From Color Sensor Space to Feasible Reflectance Spectra

Silvia Zuffi, *Member, IEEE*, Simone Santini, *Member, IEEE*, and Raimondo Schettini

Abstract—The interaction of light and object surfaces generates color signals in the visible band that are responsible for digital acquisition system outputs. Inverting this mapping from the sensor space back to the wavelength domain is of great interest for many applications. Since 1964, with the idea of Cohen to exploit the characteristic of smoothness of surface reflectance functions, a lot of work has been done in the analysis, synthesis and recovering of spectral information using linear models. The general use of such models is for the establishment of a one-to-one relationship between sensor's data and reflectance spectrum, with the requirement of ensuring the quality of the recovered spectrum in terms of physical feasibility and naturalness. In this paper, we propose a solution to correct the outcome of a generic recovery method, in order to take into account quality constrains. Our strategy assumes the smoothness of the solution of the recovery method, an assumption implicitly satisfied from the adoption of linear models to represent reflectance functions.

Index Terms—Color, color reflectance spectra, color sensors, color signal, recursive estimation, reflectance linear models, spectral recovery.

I. INTRODUCTION

THE color of objects may be communicated in many forms. Color may be described by physical samples, by color terms or by numerical parameters that encode its appearance. The description that completely characterizes the physical property of surfaces rendering color is the *reflectance spectrum*, a function defined on the domain of visible wavelengths that represents the percentage of incident light that the surface reflects at each wavelength. The product of the surface reflectance and the spectral power distribution of the illuminant defines, for given incident and viewing geometries, the *color signal* [1], which, entering the eye is filtered by the photoreceptors to determine at post receptoral and cognitive levels the perceived color. Likewise, the color signal is filtered and processed in digital cameras to give a color description in the device color space. The problem of relating measurements like those coming from a color measuring device, with spectral information—often in the form of a reflectance function—has many applications, such as estimating a reflectance function given outputs of a camera system for the characterization of such devices [2]; or inverting a spectrum-to-colorimetry mapping for a given color reproduction process for the simulation of the behavior of imaging systems [3]. A typical example of the recovery of spectral data for color surfaces is their rendering under different illuminants, a procedure known as *color correction*. Color correction methods simulate different illuminants by computing tristimulus values under a target illuminant, given the tristimulus values under a reference illuminant. Color correction may exploit the knowledge of the reference illuminant to recover a reflectance spectrum to increase the simulation accuracy [4], [5]. *Spectral recovery*, i.e., the estimation of the spectra of objects surfaces from data having lower dimensionality (i.e., from the RGB values of an image, from LMS cone signals or from a generic N-filters camera responses) is based on the limitation in bandwidth, in terms of Fourier frequency, of the reflectance spectra of the majority of natural surfaces, which are, for most natural objects, smooth and slowly varying functions of the wavelength, just like are the spectra produced in photography, printing processes, or painting [6]. There are exceptions, such as the surface reflectance spectra of some earth metals and of some animals (e.g., wings of insects, fish scales and birds feathers), but in general, the overwhelming majority of functions spectra are smooth and slowly varying functions of the wavelength in the visible range [7], [8]. Since the work of Stiles *et al.* [8], who observed the smooth spectral profiles of color signals, it has been widely demonstrated that the reflectance spectrum of natural surfaces is a smooth and low-pass function of wavelength. Band-limited functions can be adequately approximated with linear models having a small number of basis functions. As a consequence of the sampling theorem, the frequency content in the Fourier domain of a signal determines the minimum number of parameters of its linear model: if the number of parameters is limited to three, the corresponding frequency limit is approximately 0.005 cycle/nm [9]. A study of Maloney [11] investigated the frequency content of the reflectance functions in terms of Fourier analysis of natural and artificial surfaces (the Nickerson–Munsell set and natural formations collected by Krinov [12]), and related it to the number of basis functions necessary for their characterization. Values of 0.01 and 0.015 cycle/nm were reported as band limits of surface reflectance functions, which, consistently with the results of Stiles [8] of 0.01–0.02 cycle/nm, correspond to a linear model with 6 to 12 parameters. A similar frequency cut-off was observed from a more recent study conducted on a more general dataset, including natural surfaces of fruits,

flowers and leaves [10]. As already said, the color signal is the product of the reflectance function and the illuminant spectral power distribution: an upper bound for its frequency limit can be therefore evaluated by the convolution theorem in the sum of the frequency limits of the reflectance and illuminant functions. An analysis conducted by Romero *et al.* [13] considered the frequency limit of color signals, corresponding to biochrome and nonbiochrome surfaces illuminated by daylight, incandescent and fluorescent illuminants. In these signals, as far as daylight and incandescent illuminants are concerned, a limit of 0.016 cycle/nm can be considered. Fluorescent illuminants limit the possibility to adopt models with a small number of parameters due to the information content at high frequencies. On the other hand, daylights have a very low frequency limit, suggested in 0.0033 cycle/nm [10] and can be modeled with very few parameters: three are sufficient [13], [14].

In the recovery of the reflectance spectra, one may assume that the illuminant is unknown, and, considering a suitable basis set, model it with a linear model as well [15], [16]. More frequently, one assumes that the illuminant of the scene is known or can be previously estimated, and the measurements are sensors values. The problem therefore entails inverting the sensor equations (possibly with some spatial constraints). This constitutes an underdetermined, constrained problem: underdetermined because, in general, for a given illuminant there will be several (uncountably many, in fact) *metameric* functions that, under that illuminant, will produce the desired sensor response; constrained because one seeks solutions that are physically feasible, a constraint that entails, *inter alia*, that the values of the reconstructed spectra fall into a the range [0, 1]. This constraint is due to the fact that physically a reflectance spectrum is defined as the *fraction* of the incident light that the body reflects at each wavelength. Most of the standard methods for inverting the sensor's equations do not guarantee the feasibility of the solution they find. Feasible solutions can be found by associating a cost to the sensor's equations (a cost that would be zero when the equations are satisfied), and solving a nonlinear constrained optimization problem [17] or, considering the equations as a linear constraint, a linear programming problem [18]. These methods, however, are often computationally cumbersome and can have additional drawbacks. In the case of nonlinear optimization, the method may fail to find a solution even if one existed while in the case of linear programming (a method that we will consider more extensively in the following, as it will constitute part of our numerical analysis) one has little control over the properties of the solution that is found.

There is considerable research literature on methods for spectral recovery. The analysis in the frequency domain of reflectance spectra and color signals motivates the recourse to dimensionality reduction techniques involving, in the most common approach, empirical linear models. In general, when representative data are available, linear models are defined on the basis of statistical information, applying principal component analysis (PCA) [19] and independent component analysis (ICA) [20]. The number of basis functions necessary to accurately represent the reflectance spectra depends on the characteristics of the data that one is modeling, and on the characteristics of the functions used in the linear model. There are many studies on the dimensionality of such linear models based on PCA. PCA seeks the set of basis functions that minimizes the correlation among dimensions and identifies those dimensions that are most descriptive of the data set [15]. In general, for natural reflectance spectra, a number of basis functions between 6 and 9 is considered adequate, as emerged from studies from Cohen [21] on a subset of the Munsell surface spectral reflectances collected by Kelley *et al.* [22], from Maloney [11] on spectral reflectances of natural formations collected by Kirnov [12], and from Jaaskelainen [23] on samples derived from several different kinds of plants. Similar studies on skin reflectance indicate that three functions are sufficient [24]. A comparison of the linear model representation of the Munsell color chips with three basis components can be found in [25]. The methods considered, PCA, ICA, and non-negative matrix factorization [26], [27], performed similarly. Linear models based on PCA components or on functions computed with ICA assume knowledge in advance of the surfaces to model. More general representations adopt generic basis functions, such as Fourier functions. In an early study from Wandell [15], the surface reflectances of a set of 462 Munsell chips have been modeled with a three-dimensional linear system composed of Fourier basis functions, with a rather good fit between measured spectra and linear model representations. Many of the studies on the dimensionality of linear models evaluate a degree of fit between the reflectance function and its representation in the wavelength domain. But in spectral information recovery from low-dimensional data, the definition of the linear system based on a least squares cost function on reflectance representation may be inappropriate [28]. In fact, a least-square criterion ignores potential effects of the sensor sensitivities in the model fit. In selecting linear models for human vision, a weighted least-squares criterion across wavelength may be considered. According to Maloney, in fact, surface spectral reflectances fall within a linear model composed by five to seven parameters (basis reflectance vectors), but when the effect of human photoreceptors sensitivities is included, linear model with as few as three to four parameters provide excellent fit to the data sets [11]. Linear models based on the L_2 metric are optimal when the input data follow a Gaussian distribution. When the data deviate from normal, a nonlinear approach may improve linear estimation methods [29]. When it is not possible to exploit previous knowledge on data, one can use linear models with generic function components [15], [30], [31], or exploit different strategies that do not assume any *a priori* knowledge. A comparison of many of these methods to recover a reflectance spectrum of a set of textile samples from colorimetric triplets was reported by Dupont [32]. Dupont considered the simplex method, the simulated annealing method, the Hawkyard method [33], genetic algorithms, and neural networks. The simplex method, set to assign a reflectance value for each wavelength, gave very chaotic results, physically unacceptable. To solve this problem, the simplex was used to assign the proportion (between 0 and 1) of a bell curve for each wavelength. For accuracy, the simplex and Hawkyard methods were superior to the others, and the simplex coupled with the bell functions gave the most realistic shape of the curves. The high computational cost of these methods, however, was a

drawback. The Hawkyard method has been recently modified to reduce computation time [34]. The methods were evaluated according to colorimetric differences, but no information about spectral match between the real and the recovered spectra was provided. The approaches investigated by Dupont [32] were not based on training datasets. A further method that does not require the knowledge of statistical information about the reflectance spectra of the input material was proposed by Li and Luo [35]. Their work was motivated by that of van Trigt [36], [37], and uses a smoothness condition for recovering the reflectance of a set of tristimulus values, given the illuminant. The smoothness constraint corresponds to the integral of the first derivative of the reflectance curve on the whole wavelength band, and the feasibility of the solution is imposed in the optimization.

In recovery methods based on training sets, the occurrence of negative values in spectral recovery may indicate that the input color is out of the gamut defined by the recovery model. For this reason, in many of the methods based on statistical analysis of input data, the feasibility problem is not addressed, or nonfeasible solutions are omitted [30]. A method to correct reflectance functions that have negative values based on metameric black functions is described in Wyszecki and Stiles [38, p. 187]. A *metameric black* is a spectrum that, for a specific illuminant and a set of sensors, produces tristimulus values equal to zero. For any given illuminant and a set of sensors, there are infinite metameric blacks, which have the feature of being positive at some wavelength and negative at others. The procedure is to add a metameric black to a given spectral reflectance function. To ensure that the resulting curve is positive for all wavelengths, the metameric black function is multiplied by an appropriate scaling factor. The metameric black approach was exploited by Morovic *et al.* [39] to define the infinite set of metameric solutions obtainable by adding to the solution of an RGB-to-spectrum problem an arbitrarily scaled metameric black. To get the solution to the recovery problem, an appropriate selection within the metameric set is performed, in search of a function satisfying the properties of physical realisability (the belonging to the range [0, 1], a constraint that we call *feasibility*), smoothness (satisfied by linear model approach), and naturalness (property of the solution to be realizable as a convex combination of existing surface reflectances).

In this paper, we present a method of a different nature to find a feasible reflectance spectrum. Our technique consists in using a simple linear method to find a (possibly unfeasible) reflectance spectrum that generates the required sensor measurements, and then to apply repeatedly a transformation that, while maintaining metamerism, will make the spectrum converge towards a feasible one. The method we propose may contribute to render applicable recovery methods based on linear system solutions that do not address the problem of the feasibility while computing the reflectance spectrum. The method we propose is based on the assumption that the recovery procedure computes a smooth function. As previously discussed, many of the methods found in the literature are of this kind, as the smoothness of the reflectance function is the primary assumption made to estimate a reflectance function at many wavelengths points from a smaller set of sensor's responses.

II. THE MODEL

In a simplified model of light–surface interaction, a single light source is absorbed and re-emitted by just one surface, producing a color signal that impinges on the retina or on an acquisition sensor. The effect that the color signal would generate on the human eye or on the imaging device can be expressed in a model where human photoreceptors or acquisition sensors can be described by their sensitivity functions. The same model, if sensitivity functions are replaced by the color matching functions, holds for the calculus of tristimulus values, as long as normalization factors are not considered. In our work, we consider the tristimulus equations as our light–surface–sensor interaction model. The tristimulus values equations are

$$X^k = \int_\Lambda I(\lambda) x^k(\lambda) r(\lambda) d\lambda \tag{1}$$

where $I(\lambda)$ is the spectral power distribution of the illuminant, and $r(\lambda)$ is the surface reflectance function. Our goal is to solve the above equations for the unknown function $r(\lambda)$. Note that, for the sake of compactness, we use the letter Λ to denote the visible wavelength interval (in this paper, we always use $\Lambda = [\lambda_0, \lambda_1] = [400 \text{ nm}, 700 \text{ nm}]$), and write the tristimulus values as $X = [X^1, X^2, X^3]$ rather than in the customary way as $X = [X, Y, Z]$; correspondingly, the color matching functions are written as $x(\lambda) = [x^1(\lambda), x^2(\lambda), x^3(\lambda)]$. Both X and x are column vectors: throughout the paper, apices will denote row numbers and pedices will denote column numbers. The tristimulus equations are typically used in their sampled form. For $N > 2, \Delta = (\lambda_1 - \lambda_0)/(N-1)$, let $\lambda^i = \lambda_0 + (i-1)\Delta, i = 1, \ldots, N$; the tristimulus equation can then be approximated as

$$\sum_{i=1}^{N} I(\lambda^i) x^k(\lambda^i) r(\lambda^i) \Delta = X^k. \tag{2}$$

Defining $r = [r^i] = [r(\lambda^i)], \hat{I} = \text{diag}(I(\lambda^i)), S = [x_i^k] = [x^k(\lambda^i)]$, we obtain the matrix form of the tristimulus equations

$$X = S\hat{I}r\Delta. \tag{3}$$

In order to simplify the notation, define the matrix $A = S\hat{I}\Delta$ so that the tristimulus equations can be written as $X = Ar$ or

$$\sum_{i=1}^{N} a_i^k r^i = X^k. \tag{4}$$

We express reflectances as a linear combination of *basis* functions: given b functions, $\Psi_1, \ldots \ldots, \Psi_b$, the reflectance r is expressed as

$$r(\lambda) = \sum_{h=1}^{b} w^h \Psi_h(\lambda) \tag{5}$$

where w^h are suitable coefficients. Defining the matrix $\Psi = [\Psi_i^h] = [\Psi_h(\lambda_i)]$, (4) can be rewritten in terms of the coefficients w as

$$
\begin{aligned}
X^k &= \sum_{i=1}^{N} \sum_{h=1}^{b} a_i^k \Psi_h^i w^h \\
&= \sum_{h=1}^{b} \left[\sum_{i=1}^{N} a_i^k \Psi_h^i \right] w^h = \sum_{h=1}^{b} c_h^k w^h
\end{aligned}
\tag{6}
$$

or $X = A\Psi w = Cw$, with $C = A\Psi$.

In order to estimate the reflectance function given the tristimulus values and the spectrum of the illuminant, one must solve (6) for the coefficients w, and then apply (5) to derive the sampled reflectance spectrum. If the number of basis functions is limited to three, then the tristimulus (6) can be solved by inverting the 3×3 matrix C. If the number of functions is greater than three, then standard linear algebra methods can be used to find a solution with certain optimality conditions, such as the minimal norm solution or a solution with the maximum number of zero components [18]. These methods, as mentioned in the introduction, do not guarantee that the computed spectrum is feasible, that is, that all its components lie in the desired interval. One solution that has been sometimes used in practice is to *clip* the spectrum forcing it to the desired range. This solution, however, causes a distortion, possibly quite severe, in the reproduction of the color. Our method, as we mentioned briefly before, consists of using the linear method to obtain an *initial* spectrum r, which might not be in the desired range, and then to apply iteratively a correction that will produce metameric spectra closer to the desired range. Note that any method that can derive a smooth spectrum from sensor's data can be used in place of the linear method. Before presenting our strategy, however, it is opportune to spend a few words summarizing some facts about the possible ranges of the reconstructed spectra and the solutions that they can generate. The reflectance spectrum r specifies, for each wavelength λ, the fraction of incident light that an object—in a given geometrical configuration with respect to the light sources and the observer—reflects. Being a fraction, it is natural to assume that, for each $\lambda, r(\lambda) \in [0,1]$. We assume that objects are not florescent, and ignore translucency. We call a spectrum that satisfies this condition *feasible in the strong sense*, or *strongly feasible*. When solving the mathematical reconstruction problem, however, one can sometimes accept a weaker constraint. Consider a spectrum in the range [0, M], with $M > 1$. Since the spectrum satisfies the tristimulus equations, we can write

$$
\int_{\Lambda} d\lambda, I(\lambda) x^k(\lambda) r(\lambda) = \int_{\Lambda} d\lambda, MI(\lambda) x^k(\lambda) \frac{r(\lambda)}{M} = X^k
\tag{7}
$$

where, of course, $\tilde{r}(\lambda) = r(\lambda)/M \in [0,1]$ is strongly feasible. Any spectrum r in the range [0, M] can be transformed in a strongly feasible spectrum $\tilde{r}(\lambda)$ and although $\tilde{r}(\lambda)$ is not, *stricti dictu*, a metamer of $r(\lambda)$, the same color can be generated

simply by increasing the intensity of the illuminant from $I(\lambda)$ to $MI(\lambda)$. For some applications this may be acceptable, leaving on the reflectance only the constraint $r(\lambda) \geqslant 0$. We call a spectrum that satisfies this condition *feasible in the weak sense*, or *weakly feasible*. Intuitively, weak feasibility is, as the name implies, a less demanding condition to impose. For a given illuminant, the feasibility conditions on r place some constraints on the tristimulus values X^k that one can obtain, that is, on the colors that can be physically realized. Assuming for the sake of simplicity weak feasibility and the discrete model (4), let e_i be the i th element of the natural basis of the space $\Re^N$ of the discrete spectra. Each spectrum r can be written as

$$
r = \sum_{i=1}^{N} r^i e_i
\tag{8}
$$

with $r^i \geqslant 0$. If we vary the hth component of r and leave the rest of it to be zero, that is, if we take $r = r^h e_h$, we obtain, from the tristimulus equation, $X^k = r^h a_h^k$, where $r^h \geqslant 0$ because of the feasibility constraint, and $a_h^k \geqslant 0$ because $I \geqslant 0$ and all the x^k are positive [38]. The locus of the tristimulus X is in this case a half-line in the positive octant of the space X^1, X^2, X^3. Varying different components of r we obtain N such lines. Equation (4) states that, if the feasibility constraint is valid, the tristimulus values are a convex combination of values on these lines, that is, X can be realized with the given illuminant if and only if it falls within the convex hull of these lines. Such a convex hull is a cone with vertex in the origin, called the *color cone* generated by the illuminant I. Any value of the tristimulus that falls outside of this cone corresponds to a color that can't be generated under the given illuminant. In these circumstances, of course, the method for finding a feasible r would fail to converge. In practice, the tristimulus values from which one starts are often derived from a color generated under a certain illuminant and, if the illuminant I that we hypothesize for the reconstruction is not wildly different from that under which the color was generated in the first place, the desired tristimulus values will fall within the color cone. This, at least, is the hypothesis that we will make throughout this paper.

III. THE METHOD

Suppose that we have solved the system (6) obtaining a weight vector w. From the discretization of (5)

$$
r^i = \sum_{h=1}^{b} w^h \Psi_h^i
\tag{9}
$$

we obtain the sampled reflectance $r_0 = [r^1, \ldots, r^N]$. Define the clipping functions

$$
u_f(x) = \begin{cases} 1, & \text{if } x > 1 \\ x, & \text{if } 0 \leqslant x \leqslant 1 \\ 0, & \text{if } x < 0 \end{cases}
\tag{10}
$$

$$
u_+(x) = \begin{cases} x - 1, & \text{if } x \geqslant 1 \\ 0, & \text{if } x < 1 \end{cases}
\tag{11}
$$

$$
u_-(x) = \begin{cases} -x, & \text{if } x \leqslant 0 \\ 0, & \text{if } x > 0 \end{cases}
\tag{12}
$$

then, applying the functions u to a vector componentwise, one can write

$$r_0 = u_f(r_0) + u_+(r_0) - u_-(r_0)$$
$$= r_{0,f} + r_{0,+} + r_{0,-}. \tag{13}$$

Correspondingly, the tristimulus X can be decomposed from (4) as

$$X = Ar_0 = Ar_{0,f} + Ar_{0,+} + Ar_{0,-}$$
$$= X_{0,f} + X_{0,+} + X_{0,-}. \tag{14}$$

Consider the component $X_{0,+}$: we can consider it as the datum of a new spectral estimation problem for which, using (6), we find the weights

$$w_{0,+} = C^{-1}X_{0,+} \tag{15}$$

and from these the reflectance

$$\tilde{r}_{0,+} = \Psi w_{0,+} = \Psi C^{-1}X_{0,+} = \Psi C^{-1}Ar_{0,+}. \tag{16}$$

Similarly, the negative component $X_{0,-}$ is the datum of an estimation problem from which we derive the spectrum

$$\tilde{r}_{0,-} = \Psi C^{-1}X_{0,-} = \Psi C^{-1}Ar_{0,-}. \tag{17}$$

Define now the reflectance

$$r_1 = r_{0,f} + \tilde{r}_{0,+} - \tilde{r}_{0,-}. \tag{18}$$

From the previous construction, the following property follows immediately.

Proposition 3.1: The reflectances r_0 and r_1 are metameric under the given illuminant.

The procedure can be repeated by decomposing

$$r_1 = u_f(r_1) + u_+(r_1) - u_-(r_1)$$
$$= r_{1,f} + r_{1,+} + r_{1,-} \tag{19}$$

computing

$$\tilde{r}_{1,+} = \Psi C^{-1}Ar_{1,+} \tilde{r}_{1,-} = \Psi C^{-1}Ar_{1,-} \tag{20}$$

then

$$r_2 = r_{1,f} + \tilde{r}_{1,+} - \tilde{r}_{1,-} \tag{21}$$

, and so on. In this way we obtain a sequence of metameric spectra $r_n = [r_n^1, \ldots, r_n^N], n = 1, \ldots$. We claim that this sequence converges to a feasible spectrum, that is

$$\lim_{n\to\infty}\max_{1\leqslant j\leqslant N}r_{n,+}^j = \lim_{n\to\infty}\max_{1\leqslant j\leqslant N}r_{n,-}^j = 0 \tag{22}$$

leaving, in the limit, a spectrum $r_\infty = r_{\infty,f}$ which, by construction, is feasible. The reason why (22) holds will be considered in the next section but, before doing so, we should like to give an intuitive sense of why things work. Consider, for the sake of simplicity, only the portion $r_{n,+}$ of the spectrum, the reasoning for

(a)

(b)
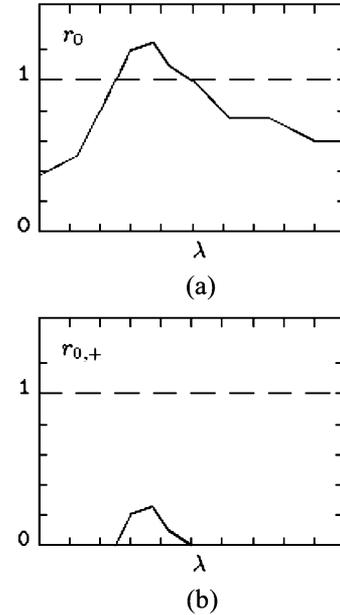
Fig. 1.   (a) Example of an unfeasible reflectance function. (b) Portion that exceed the maximum admissible value.
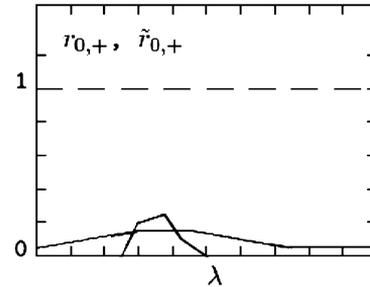
Fig. 2.   Spectrum of Fig. 1(b) and a metamer with a larger support and a smaller maximum value.

$r_{n,-}$ being analogous. Take a spectrum r_0 with values greater than 1, as in Fig. 1(a), and isolate the part $r_{0,+}$, as in Fig. 1(b).

The unknown function $\tilde{r}_{0,+}$ must be a metamer of $r_{0,+}$, that is, in the continuum, it must be

$$\int_\Lambda d\lambda, I(\lambda)x^k(\lambda)\tilde{r}_{0,+} = \int_\Phi d\lambda, I(\lambda)x^k(\lambda)r_{0,+} \tag{23}$$

where Φ is the support of $r_{0,+}$ that is, in general, smaller than Λ since $r_{0,+}$ represents only a peak in an otherwise feasible spectrum. On the other hand, $\tilde{r}_{0,+}$ is a combination of smooth basis functions and has in general the whole spectrum as support. In other words, in $\tilde{r}_{0,+}$ we distribute the same "area" over a much wider support, so that the magnitudes of $\tilde{r}_{0,+}$ will in general be smaller than those of $r_{0,+}$ (see Fig. 2).

This is only a qualitative explanation to give the reader access to the intuition that was behind the inception of the method. In the following, we will put these intuitions on a more formal footing and determine the conditions of convergence of the algorithm. The complete algorithm is given in Table I.

The function *stop* determines the point of convergence. Several criteria are possible: the simplest ones—computation-

TABLE I
THE ALGORITHM FOR THE RECOVERY OF THE REFLECTANCE FUNCTION GIVEN
THE TRISTIMULUS VALUES AND THE CORRESPONDING ILLUMINANT

$$correct(A, C, r) =$$
$$r_f = u_f(r);$$
$$r_+ = u_+(r);$$
$$r_- = u_-(r);$$

$$recover(I, X) =$$
$$\text{\underline{while}} \neg stop(r_f, r_+, r_-) \text{ \underline{do}}$$
$$A := Sdiag(I)\Delta;$$
$$r_+ := \Psi C^{-1} A r_+;$$
$$C := A\Psi;$$
$$r_- := \Psi C^{-1} A r_-;$$
$$r := \Psi C^{-1} X;$$
$$r := r + r_+ - r_-;$$
$$r_f = correct(A, C, r);$$
$$r_f = u_f(r);$$
$$\text{\underline{return}}\, r_f;$$
$$r_+ = u_+(r);$$
$$r_- = u_-(r);$$
$$\text{\underline{od}};$$
$$\text{\underline{return}}\, r_f;$$

ally speaking—compare the norm of the three components r_f, r_+, r_-, for instance

$$\text{stop}(r_f, r_+, r_-) = \frac{|r_+| + |r_-|}{|r_f|} < \varepsilon. \tag{24}$$

Other criteria consider the error in the color reconstruction, e.g.,

$$\text{stop}(r_f, r_+, r_-) = \frac{|A(r_+ + r_-)|}{|A(r_f + r_+ + r_-)|} < \varepsilon. \tag{25}$$

Smoothing the Spectrum: Reflectance spectra of natural surfaces are smooth functions of wavelength. Consequently, the basis functions computed on measured spectra are also smooth (this is in general true, and comes from the choice of the basis set). Therefore, a reflectance spectrum obtained as a solution of the linear system in (6) is a smooth function. As discussed in Section I, many studies have quantified the smoothness of natural reflectance functions in terms of content in the frequency domain. Observing the measured spectra of real surfaces, it is evident that such functions are characterized by a slowly varying and bounded curvature, as well. Estimated or synthesized spectra must exhibit this property in order to be practically realizable. Physically, it is difficult to reproduce spectra with very high curvature: each point of high curvature of the spectrum is a point where reproduction with physical means will likely fail.

If the obtained spectrum is not feasible, that is, is not in [0, 1], and the method described in the previous paragraph is applied, it may happen that the curvature of the final spectrum, which has feasible values, is unbounded or irregularly varying and will present sharp transitions between the intervals in which its values are in [0, 1] and those in which they are 0 or 1. See, for example, Fig. 3: the initial solution obtained solving the linear model produced the spectrum of Fig. 3(a), which has negative values in the region of long wavelengths. This spectrum was corrected for value feasibility, giving the metameric spectrum of Fig. 3(b). Unfortunately, the transition from positive values of reflectance to the values equal to zero does not appear to be "natural." To overcome this problem, we amend our method to
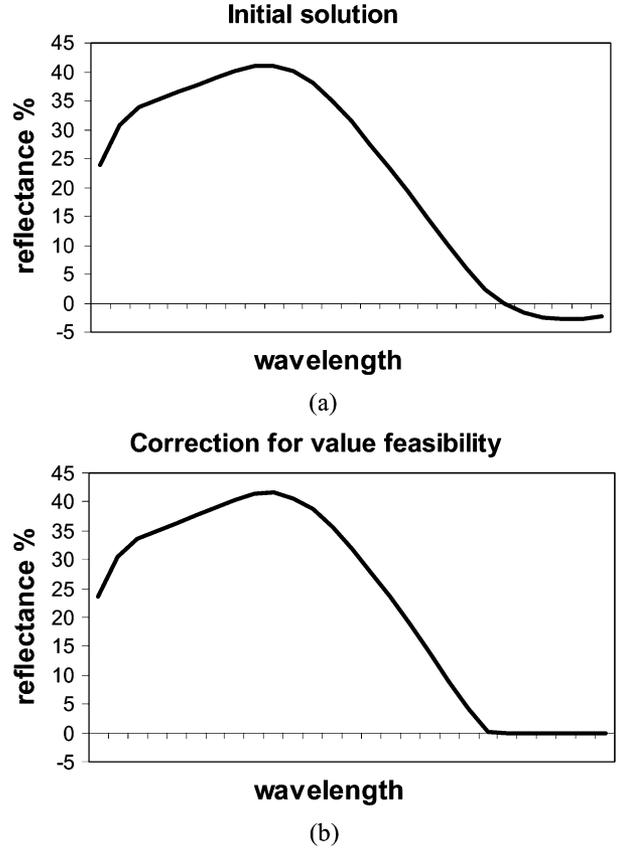
Fig. 3. (a) Recovered spectrum exhibiting negative values. (b) Metamer of (a) obtained with the application of the proposed method.

"smooth" the spectrum where it appears to have an unfeasible shape, in order to increase its naturalness and make it physically realizable.

We define *shape feasibility* as the property of a spectrum to present a smooth curvature. The shape feasibility is a further requirement we impose to the final solution in addition to the *value feasibility* discussed in Section II. We consider the reflectance curve as a line in a 2-D plane where the wavelength abscissa corresponds to x and the reflectance value corresponds to y. In order to set an appropriate scale, reflectance values must be multiplied for 300, being $700 \sim 400$ nm the wavelength domain of the curve. In this way, the ranges of values in the x and y coordinates have the same spatial extension. The procedure to round off the spectrum modifies the procedure of Table I. Inside the *while* cycle, after the calculus of r_f, r_+ and r_-, a *smooth()* procedure should be inserted. One possible definition of the procedure is listed in Table II; however, other different smoothness measures can be adopted [35], [36].

In Table II, N is the number of wavelength samples, and $\alpha_{\min}$ is the minimum angle between two segments composing the reflectance curve that ensures a smooth curvature. If we consider the example of Fig. 3 and impose an angle limit of $(3/4)\Pi$, we obtain the rounded spectrum of Fig. 4.

IV. CONVERGENCE

In this section, we put the intuitive explanation of the previous section on a more formal basis, and study the convergence of the

TABLE II
ALGORITHM TO INCREASE THE NATURALNESS OF A REFLECTANCE SPECTRUM

$smooth(r_f, r_+, r_-) = \underline{while}(i < N)\underline{do}$

$$\alpha^i = \tan^{-1}\left(\left.\frac{dr_f}{d\lambda}\right|_{i-1}\right) - \tan^{-1}\left(\left.\frac{dr_f}{d\lambda}\right|_i\right) + \Pi;$$

$\underline{if}\left(\alpha^i < \alpha_{\min}\right)$

$$r_s^i := \frac{r_f^{i-1} + r_f^{i+1}}{2};$$

$\delta = r_f^i - r_s^i;$

$if(\delta < 0)r_-^i \mathrel{-}= \delta;$

$if(\delta > 0)r_+^i \mathrel{+}= \delta;$

$\underline{fi};$

$\underline{od};$

$\underline{return}\, r_f, r_+, r_-$

Correction for shape feasibility
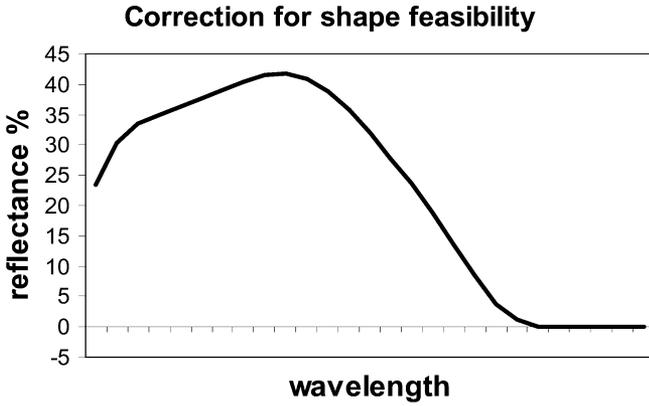
Fig. 4. Spectrum of Fig. 4(b) after the application of the procedure described in Table II.

method. We will start by introducing a class of linear transformations and their properties. Consider a linear transformation $\Re^N \to \Re^K$ ($K = 3$ being the case of interest here):

$$X^k = \sum_{i=1}^N a_i^k r^i \qquad (26)$$

and let r^i be expressed as the weighted sum of b vectors in $\Re^N$, as in (9). Then, putting together (9) and (26), one obtains

$$X^k = \sum_{i=1}^N \sum_{h=1}^b a_i^k \psi_h^i w^h = \sum_{h=1}^b l_h^k w^h \qquad (27)$$

where

$$l_h^k = \sum_{i=1}^N a_i^k \psi_h^i. \qquad (28)$$

Let $\Lambda = [1, \ldots, N]$, then one can define the matrix $[\Lambda] = \{l_h^k\}$ with $[\Lambda] : \Re^b \to \Re^K$. One can consider similar transformations

in which the summation in (28) is extended only to a subset of the indexes Λ. For instance, if $\Phi = [u, \ldots, v], \mathrm{u} \geqslant 1, \mathrm{v} \leqslant N$, one has

$$[\Phi] = \left\{\phi_h^k\right\} \quad \text{and} \quad \phi_h^k = \sum_{i=u}^v a_i^k \psi_h^i. \qquad (29)$$

The definition can be extended to any subset $\Phi \subseteq \Lambda$. Note that, for all the problems that we are considering here, $X^k \geqslant 0$, therefore

$$\|X\|_1 = \|[\Phi]w\|_1 = \sum_k ([\Phi]w)^k = \sum_{k=1}^K \sum_{h=1}^b \phi_h^k w^h. \qquad (30)$$

The following properties derive immediately from $a_i^k \geqslant 0$ and $\phi_h^i \geqslant 0$.

Lemma 4.1: For $\Phi, \Theta \subseteq \Lambda$

$$\Phi \cap \Theta = \emptyset \Rightarrow [\Phi \cup \Theta] = [\Phi] + [\Theta]. \qquad (31)$$

In particular, for $a_i^k \geqslant 0$

$$[\Lambda] = [\Phi] + [\bar{\Phi}] \qquad (32)$$

where $\bar{\Phi}$ is the complement of Φ in Λ.

Lemma 4.2: The 1-norm of $\sum_i a_i^k r^i$ is a norm for r, indicated as $\|r\|_a$ and, if $r^i = \sum_h w^h \psi_h^i$, and $[\Lambda]w = X$, then

$$\|r\|_a = \|[\Lambda]w\|_1 = \|X\|_1. \qquad (33)$$

We can now consider the normalization procedure. Let X_0 be the given color (that is, the initial tristimulus of the procedure), and consider the tristimulus (4). Writing r as in (9), the equation becomes $[\Lambda]w = X_0$. If $b = 3$, these equations can be solved for w, giving $w = [\Lambda]^{-1}X_0$ and the reflectance

$$r_0 = \Psi[\Lambda]^{-1}X_0 \qquad (34)$$

with $\Psi = \{\psi_h^i\}$. The reflectance, in general, will not have a range included in $[0, 1]$. Here, we assume, for the sake of simplicity, that $r^i \geqslant 0$, that is, that the only correction that we have to make are for those values $r^i > 1$. If there are parts with $r^i < 0$, one applies the transformation $r \leftarrow 1 - r$ to these parts, transforming them into parts with values grater than 1.

Define the set $\Phi_0 = \{i \mid r^i > 1\}$ and the functions u_f and u_+ as in (10) and (11), respectively, so that $r_0 = u_f(r_0) + u_+(r_0) = r_{0,f} + r_{o,+}$. Note that $r_{0,+} \geqslant 0$ and that the range of $r_{0,f}$ is in $[0, 1]$. With these definitions, we have

$$X_0^k = \sum_{i \in \Lambda} a_i^k r_o^i = \sum_{i \in \Lambda} a_i^k r_{0,f}^i + \sum_{i \in \Lambda} a_i^k r_{0,+}^i$$
$$= \sum_{i \in \Lambda} a_i^k r_{0,f}^i + \sum_{i \in \Phi_0} a_i^k \left(r_0^i - 1\right) \underline{\mathrm{def}} X_{0,f}^k + X_{0,+}^k \qquad (35)$$

where the symbol $\underline{\mathrm{def}}$ means "equal by definition." Note that $r_0^i > 1$ in Φ_0, and that the reflectances are limited ($r_i < M$) so that one can write

$$0 \leqslant \alpha_o^i = \frac{r_0^i - 1}{r_0^i} < \frac{M - 1}{M} < 1. \qquad (36)$$

In other words, one has

$$X_{0,+}^k = \sum_{i \in \Phi_0} a_i^k \left(r_0^i - 1 \right) = \sum_{i \in \Phi_0} \alpha_0^i a_i^k r_0^i < \alpha_0 \sum_{i \in \Phi_0} a_i^k r_o^i \tag{37}$$

with $\alpha_0 < 1$. The method now finds a reflectance function that generates $X_{0,+}$. Expressing this reflectance as

$$\tilde{r}_0^i = \sum_{h=1}^b \tilde{w}_0^h \psi_h^i \tag{38}$$

one can find the coefficients $\tilde{w}_0^h$ by solving the system $X_{0,+} = [\Lambda]\tilde{w}_0$. On the other hand, from (37) and (9), one has $X_{0,+} = \alpha_0[\Phi_0]w_0$, leading to the relation between w_0 and $\tilde{w}_0$

$$[\Lambda]\tilde{w}_0 = \alpha_0[\Phi_0]w_0. \tag{39}$$

In order to analyze this equation, we use the following property.

Lemma 4.3: Let $u^i = \sum_h x^h \psi_h^i$ and $v^i = \sum_h y^h \psi_h^i$ be two reflectance functions giving positive tristimulus values on the illuminant a_i^k, and let x and y satisfy $[\Lambda]x = [\Phi]y$ with $\Phi \subseteq \Lambda$. Then, $\|u\|_a \leqslant \|v\|_a$.

Proof: Since the functions are reconstructed from tristimulus values, which are positive by construction, (32) applies. Consider the kth term of the equality

$$([\Lambda]x)^k = ([\Phi]y)^k \leqslant ([\Phi]y)^k + ([\bar{\Phi}]y)^k = ([\Lambda]y)^k \tag{40}$$

where the last equality derives from Lemma 4.1. So

$$\|[\Lambda]x\|_a = \sum_k ([\Lambda]x)^k \leqslant \sum_k ([\Lambda]y)^k = \|[\Lambda]y\|_a. \tag{41}$$

From this and Lemma 4.2, the result follows.

We now replace $r_0 = r_{0,f} + r_{0,+}$ with $r_1 = r_{0,f} + \tilde{r}_0$ and repeat the iteration. At the end, we have

$$r_1 = r_{0,f} + \tilde{r}_0 = r_{1,f} + r_{1,+} \tag{42}$$

where $r_{1,f} = u_f(r_1)$ and $r_{1,+} = u_+(r_1)$.

Lemma 4.4: In (42), it is $\|r_{1,+}\|_a \leqslant \|\tilde{r}_0\|_a$.

Proof: All the values here are positive, so

$$\|r\|_a = \sum_{k=1}^K \sum_{i=1}^N a_i^k r^i. \tag{43}$$

From this, it follows that the proposition is true whenever $r_{1,+}^i \leqslant \tilde{r}_0^i$ for all i. We have the following cases:

i) $r_{1,f}^i < 1 \Rightarrow r_{1,+}^i = 0 \Rightarrow r_{1,+}^i \leqslant \tilde{r}_0^i$;

ii) $r_{1,f}^i = 1$ and $r_{0,f}^i = 1 \Rightarrow r_{1,+}^i = \tilde{r}_0^i$;

iii) $r_{1,+}^i = 1 \wedge r_{0,f}^i < 1 \Rightarrow r_{1,+}^i = r_1^i - r_{1,f}^i = r_{0,f}^i + \tilde{r}_0 - r_{1,f}^i < \tilde{r}_o^i$

where the symbol $\wedge$ means logical conjunction.

Lemma 4.5: In the iteration, it is $\|r_{1,+}\|_a \leqslant \alpha\|r_{0,+}\|_a$.

Proof: From Lemma 4.3 and (39), it follows that $\|\tilde{r}_0\|_a \leqslant \alpha\|r_{0,+}\|_a$. From this and Lemma 4.4, the result follows.

The main result of this section is given by the following theorem:

Theorem 4.1: Assume that a feasible spectrum exists, then the iteration

$$r(k) = u_f(r_{(k)}) + u_+(r_{(k)})\underline{def}r_{(k),f} + r_{(k),+};$$
$$r_{(k),+}^i = \sum_h w_{(k)}^h \psi_h^i; \quad [\Lambda]\tilde{w}_{(k)} = \left[\Phi_{(k)}\right] w_{(k)}$$
$$\tilde{r}_{(k)}^i = \sum_h \tilde{w}_{(k)}^h \psi_h^i; \quad r_{(k+1)} = r_{(k),f} + \tilde{r}_{(k)}$$

converges to a feasible solution.

Proof: Applying the previous lemmas to the kth iteration, one obtains

$$\|r_{(k),+}\|_a < \alpha\|r_{(k-1),+}\|_a \tag{44}$$

and, iterating

$$\|r_{(k),+}\|_a < \alpha^k\|r_{0,+}\|_a \tag{45}$$

therefore

$$\lim_{k \to \infty}\|r_{(k),+}\|_a = 0 \tag{46}$$

so that the equilibrium point of the method is

$$r_\infty = u_f(r_\infty) \tag{47}$$

which, by definition of u_f, is feasible.

The requirement that the bases be positive is too strict for many a practical situation: one commonly used basis for spectral reconstruction, for instance, is obtained by principal components analysis on a suitable training set of reflectances. In this case, of course, there is no guarantee that the basis will be positive. It is possible, however, to prove that the same convergence result holds for a wider class of basis functions.

Definition 4.1: A basis ψ_h^i is said to be positive enough with respect to an illuminant a_i^k if

$$\sum_{i=1}^N a_i^k \psi_h^i > 0 \tag{48}$$

for all k. The most important property of positive enough functions, and the basis for the convergence result, is given by the following theorem.

Theorem 4.2: Let ψ_h^i be positive enough with respect to a_i^k. Then, it is possible to write $\psi_h^i = \hat{\psi}_h^i + \tilde{\psi}_h^i$, with $\hat{\psi}_h^i \geqslant 0$, and $\sum_i a_i^k \tilde{\psi}_h^i = 0$ for all k.

Before proving this theorem, we introduce a lemma that will be used in the proof.

Lemma 4.6: Let $A = \{a_i^k\}, i = 1, \ldots, N, k = 1, \ldots, K$ be a matrix with $N > K$ and $\text{rank}(A) = K$. Then, there is a matrix $Z = \{\varsigma_k^i\}$ such that $\varsigma_k^i \geqslant 0$ and $T = \{t_h^k\}$ is nonsingular, where

$$t_h^k = \sum_i a_i^k \varsigma_k^i. \tag{49}$$

Proof: Since the matrix A has full rank, there are one or more sets of columns $\mu = [i_1, \ldots, i_k]$ such that the matrix

$$A_\mu = \begin{bmatrix} a_{i_1}^1 & \cdots & a_{i_k}^1 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{i_1}^K & \cdots & a_{i_K}^K \end{bmatrix} \tag{50}$$

is nonsingular. Let $M = \{\mu 1, \ldots, \mu M\}$ the set of indexes μ for which this is true. Define the matrix

$$\varsigma[\mu]_l = \left[\overbrace{0, \ldots, 0, 1}^{i_l}, 0, \ldots, 0 \right]^T. \tag{51}$$

Clearly, all the $\varsigma[\mu]_l$ are independent, since all the values i_l in the same index μ are different, and $AZ[\mu] = A\mu$. For $\alpha_1, \ldots, \alpha_M$ with $\alpha_\mu \geqslant 0$, define

$$Z = \sum_\mu \alpha_\mu Z[\mu] \tag{52}$$

so that

$$AZ = \sum_\mu \alpha_\mu A_\mu \tag{53}$$

any combination of α_μ for which $\sum_\mu \alpha_\mu |A_\mu| \neq 0$ will generate the required nonsingular matrix $T = AZ$.

Proof of the Theorem: Let $q_h^k = \sum_i a_i^k \psi_h^i$, and consider a basis of K ($K = 3$ in our case) sampled functions ς_l^i such that the matrix $Z = \{\varsigma_l^i\}$ satisfies the previous lemma, and define

$$\hat{\psi}_h^i = \sum_l \alpha_h^l \varsigma_l^i \tag{54}$$

where the coefficients α_h^l are determined by the equalities

$$q_h^k = \sum_i a_i^k \hat{\psi}_h^i = \sum_l \alpha_h^l \sum_i a_i^k \varsigma_l^i = \sum_l \alpha_h^l t_l^k. \tag{55}$$

Because of Lemma 4.6, it is possible to find positive functions ς_l^i such that the matrix T in the last equality is nonsingular and such that the α_h^l are positive. Defining $\tilde{\psi}_h^i = \psi_h^i - \hat{\psi}_h^i$ completes the proof. Consider now a system such as (27):

$$X^k = \sum_i a_i^k r^i = \sum_h \left(\sum_i a_i^k \psi_h^i \right) w^h. \tag{56}$$

Since $\psi_h^i = \hat{\psi}_h^i + \tilde{\psi}_h^i$, with $\sum_i a_i^k \tilde{\psi}_h^i = 0$, the system can be rewritten as

$$X^k = \sum_h \left(\sum_i a_i^k \hat{\psi}_h^i \right) w^h. \tag{57}$$

In other words, the solution of the system depends only on the positive component of the basis ψ_h: the solution that we find is the same that we would find if, instead of the basis ψ_h, we were using the positive basis $\hat{\psi}_h$. The considerations that were made

based on the positivity of ψ_h, therefore, still apply; in particular, (39) still applies. From this point on, in the original demonstration, the result depend only on the positivity of expressions of the type $\sum_i a_i^k r^i$ and $\sum_i a_i^k \psi_h^i$, which hold by construction and by the assumption that the ψ_h are positive enough. From this, therefore, we have the following result.

Theorem 4.3: The result of Theorem 4.1 still holds if the basis ψ_h is positive enough with respect to the illuminant a_i^k.

V. EVALUATION OF THE METHOD

Our method is an iterative procedure that reduces an error at every step and, from a purely mathematical point of view, two questions arise quite naturally: how fast does it converge for a desired final error, and what is the final error if a certain execution time is allocated. In this section, we report on our measurements of speed and quality comparing the results of our methods with those of another method that can be used to solve the same problem. Finally, we will try to determine the properties of the reconstructed spectrum by looking at the color that the spectrum determines under a change of illuminant. Before we explain our tests, we give a brief introduction to the use of linear programming to solve the spectral reconstruction problem, as linear programming is the method against which we measure our own.

A. Feasible Spectra by Linear Programming

Linear programming is a well-known optimization technique that can be used to solve constrained optimization problems where both the cost functional and the constraints are linear, and the solutions are required to be positive [18]. That is, problems in which one seeks a vector $x \in \Re^n$ such that

$$f = c^T x \quad c \in \Re^n \text{ is minimal}$$
$$Ax = b \quad A \in \Re^{m \times n} \quad b \in \Re^m, x_i \geqslant 0 \tag{58}$$

where A, b, and c are given constants. Note that this is not exactly the form of our reconstruction problem, since we don't quite have a cost to minimize and, in the case of strong feasibility, we have the additional constraint $x_i \leqslant 1$. Both these discrepancies can be eliminated by adopting the standard linear programming strategy of introducing additional variables. We can enforce the constraints $x_i \leqslant 1$ by introducing the variables $y_i, i = 1, \ldots, n$ and the constraints

$$x_i + y_i = 1, y_i \geqslant 0. \tag{59}$$

The introduction of these variables is equivalent to solving the linear programming problem with conditions $A'w = b', w_i \geqslant 0$, where $w = [x^T | y^T]^T \in \Re^{2n}$

$$A' = \begin{bmatrix} A & 0 \\ I_n & I_n \end{bmatrix} \in \Re^{(n+n) \times 2n} \tag{60}$$

where I_n is the $n \times n$ identity matrix, and $b' = [b^T | 1, \ldots 1]^T \in \Re^{m+n}$.

The lack of a function to be minimized can be solved by introducing the $m+n$ additional variables z_i, the conditions $z = b' - Aw$, i.e., $z + A'w = b'$, and the cost $f = \sum_i z_i = (1, \ldots 1) \cdot z$.

The cost is minimal for $z = 0$, and reaching this minimum entails finding a w for which the original condition is satisfied. In conclusion, in order to find a feasible solution, we solve the linear programming problem

$$\begin{aligned} \min c^T v \\ \bar{A}v = b' \\ v_i \geqslant 0 \end{aligned} \qquad (61)$$

where

$$v = [x^T | y^T | z^T]^T \in \Re^{3n+m}$$
$$b' = [b^T | 1, \dots 1]^T \in \Re^{m+n}$$
$$c = \left[\overbrace{0, \dots, 0}^{2n} \overbrace{1, \dots, 1}^{n+m} \right]^T \in \Re^{3n+m} \qquad (62)$$

and

$$\bar{A} = \left[\begin{array}{cc|c} A & 0 \\ I_n & -I_n \end{array} \middle| I_{n+m} \right] \in \Re^{(m+n)\times(3n+m)}. \qquad (63)$$

We solve this using the standard techniques that can be found in the literature.

B. Feasible Spectra by the Method Proposed

In all, we do four separate measurements: execution time, colorimetric reconstruction, curvature of the reconstructed spectrum, and color evaluation under change of illuminant. The latter experiment follows a rather different experimental methodology than the first three, and will be considered separately in a different sub-section. While the ultimate quality gauge for an algorithm such as this one lies in its reconstruction quality, execution time is important for applications such as computer graphics, in which the spectral reconstruction will have to be repeated tens of hundreds of thousands times per image. In order to give a general idea of what could be the advantages of the method in the area, we have included the execution time test. The two algorithms are implemented in C; for the simplex method, we used the program provided in a reference book [40]—we also tried the implementation in a second reference book [18] with virtually indistinguishable results. All reconstructions assumed the CIE D65 illuminant. The basis set adopted was obtained applying PCA to the Vrhel's dataset of objects reflectance spectra [41]. For each of the first three experiments, we considered four different vector sizes: we started with the color matching functions and the illuminant covering the range from $\lambda = 400$ nm to $\lambda = 700$ nm in steps of 10 nm, resulting in vectors of size $N = 31$, and then the experiment was repeated by down-sampling these vectors with steps 2, 3, and 4, obtaining vectors of dimension $N = 16, 11, 8$, respectively. Each experiment consisted of 30 trials, from which our statistics are derived; in each trial, we generated a random RGB color, converted it to XYZ, and run the trial on the two methods using the same color. For the *time execution experiment*, the two algorithms were compiled using the lcc compiler with all optimizations turned off. In order to avoid interference from other processes running on the same computer, the time execution experiment was done under the DOS operating system. The two algorithms were compiled with the optimizer turned off to avoid having some particularity

of the compiler used here affect the reproducibility of the results. Moreover, due to the unreliability of measuring short time intervals with the standard C timing functions, each trial consisted in the separate repetition of each algorithm 10^6 times. We indicate in the following the execution time as t_e. The *colorimetric reconstruction experiment* attempts to give a measure of the residual error in the reconstruction of the spectrum by computing back the tristimulus from which the reconstruction had started and comparing the reconstructed color with the original one. Given a color X, a method is used to compute a feasible reflectance and then to this reflectance the tristimulus (4) is applied to obtain a reconstructed color $\tilde{X}$. The outcome of the experiment is a reconstruction error:

$$\Delta_{\mathrm{XYZ}} = \left[\sum_k (X^k - \tilde{X}^k)^2 \right]^{\frac{1}{2}}. \qquad (64)$$

The last experiment attempts to give a measure of the quality of the reflectance spectra obtained using the two methods. As discussed in Section III, reflectance functions of natural surfaces or the spectra that we can physically construct are always smooth and relatively slowly varying. So, we take as a measure of the quality of the reconstructed spectra the average of the estimated curvature, that is

$$\Delta_c = \frac{1}{N-2} \sum_{i=2}^{N-1} |r^{i+1} - 2r^i + r^{i-1}|. \qquad (65)$$

In this case, we did not apply the method in Section III to ensure the naturalness of reflectance functions.

C. Results and Discussion

The results of our three tests are reported in Tables III–V. In all the tables, the first column contains the size of the vector used for that test, the second column the number of degrees of freedom for the mean comparison test (which is equal to the sum, for the two methods, of the number of observations minus one), the remaining eight columns contain the summary data for the two methods, and the p-value of the Wilcoxon rank sum test [42] for that experiment. Table III contains the results of the time execution test.

All the tests show statistically significant differences. In the case of the vector size $N = 31$ and $N = 16$, the method proposed here appears to perform more than two times faster than linear programming. The large variance is due to the fact that for a few colors both methods (the simplex and our own) fail to converge to the error specified but, when they converge, our method converges in a few iterations. Table IV contains the results of the reconstruction error test. The difference between the reconstruction errors is statistically significant, and, by the average, the method presented here appears to perform better for the majority of vector sizes.

The reason for the high variance of the results, shown also by the high values of the maxima of the errors, is due to the fact that occasionally the methods failed to converge in the allotted number of iterations. Note, however, that the maxima for the method presented here are lower than for linear programming. It appears, in other words, that our method creates relatively rapidly an acceptable solution to the reconstruction

TABLE III
RESULTS OF THE TIME EXECUTION TEST. TIMES ARE IN μs. N IS THE DIMENSIONALITY OF THE SPECTRA BEING CONSIDERED. EACH TRIAL CONSISTS OF 300
RANDOMLY GENERATED RGB COLORS, FROM WHICH SPECTRA ARE SYNTHESIZED

		Our method				Simplex				
N	d.o.f.	μ_i	σ_i	$\min_i$	$\max_i$	μ_i	σ_i	$\min_i$	$\max_i$	p-value
31	598	22.93	35.13	0.00	200.00	47.60	7.15	40.00	90.00	<<0.001
16	598	13.53	21.09	0.00	140.00	35.10	6.15	20.00	50.00	<<0.001
11	598	17.07	26.83	0.00	110.00	21.07	3.09	20.00	30.00	<<0.001
8	598	6.87	8.23	0.00	90.00	12.83	4.59	0.00	20.00	<<0.001

TABLE IV
RESULTS OF THE RECONSTRUCTION ERROR, THAT IS THE EUCLIDEAN DISTANCE BETWEEN THE ACTUAL XYZ VALUES AND THE XYZ VALUES OF THE
RECONSTRUCTED SPECTRA (64). ALL RESULTS ARE $\times 10^3$

		Our method				Simplex				
N	d.o.f.	μ_i	σ_i	$\min_i$	$\max_i$	μ_i	σ_i	$\min_i$	$\max_i$	p-value
31	598	2.51	12.87	0.00	128.67	12.92	80.07	0.00	733.54	<<0.001
16	598	4.14	17.54	0.00	167.72	3.99	42.06	0.00	657.18	0.01
11	598	0.96	7.26	0.00	111.42	1.15	19.85	0.00	343.79	<<0.001
8	598	1.80	11.46	0.00	147.01	3.52	43.10	0.00	539.30	0.02

TABLE V
RESULTS OF THE CURVATURE TEST

		Our method				Simplex				
N	d.o.f.	μ_i	σ_i	$\min_i$	$\max_i$	μ_i	σ_i	$\min_i$	$\max_i$	p-value
31	598	0.02	0.01	0.00	0.08	8.33	3.29	0.64	16.15	<<0.001
16	598	0.14	0.08	0.01	0.45	8.36	8.38	0.62	44.58	<<0.001
11	598	0.38	0.21	0.04	1.12	6.41	5.69	0.41	55.49	<<0.001
8	598	0.64	0.38	0.05	2.49	23.75	15.19	0.85	59.62	<<0.001

problem so that, even if convergence is not reached within the allotted number of iterations, a reasonable solution is still found. Note that some of these problems might be solved by allowing a greater number of iterations but, as a matter of methodology, we decided to use here the same number of iterations that we used for the execution time test.

Finally, the curvature comparison is reported in Table V. In this case, all the differences are statistically significant, and the method presented here produced results with a lower curvature, that is, smoother results, than linear programming. This was, of course, expected, since linear programming simply tries to find a vector that solves the optimization problem, without considering the vector as a sampled function and, therefore, without trying to enforce any relation between contiguous elements. The method presented here, on the other hand, derives the correction vector by sampling smooth basis functions and, therefore, produces smooth results. In order to give a visual explanation

of the advantages of the method we propose with respect to a simple clipping of the initial solution obtained by a linear method, we report the plot of an experiment performed on a subset of the Munsell samples derived by [38]. Fig. 5 shows the section of the Munsell set of samples at the same level of brightness (value = 5) in the CIE L*a*b* diagram. These data form the input for a recovery procedure based on the solution of (6) using the basis set published in [43]. The recovered spectra, when reflectance values were out of the range [0, 1] were clipped to ensure physical feasibility. The clipped spectra were then rendered under the CIE standard illuminant C, and the CIELAB coordinates of the colors computed. Results on the color opponent plane are reported in Fig. 6(a). Note that the gamut of colors has been significantly compressed. We then performed the same experiment but applying our method to correct the spectra instead of the clipping. Results are reported in Fig. 6(b). Note the difference in the gamut extension respect with the case of Fig. 6(a).
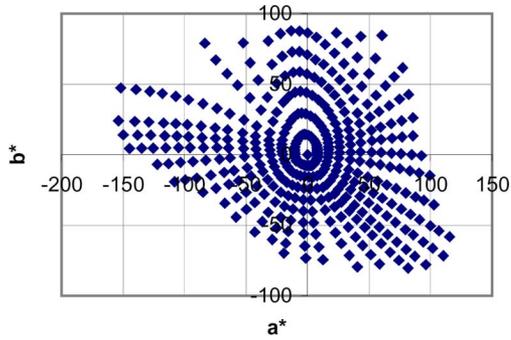
Fig. 5. Subset (value $= 5$) of the Munsell set of colors in the CIE L*a*b* color space. These data, converted in tristimulus values assuming the CIE Standard Illuminant C are the input to the reconstruction experiment.
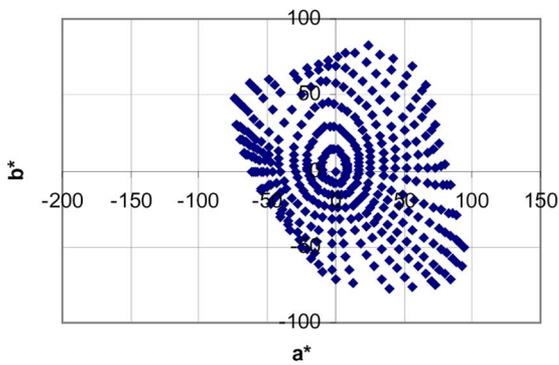
(a)
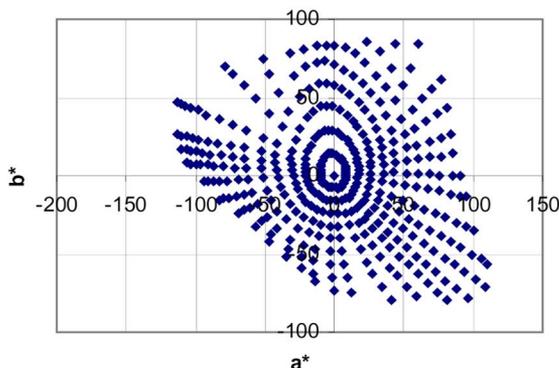
(b)

Fig. 6. Rendering of the estimated reflectance spectra of the Munsell samples under the illuminant C. (a) Reflectance values out of the range [0, 1] were clipped to ensure physical feasibility. (b) Method presented in this paper was applied. Values are in CIE L*a*b* coordinates.

In the first three experiments, for the sake of control and repeatability, we used a random set of colors and a fixed illuminant. In this situation, there is no guarantee that any natural object, or "natural" reflectance will be able to generate a given color under the assumed illuminant. There is some indications that the areas in which the method shows a slow convergence correspond to rather irregular reflectance functions: since the reflectance functions that the method builds are created by repeated addition and clipping of linear combinations of the bases, which are very regular and smooth, rendering a highly irregular reflectance is very difficult, and it may require many iterations.
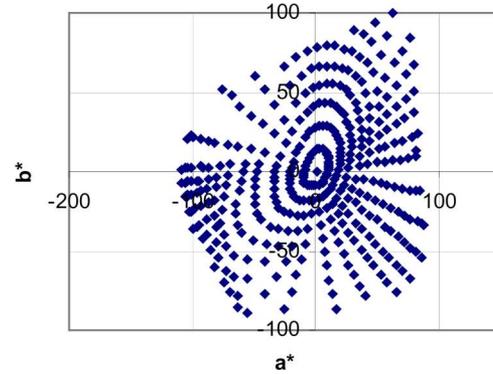
(a)
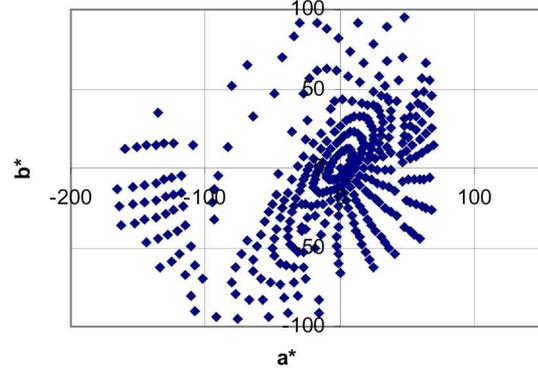
(b)

Fig. 7. (a) Colors in the Munsell plane (value $= 5$) rendered after spectral reconstruction with the method proposed under the A illuminant. (b) Colors in the Munsell plane rendered after spectral reconstruction with the simplex method and under the A illuminant.

Our method, on the other hand, starts with a nonfeasible solution and builds increasingly feasible approximations of it. If the method is interrupted before it converges, it will produce a feasible reflectance by "clipping" whatever approximation it has at that time. This behavior tends to provide reasonable (although not strictly metameric) approximations of the true reflectance and to provide acceptable solutions even in cases in which convergence is not reached. We owe to this, for instance, the fact that even the points in Fig. 7 for which the iterations were interrupted are not placed in wildly wrong positions of the plane. This seems to be the case, on the other hand, with the simplex method: whenever the method cannot converge in the allotted number of iterations, the approximation of the reflectance provided is strongly nonmetameric to the initial one, and the corresponding point can be virtually everywhere in the color plane.

D. Change of Illuminant

There is, of course, more to spectral feasibility than having the range included in [0, 1]: the spectra of natural objects tend to be smooth, for example. The curvature experiment in the previous sections tried to measure the extent to which this "physical" characteristic of naturally occurring spectra was present in the spectra reconstructed with the different methods. In order to have a better idea of the behavior of the spectra, however, it is desirable to have some form of colorimetric assessment, which is what we propose to do in this section. The idea is the following:

we consider a plane of the Munsell color system with constant value (we have taken the plane with value equal to 5) thus selecting a subset of samples in the Munsell dataset. For each selected color, we do the spectral reconstruction using a certain illuminant (the CIE standard illuminant C in our case). Then, we render the colors under a different illuminant (the CIE standard illuminant A), and compute corresponding CIE L*a*b* values. We then plot the result on a constant lightness plane. If the spectral reconstruction generated "reasonable" spectra, we should expect that the *pattern* that we observe in the Munsell plane should not change too much when the illuminant is changed. The rationale for this procedure is that changes in illuminant should not change the relative arrangement of the Munsell chips in the (a*,b*)-diagram. Results of this experiment using the proposed method and the Simplex method are displayed in Fig. 7(a) and (b), respectively.

The patterns show that the distortions caused by the simplex method are significantly larger than those caused by the present method. The main distortion of the present method, apart from translation, rotation, and shear (shear is in any case much less pronounced for our method than for the simplex) appears to be a compression of the two "long arms" of the Munsell plane corresponding to low values of b* and negative values of a*. In many cases, this is due to a slow convergence of the method: our test program would cut the method after 30 iterations, even if the point of convergence had not been reached. By comparison, the simplex method causes heavy distortion of the whole portion b∗ > 0 of the Munsell plane, to the point that the pattern is almost impossible to detect in this region.

VI. CONCLUSION

In this paper we proposed a solution to correct the outcome of reflectance recovery methods, in order to ensure the physical feasibility and naturalness of estimated reflectance functions. Our method is iterative, and converges to a feasible metamer of the initial recovered reflectance function. As future research, we would like to investigate the connection between the type of surfaces, the basis functions adopted, and our method.

REFERENCES

[1] G. Buchsbaum and A. Gottschalk, "Trichromacy, opponent colours coding and optimum colour information transmission in the retina," *Proc. Roy. Soc. London B.*, vol. 220, no. 1218, pp. 89–113, 1983.

[2] V. Cheung, S. Westland, C. Li, J. Hardeberg, and D. Connah, "Characterization of thrichromatic color cameras by using a new multispectral imaging technique," *J. Opt. Soc. Amer. A*, vol. 22, no. 7, pp. 1231–1240, Jul. 2005.

[3] G. Sharma and S. Wang, "Spectrum recovery from colorimetric data for color reproductions," *Color Imaging VII: Device-Independent Color, Color Hardcopy, and Applications, Proc. SPIE*, vol. 4663, Jan. 2002.

[4] M. J. Vrhel and H. J. Trussel, "Color correction using principal components," *Col. Res. App.*, vol. 17, no. 5, pp. 328–338, 1992.

[5] S. Zuffi and R. Schettini, "Using recovered reflectance to predict color," *Color Imaging X: Processing, Hardcopy, and Applications, Proc. SPIE*, vol. 5667, pp. 47–52, Jan. 2005.

[6] G. Sharma and H. J. Trussel, "Digital color imaging," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 901–932, 1997.

[7] W. S. Stiles and G. Wyszecki, "Counting metameric object-color," *J. Opt. Soc. Amer.*, vol. 52, pp. 313–327, 1962.

[8] W. S. Stiles, G. Wyszecki, and N. Otha, "Counting metameric object-color stimuli using frequency-limited spectral reflectance functions," *J. Opt. Soc. Amer.*, vol. 67, pp. 779–784, 1997.

[9] G. Buchsbaum and A. Gottschalk, "Chromaticity coordinates of frequency-limited functions," *J. Opt. Soc. Amer. A*, vol. 1, no. 8, pp. 885–887, Aug. 1984.

[10] V. Bonnardel and L. T. Maloney, "Daylight, biochrome surfaces, and human chromatic response in the Fourier domain," *J. Opt. Soc. Amer. A*, vol. 17, no. 4, pp. 677–686, 2000.

[11] L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with a small number of parameters," *J. Opt. Soc. Amer.*, vol. 3, pp. 1673–1683, 1986.

[12] E. L. Kirnov, "Spectral reflectance properties of natural formations," National Research Council of Canada, Ottawa, ON, Canada, Technical Translation TT-439, 1947.

[13] J. Romero, E. Valero, J. Hernàndez-Andrés, and J. L. Nieves, "Color-signal filtering in the Fourier-frequency domain," *J. Opt. Soc. Amer. A*, vol. 20, no. 9, pp. 1714–1724, Sep. 2003.

[14] D. B. Judd, D. L. McAdam, and G. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," *J. Opt. Soc. Amer.*, vol. 54, no. 8, pp. 1031–1040, 1964.

[15] B. A. Wandell, "The synthesis and analysis of color images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 1, pp. 2–13, Jan. 1987.

[16] F. Cheng, W. Hsu, and T. Chen, "Recovering colors in an image with chromatic illuminant," *IEEE Trans. Image Process.*, vol. 7, no. 11, pp. 1524–1533, Nov. 1998.

[17] R. Schettini and B. Barolo, "Estimating reflectance functions from tristimulus values," *Appl. Signal Process.*, vol. 3, pp. 104–115, 1996.

[18] W. H. Press, B. P. Flannery, S. A. Teulolsky, and W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing.* Cambridge, U.K.: Cambridge Univ. Press, 1986.

[19] D. Tzeng and R. S. Berns, "A review of principal component analysis and its applications to color technology," *Col. Res. App.*, vol. 30, no. 2, pp. 84–98, Apr. 2005.

[20] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, pp. 411–430, 2000.

[21] J. Cohen, "Dependency of the spectral reflectance curves of the munsell color chips," *Psychonomic Sci.*, vol. 1, pp. 369–370, 1964.

[22] K. L. Kelley, K. S. Gibson, and D. Nickerson, "Tristimulus specification of the Munsell Book of Color from spectrophotometric measurements," *J. Opt. Soc. Amer. A.*, vol. 33, pp. 355–376, 1943.

[23] T. Jaaskelainen, J. Parkkinen, and S. Toyooka, "A vector-subspace model for color representation," *J. Opt. Soc. Amer. A*, vol. 7, pp. 725–730, 1990.

[24] Q. Sun and M. D. Fairchild, "Statistical characterization of spectral reflectances in spectral imaging of human portraiture," in *Proc. IS&T/SID 9th Color Imaging Conf.*, Nov. 2001, pp. 73–79.

[25] R. Ramanath, R. G. Kuehni, W. E. Snyder, and D. Hinks, "Spectral spaces and color spaces," *Col. Res. Appl.*, vol. 29, no. 1, 2004.

[26] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.

[27] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," *Adv. Neural Inf. Process. Syst.*, vol. 13, pp. 556–562, 2001.

[28] D. H. Marimont and B. A. Wandell, "Linear models of surface and illuminant spectra," *J. Opt. Soc. Amer. A*, vol. 9, no. 11, pp. 1905–1913, 1992.

[29] J. M. DiCarlo and B. A. Wandell, "Spectral estimation theory: Beyond linear but before Bayesian," *J. Opt. Soc. Amer. A*, vol. 20, no. 7, pp. 1261–1270, Jul. 2003.

[30] M. S. Drew and B. V. Funt, "Natural metamers," *CVGIP: Imag. Understand*, vol. 56, no. 2, pp. 139–151, Sep. 1992.

[31] S. Zuffi and R. Schettini, "Reflectance functions estimation from tristimulus values," *Color Imaging IX: Processing, Hardcopy, and Applications, Proc. SPIE*, vol. 5293, pp. 222–231, Jan. 2004.

[32] D. Dupont, "Study of the reconstruction of reflectance curves based on tristimulus values: Comparison of methods of optimization," *Col. Res. App.*, vol. 27, no. 2, pp. 88–99, 2002.

[33] C. J. Hawkyard, "Synthetic reflectance curves by additive colour mixing," *J. Soc. Dyers Colour*, vol. 109, pp. 246–251, 1996.

[34] G. Wang, C. Li, and M. R. Luo, "Improving the Hawkyard method for generating reflectance functions," *Col. Res. App.*, vol. 30, no. 4, pp. 283–287, 2005.

[35] C. Li and M. R. Luo, "The estimation of spectral reflectances using the smoothness constraint condition," in *Proc. IS&T/SID 9th Color Imaging Conf.*, Scottsdale, AZ, Nov. 2001, pp. 62–67.

[36] C. van Trigt, "Smoothest reflectance functions. I. Definition and main results," *J. Opt. Soc. Amer. A*, vol. 7, no. 10, pp. 1891–1904, 1990.

[37] C. van Trigt, "Smoothest reflectance functions. II. Complete results," *J. Opt. Soc. Amer. A*, vol. 7, no. 12, pp. 2208–2222, 1990.

[38] G. Wyszecki and G. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed.   New York: Wiley, 1982.

[39] P. Morovic and G. D. Finlayson, "Reflectance estimation with uncertainty," presented at the 10th Congr. Int. Colour Assoc. (AIC Colour 05), Granada, Spain, May 2005.

[40] H. M. Antia, *Numerical Methods for Scientists and Engineers*, 2nd ed.   Cambridge, MA: Birkaÿser Verlag, 2001.

[41] M. J. Vrhel, R. Gershon, and L. S. Iwan, "Measurement and analysis of object reflectance spectra," *Col. Res. App.*, vol. 19, no. 1, pp. 4–9, 1994.

[42] M. Hollander and D. Wolfe, *Nonparametric Statistical Methods*, 2nd ed.   New York: Wiley, 1999.

[43] H. S. Fairman and M. H. Brill, "The principal components of reflectances," *Col. Res. App.*, vol. 29, no. 2, pp. 104–110, 2004.

Simone Santini (M'98) received the Laurea degree from the University of Florence, Italy, in 1990 and the M.Sc. and Ph.D. degrees from the University of California, San Diego (UCSD) in 1996 and 1998, respectively.

In 1990, he was a Visiting Scientist at the Artificial Intelligence Laboratory at the University of Michigan, Ann Arbor, and, in 1993, he was a visiting scientist at the IBM Almaden Research Center. He has been a Project Scientist in the Department of Electrical and Computer Engineering, UCSD, and a Researcher at Praja, Inc., San Diego, CA. Since 2004, he has been an Associate Professor at the Universidad Autónoma de Madrid, Spain. His current research interests are interactive image and video databases, behavior identification and event detection in multisensor stream, query languages for event-based multimedia databases, and evaluation of interactive database systems.

Raimondo Schettini is an Associate Professor at DISCo, University of Milano Bicocca, Milan, Italy, where he is in charge of the Imaging and Vision Laboratory. He has been associated with the Italian National Research Council (CNR) since 1987. He has been team leader in several research projects and has published more than 170 refereed papers on image processing, analysis and reproduction, and on image content-based indexing and retrieval.

He is an Associate Editor of the *Pattern Recognition Journal*. He was a co-Guest Editor of three special issues on Internet Imaging (*Journal of Electronic Imaging*, 2002), Color Image Processing and Analysis (*Pattern Recognition Letters*, 2003), and Color for Image Indexing and Retrieval (*Computer Vision and Image Understanding*, 2004). He was General Co-Chairman of the First Workshop on Image and Video Content-Based Retrieval (1998); the First European Conference on Color in Graphics, Imaging and Vision (2002); the EI Internet Imaging Conferences (2000–2006); and the EI Multimedia Content Access: Algorithms and Systems 2007 Conference (2007).

Silvia Zuffi (M'07) was born in Rimini, Italy, on July 3, 1969. She received the Laurea degree in electronic engineering from the University of Bologna, Italy, in 1995.

From 1995 to 1997, she was a Computer Software Engineer, and then she joined the Movement Analysis Laboratory at the Istituti Ortopedici Rizzoli, Bologna, Italy. Since 1999, she has been a Research Scientist in Italian National Research Council (CNR), Milan, Italy. Her research interests include applications of color in graphics, color reproduction, multispectral imaging, and color perception.