In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1:262-267, 1999, Fort Collins CO.

Separating Reflections and Lighting Using Independent Components Analysis

Hany Farid and Edward H. Adelson
Perceptual Science Group
Massachusetts Institute of Technology
Cambridge, MA 02139
{farid,adelson}@persci.mit.edu

Abstract

The image of an object can vary dramatically depending on lighting, specularities/reflections and shadows. It is often advantageous to separate these incidental variations from the intrinsic aspects of an image. This paper describes how the statistical tool of independent components analysis can be used to separate some of these incidental components. We describe the details of this method and show its efficacy with examples of separating reflections off glass, and separating the relative contributions of individual light sources.

Figure 1: Renoir's *On the Terrace*, Sheila and Sheila's reflection.

1. Introduction

The image of an object can vary dramatically depending on lighting, specularities/reflections and shadows, and yet our recognition of objects is amazingly robust despite these incidental variations. Our visual system seems to separate the various components that contribute to the formation of an image, yielding stable and reliable percepts. To facilitate such tasks as object recognition, visual-based navigation, and scene segmentation we would like to design computer systems that have a similar ability to separate the incidental from the intrinsic aspects of an image.

The combination of incidental and intrinsic components can often be approximated as a linear mixing process. Under these conditions we show how the statistical tool of independent components analysis (ICA) can be used to separate these components. One commonly occurring example is the presence of reflections on glass, for example when viewing a painting framed behind glass (Figure 1). Because it provides a simple illustration we will focus on the problem of separating reflections from glass while leaving intact the image of objects behind the glass (see [10, 4, 11, 9] for different approaches to removing specular reflections). We will show later how the same general techniques can be used to separate the relative contributions of individual light sources.

Let's first look more closely at the simple physics of reflections from planar surfaces. Shown in Figure 2 is an idealized example of a photograph of a painting behind glass. The final image is a linear combination of the light reflected by the painting and the light directly reflected by the glass. The amount of light at a single point in the image can be expressed as:

$$y_1 \quad = \quad aP \; + \; bR, \tag{1}$$

where P and R are the amount of light contributed by the painting and reflection, and a and b are multiplicative constants. We would like to remove the contribution of the reflection R from the image y_1, but the above equation provides only a single constraint in four unknowns. Additional constraints may be added by exploiting the fact that reflections are partially polarized and that by photographing through a linear polarizer the relative strength of the reflections can be adjusted. [1] With respect to Equation (1), the

The authors gratefully acknowledge support from NIH Grant EY11005-04 and MURI Grant N00014-95-1-0699.

[1] The polarizer can only completely remove the reflection when the viewing angle to the glass is at Brewster's angle, typically a severe angle resulting in significant geometric distortions [3].

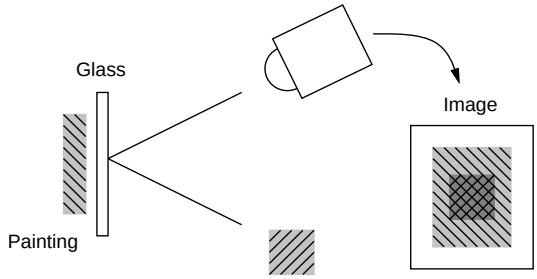

Figure 2: A photograph of a painting behind glass contains a superposition of the light reflected by the painting and the light reflected directly off the glass.

linear polarizer has the effect of manipulating the relative contributions of the painting and reflection. The second image takes the form:

$$y_2 \quad = \quad cP \; + \; dR. \tag{2}$$

The above equation provides another constraint, but two new unknowns have also been introduced, leaving us with a total of two constraints in six unknowns, and little hope of a solution without making further assumptions.

To solve the underconstrained set of equations we make the modest assumption that the image of the painting and reflection are independent. Intuitively this means that for each spatial position the pixel intensity in one image provides no predictive information about the pixel intensity in the second image. This is a perfectly reasonable assumption since there is no reason to expect a correlation between the image of objects behind the glass and the image of objects reflected by the glass. Given the linear model of image formation and this assumption of independence we perform independent components analysis (ICA) to separate the reflection from the desired image. In the next section the details of this statistical technique are outlined, and in the following section several examples of its efficacy are given.

2 Separating Images

The general problem of image separation can be stated as follows: given N distinct linear combinations of N images determine the original N images. For our application we can restrict ourselves to the case of just two images. Denoting these images in row vector form as x_1 and x_2, the linear mixing can be expressed in matrix form as follows:

$$\begin{pmatrix} \quad y_1 \quad \\ \quad y_2 \quad \end{pmatrix} \quad = \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \quad x_1 \quad \\ \quad x_2 \quad \end{pmatrix}$$

$$Y \quad = \quad MX, \tag{3}$$

where the matrix M embodies the linear mixing. Note that this form is the same as the model of reflections given in Equations (1) and (2), and that with this model it is assumed that the linear mixing is uniform over the entire image. The mixed images in Y each contain a linear combination of the source images in X. Our job is to recover the sources images from the mixed images. Of course given the full rank (i.e., invertible) matrix M it would be trivial to estimate the source images by left multiplying the mixed images with the inverse mixing matrix, $\tilde{X} = M^{-1}Y$. But we don't typically know the mixing matrix, so our job will be to estimate it from the mixed images alone.

Equation (3) provides two constraints in six unknowns and so cannot be solved without further assumptions. The first assumption we make is that the pair of source images are independent. Denoting X_1 and X_2 as the random variables from which the pixel intensities of source images x_1 and x_2 are drawn, this assumption can be expressed as $P(X_1, X_2) = P(X_1) \cdot P(X_2)$ (i.e., the joint probability distribution is separable). Although the constraint is expressed in terms of these continuous random variables we will typically work with the histograms of sampled images that, because of the dense sampling, are good approximations to the continuous probability distributions. The second modest assumption is that the mixing matrix M is full rank. With these two assumptions the general estimation problem is known as independent components analysis (ICA) of which there is a large and varied literature. Early contributors include [1, 5, 6, 8], and more recently there has been a renewed interest in the learning community (e.g., [2]). We present an analytic version of ICA based on higher-order statistical moments most similar to that of [5]. We provide a different formulation based on maximizing a series of error functions and give a simple and intuitive geometric interpretation of these steps.

We begin by attempting to gain some insight into the structure of the mixing matrix by decomposing it according to the singular value decomposition (SVD):

$$M \quad = \quad R_1 S R_2, \tag{4}$$

where R_1 and R_2 are orthonormal (rotation) matrices and S is a diagonal (scaling) matrix. Shown in Figure 3 is a geometric interpretation of the effects of each of these matrices on the joint probability distribution of the source images X. According to our assumption of independence this distribution should be separable; for illustration purposes consider the case when the marginal distributions are uniform, then the joint distribution is a square (in practice no such constraint is imposed on the marginal distributions). The mixing matrix rotates, scales, and again rotates this distribution transforming the square into a parallelogram. The estimation of the independent source images reduces to determining how to transform the parallelogram back into a square. Or more generally, transforming the two-dimensional joint distribution into a separable product of one-dimensional

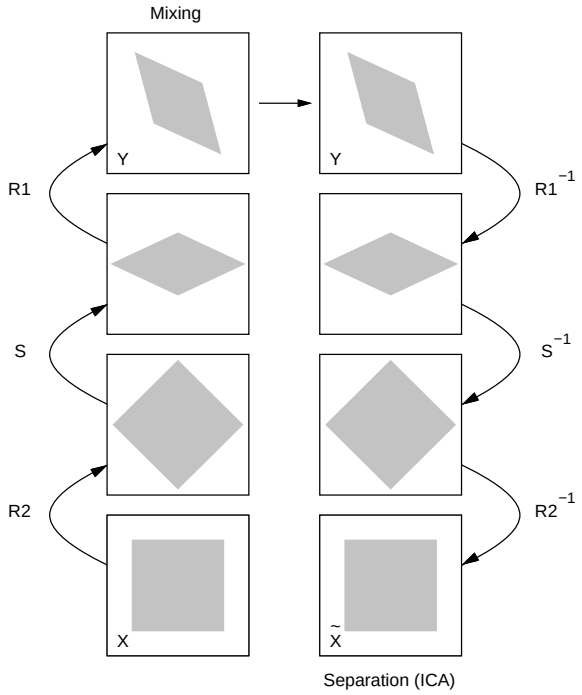

Figure 3: Shown at the bottom left is an idealized joint probability distribution for a pair of independent images, where for illustration purposes only the marginal distributions are assumed to be uniform. The linear mixing of these images, Equation (3), transforms this distribution, via a rotation, scaling, and rotation, from a square into a parallelogram (left). The goal of ICA is to transform this parallelogram back into a square, yielding the original independent images (right).

distributions. As Figure 3 suggests this may be accomplished by applying the opposite rotations and scalings in reverse order.

The first step in separating the images is to apply a rotation that aligns the long and short axis of the parallelogram with the primary axis. These are easily determined since they are the axes with the maximal/minimal variance. Assuming zero mean measurements, the variance at an arbitrary orientation is:

$$E(\theta_1) = \sum_{i=1}^{N} \left[\begin{pmatrix} y_1(i) & y_2(i) \end{pmatrix} \begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix} \right]^2. \quad (5)$$

The axis of maximal variance is determined by finding the angle θ_1 that maximizes this error function (Figure 4). The axis of minimal variance is orthogonal to this axis oriented at $\theta_1 - \pi/2$. These axes correspond to the principle axes as determined by principle components analysis (PCA). This error function can be maximized analytically by differentiating with respect to θ_1, setting equal to zero and solving, to yield:

$$\theta_1 = \frac{1}{2} \tan^{-1} \left[\frac{\sum_{i=1}^{N} r^2(i) \sin(2\phi(i))}{\sum_{i=1}^{N} r^2(i) \cos(2\phi(i))} \right], \quad (6)$$

where, because it takes on a particularly simple form, the solution is given in polar coordinates $r(i) = y_1(i)^2 + y_2(i)^2$ and $\phi(i) = \tan^{-1}(y_2(i)/y_1(i))$ (see [7] for details). Then, the first rotation matrix in the separation is:

$$\tilde{R}_1 = \begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) \\ -\sin(\theta_1) & \cos(\theta_1) \end{pmatrix}. \quad (7)$$

Following the first rotation, the now aligned parallelogram needs to be transformed into a diamond (Figure 3). More precisely, the axes need to be independently scaled so that the variance is rotationally invariant. The scaling of each axis is determined by first computing the variance along the axis of maximal and minimal variance, i.e., the axis oriented at θ_1 and $\theta_1 - \pi/2$:

$$s_1 = \sum_{i=1}^{N} \left[\begin{pmatrix} y_1(i) & y_2(i) \end{pmatrix} \begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix} \right]^2 \quad (8)$$

$$s_2 = \sum_{i=1}^{N} \left[\begin{pmatrix} y_1(i) & y_2(i) \end{pmatrix} \begin{pmatrix} \cos(\theta_1 - \pi/2) \\ \sin(\theta_1 - \pi/2) \end{pmatrix} \right]^2, \quad (9)$$

and then the scaling matrix is constructed by placing the inverse variances along the diagonal:

$$\tilde{S} = \begin{pmatrix} s_1^{-1} & 0 \\ 0 & s_2^{-1} \end{pmatrix}. \quad (10)$$

Combined, the first rotation and scaling are equivalent to PCA plus whitening. But notice from Figure 3 that this is insufficient for separating the mixed images into their independent components (the mixed images are only decorrelated, a necessary but insufficient condition). A final rotation is needed to separate the independent components.

One approach to the determination of the final rotation is to find the orientation θ_2 that maximizes the fourth statistical moment. The fourth moment at an arbitrary orientation is given by:

$$E(\theta_2) = \sum_{i=1}^{N} \left[\begin{pmatrix} y_1'(i) & y_2'(i) \end{pmatrix} \begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix} \right]^4, \quad (11)$$

where y_1' and y_2' are the result of rotating and scaling the initial mixed images y_1 and y_2 according to Equations (7) and (10). This error function cannot be solved analytically, but the following normalized fourth moment does lend itself to an analytic solution:

$$E'(\theta_2) = \frac{1}{||y_1'||^2 + ||y_2'||^2} E(\theta_2). \quad (12)$$

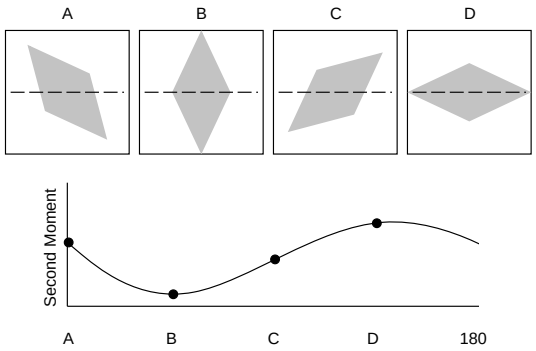

Figure 4: Shown is the variation in the second moment (variance) as the joint probability distribution is rotated and projected onto the horizontal axis (Equation (5)). The variance is minimal along the short axis and maximal at the orthogonal orientation.

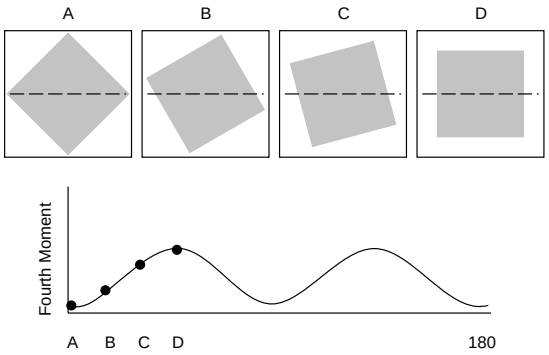

Figure 5: Shown is the variation in the fourth moment as the joint probability distribution is rotated and projected onto the horizontal axis (Equation (12)). The fourth moment is maximal at the orientation required to transform the diamond into a square, yielding the independent images.

The axis where the fourth moment is maximal is determined by finding the angle θ_2 that maximizes this error function (Figure 5). As before, we differentiate with respect to θ_2, set equal to zero and solve [7] yielding the maximal solution:

$$\theta_2 \quad = \quad \frac{1}{4}\tan^{-1}\left[\frac{\sum_{i=1}^{N} r^2(i)\sin(4\phi(i))}{\sum_{i=1}^{N} r^2(i)\cos(4\phi(i))}\right], \quad (13)$$

again, for convenience, expressed in polar coordinates. The final rotation matrix then takes the form:

$$\tilde{R}_2 \quad = \quad \begin{pmatrix} \cos(\theta_2) & \sin(\theta_2) \\ -\sin(\theta_2) & \cos(\theta_2) \end{pmatrix}. \quad (14)$$

The estimation of the source images X from the mixed images Y is now a simple matter of applying the three matrices in Equations (7), (10), and (14):

$$\tilde{X} \quad = \quad (\tilde{R}_2\tilde{S}\tilde{R}_1)Y. \quad (15)$$

There are two inherent ambiguities in the recovery of the independent components. First is the ordering ambiguity, that is, the following mixings are indistinguishable:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad = \quad \begin{pmatrix} b & a \\ d & c \end{pmatrix}\begin{pmatrix} x_2 \\ x_1 \end{pmatrix}. \quad (16)$$

Second is a scale ambiguity, that is, the independent components can only be determined within a scale factor since, for example, the following mixings are also indistinguishable.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad = \quad \begin{pmatrix} a/\gamma & b/\delta \\ c/\gamma & d/\delta \end{pmatrix}\begin{pmatrix} \gamma x_1 \\ \delta x_2 \end{pmatrix}. \quad (17)$$

For our purposes, the first of these ambiguities is not critical, and the second is dealt with by scaling the final images to fill the full intensity range.

3 Separating Reflections

In our first experiment we photographed a painting framed behind glass with the reflection of a mannequin ("Sheila"). The painting with reflection was photographed twice through a linear polarizer oriented so that the reflection was maximized, and then at the orthogonal direction to minimize the reflection. Note that even at the minimal orientation the reflection is still quite salient. We used a color digital video camera (Canon Optura DV, Canon Inc.) whose video signal was digitized through a S-video connection onto a Silicon Graphics computer. The camera was calibrated to ensure a linear response. The final three channel (RGB) color images have a spatial resolution of 640×480 pixels. In our results the entire three channel image was used in computing the independent sources.

Shown along the top row of Figure 6 are the initial images and shown along the bottom row are the results of separating the independent components. In between are the intermediate steps leading to the separation as specified by Equations (7), (10), (14) and (15). Note how the normalized joint histogram (third column) transforms in a similar manner to the idealization of Figure 3, yielding the independent components - the painting and the reflection of Sheila. Shown in Figure 7 are the results from an outdoor scene where the reflections were removed from a storefront window.

Note that even in the presence of an imperfect polarizer, a noisy sensor, imperfections in the glass, and the simple linear mixing model, we are able to separate the reflections leaving the images behind the glass largely intact.

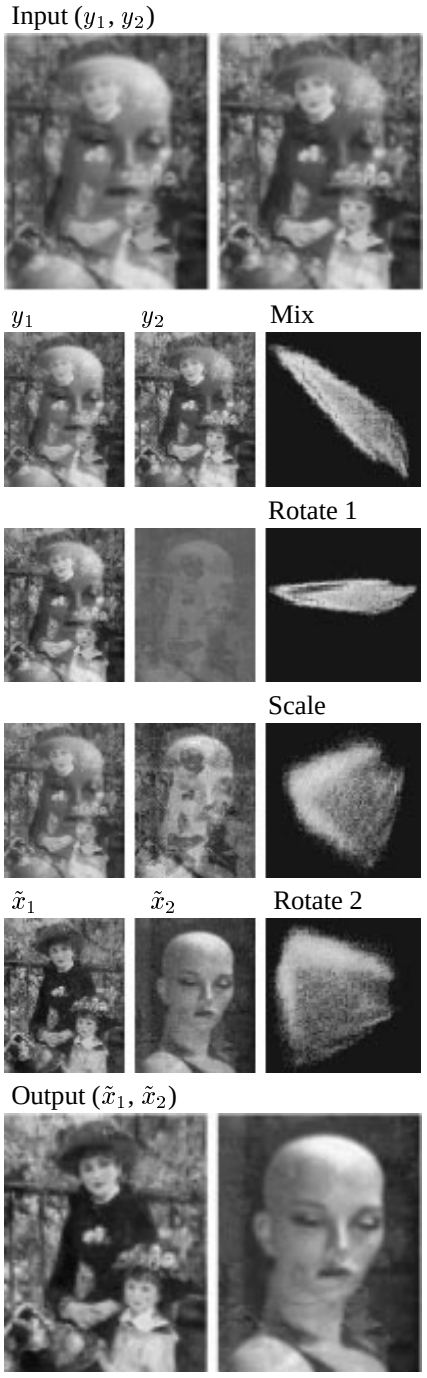

Figure 6: Along the top row are a pair of images of Renoir's *On the Terrace* with a reflection of Sheila photographed through a linear polarizer at orthogonal orientations. Along the bottom row are the independent components. Also shown are the intermediate steps leading to the separation of the independent components. The third column shows the normalized joint histogram of the pair of images to its left.

Figure 7: Shown along the top row are a pair of images photographed through a linear polarizer at orthogonal orientations, and shown along the bottom row are the separated components.

4 Separating Lighting

In addition to separating reflections, we have also used ICA to separate the relative contributions of individual light sources. A scene illuminated with a single spatially varying light source, $L(u,v)$, can be expressed as $I(u,v) = aL(u,v)R(u,v)$, where $R(u,v)$ is the reflectance function of the scene and a is a multiplicative constant that regulates the brightness of the light source. The same scene illuminated with two light sources can be expressed as,

$$\begin{aligned} y_1 &= aL_1R + bL_2R \\ &= ax_1 + bx_2, \end{aligned} \tag{18}$$

where, for convenience, the spatial parameters are dropped. Note that this image is a linear combination of the image of the scene illuminated with each light source independently. Consider now a second image taken with a different weighting of the light sources,

$$y_2 = cx_1 + dx_2. \tag{19}$$

Note that this pair of images is of the same form as that of Equation (3), and thus amenable to separation through ICA. However, the assumption of statistical independence is unlikely to strictly hold. In particular, although the same scene illuminated with different light sources can look dramatically different, it is most certainly the case that they

Figure 8: Shown along the top row are a pair of images photographed with two light sources of varying brightnesses, and shown along the bottom row are the separated components.

will *not* be independent (violating the basic assumption of ICA). Nevertheless, if there is enough independent structure in the joint distribution, then the separation outlined in the previous section will still be effective.

Shown along the top row of Figure 8 are a pair of images of Sheila. In both of these images a pair of light sources with varying brightness was used, the first light source positioned below and to the left, and second above and to the right. In the same figure are the separation results. Note that the shadows along the face and back wall are consistent with the individual light sources. Although not perfect, these results are impressive in that they show that ICA is a powerful separation tool even when the assumption of statistical independence does not strictly hold.

5 Discussion

To facilitate such tasks as object recognition, visual-based navigation, and scene segmentation it would be advantageous to separate the intrinsic aspects of an image from the incidental variations due to lighting, specularities/reflections, shadows, etc. We have shown how the statistical tool of independent components analysis (ICA) can be useful in this regard. In particular, we employed ICA in separating reflections from glass or glossy surfaces, and in separating the relative contributions of individual light sources. The separation of reflections may be useful to professional photographers, autonomous vehicles with inboard cameras where reflections from the vehicle windshield cause problems for navigation, and in the field of surveillance where activities behind a reflective window may be revealed. A real-time implementation may be realized by synchronizing the image capture with a liquid crystal polarizer (e.g, [12]).

There are of course several natural extensions to our work that will undoubtedly generalize its applicability. Most notably, the linear mixing model of Equation (3) assumes a spatially uniform linear combination of source images that is unlikely to always be true. To account for a mixing that may vary spatially across the field of view, our basic approach could be used in conjunction with a mixture model approach to fit multiple linear models. Another possible extension would be to employ a more generic version of ICA that allows for a redundant system with more measurements than unknowns.

References

[1] Y. Bar-Ness. Bootstrapping adaptive interference cancelers: Some practical limitations. In *The Globecom Conference*, pages 1251–1255, 1982.

[2] A. Bell and T. Sejnowski. An information maximisation approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.

[3] M. Born and E. Wolf. *Principles of Optics*. Pergamon, London, 1965.

[4] G. Brelstaff and A. Blake. Detecting specular reflections using lambertian constraints. In *International Conference on Computer Vision*, pages 297–302, Tampa, FL, December 1988.

[5] J. Cardoso. Source separation using higher order moments. In *International Conference on Acoustics, Speech and Signal Processing*, pages 2109–2112, 1989.

[6] P. Comon. Separation of stochastic processes. In *Workshop on Higher-Order Spectral Analysis*, pages 174–179, Vail, Colorado, June 1989.

[7] H. Farid and E. Adelson. Separating reflections from images using independent components analysis. Technical Report AI Memo 1647, MIT, 1998.

[8] C. Jutten and J. Herault. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24(1):1–10, 1991.

[9] S. Nayar, X. Fang, and B. Terrance. Separation of reflection components using color and polarization. *International Journal of Computer Vision*, 21(3):163–186, 1997.

[10] S. Shafer. Using color to separate reflection components. *Color Research and Applications*, 10:210–218, 1985.

[11] L. Wolff. Polarization-based material classification from specular reflection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(11):1059–1071, 1990.

[12] L. Wolff, T. Mancini, P. Pouliquen, and A. Andreou. Liquid crystal polarization camera. *IEEE Transactions on Robotics and Automation*, 13(2):195–203, 1997.