SEPARATION OF SEMIREFLECTIVE LAYERS USING SPARSE ICA

Alexander M. Bronstein, Michael M. Bronstein, Michael Zibulevsky and Yehoshua Y. Zeevi

Department of Electrical Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel

ABSTRACT

We address the problem of Blind Source Separation (BSS) of superimposed images and, in particular, consider the recovery of a scene recorded through a semirefective medium (e.g. glass windshield) from its mixture with a virtual reflected image. We extend the Sparse ICA (SPICA) approach to BSS and apply it to the separation of the desired image from the superimposed images, without having any *a priory* knowledge about its structure and/or statistics. Advances in the SPICA approach are discussed. Simulations and experimental results illustrate the efficiency of the proposed approach, and of its specific implementation in a simple algorithm of a low computational cost. The approach and the algorithm are generic in that they can be adapted and applied to a wide range of BSS problems involving one-dimensional signals or images.

1. INTRODUCTION

The phenomenon of a virtual image, being semireflected by a transparent medium, situated along the optical axis somewhere between the imaged scene and the observing point, and superimposed on the imaged scene, is typical of many optical setups. It may arise, for example, when photographing objects behind a glass window or windshield [9].

Approaches to reconstruction of the virtual and the real images, based on polarimetric imaging, have attracted attention during the last few years [4], [8]. Incorporation of a polarizer into the optical system is a common photographic technique allowing suppression of semireflected virtual images [9]. Several designs of such cameras, e. g. a system equipped with a liquid crystal polarizer [6], were recently proposed.

However, in most cases, the polarizer is not capable of removing the reflected component completely [5], [9]; even when the polarizer is oriented to minimize the reflected component, the virtual image is still visible.

Several signal post-processing approaches were proposed in recent studies, however, they rely mainly on motion, stereo and focus, and assume that the real and the virtual objects lie at significantly different distances from the camera [3], [10]. Other methods assume some knowledge about the scene, such as the semireflector angle and refraction index, which makes them hardly feasible in the general case [9].

Farid and Adelson [5] proposed to use an analytic version of independent component analysis (ICA) for blindly separating the reflected and the transmitted images. Such an approach does not require any prior knowledge regarding model parameters, and offers better feasibility in real-world applications. The proposed method is not general enough in that it is applicable in case having only two sources. On the other hand, iterative approaches such as the information maximization (Infomax) algorithm [2] are relatively slow, although they can handle any number of sources, provided a sufficient number of mixtures are available.

It has been recently demonstrated by Zibulevsky et al [11], [7] that the assumption of sparseness is very powerful and can significantly improve the accuracy and the computational efficiency of existing ICA algorithms. In addition, sparse decomposition allows using simple "geometric" algorithms to separation of mixed data. We adopt the Sparse ICA (SPICA) approach and show that it affords effective separation of a transmitted image from superimposed reflections [1].

2. THE BLIND SOURCE SEPARATION PROBLEM

In a typical BSS task, N mixtures are observed or received. Each of these available signals (or images), is assumed to be generated by a linear mixture of unknown *sources*, where the number of sources is usually assumed to be known to the observer, but this condition can be relaxed. Thus, an N-dimensional vector of observed signals is generated by the product of an unknown $N \times M$ *mixing matrix* A and an M-dimensional vector of unknown source signals. The task is to estimate the mixing matrix and then recover the source signals.

A typical embodiment of this problem in the context of superimposed reflections is depicted in Figure 1. The real object (a) is situated on the optical axis behind a semireflecting planar lens (d), inclined with respect to the optical axis [9]. Another object (b) is partially reflected by the lens, creating a virtual image (c). The camera (f) records a superposition of the two images. Thus, the intensity of the observed mixed image is given by:

$$m_1 = a_{11}s_1 + a_{12}s_2 \qquad (1)$$

where s_1 and s_2 are the images of two source objects (a) and (b), and a_{11}, a_{12} are constants, and their specific values depend on the optical geometry and properties of the reflective medium. It is assumed here that the problem is spatial invariant. This reasonably good approximation of the physical conditions can be relaxed.

Note that in the setup described so far we have a 1×2 mixing matrix. Since the reflected light is polarized, by introducing a linear polarizer (e), the relative weights of the two mixed images can be altered, thus yielding mixtures of the form

$$m_n = a_{n1}s_1 + a_{n2}s_2 \quad : \quad n = 1,\ldots,N \qquad (2)$$

or in matrix notation:

$$M = \mathrm{A} \cdot S \qquad (3)$$

where $m_1,...,m_n$ are the N mixed images and s_1 and s_2 are the source images represented as row vectors and A is the mixing (crosstalk) matrix.

The mixing matrix is usually unknown, unless side information regarding the physics of the mixing medium is available. In the context of the mixed reflections problem, this would require the availability of an exact optical model of the imaging system, visual scene and the medium. We, of course, do not assume any prior knowledge regarding the mixing matrix. Our goal is to determine the two source images S from the set of equations (3) with an unknown mixing matrix. This is the essence of the BSS problem.

Under the assumption that the sources are statistically independent (which is reasonable in the presented case), it is possible to recover sources s_1 and s_2 up to a permutation and multiplicative constant, by estimating the mixing matrix $\tilde{\mathrm{A}} \approx \mathrm{A}$, and estimating the sources by its inversion:

$$\tilde{S} = \tilde{\mathrm{A}}^{-1} \cdot M \qquad (4)$$

This problem is solved using the Sparse ICA method, discussed in the following section.

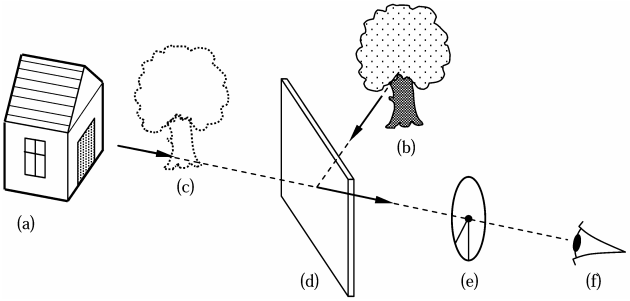

Figure 1 – A typical optical setup including a semireflector: (a) – object 1, (b) – object 2, (c) – virtual object, (d) – glass, (e) – polarizer, (f) – camera.

3. SPARSE ICA (SPICA)

Zibulevsky et al [11] have noticed that in case of *sparse* sources, their linear mixtures can be easily separated using very simple "geometric" algorithms. Given linear mixtures resulting from source images with the majority of pixels having a near-zero magnitude and under the assumption of statistical independence of the locations of the non-zero pixels in the sources, there is a high probability that only a single source will contribute to a given pixel in each mixture. Consequently, the majority of the pixels in each mixture will be influenced by one source only and have a magnitude equal to that of the source multiplied by the corresponding coefficient of the mixing matrix. In the scatter plot of one mixture versus the other these pixels will therefore lie along a line (each corresponding to a source) at a distance from the center depending on source magnitude. Hence, it is possible to reveal the ratios of each source's contribution to the mixtures by measuring the angles of each of the lines [1].

Geometric separation approaches are based on the detection of co-linearities in the distribution of the coefficients over the scatter plot. The straightforward way to recover the proper orientations from the scatter plot is by using the angular histogram. Applications of this approach are limited to low dimensions (practically, to 2D) due to the difficulty to construct the angular histogram in higher dimensions. The M points in the scatter plot (M equals the number of pixels in the image) are represented as points in N-dimensional space (in our case, $N = 2$). Since the mixtures are assumed to have a zero mean, the oriented co-linear distributions are centered at the origin. For each point $\boldsymbol{c}_k \in \mathbb{R}^2$, the angle

$$\alpha_k = \tan^{-1}\left(c_k^2 / c_k^1\right) \qquad (5)$$

is computed. Building the histogram of $\boldsymbol{\alpha}$, it is possible to detect the directions using a peak-detection algorithm.

An alternative approach is clustering along orientations of data concentration in the scatter plot. Each point $\boldsymbol{c}_k$ is projected on a unit hemisphere, by normalizing the data vectors:

$$\boldsymbol{c}_k = \boldsymbol{c}_k / \|\boldsymbol{c}_k\| \qquad (6)$$

and multiplying them by the sign of the first vector coordinate c_k^1 [7]. As the result, a number of clusters corresponding to the number of the sources is formed on the hemisphere. Applying some clustering algorithm, e.g. Fuzzy C-Means (FCM), it is possible to determine the cluster centers. The coordinates of the centers define the columns of the estimated mixing matrix, equivalent to the orientations found in the previous approach.

4. SPARSE DECOMPOSITION OF IMAGES

In Section 3 we showed simple geometric algorithms capable of separating mixtures of sparse images. It is obvious, however, that the sources in most applications, including the semireflective layer separation problem are natural images and have rather non-sparse nature. However, such images can be sparsely represented, i.e. there exists a linear transformation T such that

$$d_i = \mathrm{T}s_i \qquad (7)$$

is sparse. (Note that there is not necessary to be able to restore s_i from d_i, that is T does not necessarily have to be invertible). Application of the transformation to the mixtures in (3), due to the linearity of T, yields

$$\mathrm{T}m_i = \mathrm{T}\left(a_{i1}s_1 + a_{i2}s_2\right) = a_{i1}d_1 + a_{i2}d_2 \qquad (8)$$

Thus, the problem at hand is equivalent to separation of linearly-mixed sparse sources. This can be solved using the techniques described in Section 3.

Different classes of signals have their "natural" sparse transformations. In natural images (Figure 2a-b), for example, it is known that the edges usually have a sparse structure, hence even such a simple operation like a numeric derivative will yield a sparse image (Figure 2e-f). Figure 3 shows the scatter plot of two mixtures before (a) and after (b) applying the numeric derivative.

4.1. Multinode decomposition

Since there is no common sparse representation to different images, such a simple transformation as the derivative is usually data-dependent. Having this problem in mind, richer representations, that over a wide range of natural images lend themselves to relatively good sparse representations, such as the wavelet packet transform (WPT), were proposed [7], [16].

The task is to select only the nodes of the WPT, in which the decomposition is sparse. Kisilev et al [7] proposed an algorithm, in which the clustering procedure is first applied to each one of the nodes, but only nodes with minimal *global*

distortion (i.e. the mean-squared distance of data points to the centers of the closest clusters) are then selected for further processing. A more general approach is to assign some quality factor to each node, which determines its sparseness and then select a certain percentage of the "best" nodes in the sense of the assigned quality criterion. The choice of such a criterion will be discussed later in Section 4.2.

As an alternative to the WP decomposition, we propose to divide the image into blocks (possibly overlapping), compute some simple sparse transformation such as the first or the second order derivative (possibly concatenated) and only then to select the "best" blocks according to some sparseness criterion. Our observation is that most natural images have certain regions, in which edges and texture make such an approach efficient. Figure 4 depicts how the use of blocks can refine the sparseness and consequently the quality of the scatter plot in the previous example shown in Figure 2. The mixtures were partitioned into 16 blocks of equal size, and the same sparse transformation was applied to each block independently.

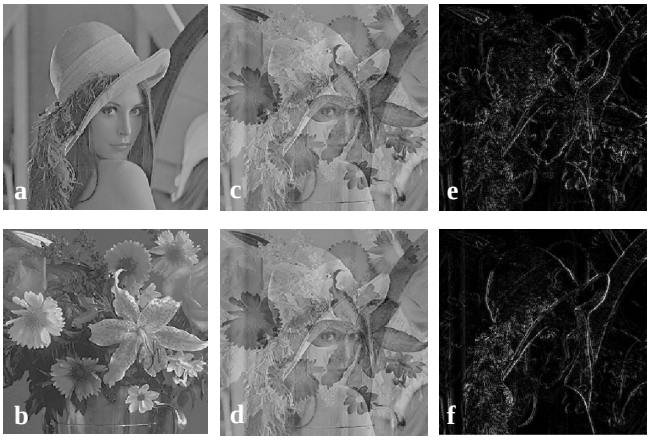

Figure 2 – (a)–(b) Non-sparse sources, (c) – (d) Synthetic mixtures, (e)–(f) The transformed mixtures obtained by the action of derivative in x-direction.

4.2. Quantitative sparseness criteria

Finding an adequate sparseness criterion is a crucial task for selecting the best nodes or blocks. The general problem of quantitative node sparseness estimation is to find such a function $q(\boldsymbol{x})$, which given a vector $\boldsymbol{x} \in \mathrm{R}^n$ returns a large value if it is sparse or a small value if it is not sparse. One of the possibilities is to use the so-called L_0 (threshold) norm, i.e. measure the number of vector coordinates, which are higher than some threshold τ:

$$q^{-1}(\boldsymbol{x}) = \tfrac{1}{n}\sum_{k=1}^{n} \mathbf{I}(x_k \geq \tau) \qquad (9)$$

where $\mathbf{I}$ is the indicator function. A natural choice of the threshold would be $\tau = \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|_2$, where $\bar{\boldsymbol{x}}$ is the mean value of $\boldsymbol{x}$. Another possible sparseness criterion is the L_p norm:

$$q^{-1}(\boldsymbol{x}) = n^{\frac{1}{2}-\frac{1}{p}} \cdot \|\boldsymbol{x}\|_p \big/ \|\boldsymbol{x}\|_2 \quad : \quad 0 < p \leq 1 \qquad (10)$$

Recent studies indicate that the L_1 norm is a more natural choice for dealing with various aspects of image quality criteria. This, normalized by the L_2 norm, as in (10), may turn out to be the best sparseness criterion. This, however, has yet to be further investigated.

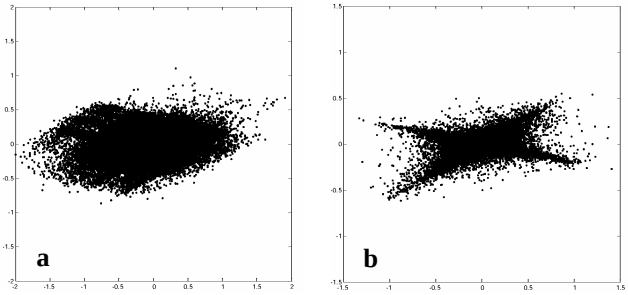

Figure 3 – Scatter plot of the mixtures m_2 vs m_1 before (a) and after (b) the sparse transformation.

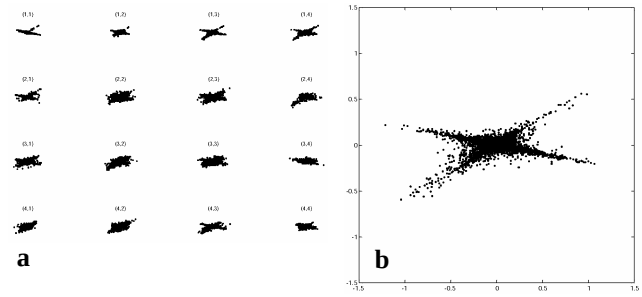

Figure 4 - Sparseness refinement by image partitioning into 16 blocks. Scatter plots of the coefficients in each block using x derivative as the sparse transformation (a) and a scatter plot resulting from merging the coefficients of blocks (1,1), (1,3) and (1,4) (b).

5. RESULTS

In this section, we apply the SPICA approach to polarization images obtained by simulation and photographed in real-world conditions. For comparison, we repeated the results of Farid and Adelson [5].

In the first experiment, the mixtures were obtained by artificially mixing two source images (Figure 5a-b). We used two SPICA methods: WP decomposition and block partitioning with second-order Sobel numeric derivative applied as sparse representation. The reconstruction was performed geometrically, using an angular histogram. The SNR values represented in Table 1 and Figure 5 reveal a considerable improvement by sparse representations, compared to the closed-form ICA [5].

To further test the performance of the SPICA algorithms and compare them to the Farid-Adelson closed-form ICA, we used the images of a painting ("Renoir"), framed behind glass, with a superimposed reflection of a mannequin ("Sheila"), photographed through a linear polarizer at orthogonal orientations[1].

The block partitioning approach is a natural way to handle the case of spatially-varying mixing coefficients, which often occur in reality. The acquired images were divided into four equally sized super-blocks and the separation problem was solved in each super-block separately. The coefficients of the

[1] Available from http://www.cs.dartmouth.edu/~farid/research/separation.html (courtesy of Hany Farid, Dartmouth College).

estimated unmixing matrix were then linearly interpolated over the entire image to produce a more accurate unmixing. Figure 6 shows the reconstruction results. The desired photographed image of Renoir's painting is recovered with high precision, without notable artifacts. The non-separated details in the reconstructed image of the mannequin, resulted from distortions due to imperfections of the optical system.

The estimated mixing matrix coefficients can be used to identify which image belongs to the real object and which to the virtual one (see [1] for details).

Table I – SNR (in dB) of the reconstructed sources A and B

	Farid & Adelson	SPICA (WP)	SPICA (blocks)
A	12.18	38.58	35.83
B	26.07	45.96	64.96

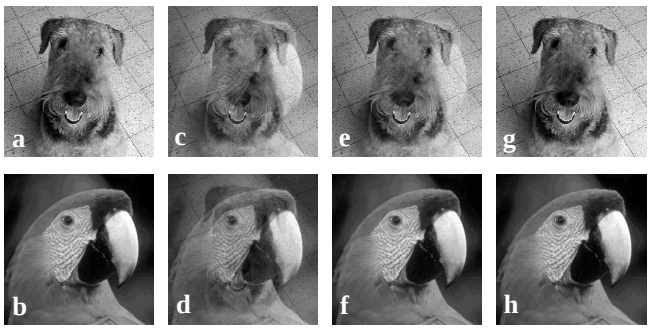

Figure 5 – Separation of synthetic mixtures. (a)-(b) Sources, (c)-(d) mixtures, (e)-(f) Reconstruction by the Farid-Adelson approach, (g)-(h) Reconstruction by SPICA with block partitioning.

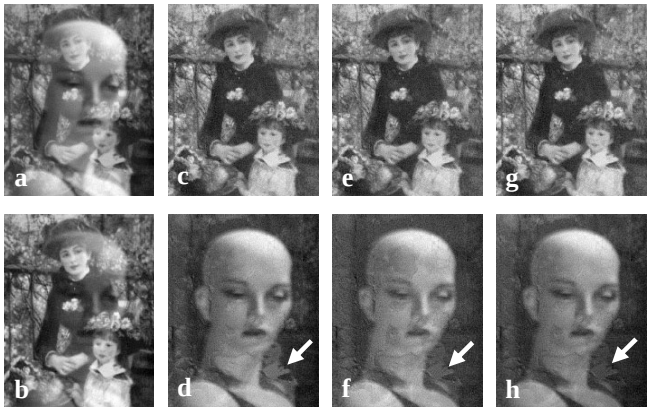

Figure 6 – Separation of synthetic mixtures. (a)-(b) Mixtures, (c)-(d) Reconstruction by the Farid-Adelson approach, (e)-(f) Reconstruction by SPICA with WP, (g)-(h) Reconstruction by SPICA with block partitioning.

6. CONCLUSIONS

The Sparse ICA approach can be effectively used in wide range of scenarios wherein various mixtures of source images are available for separation of the sources. In this study we were primarily concerned with separation of an image from virtual images superimposed on it by reflections from a semireflecting medium. The proposed novel sparse decomposition method incorporates block partitioning, suitable for nonstationary natural images, as well as for imaging systems such as polarized semireflecting media, that cannot be considered as spatial invariant systems, but can to a good approximation be dealt with as locally spatial invariant systems. Experiments conducted with simulated and photographed data show the efficiency of this approach and its advantages over previously-proposed methods.

We have assumed that only two images, acquired at perpendicular polarization angles, are available. One may extend the application to acquisition of more than two images by using principal component analysis (PCA) prior to the application of ICA.

7. ACKNOWLEDGEMENT

We thank Hani Farid for his images and mixtures data, for his codes and for his comments. This research has been supported by the Ollendorff Minerva Center, by the Fund for Promotion of Research at the Technion, by the Israeli Ministry of Science and by the ONR-MURI Grant N000M-01-1-0625, administered by Columbia University.

8. REFERENCES

[1] G. Avramovich, A. Bronstein, M. Bronstein, R. Feuerstein, Y. Y. Zeevi and M. Zibulevsky, "Sparse ICA for blind separation of transmitted and reflected images", Submitted to *JMLR*, 2002. Available at: http://visl.technion.ac.il/bron/works

[2] A. J. Bell, and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution", *Neural Comput*. 7 (6), pp. 1129–1159, 1995.

[3] J. R. Bergen, P. J. Burt, R. Hingorani, and S. Peleg, "Transparent motion analysis", *Proc. ECCV*, pp. 566-569, 1990.

[4] Cronin, T. W., Shashar, N., & Wolff, L. (1994). Portable imaging polarimeters, In *Proceedings ICPR*, Vol-A, pp. 606-609.

[5] H. Farid, and E. H. Adelson, "Separating reflections from images using independent component analysis", *JOSA*. Vol. 17, No. 9, pp. 2136-2145, 1999.

[6] H. Fujikake et al, "Electrically-controllable liquid crystal polarizing filter for eliminating reflected light", *Optical Review* 5, pp. 93-98, 1998.

[7] P. Kisilev, M. Zibulevsky, Y. Y. Zeevi, and B. A. Pearlmutter, "Multiresolution framework for blind source separation", CCIT Report # 317, Technion Press, 2000.

[8] K. S. Nayar, X. S. Fang, and T. Boult, "Separation of reflection components using color and polarization", *Int. J. Comp. Vis*. 21, pp. 163-186, 1997.

[9] Y. Y. Schechner, J. Shamir, and N. Kiryati, "Polarization and statistical analysis of scenes containing a semireflector", *JOSA*. A Vol. 17, No. 2, 2000

[10] M. Shizawa, "On visual ambiguities due to transparency in motion and stereo", *Proc. ECCV*, pp. 411-419, 1992

[11] M. Zibulevsky, and B. A. Pearlmutter, "Blind source separation by sparse decomposition, *Neural Comp*. 13(4), 2001.