SEISMIC WAVE SEPARATION BY MEANS OF ROBUST PRINCIPAL COMPONENT ANALYSIS

L. T. Duarte[1], *E. Z. Nadalin*[2], *K. Nose Filho*[2], *R. A. Zanetti*[2] *J. M. T. Romano*[2], *M. Tygel*[3]

1-School of Applied Sciences, University of Campinas (UNICAMP), Brazil
2-School of Electrical and Computer Engineering, University of Campinas (UNICAMP), Brazil
3-Department of Applied Mathematics, IMECC, University of Campinas (UNICAMP), Brazil

leonardo.duarte@fca.unicamp.br, nadalin@dca.fee.unicamp.br, knfilho@dmo.fee.unicamp.br, rzanetti@gmail.com
romano@dmo.fee.unicamp.br, tygel@ime.unicamp.br

ABSTRACT

In this work, we investigate the application of the recently introduced signal decomposition method known as robust principal component analysis (RPCA) to the problem of wave separation in seismic data. The motivation of our research comes from the observation that the elements of the decomposition performed by RPCA can be associated with particular structures that often arise in seismic data. Results obtained considering two different situations, the separation of crossing events and the separation of diffracted waves from reflected ones, confirms that RPCA is a promising tool in seismic signal processing, outperforming the classical singular value decomposition (SVD) and the extension of the SVD based on independent component analysis in most cases.

Index Terms— Seismic signal processing, robust principal component analysis, wave separation, SVD.

1. INTRODUCTION

The separation of the different types of waves present in seismic data is a very relevant task in seismic signal processing [1]. This is specially true when seismic prospecting is considered, since a reliable interpretation of individual waves is crucial to identification of key geological structures in the subsurface under analysis. For instance, the separation of diffracted waves from reflections can be of use to identify stratigraphic traps, such as geological faults, in which hydrocarbons is often accumulated [2].

Classically, wave separation (or event separation) is performed by filtering methods such as the 2-D Fourier transform (or *f-k* filtering, as is known among geophysicists) and the Radon transform [3]. A third route to wave separation, which is the one we are interested in here, is the Singular Value Decomposition (SVD) [4]. This method has been intensively applied in different contexts, such as wavefield separation of normal moveout (NMO)-corrected common-midpoint (CMP) gathers, residual static corrections [5], diffraction separation [6] and ground-roll attenuation [7].

Although computationally efficient, the use of SVD to perform wave separation has some limitations. For instance, the estimation obtained by the SVD may not be good when there are crossing events, as well as the presence of horizontal and non-horizontal events in the same data. Such a drawback can be attributed [8] to the fact that SVD imposes orthogonality to all the elements present in the decomposition, which is implicitly equivalent to enforcing decorrelation in the separation process. To overcome these limitations, extensions of the SVD have been proposed. In [8], for instance, a modified version of the SVD based on Independent Component Analysis (ICA) was introduced. In this approach, higher-order statistics of the data are also taken into account, which results in a better performance both in wave separation [8] and signal-to-noise enhancement of pre-stack seismic gathers [9].

In this work, we aim to extend the SVD approaches in seismic signal processing upon the incorporation of the recently introduced decomposition framework known as robust principal component analysis (RPCA) [10, 11]. Roughly speaking, RPCA aims at decomposing the observed multidimensional data as a sum of a low-rank matrix and a sparse matrix. The key aspect here is that such a decomposition perfectly matches some situations typical of wave separation. More specifically, we compare RPCA with SVD and the SVD-ICA methods in two situations of great interest in seismic signal separation.

The paper is organized as follows: In Section 2, we introduce the problem and briefly describe the three separation strategies considered in our work. In Section 3, a set of numerical experiments illustrates our proposed procedures. Finally, Section 4 states our conclusions.

2. WAVE SEPARATION METHODS

2.1. Preliminaries: seismic data

Seismic data comprise an ensemble of traces, i.e., signals recorded in time at a given receiver location, due to a given source location. The receivers are typically an array of sensors, which can be geophones, in land acquisition, or hydrophones, in marine acquisition. The set of traces (referred to as a seismogram) can be represented by an $n \times m$ space-time matrix, $\mathbf{X}$, where m represents the number of traces and n the number of time samples for each trace. An example of seismogram is depicted in Figure 1. In this example, wave separation methods can be applied to separate a linear event from a nonlinear one. In the sequel, we briefly describe the three separation methods that will be considered in this work.

2.2. SVD

The SVD decomposes the data $\mathbf{X}$ as follows

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \tag{1}$$

where r is the minimum between n and m, $\mathbf{U}$ and $\mathbf{V}$ are square orthogonal matrices, and $\mathbf{D}$ is a rectangular diagonal matrix of size $m \times n$. The vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ correspond to the i-th column of the matrices $\mathbf{U}$ and $\mathbf{V}$, respectively. In seismic signal processing, these vectors are known as wavelet and propagation vectors, respectively [1]. Moreover, the matrix $\mathbf{u}_i\mathbf{v}_i^T$ is usually called the i-th eigen-image of $\mathbf{X}$ [5], accounting for the influence of the i-th eigen-image on the decomposition weighted by σ_i, which corresponds to the i-th singular value of $\mathbf{X}$. The SVD of a matrix is closely related to principal component analysis (PCA), a classical multivariate statistics decomposition tool.

The SVD to process seismic data is designed to extract events characterized by a high degree of trace-to-trace correlation [5]. This is possible because such events tend to be concentrated within the signal subspace, i.e., the eigen-images associated with the highest singular values. Therefore, if one considers only these eigen-images when reconstructing the original data, it becomes possible to estimate special, for instance, horizontal events. SVD can also be used to separate linear events that are not horizontally aligned. This can be done by performing a prior time correction to the seismic section in order to align the desired event [5].

2.3. SVD-ICA

One of the limitations of SVD in seismic signal processing is that it imposes the matrices $\mathbf{U}$ and $\mathbf{V}$ to be orthogonal. While the orthogonality of the wavelet vectors is an acceptable assumption, there is no physical reason for which the propagation vectors must be orthogonal [8]. As a consequence, enforcing such unnatural constraint usually introduces artifacts in the estimated subspaces.

The SVD-ICA method was developed [8] to overcome the aforementioned limitation. The first step in SVD-ICA is to apply the ordinary SVD to the data $\mathbf{X}$. Then, an ICA method is applied to the orthogonal matrix $\mathbf{U}$ with the aim of adjusting an orthogonal matrix $\mathbf{B}$ in such a way that the columns of the matrix $\tilde{\mathbf{U}} = \mathbf{U}\mathbf{B}$ become as statistical independent as possible. This can be done, for instance, by a joint diagonalization procedure of the data cumulant matrices.

After the application of ICA, and consideration of the ordinary SVD expression, there is a rearrangement of terms so the original data can be expressed as follows:

$$\mathbf{X} = \tilde{\mathbf{U}}\tilde{\mathbf{D}}\tilde{\mathbf{V}}^T = \sum_{i=1}^{r} \tilde{\sigma}_i \tilde{\mathbf{u}}_i \tilde{\mathbf{v}}_i^T. \tag{2}$$

There are two important differences between this new decomposition and the one of SVD. Firstly, the wavelet vectors are now statistically independent, instead of being only decorrelated — this is possible because ICA considers the higher-order statistics of the data. Secondly, due to the performed data rearrangement, the matrix $\tilde{\mathbf{V}}$ is no longer orthogonal.

SVD-ICA can also be used to separate linear events from other events and from noise. This can be achieved by considering the first eigen-images related to the decomposition (2).

2.4. Robust PCA

Recently, considerable attention has been paid to a novel decomposition framework tailored to the case in which the data comprises a low-rank term and a sparse part [10, 11]. This approach can be seen as a robust extension of the SVD/PCA framework. Indeed, as opposed to PCA, these new robust PCA (RPCA) techniques are able to extract the low-rank information even when that component is corrupted by sparse errors of large magnitude. This good feature is seen to lead to impressive results in problems such as video surveillance and face recognition [10], and singing-voice separation from monaural records [12].

In mathematical terms, the decomposition performed in RPCA can be expressed as follows

$$\mathbf{X} = \mathbf{L} + \mathbf{S}, \tag{3}$$

where $\mathbf{L}$ is a low-rank matrix and $\mathbf{S}$ is a sparse matrix, meaning to be mainly composed of null elements. One of the interesting aspects of RPCA concerns the existence of theoretical results assuring the uniqueness of the decomposition (3) — this is extremely useful in a signal separation context. Candès et al. [10] have shown that if the data is indeed composed of a low-rank and a sparse matrix terms, then, under mild assumptions, these two terms can be recovered by solving the following optimization problem:

$$\begin{aligned} \underset{\mathbf{L},\mathbf{S}}{\text{minimize}} \quad & ||\mathbf{L}||_* + \lambda ||\mathbf{S}||_1 \\ \text{subject to} \quad & \mathbf{L} + \mathbf{S} = \mathbf{X}, \end{aligned} \tag{4}$$

where $||\mathbf{L}||_*$ denotes the nuclear norm of $\mathbf{L}$, i.e., the sum of the singular values of that matrix, and $||\mathbf{S}||_1$ denotes the sum of the absolute values of all elements of $\mathbf{S}$. An interesting aspect here is that the problem (4) is convex, which opens the way for the application of very efficient optimization algorithms.

Our motivation to apply the RPCA to decompose seismic data comes from the fact that, very often, the seismogram is composed by low-rank structured parts (as in the case of horizontal reflections) and structured parts in which the rank is not necessarily low, e.g. nonlinear events such as diffractions.

In practice, seismic signals are corrupted with noise, which is usually modeled as a white Gaussian process and, as such, neither low rank nor sparse. A way to overcome that difficulty is to introduce an additional term $\mathbf{N}$ in the decomposition (3), so as to accommodate the noise term. Namely, we set

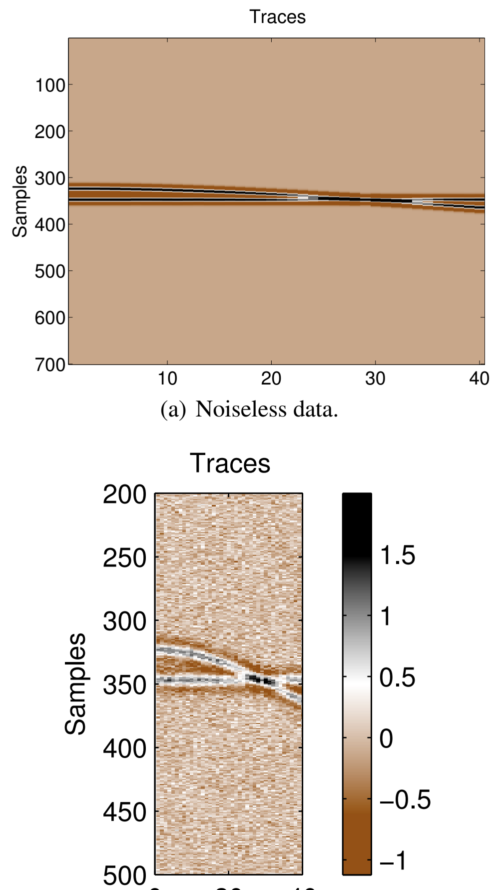
$$\mathbf{X} = \mathbf{L} + \mathbf{S} + \mathbf{N}. \quad (5)$$

The above formulation is referred to as noisy robust PCA.

In the literature, one can find several algorithms to implement the noisy PCA decomposition. Here, we will consider the Go Decomposition (GoDec), recently introduced in [13]. The GoDec algorithm considers the following optimization problem:

$$\begin{aligned} \underset{\mathbf{L},\mathbf{S}}{\text{minimize}} \quad & ||\mathbf{N}||_F^2 = ||\mathbf{X} - \mathbf{L} - \mathbf{S}||_F^2 \\ \text{subject to} \quad & rank(\mathbf{L}) \leq j \\ & card(\mathbf{S}) \leq k, \end{aligned} \quad (6)$$

where $||\mathbf{N}||_F$ denotes the Frobenius norm of $\mathbf{N}$. With respect to other noisy RPCA methods, such as the Bayesian approach proposed in [14], the GoDec algorithm provides a good solution in terms of computational effort, which is a crucial feature in seismic signal processing.



(a) Noiseless data.

(b) Zoomed version of the observed noisy data.

Fig. 1. Data considered in the first experiment.

3. RESULTS

To evaluate the RPCA decomposition in the context of seismic signals, we carry out numerical experiments in two situations that often arise in practice: separation of crossing events and separation of diffractions and reflections.

3.1. Separation of crossing events

We apply the three methods discussed before to separate the two close events depicted in Figure 1. The seismic data in this case is corrupted with additive white Gaussian noise, the signal-to-noise ratio (SNR) being 14dB. The number of traces is $m = 40$ and the number of samples is $n = 700$.

Figure 2 shows the first eigen-image obtained by the SVD method and by the SVD-ICA method, as well as the low-rank matrix $\mathbf{L}$ provided by the GoDec algorithm (we set $j = 1$ in this case). One can note that, while the SVD-ICA method was able to retrieve the horizontal event, the estimation provided by the ordinary SVD was corrupted by a strong artifact, which stems from the non-horizontal event. In the low-rank term of RPCA, we can also see an artifact created by the non-horizontal event, but with less amplitude than the one found using SVD.

Figure 3 shows the remaining terms obtained in each decomposition, that is, the remaining eigen-images (noise subspace) in SVD and SVD-ICA, as well as the sparse term $\mathbf{S}$ provided by the GoDec. It is worth noticing that, while in both SVD and SVD-ICA the retrieved data contain the non-horizontal event together with a relevant amount of noise, the sparse term of RPCA provided a much less noisy estimation of the non-horizontal event.

3.2. Separation of diffractions and reflections

We now consider the problem of separating events having a hyperbolic shape from horizontal events. More specifically, we consider in a zero-offset seismic configuration (namely coincident source-receiver pairs), the responses of five point diffractors, a broken horizontal reflection (which produces,

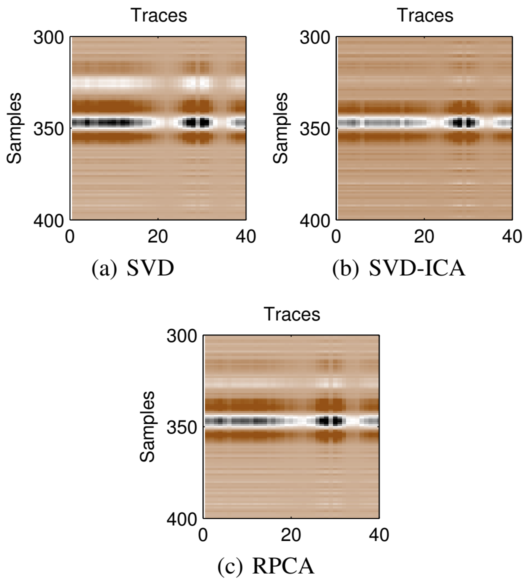


Fig. 2. First experiment: Low-rank subspaces.

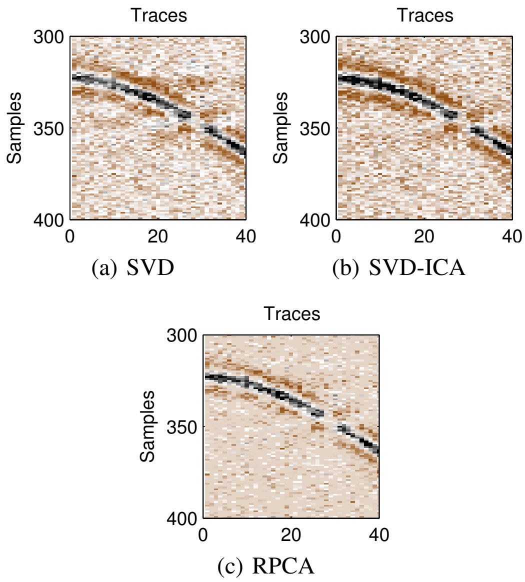


Fig. 3. First experiment: remaining eigen-images of SVD and SVD-ICA, and the sparse term provided by RPCA.

besides the reflection, also an edge diffraction) and a full horizontal reflector (which produces a horizontal reflection). One of the point diffractors lies on top of the horizontal reflector. The signals (events) considered in this experiment are presented in the seismogram of Figure 4, which contains $m = 300$ traces, each of them with $n = 300$ samples. The SNR is 20dB. One particular characteristic that makes this problem challenging is that the diffractions, which are of hyperbolic shape, have clearly smaller amplitudes than that of reflections. Moreover, the amplitude is not constant along the hyperbolas.

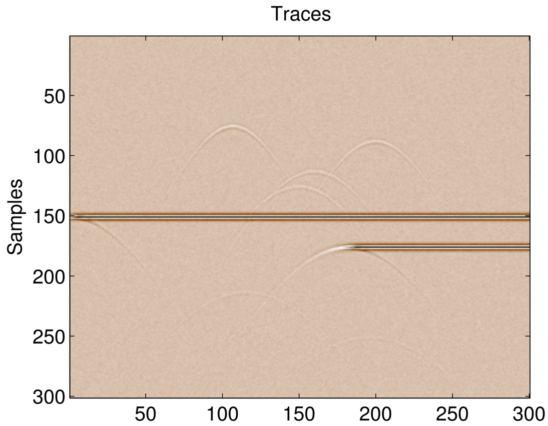


Fig. 4. Second experiment: noisy data.

Our results show that the SVD and the RPCA were able to perfectly retrieve the horizontal events in their rank-two subspaces. However, the estimation provided by the SVD-ICA presented some small horizontal artifacts (the figures are omitted here due to the lack of space).

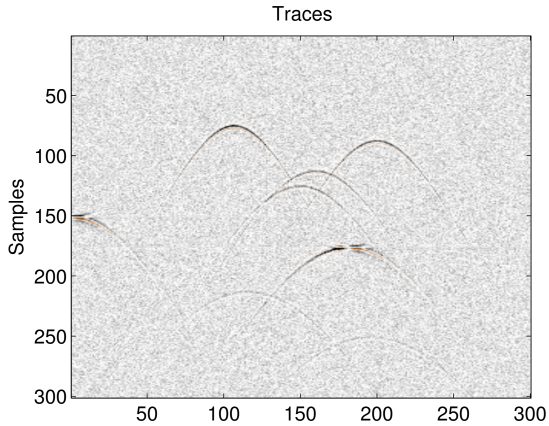
Figure 5 shows the noise subspace associated with the SVD and SVD-ICA methods and the sparse matrix obtained by RPCA. We note that SVD-ICA retrieved the non-horizontal events corrupted with noise and horizontal artifacts, which were mainly created by the horizontal events. In both SVD and RPCA, though, there was no horizontal artifacts. However, the estimation of the diffractions provided by RPCA are clearly less noisy than the ones obtained by SVD.

4. CONCLUSIONS

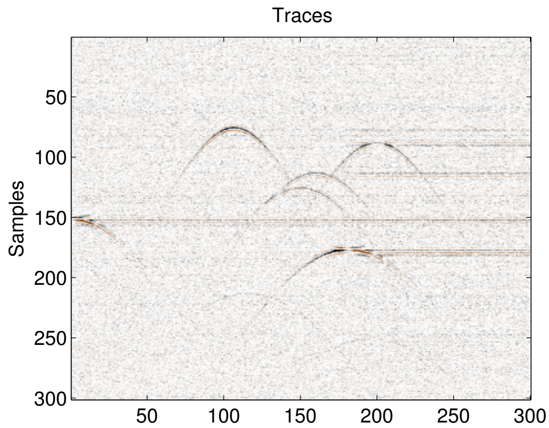
We investigated the application of an RPCA method and the GoDec algorithm, to separate different signals (events) present in seismic data. By means of numerical experiments, we observed that RPCA is a valuable tool in this context. RPCA performed better than SVD and SVD-ICA in the task of separating weak hyperbolic events from horizontal ones in noisy data. In the same way, in a second experiment related to crossing events, RPCA provided the best estimation of a non-horizontal event, as well as a better estimation of a horizontal event as compared to SVD, although SVD-ICA exhibited a better performance in this case. Besides the good performance of RPCA, these algorithms are usually more efficient in terms of computational complexity than SVD-ICA.

5. ACKNOWLEDGMENTS

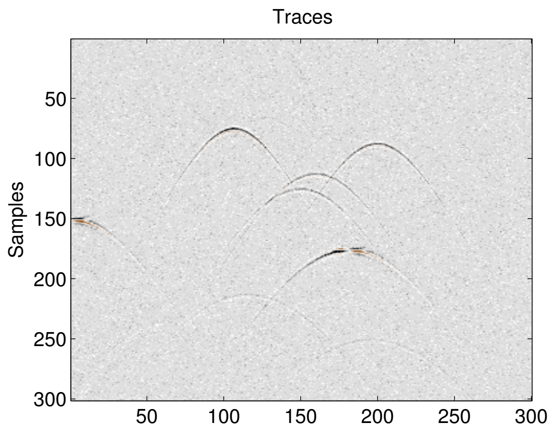
The authors would like to thank Prof. Jérôme Mars (Gipsa-lab) for providing us the SVD-ICA codes, and Tiago T. L. Barros and Jorge H. Faccipieri (UNICAMP) for proving us the data used in our experiments.



(a) SVD



(b) SVD-ICA



(c) RPCA

Fig. 5. Second experiment: remaining eigen-images of SVD and SVD-ICA, and the sparse term provided by RPCA.

6. REFERENCES

[1] F. Glangeaud and J.-L. Mari, *Wave separation*, Technip Editions, 1994.

[2] R. E. Sheriff and L. P. Geldart, *Exploratory seismology*, Cambridge University Press, 1995.

[3] O. Yilmaz, *Seismic data analysis: processing, inversion and interpretation of seismic data Volume 1*, SEG, second edition, 2001.

[4] S. L. M. Freire and T. J. Ulrych, "Application of singular value decomposition to vertical seismic profiling," *Geophysics*, vol. 53, pp. 778–785, 1988.

[5] R. L. Kirlin and W. J. Done, Eds., *Covariance analysis for seismic signal processing*, Society of Exploration Geophysicists, 1999.

[6] M. Yedlin, I. F. Jones, and B. B. Narod., "Application of the karhunen-love transform to diffraction separation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 35, no. 1, pp. 2–8, 1987.

[7] M. J. Porsani, M. G Silva, P. E. M. Melo, and B. Ursin, "SVD filtering applied to ground-roll attenuation," *Journal of Geophysics and Engineering*, vol. 7, pp. 284–289, 2010.

[8] V. D. Vrabie, J. I.Mars, and J.-L. Lacoume, "Modified singular value decomposition by means of independent component analysis," *Signal Processing*, vol. 84, pp. 645–652, 2004.

[9] M. Bekara and M. Van der Baan, "Local SVD/ICA for signal enhancement of pre-stack seismic data," in *68th EAGE Conference & Exhibition*, 2006.

[10] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of the ACM*, vol. 58, pp. 1–37, 2011.

[11] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, "Sparse and low-rank matrix decompositions," in *Proc. 47th Ann. Allerton Conf. Communication, Control, and Computing Allerton*, 2009, pp. 962–967.

[12] P.-S. Huang, S. D. Chen, P. Smaragdis, and M. Hasegawa-Johnson, "Singing-voice separation from monaural recordings using robust principal component analysis," in *Proc. of the IEEE ICASSP*, 2012.

[13] T. Zhou and D. Tao, "GoDec: Randomized low-rank & sparse matrix decomposition in noisy case," in *Proceeding of the International Conference on Machine Learning (ICML)*, 2011.

[14] X. Ding, L. He, and L. Carin, "Bayesian robust principal component analysis," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3419–3430, 2011.