Removal of Achromatic Reflections from a Single Color Image

L. Bedini, P. Savino, and A. Tonazzini*
Istituto di Scienza e Tecnologie dell'Informazione
Area della Ricerca CNR di Pisa
Via G. Moruzzi, 1, I-56124 PISA, Italy

In this paper we consider the problem of removing achromatic reflections from a picture of a scene taken through a semi-transparent medium, assuming that the reflection pattern is due to a light source or another object located in front of the object of interest. While other works assume the availability of multiple observations, we consider the more challenging problem of having as data a single color image. We suppose a data model where the virtual reflected image combines additively with the real transmitted image of the object, through unknown coefficients. This highly underdetermined problem is handled by means of a blind estimation technique that exploits the strict dependence of the gradients of the three color channels of the ideal image, and their independence from the gradient of the grayscale reflected image. The model parameters are estimated through independent component analysis, and then the component images are estimated through a regularization technique. The whole algorithm is very fast, and its performance is quantitatively evaluated on numerically generated images, and qualitatively tested on real images.

Keywords: reflection removal, color image enhancement, independent component analysis, image regularization

PACS numbers:

I. INTRODUCTION

When we take a picture through a semi-transparent medium, glass, resin or plastic, we observe an image that often is a superposition of the image of the object beyond the medium and the image of the scene or a light source located in front of the object and reflected by the medium. We call transmitted image the ideal image of the object of interest, and reflected image, or reflection, the image of this second object. The need to remove the reflected image from the observed one is common for both recreational and professional purposes, either for improving the human vision or as a first step for subsequent image processing tasks. Figure 1 shows two examples of images with reflection. The first one is a picture of an outdoor scene taken through a window glass, where the superposed image is the reflection of a house furnishing. The second one is the picture of a manuscript that reflects a light source. This example is representative of pictures taken in the field of Cultural Heritage, for digitization purposes. Indeed, important historical artworks are often conserved behind semi-transparent media. In addition, in the past, in an attempt to preserve them, manuscripts and drawings were covered with chemical substances producing a sort of semi-transparent plastic coat, which reflects light sources.

*Electronic address: anna.tonazzini@sti.cnr.it

A. Related works

Professional photographers use polarizing lens to be rotated to reduce the intensity of the reflection. Polarimetric imaging systems [1] [2], which incorporate a polarizer directly in the optics, such as cameras equipped with a liquid crystal polarizer [3], can even totally eliminate reflection [4][5]. However, this can be achieved under a condition that is difficult to be satisfied, i.e. that the viewing angle is equal to the Brewster angle [6]. In addition, although nowadays in the most important museums, libraries and archives specialized digitization equipments are available, the use of polarizers is not yet a common practice.

The majority of the proposed fully computational approaches assume that the observed image can be considered as a linear combination of the reflected and transmitted images. That is, the observed image is an unknown linear mixing of two unknown images. This model was derived in [7], by analyzing optical models. Mathematically, the problem of recovering the transmitted image from the observed image is highly ill-posed since also the coefficients of the linear combination are unknown and the number of unknowns is twice the number of equations. Blind statistical methods of independent components analysis (ICA) can be used to handle this kind of underdetermined problems. However, they need the availability of as many mixtures as the number of components to be separated. In some works ICA has been attempted to remove reflection from pairs of images of

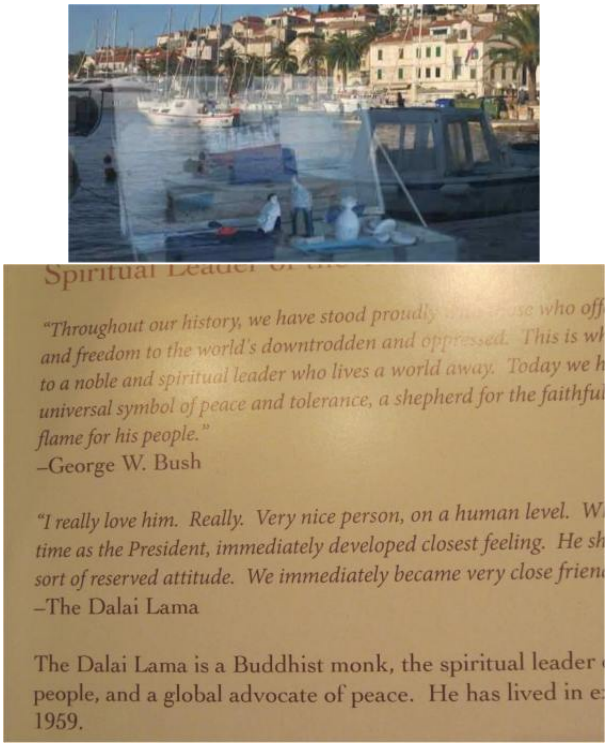


FIG. 1: Examples of images with reflection.

a same scene acquired with two different polarizations [7]. To improve separation, sparse ICA (SPICA) has also been used [8], which has the further ability to handle multiple source images with different motions [9][10][11]. For example, in [12], the relative motion between the two components in an image sequence is exploited, together with sparsity and independence of the gradient fields, to decompose the transmitted and reflected images. In [13] the physical properties of polarization for a double-surfaced glass medium are exploited within a multiscale scheme to separate the reflection from the transmitted background scene, using three polarized images, each captured from the same view point but with a different polarizer angle separated by 45 degrees.

With these approaches satisfactory removal of reflections has been obtained, however, as already mentioned, they require two or more observations. Instead, when only a single image is available, the problem remains highly ill-posed, so that more strict constraints on the problem formulation or in the component images must be exploited.

In [14], for example, it is assumed that the single acquired image is affected by shifted double reflections of the scene off a double-pane glass surface. Thus, exploiting the asymmetry between the layers, the ghosted reflection is modeled using a double-impulse convolution kernel, and the spatial separation and relative attenuation of the ghosted reflection components is automatically estimated. To separate the layers, an algorithm that uses a Gaussian Mixture Model for regularization is proposed that only requires a single input image.

In [15] the problem is handled using local features, and, in [16], using priors describing sparsity and user provided information. The dependency of the color channels of the transmitted image, and their independence from the achromatic reflection, is proposed instead as constraint in [17], where a MAP estimation approach is adopted, which takes also into account for the regularity of the images.

In all the methods above the model adopted is stationary, that is the mixing coefficients are the same for every pixel, which can be unrealistic in real-life scenarios. Thus, other methods consider spatially varying mixtures. For the case of two views, in [18] the reflected components are separated through information exchange among the various components, and, for a single view, in [19] information about the differences in the structure of the transmitted and the reflected image has been added to the basic Bayesian model of [17].

B. Our contribution

In this paper, we assume the stationary linear mixing model of [17], whose coefficients are unknown, and, as done also elsewhere, that the reflection is an achromatic, i.e. graylevel, image. The high under-determination of the problem is overcome by making use of reasonable constraints on the practical coincidence of the gradients of the three color channels of the normalized transmitted image, and the statistical independence of those gradients from the gradient of the normalized reflected image. We then propose a very fast algorithm constituted of two subsequent steps, both based on the above described constraints. The first step estimates the model parameters through an ICA algorithm. The second step, based on the now determined data model, estimates the four component images via regularization techniques.

Due to the permutation indetermination typical of ICA, the output of the first step consists of two sets of model parameters: in one the order of the coefficients is correct, in the other is inverted, without any possibility to discriminate them. Then, in the second step regularization is applied twice, and two sets of solutions are obtained. These solutions give rise to two pairs of images, the transmitted image and the reflected image. The correct pair of images is chosen by visual inspection.

The overall method is sketched in the flow chart shown in Fig. 2.

A shorter version of this work appeared in [20]. The current version extends our conference version with many further technical details, to make the method more self-contained and easily interpretable. In addition, an exhaustively discussed numerical experiment has been added, to quantitatively evaluate the method performance, and the comparison with the results of the method in [19] has been included as well.

The paper is organized as follows. Section 2 is devoted

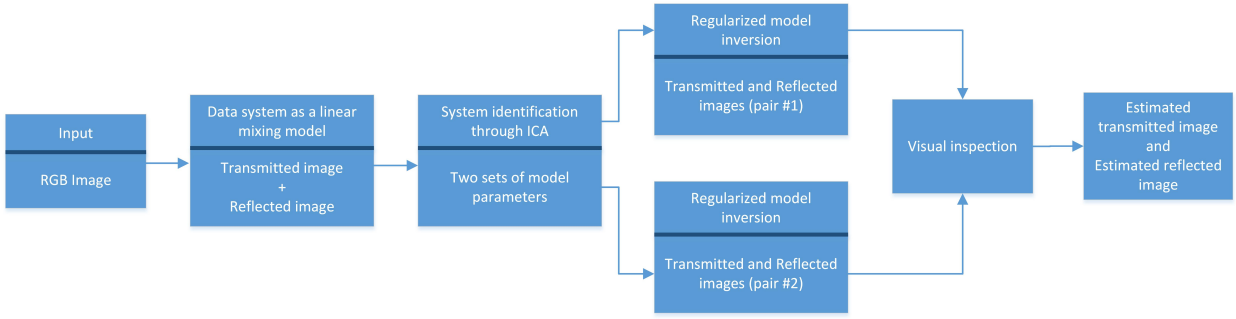


FIG. 2: Flow chart of the overall method.

to the description of the model adopted, the establishment and the justification of the a priori assumptions to reduce the indeterminacy of the problem, and, finally, the definition of our estimation strategy. Section 3 describes the method for estimating the model parameters, whereas Section 4 details the method to separate the reflected image and the transmitted image from the observations. Section 5 is devoted to the experiments, first quantitatively analyzing a simulation example, and then discussing the results obtained in the cases of real photographs of scenes taken through a transparent medium. Finally, Section 6 contains the conclusions and some ideas to improve the method.

II. DATA MODEL AND ESTIMATION STRATEGY

According to [17], we adopt the following data model:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & a_{14} \\ 0 & a_{22} & 0 & a_{24} \\ 0 & 0 & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \\ s_4(t) \end{bmatrix} \quad t = 1, ..., N \quad , \tag{1}$$

where $x_1(t)$, $x_2(t)$ and $x_3(t)$, are the Red, Green and Blue color channels of the observed image at pixel t, being N the total number of pixels in the image. At the same pixel t, $s_1(t)$, $s_2(t)$ and $s_3(t)$ are the ideal colors in the transmitted source image, and $s_4(t)$ represents the achromatic reflected component, respectively. Coefficients a_{ii}, $i = 1, 2, 3$, indicate the transparency of a semi-transparent medium to the different light colors, whereas coefficients a_{i4}, $i = 1, 2, 3$, indicate the percentage of the reflection affecting the measured colors. By calling A the 3×4 matrix above, Eq. (1) can be rewritten in vectorial form as:

$$\mathbf{x} = A\mathbf{s} \quad , \tag{2}$$

where $\mathbf{x}$ is a $3 \times N$ matrix and $\mathbf{s}$ is a $4 \times N$ matrix. For subsequent use, Eq. (2) can be partitioned as:

$$\mathbf{x} = A_c\mathbf{s}_c + A_m\mathbf{s}_m \tag{3}$$

with

$$A_c = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}, \quad A_m = \begin{bmatrix} a_{14} & 0 & 0 \\ 0 & a_{24} & 0 \\ 0 & 0 & a_{34} \end{bmatrix}, \tag{4}$$

$$\mathbf{s}_c = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \mathbf{s}_3 \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{s}_m = \begin{bmatrix} \mathbf{s}_4 & \mathbf{s}_4 & \mathbf{s}_4 \end{bmatrix}^{\mathrm{T}} \tag{5}$$

where T means transposition.

The problem is thus to estimate the ideal transmitted image $\mathbf{s}_c$ from the data $\mathbf{x}$. It is apparent that, having more unknowns than data, this problem is underdetermined. Furthermore, also the imaging system has to be identified, since the coefficients of the mixing matrix A are unknown as well. To reduce this severe ill-posedness and define a unique solution, other works (e.g., [17]) already proposed to exploit the statistical dependence/independence between the intensities of the four component images. Here, we assume instead that the normalized gradients of the three color channels $\mathbf{s}_1$, $\mathbf{s}_2$ and $\mathbf{s}_3$ coincide in the practice, i.e. they are strictly statistically dependent, whereas they are each independent of the normalized gradient of $\mathbf{s}_4$.

Rather than jointly estimating all the unknown sources and parameters, e.g. via Bayesian estimation, we split the problem into two parts: we first estimate the mixing coefficients, and then the four component images. Both steps make use of the above mentioned constraints on the normalized gradients, although within an ICA approach in the first step and a regularization approach in the second step.

Below we will give details on how the original data model of Eq. (1) can be reformulated when the normalized gradients of the sources are used instead of their intensities.

To derive the gradient maps we apply a high-pass filter H to the data, and then normalize them to unit variance by applying the two following diagonal matrices T_c and T_m:

$$T_c = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \\ 0 & 0 & \frac{1}{\sigma_3} \end{bmatrix}, \quad T_m = \begin{bmatrix} \frac{1}{\sigma_4} & 0 & 0 \\ 0 & \frac{1}{\sigma_4} & 0 \\ 0 & 0 & \frac{1}{\sigma_4} \end{bmatrix}, \tag{6}$$

being σ_1, σ_2, σ_3 and σ_4 the unknown standard deviations of the high-pass filtered versions of $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_3$ and $\mathbf{s}_4$, respectively.

Calling $\mathbf{z}_c$ and $\mathbf{z}_m$ the normalized high-pass filtered variables:

$$\mathbf{z}_c = \begin{bmatrix}\mathbf{z}_1 & \mathbf{z}_2 & \mathbf{z}_3\end{bmatrix}^{\mathrm{T}}, \quad \mathbf{z}_m = \begin{bmatrix}\mathbf{z}_4 & \mathbf{z}_4 & \mathbf{z}_4\end{bmatrix}^{\mathrm{T}} \tag{7}$$

their relationships with the original variables are:

$$\mathbf{z}_c = T_c H \mathbf{s}_c \quad , \quad \mathbf{z}_m = T_m H \mathbf{s}_m \quad . \tag{8}$$

Hence, the high-pass filtered data $\mathbf{y}$ is given by:

$$\mathbf{y} = A_c T_c^{-1} \mathbf{z}_c + A_m T_m^{-1} \mathbf{z}_m \quad , \tag{9}$$

where

$$A_c T_c^{-1} = \begin{bmatrix} a_{11}\sigma_1 & 0 & 0 \\ 0 & a_{22}\sigma_2 & 0 \\ 0 & 0 & a_{33}\sigma_2 \end{bmatrix} \tag{10}$$

and

$$A_m T_m^{-1} = \begin{bmatrix} a_{14}\sigma_4 & 0 & 0 \\ 0 & a_{24}\sigma_4 & 0 \\ 0 & 0 & a_{34}\sigma_4 \end{bmatrix} \quad . \tag{11}$$

System of Eq. (9), in expanded form, is given by:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} = \begin{bmatrix} a_{11}\sigma_1 & 0 & 0 & a_{14}\sigma_4 \\ 0 & a_{22}\sigma_2 & 0 & a_{24}\sigma_4 \\ 0 & 0 & a_{33}\sigma_3 & a_{34}\sigma_4 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \\ z_4(t) \end{bmatrix} \quad t = 1, ..., N \quad . \tag{12}$$

III. DATA SYSTEM IDENTIFICATION: ESTIMATION OF THE TRANSPARENCY COEFFICIENTS

Equation (9) (or Eq. (12)) represents a system of 3 equations and 4 unknown, whose matrix has 6 unknown coefficients. Indicating with E the expectation, the covariance matrix of the filtered data, $E\left[\mathbf{y}\mathbf{y}^{\mathrm{T}}\right]$, from the assumption of independence of $\mathbf{z}_c$ and $\mathbf{z}_m$, that is $E\left[\mathbf{z}_c\mathbf{z}_m^{\mathrm{T}}\right] = 0$, it is:

$$\begin{aligned} &E\left[\mathbf{y}\mathbf{y}^{\mathrm{T}}\right] = \\ &E\left[\left(A_c T_c^{-1}\mathbf{z}_c + A_m T_m^{-1}\mathbf{z}_m\right)\left(A_c T_c^{-1}\mathbf{z}_c + A_m T_m^{-1}\mathbf{z}_m\right)^{\mathrm{T}}\right] \\ &= A_c T_c^{-1} E\left[\mathbf{z}_c\mathbf{z}_c^{T}\right] A_c T_c^{-1} + A_m T_m^{-1} E\left[\mathbf{z}_m\mathbf{z}_m^{\mathrm{T}}\right] A_m T_m^{-1} \end{aligned} \tag{13}$$

with

$$E\left[\mathbf{z}_c\mathbf{z}_c^{\mathrm{T}}\right] = E\left[\mathbf{z}_m\mathbf{z}_m^{\mathrm{T}}\right] = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad . \tag{14}$$

In explicit form, Eq. (13), in view of eqs. (14), can be rewritten as:

$$\begin{aligned} &E\left[\mathbf{y}\mathbf{y}^{\mathrm{T}}\right] = \\ &\begin{bmatrix} a_{11}\sigma_1 & 0 & 0 \\ 0 & a_{22}\sigma_2 & 0 \\ 0 & 0 & a_{33}\sigma_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_{11}\sigma_1 & 0 & 0 \\ 0 & a_{22}\sigma_2 & 0 \\ 0 & 0 & a_{33}\sigma_3 \end{bmatrix} + \\ &\begin{bmatrix} a_{14}\sigma_4 & 0 & 0 \\ 0 & a_{24}\sigma_4 & 0 \\ 0 & 0 & a_{34}\sigma_4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a_{14}\sigma_4 & 0 & 0 \\ 0 & a_{24}\sigma_4 & 0 \\ 0 & 0 & a_{34}\sigma_4 \end{bmatrix} \quad . \end{aligned} \tag{15}$$

The left hand side of Eq. (15), i.e. the symmetric covariance matrix, has 6 coefficients that can be estimated from the data, while the right hand side contains the 6 unknown coefficients of matrices $A_c T_c^{-1}$ and $A_m T_m^{-1}$, that is $a_{11}\sigma_1$, $a_{22}\sigma_2$, $a_{33}\sigma_3$, $a_{14}\sigma_4$, $a_{24}\sigma_4$, $a_{34}\sigma_4$, so that, from the above equation, these coefficients can be estimated, for example by least mean squares and gradient descent algorithms. We initially attempted this way, but observed a slow convergence and not much precise estimates.

We then observed that, still exploiting the substantial coincidence of the three gradients of the ideal transmitted image, that is $\mathbf{z}_1 = \mathbf{z}_2 = \mathbf{z}_3$, and the independence of $\mathbf{z}_4$ from $\mathbf{z}_1$ (as well as from $\mathbf{z}_2$ and $\mathbf{z}_3$), the 3×4 system of Eq. (12) can be partitioned into the following 3 independent 2×2 systems:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} a_{11}\sigma_1 & a_{14}\sigma_4 \\ a_{22}\sigma_2 & a_{24}\sigma_4 \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_4 \end{bmatrix} \quad , \tag{16}$$

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_3 \end{bmatrix} = \begin{bmatrix} a_{11}\sigma_1 & a_{14}\sigma_4 \\ a_{33}\sigma_3 & a_{34}\sigma_4 \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_4 \end{bmatrix} \quad , \tag{17}$$

$$\begin{bmatrix} \mathbf{y}_2 \\ \mathbf{y}_3 \end{bmatrix} = \begin{bmatrix} a_{22}\sigma_2 & a_{24}\sigma_4 \\ a_{33}\sigma_3 & a_{34}\sigma_4 \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_4 \end{bmatrix} \quad . \tag{18}$$

The above systems can be solved through ICA, both for the gradient sources and the mixing coefficients. Specifically, we employed the FastICA algorithm, first proposed in [21]. In fact, to estimate all the 6 coefficients, only two of the above equations are needed, for example the first one, Eq.(16), and the second one, Eq.(17). The inconvenience that, in this way, coefficients $a_{11}\sigma_1$ and $a_{14}\sigma_4$ are estimated twice, likely obtaining slightly different values, can be trivially overcome by retaining their average value as the unique estimate. As well known, however, any ICA algorithm provides solutions that are affected by

more important, intrinsic indeterminacies, the scale and permutation indeterminacies. With respect to the scale indeterminacy, possible negative values and multiplicative factors in the obtained coefficients can be removed knowing that the coefficients must be positive and the estimated sources must be of unit variance by definition.

Due to the permutation indeterminacy, the estimated columns of the two mixing matrices of Eq.(16) and Eq.(17)can be occasionally and unpredictably inverted, either within only one matrix or in both. While the correct order of the columns cannot by no means established without additional information, we can at least assign the same order to the columns of the two matrices, the right one in both or the inverted one in both. This can be done arbitrarily assuming the solutions of the first system as our reference, and performing a similarity check between the first row of the two estimated mixing matrices or between the corresponding estimated source gradients.

Although the above mentioned strategies permit to partially overcome the indeterminacies of ICA, two extra indeterminacies are left. First of all, the 6 coefficients are estimated up to the unknown standard deviations; second, since the columns of the matrices in Eq.(16) and Eq.(17), could have been estimated in an inverted (although congruent to each other) order, we cannot discriminate which source corresponds to the red gradient $\mathbf{z}_1$ of the transmitted image and which one to the reflection gradient $\mathbf{z}_4$. The first indeterminacy is unavoidable, at least so far. As per the second indeterminacy, using the estimated coefficients, and how described in the following section, we reconstruct two different pairs of solution images assuming, the first time, the columns as they are, and, the second time, by inverting them. The two pairs of solutions are each constituted of the supposed transmitted image ($\mathbf{s}_1$,$\mathbf{s}_2$,$\mathbf{s}_3$) and the supposed reflection $\mathbf{s}_4$. We can then choose the correct pair by visual inspection.

IV. REMOVAL OF REFLECTION: ESTIMATION OF THE TRANSMITTED AND REFLECTED IMAGE

Once the transparency coefficients has been estimated up to the standard deviations with ICA, the data system of Eq. (1) can be reformulated as:

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} a_{11}\sigma_1 & 0 & 0 & a_{14}\sigma_4 \\ 0 & a_{22}\sigma_2 & 0 & a_{24}\sigma_4 \\ 0 & 0 & a_{33}\sigma_3 & a_{34}\sigma_4 \end{bmatrix} \begin{bmatrix} \mathbf{s}_1/\sigma_1 \\ \mathbf{s}_2/\sigma_2 \\ \mathbf{s}_3/\sigma_3 \\ \mathbf{s}_4/\sigma_4 \end{bmatrix} \quad , \tag{19}$$

where we assume the order of the coefficients to be correct. The linear system above has 3 equations and 4 unknowns, so that it represents an ill-posed problem with an infinity of solutions $\mathbf{s}_1/\sigma_1$, $\mathbf{s}_2/\sigma_2$, $\mathbf{s}_3/\sigma_3$, and $\mathbf{s}_4/\sigma_4$. To find a unique solution, extra information must be included into the problem.

To this end, we adopt constrained least squares, where the constraints enforce the minimum distance between the three pairs $(\mathbf{s}_1/\sigma_1, \mathbf{s}_2/\sigma_2)$, $(\mathbf{s}_1/\sigma_1, \mathbf{s}_3/\sigma_3)$ and $(\mathbf{s}_2/\sigma_2, \mathbf{s}_3/\sigma_3)$ of the color channels of the transmitted image, and the maximum orthogonality between each of such channels and the achromatic reflected image $\mathbf{s}_4/\sigma_4$. We define an energy function constituted of a data term, in the form of the squared Euclidean norm of the difference between the left hand side and the right hand side of Eq. (19), plus two quadratic stabilizer terms expressing the two constraints above, weighted by two positive regularization parameters λ_1 and λ_2, respectively.

Calling A_σ the mixing matrix in Eq. (19) and $\mathbf{s}_\sigma$ the vector of the four unknown component images, the energy function we adopt is:

$$\phi(\mathbf{s}_\sigma) = ||\mathbf{x} - A_\sigma \mathbf{s}_\sigma||^2 + \lambda_1 \mathbf{s}_\sigma^{\mathrm{T}} S_1 \mathbf{s}_\sigma + \lambda_2 \mathbf{s}_\sigma^{\mathrm{T}} S_2 \mathbf{s}_\sigma \quad , \tag{20}$$

where

$$S_1 = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix} \tag{21}$$

express the constraints defined above. Since the energy function is quadratic, its minimum can be obtained in close form. This solution obviously depends on λ_1 and λ_2, and, in vector form, can be expressed as:

$$\mathbf{s}_\sigma(\lambda_1, \lambda_2) = \left[A_\sigma^{\mathrm{T}} A_\sigma + \lambda_1 S_1 + \lambda_2 S_2 + \epsilon I_{4\times 4}\right]^{-1} A_\sigma^{\mathrm{T}} \mathbf{x} \quad , \tag{22}$$

where the term $\epsilon I_{4\times 4}$, with ϵ a small positive parameter and $I_{4\times 4}$ the 4×4 identity matrix, has been introduced to improve the problem stability. The optimal values $\hat{\lambda_1}$ and $\hat{\lambda_2}$ of λ_1 and λ_2 are assumed to be those for which the gradients of the color channels of the transmitted component of $\mathbf{s}_\sigma(\lambda_1, \lambda_2)$ essentially coincide, and are independent of the gradient of the related reflected component. This results in the following criterion:

$$(\hat{\lambda_1}, \hat{\lambda_2}) = \arg\min_{\lambda_1, \lambda_2} \left\| E\left[\mathbf{z}(\lambda_1, \lambda_2)\mathbf{z}(\lambda_1, \lambda_2)^{\mathrm{T}}\right] - B \right\|^2 \quad , \tag{23}$$

where

$$\mathbf{z}(\lambda_1, \lambda_2) = \left[\mathbf{z}_1(\lambda_1, \lambda_2) \;\; \mathbf{z}_2(\lambda_1, \lambda_2) \;\; \mathbf{z}_3(\lambda_1, \lambda_2) \;\; \mathbf{z}_4(\lambda_1, \lambda_2)\right]^{\mathrm{T}} \quad . \tag{24}$$

and

$$B = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{25}$$

is the theoretical covariance of 4 normalized gradients exactly satisfying the constraints above.

In the practice, in order to choose the optimal regularization parameters, a suitable range for λ_1 and λ_2 is empirically chosen, the corresponding $\mathbf{s}_\sigma(\lambda_1, \lambda_2)$ is computed for each point of a grid in the range, according to

(a) (b) (c)

FIG. 3: Synthetic image with a reflection: (a) original transmitted image; (b) original reflection image; (c) linear mixing of the images in (a) and (b) through the matrix of Eq. (28).

Eq. (22), and the values that minimize the cost function in Eq. (23) are selected as the optimal $\hat{\lambda_1}$ and $\hat{\lambda_2}$.

It is worth noting that the search for the optimal regularization parameters is the most expensive part of the entire algorithm, which is in any case very fast.

To obtain the final restored image we observe that, equivalently, system of Eq. (19) can be partitioned as:

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} a_{11}\sigma_1 & 0 & 0 \\ 0 & a_{22}\sigma_2 & 0 \\ 0 & 0 & a_{33}\sigma_3 \end{bmatrix} \begin{bmatrix} \mathbf{s}_1/\sigma_1 \\ \mathbf{s}_2/\sigma_2 \\ \mathbf{s}_3/\sigma_3 \end{bmatrix} + + \begin{bmatrix} a_{14}\sigma_4 \\ a_{24}\sigma_4 \\ a_{34}\sigma_4 \end{bmatrix} \mathbf{s}_4/\sigma_4 \tag{26}$$

and then it is:

$$A_c\mathbf{s}_c = \begin{bmatrix} a_{11}\sigma_1 & 0 & 0 \\ 0 & a_{22}\sigma_2 & 0 \\ 0 & 0 & a_{33}\sigma_3 \end{bmatrix} \begin{bmatrix} \mathbf{s}_1/\sigma_1 \\ \mathbf{s}_2/\sigma_2 \\ \mathbf{s}_3/\sigma_3 \end{bmatrix} = = \mathbf{x} - \begin{bmatrix} a_{14}\sigma_4 \\ a_{24}\sigma_4 \\ a_{34}\sigma_4 \end{bmatrix} \mathbf{s}_4/\sigma_4 \quad . \tag{27}$$

Note that it is impossible to separate A_c and $\mathbf{s}_c$, that is to recover the ideal transmitted image $\mathbf{s}_c$, unless extra information on the scale of $\mathbf{s}_c$ is available. Thus, we can only estimate the image free of reflection, $A_c\mathbf{s}_c$. This results in the fact that we cannot expect a full fidelity of the colors of the reconstructed image.

V. DISCUSSION OF THE EXPERIMENTAL RESULTS

Our algorithm has been designed for the general problem of removing reflection from images of any type. To quantitatively evaluate its performance, we first analyze the results for a synthetically generated image. Figure 3(a) shows the ideal transmitted image, Fig. 3(b) shows the ideal reflection, whereas Fig. 3(c) shows their linear mixing according to the model in Eq. (1), by using the following mixing matrix:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0.4 \\ 0 & 0.7 & 0 & 0.6 \\ 0 & 0 & 0.5 & 0.8 \end{bmatrix} \quad , \tag{28}$$

In this case, the normalized gradients $\mathbf{z}_1$, $\mathbf{z}_2$, $\mathbf{z}_3$, and $\mathbf{z}_4$ of the ideal sources can be computed exactly from the known sources, in such a way to verify that the constraints we assume for our problem are feasible. Indeed, Fig. 5 clearly shows that the behaviours of the three gradients of the ideal transmitted image almost coincide (top panel), whereas the behaviour of the gradient of the reflection (bottom panel) is rather different.

Furthermore, the feasibility of the constraints employed to estimate the optimal λ_1 and λ_2 (see Eq. (23)) is proven by the covariance matrix of the gradients $\mathbf{z}_1$, $\mathbf{z}_2$, $\mathbf{z}_3$, and $\mathbf{z}_4$, still computed from the known sources as:

$$E\left[\mathbf{z}\mathbf{z}^{\mathrm{T}}\right] = \begin{bmatrix} 1 & 0.9850 & 0.9845 & 0.0051 \\ 0.9850 & 1 & 0.9901 & 0.0059 \\ 0.9845 & 0.9901 & 1 & 0.0056 \\ 0.0051 & 0.0059 & 0.0056 & 1 \end{bmatrix} \quad . \tag{29}$$

As explained in Section 3, with our algorithm the indeterminacy on the correct order of transparency coefficients estimated via ICA entails the reconstruction of two different pairs of solutions ($\mathbf{s}_1$,$\mathbf{s}_2$,$\mathbf{s}_3$,$\mathbf{s}_4$), with each pair constituted of the three color channels of the image assumed as the transmitted image plus the grayscale removed reflection. The correct pair is then chosen by visual inspection. Figure 4 shows the two pairs, the inverted one (top panels) and the correct one (bottom panel). These solutions have been obtained by using $\lambda_1 = 7.84$, $\lambda_2 = 0.29$, and $\epsilon = 10^{-6}$.

For the correct pair of solutions we computed the Root Mean Square Error (RMSE) and the percentage error (i.e. the ratio between the RMSE and the maximum value in the image) of the estimated images with respect

(a) (b) (c) (d)

FIG. 4: The two pairs of solutions obtained from a synthetic image with a reflection. (a) and (b) the inverted one; (c) and (d) the correct one.

to the ideal ones. For the estimated transmitted image we obtained an RMSE of 15.32, and a percentage error of 10.2%. The relatively high RMSE error is mainly concentrated in the Blue channel, and this is reflected in the more "reddish" appearance of the estimated transmitted image with respect to the ideal one, especially in correspondence of the plumage of the hat. In other words, the Blue component of the estimated transmitted image is too low. Presumably, part of this Blue component has been retained by the estimated reflection, which, indeed, appears lighter than the ideal one, with an RMSE of 27.55 and a percentage error of 11.4%.

The estimated mixing matrix A_σ (up to the gradient standard deviations, and whose columns have been arranged according to the identified correct pair of solutions) is the following:

$$A_\sigma = \begin{bmatrix} 35.4940 & 0 & 0 & 17.8170 \\ 0 & 24.3005 & 0 & 26.9555 \\ 0 & 0 & 18.0562 & 34.6108 \end{bmatrix} . \quad (30)$$

Since this time we known by construction the standard deviations, which amount to $\sigma_1 = 35.6870$, $\sigma_2 = 36.1433$, $\sigma_3 = 36.1674$, and $\sigma_4 = 43.1189$, respectively, we can derive the coefficients of the related A matrix, which, as it can be easily verified, almost exactly coincide with the ideal mixing matrix in Eq. (28).

As per real images, we first compare the performance of our algorithm with that of the algorithm proposed in [19], by applying it to one of the images presented in that paper. This image represents a toy standing behind a transparent CD box with achromatic reflections occurring on the box surface. Figure 6 (a) shows the input image. The results of our algorithm are shown in Fig. 6 (b) and Fig. 6 (c), whereas Fig. 6 (d) and Fig. 6 (e) show the results of the algorithm in [19]. Although the ideal image of the toy free of reflection is not available, it is apparent that our algorithm is able to remove the reflection much better. With respect to color fidelity, by comparing the areas free of reflection of the input image with the same areas in our reconstruction, we might say that the color has been preserved as well.

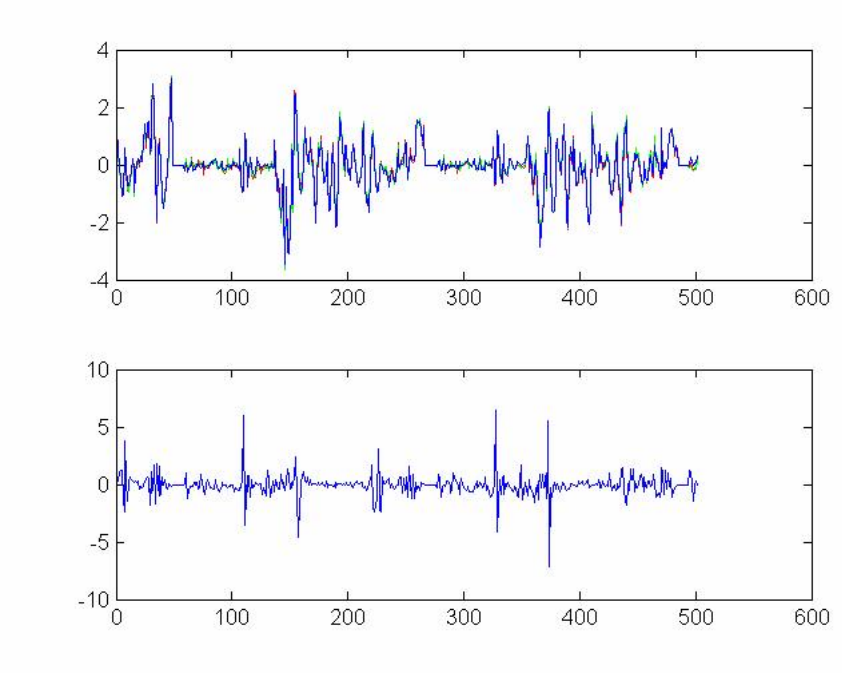


FIG. 5: Behaviours of the normalized gradients (value versus the pixel number): top panel - the three gradients of the ideal transmitted image; bottom panel - the gradient of the reflection image.

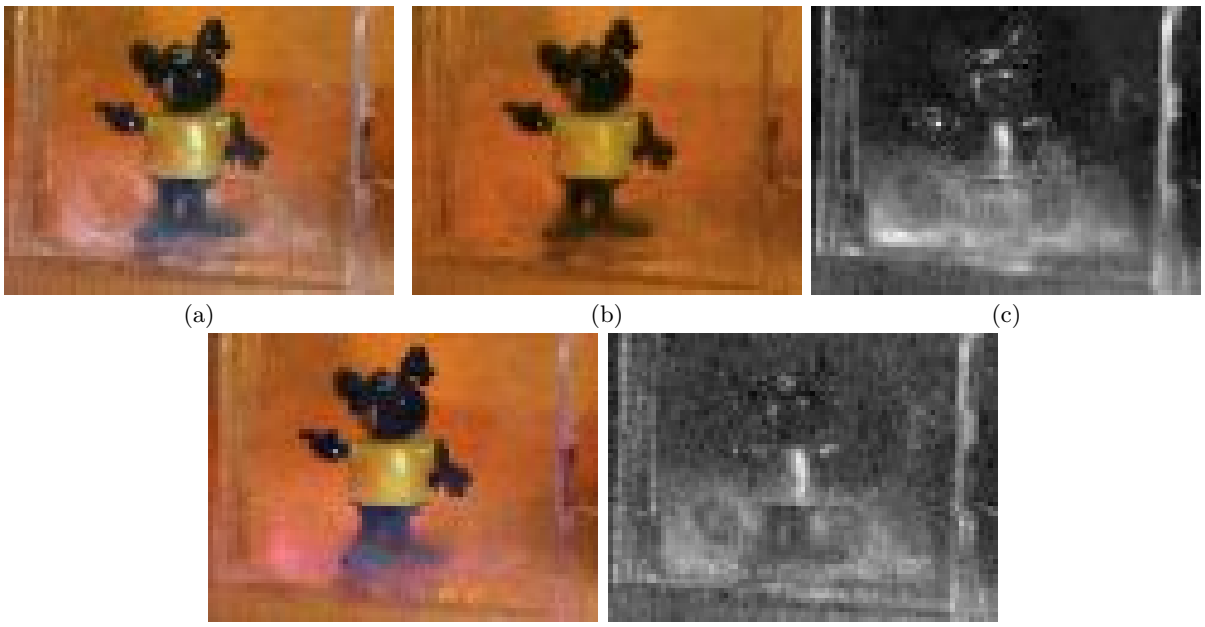


FIG. 6: Removal of reflection from a real image acquired through a transparent plastic sheet: (a) original acquired image; (b) image restored with our algorithm; (c) reflection separated with our algorithm; (d) image restored with the algorithm in [19]; (e) reflection separated with the algorithm in [19].

It is worth noting again that our algorithm is very fast, and in particular presumably much faster than the algorithm in [19], which employs usually expensive Monte Carlo methods.

Successively, we considered the application of the algorithm to real images of artworks. We start with a picture of a painting framed by a glass. Figure 7(a) shows the acquired image where the reflection of a lamp is well visible. Although separation of the two images has not been completely achieved, as it can be noticed from some residual

(a) (b)

(c)

FIG. 7: Removal of reflection from the picture of a painting: (a) original acquired image; (b) restored image; (c) separated reflection.

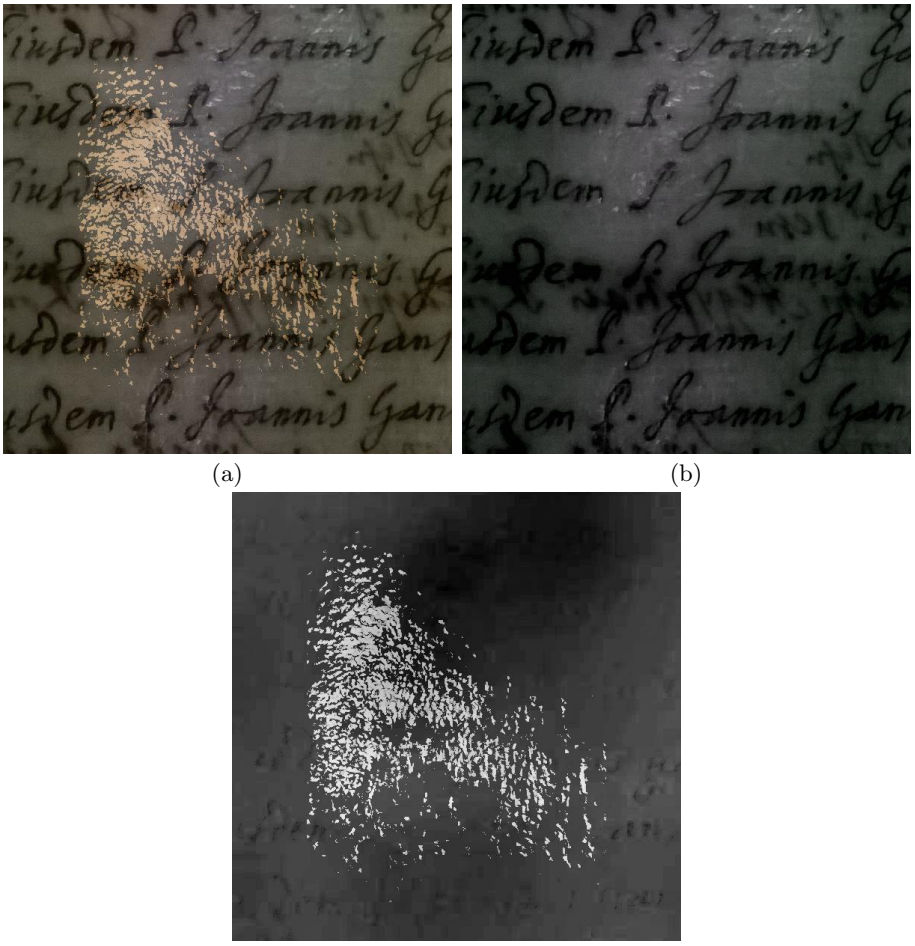
patterns of the transmitted image left in the reflected image of Fig. 7(c), Fig. 7(b) shows the transmitted image perfectly freed from reflection. However, due to the imperfect separation, along with the above mentioned scale ambiguity of the solution, the colors are not perfectly reproduced.

In a subsequent experiment, we show the results obtained when the reflection affects a document image. Figure 8(a) shows the image of a plastic coated ancient manuscript. It is apparent that the light reflection prevents reading part of the text. With our algorithm we have been able to separate the reflected image in Fig. 8(c), and recover the reflection-free image of Fig. 8(b). In this application, the loss of the true colors of the transmitted image is a minor problem, since the essential aspect is the possibility to recover the legibility of the text. Note indeed that the text under the reflection is now well visible.

VI. CONCLUSIONS

We proposed a method to remove a reflection from a single color image acquired through a semi-transparent medium. The unwanted reflection, often caused by a light source, is considered to be an achromatic image that combines additively with the real transmitted image of the object of interest. Since the mixing coefficients are unknown, we adopt a blind source separation technique exploiting the substantial coincidence of the gradients of the three color channels of the ideal image, and their independence from the gradient of the reflected image. The algorithm acts in two steps. In the first step, the model parameters are estimated through ICA, whereas, in the second step, the four component images are estimated via regularization.

The efficiency of the algorithm is quantitatively proven on a numerically generated image. As per the performance of the method on real images, we compared our results with those of a recently proposed algorithm designed for a single input image, and presented the promis-



(a) (b)

(c)

FIG. 8: Removal of reflection from the image of an ancient manuscript: (a) original acquired image; (b) restored image; (c) separated reflection.

ing results of the application of the method to the restoration of a real painting framed by a glass and a manuscript laminated for conservation purposes.

Nevertheless, some aspects of the methods still need to be investigated and improved. These mainly regard the search for an efficient way to overcome the permutation indeterminacy of ICA, and thus determine the correct order of the transparency coefficients, and the search for more effective constraints about the component images, in order to be able to estimate their scale and then, by discriminating A_c from $\mathbf{s}_c$, to obtain a more faithful color of the reconstructed images. Our future research plan regards also the study of alternative data models that better adhere to the physics of the phenomenon.

VII. ACKNOWLEDGMENT

This work has been supported by European funds, through the program POR Calabria FESR 2007-2013 - PIA Regione Calabria Pacchetti Integrati di Agevolazione Industria Artigianato Servizi, project ITACA (Innovative Tools for cultural heritage ArChiving and restorAtion).

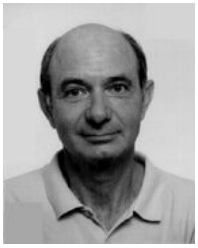
[1] T. Cronin, N. Shashar, and L. Wolff, "Portable imaging polarimeters," in *Proc. ICPR 1994*, vol. A, 1994, pp. 606–609.

[2] K. Nayar, X. Fang, and T. Boult, "Separation of reflection components using color and polarization," *Int J Comput Vis*, vol. 21, pp. 163–186, 1997.

[3] H. Fujikake, K. Takizawa, T. Aida, H. Kikuchi, T. Fujii, and M. Kawakita, "Electrically-controllable liquid crystal polarizing filter for eliminating reflected light," *Opt Rev*, vol. 5, pp. 93–98, 1998.

[4] Y. Schechner, J. Shamir, and N. Kiryati, "Polarization-based decorrelation of transparent layers: the inclination

angle of an invisible surface," in *Proc. Int. Conf. on Computer Vision 1999*, 1999, pp. 814–819.
[5] ——, "Polarization and statistical analysis of scenes containing a semireflector," *Journal of the Optical Society of America A*, vol. 17, pp. 276–284, 2000.
[6] M. Born and E. Wolf, *Principles of Optics*. London: Pergamon, 1965.
[7] H. Farid and E. Adelson, "Separating reflections and lighting using independent components analysis," in *Proc. CVPR 1999*, vol. 1, 1999, pp. 262–267.
[8] A. Bronstein, M. Bronstein, M. Zibulevsky, and Y. Zeevi, "Sparse ICA for blind separation of transmitted and reflected images," *Int. Journal of Imaging System and Technology*, vol. 15, pp. 84–91, 2005.
[9] E. Beery and A. Yeredor, "Blind separation of reflections with relative spatial shifts," in *Proc. ICASSP 2006*, vol. 5, 2006, pp. 625–628.
[10] K. Gai, Z. Shi, and C. Zhang, "Blindly separating mixtures of multiple layers with spatial shifts," in *Proc. CVPR 2008*, 2008, pp. 1–8.
[11] ——, "Blind separation of superimposed images with unknown motions," in *Proc. CVPR 2009*, 2009, pp. 1881–1888.
[12] X. Guo, X. Cao, and Y. Ma, "Robust separation of reflection from multiple images," in *Proc. Int. Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2014.
[13] N. Kong, Y.-W. Tai, and J. Shin, "A physically-based approach to reflection-separation: from physical modeling to constrained optimization," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 36, pp. 209–221, February 2014.
[14] Y. Shih, D. Krishnan, F. Durand, and W. Freeman, "Reflection removal using ghosting cues," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
[15] A. Levin, A. Zomet, and Y. Weiss, "Separating reflections from a single image using local features," in *Proc. ECCV 2004*, 2004, pp. 306–313.
[16] A. Levin and Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1647–1655, 2007.
[17] K. Kayabol, E. Kuruoglu, and B. Sankur, "Image source separation using color channel dependencies," in *Proc. 8th Int. Conf. on Independent Component Analysis and Signal Separation*, 2009, pp. 499–506.
[18] B. Sarel and M. Irani, "Separating transparent layers through layer information exchange," in *Proc. ECCV 2004*, ser. LNCS, T. Pajdla and J(G.)Matas, Eds., vol. LNCS 3024. Springer, Heidelberg, 2004, pp. 328–341.
[19] Q. Yan, E. E. Kuruoglu, X. Yang, Y. Xu, and K. Kayabol, "Separating reflections from a single image using spatial smoothness and structure information," in *Proc. LVA/ICA 2010*, ser. Lecture Notes in Computer Science, V. V. et al., Ed., vol. LNCS 6365. Springer, 2010, pp. 637–644.
[20] L.Bedini, P. Savino, and A. Tonazzini, "Removing achromatic reflections from color images with application to artwork imaging," in *Proc. 9th International Symposium on Image and Signal Processing and Analysis (ISPA 2015)*, ser. IEEE Xplore. IEEE, 2015, pp. 126–130.
[21] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.



Luigi Bedini graduated cum laude in Electronic Engineering from the University of Pisa, Italy, in 1968. Since 1970 he has been a Researcher of the Italian National Research Council, Istituto di Scienza e Tecnologie dell'Informazione, Pisa, Italy. His interests have been in modelling, identification, and parameter estimation of biological systems applied to non-invasive diagnostic techniques. At present, his research interest is in the field of digital signal processing, image reconstruction and neural networks applied to image processing. He is co-author of more than 80 scientific papers. From 1971 to 1989, he was Associate Professor of System Theory at the Computer Science Department, University of Pisa, Italy.

Pasquale Savino is senior researcher at Italian National Research Council, Institute for Information Science and Technology. He graduated in Physics at the University of Pisa, Italy, in 1979. His current research interests are multimedia information indexing and retrieval, Multimedia Content Management, Ancient document digital analysis and restoration, and Digital Libraries. He was the principal investigator in many European Research Projects in the field of multimedia content management and retrieval and in digital libraries..

Anna Tonazzini is a senior researcher at the Signals and Images Laboratory of the Institute of Information Science and Technologies, Italian National Research Council (CNR) in Pisa. She coordinated several CNR and Regional Projects for basic and applied research in Image Processing and Analysis, Neural Networks and Learning, Computational Biology and Document Analysis, and is co-author of over 100 published papers. She was also the ISTI responsible for the UE Project ISYREADET, and participated in the MUSCLE European NOE. She was Chair of the EUSIPCO2008 Special Session on Restoration of degraded document images, and Guest Editor of the Special Issue on Image and Video Processing for Cultural Heritage of the Eurasip Journal on Image and Video Processing.