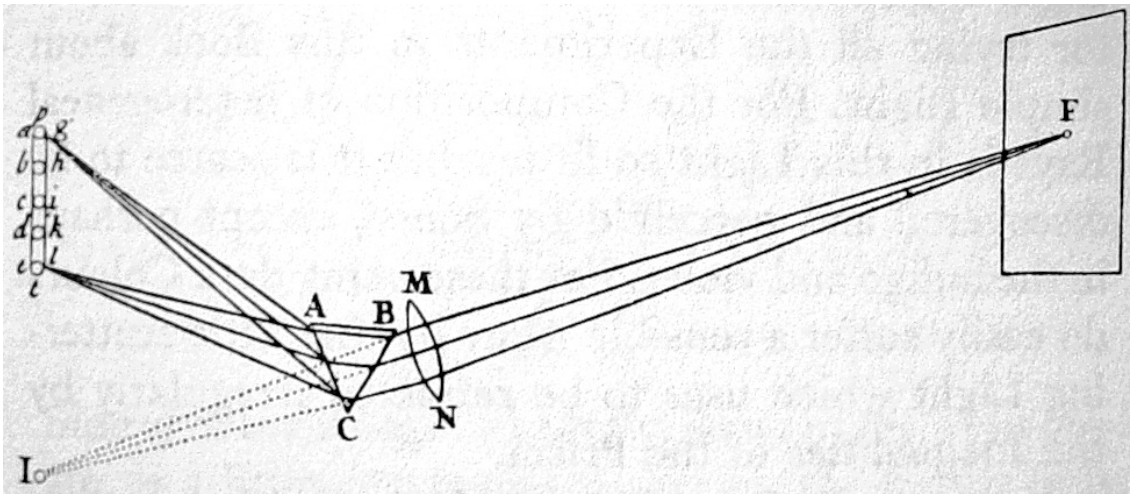


FIGURE 1.1: Sketch of Newton's experiment yielding the *spectrum* in 1666

and Aristotle; in particular, the former was sceptic about the possibility to verify experimentally how colours were made, whilst the latter saw colour as a mixture of black and white, thus disregarding the distinction of colour from light. These words from Aristotle are significant:

> Thus pure light, such as that from the sun has no colour, but is made colored by its degradation when interacting with objects having specific properties which then produce colour

Everything changed when Isaac Newton discovered the spectrum at the middle of the XVII$^{\text{th}}$ century, as explained in his treatise OPTICKS (1704). After him, THE THEORY OF COLOURS of the great German poet Johann Wolfgang von Goethe retained some points of view of Aristotle's approach to colour.

The science of colour can be dated back to 1672, when Newton published his work on PHILOSOPHICAL TRANSACTIONS containing the description of his famous experiment about the production of the *spectrum* (Figure 1.1).

To accomplish his experiment, Newton turned to a sequence of seven colours, i.e. red, orange, yellow, green, blue, indigo and violet which he recombined with a lens

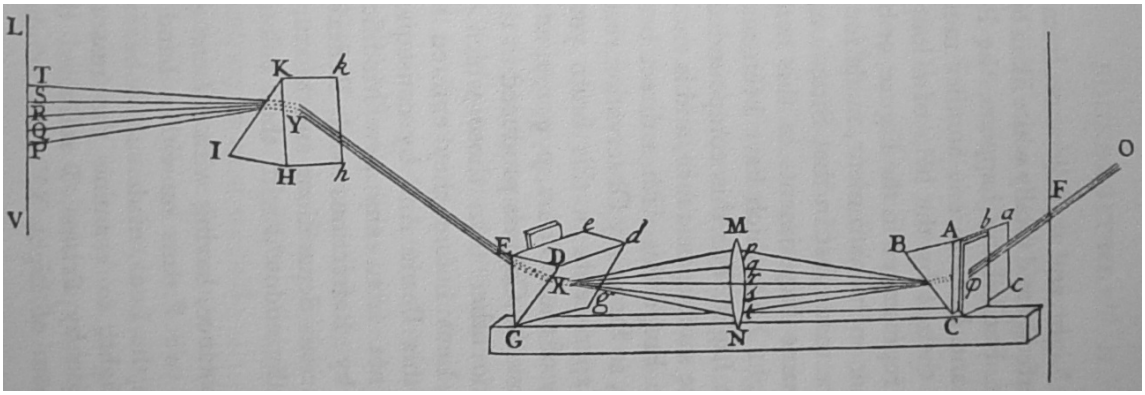


FIGURE 1.2: This is how Newton recombined the *spectrum* to get white light

or a second prism to obtain white light again (Figure 1.2). This way he proved that sunlight consisted of a mixture of colours. It was just by a curious analogy with the seven musical notes that Newton arbitrarily assumed seven fundamental colours and organized them in a color wheel fashion, where white is at the center of Figure 1.3; colours on the rim linked by a straight line through the center are said to form a *complementary pair*. Humans can perceive not only the spectral *hues* of Newton's colours, but various combinations of them as well. This is the reason why in Newton circle the pinks, purple and browns are not encountered.

It is quite clear that the brain interprets the signals, coming from the spectrum perceived by the eye, as colour. As such, the ambiguity involving colour as a physical property of objects and as a peculiar phenomenal experience of the observer, going along with the development of the knowledge of human visual system, gave rise to much dispute. However, the immediacy of colour sensation, as an external inherent quality of the material, is purely phenomenal. It has been scientifically demonstrated that the perception of colour is based upon a complex chain of interventions, which includes radiant energy absorption, the encoding of the chromatic signal and its representation within a model of the visual scene. Thus, at the same time colour is immediate for the observer as well as mediated by the observer itself. Getting into

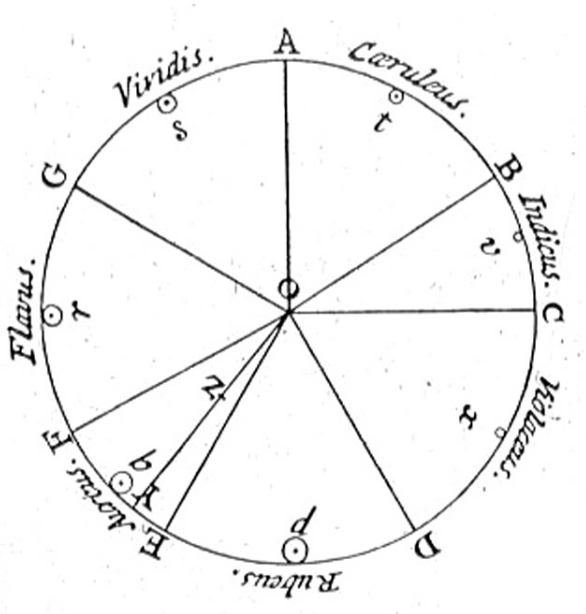


FIGURE 1.3: Colour wheel of Newton with colours shown with letters from A through G

details about the psychologic interpretation and implications of colour is beyond the scope of this dissertation.

Newton sensed that colour was not an attribute of light, and here is explained with his own words:

> And if at any time I speak of light and rays as coloured or endued with colours, I would be understood to speak not philosophically and properly, but grossly, and accordingly to such conceptions as vulgar people in seeing all these experiments would be apt to frame. For the rays to speak properly are not coloured. In them there is nothing else than a certain power, and disposition to stir up a sensation of this or that colour.

Newton's theory of light and colours has survived for more than 300 years, and constitutes the foundation of modern physics. Its most important theoretical tenets have resisted the flow of time, theoretical and empirical assaults and have emerged victorious and unchanged. The theory is written in textbooks and taught in universities that seems of uncontestable validity. But if that were the case, one would be

wrong. Newton did not always distinguish carefully what is meant as the physics underlying color, which is a physical optics of radiation, from the psychophysics of color, which constitutes the link between the kind of amount of radiation with what is seen; therefore this inevitably draws into the circle of color science matters of psychology, physiology and perception. People studying the nature of colours, philosophers and artists consider Newtons theory wanting and so they brought into question, because for the most part Newton tried to give the psychophysics of colors a purely physical basis, and thus he really did not provide a comprehensive foundation for the science of "color qua color" (Sepper, 2003).

1.3 Colour matching

Let a screen be divided into two contiguous halves, each illuminated with radiations of different spectral composition. Generally, the output colours are different but two spectrally different power distributions may give rise to identical colour sensations. Such a pair of stimuli is said to be *metameric*, and the metameric equality is indicated by the symbol $\equiv$. For example, the yellow colour of a radiation of 580 nm matches the mixing of the two radiations of 540 nm and 620 nm, each with an appropriate relative intensity. It must be pointed out that that a generic *spectral* colour, i.e. a colour relative to a wavelength belonging to the visible spectrum (classically, from 380 nm to 780 nm), cannot be matched by the mixture of two monochromatic radiations, with the exception of spectral light of medium and long wavelength. The physical parameter that allows the description of the effects on the screen is the *luminance*, which is involved into the metameric matching through the equality of the luminances of both sides. The Abney law states that the luminance produced by the sum of two radiations is the same as the sum of the luminances of the same radiations.
Given two monochromatic radiations λ_1 and λ_2, their most generic combination is

the following:

$$L(\lambda_1) + L(\lambda_2) \equiv L(\lambda) + L(\lambda_W) \tag{1.1}$$

Here $L(\lambda_W)$ is the luminance of a white source, whereas $L(\lambda_1)$, $L(\lambda_2)$ and $L(\lambda)$ are the luminances of three monochromatic sources with wavelengths λ_1, λ_2 and λ, respectively. λ is said to be the *dominant* wavelength of the given colour, which relates with the *hue*. The subjective attribute of *saturation* of a colour, namely how much it differs from from the white, can be expressed by the *purity*:

$$\frac{L(\lambda)}{L_W + L(\lambda)} \tag{1.2}$$

By definition, the colours of maximum saturation are the spectral colours and the magenta. For the sake of generalization, the sum of an arbitrary number of monochromatic radiations is represented by the following relationship:

$$\sum_{i=1}^{n} L(\lambda_i) \equiv L(\lambda_{n+1}) + L_W \tag{1.3}$$

The higher the number of radiations entering the combination, the lower the saturation of the resulting colour (i.e. the closer to the white W).

Leaving the luminance out of consideration, if only hue and saturation are taken into account, the overall range of chromatic sensations that relate to *"lights"* can be thought of as a locus, in a two-dimensional diagram (Figure 1.4), with spectral colours and the magenta on the rim and the white either in a central or nearly-central position; furthermore, a line connecting the white and a saturated colour of the rim spans the whole range of more or less saturated colours with the same hue (Oleari, 2008). As an illustration of what was said before, Figure 1.5 shows the principle of trichromatic matching that can be easily explained by the trichromatic theory of colour vision.

According to it, the human retina analyzes colours into three kinds of responses, ρ,

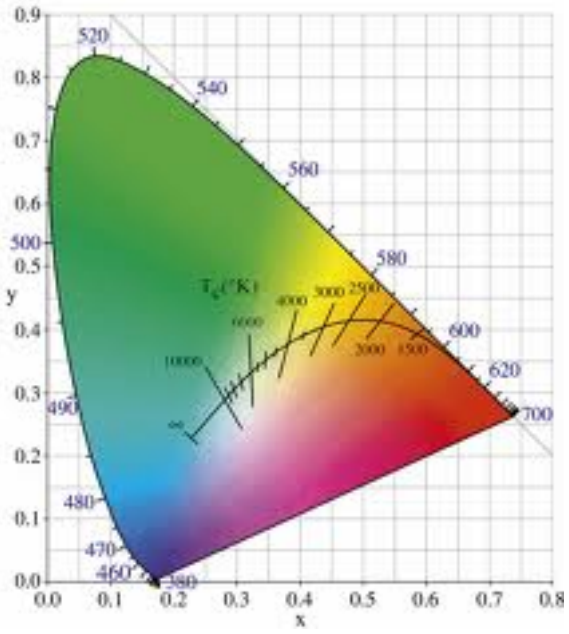


FIGURE 1.4: Chromaticity diagram (spectral locus) from CIE 1931

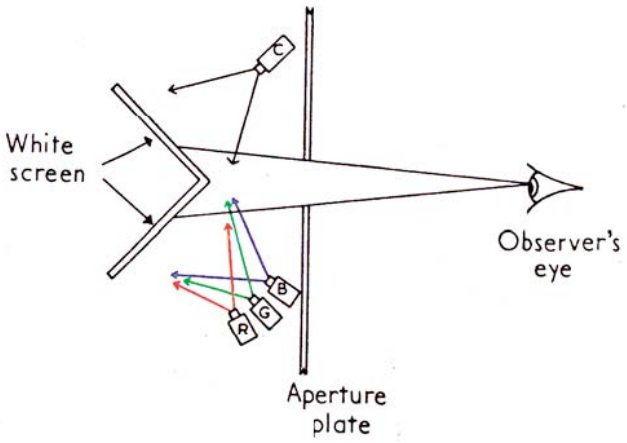


FIGURE 1.5: Sketch of the setup for the trichromatic matching

γ and β (this is the nomenclature chosen by R.W.G. Hunt in the reference (Hunt, 2005), in place of the most common $\bar{l}$, $\bar{s}$ and $\bar{m}$), each proportional to absorptions in three different photo-sensitive pigments, and the eye cannot distinguish stimuli giving rise to the same ρ, γ, β values, regardless of their spectral composition. Such stimuli, being visually identical but spectrally different, are called *metameric pairs*.

1.4 Trivariance of chromatic vision. Theory of Young-Helmholtz

The metameric equality (1.1) demonstrates the *trivariance* of chromatic vision, because the left side depends on four variables (two wavelengths and two luminances), whereas the right side only on three (the dominant wavelength and two luminances). What is called a *colour stimulus* is a function of physical quantities, such as the total luminance, the dominant wavelength and the purity. The subjective correlates are: *brightness*, *hue* and *saturation*.

Experimental evidences support the hypothesis that the whole range of chromatic sensations depends only on three variables. It follows that the colour of any radiations can be matched by the weighted sum of three arbitrarily chosen primary radiations (*primary colours* or *primaries*), provided that, with regard to colour:

- none of them results from the sum of the other two
- no pair is *complementary*, i.e. the sum of the luminances of the two radiations of the pair does not match the luminance of the white

As a rule, the choice of the primaries is made by maximizing the mutual distance of the corresponding dominant wavelengths in the visible spectrum (red, green, blue-violet). This principle underlies *additive mixing* colour systems, as is the colour television, where the combination of three primaries occurs by spatial integration of the three-colour phosphor dots in the visual system of the observer.

Since the XVIII$^{\text{th}}$ century the trivariance of chromatic vision suggested that the

human eye accommodated only three kinds of sensors, each for a particular spectral interval. This is the hypothesis formulated by Thomas Young in 1801, later continued by Hermann von Helmholtz in the second half of the XIX$^{\text{th}}$ century, according to which the three theoretical neural mechanisms, having peak sensitivities in the red, green and violet, combined linearly in proportion to the respective excitations by the radiation entering the eye. Thus the equienergetic radiation would give a sensation of white, since it equally stimulates all three mechanisms, whereas a radiation stimulating only two mechanisms would produce a sensation of yellow (red and green) or green-blue (violet and green) or magenta (red and violet), and so on.
This hypothesis, also known as the YOUNG-HELMHOLTZ theory, had to wait long for experimental validation; in fact, it was only in 1964 that two groups of american scientists demonstrated the existence of three different kinds of cones through micro-spectrophotometric absorption measurement on single cones of humans and monkeys (Marks et al., 1964; Brown and Wald, 1964). Later on these results were further confirmed through electro-physiologic experiments.

1.5 Colour vision

The colour of objects strongly depends upon their spectral reflectances[2], i.e. by the amount of incident light that the surface of the object reflects at different wavelengths. Let $l(\lambda)$ and $r(\lambda)$ be the spectral radiance of the illuminant and the spectral reflectance of a given surface point of the object, respectively. So the radiance of the light reflected from this particular surface point is $f(\lambda) = l(\lambda) \cdot r(\lambda)$, where only the dependence upon the wavelength λ has been emphasized, ignoring the effects of the spectral reflectance of an object on the angles of incidence of light and observation. The limitations of this model include the specular reflection, fluorescence, polarisation,

[2] it should be rather used the term *spectral reflectance factor* (CIE, 1989, 1986), but here any distinction is disregarded

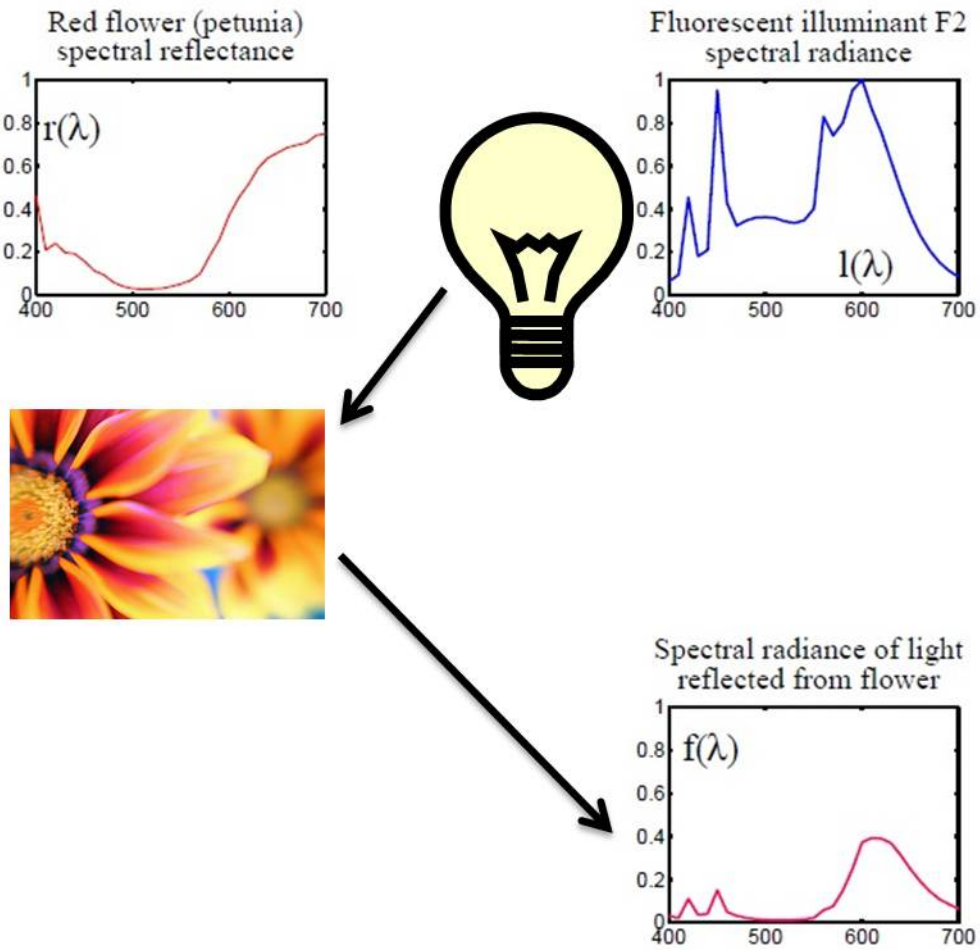


FIGURE 1.6: An oversimplified model for light-surface interaction

etc. Nonetheless, although severely limited, this model of light-matter interaction is enough for the analyses that follow.

The lens of the eye focuses light onto the retina to form images. *Rods* and *Cones* are the photoreceptors that are responsible for night (scotopic) and daylight (photopic) vision, respectively. Let $f(\lambda)$ indicate the spectral distribution of light crossing a given point in the retina, a vector $\mathbf{c} = (c_1, c_2, c_3)^T$ suffices to represent the responses of the cones, with:

$$c_i = \int_{\lambda_{min}}^{\lambda_{max}} f(\lambda) \cdot s_i(\lambda)\, d\lambda, \qquad i = 1, 2, 3 \tag{1.4}$$

where $s_i(\lambda)$ represents the sensitivity of the i^{th} cone.

Quantities under the integral are known at discrete points, therefore a vectorial formalism can be introduced:

$$\mathbf{c} = \mathbf{S^T} \cdot \mathbf{L} \cdot \rho \tag{1.5}$$

where:

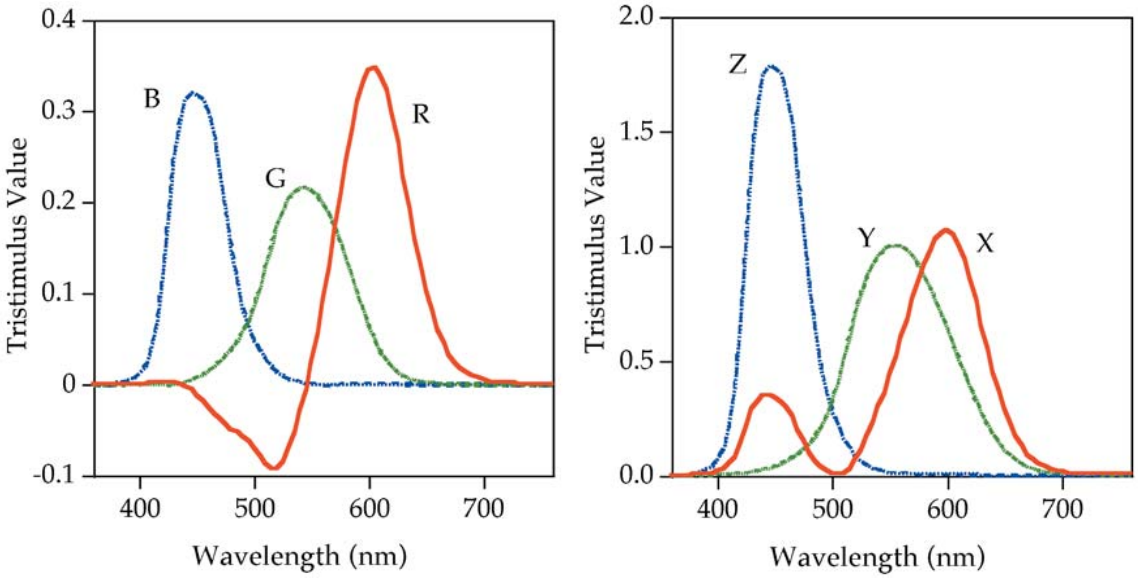
- $\mathbf{S} = (s_1, s_2, s_3)^T$ represents the matrix of sensitivity of eye's sensors, with $\mathbf{s_i} = (s_i(\lambda_1), s_i(\lambda_2), s_i(\lambda_3) \dots s_i(\lambda_N))^T$, λ_1=λ_{min}, λ_{N}=λ_{max};
- the diagonal illuminant matrix $\mathbf{L}$ has entries from the samples of $l(\lambda)$ along the diagonal;
- ρ represents the matrix of the sampled spectral reflectance of the object.

The sampling interval, namely the spectral spacing between two adjacent wavelengths, is: $\delta\lambda = \frac{1}{N-1}(\lambda_{\text{max}} - \lambda_{\text{min}})$.

The three types of cones define a Hilbert space, named *Human Visual Subspace* (HVSS), and so the mechanism ruling the cone responses can be considered as a projection of the incident spectrum onto the space spanned by the sensitivity functions $s_{\text{i}}(\lambda)$, where i=1,2,3

The vector space spanned in the sampled case by the columns of $\mathbf{S}$ is just the HVSS. Although the cone response functions are hardly measured directly, nonetheless colour-matching experiments yield non-singular linear transformations of these cone responses. In colorimetric definitions exists a standardised set of colour matching functions $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$ recommended by the CIE. The representation of a colour stimulus $\mathbf{f}$ by means of its CIE *XYZ tristimulus values* $\mathbf{t}$ is thus possible through the colour matching functions, provided that $\mathbf{A} = (\bar{x}, \bar{y}, \bar{z})$ represents the matrix of sampled colour matching functions

$$\mathbf{t} = \mathbf{A^T} \cdot \mathbf{f} \tag{1.6}$$



(a) CIE *RGB* colour matching functions (b) CIE *xyz* colour matching functions

FIGURE 1.7: Spectral tristimulus values for the CIE *RGB* system of colorimetry and of the CIE 1931 standard colorimetric observer

The linear model expressed by (1.4) does not account for the effect of *colour constancy*, namely the invariance of the perceived colours of objects as the illuminant varies, and must be considered as a description of just a small part of a more complex colour-perception process (Hering, 1905; Judd, 1933). Moreover, the attributes of *hue*, *saturation* and *lightness* are not in direct relationship with the cone responses; further information about human visual system is thoroughly detailed in the following books (Wandell, 1995; Kaiser and Boynton, 1996).

1.6 A brief history of trichromatic photography. MAXWELL's experiment

The art of colour photography and colour printing can be dated back to the XVIII$^{\text{th}}$ century, after the observations made by the german painter and engraver J.C. Le Blon who was making prints by the mezzotint process. His early attempts to make colour prints involved printing from seven plates with inks having the

hues of the seven spectral colours of Newton's ring. It was later in 1722 that he concluded that just three plates with red, yellow and blue inks sufficed to reproduce all colours. Another fundamental step in the development of three-colour processes was made by the great Scottish mathematician and physicist J.C. Maxwell in 1855, whose photographic procedure employed three negatives of the subject, each exposed through an appropriate colour filter. These colour-separation negatives served the purpose of automatically controlling the amounts of the three primaries at every single point of the positive reproduction. The two basic principles aforementioned underlie all methods of three colour reproduction devised from then on, both in photography and in the graphic arts.

Maxwell's own words here reported summarize his suggested procedure:

> "Let it be required to ascertain the colours of a landscape by means of impressions taken on a preparation equally sensitive to rays of every colour. Let a plate of red glass be placed before the camera and an impression taken. The positive of this will be transparent wherever the red light has been abundant in the landscape, opaque where it has been wanting. Let it now be put in a magic lantern along with the red glass and a red picture will be thrown on the screen. Let this operation be repeated with a green and violet glass and, by means of three magic lanterns, let the three images be superimposed on the screen. The colour of any point on the screen will then depend on that of the corresponding point of the landscape and, by properly adjusting the intensities of the lights, etc., a complete copy of the landscape, as far as visible colour is concerned, will be thrown on the screen. The only apparent difference will be that the copy will be more subdued, or less pure in tint, than the original. Here, however, we have the process performed twice; first on the screen and then on the retina."

This quotation is based on an outline of Young's theory of colour vision and the process recalled in the last sentence is just the one Young assumed to occur in the human eye.

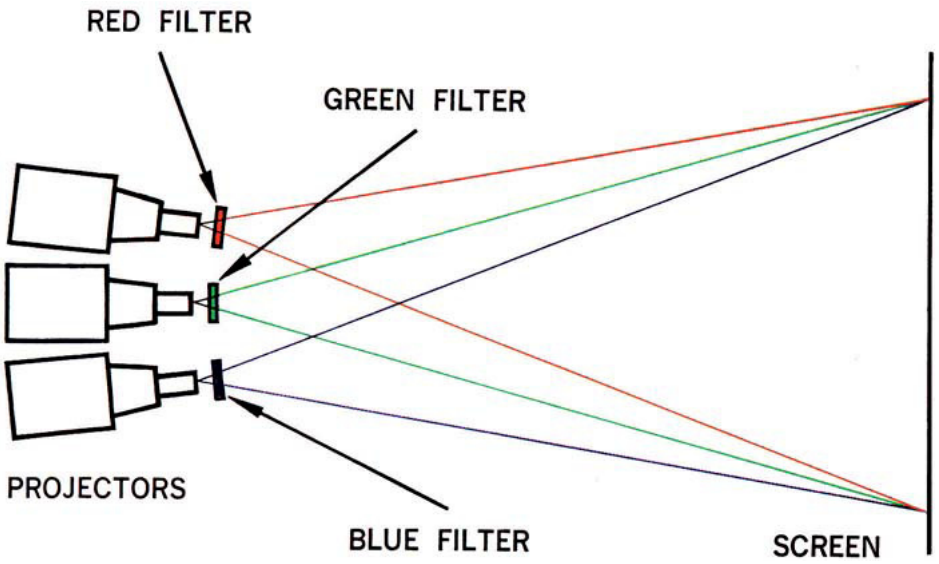


FIGURE 1.8: Triple projection for additive colour reproduction

Maxwell evidently sensed that the characteristics of the visual apparatus of the observer are involved in three-colour reproduction. So, it can be summarized that in 1861 he produced the first trichromatic colour photograph not, curiously, for its own sake but as an illustration of the triple nature of colour vision. Nowadays it is apparent that Maxwell's method is a fundamental one to all modern processes of colour reproduction. In other words, he made three positive slides from the negatives obtained after taking three photographs through a same number of coloured filters; he put these slides in three different projectors geometrically arranged so as to have the three images registered on a white screen (Figure 1.8). The colour reproduction obtained upon the screen was a result of the projection of every single slide, made from the negative taken through each coloured filter, with the same corresponding filter in place. The whole set of colours on the screen resulted from mixtures of red, green and blue light only; however, to the eye white, yellow, mauve and a whole range of pale and vivid colours were seen as well. To be more precise, since the entire light spectrum is made up of mixtures of these three primaries, red and green together yield

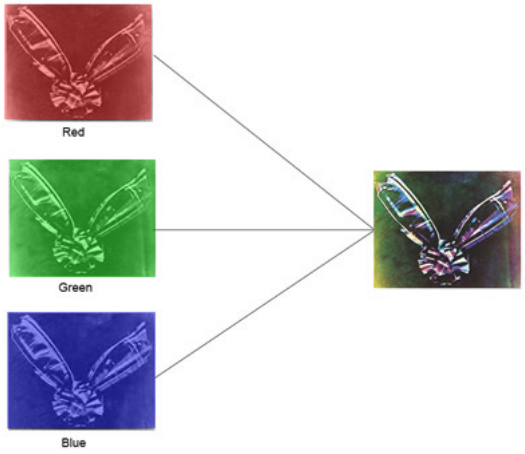


FIGURE 1.9: Maxwell's experiment tartan ribbon split into primary colour images

yellow, red and blue yield magenta, blue and green make cyan; therefore, the three primary colour images, when perfectly aligned, produced the first colour photographic image of the history (Figure 1.9).

Maxwell shot and projected a yellow image, in addition to the three images aforesaid. The reason why he used the extra filter is not fully clear, but its likely purpose was to smooth the action of the blue rays, since photographic film was particularly sensitive to blue at the time.

Current colour reproductions only vaguely recall Maxwell's method, but his principle is retained almost universally and modern resources permit very high quality results. It must be emphasized that the application of trichromatic principles to colour reproduction depends only on the experimental evidence that a wide range of colour can be produced by mixing beams of the three primaries, not by any physiological reason. The human retina contains two kinds of light-sensitive receptors, namely *rods* and *cones*. By the XX$^{\text{th}}$ century it was recognized that at very low levels of illumination (starlight, moonlight, etc.) the responsible for colourless vision was the bleaching of *visual purple*, a photo-sensitive substance contained in the rods. Accordingly, it was natural to think of colour vision as a result of bleaching of three different

photo-sensitive substances contained in the cones. Being the latter less sensitive and numerous than the rods and offering much less photo-sensitive material, the evidence of this view was difficult to come by. Measurements have been made showing that, at high levels of illumination, the retina reflects off spectral radiation through the pupil of the eye, in the same fashion that would be expected if three such pigments were being bleached (Rushton, 1957, 1958; Weale, 1959; Ripps and Weale, 1963; Brown and Wald, 1964; Mitchell and Rushton, 1971). Obviously, these pigments would absorb reddish, greenish and bluish light, respectively.

These pigments are present in separate cones and, after they absorb light, electric signals result in the form of nerve impulses; these signals convey to the brain the colour and other information concerning the images formed on the retina. The extremely intricate net of interconnections between nerve fibers, along which the electric signals travel, mostly in the region where they are joined to the rods and cones, complicates the colour information. In fact, after this stage, it cannot be thought anymore like three signals representing the absorption of light by the three pigments. In human vision (Hurvich and Jameson, 1957; Hunt, 1982), as well as for other animals, the transmitted signals are in strict connection with luminance and colour-difference signals employed in colour television (Svaetichin and MacNichol, 1958; MacNichol, 1964). Once the light has been absorbed by the retinal pigments, equal processing of the output signals correspond to equal amounts of absorption, provided the viewing conditions are the same. It is in these absorptions, occurring at the stage of retinal pigments, that the key to the success of trichromatic colour reproduction resides.

The deep comprehension of the retinal absorption stage results from the knowledge of the spectral sensitivity curve of each of the three-pigments mechanisms, that is from the ability of each of the three pigments to absorb light and be bleached, as the wavelength varies throughout the visible spectrum.

The results returned by direct methods of measuring them, based on the evalua-

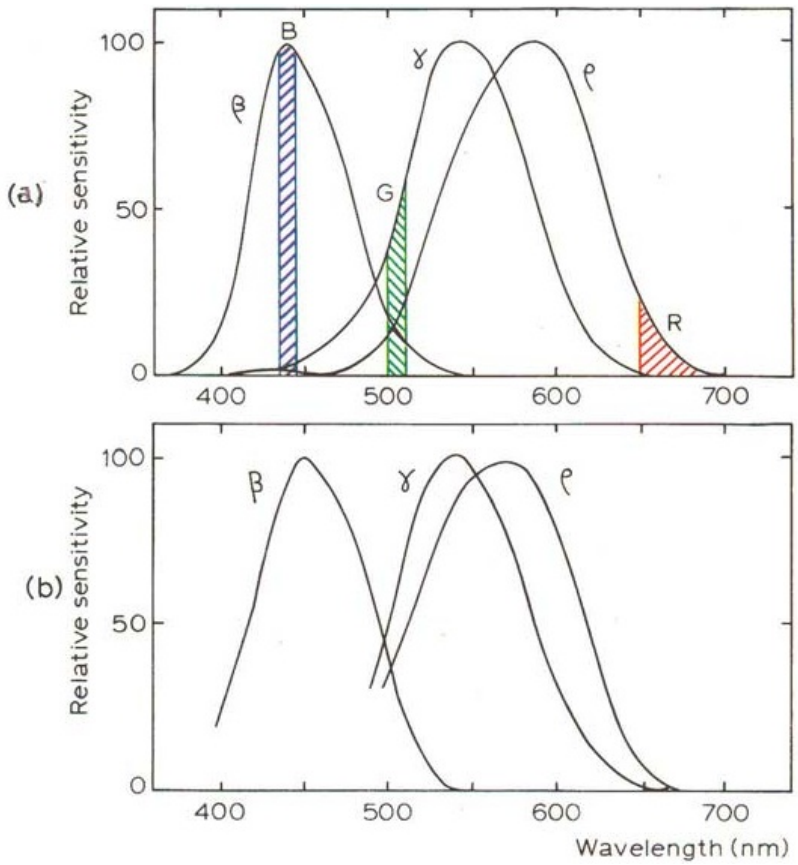


FIGURE 1.10: (a) The probable sensitivity curves α, β and ρ of the three types of light receptors believed to be responsible for colour vision as determined by indirect methods, together with the spectral quality of the three best lights R, G and B for additive colour reproduction (b) Spectral sensitivity curves typical of those found from bleaching experiments on pigments in the human retina

tion of light bouncing back through the pupil of the eye, slightly differ from those obtained by indirect methods, which have provided good approximations to these curves for many years. The complications arise in the blue part of the spectrum where absorption, in variable amounts, occurs by some of the products of bleaching, though not generating corresponding colour responses. That is why the results for the blue-absorbing pigments are less certain than those for the red-absorbing and green-absorbing pigments, and sometimes the results for the blue-absorbing pigments are not communicated.

The best set of curves representing the visual response to light of different wavelengths

(*action spectra*), for light crossing the cornea of the human eye, is not clearly defined yet. In the literature, such curves are peaked at about 440 nm, 535 nm and 565 nm (Smith and Pokorny, 1975); in Figure 1.10 is reported a set peaked at about 440 nm, 545 nm and 580 nm. Even though they exhibit some differences, these set of curves are similar enough for discussion without loss of generality. It must be pointed out that the sensitivity curves overlap to a relevant extent; thus one type, covering mainly the red, orange and yellow parts of the spectrum, another the orange, yellow, green and blue-green parts, and the third covering the blue-green, blue and violet parts. These three kinds of sensitivities can be regarded as belonging to three different types of cone, i.e. ϱ, γ and β, respectively. Physiological studies of the retina demonstrate that only one out of the three types of pigment is contained within each cone (Marks et al., 1964) but if there were some cones with more than one pigment, having sensitivity curves as mixtures of those shown in Figure 1.10, the considerations would be unaffected.

Returning to Maxwell's method, for a correct reproduction the photographic film must be able to analyze a scene just like the eye would do, when taking three negatives. Therefore, it is apparent that the curves of Figure 1.10 should represent the three film-filter combinations. The real looks of red and green filters is orange and pale-green, respectively, due to the broad nature of ϱ and γ curves. Looking at Figure 1.10, it is clearly seen that no spectral regions exist to which only the γ cones respond, so no filter can be found that transmits light stimulating the γ cones only. A good solution for the green would be the choice of a narrow-band transmission filter centered at a wavelength of about 510 nm, whereas B and R for the blue and red, respectively (shown as G, B and R in Figure 1.10). However, this choice of the green filter carries an undesired amount of ϱ and β responses which causes the greens to look paler and the whites to exhibit a magenta tinge. The choice of the filters reported in Figure 1.10 would be inefficient in terms of transmitted energy because

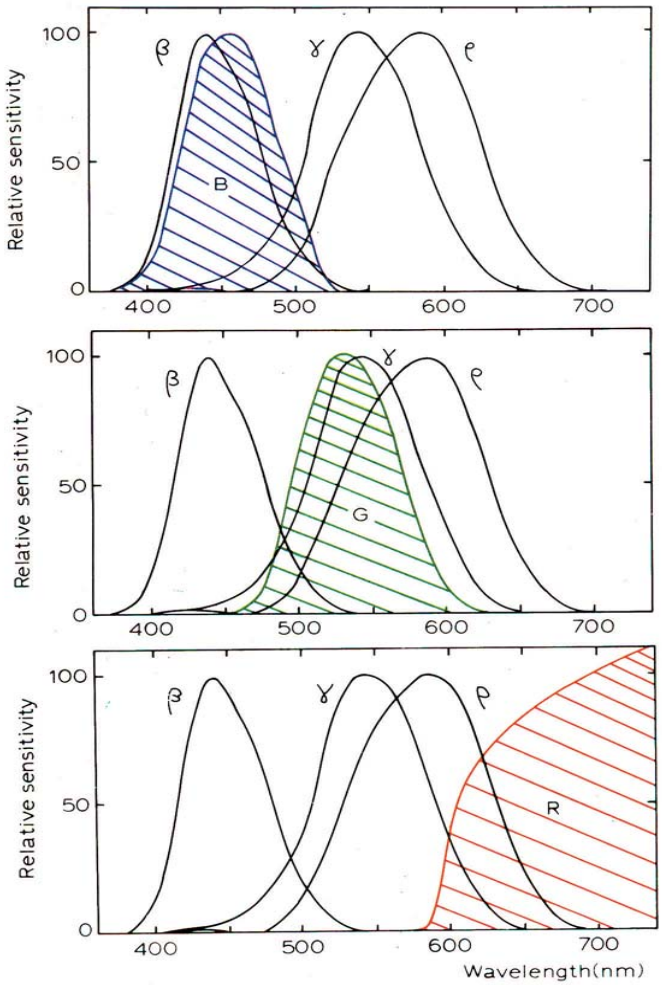


FIGURE 1.11: The ρ, γ and β sensitivities of the eye, along with the spectral powers of light transmitted by red (R), green (G) and blue (B) filters (like those typically used in additive colour reproduction)

of absorption losses. To produce a brighter picture on the screen by throwing a higher amount of light with each projector, a good practical compromise is to employ broader transmission filter sets, like those shown in Figure 1.11; this strategy has the drawback of introducing further inaccuracies in colour rendering, since the light from each projector will introduce a higher amount of the unwanted cone responses than in the case of the filter set of Figure 1.10(a). When the primary lights are produced, as in colour television, by the excitation of phosphors the situation is not very different

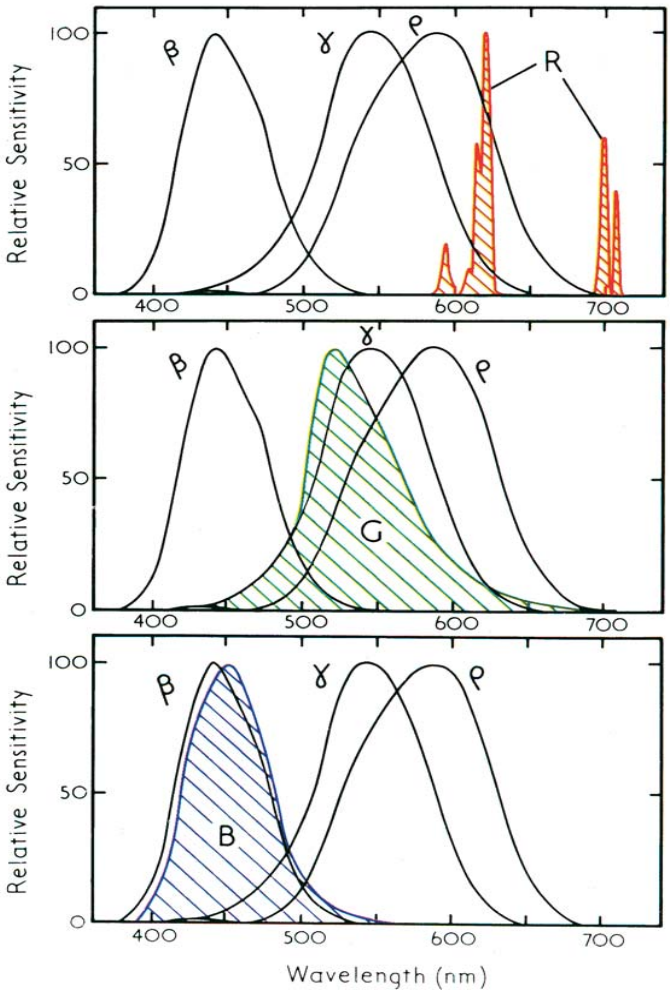


FIGURE 1.12: The ρ, γ and β sensitivities of the eye, along with the spectral powers of light emitted by red (R), green (G) and blue (B) phosphors typical of colour television

from that in Figure 1.11, as illustrated in Figure 1.12. Now it is apparent that retinal cones cannot be stimulated separately by any red, green and blue beams and this constitutes a problem for trichromatic colour reproduction. On the one hand, the inherent overlapping of sensitivity curves helps in detecting changes in hue throughout the spectrum, for colour vision, on the other hand simple trichromatic methods do not achieve correct colour reproduction of all colours. So the unwanted stimulations cannot be eliminated, as they stem from the basic nature of human colour vision.

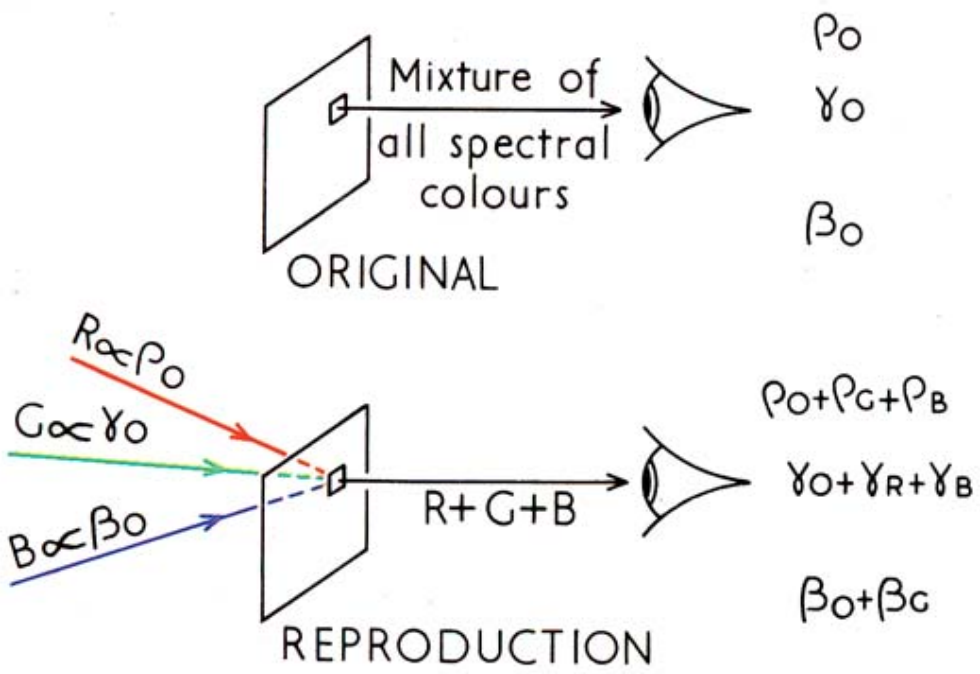


FIGURE 1.13: Representation of how unwanted stimulations affect three-colour reproduction of an original scene

In Figure 1.13 a particular part of the original beam induces the responses ρ_0, γ_0 and β_0 and if the strengths (R,G,B) of the primaries of this part of the reproduction are proportional to these responses, then the reproduction is altered because of consequent unwanted responses (γ_R from R, ρ_G and β_G from G, ρ_B and γ_B from B). As a consequence, instead of producing the original responses (ρ_0, γ_0, β_0), the amounts of R, G and B induce:

$$\begin{cases} \rho = \rho_0 + \rho_G + \rho_B \\ \gamma = \gamma_0 + \gamma_R + \gamma_B \\ \beta = \beta_0 + \beta_G \end{cases}$$

The amounts of red, green and blue light are then reduced appropriately to reproduce neutral colours correctly, namely when ρ, γ and β are equal one another. Other colours turn their appearance to whites or greys (i.e. become less colourful) as a result of the presence of unwanted stimulations. It is because of the overlapping of

cone sensitivity curves (Figure 1.10) that correct colour reproduction by trichromatic methods is made impossible for all colours.

1.7 Colorimetry

Nowadays colorimetry, namely the science of measuring colour, is still an open chapter of colour science and its historical phases have to be presented individually. For the sake of clarity, a summarizing scheme (Table 1.1) is proposed, which shows that every moment of colorimetry is associated to its own historical phase, to a vision stage, to a standard observer and to CIE systems.
The ultimate task of colorimetry is to provide colour specification by appropriate numbers; this may be achieved through different ways and meanings.
Here is a list of colour presentation modes of both practical and colorimetric interest:

- *Illuminant-mode*: for the direct vision of a light source
- *Object-mode*: for the vision of an illuminated object characterized by either diffuse reflection or transmission of radiation
- *Aperture-mode*: for the vision through an aperture that ignores whether the radiation entering the eye is associated to an object or a source

The colour of an emitting body is said *self-luminous*, otherwise *surface* colour. In aperture-mode is possible to specify colour numerically (*psychophysical* colour) if only the intensity and the spectral decomposition of the radiation are considered. Whenever colour results from perception, not from instrumental measurements, it is named *perceived* colour. Vision is a rather complex phenomenon and includes colour sensation; besides colour, there are shapes, movements, shades and lights, surface textures, light polarization, gloss and others. It is the mutual interaction between the aforesaid factors that complicates the understanding of human vision. The aim of

colorimetry is studying colour as an isolated phenomenon, so the observation mode plays a relevant role; in colour science it allows the quantitative specification of the colour of a physical visual stimulus, following these fundamental bases (Wyszecki and Stiles, 1982):

1. stimuli having the same specification must appear the same, under the same viewing conditions, to an observer without anomalies in colour vision

2. stimuli looking alike have the same specification

3. the numbers entering the specification of a colour stimulus are continuous functions of the physical parameters defining the spectral radiant power distribution.

As already mentioned in this chapter, the visual trivariance that supports the definition of colour is based on the laws of Grassmann (1853), that have been reviewed by different scientists (Krantz, Wyszecky et al.):

1. *Additive property*: when a radiation is identically added to both sides of a color mixture field, the metamerism is unchanged

2. *Scalar property*: when both sides of the color mixture field are changed in radiance by the same proportion, the metamerism is unaltered

3. *Associative property*: a metameric mixture may be substituted for a light without changing the metameric property of the color fields

It is thanks to these laws that the set of colour stimuli forms a real-linear vector space, named the *tristimulus space*. In order to better understand how it is made, the experimental conditions require colour proposed in aperture-mode with the color matching experiment with an observer that, for example, is required to evaluate the equality of contiguous visual fields. The output color is within the domain of psychophysical colour, described by the *luminance* and *chromaticity*. A *colour*

stimulus is the luminous radiation entering the eye with a known spectral power distribution. It is then filtered after passing the eyeball and reaches the retina where activates the photoreceptors. Two different angular extent of the retina can be considered, namely 2° (*foveal* vision) and 10° (*extrafoveal* vision). Colour specification is achieved by the activation of the photoreceptors, that form the basis for tristimulus space. Even though an infinity of reference frames can be chosen within this space, the most interesting are essentially three:

- **the *PDT* fundamental reference frame** where the components of its vectors are in direct relationship with the cones' activations [3] (positive components)
- **the *RGB* reference frame** that can be put in direct relationship with the trivariance of chromatic vision. Here the stimuli are a consequence of three linearly-independent monochromatic primary radiations (not necessarily positive components)
- **the CIE1931 *XYZ* reference frame** where the component Y is proportional to the luminance of the stimulus (positive components)

1.7.1 The fundamental reference frame

It was already made clear in the text that human retina accommodates two kinds of photoreceptors, each kind for a different visual condition, i.e rods for scotopic vision and cones for photopic vision. Every single type of cones has only one pigment absorbing light in different amounts.

Rods are far more than cones (120 to 6-7 millions) and cones come in three varieties indicated as S (short), M (medium), L (long), sensitive to short, medium, and long wavelengths of light, respectively. The absorbance of the pigments of each of

[3] The fundamental reference frame is generally called *LMS*, but here is called *PDT* to distinguish this reference frame, related to the 1931 CIE observer, from that corrected by Vos and used by physiologists, called *LMS*

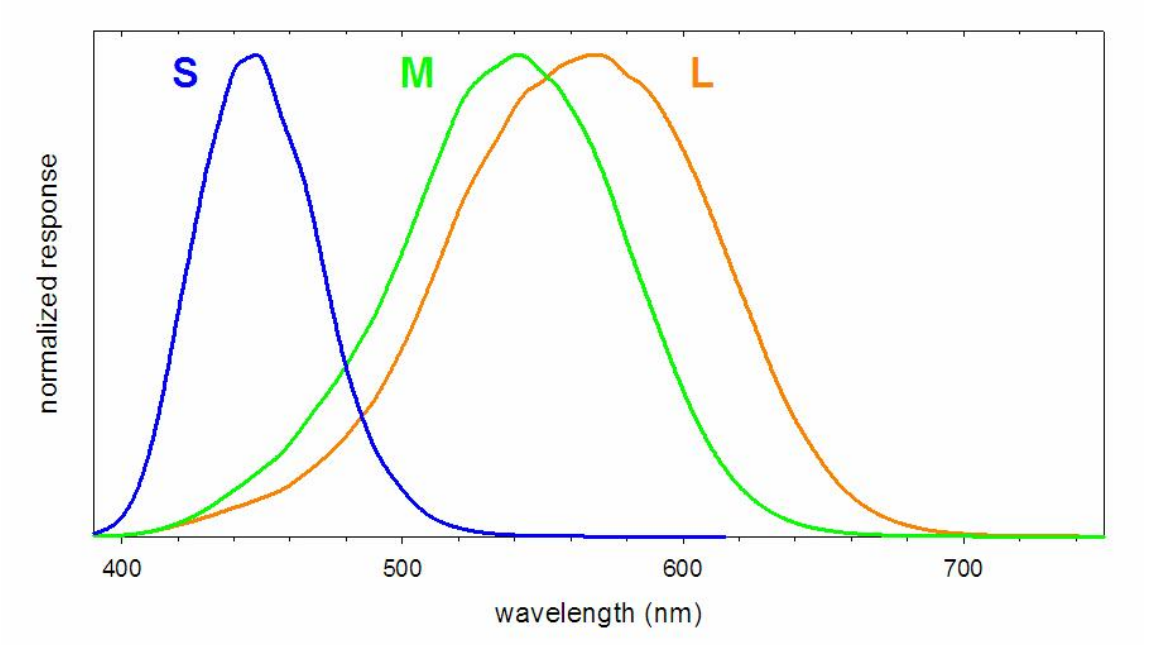


FIGURE 1.14: Normalised absorption spectra for the three types of cones in the visible

them with visible wavelengths is shown in Figure 1.14. The response of the S cone is centered at 420 nm, so light that triggers primarily the S cone is perceived as blue. There is a substantial overlap in the responses of the other cones, which have maximum responses at 530 nm and 560 nm, for the M and L cones. The M and L cones are most responsive to green and yellow-green light, but each retains tails far in the red region above 600 nm. The percentages of abundance are 63 %, 31 % and 6 % for the L, M and S cones, respectively. If these data are taken as weighting factors for the normalised curves of Figure 1.14, then the abundance of each type of cone is obtained (Figure 1.15). It is straightforward that the low abundance of S cones is the relatively poor ability of human eyes to discern the color blue; so, the eye is most responsive to green, yellow and orange light. Light outside the visible range of wavelengths is undetected by the cones and thus does not produce the perception of color. No type of cones can provide any information about the spectral quality of the absorbed radiation; that is because the amplitude of the electrical responses

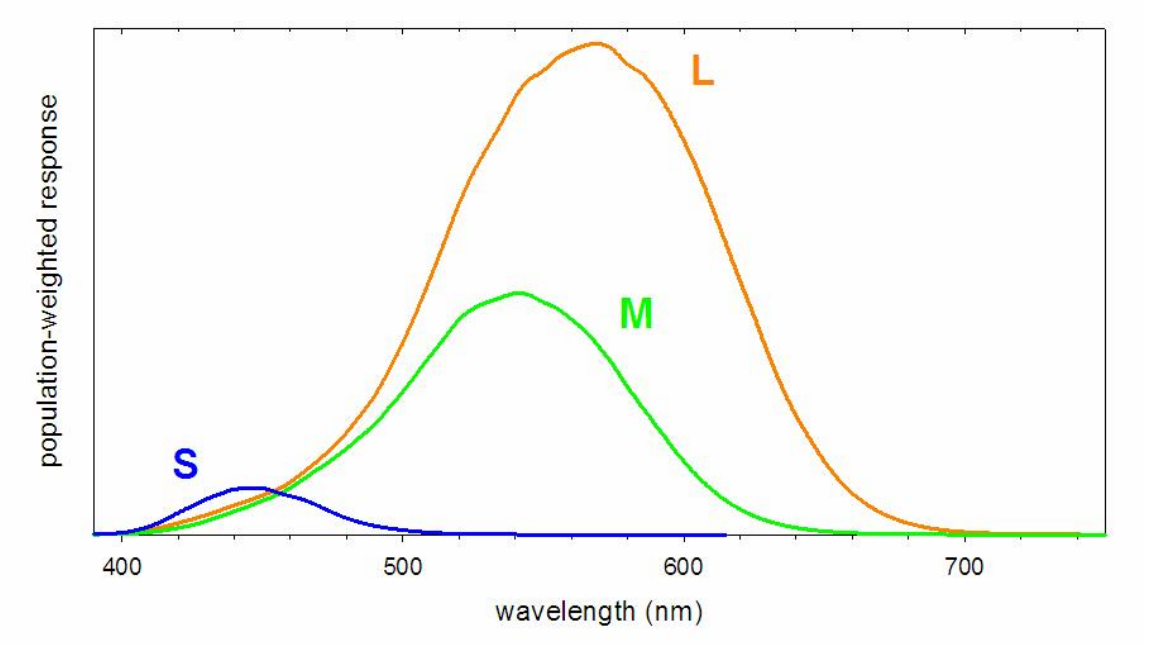


FIGURE 1.15: Weighted values of the curves of Figure 1.14

depends on the number of absorbed photons, but not on the wavelength of photons keeping constant the number of photons absorbed. A cone ultimately responds with a signal of larger or smaller amplitude, according to the number of photons absorbed (*colour blindness* - RUSHTON's principle of *univariance*). Another relevant feature of human visual system is that L and M cones enter the mechanisms that set how much luminous a stimulus appears, whereas S cones play only a relevant role in chromatic vision.

The fundamental space is based on the independency of the behavior of cones, on their *activations* (quantities proportional to the numbers of absorbed photons) and on RUSHTON's univariance principle. The mathematical form of the activation is the

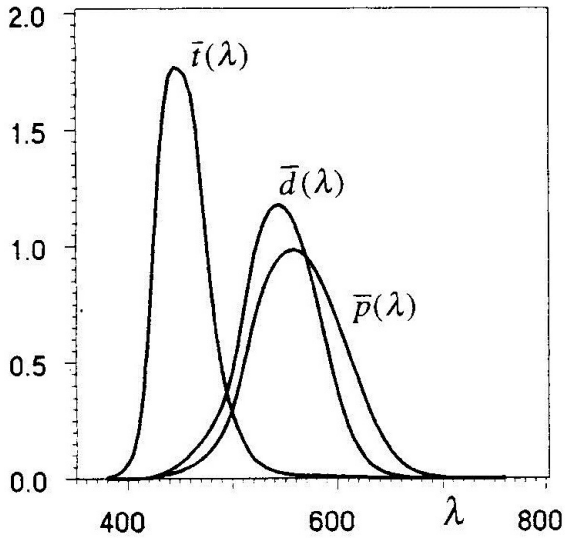


FIGURE 1.16: Colour matching functions in the fundamental tristimulus space (λ values in nm)

following:

$$
\begin{aligned}
P &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{p}(\lambda)\, d\lambda \\
D &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{d}(\lambda)\, d\lambda \\
T &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{t}(\lambda)\, d\lambda
\end{aligned}
\tag{1.7}
$$

where $L_{e,\lambda}$ is the spectral radiance and $(\bar{p}(\lambda), \bar{d}(\lambda), \bar{t}(\lambda))$ are the *colour matching functions* which are in direct relationship with the transmittance of cornea and the quantum efficiency of cones. The totality of (P,D,T) values constitutes a real-linear vector space (*cones' activations space* or *tristimulus space* in the *fundamental reference frame*). Every single vector of this space $\mathbf{Q} = (P, D, T)$ is said *tristimulus vector* has its components said *tristimulus values*, that correspond to the activations of cones. Any tristimulus value is positive and, except for the outer wavelengths, because of the spectral overlapping of colour matching functions in Figure 1.16, it is not possible to realize the activation of a single cone only.

The tristimulus space is wholly described by the knowledge of the tristimulus vectors associated to the monochromatic radiances of unit value. Thus, in this space the colour matching functions represent the components of the tristimulus vector $\mathbf{E}(\lambda) = (\bar{p}(\lambda), \bar{d}(\lambda), \bar{t}(\lambda))$, associated to the monochromatic components of the *equienergetic stimulus*. Accordingly, in correspondence of the equienergetic radiance ($E(\lambda) = 1$) the following relationship must hold true: $P = D = T = 1$.

By analogy with the tristimulus values and following the ABNEY law, the relationship between the *luminance* and the spectral heterochromatic radiance $L_{e,\lambda}$ of a stimulus is:

$$L_\nu = K_m \int_{380}^{780} L_{e,\lambda} \cdot V(\lambda)\, d\lambda \tag{1.8}$$

where $V(\lambda)$ represents the *relative luminous efficiency function*, with the attribute "relative" indicating that, as a rule, the peak value is assigned unit value. This function is the link between photometry and radiometry; it was historically defined for the *foveal* vision, i.e. for a field-of-view subtended by a solid angle centered at the observer's pupil and with a plane angular section less than 2°. Its experimental measurement is difficult and the procedure is still debated. The constant K_m=683 $\mathrm{lm\,W^{-1}}$, so the number 683 is just the conversion factor of a radiometric quantity into a photometric quantity.

At *photopic* levels, namely for the range of luminances higher than 10 $\mathrm{cd\,m^{-2}}$, daily vision is characterized by the *photopic luminous efficiency* curve, which describes the sensitivity of the eye in correspondence with monochromatic stimuli. The function is normalized at the peak value of 555 nm and is referred to the *standard photopic observer*, defined by the CIE in 1924. In the ranges below 400 nm and above 700 nm it approaches zero, so these limits are traditionally taken as the borders of visible spectrum. A *scotopic* equivalent exists for luminance levels lower than 0.1 $\mathrm{cd\,m^{-2}}$, $V'(\lambda)$, issued by the CIE in 1951 and defined as *standard scotopic observer*.

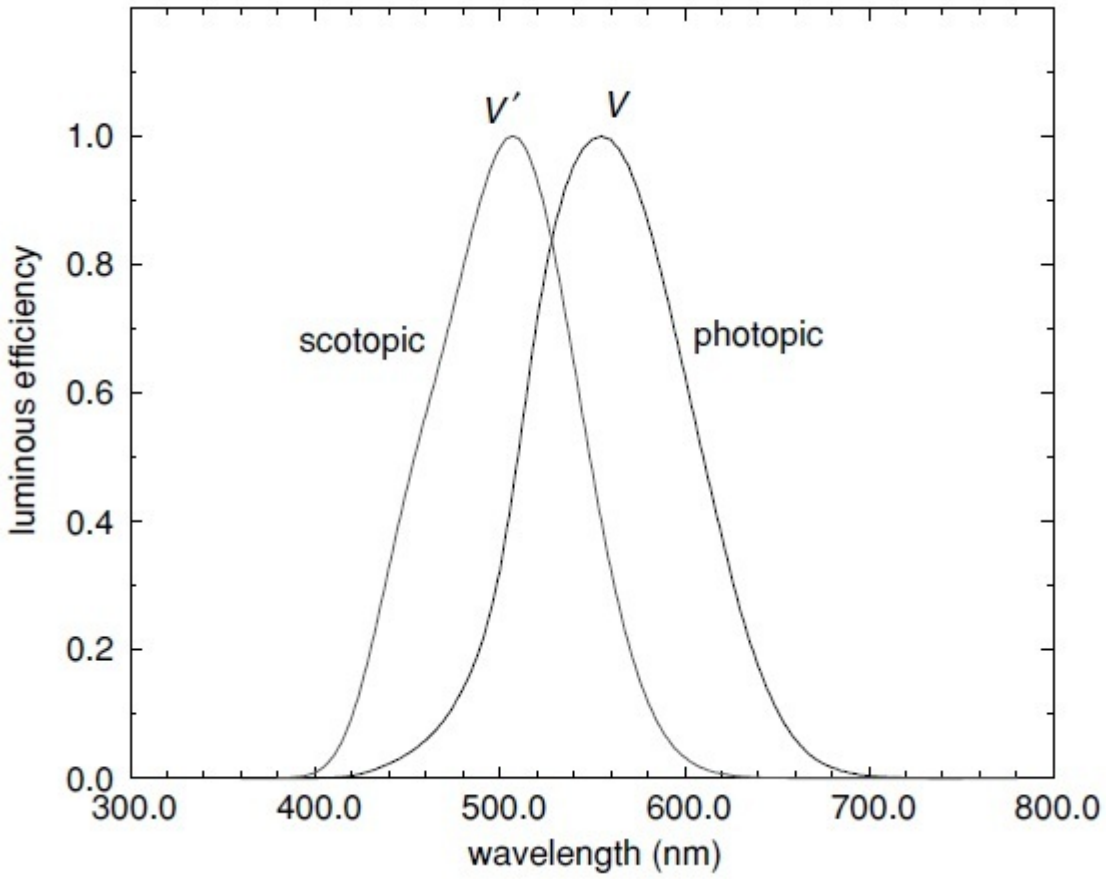


FIGURE 1.17: The CIE luminous efficiency functions; $V(\lambda)$ for the *photopic* vision, $V'(\lambda)$ for the *scotopic* vision

For the scotopic function the peak value is at 507 nm with fairly higher sensitivity values (Figure 1.17). The *mesopic* vision lies between photopic and scotopic luminance levels and plays a fundamental role in describing the vision in twilight conditions, with the contribution from both rods and cones. Experimentally, it was demonstrated that the responses of the three kinds of cones give rise to luminance additively and in proportion to their activations:

$$L_\nu = \tilde{\mathbf{\Lambda}}\mathbf{Q} \tag{1.9}$$

with $\tilde{\mathbf{\Lambda}} = K_m(L_P, L_D, L_T)$ that is a quantity proportional to the weights with which the cones participate to the luminance and $\mathbf{Q} = (P, D, T)$ representing the tristimulus

vector. After a simple substitution it turns out that:

$$V(\lambda) = L_p\bar{p}(\lambda) + L_d\bar{d}(\lambda) + L_t\bar{t}(\lambda) = \tilde{\mathbf{\Lambda}}\mathbf{E}(\lambda) \tag{1.10}$$

that can undergo further simplification since L_T=0, due to the insignificant contribution of S cones to luminance.

1.7.2 The RGB reference frame

Experimental difficulties discourage the construction of tristimulus space as previously shown and an indirect way is followed, instead. The linear attribute of the tristimulus vector space makes possible the choice of three convenient independent monochromatic vectors, like Maxwell did for the construction of the first tristimulus space and later followed also by the CIE for the 1931 standard observer. In this scenario, the monochromatic stimulus of unit power can be expressed as a linear combination:

$$\mathbf{E}(\lambda) = \bar{r}(\lambda)\mathbf{R} + \bar{g}(\lambda)\mathbf{G} + \bar{b}(\lambda)\mathbf{B} \tag{1.11}$$

where $\mathbf{R}, \mathbf{G}, \mathbf{B}$ represent the new primaries (colour stimuli of unit power) and with $(\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda))$ corresponding to the new colour-matching functions, shown Figure 1.7(a) (which can assume negative values).

The primary stimuli for the definition of a standard colorimetric observer are monochromatic vectors associated to λ_r=700.0 nm (red, $\mathbf{R}$), λ_g=546.1 nm (green, $\mathbf{G}$), λ_b=435.8 nm (blue, $\mathbf{B}$).

Similarly, as with the 1.7, the activation of cones are now:

$$\begin{aligned} R &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{r}(\lambda)\, d\lambda \\ G &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{g}(\lambda)\, d\lambda \\ B &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{b}(\lambda)\, d\lambda \end{aligned} \tag{1.12}$$

The units of the colour matching functions are chosen in such a way that the equienergetic stimulus $\mathbf{E}(\lambda) = (\bar{r}(\lambda), \bar{g}(\lambda), \bar{b}(\lambda))$, represented by a radiance $L_e(\lambda)$=1, has equal components:

$$\int_{380}^{780} \bar{r}(\lambda)\, d\lambda = \int_{380}^{780} \bar{g}(\lambda)\, d\lambda = \int_{380}^{780} \bar{b}(\lambda)\, d\lambda \tag{1.13}$$

The equivalent of 1.10 can be calculated once the luminances of the primaries (with unit radiance) $L_{\nu,red}$, $L_{\nu,green}$, $L_{\nu,blue}$ are known:

$$V(\lambda) = \frac{1}{K_m}[L_{\nu,red}\bar{r}(\lambda) + L_{\nu,green}\bar{g}(\lambda) + L_{\nu,blue}\bar{b}(\lambda)] = \tilde{\mathbf{\Lambda}}\mathbf{E}(\lambda) \tag{1.14}$$

The luminances can be retrieved from the equations:

$$\begin{aligned} L_{\nu,red} &= K_m L_{e,red} V(\lambda_r) \\ L_{\nu,green} &= K_m L_{e,green} V(\lambda_g) \\ L_{\nu,blue} &= K_m L_{e,blue} V(\lambda_b) \end{aligned} \tag{1.15}$$

with $(L_{e,red}, L_{e,green}, L_{e,blue}) = (3.4432, 4.7485, 248.2388)$ obtained from calculations involving the metameric match of the equienergetic stimulus.

1.7.3 The XYZ CIE *1931 reference frame*

D.B. Judd's linear transformation in the tristimulus space, to make luminance represent one of the components of that vector space, was accepted by the CIE in 1931 and gave birth to the XYZ tristimulus space. The transformation $RGB \rightarrow XYZ$ is based by imposing that the axes are orthogonal and that Y represents the luminance L_ν. The equienergetic stimulus must have equal components and the $Y = 0$ $[XZ]$ plane is the locus of vectors with null luminance. The $X = 0$ and $Z = 0$ planes must be tangent to the *spectral locus*[4] at short and long wavelengths, respectively.

[4] SPECTRUM LOCUS is the intersection of monochromatic vectors with the plane: $X + Y + Z = 1$

Summarizing, the following relationships hold true:

$$\begin{aligned} L_\nu &= K_m Y \\ \tilde{\mathbf{\Lambda}} &= (0,1,0) \\ \bar{y}(\lambda) &= V(\lambda) \end{aligned} \tag{1.16}$$

The colour matching functions have positive values, are indicated with $(\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda))$ and their plots are reported in Figure 1.7(b). Within this reference frame a generic tristimulus vector $\mathbf{Q} = (X, Y, Z)$ with radiance $L_{e,\lambda}$ has the following components:

$$\begin{aligned} X &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{x}(\lambda)\, d\lambda \\ Y &= \int_{380}^{780} L_{e,\lambda} \cdot V(\lambda)\, d\lambda = \frac{L_\nu}{K_m} \\ Z &= \int_{380}^{780} L_{e,\lambda} \cdot \bar{z}(\lambda)\, d\lambda \end{aligned} \tag{1.17}$$

The linear-transformations between RGB and XYZ reference frames are ruled by the following matrices:

$$\mathbf{A} = \begin{pmatrix} 0.4900 & 0.3100 & 0.2000 \\ 0.1770 & 0.8124 & 0.0106 \\ 0 & 0.0100 & 0.9900 \end{pmatrix} \qquad RGB \rightarrow XYZ$$

$$\mathbf{A^{-1}} = \begin{pmatrix} 2.3646 & -0.8965 & -0.4681 \\ -0.5152 & 1.4264 & 0.0888 \\ 0.0052 & -0.0144 & 1.0092 \end{pmatrix} \qquad XYZ \rightarrow RGB$$

1.7.4 The XYZ CIE *1964 reference frame*

Chromatic discrimination and colour matching show up the differences between foveal and extrafoveal vision. The latter is a wide-field vision, involving a larger angular extent of the retina (10°), in which the inert pigments of the *macula* are

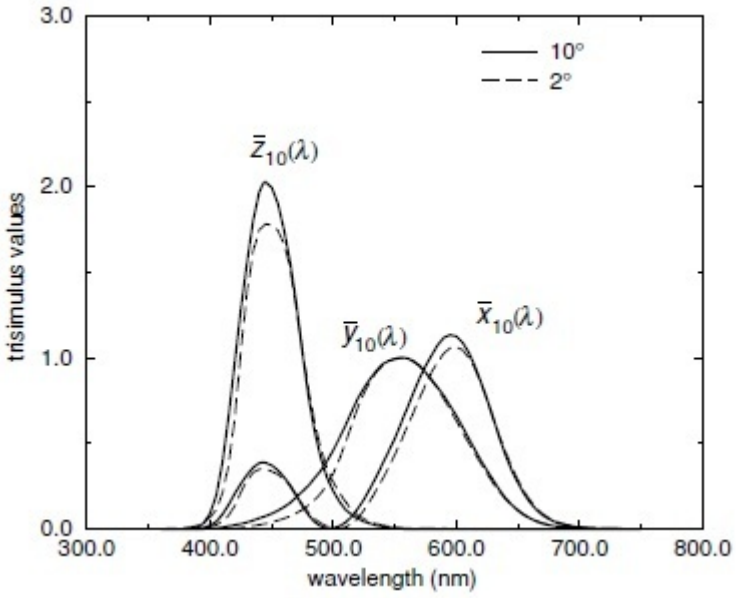


FIGURE 1.18: CIE 1964 (solid) and CIE 1931 (dashed) colour matching functions

responsible for the discrepancies aforementioned, even though under ordinary situations they are not sensed by the observer. Disregarding the role of rods, in 1964 the CIE releases a new standard observer, namely the *supplementary standard observer*, purposely tailored for extrafoveal vision. This is of widespread use in those fields where colour assessments have to be made on large visual areas, whereas the CIE 1931 standard observer is employed for foveal vision. In this brand-new reference frame the quantities have the subscript 10 and the experiments, mostly with the contributions from W.S. Stiles and J.M. Burch (Stiles and Burch, 1955) and by N.I. Speranskaya (Speranskaya, 1959), underlying its construction exploit the same wavelengths as for the *RGB* 1931 reference frame. Note that $\bar{y}_{10}(\lambda) = V_{10}(\lambda)$ (Figure 1.18). The $X_{10}Y_{10}Z_{10}$ is the result of a linear transform of the $R_{10}G_{10}B_{10}$ that minimizes the difference with the spectral locus of the 2° standard observer and that preserves the equality of the components of the equienergetic stimulus.

1.8 The attributes of colour

Colour attributes are classified as *base* and *relative*. The distinction is made clear when moving from psychophysics to psychometric specifications and in the construction of colour atlases. Moreover, base attributes are more frequently used for the presentation of colour in illuminant-mode, whereas relative attributes in object-mode, although some exceptions exist to this schematic classification.

Base colour attributes are *brightness*, *hue* and *colourfulness*. The latter is in relation to the intensity with which the hue appears (Hunt, 1991). Relative attributes make sense whenever colour has to be judged in relation to other colours. The need to define surface colour, thus disregarding the effect of source power, implies a comparison with another colour, usually the equally illuminated ideal diffuser. The radiation bounced back from this "reference" colour has the same spectral composition as the illuminating source and so both source and ideal diffuser have colours with the same base attributes. They send to the observer radiations with the same spectral composition. The choice of the proper illuminating source can render the ideal diffuser a "perfect" white.

Relative colour attributes (related to the ideal diffuser) are *lightness*, *chroma* and *saturation*.

A really valuable alternative to CIE colour specification is given by the colour ordering systems, whose treatment strictly follows CIE specification for a more complete mutual comparison.

1.9 Colorimetric computation

Tristimulus space is able to represent psychophysical colour and, as such, it cannot account for the multitude of phenomena following the activation of cones. At the same time, colour differences are not correctly represented within this domain, because

Table 1.1: Scheme of the historical phases of colorimetry, vision stages, visual fields with reference to CIE standard systems

		Historic evolution		
		1st historic phase *"colour matching"*	2nd historic phase *"colour difference"*	3rd historic phase *"colour appearance"*
		1st vision stage *transduction*	2nd vision stage *adaptation*	3rd vision stage
		psychophysic system	psychometric system	...
field of view	foveal vision	CIE 1931 std observer JUDD-VOS observer	CIELUV/LAB systems	CIECAM Retinex
	extra-foveal vision	CIE 1964 supplemental observer	CIELUV/LAB systems	...

of the inadequacy to make correspond equal perceived colour difference to colour distances. So far, a spectral radiance $L_{e,\lambda}$ has been associated to a colour stimulus but the common practice in colorimetry disregards the absolute attributes of stimuli, thus considering relative ones such as the relative spectral power distribution of a stimulus. The observer collects the light scattered from an ideal diffusing reflector illuminated by the source, whose reflected luminance is 100 or 1 , as a rule. The colour of non self-luminous objects seen in reflection is relative to that of the perfect (ideal) reflector. The CIE has recommended a number of geometries for colour measurement, because surface effects of objects can greatly affect both visual and instrumental colour specification. Regardless of whether sources or non self-luminous objects are considered, colour specifications must be performed after stating:

- the observer (CIE 1931, CIE 1931, VOS, etc)
- the illuminant (for non self-luminous objects only)
- the particular illuminating-viewing geometry
- the spectral data type such as bandpass and sampling interval
- the particular technique employed in the calculation

Because the appearance of a colour is strongly affected by the colour of the illuminating source, a crucial step towards the specification of colour is an accurate definition of the illuminants involved. In 1931, the CIE promulgated and recommended the use of three standard *illuminants*, indicated as A, B and C. Note that these are also physical sources since A illuminant has the same spectral power distribution of a blackbody at nearly 2856 K (experimentally reproducible with a tungsten filament lamp at that temperature), B is close to the average daylight (direct sunlight at noon) at nearly 4900 K, C is intended to represent the average daylight (with overcast sky)

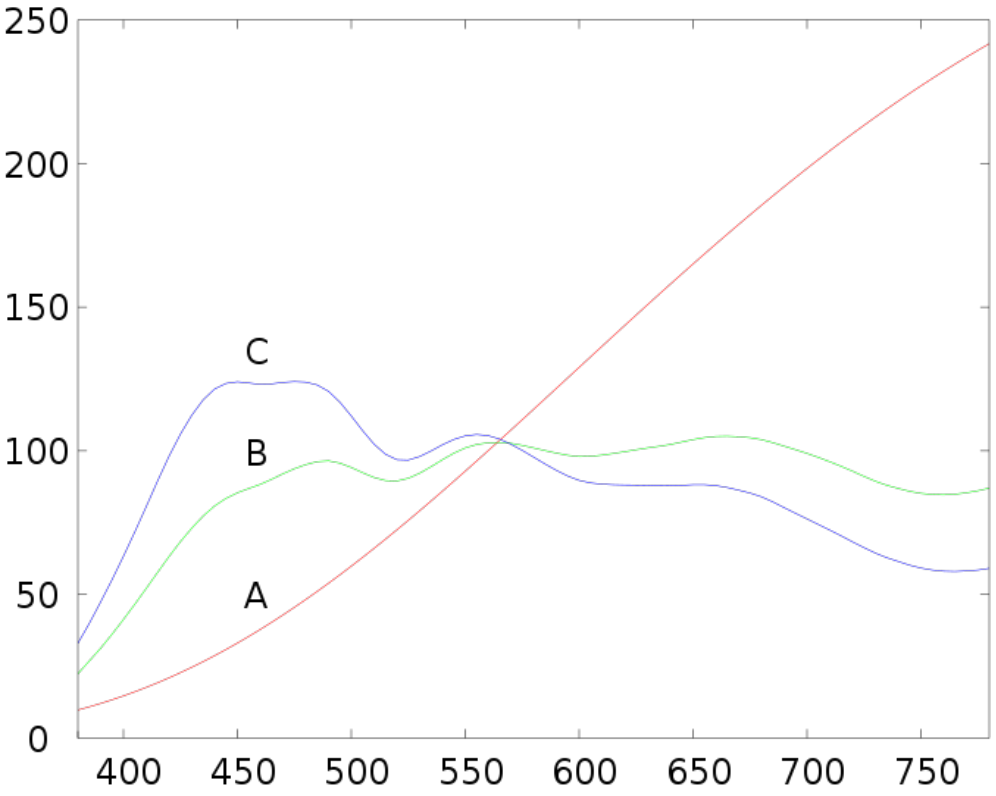


FIGURE 1.19: Relative spectral power distributions of the CIE standard illuminants A, B and C

at nearly 6800 K. The application of liquid colour filters (Wyszecki and Stiles, 1982) to A illuminant yields the spectral features of B and C illuminants (Figure 1.19). In 2005 both B and C illuminants have been declared out-of-date by the CIE because of their serious deficiency at wavelengths below 400 nm. A more precise information is required to demystify the significance of *sources* and *illuminants*:

- A *source* is the equivalent of the radiation physically emitted by a source
- An *illuminant* is every radiation with numerically defined and agreed spectral power distribution

Both sources and illuminants are specified through their emitted spectral power distribution, either absolute (S_λ measured in $\mathrm{W\,nm^{-1}}$) or relative ($S(\lambda)$). In the latter

case, the reference wavelength is 560 nm where, as a rule, $S(560) = 100$. Another notable remark is that illuminants are numerically specified in the interval 300 nm to 800 nm, which extends down to the UV range when measuring colour of fluorescent objects. The CIE also defines a number of standard illuminants (not physical sources) that are meant to reproduce daylight and are indicated with the letter D, which stands for "daylight", followed by two numbers. These numbers are in connection with the *correlated colour temperature* [5] as they represent the value in kelvin divided by 100. Their widespread use in colorimetry is ascribable to their flexibility in the description of real situations, although within a certain degree of tolerance. Their spectra follow the blackbody's footprints at the same temperature (Figure 1.20), except for the presence of absorption features due to the terrestrial atmosphere. D75 represents the typical north-sky light, D65 the standard daylight for general use, D55 the typical sunlight with sky-light and D50 a light somewhat more yellowish than D65. Other daylight illuminants are recommended by the CIE when D65 and others do not fit the particular needs. These standard illuminants are of common use in colour systems and standards; in fact, on the one hand D65 is the reference white for PAL and C for the NTSC television sytems, on the other hand D50 is the rule in the graphic arts industry.

1.9.1 The colour of sources

The characteristic feature of a source of radiation is, as already discussed throughout the text, its relative spectral power distribution $S(\lambda)$, calculated with reference to the relative spectral power distribution of a certified source $S_c(\lambda)$. The tristimulus

[5] CORRELATED COLOUR TEMPERATURE is the temperature of the Planckian radiator whose perceived colour most closely matches that of a given stimulus seen at the same brightness and under specified viewing conditions (Hunt, 1991)

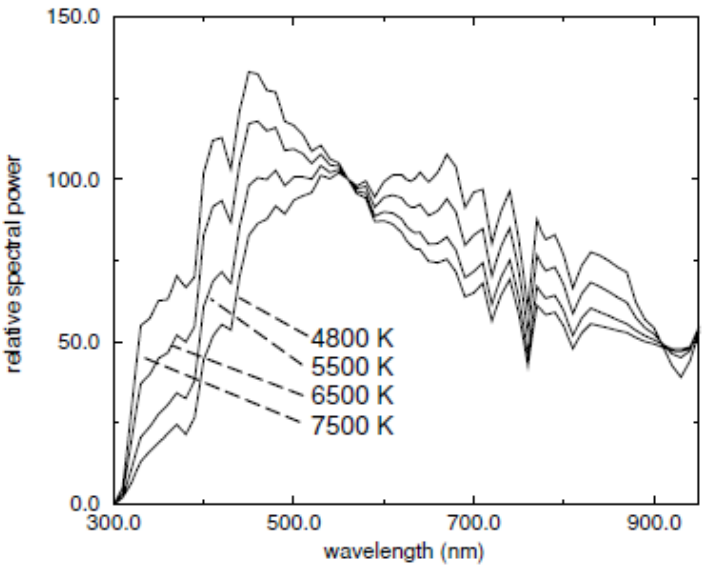


FIGURE 1.20: Relative spectral power distributions of some CIE daylight standard illuminants

values of the source are:

$$
\begin{aligned}
X &= K \int_{380}^{780} S(\lambda) \cdot \bar{x}_i(\lambda)\, d\lambda \\
Y &= K \int_{380}^{780} S(\lambda) \cdot \bar{y}_i(\lambda)\, d\lambda \\
Z &= K \int_{380}^{780} S(\lambda) \cdot \bar{z}_i(\lambda)\, d\lambda
\end{aligned}
\tag{1.18}
$$

$$
K = \frac{100}{\int_{380}^{780} S(\lambda) \cdot \bar{y}_i(\lambda)\, d\lambda}
$$

with i taking either no value (CIE 1931 standard observer) or 10 (CIE 1964 standard observer). The *chromaticity* of the source is given by:

$$
x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z} \tag{1.19}
$$

1.9.2 The colour of non self-luminous objects

Everyday's common experience proposes lots of non self-luminous objects that are visible under a lot of illuminating-viewing geometries and under several physical sources of illumination. Once these elements in a scene are fixed, then a determined visual situation is identified. The chromaticities of standard illuminants represent a good approximation of a great number of real visual situations and, at the same time, the luminance of a surface has no interest when specified in a single visual situation. For these reasons, it is convenient to specify the colour of the surfaces of non self-luminous objects for every particular illuminant, with reference to its relative spectral distribution. The ideal reflecting diffuser is taken as the reference for the luminance of a stimulus. In the case of colour seen in *reflection* the tristimulus values are:

$$\begin{aligned} X &= K \int_{380}^{780} S(\lambda) \cdot R(\lambda) \cdot \bar{x}_i(\lambda)\, d\lambda \\ Y &= K \int_{380}^{780} S(\lambda) \cdot R(\lambda) \cdot \bar{y}_i(\lambda)\, d\lambda \\ Z &= K \int_{380}^{780} S(\lambda) \cdot R(\lambda) \cdot \bar{z}_i(\lambda)\, d\lambda \end{aligned} \tag{1.20}$$

$$K = \frac{100}{\int_{380}^{780} S(\lambda) \cdot \bar{y}_i(\lambda)\, d\lambda}$$

with Y known as *percentage luminance factor* that equals 100 on the ideal reflecting diffuser. $R(\lambda)$ is the *spectral reflectance factor*, which may also take other names in particular viewing geometries. It is a relative quantity expressed by the ratio of the luminous flux reflected off a point on the surface of the object, within a given solid angle, to the flux reflected off the ideal reflecting diffuser in the same direction (in practice, a standard tile with spectral reflection factor certified by a metrological

laboratory) when it is equally illuminated. The chromaticity is the same as for the sources (1.19).

1.10 The search for uniform colour spaces

As a matter of fact, the human eye responds to light in a non-linear fashion, whereas in the previous sections the CIE *XYZ* and *RGB* colour spaces show a linear relationship with respect to the spectral power distribution of the stimuli. Depending on the particular position of the chromaticity of a stimulus in the spectral locus of the tristimulus space, any variation of trichromatic coordinates could be susceptible not to be noticed by the eye unless it exceeds a certain amount, named *Just Noticeable Difference* and shortened with the acronym *JND*. Thus, the tristimulus spaces are not uniform, in the sense that equal distances do not correspond with equally perceived colour differences.

In 1976 the CIE enters the second phase of colorimetry, where a *psychometric* specification of colours is proposed in the attempt of specifying colour as it is perceived in everyday experience. CIE proposed the CIELAB and CIELUV systems, the former for non self-luminous objects (reflective media) the latter for self-luminous objects (displays). Moreover, CIELAB has become a standard colour space for several fields, as in graphic arts (McDowell, 1999), multimedia (IEC, 1999), colour facsimile (ITU-T T.42, 1994). However, in both cases a relative specification is provided of the perceived colour, since proximal field, background and illuminant influence the perception of colours significantly. These new spaces are said *pseudo-uniform*[6].

[6] UNIFORM is a colour space inside of which equal visually perceptible differences are represented with equi-spaced points. It yields the conservation of the *JND*. The practice demonstrated that this condition was only partially satisfied, from which the term PSEUDO-UNIFORM follows. In these colour spaces relative colour is not specified anymore with just the tristimulus values, but some *visual response functions* are used, instead. Another fundamental idea is that the visual system of the observer is considered adapted to the illuminant chosen as reference. It is noteworthy to observe that CIELUV was partially and CIELAB was fully derived form tristimulus space by means of non-linear transformations; as a consequence, the vectorial sum is of no significance within these

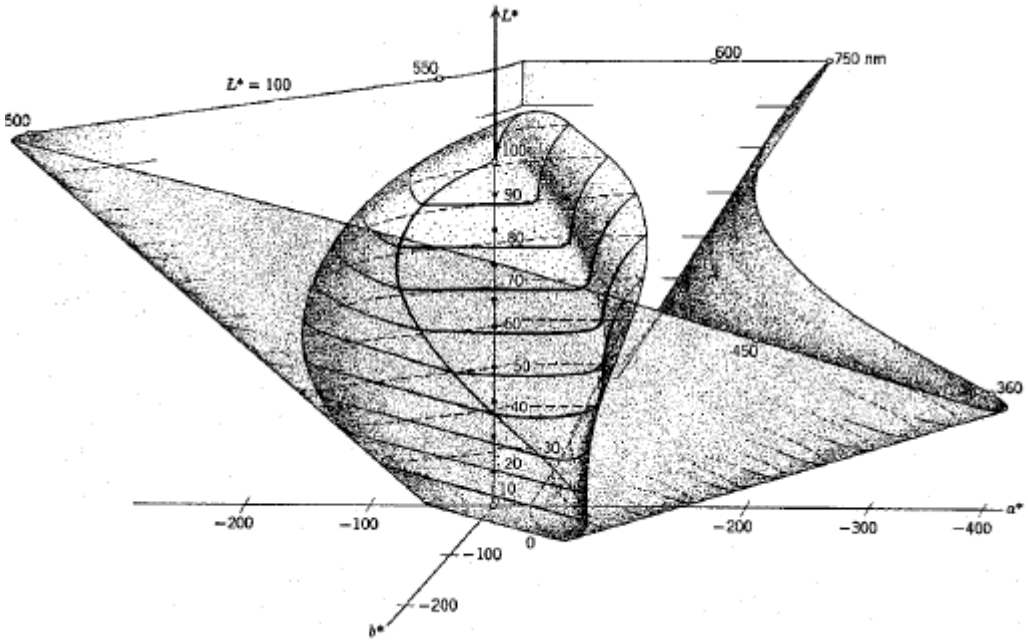


FIGURE 1.21: Perspective view of the CIELAB object colour space for the CIE 1964 standard observer and illuminant D65 (inner contour) and the spectrum locus

The CIELAB colour space is defined by the following quantities:

$$
\begin{aligned}
L^* &= 116 f(\frac{Y}{Y_n}) - 16 \\
a^* &= 500 \left[f(\frac{X}{X_n}) - f(\frac{Y}{Y_n}) \right] \\
b^* &= 200 \left[f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n}) \right]
\end{aligned}
\tag{1.21}
$$

Here the function f may take different forms:

$$
f(\gamma) = \begin{cases} \gamma^{1/3} & \text{if } \gamma \geqslant 0.008856 \\ 7.787\gamma + \dfrac{16}{116} & \text{otherwise} \end{cases}
$$

Here the considerations about colour are independent of the luminance level; the adaptation of the observer to the illuminant, not too much different from the day-

metric colour spaces. Due to its property to represent colours and to allow the calculation of colour differences for reflective objects, only details for the CIELAB are reported (Figure 1.21)

light, is accounted for considering that the visual field contains a surface colour surrounded by a mid-grey colour. The reference white has its percentage *luminance factor*[7] Y_{n}=100 and the chromaticity is the same as that of the particular illuminant under consideration. L^* is the *lightness* of the colour, also known as the CIE 1976 psychometric lightness, that ranges from 0 (ideal black) to 100 (reference white). It represents the ratio of light reflected off the object in a diffuse way with respect to the reference white.

The chromaticity of a colour is represented in a two-dimensional (a^*, b^*) diagram, with a^* representing the red-green coordinate and b^* the yellow-blue coordinate. Sometimes can be convenient to turn to an alternative representation of colours in the CIELAB space, the cylindrical coordinates, defining the CIE 1976 *chroma*, designating the distance from the L^*-axis:

$$C^*_{ab} = \sqrt{a^{*2} + b^{*2}} \tag{1.22}$$

and the CIE 1976 *hue-angle* measured in degrees:

$$h_{ab} = \arctan\left(\frac{b^*}{a^*}\right) \tag{1.23}$$

These quantities may help in the comprehension of the CIELAB colour space because are more related to the perceptual attributes of colour (Figure 1.22). The Euclidean distance between the CIELAB coordinates of two objects provides a rough guide to their discrimination. The symbol ΔE^*_{ab} is used to denote distance in the uniform color space and is defined as:

$$\Delta {E_{ab}}^* = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} = \sqrt{(\Delta L^*)^2 + (\Delta C^*_{ab})^2 + (\Delta H^*_{ab})^2} \tag{1.24}$$

[7] LUMINANCE FACTOR of a stimulus is the ratio of its luminance to the luminance of the ideal reflecting diffuser equally illuminated

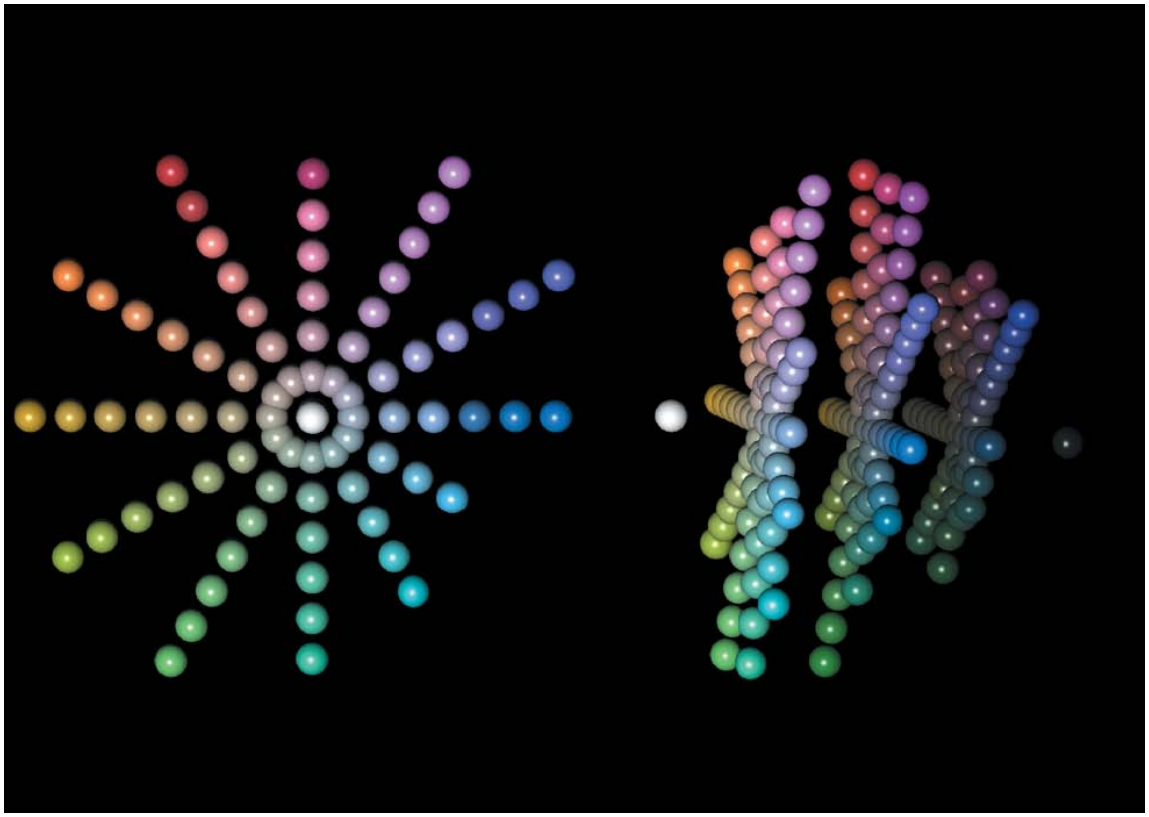


FIGURE 1.22: Views of a 3-D rendering of a sampling of the CIELAB color space along L^*, C^* and h axes

where the Δ quantities on the right represent the differences between the corresponding coordinates of the two stimuli and

$$\Delta H^*_{ab} = \sqrt{(\Delta E^*_{ab})^2 - (\Delta L^*)^2 - (\Delta C^*_{ab})^2} \tag{1.25}$$

For small colour differences away from the L^*-axis, $\Delta H^* \approx C^*_{ab} \cdot \Delta h_{ab}(\pi/180)$.
Using the CIELAB system is an improvement over using the Euclidean distance between XYZ tristimulus coordinates as a color difference metric. Since its initial standardization, predictions of the CIELAB system have been compared to new measurements, which have have led to a revision of the original system. The CIE has recommended one of these revisions, CIE94 (Hung and Berns, 2007; CIE, 1995). A second revision in widespread use is the CMC formula (Clarke et al., 1984). In the

CIE94 system, a distance measure ΔE^*94 is substituted for ΔE^*_{ab} and in the year 2000 a new more elaborate formula was derived CIEDE2000 (Luo et al., 2001).

1.11 The chromaticity diagram

A *chromaticity diagram* is a kind of colour map that accommodates colours in a systematic way by points in a triangle-like shape, which greatly facilitates the considerations about colour. The first of such diagrams in history dates back to Newton (1704), based on the centre of gravity rule; it was a circular figure (Figure 1.3) useful for the reproduction of colours of lights by light colour mixing and the same technique is employed today with updated both mathematics and measuring instruments.

It was thanks to the work of G. Palmer (1777) and T. Young (1802) who postulated the existence of three kinds of transducers (fibers) of the visible light in colour sensations, to the work of H.G. Grassmann (1853) who provided mathematical formalisation to Newton's ideas and to the work of H.L.F. von Helmholtz (1866) who produced a theory, based on the three fibers of Young and on the mathematics of Grassmann, that the chromaticity diagram took on a half-moon-like shape.

In 1857 J.C. Maxwell gave numbers to the chromaticity diagram that made possible to measure colours and to specify correctly any mixing of lights (Mollon, 2003; Wright et al., 1944).

Nowadays, the chromaticity diagrams are based on the experimental data of W.D. Wright and J. Guild and have become standardized by the CIE in 1931, under the guidance of D.B. Judd, in a reference frame that is imaginary, i.e. not related to physical lights or to photoreceptor activations, where the luminance enters with a conditioning role.

Newton had to fight against the general view that coloured light was some kind of contamination of pure light and in 1672 with the "Experimentum Crucis" (Figure 1.2) he demonstrated that:

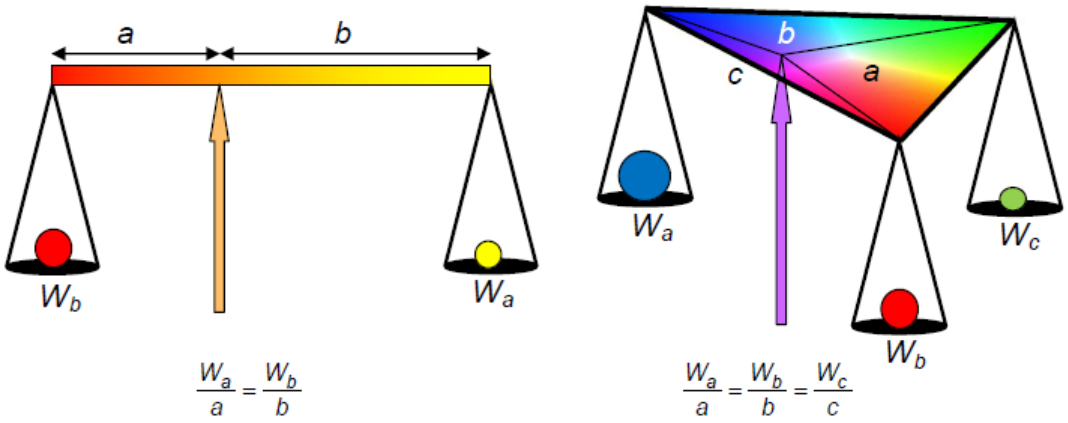


FIGURE 1.23: Barycentric coordinates for a one-dimensional yoke and two plates (left) and for a two-dimensional yoke and three plates (right)

> "white light is a heterogeneous mixture of differently refrangible rays. No one sort of rays alone... can exhibit whiteness. As the rays of light differ in the degrees of refrangibility, so they also differ in their disposition to exhibit this or one particular colour. Colours are not qualifications of light,... but original and connate properties, which in diverse rays are diverse."

Once the novelties of the Experimentum Crucis were accepted, Newton gave an empirical basis for his geometrical representation of compound colours (Newton, 1730).

In order to easily grasp the new ideas for producing compound coloured lights and represented by the centre of gravity rule, Figure 1.23 considers the way that two coloured lights combine to form a mixture; every single points of a bar shows the result of mixing components in variable ratios. Thus, a general point on the bar bears a ratio W_a/W_b, with W_a and W_b representing the amounts of the components of the mixture. Any colour on the bar creates two segments that are fixed proportional to weights W_a and W_b, respectively. Obviously, the balance is given by $W_a/a = W_b/b$. A triangular yoke serves for the generalization to a mixture of three independent coloured lights, where the amounts of these lights are proportional to the areas a, b

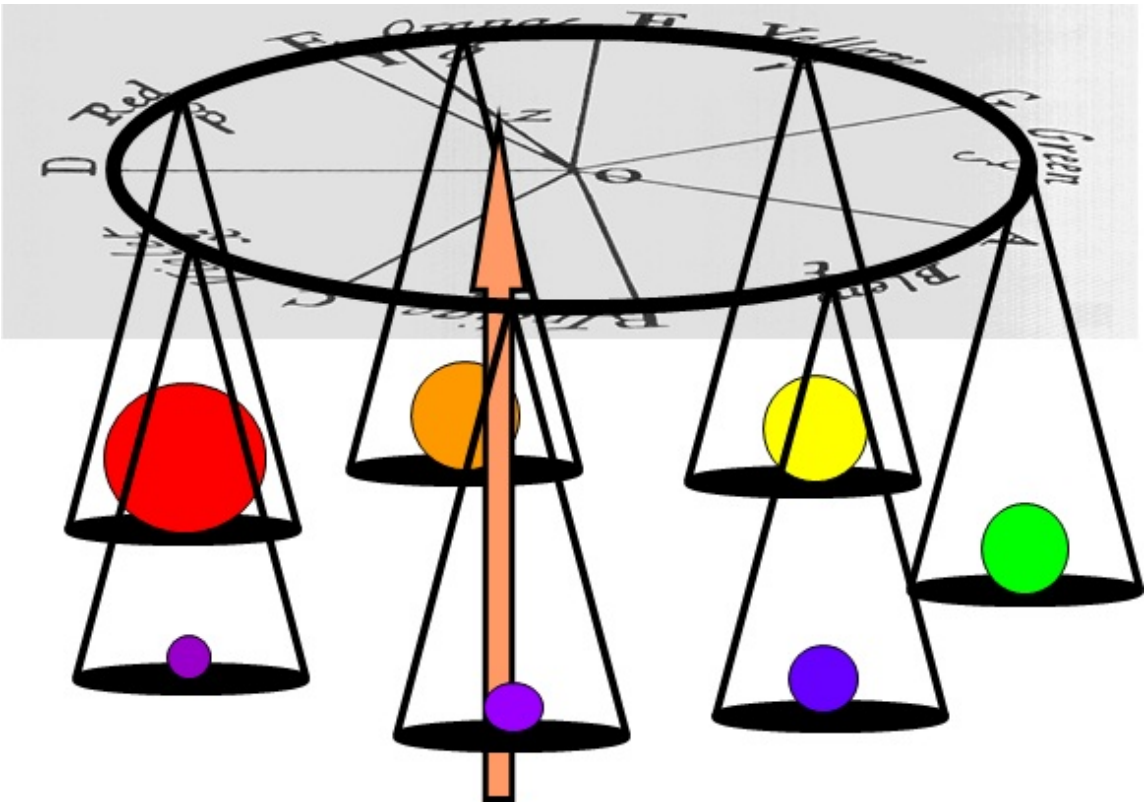


FIGURE 1.24: Bidimensional yoke, shaped like Newton's ring, balanced with seven plates corresponding to Newton's colours

and c of the triangles meeting at the equilibrium point, i.e. the compound colour. By analogy with the one-dimensional yoke, $W_a/a = W_b/b = W_c/c$, with the *barycentric coordinates* (a,b,c) representing the coordinates of the equilibrium point.

Now it is apparent that Newton's circle of Figure 1.3 is equivalent to some sort of subsampling, to only seven colours, of a two-dimensional yoke (Figure 1.24). It is because of the centre of gravity rule that Newton's ring is not just a simple colour wheel; in fact, it is rather a chromaticity diagram.

The center of gravity rule yields two really interesting facts:

- the same equilibrium point can be attained with different mixtures (*metamerism*)
- proper amounts of spectral lights on the rim yield white light (*complementary*

lights)

At his times, Newton was not fully convinced about the content of his centre of gravity rule, for three fundamental reasons basically:

1. colour complementarity was not met and Newton placed primary colours on the circle according to the musical notes

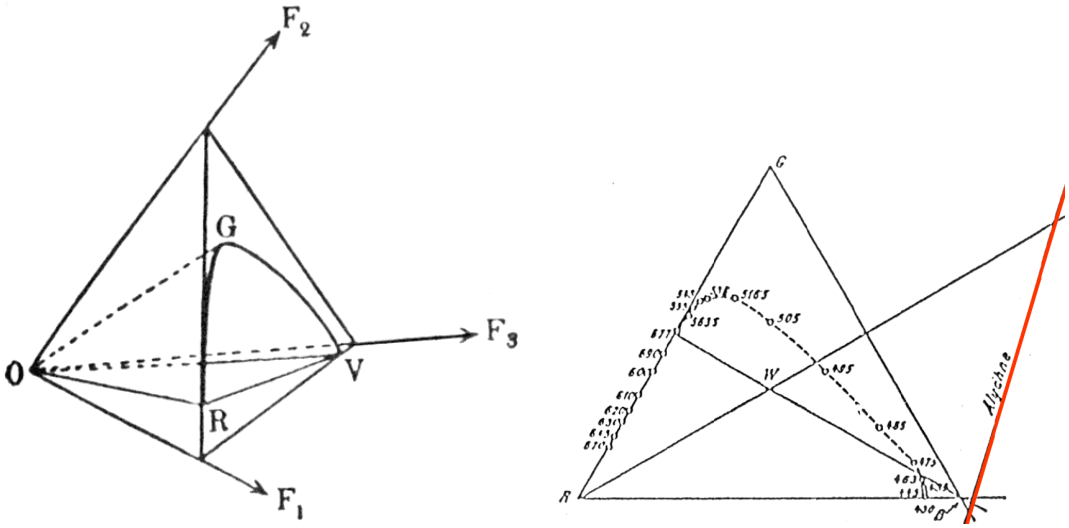
2. the circle disregarded *magenta* hues, because Newton decided to place only spectral lights on the rim

3. circular shape is only an approximation, as Newton was not able to perform any radiometric measurements

The years following the middle of the XIX$^{\text{th}}$ shed light on Newton's circle drawbacks. It is of great importance to recall Newton's words employed to clarify the context of application of his centre of gravity rule, with reference to Figure 1.3:

> "...it is such an orange as may be made by mixing an homogeneous orange with a white in proportion of the line OZ to the line ZY, this proportion being not of the quantities of mixed orange and white powders, but the quantities of the lights reflected from them."

Newton was very clear when he limited the field of applicability of his rule to light compounds, not to pigment mixtures; nonetheless, he was misunderstood by many people. In 1855 J.C. Maxwell tested the correctness of Newton's centre of gravity rule (Maxwell, 1855, 1885) and in 1860 (Maxwell, 1860) he produced a chromaticity diagram, after measurements in a reference frame (the *laboratory* reference frame) referred to three "standard" spectral lights.
The efforts of Maxwell made the centre of gravity rule a mathematical tool to specify colour, based on measurements of spectral lights reflected or transmitted by coloured



(a) (b)

FIGURE 1.25: (a) Perspective view of the tristimulus space and chromaticity diagram in the fundamental reference frame and (b) normal view of the chromaticity diagram

objects or emitted by any physical light sources.

As already mentioned in a previous section, in 1931 the CIE defined the colorimetric standard observer and the chromaticity diagram was referred to a set of three imaginary primaries in the *tristimulus space*. So, the chromaticity diagram arises from the intersection of the tristimulus vectors with a conveniently chosen plane in the tristimulus space (Figure 1.25). These are Schrödinger's original figures (Schrödinger, 1920) where the *alychne*[8] line is drawn. Such a line was chosen as abscissa line in the CIE 1931 (x, y) chromaticity diagram (Figure 1.4). As it would have been impractical to work with three coordinates, CIE chromaticity diagrams were obtained by projecting the diagram of Figure 1.25(a) onto the plane $Z = 0$; thereby a simple cartesian set of coordinates is introduced to describe the chromaticities of colour stimuli (shown above with equation 1.19).

If the centre of gravity rule is applied in the CIE 1931 (x, y) chromaticity diagram,

[8] SCHRÖDINGER'S ALYCHNE is a line originating from the intersection of the chromaticity diagram with the zero luminance plane. It represents the chromaticities of non-physical points with null luminance

considering two stimuli $\mathbf{Q_1} = (X_1, Y_1, Z_1)$ and $\mathbf{Q_2} = (X_2, Y_2, Z_2)$ of chromaticities $q_1 = (X_1/W_1, Y_1/W_1)$ and $q_2 = (X_2/W_2, Y_2/W_2)$ (where $W_{1,2} = X_{1,2} + Y_{1,2} + Z_{1,2}$), then the sum of these two stimuli will have the chromaticity (Oleari, 2008):

$$q = (x, y), \qquad x = \frac{x_1 W_1 + x_2 W_2}{W_1 + W_2}, \qquad y = \frac{y_1 W_1 + y_2 W_2}{W_1 + W_2} \tag{1.26}$$

It follows that point $\mathbf{q}$ lies in between the segment $\mathbf{q_1 q_2}$ and that the lengths of the segments $\mathbf{q q_1}$ and $\mathbf{q q_2}$ are such that:

$$| \mathbf{q q_1} | W_1 = | \mathbf{q q_2} | W_2 \tag{1.27}$$

where the 1.27 holds true for every single component x, y and z, with $z = 1 - x - y$. The previous description is in close analogy with the yoke of Figure 1.24 and here the equilibrium point $\mathbf{q}$ lies along the yoke $\mathbf{q_1 q_2}$, with weights W_1 and W_2 (Figure 1.26). If white light is added in increasing amounts to any colour on the chromaticity diagram, then the mixture moves towards the white point along a straight line from the original colour. In Figure 1.27 is shown the CIE *RGB* 1931 chromaticity diagram (r, g), with the intersection of the *XYZ* primaries. Great attention must be paid to make a correct distinction between a chromaticity diagram and a colour triangle; in fact, the former is the result of a theoretical formulation, whereas the latter is of empirical definition and of frequent use in "photographic" contexts. A colour triangle, also named *gamut*, is defined by the position of the chromaticities of the particular primaries of a device in the chromaticity diagram. So, it is only a subset of the chromaticity diagram that can be either small or large, depending on the instruments. It is obvious to desire a triangle covering as much as possible of the domain of all colours, but in practice a compromise must be made due to the difficulty of obtaining very bright beams of purely monochromatic light.

Based on this considerations, many additive colour photographic systems employed reproduction stimuli made of red, green and blue filters, either of dyed gelatin or

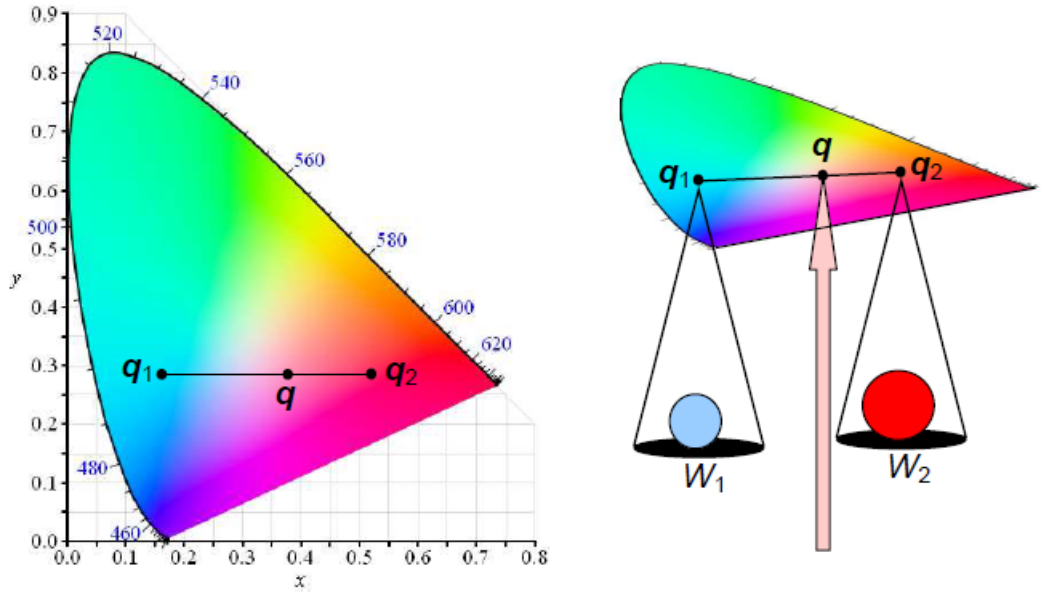


FIGURE 1.26: Centre of gravity rule in the CIE 1931 chromaticity diagram (left) and equivalent representation with yoke and weights (right)

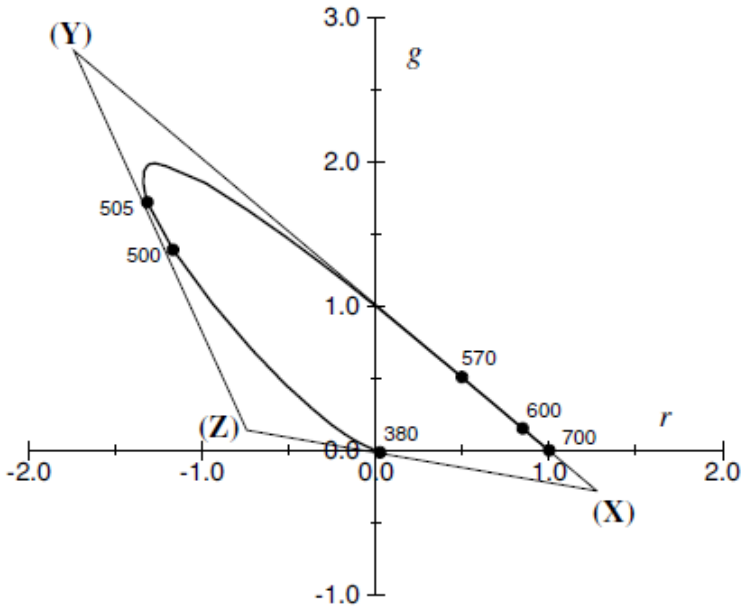


FIGURE 1.27: The (r, g) chromaticity coordinates of the XYZ primaries

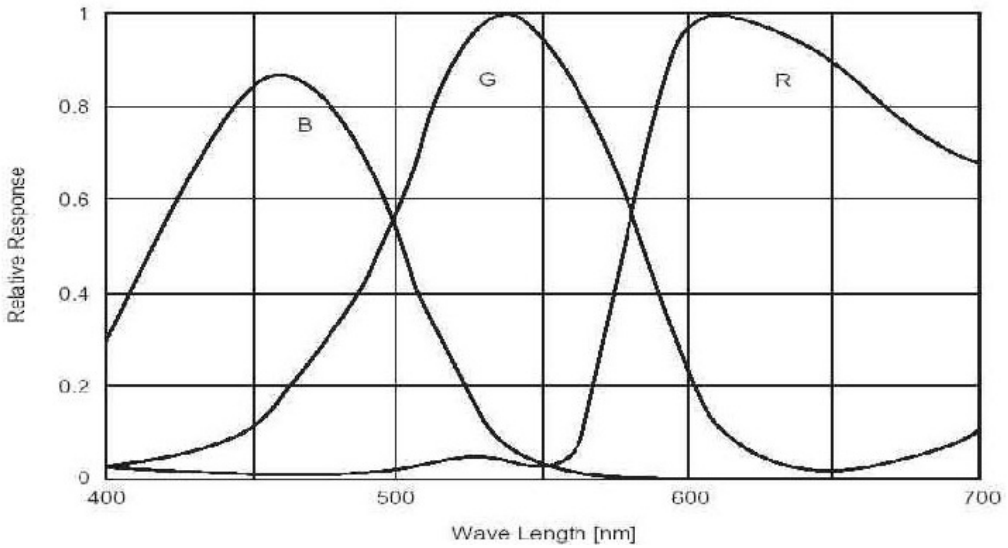


FIGURE 1.28: Spectral sensitivity of a typical electronic camera

glass, illuminated by a tungsten light source. It is quite apparent that the choice of reproduction stimuli is strategic in colour photography. Another feature that is rightful to point out is the extreme difficulty to realize the negative parts of the sensitivity curves, that leads to ignore them; that results in practical incorrect reproduction of all colours. The departures of the sensitivity curves from the theoretical colour matching functions have extensively been studied (Hunt, 2005).

The photosensitive elements of a camera have their own set of three spectral sensitivities and the camera itself has its own colour space, that should approach at best that of a standard observer. The chromatic information results from the comparison between the three signals. To increase the signal-to-noise ratio of the chromaticity signals, it is customary to use sensors with evenly spaced spectral sensitivity curves in the visible range (Figure 1.28). If the colour matching functions of the standard observer can be expressed as linear combination of the spectral sensitivities of the three sensors, then the output signals from the camera are:

- in one-to-one relationship with the colour stimuli of the tristimulus space of the

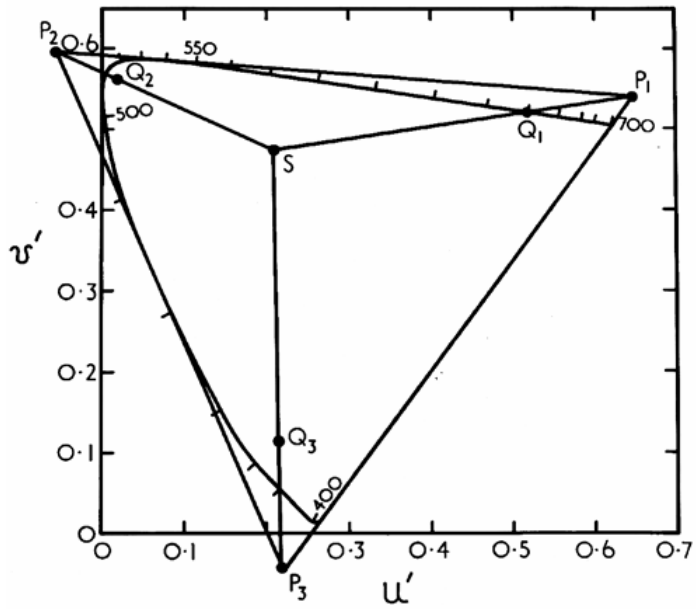


FIGURE 1.29: The IVES-ABNEY-YULE compromise

standard observer

- easily represented with tristimulus vectors in a fundamental reference frame of the camera, with primaries $\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}$ having chromaticities at the apexes of a triangle with sides tangent to the spectral locus

If such *super-saturated* stimuli were available (Figure 1.29), all colours would be matched by fully positive mixtures of them.

Commonly, colours in real life vary by a uniform addition of white to all colours; for this reason, whenever an error is made in colour reproduction that is equivalent to the addition of a little amount of white to all colours, this turns out to be almost unnoticeable. In the common practice, it is customary to follow the *compromise of* IVES-ABNEY-YULE whenever colour reproduction as a whole is considered. If the three new reproduction stimuli $\mathbf{Q_1}, \mathbf{Q_2}, \mathbf{Q_3}$ were taken, then there would be just a uniform reduction in colour purity (Figure 1.29), without affecting the dominant wavelengths, to which the human visual system is more sensitive. The effect on

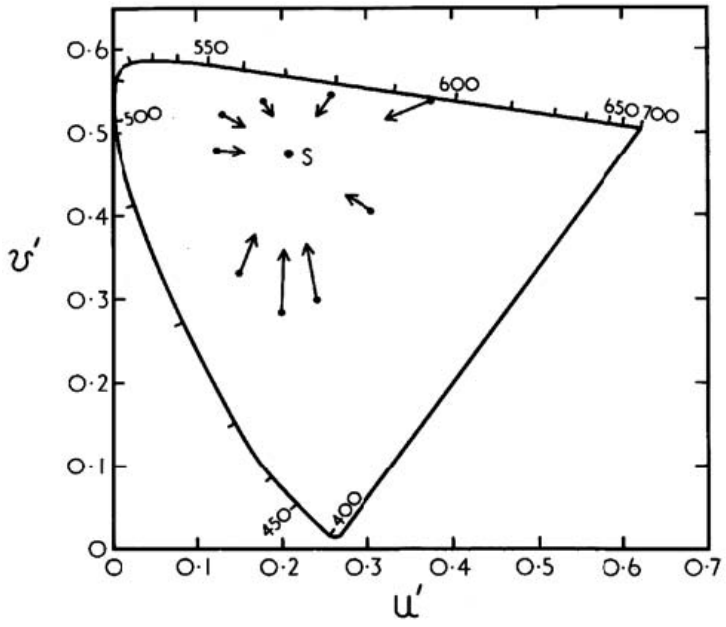


FIGURE 1.30: Errors resulting after the application of the IVES-ABNEY-YULE compromise

a chromaticity diagram is the shifting of points inwards, towards the white point (Figure 1.30), and represents the typical way in which the purities of colours vary in real scenes. In order to make a camera work as a colorimeter, the compromise of IVES-ABNEY-YULE must be abandoned because introduces unacceptable approximations. Serious problems of metamerism occur whenever the colour matching functions of the camera differ from those of a standard observer. The metamerism involved is thus twofold, i.e. the observer metamerism and the instrumental metamerism. Observer metamerism is difficult to quantify and is caused by the normal variations in the color responsivity of various observers. The metamerism of the camera opens the way to further complications since stimuli that are metameric for the standard observer may not be such for the camera, and viceversa.

1.12 How to measure the colour matching functions of a camera

Every single camera has its own colour space and the characterization of the camera is effected through the measurement of its colour matching functions. There are at least two reference frames of interest, i.e. the fundamental reference frame and the *RGB* reference frame, where the primary vectors are the same of a reference monitor. The definition of the former is obtained from that for the standard observer, except for the fact that the quantities are referred to the camera. Since they are electrical signals, it is not difficult to measure the activations of the photosensors of a camera.

A monochromator with a light source is used to select light, with a small but finite bandpass, that is made to enter an integrating sphere. The inner part of the entrance port of the sphere, where monochromatic lambertian light is uniformly distributed, is brought to focus on the sensor of the camera. A calibrated photodetector records the signal T_λ at the exit port of the sphere. After setting γ=1[9], the signals of the matrix of photodetectors of the camera ($S_{R,\lambda}$, $S_{G,\lambda}$ and $S_{B,\lambda}$) are taken (Figure 1.31). If each type of photodetector returns a signal that is proportional to the incident power, then the requested colour matching functions of the camera in the fundamental reference frame are:

$$r_F(\lambda) = \frac{S_{R,\lambda}}{T_\lambda} \qquad g_F(\lambda) = \frac{S_{G,\lambda}}{T_\lambda} \qquad b_F(\lambda) = \frac{S_{B,\lambda}}{T_\lambda} \tag{1.29}$$

[9] GAMMA γ is a parameter, historically defined for Cathode Ray Tube (CRT) monitor displays and nowadays adapted to modern Liquid Crystal Displays (LCD), which accounts for the non-linear response of the luminance emitted L_ν as the electric driving signal E changes.

$$(L_\nu / L_{\nu,\max}) = (E/E_{\max})^\gamma \tag{1.28}$$

with $L_{\nu,max}$ and E_{max} representing the maximum values of the luminance and of the electric driving signal, respectively. So, in order to have the same luminance as the scene reproduced on a display, the signal is distorted at the source. For a camera, it is mandatory to set γ=1 because the camera itself must not introduce any distortions to its output signals entering a display

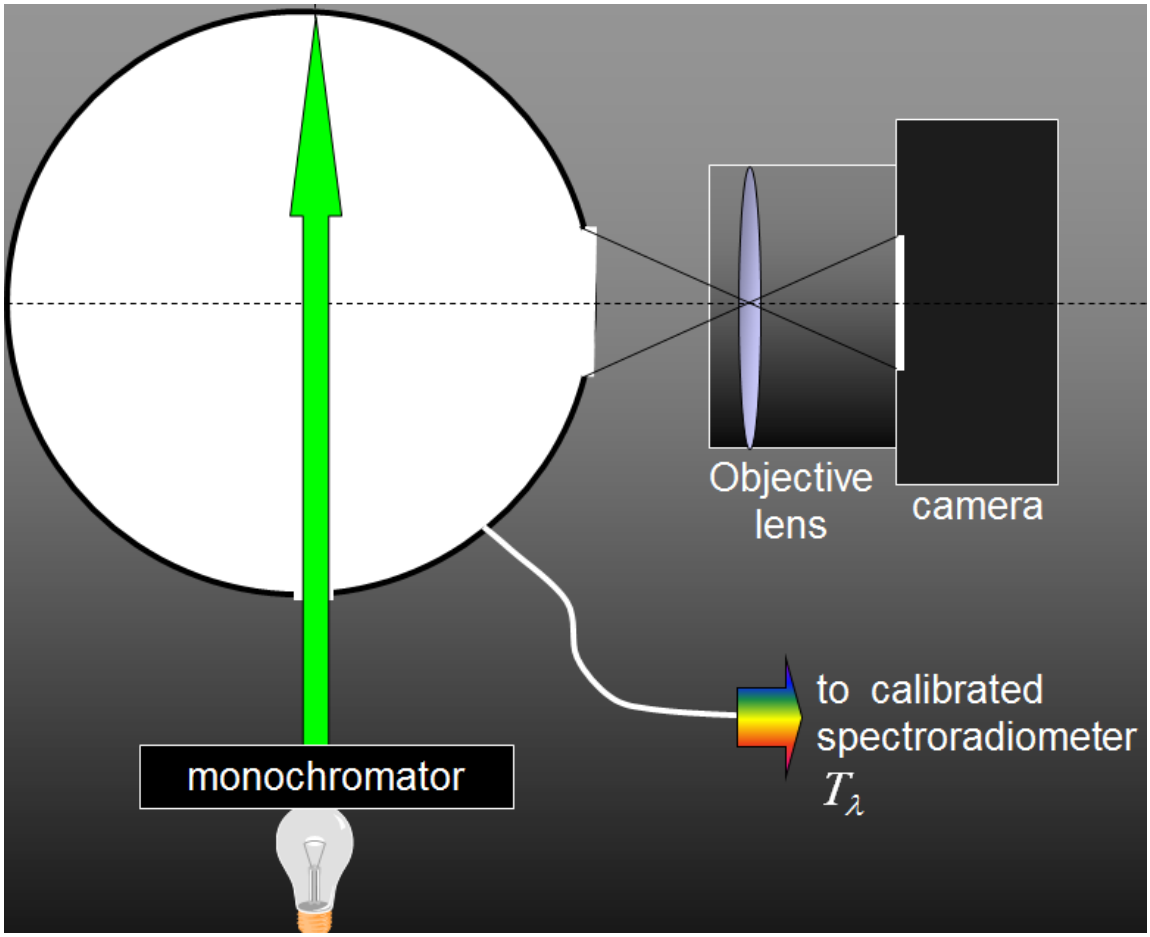


FIGURE 1.31: Set-up for the measurement of the colour matching functions of a camera in the fundamental reference frame

Such functions are always positive within this reference frame and their values strongly depend on the gain settings, acting as multiplicative factors.

Moving to the instrumental *RGB* reference frame, defined by the primaries of a trichromatic monitor, a convenient way to replace the three primary sources is by the use of a trichromatic monitor, whose $\mathbf{R}, \mathbf{G}$ and $\mathbf{B}$ primary lights are the primaries of the reference frame. Preliminarily, the camera is focused on the inner part of the aperture of an integrating sphere; the light emerging from such aperture is monochromatic lambertian and is uniformly distributed; a white light of the monitor, $\mathbf{W} = \mathbf{R} + \mathbf{G} + \mathbf{B}$, is let into the sphere and the three signals $S_{R,W}$, $S_{G,W}$, $S_{B,W}$ are

recorded (these can be brought to equality by adequate gain tuning). A calibrated photodetector is used to measure the fluxes (r_W, g_W, b_W), exiting the sphere from another port, due to the primaries.

As in the fundamental reference frame, a monochromator selects the small finite bandpass light, with central wavelength λ; this light, together with the primaries of the monitor, are sent into the integrating sphere and the intensities of these lights are tuned until they match the signals read with the white. During this tuning, the intensity of one of the primary lights falls to zero. From another port of the sphere, a calibrated photodetector acquires the signals due to the fluxes of radiation generated by the three primaries, r_λ, g_λ and b_λ, and by the monochromatic light e_λ (Figure 1.32). The resulting colour matching functions, that can also take negative values, in the *RGB* reference frame of the monitor are (Oleari, 2008):

$$\begin{aligned} r_{RGB}(\lambda) &= \frac{r_W - r_\lambda}{e_\lambda} \\ g_{RGB}(\lambda) &= \frac{g_W - g_\lambda}{e_\lambda} \\ b_{RGB}(\lambda) &= \frac{b_W - b_\lambda}{e_\lambda} \end{aligned} \tag{1.30}$$

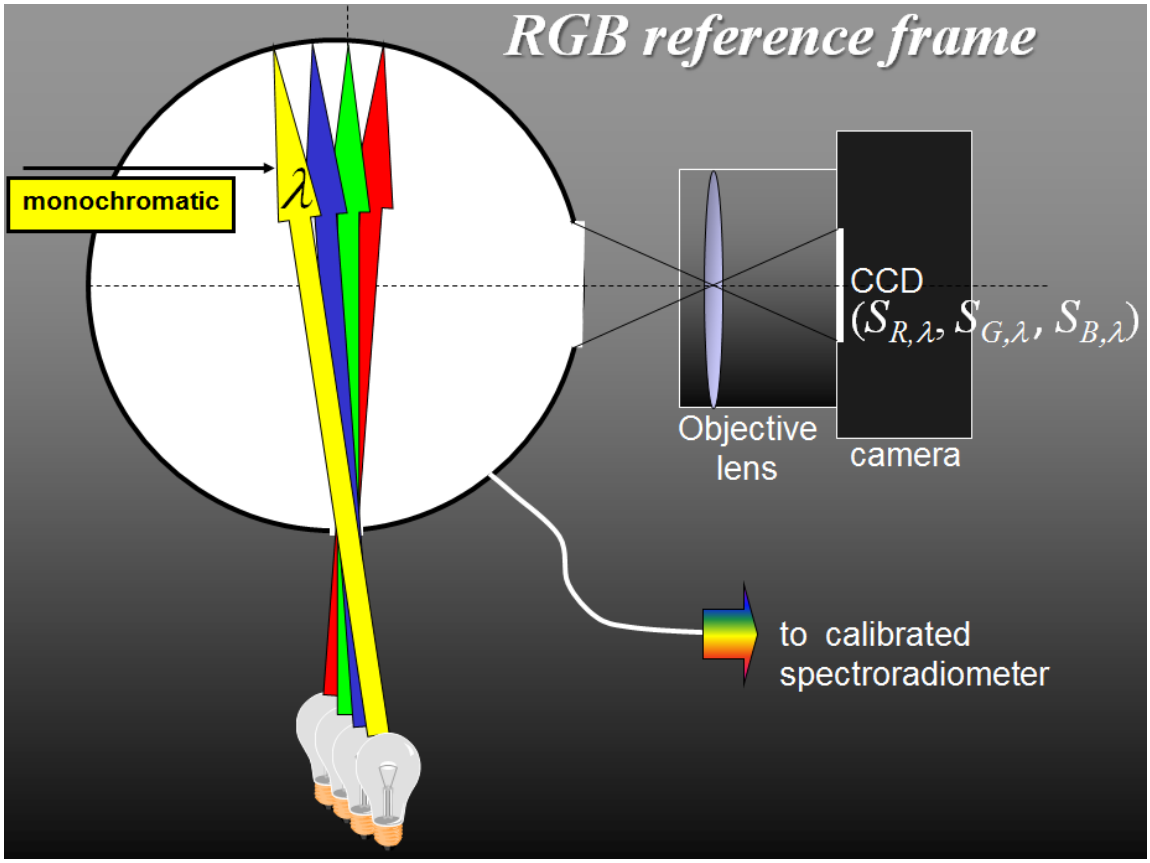


FIGURE 1.32: Set-up for the measurement of the colour matching functions of a camera in the RGB reference frame

2

The spectral imaging

With the limits of trichromatic photography clear in mind, great efforts have been made within the scientific community to get rid of the drawbacks introduced by metameric matching, in favour of spectral matching; in fact, spectral matching between the original and their reproductions have attracted many researchers during the last two decades for being the only way to assure a color match for all observers across changes of illumination (Imai and Berns, 1999). Image analysis techniques include technologic strategies aimed at characterizing either a volume or a surface, by exploiting some physical quantities. As a consequence, this new added dimension makes much more appealing the acquisitions, as they can readily and easily be output through one or multiple images of the investigated object. It is often customary to read in the literature or to hear that a spectral image is a measure of the reflectance of an object; of course, this is a misnomer because what is actually measured is the Spectral Reflectance Factor (SRF) and this fact should be more thoroughly stressed in the scientific community for broad acceptance. The former is an absolute quantity, whereas the latter is a quantity relative to a reference surface (CIE, 1989). In this way the spectral characterization is accomplished throughout the extension of the

surface in a spectral interval that typically ranges from the visible (*Vis*) until the near-infrared range (*NIR*). Among the spectral images a distinction is made between related but distinct techniques, namely multispectral and hyperspectral images; in the former, a sensor images at a number of discrete wavelength bands, whereas in the hyperspectral imaging the sensor still images at discrete bands, but over a continuous wavelength range. Some authors (Klein et al., 2008) fix an upper limit of ten to the number of optical wavelengths, for the acquisition of digital images, necessary to discriminate between multispectral and hyperspectral techniques. However, this statement is somewhat faint and in the common practice not a sharp limit to the number of imaging wavelengths can be fixed.

Multispectral imaging, on the other hand, might measure at specific wavelength bands, with the incoming light split with adequate filters, though not over a continuous range. It can be stated that multispectral imaging combines spectral information with machine vision information and it could be thought of as some kind of 3D imaging, with the third dimension associated not to a height profile but rather to a spectral feature.

For the contents aforesaid and regardless of the particular kind, the luxury of spectral images lies in their ability to return information about the *SRF* for every single elementary unit (pixel) of the imaged scene. The information may range from three components per pixel, in the case of the *RGB* colour images, up to several hundreds of bands, as in the case of hyperspectral imaging. Multispectral imaging plays a fundamental role in several domains of application, such as remote sensing (Davis et al., 1978), mineralogy, oil industry, astronomy (Rosselet et al., 1995), medical imaging, artworks analysis and archaeological finds (Carcagni et al., 2007; Maitre et al., 1997; Haneishi et al., 1997), cosmetics, colour printing (Berns, 1998) and computer graphics. Although of very high accuracy, so far the complexity and the high cost of hyperspectral image acquisition systems have limited their use mainly to

remote sensing applications.

In the field of cultural heritage (Havermans et al., 2003; Easton Jr et al., 2003), hyperspectral imaging is greatly appreciated for its non-destructive impact to get information unavailable with conventional colour or multispectral photography; within this particular and fascinating field of application, it is successfully employed to detect and characterize materials, to unveil features that are either obscured or invisible to the naked eye, to sense and study degradation phenomena and to assess possible detrimental effects on the object.

When dealing with the domain of spectral imaging, a really highly desirable way to operate is to employ device-independent instruments, namely with technologies that do not depend upon the spectral features of both the particular illuminant and acquisition system used. In order to achieve such result, the spectral characteristics of the components of the image acquisition process have to be known.

This chapter will concentrate on the presentation of the most successful instrumentation for spectral imaging in the field of conservation of artworks, i.e. for the sake of extracting important information from painted layers. This is made to stay in line with the main topic of this dissertation, that is the project of a hyperspectral camera that, among its several fields of application, encompasses also the imaging of artworks.

2.1 The image acquisition process

Here is presented a general spectral model of an image acquisition system that is suitable for systems using solid-state sensors, whose acquisition process is schematically depicted in Figure 2.1. The relative spectral power distribution of the illuminant is indicated with $S(\lambda)$, the spectral reflectance of the object with $R(\lambda)$, the spectral transmittance of the optical elements in the optical path with $\tau(\lambda)$, the spectral transmittance of one of the coloured filters (say the k^{th}) with $\varphi_k(\lambda)$ and the spec-

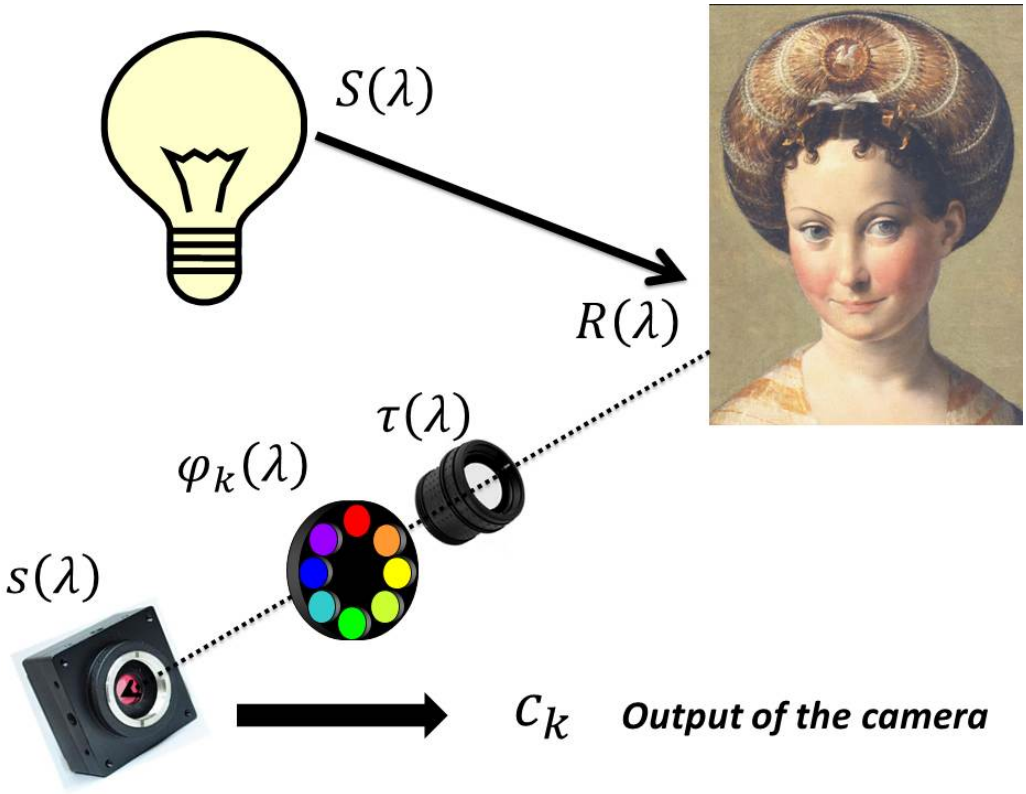


FIGURE 2.1: Image acquisition process

tral sensitivity of the sensor with *s(λ)*. Supposing for a while to neglect any noise contribution at the sensor, the camera response c_k at each pixel is the following:

$$c_k = \int_{\lambda_{min}}^{\lambda_{max}} R(\lambda) \cdot \kappa_k(\lambda)\, d\lambda \tag{2.1}$$

where: $\kappa_k(\lambda) = S(\lambda) \cdot \tau(\lambda) \cdot \varphi_k(\lambda) \cdot s(\lambda)$ represents the convolution kernel of the k^{th} channel. If some kind of noise δ_k is taken into account, then with a CCD sensor the system is linear, i.e. the noise adds linearly to c_k. In other situations such a simple circumstance may not hold true (the assumption of a noise-free system is unrealistic), because of possible stray-light contributions or electronic amplification in the camera that introduce non-linearities in the response (Farrell and Wandell, 1993). Thus, if

the the non-linearities of a linear camera are modeled as (Hardeberg, 1999):

$$\check{c}_k = \Gamma(\int_{\lambda_{min}}^{\lambda_{max}} R(\lambda) \cdot \kappa_k(\lambda)\, d\lambda + \delta_k) \tag{2.2}$$

then it is easy to obtain the response by inverting the 2.2: $c_k = \Gamma^{-1}(\check{c}_k)$. Putting the 2.1 in matrix notation demands uniform sampling of colour spectra (Trusseli and Kulkarni, 1996):

$$c_k = \mathbf{R^T}\kappa_k \tag{2.3}$$

with possible additive contribution of noise δ_k. Here N different equal wavelength intervals are considered, with:

$\kappa_k = (\kappa_k(\lambda_1), \kappa_k(\lambda_2), \ldots, \kappa_k(\lambda_N))^T$ and $\mathbf{R} = (R(\lambda_1), R(\lambda_2), \ldots, R(\lambda_N))^T$ representing the vectors of the spectral sensitivity of the *k-th* channel and the spectral reflectance of sampled data, respectively.

In hyperspectral imaging it is mandatory to record a so-called spectral data cube, where the added dimension is obviously referred to the spectral information. This is typically achieved by either spatial or spectral multiplexing of time-sequential images on the surface of a two-dimensional detector array. Sometimes it is necessary to record hyperspectral content from images of time-varying phenomena, but in these particular situations time-sequential techniques relentlessly fail (Fletcher-Holmes and Harvey, 2005) because of the onset of spatial misalignment of the images and distortion in the recorded spectra (Harvey et al., 2000). So, traditional techniques are fundamentally inadequate for recording phenomena that vary on a timescale shorter than the time-interval necessary for the reconstruction of the data cube.

Having the spectral dimension returned in time, as occurs in filtered cameras, with simultaneous capture of both spatial dimensions is not the only way to retrieve a spectral image, as another widely accepted alternative is represented by the use of scanners that include optical systems to acquire images in a line-like fashion; more precisely,

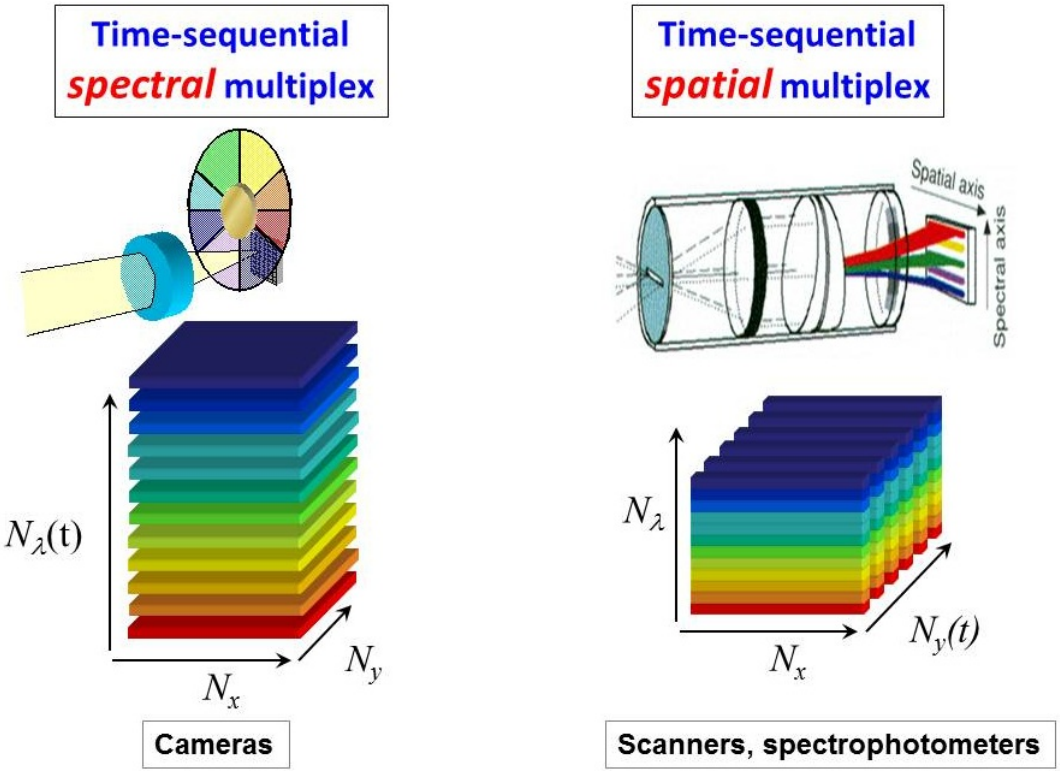


FIGURE 2.2: Schematic representation of the two types of time-sequential multiplex: spectral (left) and spatial (right)

in a single acquisition a line of the object of a fixed length in one spatial direction is imaged while simultaneously recording the *SRF* of each pixel. The remaining spatial dimension is recovered by replicating the aforementioned procedure through suitable mechanical scanning of the instrument to span the entire object. The first procedure is named *Time-sequential spectral multiplex* and is typical of multispectral cameras employing filter wheels, whereas the other procedure is commonly encountered in scanners and spectrophotometers and is named *Time-sequential spatial multiplex* (Figure 2.2). For applications where the luminance levels are very low, a common approach to increase the signal-to-noise ratio is to employ time-sequentially scanned Fourier-transform imaging spectrometers (Fletcher-Holmes and Harvey, 2002).

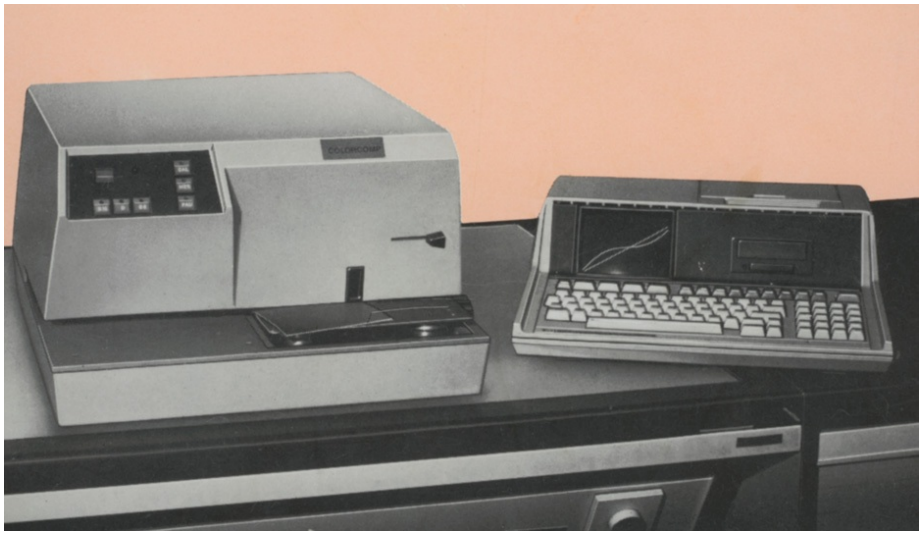


FIGURE 2.3: The COLORCOMP™ instrument

2.2 Multispectral and hyperspectral instrumentation

2.2.1 Instrumentation at the Aachen University of Technology

Among the earliest contributors to the establishment and consolidation of multispectral color technology, the Aachen University of Technology in Germany has played a pioneering role since 1974 when the first industrial multispectral spectrophotometer, called COLORCOMP™ (Figure 2.3) and launched into the market ten years later, was devised and designed for fast parallel colour capture (Heydt, 1984). For the application it was required, the measurement of printed colours in the reproduction process of a printing factory, it employed 16 channels for accurate description of the colour stimuli of surface colours. The light reflected from a colour patch is delivered in front of an equal number of photodiodes, each equipped with narrowband spectral filters to sense the light. Even though instrumental details have remained strictly confidential, the COLORCOMP™ has gained so much success that, with good reason, can be considered as the source of multispectral technology in Aachen. In the same years, with the awareness of the considerable drawbacks arising from metamerism in

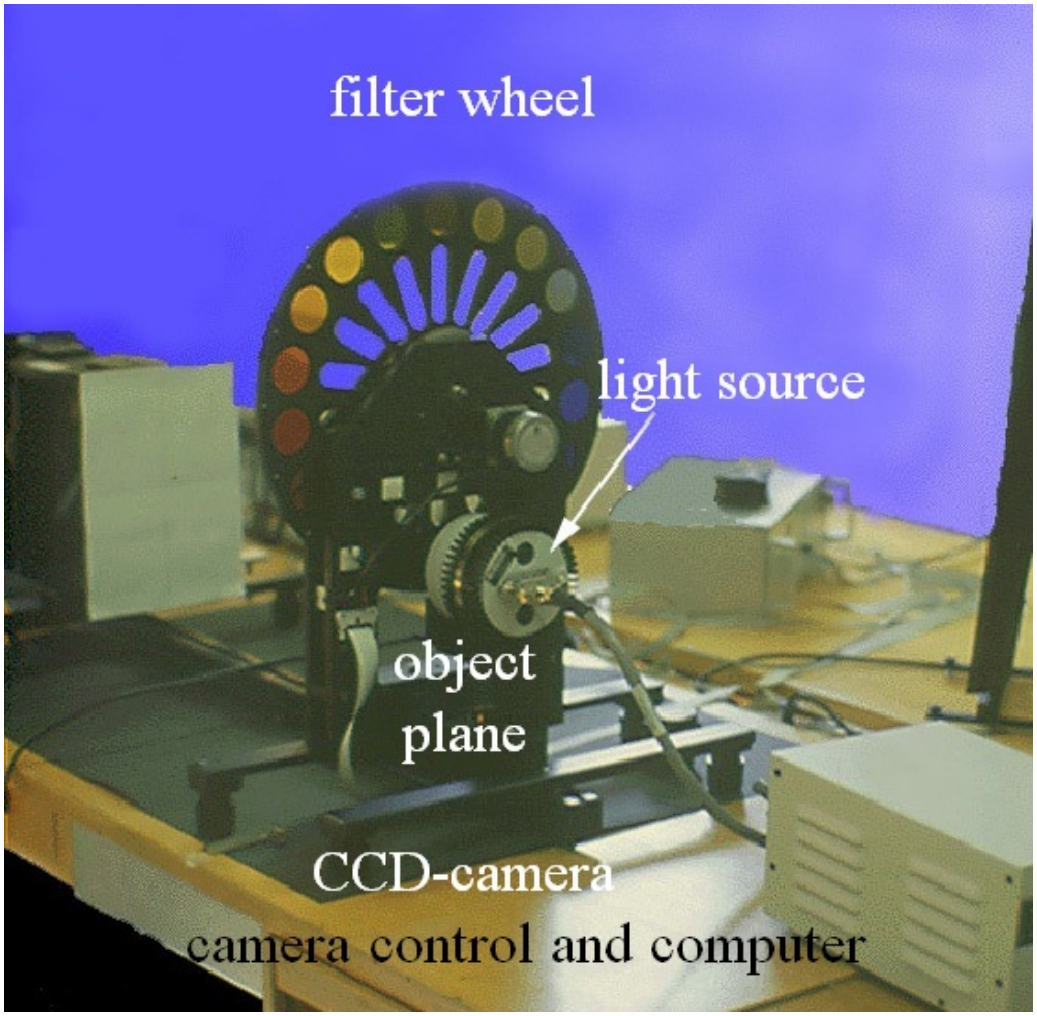


FIGURE 2.4: The first multispectral camera setup at Aachen

imaging by trichromatic cameras and image scanners, scientists at Aachen decided to export the technological solution of COLORCOMP™ to the capture of full images, i.e. they used 16 spectral scans to reproduce the colour of images. The first experimental setup of multispectral camera was released in 1991 and schematically consisted of a monochromatic CCD camera with white halogen light illuminating the image plane. A 16 interference filters wheel was used to provide narrow band spectral illumination at the image plane (Figure 2.4). Almost 10 years passed after prototype definition before

a commercial instrument was made available on the market by "Color Aixperts"[1] in 2000 (Figure 2.5); during this lag, the scientists at Aachen had to concentrate on a number of issues that affected the prototype. They worked to provide uniformity of the illumination, to get rid of multireflection problems between optical surfaces and interference filters, to enhance mechanical stability, to address sensitivity and spectral problems, etc. (Keusen, 1996; König and Praefcke, 1998). The camera provides various image resolution up to 3000×2000 pxl and has reduced the acquisition times down to 10 s, with colour differences as low as 1.2 CIEΔE94 units. A suitable color management processor is the link between the commercial system and the cathode ray tube used to display the image output. It is thus possible for the user to compare the result of the reproduction on the display with the original on a dedicated viewer (Figure 2.6), under a number of selected light sources and with a purposely chosen background.

2.2.2 The multispectral scanners at London National Gallery

The VASARI *instrument*

By the late 1980s, at the National Gallery of London a European project was funded within the ESPRIT II research framework, involving the most renowned gallery scientific departments, universities and companies from around Europe (Birkbeck College's Art History Department, Munich's Doerner Institute, Louvre's Research Laboratory, Telecom Paris' Ecole Nationale Superior de Telecommunication, TÜV Bayern, etc.). This project was named VASARI (*Visual Arts System for Archiving & Retrieval of Images*), from a suggestion in honor of GIORGIO VASARI the "Father of Art History". From its early planning, it was aimed at developing technology to solve problems and meet needs in the real world of users of art and art history. The

[1] `http://www.caddon.com`

FIGURE 2.5: Multispectral camera from "Color Aixperts"

most important result was the development of a colorimetric scanner system for direct digital imaging of paintings, intended to replace film photography due to its higher colour accuracy and resolution than conventional film. In other words, it was perhaps the earliest attempt to demonstrate the feasibility of high resolution of colorimetric imaging directly from paintings.

The great advantages offered by digital imaging were, at that time, the increased accuracy in terms of colour and improved stability with aging than conventional photographic materials; moreover, the scanner was intended to create more accurate

FIGURE 2.6: Viewing station associated with the multispectral camera of Figure 2.5

permanent records for research, conservation, teaching, printing an so on (Martinez, 1991). For the sake of acquiring images that could clearly show the net of cracks (*craquelures*) and brushstrokes on the painting, the inventors of the scanner made a careful study about the spatial resolution needed, based on the average dimensions of the cracks (typically 100 µm), which turned out to be $10\,\mathrm{pxl\,mm^{-1}}$ to $20\,\mathrm{pxl\,mm^{-1}}$ (Martinez and Hamber, 1989) ($40\,\mathrm{pxl\,mm^{-1}}$, in case of extremely fine patterns). For these reasons, commercial cameras presented severe limitations because accurate imaging took approximately $10^8\,\mathrm{pxl\,m^{-2}}$, which none of them could ever fulfil. These shortcomings were overcome by taking as many frames as necessary, by a commercially available camera, and then joining them with the *mosaicing* technique. The system underwent several improvements and technical refinements with time and here the ultimate layout is presented.

A $500 \times 290\,\mathrm{pxl}$ monochrome CCD camera, purposely designed at the University of Munich (Lenz and Lenz, 1990a), with the sensor masked so as to reduce the size

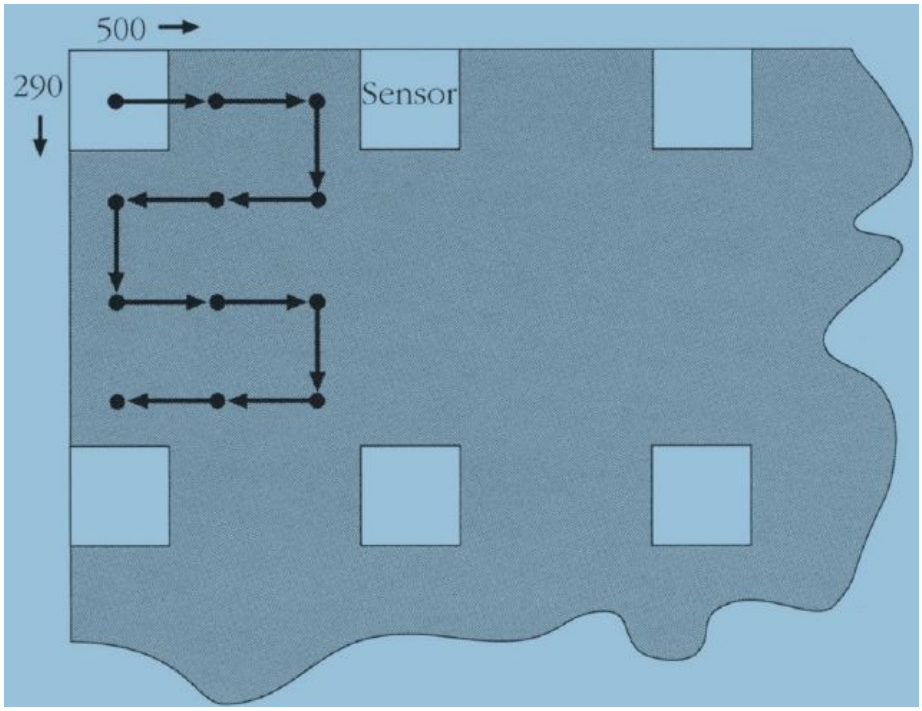


FIGURE 2.7: Scheme of the masking pattern in the PAD technique; the sensor chip translates as shown

of the active photosites, therefore unable to cover the whole area of the chip, was employed to acquire 2994×2320 pxl images. Such a large image was formed by micro-scanning the intermediate grid positions of a low resolution CCD array (Lenz and Lenz, 1990b); the final image was then constructed by interlacing sub-images (Hardeberg, 2001). This micro-scanning technique is called *Piezo-controlled Aperture Displacement* (PAD), with piezoelectric ceramic crystal actuators, mounted around the sensor, that shift the chip by minute amounts (a fraction of the active site separation), as shown in Figure 2.7.

For full resolution 48 images sufficed; a lower resolution image could be taken to facilitate focusing. The camera, in its most recent release, was supplied with an A-D converter that provided 256 levels, i.e. $8\,\text{bit}\,\text{pxl}^{-1}$.

The mechanics of the entire system has a XYZ positioning system placed on a rigid steel base with vibration damping blocks on a concrete floor (Figure 2.8). Two stain-

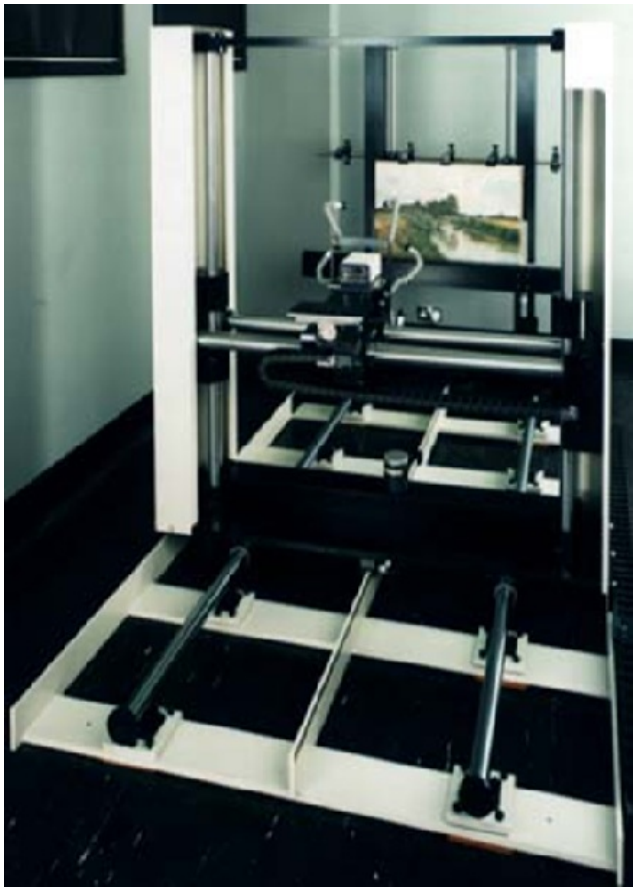


FIGURE 2.8: Overall view of the mechanics of the VASARI system

less steel rails, each 2.5 m long, allow the portal movement in front of the painting. The horizontal axis is mounted between two vertical rods on the portal. Both axes are motorised and computer controlled and can run up to 1 m in steps of 50 µm, in both directions. An easel, mounted at the far end of the base, is used to keep the paintings in place. The third motorised axis (Z) allows the movement of the camera, orthogonally to the painting, for fine automatic focusing.

The repositioning accuracy was better than 10 µm, whereas the absolute accuracy was about 30 µm. Details about the overall alignment of the system are reported in (Saunders and Cupitt, 1993).

The collection optical head was a commercial high-quality lens (Figure 2.9), intro-

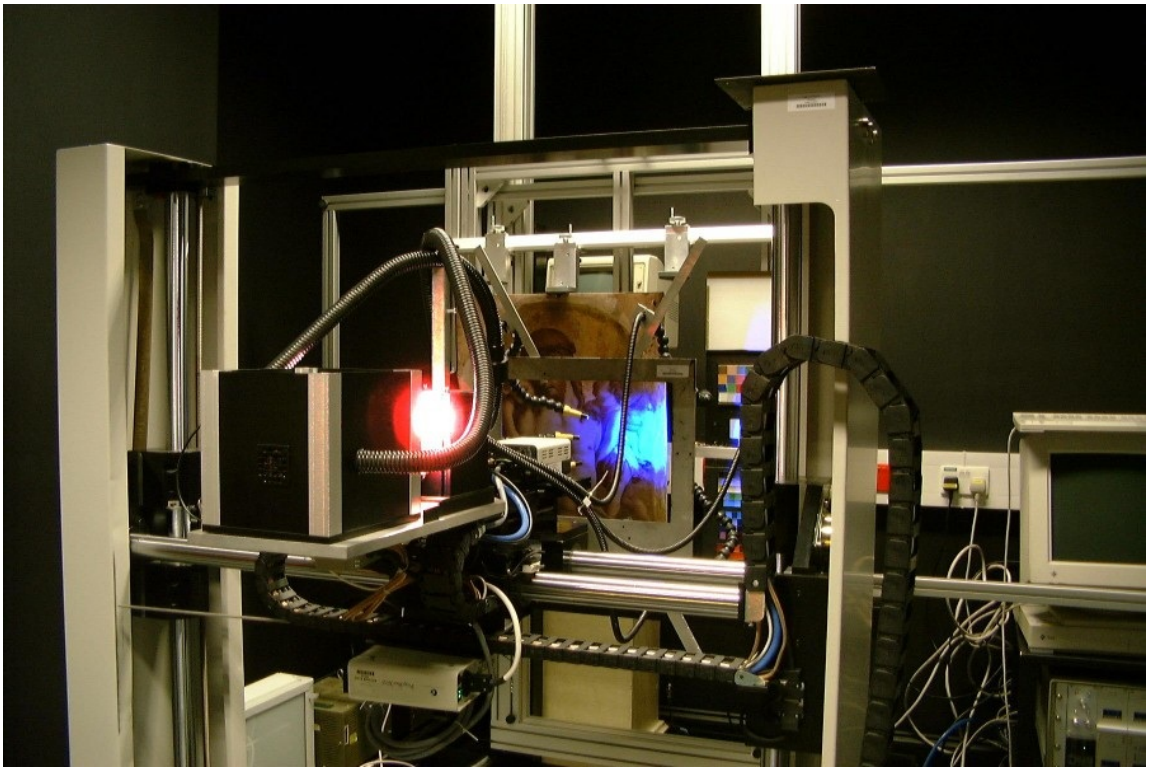


FIGURE 2.9: The VASARI scanner imaged during a working session

ducing very low geometric and radiometric distortion, if none, and having a very high modulation transfer function.

In spite of the most common trichromatic approach for colour acquisition, the colour of each point of the image was measured in terms of a generally recognised colour notation, the CIE1931XYZ standard colour space, using a multispectral approach, i.e. through seven filters (Saunders and Hamber, 1990) arranged on a wheel, with roughly Gaussian characteristics, with 70 nm bandwidth (FWHM) and their peak transmittance ranging from 400 nm to 700 nm, with 50 nm steps.

The lighting system moves as a whole with the camera, thus ensuring identical light distribution anywhere on the painting. It is composed of two 24 V, 250 W DC halogen lamps, each in a separate enclosure with a set of collimating optics and IR-blocking filters. A Y-shaped fibre optic guide receives light from the lamps and guides it to the

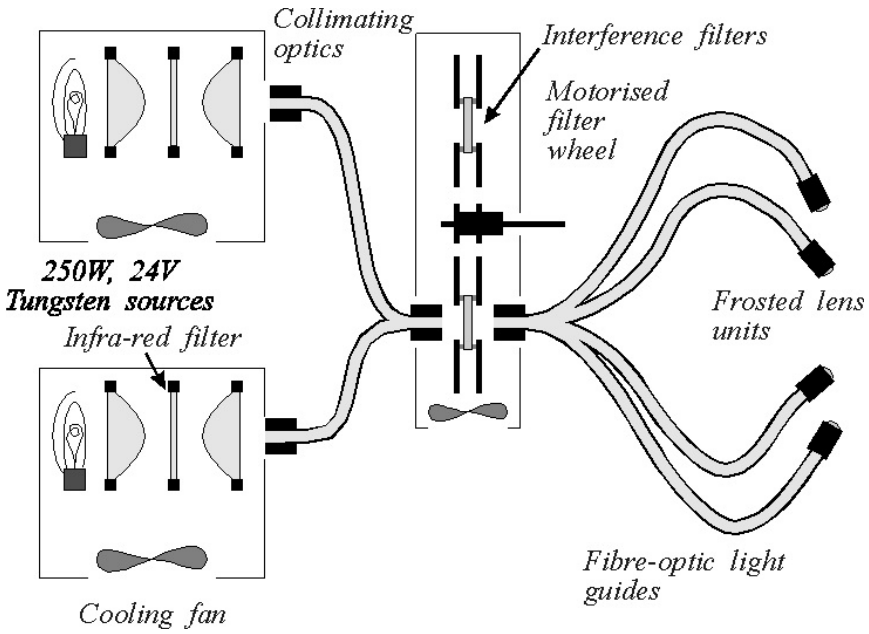


FIGURE 2.10: Sketch of the light projector of the VASARI system

filter box, where a motorised filter wheel is placed; there it crosses one of the seven interference filters and lastly passes into a second fibre optic guide, which divides into six terminations that provide even illumination over the field-of-view (Figure 2.10).

Although it was possible to record the reconstructed spectrum for each pixel, in the early 90's unrealistically large requirements with digital data storage existed. Therefore spectral information was abandoned in favour of colorimetric data alone, which were rather calculated directly from the response of the camera through every single filter to a set of a number of colour standards, than from spectral data. Since the calculation involved the use of a set of 7 colour standards (objects with known XYZ values) or more (Saunders and Cupitt, 1993), it was demonstrated that the 24 colour patches of the MACBETH COLORCHECKER CHART™ [2] worked better than other commercial counterparts. The British standard colour difference formula ΔE_{CMC} developed by the Society of Dyers and Colourists Colour Measurement Committee

[2] `http://www.xrite.com`

(CMC), was preferred to quantify colour differences throughout the VASARI project. The imaging system requires a calibration procedure after the acquisition, in order to perform colour calculation for every pixel and to generate an image suitable for display on a workstation's screen. The easel accommodates also a set of resolution targets, namely one for the accurate determination of scanning resolution, one for the detection of non-linearities in the CCD, one to calibrate colour (MACBETH COLORCHECKER CHART) and a white reference to measure non-uniformities in light distribution.

The VASARI instrument was successfully used for a number of research programs, to monitor changes both in colour and in surface texture, and to support this work a package of image processing programs was purposely written (Cupitt and Martinez, 1996).

The CRISATEL *instrument*

In the framework of the IST-CRISATEL european project, started in 2001 in the field of conservation and restoration of canvas paintings, the National Gallery of London together with a group of private enterprises, research institutes and museum laboratories, was involved in the development of a multispectral scanning imaging system (scanner) allowing direct fast capture and processing of paintings at ultra-high resolution and with a homogeneous lighting system, from the ultraviolet to the very near infrared region. The name CRISATEL is the acronym of *Conservation Restoration Innovation Systems for image capture and digital Archiving to enhance Training, Education and lifelong Learning.* Among the most active contributors were Telecom Paris' Ecole Nationale Superior de Telecommunication, Laboratoire des Instruments et Systemes, Centre de Recherche et des Restaurations des Musees de France, Université Pierre et Marie Curie, Lumiére Technology, Pébéo Industries, University of Cyprus,

etc.).

Within the goals of the project was a research activity aimed at analysing the spectral characteristics of varnish layers, in order to discount the effect of an aged varnish from an image of a painting. Moreover, the colour-accurate prints of large-format output were intended to assist the restoration of paintings, to substitute for the originals in exhibitions, in case of fragile works of art, and to create an archive of digital images for long-term preservation. With the disadvantage of the large fixed apparatus of the VASARI system, the new instrument had to keep into account and overcome this limitation by considering the fundamental requisite of effortless portability.

The CRISATEL multispectral camera, JUMBOSCAN™ developed by Lumiére Technology[3] in Paris, has a vertically mounted 12 000 pxl linear array CCD sensor, with square pixels of 6.5 µm without inter-pixel space, usually employed in the satellite and aerospace industry.

It is mechanically displaced by a stepper motor up to 20 000 horizontal positions, corresponding to 13 cm, with a full displacement mechanical precision of 12 µm and a repeatability of 3 µm in the positioning of the CCD (Cotte and Dupouy, 2003). This means that images up to 360 Mpxl can be generated. The camera box is equipped with a cask (Figure 2.11) that accommodates 13 interference filters, specially made for the CRISATEL instrument, ten covering the UV and visible spectrum and the remaining the very near infrared domain, for an overall range from 380 nm to 780 nm, equally spaced with steps of 40 nm; the former have 40 nm bandwidth (FWHM), whereas the latter a broader 100 nm bandwidth (Figure 2.12). They are all rectangular in shape ($78 \times 7\,\mathrm{mm}^2$) and are positioned 10 mm in front of the CCD, where they are switched by means of rotating half-barrel mechanism (Figure 2.13).

The camera has a displacement system for fine focusing, controlled by a stepper motor with a total excursion of 10 cm with a precision of 20 µm, of the full body

[3] `http://www.lumiere-technology.com/`

FIGURE 2.11: Cask containing the 13 filters of the CRISATEL instrument

to compensate for differences in scale, for every single spectral channel, due to lens' inherent chromatic aberrations that introduce a dependence of the focal length on the wavelength (Ribés et al., 2003).

The electronic architecture of the CCD has a two parallel channels current system, i.e. one for even and one for odd located photosites; the raw signals from each channel pass through an analog amplifier. These amplifiers have offset, gain and exposure time control parameters each, that are properly adjusted during the calibration stage before a multispectral acquisition. An A-D converter outputs 12 bit signals for each photosite, ensuring high-dynamic range to the high-resolution images.

The need to protect archived objects by preventing flood of light on the documents during the scanning process, motivated the Lumiére Technology to devise a brand new projector of light, the JUMBOLUX™. It is a motorized projector, producing just a

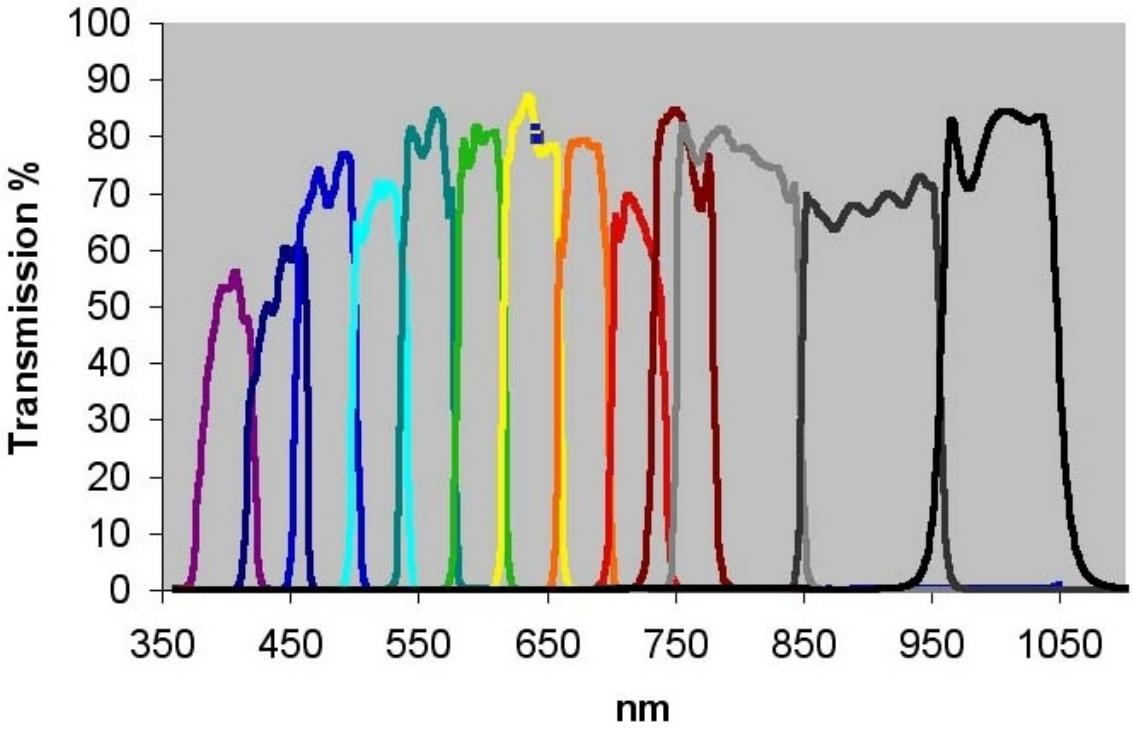


FIGURE 2.12: Transmittance curves of the 10 UV-Vis-NIR filters of the CRISATEL instrument

band of light on the object synchronized with the linear motion of the JUMBOSCAN™ sensor. It consists of two banks of elliptic lights (Hydrargyrum quartz iodide, 150 W), positioned at left and right sides of the camera, that deliver a 7.5 cm wide and 3 m long 10^5 lm homogeneous strip of light at 2 m distance, for momentary illumination of the object (3 s), that moves on the surface of the imaged plane. The combination of higher light intensity and larger CCD signal enables fast capturing, while protecting light-sensitive watercolors and chemical-based papers drawings that are easy to fade out.

Instrumental portability allow the users to operate effortlessly almost anywhere, to scan reflective and transparent objects on copyboards, on easels and on walls (Figure 2.14).

With a white homogeneous diffuse tile in place of the object, the via-software auto-

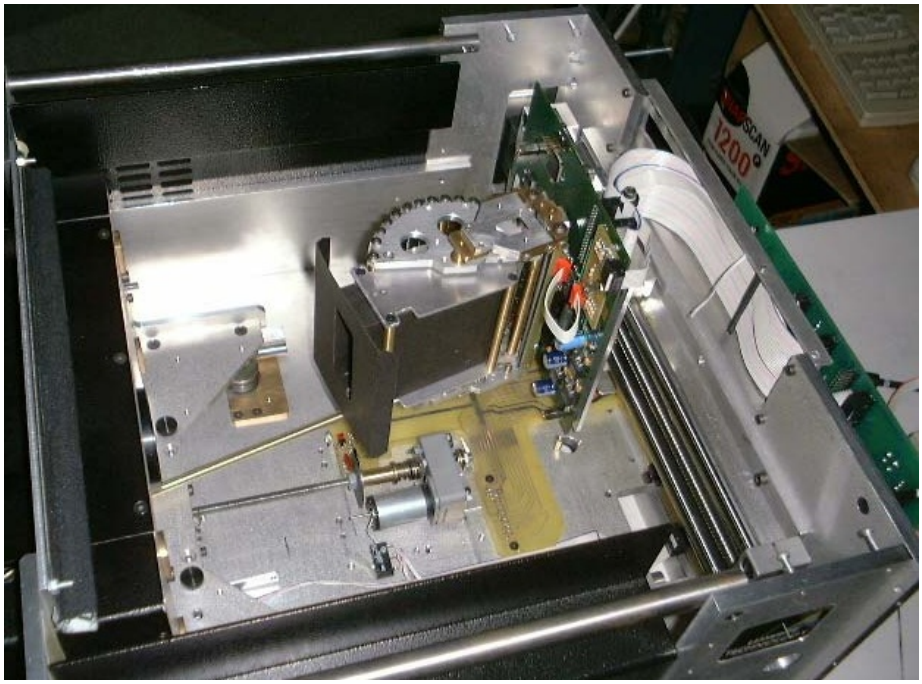


FIGURE 2.13: Interior view of the camera of the CRISATEL instrument

matic calibration procedure adjusts the gains and offsets of the two amplifiers and the exposure time of the sensor, aimed at maximizing the signal-to-noise ratio. Then a measurement of the non-homogeneity of the spatial distribution of the illumination on the white tile follows, with corrections to ensure good uniformity. Furthermore, the non-homogeneity of every single photosite of the CCD is measured by operating the camera with an opaque cap covering the objective lens (Ribés et al., 2005).

The lack of commercial charts containing colour patches made of pigments that are stable to illumination and with continuous curves of the SRF was resolved with a custom-made chart for image colour calibration and varnish thickness measurement. It was made at Pébéo Industries in Marseille and it consists of three parts containing 81 coloured patches and 36 neutral patches, made with pure pigments mixed with acrylic medium. It is divided in three blocks with the first uncovered, whereas the remaining are covered with a matt and a glossy varnish, respectively (Figure 2.15)

FIGURE 2.14: Complete setup of the CRISATEL instrument during a working session

(Cortés, 2003).

In 2004 the scientist who developed the CRISATEL system had the rare privilege to digitize the world masterpiece "*Mona Lisa*" by LEONARDO DA VINCI, capturing 13 digital files of 450 MB each (Elias and Cotte, 2008; Cotte and Dupraz, 2006).

The images of all the investigated paintings were fully processed using VIPS (the acronym of VASARI Image Processing System), an Open Source image processing software developed at the National Gallery of London, for the manipulation of extremely large images (Cupitt and Martinez, 1996).

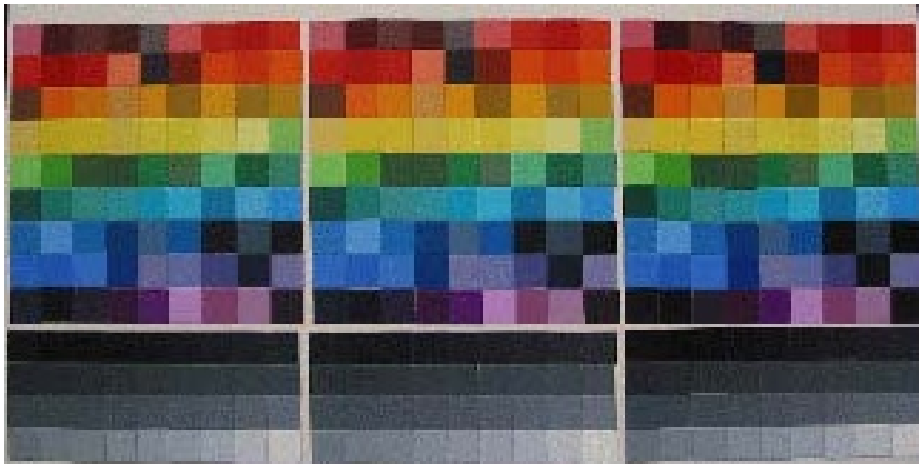


FIGURE 2.15: The CRISATEL colour chart

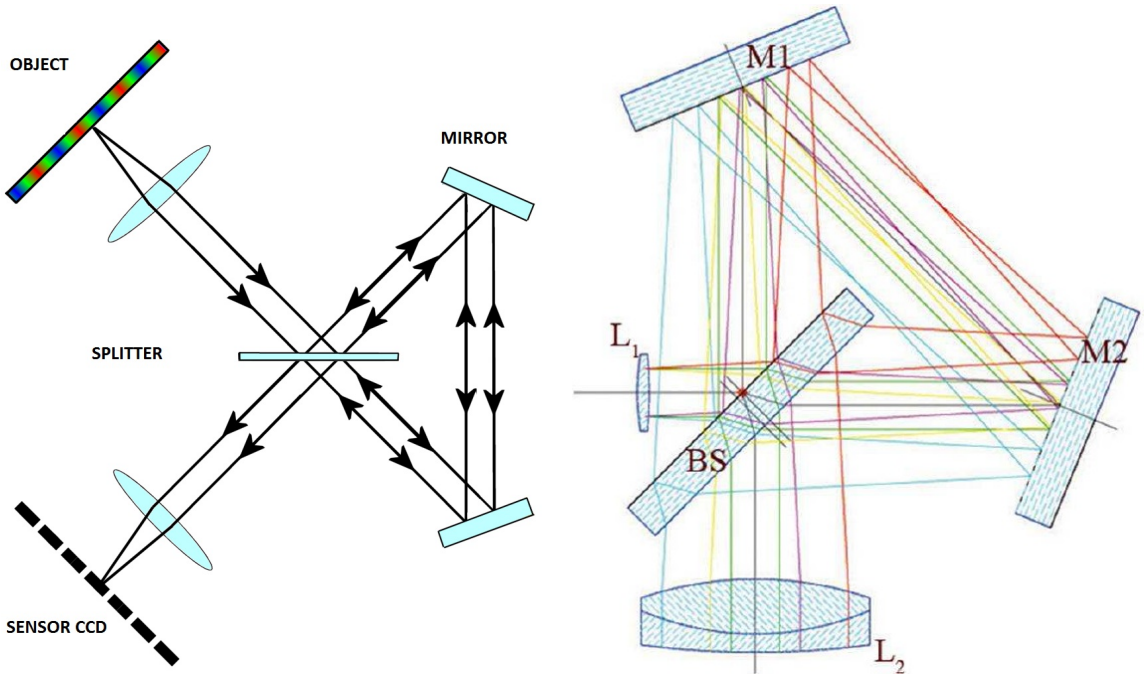
2.2.3 The SpectraCube™ interferometric camera

An interesting camera based on the Fourier-transform (FT) interferometry is the SPECTRACUBE™ by Applied Spectral Imaging[4]. In spite of the most commonly employed scheme in molecular spectroscopy equipments, the Michelson type, this device was designed following the Sagnac scheme (Figure 2.16).

The *Sagnac interference* (Sagnac, 1913) is one form of interferometry in which a rotation results in a phase shift between clockwise and counterclockwise propagating optical fields (Figure 2.17). This optical configuration encompasses a class of instruments having a common path for the two rays of the interferometer and this is the reason why they are also known as *common path interferometers*, as opposed to *tilted interferometers*, like the Michelson-Morley and Mach-Zender instruments.

As a consequence, the Sagnac interferometer is extremely stable and easy to align, even while using an extended broadband source. In the classic Sagnac system a beam of light enters the system and is split into two paths having opposite directions. A number of mirrors is usually employed to accomplish either triangular or square trajectories. When the light exits the apparatus it forms an interference pattern,

[4] `http://www.spectral-imaging.com`



(a) (b)

FIGURE 2.16: (a) SPECTRACUBE™ imaging camera layout (b) Ray-paths coloured as a function of the incidence angle in the layout of a triangular Sagnac interferometer

that depends on the angular velocity of the platform. In fact, the lines of the interference pattern undergo a shift, with respect to their position at null rotation, as a consequence of the displacement of the entry/exit point during the transit time of the light. If the angular velocity of the platform is varied, then the corresponding interference pattern exhibits a phase shift that is particular to that velocity. In Figure 2.16(b) a collimating lens L1 creates a plane wave from every object point with a direction that depends on the viewing angle of every particular point (coloured rays). Light enters a beam-splitter BS and is reflected at the folding mirrors M1 and M2. The second focusing lens L2 focuses the emerging light onto the output focal plane that holds a sensor, where a stationary pattern of interference fringes of equal thickness, that are localized at infinity, is formed (Barducci et al., 2010). It has been demonstrated that the BS accounts only for a part of the phase-delay, the remaining

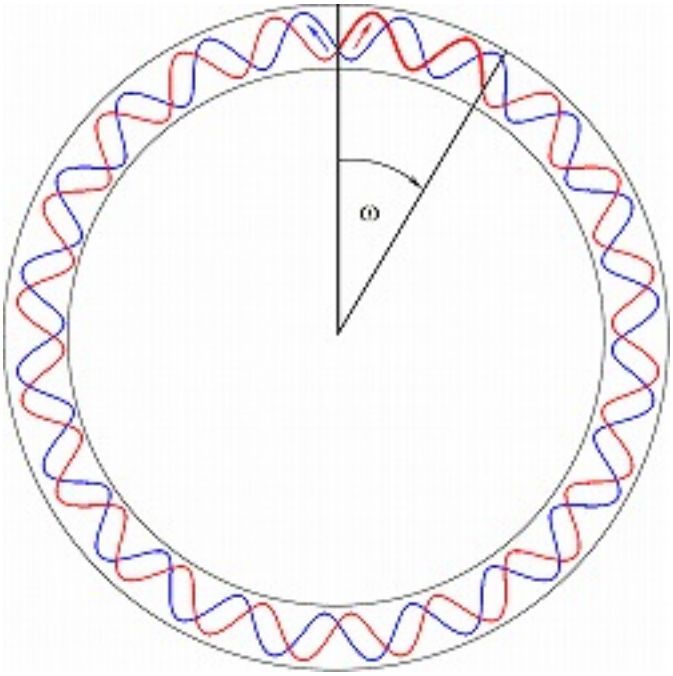


FIGURE 2.17: Sketch of the frequency shift between counterpropagating fields in a Sagnac interferometer

being originated outside, and that the overall optical-path difference varies linearly as the incidence angle, with respect to the instrumental optical axis, of the incoming light changes (Barducci et al., 2001, 2002). The lack of an entrance slit has the effect of superimposing the interference pattern to the image of the object. Moreover, the introduction of a relative motion between the object and the sensor makes every pixel exploit the whole interference pattern, so that both the interferogram and the spectrum can be inferred through algorithms based on the Fast Fourier-transform (FFT).

Most times in real instruments, the BS is rotated in a stepwise manner to acquire hyperspectral images; this mechanical feature makes the Sagnac interferometers very stable, easy to align and almost insensitive to vibrations, even at the shortest wavelengths. So, this advantage extends the working range of the traditional FT interference spectroscopy from the infrared to the visible and short-wavelength near-infrared regions

(Sun, 2010). In order to provide a reasonably high signal-to-noise ratio, an ideal interferometric instrument should have a sensor with a uniform spectral response, because the signal collected at every position of the mirror is the superposition of all the wavelengths composing the spectrum under analysis. Some noteworthy shortcomings are represented by the varying sensitivity in the entire spectral region, the amount of vignetting that affects fringe visibility and the need for for heavy data pre-processing to restore the correct radiometric level of the signal in the interferogram.

The SpectraCube™ camera can be considered as an extensive enhancement of the famous Fourier-transform infrared (FTIR) spectrometer. While a FTIR measures a single spectrum by using an interferometer, the SpectraCube™ uses the same principle to measure many spectra, one for each point in the image. The system can be attached to a microscope, telescope, lens or any other fore-optics (Figure 2.18).

Unlike the classic Sagnac system which rotates as a whole, the SpectraCube™

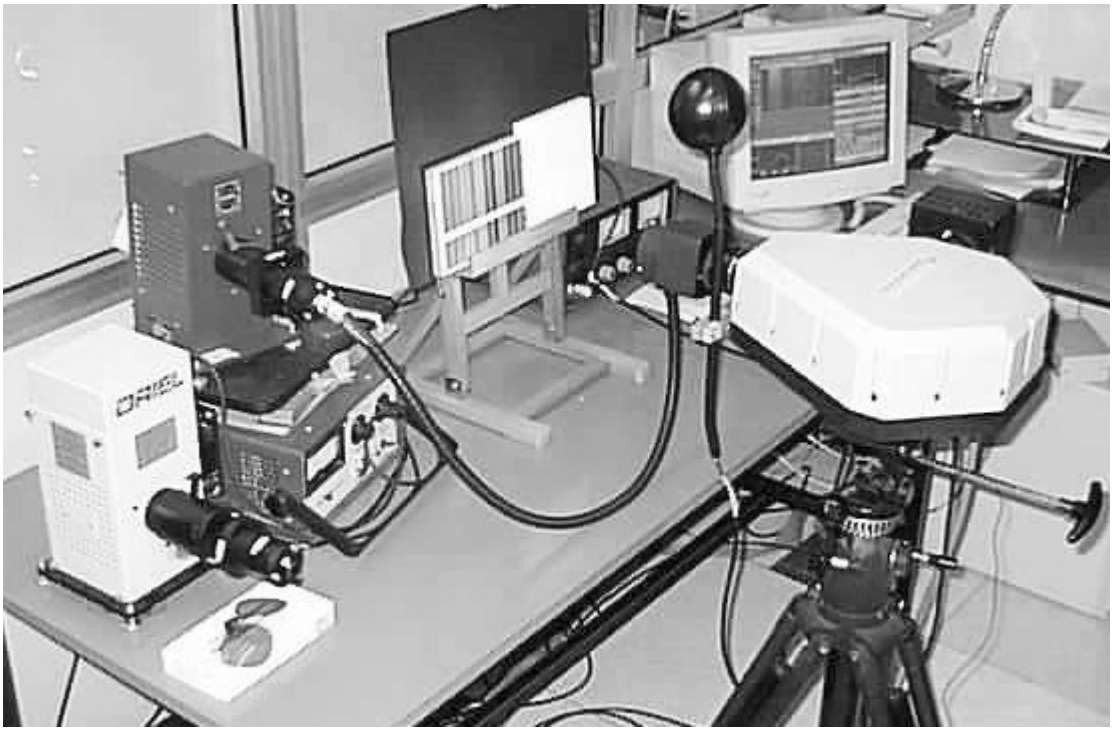


Figure 2.18: The SpectraCube™ camera mounted on a tripod in a laboratory

layout is a somewhat modified Sagnac layout in that employs a double-face rotating BS, actuated through a computer controlled rotating stage, and a CCD sensor with 512×512 pxl. Such a BS is the counterpart of the translating mirror in the Michelson scheme, because it allows to scan of the optical path differences with constant angular steps. In such a way, every single monochromatic component of the radiation hitting the sensor gives rise to a periodic signal ranging between a maximum and zero. The spectral components participate to the overall signal by superposition of sinusoidal signals with a period that depends on the corresponding frequency and with an amplitude that is proportional to the corresponding intensity.

It is well known that a signal originating from the superposition of sinusoidal components can be decomposed into the single components by Fourier analysis; therefore, the application of a suitable FFT algorithm to each point of the images recorded by the sensor allows to build the *spectral image cube* with the radiation coming from the framed object.

SpectraCube™ 's manufacturer specifications report FWHM bandwidths of 0.8 nm at 407 nm, broadening up to 7.85 nm at 1093 nm, for a measurement effected with 185 bands, provided that the modulation of the angular amplitude and the correct number of angular steps of the BS have been set. The accuracy is better than 2 nm. The instrument sometimes needs optical re-alignment, especially after heavy temperature excursions.

2.2.4 The Fabry-Perot hyperspectral imaging spectrometer at the italian National Institute of Metrology (National Research Council)

A novel interesting compact hyperspectral imaging device, based on the the Fourier-transform, was devised in 2010 at the National Institute of Metrology (INRIM) of the National Research Council (CNR) in Turin (Italy). It was intended for fluorescence microscopy to allow multilabeled analysis without the use of

selective filters, leading to a faster, simpler and more efficient instrument with high throughput with respect to classical instruments based on dispersive means.
The idea behind the instrument was the insertion of a Fabry-Perot interferometer (Figure 2.19) in between a photographic objective and a monochrome matrix camera (Figure 2.20).

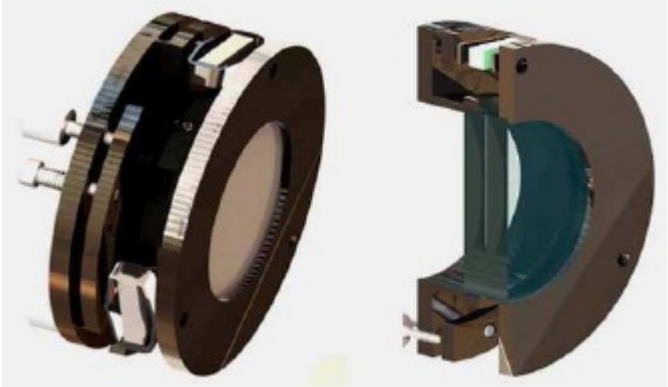


FIGURE 2.19: The Fabry-Perot interferometer and its section

FIGURE 2.20: The hyperspectral imaging device developed at the CNR-INRIM

The interferometer consists of two glass mirrors coated with a thin layer of aluminium

with a reflectivity around 25 %. The choice is in accordance with the requirements for the Fabry-Perot interferogram (reflectivity less than 30 % (Pisani and Zucco, 2011)) to approach a cosine form (2.4), in order to apply FT-based algorithms to calculate the spectrum (Jiao et al., 2003). The distance of the cavity is varied from 0 µm to 40 µm through three piezoelectric actuators. A video is captured while the cavity distance is scanned and so the light intensity on each photosite of the camera is recorded. The signals from the photosites are then processed to obtain an interferogram from which is possible to calculate the spectral composition of each pixel of the scene. A calibration procedure is mandatory for correct quantification of the distance between mirrors before to apply the FT algorithms, effected by the analysis of the interferograms of pixels illuminated with laser light (Pisani and Zucco, 2009), sent through a folding mirror on the surface of a holographic diffuser and then shining a beam-splitter placed in front of the objective (Figure 2.21).

In such a way, it is possible to associate to any voltage, driving the piezoelectric actuators, a distance between the mirrors of the cavity. These distances are necessary for the reconstruction of the interferogram. This calibration is required prior to any image captures.

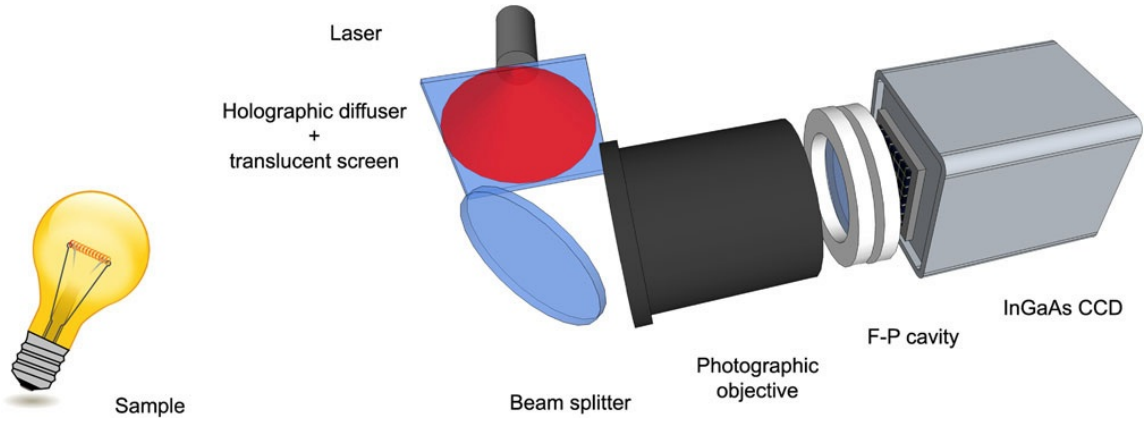


FIGURE 2.21: Schematic layout of the Fabry-Perot imaging spectrometer

The mathematical expression of the interferogram *I(δ)* with a broadband and contin-

uum source is, for the particular case of very low R:

$$I(\delta) \approx \int_{-\infty}^{+\infty} [S(\nu)(1-2R) + S(\nu)2R\cos(\pi\nu\frac{d}{c})]\, d\nu \tag{2.4}$$

where δ is the optical path difference of the interfering beams, ν is the frequency of the light, R the reflectance of the aluminium coating (approximately 0.25), d the distance between mirrors and $S(\lambda)$ represents the spectrum of the source.
If only the modulated part is considered, then a discrete cosine transform can be used to retrieve the spectrum $S(\nu)$, provided that two points per fringe are acquired. In this case, sampling the signal at finite intervals Δs with N discrete equidistant points, the spectrum is:

$$S[k \cdot \Delta\nu] = \sum_{n=0}^{N-1} I[n \cdot \Delta s] \cos(2\pi\frac{nk}{N}) \tag{2.5}$$

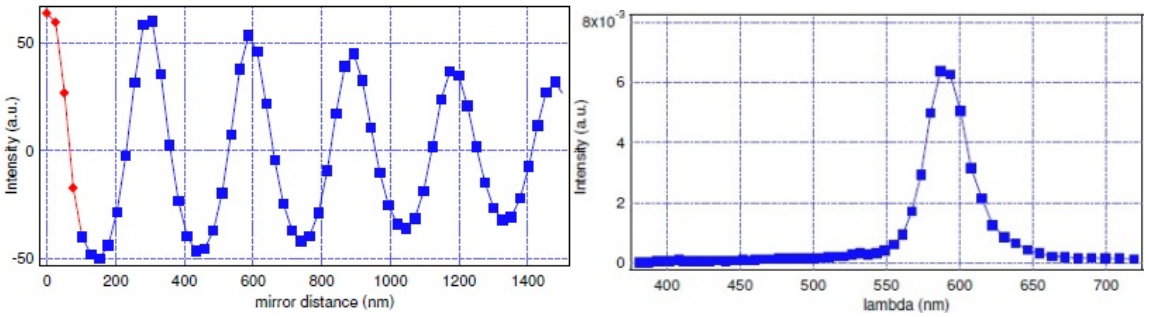
The resolution $\Delta\nu$ is inversely proportional to the maximum retardation $N\Delta s$:

$$\Delta\nu = \frac{c}{N\Delta s}$$

Because of the penetration depth of the radiation in the metallic coating when the mirrors are in contact, the interferogram never starts from the zero fringe and this can be exploited for an estimation of the penetration depth of the metallic layer. An algorithm and the use of an optical bandpass filter placed in the optical system are used to reconstruct the missing data points in the interferogram of Figure 2.22 (Pisani and Zucco, 2011). In the visible range from 400 nm to 720 nm the instrument has demonstrated hyperspectral imaging capability with a spectral resolution of 2 nm at 532 nm and a spatial resolution limited by the CCD sensor used (696 × 512 photosites).

2.2.5 Spectrophotometric scanner at Parma University

During the first years of the previous decade, the Department of Physics at Parma University in Italy made a very interesting optical scanner. It was intended for non



(a) (b)

FIGURE 2.22: (a) Corrected interferogram with reconstructed data points (red) (b) emission spectrum of a yellow light-emitting diode calculated from the interferogram aside

invasive applications concerning the conservation and restoration of paintings and other artworks' plane surfaces, since the comparison of images taken at different times is of great help for the diagnosis of the pictorial layer and allows to follow the historic evolution of the paintings. In order to fit such a purposes, these images must be of high spatial resolution without geometrical distortion and, moreover, have to contain device independent and high fidelity colour measurements. Since the *SRF* of a pictorial layer is a device independent physical quantity which enters colour calculation, the scanner returns a digital image with the *SRF* associated to each pixel. It is thus possible to:

- associate CIE colour to every single pixel of the image
- overcome the limits of trichromatic reproduction methods
- exploit the images in the future, because the spectral reflectance factor will be the reference quantity also for new colorimetric observers and models
- assess the conservation state of a painting as changes in the spectral reflectance factor are in connection with physic-chemical changes of the pictorial layer
- infer the materials used by an artist to overcome metamerism in the restoration

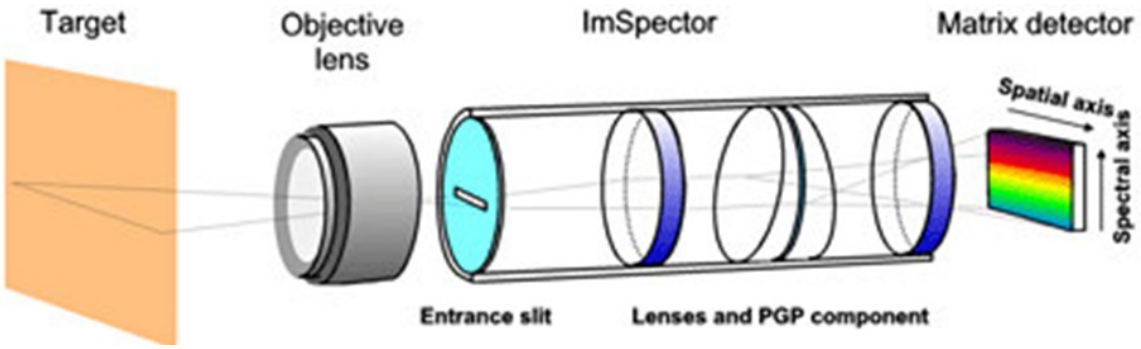


FIGURE 2.23: Drawing of the main constituents of the spectrophotometric scanner

The optical arrangement consists of a commercial transmission spectrometer, tailored for a ⅔ in CCD sensor, which accommodates a 25 µm entrance slit. It covers the spectral range from 400 nm to 780 nm with a nearly constant spectral resolution of about 2 nm. A monochromatic ⅔ in CCD camera (12 bit, 1280 × 1024 pxl, 9 fps) is coupled to the spectrometer; an objective creates an image of the painting on the plane of the entrance slit (Figure 2.23). The illumination is provided by two 150 W each halogen lamps, filtered beyond 750 nm, coupled to two optical fibers that are oriented at 45° with respect to the optical axis of the scanner, to accomplish the measurement geometry 0°/45°. The systems is secured to a rigid and massive platform that translates it during the scan.

In a single scan the instrument takes a vertical strip painting area of $7 \times 60\,\mathrm{cm}^2$, whereas in several scans it covers a total area of $120 \times 140\,\mathrm{cm}^2$. The different strips are then reassembled to return a single image with a dedicated software. To understand the way the optical system works, from Figure 2.23 it is clear that the presence of the entrance slit, between the objective and the dispersive system, tailors the field of view as a line. The plane of the object and the plane of the entrance slit are conjugate, so the image of the object-line is admitted into the spectrometer via the slit. Since the spectrometer has unit magnification, the rows of its sensor display the object-line, whereas the columns display the spectral values in the 400 nm to 780 nm

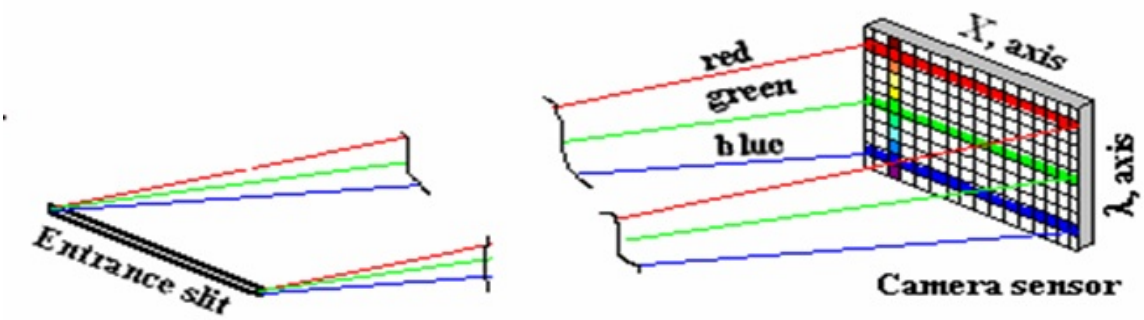


FIGURE 2.24: Spatial and spectral separation on the matrix-sensor of the spectrometer

range (Figure 2.24). Hence, the sensor is fully filled with white light entering the slit. This is made possible thanks to a particular optical element contained in the spectrometer, a CARPENTER *prism* or *grism*, which originates form the terms *grating* and *prism* (Figure 2.26). It is a dispersing element that provides in-line viewing for one wavelength, obtained by replicating a transmission grating onto the hypotenuse face of a right-angle prism. The prism refracts in-line the light diffracted by the grating (Loewen and Palmer, 2005).

The acquisition of one frame is nothing but the acquisition of the reflectance spectra of the single pixels of the imaged strip of the painting. Given the spectral range of colour measurements (400 nm to 730 nm) and the spectral resolution of the spectrometer (2 nm, approximately), then the reflectance spectrum of a single pixel consists of nearly 170 points (Figure 2.25).

The scanning system performs the measurement of a spectral image of a painting in four steps, as follows:

- the illuminant and the CIE standard observer are chosen
- after shuttering the lens, a number of black reference signal frames is taken and averaged to minimize noise
- a white reference signal frame is taken on a white reference tile; several frames

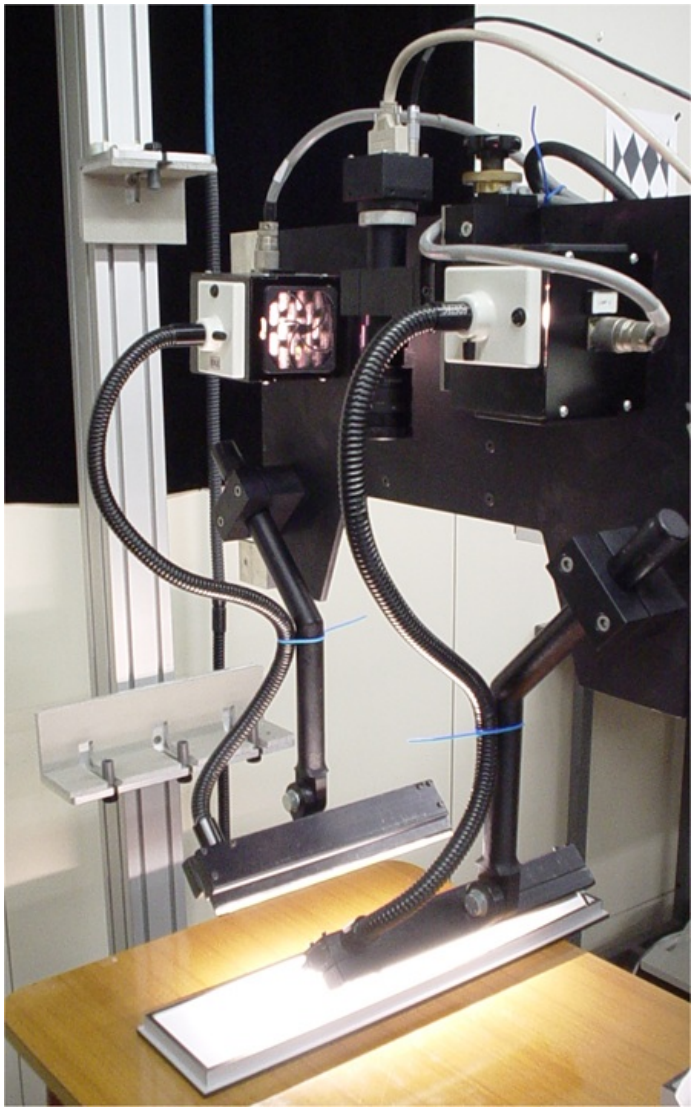


FIGURE 2.25: Close-up view of the illumination and optical collection system

are averaged in space and time for the twofold purpose of compensating the nonuniformities of the standard white radiance factor and to have the noise decreased

- the scan speed and the number of frames are selected

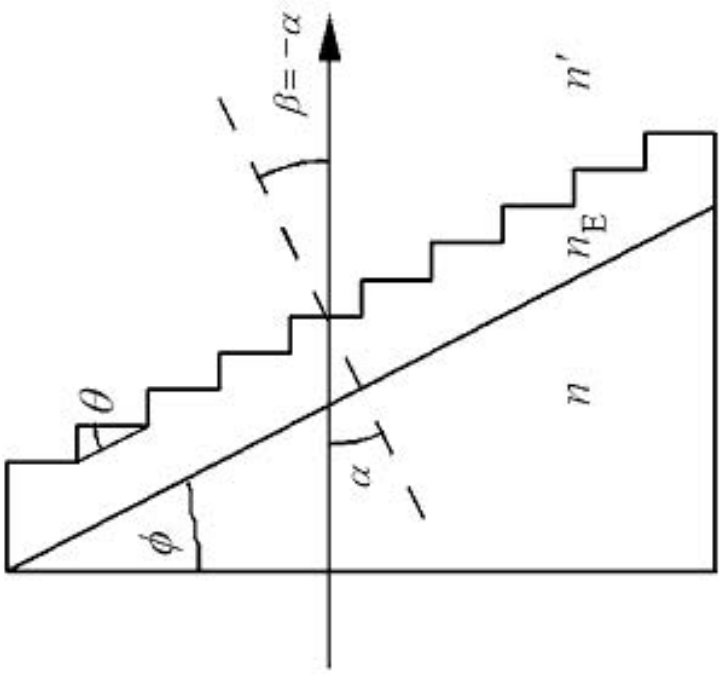


FIGURE 2.26: Cross-section of a grism

A dedicated software calculates the spectral reflectance of the object for any pixel, according to the following relationship:

$$R_m(\lambda_n) = R_W(\lambda_n) \cdot \frac{V_{nm} - D_{nm}}{W_{nm} - D_{nm}} \tag{2.6}$$

where nm are the indices of the rows and columns of the matrix sensor. Different wavelengths are located on separate columns of the sensor, so they are characterized with different values of the index n. In the 2.6, $R_m(\lambda_n)$ is the SRF at λ_n for the m^{th} pixel, $R_W(\lambda_n)$ is the spatially average spectral radiance at λ_n of the white reference tile. Moreover, V_{nm}, W_{nm} and D_{nm} are respectively the signals reflected by the object, the white standard tile and the black, as measured at the pixel (n,m).

The linearity of the device has been ascertained by taking spectral images of a suitable set of neutral tiles, encompassing an intensity ratio of nearly 1:250. This technique ensures that the measurement of the SRF is not a function of the spectral distribution

of the particular illumination source, of the spatial nonuniformity of irradiance, of the spectral transmission of the optics in the system nor of the spectral responsivity of the scanner (Antonioli et al., 2004).

The instrument has been shown to exhibit excellent agreement between the measured and certified values of the *SRF* of a set of reference coloured tiles. As a consequence, colour reproduction accuracy is good since colour differences are below the unit value both with the CIE 1976 and 1994 colour difference formulas. For all these reasons, the instrument has been successfully applied on a number of extremely valuable paintings by Leonardo da Vinci, Annibale Carracci, Jan Provost, etc. kept at the National Gallery of Parma.

2.2.6 The multispectral scanner at the italian National Institute of Optics (National Research Council)

Following the growing interest in the scientific world on multispectral imaging, as a non-destructive technique for the diagnostics of paintings, in 2003 the National Institute of Optics (INO) of the National Research Council (CNR) in Florence and Lecce (Italy) started the design of a multispectral scanner to document the conservation state of artworks, basically paintings, through spectral and colorimetric characterization of the whole paint layers in the 380 nm to 800 nm range. The instruments employing filter wheels or imaging spectrographs, to collect both spectral and spatial data, need suitable calibration procedures to correct for nonuniform irradiation and for geometrical distortion.

The instrument consists of a lighting system and a collection head, to fulfil the CIE standard $45^\circ/0^\circ$ for reflectance measurements (CIE, 1977), moved by two orthogonally mounted motorized translation stages.

An optical head collects the light scattered from the illuminated object point and focuses it on the end of a multimode optical fibre. Two low-voltage commercial halogen

lamps (current-stabilized to prevent unwanted changes in the spectral characteristics of the emitted radiation) irradiate a nearly 5 cm^2 object area. The reflectors deliver light with 10° divergence and are coated with a dichroic coating to filter-off a large amount of IR light in order to minimize surface heating.

The collection head is a catoptric arrangement of two concave spherical mirrors, namely a finite-conjugates Cassegrain-like telescope. The use of mirrors removes any chromatic aberration, whereas the curvature radii and the vertex-to-vertex axial displacement of the mirrors define the magnification of the optical head. In particular, the working *relative aperture* (or *f-number*) $f_{\#}$ is 2.4 object side and 3.6 image side, to obtain a magnification factor of -1.5 (Figure 2.27). Since the collecting fibre has a diameter of 200 µm, the object point that is imaged on the fibre end has a diameter of 130 µm approximately, with a good margin before the sampling step, fixed by the translating mechanisms at 250 µm, is exceeded. Radiometric simulations have demonstrated that the depth of field is about ± 250 µm.

There are other valuable features that make this instrument really appreciated:

- low load on the scanning system limited to the sources and the optical heads
- limited surface heating due to the simultaneous motion of collecting optics and sources
- uniform lighting due to point-by-point surface sampling
- the very narrow field of view ensures off-axis and aberration-free images (absence of geometrical distortion in the reconstructed image)

The detection system is a customized version of a commercial 32 channels photo-multiplier array, each of the channels filtered with either 10 nm or 20 nm bandwidth (FWHM) interference filters covering the 380 nm to 800 nm spectral interval. The sensitivity of the cathode is higher in the blue region, thus compensating the lack of

FIGURE 2.27: The finite conjugates catoptric optical head of the scanner

power of the lamps in that range. The frequent low reflectance of pictorial materials (2 % to 5 %) entailed the choice of a sensor which was made to operate at a sampling frequency of 50 Hz and with a high anode sensitivity (1×10^2 to 1×10^6). A 12 bit A-D data acquisition board permits high tonal dynamics.

The modular composition of the whole system allows the users to separate it (Figure 2.28) and make it ready to move for in-situ operation (Figure 2.29) for multispectral imaging up to 1.5 m^2, with a spatial resolution of 16 dots/mm^2.

The digitized output data of the acquisition are displayed as 32 monochromatic images, one for each channel. The calculation of the *SRF* is effected as in the case

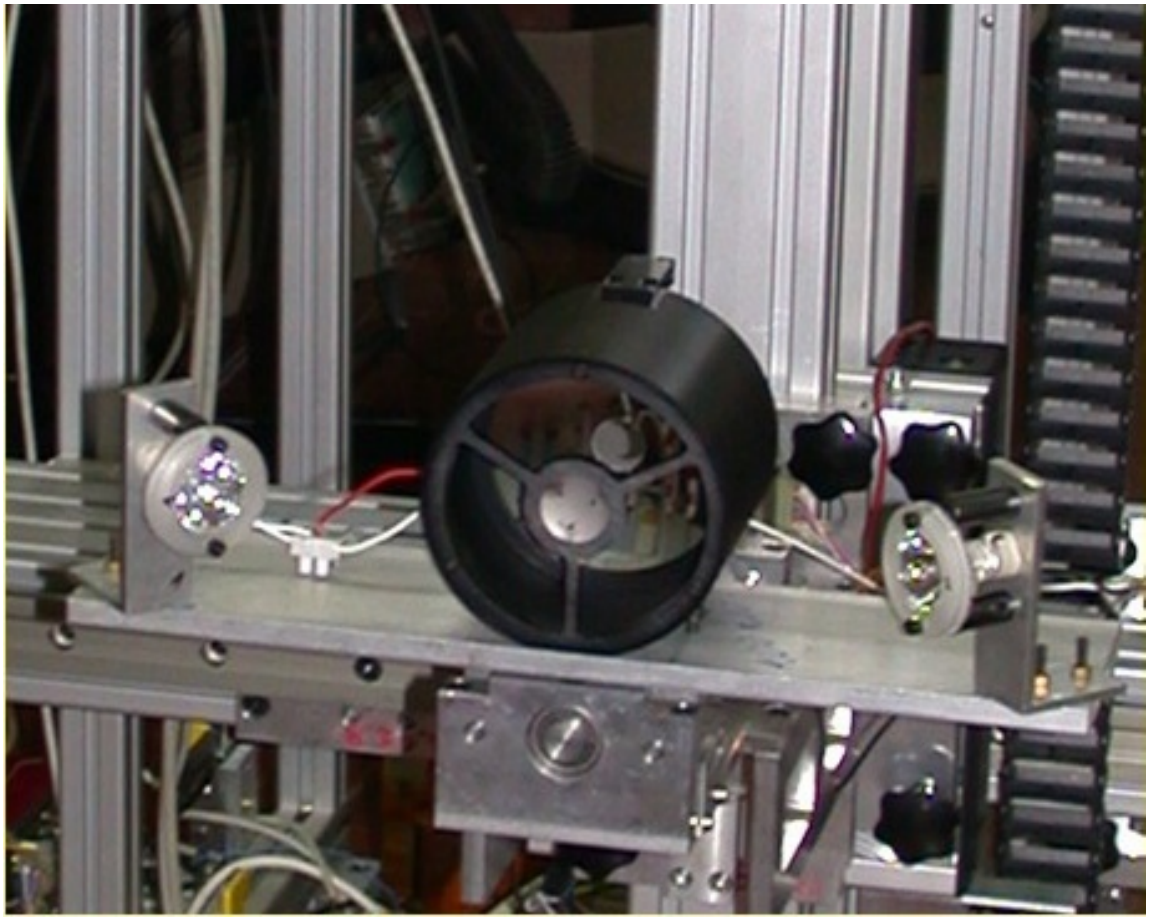


FIGURE 2.28: Detail of the unit containing the optical head and the light sources

of the instrument of 2.2.5. The *SRF* of seven coloured certified tiles was measured to certify instrumental accuracy, resulting in a mean absolute deviation of 1.3 % approximately (Carcagni et al., 2007).

The scanning instrument was applied on a great variety of paintings, with great success, in the most famous museums of the world. With a sampling step of 4 dots/mm, the instruments permits to unveil the tiniest details both in the painting layer and in the underdrawing arising from IR illumination. Figure 2.30 was taken on a painting of a Florentine anonymous from the XV$^{\text{th}}$ century, representing the portrait of S. Nicholas. The colour image was obtained from spectral data first in the *XYZ* reference frame and then converted in the *RGB* space. The very high spatial and colour

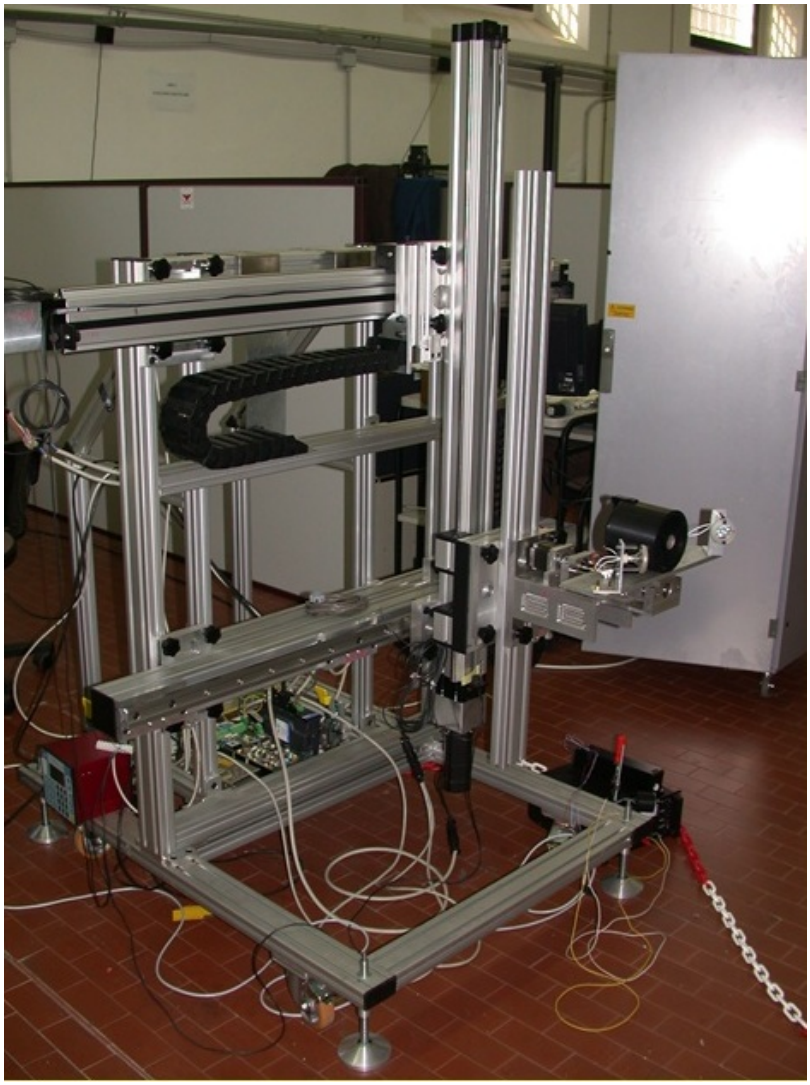


FIGURE 2.29: Global view of the scanner during a working session

resolution make visible the intricate net of *craquelures*. Spectral information were exploited to make hypotheses about the nature of the pigments used by the artist.

2.2.7 Snapshot spectral imaging

As it has been already made clear, spectral imaging is effected through the acquisition of a 3D-*data cube*, also known as *image cube* or *hypercube*, consisting of two spatial and one spectral dimensions. The vast majority of spectral imagers acquire

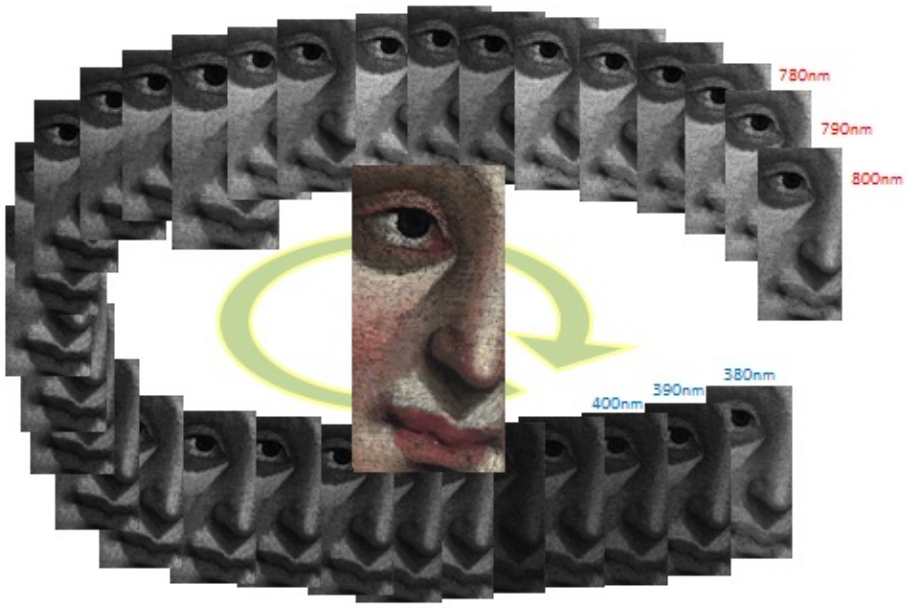


FIGURE 2.30: The 32 monochromatic bands of a multispectral acquisition and colour image after processing of the same bands

only two dimensions simultaneously, the remaining being acquired sequentially. A broad classification of such instruments identifies "staring" imagers and "pushbroom" imagers, the former performing simultaneous acquisition of both spatial dimensions, the latter performing simultaneous acquisition of one spatial and one spectral dimension. The entire hypercube is accomplished after temporal scanning.

Another really interesting class of spectral imagers is that of the "snapshot" imagers, namely instruments performing simultaneous and instantaneous acquisition of all the dimensions of the hypercube, without turning to scanning and with the clear advantage of being able to capture fast-changing objects with an unpredictable development along the time scale. Most of the snapshot imagers are of the CTIS type (*Computed-Tomography Imaging Spectrometer*) (Okamoto and Yamaguchi, 1991; Okamoto et al., 1993; Descour and Dereniak, 1995) and make use of a dispersing element (prism or grating), either rotating or still, that must provide multiple diffrac-

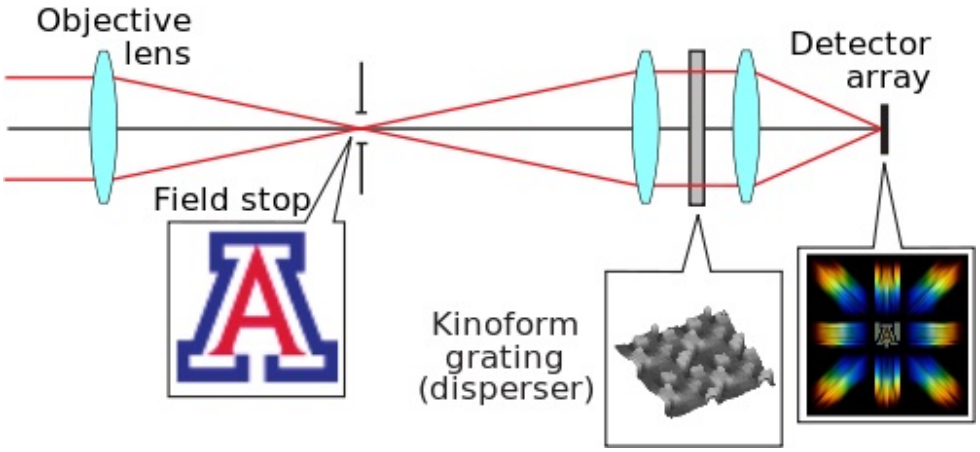


FIGURE 2.31: Layout of a typical CTIS imager with a still two-dimensional grating

tion orders of the imaged object on the sensor (Wilson et al., 2003; Schau, 2006) (Figure 2.31). An image from any primary optical system serves as input for the CTIS imager and a field stop is placed right at the image plane of the primary system. After that, light enters the CTIS imager where collimating optics prepare parallel rays for the transmissive disperser; imaging optics re-image the diffracted pattern on the camera sensor. From the captured intensity pattern, computed-tomography algorithms are used to reconstruct the scene (Figure 2.32) into a 3D-data cube. Figure 2.33(a) shows how the object pattern (b) is diffracted across the focal plane at the first-order, with zero-order (undiffracted object) being imaged unaltered in the center which greatly helps in focusing, a difficult task for many spectral imagers. The acquisition time is only constrained by the exposure time of the monochrome camera, usually on the order of milliseconds, and by the required signal-to-noise ratio. Besides the possibility to include moving parts, there are other disadvantages in CTIS imagers, like the presence of the field stop and the need of an intermediate image, which lead to more complicated and expensive optics. Two interesting instruments of this kind are noteworthy; the first was used in the field of ophthalmology for retinal hyperspectral imaging (Johnson et al., 2007), the other is a patented instrument, for

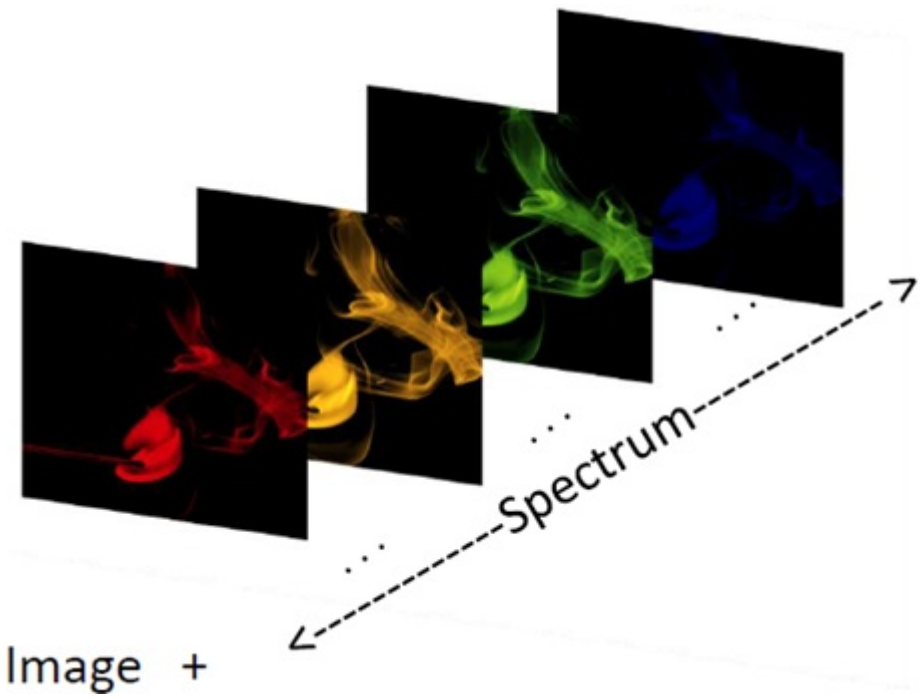


FIGURE 2.32: Demultiplexed monochromatic images (false colours) of a CTIS imager's acquisition

use in airborne and satellite applications, of particular significance as it is all-reflective thanks to the integration of a two-dimensional computer-generated hologram grating on the surface of a convex spherical substrate, by direct-write electron-beam lithography (Wilson et al., 2003; Johnson et al., 2005).

A commercial product currently available on the market is the VERDE™ snapshot hyperspectral camera from "Horiba"[5] (Figure 2.34), of wide application for field work and industrial quality control environments, involving color monitoring and control.

Another layout of snapshot hyperspectral camera does not include any moving parts and can come with different embodiments; the basic layout of Figure 2.35 consists of a lens, an optional bandpass filter, a chromatic disperser preferably one-dimensional (either a prism, as shown in the figure, or a diffraction grating) placed in the vicinity of a system aperture or any plane that is image-conjugate with the system aperture

[5] http://www.horiba.com/

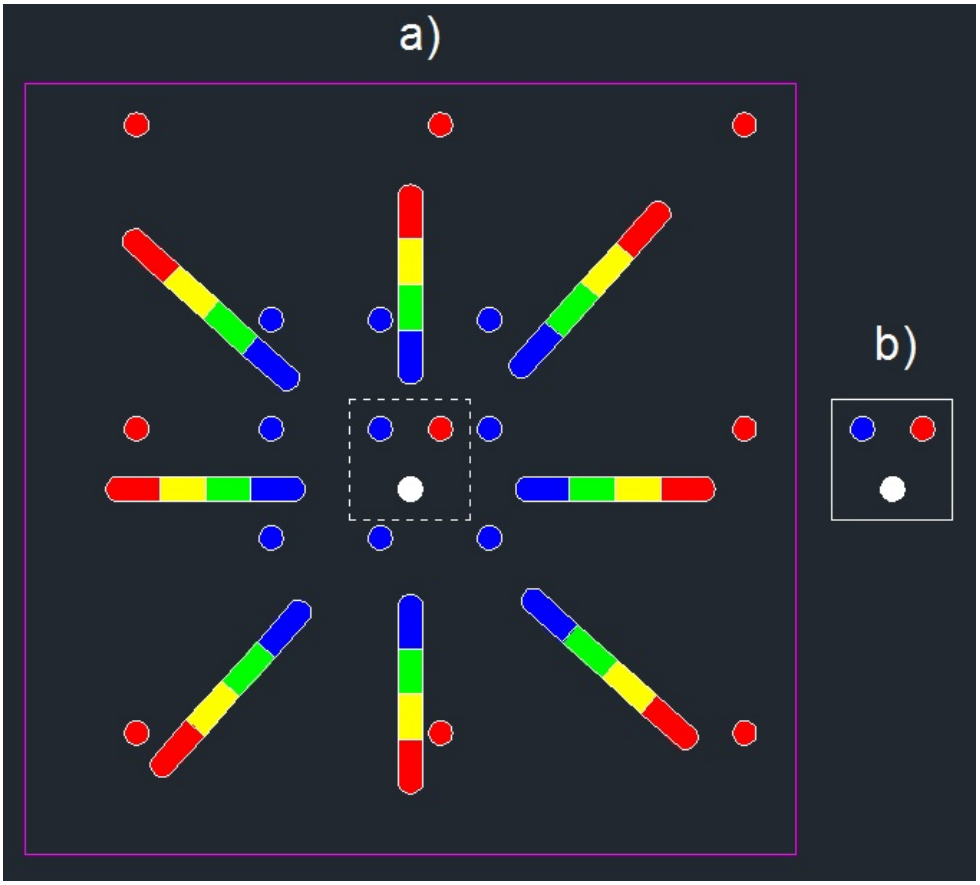


FIGURE 2.33: (a) First-order diffraction pattern on the image sensor of the polychromatic object in (b)

FIGURE 2.34: The VERDE™ snapshot hyperspectral camera from "Horiba"

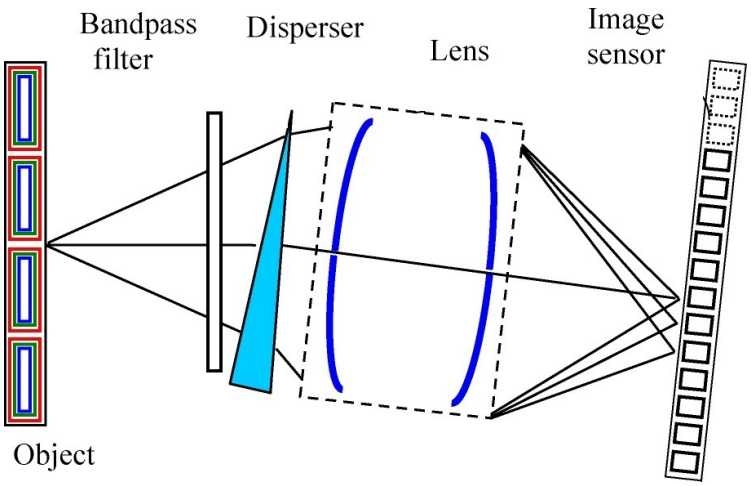


FIGURE 2.35: Basic layout of a snapshot hyperspectral camera with a prism disperser

of the lens (including the entrance and exit pupil planes), and an image sensor array located at a focused image plane. The particular placement of the disperser allows it to function in a plane light-wave mode, in order to minimize inherent aberrations while improving image quality. The design of the component optical elements is matched with the particular dispersing properties of the disperser, in order to create a diffraction pattern fitting a predetermined number of photosites (macropixel) on the matrix sensor. In general, the direction of the chromatic dispersion may not coincide with discrete raster directions.

Other layouts exist in which the disperser is coupled with a blurring optical element (weak spherical lens, cylindrical lens, aspherical lens with non-radial symmetrical surfaces, beam shaper, etc.), sometimes substituted with a physical intentional small defocusing, resulting in an engineered blurred point spread function. The blurred spot on the sensor can be made of rectangular shape and also apodized for equally weighted spectral multiplexing. Each pixel of a dispersed image receives light with different wavelengths coming from at least two different object points; the consequent spectral mixing at each photosite is termed *multiplexing* (Golub et al., 2011).

In the illustrative case in which the chromatic dispersion occurs along a raster di-

rection (row-wise), macropixels lay along rows and do not interfere between those belonging to adjacent rows. Figure 2.36 shows the formation of the spectrum image in the non limiting situation in which only three bands are present. The dispersing element separates the light with different spectral content in every macropixel by exactly one detector photosite, with respect to an adjacent band (Golub et al., 2009).

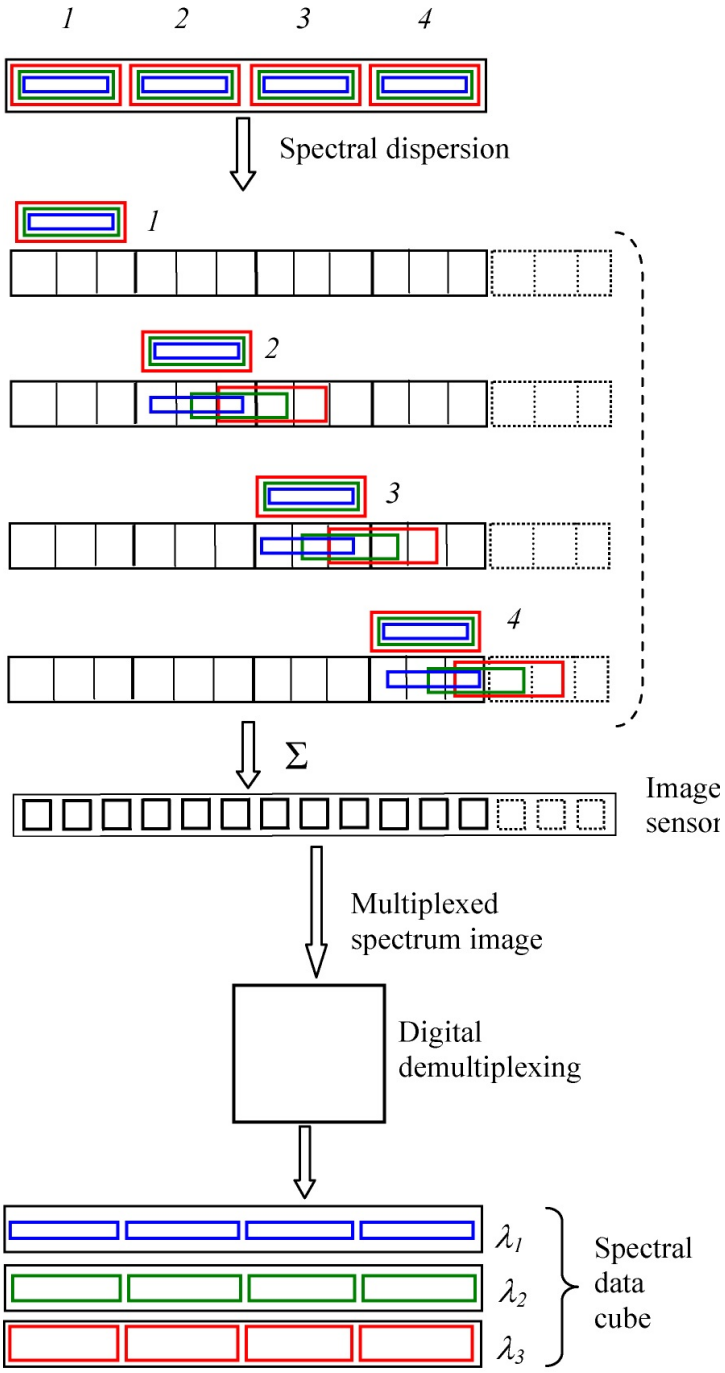


FIGURE 2.36: Formation of a row-wise multiplexed spectrum image, in the simple case of three spectral bands (shown as λ_1, λ_2 and λ_3)

3

The design of a compact hyperspectral camera

Trichromatic cameras, only in the case of a very accurate calibration, measure the color of objects illuminated by a particular light source and this quantity cannot be considered as a measurement of the spectral reflectance factor or of the radiance of a scene. Today, this limit is overcome by multispectral and hyperspectral scanners and big effort is dedicated to design new scanners and cameras. Multispectral and hyperspectral imaging are widely used in many fields such as agriculture, mineralogy, surveillance, forensics, cultural heritage conservation, environmental monitoring, biomedical, remote sensing, food quality and many others. This dissertation illustrates the design of a compact hyperspectral camera, that can be used in both portable and in-situ applications, with two main characteristics:

1. a *Linearly Variable Filter* (LVF) on the image plane of the objective lens, shifted in the direction orthogonal to the optical axis of the camera, realizes the image spectral scanning and allows the measure of the spectral radiance of the objects of a scene and/or the spectral reflectance factor

2. a relay lens transports the image produced by the objective lens and filtered by

the LVF to the image-matrix sensor

3. once the LVF is shifted out, the same image, used for measuring the spectral radiance of the objects of a scene, is used for the image framing and focusing

The camera can be either linked to a laptop, for acquiring and processing a fairly large amount of output data, or can be operated standalone by means of an embedded programmable processing unit, associated with a touchscreen display interface.
The requisite of compactness is made possible through the use of a suitable image spectral scanning based on a LVF, namely a wedge-shaped optical filter obtained by a thin-film interference coating with wavelength selective transmittance along one direction and uniform behavior in the perpendicular direction. The fundamental difference between a traditional scanner and the proposed camera lies in the fact that in the scanner the whole apparatus is moved along a path as long as the scene, whereas in this project the camera body is still and a LVF inside of it is the only moving part.
This movement is supplied by a stepper motor. This solution allows a compact and easily portable instrument. The final layout of the proposed camera is illustrated in the following sections, where emphasis is put on the calculations and physical considerations that supported the acceptance of this opto-mechanical scheme; in fact, it took three years, approximately, to devise the accepted layout, since the prior design underwent a number of successive refinements aimed at reducing the number of optical components and also to ensure a light weight to the final assembly. Firstly, the way the instrument works is explained and later the various elements composing the camera are detailed for.

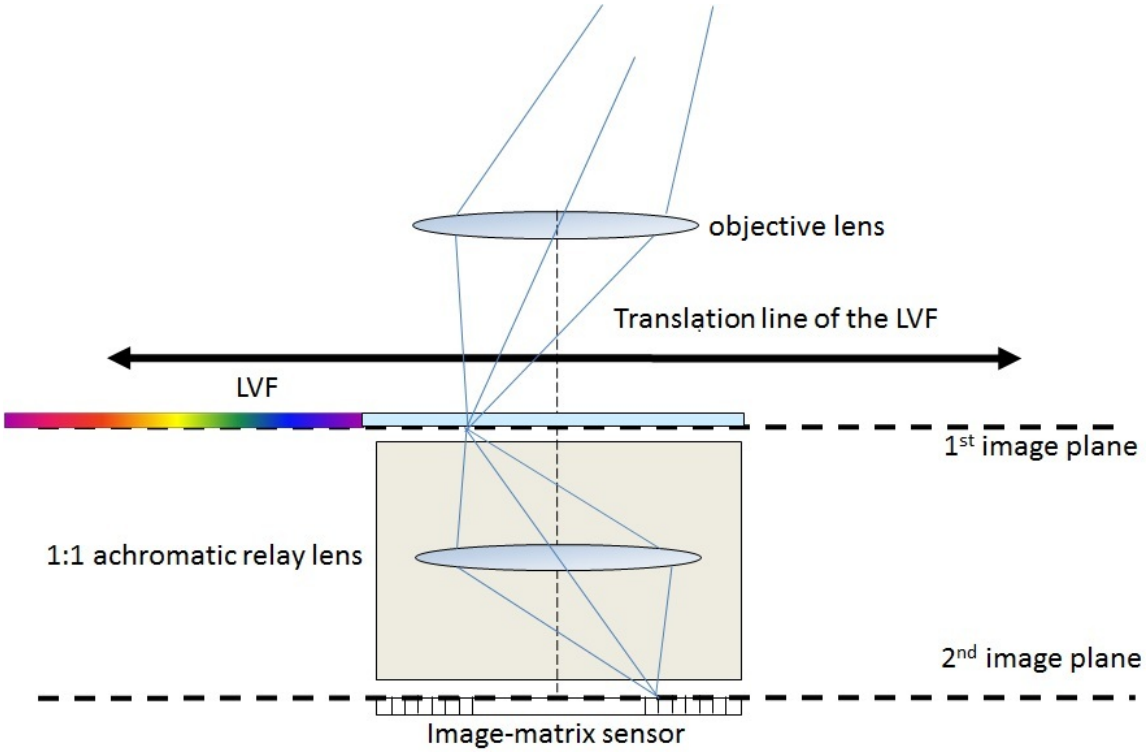


FIGURE 3.1: Sketch of the opto-mechanical layout of the instrumental proposal

3.1 The working principle

Figure 3.1 shows the schematic opto-mechanical layout of the hyperspectral camera, exactly as it was conceived for construction.

The objective lens of the camera focuses the image of a scene on a plane, indicated as 1st image plane, where a LVF is positioned. The filter is translated stepwise by a stepper motor in a direction orthogonal to the optical axis of the system, in order to achieve an image spectral filtering because the translation is made to occur just along the dispersion direction of the LVF. A second optical system, a relay lens, transports the filtered image onto the 2nd image plane, with a 1:1 magnification ratio. The relay lens and the LVF as a whole work as a *Wavelength Selective Optical System* (WSOS), in that any image strip is orthogonal to the translation direction is selected at a wavelength λ. Such a wavelength is typical of the corresponding vertical line on

the LVF. Further selection of different wavelength values is accomplished through shifting, thus implementing the measurement of the spectral radiance of the entirety of the image strips (Piegari et al., 2011; Della Patria et al., 2012, 2011).

3.2 The objective lens

The choice of the objective lens was driven by geometrical considerations, with the necessary tradeoff between size, weight and optical specifications. An activity of lens design, with a commercial dedicated software ZEMAX-EE™ [1], was carried out during the earlier stages of definition of the instrumental layout, in order to address the specific requirements of the system. In fact, in its earliest conception the camera was intended to frame an object located at infinity and so a lens was designed for the purpose, with a strong constraint on geometrical distortion, which had to be as low as possible for exact correspondence between points of the scene and photosites on the image-matrix sensor. For this purpose, the lens layout of Figure 3.2 shows the 5 elements system in three groups of lenses. There are two external achromatic doublets, with cemented elements, in a crown-out arrangement, namely with the crown glasses outside (Smith and Genesee Optics Software, 1992). The lens has a focal length f=60 mm and is designed to work with a ⅔ inch image-matrix sensor. It has an angular field of view (half-angle) of 5° and is well corrected, with respect to aberrations, throughout the visible range. It is worthwhile to notice in Figure 3.3 that the percent barrel distortion is very low (less than 0.01 %) throughout the visible, far below any reasonable tolerance for pixel-to-scene geometrical correspondence. The percentage distortion is defined by:

$$D = \frac{y_{\text{chief}} - y_{\text{ref}}}{y_{\text{ref}}} \cdot 100 \tag{3.1}$$

[1] `http://www.radiantzemax.com`

where y_{chief} and y_{ref} are the image surface radial coordinates (Y axis) of the chief ray and of a reference ray, respectively. The reference height for an undistorted ray in a rotationally symmetric system at paraxial focus[2] is given by:

$$y_{\text{ref}} = f \cdot \tan\theta \tag{3.2}$$

with f representing the focal length and ϑ is the angle the ray forms with the optical axis in object space.

Moreover, the lens was also designed to have relative aperture $f_{\#}$=5.2 to limit the incidence angle on the image plane for reasons that will be detailed in the following sections. The overall performance of the objective can be inferred from the spot diagram in Figure 3.4 where the the Airy disk[3] (Geary, 2002) is superimposed as a black circle and the coloured dithered patterns correspond to the wavelengths 486 nm (blue), 587 nm (green) and 656 nm (red). The left graph is for the on-axis field and the right graph for the upper off-axis field, both within a 40 µm square-frame. The r_{Airy}=3.7 µm and the RMS-spotradius[4] is 3 µm and 3.5 µm for the on-axis and off-axis fields, respectively. The modulation transfer function (MTF) (Williams and Becklund, 1989), shown in Figure 3.5 as the modulus of the optical transfer function[5]

[2] PARAXIAL FOCUS is the point of the optical axis where paraxial rays converge. Paraxial rays make small angles with the optical axis and paraxial optics are based upon the assumption of linearity described by the Snell's law. Hence, paraxial properties of optical systems are often considered the properties the system has in the absence of aberrations

[3] The AIRY DISK of an imaging system with a uniformly-illuminated circular aperture, in the absence of aberrations other than diffraction (*diffraction-limited* system), is the circle passing in the middle of the first dark ring of the Fraunhofer diffraction irradiance pattern and encompasses approximately the 84 % of the incoming energy. If Lord Rayleigh's criterion is applied, then two object points are said to be *just resolved* whenever the center of one Airy disk coincides with the first minimum of the Airy pattern of the other point. The Airy radius is expressed as: $r_{\text{Airy}} = 1.22\lambda f_{\#}$

[4] The RMS-SPOT RADIUS is the root-mean-square radial size; i.e. the distance between each ray and a reference point (either the *chief-ray* or the *centroid*) is squared, and averaged over all the rays, and then the square root is taken. The RMS-spot radius gives a rough idea of the spread of the rays, since it depends upon every ray

[5] The OPTICAL TRANSFER FUNCTION is the function that performs the transformation of the individual Fourier components of an object in the imaging process. It is a spatial frequency-dependent complex quantity whose modulus is the MTF and whose phase is the phase transfer function (PTF).

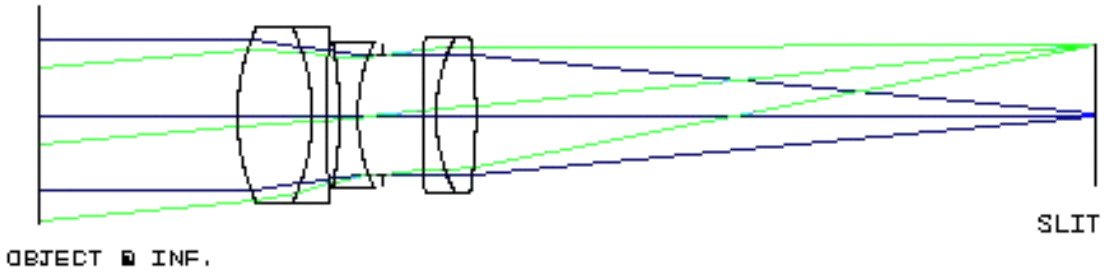


FIGURE 3.2: Two-dimensional layout of the objective for use at infinity

(OTF), exhibits a fairly slow drop-off and is quite close to the limit of diffraction all over the field of view. These considerations ensure that the lens system is well corrected for use in the visible range.

As the layout was changed with time until it took the form of Figure 3.1, the fore lens was intended to work at finite conjugates so a commercial solution with fixed focal length and with low distortion was the choice. The objective lens (model NT68-685) was eventually selected from the EDMUND OPTICS catalog[6] among the set of ultra-high resolution lenses, employing aspheric surfaces (Fischer et al., 2000). The possibility to get the lens recipe either from vendor, would have been an unquestionable advantage, because every single optical parameter could have been deeply analysed at convenience; unfortunately, as in most cases for fine commercial optical lens systems, the layout and recipe are some kind of black-box for the user, as everything is kept strictly confidential by the manufacturer for reasons related to the intellectual property. The only elements to drive the choice are listed in Table 3.1. The objective is endowed with opto-mechanical solutions that provide fine floating focus capability.

The former measures the loss in contrast over the spectrum in the imaging process, the latter represents the commensurate relative phase shift (Williams and Becklund, 1989; Hecht, 2001; Boreman, 2001)

[6] `http://www.edmundoptics.com`

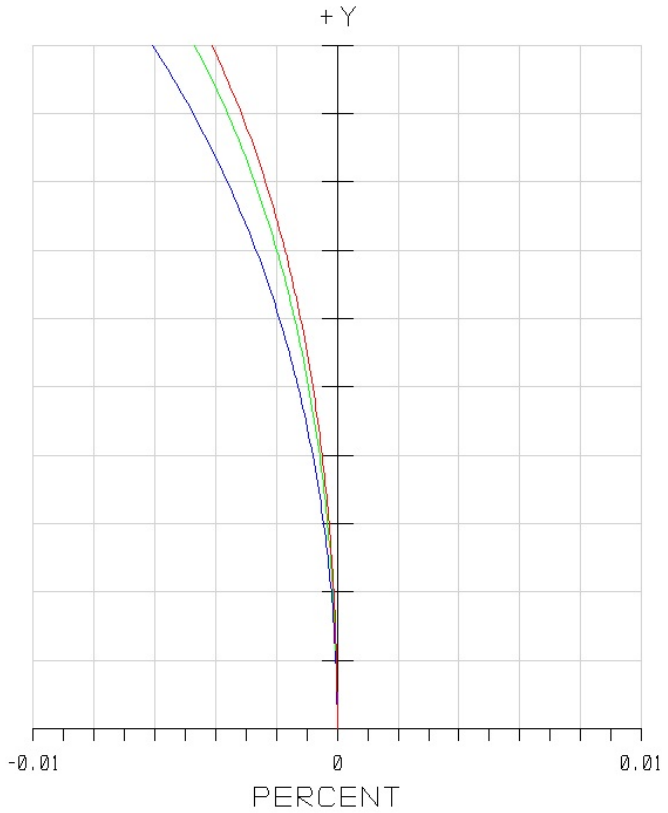


FIGURE 3.3: Percentage distortion of the lens in the visible, for the wavelengths 486 nm (blue line), 587 nm (green line), 656 nm (red line)

Table 3.1: Specifications of the objective lens NT68-685

Focal length (mm)	25
Relative aperture $f_{\#}$	1.8 - 16
Working distance (mm)	100 - ∞
Field of view @ min. working distance (mm)	26.7
Center field resolution (lp/mm)	200
Corner field resolution (lp/mm)	160
Distortion (%)	<-0.09
Filter thread (mm)	M25.5×0.5
Mount type	C-mount
Diameter (mm)	33
Length (mm)	45.5
Weight (g)	95

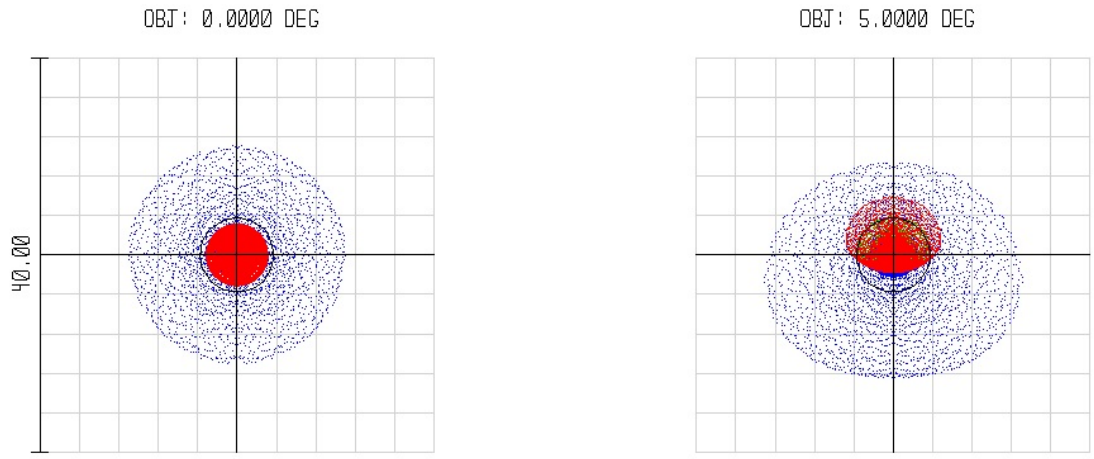


FIGURE 3.4: Spot diagram in the visible for on-axis (left) and extreme off-axis (right) fields at 486 nm (blue), 587 nm (green), 656 nm (red); the small black centered circle is the Airy disk

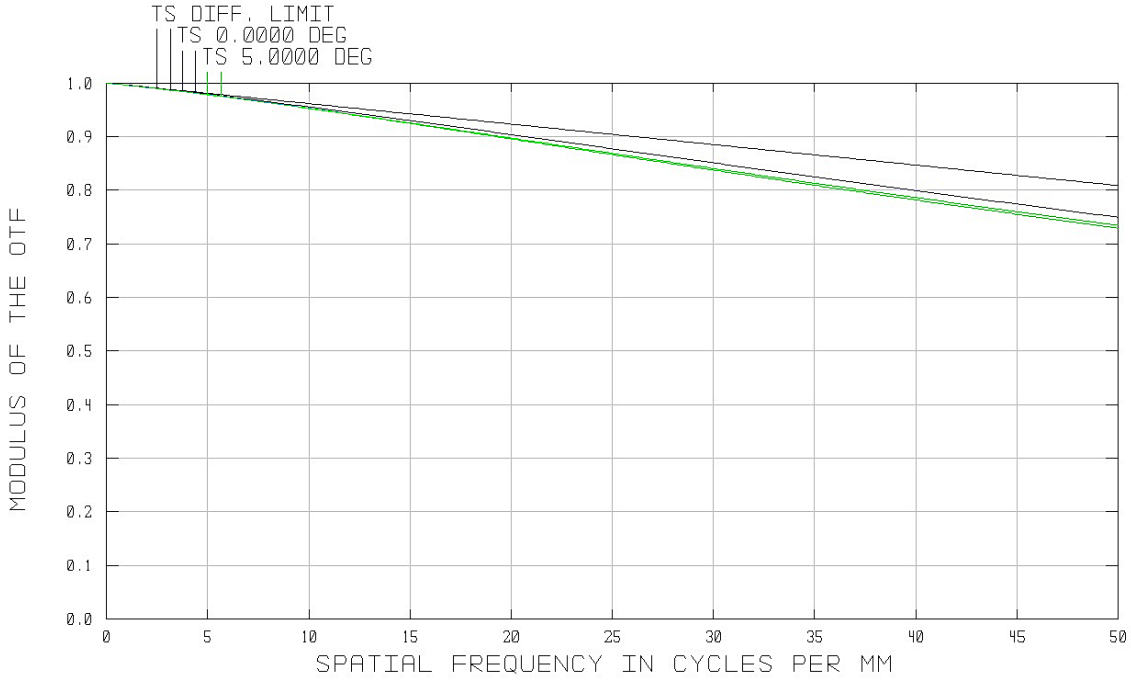


FIGURE 3.5: Modulation transfer function of the lens for on-axis (blue line) and extreme off-axis (green line) fields; the black line is the diffraction limited performance

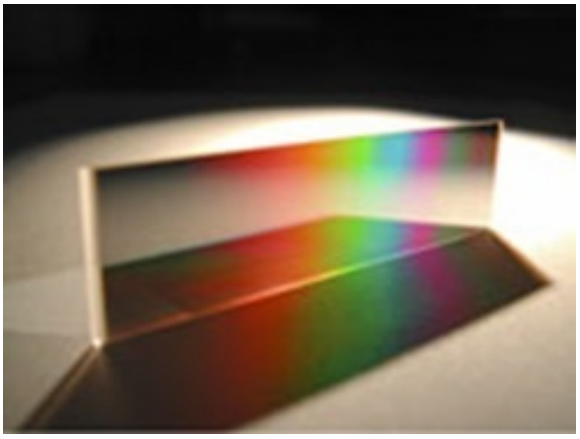
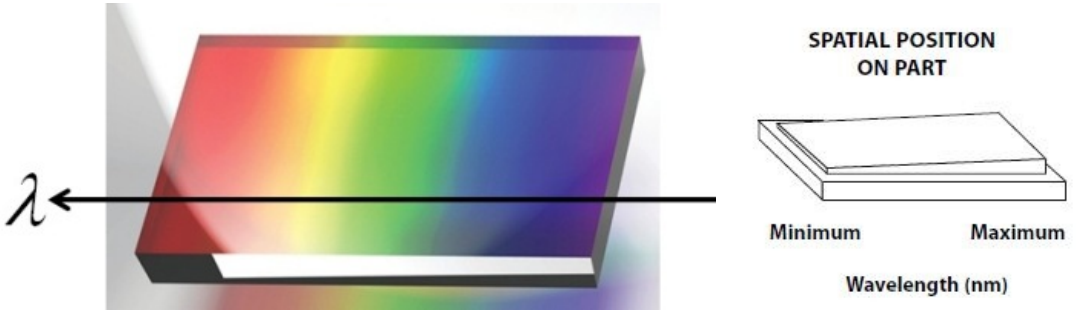


FIGURE 3.6: Picture of a LVF deposited on a rectangular glass substrate



(a) (b)

FIGURE 3.7: (a) The LVF with the dispersion line superimposed (b) Wedge-shaped profile of the coating thickness of the LVF and positions of lower and higher transmitted wavelengths

3.3 The Linearly variable filter

What in the previous section was called a *Linearly Variable Filter* (Figure 3.6) is a narrow-band transmission filter whose peak wavelength is displaced over the surface of the filter itself, along one direction (Morelli, 2006). Such a directional behavior is made possible by a suitably engineered interference coating deposition on a glass substrate. The coating is intentionally wedged in the direction of wavelength dispersion (Figure 3.7).

Since the transmitted (peak) center wavelength (CWL) depends on the coating

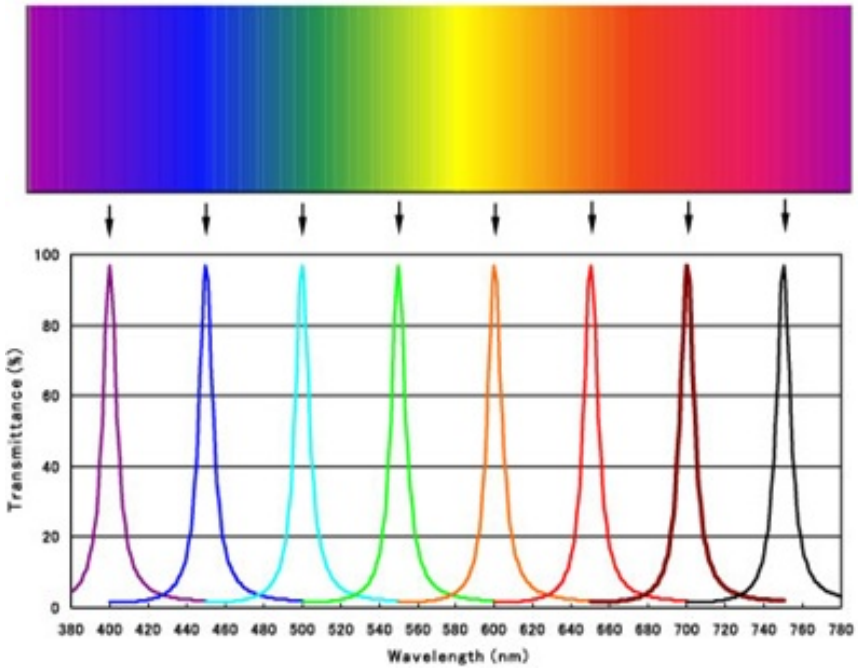


FIGURE 3.8: Displacement of the transmitted central wavelength along the direction of dispersion in a LVF

thickness, the CWL that traverses the filter can be made to vary in a linear fashion along the wedge, following either a linear (or logarithmic) monotonic function. The particular wavelength of interest may be selected by adjusting the filter to the corresponding appropriate position (Figure 3.8); as a consequence, a raster scanning of the filter in front of a source implements a scan of the whole wavelength working range of the filter itself, with uniform behavior along the other direction.

The spatial profile of coating thickness influences the dispersion gradient, i.e. the velocity at which the CWL changes over a unit length of the filter along the dispersion line.

One way to make a transmission narrow-band LVF is to deposit all-dielectric FABRY-PEROT filters, provided that the operation wavelength interval, i.e. the ratio between the minimum and maximum wavelengths of the rejection band, is not too large. All-dielectric FABRY-PEROT filters typically exhibit a ratio lower than 2:1. In order

to increase such ratio, it is mandatory to deposit supplementary blocking filters, which might increase the total amount of layers up to a hundred or even more.

The blocking filters on the one hand enlarge the rejection band, on the other hand make the deposition process unpractical due to the high number of layers. Moreover, no absorption filters can be employed for the blocking action if the complete filter performance must be varied over the surface (Piegari and Bulir, 2006). In this case the rejection range should exceed the traditional 2 :1 ratio for uniform filters, because as the CWL shifts the out-of-band rejection needs not be altered. A ratio of 2.5 :1 is considered to be very high for a transmission narrow-band LVF, therefore another architecture that permits to achieve this ratio, with the added value of a sensible cut in the number of deposited layers, employs metal-dielectric coatings (Berning and Turner, 1957; Macleod, 2001) to fabricate the so-called INDUCED-TRANSMISSION filters (Appendix C).

With this alternative solution, narrow-band high-transmission filters can be obtained through suppression of the reflectance of a metal layer, achieved with a matching of the admittance of the metal layer to the substrate and air side by adding a number of alternate dielectric layers on both sides. In the 400 nm to 1000 nm wavelength range the best choice is silver. The overall filter performance is quite unaffected by the substitution of dielectric materials, but the parameters of the metal are the most critical.

Typically a symmetric layout is used, which consists of the metal layer surrounded by equal quarter-wave stacks of dielectric layers, with the only exception for the layers in contact with the metal (Macleod, 2001). In this way a transmission peak as high as 70 % or more can be achieved at the particular reference wavelength used for calculations, although the limit to better performances is posed by absorption in the thickness of the metal. Despite the all-dielectric FABRY-PEROT filters, which have higher peak transmittance and narrower bandwidths, INDUCED-TRANSMISSION filters

exhibit broader rejection bands. Optimization processes can reduce the bandwidth at the expense of the aforementioned symmetry.

A drawback with metal-dielectric filters is the drop in performance as the CWL is advanced in its operating domain, due to the high dispersion of the complex refractive index of the metal layer; every single CWL would entail its own match with the surrounding media and the optimization process needs be carried out only for a selected reference wavelength (Piegari and Janicki, 2003).

Once the number and sequence of layers has been chosen for a reference CWL, the only degrees of freedom are the thickness of the layers over the surface. Furthermore, the ratio of each thickness to each CWL must be invariant with the displacement of the CWL. Because of the thickness variation, it must be expected that the overall filter performance decreases at the extreme wavelengths with respect to the reference wavelength.

The possibility to operate at higher peak transmission is satisfied through a broadening of the bandwidth. A good control of the thicknesses, through optical monitoring and reverse engineering, is of great help during the deposition process for prompt online correction (Piegari et al., 2008). While the thickness of the metal is kept uniform, the wavelength shifts along the dispersion line because of the change of thickness t of each dielectric layer of the stack, according to the following relationship:

$$nt = q\frac{\lambda_0}{4} \tag{3.3}$$

Here n is the refractive index, λ_0 is the reference wavelength, q is the optical thickness (which is its geometric thickness multiplied by the refractive index) expressed in terms of quarter-waves[7]. For the CWL to shift linearly, both high-index and low-index

[7] Quarter-wave films are of common use in coating technology when it is necessary to maximize transmitted light (anti-reflection coatings) because, if incident light strikes the quarter-wave layer at normal incidence, then the reflected waves will be completely out-of-phase and will interfere destructively. If more layers are added, each designed to match a particular wavelength, further reduction in reflection is accomplished

dielectric layers have to vary their thickness almost linearly, provided the dispersion of their refractive index in the operating range is negligible. The extreme thickness values, i.e. the minimum t_m and the maximum t_M, depend on the values of both the CWL and refractive index at the extreme spatial positions, as follows:

$$\frac{t_M}{t_m} = \frac{\lambda_M}{\lambda m} \frac{n_m}{n_M} \tag{3.4}$$

This relationship is insensitive to q but changes with the particular materials used, because of the presence of n at different wavelengths.

In the application here presented an INDUCED-TRANSMISSION LVF was fabricated by radio frequency sputtering in the facilities of the Optical Coatings Laboratory at ENEA-CASACCIA in Rome. The graded multilayer coating was deposited over a 1 inch diameter circular substrate, with a rectangular shape (Figure 3.9). The short side of the filter was sized to match the length of the long side of a ⅔ inch image-matrix sensor; the working wavelength is 430 nm to 950 nm because of the presence of the thin silver layer (50 nm) that absorbs from 430 nm below. The full-length is 16 mm (along the direction of dispersion) so the gradient is given by the operating wavelength range divided by the long filter size, i.e. $32\,\mathrm{nm\,mm^{-1}}$ approximately. The dielectric layers employed were silicon oxide (SiO_2) as low-index material and tantalum oxide (Ta_2O_5) as high-index material. The geometry of the 21 layers symmetrical stack, deposited on a quartz substrate, is shown in Figure 3.10. A metal layer M is sandwiched between two equal blocks (10 layers each) of alternating quarter-wave H (tantalum oxide) and L (silicon oxide). Very thin nitrogen-enriched buffer layers of silicon oxide (L') were grown at the interface with silver, to prevent oxidations or other problems.

By virtue of 3.4 and using the measured refractive index values of H and L dielectric layers reported in Table 3.2 at the two extreme (430 nm and 940 nm) and at the middle (685 nm) wavelengths of the operating range, then the thickness ratios for SiO_2 and Ta_2O_5 are about 2.2 and 2.3 , respectively. By knowing the sequence of layers in

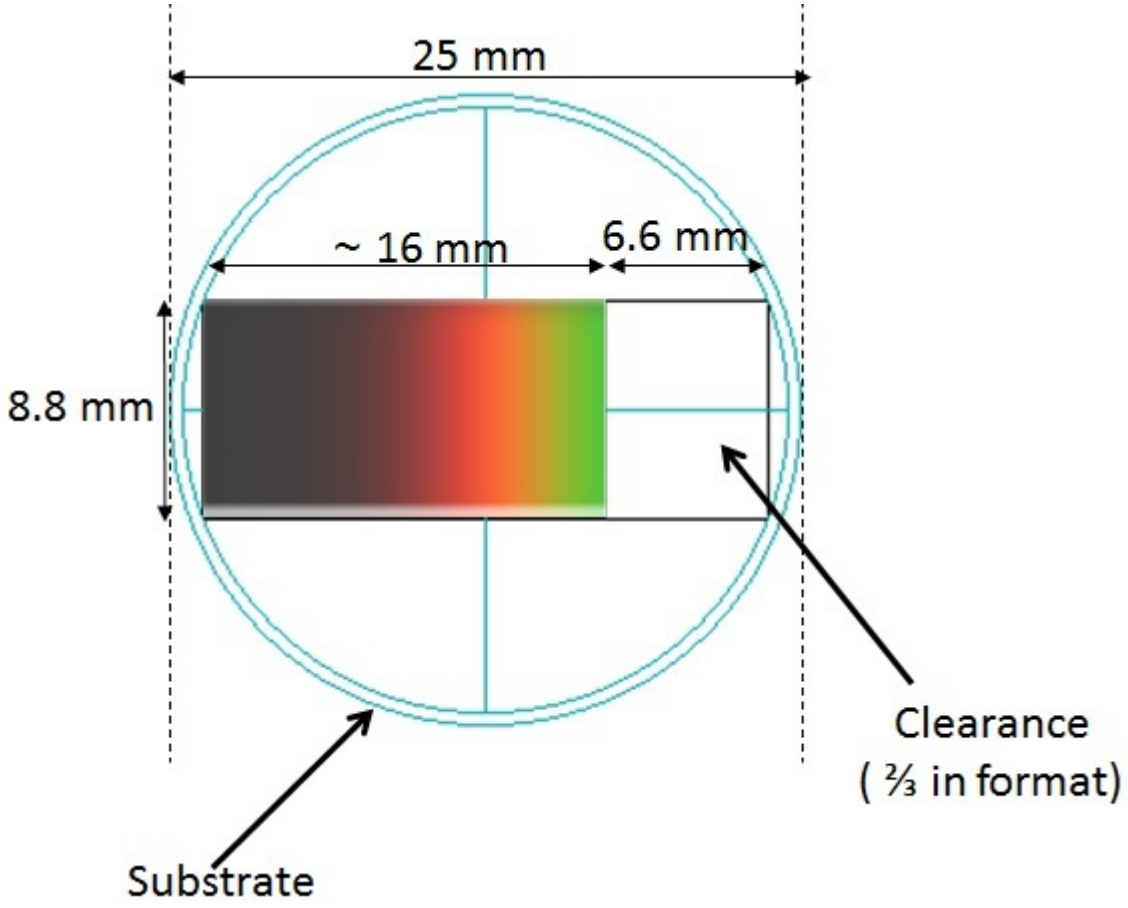


FIGURE 3.9: Size of LVF, clearance and substrate

FIGURE 3.10: Schematic architecture of the stack of layers forming the INDUCED-TRANSMISSION LVF

the stack and using the 3.3 (with q=1), the filter thickness T at any wavelength can be calculated by the following relationship:

$$T(\lambda) = t_{\mathrm{M}} + 2.5\lambda\left(\frac{1}{n_{\mathrm{L}}(\lambda)} + \frac{1}{n_{\mathrm{H}}(\lambda)}\right) \tag{3.5}$$

Table 3.2: Measured refractive index values of SiO_2 (L-index) and Ta_2O_5 (H-index) and total thickness of the filter at different wavelengths

λ (nm)	n_{L}	n_{H}	T (nm)
430	1.46672	2.17409	1277
685	1.45552	2.08608	2048
940	1.45108	2.06733	2806

Table 3.2 also reports the values of T, calculated in correspondence to the wavelengths listed in the leftmost column.

The measured transmission curves (Krasilnikova et al., 2005) of the filter, at normal incidence, are plotted in Figure 3.11; good average transmittance peaks of about 50 % and bandwidth values (FWHM) of 16 nm. Filter performance is not significantly affected for changes of the incidence angle up to about 10°, as shown in Figure 3.12. A masking system, inside the deposition chamber, with accurately controlled mask shape and movement allows the fabrication of wedged layers.

Essentially two different masking methods are commonly used: the first employs apertures of different shapes on the mask, depending on the particular material to be deposited; the second uses a blade that works as an obscuration at determined positions on the samples for programmable exposure times (Piegari et al., 2008).

3.4 The relay lens

The relay lens (model NT45-759), as with the objective lens, was chosen from the Edmund Optics catalog and its specifications are listed in Table 3.3. It was intended

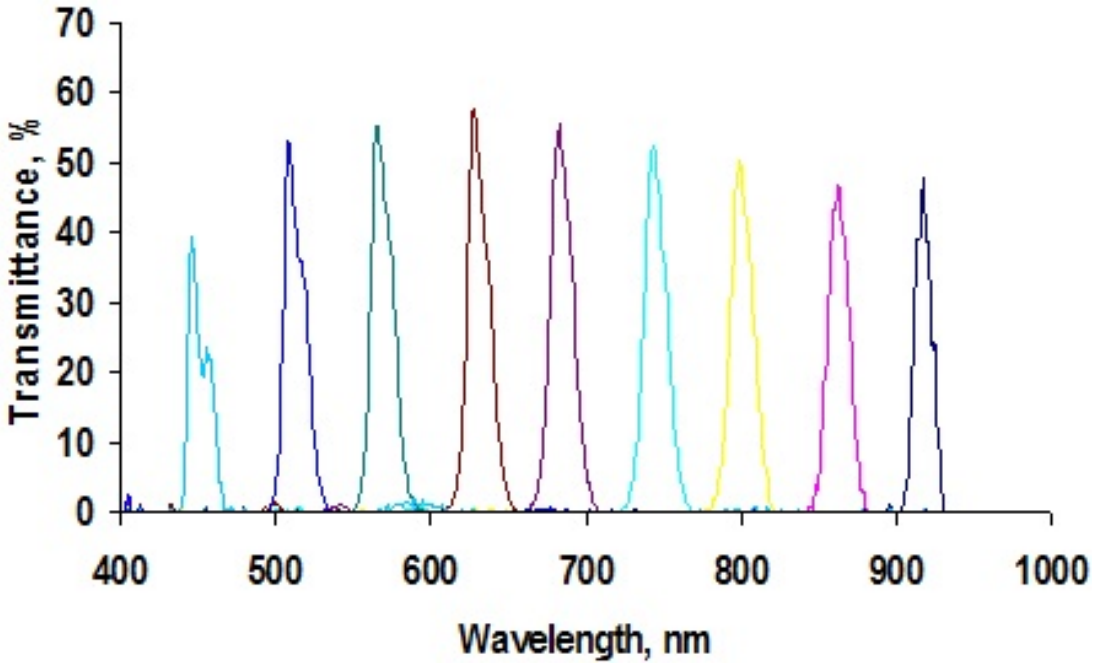


FIGURE 3.11: Measured transmittance of the 21 layers stack along the dispersion direction

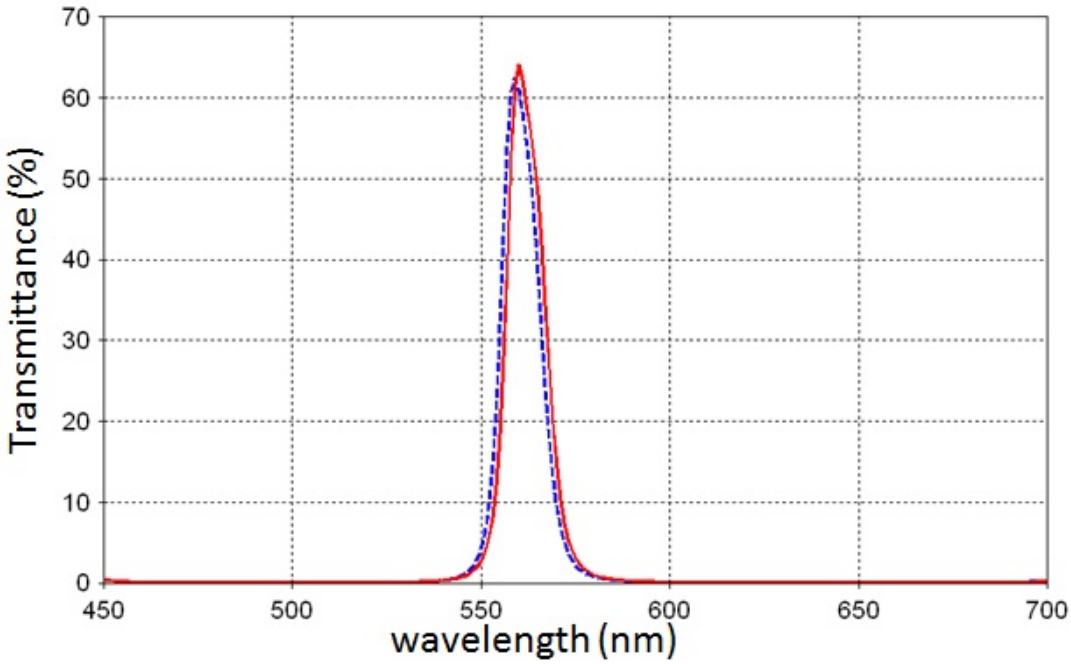


FIGURE 3.12: Transmittance band of the INDUCED-TRANSMISSION LVF at 560 nm for normal (solid) and 10° (dashed) incidence

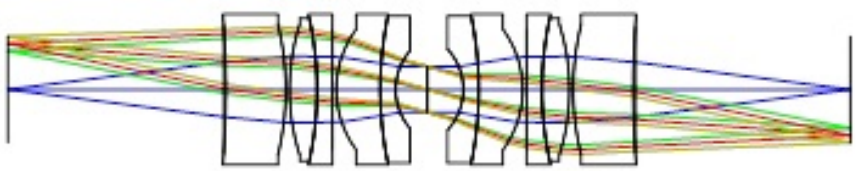


FIGURE 3.13: Two-dimensional layout of the relay lens system NT45-759

for distortion-free imaging with 1:1 magnification ratio. In the cross-sectional layout Figure 3.13 it is clearly seen that there are 10 lenses symmetrically arranged about the aperture stop. Each specular group consists of two achromatic doublets and a thick external meniscus. Symmetry and 1:1 magnification together ensure the absence of coma, distortion (Figure 3.14) and lateral color (Smith, 1997).

The good performance of the relay lens system can be inferred from the plot of the

Table 3.3: Specifications of the relay lens NT45-759

Diameter (mm)	21
Aperture $f_{\#}$	4
Effective focal length (mm)	30
Length (mm)	48.9
Image distance (mm)	19.2
Object distance (mm)	19.2
Obj.-to-img. length (mm)	87.3
Number of lenses	10

RMS-spot radius versus the field position (half the diagonal of a ⅔ inch image-matrix sensor), where it is easy to see that the polychromatic behaviour is below the diffraction limit almost all through the diagonal length (Figure 3.15). Moreover, the plot of the polychromatic modulation transfer function of the relay lens, shown at discrete field values in the right box, is close to diffraction limit (black line) with little loss of modulation until 50 lp/mm (Figure 3.16). The role of the relay lens is simply to translate unscaled the image produced by the objective lens, taken at the 1st image plane, straight to the sensor surface (2nd image plane).

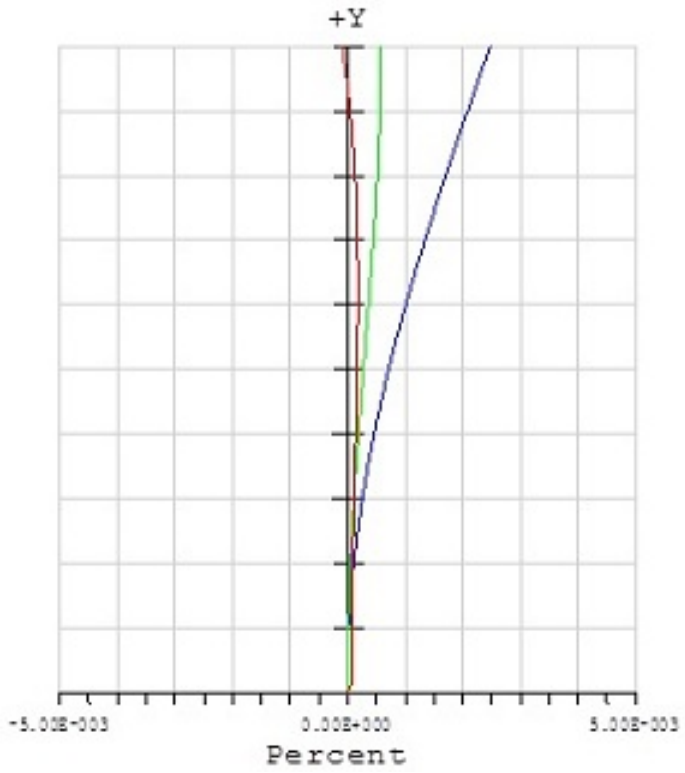


FIGURE 3.14: Percentage distortion of the relay lens in the visible, for the wavelengths 486 nm (blue line), 587 nm (green line), 656 nm (red line)

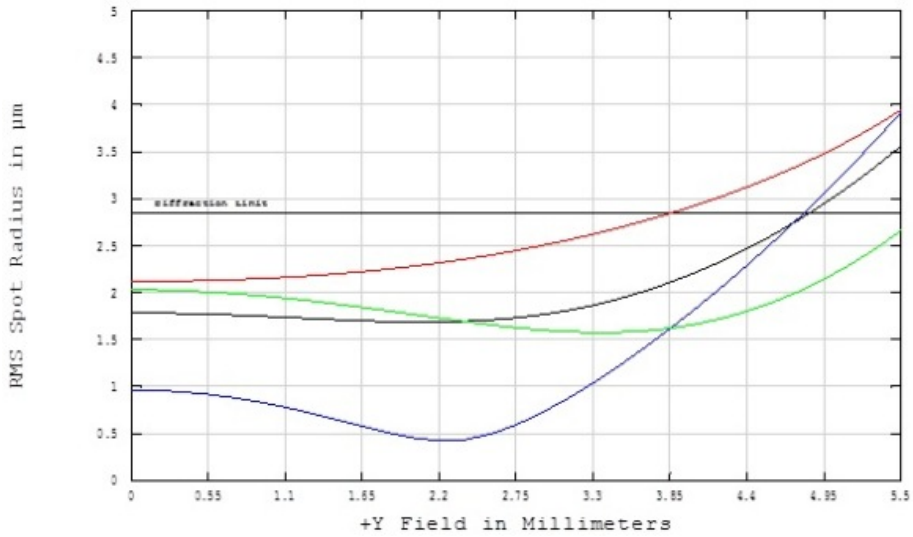


FIGURE 3.15: The RMS-spot radius versus the field for three wavelengths (black line is for polychromatic)

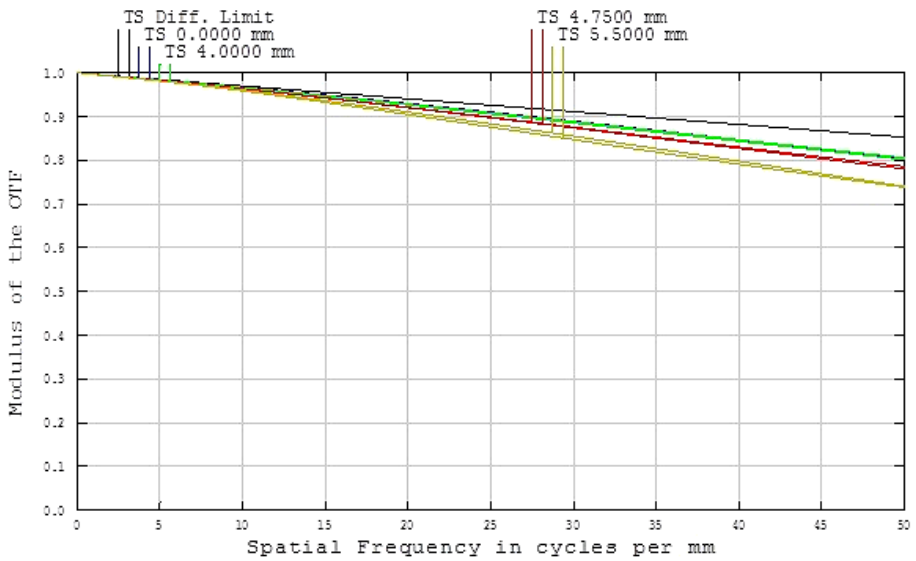


FIGURE 3.16: Modulation transfer function for axial and extreme off-axis fields (black line is the diffraction limited performance)

3.5 The image-matrix sensor KAI-02050™ Eastman Kodak

The sensor of choice was a commercial monochrome CCD (*Charge-Coupled Device*), model KAI-02050 from EASTMAN KODAK[8]. It is a 2 Mpxl matrix in a standard ⅔ inch format ($6.6 \times 8.8\,\mathrm{mm}^2$). It has a 1240×1640 pxl matrix of 5.5 µm size square photosites. It is based on an *Interline Transfer* architecture, which exploits the electronic shutters in place of the more cumbersome electromechanical shutters. The active area is fully covered with a telecentric microlens-array (Figure 3.17) in order to maximize the light gathering ability, and the resulting stack is protected with a clear cover glass with anti-reflection coating on both sides to minimize the onset of stray light. Figure 3.18 is the plot of the sensitivity of the sensor as a function of wavelength, where it is easily inferred that it is a decreasing function in the operating spectral range of the LVF; furthermore, Figure 3.19 is a graph of the percent relative quantum efficiency of the sensor as a function of the incidence angle, both in the

[8] `http://www.kodak.com/go/imagers`

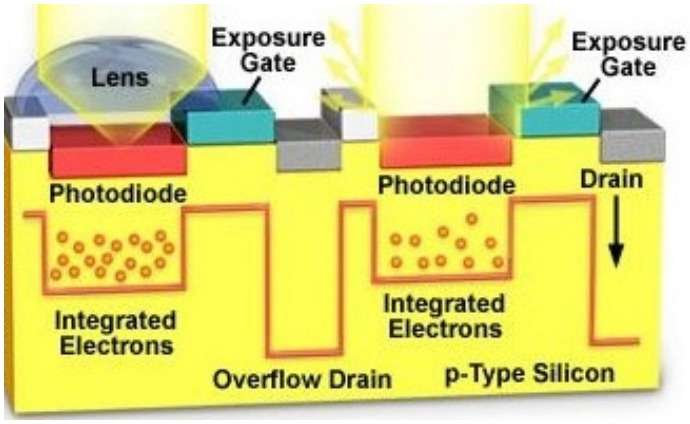


FIGURE 3.17: Schematic illustration of the working principle of a microlens over a matrix sensor

horizontal and vertical directions, which demonstrates that, for the range of incidence angles allowed by the LVF ($\pm 10°$), the sensor is insensitive to the particular direction of incidence of the rays.

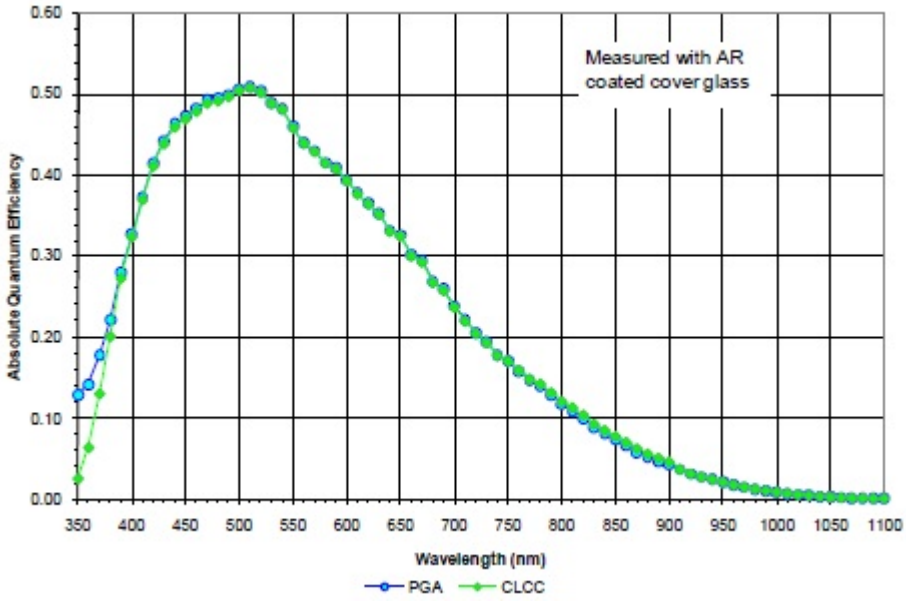


FIGURE 3.18: Sensitivity (absolute quantum efficiency) vs wavelength of the KAI-02050™ sensor

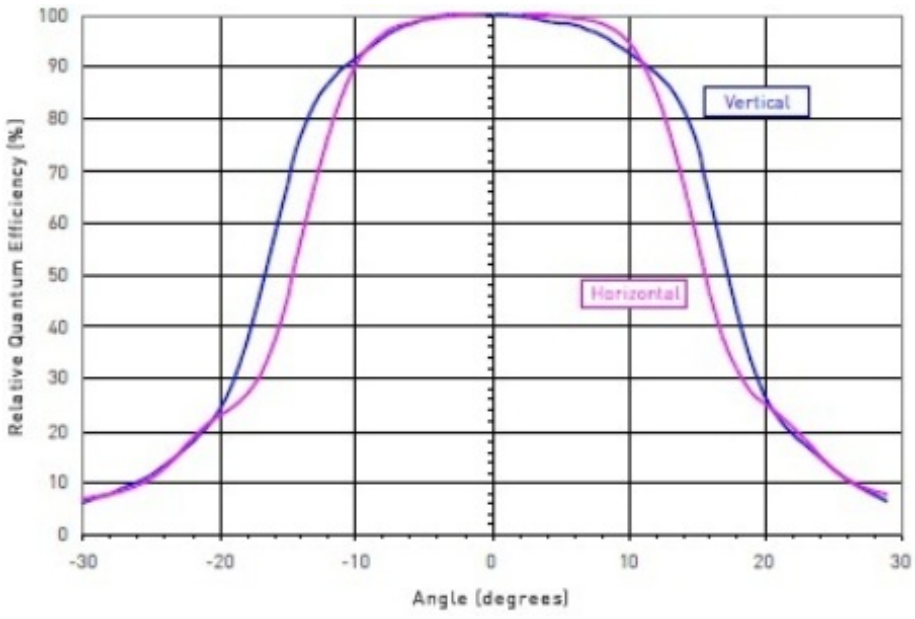


FIGURE 3.19: Angular quantum efficiency of the KAI-02050™ sensor for vertical and horizontal incidence

3.6 Incidence on the LVF

As shown in Figure 3.20 experimental evidences, based both on energetic and spectral considerations, demonstrated that the incidence angle should be constrained to within ±10°, approximately. The on-axis converging cone is defined by its incidence angle α through the objective's relative aperture $f_{\#}$:

$$\alpha = \arctan\left(\frac{1}{2f_{\#}}\right) \tag{3.6}$$

Of course, this constraint is tightened for extreme off-axis incidence, since the chief-ray itself has a nonzero incidence angle β:

$$\beta = \arctan\left(\frac{Y}{f}\right) \tag{3.7}$$

with f the focal length of the objective lens and Y the radial position, with respect to the optical axis, on the LVF plane.

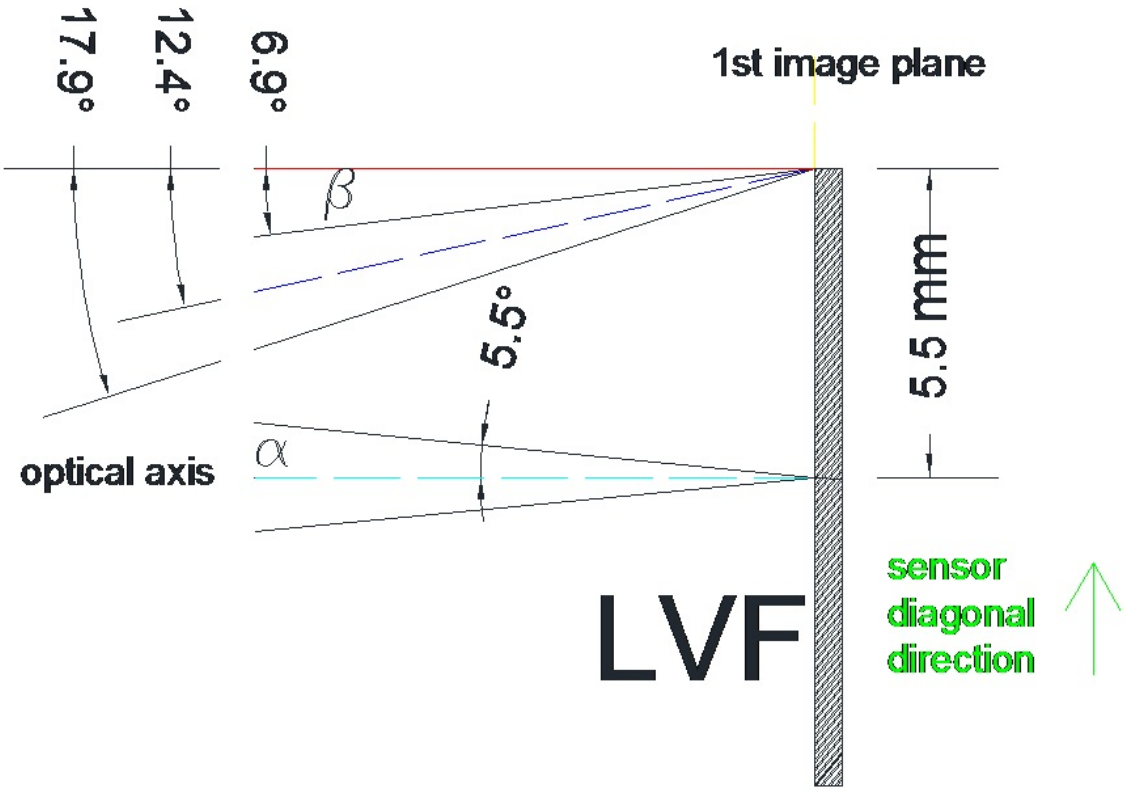


FIGURE 3.20: Geometry of the converging beams for $f_{\#}$=5.2 on the LVF in the 1st focal plane; the points of convergence are the center point (on-axis cone) and the corner (off-axis cone) of the sensor, which are 5.5 mm apart (half-diagonal for a ⅔ inch camera format)

The worst case is Y=5.5 mm, i.e. when Y equals the half-diagonal length of the sensor (Figure 3.20); in fact, although the on-axis converging cone of rays is within specification up to $f_{\#}$=2.8, the limit imposed by the focal length of the current objective prevents the extreme off-axis cone from meeting the incidence constraint (β=12.4°).

Much worse are matters with the most oblique ray, which impinges with an angle greater than β. Supposing to work with $f_{\#}$=5.2, the highest incidence angle in the extreme off-axis cone is 17.9°.

The choice of an objective with f=25 mm was a result of a tradeoff between the need to ensure good spectral performance and to keep the system compact; should testing

demonstrate, as expected, that system's performance are not altered significantly, this would allow some clearance to employ even shorter focal length objectives.

3.7 The *spatial* resolution of the camera

The spatial resolution of the camera is the minimum distance between two object points that the camera is able to discriminate.

There are different factors that contribute to the spatial resolution, both optical and electronic.

The optical contributions come from the objective and relay lenses, to a large extent, and from the LVF, to a small extent.

It is well known that the imaging properties of a lens depend on the point-spread function (Hecht, 2001; Boreman, 2001) (PSF), which represents the two-dimensional image irradiance distribution (diffraction pattern from a finite aperture) in response to a incoherent truly point-like source. The Airy disk is the circular spot that encompasses the 84 % energy of the PSF and its radius r_{Airy} measures the distance from the center of the pattern until the 1$^{\mathrm{st}}$ dark ring (Figure 3.21). The Airy disk diameter d_{Airy}=$2r_{\mathrm{Airy}}$ is:

$$d_{\mathrm{Airy}} = 2.44\lambda f_{\#} \tag{3.8}$$

where λ is the wavelength of the rays and $f_{\#}$ is the lens relative aperture.

According to Rayleigh's criterion, two incoherent point-like sources are said to be *just resolved* when the center of one Airy disk falls on the first minimum of the other point, i.e. when the two peaks are one r_{Airy} far apart.

A more appropriate criterion was proposed by C. Sparrow according to which two point-like imaged sources are *just resolved* when the central dip between the irradiance patterns disappears.

Unlike the Rayleigh rule, which assumes incoherence, the Sparrow's condition can be

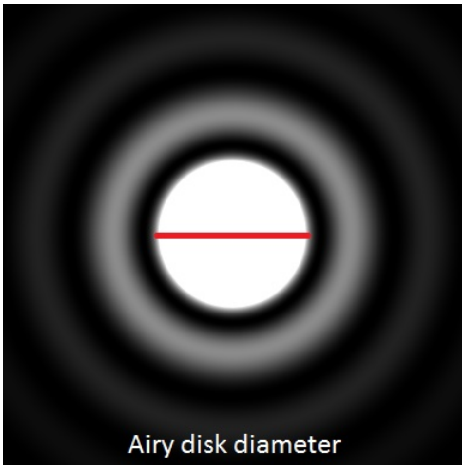


FIGURE 3.21: The irradiance pattern from a circular aperture of a very distant object (the Airy disk diameter is the red segment)

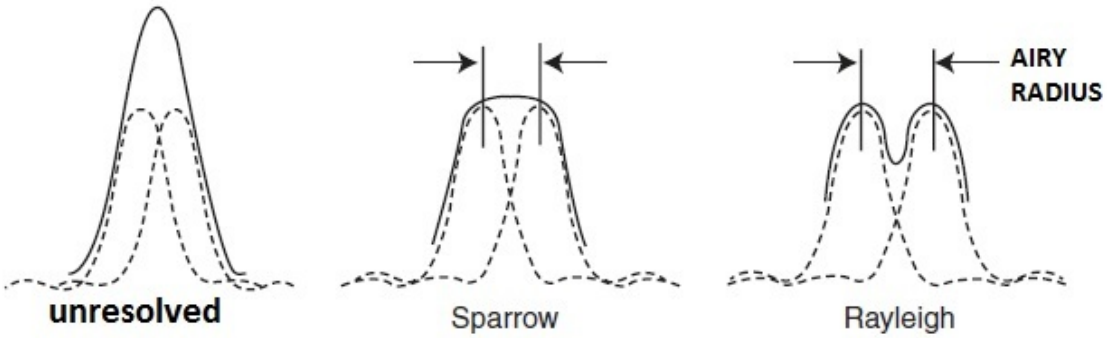


FIGURE 3.22: Unresolved incoherent monochromatic point-like irradiance patterns (left), Sparrow's criterion (center) and Rayleigh criterion (right)

readily generalized to coherent sources (Hecht, 2001).

In Figure 3.22 it is clearly visible that if the points are too close they are unresolved; an increase in the separation at $r_{\text{Airy}}/1.22$ creates a flat envelope (Sparrow's criterion), i.e. no minimum between the maxima, and a further increase at r_{Airy} introduces a saddle point (Rayleigh's criterion) (Smith, 2000; Hecht, 2001). An incorrect placement of the LVF on the 1$^{\text{st}}$ image plane, due to a tilt angle of the filter normal with the optical axis, would eventually deteriorate image quality through the onset of an uncontrolled amount of blurring, which adds to the spherical aberration arising from

the insertion of a glass plate in the converging beam.

However, on first approximation it can be considered negligible, due both to the very thin thickness of the substrate (less than 1 mm) and to the fine-focusing mechanisms included in the opto-mechanical design of the camera that can help to precisely bring to focus the objective lens on the coated surface of the LVF.

The matrix-sensor plays a fundamental role in determining the overall spatial resolution in the final image. The key-parameter is the size of the photosite, which has to be accounted for with reference to the overall d_{Airy} of the system. In Figure 3.23 is depicted the case in which the Airy disk diameter, namely the period of the black and white pattern, is twice the size of the sensor's photosite S, i.e. $d_{\text{Airy}} = 2S$. Here, the averaged sampled signal might as well be correctly (good case) reconstructed or flattened (bad case) in an uncontrolled fashion. Hence, the sampling frequency is too low and the image is said to be *aliased*[9], thereby containing spurious details, and the sensor undersamples the image.

The maximum spatial frequency of an image, which can effectively be resolved by a pixelated sensor without aliasing, is the spatial frequency where the bright and dark bars line up with adjacent rows or columns of the sensor, as shown in Figure 3.24. This frequency is called the *Nyquist frequency* (NF), measured in line-pairs per millimetre (where a line-pair means a white and a black line):

$$NF = \frac{1}{2 \cdot (2S)} \tag{3.9}$$

with $2S$ the size of a pair of photosites.

The resolution of an image is possible when the photosites-pair frequency is higher than twice the image frequency. In the proposed camera, if the objective lens is stopped down to $f_{\#}$=5.2 and the central wavelength of the visible is considered

[9] ALIASING consists of output image artifacts that occur in a sampled imaging system for input images having significant energy at frequencies greater than the *Nyquist frequency* of the system (Lee, 2005)

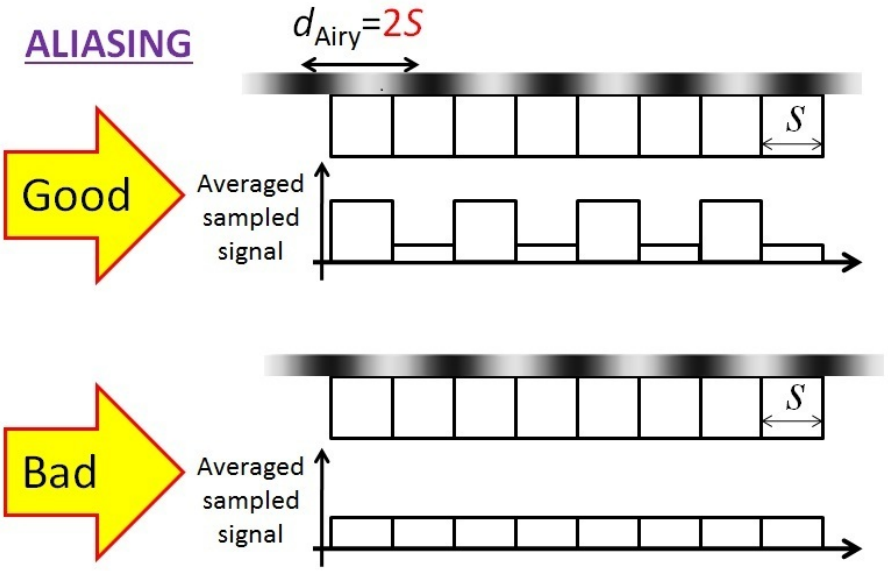


FIGURE 3.23: Undersampled image: the phase of the image pattern between good and bad cases plays a fundamental role in determining the reconstruction

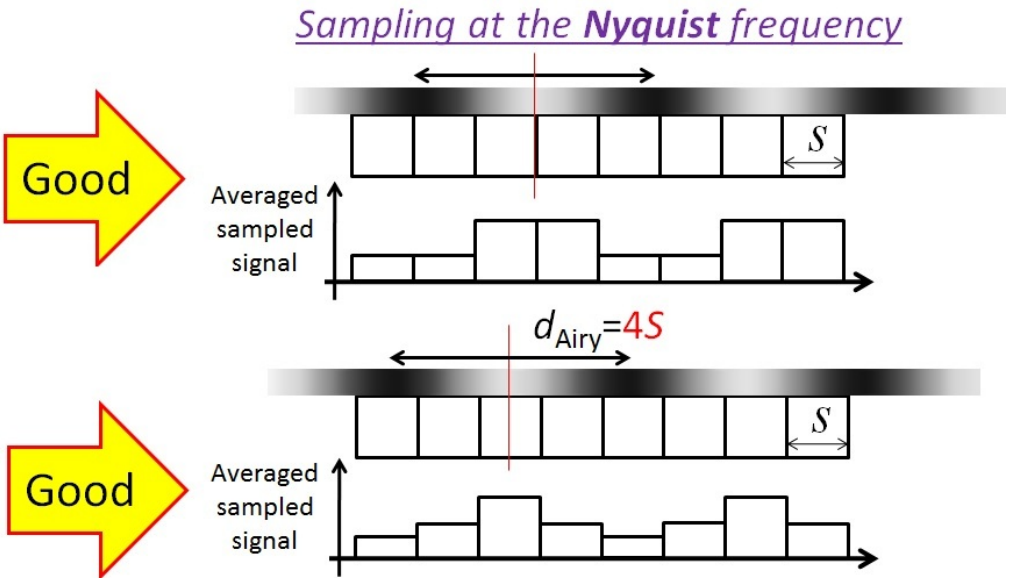


FIGURE 3.24: Image sampled at the Nyquist frequency ($d_{\mathrm{Airy}}=4S$): the success of the reconstruction does not depend on the initial phase

λ=0.55 µm), it follows that the Airy disk diameter d_{Airy}=7 µm, approximately.
Since the photosite size is S=5.5 µm, then $d_{\text{Airy}} < 4S$ and the sensor is too coarse to resolve the fine details of the image proposed by the lenses.
The sensor will resolve image features down to $4S$ line-pairs per millimetre (22 µm), but is unable to resolve frequencies higher than NF (50 line-pairs/mm), i.e. finer details. For full resolution with the current Airy disk diameter, a sensor having photosites of $d_{\text{Airy}}/4$=1.75 µm size should have been the choice.
In other words, the optical elements of the camera can tolerate a degradation of the Airy disk diameter up to 22 µm, without significantly affecting the quality of the reproduced details.

3.8 The *spectral* resolution of the camera

The spectral resolution of the camera can be expressed in two ways. The first is by means of the measured Full-Width at Half Maximum (FWHM) of the LVF instrument-spectral-bandpass function, which is almost constant over the operating range and whose average value is 16 nm; the second is through the spectral separation of two adjacent wavelength lines on the sensor, in accordance with Rayleigh's criterion, and has to be evaluated experimentally.
The instrument-spectral-bandpass function is eventually determined by the LVF bandpass, by the size and the position of the photosite and by its spectral sensitivity. The spectrum recorded with a single snapshot is taken at constant spectral step and at wavelength values that change with the columns of the matrix-sensor. Colour spectral data have to be available at 1 nm bandpass and 1 nm spectral step for a correct colorimetric computation. If the measured spectral data do not possess these properties then deconvolution is necessary, if the instrument-spectral-bandpass is larger than 1 nm, and interpolated with Lagrange (Oleari, 2008) or Sprague (Westland et al., 2012) formula, if the measurement steps are larger than 1 nm (Hunt, 1991;

Oleari, 2008).

If the spectrum is with good approximation constant throughout each spectral interval, where the measuring instrument performs the convolution, there is no reason to perform the deconvolution of the spectrum.

This clearly demonstrates why it is mandatory to illuminate with broadband sources, provided that the spectral sensitivity of the sensor can be considered constant likewise.

In the vast majority of cases, objects do not exhibit too elaborate reflection spectra and it is thanks to their smoothness that no deconvolution is ordinarily required.

In spite of the difficulties to adopt an unambiguous definition, considering the spectral sampling step as the definition of spectral resolution is certainly wrong.

By virtue of the previous considerations, the spectral resolution is constant throughout the whole spectral range and, therefore, does not depend on the particular position on the sensor, because every column "sees the world" exactly like others do, even if this occurs at different times.

Previous considerations apply when the filter is perfectly at the 1st focal plane. This is, of course, simply an abstraction, singe both a defocus or a tilt of the LVF about a transversal direction can affect the instrument-spectral-bandpass function.

The objective lens focuses the light on the LVF (Figure 3.25(a)) and, neglecting to represent the relay lens, the instrument-spectral-bandpass function is determined by the position of the photosite, therefore just behind the filter (Figure 3.25(b)).

If the filter is given a tilt angle, then the maximum displacement of its surface along the optical axis is a or, equivalently, the same parameter can be interpreted as a displacement of the photosite from the 1st focal plane (Figure 3.25(c)).

Now the photosite subtends a half-angle α, fixed by the working $f_{\#}$, to the filter (α=5.5° for $f_{\#}$=5.2); therefore, the collection area from the filter is increased by an amount d each side, that can be thought of as a photosite enlargement. It is clear that the contribution ΔBP to the spread of the instrument-spectral-bandpass function is

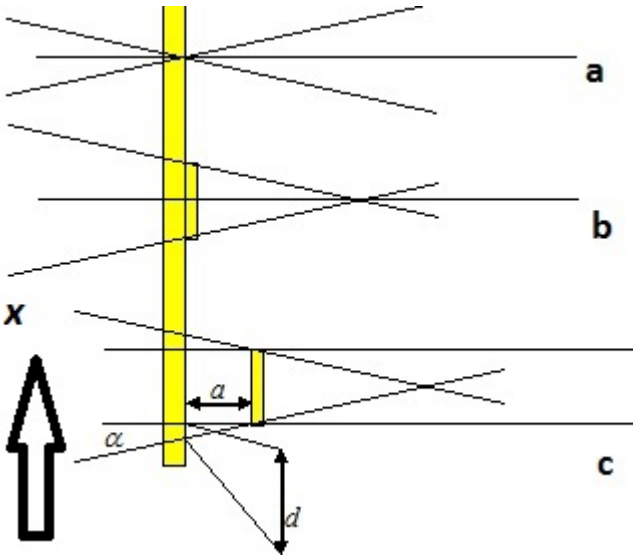


FIGURE 3.25: Geometric representation of the quantities involved in the calculation of the variation of the instrument-spectral-bandpass function, due to a tilting of the LVF

given by:

$$\Delta BP = 2d \cdot \frac{\delta\lambda}{\delta x} = 2a \tan\alpha \cdot \frac{\delta\lambda}{\delta x} \tag{3.10}$$

where x is the direction of wavelength dispersion and $\frac{\delta\lambda}{\delta x}$ ($\mathrm{nm\,mm^{-1}}$) is the wavelength gradient of the filter (Figure 3.26). To get a rough idea of the negligible contribution to the spread of the instrument-spectral-bandpass function, it is enough to consider the limit case in which a=0.1 mm results from a tilt of the LVF, which induces a very low ΔBP=0.6 nm, approximately.

3.9 The opto-mechanical alignment of the camera

Closely related to the performance are the multilevel budgets on allowable deviations from perfection of component dimensions and alignments relative to other components in the instrument. Tolerances strongly influence how an opto-mechanical system will perform and the life cycle cost of that instrument.

Generally speaking, tolerances apply primarily to errors in dimensions, locations, and

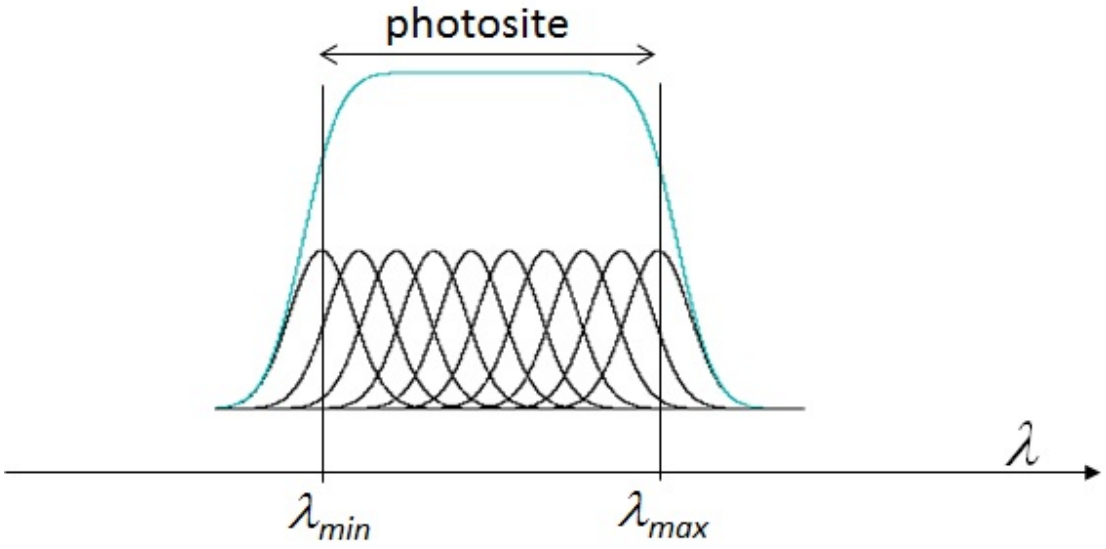


FIGURE 3.26: Transmittance bands centered about different wavelengths (black curves) and their envelope (blue curve) in the spectral range of an enlarged photosite. λ_{min} and λ_{max} are the wavelengths corresponding to the extreme limits of the original size of the photosite. Wavelengths outside contribute to the increase of the instrument-spectral-bandpass function

orientations of optical components. Mechanical part designs also need tolerances in order to fabricate, inspect, and assemble those parts. Usually, broad tolerances mean lower production costs because less expensive fabrication methods can be used and fewer inspections are required. In some cases, broad tolerances on alignment and positioning of mechanical components reduce labor during assembly. On the other hand, broad tolerances introduce greater variability into the parts and their compatibility during assembly. These factors may also reduce performance of the completed instrument (Yoder Jr, 2005).

In the proposed instrument, the two imaging elements (the objective and the relay lenses) are individually aligned by the manufacturer, whereas the LVF and the matrix-sensor (with its glass slip in front of the photosensitive elements) have to be kept orthogonal to the optical axis and mutually parallel. The optical axis is an imaginary straight line passing through the center of each optical surface, thereby

two such axes can be identified for the single imaging lenses.

Here is a schematic list of operations that are of fundamental importance for the correct alignment of the system:

1. bringing the optical axes of the two imaging lenses into coincidence and securing the block to a reference surface

2. positioning the filter and the sensor surfaces perpendicular to the reference surface

3. ensuring the parallelism between the filter and the sensor surfaces

4. placing the block of vertical surfaces orthogonal to the overall optical axis

Even if many strategies can be devised to accomplish each aforesaid task, the simplest method is to employ a laser pointer which is let into the system where it bounces back and forth between the refracting surfaces of the imaging elements. Great care must be taken to ensure that the laser is kept parallel to the reference surface and, once this job is done, after verification that the reflected spot is back to the injection point, the combined use of mechanical translation stages both for the laser pointer and the element to be aligned completes the alignment. Alignment is always a time-consuming job which is accomplished by trial-and-error practice.

After completion of item 1, the perpendicularity of the elements of item 2 is ensured by hitting each surface in turn with the laser on three points or more, to check that the reflected spot has been reflected in line.

The difficulty to perform a correct calculation of the tolerances for the tilt angles of the LVF and the sensor arises from the assumption that they are really planar surfaces, whereas the contrary is probably true. This can be more severe for the substrate where the LVF has been deposited, which may have undergone sensible stress yielding unexpected bending and warping, that would have made it depart

from the planar figure.

The parallelism between filter and sensor can be controlled by illuminating with the laser from the coated face of the filter, thereby exploiting the transmission window at the laser wavelength, and tuning the tilt angles of the two elements until the laser beam is back on the entrance point.

A little more accurate approach employing an autocollimator (Smith, 2000; Hecht, 2001; Laikin, 2006) can be adopted to ensure the orthogonality of the LVF and of the sensor surface to the optical axis of the imaging elements.

The optical autocollimator (Figure 3.27) projects a beam of collimated light. An external reflector reflects the beam back into the instrument where the beam is focused and detected. The autocollimator measures the deviation between the emitted beam and the reflected beam. A reflector is used to direct the emitted beam to the eyepiece graticule of the instrument or, with an alternative layout, to a concentric target through a laser structured light pattern projector, based on a holographic lens that creates a concentric rings' pattern (Figure 3.28). If the surface on which the reflector is placed is perpendicular to the projected light, the beam is reflected back to its point of origin. If, however, the surface is tilted relative to the optical axis of the collimator, the reflected light is displaced (visualized against a graticule, usually a crossline). Any deviation between the projected and reflected beam is measured against a scale and measured in arc seconds. The fourth item is perhaps the most delicate task because the filter and the sensor, in turn, must substitute for the plane dummy mirrors used as external reflectors and this is usually done by means of precision linear translation stages.

3.10 The opto-mechanical set-up

The camera-box consists of 3 plane plates and a top-cover, machined with a hot press, made of anodized polished aluminium. Anodization is required both to

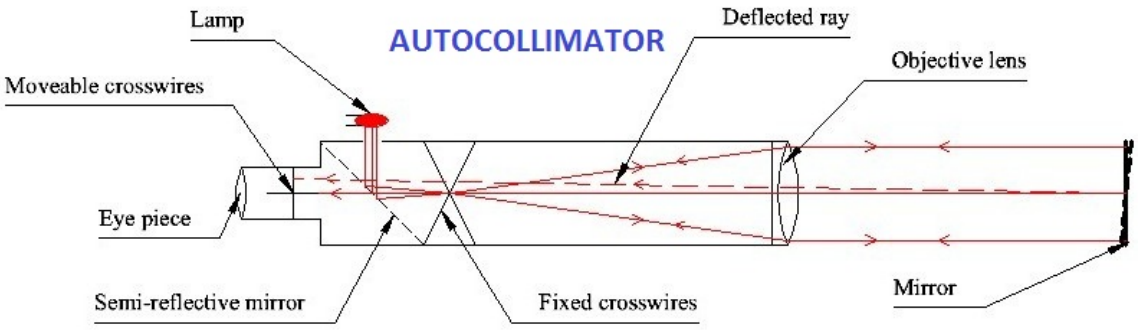


FIGURE 3.27: The optical layout of an autocollimator employing an eyepiece

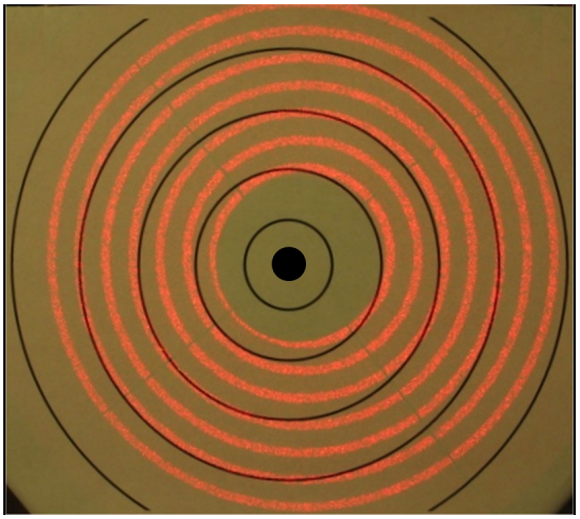


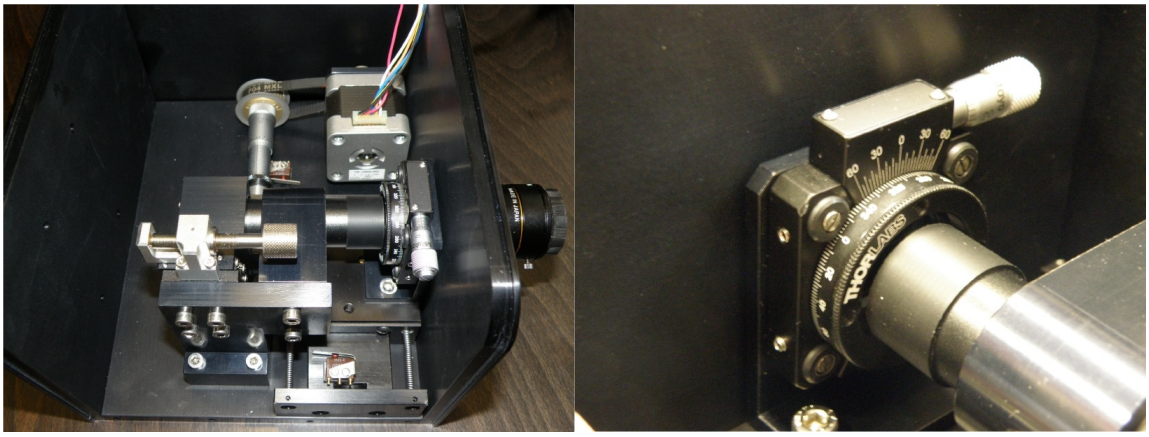
FIGURE 3.28: Decentered concentric rings' pattern on a target, showing the effect of a tilt angle, generated by a laser source coupled with a holographic filter

electrically insulate the surfaces of the box and to reduce the effect of radiation leaks entering the instrument, either from the outside or the inside.

The LVF is locked inside a precision cage rotation mount (model CRM1P/M), selected from the THORLABS catalog[10].

This opto-mechanical component allows the rotation of the filter, about the optical axis of the camera, through a high-precision, backlash-free micrometer, with an angular resolution of 5′. This will be of fundamental importance when the spectral

[10] `http://www.thorlabs.com/`



(a) (b)

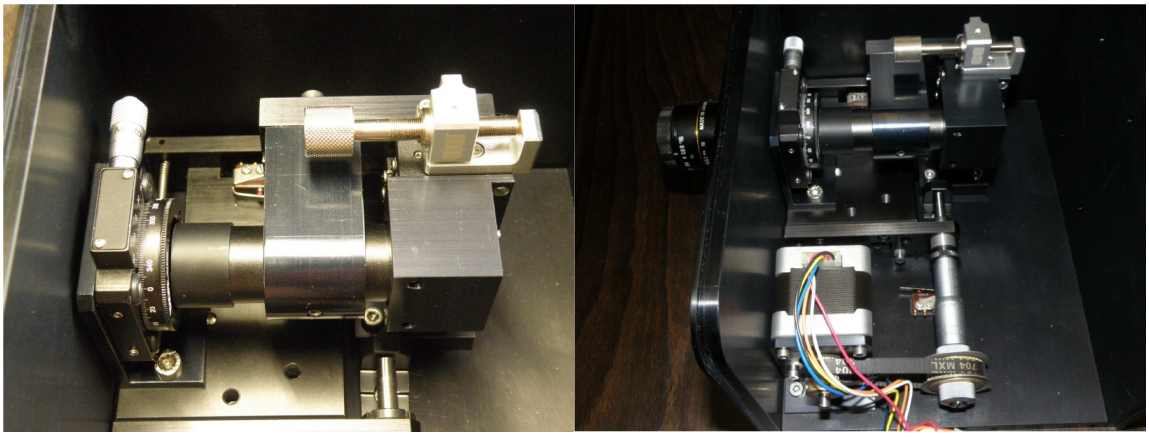
FIGURE 3.29: (a) Top-view of the interior of the camera (b) close-up view of the cage rotation mount

transmission windows of the filter have to be aligned parallel to the columns of the matrix-sensor.

The cage mount is fixed on the translation stage, just behind the aperture where the objective lens is screwed (Figure 3.29). A tubular shielding-sleeve is fitted in the relay lens' cylindrical barrel and protrudes until it nearly shaves the front surface of the cage mount, in order to prevent unwanted rays from entering the imaging path. The base plate is the reference surface for optical alignment (3.9) and on it are loaded the translation stage, the stepper motor and the mounting block, which encloses the sensor and the relay lens.

A single-axis crossed roller translation stage with 13 mm travel length (model NT66-392 from EDMUND OPTICS), allows translations of the sensor and the relay lens as a whole, for focusing along the optical axis (Figure 3.30).

The stepper motor 103-H548-04500 is manufactured by SANYO-DENKY and is coupled with a precision translation axis to supply the motion to the LVF, by a toothed belt drive. The "speed" can be expressed in terms of the number of steps required by the translation axis to shift the filter by the physical length of the photosite, which is



(a) (b)

FIGURE 3.30: Top-views of the interior of the camera showing the positions of the (a) single-axis translation stage for focusing and of (b) the stepper motor used to shift the LVF

110 step/pxl or, equivalently, 20 step/µm.

Because of possible memory-effects induced by hysteresis that could induce repeatability issues, it is recommended to acquire images when the motion is supplied in one-way only.

3.11 The calibration of the camera

The camera must undergo a number of delicate calibration steps, which are mandatory for its correct operation. In fact, the camera must ensure both repeatability, heavily dependent on the geometrical stability, and suitable correspondence with radiometric quantities. The former is fulfilled by the *wavelength calibration*, the latter by the *radiometric calibration*.

3.11.1 Wavelength calibration

Here is a list of the steps required for *wavelength calibration*:

1. vertical alignment of lines of constant wavelength on the sensor

2. minimization of the spot size over the sensor, under monochromatic and broad-band illumination

3. creation of a correspondence between wavelengths and points of the LVF

4. creation of a correspondence between photosites of the sensor and points of the LVF

The first alignment can be carried out by employing suitable sources which emit characteristic spectral lines. In Table 3.4 are reported nominal wavelength values of characteristic lines belonging to lamps available in our laboratory (i.e. Sodium, Mercury, Cadmium and Helium lamps), whereas in Table 3.5 are listed the measured values. The difference in significant digits between nominal and experimental values can be ascribed to the poorer wavelength accuracy (poorer in absolute sense, nonetheless sufficient for efficient instrumental calibration) of the particular spectroradiometer used (PMA-11 by HAMAMATSU[11]). In Figure 3.31 are reported the emission spectra, measured with the spectroradiometer, of all the aforementioned lamps. Each of these lines has to be imaged parallel to the matrix-sensor; this is effected by rotating the filter about the optical axis with the cage rotation stage. Moreover, the spot size of each imaged monochromatic line must be minimized by operating with the single-axis crossed roller translation stage that moves the relay lens and the sensor, as a whole, along the optical axis for best focus.

This condition must be met with every single emission line throughout the full length of the LVF. As an example, supposing to operate with a particular line (such a line can be provided either by the lamp or by a monochromator), at first the LVF is shifted to have that line imaged at the central column of the sensor. After the best focal plane has been determined, the LVF is shifted again to check the focus at different intermediate positions on the sensor up to the extreme ends. If focus

[11] `http://www.hamamatsu.com/`

Table 3.4: Nominal wavelengths (in nm) at characteristic emission lines of four different lamps

Sodium	Mercury	Cadmium	Helium
588.995	404.66	467.81	447.15
589.590	435.83	479.99	501.57
	546.07	508.58	587.56
	579.07	643.65	667.82
			706.52

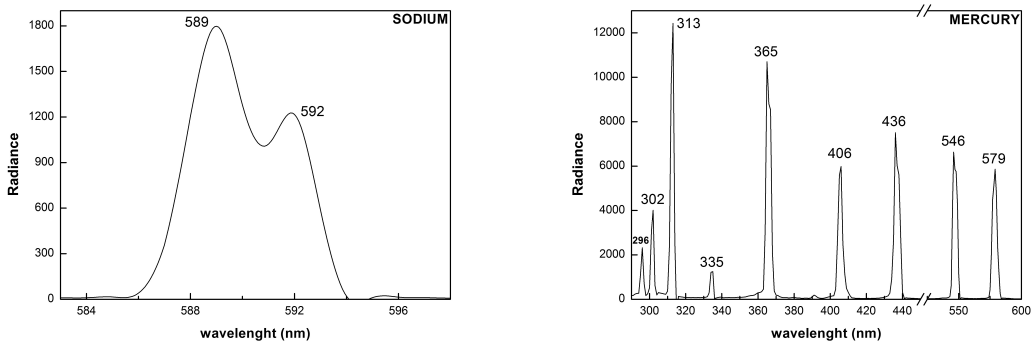
Table 3.5: Measured wavelengths (in nm) at characteristic emission lines of the same lamps

Sodium	Mercury	Cadmium	Helium
589	406	468	447
592	436	480	502
	546	509	587
	579	644	668
			707

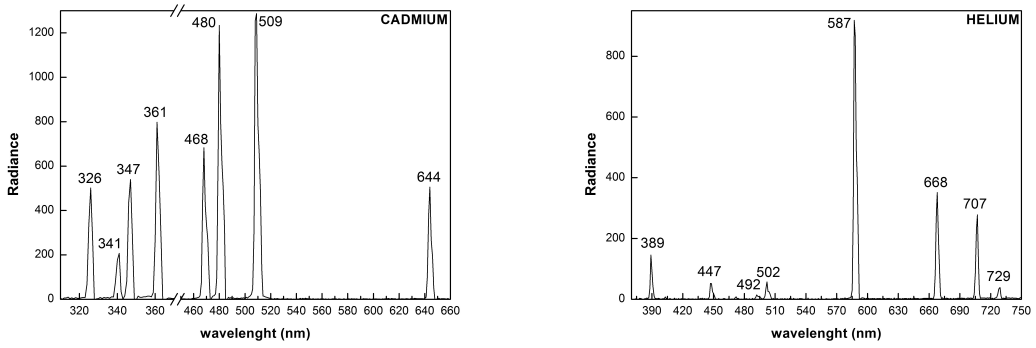
does not degrade with position from end to end the job is complete, otherwise the alignments inside the camera should be revised.

After having associated the wavelengths of spectral lines with positions on the LVF, the wavelength calibration is carried out by establishing the correct wavelength values for filter positions intermediate between points previously assigned with spectral lines. Since the light is detected by the photosites of the matrix-sensor, the mutual geometric relationship between photosites on the sensor and points on the LVF during the acquisition of an image must be considered.

In the practice, the constraint on unit magnification between LVF and sensor can be relaxed a little bit, therefore a projection of the photosite on the filter that slightly exceeds its physical size can be tolerated and the LVF can be ideally partitioned into cells corresponding to the photosites of the matrix-sensor. If both the photosites and the cells on the LVF are indexed, then motion of the LVF parallel to the sensor creates a correspondence between these two different elementary units, which allows



(a) (b)



(c) (d)

FIGURE 3.31: Measured emission spectra of (a) sodium (b) mercury (c) cadmium and (d) helium lamps

to associate a wavelength value to every photosite.

In the case of constant wavelength gradient g over the filter (*linear* relationship), with $g = \dfrac{\Delta\lambda}{\Delta i}$, where i is the index associated to the cells, the wavelength associated with each cell is:

$$\lambda_i = \lambda_0 + gi \tag{3.11}$$

If the linear interpolation is inadequate, then *quadratic* or *cubic* interpolation could be necessary:

$$\lambda_i = \lambda_0 + ai + bi^2 + ci^3 + \ldots \tag{3.12}$$

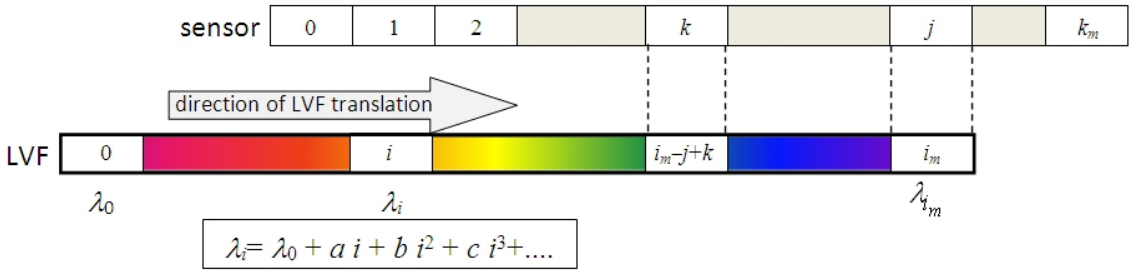


FIGURE 3.32: Spatial correspondence between photosites (on the sensor) and cells on the LVF, for a shift of the LVF of j-times the cell size

The wavelength calibration establishes the values of the coefficients g or a, b, c, ... This correspondence has to be made for any of the cells of the LVF, with respect to a photosite sensor. For example, if the LVF is shifted from left to right (Figure 3.32) j (positive integer) times the cell size (the size of the photosite on the sensor), then the indexed correspondence yields:

- photosite j corresponds to the cell i_m
- photosite k corresponds to the cell i_m-j+k
- photosite 0 corresponds to the cell i_m-j

In this way, the wavelength associated with the cell index is associated with a corresponding photosite index.

The wavelength calibration involves the LVF-sensor mutual position and depends on the relay lens, while is independent of the objective lens. The quality of the relay lens ensures an image distortion that does not exceed the size of a photosite.

Distortion verification

Once the wavelength calibration is finished, the distortion of the whole optical system (objective and relay lenses) should be verified. A pedestrian verification of best focus position, through the minimization of spot size on the sensor, at

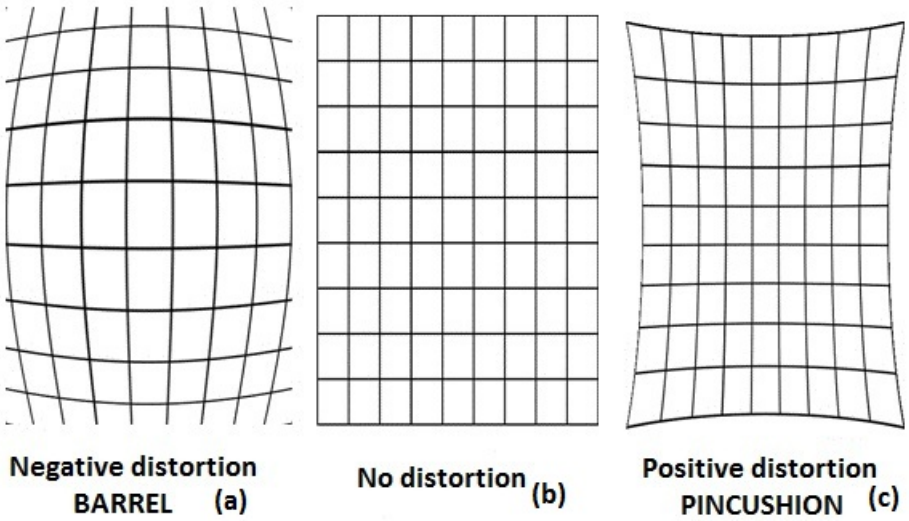


FIGURE 3.33: Grid pattern with (a) negative distortion, (b) undistorted, and (c) with positive distortion

different filter spatial positions for different spectral line should confirm that best focus position does not vary with the wavelength for a high-quality objective lens (that is well-corrected against chromatic aberrations).

Further verification can be carried out by acquiring the image of an object under broadband illumination (a halogen lamp, a Xenon lamp, sunlight), with the clearance (unfiltered portion) of the glass substrate of the LVF (Figure 3.9) fully parallel to the area of the matrix-sensor.

With this configuration, a reflecting target with line grid patterns, i.e. offering repeated parallel lines in orthogonal directions, can be imaged to check the presence of possible distortion issues (this is particularly true for low-quality optics). Distortion has two different forms (Figure 3.33), one with the imaged lines displaced inward from the correct position (negative or *barrel* distortion) and the other outward (positive or pincushion distortion).

The quality of the chosen optics certainly avoids these defects.

3.11.2 Radiometric calibration

The *radiometric calibration* follows the *wavelength calibration* and is performed to provide radiometric significance to the output from every single photosite of the hyperspectral camera.

It can be defined as the process of adjusting an instrument's output using a transfer standard, thereby transferring a national or international measurement scale (Berns, 2000).

An instrument calibrated in such a way can operate in *absolute* mode or in *relative* mode. The former is typical of the spectroradiometers, because the output quantities are absolute quantities, and the latter mode is typical of the spectrophotometers, because they return quantities which are measured in comparison with a reference surface.

When the camera operates in relative mode, the *Spectral Reflectance Factor* (*SRF*) is measured at a region of the scene, whose image is produced on a single photosite of the matrix-sensor, under defined llumination/light-collection geometry, and is defined by the ratio of the spectral flux density of the object-sample ($\Phi_S(\lambda)$) to that of the lambertian and non absorbing ideal surface ($\Phi_{\text{ideal}}(\lambda)$):

$$R(\lambda) = \frac{\Phi_S(\lambda)}{\Phi_{\text{ideal}}(\lambda)} \tag{3.13}$$

Because the ideal surface is only an abstraction, a real surface is used instead, whose *SRF* ($R_{\text{standard}}(\lambda)$) is certified by a metrological laboratory.

Then the measured *SRF* of a surface is:

$$R(\lambda) = \frac{\Phi_S(\lambda)}{\Phi_{\text{standard}}(\lambda)} \cdot \frac{\Phi_{\text{standard}}(\lambda)}{\Phi_{\text{ideal}}(\lambda)} = \frac{\Phi_S(\lambda)}{\Phi_{\text{standard}}(\lambda)} \cdot R_{\text{standard}}(\lambda) \tag{3.14}$$

Generally, in the laboratories a duplication of the certified surface is used, named *working standard*.

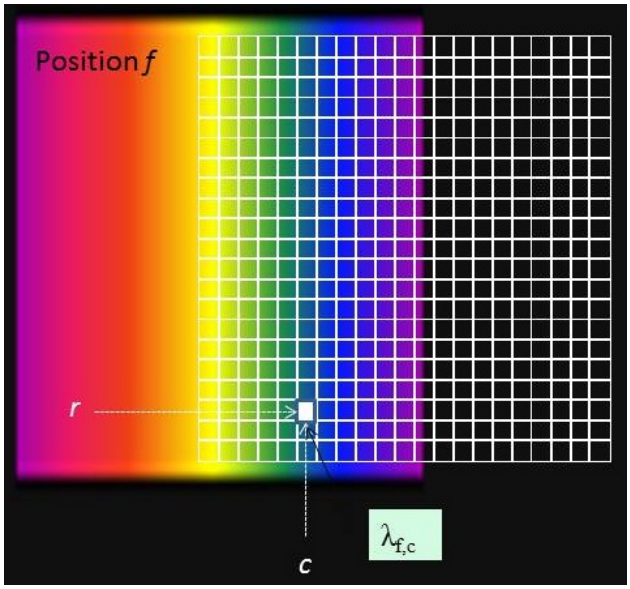


FIGURE 3.34: A general mutual position of the LVF with respect to the matrix-sensor

The photosite produces an electric signal that has to be related to the flux crossing the photosite. The useful signal has to be a quantity proportional to the crossing flux, therefore it must be reduced by the dark signal, i.e. the photosite signal D when the objective lens is obscured, and multiplied by a linearization factor[12], that normally is a constant.

In the case of linearization factor equal to 1, the *SRF* measured by the photosite (r,c), at the r^{th}-row and c^{th}-column, with the LVF at a position f with respect to the sensor (Figure 3.34), is:

$$R_{r,c}(\lambda_{c,f}) = \frac{S_{r,c,f} - D_{r,c}}{S_{\text{standard},r,c,f} - D_{r,c}} \cdot R_{\text{standard},r,c}(\lambda_{c,f}) \tag{3.15}$$

[12] It is necessary for the electrical signals from the detector to be proportional (linear) to the incident radiant power. *Linearization* can be verified by using a set of neutral lambertian *certified surfaces* of different average reflectance and calculating the *SRF* of 3.15 using each time a different neutral reference as standard. In case of non-linearity, the data need be linearized by suitable algorithms. Normally, non-linearity in the response of a detector, or in its associated circuitry, would result in a non-linear error (provided that it is the only source of error) that is proportional to the square of the measured values of the *SRF*. Moreover, it is always necessary to specify the range of radiances where the detector is linear (Hunt, 1991).

where $S_{r,c,f}$ represents the (r,c) photosite signal produced at position f of the LVF and $D_{r,c}$ is the signal produced by the photosite when the objective lens is obscured (*dark signal*).

Once the LVF has scanned all the wavelengths related to the chosen spectral sampling step, in front of the c^{th}-column of the sensor, the collection of the values of 3.15 related to the photosite (r,c) represents the SRF.

This relative quantity is independent of the particular illuminating source and of the radiometric calibration.

In the case of absolute calibration, the quantity measured by the (r,c) photosite is put in correspondence with the radiance scale (units $\mathrm{W\,m^{-2}\,sr^{-1}}$) by a calibration factor $C_{r,c,f}$:

$$L_{\text{certified}}(\lambda_f) = C_{r,c,f} \cdot (S_{r,c,f} - D_{r,c}) \tag{3.16}$$

The calibration factor $C_{r,c,f}$ is obtained by associating to any photosite (r,c) the spectral radiance of a light source certified by a metrological laboratory (usually with emission spectrum similar to the spectrum of a blackbody radiator).

Often, in this calibration, the correspondence with the radiance scale is made up to a factor and the unit is named *arbitrary unit* (au).

The calculation of the calibration factor for every photosite can be performed by illuminating an integrating sphere with a certified broadband light source and then capturing with the camera the image of a region internal to the exit-port, that has a uniform and lambertian emission.

An unfiltered image of a region internal to the sphere port is captured and analyzed to correct spatial non-uniformities, thereby a correction factor $p_{r,c}$ (which accounts for illumination drop-off at the edge of the matrix-sensor and for the inherent characteristics of the response of individual photosites) is associated to every single photosite to remove spatial non-uniformities.

Furthermore, images filtered by the LVF at different wavelengths are captured as well; the linearized signal $(S_{r,c,f} - D_{r,c})$, read from a generic photosite, depends on the illumination source, on the the transmittance of the LVF, on the position of the photosite and on the spectral sensitivity $\varepsilon(\lambda_f)$ of the photosite.

The spectral transmittance $t(\lambda_f)$, of the LVF at the wavelength λ_f, and a factor K with radiometric dimension are introduced such that:

$$L_{\text{certified}}(\lambda_f) \equiv K \cdot \frac{S_{r,c,f} - D_{r,c}}{p_{r,c}\varepsilon(\lambda_f) \cdot t(\lambda_f)} \tag{3.17}$$

The LVF transmittance $t(\lambda_f)$ is defined by:

$$t(\lambda_f) \equiv \frac{S_{r,c,f,\text{unfiltered}} - D_{r,c}}{S_{r,c,f,\text{filtered}} - D_{r,c}} \tag{3.18}$$

where $S_{r,c,f,\text{unfiltered}}$ and $S_{r,c,f,\text{filtered}}$ are the linearized signals produced by the photosite (r,c) when the monochromatic radiation λ_f crosses the photosite without and with the LVF, respectively. Since the LVF should have no irregularities, the result is expected to be independent of the photosite (r,c) and the spectral transmittance $t(\lambda_f)$ to be constant orthogonally to the direction of wavelength dispersion. In this way the radiometric calibration dependence on the LVF and on the photosite are factorized and the calibration factor is:

$$C_{r,c,f} \equiv \frac{K}{p_{r,c}\varepsilon(\lambda_f)t(\lambda_f)} \tag{3.19}$$

As already mentioned, the response of the photosites should be linear with the incident radiant power, otherwise a number of unfiltered images captured with different exposure-times is used to check that the signals are proportional to the exposure-time.

The *accuracy* of an instrument is the degree to which measurements agree with those made by a standard instrument or procedure in which all possible errors are at

minimum; the *precision* is the consistency with which measurements are made. The former is affected by *systematic errors*, the latter by *random errors*.
The accuracy is usually determined by comparing the results with results obtained by metrological laboratories using the best possible instrumentation and standard methodologies, the precision is determined through statistical analysis on a set of measurements of a certain quantity.
In practice, measurements on the white working standard, neutral certified surfaces and certified blackbody-like sources have to be thoroughly repeated and averaged; the accuracy of the camera is assessed through a comparison of the average measurements of the *SRF* or the spectral radiance with the corresponding values certified by metrologic laboratories.
The precision of the measurements of spectral quantities will be expressed by the standard deviation of the measurements.

Stray-light

An important remark must be made about the sources of systematic errors leading to a lack of accuracy, in particular about the *stray-light*.
Stray-light is the light that follows paths other than intended in the instrumental design and crosses the detector altering the intended observation.
Often the elimination or reduction of stray-light is difficult and the origin of stray-light is in the opto-mechanical design.
The LVF only transmits or reflects light, and this reflection is the only source of stray-light in the hyperspectral camera. This reflected light is present only in the region between LVF and objective lens and is a stray-light (Figure 3.35).
All the metallic parts in this region are absorbing black coated and the rear face of the objective lens has anti-reflex coating. These properties prevent the existing stray-light from crossing the sensor.

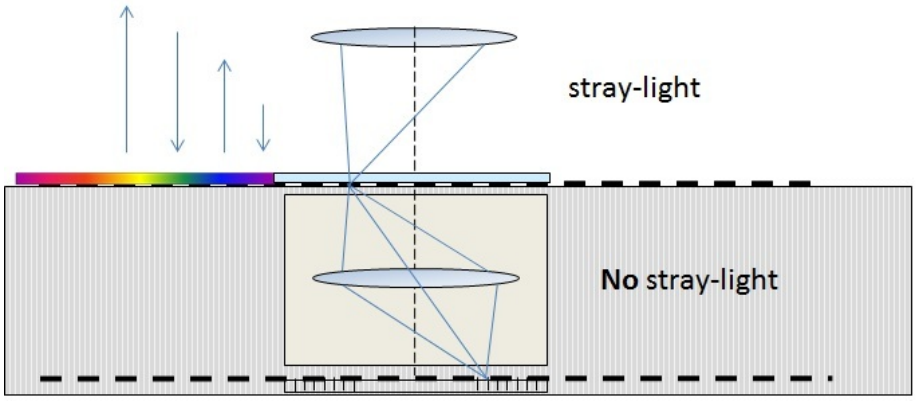


FIGURE 3.35: Layout of the hyperspectral camera with focus on the possible sources of stray-light

3.12 Camera output

With the numerical information collected in Table 3.6 it is possible to calculate some useful information about the output of the hyperspectral camera. Supposing to translate the LVF at a convenient speed (here a reasonable choice was 20 step/µm), the results depend on the value of the required spectral sampling step (p). Once a

Table 3.6: Specifications of the LVF and the matrix-sensor, useful for hyperspectral camera's output calculations

LVF operating range - Δ_s (nm)	510
Filter size - l (mm)	16
Sensor size - L (mm)	6.6
Stepper motor "*speed*" (step/µm)	20

value of p has been selected, the translation step of the LVF can be calculated by:

$$P_p = \frac{l}{INT\left(\frac{\Delta_s}{p} + 1\right)} \tag{3.20}$$

and the number of output images by:

$$N_I = INT\left(\frac{\Delta_s}{p} + 1\right) + \frac{L}{P_p} = \frac{l + L}{P_p} \tag{3.21}$$

A collection of estimated output data, according to the particular value of p, is reported in Table 3.7: The full acquisition time is expected to range from 1 min to

Table 3.7: Estimated output information for the hyperspectral camera

p (nm)	*5*	*10*	*20*
P_p (µm)	155	308	192
N_I	146	74	38
Size (MB)	300	150	75

2 min, depending on the selected speed of the motor and on the integration time at every single position of the LVF in front of the matrix-sensor.

Conclusions

The focus of this thesis has been on the design and construction of a compact portable hyperspectral camera for the capture of still images.

Its strong point is the use of a *Linearly-Variable Filter* (LVF), based on the *Induced-Transmission* technology. This filter is the only moving part of the instrument and allows every single column of the matrix-sensor to be filtered, with respect to incoming light, with the whole set of spectral bandpass as the LVF is scanned parallel to its dispersion line.

The path from the original idea to the final layout was full of challenges, the most demanding being the minimization of both the optical elements and moving parts. In fact, the design went through continuous refinements on a regular basis.

The design has arrived to an essential structure, whose simplicity is evident from the comparison with other commercial solutions that are summarized in Chapter 2.

This project, object of communications at international conferences on Optical Metrology (*Optics for Arts, Architecture and Archaeology* - Munich (Germany) 2011) (Piegari et al., 2011) and, particularly, on Spectral-Imaging (*13th International Symposium on Multispectral Color Science* - San José (USA) 2011, *3rd Annual Hyperspectral Imaging Conference* - Rome (Italy) 2012) (Della Patria et al., 2011, 2012), has been approved and very well considered.

What actually represents a source of pride for us is the interest shown by experts in the field of spectral imaging about purchasing the proposed instrument.

Great attention was paid to the balance between optical performances and the requisite of compactness, that in some circumstances inevitably led to a compromise. Some activities have been committed to external professionalism, indeed really difficult to find out in Italy by now, that protracted excessively the manufacturing of the opto-mechanical assembly, and two really demanding activities still remain before the camera can be fully operating:

- the wavelength and radiometric calibration, whose procedures have been discussed in Chapter 3;
- the project and commission of the software for the spectra acquisition and colour image visualization.

There is still room to provide more compactness to the camera, and this will constitute the object of future efforts.

Among the many field of applications, this camera is really promising in forensics and artworks conservation. In the former, it would be of great support in expert reports on handwriting, in the latter it would allow effortless in-situ operation. Nonetheless, it is expected that it could be successfully employed in the biomedical field, where there is urgent need to adopt non-invasive diagnostic techniques for rapid identification of potential anomalies.

Appendix A

Compression of spectral data: a short review

In colour research, three-dimensional colour coordinate systems have been largely used which, as is well-known, suffer from *metamerism*, i.e. several different spectra result in the same set of coordinates. Differently, spectra made of many components corresponding to each wavelength range are not affected by such a drawback, and permit high accuracy for the quantification of colour.

Whatever the technique used to acquire spectral images, a great number of images may have to be taken according to the desired degree of spectral resolution; accordingly, a copious amount of image wold demand processing and storage.

The compression of spectral data is therefore of great importance and much effort was made to find out the minimum number of information needed for accurate spectra reconstruction (Maloney and Wandell, 1986; Parkkinen et al., 1989; Miyazawa et al., 2001; Tzeng and Berns, 2005).

Basically there are two different approaches for compression, either by obtaining the measured spectra and then compressing by software (Parkkinen et al., 1989; Usui

et al., 1992) or by designing the low dimensional spectral imager[1] so that the output is already the optimal component images for spectral reconstruction.

Finding the minimum number of color filters (*basis spectra*) for accurate spectral reconstruction has been the focus of several researches in the last thirty years and, for this purpose, the data sets of the MUNSELL *book of colours* (Munsell, 1979), of the Swedish NATURAL COLOR SYSTEM (NCS) (Hård and Sivik, 2007) and of colour spectra of natural samples (Westland et al., 2000) have been considered.

Here is given a short review about the most important strategies of spectral images' compression, with focus on the methods used to find out the optimum number of primaries.

Around the mid-80's several works were published which represented chromatic data by the sum of some basis functions (Sobagaki, 1984; DZmura and Lennie, 1986; Maloney and Wandell, 1986; Young, 1986), but emphasis was on the trichromatic representation of colour based on the tristimulus values.

Before that time, the first work which investigated the reflectance curves of the Munsell color chips by using PCA (Appendix B) was by J. Cohen in 1964 (Cohen, 1964). The author used the correlation matrix of 150 spectra taken at 10 nm intervals and concluded that the first four principal components were enough to reconstruct the measured curves. In 1986 the same analysis was repeated by L.T. Maloney *et al.* (Maloney and Wandell, 1986) on a set of 462 samples, but the authors never did publish their vectors.

In 1989 a work was published (Parkkinen et al., 1989) which enlarged the investigation on the MUNSELL *book of colours* to a set of 1257 colour chips. The authors found that, after using the PCA, 8 eigenvectors sufficed for accurate reconstruction, in the range from 400 nm to 700 nm at 5 nm intervals, of the reflectance spectra of chips

[1] An IMAGER is a common term which indicates a general device used to capture images, either trichromatic or spectral

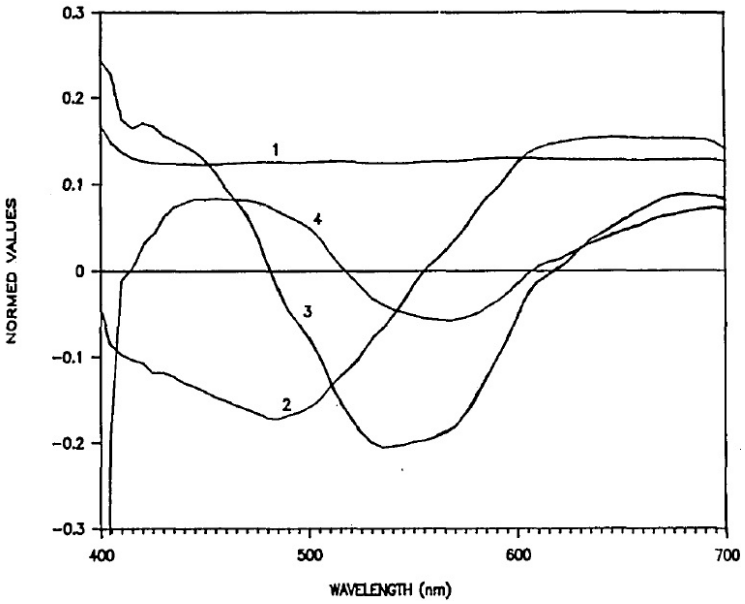


FIGURE A.1: The four first eigenvectors of a set of 1257 colour chips spectra from the MUNSELL *book of colours*

with different hues, saturation and values. In Figure A.1 are plotted the the first four eigenvectors, with the first one representing the mean of the measured vectors and whose flatness indicates uniform covering of the colour space. Therefore, an arbitrary colour spectrum $s_a(\lambda)$ could be represented by the following reconstruction equation by n eigenvectors $\boldsymbol{e_j(\lambda)}$:

$$s_a(\lambda) = \sum_{j=1}^{n} [s_a(\lambda)|\mathbf{e_j}(\lambda)]\mathbf{e_j}(\lambda) \tag{A.1}$$

where the notation $[\bullet|\bullet]$ denotes the inner product. It is straightforward that the A.1 reconstructs a spectrum exactly only if it belongs to the original data set.

The A.1 is independent of the choice of the sign of the vector, so the decision of the sign fixes the form of the principal colours corresponding to the eigenvectors.

The accuracy of reconstruction, as a function of the number of elements of the reconstruction basis, was also evaluated through the calculation of both the maximum and average errors of colour coordinates. As a result the authors established that 7

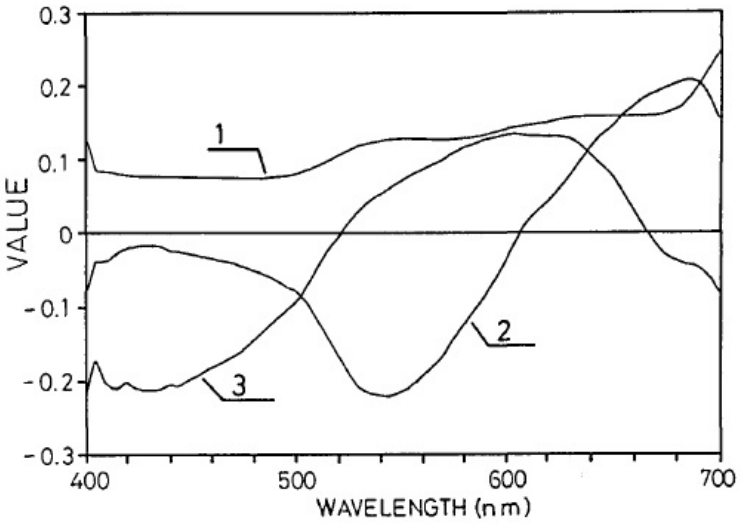


FIGURE A.2: The first three basis vectors of the set of 218 natural colours

to 10 eigenvectors sufficed.

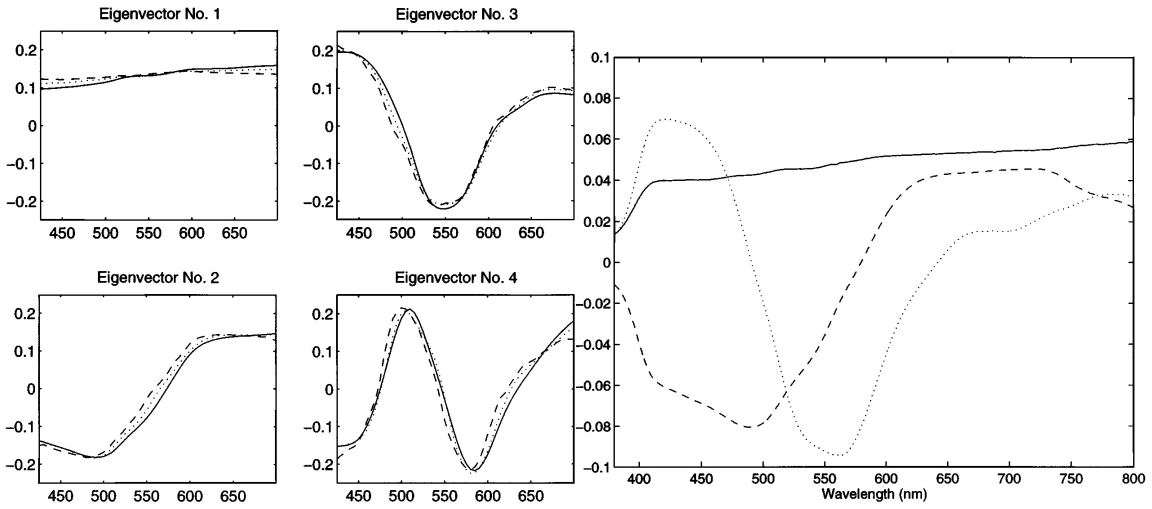
A work published in 1990 extended the previous one to natural colors (Jaaskelainen et al., 1990), including the analysis on 218 spectra of natural-colour samples. As with the Munsell samples, such spectra were taken in the 400 nm to 700 nm at 5 nm intervals on flowers, flower clusters, leaves and berries. The eigenvalues and eigenvectors (forming an orthonormal basis) of the correlation matrix were determined and Figure A.2 shows the plots of the first three vectors of this natural basis. Since colour spectra of natural samples exhibit smooth shapes as well as the samples of the MUNSELL *book of colours*, the authors compared the reconstructive ability of the two bases by representing natural colours both in the subspace spanned by the Munsell basis vectors and in the subspace spanned by the natural colour basis vectors. The Munsell basis produced a good reconstruction of all spectra, including natural data, even if the natural basis was a little better in reconstructing its original data set. The authors concluded that the Munsell basis vectors were able to reconstruct natural colours as accurately as the natural basis did, but it took more vectors in the Munsell basis to match the accuracy in both subspaces.

An extremely interesting work (Lenz et al., 1996) published in 1996 described some investigations of the space of colour spectra that are relevant for human colour vision. The authors started from raw colour spectra of three databases of spectral data; more precisely, two of them are based on the MUNSELL *book of colours* and one on the NATURAL COLOR SYSTEM developed at the Scandinavian Color Institute AB. The first Munsell database (*Munsell-I*) consisted of 1253 spectra represented by 61 measurement each, from 425 nm to 700 nm with a step length of 5 nm. The second Munsell database (*Munsell*) consisted of 1269 spectra of 421 measurements each, in the range from 380 nm to 800 nm in 1 nm steps. The NCS database consisted of 1513 spectra in the wavelength range from 380 nm to 780 nm in 5 nm steps.

The eigenvectors and the eigenvalues of the correlation matrix were calculated in the band from 425 nm to 700 nm and Figure A.3(a) shows the first four eigenvectors from the three different databases, whereas Figure A.3(b) displays only the first three from the *Munsell* database in the full spectral range. It was demonstrated that the first few eigenvectors were almost identical for all databases, therefore the databases had very similar statistical properties. In the same paper the authors tested a different way to find the filter functions (basis of eigenvectors), based on unsupervised learning procedures of the theory of *neural networks*.

From the reference (Johnson and Wichern, 1992):

> "A *neural network* (NN) is a computer-intensive, algorithmic procedure for transforming inputs into desired outputs using highly connected networks of relatively simple processing units (*neurons*). The three essential features of an NN are:
>
> - the *neurons*
> - the network architecture describing the connections between the *neurons*
> - the training algorithm (*learning rule*) used to find values of the network parameters (*weights*) for performing a particular task



(a) (b)

FIGURE A.3: (a) First four eigenvectors for the three databases: *Munsell-I* (dashed), *Munsell* (dotted), NCS (solid) (b) first three eigenvectors: 1st-solid, 2nd-dashed, 3rd-dotted from the *Munsell* database

> *Neurons* are connected to one another in the sense that the output from one *neuron* can serve as part of the input to another one. Each *neuron* transforms an input to an output using some prespecified function that is typically monotone, but otherwise arbitrary. This function depends on parameters whose values must be determined with a training set of inputs and outputs.
>
> Network architecture is the organization of *neurons* and the types of connections permitted. In statistical applications, the *neurons* are arranged in a series of layers with connection between nodes in different layers, but not between nodes in the same layer. The layer receiving the initial input is called the *input layer*. The final layer is called the *output layer*. Any layers between the *input* and *output* layers are called *hidden layers* (Figure A.4).
>
> NN can be used for discrimination and classification. When they are so used, the input variables are the measured group characteristics X_1, X_2,...,X_p, and the output variables are categorical variables indicating group membership..."

The property that is of primary significance fora an NN is its ability to *learn* from its environment and to improve its performance through learning[2] The starting

[2] The notion of LEARNING is too difficult to be defined in precise manner, since too many are the

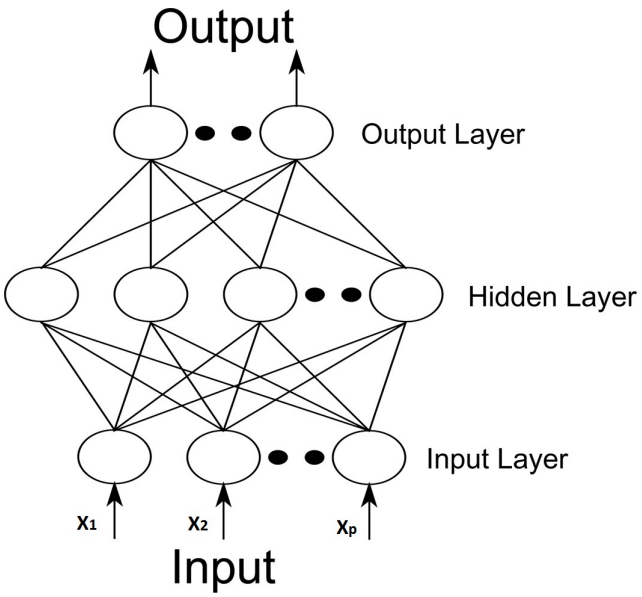


FIGURE A.4: A neural network with one hidden layer

requirement for the extracted feature vectors was that they should minimize as possible the mean squared error compared with the input data; the quality function chosen was $Q = Q_D Q_A$, where:

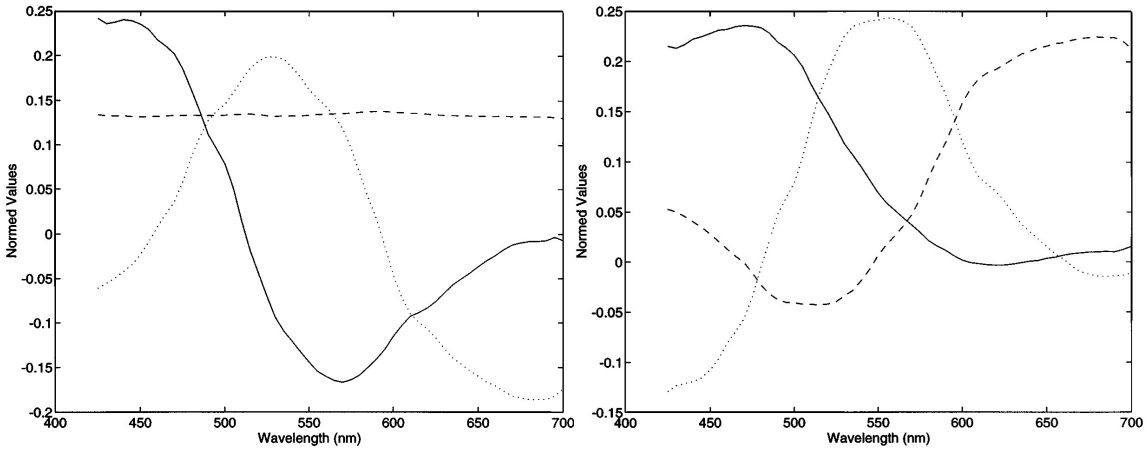
$$Q_D = -det(C_0) \tag{A.2}$$

$$Q_A = Q_V = \sum_{n=1}^{N} var(o_n^2) \tag{A.3}$$

$$\frac{1}{Q_A} = Q_4 = \sum_{n=1}^{N} E\left[o_n^2(1 - o_n^2)\right] \tag{A.4}$$

The patterns in the database are given by the vectors p and so the correlation matrix of the input signals is $C_p = E(pp^T)$, the symbol E meaning the expected value of the stochastic variable in parentheses and the superscript T indicating the transpose

activities associated with it. In the particular case of a NN, here is reported the definition from the reference (Mendel and McLaren, 1970): *"Learning is a process by which the free parameters of an NN are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter changes take place"*



(a) (b)

FIGURE A.5: Learned filters after optimization of quality function (a) Q_4 (b) Q_V using the *Munsell-I* database

operation of linear algebra. For an input $\boldsymbol{p}$, the system output is the vector $\mathbf{o} = \mathbf{f}(\mathbf{p})$, so the correlation matrix of the output is $C_{\mathbf{o}} = E(\mathbf{o}\mathbf{o}^{\mathbf{T}})$. Note that when $\boldsymbol{f}$ is linear, there exists a *weight matrix* $\boldsymbol{W}$ such that $\mathbf{o} = \mathbf{W}\mathbf{p}$. It does not change after iterative optimization procedures to find the filters have been completed.

Very efficient optimizations were effected through the use of optimization programs (Shanno and Phua, 1980; Dennis and Schnabel, 1987; Nocedal, 1992; Moré and Thuente, 1994). In a first series of experiments, a system consisting of three units was trained and the resulting learned filters after optimization are reported in Figure A.5. Here it is clear that the system based on the Q_4 quality function stabilizes in a state that is similar to the three eigenvectors of the correlation matrix, whereas the system based on the Q_V splits the domain into three separate spectral bands. In a second series of experiments where positive weights were enforced, the resulting filter vectors after optimization through the Q_V quality function (Figure A.6) can be characterized as *blue*, *green* and *red* filters. This demonstrated that such a systems could be used for accurate reconstruction of the colour space used by the human visual system. Since the statistical properties of the three databases were the same,

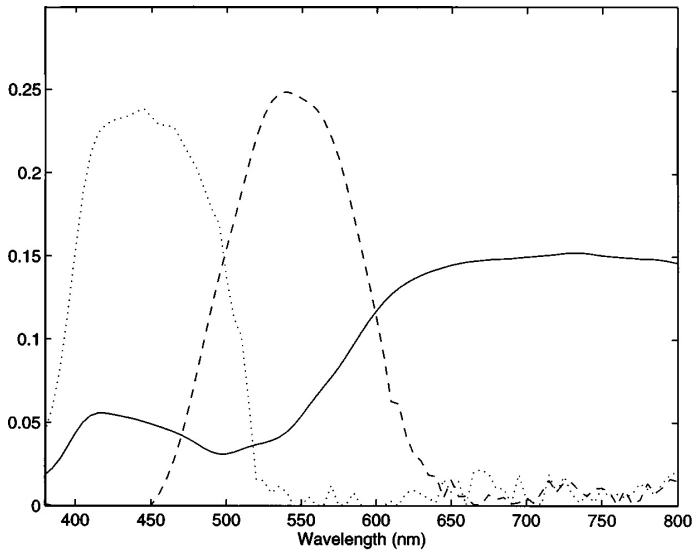


FIGURE A.6: Learned filters after optimization of quality function Q_V using the *Munsell* database

the similarity is reflected in the structure of the filter systems computed from the three databases. By increasing the number of filter functions in the training set, the learned filters exhibited bandpass characteristic although often with a number of multi-modal functions. The orthogonality of the eigenvectors resulting from the PCA does not prevent principal components from having negative sign, which poses problems for direct implementation in optical components. The search for a vector filter set with positive coefficients, which could be directly used in optical pattern recognition, was the goal of another interesting work (Hauta-Kasari et al., 1997) that exploited the potentialities of unsupervised neural networks[3]. The authors investigated the clustering properties of competitive learning and self-organization methods (Haykin, 1994; Kohonen, 2001), trying to cluster the input colour spectra of the MUNSELL *book of colours* and use the centers of these clusters for a representation

[3] In UNSUPERVISED or SELF-ORGANIZED learning there is no external teacher that oversees the learning process. The free parameters of the NN are optimized with respect to a *task-independent measure* of the quality of the representation to be learned. Once the NN has become tuned to the input data, it develops the ability to form internal representations for encoding features of the input and thereby to create new classes automatically (Becker, 1991)

of the whole colour space. Competitive learning[4], which is the mechanism behind unsupervised NN clusters the data without a-priori knowledge of the right cluster distribution. The optimization learning algorithm[5] (Grossberg, 1982) incorporated a particular kind of self-organizing maps (Kohonen, 2001).

The NN clusters the input data so that the weight vectors are the centers of these clusters. Since input data are the measured colour spectra (from 400 nm to 700 nm at 5 nm intervals), which contain only positive coefficients, as a consequence the colour filters will be positive as well. Furthermore, the spectra will be separated from each other because the winner neuron is the only one that learns during each learning cycle. The filters were used to reconstruct the *Munsell* database by the *singular value decomposition* technique (SVD) (Pratt, 1978; Jolliffe, 2005) which, in the general case of non-orthonormal vectors, reconstructs the spectrum $\boldsymbol{s}$ by the spectrum $\boldsymbol{s}'$, as follows:

$$\mathbf{s}' = \mathbf{W}(\mathbf{W^T W})^{-1}\mathbf{W^T s} \tag{A.5}$$

with W as the matrix of the filter set. The first eight eigenvectors obtained by the PCA (the first five of which have a fidelity value of 99.93 %) and the eight filters returned by the competitive neural network with 8 neurons, after 50000 learning cycles, are shown in Figure A.7 and Figure A.8, respectively.

The filter set in Figure A.8 represents regularly the spectral range from 400 nm to 700 nm. After they have been orthogonalized by the SVD technique, it was demonstrated that were highly correlated to the basis vectors of the PCA.

Some of the same authors also demonstrated, following the same procedure as before,

[4] In Competitive Learning the output neurons of a NN compete among themselves to become active. Only a single output neuron is active at any one time. For this reason, competitive learning is highly suited to unveil statistically salient features that can be used to classify a set of input patterns. The winner neuron is called a *winner-takes-all* neuron. Also, there is substantial evidence for competitive learning playing an important role in the formation of topographic maps in the brain (Durbin and Rumelhart, 1989; Ambros-Ingerson et al., 1990)

[5] A learning algorithm is a prescribed set of well-defined rules for the solution of a learning problem

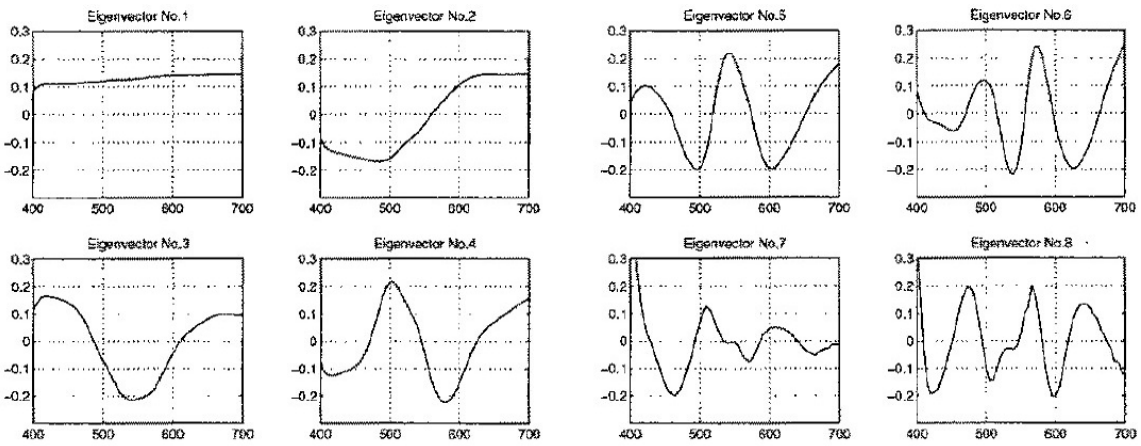


FIGURE A.7: The first 8 eigenvectors from the PCA

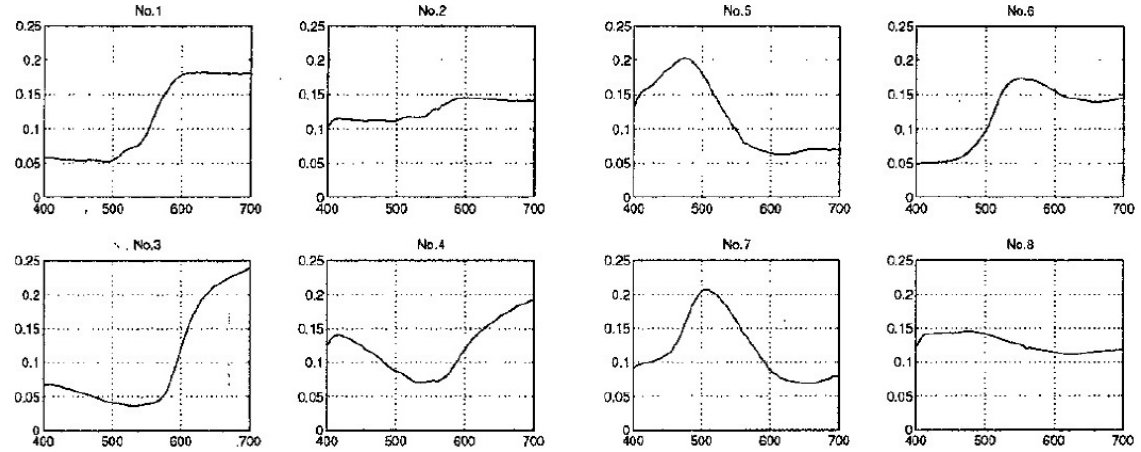


FIGURE A.8: Learned filters functions after competitive neural network

that the optimal number of filter for spectral reconstruction was four (Hauta-Kasari et al., 1999; Miyazawa et al., 1999). They implemented the previous filter set into a liquid-crystal spatial light modulator and found that the lowest reconstruction error in terms of average colour difference, $\Delta E^*(1976)$=10.13, occurred with four filters (Figure A.9). Lastly, the application of PCA for dimensionality reduction to a set of 125 colour patches, printed with a continuous-tone dye-diffusion printer, demonstrated that the best results for reflective samples could be achieved in the absorptance space ($\Phi(\lambda)$), rather than spectral reflectance factor ($R(\lambda)$) space (Tzeng

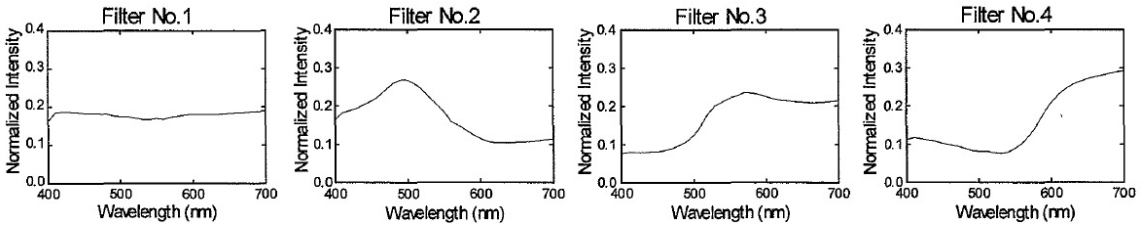


FIGURE A.9: The optimal four filter basis from (Hauta-Kasari et al., 1999; Miyazawa et al., 1999)

and Berns, 2005), thus confirming some theoretical formulations made by some of the authors themselves (Imai et al., 2000; Berns, 2002).

Here the spectral reflectance factor data $R(\lambda)$ are transformed into spectral absorptance data by using the transparent form of the Kubelka-Munk theory (Nobbs, 1985; Oleari, 2008), as follows:

$$\Phi(\lambda) = -0.5 \ln \left(\frac{R(\lambda)}{R_g(\lambda)} \right) \tag{A.6}$$

where $R_g(\lambda)$ is the spectral reflectance factor of the substrate.

The filter sets are shown in Figure A.10 and Figure A.11. It resulted that 3 eigenvectors sufficed to explain almost entirely the total variance of the set in the $\Phi(\lambda)$ space. The authors concluded that the $R(\lambda)$ data required more eigenvectors to match the same amount of variance. They also demonstrated that as the number of colorants increased to form large data sets, the application of PCA was more effective in the $R(\lambda)$ space.

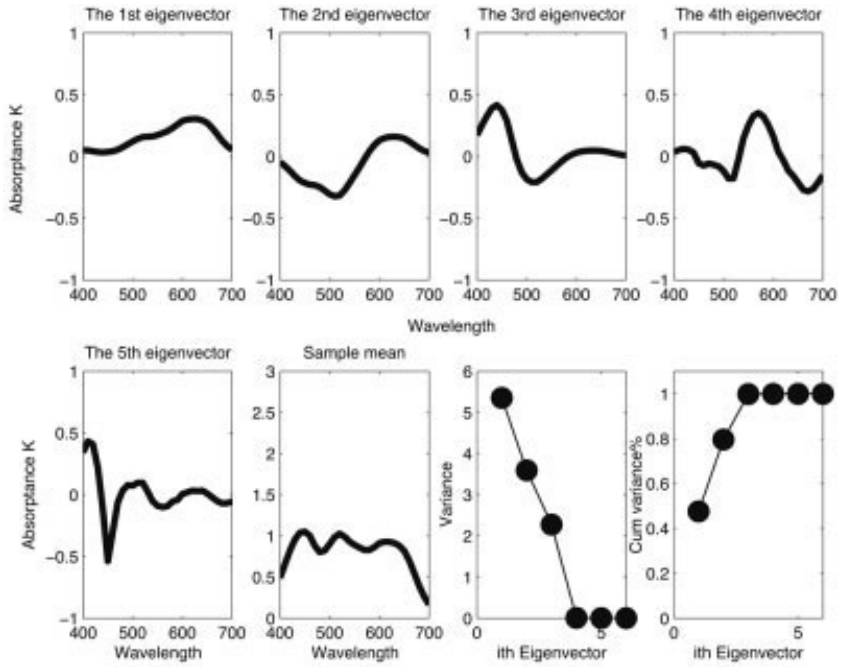


FIGURE A.10: First five eigenvectors, mean and score plots from the absorption spectra of the dye-diffusion-thermal-transfer printer sample set

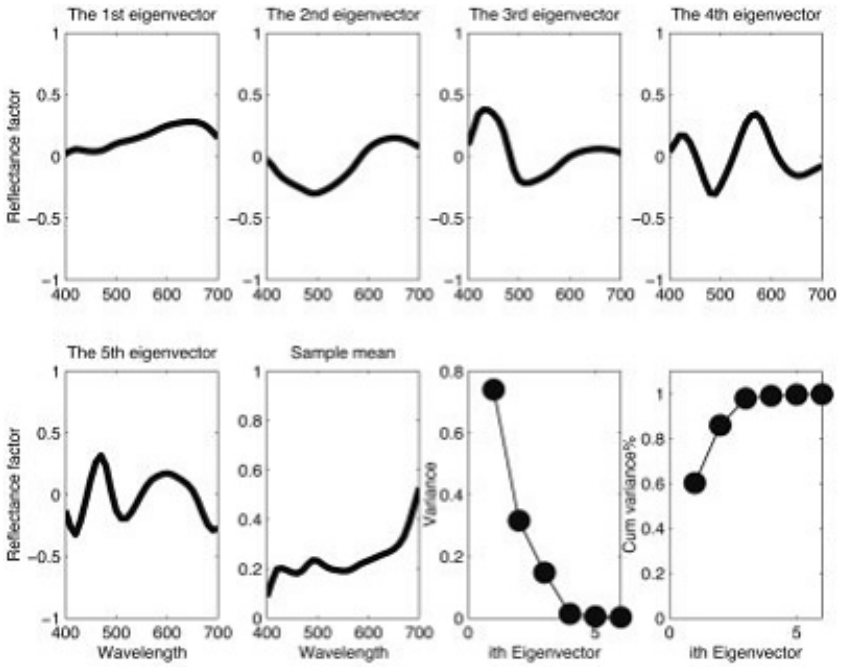


FIGURE A.11: First five eigenvectors, mean and score plots from the reflectance spectra of the dye-diffusion-thermal-transfer printer sample set

Appendix B

Principal Component Analysis (PCA)

Spectral imaging combines conventional imaging and spectroscopy to simultaneously acquire both spatial and spectral information from an object. Determining the intrinsic dimensionality of a hypercube through statistical methods is the process of reducing the number of bands of a spectral image while preserving the main features of the original data.

The Principal Component Analysis (PCA) is one among many statistical methods of *multivariate analysis*, initially described by K. Pearson (Pearson, 1901) but that usually dates back to 1933, with the work of H. Hotelling (Hotelling, 1933). A statistical analysis is said to be *multivariate* when, considering several variables from statistical units, they are analyzed in pairs or in groups of more than two, in order to infer their mutual relationship. Multivariate analysis is such a vast sphere that its boundaries almost coincide with those of statistics.

The problem of the reduction of dimensionality in analyzing multi-response data is, on the one hand the attainment of simplicity for understanding, visualization and interpretation, on the other hand the conservation of enough details for suitable

representation.

From the point of view of description, it must be pointed out that excessive reduction is almost always undesirable, because a meaningful statistical analysis is such only when there has not been severe elimination. Therefore, it should be kept in mind that using statistical approaches for the reduction of dimensionality must lead to the interpretability of the lower dimensional representations.

PCA explores the correlation structure of a sample set in vector form. Data reduction is effected by disregarding the unimportant directions along which the variance of the samples are insignificant. It has been an important and really useful tool in colour technology, as it helped to define tolerance intervals and ellipsoidal regions, to estimate spectral properties of colorants from mixtures, to derive CIE daylight and to reduce huge amounts of data from spectral imaging (Tzeng and Berns, 2005). Data reduction may either derive a set of spectra (*statistical colorants*), whose additive mixtures reproduce the spectral properties of a large set of spectra, or a set of images that emphasize details otherwise difficult to highlight (Bonifazzi et al., 2007).

B.1 Background mathematics of PCA

Whenever one deals with a great three-dimensional (3D) data set, it is fairly straightforward to infer possible relationships between variables, as just three projections along each coordinate axis are readily available for two-dimensional (2D) interpretation. The situation becomes more complicated as the dimensionality of the data set increases, because N-dimensional data points have not the luxury of graphical representation. Here the only strategy for data reduction is to limit the dimensionality of the original set in order that its ND points are projected into a 2D space.

During this dimensional shrinkage the issue of the reduction of distances arises and so the aim of the PCA is to define a specific plane where the distances of data points are

almost unaffected. The guiding principle is the search for the plane that approximates more closely the cloud of data points, i.e. that minimizes the mean of the squares of distances between the original points and their projections on this reference plane. Such projections will be the linear combination of the original variables obtained after multiplying the matrix of the original data with suitable coefficients.

Here is illustrated the way the PCA mathematically operates to reduce the dimensionality of a sample set.

Let the matrix $\mathbf{V}$ $(n \times q)$ represent the sample set, where q is the number of samples, each with n features (the wavelengths, for example). The overall matrix can be thought to be the collection of column vectors $(x_1, x_2, x_3, \ldots, x_n)_i$ with $i = 1, 2, 3, \ldots, q$.

The *mean* vector is defined as:

$$\bar{\mathbf{V}}^T = \frac{1}{q} \sum_{i=1}^{q} (x_1, x_2, x_3, \ldots, x_n)_i \tag{B.1}$$

The sample *variance-covariance* $(n \times n)$ matrix $\mathbf{S}$ is calculated as follows:

$$\mathbf{S} = \frac{1}{q-1} \sum_{i=1}^{q} \left(\mathbf{V_i} - \bar{\mathbf{V}}\right) \left(\mathbf{V_i} - \bar{\mathbf{V}}\right)^T \tag{B.2}$$

A remark must be made about the diagonal elements of $\mathbf{S}$, because they represent sample variances measured along the standard basis vectors, whereas off-diagonal elements are the sample covariance.

In other words, the role of PCA is to make the matrix $\mathbf{S}$ diagonal, in order to get a m-dimensional eigenspace, with $m<n$, of mutually orthogonal vectors (*eigenvectors*) which summarize at best the variance of $\mathbf{V}$ along their directions, as follows:

$$\mathbf{\Lambda} = \mathbf{E}^T \mathbf{S} \mathbf{E} \tag{B.3}$$

with $\mathbf{E}$ $(n \times n)$ the matrix of eigenvectors (columns) of $\mathbf{S}$. As a consequence of B.3, the new sample representations are uncorrelated and the corresponding variances

are maximized along the basis $\boldsymbol{e_1}$,$\boldsymbol{e_2}$,$\boldsymbol{e_3}$,...,$\boldsymbol{e_n}$ in the eigenvector coordinate system. Hence, $\mathbf{\Lambda}$ is diagonal with the eigenvalues L_1,L_2,L_3,...,L_n as nonzero terms, meaning the variances measured with respect to the eigenvectors.

The eigenvalues of $\mathbf{V}$ are the solution of the equation:

$$det\,(L \cdot \mathbf{I} - \mathbf{S}) = 0 \tag{B.4}$$

where L is a generic eigenvalue and $\mathbf{I}$ is the $(n \times n)$ identity matrix.

Once the eigenvalues have been determined, the eigenvectors are found by the B.3 commonly by numerical methods in specialized software.

When $\mathbf{V}$ is distributed in a lower dimensional subspace, then at least one of the eigenvalues is zero. Thus, if an arbitrary sample $\mathbf{V}_{\text{sample}}$ is taken out of the original set $\mathbf{V}$, it can be exactly described by those m eigenvectors having nonzero eigenvalues:

$$\mathbf{V_{sample}} = \sum_{i=1}^{m} b_i \mathbf{e_i} \tag{B.5}$$

The coefficients b_i are calculated by inversion of the B.5, provided that the other quantities are known.

A convenient way to estimate the dimensionality of a sample set is to calculate the *total percent variance* explained by the m eigenvectors:

$$Percent \quad Variance = 100 \frac{\sum_{i=1}^{m} L_i}{\sum_{i=1}^{n} L_i} \tag{B.6}$$

If in a N-dimensional sample set the 99.00 % of total variance is explained by means of m eigenvectors, then an approximate reconstruction for the sample set is:

$$\mathbf{V_{sample}} \cong \bar{\mathbf{V}} + \sum_{i=1}^{m} b_i \mathbf{e_i} \tag{B.7}$$

Although the percentage variance can be undoubtedly taken as a guideline for the selection of the eigenvectors, in colour science applications there are other metrics, like

colour difference, spectral fit, degree of metamerism, that are frequently considered for the assessment of accuracy of reconstruction (Imai et al., 2002).

It is worthwhile to remark that the term *principal components* has been misused; in fact, too often it has been used interchangeably with *eigenvectors* (Tajima, 1998; Vrhel and Trussell, 1992; Haneishi et al., 1997; Baronti et al., 1998; Fairman and Brill, 2004).

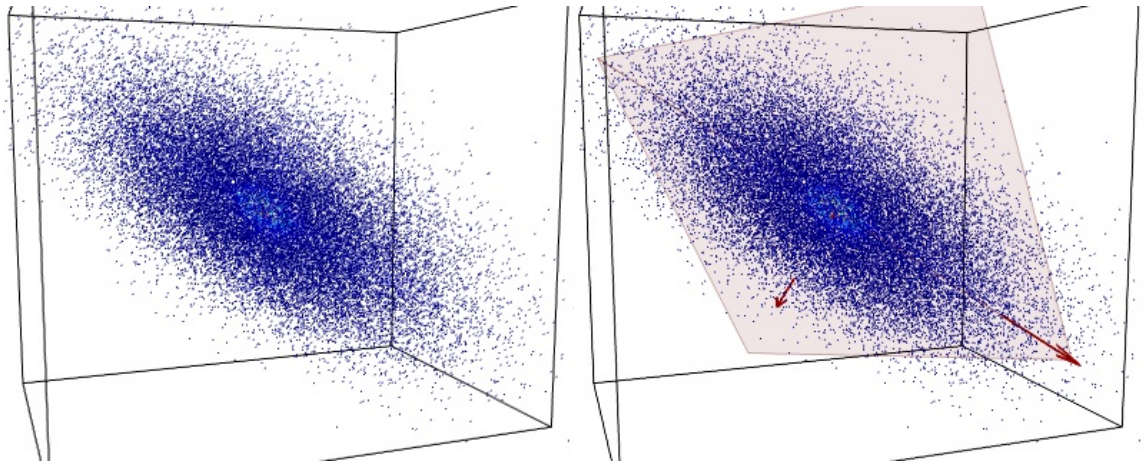
In order to clarify, *principal components* (the coefficients b_i in B.5 and B.7) are the coordinates in the *eigenvector* coordinate system with maximized variance, and as such they are linear combinations of the *eigenvectors*. The *eigenvalues* are the variances of the *principal components*. From B.7, it becomes apparent that the set of the useful eigenvectors (i.e. those which explain the most of the overall variance) is very small and can be expressed by a 3D rotation of an m-dimensional space about $\bar{\mathbf{V}}$. Along the new axes, defined by principal components, the variance is a very rapidly decreasing function:

$$L_1 \geqslant L_2 \geqslant L_3 \cdots \geqslant L_m$$

Hence, the first principal component explains the maximum variance that can be explained through reduction to 1D, whereas the first two principal components explain the maximum variance that can be explained through reduction to 2D, and so forth.

A simple application of PCA

As an illustrative example (Varoquaux, 2012), in Figure B.1(a) is plotted a cloud of points representing a certain 3D very large data set, the simultaneous measurements made by three thermometers positioned at different locations in a room, for instance. If each point depends on the setting of a heater and on the outside temperature, then it is clear that the three set of measurements are not independent one another. Therefore, it is reasonable to find the data mostly on a plane. After the PCA is



(a) (b)

FIGURE B.1: (a) The cluster of points of a 3D very large date set (b) the three eigenvectors (red lines) identified by PCA; highlighted in pink is the plane defined by the first two eigenvectors

applied, it identifies the orthogonal eigenvectors that point in the directions in which the data variance is maximum (Figure B.1(b)). In this particular case, it is expected that the variance is almost completely explained by the first two eigenvectors. Looking at the data set in the plane found by the PCA, its two-dimensional nature is stressed (Figure B.2).

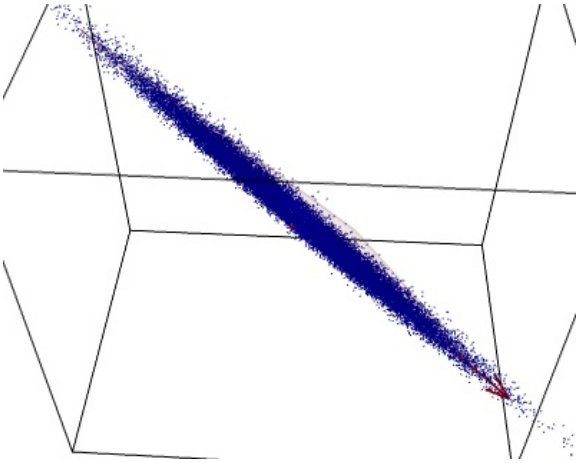


FIGURE B.2: View of the data set in the plane identified by PCA

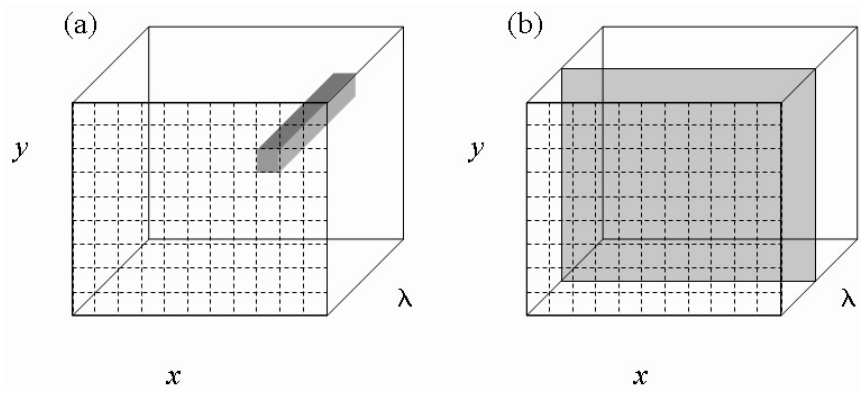


FIGURE B.3: Single spectrum (a) and single image (b) in the hyper-cube

Multivariate image analysis by PCA

The three-dimensional spectral hyper-cube acquired by a spectral imager can be interpreted either as a set of monochromatic images (one for each spectral channel) or as a set of *spectral reflectance factors* (one for each pixel of the colour image) (Figure B.3). Traditionally, PCA is performed on a two-dimensional matrix with the samples of the data set arranged in one dimension and the different variables in the other (Geladi and Grahn, 1996; Bonifazzi et al., 2006). Such matrix is obtained by unfolding each x by y image in the spatial dimension into an $(n \times q)$-dimensional vector. The columns accommodate images taken at a particular wavelength, whereas rows represent the spectral content of a particular pixel.

An interesting application of PCA as a dimensionality reduction method is aimed at assessing the number of "important" bands within multispectral image stacks. A successful application in histopathology (Boucheron et al., 2007) exploited 29 multispectral images (Figure B.4), spaced 10 nm apart, ranging within the visible spectrum from 420 nm to 700 nm. Each image in the stack is a 768×896 pxl grayscale image, thus the overall two-dimensional matrix is 688128×29

After performing the PCA, the logarithmic plot of the normalized eigenvalues (*variance plot*) for each image in Figure B.5 indicates that there is one dominant eigenvalue;

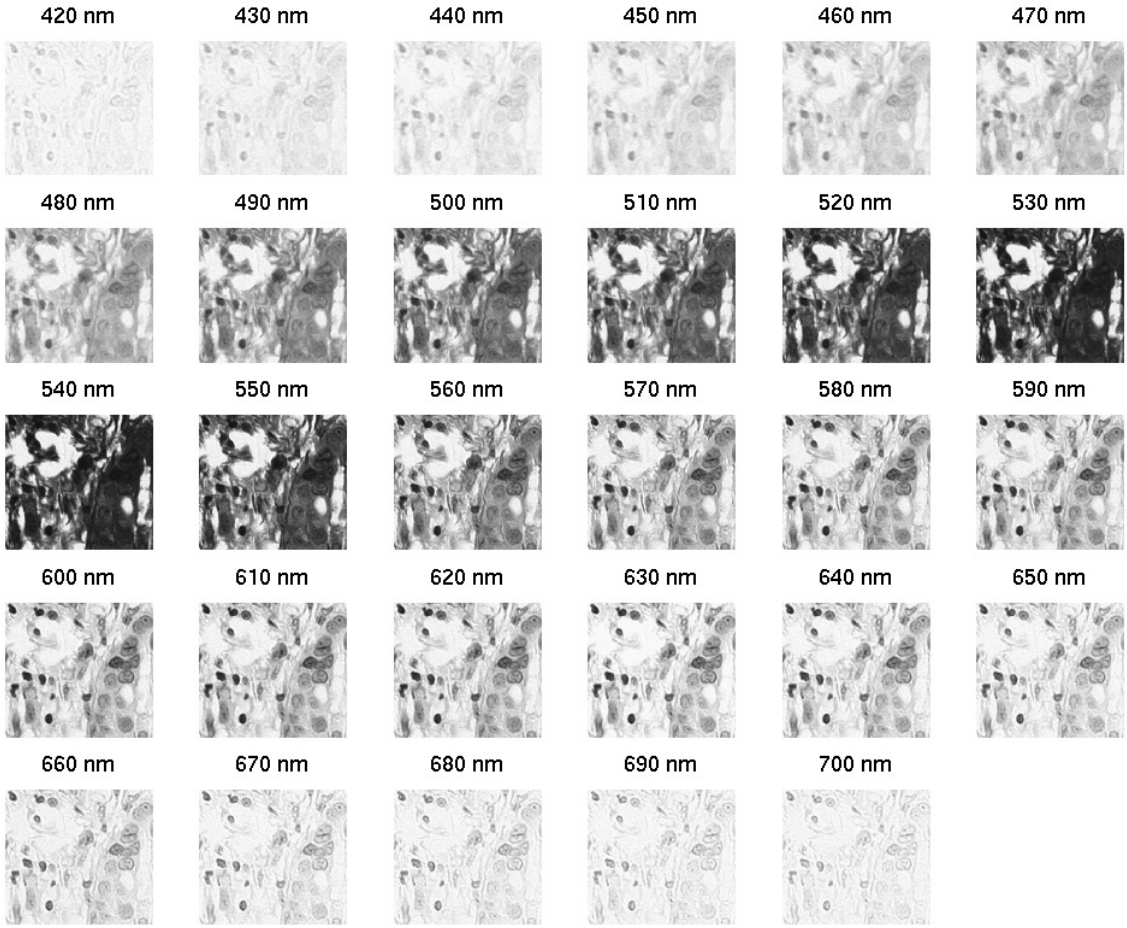


FIGURE B.4: Multispectral stack of 29 monochromatic images from 420 nm to 700 nm

in fact, the second is nearly one-tenth the value of the first and after that a precipitous drop-off is encountered. So the most interesting dynamics occur only on the first two dimensions which are as many as the stains used in the histopathology imagery, as expected by authors. In short, the PCA indicates the presence of two eigenvalues that account for the 97 % of the variation in the data. Looking at Figure B.6, the projection of the multispectral image (A) on the first eigenvector emphasizes nuclear regions (B) (stained with one substance), whereas the projection on the second eigenvector (C) highlights connective tissue and cytoplasm (stained with the other substance); the third projection (C) and the following are of no straightforward interpretation,

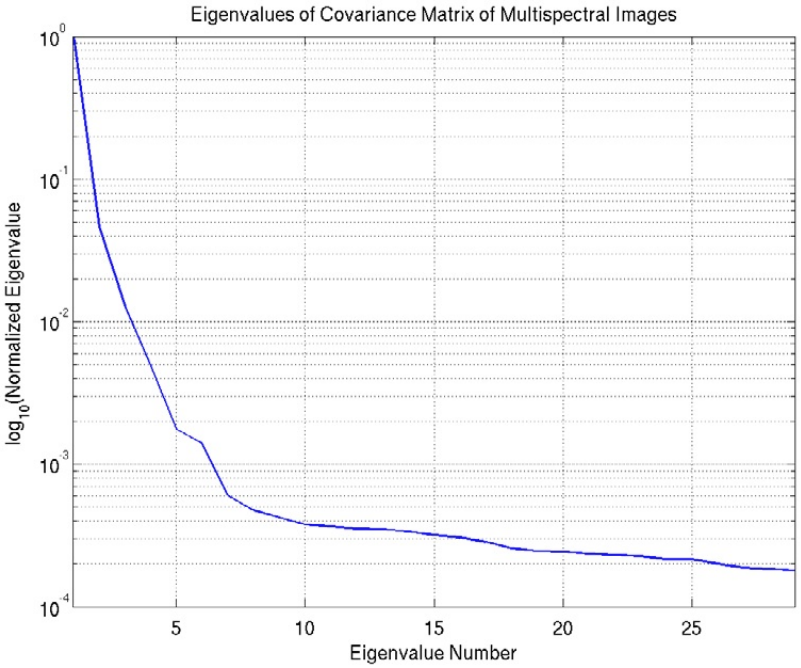


FIGURE B.5: Logarithmic plot of the normalized eigenvalues of multispectral imagery in Figure B.4

because they have no correlation with the tissue stains. As a consequence, there are two information bearing bands in the imagery for the nuclear classification task. An important note about the object of PCA: it is a compression method that often gathers data points in clusters on two-dimensional plots, but in no way it can be accounted as a classification tool (at most as the basis of a classification).

The PCA is just a tool to reduce multidimensional data to lower dimensions, while retaining most of the information.

*B.1.1 Singular Value Decomposition (*SVD*)*

Theorem (Singular value decomposition)**.** *Let* $\mathbf{V}$ *be a real matrix* $n \times q$*, with rank* m*. There exist an* $n \times m$ *matrix* $\mathbf{U}$ *such that* $\mathbf{U^T U} = \mathbf{I_m}$*, an* $q \times m$ *matrix* $\mathbf{X}$ *such that* $\mathbf{X^T X} = \mathbf{I_m}$*, and an* $m \times m$ *diagonal matrix* $\mathbf{W}$ *with positive diagonal elements, such that* $\mathbf{V} = \mathbf{UWX^T}$*, where* $\mathbf{I_m}$ *is an* $m \times m$ *identity matrix (Magnus and Neudecker,*

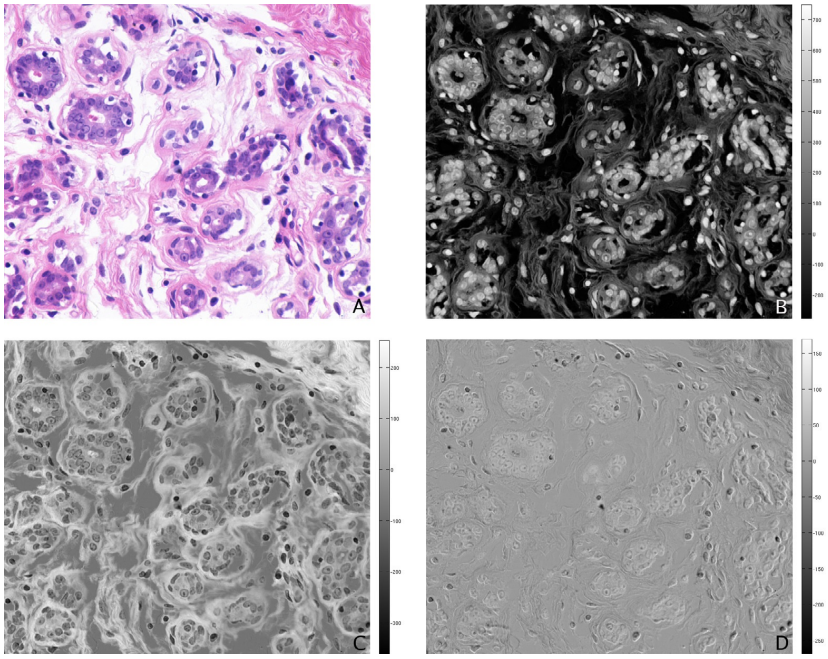


FIGURE B.6: (A) Histopathology multispectral colour image and its projection on the (B) 1st (C) 2nd (D) 3rd eigenvector

1988)

An important property of the SVD is that it provides a computationally efficient method of finding the principal components.

In fact, once the matrixes **U**, **W** and **X** satisfying the above theorem are found, then **U** and **W** gives us the eigenvectors and the square roots of the eigenvalues of the matrix $\mathbf{VV^T}$, and hence the principal components and the variances for the sample covariance matrix (Hardeberg, 2001).

Appendix C

Induced-Transmission filters

In 1956 two American scientists at the Bausch & Lomb optical company in Rochester (New York), P.H. Berning and A.F. Turner, demonstrated both theoretically and experimentally, on the grounds of classic electromagnetic theory, that metallic absorbing layer can enter the stack of dielectric multilayer systems in order to provide narrow-band interference filters, with really attractive and interesting possibilities (Berning and Turner, 1957). In fact, a reasonably thick film of certain metals could be made to transmit a high amount of energy of a particular wavelength, if adequately surrounded by combinations of interference films. Such filters are therefore called Induced-Transmission.

Induced-transmission filters, with respect to all-dielectric filters:

1. are equally free of troublesome side-bands;
2. generally exhibit lower peak transmittance for a given wavelength;
3. generally exhibit wider transmittance band for a given wavelength;
4. have sharper transition between pass-band and stop-band, and better rejection

in the stop-band.

The theory of induced-transmission is based on the concept of *potential transmittance* (Ψ), namely the rate of transfer of energy across unit area of the last boundary of the system divided by that of the first boundary. In different words, potential transmittance can be considered as the transmittance that the system would perform in the case of null reflectance at the first surface of incidence.

The idea behind the possibility to induce transmission at a particular wavelength can be considered as a problem of maximizing the potential transmittance, and is therefore part of a much wider problem of designing multilayer coatings with absorbing layers to perform specified optical functions.

In the design of optical coatings, any solution to a problem must conform to the fundamental laws of physics[1]. The conservation of energy principle implies that, if T, R and A refer to the transmittance, reflectance and absorptance of a multilayer system, at the front (subscript 1) and rear (subscript 2) surfaces (Figure C.1) the following relationships hold true:

$$T_1 + R_1 + A_1 = 1 \tag{C.1}$$

$$T_2 + R_2 + A_2 = 1 \tag{C.2}$$

Moreover, the second law of thermodynamics (Holloway and Lissberger, 1969; Landau and Lissberger, 1972) ensures that $T_1 = T_2$, hence from C.1 and C.2:

$$R_1 + A_1 = R_2 + A_2 \tag{C.3}$$

If the multilayer system were perfectly symmetrical in its structure of layers, then additional constraints would be introduced:

$$R_1 = R_2 \tag{C.4}$$

$$A_1 = A_2 \tag{C.5}$$

[1] Throughout the text the discussion is restricted to perfect systems, namely ones that consist of homogeneous and isotropic materials with plane-parallel boundaries

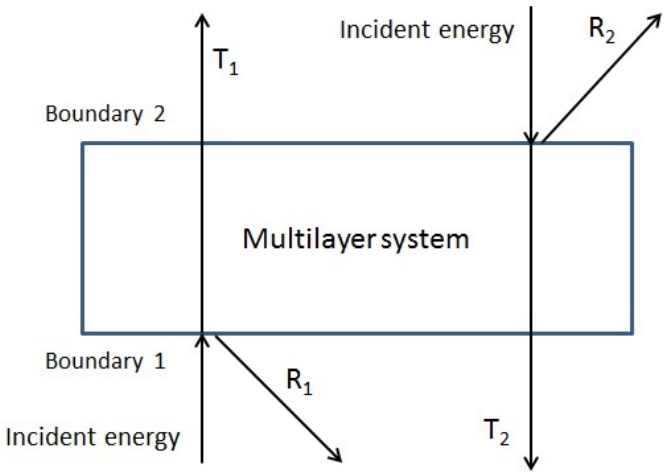


FIGURE C.1: Transmittance and reflectances of a multilayer coating

which retain their validity even without the requirement of perfect symmetry, as occurs in induced-transmission filters. Furthermore, with such induced-transmission filters another simplification holds true (Lissberger, 1981) in limited conditions of symmetry:

$$R_1 = R_2 = 0 \tag{C.6}$$

In this case there is symmetry in the basic optical functions, not necessarily in the structure of the layers.

The potential transmittance Ψ of a layer or multilayer system is:

$$\Psi = \frac{T}{1-R} = \frac{T}{T+A} = 1 - \frac{A}{1-R} \tag{C.7}$$

where T, R and A indicate the basic optical functions of the layer or of the multilayer system.

The potential transmittance function has some interesting properties:

1. it can assume different values for radiation crossing the coating in opposite directions;

2. $0 \leqslant \Psi \leqslant 1$; in particular $\Psi = 0$ for an optically opaque absorbing layer or system and $\Psi = 1$ in the case of perfect all-dielectric materials;

3. in the case of a multilayer system of N layers, it equals the product of the potential transmittances of the constituent layers[2]: $\Psi_{system} = \prod_{j=1}^{N} \Psi_j$;

4. for a subsystem of a multilayer coating, the potential transmittance is independent of the parameters of that part of the whole system on the side of incidence of the subsystem;

5. Ψ_{j} depends explicitly only on the thickness and complex refractive index of the j^{th} layer;

6. Ψ is continuous across a single boundary.

Supposing to insert a thin metallic layer between two all-dielectric multilayer subsystems (Figure C.2), it results that the absorption of metal can depend very markedly on the optical environment of that layer. Maximizing transmittance of the metallic film by the appropriate choice of environment is described as *inducing transmission*. When light is incident from the top of the whole system, appropriate design of subsystem II can maximize potential transmittance Ψ_1:

$$\Psi_1 = \Psi_{max} = \frac{T}{1 - R_1} \tag{C.8}$$

If now subsystem I is designed to work as an anti-reflection, then $R_1 = 0$ and so $T = T_{max} = \Psi_{max}$ from C.8, considering that the architecture of subsystem I does not affect Ψ_1.

Designing both subsystems to meet the criteria stated above, then $\Psi_2 = \Psi_{max}$ and

[2] The potential transmittance of the j^{th} layer, Ψ_{j}, can be expressed by the ratio of the average numerical magnitude of the Poynting vector at the j^{th} and at the $(j\text{-}1)^{\mathrm{th}}$ surface. This is because it determines the ratio of net luminous flux transmitted past the j^{th} boundary to that transmitted past the $(j\text{-}1)^{\mathrm{th}}$ boundary

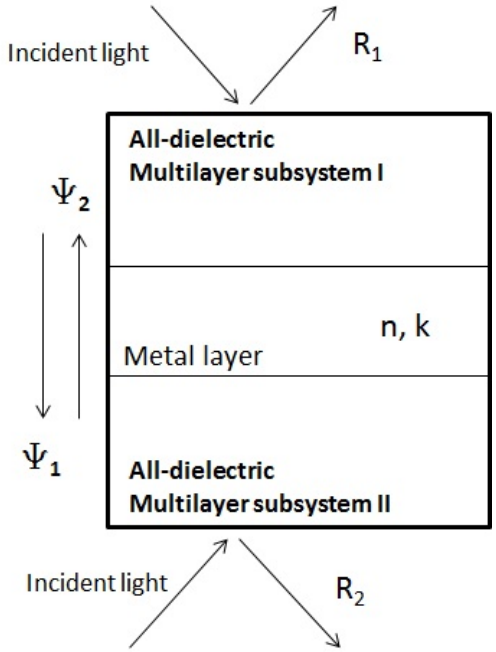


FIGURE C.2: Optical functions of a coating containing a single metal layer

$R_2 = 0$ follow, simultaneously.

It is quite apparent that inducing transmission introduces elements of symmetry into the system as a whole, even though it is only an approximation since the presence of the substrate in the real case destroys full-symetry.

For the system to have maximum transmittance, rather than symmetry about the metal layer, the complex amplitude reflectance at the boundaries of the metal must be the same and the values of its amplitude and phase, that maximize the potential transmittance of the metal as well as the value of the maximum potential transmittance, are unique and depend on the optical parameters of the metal layer alone (Berning and Turner, 1957; Landau and Lissberger, 1972).

Such values of amplitude and phase can be achieved in a variety of ways, but in practice they are obtained by a combination of a dielectric quarter-wave system with a single dielectric non quarter-wave layer (*phase-adjustment* layer) in contact with the metal. There are some varieties of dielectric combinations which satisfy the

requirements for inducing transmittance at a given wavelength λ. In the visible range the only metal that can be used is silver ($\hat{n}_{Ag}$=0.055-i3.32 (Macleod, 2001)).

As a rule of thumb, inducing transmission at λ implies a marked extinction of the potential transmittance on both sides; this is a desirable behavior for bandpass filters because ensures high peak transmittance and good refection at the same time.

Appendix D

Lookup Tables (LUT) in colour calculation

Electronic trichromatic imaging devices, as can be digital cameras, film and paper scanners, colour monitors and printers, are not closed systems, in the sense that they require user intervention at some stages of the workflow (Figure D.1) from the scene to the reproduction. In fact, they are not manufactured to the same system specifications and therefore demand proper calibration to ensure that they work together for good reproduction of colour images. Two stages of calibration are mandatory, one for the input and one for the output devices. The former prompts the output of the inputs device for successive processing algorithms' manipulation, whereas the latter ensures optimal colour reproduction of the image by the output device.

Since each device outputs numerical data, it must be understood what these numbers

FIGURE D.1: Synthetic workflow for a colour imaging application

represent. In order to ensure consistent tone and colour reproduction, the relationship between the colour in the original picture and the numbers in the output device must be known. This is the task of *device calibration*, which involves colour measurements on each device.

D.1 Colorimetric calibration

Basically three are the calibration steps usually performed in colour imaging applications:

1. *device adjustment*: sets-up the device at the preferred operating conditions and is usually performed by skilled personnel, as it requires professional knowledge

2. *target measurement*: is the stage of measurement of some quantities of the output device in response to some input values

3. *signal transformation*: is a set of processing aimed at producing colorimetrically correct output

Sometimes the term *calibration* is found in the literature (Brown, 1993; Martens and Naes, 1992) to mean the three stages above as a whole, sometimes only the second and third, some other times the third alone.

D.1.1 The calibration of input and output devices

For an input device a reference set of colour coordinates of some physical objects must be available to serve as reference colour set of data points. In fact, the images of these target objects are captured under a known illuminant. Most of the times such targets are found in form of board like the MACBETH COLORCHECKER™ [1]. A snapshot is taken to establish the correspondence between the measured colours on the

[1] `http://www.xrite.com`

targets and the responses from the camera. The final goal is to determine the colour metrics, in a device-independent colour space, of an unknown object, corresponding to a future camera digital response of a pixel of the image of that object, expressed in device-dependent coordinates (say, in terms of (R,G,B)).
A number of difficulties arise in the calibration of an imaging device:

- *spectra sampling*: a trichromatic imaging device undersamples the spectral signals whenever its spectral responsivities are not a linear combination of the human cone sensitivities, in which case in no way it can be calibrated to see colour as an average human observer

- *scene balance*: colour differences may arise when the scene illuminant differs from that employed in the calibration, so an adjustment is required either for the calibration table or for the captured image. In video imaging applications this is known as the *white balance* problem

- *colour interpolation*: it is not difficult to figure out that just an extremely small number of possible colours can be accommodated in a colour chart, therefore a model must be found that is forced to pass through all measured data points on the calibration chart

Figure D.2 displays the steps necessary for input device calibration. Initially, a linearization of the sensor response is carried out through a conversion into a metric that is linear with the image irradiance; this task is accomplished with a one-dimensional lookup table[2] (1D-LUT), for every single colour channel. Figure D.3 shows the linearization curves of a scanner, each for one colour channel, with the scanner code values taken on a colour chart versus the reflectance factor of the neutral

[2] A lookup table is an array that replaces runtime computation with a simpler array indexing operation. LUTs are used extensively to validate input values by matching against a list of valid (or invalid) items in an array. In image processing, a LUT is used to transform the input data in a more desirable output format

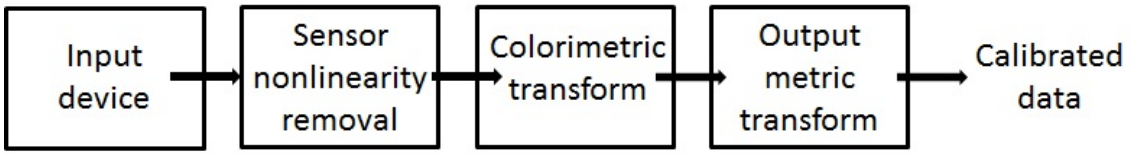


FIGURE D.2: Synthetic workflow for input device calibration

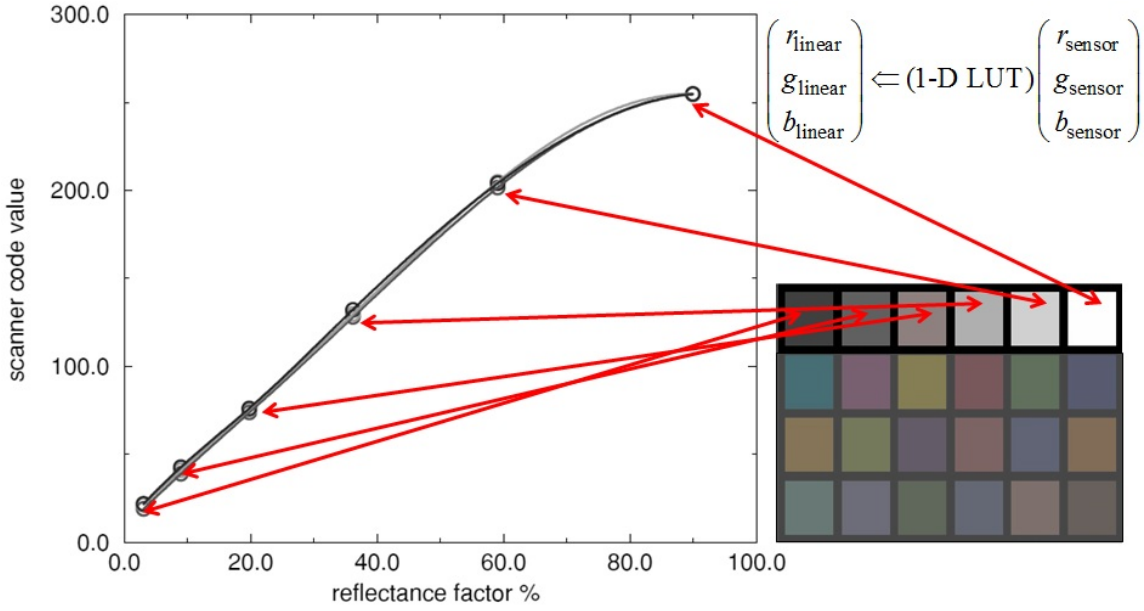


FIGURE D.3: Linearization curves of the chromatic channels of a scanner and the MACBETH™ COLORCHECKER with the neutral patches highlighted

patches on the same chart. A 1D-LUT ensures the calculation of the linear (r,g,b) coordinates.

After that, image colour data are transformed to a device-independent colour space, such as CIEXYZ or CIELAB, with either a 3×3 matrix or a 3D-LUT.

This colour conversion matrix is found through the *pseudoinverse* calculation, which is a common practice when dealing with overdetermined problems, as is of course the case with the number of colour patches in a colour chart.

Let $\mathbf{X}$ $(3 \times k)$ represent the matrix of the linearized (r,g,b) coordinates of a set of k

colour patches

$$\mathbf{X} = \begin{bmatrix} r_1 & \dots & r_k \\ g_1 & \dots & g_k \\ b_1 & \dots & b_k \end{bmatrix}$$

and let $\mathbf{Y}$ $(3 \times k)$ represent the matrix of the corresponding tristimulus values

$$\mathbf{Y} = \begin{bmatrix} X_1 & \dots & X_k \\ Y_1 & \dots & Y_k \\ Z_1 & \dots & Z_k \end{bmatrix}$$

The *pseudoinverse* calculation is defined by the matrix of the coefficients $\mathbf{b}$ (3×3) that minimize the sum-of-square error between measured and estimated tristimulus values

$$\mathbf{b} = \mathbf{Y}\mathbf{X}^{\mathbf{T}}(\mathbf{X}\mathbf{X}^{\mathbf{T}})^{-\mathbf{1}} \tag{D.1}$$

where

$$\mathbf{b} = \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \beta_{1,3} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} \\ \beta_{3,1} & \beta_{3,2} & \beta_{3,3} \end{bmatrix}$$

and that defines the relationship between measured $\mathbf{X}$ and estimated values $\hat{\mathbf{Y}}$ (Berns, 2000)

$$\hat{\mathbf{Y}} = \mathbf{b}\mathbf{X} \tag{D.2}$$

Finally as a result of the calibration, data are ready to be expressed in the desired output metric (Figure D.4).

Besides scene balance and colour interpolation problems, in the calibration of an output device (Figure D.5) other drawbacks may arise:

- *colour matching stability*: inks and dyes used in colour hard-copy output devices perform metameric colour matching to the original colour stimuli, under a given viewing illuminant. Therefore a strong dependence on the illuminant exists, which prevents the reproduced color from matching those of the original.

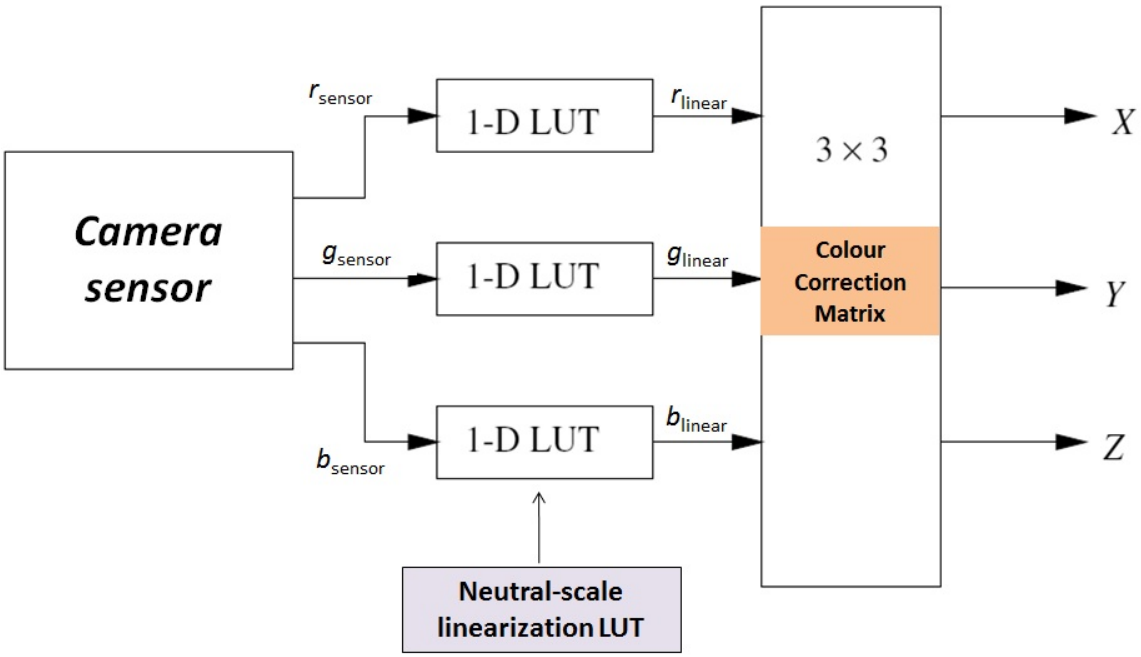


FIGURE D.4: Colour data processing flow for the output of a trichromatic digital camera

The problem is only made less severe by an adequate choice of broadly reflecting or transmitting inks or dyes, but not suppressed, and this limits the range of reproducible colours

- *colour gamut*: a colour image that has to be rendered on an output device requires a systematic procedure (*gamut mapping algorithm*) to handle the out-of-gamut colours, namely those colours that cannot be reproduced by the output device

- *media-dependency*: calibration of hard-copy devices is a function of the specific output medium and how it is processed

- *viewing condition*: the viewing of a hard-copy reproduction is affected by the difference in the measurement and viewing geometries, by the luminance level difference between the illumination of the scene and that of the viewing environment and by the surround of the image, which affects the appearance of

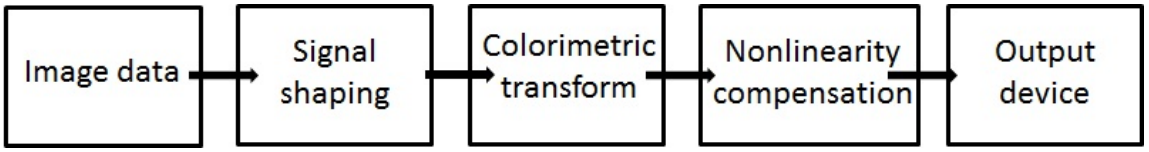


FIGURE D.5: Synthetic workflow for output device calibration

colour

Usually a signal shaping on the image data is performed to account for the high nonlinearities between the incoming image data and the input data expected by the output device. This is accomplished by using 1D-LUTs that convert the input image data into a shape that is more nearly linear with respect to the output metric. Either a 3D-LUT or a 3×3 matrix can convert the image data into the output device's colour primaries and, once in this state, the colour data are pre-distorted to compensate the nonlinearities of the output device or medium. LUTs are both affordable tools and the preferred method for device calibration, returning more accurate calibration results than device models, mostly in the case of complex behaviors. Moreover, the availability of automated measuring instruments has made exclusive the use of large and dense LUTs for the calibration of imaging devices.

However relying on large 3D-LUTs to handle every type of device calibration is an impractical job indeed, since the table needed for a nonlinear device would be unacceptably large. The solution is a hybrid approach employing a device model at a preliminary stage, in order to remove most of the nonlinearity before the 3D-LUT is constructed.

D.2 Building a three-dimensional lookup table

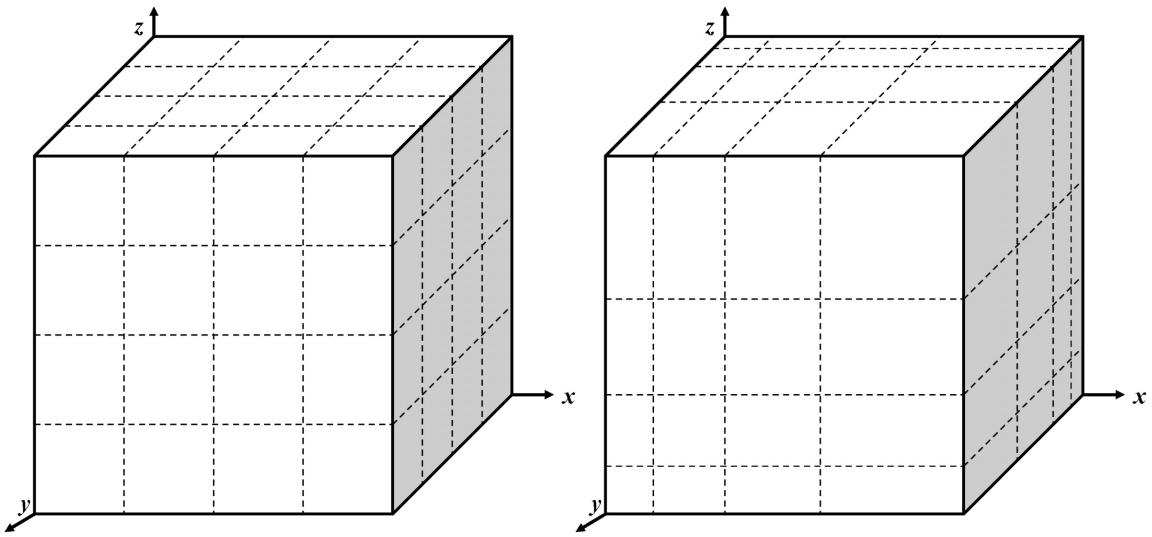
Here is a list of the parts forming the structure of a uniform 3D-LUT:

- *Partition* or *Packing*: the source space is partitioned and becomes populated with sample points for the purpose of building a LUT. For a uniform LUT

an equal sampling step is assumed along each axis of the source space, which yields (n-1)^3 elementary cubes and n^3 lattice points, where n is the number of levels. Generally, every single lattice point of the source space corresponds to a colour patch, thereby generating the destination colour specifications by measurement. The corresponding values between equal patches in both spaces are stored in the LUT

- *Find* or *Extraction*: a search is performed to select the lattice points used to locate correctly an input point. Well-packed spaces perform simple searches, whereas unequally spaced packing are more time consuming because comparisons are needed to locate the nearest lattice points. Colour calculation of nonlattice points could demand further comparisons, depending on the particular interpolation technique

- *Computation* or *Interpolation*: the input signals and the extracted lattice points are used to calculate the destination colour specifications for an input point. Interpolation techniques are performed through mathematical computations that employ either geometrical relationships or cellular regression. Those of the former kind rely on the various ways of subdividing a cube and require inequality comparisons to find the subdivided structure where the point falls. The differences among various geometrical techniques (bilinear, trilinear (Pugsley, 1975), prism (Kanamori et al., 1992), pyramid (Franklin, 1982), tetrahedral (Iri, 1967) and their derivatives and extensions (Sakamoto and Itooka, 1981; Kang, 1997)) is in how every single technique slices the cube. The computations based on the cellular regression do not perform searches within the cube; they apply regression locally in the lattice for the sake of accuracy enhancement, thereby providing additional flexibility and accuracy not offered by geometrical interpolations (Kang, 1997, 1995a,b)



(a) (b)

FIGURE D.6: Uniform (a) and nonuniform (b) five-levels 3D partition (Kang, 2006)

Compared to other transformation techniques, with 3D-LUTs the colour space can be further divided by increasing the sampling rate, to improve accuracy, and if such sampling is high enough the linear interpolation can be successfully employed to compute a nonlattice point with excellent approximation.

When nonlinear colour conversions are involved, in spite of the source space regularity an uneven and irregular destination space results; as a result, efforts have been made to devise techniques working on nonuniform data (Kasson et al., 1993; Bell and Cowan, 1993; Allebach et al., 1993; Chang et al., 1994). Figure D.6 shows both equal and unequal sampling of an input colour space.

Nonuniform 3D-LUT are used extensively, for their versatility, for their enhanced interpolation accuracy and because they make the error more uniformly distributed. Nonetheless, the most serious drawback when using a 3D-LUT is the inefficient use of the available colour space. Solutions have been proposed that optimally allocate grid points according to the characteristics of the particular function (Allebach et al., 1993; Chang et al., 1994).

Appendix E

CCD array architecture

Array architecture of CCD (*Charge-Coupled Device*) (Millman and Grabel, 1987) is heavily dependent on the particular application. Full-frame and frame transfer devices are more frequently encountered in scientific applications, whereas interline transfer devices are well-established in consumer camcorders and professional television systems. Industrial applications require mostly linear arrays and progressive scan based solutions. Although colour cameras are becoming increasingly attractive, monochrome cameras are often the preferred choice in scientific and industrial applications.

E.1 The physics of CCD sensors

Digital cameras employ solid-state photosensitive elements (*photosites*) made of silicon for applications in the visible range. Hence, the key element in a digital camera is charge-transfer optoelectronic component. The inventors of the CCD were Boyle and Smith in 1970 (Boyle and Smith, 1970; Amelio et al., 1970); since then a considerable amount of literature was produced on the physics, fabrication

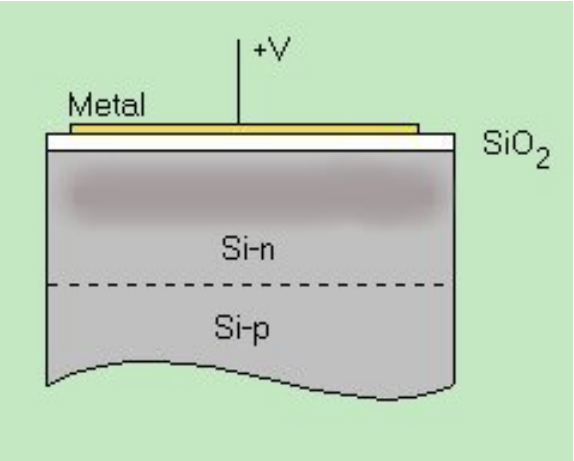


FIGURE E.1: Schematic arrangement of layers forming a CCD stack

and operation of the CCD. It consists of a set of light-sensitive elementary units (*photosites*), regularly arranged in a matrix fashion along rows and columns. Their shape can be either linear, square or rectangular and can accommodate a variable number of photosites.

Every single photosite usually consists of one electrode (*gate*) or more, driven by synchronized electric signals. Most commonly the number of gates in scientific instruments is three; the photosensitive element is driven by the central gate, whereas the other two gates are shielded and activate the transfer of the accumulated charges. A CCD sensor is a stack made of a p-type silicon substrate with a n-type silicon layer grown on it. A thin layer of silicon oxide follows, with either a metal electrode or a heavily doped poly-crystalline silicon layer on top (Figure E.1).

The electrons are distributed in the substrate and in part of the n-type silicon with a minimum energy, whose particular value is a function of both the silicon doping and the thickness of the oxide layer. A positive voltage applied to the electrode increases the depth of potential energy in the region beneath the electrode itself, whereas a negative voltage has the opposite effect.

If different electrodes are deposited on the oxide layer and then positively polarized,

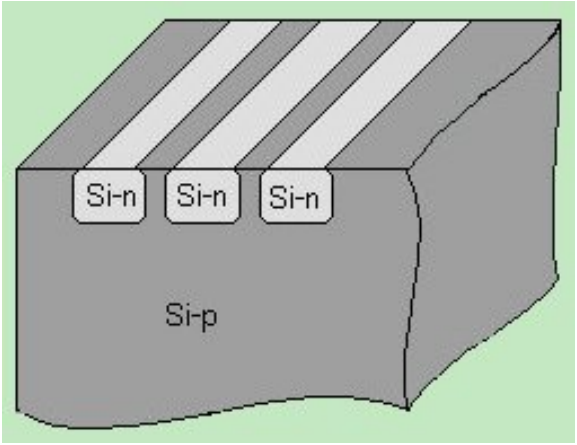


FIGURE E.2: Sketch of n-type silicon strips obtained on a p-type silicon substrate

then the potential energy profiles that arise beneath can trap electrons within wells; trapped electrons can then be released by switching-off the polarization. Hence, with proper modulation of the voltage charges can be moved among different wells.
A matrix CCD can be thought of as a sequence of side-to-side independent linear CCDs on the same substrate. In order to have independent channels of potential energy, the n-type silicon layer is heavily doped to turn it into a p-type (by ion implantation, for example). The result is an array of n-type silicon strips embedded in the p-type silicon substrate (Figure E.2).

The substrate is a barrier (*channel stop*) for the electrons, whose efficiency can be further increased by growing a layer of silicon oxide with suitable thickness. This makes the electrostatic potential within this region almost unaffected by variations in gate voltages while the CCD is working; as a consequence, the separation of channels is always ensured (Dereniak and Crowe, 1984).

E.1.1 The charge transfer process

The charge transfer process is illustrated in Figure E.3 in the simple case of a linear CCD with three gates. At instant t_1 central electrodes of each photosite are supplied a positive voltage in order to form a well of potential energy beneath their

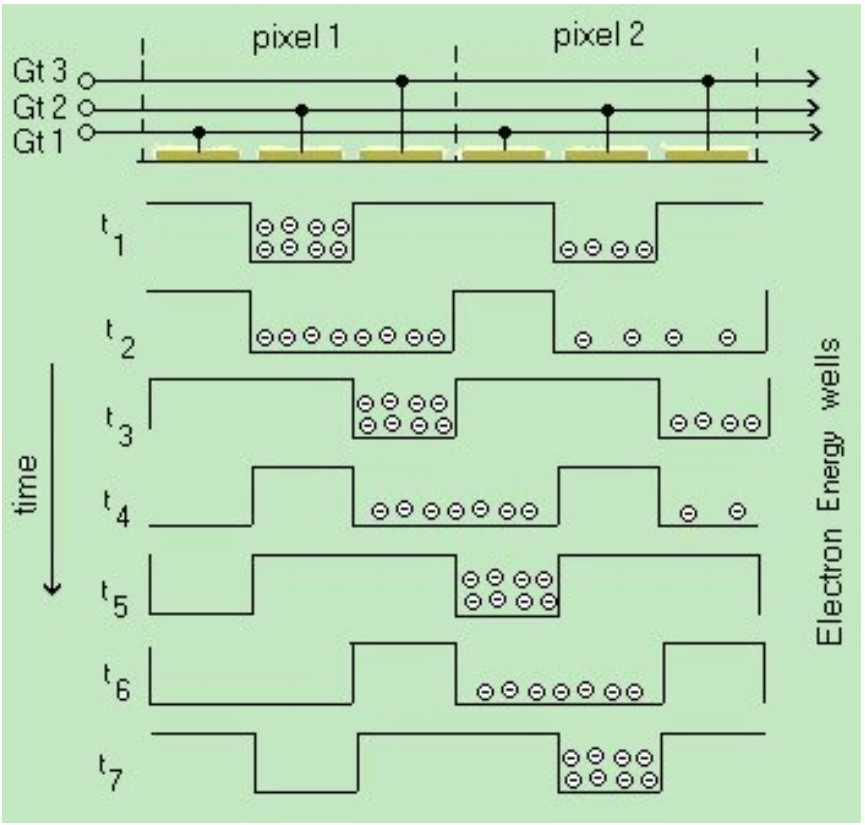


FIGURE E.3: Schematic illustration of the charge transfer process for a linear CCD

area. Side electrodes are not biased. Whenever an incident photon is absorbed in the material underneath the electrode, then an electron-hole pair is generated due to the photoelectric effect. Such a pair is broken under the effect of the inner electric field, so the electron reaches the bottom of the well, whereas the hole vanishes deep in the substrate. After separation, electrons and holes cannot recombine anymore; therefore the higher the number of photons absorbed, the higher the number of electron-hole pairs generated that are gathered in the well. The accumulated charge is a function of the intensity of the incident light, of its spectral composition and of the exposure time of the photosite.

When the exposure to light ends at time t_2, the voltage of the gate at the right side of the photosite is increased to enlarge the well underneath and the electrons

redistribute homogeneously in the well.

At time t_3 the central gate voltage is suppressed; it follows that the well is smaller and is located beneath the gate at right. The net effect is a shift at right of the accumulated charge.

If at time t_4 the voltage of the left gate of an adjacent photosite is increased, the well is again enlarged and the suppression of the right gate voltage (at time t_5) of the first photosite will confine the electrons under the first electrode of the second photosite. It is quite apparent that, in the case of a row of adjacent photosites, a suitable synchronization of the three gate voltages will make the charge travel accordingly. Such a transfer process is extremely efficient since it can reach the 99.999 % within a single step.

E.2 The architecture of CCD sensors

The readout scheme of a CCD depends on the role played by the CCD itself. Different technical solutions exist according as either the spatial resolution, the ability to reproduce fine details of an image or the image acquisition speed is the goal. Generally, a matrix with many photosites ensures high spatial resolution but entails long readout times. There are basically three kinds of technical solutions whose main characteristics are described in the following subsections.

E.2.1 Full-frame transfer CCD

In Figure E.4 light sensitive vertical CCD columns are separated by channel stops in between. A transfer gate sets the flow of charge packets to the linear horizontal CCD, which in turn transfers such packets to the output amplifier. The signals returned by the amplifier are sent to the frame grabber between the digital camera and the computer. Both the transfer gate and the CCD are shielded while the light excites the photosites, thus permitting to accumulate the charge. After the exposure

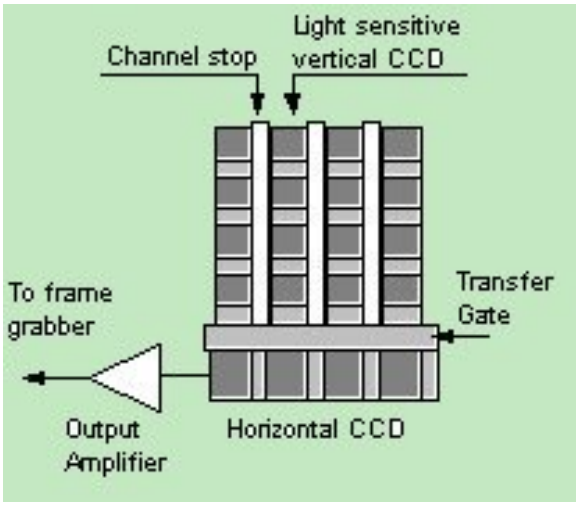


FIGURE E.4: Schematic illustration of the architecture of a full-frame transfer CCD

the readout process starts. The transfer gate permits the discharging of the row of photosites that is close to the horizontal CCD and, at the same time, all the other rows shift away one position downwards. The process stops for as long as it takes to sequentially transfer the packets from the horizontal CCD to the output amplifier. The sequence of operations then restarts with the second row until the charge packets are entirely transferred.

Once the transfer is complete, a new exposure process can be started. The transfer charge process along columns and the horizontal CCD is effected with the procedure described in E.1.1. In this case, the photosites act first as light sensors, then like shift registers as well as the horizontal CCD.

For a 1000×1000 square matrix, less than 100 ms are requested for a charge transfer but it allows a high spatial resolution, since it exploits the full matrix area. It follows that, because the time interval between two consecutive exposures is not lower than 100 ms, the highest acquisition speed will be 10 fps.

During the charge transfer process the photosites are exposed to light and so a distortion of the image can arise, due to the longer exposure time of the photosites

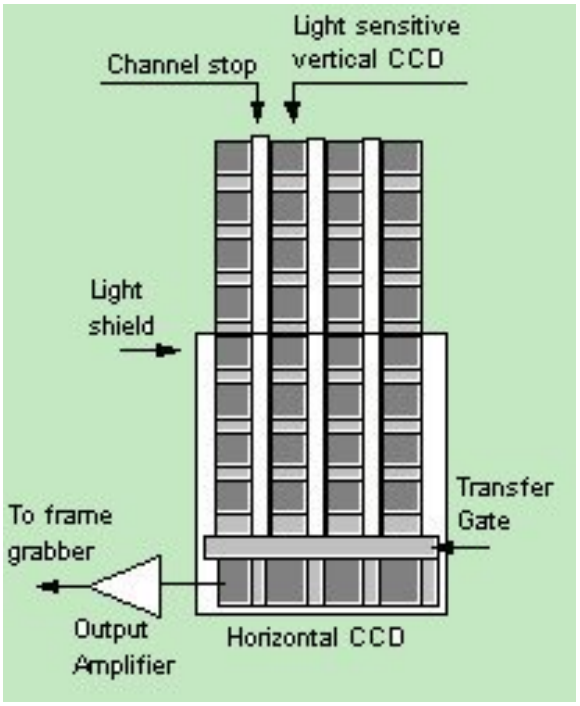


FIGURE E.5: Schematic illustration of the architecture of a frame transfer CCD

discharged at the last. Hence, it becomes necessary to employ an external light shield to obscure the CCD during the transfer process.

E.2.2 Frame transfer CCD

Here the matrix is divided into two equal halves; the former is sensible to light and serves for light capture, the latter works as a memory of the first half and is located between the first half and the horizontal CCD (Figure E.5). After exposure, each column of charge packets is scrolled down to the memory matrix in a time interval of 0.1 ms to 0.8 ms. Then it is transferred to the frame grabber, with the modality discussed in E.2.1.

It must be pointed out that a new exposure can start only after the charge transfer to the memory matrix process is complete. The readout of the memory matrix is effected during the new exposure of the upper matrix. Hence, the time lag between

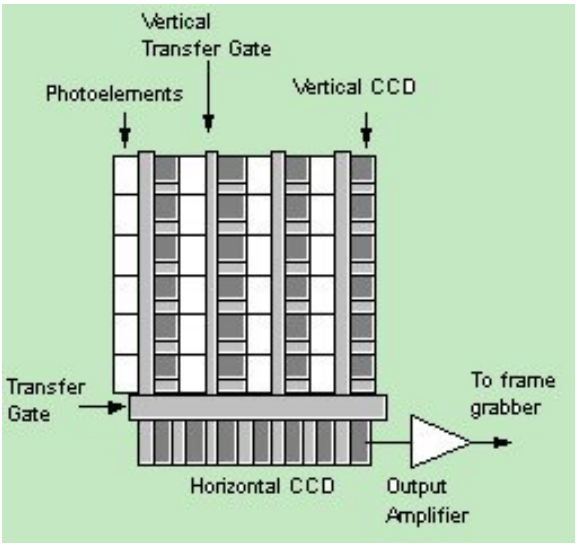


FIGURE E.6: Schematic illustration of the architecture of an interline transfer CCD

two subsequent exposures is less than 1 ms whereas the time the frame grabber takes to read the full matrix is the same as for the full-frame transfer architecture (i.e. 100 ms for a 1000×1000 matrix).
However, the higher speed of this architecture is achieved through a double-matrix that, obviously, entails the need of an external light shield.

E.2.3 Interline transfer CCD

In this particular CCD architecture, for every column of photosites of the matrix a vertical shift register (vertical CCD) exists with a vertical transfer gate in between, both obscured to light by an opaque metal shield (Figure E.6). After the exposure, the charges are transferred to the shift register in a single run, then are passed to the horizontal CCD and lastly to the frame grabber, as for the other architectures. With this kind of CCD a new exposure can start soon after the charge package have been transferred to the vertical shift register, namely after a very short time (nearly 1 µs) with a considerable reduction of the time interval between two consecutive exposures. The readout time is unchanged with respect to the other architectures.

The undeniable benefit in terms of speed is achieved at the cost of a considerable reduction in the photosensitive area. In order to increase the area exposed to light, it is mandatory the use of an array of microlenses (Holst, 1998); moreover, with such an architecture there is no need of electromechanical shields (shutters).
Commonly, two subcategories of interline transfer CCD sensors can be identified:

- *Interlaced interline transfer* CCD: normally employed in video-cameras and television systems, particularly with the american television standard NTSC. Charge transfer occurs through the readout of odd rows during an exposure cycle, whereas even rows are processed during the next exposure cycle. Therefore, the image on a screen is obtained through optical superposition of two fields (odd and even rows) acquired with a known delay

- *Progressive scan interline transfer* CCD: here the matrix readout is accomplished in row-by-row fashion. After the exposure time the charge packets are transferred from the photosites to the vertical shift registers, thus allowing a new exposure cycle. As the new image acquisition is in progress, the previous is sent to the frame grabber. only at the completion of this last operation a new transfer will start. The progressive scan interline transfer CCD is particularly indicated for the acquisition of images of rapidly evolving scenes

The technology of CCD sensors proposes so many further details that are beyond the scope of this dissertation, nonetheless some considerations about the readout speed are worthy of remark. In fact, the speed of a digital camera is of crucial importance in the acquisition of moving objects and great attention is paid by manufacturers to decrease the acquisition time. The upper speed limit is set by the huge number of photosites forming a CCD. The transfer speed can be increased in two ways:

1. by increasing the clock[1] speed (i.e. the frequency of the voltage used for the

[1] The CLOCK or CLOCKING SIGNAL is the control voltage, associated to each gate, which is varied

charge transfer). A good frequency is 10 MHz, but in the most recent CCD it can be as high as 40 MHz. An increment of the clock frequency reduces the efficiency of the charge transfer and increases the noise

2. by dividing the matrix into smaller subregions, each equipped with its own horizontal shift register. The different regions are eventually readout at once after exposure, thus decreasing the readout time of the full sensor to the time necessary to accomplish the readout of a single subregion

---

as a function of time to accomplish the charge transfer

Bibliography

Allebach, J. P., Chang, J. Z., and Bouman, C. A. (1993), "Efficient implementation of nonlinear color transformations," in *Conference Proceedings: The First IS&T/SID Color Imaging Conference*, pp. 143–148.

Ambros-Ingerson, J., Granger, R., and Lynch, G. (1990), "Simulation of paleocortex performs hierarchical clustering." *Science*.

Amelio, G. F., Tompsett, M. F., and Smith, G. (1970), "Experimental verification of the charge-coupled concept," *Bell Syst. Tech. J.*, 49, 593–600.

Antonioli, G., Fermi, F., Oleari, C., and Reverberi, R. (2004), "Spectrophotometric scanner for imaging of paintings and other works of art," in *Proceedings of CGIV*, pp. 219–224.

Barducci, A., Marcoionni, P., Pippi, I., and Poggesi, M. (2001), "Simulation of the performance of a stationary imaging interferometer for high-resolution monitoring of the Earth," in *International Symposium on Remote Sensing*, pp. 112–121, International Society for Optics and Photonics.

Barducci, A., Casini, A., Castagnoli, F., Marcoionni, P., Morandi, M., and Pippi, I. (2002), "Performance assessment of a stationary imaging interferometer for high-resolution remote sensing," in *AeroSense 2002*, pp. 547–555, International Society for Optics and Photonics.

Barducci, A., Guzzi, D., Lastri, C., Marcoionni, P., Nardino, V., and Pippi, I. (2010), "Theoretical aspects of Fourier transform spectrometry and common path triangular interferometers," *Optics Express*, 18, 11622–11649.

Baronti, S., Casini, A., Lotti, F., and Porcinai, S. (1998), "Multispectral imaging system for the mapping of pigments in works of art by use of principal-component analysis," *Applied optics*, 37, 1299–1309.

Becker, S. (1991), "Unsupervised learning procedures for neural networks." *International Journal of Neural Systems*, 2, 17–33.

Bell, I. E. and Cowan, W. (1993), "Characterizing printer gamuts using tetrahedral interpolation," in *Proceedings of IS&T and SIDs Color Imaging Conference: Transforms and Transportability of Color*, pp. 108–113.

Berning, P. H. and Turner, A. F. (1957), "Induced transmission in absorbing films applied to band pass filter design," *Journ. Opt. Soc. Am*, 47, 230–239.

Berns, R. S. (1998), "Challenges for color science in multimedia imaging," *Colour Imaging: Vision and Technology, John Wiley & Sons, Chichester*, pp. 99–127.

Berns, R. S. (2000), "Principles of colour technology," *Wiley, NY*.

Berns, R. S. (2002), "Visible-spectrum imaging techniques: an overview," in *9th Congress of the International Color Association*, pp. 475–480, International Society for Optics and Photonics.

Bonifazzi, C., Carcagnì, P., Della Patria, A., Ferriani, S., Fontana, R., Greco, M., Mastroianni, M., Materazzi, M., Pampaloni, E., and Romano, A. (2006), "A scanning device for multispectral imaging of paintings," in *Spectral Imaging: Eighth International Symposium on Multispectral Color Science*, vol. 6062, pp. 1–10.

Bonifazzi, C., Carcagnì, P., Della Patria, A., Fontana, R., Greco, M., Mastroianni, M., Materazzi, M., Pampaloni, E., Pezzati, L., and Romano, A. (2007), "Multispectral imaging of paintings: instrument and applications," in *Optical Metrology*, pp. 661816–661816, International Society for Optics and Photonics.

Boreman, G. D. (2001), *Modulation transfer function in optical and electro-optical systems*, SPIE Press Bellingham, WA.

Boucheron, L. E., Bi, Z., Harvey, N. R., Manjunath, B. S., and Rimm, D. L. (2007), "Utility of multispectral imaging for nuclear classification of routine clinical histopathology imagery," *BMC Cell Biology*, 8, S8.

Boyle, W. S. and Smith, G. E. (1970), "Charge coupled semiconductor devices," *Bell Syst. Tech. J.*, 49, 587–593.

Brown, P. J. (1993), *Measurement, Regression, and Calibration*, Oxford Statistical Science Series, Clarendon Press.

Brown, P. K. and Wald, G. (1964), "Visual pigments in single rods and cones of the human retina." *Science; Science*.

Carcagni, P., Della Patria, A., Fontana, R., Greco, M., Mastroianni, M., Materazzi, M., Pampaloni, E., and Pezzati, L. (2007), "Multispectral imaging of paintings by optical scanning," *Optics and lasers in engineering*, 45, 360–367.

Chang, J. Z., Bouman, C. A., and Allebach, J. P. (1994), "Recent results in color calibration using sequential linear interpolation," in *IS&Ts 47th Ann. Conf., ICPS94: Physics and Chemistry of Imaging Systems*, vol. 2, pp. 500–505.

CIE (1977), *Radiometric and photometric characteristics of materials and their measurement*, Wien, Technical Report 38.

CIE (1986), *Colorimetry, No. 15.2*.

CIE (1989), *International Lighting Vocabulary*, CIE Publications,.

CIE (1995), *116-1995. Industrial Colour-Difference Evaluation*.

Clarke, F. J. J., McDonald, R., and Rigg, B. (1984), "Modification to the JPC79 Colour–difference Formula," *Journal of the Society of Dyers and Colourists*, 100, 128–132.

Cohen, J. (1964), "Dependency of the spectral reflectance curves of the Munsell color chips," *Psychonomic Science*.

Cortés, A. (2003), "Multispectral analysis and spectral reflectance reconstruction of art paintings," Ph.D. thesis, École Nationale Supérieure des Télécommunications.

Cotte, P. and Dupouy, M. (2003), "Crisatel high resolution multispectral system," in *IS & TS PICS Conference*, pp. 161–165, Society for Imaging Science & Technology.

Cotte, P. and Dupraz, D. (2006), "Spectral imaging of Leonardo Da Vincis Mona Lisa: A true color smile without the influence of aged varnish," *Proceeding IS & T CGVI*, 6, 311.

Cupitt, J. and Martinez, K. (1996), "VIPS: an image processing system for large images," in *Electronic Imaging: Science & Technology*, pp. 19–28, International Society for Optics and Photonics.

Davis, S. M., Landgrebe, D. A., Phillips, T. L., Swain, P. H., Hoffer, R. M., Lindenlaub, J. C., and Silva, L. F. (1978), "Remote sensing: the quantitative approach," *New York, McGraw-Hill International Book Co., 1978. 405 p.*, 1.

Della Patria, A., Fermi, F., Oleari, C., Piegari, A., and Sytchkova, A. (2011), "A portable spectro-photo/radio-metric camera with spatial filtering for VIS-NIR imaging," *Proceeding 13th IS & T CIC19*, 19, 285–289.

Della Patria, A., Fermi, F., Oleari, C., Piegari, A., and Sytchkova, A. (2012), "Visible-Infrared imaging by a portable spectrometer with linearly variable filter," in *Third Annual Hyperspectral Imaging Conference*, pp. 51–55, National Insitute of Geophysics and Volcanology.

Dennis, J. E. and Schnabel, R. B. (1987), *Numerical methods for unconstrained optimization and nonlinear equations*, vol. 16, Society for Industrial Mathematics.

Dereniak, E. L. and Crowe, D. G. (1984), *Optical radiation detectors*, Wiley and Sons Inc., New York.

Descour, M. and Dereniak, E. (1995), "Computed-tomography imaging spectrometer: experimental calibration and reconstruction results," *Appl. Opt.*, 34, 4817–4826.

Durbin, R. and Rumelhart, D. E. (1989), "Product units: A computationally powerful and biologically plausible extension to backpropagation networks," *Neural Computation*, 1, 133–142.

DZmura, M. and Lennie, P. (1986), "Mechanisms of color constancy," *JOSA A*, 3, 1662–1672.

Easton Jr, R. L., Knox, K. T., and Christens-Barry, W. A. (2003), "Multispectral imaging of the Archimedes palimpsest," in *Applied Imagery Pattern Recognition Workshop, 2003. Proceedings. 32nd*, pp. 111–116, IEEE.

Elias, M. and Cotte, P. (2008), "Multispectral camera and radiative transfer equation used to depict Leonardo's sfumato in¡ i¿ Mona Lisa¡/i¿," *Applied optics*, 47, 2146–2154.

Fairman, H. S. and Brill, M. H. (2004), "The principal components of reflectances," *Color Research & Application*, 29, 104–110.

Farrell, J. E. and Wandell, B. A. (1993), "Scanner linearity," *Journal of Electronic Imaging*, 2, 225–230.

Fischer, R. E., Tadic-Galeb, B., Yoder, P. R., and Galeb, R. (2000), *Optical system design*, McGraw Hill.

Fletcher-Holmes, D. W. and Harvey, A. R. (2002), "Birefringent Fourier-transform imaging spectrometer," in *Proc. SPIE*, vol. 4816, pp. 397–406.

Fletcher-Holmes, D. W. and Harvey, A. R. (2005), "Real-time imaging with a hyperspectral fovea," *Journal of Optics A: Pure and Applied Optics*, 7, S298.

Franklin, P. E. (1982), "Interpolation methods and apparatus," US Patent 4,334,240.

Geary, J. M. (2002), *Introduction to lens design: with practical ZEMAX examples*, Willmann-Bell.

Geladi, P. and Grahn, H. (1996), *Multivariate image analysis*, Wiley Chichester, UK.

Golub, M., Nathan, M., Averbuch, A., et al. (2011), "Snapshot spectral imaging systems and methods," US Patent 8,081,244.

Golub, M. A., Nathan, M., Averbuch, A., Lavi, E., Zheludev, V. A., and Schclar, A. (2009), "Spectral multiplexing method for digital snapshot spectral imaging," *Applied optics*, 48, 1520–1526.

Grossberg, S. (1982), *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control*, D. Reidel Publishing Company.

Haneishi, H., Hasegawa, T., Tsumura, N., and Miyake, Y. (1997), "Design of color filters for recording artworks," in *IS&T annual conference*, pp. 369–372, The society for imaging science and technology.

Hård, A. and Sivik, L. (2007), "NCSNatural Color System: a Swedish standard for coloer notation," *Color Research & Application*, 6, 129–138.

Hardeberg, J. Y. (1999), "Acquisition et reproduction d'images couleur: approches colorimétrique et multispectrale," Ph.D. thesis, Télécom ParisTech.

Hardeberg, J. Y. (2001), *Acquisition and reproduction of color images: colorimetric and multispectral approaches*, Dissertation. com.

Harvey, A. R., Beale, J., Greenaway, A. H., Hanlon, T. J., and Williams, J. (2000), "Technology options for imaging spectrometry," in *Proc. SPIE*, vol. 4132, pp. 13–24.

Hauta-Kasari, M., Wang, W., Toyooka, S., Parkkinen, J., and Lenz, R. (1997), "Unsupervised filtering of Munsell spectra," *Lecture notes in computer science*.

Hauta-Kasari, M., Miyazawa, K., Toyooka, S., and Parkkinen, J. (1999), "Spectral vision system for measuring color images," *JOSA A*, 16, 2352–2362.

Havermans, J., Aziz, H. A., and Scholten, H. (2003), "Non destructive detection of iron gall inks by means of multispectral imaging," *Restaurator, XXIV*, pp. 55–60.

Haykin, S. (1994), "Neural networks: A comprehensive approach," *IEEE Computer Society Press*.

Hecht, E. (2001), "Optics 4th edition," *Optics 4th edition by Eugene Hecht Reading, MA: Addison-Wesley Publishing Company, 2001*, 1.

Hering, E. (1905), "Grundzüge zur Lehre vom Lichtsinn, Gräfe-Sämisch: Handbuch der Augenheilkunde, I," *Engelmann, Leipzig*.

Heydt, S. (1984), "Klimsch COLORCOMP, description of a new system," Klimsch report no. 30518, Frankfurt, Germany.

Holloway, R. J. and Lissberger, P. H. (1969), "The design and preparation of induced-transmission filters," *Applied optics*, 8, 653–660.

Holst, G. C. (1998), *CCD arrays, cameras and displays*, SPIE Optical Engineering Press.

Hotelling, H. (1933), "Analysis of a complex of statistical variables into principal components," *Journal of educational psychology*, 24, 417.

Hung, P. C. and Berns, R. S. (2007), "Determination of constant hue loci for a CRT gamut and their predictions using color appearance spaces," *Color Research & Application*, 20, 285–295.

Hunt, R. W. G. (1982), "A model of colour vision for predicting colour appearance," *Color Research and Application*, 7, 95–112.

Hunt, R. W. G. (1991), "Measuring colour," *Ellis Horwood series in applied science and industrial technology*.

Hunt, R. W. G. (2005), *The Reproduction of Colour*, The Wiley-IS&T Series in Imaging Science and Technology, Wiley.

Hurvich, L. M. and Jameson, D. (1957), "An opponent-process theory of color vision." *Psychological Review; Psychological Review*, 64, 384.

IEC (1999), "Multimedia systems and equipment-colour measurement and management-part 2-1: Colour management-default RGB colour space-sRGB," Tech. rep., IEC 61966-2-1.

Imai, F. H. and Berns, R. S. (1999), "Spectral estimation using trichromatic digital cameras," in *Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, vol. 42.

Imai, F. H., Berns, R. S., and Tzeng, D. Y. (2000), "A comparative analysis of spectral reflectance estimated in various spaces using a trichromatic camera system," *Journal of imaging science and technology*, 44, 280–287.

Imai, F. H., Rosen, M. R., and Berns, R. S. (2002), "Comparative study of metrics for spectral match quality," in *Proceedings of the First European Conference on Colour in Graphics, Imaging and Vision*, pp. 492–496.

Iri, M. (1967), "A method of multi-dimensional linear interpolation," *J. Inform. Processing Soc. of Japan*, 8, 211–215.

Jaaskelainen, T., Parkkinen, J., and Toyooka, S. (1990), "Vector-subspace model for color representation," *JOSA A*, 7, 725–730.

Jiao, Y., Bhalotra, S. R., Kung, H. L., and Miller, D. A. B. (2003), "Adaptive imaging spectrometer in a time-domain filtering architecture," *Optics express*, 11, 1960–1965.

Johnson, R. A. and Wichern, D. W. (1992), *Applied multivariate statistical analysis*, vol. 4, Prentice hall Englewood Cliffs, NJ.

Johnson, W. R., Wilson, D. W., and Bearman, G. (2005), "All-reflective snapshot hyperspectral imager for ultraviolet and infrared applications," *Optics letters*, 30, 1464–1466.

Johnson, W. R., Wilson, D. W., Fink, W., Humayun, M., and Bearman, G. (2007), "Snapshot hyperspectral imaging in ophthalmology," *Journal of biomedical optics*, 12, 014036–014036.

Jolliffe, I. (2005), *Principal component analysis*, Wiley Online Library.

Judd, D. B. (1933), "Sensibility to Color-Temperature Change as a Function of Temperature," *J. Opt. Soc. Am.*, 23, 7–14.

Kaiser, P. K. and Boynton, R. M. (1996), *Human color vision*, vol. 287, Optical Society of America Washington DC.

Kanamori, K., Fumoto, T., and Kotera, H. (1992), "A Color Transformation Algorithm Using PrismInterpolation," in *Proceedings of IS&Ts Eighth International Congress on Advances in Non-Impact Printing Technologies, Springfield, VA: Society of Photographic Science Engineers*, pp. 477–482.

Kang, H. (2006), *Computational color technology*, vol. 159, Society of Photo Optical.

Kang, H. R. (1995a), "Comparisons of three-dimensional interpolation techniques by simulations," pp. 104–114.

Kang, H. R. (1995b), "Printer-related color processing techniques," pp. 410–419.

Kang, H. R. (1997), "Color technology for electronic imaging devices," SPIE-International Society for Optical Engineering.

Kasson, J. M., Plouffe, W., and Nin, S. I. (1993), "Tetrahedral interpolation technique for color space conversion," in *IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology*, pp. 127–138, International Society for Optics and Photonics.

Keusen, T. (1996), *Multispektrale Farbanalyse*, Shaker.

Klein, M. E., Aalderink, B. J., Padoan, R., De Bruin, G., and Steemers, T. A. G. (2008), "Quantitative hyperspectral reflectance imaging," *Sensors*, 8, 5576–5618.

Kohonen, T. (2001), *Self-organizing maps*, vol. 30, Springer Verlag.

König, F. and Praefcke, W. (1998), "Practice of multispectral image acquisition," in *SYBEN-Broadband European Networks and Electronic Image Capture and Publishing*, pp. 34–41, International Society for Optics and Photonics.

Krasilnikova, A., Piegari, A., Dami, M., Abel-Tiberini, L., Lemarquis, F., Lequime, M., et al. (2005), "Spatially resolved spectroscopy for non-uniform thin film coatings: comparison of two dedicated set-ups," in *Proc. SPIE*, vol. 5965, pp. 573–580.

Laikin, M. (2006), *Lens design*, vol. 27, CRC.

Landau, B. V. and Lissberger, P. H. (1972), "Theory of induced-transmission filters in terms of the concept of equivalent layers," *JOSA*, 62, 1258–1264.

Lee, H. C. (2005), *Introduction to color imaging science*, Cambridge University Press.

Lenz, R. and Lenz, U. (1990a), "Calibration of a color CCD camera with 3000x2300 picture elements," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 1395, p. 104.

Lenz, R. and Lenz, U. (1990b), "ProgRes 3000: A digital color camera with a 2D array CCD sensor and programmable resolution up to 2994 x 2320 picture elements,"," *MedTech*, 89, 204–209.

Lenz, R., Österberg, M., Hiltunen, J., Jaaskelainen, T., and Parkkinen, J. (1996), "Unsupervised filtering of color spectra," *JOSA A*, 13, 1315–1324.

Lissberger, P. H. (1981), "Coatings with induced transmission," *Applied optics*, 20, 95–104.

Loewen, E. and Palmer, C. (2005), *Diffraction Grating Handbook*, Newport Corporation.

Luo, M. R., Cui, G., and Rigg, B. (2001), "The development of the CIE 2000 colour-difference formula: CIEDE2000," *Color Research & Application*, 26, 340–350.

Macleod, H. A. (2001), *Thin film optical filters*, Taylor & Francis.

MacNichol, E. F. (1964), "Retinal mechanisms of color vision," *Vision Research*, 4, 119 – 133, IN15.

Magnus, X. and Neudecker, H. (1988), "Matrix differential calculus," *New York*.

Maitre, H., Schmitt, F. J. M., and Crettez, J. P. (1997), "High-quality imaging in museum: from theory to practice," in *Electronic Imaging'97*, pp. 30–39, International Society for Optics and Photonics.

Maloney, L. T. and Wandell, B. A. (1986), "Color constancy: a method for recovering surface spectral reflectance," *JOSA A*, 3, 29–33.

Marks, W. B., Dobelle, W. H., and MacNichol Jr, E. F. (1964), "Visual pigments of single primate cones." *Science*.

Martens, H. and Naes, T. (1992), *Multivariate calibration*, Wiley.

Martinez, K. (1991), "High resolution digital imaging of paintings: the VASARI project," *Microcomputers for Information Management*, 8, 277–283.

Martinez, K. and Hamber, A. (1989), "Towards a colorimetric digital image archive for the visual arts," in *Proceedings of the SPIE*, vol. 1073, pp. 114–121, SPIE.

Maxwell, J. C. (1855), "On the Theory of Colours in Relation to Colour-blindness," *Transactions of the Royal Society of Arts*, IV.

Maxwell, J. C. (1860), "On the theory of compound colours, and the relations of the colours of the spectrum," *Philosophical Transactions of the Royal Society of London*, 150, 57–84.

Maxwell, J. C. (1885), "Experiments on colour," *Transactions of the Royal Society Edinburgh*, 21, 275–298.

ITU-T T.42 (1994), "Continuous-tone colour representation method for facsimile," Tech. rep., ITU-T Recommendation, T.42.

McDowell, D. Q. (1999), "Graphic Arts Color Standards," in *TAGA*, pp. 661–670, TAGA; 1998.

Mendel, J. M. and McLaren, R. W. (1970), "Reinforcement-Learning Control and Pattern Recognition Systems," *Mathematics in Science and Engineering*, 66, 287–318.

Millman, J. and Grabel, A. (1987), *Microelectronics*, McGraw-Hill.

Mitchell, D. E. and Rushton, W. A. H. (1971), "The red/green pigments of normal vision," *Vision Research*, 11, 1045 – 1056.

Miyazawa, K., Hauta-Kasari, M., and Toyooka, S. (1999), "Rewritable broad-band filters for color image analysis," in *Proc. SPIE*, vol. 3740, pp. 468–471.

Miyazawa, K., Hauta-Kasari, M., and Toyooka, S. (2001), "Rewritable broad-band color filters for spectral image analysis," *Optical review*, 8, 112–119.

Mollon, J. D. (2003), *The origins of modern color science*, vol. 2, Optical Society of America, Oxford.

Moré, J. J. and Thuente, D. J. (1994), "Line search algorithms with guaranteed sufficient decrease," *ACM Transactions on Mathematical Software (TOMS)*, 20, 286–307.

Morelli, D. (2006), "Interference Filter Handbook," *JDS Uniphase Corporation*, pp. 39–48.

Munsell, A. H. (1979), "Munsell Book of Color: Matte Finish Collection," .

Newton, S. I. (1730), *Opticks, Or a Treatise of the Reflections, Refractions, Inflections and Colours of Light*, vol. 678, William Innys.

Nobbs, J. H. (1985), "KubelkaMunk Theory and the Prediction of Reflectance," *Review of Progress in Coloration and Related Topics*, 15, 66–75.

Nocedal, J. (1992), "Theory of algorithms for unconstrained optimization," *Acta numerica*, 1, 199–242.

Okamoto, T. and Yamaguchi, I. (1991), "Simultaneous acquisition of spectral image information," *Opt. Lett.*, 16, 1277–1279.

Okamoto, T., Takahashi, A., and Yamaguchi, I. (1993), "Simultaneous Acquisition of Spectral and Spatial Intensity Distribution," *Appl. Spectrosc.*, 47, 1198–1202.

Oleari, C. (2008), "Misurare il colore, spettrofotometria, fotometria e colorimetria: Fisiologia e percezione," *ed. Hoepli, Milano*.

Parkkinen, J., Hallikainen, J., and Jaaskelainen, T. (1989), "Characteristic spectra of Munsell colors," *JOSA A*, 6, 318–322.

Pearson, K. (1901), "On lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2, 559–572.

Piegari, A. and Bulir, J. (2006), "Variable narrowband transmission filters with a wide rejection band for spectrometry," *Applied optics*, 45, 3768–3773.

Piegari, A., Bulir, J., and Krasilnikova Sytchkova, A. (2008), "Variable narrow-band transmission filters for spectrometry from space. 2. Fabrication process," *Applied optics*, 47, C151–C156.

Piegari, A., Sytchkova, A., Della Patria, A., Fermi, F., and Oleari, C. (2011), "Small-dimension portable instrument for in-situ multispectral imaging," in *SPIE Optical Metrology-O3A*, vol. 8084, pp. 80840O1–80840O6, International Society for Optics and Photonics.

Piegari, A. M. and Janicki, V. (2003), "Metal-dielectric coatings for variable transmission filters with wide rejection bands," in *Proceedings of SPIE*, vol. 5250, pp. 343–348.

Pisani, M. and Zucco, M. E. (2009), "Compact imaging spectrometer combining Fourier transform spectroscopy with a Fabry-Perot interferometer," *Optics express*, 17, 8319–8331.

Pisani, M. and Zucco, M. E. (2011), *Fourier Transform Based Hyperspectral Imaging*, InTech.

Pratt, W. K. (1978), *Digital Image Processing*, John Wiley & sons.

Pugsley, P. C. (1975), "Image reproduction method and apparatus," US Patent 3,894,178.

Ribés, A., Brettel, H., Schmitt, F., Liang, H., Cupitt, J., and Saunders, D. (2003), "Color and multispectral imaging with the CRISATEL multispectral system," in *IS & TS PICS Conference*, pp. 215–219, Society for Imaging Science & Technology.

Ribés, A., Schmitt, F., Pillay, R., and Lahanier, C. (2005), "Calibration and spectral reconstruction for CRISATEL: an art painting multispectral acquisition system," *Journal of Imaging Science and Technology*, 49, 563–573.

Ripps, H. and Weale, R. A. (1963), "Cone pigments in the normal human fovea," *Vision Research*, 3, 531 – 543.

Rosselet, A. C., Graff, W., Wild, U. P., Keller, C. U., and Gschwind, R. (1995), "Persistent spectral hole burning used for spectrally high-resolved imaging of the sun," in *SPIE's 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics*, pp. 205–212, International Society for Optics and Photonics.

Rushton, W. A. H. (1957), "Visual problems of colour," N.P.L. Symposium, p. 73, H.M.S.O.

Rushton, W. A. H. (1958), "Visual pigments in the colour blind." *Nature*, 182, 690–2.

Sagnac, G. (1913), "L'éther lumineux démontré par l'effet du vent relatif d'éther dans un interféromètre en rotation uniforme," *CR Acad. Sci.*, 157, 708–710.

Sakamoto, T. and Itooka, A. (1981), "Linear interpolator for color correction," US Patent 4,275,413.

Saunders, D. and Cupitt, J. (1993), "Image processing at the National Gallery: The VASARI project," *the National Gallery technical bulletin*, 14, 72–85.

Saunders, D. R. and Hamber, A. (1990), "From pigments to pixels: Measurement and display of the color gamut of paintings," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 1250, pp. 90–102.

Schau, H. C. (2006), "Field multiplexed dispersive imaging spectrometer," US Patent 7,092,088.

Schrödinger, E. (1920), "Grundlinien einer Theorie der Farbenmetrik im Tagessehen," *Annalen der Physik*, 368, 481–520.

Sepper, D. L. (2003), *Goethe contra Newton: Polemics and the project for a new science of color*, Cambridge University Press.

Shanno, D. F. and Phua, K. H. (1980), "Remark on Algorithm 500: Minimization of unconstrained multivariate functions [e4]," *ACM Transactions on Mathematical Software (TOMS)*, 6, 618–622.

Smith, V. C. and Pokorny, J. (1975), "Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm," *Vision Research*, 15, 161 – 171.

Smith, W. J. (1997), *Practical optical system layout and use of stock lenses*, McGraw-Hill.

Smith, W. J. (2000), *Modern optical engineering*, Tata McGraw-Hill Education.

Smith, W. J. and Genesee Optics Software, I. (1992), *Modern Lens Design: A Resource Manual*, McGraw-Hill.

Sobagaki, H. (1984), "New approach to the colorimetric standardization for object colors," *Bull. Electrotechnical Lab. Jpn*, 48, 785–792.

Speranskaya, N. I. (1959), "Determination of spectrum color co-ordinates for twenty-seven normal observers," *Optics and Spectroscopy*, 7, 424 – 428.

Stiles, W. S. and Burch, J. M. (1955), "On the National Physical Laboratory's investigation of colour-matching," vol. 2 of *Interim report to the* CIE, pp. 168 – 181, Optica Acta.

Sun, D. W. (2010), *Hyperspectral imaging for food quality analysis and control*, Academic Press.

Svaetichin, G. and MacNichol, E. F. (1958), "Retinal mechanisms for chromatic and achromatic vision," *Annals of the New York Academy of Sciences*, 74, 385–404.

Tajima, J. (1998), "A huge spectral characteristics database and its application to color imaging device design," in *The sixth Color Imaging Conference*, pp. 86–89.

Trusseli, H. J. and Kulkarni, M. S. (1996), "Sampling and processing of color signals," *IEEE Trans Image Process*, 5, 677–81.

Tzeng, D. Y. and Berns, R. S. (2005), "A review of principal component analysis and its applications to color technology," *Color Research & Application*, 30, 84–98.

Usui, S., Nakauchi, S., and Nakano, M. (1992), "Reconstruction of Munsell color space by a five-layer neural network," *JOSA A*, 9, 516–520.

Varoquaux, G. (2012), "Dimension reduction with PCA," `http://gael-varoquaux.info/scientific_computing/ica_pca/index.html`.

Vrhel, M. J. and Trussell, H. J. (1992), "Color correction using principal components," *Color Research & Application*, 17, 328–338.

Wandell, B. A. (1995), *Foundations of vision.*, Sinauer Associates.

Weale, R. A. (1959), "Photo-sensitive Reactions in Foveae of Normal and Cone-monochromatic Observers," *Optica Acta: International Journal of Optics*, 6, 158–174.

Westland, S., Shaw, J., and Owens, H. (2000), "Colour statistics of natural and man-made surfaces," *Sensor Review*, 20, 50–55.

Westland, S., Ripamonti, C., and Cheung, V. (2012), *Computational Colour Science Using MATLAB*, The Wiley-IS&T Series in Imaging Science and Technology, Wiley.

Williams, C. S. and Becklund, O. A. (1989), *Introduction to the optical transfer function*, vol. 112, Society of Photo Optical.

Wilson, D. W., Maker, P. D., Muller, R. E., and Mouroulis, P. Z. (2003), "Computed tomography imaging spectrometer (CTIS) with 2D reflective grating for ultraviolet to long-wave infrared detection especially useful for surveying transient events," US Patent 6,522,403.

Wright, W. D. et al. (1944), *The measurement of colour.*, Adam Hilger Ltd.

Wyszecki, G. and Stiles, W. S. (1982), *Color science*, Wiley New York.

Yoder Jr, P. R. (2005), *Opto-mechanical systems design*, vol. 105, CRC.

Young, R. A. (1986), "Principal-component analysis of macaque lateral geniculate nucleus chromatic data," *JOSA A*, 3, 1735–1742.