BLIND SEPARATION OF REFLECTIONS WITH RELATIVE SPATIAL SHIFTS

Efrat Be'ery and *Arie Yeredor*

School of Electrical Engineering
Tel-Aviv University, Tel-Aviv, 69978, ISRAEL
e-mail: `<efratb,arie>@eng.tau.ac.il`

ABSTRACT

We address the problem of blind separation of image mixtures (resulting, e.g, from reflections through window glass) consisting of pure unknown relative spatial-shifts in addition to scalar mixing coefficients. Most (and maybe all) existing approaches to the problem assume static mixtures, i.e., the spatial positions of the source images are assumed to remain fixed between snapshots. We propose an integrated method to estimate both the mixing coefficients and the spatial-shifts using a second-order statistics based algorithm, which uses specially parameterized approximate joint diagonalization of two-dimensional-spectra matrices. The accommodation of spatial shifts allows the exploitation of further diversity between snapshots, and thus enables to attain improved separation, especially when the static mixing coefficients are ill-conditioned - as we demonstrate using simulations results.

1. INTRODUCTION

Quite commonly in photography, when scenes are recorded through window glass under poorly balanced lighting conditions, a reflection of the room interior appears superimposed over the outside scene. Separating such reflections from the scene of interest is quite difficult when a single image (snapshot) is available (although it has been attempted, e.g., in [1]). However, when (at least) two such snapshots are available, each taken under slightly different conditions, it is conceivable that successful separation can be attained by proper exploitation of the diversity in the different snapshots. Such an approach has been considered from different aspects, e.g., optical (based on diverse polarization [2] or on convolutive transformations [6] such as focusing [3]), or intensity-profile based source separation, based on diversity in the lighting conditions [4], [5].

Different relative spatial-shifts of the sources between snapshots are the inevitable outcome of possible movement of the camera / reflective medium / recorded objects from one snapshot to another. Moreover, such relative shifts can be introduced deliberately (e.g., by slightly slanting the window between snapshots), so as to introduce yet another source of diversity. To the best of our knowledge, the separation of images with different relative shifts has not yet been considered explicitly in open literature. The approaches mentioned above, as well as some other methods, assume either a model of static (unshifted) linear mixing, or a more general convolutive mixture model, which is generally too over-parameterized to address the relative-shifts model.

For single-dimensional signals, the extended AC-DC algorithm in [7] demonstrated a solution for the problem of mixed time-domain sources with different relative time-delays. In this paper we address a model similar to the single-dimensional one considered in [7], but extend the approach of [7] to accommodate two-dimensional (2D) image signals, with time-delays substituted by spatial-shifts.

The shifts are assumed to have occurred prior to the sampling (digital imaging) process, and are therefore not necessarily an integer multiple of pixels. The (pre-sampled, continuous-space) source signals (images) are assumed to be mutually uncorrelated, wide-sense stationary (WSS) with unknown spectra. Yet, it has to be stressed, that the stationarity assumption is only used for simplifying the derivation, and is not instrumental for the resulting performance. The more important condition for successful separation in this context, is that the sources be mutually uncorrelated. Luckily, this property is widely satisfied by images of independent sources, unlike the stationarity assumption, which is rarely satisfied in natural images.

We consider the following L sources - L sensors model (to be later reduced to $L = 2$):

$$x_p(u,v) = \sum_{q=1}^{L} a_{pq} s_q(u - d_{pq}^u, v - d_{pq}^v) \quad p = 1, 2, \dots L \tag{1}$$

where $s_q(u,v)$ are 2D zero-mean, WSS processes with unknown spectra, $x_p(u,v)$ are the observations, a_{pq} are the mixing coefficients and d^u_{pq}, d^v_{pq} are the shifts in u, v directions of source image q in mixture p with respect to its position in mixture 1. Without loss of generality, we use as a "working assumption" zero shifts of the source images in the first mixture, i.e., $d^u_{1q} = d^v_{1q} = 0$ for $q = 1, 2, \ldots L$.

The available data are samples of the continuous-space observations, $x_p[m, n] = x_p(m\Delta, n\Delta) \quad 1 \leq m \leq M,\ 1 \leq n \leq N$, where Δ is the sampling interval, assumed to comply with the Nyquist rate (we shall use parentheses / brackets to enclose continuous- / discrete- time indices, respectively).

The paper is organized as follows: In the next section we formulate the estimation problem as a specially parameterized approximate joint diagonalization (AJD) problem in the frequency domain; In section 3 we propose an iterative algorithm for the AJD, based on an extension of the AC-DC algorithm for non-orthogonal AJD [8]; In section 4 we present some simulation results.

2. FORMULATION AS A JOINT DIAGONALIZATION PROBLEM

The observations' correlation functions are given by

$$R^x_{mn}(\xi, \eta) \triangleq E[x_m(u+\xi, v+\eta)x_n(u, v)] = \sum_{p,q=1}^{L} a_{mp}a_{nq} \cdot$$
$$\cdot E[s_p(u - d^u_{mp} + \xi, v - d^v_{mp} + \eta)s_q(u - d^u_{nq}, v - d^v_{nq})]$$
$$= \sum_{q=1}^{L} a_{mq}a_{nq}R^s_q(\xi + d^u_{nq} - d^u_{mq}, \eta + d^v_{nq} - d^v_{mq})$$
$$1 \leq m, n \leq L \quad (2)$$

where $R^x_{mn}(\xi, \eta)$ denotes the correlation between the m-th and the n-th mixed images, and $R^s_q(\xi, \eta)$ denotes the autocorrelation of the q-th source image. Fourier-transforming (2), we obtain $S^x_{mn}(\omega, \theta)$, the cross-spectrum between the m-th and n-th mixtures

$$S^x_{mn}(\omega, \theta) = \sum_{q=1}^{L} a_{mq}a_{nq}S^s_q(\omega, \theta)e^{-j\omega(d^u_{mq}-d^u_{nq})-j\theta(d^v_{mq}-d^v_{nq})}$$
$$1 \leq m, n \leq L \quad (3)$$

where $S^s_q(\omega, \theta)$ is the q-th source's unknown spectrum. Eq. (3) can also be expressed in matrix-form as

$$\boldsymbol{S}_x(\omega, \theta) = \boldsymbol{B}(\omega, \theta)\boldsymbol{S}_s(\omega, \theta)\boldsymbol{B}^H(\omega, \theta) \quad (4)$$

where $\boldsymbol{S}_x(\omega, \theta)$ is an $L \times L$ matrix consisting of $S^x_{mn}(\omega, \theta)$ as the m, n-th element, $\boldsymbol{S}_s(\omega, \theta)$ is an $L \times L$ diagonal matrix consisting of $S^s_q(\omega, \theta)$ as its (q, q)-th elements, and $\boldsymbol{B}(\omega, \theta)$ is the $L \times L$ matrix given by $\boldsymbol{B}(\omega, \theta) = \boldsymbol{A} \odot \boldsymbol{D}^u(\omega) \odot \boldsymbol{D}^v(\theta)$, where $\odot$ denotes Hadamard's (element-wise) product, $\boldsymbol{A}$ is the constant matrix of mixing coefficients, whose m, n-th element is a_{mn}, and $\boldsymbol{D}^u(\omega)$,$\boldsymbol{D}^v(\theta)$ contain the shifts, such that their m, n-th elements are given by

$$D^u_{mn} = e^{-j\omega d^u_{mn}}, D^v_{mn} = e^{-j\theta d^v_{mn}} \quad 1 \leq m, n \leq L. \quad (5)$$

The cross-spectral matrices $\boldsymbol{S}_x(\omega, \theta)$ are unknown, but can be estimated from the available data, possibly by using the 2D Discrete-Space Fourier Transform (DSFT) of a truncated series of biased cross-correlations estimates (Blackman-Tuckey's method, e.g., [9]). Specifically, to estimate the m, n-th element of $\boldsymbol{S}_x(\omega)$, we may use

$$\hat{S}^x_{mn}(\omega, \theta) = \sum_{\ell=-P_u}^{P_u} \sum_{k=-P_v}^{P_v} \hat{R}_{mn}[\ell, k]e^{-j\omega\ell-j\theta k} \quad (6)$$

where P_w is the truncation-window length in direction $w = u, v$ and

$$\hat{R}_{mn}[\ell, k] = \frac{1}{M}\frac{1}{N} \sum_{p=1}^{M-|\ell|} \sum_{q=1}^{N-|k|} x_m[p + \ell, q + k]x_n[p, k]$$
$$-P_u \leq \ell \leq P_u\ ,\ -P_v \leq k \leq P_v \quad (7)$$

When estimated values, rather than true values, of $\boldsymbol{S}_x(\omega, \theta)$ are used, the equations (4) usually can no longer be satisfied simultaneously at all frequencies. Nevertheless, once $\boldsymbol{S}_x(\omega, \theta)$ is estimated at several frequencies-pairs $(\omega_0, \theta_0), (\omega_0, \theta_1), \ldots (\omega_k, \theta_j), \ldots (\omega_K, \theta_J)$, an estimate of the unknown parameters of interest can be obtained by resorting to AJD (see e.g. [8]), seeking to minimize the following least-squares (LS) criterion:

$$\min_{\boldsymbol{A},\Delta_u,\Delta_v,\boldsymbol{\Gamma}} C_{LS} \triangleq \sum_{k=0}^{K} \sum_{j=0}^{J} ||\boldsymbol{S}_x(\omega_k, \theta_j) -$$
$$- \boldsymbol{B}(\omega_k, \theta_j)\boldsymbol{S}_s(\omega_k, \theta_j)\boldsymbol{B}^H(\omega_k, \theta_j)||^2_F \quad (8)$$

where Δ_w $(w = u, v)$ is an $L \times L$ matrix containing the shift parameters d^w_{mn}, $\boldsymbol{\Gamma}$ is an $L \times (K + 1) \times (J + 1)$ (3-dimensional) matrix containing the sources' spectra, $\gamma_{mkj} = S^s_m(\omega_k, \theta_j)\ \ 1 \leq m \leq L;\ 0 \leq k \leq K;\ 0 \leq j \leq J$ and $|| \cdot ||^2_F$ denotes the squared Frobenius norm. Note that it is also possible to use a weighted LS criterion by introducing some positive weights w_{kj} into the sum; however, to simplify the exposition, we shall not pursue this possibility in here.

Several AJD algorithms have been proposed in recent years, however they all assume a fixed diagonalization matrix $\boldsymbol{B}$, rather than $\boldsymbol{B}(\omega_k,\theta_j)$ which depends on the indices (k,j). In [7] an extension of one particular AJD algorithm (AC-DC, [8]) was proposed to address the one-dimensional problem, i.e., $\boldsymbol{B}(\omega_k)$. In the next section we propose further extension of the extended AC-DC, adapted to this 2D minimization problem.

3. AJD VIA EXTENDED 2D-AC-DC

AC-DC ("Alternating Columns / Diagonal Centers", [8]) is an alternating-directions minimization algorithm, originally intended for the case of a fixed diagonalizing matrix $\boldsymbol{B}$. In our case the matrix $\boldsymbol{B}$ is not constant, but can be factored so as to depend on three constant matrices, $\boldsymbol{A}$, Δ_u and Δ_v. It is then possible to minimize with respect to (w.r.t.) each column of $\boldsymbol{A}$ and pair of matching columns (columns with the same index) of Δ_u and Δ_v separately, thus alternating between minimizations w.r.t.: $\boldsymbol{\Gamma}$ (DC phase); each column of $\boldsymbol{A}$ (AC-1 phase); pair of matching columns of Δ_u and Δ_v (AC-2 phase). Due to limited space, only final expressions will be given here for the DC and AC-1 phases, based on the algebraic work done in [7]. The AC-2 phase, being the fundamental innovation of this paper, will be described more extensively.

3.1. The "DC" phase

In the DC phase we wish to minimize C_{LS} w.r.t. $\boldsymbol{\Gamma}$, with $\boldsymbol{A}$, Δ_u and Δ_v fixed. As described in [7] the minimization can be decomposed into $(K+1)\times(J+1)$ distinct minimization problems which are all linear in the unknown parameters, and thus admit the well-known linear LS solution. We define $\boldsymbol{\gamma}_{kj}$ to be the (k,j)-th column of $\boldsymbol{\Gamma}$, $\boldsymbol{y}_{kj} \triangleq \mathrm{vec}\{\boldsymbol{S}_x(\omega_k,\theta_j)\}$ ($\mathrm{vec}\{\cdot\}$ denoting the concatenation of the matrix' columns into one vector), and $\boldsymbol{H}_{kj} = (\boldsymbol{B}(\omega_k,\theta_j)^* \otimes \boldsymbol{1}) \odot (\boldsymbol{1} \otimes \boldsymbol{B}(\omega_k,\theta_j))$ where $\boldsymbol{1}$ denotes an $L\times 1$ vector of 1-s, $\otimes$ denotes Kronecker's product and the superscript * denotes conjugation. The minimizer of the linear LS problem is

$$\boldsymbol{\gamma}_{kj} = [\boldsymbol{H}_{kj}^H \boldsymbol{H}_{kj}]^{-1} \boldsymbol{H}_{kj}^H \boldsymbol{y}_{kj}. \tag{9}$$

3.2. The "AC-1" phase

We now wish to minimize C_{LS} w.r.t. $\boldsymbol{a}_\ell$, the ℓ-th ($\ell = 1,2,\dots L$) column of $\boldsymbol{A}$, assuming the other columns, as well as Δ_u, Δ_v and $\boldsymbol{\Gamma}$, are fixed. We Define

$$\tilde{\boldsymbol{S}}(\omega_k,\theta_j) \triangleq \boldsymbol{S}_x(\omega_k,\theta_j) - \sum_{\substack{n=1\\ n\neq\ell}}^{L} S_n^s(\omega_k,\theta_j)\boldsymbol{b}_n(\omega_k,\theta_j)\boldsymbol{b}_n^H(\omega_k,\theta_j) \tag{10}$$

where $\boldsymbol{b}_n(\omega_k,\theta_j)$ is the n-th column of $\boldsymbol{B}(\omega_k,\theta_j)$. Observe now, that $\boldsymbol{b}_\ell(\omega_k,\theta_j)$ can be written as $\boldsymbol{b}_\ell(\omega_k,\theta_j) = \boldsymbol{\Lambda}_\ell(\omega_k,\theta_j)\boldsymbol{a}_\ell$ where $\boldsymbol{\Lambda}_\ell(\omega_k,\theta_j) = \mathrm{diag}\{e^{-j\omega_k d_{1\ell}^u - j\theta_j d_{1\ell}^v}, e^{-j\omega_k d_{2\ell}^u - j\theta_j d_{2\ell}^v}, \dots, e^{-j\omega_k d_{L\ell}^u - j\theta_j d_{L\ell}^v}\}$. The obtained minimizing solution (see [7]) is

$$\boldsymbol{a}_\ell = \sqrt{\lambda_{max}/f} \cdot \boldsymbol{\alpha}_{max} \tag{11}$$

Where $f \triangleq \sum_{j,k=0}^{J,K} (S_\ell^s(\omega_k,\theta_j))^2$ and $\boldsymbol{\alpha}_{max}$ is the eigenvector of $\mathrm{Real}\{\boldsymbol{F}\}$ associated with the largest (*positive*) eigenvalue, λ_{max}. $\boldsymbol{F}$ is the Hermitian matrix: $\boldsymbol{F} \triangleq \sum_{j,k=0}^{J,K} S_\ell^s(\omega_k,\theta_j)\boldsymbol{\Lambda}_\ell^*(\omega_k,\theta_j)\tilde{\boldsymbol{S}}(\omega_k,\theta_j)\boldsymbol{\Lambda}_\ell(\omega_k,\theta_j)$.

3.3. The "AC-2" phase

It is now desired to minimize C_{LS} w.r.t. $\boldsymbol{d}_\ell^u$ and $\boldsymbol{d}_\ell^v$, the ℓ-th ($\ell = 1,2,\dots L$) columns of Δ_u and Δ_v respectively, assuming the other columns, as well as $\boldsymbol{A}$ and $\boldsymbol{\Gamma}$ are fixed. Using the definitions from the AC-1 phase we can obtain, through some algebraic manipulations

$$C_{LS} = \tilde{C} - 2\boldsymbol{a}_\ell^T \boldsymbol{F} \boldsymbol{a}_\ell + (\boldsymbol{a}_\ell^T \boldsymbol{a}_\ell)^2 f. \tag{12}$$

where $\tilde{C}$ is a constant. Since the dependence on the shifts $\boldsymbol{d}_\ell^u$, $\boldsymbol{d}_\ell^v$ appears only through $\boldsymbol{F}$, minimization of C_{LS} requires maximization of $\boldsymbol{a}_\ell^T \boldsymbol{F} \boldsymbol{a}_\ell$, or, more explicitly, maximization w.r.t. $\{d_{p\ell}^u, d_{p\ell}^v\}_{p=2}^L$ of

$$\sum_{p,q=1}^{L} a_{p\ell} \left[\sum_{j,k=0}^{J,K} G_{pq}^\ell(\omega_k,\theta_j) e^{-j(d_{q\ell}^u - d_{p\ell}^u)\omega_k - j(d_{q\ell}^v - d_{p\ell}^v)\theta_j} \right] a_{q\ell} \tag{13}$$

where $\boldsymbol{G}^\ell(\omega_k,\theta_j) \triangleq S_\ell^s(\omega_k,\theta_j)\tilde{\boldsymbol{S}}(\omega_k,\theta_j)$. If we focus on the (relatively) simple $L=2$ case: Note that in this case, only two elements (out of four) in the outer summation in (13) depend on the unknown shifts - the elements corresponding to $p \neq q$. Thus, for $\ell = 1$ we seek to maximize, w.r.t. d_{21}^u and d_{21}^v:

$$\begin{aligned} & a_{11}a_{21} \left[\sum_{k=0}^{K}\sum_{j=0}^{J} G_{12}^1(\omega_k,\theta_j) e^{-jd_{21}^u\omega_k - jd_{21}^v\theta_j} \right] + \\ & + a_{11}a_{21} \left[\sum_{k=0}^{K}\sum_{j=0}^{J} G_{21}^1(\omega_k,\theta_j) e^{+jd_{21}^u\omega_k + jd_{21}^v\theta_j} \right] \end{aligned} \tag{14}$$

Observing that due to conjugate-symmetric structure of $\boldsymbol{G}^1(\omega,\theta)$ these two terms are a conjugate pair and that $a_{ij} \geq 0$, we end up maximizing

$$\max_{u,v} \mathrm{Real}\{g_1(u,v)\} \tag{15}$$

where $g_1(u,v) \triangleq \sum_{j,k=0}^{J,K} G_{21}^1(\omega_k,\theta_j) e^{ju\omega_k + jv\theta_j}$. If the frequencies $\{\omega_k\}$ and $\{\theta_j\}$ are chosen as $\omega_k = k\Omega,\ \ k =$

$0, 1, \ldots K$ and $\theta_j = j\Theta,\ \ j = 0, 1, \ldots J$ (with Ω and Θ selected constants), then $g_1(u, v)$ is the IDFT of the sequence $G^1_{21}(\omega_k, \theta_j)$. The value of the constants Ω and Θ determines the resolution of the search. The search may be conducted in a pre-defined range of the possible images' shifts. The same way, for $\ell = 2$ (looking for d^u_{21}, d^v_{21}) we need to maximize $\text{Real}\{g_2(u, v)\}$ with $g_2(u, v) \triangleq \sum_{j,k=0}^{J,K} G^2_{21}(\omega_k, \theta_j) e^{ju\omega_k + jv\theta_j}$. A rough "intelligent" initial guess for the shifts may be obtained from observing the locations of the (two) peaks of the cross-correlation between the mixtures ($R^x_{21}(\xi, \eta)$).

3.4. Pre-Processing Stage: Edge Sharpening

Since the important condition of uncorrelated sources is generally better satisfied between edge-enhanced images than by their original counterparts, we attained considerably better estimation (hence, separation) when a linear shift-invariant (zero-phase, 9×9) edge-sharpening filter was applied to both mixtures as a preprocessing stage. Such filtering is merely equivalent to the introduction of spectral shaping, resulting in special weighting of the LS criterion (8), which attributes more weight to respective frequencies in the AJD process.

4. SIMULATIONS RESULTS AND CONCLUSIONS

We present two experiments of synthetic mixture separation. In the first experiment we artificially mixed the two source images with a non-singular mixing matrix $\boldsymbol{A} = \left[\begin{smallmatrix} 0.65 & 0.35 \\ 0.42 & 0.58 \end{smallmatrix}\right]$. In the second experiment we simulated a shifts-only mixture scenario, therefore we used a singular matrix $\boldsymbol{A} = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, so as to demonstrate the attainable separability induced by the shifts in a case that would otherwise be inseparable. In both experiments the true shifts matrices were $\Delta_u = \left[\begin{smallmatrix} 0 & 0 \\ -1 & 2 \end{smallmatrix}\right]$ $\Delta_v = \left[\begin{smallmatrix} 0 & 0 \\ 3 & 1 \end{smallmatrix}\right]$. We applied 2D-AC-DC to both, using, as an initial guess, the identity matrix for $\boldsymbol{A}$ and all-zeros matrices for Δ_u, Δ_v.

In the non-singular case we also compared to the result of separating with the true (inverse) mixing matrix while ignoring the spatial-shifts: Images (a),(b) in Figure 2 demonstrate the importance of shifts-estimation. Even when the true mixing matrix is known, the separation results are poor if the spatial-shifts are not considered. Shifts-estimation also enables separation in the shifts-only ("singular") mixtures case, as can be observed from Images (e),(f) in Figure 2 (note, however, that some related tolerable artifacts can also be observed in that case).

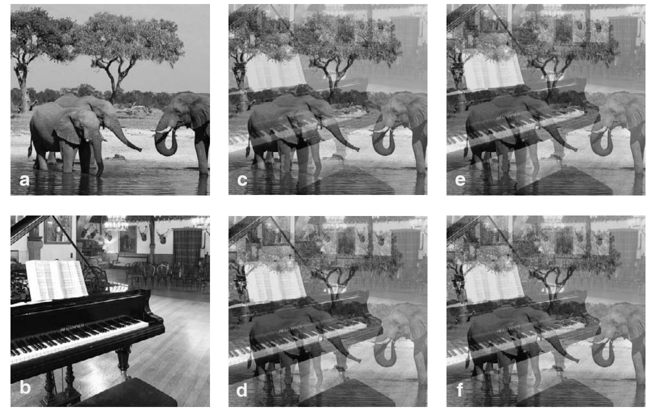

Fig. 1: (a)(b) Sources, (c)(d) Mixture with non-singular coefficients, (e)(f) Mixture with a singular coefficients matrix

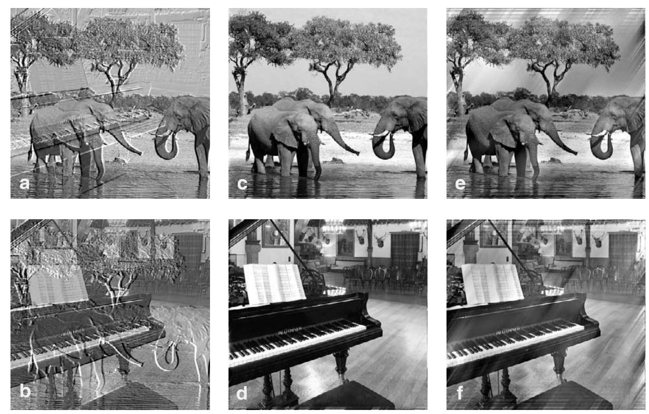

Fig. 2.: (a)(b) Zero-shift clairvoyant separation, (c)(d) Non-singular case separation, (e)(f) Singular case separation

5. REFERENCES

[1] A. Levin, A. Zomet, and Y. Weiss, "Separating reflections from a single image using local features", *Proceedings, CVPR 2004*, vol. 1, pp. 306–313, 2004.

[2] Y.Y. Schechner, J. Shamir, and N. Kiryati, "Polarization and statistical analysis of scenes containing a semireflector", *Journal of the Optical Society of America*, vol. 17, no. 2, pp. 276–284, 2000.

[3] Y. Y. Schechner, N. Kiryati, and R. Basri, "Separation of transparent layers using focus", *Int. Journal of Computer Vision*, vol. 39, no. 1, pp. 25–39, 2000.

[4] K.I. Diamantaras and T. Papadimitriou, "Blind separation of reflections using the image mixtures ratio", *Proceedings, ICIP2005*, pp. 1034–1037, 2005.

[5] A.M. Bronstein, M.M. Bronstein, M. Zibulevski, and Y.Y. Zeevi, "Blind separation of reflections using sparse ica", *Proceedings, ICA2003*, pp. 227–232, 2003.

[6] M. Castella and J.-C. Pesquet, "An iterative blind source separation method for convolutive mixtures of images", *Proceedings of ICA2004*, pp. 922–929, 2004.

[7] A. Yeredor, "Blind source separation with pure delays mixture", *Proceedings, ICA2001*, 2001.

[8] A. Yeredor, "Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation", *IEEE Trans. Signal Processing*, vol. 50, no. 7, pp. 1545–1553, 2002.

[9] S.M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice-Hall, 1993.