14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

CONTRIBUTIONS TO ICA OF NATURAL IMAGES

Rubén Martín-Clemente, Susana Hornillo-Mellado

Dpto. de Teoría de la Señal y Comunicaciones
Escuela Superior de Ingenieros, University of Seville, Seville, Spain
phone: + 34 954 487 335, fax: + 34 954 487 341, emails: ruben@us.es, susanah@us.es
webs: www.personal.us.es/ruben, www.personal.us.es/susanah

ABSTRACT

In this paper we analyze the results provided by the popular algorithm FastICA when it is applied to natural images, using the kurtosis as non-linearity. In this case show that the so-called ICA filters can be expressed in terms of the eigenvectors associated to the smallest eigenvalues of the data correlation matrix, meaning that these filters are all high-pass. From this property emerges the sparse distribution of the independent components. On the other hand, the use of the kurtosis as contrast function causes the appearance of "spikes" in the independent components that make that the ICA bases are very similar to patches of the images analyzed. Some experiments are included to illustrate the results.

1. INTRODUCTION

Independent Component Analysis (ICA) is a technique for analyzing multivariate data that has received a great interest in the last years [6, 7]. Let the observed multidimensional data be represented as a matrix $\mathbf{X}$ with N rows and T columns, where N is the number of variables and T is the number of observations recorded on each variable. The goal of ICA is to calculate the square matrix $\mathbf{B}$ that linearly transforms $\mathbf{X}$ into a matrix:

$$\mathbf{Y} = \mathbf{B}\mathbf{X} \tag{1}$$

The goal of ICA is to calculate the $N \times N$ matrix $\mathbf{B}$ that linearly transforms $\mathbf{X}$ into $\mathbf{Y}$. Different approaches have been proposed to characterize this model (see, for example, [1, 2, 4, 15]).

Bell and Sejnowski showed in [3] that when the observed data are patches of natural images, ICA provides a sparse code in terms of a set of *ICA bases* that present oriented and localized lines or "edges". In particular, they proposed an unsupervised learning algorithm based on information maximization ("infomax" network) to perform the independent component analysis [2] and used a training set of patches taken from different natural images. In this case, each column of matrix $\mathbf{X}$ represents a (square) patch of a natural image and can be expressed as follows:

$$\mathbf{x}_{:k} = \mathbf{a}_{:1} y_{1k} + \mathbf{a}_{:2} y_{2k} + \cdots + \mathbf{a}_{:N} y_{Nk} \tag{2}$$

Using a `Matlab`-like notation, $\mathbf{x}_{:k}$ is the kth column of $\mathbf{X}$, y_{jk} is the (j,k)th element of matrix $\mathbf{Y}$ and $\mathbf{a}_{:j}$ is the jth column of matrix $\mathbf{A}$, with $j = 1,2,...,N$ and $k = 1,2,...,T$. Bell and Sejnowski showed that the ICA bases $\mathbf{a}_{:j}$ *looked like edges*, while the distribution of the weighting coefficients, y_{jk}, was *sparse*. The connection between this result and the behaviour of some neurons of the primary visual cortex (V1) was argued by Bell and Sejnowski [3] (see also the previous works [8, 18] and the references therein).

Similar results were obtained with the FastICA algorithm [15], as shown in [9, 14, 17], under certain operational conditions (in particular, using the nonlinearity $g_1(y) = \tanh(y)$ – see [14] for details).

Our aim is to analyze both the independent components and the ICA bases of natural images obtained with the FastICA algorithm, but now using its original version based on the maximization of the *kurtosis* (i.e., using the nonlinearity $g_3(y) = y^3$) [14, 15]. As Hyvärinen argued in [14], the results obtained with these two approaches are basically the same, but the kurtosis has two advantages compared to the hyperbolic tangent. First, from a theoretical point of view, it leads to a mathematically more tractable analysis. Second, from a practical point of view, the number of flops required are much lower (about nine times lower)[1].

The analysis presented in this paper shows that the so-called "ICA filters" (rows of matrix $\mathbf{B}$ – see (1)) can be decomposed as a *linear combination of the eigenvectors associated to the smallest eigenvalues of the data correlation matrix*, meaning that these filters are all *high-pass*. This property leads to a *sparse* distribution of the independent components, whose non-negligible elements correspond to patches that contain lines or *edges* (i.e., high frequency content). Besides, the use of the kurtosis as contrast function causes the appearance of *spikes*, making the ICA bases very similar to patches of the natural images analyzed, also corresponding to *edges*.

The "high-pass nature" of the ICA filters was already observed in [14, 20], but was related to the "high-pass appearance" of the ICA bases. The innovation that we introduce here is that *the ICA bases "look like edges" because the ICA filters are high-pass, not the inverse*.

The paper is structured as follows. We first introduce the usual preprocessing in Section 2. In Section 3 we derive a closed-form expression in which the first ICA filter is expressed as the linear combination of the eigenvectors associated to the smallest eigenvalues of the data correlation matrix. The implications of this result in the case of natural images are discussed in Section 4, in which we argue the convenience of interpreting ICA as a two-dimensional filtering. The extension to the rest of the independent components is done in Section 5. Experiments and discussion is given in Section 6. Final conclusions are provided in Section 7.

[1] The kurtosis is very sensitive to outliers, something that could be a problem in many applications [14, 11]. This is not our case, because our aim is to provide a sparse representation of natural images (to be used in compression applications, for example), and the kurtosis is as valid as any other method.

2. PREPROCESSING

There are two quite standard preprocessing steps in ICA. First, the mean of the data is usually subtracted to center the data on the origin. Using our notation:

$$\frac{1}{T}\sum_{n=1}^{T} x_{kn} = 0 \quad \forall k \tag{3}$$

This does not alter the ICA model (1) except that now the independent components are also zero-mean.

Secondly, the data are usually transformed into uncorrelated variables by means of a *whitening* matrix, $\mathbf{W}$:

$$\bar{\mathbf{X}} = \mathbf{W}\,\mathbf{X} = \mathbf{D}^{-1/2}\,\mathbf{V}^t\,\mathbf{X} \tag{4}$$

where the superscript t means "transpose". Here, $\mathbf{D}^{-1/2} = \mathrm{diag}(\lambda_1^{-1/2},...,\lambda_N^{-1/2})$, being $\lambda_1 \geq \ldots \geq \lambda_N$ the eigenvalues of the sample data correlation matrix, $\mathbf{R}_x = \frac{1}{T}\mathbf{X}\mathbf{X}^t$; $\mathbf{V} = (\mathbf{v}_1|\ldots|\mathbf{v}_N)$ is the matrix containing, by columns, the corresponding eigenvectors ($\mathbf{R}_x\,\mathbf{v}_j = \lambda_j\,\mathbf{v}_j$). The ICA model (1) can be now written as

$$\mathbf{Y} = \bar{\mathbf{B}}\,\bar{\mathbf{X}} \tag{5}$$

where $\bar{\mathbf{B}} = \mathbf{B}\,\mathbf{W}^{-1}$. It is straightforward to show that matrix $\bar{\mathbf{B}}$ is orthogonal, i.e., $\bar{\mathbf{B}}\bar{\mathbf{B}}^t = \mathbf{I}$ ($\mathbf{I}$ is the identity matrix).

3. THE ICA FILTERS OF FASTICA

Let $\mathbf{y}_{1:} = [y_{11},\ldots,y_{1T}]$ be the first independent component and define $\mathbf{y}^3_{1:} = [y^3_{11},\ldots,y^3_{1T}]$ as the vector which contains the entries of $\mathbf{y}_{1:}$ raised to the third power. Similarly, let $\bar{\mathbf{b}}_{1:}$ be the first row of matrix $\mathbf{B}$. With this notation, the basic `FastICA` iteration can be expressed as follows [14]:

1. $\mathbf{y}_{1:} \longleftarrow \bar{\mathbf{b}}_{1:}\,\bar{\mathbf{X}}$
2. $\bar{\mathbf{b}}_{1:} \longleftarrow \frac{1}{T}\mathbf{y}^3_{1:}\,\bar{\mathbf{X}}^t - 3\,\bar{\mathbf{b}}_{1:}$
3. $\bar{\mathbf{b}}_{1:} \longleftarrow \bar{\mathbf{b}}_{1:}/\|\bar{\mathbf{b}}_{1:}\|$

These three steps are repeated until convergence ($\longleftarrow$ means "update"). The update logically stops when

$$\bar{\mathbf{b}}_{1:} \propto \frac{1}{T}\mathbf{y}^3_{1:}\,\bar{\mathbf{X}}^t - 3\bar{\mathbf{b}}_{1:} \tag{6}$$

where "$\propto$" means "proportional to". Post-multiplying both sides with the whitening matrix $\mathbf{W}$, and using that $\bar{\mathbf{b}}_{1:} = \mathbf{b}_{1:}\,\mathbf{W}$, we get

$$\bar{\mathbf{b}}_{1:}\,\mathbf{W} \propto \mathbf{y}^3_{1:}\,\bar{\mathbf{X}}^t\,\mathbf{W} \quad \Rightarrow \quad \mathbf{b}_{1:} \propto \mathbf{y}^3_{1:}\,\bar{\mathbf{X}}^t\,\mathbf{W} \tag{7}$$

Using the definition of $\mathbf{W}$ (see (4)) in (7) we get:

$$\mathbf{b}_{1:} \propto \mathbf{y}^3_{1:}\,\bar{\mathbf{X}}^t\,\mathbf{D}^{-1/2}\,\mathbf{V}^t \tag{8}$$

from which we get, after some algebra, the essential relation:

$$\mathbf{b}^t_{1:} \propto \sum_{n=1}^{N} \gamma_n \mathbf{v}_n \tag{9}$$

where $\mathbf{v}_n$ is the nth eigenvector of the data correlation matrix $\mathbf{R}_x$ and

$$\gamma_n = \frac{\sum_{k=1}^{T} \bar{x}_{nk}\,y^3_{1k}}{\sqrt{\lambda_n}} \overset{(a)}{\propto} \frac{\bar{b}_{1n}}{\sqrt{\lambda_n}} \tag{10}$$

Here, $\bar{b}_{1n}$ is the $(1,n)$th entry of $\bar{\mathbf{B}}$ and (a) follows from (6).

4. APPLICATION TO NATURAL IMAGES

Equation (9) indicates that the first ICA filter can be written as a linear combination of the eigenvectors of the correlation data matrix. This is not surprising *a priori*, since these eigenvectors always form an orthogonal basis of the space. When the data analyzed are natural images, it is well known that due to the great correlation among neighboring pixels, most of the eigenvalues are very close to zero (this is the basis of image compression [16]). The opposite happens with the corresponding factors $\frac{1}{\sqrt{\lambda_n}}$. Hence, considering that the norm of $\bar{b}_{1n}$ is bounded to one, we can approximate the first ICA filter by the weighted sum of the eigenvectors associated to the *smallest eigenvalues* of the data correlation matrix:

$$\mathbf{b}^t_{1:} \propto \sum_{n=m}^{N} \gamma_n \mathbf{v}_n, \quad 1 << m \leq N \tag{11}$$

For natural images, the eigenvectors associated to the smallest eigenvalues contain the higher-frequency information (edges, textures, etc.). Consequently, the first ICA filter will have high-pass characteristics.

4.1 ICA as a 2-D "filtering-sampling" process

Taking into account that each column of this matrix $\mathbf{X}$ contains the pixels of a patch of an image, it is clear to see that (1) implicitly represents a two-dimensional filtering. In this section we analyze this question and obtain an explicit expression of the first independent component as the sampling of the result of a 2-D filtering of the original image (the extension to the rest of independent components will be done in the following section). This alternative interpretation of ICA seems to be an unnecessary complication, but it is of great relevance for our analysis.

For convenience, we can also represent the kth image patch as a 2-D sequence $x_k(n_1,n_2)$, with $n_1,n_2 = 0,...,\sqrt{N}-1$ (in other words, $x_k(n_1,n_2)$ is the mapping of the $N\times 1$ vector $\mathbf{x}_{:k}$ to a $\sqrt{N}\times\sqrt{N}$ matrix). Similarly, $b_1(n_1,n_2)$ represents the first ICA filter, so that

$$y_{1k} = \sum_{i=1}^{N} b_{1i}\,x_{ik} = \sum_{n_1=0}^{\sqrt{N}-1}\sum_{n_2=0}^{\sqrt{N}-1} b_1(n_1,n_2)\,x_k(n_1,n_2) \tag{12}$$

The most-right part of this identity clearly reminds *a convolution*. To state it more properly, let us define the 2-D sequence $b_1^R(n_1,n_2)$ as $b_1(n_1,n_2)$ rotated 180° counterclockwise:

$$b_1^R(n_1,n_2) = b_1(\sqrt{N}-1-n_1,\sqrt{N}-1-n_2) \tag{13}$$

After some cumbersome algebra one finds that

$$y_{1k} = z(\sqrt{N}-1,\sqrt{N}-1) \tag{14}$$

where $z(n_1,n_2)$ is the 2-D convolution between the patch $x_k(n_1,n_2)$ and $b_1^R(n_1,n_2)$.

In other words, each element of $\mathbf{y}_{1:}$ is the filtering of an image patch with the corresponding ICA filter rotated 180° counterclockwise, followed by the sampling of the filter output at $n_1 = \sqrt{N}-1, n_2 = \sqrt{N}-1$. It is straightforward to show that an equivalent result will be obtained by filtering the whole image (rather than filtering isolated patches).

Finally, observe that the magnitude responses of the filters $b_1^R(n_1,n_2)$ and $b_1(n_1,n_2)$ are the same because the rotation only affects their phase response. In particular, if $b_1(n_1,n_2)$ is a high-pass filter then $b_1^R(n_1,n_2)$ will be also high-pass. In conclusion: *ICA is equivalent to a high-pass filtering and a sampling of the image*.

5. EXTENSION TO SEVERAL INDEPENDENT COMPONENTS

In practice, the `FastICA` algorithm is executed as many times as the number of desired independent components [15]. In the jth iteration it is imposed that the corresponding $\mathbf{b}_{j:}$ has to be orthogonal to the $\mathbf{b}_{k:}$, $k = 1,...,j-1$ previously obtained (deflationary orthogonalization). This method has the drawback that estimation errors in the first vector are accumulated in the subsequent ones by the orthogonalization [14]. The implications of this problem depends on the particular application. In our case, as we will see in the following section, the last independent components are not so sparsely distributed as the others.

6. EXPERIMENTS AND DISCUSSION

In this section we show and discuss the results obtained by applying the FastICA algorithm to natural images. We present two experiments: in the first one we analyze the results obtained considering the patches of only one natural grey-scale image; in the second experiment a great ensemble of natural grey-scale images is considered. In both cases we have selected the approach by deflation and the kurtosis as nonlinearity.

6.1 Results with only one natural image

Consider the natural, grey-scale image shown in Fig. 1. We divide it into 12×12 blocks or patches to compose the data matrix $\mathbf{X}$ ($\mathbf{X}$ is of size $144 \times 1\,764$). Next we run `FastICA`.

Figure 1: The "Barb" image (504×504) (available in [22]).

To analyze the frequency content of the ICA filters, we map the rows of $\mathbf{B}$ into 12×12 matrices and represent the magnitudes of their 2-D Fourier Transform. In Fig. 2 we show these magnitudes. As predicted, all of them are high-pass.

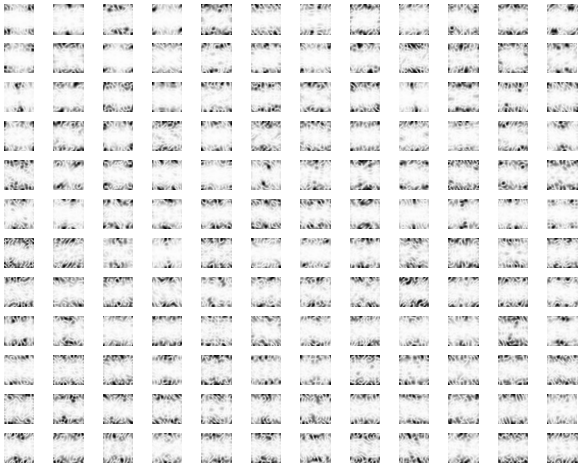

Figure 2: Amplitude spectra of the 12×12 ICA filters corresponding to the "Barb" image. The darker grey values refers to larger amplitudes; zero spatial frequency is at the centre of each patch.

Fig. 3 represents some of the independent components (i.e., the rows of matrix $\mathbf{Y}$) as well as their histograms (we can not represent all the independent components due to the lack of space). It is clear to see that the distribution of these independent components is sparse (most of the elements close to zero), which is in agreement with the results obtained by other authors [12, 14]. As we said in the previous Section, the last independent component is "not so sparse", due to the accumulated errors.

This sparse distribution of the independent components can be explained as follows. Our previous derivation indicates that ICA performs a high-pass filtering of the image data. Carrying out a high-pass filtering of a natural image, only the edges are enhanced whereas most of the image is attenuated or even eliminated[2]. So, when we sample the filter output most samples of the filtered image are expected to be small: only those samples at the edges will take significative values and the resulting independent component will be, consequently, "sparse". It also means that the spikes of the independent components are not randomly placed: they correspond to "edges" of the image.

Furthermore, most of the independent components have *one* very large element or *spike*. This is expected because *such a solution maximizes the kurtosis*, and we can not forget that this is the objective of the FastICA algorithm[3] [10, 14]. Recall the decomposition of each image patch $\mathbf{x}_{:k}$ in terms of the ICA bases represented by (2), re-written here for convenience:

$$\mathbf{x}_{:k} = \mathbf{a}_{:1}\, y_{1k} + \mathbf{a}_{:2}\, y_{2k} + \cdots + \mathbf{a}_{:N}\, y_{Nk} \qquad (15)$$

Consider that the jth independent component, $\mathbf{y}_{j:}$, has a spike in its mth sample, i.e., at y_{jm}. This position will be different from one independent component to another, due to

[2] Consequently, the filtered image will have a very sparse distribution, characterized by a sharp histogram centered around zero

[3] It is easy to show that the unity-variance signal with maximum kurtosis is the signal in which all its elements vanish excepting one.

the orthogonal relation among the ICA filters, meaning that:

$$\mathbf{x}_{:m} \simeq y_{jm}\,\mathbf{a}_{:j} \tag{16}$$

That is, each ICA basis, $\mathbf{a}_{:j}$ is approximately proportional to an image patch $\mathbf{x}_{:m}$ that corresponds to an "edge". To illustrate this point, in Fig. 4 we show the ICA bases corresponding to the "Barb" image. Almost all of them are like patches of the image corresponding to the edges, due to the dominant element present in each independent component.

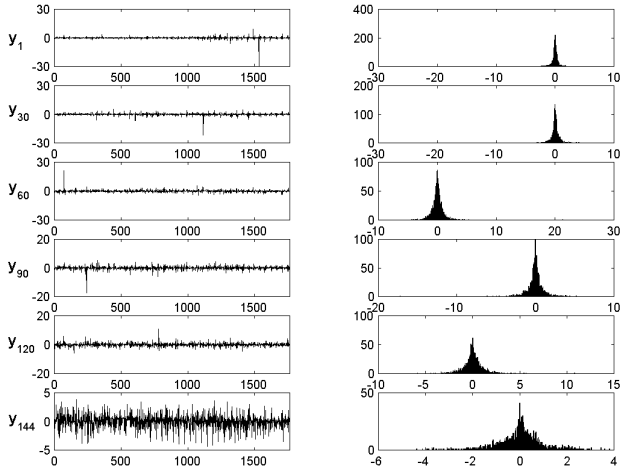

Figure 3: Some of the independent components corresponding to the "Barb" image. The scales have been adjusted to a better visualization.

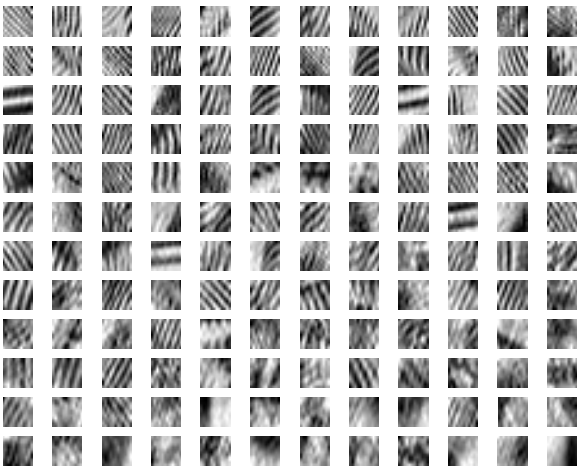

Figure 4: "ICA bases" (12×12) corresponding to the "Barb" image.

6.2 Results with an ensemble of natural images

For this experiment we have chosen a set of images from the database available in [21], to obtain a total of 24000 12×12 patches. These images are characterized by the presence of trees, bushes, leaves, etc. In Fig. 5, Fig. 6 and Fig. 7 we show the ICA filters, some of the independent components and the ICA bases, respectively. The conclusions related to the previous experiment still hold in this case. Regarding the ICA bases, they are very different from the ones obtained in the previous experiments

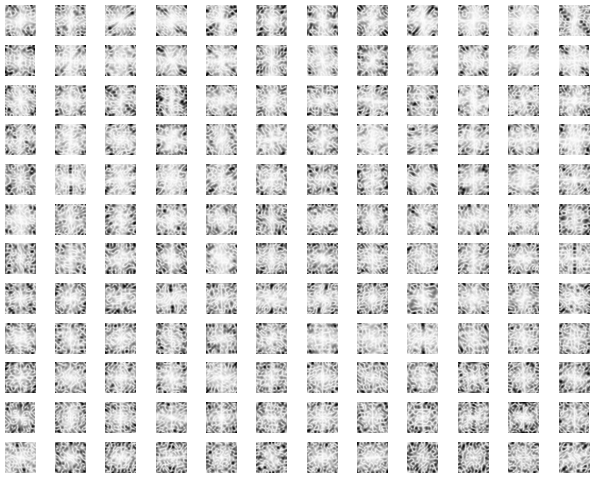

Figure 5: Amplitude spectra of the 12×12 ICA filters corresponding to an ensemble of natural images (24000 patches in total). The darker grey values refers to larger amplitudes; zero spatial frequency is at the centre of each patch.

Figure 6: 12×12 ICA bases corresponding to an ensemble of natural images (24000 patches in total). The branches and leaves that characterize the images appear in the bases.

7. CONCLUSIONS

In this paper we have analyzed, both mathematical and experimentally, the results obtained by performing an Independent Component Analysis (ICA) to natural images by means of the FastICA algorithm. In particular, we have used the kurtosis as contrast function. We have shown that the ICA filters can be decomposed in terms of the eigenvectors associated to the smallest eigenvalues of the data correlation matrix, meaning that these filters are all high-pass. This leads to a sparse distribution of the independent components. Moreover, most of the independent components present a quite marked spike

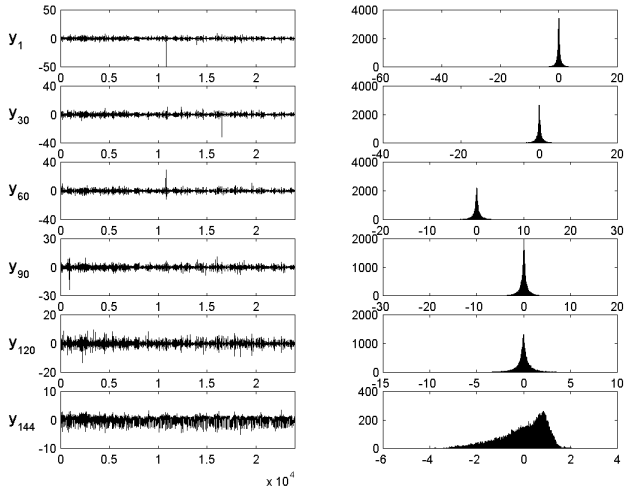

Figure 7: Some of the independent components (and their associated histograms) corresponding to an ensemble of natural images (24 000 patches in total). The scales have been adjusted to a better visualization.

that are responsible of the fact that the ICA basis are similar to the edges of the original image or images. As a conclusion, the ICA basis are very dependent of the images analyzed.

REFERENCES

[1] Amari, S. I., Cichocki, A., Yang, H.: *A new learning algorithm for blind source separation*. In Advances in Neural Information Processing Systems 8. MIT Press (1996) 757–763

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

[3] Bell, A. J., Sejnowski, T. J.: *The independent component of natural images are edge filters*. Vision Research, vol. 37, no. 23 (1997) 3327–3338

[4] Cardoso, J. F., Laheld, B. H.: *Equivariant adaptive source separation*. IEEE Trans. on Signal Processing, 44(12) (1996) 3017–3030

[5] Caywood, M., Willmore, B., Tolhurst, D.: *Independent Components of Color Natural Scenes Resemble V1 Neurons in Their Spatial and Color Tuning*. Journal of Neurophysiology, vol. 91 (2004) 2859–2873

[6] Cichocki, A., Amari, S. I.: *Adaptive blind signal and image processing*. John Wiley & Sons (2002)

[7] Comon, P.: *Independent component analysis, a new concept?*. Signal Processing, vol. 36, no. 3 (1994) 287–314

[8] Field, D. J.: *Relations between the statistics of natural images and the response properties of cortical cells*. Journal of the Optical Society of America A, vol. 4 (1987) 2379–2394

[9] Hoyer, P. O., Hyvärinen, A.: *Independent Component Analysis Applied to Feature Extraction from Colour and Stereo Images*. Computation in Neural Systems, 11 (2000) 191–210

[10] Hornillo-Mellado, S., Martín-Clemente, R., Puntonet, C.G., Acha, J. I., Górriz, J. M.: *Maximization of Statistical Moments for Blind Separation of Sources Revisited*. Accepted for publishing in *Neurocomputing*.

[11] Hyvärinen, A. : *Fast and robust fixed-point algorithms for independent component analysis*. IEEE Trans. on Neural Networks, 10(3) (1999) 626–634

[12] A. Hyvärinen, M. Gutmann and P.O. Hoyer. *Statistical model of natural stimuli predicts edge-like pooling of spatial frequency channels in V2*. BMC Neuroscience, (2005) 6–12

[13] Hyvärinen, A., Hoyer, P. O.: *A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images*. Vision Research 41 (2001) 2413–2423

[14] Hyvärinen, A., Karhunen, J., Oja, E.: *Independent component analysis*. John Wiley & Sons (2001)

[15] Hyvärinen, A., Oja, E.: *A fast fixed-point algorithm for independent component analysis*. Neural Computation, vol. 6 (1997) 1484–1492

[16] Jain, A.: *Fundamentals of Digital Image Processing*. Prentice Hall (1989)

[17] Lindgren, J.T., Hyvärinen, A.: *Learning High-level Independent Components of Images through a Spectral Representation*. Proceedings of the International Conference on Pattern Recognition (ICPR2004) (2004) 72–75

[18] Olshausen, B. A., Field, D. J.: *Sparse coding with an overcomplete basis set: a strategy employed by V1?* Vision Research, vol. 37 (23) (1997) 3311–3325

[19] Olshausen, B. A.: *Principles of image representation in visual cortex*. The Visual Neurosciences, L.M. Chalupa, J.S. Werner, Eds. MIT Press (2003) 1603–1615

[20] van Hateren, J. H., van der Schaaf, A.: *Independent component filters of natural images compared with simple cells in primary visual cortex*. Proc. Roy. Soc. Lond. B 265 (1998) 359–366

[21] *Natural stimuli collection*. J. H. van Hateren. `http://hlab.phys.rug.nl/imlib/index.html`

[22] *Waterloo Repertoire*. `http://links.uwaterloo.ca`