

Super-Resolution Using Hidden Markov Model and Bayesian Detection Estimation Framework

Fabrice Humblot^{1,2} and Ali Mohammad-Djafari²

¹DGA/DET/SCET/CEP/ASC/GIP, 94114 Arcueil, France

²LSS/UMR8506 (CNRS-Supélec-UPS), 91192 Gif-sur-Yvette Cedex, France

Received 12 December 2004; Revised 22 May 2005; Accepted 27 May 2005

This paper presents a new method for super-resolution (SR) reconstruction of a high-resolution (HR) image from several low-resolution (LR) images. The HR image is assumed to be composed of homogeneous regions. Thus, the a priori distribution of the pixels is modeled by a finite mixture model (FMM) and a Potts Markov model (PMM) for the labels. The whole a priori model is then a hierarchical Markov model. The LR images are assumed to be obtained from the HR image by lowpass filtering, arbitrarily translation, decimation, and finally corruption by a random noise. The problem is then put in a Bayesian detection and estimation framework, and appropriate algorithms are developed based on Markov chain Monte Carlo (MCMC) Gibbs sampling. At the end, we have not only an estimate of the HR image but also an estimate of the classification labels which leads to a segmentation result.

Copyright © 2006 F. Humblot and A. Mohammad-Djafari. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

This paper concerns super-resolution (SR) reconstruction of an image from a few low-resolution (LR) images in order to gain a better detection on faintly contrasted and isolated specks having a size of almost one pixel in a picture. In this paper we deal more particularly with the SR process.

The SR reconstruction consists in producing a high-resolution (HR) image from a set of LR images. These LR images can be taken from a video sequence. They must come out from a moving scene: *the movement and nonredundant information are what make SR process possible*. So, the obtained HR picture contains more useful information than one of the LR pictures taken in the initial video sequence.

One of the main steps in a SR reconstruction process is the registration of the different LR images. It is defined as the way of matching two or more pictures showing the same scene from different viewpoints, from different sensors, or at different times. To get a good SR reconstruction, it is essential to know accurately the transformation that enables to go from one LR picture to another. Our work context allows us to limit the field of possible transformations between two pictures to the global translatory motion. Indeed, it does not seem unrealistic to have a stabilized and controlled camera to obtain the initial LR video sequence. We can imagine that the movement of the camera during the image acquisition

is limited to global translational move, and that there is no zoom effect (equivalent to homothety transformation) and no rotation of the camera axis. To deal with this problem, we use the image registration by phase correlation. This method and its extension using the gradients of the images are well explained in [1, 2]. We also evaluated both methods in our previous work [3]. These methods are interesting for the easiness of their implementation, the speed of their execution on a computer, and their good subpixel accuracy.

SR methods may be categorized into two main divisions: frequency domain and spatial domain techniques. Our proposed methods use the linear spatial domain observation model. They are stochastic methods, and they use the Bayesian framework. Other works used the same Bayesian approach, we can cite for instance [4, 5].

The problem of SR has been addressed for the first time in [6], with a frequency domain approach. It would be too long to describe all the different SR approaches, so we suggest [7, 8] that give very good overviews of them.

Due to the fact that SR reconstruction is an ill-posed problem, Tikhonov-regularized SR methods have been examined [9–12]. This method is a deterministic approach utilizing regularization functions to impose smoothness constraints on the space of feasible solutions. The regularization functions used in [9] is equivalent to a Gaussian Markov random field (MRF) prior in the Bayesian maximum a posteriori

(MAP) framework. Gaussian MRF priors are well known to produce overly smoothed restorations. The Tikhonov regularization approach is a special case of the more general Bayesian framework under the assumptions of Gaussian noise and prior.

Many works are based basically on modeling the link between the LR images and the HR image through a lowpass filtering, decimating, and translation movement. This is also the model we choose. But to our knowledge, all the existing HR reconstruction methods are based either on a least square estimation which leads to an interpolation, registration, and summation [13], or at the best, on the regularization approach [9–12]. Here, we propose a Bayesian estimation framework which gives the possibility to account for a wider class of a priori model for the distribution of the pixels of the HR image. The method we propose here assumes that the HR image is composed of homogeneous regions. Thus, the a priori distribution of the pixels is modeled by a finite mixture model (FMM) to let their classification in a finite number of classes, and a Potts Markov model (PMM) for the labels. The FMM is a common tool for classification, but in general, the discrete variable which represents the classes is assumed to be i.i.d. Our Potts Markov model for these variables gives the possibility to account for spatial correlation of the pixels. In fact, the PMM parameter controls the mean value of the size of the agglomerated pixels in different classes and thus the mean value of the segmented regions in the image.

The method proposed in [4] concerns the image restoration using Gauss-Markov random fields and line process. Our work differs from that work in two main aspects. First, the purpose of that work is an image restoration which is conceptually different of SR methods. Next, authors of this paper use a line process where we use a label process. More, the novelty of our work in comparison to all known methods is that our framework allows us to obtain not only an estimate of the HR image, but also an estimate of the classification labels which leads to a segmentation result. This last result is useful for possible geometrical features estimation of the results, for example, for a tracking of feature in satellite images. The Bayesian probabilistic framework gives a good estimation of the real classification of the scene.

The authors in [5] consider the SR problem and use a classical approach, also used for instance in [12]. More specifically, in this article, authors focus on an estimation stage of the SR parameters. Our method differs from this work in the model relating the LR images to HR image which is not exactly the same, and in the prior modeling of the HR image which is a Gauss-Markov and not a compound Gauss-Markov random fields, with hidden label process.

In summary, the proposed a priori model for the distribution of the pixels of the HR image results in a hierarchical Markov model which gives the possibility to estimate jointly the HR image, the classification labels which can be used as a segmentation result and also the parameters of the a priori, and the noise model which results in a totally unsupervised HR image estimation and segmentation.

Indeed, the hierarchical structure of the model can be appropriately implemented using the Markov chain Monte Carlo (MCMC) Gibbs sampling.

The paper is organized as follows: in Section 2, the forward model linking the HR image to LR images is detailed, and the basics of the Bayesian estimation framework for SR reconstruction is presented. In Section 3, we give the details of the a priori models for the HR image pixels distribution which is composed of a FMM and Potts MRF. In Section 4, we give the expressions of all the posterior laws which are necessary for the implementation of the MCMC Gibbs sampling. In Section 5, we give details of implemented MCMC Gibbs sampling algorithm. In Section 6, we first present the results which we can obtain with the proposed method and then compare its performances with a classical interpolation, registration, and summation [13], and with another classical method based on popular Tikhonov regularized approach [9]. Section 7 shows results obtained with real video sequences. Finally, we present conclusions and perspectives of this work in Section 8.

2. FORWARD MODEL AND THE BAYESIAN ESTIMATION FOR SR RECONSTRUCTION

A simple model which links the LR images and the HR image is

$$\mathbf{g}_i = \mathbf{D}\mathbf{M}_i\mathbf{B}\mathbf{f} + \boldsymbol{\epsilon}_i = \mathbf{H}_i\mathbf{f} + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, M, \quad (1)$$

where \mathbf{f} is the HR image, $\{\mathbf{g}_i, i=1, \dots, M\}$ represent the M LR images, \mathbf{B} represents the lowpass filter operator needed before sampling a HR image, \mathbf{M}_i are operators representing the translation movements, \mathbf{D} represents a decimation operator, $\boldsymbol{\epsilon}_i$ are the additive noises representing all the errors, and finally, $\mathbf{H}_i = \mathbf{D}\mathbf{M}_i\mathbf{B}$ represents the composite operator linking the HR image \mathbf{f} to the LR image \mathbf{g}_i . This is the forward model.

Noting that \mathbf{g}_i , $\boldsymbol{\epsilon}_i$, and \mathbf{f} represent the images, we also note them by $g_i(\mathbf{r})$, $\epsilon_i(\mathbf{r})$, and $f(\mathbf{r})$, where $\mathbf{r} = (x, y) \in \mathbb{N}^2$ represents the pixel position, and $g_i(\mathbf{r}) = [\mathcal{D}\mathcal{M}_i\mathcal{B}f](\mathbf{r}) + \epsilon_i(\mathbf{r})$, where \mathcal{B} , \mathcal{M}_i , and \mathcal{D} are respectively the equivalent continuous operators of \mathbf{B} , \mathbf{M}_i , and \mathbf{D} .

Note that we may use the following combination:

$$\mathbf{g} = \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_M \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_M \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \boldsymbol{\epsilon}_1 \\ \vdots \\ \boldsymbol{\epsilon}_M \end{bmatrix} \quad (2)$$

to rewrite (1) as

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}, \quad (3)$$

where, given \mathbf{f} and \mathbf{H} , computing \mathbf{g} is the forward model, and estimating \mathbf{f} , given \mathbf{H} and \mathbf{g} , is the corresponding inverse problem. The direct problem is to obtain the LR images from a HR image.

An illustration of operator's role is given with Figure 1. We consider this figure as being made of 3 parts. Each part can be seen as a row of the full Figure 1. The first one

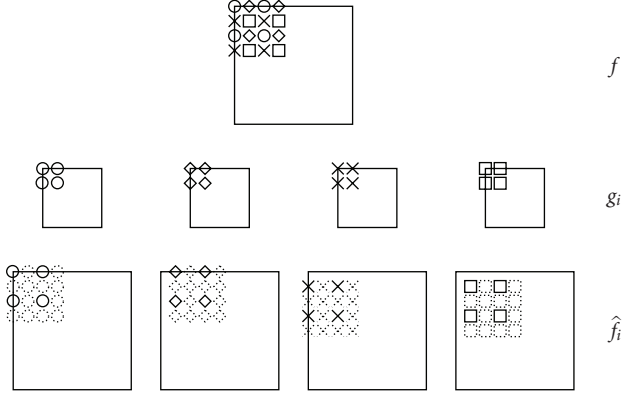


FIGURE 1: HR image \mathbf{f} , LR images $\mathbf{g}_i = \mathbf{D}\mathbf{M}_i\mathbf{f}$, $i = 1, \dots, 4$, registered and interpolated images $\hat{\mathbf{f}}_i = \mathbf{M}_i^t\mathbf{D}^t\mathbf{g}_i$, which can be used to obtain a HR image by a classical method consisting in $\hat{\mathbf{f}} = (1/M) \sum_i \hat{\mathbf{f}}_i$.

represents a HR image, each symbol representing a different pixel of the image. The second shows a representation of all the LR images obtained from the previous HR image and for a decimation factor $d = 2$. Finally, all the registered and interpolated LR pictures on the HR grid can be seen on the third and last row. The symbols with full contours are the original pixel values from LR images, and symbols with dotted contours are interpolated pixel values.

Next, the fact of taking one pixel over two in the original HR image to create the LR images shows the effect of operator \mathbf{D} for a decimation factor $d = 2$. And the fact that two neighboring pixels of the HR image, for instance, the upper-left circle and diamond, are located at the same position in the corresponding LR images (the upper-left pixel) shows the effect of operators \mathbf{M}_i . Thus, the second row shows how to go from \mathbf{f} to \mathbf{g}_i , $i=1, \dots, 4$.

Lastly, the presence of interpolated pixels values on the third row of Figure 1 shows the inverse effect of operator \mathbf{D} . And the fact that the upper-left circle and diamond have found back their correct initial positions (in comparison to their positions in the initial HR image) shows the inverse effect of operator \mathbf{M}_i .

Through an example of size 125×125 pixels², Figure 2 shows the forward process of generation of LR images from a HR image \mathbf{f} that can be seen on image (a). It starts by a lowpass filtering, giving $\mathbf{B}\mathbf{f}$ shown on (b), sampling, translation, and decimation by a factor d of 5 give $\mathbf{H}\mathbf{f}$ and it is shown on (c), and finally, alteration by a random noise representing measurement noise and all the other errors of modeling gives \mathbf{g} , which is shown on (d) (we used a noise centered, white, and Gaussian obtained with a SNR—see (35)—of 5 dB).

The Bayesian estimation framework for SR reconstruction can be summarized as follows.

- (i) Use the forward model (1) and some assumptions on the noise to obtain the likelihood $p(\mathbf{g} | \mathbf{f}, \theta_\epsilon)$, where θ_ϵ represents the parameters of the probability distribution of the noise.

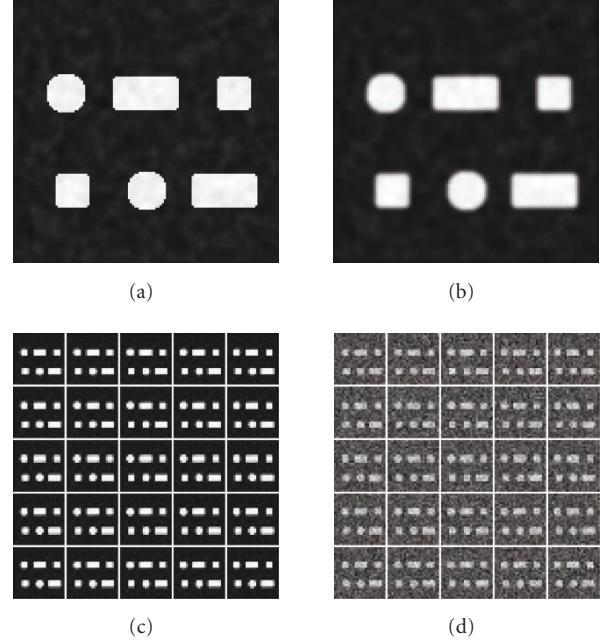


FIGURE 2: (a) A HR image \mathbf{f} , (b) its lowpass filtered $\mathbf{B}\mathbf{f}$, (c) LR images without noise $\mathbf{g}_i = \mathbf{H}_i\mathbf{f}$, and (d) LR images with additive noise $\mathbf{g}_i = \mathbf{H}_i\mathbf{f} + \epsilon_i$.

- (ii) Use all the prior information or the desired properties for the solution to assign an a priori probability distribution $p(\mathbf{f} | \theta_f)$, where θ_f is its parameters.
- (iii) Use the Bayesian approach to obtain

- (a) the a posteriori probability distribution

$$p(\mathbf{f} | \mathbf{g}, \theta) \propto p(\mathbf{g} | \mathbf{f}, \theta_\epsilon) p(\mathbf{f} | \theta_f), \quad (4)$$

where $\theta = (\theta_\epsilon, \theta_f)$ if θ is known (supervised case), or

- (b) the joint a posteriori probability distribution

$$p(\mathbf{f}, \theta | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \theta_\epsilon) p(\mathbf{f} | \theta_f) p(\theta_f) p(\theta_\epsilon) \quad (5)$$

if the θ is unknown (unsupervised case).

- (iv) Finally, define an estimator $\hat{\mathbf{f}}$ for \mathbf{f} , and $\hat{\theta}$ for θ , based on these posterior probability laws.

The next section defines more in detail the prior laws which are needed to obtain the expressions of these posterior laws, and proposes different estimators based on them.

3. A PRIORI MODEL OF THE HR IMAGE PIXELS

3.1. HMM modeling of HR image

The main idea in this modeling is to assume that the image $f(\mathbf{r})$, $\mathbf{r} \in \mathcal{R}$, is composed of a finite set K of homogeneous regions \mathcal{R}_k with given labels $Z(\mathbf{r}) = k$, $k = 1, \dots, K$, such that $\mathcal{R}_k = \{\mathbf{r} : z(\mathbf{r}) = k\}$, $\mathcal{R} = \cup_k \mathcal{R}_k$, and the corresponding pixel values $\mathbf{f}_k = \{f(\mathbf{r}) : \mathbf{r} \in \mathcal{R}_k\}$, and $\mathbf{f} = \cup_k \mathbf{f}_k$. The

hidden Markov modeling (HMM) is a very general and efficient way to model appropriately such images. The main idea is to assume that all the pixel values \mathbf{f}_k of a homogeneous region k follow a given probability law, for example, a Gaussian $\mathcal{N}(m_k \mathbf{1}, \Sigma_k)$, where $\mathbf{1}$ is a generic vector of ones of size $n_k = |\mathcal{R}_k|$ which is the number of pixels in region k with $\sum_k n_k = |\mathcal{R}| = n$, the total number of pixels of the HR image.

In the following, we consider two cases.

- (i) The pixels in a given region are assumed i.i.d.:

$$p(f(\mathbf{r}) | z(\mathbf{r}) = k) = \mathcal{N}(f(\mathbf{r}) | m_k, \sigma_k^2), \quad k = 1, \dots, K, \quad (6)$$

and thus

$$p(\mathbf{f}_k) = p(f(\mathbf{r}), \mathbf{r} \in \mathcal{R}_k) = \mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}, \sigma_k^2 \mathbf{I}), \quad (7)$$

with \mathbf{I} the identity matrix of size n_k .

- (ii) The pixels in a given region are assumed to be locally dependent:

$$p(\mathbf{f}_k) = p(f(\mathbf{r}), \mathbf{r} \in \mathcal{R}_k) = \mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}, \Sigma_k), \quad (8)$$

where Σ_k is an appropriate covariance matrix whose shape and structure depend on the modeling of this dependency. We propose a first-order Markov model with the four nearest neighbors.

In both cases, the pixels in different regions are assumed to be independent:

$$p(\mathbf{f} | \mathbf{z}) = \prod_{k=1}^K p(\mathbf{f}_k) = \prod_{k=1}^K \mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}, \Sigma_k). \quad (9)$$

Note that $f(\mathbf{r})$ is a scalar and $p(f(\mathbf{r}) | z(\mathbf{r}) = k)$ is its conditional probability density function, but \mathbf{f}_k is a vector and $p(\mathbf{f}_k)$ is the joint probability density function of all the pixels in region k .

3.2. Modeling the labels

Noting that the two models (7) and (8) are conditioned on the value of $z(\mathbf{r}) = k$, they can be rewritten in the following general form:

$$p(f_k(\mathbf{r})) = \sum_{k=1}^K P(z(\mathbf{r}) = k) \mathcal{N}(f_k(\mathbf{r}) | m_k, \sigma_k^2). \quad (10)$$

Now, we need also to model the probability distribution $P(Z(\mathbf{r}) = z(\mathbf{r}), \mathbf{r} \in \mathcal{R})$ of the vector random variables $\mathbf{Z} = \{Z(\mathbf{r}) : \mathbf{r} \in \mathcal{R}\}$ which we note hereafter $p(\mathbf{z})$. For this too, we consider two cases.

- (i) Independent Gaussian mixture (IGM) model, where $\{Z(\mathbf{r}), \mathbf{r} \in \mathcal{R}\}$ are assumed to be independent and

$$P(z(\mathbf{r}) = k) = p_k, \quad (11)$$

$$\text{with } \sum_{k=1}^K p_k = 1, \quad p(\mathbf{z}) = \prod_{k=1}^K p_k.$$

- (ii) Contextual Gaussian mixture (CGM) model that we also call hidden Markov model (HMM), where $\{Z(\mathbf{r}), \mathbf{r} \in \mathcal{R}\}$ are assumed to be Markovian:

$$p(\mathbf{z}) \propto \exp \left[\alpha \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{s} \in \mathcal{V}(\mathbf{r})} \delta(z(\mathbf{r}) - z(\mathbf{s})) \right], \quad (12)$$

which is the Potts Markov random field (PMRF). $\mathcal{V}(\mathbf{r})$ represents neighboring pixels of \mathbf{s} with $|\mathcal{V}(\mathbf{r})| = 4$. The Markovian modeling with a connectivity of 4 considers that each label value $z(\mathbf{r})$ of a pixel at position \mathbf{r} is function of the values of its four closest pixels $z(\mathbf{s}), \mathbf{s} \in \mathcal{V}(\mathbf{r})$ (neighboring pixels). They are the one above, the one on its right, the one under, and the one on its left. The parameter α controls the mean value of the regions' sizes. Here, it controls the mean value of the sizes of the classes, that is, increasing α results in a realization where the different classes become more homogeneous. Using the Hammersley-Clifford equivalence of the Gibbs and Markov random fields, we can also write

$$p(z(\mathbf{r}) | z(\mathbf{s}), \mathbf{s} \in \mathcal{R}) \propto \exp \left[\alpha \sum_{\mathbf{s} \in \mathcal{V}(\mathbf{r})} \delta(z(\mathbf{r}) - z(\mathbf{s})) \right] \quad (13)$$

which shows that the probability of obtaining a label $z(\mathbf{r})$ for a pixel is related to the number of neighboring pixels having the same label. We remind that $\delta(\cdot)$ is the Dirac function defined by $\delta(0) = 1$, and $\delta(t) = 0$ for $t \neq 0$.

3.3. Hyperparameters prior law

The final point before obtaining an expression for the posterior probability law of all the unknowns, that is, $p(\mathbf{f}, \boldsymbol{\theta} | \mathbf{g})$ is to assign a prior probability law $p(\boldsymbol{\theta})$ to the hyperparameters $\boldsymbol{\theta}$. Even if this point has been one of the main discussing points between Bayesian and classical statistical research community, and still there are many open problems, we choose here to use the conjugate priors for simplicity. The conjugate priors have at least two advantages:

- (i) they can be considered as a particular family of a differential geometry-based family of priors [14–16],
- (ii) and they are easy to use because the prior and the posterior probability laws stay in the same family.

In our case, we need to assign prior probability laws to the means m_k , to the variances σ_k^2 or to the covariance matrices Σ_k , and also to the covariance matrices of the noises Σ_{ϵ_i} .

The conjugate priors for the means m_k are, in general, the Gaussians $\mathcal{N}(m_k | m_{k0}, \sigma_{k0}^2)$, those of variances σ_k^2 are the inverse Gammas $\mathcal{IG}(\sigma_k^2 | \alpha_{k0}, \beta_{k0})$, and those for the covariance matrices Σ_k are the inverse Wishart's $\mathcal{IW}(\Sigma_k | \alpha_{k0}, \Lambda_{k0})$. See the appendix for a detailed expression of these probability density functions.

4. A POSTERIORI PROBABILITY LAWS

We now have all the elements for writing the expressions of the posterior laws. We are going to summarize them here.

(i) Likelihood.

The expression of the likelihood depends on the observation model (1). Then, the expression is

$$\begin{aligned} p(\mathbf{g} | \mathbf{f}, \boldsymbol{\theta}_\epsilon) &= \prod_{i=1}^M p(\mathbf{g}_i | \mathbf{f}, \boldsymbol{\Sigma}_{\epsilon_i}) \\ &= \prod_{i=1}^M \mathcal{N}(\mathbf{g}_i | \mathbf{H}_i \mathbf{f}, \boldsymbol{\Sigma}_{\epsilon_i}), \end{aligned} \quad (14)$$

where we assumed that the noises $\boldsymbol{\epsilon}_i$ are independent, centered, and Gaussian with covariance matrices $\boldsymbol{\Sigma}_{\epsilon_i}$ which, hereafter, are also assumed to be diagonal $\boldsymbol{\Sigma}_{\epsilon_i} = \sigma_{\epsilon_i}^2 \mathbf{I}$. We note $\boldsymbol{\theta}_\epsilon = \{\sigma_{\epsilon_i}^2, i = 1, \dots, M\}$.

(ii) HMM for the images:

$$\begin{aligned} p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}_f) &= \prod_{k=1}^K p(\mathbf{f}_k | m_k \mathbf{1}, \boldsymbol{\Sigma}_k) \\ &= \prod_{k=1}^K \mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}, \boldsymbol{\Sigma}_k), \end{aligned} \quad (15)$$

where $\boldsymbol{\Sigma}_k$ is characterized either by σ_k^2 assuming $\boldsymbol{\Sigma}_k = \sigma_k^2 \mathbf{I}$, or by an extra parameter ρ_k which controls the correlation between the neighboring pixels. Assuming that the pixels in a homogeneous region are modeled with a homogeneous Gauss-Markov field,

$$p(\mathbf{f}_k) \propto \exp \left[\sum_{\mathbf{r} \in \mathcal{R}_k} \left(f(\mathbf{r}) - \beta(\mathbf{r}) \sum_{\mathbf{s} \in \mathcal{V}(\mathbf{r})} f(\mathbf{s}) \right)^2 \right] \quad (16)$$

considering the four nearest pixels neighborhood.

Thus, here, the hyperparameters become $\boldsymbol{\theta}_f = \{(m_k, \sigma_k^2, \rho_k), k = 1, \dots, K\}$.

(iii) PMRF for the labels:

$$p(\mathbf{z}) \propto \exp \left[\alpha \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{s} \in \mathcal{V}(\mathbf{r})} \delta(z(\mathbf{r}) - z(\mathbf{s})) \right], \quad (17)$$

where we used the simplified notation $p(\mathbf{z}) = P(Z(\mathbf{r}) = z(\mathbf{r}), \mathbf{r} \in \mathcal{R})$.

(iv) Conjugate priors for the hyperparameters:

$$\begin{aligned} p(m_k) &= \mathcal{N}(m_k | m_{k0}, \sigma_{k0}^2), \\ p(\sigma_k^2) &= \mathcal{I}\mathcal{G}(\sigma_k^2 | \alpha_{k0}, \beta_{k0}), \\ p(\boldsymbol{\Sigma}_k) &= \mathcal{I}\mathcal{W}(\boldsymbol{\Sigma}_k | \alpha_{k0}, \boldsymbol{\Lambda}_{k0}), \\ p(\sigma_{\epsilon_i}^2) &= \mathcal{I}\mathcal{G}(\sigma_{\epsilon_i}^2 | \alpha_0^{\epsilon_i}, \beta_0^{\epsilon_i}). \end{aligned} \quad (18)$$

(v) Joint posterior law of \mathbf{f} , \mathbf{z} , and $\boldsymbol{\theta}$:

$$p(\mathbf{f}, \mathbf{z}, \boldsymbol{\theta} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \boldsymbol{\theta}_\epsilon) p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}_f) p(\mathbf{z}) p(\boldsymbol{\theta}). \quad (19)$$

The forward model and the priors for this case can be summarized as follows:

$$\begin{aligned} \mathbf{g}_i &= \mathbf{H}_i \mathbf{f} + \boldsymbol{\epsilon}_i \longleftrightarrow \mathbf{g} = \mathbf{H} \mathbf{f} + \boldsymbol{\epsilon}, \\ p(\mathbf{g} | \mathbf{f}) &= \mathcal{N}(\mathbf{g} | \mathbf{H} \mathbf{f}, \boldsymbol{\Sigma}_\epsilon) \quad \text{with } \boldsymbol{\Sigma}_\epsilon = \text{diag}[\boldsymbol{\Sigma}_{\epsilon_1}, \dots, \boldsymbol{\Sigma}_{\epsilon_M}], \\ p(\mathbf{g}_i | \mathbf{f}) &= \mathcal{N}(\mathbf{g}_i | \mathbf{H}_i \mathbf{f}, \boldsymbol{\Sigma}_{\epsilon_i}) \quad \text{with } \boldsymbol{\Sigma}_{\epsilon_i} = \sigma_{\epsilon_i}^2 \mathbf{I}, \\ p(f(\mathbf{r}) | z(\mathbf{r}) = k) &= \mathcal{N}(f(\mathbf{r}) | m_k, \sigma_k^2), \quad k = 1, \dots, K, \\ \mathcal{R}_k &= \{\mathbf{r} : z(\mathbf{r}) = k\}, \quad \mathbf{f}_k = \{f(\mathbf{r}) : \mathbf{r} \in \mathcal{R}_k\}, \\ p(\mathbf{f}_k) &= \mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}_k, \boldsymbol{\Sigma}_k) \quad \text{with } \boldsymbol{\Sigma}_k = \sigma_k^2 \mathbf{I}_k, \\ p(\mathbf{z}) &= p(z(\mathbf{r}), \mathbf{r} \in \mathcal{R}) \propto \exp \left[\alpha \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{s} \in \mathcal{V}(\mathbf{r})} \delta(z(\mathbf{r}) - z(\mathbf{s})) \right], \\ p(\mathbf{f} | \mathbf{z}) &= \prod_k \mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}_k, \boldsymbol{\Sigma}_k) = \mathcal{N}(\mathbf{f} | \mathbf{m}_z, \boldsymbol{\Sigma}_z), \\ \text{with } \mathbf{m}_z &= [m_1 \mathbf{1}'_1, \dots, m_K \mathbf{1}'_K]', \quad \boldsymbol{\Sigma}_z = \text{diag}[\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K], \\ p(m_k) &= \mathcal{N}(m_k | m_{k0}, \sigma_{k0}^2), \\ p(\sigma_k^2) &= \mathcal{I}\mathcal{G}(\sigma_k^2 | \alpha_{k0}, \beta_{k0}), \\ p(\sigma_\epsilon^2) &= \mathcal{I}\mathcal{G}(\sigma_\epsilon^2 | \alpha_0^\epsilon, \beta_0^\epsilon). \end{aligned} \quad (20)$$

5. MAXIMUM A POSTERIORI AND MCMC GIBBS SAMPLING

5.1. Maximum a posteriori

First, assume that $\boldsymbol{\theta}$ and \mathbf{z} are known. Then, consider the maximum a posteriori (MAP) estimate

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \{p(\mathbf{f} | \mathbf{g}, \mathbf{z}, \boldsymbol{\theta})\} = \arg \min_{\mathbf{f}} \{J(\mathbf{f} | \mathbf{g}, \mathbf{z}, \boldsymbol{\theta})\}. \quad (21)$$

Then, it is easy to show that

(i) when pixels in given regions are assumed to be i.i.d. (7), we have

$$\begin{aligned} J_1(\mathbf{f} | \mathbf{g}, \mathbf{z}, \boldsymbol{\theta}) &= \|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2 + \lambda \|\mathbf{f} - \mathbf{m}\|_{\boldsymbol{\Sigma}}^2 \\ &= \|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2 + \lambda (\mathbf{f} - \mathbf{m})^t \boldsymbol{\Sigma}^{-1} (\mathbf{f} - \mathbf{m}) \\ &= \sum_{i=1}^M \|\mathbf{g}_i - \mathbf{H}_i \mathbf{f}\|^2 + \lambda \sum_{k=1}^K \frac{\|\mathbf{f}_k - m_k \mathbf{1}\|^2}{\sigma_k^2} \\ &= \sum_{i=1}^M \|\mathbf{g}_i - \mathbf{H}_i \mathbf{f}\|^2 + \lambda \sum_{k=1}^K \sum_{\mathbf{r} \in \mathcal{R}_k} \frac{\|f(\mathbf{r}) - m_k\|^2}{\sigma_k^2}, \end{aligned} \quad (22)$$

where we noted by $\boldsymbol{\Sigma}^2 = \text{diag}[\sigma_1^2, \dots, \sigma_K^2]$,

(ii) when pixels in given regions are assumed to be locally dependent (8) with a local Markovian model, we have

$$\begin{aligned}
J_2(\mathbf{f} \mid \mathbf{g}, \mathbf{z}, \boldsymbol{\theta}) &= \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda \|\mathbf{f} - \mathbf{m}\|_{\Sigma}^2 \\
&= \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda(\mathbf{f} - \mathbf{m})^t \boldsymbol{\Sigma}^{-1} (\mathbf{f} - \mathbf{m}) \\
&= \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda \|\mathbf{D}(\mathbf{f} - \mathbf{m})\|^2 \\
&= \sum_{i=1}^M \|\mathbf{g}_i - \mathbf{H}_i \mathbf{f}\|^2 + \lambda \sum_{k=1}^K \frac{\|\mathbf{D}\tilde{\mathbf{f}}_k\|^2}{\sigma_k^2} \\
&= \sum_{i=1}^M \|\mathbf{g}_i - \mathbf{H}_i \mathbf{f}\|^2 \\
&\quad + \lambda \sum_{k=1}^K \frac{1}{\sigma_k^2} \sum_{\mathbf{r} \in \mathcal{R}_k} \left(\tilde{f}(\mathbf{r}) - \beta_{\mathbf{r}} \left(\sum_{\mathbf{s} \in (\mathcal{V}(\mathbf{r}) \cap \mathcal{R}_k)} \tilde{f}(\mathbf{s}) \right) \right)^2,
\end{aligned} \tag{23}$$

where $\boldsymbol{\Sigma}^{-1} = \text{diag}[\sigma_1^2, \dots, \sigma_K^2] \mathbf{D}\mathbf{D}^t$, $\tilde{f}(\mathbf{r}) = f(\mathbf{r}) - m(\mathbf{r})$, and $\beta_{\mathbf{r}}$ a coefficient depending on pixel \mathbf{r} . Let $n_{\mathbf{r}}$ be $\text{Card}(\mathcal{V}(\mathbf{r}) \cap \mathcal{R}_k)$, which is the number of neighboring pixels of the pixel \mathbf{r} that belong to the same region \mathcal{R}_k . $\beta_{\mathbf{r}}$ equals to $1/n_{\mathbf{r}}$ if $n_{\mathbf{r}} \neq 0$ and equals to 0 otherwise.

Note that, when we assume that the whole image is composed of one statistically homogeneous region ($K = 1$), these two criteria become

$$\begin{aligned}
J_1(\mathbf{f}) &= \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \frac{\lambda}{\sigma^2} \|\mathbf{f} - \mathbf{m}\|^2, \\
J_2(\mathbf{f}) &= \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \frac{\lambda}{\sigma^2} \|\mathbf{D}(\mathbf{f} - \mathbf{m})\|^2,
\end{aligned} \tag{24}$$

which can then be compared to the regularization criterion used by [9]. It is the classical Tikhonov regularization approach without σ^2 and \mathbf{m} , values that we get from our classification.

So, compared to the classical regularization methods and, in particular, the SR method firstly proposed by [9], here, we go farther in details of modeling the HR image. Indeed, we model the HR image as to be a piecewise homogeneous regions, each characterized by a Gaussian process with mean m_k , variance σ_k^2 , or covariance $\boldsymbol{\Sigma}_k$. In the last case, we assume a Gauss-Markov process with the four nearest neighbors.

5.2. MCMC Gibbs sampling

One more advantage with this model is that we also can go farther and try to estimate both the shape of each homogeneous region modeled through $z(\mathbf{r})$ and estimate also the corresponding hyperparameters m_k and σ_k^2 through an iterative process (unsupervised) using either any alternate optimization such as Expectation-Maximization or a more general MCMC Gibbs sampling process. For this purpose, we propose the following general iterative algorithms.

(i) Joint MAP (Algorithm 1):

$$\begin{aligned}
\hat{\mathbf{f}} &= \arg \max_{\mathbf{f}} \{p(\mathbf{f} \mid \mathbf{z}, \boldsymbol{\theta}, \mathbf{g})\}, \\
\hat{\boldsymbol{\theta}} &= \arg \max_{\boldsymbol{\theta}} \{p(\boldsymbol{\theta} \mid \mathbf{f}, \mathbf{z}, \mathbf{g})\}, \\
\hat{\mathbf{z}} &= \arg \max_{\mathbf{z}} \{p(\mathbf{z} \mid \mathbf{f}, \boldsymbol{\theta}, \mathbf{g})\}.
\end{aligned} \tag{25}$$

(ii) MAP-Gibbs sampling (Algorithm 2):

$$\begin{aligned}
\hat{\mathbf{f}} &= \arg \max_{\mathbf{f}} \{p(\mathbf{f} \mid \mathbf{z}, \boldsymbol{\theta}, \mathbf{g})\}, \\
\text{sample } \hat{\boldsymbol{\theta}} &\text{ using } p(\boldsymbol{\theta} \mid \mathbf{f}, \mathbf{z}, \mathbf{g}), \\
\text{sample } \hat{\mathbf{z}} &\text{ using } p(\mathbf{z} \mid \mathbf{f}, \boldsymbol{\theta}, \mathbf{g}), \\
&\text{or using } p(\mathbf{z} \mid \boldsymbol{\theta}, \mathbf{g}),
\end{aligned} \tag{26}$$

or still

(ii) MAP-Gibbs sampling (Algorithm 3):

$$\begin{aligned}
\hat{\mathbf{f}} &= \arg \max_{\mathbf{f}} \{p(\mathbf{f} \mid \mathbf{z}, \boldsymbol{\theta}, \mathbf{g})\}, \\
\hat{\boldsymbol{\theta}} &= \arg \max_{\boldsymbol{\theta}} \{p(\boldsymbol{\theta} \mid \mathbf{f}, \mathbf{z}, \mathbf{g})\}, \\
\text{sample } \hat{\mathbf{z}} &\text{ using } p(\mathbf{z} \mid \mathbf{f}, \boldsymbol{\theta}, \mathbf{g}), \\
&\text{or using } p(\mathbf{z} \mid \boldsymbol{\theta}, \mathbf{g}).
\end{aligned} \tag{27}$$

In all cases, we need to initialize the algorithm. For this, we propose to start by assuming $K = 1$, and thus $\mathbf{z} = \mathbf{1}$ and $\boldsymbol{\theta} = [m_1, \sigma_1^2] = [0, 1]$. This means that we try to obtain a regularized solution $\mathbf{f}^{(0)}$ from which we can use any classical histogram-based segmentation to obtain $\mathbf{z}^{(0)}$, and a classical maximum likelihood estimation approach to obtain a first estimation $\boldsymbol{\theta}^{(0)}$ for $\boldsymbol{\theta}$. Then, we can continue the iterations using any of the proposed algorithms.

Between the three proposed algorithms, we may note that there is not any theoretical guaranty of the convergence for either of them. However, in Algorithms 1 and 3, the critical point in the segmentation part is to compute $\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta} \mid \mathbf{f}, \mathbf{z}, \mathbf{g})$, because it can give values of σ_k which are very small (10^{-6} or even less) if pixels of a region \mathcal{R}_k have almost all the same values. This difficulty is attenuated in Algorithm 2, which is the one we used for all our simulations.

The following relations summarize all the posterior probability laws that are needed to implement these algorithms:

$$p(\mathbf{f} \mid \mathbf{z}, \boldsymbol{\theta}, \mathbf{g}) = \mathcal{N}(\mathbf{f} \mid \hat{\mathbf{f}}, \hat{\boldsymbol{\Sigma}}) \tag{28}$$

with

$$\begin{aligned}
\hat{\boldsymbol{\Sigma}} &= (\mathbf{H}^t \boldsymbol{\Sigma}_{\epsilon}^{-1} \mathbf{H} + \boldsymbol{\Sigma}_z^{-1})^{-1}, \quad \hat{\mathbf{f}} = \hat{\boldsymbol{\Sigma}} (\mathbf{H}^t \boldsymbol{\Sigma}_{\epsilon}^{-1} \mathbf{g} + \boldsymbol{\Sigma}_z^{-1} \mathbf{m}_z), \\
p(\mathbf{z} \mid \mathbf{g}, \boldsymbol{\theta}) &\propto p(\mathbf{g} \mid \mathbf{z}, \boldsymbol{\theta}) p(\mathbf{z})
\end{aligned} \tag{29}$$

with

$$\begin{aligned}
p(\mathbf{g} \mid \mathbf{z}, \boldsymbol{\theta}) &= \mathcal{N}(\mathbf{g} \mid \mathbf{H}\mathbf{m}_z, \boldsymbol{\Sigma}_g) \quad \text{with } \boldsymbol{\Sigma}_g = \mathbf{H}\boldsymbol{\Sigma}_z\mathbf{H}^t + \boldsymbol{\Sigma}_{\epsilon}, \\
p(m_k \mid \mathbf{z}, \mathbf{f}) &= \mathcal{N}(m_k \mid \mu_k, \nu_k)
\end{aligned} \tag{30}$$

with

$$v_k = \left(\frac{n_k}{\sigma_k^2} + \frac{1}{\sigma_{k_0}^2} \right)^{-1}, \quad \mu_k = v_k \left(\frac{n_k \bar{f}_k}{\sigma_k^2} + \frac{m_{k_0}}{\sigma_{k_0}^2} \right), \quad (31)$$

$$p(\sigma_k^2 | \mathbf{f}, \mathbf{z}) = \mathcal{L}\mathcal{G}(\sigma_k^2 | \alpha_k, \beta_k)$$

with

$$\alpha_k = \alpha_{k_0} + \frac{n_k}{2}, \quad \beta_k = \beta_{k_0} + \frac{n_k \bar{s}_k}{2}, \quad (32)$$

where

$$\bar{f}_k = \frac{1}{n_k} \sum_{\mathbf{r} \in \mathcal{R}_k} f_i(\mathbf{r}), \quad \bar{s}_k = \sum_{\mathbf{r} \in \mathcal{R}_k} (f(\mathbf{r}) - m_k)^2, \quad (33)$$

$$p(\sigma_\epsilon^2 | \mathbf{f}, \mathbf{g}) = \mathcal{L}\mathcal{G}(\sigma_\epsilon^2 | \alpha^\epsilon, \beta^\epsilon)$$

with

$$\alpha^\epsilon = \frac{n}{2} + \alpha_0^\epsilon, \quad \beta^\epsilon = \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \beta_0^\epsilon, \quad (34)$$

$$n_k = \text{number of pixels in } \mathcal{R}_k,$$

$$n = \text{total number of pixels}.$$

Some details about the way we obtained these relations are given in the appendix.

6. SIMULATION RESULTS AND PERFORMANCES OF THE PROPOSED METHOD WITH ARTIFICIALLY GENERATED LR IMAGES

To have a good evaluation of this new method, and for the evaluation of its performances, we constructed our own LR pictures from a supposed HR image of size 125×125 pixels². Because of the classification feature of our technique that leads to a segmentation of the HR image, we began this construction with a segmented picture composed of only two labels. It is made of two circles, two rectangles, and two squares. This picture constitutes our reference segmented image \mathbf{z}_0 and can be seen on Figure 3(a). Let $\epsilon_n(\mathbf{r})$ be a noise, we define here the signal-to-noise ratio (SNR) in dB as follows:

$$\text{SNR} = 10 \times \log_{10} \left(\frac{\sum_{\mathbf{r} \in \mathcal{R}} (f(\mathbf{r}) + \epsilon_n(\mathbf{r}))^2}{\sum_{\mathbf{r} \in \mathcal{R}} \epsilon_n(\mathbf{r})^2} \right). \quad (35)$$

Our idea about this construction is to start from a known discrete-value image \mathbf{z}_0 representing the labels of homogeneous regions in the image and shown on Figure 3(a), and simulate a HR image \mathbf{f} . For this step, we generated a colored noise with a SNR of 20.5 dB and added it to \mathbf{z}_0 to obtain \mathbf{f} . We used this image, which can be seen on Figure 3(b), as our reference HR image \mathbf{f}_0 . Then we apply the forward transformation equation (1), applying first a 5×5 Gaussian kernel lowpass filter \mathbf{B} to \mathbf{f}_0 . Resulting image is shown on Figure 3(c). We choose a decimation factor $d = 5$, and we construct $M = d^2 = 25$ LR images using the scheme illustrated on Figure 1 in the case $d = 2$. Thus, we are in an ideal

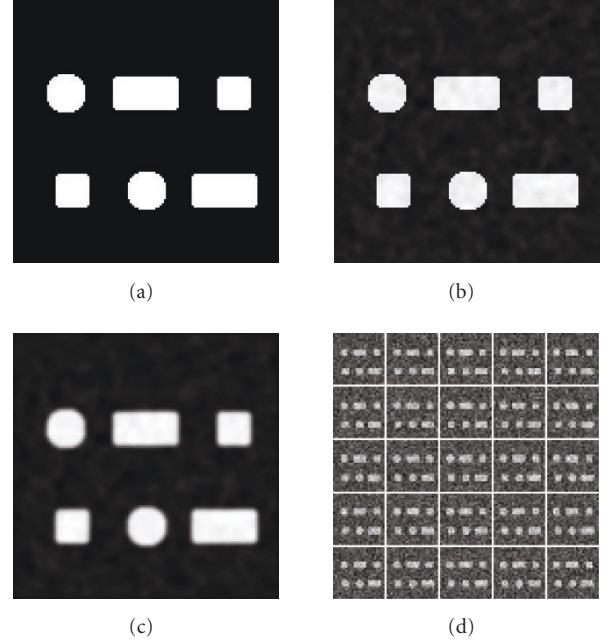


FIGURE 3: (a) \mathbf{z}_0 to obtain HR image \mathbf{f}_0 , (b) HR image \mathbf{f}_0 , (c) lowpass filtered image $\mathbf{B}\mathbf{f}_0$, and (d) noisy LR images $\mathbf{g}_i = \mathbf{H}_i\mathbf{f}_0 + \epsilon_i$.

case where we can observe M different LR images \mathbf{g}_i showing different views of a same HR image \mathbf{f}_0 . These LR images can be seen on Figure 2(c). Finally we add to each LR image \mathbf{g}_i independently a centered, white and Gaussian noise, obtained for a SNR of 5 dB. These LR noisy images, which constitute our simulated data, are on Figure 3(d). This choice of value for the noise parameters was a way of showing the robustness of our method in noise conditions. If it is working quite well with hard noise conditions, it will be working too with less noise.

We used the gradient phase correlation method that we evaluated in [3] for the registration process. For this ideal case, and without noise (Figure 2(c)), the registration process gives an accurate estimation of the shifts between all LR pictures. These information, and for a given interpolation factor d , allow to register and interpolate linearly pixel values on the HR grid. This step is illustrated on the second and third lines of Figure 1. We can also use one or two reference images to fill the unknown areas (e.g., the first and the last registered and interpolated LR pictures).

For the noisy case, the registration method does not give the same accuracy in the estimation of movement between LR images. The registration method evaluates correctly 28% of all shifts between pictures. For 30% of them, the wrong values were very close ($\pm 1/d$) to the exact move values.

For the choice of parameters and hyperparameters, we choose the PMRF parameter α equals to 2 in all the following simulations of our methods. We normalize all the original HR images \mathbf{f}_0 between 0 and 1, and we choose for all k : $m_{k_0} = 0.5$, $\sigma_{k_0}^2 = 0.1$, $\alpha_{k_0} = \alpha_0^\epsilon = 1$, and $\beta_{k_0} = \beta_0^\epsilon = 2$.

In the case of our method, we used a CGM model (12) for the labels because it gives better results than for a IGM

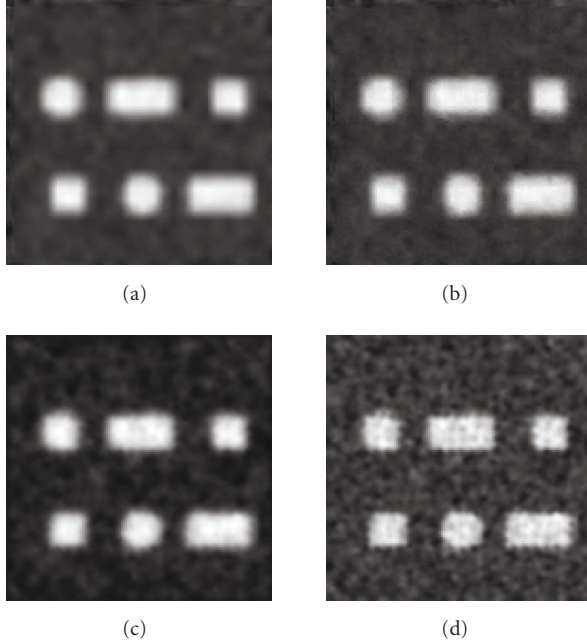


FIGURE 4: (a) $\hat{\mathbf{f}}_{\text{mean}}$, (b) $\hat{\mathbf{f}}_{\text{median}}$, (c) $\hat{\mathbf{f}}_{\text{Tikhonov } 1}$, $\lambda = 1$, and (d) $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$, $\lambda = 1$.

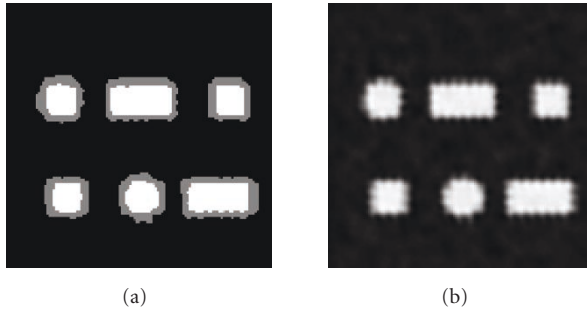


FIGURE 5: Results of segmentation $\hat{\mathbf{z}}_3$ (a), and reconstructed HR image $\hat{\mathbf{f}}_3$ (b) with $K = 3$ and for $\lambda = 1$, when pixels in given regions are assumed to be i.i.d.

model (11). Concerning the three algorithms presented in Section 5.2, we used Algorithm 2 with MAP-Gibbs sampling which provided us the better results. Algorithm 3 was particularly critical because, at that stage of our works, we imposed the number of classes to use in the segmentation part. So, a class becoming unrepresented can give a σ_k^2 very close to zero, making (22) or (23) to explode.

We compared the performances of our method with two other SR schemes. The first one uses a registration, a classical linear interpolation, and a summation [13]. If $\hat{\mathbf{f}}_i$ are the LR images \mathbf{g}_i ; registered and interpolated, we have $\hat{\mathbf{f}}_{\text{mean}}(\mathbf{r}) = (1/M) \sum_{i=1}^M \hat{\mathbf{f}}_i(\mathbf{r})$. On the same model, we can also replace the median of pixels instead of the mean giving $\hat{\mathbf{f}}_{\text{median}}(\mathbf{r})$. The second one is another classical method based on popular Tikhonov regularized approach [9–12]. This method is a deterministic approach utilizing regularization functions

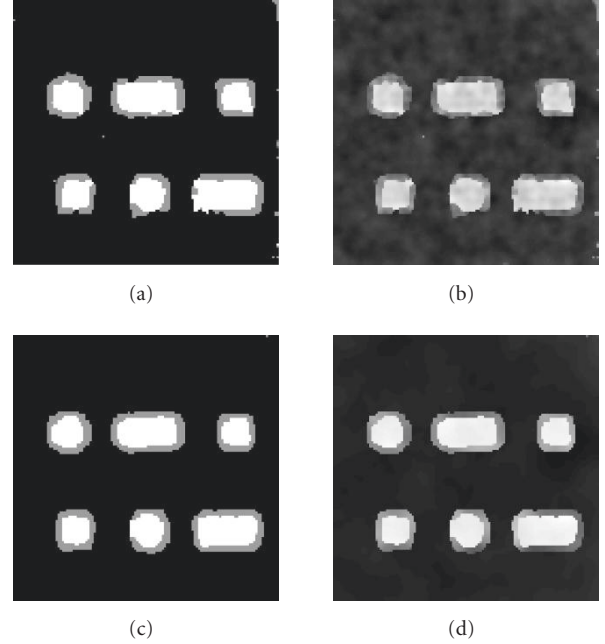


FIGURE 6: Results of segmentation $\hat{\mathbf{z}}_3$ with $K = 3$ (a) and (c), and reconstructed HR image $\hat{\mathbf{f}}_3$ (b) and (d), respectively, for $\lambda = 10$ and $\lambda = 10^2$, when pixels in given regions are assumed to be locally dependent.

to impose smoothness constraints on the space of feasible solutions. We implemented two versions of this method: $\hat{\mathbf{f}}_{\text{Tikhonov } 1}$, where we use classical derivative highpass kernel operator of size 3×3 to compute $\mathbf{D}\mathbf{f}$. $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$ is almost the same thing, except that it adapts \mathbf{D} for the pixels located on the corner of the pictures.

We define evaluation criteria using the difference image Δ between our reference HR image \mathbf{f}_0 and any HR image $\hat{\mathbf{f}}$ reconstructed from our noisy LR images \mathbf{g}_i . We used the estimation of the shifts obtained from these images. Thus, $\Delta(\hat{\mathbf{f}}, \mathbf{f}_0) = \hat{\mathbf{f}} - \mathbf{f}_0$. We also define $\Delta_\alpha(\hat{\mathbf{f}}, \mathbf{f}_0) = \sum_{\mathbf{r} \in \mathcal{R}} |\Delta(\hat{\mathbf{f}}(\mathbf{r}), \mathbf{f}_0(\mathbf{r}))|^\alpha / \sum_{\mathbf{r} \in \mathcal{R}} |\mathbf{f}_0(\mathbf{r})|^\alpha$, for $\alpha = \{1, 2\}$, as being L1 and L2 normalized relative error measures.

Because of our choice of HR image, which have regions with frank discontinuities, it seems more interesting to look at $\Delta_1(\hat{\mathbf{f}}, \mathbf{f}_0)$. So, for each case, we computed Δ_1 , the mean (which is the bias of our HR estimator) and standard deviation of the difference image $\Delta(\hat{\mathbf{f}}, \mathbf{f}_0)$: Δ_{mean} and Δ_{SD} . Moreover, for our methods, we will give percentage of error on $\hat{\mathbf{z}}$ label estimation.

We first experimented our method with a synthetic image obtained from a segmented image using 2 labels only. Figure 4 shows the results obtained for $\hat{\mathbf{f}}_{\text{mean}}$ (a), $\hat{\mathbf{f}}_{\text{median}}$ (b), $\hat{\mathbf{f}}_{\text{Tikhonov } 1}$ ($\lambda = 1$) (c), and $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$ ($\lambda = 1$) (d). Figure 5 shows the results obtained with our method when pixels in given regions are assumed to be i.i.d. and using (7) and (22). (a) and (b) show respectively $\hat{\mathbf{z}}$ and $\hat{\mathbf{f}}$, for $\lambda = 1$ and $K = 3$. Figure 6 shows the results obtained with our method when we consider a local dependency between pixels of a same

TABLE 1: Synthetic image made of 2 labels: evaluation criteria of SR methods.

Methods	$\Delta_{\text{mean}} (\times 10^{-2})$	Δ_{SD}	$\Delta_1 (\times 10^{-1})$	Error on estimation of labels (%)
Mean of pixel values, $\hat{\mathbf{f}}_{\text{mean}}$	0.74	0.125	3.36	×
Median of pixel values, $\hat{\mathbf{f}}_{\text{median}}$	0.63	0.124	3.40	×
$\hat{\mathbf{f}}_{\text{Tikhonov } 1}, \lambda = 1$	0.67	0.117	3.56	×
$\hat{\mathbf{f}}_{\text{Tikhonov } 2}, \lambda = 1$	0.06	0.128	4.24	×
$(\hat{\mathbf{f}}_3, \hat{\mathbf{z}}_3)$, i.i.d., $\lambda = 1$	0.41	0.105	2.04	9.09
$(\hat{\mathbf{f}}_3, \hat{\mathbf{z}}_3)$, dependent, $\lambda = 10$	1.03	0.134	3.45	7.17
$(\hat{\mathbf{f}}_3, \hat{\mathbf{z}}_3)$, dependent, $\lambda = 10^2$	3.86	0.118	3.31	7.58
$(\hat{\mathbf{f}}_2, \hat{\mathbf{z}}_2)$, dependent, $\lambda = 10$	3.35	0.135	2.80	2.67

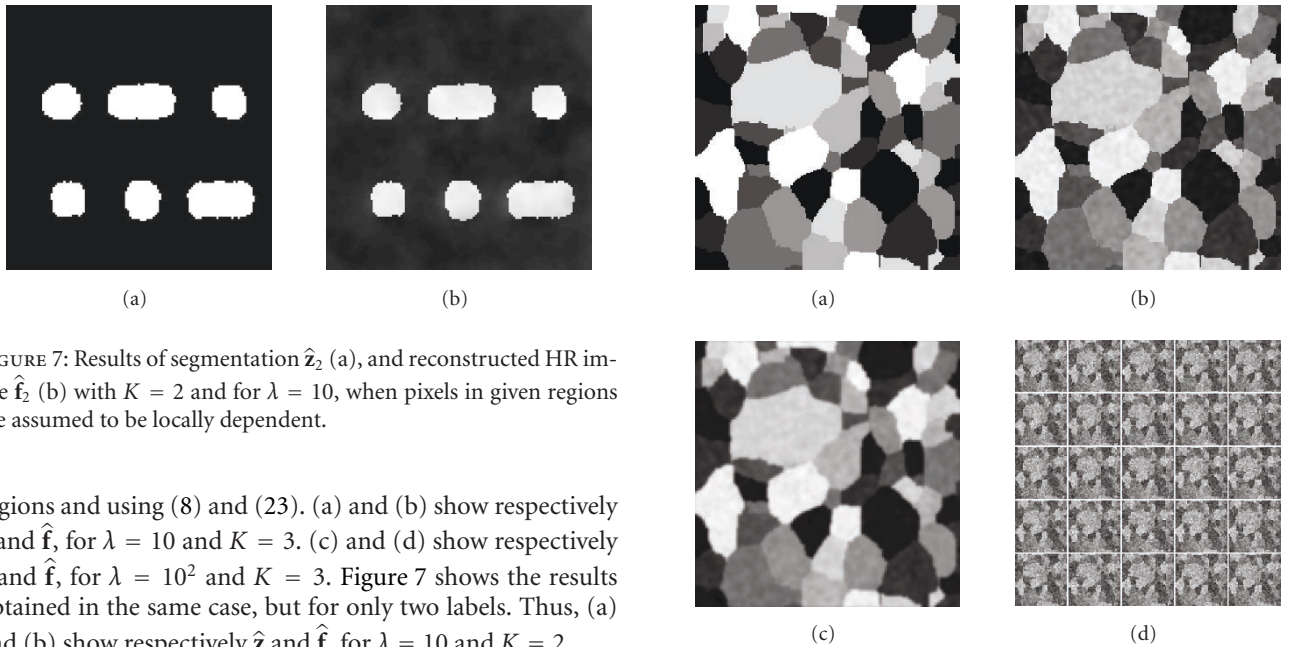


FIGURE 7: Results of segmentation $\hat{\mathbf{z}}_2$ (a), and reconstructed HR image $\hat{\mathbf{f}}_2$ (b) with $K = 2$ and for $\lambda = 10$, when pixels in given regions are assumed to be locally dependent.

regions and using (8) and (23). (a) and (b) show respectively $\hat{\mathbf{z}}$ and $\hat{\mathbf{f}}$, for $\lambda = 10$ and $K = 3$. (c) and (d) show respectively $\hat{\mathbf{z}}$ and $\hat{\mathbf{f}}$, for $\lambda = 10^2$ and $K = 3$. Figure 7 shows the results obtained in the same case, but for only two labels. Thus, (a) and (b) show respectively $\hat{\mathbf{z}}$ and $\hat{\mathbf{f}}$, for $\lambda = 10$ and $K = 2$.

Table 1 presents the values of our evaluation criteria for all the previous cases shown on Figures 4, 5, 6, and 7. Firstly, it is important to remind that our method is stochastic. So, another realization of our algorithm can give some other, but close, results. Next, our methods give a piece of information that usual methods do not give: a segmented image $\hat{\mathbf{z}}$. This knowledge about the HR image could be useful to a detection purpose. It is also useful to the reconstruction of the HR image. If we look at all these different reconstructed HR images $\hat{\mathbf{f}}$, it seems that our method considering a local dependency between pixels of the same region, with $K = 3$, $\lambda = 10^2$, and shown on Figure 6(d), highlights the contours of the shapes. These contours seem to be more sharp and contrasted with our methods than for the other compared methods. Indeed, the methods of comparison give images without sharp edges, but with blurred object's contours. Also, our method allows to distinguish the squares from the circles.

For our methods, on the point of view of the percentage of error on labels' estimation, it is not surprising to have less errors when we use 2 labels than when we use more labels since we made our original HR image from 2 labels.

FIGURE 8: (a) \mathbf{z}_0 to obtain HR image \mathbf{f}_0 , (b) HR image \mathbf{f}_0 , (c) lowpass filtered image \mathbf{Bf}_0 , and (d) noisy LR images $\mathbf{g}_i = \mathbf{H}_i \mathbf{f}_0 + \epsilon_i$.

We can also notice that our reconstruction seems robust to error of the estimation of the movement since we use an estimated movement vector which included errors. Finally, we remark that no one of our chosen criteria seems to be adapted to evaluate the quality of the reconstructed HR picture.

Next, we did the same kind of evaluation with a synthetic image of size 200×200 pixels² originally made of 8 different labels. So, we constructed our own LR images as in the previous case. We built \mathbf{f}_0 (Figure 8(b)) from a segmented image \mathbf{z}_0 (Figure 8(a)) to which we added a colored noise with a SNR of 27.2 dB. The blurred image \mathbf{Bf}_0 is on Figure 8(c), and the noisy $M = 25$ LR pictures, with a noise of 5 dB, can be seen on Figure 8(d). The registration method evaluates correctly 30% of all shifts between noisy pictures. All the results of the used methods are shown on Figures 9 and 10. Table 2 gives the evaluation criteria values.

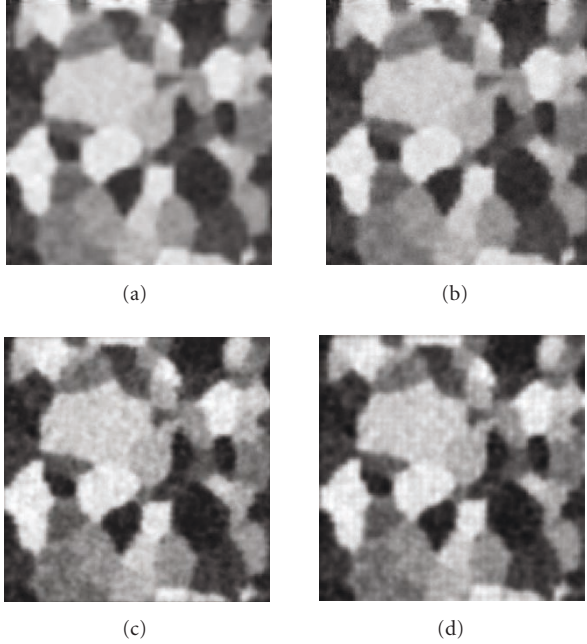


FIGURE 9: (a) $\hat{\mathbf{f}}_{\text{mean}}$, (b) $\hat{\mathbf{f}}_{\text{median}}$, (c) $\hat{\mathbf{f}}_{\text{Tikhonov } 1}, \lambda = 1$, and (d) $\hat{\mathbf{f}}_{\text{Tikhonov } 2}, \lambda = 1$.

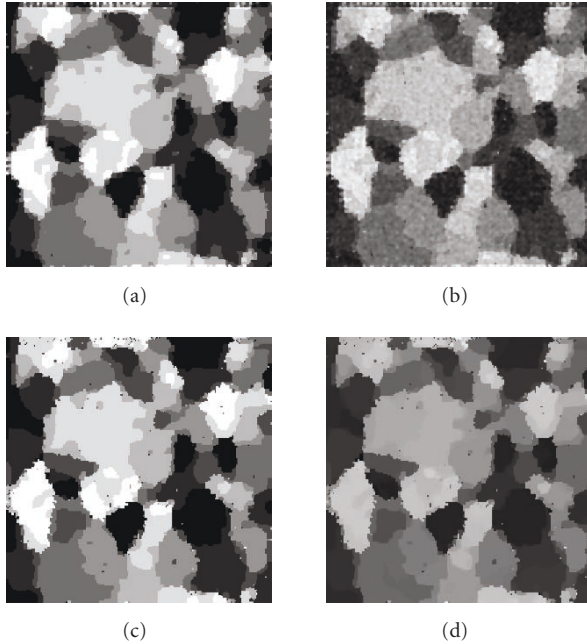


FIGURE 10: Results of segmentation $\hat{\mathbf{z}}_8$ with $K = 8$ (a) and (c), and reconstructed HR image $\hat{\mathbf{f}}_8$ (b) and (d), respectively, for pixels assumed to be i.i.d. and $\lambda = 1$, and for pixels assumed to be locally dependent and $\lambda = 10^3$.

Here again, our method that considers a local dependency between pixels of the same region, with $K = 3$ and $\lambda = 10^3$, gives a better reconstructed HR image if we look at the homogeneity of regions and their sharp contours.

Finally, we did the same experiment on a real image of size 250×250 pixels² taken from the sky, and that can be seen on Figure 11(b). We did a segmentation of this picture with $K = 8$ labels and used evenly distributed thresholds. It is shown on (a) of the same figure. Except for the colored noise that we did not add, we constructed our own LR images as before. The blurred image \mathbf{Bf}_0 is on Figure 11(c), and the noisy $M = 25$ LR pictures, with a noise of 5 dB, can be seen on Figure 11(d). The registration method evaluates correctly 28% of all shifts between noisy pictures. All the results of the used methods are shown on Figures 12 and 13. Table 3 gives the evaluation criteria values.

For this real and more detailed HR image, our method that assumes pixels are i.i.d., $K = 8$, $\lambda = 1$, and shown on Figure 13(b), seems to give a good reconstruction of the initial HR image $\hat{\mathbf{f}}$.

7. SIMULATION RESULTS OBTAINED WITH REAL VIDEO SEQUENCES

We used a Philips ToUCam Video WebCam to obtain three real video sequences. For each case, we used a super-resolution factor of $d = 4$. First, we did a 208 frames movie using Figure 14(a) as the main image of the video sequence. The registration process found all the 16 over 16 possible subpixel moves in all the frames of the sequence. Figure 14(b) shows one of the 38×64 pixels² LR real frames we used to compute HR images, (c) is the $\hat{\mathbf{f}}_{\text{mean}}$ solution, and (d) is the $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$ with $\lambda = 1$ solution.

Figure 15(a) shows the result of segmentation $\hat{\mathbf{z}}_2$, and (b) the reconstructed HR image $\hat{\mathbf{f}}_2$ when we use our method with $K = 2$ and $\lambda = 10^{-1}$. (c) shows the result of segmentation $\hat{\mathbf{z}}_3$, and (b) the reconstructed HR image $\hat{\mathbf{f}}_3$ when we use our method with $K = 3$ and $\lambda = 10^{-1}$.

The HR images obtained with our method are quite similar to the one obtained with the two other methods. However, if we look closely, it appears that object's contours are more accurate with our method (Figures 15(b) and 15(d)). Indeed, we also have the segmented images (a) and (c) of Figure 15 that show more precisely the object's forms, in particular, image (c) which is the segmentation's result when we use 3 labels.

The second movie is filmed using Figure 16(a) as the main image of the video sequence. The registration process found 14 over 16 possible subpixel moves in the 120 frames of the sequence. Figure 16(b) shows one of the 53×57 pixels² LR real frames we used to compute HR images, (c) is the $\hat{\mathbf{f}}_{\text{mean}}$ solution, and (d) is the $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$ with $\lambda = 1$ solution.

Figure 17(a) shows the result of segmentation $\hat{\mathbf{z}}_3$, and (b) the reconstructed HR image $\hat{\mathbf{f}}_3$ when we use our method with $K = 3$ and $\lambda = 10^{-1}$. (c) shows the result of segmentation $\hat{\mathbf{z}}_7$, and (b) the reconstructed HR image $\hat{\mathbf{f}}_7$ when we use our method with $K = 7$ and $\lambda = 10^{-1}$.

As in the previous example, we can notice that object's contours are more accurate with our method (Figures 17(b) and 17(d)). The segmented images (a) and (c) of Figure 17 show more precisely the object's real forms, in particular,

TABLE 2: Synthetic image using 8 labels: evaluation criteria of SR methods.

Methods	$\Delta_{\text{mean}} (\times 10^{-2})$	Δ_{SD}	$\Delta_1 (\times 10^{-1})$	Error on estimation of labels (%)
Mean of pixel values, $\hat{\mathbf{f}}_{\text{mean}}$	0.14	0.108	1.33	×
Median of pixel values, $\hat{\mathbf{f}}_{\text{median}}$	0.21	0.110	1.35	×
$\hat{\mathbf{f}}_{\text{Tikhonov } 1}, \lambda = 1$	0.34	0.114	1.48	×
$\hat{\mathbf{f}}_{\text{Tikhonov } 2}, \lambda = 10$	1.85	0.108	1.48	×
$(\hat{\mathbf{f}}_8, \hat{\mathbf{z}}_8)$, i.i.d., $\lambda = 1$	4.54	0.132	2.03	32.7
$(\hat{\mathbf{f}}_8, \hat{\mathbf{z}}_8)$, dependent, $\lambda = 10^3$	8.16	0.155	2.82	29.7

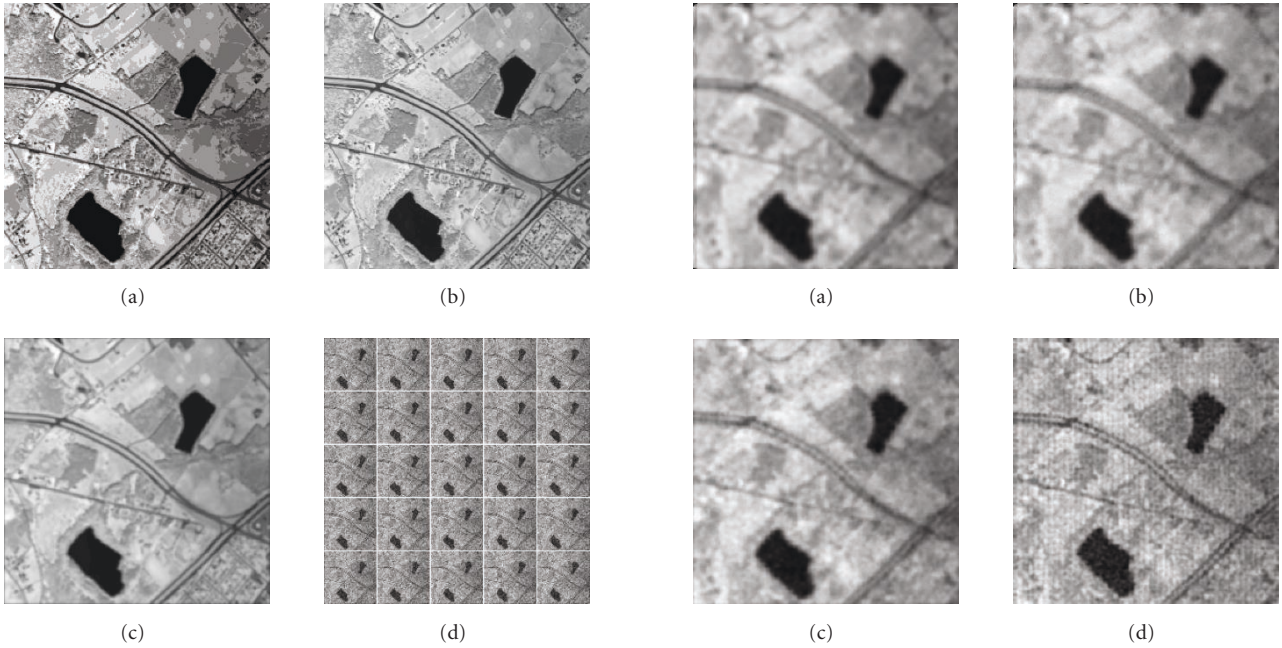


FIGURE 11: (a) \mathbf{z}_0 obtained by thresholding of the HR image \mathbf{f}_0 and $K = 8$ labels, (b) HR image \mathbf{f}_0 , (c) lowpass filtered image \mathbf{Bf}_0 , and (d) noisy LR images $\mathbf{g}_i = \mathbf{H}_i\mathbf{f}_0 + \epsilon_i$.

FIGURE 12: (a) $\hat{\mathbf{f}}_{\text{mean}}$, (b) $\hat{\mathbf{f}}_{\text{median}}$, (c) $\hat{\mathbf{f}}_{\text{Tikhonov } 1}, \lambda = 1$, and (d) $\hat{\mathbf{f}}_{\text{Tikhonov } 2}, \lambda = 10$.

image (c) which is the segmentation’s result when we use 7 labels.

The last video sequence is made of 313 frames. The registration process found all the 16 over 16 possible subpixel moves in all the frames of the sequence. Figure 18(b) shows one of the 70×69 pixels² LR real frames we used to compute HR images, (c) is the $\hat{\mathbf{f}}_{\text{mean}}$ solution, and (d) is the $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$ with $\lambda = 1$ solution.

Figure 19(a) shows the result of segmentation $\hat{\mathbf{z}}_3$, and (b) the reconstructed HR image $\hat{\mathbf{f}}_3$ when we use our method with $K = 3$ and $\lambda = 10^{-2}$. (c) shows the result of segmentation $\hat{\mathbf{z}}_6$, and (b) the reconstructed HR image $\hat{\mathbf{f}}_6$ when we use our method with $K = 6$ and $\lambda = 10^{-2}$.

This last video, showing more complicated real images, allows to see that our method works well. For example, the text “ PHILIPS” on both batteries can be clearly read on the HR images obtained with our method (Figures 19(b)

and 19(d)), but it is not so clear on the HR images obtained with other methods. Even the time on the digital clock can be read on the HR images obtained with our method, which is not the case on the other HR images. More, the segmentation’s result shown on images (a) and (c) of Figure 19 allows to have information about the different homogeneous regions of the HR image, which is not the case for the HR images obtained with the other methods.

From a computational point of view, our methods are more expensive in comparison to classical methods presented in this work. Indeed, in each main iteration of the algorithm we have chosen in Section 5.2, we have 3 steps:

- (1) estimation of $\mathbf{f} \mid \mathbf{z}, \boldsymbol{\theta}, \mathbf{g}$, which needs an optimization and costs approximately the same as any regularization technique and which can be estimated as about 50 times the cost of a forward model (\mathbf{Hf}),

TABLE 3: Image taken from the sky: evaluation criteria of SR methods.

Methods	$\Delta_{\text{mean}} (\times 10^{-2})$	Δ_{SD}	$\Delta_1 (\times 10^{-1})$	Error on estimation of labels (%)
Mean of pixel values, $\hat{\mathbf{f}}_{\text{mean}}$	0.24	0.104	1.41	×
Median of pixel values, $\hat{\mathbf{f}}_{\text{median}}$	0.27	0.106	1.44	×
$\hat{\mathbf{f}}_{\text{Tikhonov } 1}, \lambda = 1$	0.26	0.113	1.48	×
$\hat{\mathbf{f}}_{\text{Tikhonov } 2}, \lambda = 1$	0.20	0.102	1.44	×
$(\hat{\mathbf{f}}_8, \hat{\mathbf{z}}_8)$, i.i.d., $\lambda = 1$	0.90	0.117	1.61	13.0
$(\hat{\mathbf{f}}_8, \hat{\mathbf{z}}_8)$, dependent, $\lambda = 10^3$	8.76	0.122	2.27	13.8

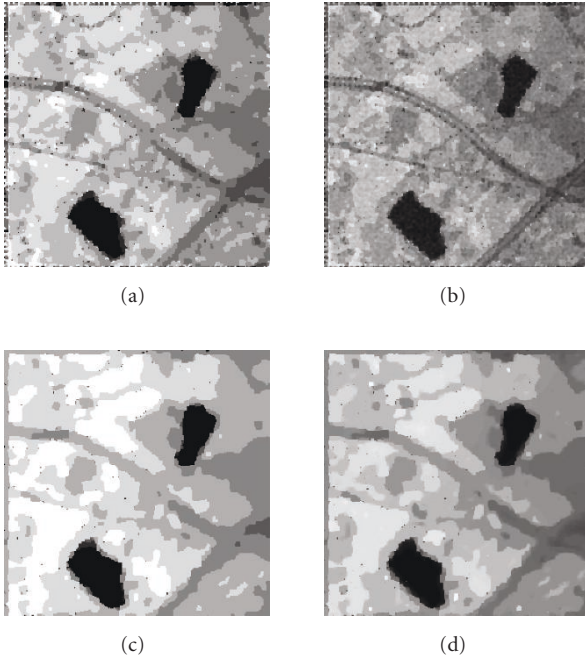


FIGURE 13: Results of segmentation $\hat{\mathbf{z}}_8$ with $K = 8$ (a) and (c), and reconstructed HR image $\hat{\mathbf{f}}_8$ (b) and (d), respectively, for pixels assumed to be i.i.d. and $\lambda = 1$, and for pixels assumed to be locally dependent and $\lambda = 10^3$.

- (2) sampling of \mathbf{z} which is not really too expensive, mainly the cost of the computation of a forward problem,
- (3) sampling of hyperparameters which is not expensive neither, mainly the cost of computation of the means and variances of the pixel values in each region.

More precisely, if we let X and Y be the size of LR images in pixels (width and height), M the number of LR images used in the SR process, K the number of labels used for the segmentation of the HR image $\hat{\mathbf{f}}_K$, and d the SR factor, then the calculation complexity of methods giving $\hat{\mathbf{f}}_{\text{mean}}$ and $\hat{\mathbf{f}}_{\text{median}}$ is $O(X, Y, M, d^2)$. It is the same complexity when $\hat{\mathbf{f}}_{\text{Tikhonov } 1}$ and $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$ are computed, even if an iterative process is performed.

For our methods, the calculation complexity is $O(X, Y, M, d^2, K)$. However, it is not easy to give an exact comparative cost computation. For this reason, we performed some experimental time comparison using the CPU time for the execution of these different methods. Table 4 gives the computation times (in seconds) we obtained using Matlab, on a Windows XP OS, with a Pentium IV processor at 2.6 GHz. We did these time estimations for the clock real sequence (whose results are shown on Figures 18 and 19), and the artificially generated sequence showing a view of the earth took from the sky (results shown on Figures 12 and 13). Table 4 also gives in each case the overall number of iterations realized to obtain $\hat{\mathbf{f}}$ in the full optimization process (first number between brackets), and the number of iterations done in our SR methods (second number between brackets). The two last lines of this table show that, with the absolute same conditions, the number of iterations in our algorithm can be very different, even if the obtained results are almost the same. This is due to the probabilistic approach of our methods.

We first noticed that there is a multiplicative factor ranging between 5 and 10 as regards the necessary computing time between the computation time of classical method $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$ and our method giving $(\hat{\mathbf{f}}_K, \hat{\mathbf{z}}_K)$. More, we can also see that the computational time needed to obtain the final result is a function of the number of iteration realized.

8. CONCLUSION

In this paper, we presented a new approach to the SR reconstruction problem. The forward model relating a LR image \mathbf{g}_i to the HR image \mathbf{f} is a classical one composed of lowpass filtering, a translational movement, decimation and corruption by a noise. The main idea in the inversion process is to propose a compound Gauss-Markov intensity process with a hidden Potts random field for the labels of the homogeneous regions in the image, and to use the fusion process which has been presented in [17], and in a more general way in [14], to obtain a HR image which can be considered as the fusion result of the LR images. This HR image is computed thanks to an optimization criterion, which uses information taken from the segmentation result. The proposed method, based on the Bayesian estimation framework, estimates jointly the HR image, its segmentation, and all the hyperparameters of

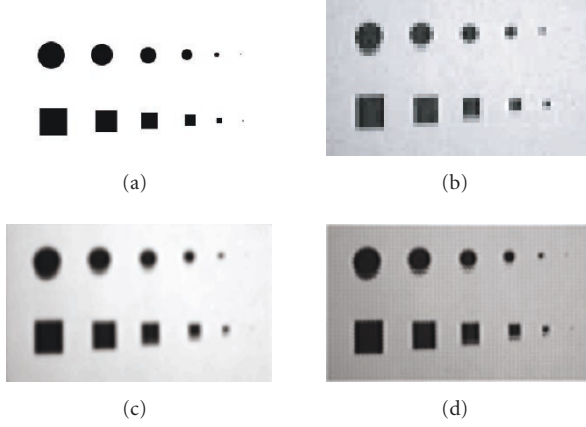


FIGURE 14: (a) Image used to obtain the video sequence, (b) one LR image (38×64 pixels²), (c) $\hat{\mathbf{f}}_{\text{mean}}$ (152×256 pixels²), and (d) $\hat{\mathbf{f}}_{\text{Tikhonov } 2}$, $\lambda = 1$.

the problem through a global Gibbs sampling scheme. One aspect of the originality of our work concerns the different kind of results we obtain with our stochastic SR method: we do not only obtain a reconstructed HR image $\hat{\mathbf{f}}$, but also a segmented image $\hat{\mathbf{z}}$ which can be useful for geometrical feature extraction and their tracking.

The perspectives of our future works on SR reconstruction methods are to first apply these methods on more complex real video sequences. Also, we want to use a more complete movement scheme to estimate the affine motion between LR images, including homothety and rotation. The authors of [18] propose such a more sophisticated movement model. Finally, we may also look to an adaptation of our SR methods for color images.

For now, the number of labels used in the segmentation process must be fixed. So, another research axis could concern an automated and clever evolution of the number of necessary labels to obtain the best reconstructed HR picture.

To end, it would be interesting to find a good evaluation criterion that will permit to characterize a good HR reconstructed image.

APPENDIX

A. EXPRESSIONS OF THE PRIOR PROBABILITY LAWS MENTIONED IN SECTION 3.3

$$\mathcal{N}(m_k | m_{k0}, \sigma_{k0}^2) = \frac{1}{\sqrt{2\pi\sigma_{k0}^2}} \exp\left[-\frac{1}{2}\left(\frac{m_k - m_{k0}}{\sigma_{k0}}\right)^2\right],$$

$$\mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}, \Sigma_k)$$

$$= \frac{|\Sigma_k|^{1/2}}{(2\pi)^{n_k/2}} \exp\left[-\frac{1}{2}(\mathbf{f}_k - m_k \mathbf{1})^t \Sigma_k^{-1} (\mathbf{f}_k - m_k \mathbf{1})\right], \quad (\text{A.1})$$

with n_k the number of pixels of \mathbf{f} that belong to the region

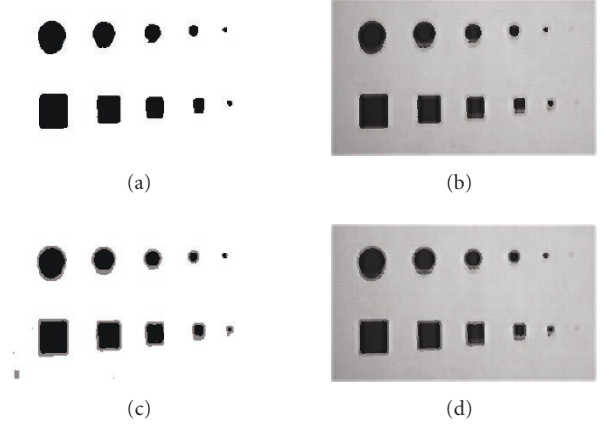


FIGURE 15: (a) Results of segmentation $\hat{\mathbf{z}}_2$ and (b) reconstructed HR image $\hat{\mathbf{f}}_2$ with $K = 2$ and $\lambda = 10^{-1}$, (c) results of segmentation $\hat{\mathbf{z}}_3$ and (d) reconstructed HR image $\hat{\mathbf{f}}_3$ with $K = 3$ and $\lambda = 10^{-1}$.

\mathcal{R}_k , $|\Sigma_k|$ the determinant of the matrix Σ_k , and $\mathbf{1}$ a vector of size n_k with all components equal to 1.

$$\mathcal{I}\mathcal{G}(\sigma_k^2 | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\sigma_k^2}\right)^{\alpha-1} \exp\left[-\beta \frac{1}{\sigma_k^2}\right], \quad (\text{A.2})$$

$$\mathcal{I}\mathcal{W}(\Sigma_k | \alpha, \Lambda) = \frac{|\Lambda|^\alpha}{\Gamma(\alpha)} (\Sigma_k)^{1-\alpha} \exp[-\Lambda \Sigma_k^{-1}].$$

B. EXPRESSIONS OF THE POSTERIOR PROBABILITY LAWS MENTIONED IN SECTIONS 5.1 AND 5.2

(i) Posterior law of $\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}, \mathbf{g}$:

$$p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta}, \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \mathbf{z}, \boldsymbol{\theta}) p(\mathbf{f} | \mathbf{z}, \boldsymbol{\theta})$$

$$\propto \mathcal{N}(\mathbf{g} | \mathbf{f}, \sigma_\epsilon^2 \mathbf{I}) \prod_{k=1}^K p(\mathbf{f}_k | \mathbf{z}_k, \boldsymbol{\theta})$$

$$\propto \mathcal{N}(\mathbf{g} | \mathbf{H}\mathbf{f}, \sigma_\epsilon^2 \mathbf{I}) \prod_{k=1}^K \mathcal{N}(\mathbf{f}_k | m_k \mathbf{1}, \sigma_k^2 \mathbf{I}_k) \quad (\text{B.1})$$

$$\propto \mathcal{N}(\mathbf{g} | \mathbf{H}\mathbf{f}, \sigma_\epsilon^2 \mathbf{I}) \mathcal{N}(\mathbf{f} | \mathbf{m}_z, \Sigma_z)$$

$$\propto \mathcal{N}(\mathbf{f} | \hat{\mathbf{f}}, \hat{\Sigma}),$$

with $\hat{\Sigma} = (\mathbf{H}^t \Sigma_\epsilon^{-1} \mathbf{H} + \Sigma_z^{-1})^{-1}$, and $\hat{\mathbf{f}} = \hat{\Sigma}(\mathbf{H}^t \Sigma_\epsilon^{-1} \mathbf{g} + \Sigma_z^{-1} \mathbf{m}_z)$, and where \mathbf{m}_z is a vector of the size of the image defined as $\mathbf{m}_z = [m_1 \mathbf{1}'_1, \dots, m_K \mathbf{1}'_K]'$, and where $\Sigma_z = \text{diag}[\Sigma_1, \dots, \Sigma_K]$. $\mathbf{1}'_k$ is a vector of size n_k with all components equal to 1, and Σ_k is a diagonal matrix with σ_k^2 as its diagonal elements.

(ii) Posterior laws of $\mathbf{z} | \mathbf{f}, \boldsymbol{\theta}, \mathbf{g}$, and $\mathbf{z} | \mathbf{g}, \boldsymbol{\theta}$:

$$p(\mathbf{z} | \mathbf{f}, \boldsymbol{\theta}, \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \mathbf{z}, \boldsymbol{\theta}) p(\mathbf{z})$$

$$\propto \mathcal{N}(\mathbf{g} | \mathbf{H}\mathbf{f}, \sigma_\epsilon^2 \mathbf{I}) p(\mathbf{z}), \quad (\text{B.2})$$

$$p(\mathbf{z} | \mathbf{g}, \boldsymbol{\theta}) \propto p(\mathbf{g} | \mathbf{z}, \boldsymbol{\theta}) p(\mathbf{z})$$

$$\propto \mathcal{N}(\mathbf{g} | \mathbf{H}\mathbf{m}_z, \mathbf{H}\Sigma_z \mathbf{H}^t + \Sigma_\epsilon) p(\mathbf{z}).$$

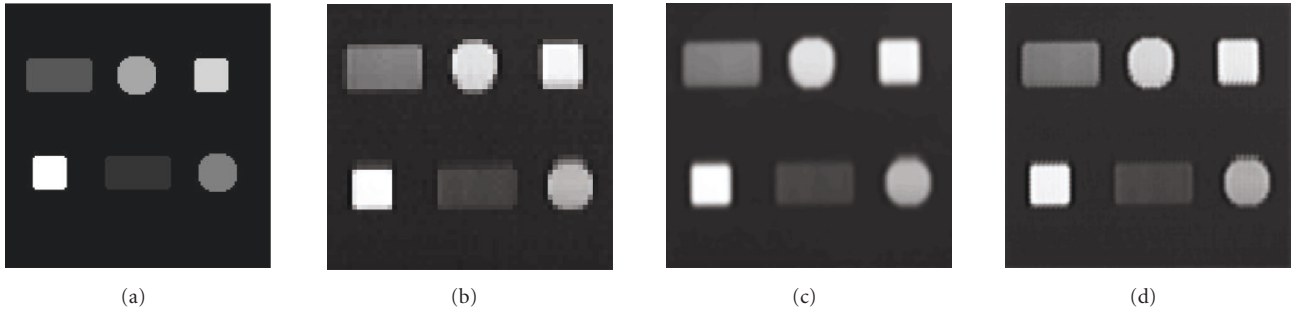


FIGURE 16: (a) Image used to obtain the video sequence, (b) one LR image (53×57 pixels²), (c) \hat{f}_{mean} (212×228 pixels²), and (d) $\hat{f}_{\text{Tikhonov2}}$, $\lambda = 1$.

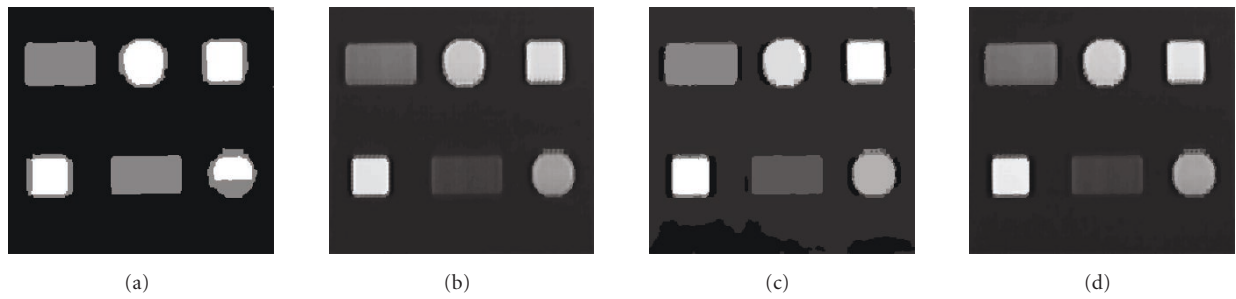


FIGURE 17: (a) Results of segmentation \hat{z}_3 and (b) reconstructed HR image \hat{f}_3 with $K = 3$ and $\lambda = 10^{-1}$, (c) results of segmentation \hat{z}_7 and (d) reconstructed HR image \hat{f}_7 with $K = 7$ and $\lambda = 10^{-1}$.

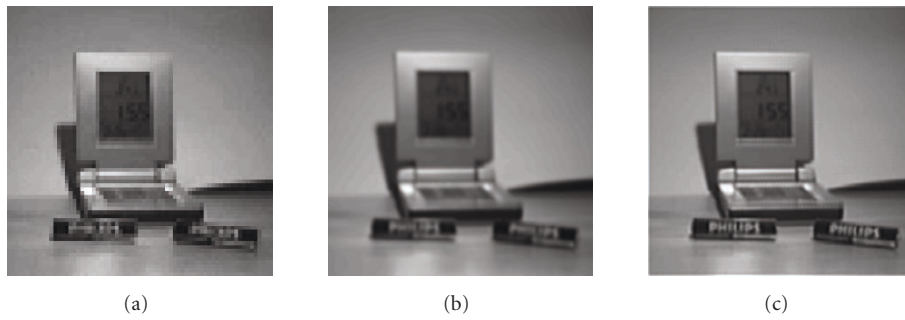


FIGURE 18: (a) One LR image (70×69 pixels²), (b) \hat{f}_{mean} (280×276 pixels²), and (c) $\hat{f}_{\text{Tikhonov2}}$, $\lambda = 1$.

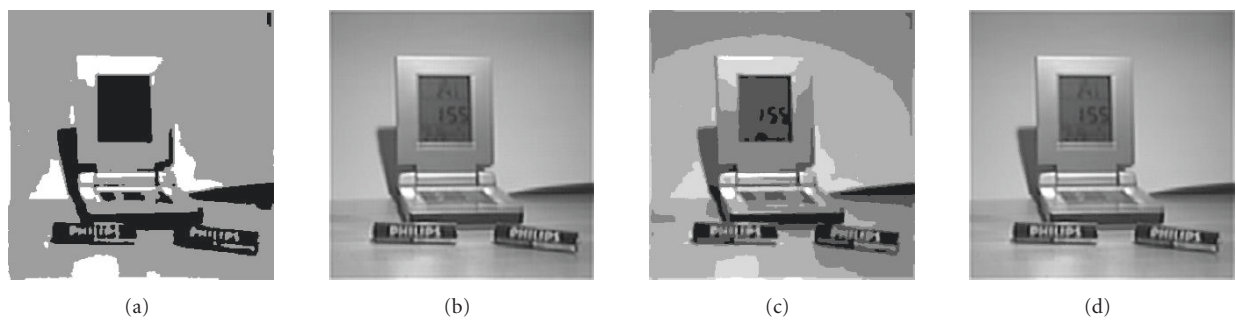


FIGURE 19: (a) Results of segmentation \hat{z}_3 and (b) reconstructed HR image \hat{f}_3 with $K = 3$ and $\lambda = 10^{-2}$, (c) results of segmentation \hat{z}_6 and (d) reconstructed HR image \hat{f}_6 with $K = 6$ and $\lambda = 10^{-2}$.

TABLE 4: Computation times (in seconds) and number of iterations (to compute $\hat{\mathbf{f}}$, in the iterative algorithm) obtained using different SR methods.

Methods	Clock sequence (Figures 18, 19) (X, Y, M, d) = (70, 69, 16, 4)		View from sky (Figures 12, 13) (X, Y, M, d) = (50, 49, 13, 4)	
	Computation time	Iterations (No.)	Computation time	Iterations (No.)
[2pt] Mean of pixel values, $\hat{\mathbf{f}}_{\text{mean}}$	6.3	(×, ×)	3.0	(×, ×)
Median of pixel values, $\hat{\mathbf{f}}_{\text{median}}$	15.6	(×, ×)	7.5	(×, ×)
$\hat{\mathbf{f}}_{\text{Tikhonov } 2}, \lambda = 1$	8.1	(26, ×)	3.2	(24, ×)
$(\hat{\mathbf{f}}_3, \hat{\mathbf{z}}_3)$, i.i.d., $\lambda = 10^{-3}$	22.0	(41, 17)	8.1	(37, 11)
$(\hat{\mathbf{f}}_6, \hat{\mathbf{z}}_6)$, i.i.d., $\lambda = 10^{-3}$	24.3	(36, 12)	11.4	(37, 11)
$(\hat{\mathbf{f}}_3, \hat{\mathbf{z}}_3)$, dependent, $\lambda = 10^{-2}$	42.2	(41, 16)	13.6	(34, 6)
$(\hat{\mathbf{f}}_6, \hat{\mathbf{z}}_6)$, dependent, $\lambda = 10^{-2}$	32.0	(29, 4)	20.4	(36, 6)
$(\hat{\mathbf{f}}_6, \hat{\mathbf{z}}_6)$, dependent, $\lambda = 10^{-2}$	53.0	(38, 13)	30.0	(44, 14)

(iii) Posterior laws of $\boldsymbol{\theta} \mid \mathbf{z}, \mathbf{f}, \mathbf{g}$:

$$p(\boldsymbol{\theta} \mid \mathbf{z}, \mathbf{f}, \mathbf{g}) = \prod_{i=1}^M p(\sigma_{\epsilon_i}^2 \mid \mathbf{f}, \mathbf{g}_i) \prod_{k=1}^K p(m_k \mid \sigma_k^2, \mathbf{f}, \mathbf{z}) p(\sigma_k^2 \mid m_k, \mathbf{f}, \mathbf{z}),$$

$$p(m_k \mid \sigma_k^2, \mathbf{f}, \mathbf{z}) \propto p(\mathbf{f} \mid m_k, \sigma_k^2, \mathbf{z}) p(m_k \mid \sigma_k^2, \mathbf{z}) \\ \propto \mathcal{N}(\mathbf{f}_k \mid m_k \mathbf{1}, \sigma_k^2 \mathbf{I}) \times \mathcal{N}(m_k \mid m_{k0}, \sigma_{k0}^2) \\ \sim \mathcal{N}(m_k \mid m_{k0}^{\text{apost}}, \sigma_{k0}^{2 \text{apost}})$$

$$\propto \left(\frac{1}{2\pi\sigma_k^2} \right)^{n_k/2} \exp \left[-\frac{1}{2\sigma_k^2} \|\mathbf{f}_k - m_k \mathbf{1}\|^2 \right] \\ \times \frac{1}{\sqrt{2\pi\sigma_{k0}^2}} \exp \left[-\frac{1}{2\sigma_{k0}^2} |m_k - m_{k0}|^2 \right],$$

$$\frac{\delta p(m_k \mid \sigma_k^2, \mathbf{f}, \mathbf{z})}{\delta m_k} = 0 \Rightarrow \frac{1}{\sigma_k^2} \sum_{\mathbf{r} \in \mathcal{R}_k} f(\mathbf{r}) + \frac{m_{k0}}{\sigma_{k0}^2} \\ = m_k \left(\frac{n_k}{\sigma_k^2} + \frac{1}{\sigma_{k0}^2} \right), \\ \sigma_{k0}^{2 \text{apost}} = \left(\frac{n_k}{\sigma_k^2} + \frac{1}{\sigma_{k0}^2} \right)^{-1},$$

$$m_{k0}^{\text{apost}} = \sigma_{k0}^{2 \text{apost}} \left(\frac{1}{\sigma_k^2} \sum_{\mathbf{r} \in \mathcal{R}_k} f(\mathbf{r}) + \frac{m_{k0}}{\sigma_{k0}^2} \right),$$

$$p(\sigma_k^2 \mid m_k, \mathbf{f}, \mathbf{z})$$

$$\propto p(\mathbf{f} \mid \sigma_k^2, m_k, \mathbf{z}) p(\sigma_k^2 \mid m_k, \mathbf{z}) \\ \propto \mathcal{N}(\mathbf{f}_k \mid m_k \mathbf{1}, \sigma_k^2 \mathbf{I}) \times \mathcal{IG}(\sigma_k^2 \mid \alpha_{k0}, \beta_{k0}) \\ \sim \mathcal{IG}(\sigma_k^{2 \text{apost}} \mid \alpha_{k0}^{\text{apost}}, \beta_{k0}^{\text{apost}}) \\ \propto \left(\frac{1}{2\pi\sigma_k^2} \right)^{n_k/2} \exp \left[-\frac{1}{2\sigma_k^2} \|\mathbf{f}_k - m_k \mathbf{1}\|^2 \right]$$

(B.3)

$$\times \left(\frac{1}{\sigma_k^2} \right)^{\alpha_{k0}-1} \exp \left[-\beta_{k0} \frac{1}{\sigma_k^2} \right] \\ \propto \left(\frac{1}{\sigma_k^2} \right)^{\alpha_{k0}+(n_k/2)-1} \exp \left[-\frac{1}{\sigma_k^2} \left(\beta_{k0} + \frac{\|\mathbf{f}_k - m_k \mathbf{1}\|^2}{2} \right) \right], \\ \alpha_{k0}^{\text{apost}} = \alpha_{k0} + \frac{n_k}{2},$$

$$\beta_{k0}^{\text{apost}} = \beta_{k0} + \frac{1}{2} \sum_{\mathbf{r} \in \mathcal{R}_k} |f(\mathbf{r}) - m_k|^2,$$

$$p(\sigma_{\epsilon_i}^2 \mid \mathbf{f}, \mathbf{g}_i) \propto p(\mathbf{g}_i \mid \sigma_{\epsilon_i}^2, \mathbf{f}) p(\sigma_{\epsilon_i}^2 \mid \mathbf{f}) \\ \propto \mathcal{N}(\mathbf{g}_i \mid \mathbf{H}_i \mathbf{f}, \sigma_{\epsilon_i}^2 \mathbf{I}) \times \mathcal{IG}(\sigma_{\epsilon_i}^2 \mid \alpha_0^{\epsilon_i}, \beta_0^{\epsilon_i}) \\ \sim \mathcal{IG}(\sigma_{\epsilon_i}^2 \mid \alpha_0^{\epsilon_i \text{apost}}, \beta_0^{\epsilon_i \text{apost}}) \\ \propto \left(\frac{1}{2\pi\sigma_{\epsilon_i}^2} \right)^{n/2} \exp \left[-\frac{1}{2\sigma_{\epsilon_i}^2} \|\mathbf{g}_i - \mathbf{H}_i \mathbf{f}\|^2 \right] \\ \times \left(\frac{1}{\sigma_{\epsilon_i}^2} \right)^{\alpha_0^{\epsilon_i}-1} \exp \left[-\beta_0^{\epsilon_i} \frac{1}{\sigma_{\epsilon_i}^2} \right] \\ \propto \left(\frac{1}{\sigma_{\epsilon_i}^2} \right)^{\alpha_0^{\epsilon_i}+(n/2)-1} \exp \left[-\frac{1}{\sigma_{\epsilon_i}^2} \left(\beta_0^{\epsilon_i} + \frac{\|\mathbf{g}_i - \mathbf{H}_i \mathbf{f}\|^2}{2} \right) \right], \\ \alpha_0^{\epsilon_i \text{apost}} = \alpha_0^{\epsilon_i} + \frac{n}{2}, \\ \beta_0^{\epsilon_i \text{apost}} = \beta_0^{\epsilon_i} + \frac{1}{2} \sum_{\mathbf{r} \in \mathcal{R}} |g_i(\mathbf{r}) - [\mathcal{H}_i f](\mathbf{r})|^2.$$

(B.4)

ACKNOWLEDGMENTS

This work was supported by the French Department of Defense (DGA) with the Department *Géographie Imagerie Perception* (GIP), and the French Department of Research (CNRS) with the *Laboratoire des Signaux et Systèmes* (L2S).

REFERENCES

- [1] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, 2002.
- [2] V. Argyriou and T. Vlachos, "Sub-pixel motion estimation using gradient cross-correlation," in *Proceedings of 7th IEEE International Symposium on Signal Processing and its Applications (ISSPA '03)*, vol. 2, pp. 215–218, Paris, France, July 2003.
- [3] F. Humblot, B. Collin, and A. Mohammad-Djafari, "Evaluation and practical issues of subpixel image registration using phase correlation methods," in *Proceedings of Physics in Signal and Image Processing (PSIP '05)*, Toulouse, France, January–February 2005.
- [4] R. Molina, J. Mateos, A. K. Katsaggelos, and M. Vega, "Bayesian multichannel image restoration using compound Gauss-Markov random fields," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1642–1654, 2003.
- [5] R. Molina, M. Vega, J. Abad, and A. K. Katsaggelos, "Parameter estimation in Bayesian high-resolution image reconstruction with multisensors," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1655–1667, 2003.
- [6] R. Y. Tsai and T. S. Huang, "Multi-frame image restoration and registration," in *Advances in Computer Vision on Image Processing*, vol. 1, pp. 317–339, 1984.
- [7] T. J. Schulz, "Multiframe image restoration," in *Handbook of Image and Video Processing*, A. Bovik, Ed., chapter 3.8, pp. 175–189, Academic Press, New York, NY, USA, 2000.
- [8] S. Borman, *Topics in multiframe super-resolution restoration*, Ph.D. thesis, University of Notre Dame, Notre Dame, Ind, USA, May 2004.
- [9] M.-C. Hong, M. G. Kang, and A. K. Katsaggelos, "A regularized multichannel restoration approach for globally optimal high resolution video sequence," in *Visual Communications and Image Processing (VCIP '97)*, vol. 3024 of *Proceedings of SPIE*, pp. 1306–1316, San Jose, Calif, USA, February 1997.
- [10] M. Elad and A. Feuer, "Restoration of a single super-resolution image from several blurred, noisy, and undersampled measured images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1646–1658, 1997.
- [11] N. Nguyen, P. Milanfar, and G. Golub, "A computationally efficient super-resolution image reconstruction algorithm," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 573–583, 2001.
- [12] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super-resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.
- [13] J. C. Gillette, T. M. Stadtmiller, and R. C. Hardie, "Aliasing reduction in staring infrared imagers utilizing subpixel techniques," *Optical Engineering*, vol. 34, no. 11, pp. 3130–3137, 1995.
- [14] H. Snoussi and A. Mohammad-Djafari, "Fast joint separation and segmentation of mixed images," *Journal of Electronic Imaging*, vol. 13, no. 2, pp. 349–361, 2004.
- [15] H. Snoussi and A. Mohammad-Djafari, "Information geometry and prior selection," in *Proceedings of 22nd International workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt '02)*, C. J. Williams, Ed., pp. 307–327, Moscow, Idaho, USA, August 2002.
- [16] H. Snoussi, *A Bayesian approach to source separation. Applications in imagery*, Ph.D. thesis, University of Paris-Sud, Orsay, France, September 2003.
- [17] O. Féron and A. Mohammad-Djafari, "Image fusion and unsupervised joint segmentation using a HMM and MCMC algorithm," *Journal of Electronic Imaging*, vol. 14, no. 2, pp. 1–12, 2005.
- [18] G. Rochefort, F. Champagnat, G. Le Besnerais, and J.-F. Giovannelli, "Super-resolution from a sequence of undersampled image under affine motion," submitted to *IEEE Transactions on Image Processing*, March 2005.

Fabrice Humblot received the B.S. degree in computer science from the *École Supérieure de Technologie Électrique*, Noisy-le-Grand, France, in 1999 and he studied during the 1998-99 academic year in Harvey Mudd College, Claremont, California. He received the Engineering degree in electronic and computer science from the *École Supérieure d'Ingénieurs en Électronique et Électrotechnique (ESIEE)*, Noisy-le-Grand, France, in 2002; the M.S. degree in digital communications systems from the *École Nationale Supérieure des Télécommunications (ENST)*, Paris, France, in 2002; and the Ph.D. degree in physics from the *Université Paris-Sud (UPS)*, Orsay, France, in 2005. Since then, he has been working at the *Laboratoire des Signaux et Systèmes (L2S)*, *École Supérieure d'Électricité (Supélec)*, Gif sur Yvette, France. His main research fields include signal and image processing, detection in imagery, image registration, Bayesian techniques for inverse problems, and super-resolution image reconstruction methods.



Ali Mohammad-Djafari received the B.S. degree in electrical engineering from Polytechnique of Teheran in 1975, the Diploma degree (M.S.) from *École Supérieure d'Électricité (Supélec)*, Gif sur Yvette, France, in 1977, and the "Docteur Ingénieur" (Ph.D.) degree and "Doctorat d'Etat" in physics from the *Université Paris-Sud (UPS)*, Orsay, France, in 1981 and 1987, respectively. He was Associate Professor at UPS for two years (1981–1983). Since 1984, he has had a permanent position at "Centre National de la Recherche Scientifique (CNRS)" and works at "Laboratoire des Signaux et Systèmes" (L2S) at Supélec. From 1998 to 2002, he has been the Head of Signal and Image Processing Division at this laboratory. Presently, he is "Directeur de Recherche," and his main scientific interests are in developing new probabilistic methods based on information theory, maximum entropy, and Bayesian inference approaches for inverse problems in general, and more specifically; image reconstruction, signal and image deconvolution, blind source separation, data fusion, multi- and hyperspectral image segmentation. The main application domains of his interests are computed tomography (X-rays, PET, SPECT, MRI, microwave, ultrasound, and eddy current imaging) either for medical imaging or for nondestructive testing (NDT) in industry.

