

A COMPARISON OF DICTIONARY BASED APPROACHES TO INPAINTING AND DENOISING WITH AN EMPHASIS TO INDEPENDENT COMPONENT ANALYSIS LEARNED DICTIONARIES

MARKO FILIPOVIĆ AND IVICA KOPRIVA

Division of Laser and Atomic Research and Development
Rudjer Bošković Institute
Bijenička cesta 54, P. O. Box 180, 10002, Zagreb, Croatia

(Communicated by Naoki Saito)

ABSTRACT. The first contribution of this paper is the comparison of learned dictionary based approaches to inpainting and denoising of images in natural scenes, where emphasis is given on the use of complete and overcomplete dictionary learned by independent component analysis. The second contribution of the paper relates to the formulation of a problem of denoising an image corrupted by a salt and pepper type of noise (this problem is equivalent to estimating saturated pixel values), as a noiseless inpainting problem, whereupon noise corrupted pixels are treated as missing pixels. A maximum a posteriori (MAP) approach to image denoising is not applicable in such a case due to the fact that variance of the impulsive noise is infinite and the MAP based estimation relies on solving an optimization problem with an inequality constraint that depends on the variance of the additive noise. Through extensive comparative performance analysis of the inpainting task, it is demonstrated that ICA-learned basis outperforms K-SVD and morphological component analysis approaches in terms of visual quality. It yielded similar performance as a field of experts method but with more than two orders of magnitude lower computational complexity. On the same problems, Fourier and wavelet bases as representatives of fixed bases, exhibited the poorest performance. It is also demonstrated that noiseless inpainting-based approach to image denoising (estimation of the saturated pixel values) greatly outperforms denoising based on two-dimensional myriad filtering that is a theoretically optimal solution for this class of additive impulsive noise.

1. Introduction. This paper presents comparative performance analysis of several approaches to dictionary learning, with an emphasis on complete and overcomplete dictionary learned by independent component analysis (ICA), motivated by practical problems in the inpainting of natural images as well as in denoising of natural images corrupted by a salt and pepper noise. The inpainting refers to the recovery of lost or corrupted parts of images that occurs, as an example, in image or video restoration in reverting deterioration, removing selected elements or filling in missing pixels [5], [14]. Salt and pepper type of impulsive noise refers to the situation when corrupted pixels take either maximum or minimum gray values. Median and

2000 *Mathematics Subject Classification.* Primary: 68U10.

Key words and phrases. Inpainting, denoising, learned basis, independent component analysis, K-SVD, sparse representation.

This work was supported through grant 098-0982903-2558 funded by the Ministry of Science, Education and Sports, Republic of Croatia.

myriad nonlinear filtering approach to denoising works well when a small number of pixels is corrupted but at high noise densities, when the window size has to be increased, it leads to blurring [2]. This paper proposes an inpainting-based approach for denoising an image corrupted by a salt and pepper noise that adds up to the estimation of the saturated pixel values, whereupon saturated pixels are treated as missing pixels [48], [36]. This virtually converts an image denoising problem into the noiseless inpainting problem. Thus, inpainting and denoising are unified within the same common denominator referring to learning a basis that provides a sparse representation of an image. Therefore, it is expected that the proposed denoising method will enable a high quality of the restored image even when a large amount of pixels is saturated i.e. missing. A maximum a posteriori (MAP) approach to image denoising is not applicable in such cases due to the fact that the variance of the impulsive noise is infinite. The MAP based estimation relies on solving an optimization problem with an inequality constraint that depends on the variance of the additive noise (please see discussion related to eq. (2)), [15], [34]. As opposed to the approach to the estimation of the saturated pixel values proposed in [48], the method proposed herein is aimed for gray scale images and does not rely on any specific assumption such as correlation between the color channels as it has been used in [48]. As opposed to the approach proposed in [36], which estimates saturated pixel values through sparse reconstruction in fixed basis, the approach proposed herein solves the related inpainting problem in a learned basis. As it is demonstrated in section 3 that yields significantly better quality of reconstructed images than when related inpainting problem is solved in fixed bases.

The inpainting problem is casted into the following mathematical framework. The vectorized (column-wise) image $x \in \mathbf{R}^n$ is to be reconstructed from the vector of known pixels $y \in \mathbf{R}^l$, where $l < n$. It is assumed that the unknown image x can be represented by $k \ll n$ significant coefficients over the known m -dimensional basis¹ $D \in \mathbf{R}^{n \times m} : x \approx Dc$, $\|c\|_0 = k$ and $m \geq n$. Here, $\|c\|_0$ stands for the ℓ_0 -quasi-norm that counts the number of the nonzero coefficients of the vector c . The vector of known pixels y is related to the unknown vectorized image x through $y = Mx$, where $M \in \mathbf{R}^{l \times n}$ has a special structure determined by the layout of the missing pixels. Thus, the following relation describes the inpainting problem:

$$(1) \quad y = Mx = MDc = \Phi c$$

Hence, the solution of the inpainting problem is obtained as the sparse solution of the underdetermined system of linear equations (1). Once the estimate of c , denoted as \hat{c} , is obtained, the estimate of x is obtained from: $\hat{x} = D\hat{c}$ ². In the noiseless scenario, underdetermined system of linear equations (1) has a unique solution if the number of known pixels l satisfies: $l \geq 2k$, see [45] for a review of this and related results. Theoretical results defining conditions under which the problem (1) is solvable are discussed in slightly more details in section 2. There, the nonlinear

¹Throughout the paper we will use the term ‘‘basis’’ for matrix D both when $m = n$ and $m > n$. When $m = n$ the basis is critically sampled or complete while for $m > n$ the basis is overcomplete and is also referred to as a frame, [28], [29].

²The approach to the inpainting proposed in this paper presumes a learning of the basis D using patches collected from the training images. In the actual inpainting phase, the learned basis D is used to recover the original image x . Formally, if $I \in \mathbf{R}^{\sqrt{n} \times \sqrt{n}}$ is an image patch, then x denotes vectorized version of the image patch I .

signal reconstruction algorithm used for the solution of the inpainting problem (1) is discussed in more details as well.

Choice of Φ in (1) is crucial for the success of inpainting. Because $\Phi = MD$ the choice of Φ is dictated by the choice of the matrix M and the basis matrix D , whereas the matrix M has special structure defined by the layout of missing pixels, that is also observed in [23], [24]. Thus, it is the basis matrix D that governs the quality of the image reconstruction, whereas the basis which enables representation of the signal vector x with $k \ll n$ coefficients is of special interest. Such basis can be either selected from a specified set of linear transforms (such as Fourier or wavelet transforms), or learned from a set of training signals. A representative example for the first case is the paper [14], where images are modeled as a combination of piecewise smooth (“cartoon”) and texture part, wherein the cartoon part is sparsely represented by wavelet-like transforms, and the texture part is sparsely represented by local DCT or related transforms. On the other hand, better results are obtained by using bases learned from a training set by using K-SVD like algorithms [1], introduced and discussed later in this paper. Our approach to basis learning is *example-based*: we select a set of training images, randomly extract large number of patches from each image in the training set, and learn a basis that yields sparse representation of the selected patches³. Basis learning is accomplished using the fast independent component analysis (FastICA) algorithm, [27], with the tanh nonlinearity. The FastICA algorithm is chosen due to two reasons: (1) it has information-theoretic interpretation that enables to cast the basis learning problem within a probabilistic framework⁴; (2) use of the FastICA enables to learn an overcomplete basis, which is not the case with the ICA-based approach to basis learning presented in [31]⁵. Since the level of sparseness achieved by basis D is of crucial importance for the quality of the solution of the inpainting problem, it is more important that nonlinear function used in the ICA learning rule yields code that is sparse than statistically independent⁶.

According to [43] three major approaches to inpainting rely on: physical simulation, where images are generated via simulation of the underlying physical process, random fields, and function spaces. One group includes partial differential equations based methods [5], [6]. The other group is within Bayesian framework relying of

³It should be noted that in [15] the basis (i.e. the “dictionary”) is trained directly on the corrupted image, which gives better results since the learned basis is better adapted to the image and the inpainting problem (since training takes into account the pattern of missing pixels). However, that approach is much slower compared to the training set-based approach.

⁴For example, it is described in the section 3.2 that choice of the tanh nonlinearity corresponds to the probability density function of the code in the linear mixture model (1), which is sparse. To be more specific, choice of the tanh(5c) nonlinearity induces code with the Laplacian distributions, please see the Figure 3, that is commonly used to model sparse distributions of speech and audio signals in general.

⁵Incapability to learn an overcomplete basis or frame was a formal objection for not using the ICA for basis learning. This, in part, was motivation for development of fields of experts approach to basis learning, [40]. However, it has been shown in the early work in [25] how to estimate an overcomplete basis of independent components from an image data using the concept of quasi-orthogonality. Therefore, in section 3 we discuss how FastICA can be used to learn both complete and overcomplete basis.

⁶Results presented in [7] are not related to the performance of the inpainting problem in the ICA-learned basis. They state that statistical independence of the code obtained by the ICA algorithm under linear model of the natural image is not significantly smaller than the one obtained by second order statistical methods. What is important for the quality of image reconstruction is, however, sparseness and not statistical independence of the code.

statistical models of image priors, [40], [42]. Basically all approaches are connected with fundamental problem of image modeling. The approach presented here is most similar to the inpainting algorithm proposed in [34].

The rest of the paper is organized as follows. Section (2) relates inpainting problem to the underlying problem of solving underdetermined system of linear equations with sparseness constraints. Section 3 elaborates on the ICA-based approach to the learning of a complete and overcomplete bases. The K-SVD approach to basis learning is also briefly described there. Results of comparative performance analysis between ICA- and K-SVD learned bases, morphological component analysis (MCA) that relies on a union of fixed bases⁷, fields of experts approach [40], [41] that belongs to the class of Markov random field models of images, as well as Fourier and wavelet bases, are presented in section 4. The comparative analysis has been performed on the problem of inpainting of natural images in several scenarios related to different structures of the missing pixels. It is demonstrated that ICA-learned basis outperforms other bases in terms of reconstruction quality, measured through peak-signal-to-noise-ratio and structural similarity index, [46], [47]. It is additionally demonstrated in section 4 how inpainting can be used for denoising purpose when image is corrupted by the salt and pepper type of additive impulsive noise, i.e. how to estimate the values of the saturated pixels, whereupon saturated pixels are treated as missing pixels. There, it is demonstrated that the inpainting-based approach to image denoising greatly outperforms the one based on 2D myriad filtering which is a theoretically optimal solution within the filtering framework, [21], [2]. The conclusion is presented in section 5.

2. Preliminaries. Methods for finding sparse solution of problem (1) are briefly discussed in this section. What is known is vector $y \in \mathbf{R}^l$ and matrix $\Phi \in \mathbf{R}^{l \times m}$ which is overcomplete, i.e. $l < m$. The aim is to find the sparsest solution of this *underdetermined* linear system; this is motivated by supposed sparsity of the vector of coefficients $c \in \mathbf{R}^m$. Sparsity of a vector $c \in \mathbf{R}^m$ is usually measured by its ℓ_0 -quasi-norm, denoted by $\|c\|_0$ and defined as $\|c\|_0 = |\{i : c_i \neq 0\}|$, i.e. as the number of its nonzero coefficients. This leads to the following formulation of the problem:

$$\arg \min_c \|c\|_0 \quad \text{subject to} \quad \Phi c = y$$

A more practical formulation, one that allows some noise or error, is

$$(2) \quad \arg \min_c \|c\|_0 \quad \text{subject to} \quad \|\Phi c - y\|_2 \leq \epsilon$$

where ϵ represents an error tolerance that depends on the noise variance. Related formulations are

$$\arg \min_c \{\|\Phi c - y\|_2^2 + \lambda \|c\|_0\}$$

and

⁷The MCA models an image as a combination of piecewise smooth (“cartoon”) and texture parts. The image decomposition part and the filling-in part are integrated using a union of fixed bases consisting of one adapted for cartoon and the other adapted for texture part. Inpainting is solved by combining sparse representation of each part. As opposed to that, the ICA-based basis learning approach, although limited to small areas of missing pixels, solves the inpainting problem by learning a single basis only.

$$\arg \min_c \|\Phi c - y\|_2^2 \quad \text{subject to} \quad \|c\|_0 \leq s$$

where parameters λ and s depend on error tolerance ϵ . The difficulty with this formulation is its complexity, which is exponential in m (since (2) is combinatorial optimization problem). This makes solving (2) computationally intractable⁸. However, computationally a more feasible approach is found by intuitively replacing ℓ_0 with its continuous approximation, like ℓ_p -quasi-norm, defined as

$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{1/p}$$

For $0 < p < 1$, this is not a convex function. Algorithms minimizing ℓ_p -norm, for $p < 1$, include [11], [12], [19], [22]. These approaches were shown to outperform ℓ_1 -norm minimization in practice. In [37] authors approximated ℓ_0 -norm of x by $\|x\|_0 \approx m - F_\sigma(x)$, where $F_\sigma(x) = \sum_i f_\sigma(x_i)$ and

$$(3) \quad f_\sigma(x) = \exp\left(\frac{-s^2}{2\sigma^2}\right)$$

is an approximation of the indicator function of the set $\{0\}$. The smaller the parameter σ , the better approximation of ℓ_0 we get; the trade-off is that the larger σ is, the smoother function we get, which therefore is easier to optimize. Note that minimizing this approximation of ℓ_0 is equivalent to maximizing F_σ . The idea is to maximize F_σ for larger values of parameter σ by only a few iterations of gradient ascent, and use this approximation as the initial point to maximize F_σ for smaller values of the parameter σ . After each iteration the computed approximation of the solution c needs to be projected onto the constraining set $\Phi c = y$. This simple approach seems to work *very well* in practice and outperforms more commonly used ℓ_1 -norm minimization in terms of accuracy and computational efficiency. It is up to two orders of magnitude faster than the methods based on ℓ_1 -norm minimization. Due to these reasons, this method has been used in implementation of inpainting and denoising algorithms in comparative performance analysis presented in section Results. The condition that guarantees uniqueness of the solution of the problem (2) is given in terms of *mutual coherence* of the matrix Φ , denoted as $\mu(\Phi)$, which is defined as the maximal absolute normalized inner product between the columns of Φ . The main result [10] states that, if a solution c_0 of (2) satisfies $\|c_0\|_0 < (1 + 1/\mu(\Phi))/2$, then every other solution is close to c_0 . Mutual coherence is theoretically a worst-case property and gives very pessimistic bounds on the sparsity of the solution. However, it is computed easily and can be used for rough prediction of the degree of sparsity of c necessary to guarantee that the solution of (2) is recovered up to a small error. Thus, according to Figure 1, which shows that on a given training sample the ICA-learned basis has smaller mutual coherence than K-SVD learned basis, ICA-learned basis would require less degree of sparsity of c than K-SVD learned basis. This is important in the light of

⁸Some of the heuristic approaches that try to directly solve the problem (2) are greedy algorithms like orthogonal matching pursuit (OMP) and iterative hard thresholding approaches [8]. We haven't used them in the inpainting stage since the obtained results were inferior. OMP was used as a part of the K-SVD algorithm for basis learning, see text for details.

the fact that natural images are not exactly sparse, but coefficients of image patches in chosen basis decay slowly, see Figure 2.

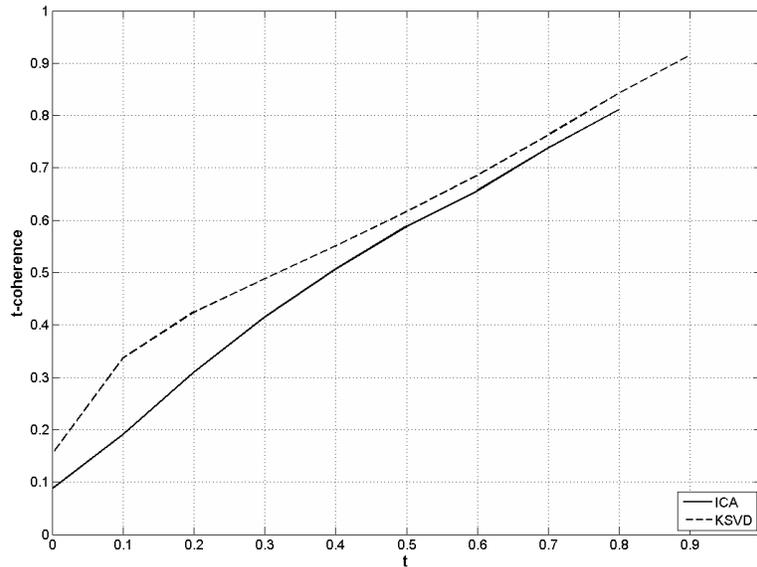


FIGURE 1. Comparison of t -coherence of ICA and K-SVD learned bases as a function of t . For a given t , $0 \leq t < 1$, t -coherence of a matrix is defined as the mean value of all absolute normalized inner products between different columns of the matrix that are above or equal to t . For $t \rightarrow 1$, t -coherence approaches a mutual coherence measure which is defined as maximal absolute normalized inner product between the columns of a matrix.

3. Basis learning algorithms. We now briefly review important work done in the bases learning domain and present the ICA-based approach to basis learning. In [38], the method has been derived that yields an overcomplete dictionary trained on the patches of natural images by maximizing sparseness based cost function. Similar results were presented in [4] using ICA, and in [31], [32] using probabilistic methods. In a deterministic setting, with data matrix X of size $n \times T$ with columns that represent the *training set* for the signals of interest, the dictionary learning problem is stated as

$$(4) \quad \arg \min_{D, C} \|X - DC\|^2 \quad \text{subject to} \quad \|c_i\|_0 \leq K$$

where c_i denotes columns of the coefficients matrix C , and K is a bound on the sparsity of representation. The usual method for solving (4) is alternating minimization: first, D is fixed and C is approximated by some of the algorithms reviewed in the previous section. This stage is called sparse coding. In the next step, C is fixed and D is updated. This alternating minimization is repeated until convergence to a local minimum (it is also possible that algorithm gets stuck to a saddle point of the problem). The MOD [16] algorithm follows this approach. The dictionary

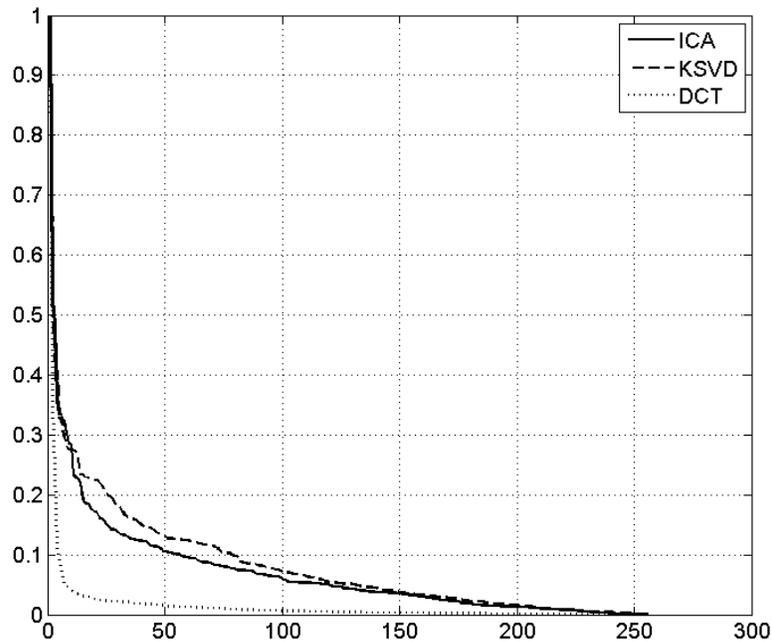


FIGURE 2. Coefficients in different bases of patch chosen randomly from natural image. Coefficients were normalized for comparison purpose. Without normalization coefficients in K-SVD and DCT bases are much larger in absolute value (up to three orders of magnitude) than those in ICA basis.

update step in MOD is calculated as $D_{k+1} = XC_k^\dagger$, where subscript denotes iteration number and C^\dagger is pseudoinverse of C . Calculating the pseudoinverse for each iteration of the algorithm makes MOD slow. K-SVD algorithm, reviewed in subsection 3.1, fixed this partially. It should be noted that the dictionary learning problem could be stated more generally than in (4), by using some other measures of representation error and sparsity of coefficients. In the two following subsections, we discuss ICA- and K-SVD-based approaches for approximating the local solution of the problem (4).

3.1. K-SVD algorithm. The K-SVD algorithm is generalization of the K-means clustering approach to basis learning with sparseness constraints when number of clusters (that corresponds to the level of sparseness K) is greater than 1. Hence, philosophy in basis learning by the K-SVD algorithm reflects the knowledge that clustering yields the most efficient signal representation, i.e. representing the signal by one coefficient only. Its development was inspired in part by computational inefficiency of the MOD basis learning algorithms [16], [17], [44]. The K-SVD algorithm starts with the formulation (4). The improvement upon the MOD is in the dictionary update stage. Contrary to the MOD, the K-SVD updates D column by column. The objective function in (4) can be written as

$$\|X - DC\|_F^2 = \left\| \left(X - \sum_{j \neq k} d_j c^j \right) - d_k c^k \right\|_F^2$$

where d_j denote columns of D , and c^j denote rows of C . We denote by ω_k a set of indices i for which $c^j(i) \neq 0$, and by Ω_k a matrix of size $N \times |\omega_k|$ with ones at positions $(\omega_k(i), i)$ and zeros elsewhere. If we define the error matrix E_k by $E_k = X - \sum_{j \neq k} d_j c^j$, then K-SVD minimizes $\|E_k \Omega_k - d_k c^k \Omega_k\|_F^2$ with respect to d_k and c^k . Therefore, updated d_k and c^k are first, scaled, left and right singular vectors of the restricted matrix $E_k^R = E_k \Omega_k$, respectively. By restricting E_k and c_k in this way, updated c^k is forced to have the same or smaller support than the previous one. Provided that sparse coding stage is solved exactly, this guarantees convergence to a local minimum of objective function in (4) on the constraints set. Since it is solved only approximately, convergence is not guaranteed. However, it seems that convergence occurs in practice. Usually, algorithm used in sparse coding stage of the K-SVD is the OMP, for the following reasons: first, it naturally finds an approximation with fixed number of nonzero coefficients, which fits into the dictionary learning framework (4), and second, it is fast, especially when compared to convex relaxation methods. We discuss details related to the K-SVD algorithm, such as choice of parameter K , in the Results section. Since, it has been demonstrated in [1] that the K-SVD outperforms the MOD algorithm, in the Result section the MOD is not included in the comparative performance analysis between the bases learning algorithms.

3.2. ICA for complete and overcomplete basis. In comparison with the probabilistic framework to basis learning in [31], that in part is also based on the use of ICA, the use of ICA proposed here is motivated by two reasons: a) it extends the probabilistic framework to learn the overcomplete basis. This is achieved through the use of the FastICA algorithm, [27], that works in sequential mode; b) in regard to the probabilistic framework to basis learning presented in ref. [31], the adopted ICA approach is more flexible. This is due to the fact that proper selection of the nonlinear functions (that are related to parameterized form of the probability density functions of the code) enables basis learning that is tied with a code with the prespecified level of sparseness without affecting the structure of the basis learning equation (by ICA the basis inverse is actually learned). As opposed to that, in the Bayesian paradigm to the basis learning presented in [31], the structure of the basis learning equation depends on the choice of what was previously imposed on the probability density function of the code.

We suppose that the linear model $x = Dc$ is valid, where x and c are random vectors (we interpret columns of the data matrix X , denoted as x_i , as realizations of x), and D is the basis matrix we want to estimate. For now we consider only the complete case (D is a $n \times n$ square matrix, and x and c are n -dimensional). Hence, the basis D is what in blind source separation is referred to as a mixing matrix, [26]. Extraction of the code matrix C (also referred to as a source matrix in blind source separation) can be performed by means of the ICA algorithms. Herein, we are interested in the ICA algorithm that: a) can be casted into the probabilistic framework tied with the linear generative model as in [31]; b) can be extended for

learning the overcomplete basis⁹. In this regard, assuming the linear generative model $x = Dc$, the minimization of the mutual information $I(c)$ is used:

$$(5) \quad I(c) = \sum_{i=1}^n H(c_i) - H(x) - \log |\det D^{-1}|$$

where $H(c_i)$ stands for the differential entropy of the code and $H(x)$ stands for the joint entropy of data. The ICA algorithms that maximize information flow through nonlinear network (Infomax algorithm) [3], maximize likelihood (ML) of the ICA model $x = Dc$ [39], or minimize mutual information between components of $c = D^{-1}x$ [18], are equivalent in a sense that all minimize $I(c)$ and yield the same learning equation for D^{-1} :

$$(6) \quad D^{-1}(k+1) \leftarrow D^{-1}(k) + \eta [I - \phi(c(k))c(k)^T] D^{-1}(k)$$

where ϕ represents the *score function* defined as

$$(7) \quad \phi_i = -\frac{1}{p_i} \frac{dp_i}{dc_i}$$

and p_i stands for the probability density function of the code c_i . Hence, by proper selection of the p_i , the code c_i with a distribution with the prespecified level of sparseness can be generated. This virtually learns the basis D that efficiently represents data x . For this purpose it is useful to represent the density function p_i in (7) by generalized Gaussian density, which is parameterized by one parameter β [13], [49]:

$$p_i(c_i) = \frac{\beta_i}{2\sigma_i\Gamma(1/\beta_i)} \exp\left(-\frac{1}{\beta_i} \left|\frac{c_i}{\sigma_i}\right|^{\beta_i}\right)$$

where β_i is a Gaussian exponent, Γ stands for the Gamma function and σ_i represents the variance of c_i . Inserting above equation in (7) yields the score function in a parameterized form

$$\phi_i(c_i) = \text{sgn}(c_i) |c_i|^{\beta_i-1}$$

Hence, the selection of the generalized Gaussian exponent β_i enables modeling of some pre-specified density function. Super-Gaussian (i.e. sparse) densities are modeled by choosing $\beta_i < 2$. This enables learning the basis matrix D that gives sparse representation for x_i . Due to its capability to learn the overcomplete basis, the FastICA algorithm was selected herein for the basis learning purpose. It also has the information-theoretic interpretation, but, as opposed to ML algorithm (6), it

⁹To avoid misunderstandings regarding the possible application of sparse component analysis (SCA) [9], [20] for the basis learning problem, we discuss the main differences between the SCA and ICA from the basis learning standpoint. The SCA methods are focused on the solution of the underdetermined blind source separation (uBSS) problems that are characterized by having more sources than sensors. To solve the uBSS problem the SCA methods first estimate the basis matrix by means of clustering and afterwards estimate sources through the solution of the underdetermined system of linear equations. The basis learning problem that is based on the linear generative model reverses this sequence by focusing on the generation of the code (sources in the BSS terminology) that is sparse and obtaining, virtually, the associated basis matrix as a byproduct. Hence the SCA methods are not applicable to the basis learning problem.

employs fixed-point instead of gradient-based minimization of the contrast function that approximates negentropy of the code

$$(8) \quad J(c_i) \propto [\langle \{G(c_i)\} \rangle - \langle \{G(\nu)\} \rangle]^2$$

where G is some non-quadratic function, $\langle \cdot \rangle$ denotes mathematical expectation, and ν represents standardized Gaussian random variable. The approximation (8) is related to mutual information $I(c)$, derived for the assumed model $x = Dc$ i.e. $c = D^{-1}x$, when $G(c_i) = -\log p_i$ is substituted in marginal entropies in (5). Hence, the contrast function (8) is related to contrast functions used by ML [39], information maximization [3], and minimum mutual information [18] ICA algorithms. Thus, by the virtue of same logic as in (6) and (7), nonlinear function in the FastICA algorithm can be also chosen to generate sparse distribution of the code c_i . In the experiments reported in section 4 we have used the tanh function that is associated with $G(c_i) = (1/a) \log \cosh ac_i$. Such a choice for G approximates density function of the type

$$p_i \sim \left(\frac{1}{\cosh ac_i} \right)^{1/a}$$

that models sparse or super-Gaussian distributions¹⁰. See Figure 3 for the comparison between the generalized Gaussian and this density.

The FastICA algorithm was extended to overcomplete case in [25] using the concept of quasi-orthogonality. Here, the basis matrix D has more columns than rows, and the idea is to make its columns as orthogonal as possible. In the sequential mode of the FastICA, basis vectors are estimated one at a time. After every iteration, the basis vector is orthogonalized with respect to previously estimated basis vectors using the Gram-Schmidt orthogonalization. This idea can be extended to overcomplete case as follows

$$(9) \quad d_i \leftarrow d_i - \alpha \sum_{j=1}^{i-1} (d_i^T d_j) d_j$$

where $0 < \alpha < 1$ is a parameter ($\alpha = 1$ corresponds to the ordinary orthogonalization) and is usually chosen heuristically. We discuss our choice of α in section 4.

4. Results. This section presents numerical results related to the comparative performance analysis of the basis learning methods on the problems of inpainting and denoising of natural images. All reported numerical simulations were done in MATLAB 7.7 on a 3 GHz Dual-Core Windows XP PC with 2GB memory. The section is organized as follows. In subsection 4.1 we describe the basis learning procedure. Subsection 4.2 describes parameter selection of sparse recovery algorithms used for the inpainting. There, the inpainting methods used in the comparative performance analysis are described as well. Subsection 4.3 presents comparative results of denoising of natural images corrupted by the salt and pepper noise.

¹⁰Experimental results reported in section 4 were obtained with scale parameter $a = 5$. Other choices ($a = 1$ and $a = 10$) were tried, but this one gave best results.

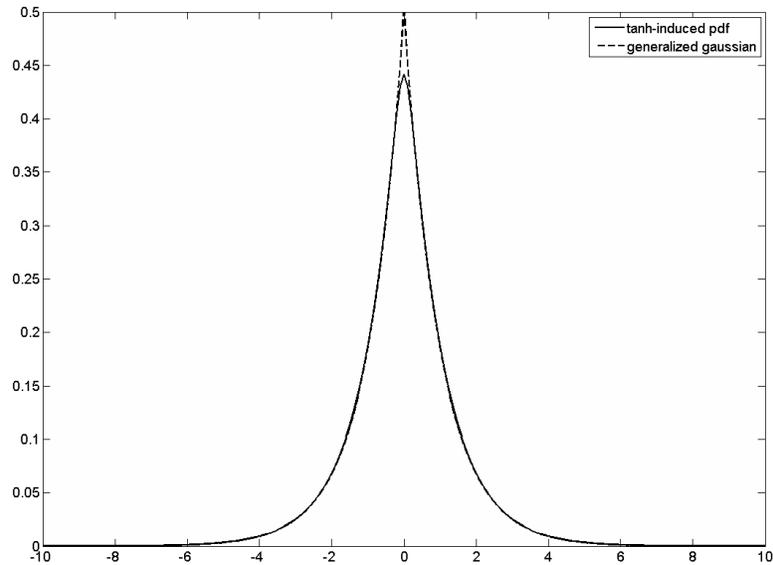


FIGURE 3. Comparison of the probability density function (pdf) induced by using tanh nonlinearity with the scale parameter $\alpha = 5$ in the FastICA algorithm, and generalized Gaussian pdf with parameter $\beta = 1$, which models Laplacian pdf.

4.1. Bases learning. Six images of natural scenes, shown in Figure 4, were used as the training set for learning the basis matrix. Images were taken from a publicly available database¹¹ and converted to grayscale. Training images were of the size 576×768 pixels. We randomly extracted 18000 patches of the size 16×16 pixels from six training images (3000 patches per training image) and organized them as columns of the 256×18000 data matrix X . Mean value was subtracted from every patch: this is a very important preprocessing step. Then, the ICA and the K-SVD were used for basis learning. In many papers in the literature smaller patches were used, and hence also smaller number of atoms in the K-SVD training stage. For example, in [15] the authors used patches of the size 8×8 pixels, and the number of atoms in the training stage was 6, whereas in the inpainting experiments in [34] (although they worked with color images) patches of the size 9×9 pixels were used and the number of atoms in the training stage was 20. Therefore, for comparison, below we also present results obtained using smaller patches, of the size 8×8 pixels, and smaller number of atoms in the K-SVD training stage, namely 4. It should be noted that one of the seminal papers [38] also used patches of size 16×16 pixels, as well as recent papers [33] and [35] (in [35], few patch sizes were used for comparison). Another important point is that the same patch sizes are used for both bases, which allows fair comparison between them.

¹¹A. Olmos, F. A. A. Kingdom, McGill calibrated color image database, 2004., <http://pirsquared.org/research/mcgilldb/>



FIGURE 4. Six images from the training set used for bases learning. Images were randomly selected from¹¹.

MATLAB implementations of the FastICA algorithm, available at¹², and the K-SVD, available at¹³, were used. As discussed in section 3.2, tanh nonlinearity in the FastICA algorithm has been used because it yields components with sparse (super-Gaussian) distributions. Sparse coding stage in the K-SVD was done using the OMP algorithm with 40 nonzero coefficients of a solution. Smaller number of nonzero coefficients would speed up the K-SVD algorithm and possibly find sparser representation, but simulations showed that this doesn't yield a better performance. Namely, if the training set doesn't allow such a sparse representation, which seems to be the case with the real signals such as natural images, then this model is not appropriate. Therefore, this number was chosen heuristically to obtain sparsity, but only at a reasonable, realistic level. For comparison, we have also tried running the K-SVD algorithm with 4 nonzero coefficients of a solution, see later text. It is expected that better performance can be obtained by using some other reconstruction algorithm, but the reason for choosing the OMP is its speed. Basis learning with the K-SVD algorithm took around 5 hours for 100 iterations, while basis learning for the complete case with the FastICA took around 3 hours. It should be noted that the K-SVD is much faster when using smaller number of atoms, but with larger number of atoms better performance was obtained. Figure 5 shows basis vectors learned by the FastICA and K-SVD algorithms with the patches of the size 16×16 pixels. The FastICA and K-SVD were also used for learning overcomplete basis. Sequential version of the FastICA algorithm was modified to perform quasi-orthogonalization after every step, as explained in Section 3. Parameter α in (9) was set to 0.5. The K-SVD takes the number of basis vectors directly as a parameter.

¹²The FastICA package for MATLAB, <http://www.cis.hut.fi/projects/ica/fastica/index.shtml>

¹³The KSVD-Box MATLAB toolbox, <http://www.cs.technion.ac.il/~ronrubin/software.html>

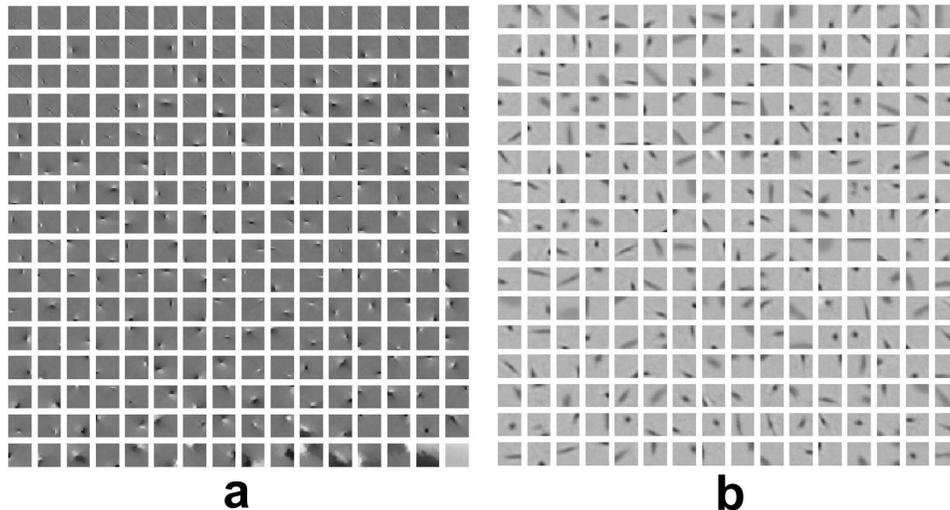


FIGURE 5. 256 matricized basis vectors of the size 16×16 pixels learned by ICA, (a), and K-SVD, (b), algorithms. Basis vectors are columns of the learned basis matrix.

4.2. Inpainting. We have used freely available MATLAB implementation of the Smoothed ℓ_0 (SL0) algorithm¹⁴ for image reconstruction, i.e. inpainting. Justification of this choice was already discussed in Section 2. The parameter σ in (3) was chosen as suggested by the authors: we randomly selected an image patch and computed its coefficients in the learned basis by applying a direct transformation. Then the absolute values of these coefficients were sorted in descending order and the smallest 80 percent of them were interpreted as noise (see discussion in Section 2). Parameter σ was selected to be few times larger than the standard deviation of the vector of these smallest coefficients. In this way, since natural image patches are not exactly sparse signals, smallest coefficients in selected basis were interpreted as a noise/error. We have also experimented with the other values of σ and found that this choice yields best results. In our simulations the SL0 algorithm worked better and was much faster than other approaches, especially those using ℓ_1 minimization. For example, we have also tested the ℓ_1 ls algorithm for the problem (4) (MATLAB implementation is available at¹⁵). On average, reconstruction using the ℓ_1 ls took 10 to 15 minutes per image, while the SL0 took about 30 seconds per image only. The OMP performed worse (in terms of the measures of the quality of reconstructed images, see below) than the SL0 and the ℓ_1 ls, with a computational complexity of the order of few minutes. For every patch, before the reconstruction, mean value of the observed pixels in the patch was subtracted from the vector of the observed pixels and added back after the reconstruction. Thus, DC component was artificially added in the reconstruction yielding better results than when the DC component was a part of the basis. To prevent border effects, reconstruction was done with two rows, i.e. columns, of adjacent patches overlapping. After reconstruction, overlapping regions were averaged.

¹⁴<http://ee.sharif.ir/~SLzero/>

¹⁵http://www.stanford.edu/~boyd/l1_ls/

For measuring the quality of the reconstructed images, we have used structural similarity index (SSIM) [46], [47] and peak signal-to-noise ratio (PSNR). We have noted that PSNR can give higher values (that should correspond to higher image quality) despite obvious visual quality degradation. That is in line with the objection already pointed out in [46] that high PSNR value does not always correspond with the high quality of visual perception. It was demonstrated that the SSIM is the metric that better corresponds to subjective quality of visual perception. The SSIM index is computed locally on image patches, within a sliding window that moves pixel-by-pixel across the image; local SSIM measures the similarity of local patch brightness values, contrasts and structures. For more details, we refer the interested reader to the paper [47]. Global SSIM is computed as an average of the SSIM values across the image. It has values between -1 and 1 , achieving maximum value 1 if and only if the images being compared are equal. MATLAB code for computing the SSIM index is available at¹⁶. For comparison, we have also used the MCA and the Fields of experts (FoE) [41] methods for image inpainting. MATLAB implementation of the MCA is available as a part of the MCALab package¹⁷. Dictionaries used in the MCA were curvelets for the cartoon part and two-dimensional cosine packets for the texture part (see Introduction for a brief explanation of the MCA). Parameters of the MCA were the same as in the Barbara image inpainting example that is available as a part of the MCALab package. Namely, window width for the cosine packets was 32 pixels, and coarsest scale for the curvelets was 2. The hard thresholding was used with linear decrease schedule, and the number of iterations was 300. MATLAB implementation of the FoE method is available at¹⁸. Default values of the parameters were used.

Six images shown in Figure 6, also taken from¹¹, were used as the validation set. These images were reshaped to the size of 512×512 pixels. Since the distribution of missing pixels was generated randomly, we repeated the inpainting experiment 10 times for every image in the validation set, every time randomly generating missing pixels distribution. We did not repeat inpainting experiments with the MCA and FoE 10 times because reconstruction was slow (the MCA took about 50 minutes for one image, while the FoE took about 5 hours for one image). Table 1 shows detailed results of inpainting of six natural images in the validation set using: the ICA and K-SVD learned complete bases and DCT and symmlet 4 wavelet fixed bases. Also presented are results using the MCA and the FoE. Numbers in the table stand for mean values and standard deviations of the SSIM metric of the reconstructed images after 10 runs.

It is clear that the learned bases greatly outperformed fixed bases. It is also clear that the ICA outperformed the K-SVD, although not by that large margin. One reason for better performance of the ICA-learned basis is its smaller coherence, see Figure 1. Table 2 shows the corresponding results in the PSNR metric. It can be seen that results using both metrics are consistent. For comparison, in Table 3 we also show results obtained by the ICA and K-SVD bases learned on patches of the size of 8×8 pixels, wherein the number of atoms used in the K-SVD training phase was 4. It can be seen that the comparative performance of the two bases is similar. The K-SVD basis performed even worse than when using larger patches

¹⁶<http://www.ece.uwaterloo.ca/~z70wang/research/ssim/>

¹⁷<http://www.greyc.ensicaen.fr/~jfadili/demos/WaveRestore/downloads/mcalab/Home.html>

¹⁸<http://www.gris.informatik.tu-darmstadt.de/research/visinf/software/index.en.htm>

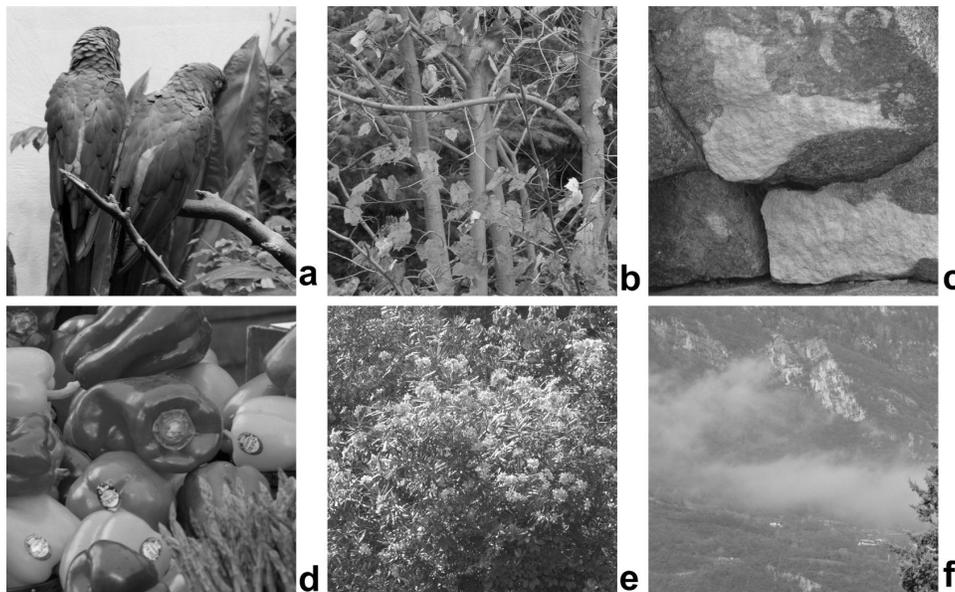


FIGURE 6. Images used for validation purpose. Images were randomly selected from¹¹.

TABLE 1. Inpainting results for the complete bases in terms of the SSIM metric for the complete bases learned on patches of size 16×16 pixels.

	ICA	K-SVD	DCT	Symmlet 4 wavelet	MCA	FoE
Fig. 6a	0.907 \pm 0.0008	0.905 \pm 0.001	0.75 \pm 0.0008	0.736 \pm 0.0022	0.789	0.92
Fig. 6b	0.76 \pm 0.0016	0.749 \pm 0.0012	0.55 \pm 0.0015	0.503 \pm 0.0015	0.682	0.77
Fig. 6c	0.773 \pm 0.0007	0.766 \pm 0.0011	0.617 \pm 0.0011	0.562 \pm 0.003	0.644	0.78
Fig. 6d	0.944 \pm 0.0005	0.94 \pm 0.0004	0.81 \pm 0.0007	0.81 \pm 0.0017	0.854	0.95
Fig. 6e	0.6 \pm 0.002	0.577 \pm 0.0015	0.434 \pm 0.0015	0.35 \pm 0.0022	0.491	0.6
Fig. 6f	0.919 \pm 0.0006	0.917 \pm 0.0005	0.84 \pm 0.0003	0.812 \pm 0.0009	0.852	0.92
Mean	0.817 \pm 0.001	0.809 \pm 0.0009	0.666 \pm 0.0008	0.63\pm0.002	0.719	0.824

and larger number of atoms in the training phase. Figure 7 shows degraded and reconstructed versions of the two images from the validation set, whereupon 80 percent of pixels were removed randomly from each image in the validation set. Images reconstructed using fixed basis are not shown because they are inferior.

TABLE 2. Inpainting results in terms of the PSNR metric for the complete bases learned on patches of size 16×16 pixels. The values are in dB.

	ICA	K-SVD	DCT	Symmlet 4 wavelet	MCA	FoE
Fig. 6a	30.2 \pm 0.07	30.2 \pm 0.08	25.4 \pm 0.02	23.4 \pm 0.1	27.1	30.9
Fig. 6b	24.4 \pm 0.04	24.1 \pm 0.04	21.6 \pm 0.02	19.8 \pm 0.03	23.6	24.7
Fig. 6c	29.7 \pm 0.01	29.3 \pm 0.02	27.1 \pm 0.02	25.6 \pm 0.05	27.6	29.8
Fig. 6d	34.3 \pm 0.08	34.4 \pm 0.06	28.4 \pm 0.05	27 \pm 0.1	30.3	35.5
Fig. 6e	19.5 \pm 0.03	18.8 \pm 0.03	18.2 \pm 0.01	16 \pm 0.01	18.4	19.5
Fig. 6f	33.4 \pm 0.09	32.9 \pm 0.07	30.5 \pm 0.02	28.6 \pm 0.09	30.7	33.1
Mean	28.6 \pm 0.05	28.3 \pm 0.05	25.2 \pm 0.03	23.4 \pm 0.07	26.3	28.9

TABLE 3. Inpainting results in terms of the SSIM metric for the complete bases learned on patches of size 8×8 pixels.

	ICA	K-SVD
Fig. 6a	0.9 \pm 0.002	0.9 \pm 0.0007
Fig. 6b	0.75 \pm 0.003	0.74 \pm 0.0016
Fig. 6c	0.77 \pm 0.001	0.76 \pm 0.0012
Fig. 6d	0.936 \pm 0.0004	0.935 \pm 0.0007
Fig. 6e	0.58 \pm 0.004	0.575 \pm 0.0015
Fig. 6f	0.91 \pm 0.001	0.914 \pm 0.0008
Mean	0.809 \pm 0.002	0.804 \pm 0.0011

Table 4 shows detailed results of inpainting for learned overcomplete bases in term of the SSIM values, while Table 5 shows corresponding PSNR values. Both bases were twice overcomplete, i.e. of the size 256×512 . Again, for comparison, in Table 6 we also show results obtained with bases learned on patches with the size of 8×8 pixels, wherein number of atoms used in K-SVD training was 4. It can be seen that smaller patches and smaller number of atoms in the K-SVD training phase did not bring performance improvement. It can be seen from the Tables 4 and 5 that the ICA basis again outperformed the K-SVD learned basis. It is also clear that the use of overcomplete bases did not make significant performance improvement with respect to the complete case, which seems to be consistent with the conclusion already presented in [32] in the speech coding problem.

It should be said that the K-SVD algorithm is designed to minimize the mean squared error (MSE) of the representation, and thus does not necessarily give good performance in terms of the SSIM. The authors in [34] themselves noted that using

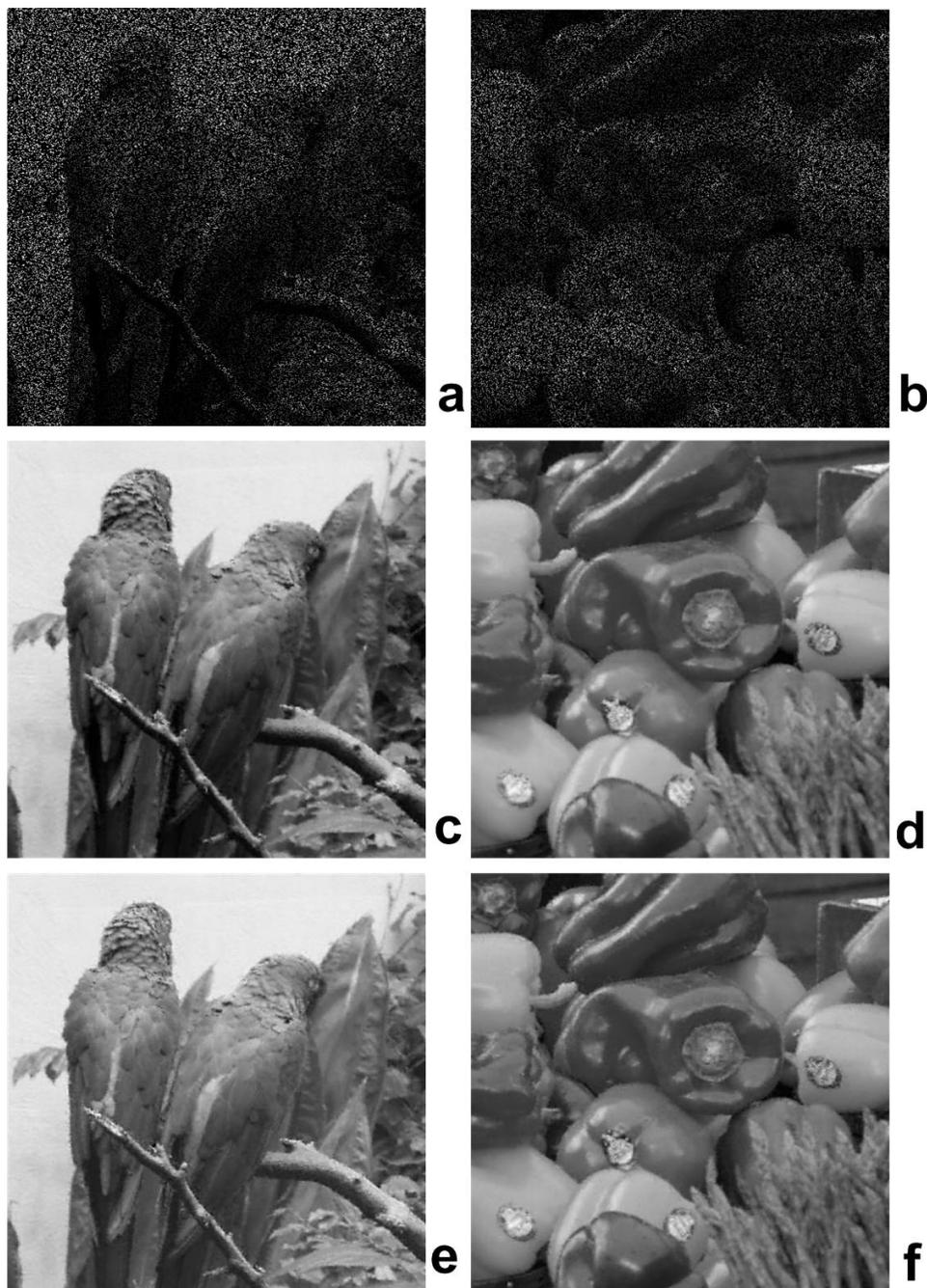


FIGURE 7. Examples of two images from the validation set with 80 percent of missing pixels: a) and b). Inpainting of two degraded images using ICA learned basis: c) and d). Inpainting of two degraded images using K-SVD learned basis: e) and f).

TABLE 4. Inpainting results in terms of the SSIM metric for the overcomplete bases learned on patches of size 16×16 pixels.

	ICA	K-SVD
Fig. 6a	0.907 ± 0.0005	0.903 ± 0.0011
Fig. 6b	0.76 ± 0.0008	0.754 ± 0.0011
Fig. 6c	0.773 ± 0.0011	0.772 ± 0.001
Fig. 6d	0.944 ± 0.0004	0.94 ± 0.0004
Fig. 6e	0.6 ± 0.0009	0.596 ± 0.0014
Fig. 6f	0.919 ± 0.0005	0.918 ± 0.0005
Mean	0.817 ± 0.0007	0.814 ± 0.0009

TABLE 5. Inpainting results in terms of the PSNR metric for the overcomplete bases learned on patches of size 16×16 pixels. The values are in dB.

	ICA	K-SVD
Fig. 6a	30.2 ± 0.06	30 ± 0.1
Fig. 6b	24.5 ± 0.03	24.3 ± 0.03
Fig. 6c	29.7 ± 0.03	29.6 ± 0.02
Fig. 6d	34.4 ± 0.06	34.1 ± 0.06
Fig. 6e	19.5 ± 0.03	19.3 ± 0.02
Fig. 6f	33.3 ± 0.09	33.2 ± 0.05
Mean	28.6 ± 0.05	28.4 ± 0.05

TABLE 6. Inpainting results in terms of the SSIM metric for the overcomplete bases learned on patches of size 8×8 pixels.

	ICA	K-SVD
Fig. 6a	0.899 ± 0.0006	0.896 ± 0.0013
Fig. 6b	0.745 ± 0.0013	0.74 ± 0.0011
Fig. 6c	0.765 ± 0.0007	0.765 ± 0.0011
Fig. 6d	0.936 ± 0.0006	0.931 ± 0.0004
Fig. 6e	0.593 ± 0.0016	0.59 ± 0.0017
Fig. 6f	0.915 ± 0.0006	0.915 ± 0.0004
Mean	0.809 ± 0.0009	0.807 ± 0.001

some metric other than the MSE in the K-SVD could be an interesting direction to study. However, such extensions of the K-SVD are not straightforward.

We also include results of the inpainting experiments with the stronger structure of the missing pixels, namely lines, blocks and text. We have modified the patch based algorithm presented above for the case when missing regions are larger than the used patch size: only patches that overlap with the missing region and have percentage of known pixels greater than the predefined threshold are inpainted; after the whole image is processed, this procedure is repeated. Similar iterative approaches were used in [23], [24]. Threshold of 0.8 was used in our experiments. Table 7 shows results of inpainting, using the ICA and K-SVD bases, where missing

pixels have block structure. The ICA basis again performed better compared to the K-SVD. Figure 8 shows two degraded and reconstructed images from the validation set, whereupon images are corrupted by the block pattern of missing pixels. It can be seen that the inpainted regions are blurry, but this is a known effect when large missing regions are being inpainted (for an example, see [14]). Figure 9 shows another, often used example. It should be noted that our heuristic approach in this case can not compete with the more specialized methods like [24], [40]. Table 8 shows results of inpainting of images corrupted by the line structures. Figure 10 shows two degraded and reconstructed images from the validation set, whereupon images are corrupted by the lines pattern of missing pixels. Figure 11 shows another often used example. We also show results of the text inpainting in Figure 12 (the images and the corresponding masks were taken from¹⁹). These examples should illustrate that ICA-based approach to basis learning performs reasonably well on realistic inpainting problems.

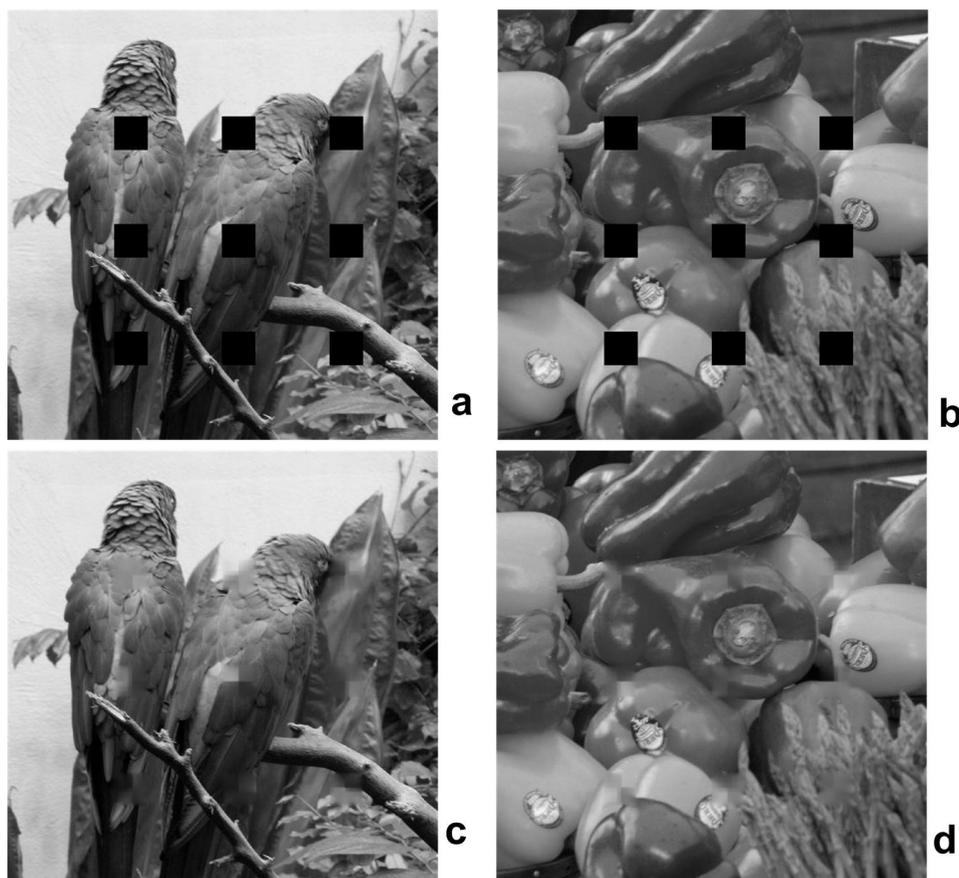


FIGURE 8. Examples of two images from the validation set with the block pattern of missing pixels: a) and b). Inpainting of two degraded images using ICA learned basis: c) and d).

¹⁹<http://www.dtic.upf.edu/~mbertalmio/restoration0.html>

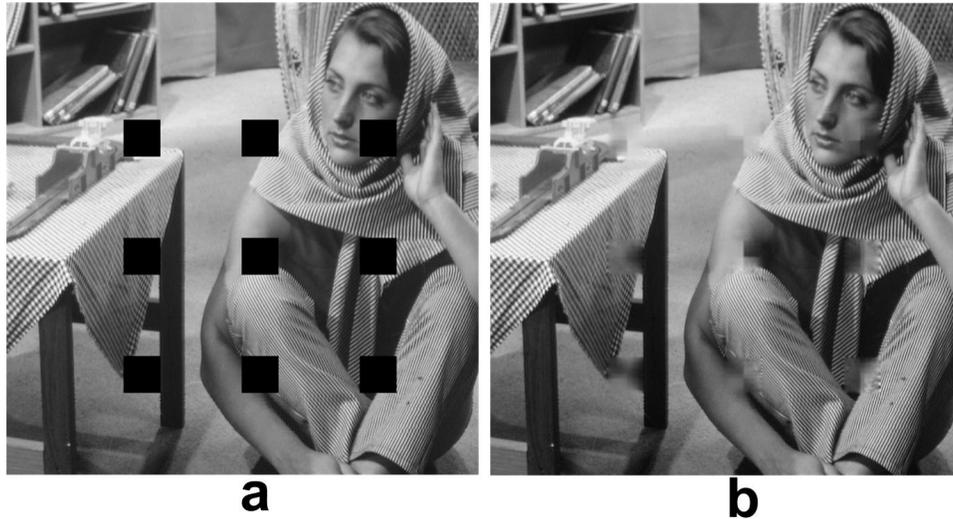


FIGURE 9. a) The Barbara image corrupted by the block structure of missing pixels. b) Inpainting using ICA learned basis.

TABLE 7. Inpainting results in terms of the SSIM metric for the block pattern of missing pixels.

	initial	ICA	K-SVD
Fig. 6a	0.9284	0.9746	0.963
Fig. 6b	0.9303	0.9658	0.9488
Fig. 6c	0.9254	0.9689	0.9554
Fig. 6d	0.9265	0.9775	0.97
Fig. 6e	0.9337	0.9545	0.942
Fig. 6f	0.9205	0.9885	0.98
Mean	0.927	0.9716	0.96

TABLE 8. Inpainting results in terms of the SSIM metric for the lines pattern of missing pixels.

	initial	ICA	K-SVD
Fig. 6a	0.939	0.997	0.9965
Fig. 6b	0.9439	0.9935	0.9887
Fig. 6c	0.935	0.9943	0.9913
Fig. 6d	0.935	0.9982	0.9983
Fig. 6e	0.952	0.9895	0.9825
Fig. 6f	0.924	0.9975	0.9964
Mean	0.938	0.995	0.9923

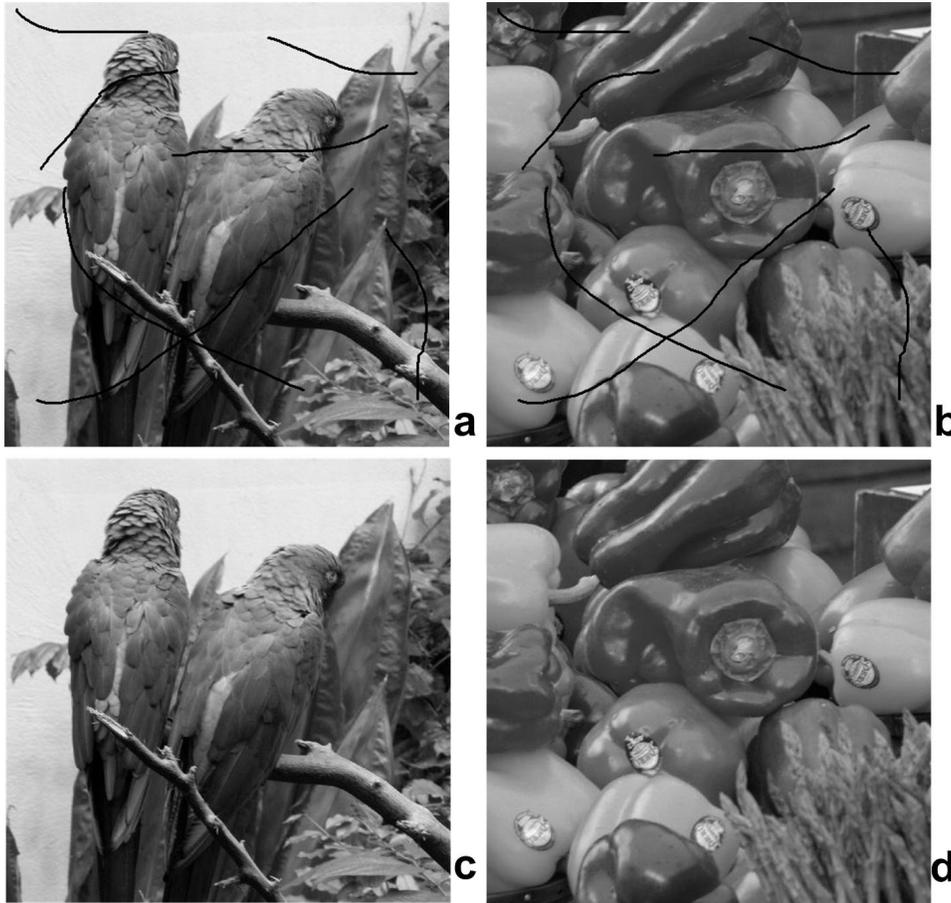


FIGURE 10. Examples of two images from the validation set with the lines pattern of missing pixels: a) and b). Inpainting of two degraded images using ICA learned basis: c) and d).

4.3. Denoising. As pointed out previously, presented approach to image inpainting can be used for image denoising purpose as well when image is corrupted by additive impulsive noise such as salt and pepper noise. All pixels with maximal intensity in the given resolution (when salt noise is considered) or with zero intensity (when pepper noise is considered) are declared as missing pixels. Due to the high efficiency of nonlinear image reconstruction methods and the learned basis that provides sparse representation of an image, the small amount of correct pixels that is possibly mistakenly declared as corrupted will not significantly influence the quality of denoising. Yet, such noise corrupted pixel detection scheme is very simple. Impulsive noise belongs to the class of alpha stable processes, has infinite variance and is optimally filtered out by means of myriad filters, [2], [21]²⁰. In the comparative performance analysis presented below we have used two-dimensional

²⁰Due to the infinite variance of impulsive additive noise, the MAP approach to image denoising [15], [34], that relies on solution of optimization problem with inequality constraint that depends on the noise variance, see eq.(2), is not applicable to the considered denoising scenario.

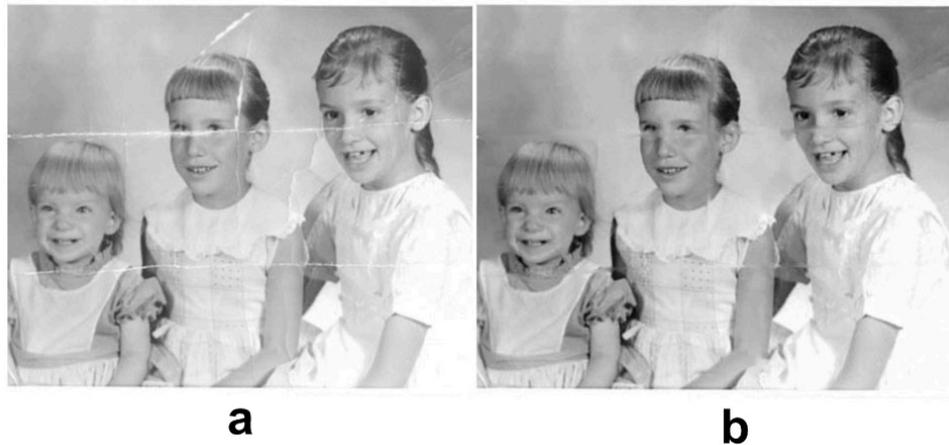


FIGURE 11. a) The Girls image. b) Inpainting using ICA learned basis.

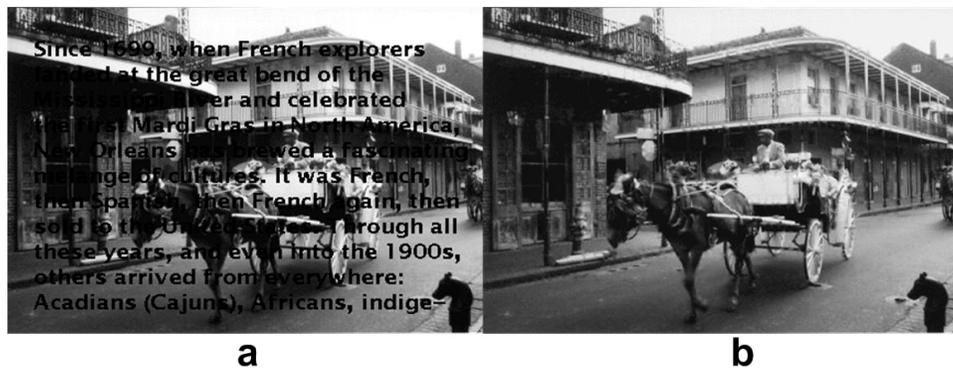


FIGURE 12. Inpainting for text removal. a) Image with text. b) Inpainting using ICA learned basis.

myriad filter with the sliding window of size 5×5 pixels. Larger window would filter out impulsive noise better, but would also cause a loss of details in filtered image. We refer the interested reader to ref. [2], page 337, for other details related to parameter settings for myriad filtering based denoising of gray scale images. Figures 13a and 13b show an image chosen from the validation set and corrupted with impulsive noise. Corrupted pixels were selected randomly to respectively occupy 5 and 20 percent of the image. Denoising results obtained by myriad filtering are respectively shown in Figures 13c and 13d, while denoising results obtained by inpainting in ICA learned basis are shown in Figures 13e and 13f. It can be seen that myriad filtering failed to denoise the image well, especially when 20 percent of the image pixels were corrupted. On the contrary, quality of visual perception of the image denoised through inpainting approach with the ICA learned basis is very good. Numerical results for all six images from the validation set are shown in Tables 9 and 10 for the corruption level of 5 and 20 percent respectively. It is evident that denoising using myriad filtering becomes very poor when higher percentage of

TABLE 9. Denoising results in terms of the SSIM metric when 5 percent of pixels were corrupted by impulsive noise.

	ICA	Myriad filtering
Fig. 6a	0.998	0.873
Fig. 6b	0.983	0.93
Fig. 6c	0.986	0.865
Fig. 6d	0.999	0.909
Fig. 6e	0.871	0.854
Fig. 6f	0.998	0.771
Mean	0.973	0.867

TABLE 10. Denoising results in terms of SSIM metric when 20 percent of pixels were corrupted by impulsive noise.

	ICA	Myriad filtering
Fig. 6a	0.991	0.643
Fig. 6b	0.977	0.691
Fig. 6c	0.974	0.569
Fig. 6d	0.997	0.674
Fig. 6e	0.916	0.743
Fig. 6f	0.993	0.287
Mean	0.975	0.601

the pixels is corrupted. On the contrary, denoising based on inpainting approach with learned basis remains robust even when as much as 80 percent of the pixels are corrupted, see Figure 7. Obtained results are explained by the fact that filtering is trying to smooth the image corrupted by additive impulsive noise while inpainting is based on nonlinear signal reconstruction approach whereupon pixels corrupted by additive noise are treated as missing pixels. Hence, provided that the learned basis yields sparse representation of the image, good reconstruction is possible even when a large number of pixels is corrupted. In the filtering approach, this requires a smoothing window with a large support that causes severe loss of details in the image.

5. **Conclusion.** This paper compared dictionary based approaches to inpainting and denoising of natural images, whereas the emphasis has been given on the ICA-learned complete and overcomplete basis. The paper also presented an algorithm for the reconstruction of the saturated pixels in the image corrupted by the salt and pepper noise, whereas saturated pixels are treated as missing pixels. Thus, image denoising problem has been virtually converted into the noiseless image inpainting problem. MAP approach can not solve this type of denoising problem due to the fact that the noise variance is infinite, while the filtering approach yields an image with degraded resolution. The FastICA algorithm with tanh nonlinear function has been used to learn the basis from patch-based representation of randomly chosen images of natural scenes that served as a training set. In the extensive comparative performance analysis in the inpainting problem of randomly chosen images of natural scenes from a validation set, ICA-learned bases outperformed bases learned

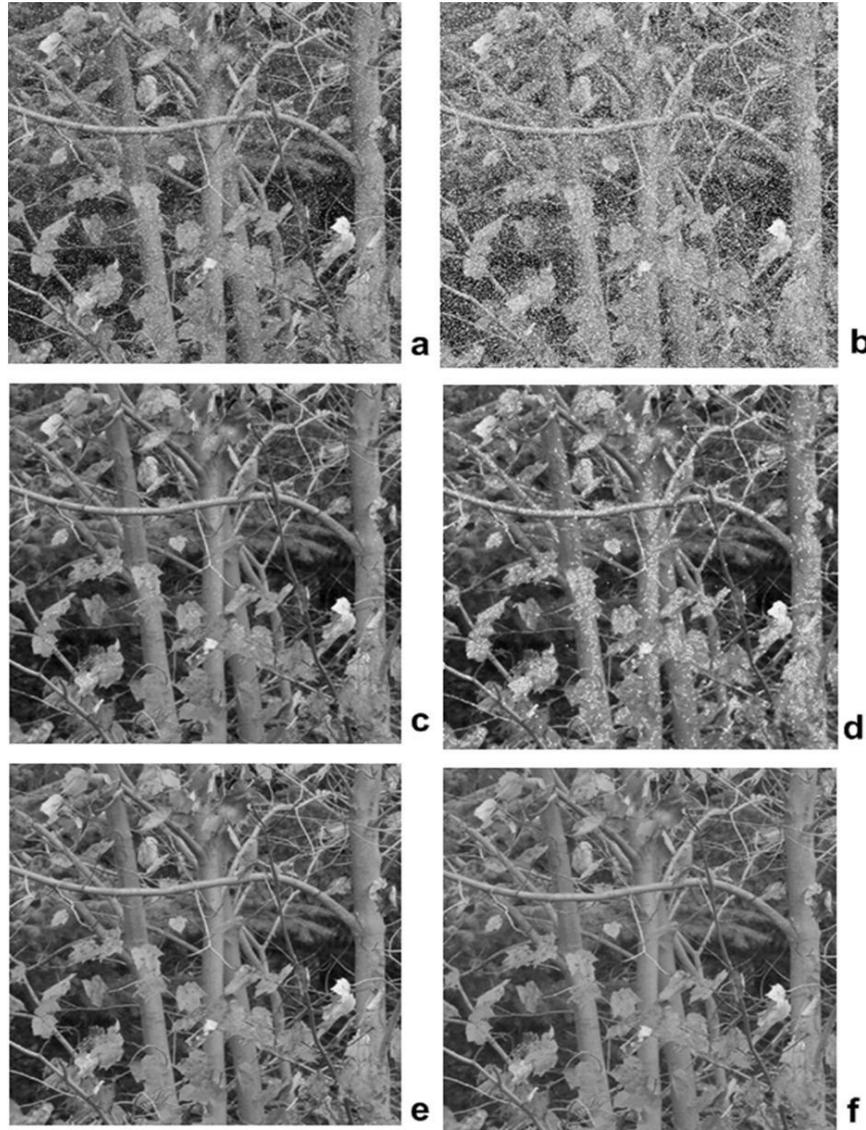


FIGURE 13. Examples of noisy image from the validation set with 5 and 20 percent of corrupted pixels, respectively: a) and b). Denoising using two-dimensional myriad filtering with the sliding window of size 5×5 pixels: c) and d). Denoising using inpainting and ICA learned basis: e) and f).

by K-SVD and MCA algorithm in term of quality of reconstruction. It yielded similar result as the FoE method with more than two orders of magnitude smaller computational complexity. On the same problems, Fourier and wavelet bases as representatives of fixed bases exhibited the poorest performance. It is also demonstrated that noiseless inpainting-based approach to image denoising (estimation of the saturated pixel values) greatly outperforms denoising based on two-dimensional

myriad filtering that is theoretically optimal solution for this class of additive impulsive noise.

Acknowledgments. This work was supported through grant 098-0982903-2558 funded by the Ministry of Science, Education and Sports, Republic of Croatia. We also thank the anonymous reviewers for their very constructive comments. Professor Vojislav Kecman's help in proofreading the manuscript is also gratefully acknowledged.

REFERENCES

- [1] M. Aharon, M. Elad and A. M. Bruckstein, *K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation*, IEEE Trans. Signal Process., **54** (2006), 4311–4322.
- [2] G. R. Arce, "Nonlinear Signal Processing - A Statistical Approach," John Wiley & Sons, 2005.
- [3] A. J. Bell and T. J. Sejnowski, *An information-maximization approach to blind separation and blind deconvolution*, Neural Comput., **7** (1995), 1129–1159.
- [4] A. J. Bell and T. J. Sejnowski, *The 'independent components' of natural scenes are edge filters*, Vision Research, **37** (1997), 3327–3338.
- [5] M. Bertalmio, G. Sapiro, V. Caselles and C. Ballester, *Image inpainting*, in "Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques," ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, (2000), 417–424.
- [6] M. Bertalmio, L. Vese, G. Sapiro and S. Osher, *Simultaneous structure and texture image inpainting*, IEEE Trans. Image Process., **12** (2003), 882–889.
- [7] M. Bethge, *Factorial coding of natural images: how effective are linear models in removing higher-order dependencies?*, J. Opt. Soc. Amer. A, **23** (2006), 1253–1268.
- [8] T. Blumensath and M. E. Davies, *Normalised iterative hard thresholding; guaranteed stability and performance*, IEEE J. Sel. Top. Signal Process., **4** (2010), 298–309.
- [9] P. Bofill and M. Zibulevsky, *Underdetermined blind source separation using sparse representations*, Signal Proc., **81** (2001), 2353–2362.
- [10] A. M. Bruckstein, M. Elad and D. L. Donoho, *From sparse solutions of systems of equations to sparse modeling of signals and images*, SIAM Rev., **51** (2009), 34–81.
- [11] E. Candès, M. B. Wakin and S. Boyd, *Enhancing sparsity by reweighted ℓ_1 minimization*, J. of Fourier Anal. Appl., **14** (2008), 877–905.
- [12] R. Chartrand, *Exact reconstructions of sparse signals via nonconvex minimization*, IEEE Signal Process. Lett., **14** (2007), 707–710.
- [13] S. Choi, A. Cichocki and S.-i. Amari, *Flexible independent component analysis*, J. VLSI Signal Process. Sys., **26** (2000), 25–38.
- [14] M. Elad, J.-L. Starck, P. Querre and D. L. Donoho, *Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)*, Appl. Comput. Harmon. Anal., **19** (2005), 340–358.
- [15] M. Elad and M. Aharon, *Image denoising via sparse and redundant representations over learned dictionaries*, IEEE Trans. Image Proc., **15** (2006), 3736–3745.
- [16] K. Engan, S. O. Aase and J. H. Husoy, *Multi-frame compression: Theory and design*, Signal Process., **80** (2000), 2121–2140.
- [17] K. Engan, K. Skretting and J. H. Husoy, *Family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation*, Dig. Signal Process., **17** (2007), 32–49.
- [18] D. Erdogmus, K. E. Hild II, Y. N. Rao and J. C. Principe, *Minimax mutual information approach for independent component analysis*, Neural Comput., **16** (2004), 1235–1252.
- [19] S. Foucart and M.-J. Lai, *Sparsest solution of underdetermined linear systems via ℓ_q minimization for $0 < q \leq 1$* , Appl. Comput. Harmon. Anal., **26** (2009), 395–407.
- [20] P. Georgiev, F. Theis and A. Cichocki, *Sparse component analysis and blind source separation of underdetermined mixtures*, IEEE Trans. Neural Netw., **16** (2005), 992–996.
- [21] J. G. Gonzalez and G. R. Arce, *Statistically-efficient filtering in impulsive environments: Weighted myriad filters*, EURASIP J. Appl. Signal Process., **1** (2002), 4–20.

- [22] I. F. Gorodnitsky and B. D. Rao, *Sparse signal reconstruction from limited data using FOCUSS: a re-weighted norm minimization algorithm*, IEEE Trans. Signal Process., **45** (1997), 600–616.
- [23] O. G. Guleryuz, *Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising: Part I - Theory*, IEEE Trans. Image Process., **15** (2006), 539–554.
- [24] O. G. Guleryuz, *Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising: Part II - Adaptive algorithms*, IEEE Trans. Image Process., **15** (2006), 555–571.
- [25] A. Hyvärinen, R. Cristescu and E. Oja, *A fast algorithm for estimating overcomplete ICA bases for image windows*, in “Proc. Int. Joint Conf. on Neural Networks,” Washington, D.C., USA, (1999), 894–899.
- [26] A. Hyvärinen, J. Karhunen and E. Oja, “Independent Component Analysis,” John Wiley & Sons, 2001.
- [27] A. Hyvärinen and E. Oja, *A fast fixed-point algorithm for independent component analysis*, Neural Comput., **9** (1997), 1483–1492.
- [28] J. Kovačević and A. Chebira, *Life beyond bases: The advent of frames (Part I)*, IEEE Signal Process. Mag., **25** (2007), 86–104.
- [29] J. Kovačević and A. Chebira, *Life beyond bases: The advent of frames (Part II)*, IEEE Signal Process. Mag., **25** (2007), 115–125.
- [30] K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T. W. Lee and T. J. Sejnowski, *Dictionary learning algorithms for sparse representation*, Neural Comput., **15** (2003), 349–396.
- [31] M. S. Lewicki and B. A. Olshausen, *Probabilistic framework for the adaptation and comparison of image codes*, J. Opt. Soc. Amer. A, **16** (1999), 1587–1601.
- [32] M. S. Lewicki and T. J. Sejnowski, *Learning overcomplete representations*, Neural Comput., **12** (2000), 337–365.
- [33] L. Ma and Y. Zhang, *Bayesian estimation of overcomplete independent feature subspaces for natural images*, in “Proceedings of the 7th International Conference on Independent Component Analysis and Signal Separation,” London, UK, (2007), 746–753. Available from: <http://portal.acm.org/citation.cfm?id=1776684.1776783>.
- [34] J. Mairal, G. Sapiro and M. Elad, *Sparse representation for color image restoration*, IEEE Trans. Image Process., **17** (2008), 53–69.
- [35] J. Mairal, F. Bach, J. Ponce and G. Sapiro, *Online learning for matrix factorization and sparse coding*, J. Mach. Learn. Res., **11** (2010), 19–60.
- [36] H. Mansour, R. Saab, P. Nasiopoulos and R. Ward, *Color image desaturation using sparse reconstruction*, in “Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing,” Dallas, TX, USA, (2010), 778–781.
- [37] H. Mohimani, M. Babaie-Zadeh and C. Jutten, *A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0 norm*, IEEE Trans. Signal Process., **57** (2009), 289–301.
- [38] B. A. Olshausen and D. J. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, Nature, **381** (1996), 607–609.
- [39] D. T. Pham and P. Garat, *Blind separation of mixtures of independent sources through a quasimaximum likelihood approach*, IEEE Trans. Signal Process., **45** (1997), 1712–1725.
- [40] S. Roth and M. J. Black, *Fields of experts*, Int. J. Computer Vision, **82** (2009), 205–229.
- [41] U. Schmidt, Q. Gao and S. Roth, *A generative perspective on Markov random fields in low-level vision*, in “Proc. Of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR),” San Francisco, CA, USA, (2010), 1751–1758.
- [42] I. W. Selesnick, R. V. Slyke and O. G. Guleryuz, *Pixel recovery via ℓ_1 minimization in wavelet domain*, in “Proc. IEEE Int. Conf. Image Process.,” Singapore, (2004), 1819–1822.
- [43] J. Shen, *Inpainting and the fundamental problem of image processing*, SIAM News, **36** (2003).
- [44] K. Skretting, J. H. Husoy and S. O. Aase, *General design algorithm for sparse frame expansions*, Signal Process., **86** (2006), 117–126.
- [45] J. Tropp and S. J. Wright, *Computational methods for sparse solution of linear inverse problems*, Proc. of the IEEE, **98** (2010), 948–958.
- [46] Z. Wang and A. Bovik, *Mean squared error: Love it or leave it? A new look at signal fidelity measures*, IEEE Signal Process. Mag., **26** (2009), 98–117.
- [47] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, *Image quality assessment: From error visibility to structural similarity*, IEEE Trans. Image Process., **13** (2004), 600–612.

- [48] X. Zhang and D. H. Brainard, *Estimation of saturated pixel values in digital color imaging*, J. Opt. Soc. Amer. A, **21** (2004), 2301–2310.
- [49] L. Zhang, A. Cichocki and S.-i. Amari, *Self-adaptive blind source separation based on activation function adaptation*, IEEE Trans. Neural Net., **15** (2004), 233–244.

Received February 2010; revised July 2011.

E-mail address: filipov@irb.hr

E-mail address: ikopriva@irb.hr