Separation of Nonlinear Image Mixtures by Denoising Source Separation

Mariana S.C. Almeida[1], Harri Valpola[2], and Jaakko Särelä[3]

[1] Instituto de Telecomunicações
Av. Rovisco Pais - 1, 1049-001 Lisboa, Portugal
mariana.almeida@lx.it.pt
[2] Laboratory of Computational Engineering
Helsinki University of Technology
P.O.Box 9203, FI-02015 HUT, Espoo, FINLAND
harri.valpola@tkk.fi http://www.lce.hut.fi/~harri/
[3] Adaptive Informatics Research Centre
Laboratory of Computer and Information Science
Helsinki University of Technology
P.O.Box 5400, FI-02015 HUT, Espoo, FINLAND
jaakko.sarela@tkk.fi http://www.cis.hut.fi/jaakkos/

Abstract. The denoising source separation framework is extended to nonlinear separation of image mixtures. MLP networks are used to model the nonlinear unmixing mapping. Learning is guided by a denoising function which uses prior knowledge about the sparsity of the edges in images. The main benefit of the method is that it is simple and computationally efficient. Separation results on a real-world image mixture proved to be comparable to those achieved with MISEP.

1 Introduction

Nonlinear source separation refers to separation of sources from their nonlinear mixtures (for reviews, see [1, 2]). It is much harder than linear source separation because the problem is highly ill-posed. In practice, some type of regularisation is needed. It is, for instance, possible to require that the nonlinear mixing or unmixing mapping is smooth or belongs to a restricted class of nonlinear functions. Alternatively, it is possible to impose restrictions on the extracted sources. In any case, it is important to reduce the number of available degrees of freedom.

Denoising source separation (DSS, [3]) has been introduced as a framework for source separation algorithms, where separation is constructed around denoising procedures. DSS algorithms can range from almost blind to highly tuned separation with detailed prior knowledge. The framework has already been successful in several applications such as biomedical signal processing [3], CDMA signal recovery [4] and climatology [5].

So far, DSS has been applied to linear separation only, but in this paper we show that nonlinear separation is possible, too. In the DSS framework, it is easy to use detailed prior information. This means that separation becomes possible even if the nonlinear mappings are not carefully regularised. This is a significant benefit because this translates to significant savings in computational complexity, particularly in large problems with many sources and mixtures.

The rest of the paper is organised as follows. The nonlinear DSS method is introduced in Sec. 2. In many respects the separation procedure is exactly like linear separation except that decorrelation and scaling of the sources need to be embedded into the denoising whereas in linear separation this can be implemented by orthogonalising the mixing matrix.

In the rest of the paper, we demonstrate the nonlinear DSS in a real-world nonlinear separation problem introduced by [6]. The problem is to separate two images which have been printed on opposite sides of a paper. Due to partial transparency of the paper, both images are visible from each side, corresponding to two nonlinear mixtures of the source images. In the DSS framework, separation is built around a denoising procedure which can be tailored to the problem at hand. A suitable denoising function which utilises the sparsity of image edges is introduced in Sec. 3 and separation results are reported in Sec. 4.

Finally, in Sec. 5, we discuss the relation of the proposed nonlinear DSS framework with other nonlinear separation methods and also discuss possible future research directions.

2 Nonlinear DSS method

In DSS, separation consists of the following steps:

1. estimation of the current sources using current mapping,
2. denoising of the sources and
3. adaptation of the mapping to match the denoised sources.

Note that the procedure bears resemblance to the EM algorithm: the first two steps correspond roughly to the E-step and the last step to the M-step. The main difference is that the EM algorithm is a generative approach where the mixing mapping is estimated. With generative models assuming uncorrelated sources, the sources will automatically become approximately uncorrelated due to the so-called explaining-away phenomenon. This needs to be emulated in DSS using some type of competition mechanism (see, e.g., [7] for discussion about emulating explaining away by lateral inhibition).

In linear separation, decorrelation and scaling can be realised by prewhitening the data and orthogonalising and scaling the projection vectors in the last step. In nonlinear DSS, this option is not available as there is, in general, no easy way to make sure that the outputs of a nonlinear mapping are orthogonal and suitably scaled. Instead, the decorrelation and scaling must be embedded in the denoising step. Besides this, the basic principle in nonlinear DSS is exactly the same as in linear DSS.

The method that we have used for nonlinear separation is illustrated in Fig. 1, for the case of separation of a two-source mixture.

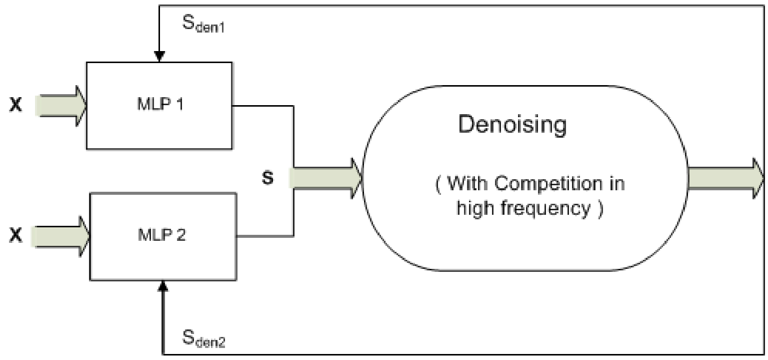


Fig. 1. Schematic representation of the nonlinear DSS method.

The principle of operation is as follows. The mixture vector $\boldsymbol{X}$ is fed into two multilayer perceptrons (MLP1 and MLP2), which yield the current estimates of the sources $\boldsymbol{S}$ as their outputs (step 1). The estimates are then denoised (step 2). Finally the MLP networks are adapted to the denoised source estimates $\boldsymbol{S}_{den1}$ and $\boldsymbol{S}_{den2}$ (step 3). Provided that the denoising step is well chosen, iterating these three steps will result in the separation of the mixed sources.

3 Denoising for image separation

The crucial element in DSS, with linear or nonlinear mapping, is the choice of the denoising function. A lengthy discussion of denoising functions and their properties can be found in [3]. In brief, removing noise helps identify the signal subspace and removing the interference from other sources promotes separation. In this paper, we focus on the case where there is an equal number of sources and mixtures. Therefore, the most important thing is to reduce the interference from other sources.

3.1 Mixing process

The image mixtures that were studied correspond to a well known practical situation: when an image of a paper document is acquired, the back page sometimes shows through. The paper that was used was onion skin, which leads to a strong mixture, which is significantly nonlinear. This separation problem has first been introduced by [6]. We show the effectiveness of the proposed DSS method using the first, the second and the fifth mixtures from that paper. The source images, which were printed on the onion skin paper, are shown in Fig. 3a. The acquired images (mixtures) are shown in Fig. 3b. For more detailed description of the data aquisition, see [6].

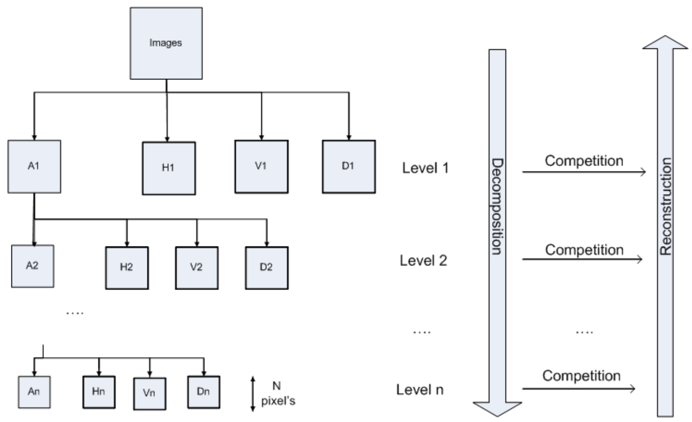


Fig. 2. Diagram of the wavelet-based denoising operation.

3.2 Edge denoising

Looking at the last two pair of mixtures in Fig. 3b, it is evident that despite strong nonlinear mixtures, a human being can easily separate the images, i.e., can tell which features or objects belong in which image, even without knowing the original images. What features could be used for separating, i.e., denoising, the images?

A characteristic feature of most natural images is the sparsity of edges. When an edge is found in the same place in both mixtures, it probably originates from only one of the source images. Decision about which source the edge belongs to can be based on the relative strength of the edges in the mixtures. Hence, we suggest the following denoising scheme:

1. Represent each of the source estimates by their edges.
2. Induce a competition between the edges in different images in such a way that stronger edges tend to eliminate weaker ones.

Note that the edge features in different natural images are usually almost independent, which is not necessarily true for low-frequency features. Consider for instance natural images of faces.

Edge detection in images A crude approach for edge detection that already leads to somewhat acceptable results, is to use simple high-pass filtering to extract the edges. Another, more advanced possibility is to use wavelet analysis. We decided to use a wavelet family that forms a spatio-frequency representation of an image separately with horizontal, vertical and diagonal components (H, V and D). The representation results in a hierarchy of increasing frequencies. A schematic illustration of the wavelet transform that was used, is depicted on the left side of Fig. 2.

Competition between the edges Once the edges of both source estimates have been extracted, one should decide which edge belongs to which image. On average, edges of the foreground image appear stronger on the foreground mixture. Hence strong edges on the foreground images should be privileged for the foreground source estimate. This has been achieved by using a soft winner-take-all operation, which assigned most of the energy to the stronger component. The competition was induced in each level of the wavelet transform, except for the first one that represents the slowest frequencies (see the right side of Fig. 2).

Additionally, the artificial nature of the first pair of mixtures (Fig. 3c, top-row), was taken into account. Since one of the source images contained only vertical and the other one only horizontal edges, the horizontal (H) components were set to zero in one of the images, prior to reconstruction, and the vertical (V) ones in the other image.

4 Results

The multilayer perceptrons that were used had a hidden layer with five sigmoidal units. They also had direct ”shortcut” connections between inputs and output, and their output units were linear. With this structure they were able to implement linear operations. These perceptrons were initialised to perform an approximate linear whitening (also called sphering) of the mixture data, subject to the restriction of being symmetrical (processing the two input components equally). Training was performed with the adaptive step sizes speedup method [8]. Fifty training epochs were performed, within each iteration of the global nonlinear DSS procedure. Two-level description was used in the wavelet decomposition.

Figure 3c shows the results obtained after 10 iterations of the nonlinear DSS. For comparison, the results obtained with the MISEP technique of nonlinear ICA can be consulted in [6].

For an objective quality assessment, the four quality measures defined in [6] were computed. Q_1 is simply the signal-to-noise ratio (SNR). Q_2 is also an SNR measure, but with a correction for possible nonlinear distortions of the intensity scale of the separated images. Q_3 is the mutual information between each separated component and the corresponding source. Finally, Q_4 is the mutual information between each separated component and the opposite source. For Q_1, Q_2 and Q_3, higher values are better, while for Q_4 lower values are better. See [6] for more details. Table 1 shows the results, together with the results obtained with the MISEP method, for comparison (the latter were obtained from [6]).

In the first pair, nonlinear DSS performed better than MISEP. This is probably due to the specific denoising operation that was used, which is very well suited to this pair of sources. In the second image pair, nonlinear DSS and MISEP performed approximately equally on the right-hand image, and MISEP performed better on the left-hand image. In the third pair, nonlinear DSS performed globally better. This pair of sources is not independent (see [6]), and

Image pair	Quality measure	Nonlinear DSS		MISEP	
		source 1	source 2	source 1	source 2
1	Q_1 (dB)	**14.6**	**14.1**	13.8	13.1
	Q_2 (dB)	**15.3**	**14.7**	14.7	14.2
	Q_3 (bit)	**2.57**	**2.50**	2.45	2.39
	Q_4 (bit)	0.29	0.27	**0.23**	**0.26**
2	Q_1 (dB)	6.4	13.6	**9.3**	**13.9**
	Q_2 (dB)	9.5	**15.1**	**11.0**	15.0
	Q_3 (bit)	1.62	1.93	**1.83**	**1.95**
	Q_4 (bit)	0.44	**0.39**	**0.24**	0.40
3	Q_1 (dB)	13.5	**9.2**	**14.2**	6.4
	Q_2 (dB)	**15.5**	**9.9**	15.3	7.8
	Q_3 (bit)	**2.23**	**1.62**	2.19	1.29
	Q_4 (bit)	0.74	0.56	**0.56**	**0.49**

Table 1. Quality measures. For each pair (Nonlinear DSS and MISEP, for the same source), the best result is shown in bold. For Q_1, Q_2 and Q_3 higher results are better, while for Q_4 lower results are better.

therefore nonlinear DSS is probably more suited to handle it than MISEP, which is an independence based method.

5 Discussion

In this paper, we reported the first results about nonlinear separation with DSS. As the results show, separation was relatively successful but still far from perfect. For instance, from the extracted image pair in the middle of Fig. 3c, it is evident that the contrast on the image on the left depends on the intensity of the image on the right (lighter on the right implies better contrast on the left). Furthermore, we had to resort to early stopping in the separation of the mixtures of natural images. Such problems could be avoided by improving the denoising function, for example by introducing a local normalisation of image contrast, or by using more prior information about the mixing process to restrict the parametric form of the unmixing mapping.

Of the existing nonlinear separation techniques, MISEP is similar to the one proposed here in that it, too, estimates a separating MLP network. The main advantage over MISEP is that the learning procedure is simpler and computationally more efficient. In MISEP, the Jacobian matrix of the nonlinear mapping needs to be computed for every sample, inverted and then propagated back through the MLP network. For two-dimensional case this is not of importance and MISEP was actually faster in these simulations. However, it means that MISEP cannot be extended to problems with a large number of sources.

Slow-feature analysis (SFA, [9]) resembles nonlinear DSS in its use of denoising for guiding separation. In SFA, the denoising is implemented by low-pass filtering (see [3] for details) and therefore assumes that the sources have slowly changing temporal or spatial structure. In DSS, the denoising can be more general and tuned to the problem at hand, such as the presented edge-based denoising for separating images. Interestingly, SFA has been shown to be applicable to very large problems when the set of nonlinearities is fixed and only a linear mapping is learned [9]. It should therefore be possible to apply nonlinear DSS in very large problems using a similar restricted mapping.

6 Conclusion

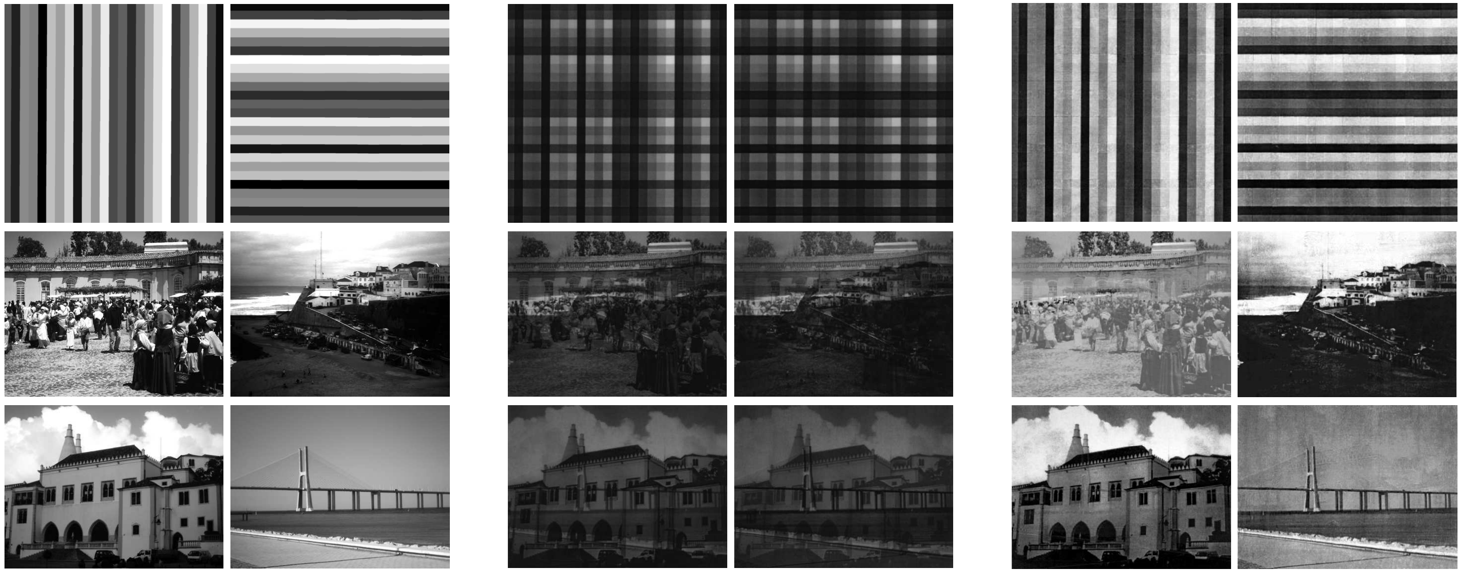
We have presented a nonlinear separation method based on the denoising source separation framework. The method uses a competition-based denoising stage which performs a partial separation of the sources, the partially separated components being used to iteratively re-train a set of nonlinear separators. The method was applied to real-life nonlinear mixtures of images, and proved to be competitive with ICA-based nonlinear separation.

7 Acknowledgements

We are grateful to L. B. Almeida for helpful comments on this manuscript and for providing the images and related information.

References

1. Jutten, C., Karhunen, J.: Advances in blind source separation (BSS) and independent component analysis (ICA) for nonlinear mixtures. International Journal of Neural Systems **14**(5) (2004) 267–292
2. Almeida, L.: Nonlinear Source Separation. Morgan and Claypool (2005) in press.
3. Särelä, J., Valpola, H.: Denoising source separation. Journal of Machine Learning Research **6** (2005) 233–272
4. Raju, K., Särelä, J.: A denoising source separation based approach to interference cancellation for DS-CDMA array systems. In: Proceedings of the 38th Asilomar Conference on Signals, Systems, and Computers, Pacific grove, CA, USA (2004) 1111 – 1114
5. Ilin, A., Valpola, H., Oja., E.: Semiblind source separation of climate data detects El Niño as the component with the highest interannual variability. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN 2005), Montréal, Québec, Canada (2005) 1722 – 1727
6. Almeida, L.: Separating a real-life nonlinear image mixture. Journal of Machine Learning Research **6** (2005) 1199–1229
7. Ghahramani, Z., Hinton, G.E.: Hierarchical non-linear factor analysis and topographic maps. In Jordan, M.I., Kearns, M.J., Solla, S.A., eds.: Advances in Neural Information Processing Systems 10. The MIT Press, Cambridge, MA, USA (1998) 486–492
8. Almeida, L.: Multilayer perceptrons. In Fiesler, E., Beale, R., eds.: Handbook of Neural Computation, Institute of Physics, Oxford University Press (1997)
9. Wiskott, L., Sejnowski, T.: Slow feature analysis: Unsupervised learning of invariances. Neural Computation **14** (2002) 715 – 770



(a) (b) (c)

Fig. 3. a) Source images b) mixture (acquired) images c) separation results.