

A non-stationary image prior combination in super-resolution [☆]



S. Villena ^{a,*}, M. Vega ^a, R. Molina ^b, A.K. Katsaggelos ^c

^a Dept. de Lenguajes y Sistemas Informáticos, Universidad de Granada, 18071 Granada, Spain

^b Dept. de Ciencias de la Computación e I. A., Universidad de Granada, 18071 Granada, Spain

^c Dept. of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208-3118, United States

ARTICLE INFO

Article history:

Available online 6 June 2014

Keywords:

Super-resolution
Total variation
Variational methods
Parameter estimation
Bayesian methods

ABSTRACT

A new Bayesian *Super-Resolution* (SR) image registration and reconstruction method is proposed. The new method utilizes a prior distribution based on a general combination of spatially adaptive, or non-stationary, image filters, which includes an adaptive local strength parameter able to preserve both image edges and textures. With the application of variational techniques, the proposed method allows for the automatic estimation of all problem unknowns. An experimental comparison between state of the art methods and the proposed SR approach has been performed on both synthetic and real images.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Image SR is the post-processing image enhancement technique making it possible to infer a spatially *High Resolution* (HR) image of a scene, from multiple *Low Resolution* (LR) images affected by warping, blurring, and the noise inherent to the capture process. The basic principle of SR is that changes in LR images caused by the acquisition process and the camera (and/or scene) motion, provide additional information which can be utilized to reconstruct the HR image. Currently image SR is an active research field (see [1,2] for a review and [3] for a comparative study of recent SR methods).

Usually SR methods comprise two processing blocks: registration, where the motion between LR images, or an upsampled version of them, is estimated, and image reconstruction, where the HR image is recovered from the LR images. In this paper, both registration and reconstruction are studied within the Bayesian framework.

In the Bayesian framework a prior model on the HR image to be reconstructed has to be introduced, which allows to encapsulate prior image knowledge and to avoid the ill-posedness of the image reconstruction problem. The selection of this Bayesian prior model is critical. Prior models imposing image smoothness, like the *Simultaneous Auto Regressive* (SAR) image model (see [4]), are known to oversmooth edge regions. More sophisticated edge pre-

serving prior models based on wavelets [5,6], compound *Markov Random Fields* (MRF) [7,8], and sparse image priors as TV [9] or ℓ_1 [10] have also been applied. Some of these edge preserving priors oversmooth non-edge textured regions, and to prevent this, a combination of sparse and non-sparse prior models exploiting the ability of sparse priors to recover image edges while avoiding their tendency to oversmooth inner regions by combining them with a smoothness promoting prior model was proposed in [10]. However, since it is not straightforward to determine the optimal contribution of each prior to the mixture, the weights were determined experimentally.

Accurate registration of displaced and rotated images, is vital in SR image reconstruction. There are two major approaches to registration in SR, which differ in the stage where registration is performed. In the first approach the motion parameters are previously estimated from the observed LR images, in a preprocessing step, and then used in a separate image estimation process (see [11–14]). The limited accuracy inherent to HR registration from LR images is a shortcoming of this first approach. The second approach is to alternate between LR image registration and HR image estimation (see [15–20,7,9,21–23,8,10]).

In this paper, we extend the work in our previous conference paper [24], in which the application of a spatially adaptive general linear filter combination prior model to the SR problem was proposed. This kind of prior model includes an adaptive, local strength parameter, able to preserve both image edges and textures, and has been successfully applied to image restoration [25]. This paper is more than an enhanced version of our previous conference paper [24]. The algorithm proposed in this paper is the same as in [24], but both Bayesian model and inference have been reformulated in a clearer way, not relying on the representation, utilized in [24],

[☆] This work was supported in part by the Comisión Nacional de Ciencia, Tecnología under contract TIN2010-15137, microproyct cod.21 and proyct cod.25 CEI BioTic of University of Granada and by a grant from the Department of Energy (DE-NA0000457).

* Corresponding author.

E-mail addresses: svillena@ugr.es (S. Villena), mvega@ugr.es (M. Vega), rms@decsai.ugr.es (R. Molina), aggk@eecs.northwestern.edu (A.K. Katsaggelos).

of the prior model as a Scale Mixture of Gaussians. This paper differs from the previous paper [25], not only because in [25] the addressed problem was image deconvolution and denoising, while in this paper we are studying SR, but also because of the different Bayesian inference methods applied.

In our proposed approach, the entire SR pipeline, that is HR image reconstruction, registration, and the estimation of the model parameters, usually called hyperparameters, a critical issue in SR (see for instance [9,20,23,10]), is approached from a Bayesian perspective. The proposed framework provides uncertainties of the estimates during the restoration process, which helps to prevent error-propagation and improves robustness. All required algorithmic parameters are estimated along with the HR image and the motion parameters, and therefore the proposed algorithms do not require user supervision. This is a very important characteristic of our method, marking the difference with respect to the method in [10], which used a different prior models combination, whose relative weights had to be determined by a long time consuming experimental procedure.

The rest of this paper is organized as follows. Section 2 provides the mathematical model for the LR image acquisition process. We provide the description of the hierarchical Bayesian framework modeling the unknowns in Section 3. The inference procedure to develop the proposed methods is presented in Section 4. We demonstrate the effectiveness of the proposed methods with experimental results in Section 5 and conclusions are drawn in Section 6.

2. Problem formulation

In this paper we study the image SR problem, i.e.: the reconstruction of an HR image \mathbf{x} of a scene, from a sequence of L LR observed images \mathbf{y}_k , $k = 1, \dots, L$, of the same scene. The LR images \mathbf{y}_k consist of $N = N_h \times N_v$ pixels (where N_h and N_v are the observations pixel numbers in horizontal and vertical directions, respectively) while the HR image \mathbf{x} consists of PN pixels, where $\sqrt{P} \in \mathbb{N}$ is the factor of increase in resolution. In this paper we adopt the matrix-vector notation such that images \mathbf{y}_k and \mathbf{x} are arranged as $N \times 1$ and $PN \times 1$ column vectors, respectively. The imaging process introduces shifting, blurring and downsampling, which is modeled as

$$\mathbf{y}_k = \mathbf{A}\mathbf{H}_k\mathbf{C}(\mathbf{s}_k)\mathbf{x} + \mathbf{n}_k = \mathbf{B}_k(\mathbf{s}_k)\mathbf{x} + \mathbf{n}_k, \quad (1)$$

with the system $N \times PN$ matrix $\mathbf{B}_k = \mathbf{A}\mathbf{H}_k\mathbf{C}(\mathbf{s}_k)$, where \mathbf{A} is the $N \times PN$ downsampling matrix, \mathbf{H}_k is the $PN \times PN$ blurring matrix, $\mathbf{C}(\mathbf{s}_k)$ is the $PN \times PN$ warping matrix generated by the motion vector \mathbf{s}_k , and \mathbf{n}_k is the $N \times 1$ acquisition noise. A detailed description of the explicit form of the warping matrices $\mathbf{C}(\mathbf{s}_k)$ in Eq. (1) can be found in [10]. Note that the matrices \mathbf{H}_k and $\mathbf{C}(\mathbf{s}_k)$ and the noise \mathbf{n}_k can be different for each LR image \mathbf{y}_k .

In this work, we assume that the blurring matrices \mathbf{H}_k are known and we consider a motion model consisting of translational and rotational motion, so that $\mathbf{s}_k = (\theta_k, c_k, d_k)^t$, where θ_k is the rotation angle, and c_k and d_k are the horizontal and vertical translations of the k th HR image with respect to the reference frame \mathbf{x} . We also assume that the observation set $\{\mathbf{y}_k\}$ and HR image \mathbf{x} are normalized between zero and one.

The effects of downsampling, blurring, and warping are combined into the system matrix $\mathbf{B}_k(\mathbf{s}_k)$, from which each row maps the pixels of the HR image \mathbf{x} to a given pixel in the LR image \mathbf{y}_k . Given Eq. (1), the SR problem is expressed as the search of an estimate of the HR image \mathbf{x} from the set of LR images $\{\mathbf{y}_k\}$ using our prior knowledge about $\{\mathbf{C}(\mathbf{s}_k)\}$, $\{\mathbf{n}_k\}$, and \mathbf{x} .

3. Hierarchical Bayesian model

It is well known the ill-posed nature of the SR problem, that has been traditionally circumvented by mean of regularization terms in the optimization approach, or using prior distributions in the Bayesian approach (see [1]). In this work, we adopt a hierarchical Bayesian framework consisting of two stages. The first stage is used to model the acquisition process, the unknown HR image \mathbf{x} and the motion vectors $\{\mathbf{s}_k\}$. For the unknown \mathbf{x} we adopt a general non-stationary image prior combination. Prior distributions $p(\mathbf{s}_k)$ are assigned to the unknowns motion parameters \mathbf{s}_k , for $k = 1, \dots, L$. The observation $\mathbf{y} = \{\mathbf{y}_k\}$ is also a random process with corresponding conditional distribution $p(\mathbf{y}|\mathbf{x}, \{\mathbf{s}_k\}, \{\beta_k\})$. This distribution depends on a set of additional parameters (called *hyperparameters*). They are modeled together with the hyperparameters of the distributions of \mathbf{x} and $\{\mathbf{s}_k\}$ by assigning *hyperprior* distributions in the second stage of the hierarchical model.

In the following subsections we provide the description of the individual distributions used to model the unknowns.

3.1. Observation model

Using the model of Eq. (1) and assuming that \mathbf{n}_k is zero-mean white Gaussian noise with inverse variance (precision) β_k , the conditional distribution of the LR image \mathbf{y}_k is given by

$$p(\mathbf{y}_k|\mathbf{x}, \mathbf{s}_k, \beta_k) \propto \beta_k^{N/2} \exp\left[-\frac{\beta_k}{2} \|\mathbf{y}_k - \mathbf{B}_k(\mathbf{s}_k)\mathbf{x}\|^2\right]. \quad (2)$$

Assuming statistical independence of the noise among the LR image acquisitions, the conditional probability of the set of LR images \mathbf{y} given \mathbf{x} can be expressed as

$$p(\mathbf{y}|\mathbf{x}, \{\mathbf{s}_k\}, \{\beta_k\}) = \prod_{k=1}^L p(\mathbf{y}_k|\mathbf{x}, \mathbf{s}_k, \beta_k). \quad (3)$$

The independent Gaussian model in Eq. (3) is used in most of the existing super-resolution methods [26,17,20,7,8].

3.2. A non-stationary image prior combination

In this paper the following prior distribution, on the unknown HR \mathbf{x} image, will be used

$$p(\mathbf{x}|\boldsymbol{\eta}) = \prod_{j=1}^d p(\mathbf{z}_j|\boldsymbol{\eta}_j) = \prod_{j=1}^d \prod_{i=1}^{PN} p(z_j(i)|\eta_j(i)), \quad (4)$$

with

$$p(z_j(i)|\eta_j(i)) \propto \sqrt{\eta_j(i)} \exp\left(-\frac{1}{2}\eta_j(i)\|z_j(i)\|^2\right). \quad (5)$$

In (4) $\mathbf{z} = \{\mathbf{z}_1, \dots, \mathbf{z}_d\}$ is the set of d unknown filtered images $\mathbf{z}_j = \mathbf{F}_j\mathbf{x}$, with \mathbf{F}_j being convolution operators and $z_j(i)$ the i component of the \mathbf{z}_j filtered image vector. As previously stated, our prior model takes the form of a general combination of d linear filters. We can use for \mathbf{F}_j *first order difference* (f.o.d.) operators, in different directions, although we are not limited to it.

In (4) $\boldsymbol{\eta} = \{\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_d\}$ is the set of unknown prior parameter vectors $\boldsymbol{\eta}_j$, which has to be estimated, and $\eta_j(i)$ is the i component of $\boldsymbol{\eta}_j$. Through the presence of the prior parameter $\eta_j(i)$ in Eq. (5), our prior model becomes spatially adaptive. Notice that in Eq. (4) we are approximating the partition function as an independent product of partition functions.

3.3. Modeling the registration parameters

In this paper the uncertainties in the registration parameters are modeled as in our previous paper [10]. We denote by $\{\bar{\mathbf{s}}^p\}$ the estimate of $\{\mathbf{s}_k\} = \{\mathbf{s}_1, \dots, \mathbf{s}_L\}$ obtained from LR observations in a preprocessing step, using registration algorithms, such as the ones reported in [27]. As these estimates are in general inaccurate, we model the motion parameters as stochastic variables following Gaussian distributions with *a priori* means the preliminary motion parameters $\bar{\mathbf{s}}_k^p$, that is,

$$p(\{\mathbf{s}_k\}) = \prod_{k=1}^L \mathcal{N}(\mathbf{s}_k | \bar{\mathbf{s}}_k^p, \Xi_k^p), \quad (6)$$

with Ξ_k^p the *a priori* covariance matrix. The parameters $\bar{\mathbf{s}}_k^p$ and Ξ_k^p incorporate prior knowledge about the motion parameters into the estimation procedure. If such knowledge is not available, $\bar{\mathbf{s}}_k^p$ and $(\Xi_k^p)^{-1}$ can be set to zero, which makes the observations solely responsible for the estimation process.

3.4. Hyperpriors on the hyperparameters

In this paper we assume flat hyperpriors for the $\eta_j(i)$ hyperparameters in Eq. (5). That is, $p(\eta_j(i)) \propto \text{const}$. For modeling the $\{\beta_k\}$ hyperparameters in Eq. (3), we employ Gamma distributions

$$p(\{\beta_k\}) = \prod_{k=1}^L \Gamma(\beta_k | a_{\beta_k}^0, b_{\beta_k}^0), \quad (7)$$

where $a_{\beta_k}^0 > 0$ and $b_{\beta_k}^0 > 0$ are the shape and scale parameters, respectively. The hyperpriors are chosen as Gamma distributions since they are conjugate priors for the Gaussian distribution.

3.5. Joint model

Combining Eqs. (3), (4), (6) and (7) we obtain the following joint probability distribution

$$p(\Theta, \mathbf{y}) = p(\mathbf{y} | \mathbf{x}, \{\mathbf{s}_k\}, \{\beta_k\}) p(\mathbf{x} | \boldsymbol{\eta}) p(\{\beta_k\}) p(\{\mathbf{s}_k\}) p(\boldsymbol{\eta}), \quad (8)$$

where $\Theta = \{\boldsymbol{\eta}, \mathbf{x}, \{\mathbf{s}_k\}, \{\beta_k\}\}$ denote the set of all unknowns.

4. Variational Bayesian inference

Bayesian inference is based on the posterior distribution $p(\Theta | \mathbf{y}) = \frac{p(\Theta, \mathbf{y})}{p(\mathbf{y})}$. As $p(\mathbf{y})$ cannot be obtained, we approximate $p(\Theta | \mathbf{y})$ by the $q(\Theta)$ distribution using the Kullback–Leibler divergence. This is the well known variational approximation, which is very well described in [28] (see also [29,7,9]). Within the mean field approximation, $q(\Theta)$ is approximated by the distribution $q(\Theta) = \prod_{\zeta \in \Theta} q(\zeta)$. For each $\zeta \in \Theta$ where we assume that $q(\zeta)$ is a degenerate distribution, its value is obtained by calculating

$$\hat{\zeta} = \underset{\zeta}{\text{argmax}} \log q(\zeta) = \underset{\zeta}{\text{argmax}} \langle \log p(\Theta, \mathbf{y}) \rangle_{\Theta_\zeta}, \quad (9)$$

where Θ_ζ denotes the set Θ with ζ removed, and $E_{q(\Theta_\zeta)}[\cdot] = \langle \cdot \rangle_{\Theta_\zeta}$. In the following, the subscript of the expected value will be removed when it will be clear from the context. For non-degenerate distributions we have

$$q(\zeta) \propto \exp(\langle \log p(\Theta, \mathbf{y}) \rangle_{\Theta_\zeta}). \quad (10)$$

We assume a degenerate distribution only for $\eta_j(i)$, $j = 1, \dots, d$, $i = 1, \dots, PN$. From Eq. (10), we obtain for $q(\mathbf{x})$

$$q(\mathbf{x}) \propto \exp\left\{\langle \log(p(\{\mathbf{y}_k\} | \mathbf{x}, \{\mathbf{s}_k\}, \{\beta_k\})) \rangle_{\{\mathbf{s}_k, \beta_k\}} + \log(p(\mathbf{x} | \hat{\boldsymbol{\eta}}))\right\}, \quad (11)$$

which is the multivariate Gaussian

$$q(\mathbf{x}) = \mathcal{N}(\mathbf{x} | \hat{\mathbf{x}}, \Xi_{\mathbf{x}}), \quad (12)$$

with

$$\Xi_{\mathbf{x}}^{-1} = \sum_{j=1}^d \mathbf{F}_j^t \text{diag}(\hat{\boldsymbol{\eta}}_j) \mathbf{F}_j + \sum_{k=1}^L \langle \beta_k \rangle \langle \mathbf{B}(\mathbf{s}_k)^t \mathbf{B}(\mathbf{s}_k) \rangle_{\mathbf{s}_k}, \quad (13)$$

and

$$\hat{\mathbf{x}} = \Xi_{\mathbf{x}} \sum_{k=1}^L \langle \beta_k \rangle \langle \mathbf{B}(\mathbf{s}_k) \rangle_{\mathbf{s}_k}^t \mathbf{y}_k, \quad (14)$$

this Eq. (14) can be solved iteratively utilizing a *Conjugate Gradient* method.

Also, from Eq. (10), we find the following distribution for the registration parameters

$$q(\mathbf{s}_k) \propto \exp\left[-\frac{1}{2} \langle \beta_k \rangle \langle \|\mathbf{y}_k - \mathbf{B}_k(\mathbf{s}_k) \mathbf{x}\|^2 \rangle_{\mathbf{x}} + (\mathbf{s}_k - \bar{\mathbf{s}}_k^p)^t (\Xi_k^p)^{-1} (\mathbf{s}_k - \bar{\mathbf{s}}_k^p)\right]. \quad (15)$$

The explicit form of the distribution $q(\mathbf{x})$ in Eq. (12) depends on the expectation values $\langle \mathbf{B}(\mathbf{s}_k)^t \mathbf{B}(\mathbf{s}_k) \rangle_{\mathbf{s}_k}$ and $\langle \mathbf{B}(\mathbf{s}_k) \rangle_{\mathbf{s}_k}$, and $q(\mathbf{s}_k)$ in Eq. (15) depends on $\langle \|\mathbf{y}_k - \mathbf{B}_k(\mathbf{s}_k) \mathbf{x}\|^2 \rangle_{\mathbf{x}}$. These calculations are not straightforward since $\mathbf{C}(\mathbf{s}_k)$, in Eq. (1), is nonlinear with respect to \mathbf{s}_k . Therefore, we expand $\mathbf{C}(\mathbf{s}_k)$ using its first-order Taylor series around the mean value $\bar{\mathbf{s}}_k = \langle \mathbf{s}_k \rangle = (\bar{\theta}_k, \bar{c}_k, \bar{d}_k)^T$ of the distribution $q(\mathbf{s}_k)$, in Eq. (15); details can be found in [10].

4.1. Estimation of the hyperparameter distributions

Using Eq. (9), the following values are obtained for the components of $\hat{\boldsymbol{\eta}}_j$ in Eq. (13),

$$\hat{\eta}_j(i) = \frac{1}{\langle z_j^2(i) \rangle_{\mathbf{x}}}, \quad (16)$$

with

$$\langle z_j^2(i) \rangle_{\mathbf{x}} = \hat{\mathbf{x}}^t \mathbf{F}_j^t \mathbf{J}^{ii} \mathbf{F}_j \hat{\mathbf{x}} + \text{tr}(\Xi_{\mathbf{x}} \mathbf{F}_j^t \mathbf{J}^{ii} \mathbf{F}_j), \quad (17)$$

where \mathbf{J}^{ii} is a single-entry $PN \times PN$ matrix with zeros everywhere except at the entry (i, i) , which is equal to one. In this paper, $\Xi_{\mathbf{x}}$ in Eq. (17) is calculated by applying the Jacobi approximation.

Finally we obtain the distributions for the hyperparameters $\{\beta_k\}$, which are found to be Gamma distributions. For the $\{\beta_k\}$ hyperparameters, using Eq. (10), we obtain

$$q(\beta_k) \propto \beta_k^{N/2 - 1 + a_{\beta_k}^0} \exp\left[-\beta_k \left(b_{\beta_k}^0 + \frac{\langle \|\mathbf{y}_k - \mathbf{B}_k(\mathbf{s}_k) \mathbf{x}\|^2 \rangle_{\mathbf{x}, \mathbf{s}_k}}{2}\right)\right] \quad (18)$$

with

$$\langle \beta_k \rangle = \frac{N + 2a_{\beta_k}^0}{\|\mathbf{y}_k \mathbf{B}(\mathbf{s}_k) \hat{\mathbf{x}}\|^2 + \text{tr}(\Xi_{\mathbf{x}} \mathbf{B}^t(\mathbf{s}_k) \mathbf{B}(\mathbf{s}_k)) + 2b_{\beta_k}^0}. \quad (19)$$

We summarize below the proposed iterative SR [Algorithm 1](#), which comprises the estimation of the HR image, registration parameters, HR prior parameters and model hyperparameters:

Algorithm 1 Variational Bayesian super-resolution.

Require: Values $\{a_{\beta_k}^0\}$, $\{b_{\beta_k}^0\}$, $\{\hat{s}_k^p\}$, $\{\Xi_k^p\}$ and initial HR image value $\hat{\mathbf{x}}^{(0)}$.

Set $n = 1$, $q^{(0)}(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\mathbf{x}^{(0)}, \mathbf{0})$ and $q^{(0)}(\{\mathbf{s}_k\}) = p(\{\mathbf{s}_k\})$, with $p(\{\mathbf{s}_k\})$ given in Eq. (6).

while convergence criterion is not met **do**

1. Given $q^{(n-1)}(\mathbf{x})$ and $q^{(n-1)}(\{\mathbf{s}_k\})$, obtain $q^{(n)}(\{\beta_k\})$ using Eq. (18).
 2. Given $q^{(n-1)}(\mathbf{x})$ and $q^{(n)}(\{\beta_k\})$, obtain $q^{(n)}(\{\mathbf{s}_k\})$ using Eq. (15).
 3. Given $q^{(n-1)}(\mathbf{x})$, compute $\hat{\eta}^{(n)}$ using Eq. (16).
 4. Given $q^{(n)}(\{\beta_k\})$, $q^{(n)}(\{\mathbf{s}_k\})$ and $\hat{\eta}^{(n)}$, obtain $q^{(n)}(\mathbf{x})$ using Eq. (12).
-

5. Experimental results

A number of experiments, both on simulated and real images, has been carried out in order to assess the performance of the new proposed method and compare it with: 1) bicubic interpolation (denoted by *BBC*), 2) the SR method in [12] (denoted by *ZMT*), which is based on backprojection with median filtering, 3) the robust SR method in [13] (denoted by *RSR*), which is based on bilateral TV priors, 4) the variational SR method using a TV prior in [9] (denoted by *TV*), 5) the variational SR method in [10] based on a combination of ℓ_1 and SAR priors (denoted by *LTS*).

In all experiments reported below, the observed images have been normalized to the interval $[0, 1]$. The initial HR image value $\hat{\mathbf{x}}^{(0)}$ in Algorithm 1 has been obtained from the bicubic interpolation of observation \mathbf{y}_1 . The inverse covariance matrices (Ξ_k^p)⁻¹ are set to zero, that is, no prior information is utilized on the uncertainty of the motion vectors. Setting $a_{\beta_k}^0 = 0$ and $b_{\beta_k}^0 = 0$ in Eq. (19), we have used non-informative prior for β_k in all experiments. The rest of the algorithm parameters are automatically calculated using the algorithmic steps in Algorithm 1. As convergence criterion we used $\|\hat{\mathbf{x}}^{(n)} - \hat{\mathbf{x}}^{(n-1)}\|^2 / \|\hat{\mathbf{x}}^{(n-1)}\|^2 < 10^{-5}$, where $\hat{\mathbf{x}}^{(n)}$ and $\hat{\mathbf{x}}^{(n-1)}$ are the HR image estimates at iterations n and $(n - 1)$, respectively. In this paper, color LR images are first converted to grayscale, then Algorithm 1 is applied to them and the obtained motion parameter values $\{\hat{s}_k\}$ retained. Next Algorithm 1 is separately applied to each color channel using those motion parameters, without executing the registration step 2.

5.1. Experiments with simulated images

Let us first present the experimental results obtained using simulated images. Synthetic sequences with five LR images have been generated from the 132×132 images shown in Fig. 1, through warping, blurring and downsampling by a factor $\sqrt{P} = 2$. The warping of the images in the sequences consisted of translations of $(0, 0)^t$, $(0, 0.5)^t$, $(0.5, 0)^t$, $(1, 0)^t$ and $(0, 1)^t$ pixels respectively, and rotations of 0° , 3° , -3° , 5° and -5° . A 3×3 uniform PSF has been used for blurring. The LR images obtained after warping, blurring and downsampling operations are further degraded by additive white Gaussian noise at SNR levels of 10 dB, 15 dB, 20 dB, 25 dB and 30 dB. At each SNR level, ten noise realizations per sequence have been obtained. The quality of the reconstructed HR images has been quantitatively measured in terms of the *Peak Signal-to-Noise Ratio* (PSNR) and the *Structural Similarity Index Measure* (SSIM) [30]. The SSIM index measures visual similarity between images, while PSNR measures denoising effectiveness.

The proposed method uses the prior distribution on the HR image in Eq. (4), which is based on a general combination of image filters $\{\mathbf{F}_j\}$. We have used the following *first order difference* (f.o.d.) filters: horizontal ($\mathbf{f1}$), vertical ($\mathbf{f2}$), upper-right diagonal ($\mathbf{f3}$), and lower-right diagonal ($\mathbf{f4}$). We have also considered the second order Laplacian filter ($\mathbf{f5}$), and the following *second order difference* (s.o.d.) filters: horizontal ($\mathbf{f6}$), vertical ($\mathbf{f7}$), upper-right diagonal ($\mathbf{f8}$), and lower-right diagonal ($\mathbf{f9}$). Finally, the first order Sobel horizontal ($\mathbf{f10}$), and vertical ($\mathbf{f11}$), and the Prewitt horizontal

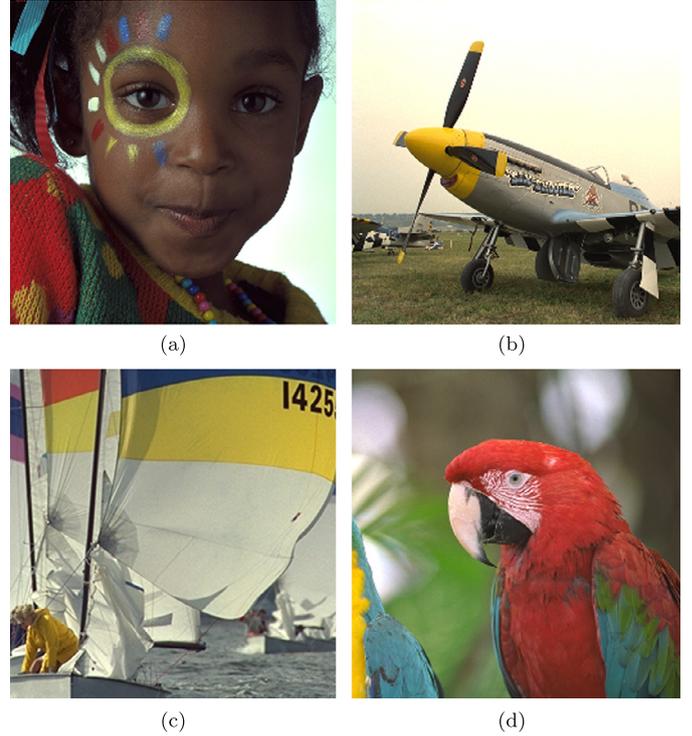


Fig. 1. Images used in the synthetic experiments.

($\mathbf{f12}$), and vertical ($\mathbf{f13}$) filters, have also been considered. We illustrate the observed dependence of the method performance on the selected combination of image filters $\{\mathbf{F}_j\}$, by presenting next results for the image in Fig. 1(c).

Fig. 2 shows a quantitative comparison in terms of grayscale PSNR and SSIM, of the reconstructions of the image in Fig. 1(c), at different noise levels, obtained using different first order filter combinations. The filter combinations compared in Fig. 2, are coupled by considering the distance between the positive and negative values of the filters: distance 1 1) $\{\mathbf{f1}, \mathbf{f2}\}$, distance $\sqrt{2}$ 2) $\{\mathbf{f3}, \mathbf{f4}\}$, distance 2 3) $\{\mathbf{f10}, \mathbf{f11}\}$, and also the filters 4) $\{\mathbf{f12}, \mathbf{f13}\}$. It can be observed in Fig. 2, how the method performance degrades as the distance between the values of filters increases. Fig. 3 shows a similar comparison for second order filters, with distances between positive and negative values of the corresponding first order filter 1 pixel for 1) $\{\mathbf{f6}, \mathbf{f7}\}$, $\sqrt{2}$ for 2) $\{\mathbf{f8}, \mathbf{f9}\}$, and 3) $\{\mathbf{f5}\}$. The performance of the proposed method when using second order filters also decreases as the distance between positive and negative values increases.

The results shown in Figs. 2 and 3 reveal that the proposed method performs better for the first order filters than for the second order ones. Fig. 4 shows a comparison of the results obtained using first order–second order combinations. It can be concluded from a comparison between Figs. 2, 3, and 4, that combining first and second filters slightly deteriorates the performance in PSNR terms, while slightly improving SSIM performance. A detailed observation of the distances between compared pixels of the filters combinations in Fig. 4, also shows that when combining first and second order filters, performance degrades as distance increases, and that the distances corresponding to the first order filters have more influence on the method performance, than the corresponding to the second order ones.

Fig. 5 shows a quantitative comparison in terms of grayscale PSNR and SSIM, of the reconstructions of the images in Fig. 1 at different noise levels, obtained using the following filter combinations: 1) $NF2 = \{\mathbf{f1}, \mathbf{f2}\}$, 2) $NF3 = \{\mathbf{f1}, \mathbf{f2}, \mathbf{f5}\}$, 3) $NF4 = \{\mathbf{f1}, \mathbf{f2}, \mathbf{f3}, \mathbf{f4}\}$ and 4) $NF5 = \{\mathbf{f1}, \mathbf{f2}, \mathbf{f3}, \mathbf{f4}, \mathbf{f5}\}$.

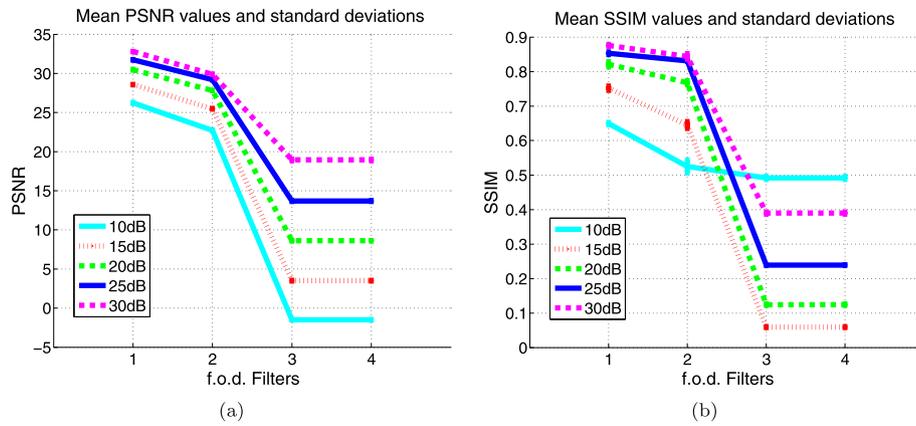


Fig. 2. (a) PSNR and (b) SSIM for the reconstruction of the image in Fig. 1(c) with our proposed method, using different first order filters combinations: 1) $\{f1, f2\}$, 2) $\{f3, f4\}$, 3) $\{f10, f11\}$, and 4) $\{f12, f13\}$.

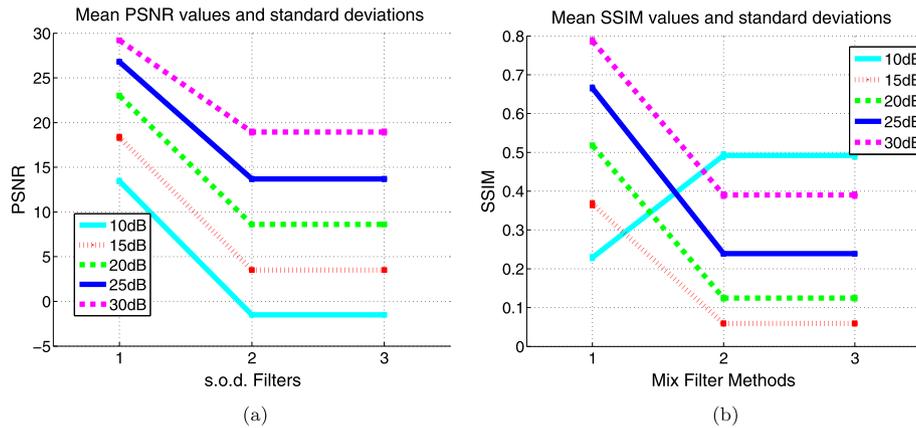


Fig. 3. (a) PSNR and (b) SSIM for the reconstruction of the image in Fig. 1(c) with our proposed method, using different second order filters combinations: 1) $\{f6, f7\}$, 2) $\{f8, f9\}$, and 3) $\{f5\}$.

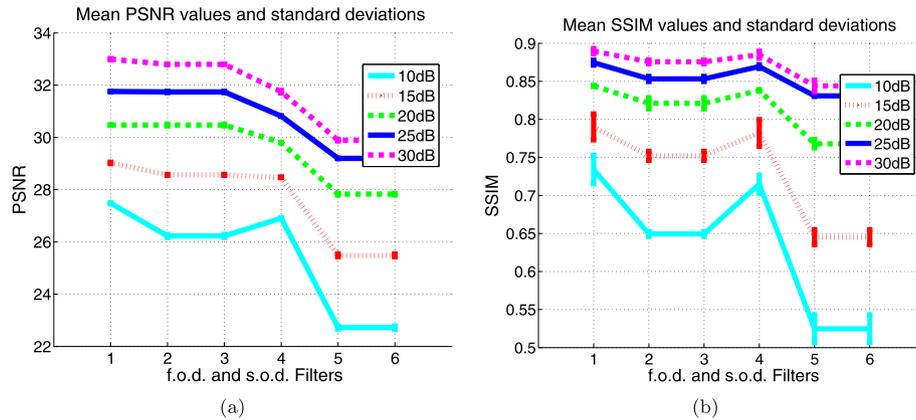


Fig. 4. (a) PSNR and (b) SSIM for the reconstruction of the image in Fig. 1(c) with our proposed method, using different first order and second order filters combinations: 1) $\{f1, f2, f6, f7\}$, 2) $\{f1, f2, f8, f9\}$, 3) $\{f1, f2, f5\}$, 4) $\{f3, f4, f6, f7\}$, 5) $\{f3, f4, f8, f9\}$, and 6) $\{f3, f4, f5\}$.

It can be observed in Fig. 5, that the best overall performance both in PSNR and SSIM terms is achieved by our proposed method with the first and second order filter combination NF3. If only PSNR is considered, the *L1S* combination of sparse and non-sparse prior models of [10] obtains similar results, but the proposed method performs better than the method in [10] in SSIM terms. All the filter combinations considered in Fig. 5 for our method give good results, except for the NF2 filter combination, which performs very well at low noise levels, but worse at high noise levels.

The difference between our proposed method and the others is higher in terms of SSIM than of PSNR. The results obtained using our proposed method, TV, and L1S, are better than the obtained using *BBC* and *ZMT*. The *RSR* method performs very well in PSNR terms, but it achieves the worst SSIM performance for this synthetic image set, specially for high noise levels.

5.2. Experiments with real images

Let us finally perform a qualitative comparison of the results obtained in the SR reconstruction of real LR image sequences. The

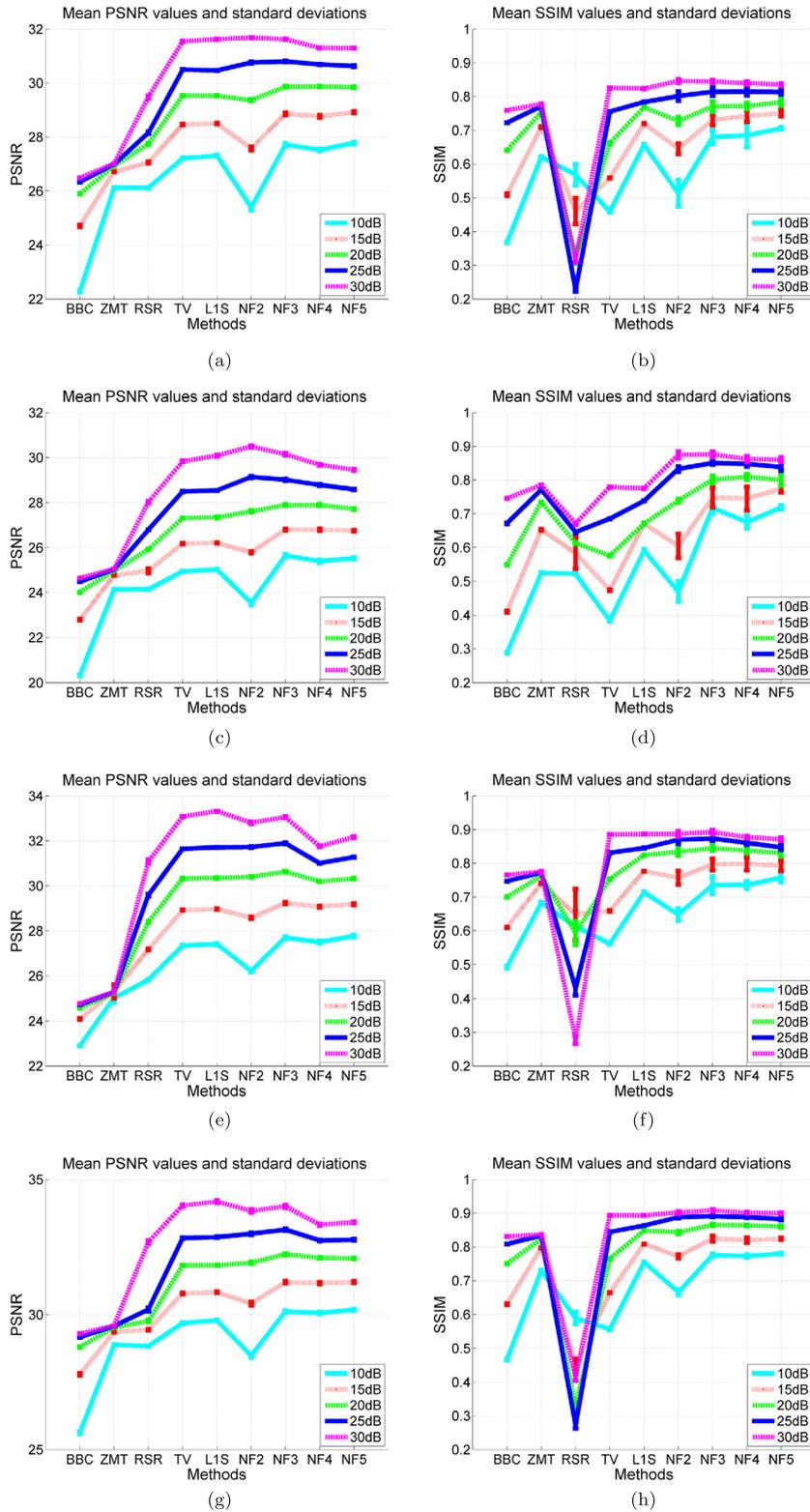


Fig. 5. Mean PSNR and SSIM values, and their standard deviations, at different noise levels, corresponding to the proposed method using $NF2$, $NF3$, $NF4$ and $NF5$ prior combinations, and to the other methods: (a)–(b) for the image in Fig. 1(a), (c)–(d) for the image in Fig. 1(b), (e)–(f) for the image in Fig. 1(c) and (g)–(h) for the image in Fig. 1(d).

LR image sequences have been obtained using a Sony Nex5 digital camera. Three 19 image sequences of 100×100 RAW images have been obtained using 6400 ISO sensitivity. The real observed LR sequences have been first superresolved by a factor $\sqrt{P} = 2$, using the different methods, assuming a 5×5 Gaussian integration PSF of variance 1.

Figs. 6, 7, and 8 show, respectively for the three LR sequences, the first four observations, and the HR reconstructions obtained using different methods. The reconstructions obtained using RSR , ZMT , TV , $L1S$ and our proposed method, with the $NF3$ filter combination, are shown. The reconstructions obtained using ZMT and the proposed method are smoother, and visually more pleasant than

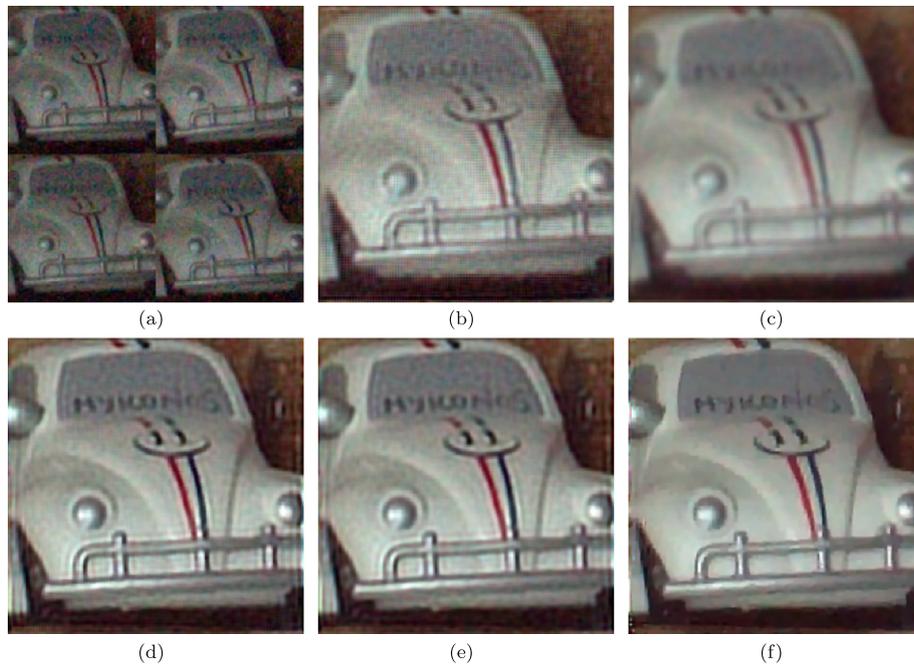


Fig. 6. Real sequence of 19 LR images, superresolved by a factor $\sqrt{P} = 2$. (a) First four LR observations. HR reconstruction using the following methods: (b) RSR, (c) ZMT, (d) TV, (e) LIS and (f) NF3.

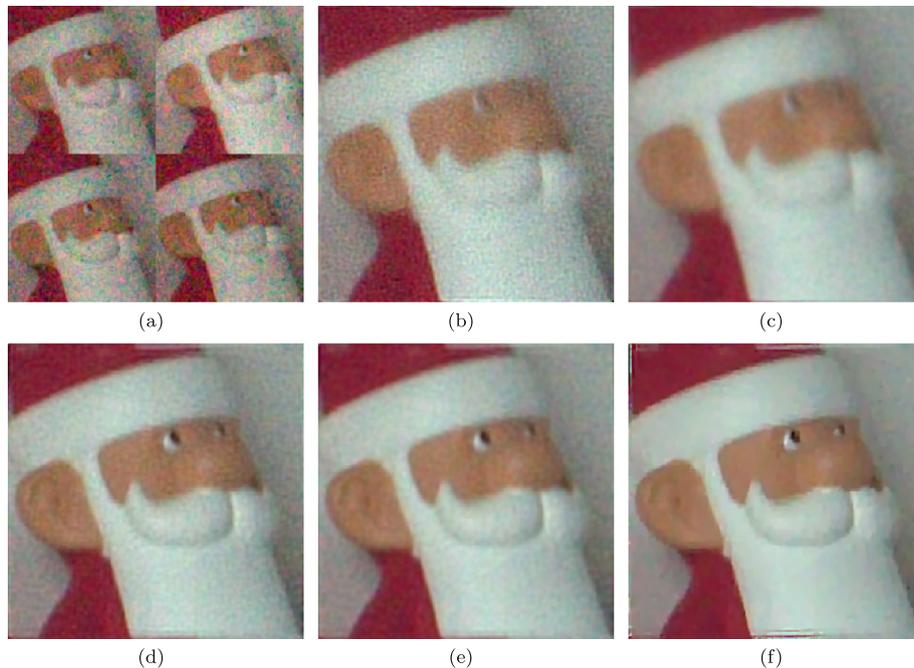


Fig. 7. Real sequence of 19 LR images, superresolved by a factor $\sqrt{P} = 2$. (a) First four LR observations. HR reconstruction using the following methods: (b) RSR, (c) ZMT, (d) TV, (e) LIS and (f) NF3.

the obtained using RSR, TV and the LIS combination of sparse and non-sparse prior models of [10]. This is consistent with the results obtained for the synthetic images set, which were better in SSIM terms for our method than for the others. It can be observed in Figs. 6, 7, and 8, that our method preserves image details better than ZMT.

Experiments with a lower number of observations have also been performed. Fig. 9 shows the results obtained using the different methods for $\sqrt{P} = 2$, on the sequence obtained by taking only the first four LR images of the sequence of Fig. 6. As ex-

pected, the restorations are worse with 4 than with 19 LR observations, but this decrease of quality affects less to our proposed method.

Finally the first 8 images of the sequence of Fig. 8 have been superresolved by a factor $\sqrt{P} = 4$, using the different methods, and the results shown in Fig. 10. Once again, the proposed method suppresses noise better than the other methods, and provides superior reconstructions.

We have used MATLAB implementations of all methods considered our code can be downloaded from <http://decsai.ugr.es/pi/>

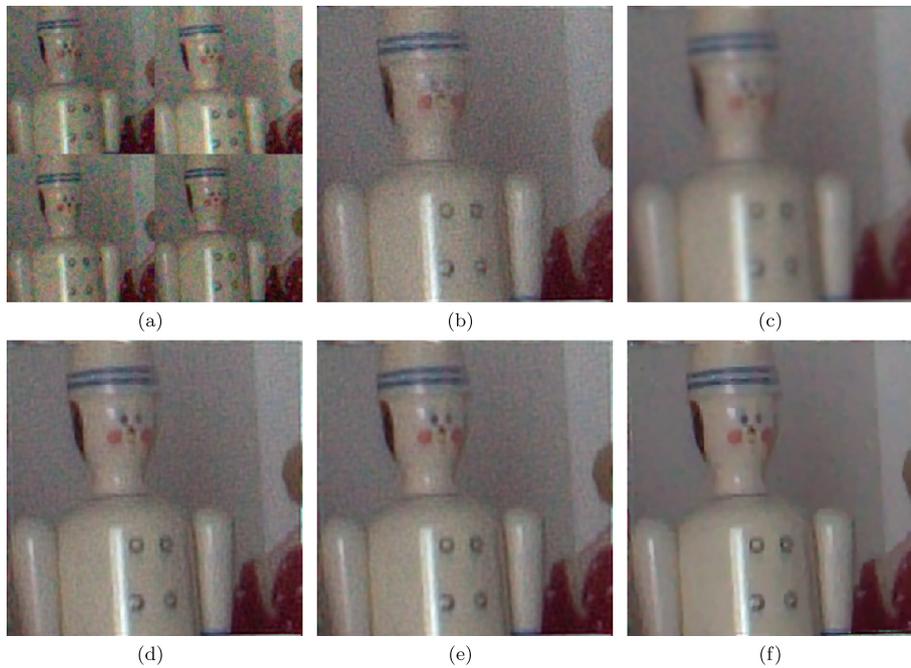


Fig. 8. Real sequence of 19 LR images, superresolved by a factor $\sqrt{P} = 2$. (a) First four LR observations. HR reconstruction using the following methods: (b) RSR, (c) ZMT, (d) TV, (e) LIS and (f) NF3.

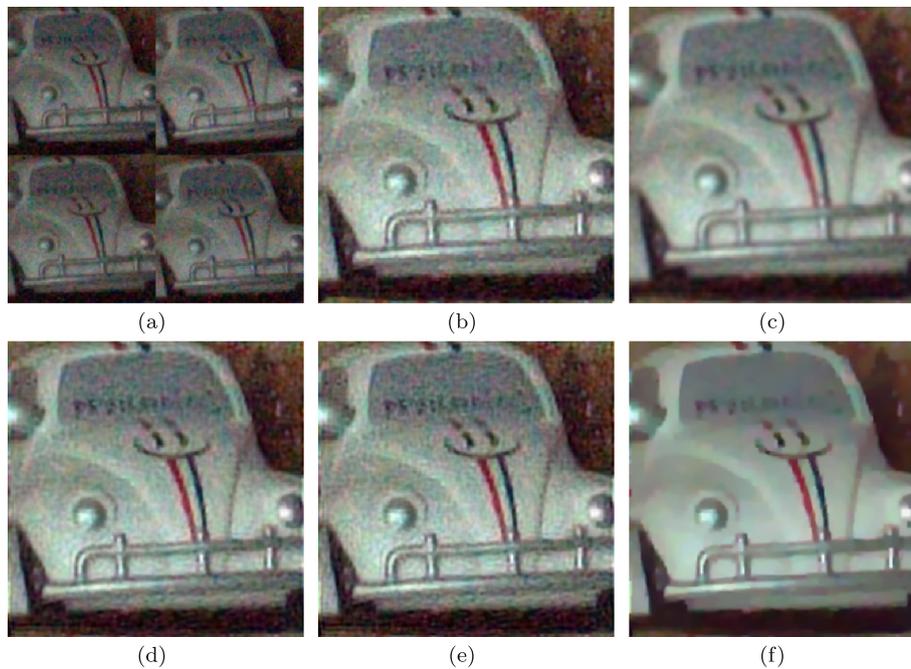


Fig. 9. Real sequence of 4 LR images, superresolved by a factor $\sqrt{P} = 2$. (a) The four LR observations. HR reconstruction using the following methods: (b) RSR, (c) ZMT, (d) TV, (e) LIS and (f) NF3.

superresolution/software.html. All experiments have been run on an @Intel(R) Core(TM) i7CPU 950 at 3.07 GHz processor. The efficiency of the proposed method depends on the characteristics of the B_k system matrix of Eq. (1), and on the number of LR observations. For the simulated image experiments in this section, an iteration takes 35.94 sec. of CPU time, and 5 iterations were necessary. For $\sqrt{P} = 2$ each iteration took 28.3 secs. and 5 iterations were necessary when using 4 LR observations, 231.2 secs. and 4 iterations were necessary when using 19 LR observations. Notice how the CPU time depends on the number of observations used.

6. Conclusions

In this paper the SR image registration and reconstruction problem has been studied within the Bayesian framework, using a general non-stationary HR image prior combination. A new SR method has been proposed, which automatically estimates all the problem unknowns using variational techniques. The influence on the proposed method of the selected prior combination has been analyzed. The proposed method outperforms state of the art SR methods, in the reconstruction of both synthetic and real LR image sequences.

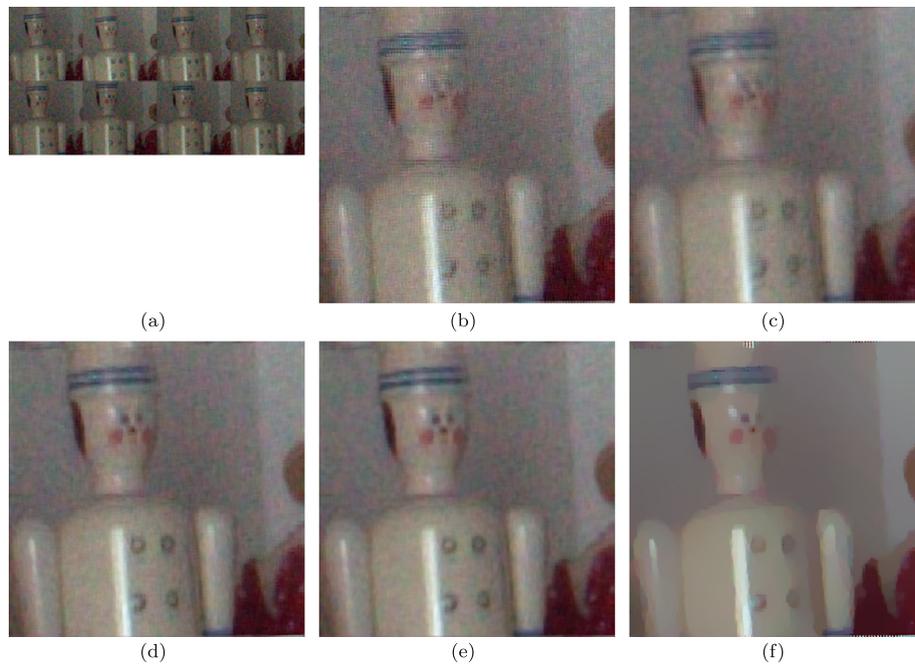


Fig. 10. Real sequence of 8 LR images, superresolved by a factor $\sqrt{P} = 4$. (a) The eight LR observations. HR reconstruction using the following methods: (b) *RSR*, (c) *ZMT*, (d) *TV*, (e) *LIS* and (f) *NF3*.

References

- [1] A.K. Katsaggelos, R. Molina, J. Mateos, Super-resolution of Images and Video, Morgan and Claypool, 2007.
- [2] P. Milanfar, Super-Resolution Imaging, Digital Imaging and Computer Vision, Taylor&Francis/CRC Press, 2010.
- [3] S.N. Kyle Nelson, Asim Bhatti, Performance evaluation of multi-frame super-resolution algorithms, in: International Conference on Digital Image Computing: Techniques and Applications (DICTA), vol. 2012, Centre for Intelligent Systems Research, Deakin University, Geelong, Victoria, Australia, Fremantle, Australia, 2012.
- [4] R. Molina, J. Núñez, F.J. Cortijo, J. Mateos, Image restoration in astronomy. A Bayesian perspective, IEEE Signal Process. Mag. 18 (2) (2001) 11–29.
- [5] N. Nguyen, P. Milanfar, A wavelet-based interpolation–restoration method for superresolution, Circuits Syst. Signal Process. 19 (2000) 321–338.
- [6] G. Anbarjafari, H. Demirel, Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image, ETRI J. 32 (3) (2010) 1642–1654.
- [7] A. Kanemura, S.-I. Maeda, S. Ishii, Superresolution with compound Markov random fields via the variational EM algorithm, Neural Netw. 22 (7) (2009) 1025–1034.
- [8] T. Katsuki, A. Torii, M. Inoue, Posterior-mean super-resolution with a causal Gaussian Markov random field prior, IEEE Trans. Image Process. 21 (7) (2012) 3182–3193.
- [9] S.D. Babacan, R. Molina, A. Katsaggelos, Variational Bayesian super resolution, IEEE Trans. Image Process. 20 (4) (2011) 984–999.
- [10] S. Villena, M. Vega, D. Babacan, R. Molina, A. Katsaggelos, Bayesian combination of sparse and non sparse priors in image super resolution, Digit. Signal Process. 23 (2) (2013) 530–541.
- [11] M. Elad, A. Feuer, Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images, IEEE Trans. Image Process. 6 (1997) 1646–1658.
- [12] A. Zomet, A. Rav-Acha, S. Peleg, Robust super-resolution, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), 2001, pp. 645–650.
- [13] S. Farsiu, M.D. Robinson, M. Elad, P. Milanfar, Fast and robust multiframe super resolution, IEEE Trans. Image Process. 13 (10) (Oct. 2004) 1327–1344.
- [14] P. Vandewalle, L. Sbaiz, J. Vandewalle, M. Vetterli, Super-resolution from unregistered and totally aliased signals using subspace methods, IEEE Trans. Signal Process. 55 (7) (2007) 3687–3703.
- [15] R. Hardie, K. Barnard, E. Armstrong, Joint MAP registration and high-resolution image estimation using a sequence of undersampled images, IEEE Trans. Image Process. 6 (12) (1997) 1621–1633.
- [16] M. Ng, J. Koo, N. Bose, Constrained total least-squares computations for high-resolution image reconstruction with multisensors, Int. J. Imaging Syst. Technol. 12 (1) (2002) 35–42.
- [17] M.E. Tipping, C.M. Bishop, Bayesian image super-resolution, in: Advances in Neural Information Processing Systems (NIPS), vol. 15, MIT Press, 2003, pp. 1303–1310.
- [18] F. Šroubek, J. Flusser, Multichannel blind deconvolution of spatially misaligned images, IEEE Trans. Image Process. 7 (2005) 45–53.
- [19] N. Woods, N. Galatsanos, A. Katsaggelos, Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images, IEEE Trans. Image Process. 15 (2006) 201–213.
- [20] L.C. Pickup, D.P. Capel, S.J. Roberts, A. Zisserman, Bayesian methods for image super-resolution, Comput. J. 52 (1) (2009) 101–113.
- [21] H. Takeda, P. Milanfar, M. Protter, M. Elad, Superresolution without explicit subpixel motion estimation, IEEE Trans. Image Process. 18 (9) (2009) 1958–1975.
- [22] M. Protter, M. Elad, H. Takeda, P. Milanfar, Generalizing the nonlocal-means to super-resolution reconstruction, IEEE Trans. Image Process. 18 (1) (2009) 36–51.
- [23] M.V. Zibetti, F.S. Bazán, J. Mayer, Estimation of the parameters in regularized simultaneous super-resolution, Pattern Recognit. Lett. 32 (1) (2011) 69–78.
- [24] S. Villena, M. Vega, R. Molina, A. Katsaggelos, A general sparse image prior combination in super-resolution, in: 18th International Conference on Digital Signal Processing (DSP 2013), Santorini, Greece, 2013, pp. 1–6.
- [25] E. Vera, M. Vega, R. Molina, A.K. Katsaggelos, Iterative image restoration using nonstationary priors, Appl. Opt. 52 (10) (2013) D102–D110.
- [26] E.S. Lee, M.G. Kang, Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration, IEEE Trans. Image Process. 12 (7) (2003) 826–837.
- [27] B. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: Proceedings of Imaging Understanding Workshop, 1981, pp. 121–130.
- [28] C. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
- [29] M. Beal, Variational algorithms for approximate Bayesian inference, Ph.D. thesis, The Gatsby Computational Neuroscience Unit, University College London, 2003.
- [30] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Perceptual image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

Salvador Villena was born in 1953, in Spain. He received his Bachelor Physics degree from Universidad de Granada (1979) and Ph.D. degree from Universidad de Granada (Departamento de Lenguajes y Sistemas Informáticos), in 2011. He is a staff member (1974–1983) of Computing Center Facility of Universidad de Granada. He is a Lecturer, from 1983 till now, in the ETS Ingeniería informática of Universidad de Granada (Departamento de Lenguajes y Sistemas Informáticos). He teaches language processing. His research focus on image processing (super resolution image reconstruction). He has collaborated at several projects from the Spanish Research Council. He is the recipient of an ISPA Best Paper Award (2009).

Miguel Vega was born in 1956, in Spain. He received the Bachelor Physics degree from the Universidad de Granada, Granada, Spain, in 1979, and the Ph.D. degree from the Universidad de Granada, in 1984. He is with the E.T.S. Ing. Informática of the Universidad de Granada (Departamento de Lenguajes y Sistemas Informáticos), where he is an Associate Professor and teaches software engineering. His research focuses on image processing (multichannel and superresolution image reconstruction). He has collaborated at several projects from the Spanish Research Council.

Rafael Molina was born in 1957. He received the degree in mathematics (statistics), in 1979, and the Ph.D. degree in optimal design in linear models, in 1983. He became Professor of Computer Science and Artificial Intelligence at the University of Granada, Granada, Spain, in 2000. Former Dean of the Computer Engineering School at the University of Granada (1992–2002) and head of the Computer Science and Artificial Intelligence department of the University of Granada (2005–2007).

His research interest focuses mainly in using Bayesian modeling and inference in problems like image restoration (applications to astronomy and medicine), super resolution of images and video, blind deconvolution, computational photography, source recovery in medicine, compressive sensing, low rank matrix decomposition, active learning and classification. See <http://decsai.ugr.es/~rms> for publications, funded projects and grants. Dr. Molina serves the IEEE and other Professional Societies: Applied Signal Processing, Associate Editor (2005–2007), IEEE Trans. on Image Processing, Associate Editor (2010–), Progress in Artificial Intelligence, Associate Editor (2011–) and Digital Signal Processing, Area Editor (2011–). He is the recipient of an IEEE ICIP Paper Award (2007), an ISPA Paper Award (2009), an EUSIPCO Paper Award (2013), and coauthor of a paper awarded the runner-up prize at Reception for early-stage researchers at the House of Commos.

Aggelos K. Katsaggelos received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Greece, in 1979, and the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, in 1981 and 1985, respectively. In 1985, he joined the Department of Electrical Engineering and Computer Science at Northwestern University, where he is currently a Professor holder of the AT&T chair. He was previously the holder of the Ameritech Chair of Information Technology (1997–2003). He is also the Director of the Motorola Center for Seamless Communications, a member of the Academic Staff, NorthShore University Health System, an affiliated faculty at the Department of Linguistics and he has an appointment with the Argonne National Laboratory. He has published extensively in the areas of multimedia signal processing and communications (over 200 journal papers, 500 conference papers and 45 book chapters) and he is the holder of 23 international patents. He is the co-author of Rate-Distortion Based Video Compression (Kluwer, 1997), Super-Resolution for Images and Video (Claypool, 2007) and Joint Source-Channel Video Transmission (Claypool, 2007). Among his many professional activities Professor Katsaggelos was Editor-in-Chief of the IEEE Signal Processing Magazine (1997–2002), a BOG Member of the IEEE Signal Processing Society (1999–2001), and a member of the Publication Board of the IEEE Proceedings (2003–2007). He is a Fellow of the IEEE (1998) and SPIE (2009) and the recipient of the IEEE Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), the IEEE Signal Processing Society Technical Achievement Award (2010), an IEEE Signal Processing Society Best Paper Award (2001), an IEEE ICME Paper Award (2006), an IEEE ICIP Paper Award (2007), an ISPA Paper Award (2009), and EUSIPCO Paper Award (2013). He was a Distinguished Lecturer of the IEEE Signal Processing Society (2007–2008).