The Computer Journal Advance Access published March 7, 2008

doi:10.1093/comjnl/bxn005

# Super-Resolution: A Short Review, A New Method Based on Hidden Markov Modeling of HR Image and Future Challenges

ALI MOHAMMAD-DJAFARI*

*Laboratoire des signaux et systèmes (L2S), UMR 8506 du CNRS-SUPELEC-Univ Paris Sud 11, Supélec, plateau de Moulon, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France*
**Corresponding author: djafari@lss.supelec.fr*

**Super-resolution (SR) is the area of research and development which produces one or a set of high-resolution images from one or a set of low-resolution frames. In this paper, first, a short review of a variety of SR problems is presented. Then, starting by a single input single output case, we present different forward modeling of 1D or 2D SR problems. We focus then on the multi input single output and multi input multi output SR problems and provide a summary of recent contributions to them. Then, the SR problem is considered as an inverse problem. A general forward-modeling and inversion framework is presented, which gives the possibility to understand the basics of several classical SR methods and to discuss some important open problems of SR. Specifically, we discuss a different forward modeling, which leads to different classical methods and present our recent inversion methods based on the Bayesian estimation with different prior modeling. In particular, we give the details of a new method, particularly appropriate for piecewise homogeneous images, which provides not only an SR image, but also simultaneously an optimal segmentation of an HR image. Some comparisons of the relative performances of these methods are also presented. Finally, some future challenges in SR are outlined and discussed.**

*Keywords: super-resolution; inverse problem; motion estimation; robust estimation; regularization; MISO; MIMO*

*Received 28 December 2006; revised 9 January 2008*

## 1. INTRODUCTION

The quest for obtaining high-resolution (HR) fixed images or image sequences from one or a set of low-resolution (LR) acquisition systems is a challenge in both hardware and software of electrical engineering and computer science [1–18]. While in many imaging systems, hardware advances in producing higher and higher-resolution sensors, and greater and greater capacity for memories, the progress of software and appropriate algorithms to handle those data requires still more research and development. This is due to the fact that, even if the prices of HR sensors decreases, the transmission and processing of such images may still cost enough. Also, in many medical diagnostic systems or in industrial non-destructive testing (NDT) systems, the acquisition of HR images may still cost enough to consider the problem of SR as an important area of research. Finally, in any situation, we always want to extract more and more details from the available images, whatever their resolutions.

Examples of applications where super-resolution (SR)-based techniques have become a focus of research in image processing are:

- Embedded LR imaging devices, such as hand-held computers and mobile phones, where we may need to reconstruct an HR image from an LR sequence of images accurately and quickly [19];
- Multi-camera and multi-view recording in aerial or satellite imaging [20–23];
- Many medical and biological imaging systems where we always want to obtain higher-resolution 2D or 3D images by combining different images obtained from the same object in different contexts (at different times, different viewing angles or different energy levels) [15, 22, 24];

– Holographic and 3D TV imaging where the data transmission limits the maximum resolution, and so the reconstruction of an HR 3D scene from LR data becomes crucial;
– 3D photography and surface modeling for 3D scenes [25, 26].

Accordingly SR problems can be classified into:

– single input single output (SISO) SR which can be considered as an interpolation problem;
– multi input single output (MISO) SR which is the classical SR problem on which we are going to focus and
– multi input multi output (MIMO) SR which is, for example, the case of video SR reconstruction and an example of MISO 3D SR problems in NDT applications.

The SISO SR problem can either be considered as an interpolation or, more generally, as an image restoration (deconvolution) problem. However, as this SISO SR can be considered as a particular case of the MISO SR problem, we focus on this last problem which concerns the integration of multiple LR frames to estimate one HR image. The extension to MIMO can also be done easily.

In MISO SR problem, the main idea is based on the fact that each image in the sequence provides small amount of additional information about the original HR image. This means that we assume that the LR images are obtained either by one camera with a slightly changing scene or with a moving camera focusing on a fixed scene, or by different cameras with different viewing positions and angles. Although other situations in medical, astronomical, electronic microscopy or NDT imaging may be present, in this paper we focus on translational movement cases.

The organization of this article is as follows. In Section 2 we detail the forward modeling of the SR problem, trying to model the main operations applied to an HR image to obtain LR images. Also, for each operator we give the expression of its adjoint operator. Using this forward and adjoint operators, we will see in Section 3 that many classical SR methods are based on different combination of these adjoint operators. In Section 4, we use again the general structure of the forward problem to consider the SR as an inverse problem and summarize the main classical methods based on least square (LS), constrained least square, quadratic regularization (QR) and robust regularization (RR) criteria optimization. In addition, we will see that the Bayesian maximum *a posteriori* (MAP) estimation method generalizes all these methods. We will also see how to handle the two main difficult tasks in SR, which are the estimation of the blurring point spread function (PSF) and the parameters of movement and registration of images. In Section 5, we summarize a more advanced Bayesian estimation method with a more sophisticated (and so more appropriate and more accurate) prior modeling of the HR image which accounts for the fact that, in general, all images are composed of statistically homogeneous regions. A compound intensity-regions Markov model is presented to account for this fact. Even if this method has been recently presented elsewhere [27, 28], we give here new extensions and new implementation algorithms. In Section 6, we discuss limitations of the existing forward models, inversion methods and new challenges for SR problems.

## 2. FORWARD MODELING

In any SR problem, there are three main operations which link an HR image to LR images which are: sampling, movement or other geometrical transformation and blurring. In this section, we give a brief description of these operations and their corresponding adjoint operators.

### 2.1. Sampling basis functions

To be able to explain a great number of SR methods, we consider first the simplest case which is the SISO case and start by modeling the process of the transformation $\mathcal{H}$ of an HR image $f(\boldsymbol{r})$ to an LR image $g(\boldsymbol{r})$. As we work with discretized images, let us denote the HR image $f(\boldsymbol{r})$ as

$$f(\boldsymbol{r}) = \sum_{j=1}^{n} f_j \, \delta(\boldsymbol{r} - \boldsymbol{r}_j) \tag{1}$$

where $\boldsymbol{r} = (x, y)$ is any position in space, $\boldsymbol{r}_j = (x_j, y_j)$ is the central position of the pixel $j$ assuming to have $(\Delta x \times \Delta y)$ as its size, $\delta(\boldsymbol{r} - \boldsymbol{r}_j)$ is a basis function of the form

$$\delta_1(\boldsymbol{r} - \boldsymbol{r}_j) = \begin{cases} 1 & \text{if } \boldsymbol{r} = \boldsymbol{r}_j \\ 0 & \text{else} \end{cases} \tag{2}$$

or

$$\delta_2(\boldsymbol{r} - \boldsymbol{r}_j) = \begin{cases} 1 & \text{if} \quad \begin{matrix} |x - x_j| < \frac{\Delta x}{2} \\ |y - y_j| < \frac{\Delta y}{2} \end{matrix} \\ 0 & \text{else} \end{cases} \tag{3}$$

and $f_j$ represents either the sample value $f(\boldsymbol{r}_j)$

$$f_j = f(\boldsymbol{r}_j) \tag{4}$$

or the mean value of $f(\boldsymbol{r})$ over the pixel surface

$$f_j = \int_{x_j - (\Delta x/2)}^{x_j + (\Delta x/2)} \mathrm{d}x \int_{y_j - (\Delta y/2)}^{y_j + (\Delta y/2)} \mathrm{d}y \, f(x, y) \tag{5}$$

depending on the choice of the basis function $\delta_1(\boldsymbol{r} - \boldsymbol{r}_j)$ or $\delta_2(\boldsymbol{r} - \boldsymbol{r}_j)$.

From now on, we consider that the HR image is discretized at the best possible resolution ($\Delta x = 1$, $\Delta y = 1$) and that the LR images are discretized with a ($\Delta x = k$, $\Delta y = k$), where $k > 1$ is a real or integer factor. Now, depending on the type of this factor (real or integer), the type of the basis function and the relative size of the LR and HR imaging sensors, we will have different forward modeling for the SR inverse problem.

### 2.2. Single input single output case

The first model for LR image $g(\boldsymbol{r})$ is just an integer value $k$ down-sampling (DS) of $f(\boldsymbol{r})$

$$g(\boldsymbol{r}') = \sum_{j'=1}^{n/k} g_{j'}\, \delta(\boldsymbol{r}' - \boldsymbol{r}'_{j'}) \tag{6}$$

where

$$g_{j'} = f_j \quad \text{with} \quad j' = k\, j \tag{7}$$

which means that the LR samples $g_{j'}$ are obtained just by DS of the HR samples $f_j$. This forward model, which will be presented by the operator $\mathcal{D}_0$ is shown in Fig. 1.

The second simple model for LR image $g(\boldsymbol{r})$ is just an integer value ($k$) DS of $f(\boldsymbol{r})$ but with considering the LR sensor size which is assumed to be the same ($k$) times the size of the original image pixel size. This forward model which is presented by the operator $\mathcal{D}_1$ is shown in Fig. 2.

The third simple model for LR image $g(\boldsymbol{r})$ is where the sub-sampling factor ($k$) is no longer an integer value. Moreover, we consider not only the LR sensor size but also a more general blurring effect $\mathcal{B}$ due to LR pixel size and any other imaging system blurring effects. In this case, the relation between $g(\boldsymbol{r})$ and $f(\boldsymbol{r})$ can be modeled as a blurring and a DS. The support of the blur PSF is larger than LR pixel size, because the total blurring PSF is obtained by the convolution of the PSF of the sensor with other blurring effect. Figure 3 summarizes, in one dimension, the shapes of the different PSF in different cases. In this figure, we may note that the

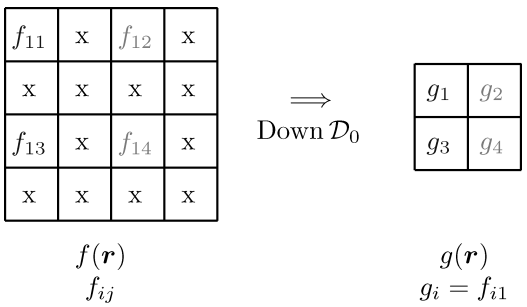


**FIGURE 1.** A first very simple SISO SR model; this forward model is denoted $\mathcal{D}_0$.

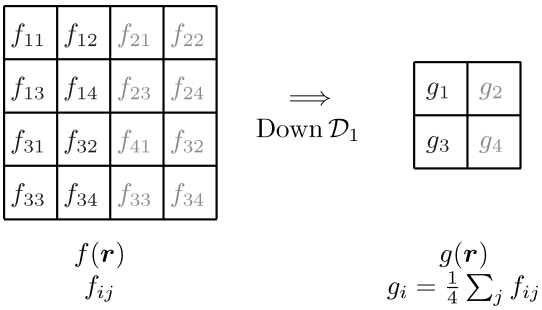


**FIGURE 2.** A second simple SISO SR model; this forward model is denoted $\mathcal{D}_1$.

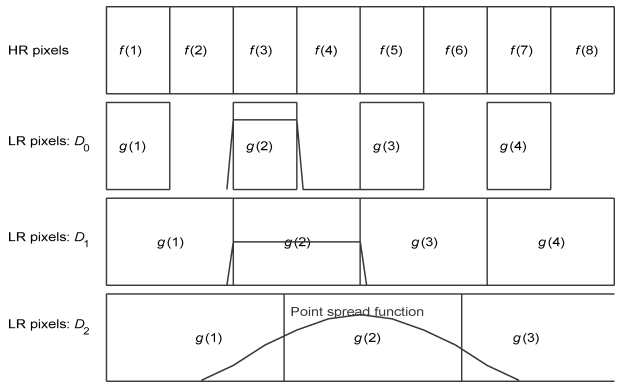


**FIGURE 3.** PSF related to different forward models of sub-sampling: $\mathcal{D}_0$, $\mathcal{D}_1$ and $\mathcal{D}_2$.

PSF associated to $\mathcal{D}_0$ is a gate function with one HR sampling interval, the PSF associated to $\mathcal{D}_1$ is a gate function with $k$ HR sampling intervals and the PSF associated to $\mathcal{D}_2$ is the result of the other blurring effect PSF and a gate function with $k$ HR sampling interval, which in this case may not need to be an integer value.

### 2.3. Multi input single output case

Now, we consider the MISO case, where the main idea is that different LR images $g_i(\boldsymbol{r})$ have, in some sense, complementary informations. These LR images may have been obtained by:

- A moving camera with fixed scene;
- A fixed camera but focusing on a slightly moving scene;
- Multiple fixed cameras focusing on a same fixed scene;
- Multiple fixed cameras but focusing on a slightly moving scene, etc.

In any of these cases, the main idea is that the LR images are not registered. The simplest model is then assuming a translational movement between these images. Now, depending on the hypothesis whether these translational movements are integer factors of sampling interval or not, we may change the formulation of the forward problem.

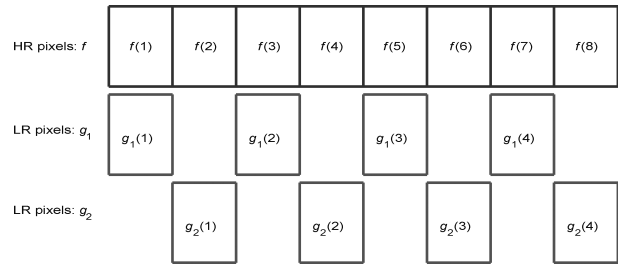


**FIGURE 4.** A first very simple 1D MISO SR model; the relation between LR data $g_i$ and the HR unknown $f$ can be written as $g_k = \mathcal{D}_0\mathcal{M}_k f$.

The simplest model is a translational movement $\mathcal{M}_k$ and an integer ratio DS $\mathcal{D}_0$. Indeed, if we assume that the translational movements are also integer sub-pixel movements of the same DS ratio, then the SR forward model and its inversion become very easy. Figure 4 shows such a simple forward model.

This simple model is however unrealistic, because it does not account for the integration of LR sensor size. The second model that accounts for this is shown in Fig. 5. This model is still unrealistic, because it does not account for the blurring effects of the measurement system and also for the fact that the LR sensor size may not be an integer factor of the size of the HR pixel size. Figure 6 shows such a forward model.

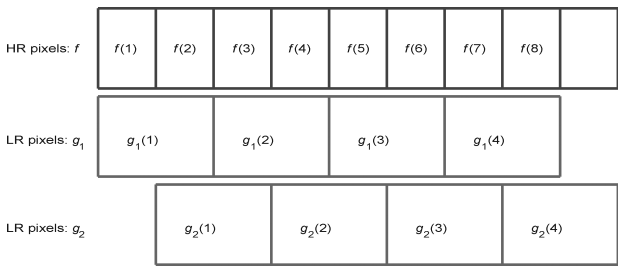


**FIGURE 5.** A second very simple 1D MISO SR model; the relation between LR data $g_i$ and the HR unknown $f$ can be written as $g_k = \mathcal{D}_1\mathcal{M}_k f$.

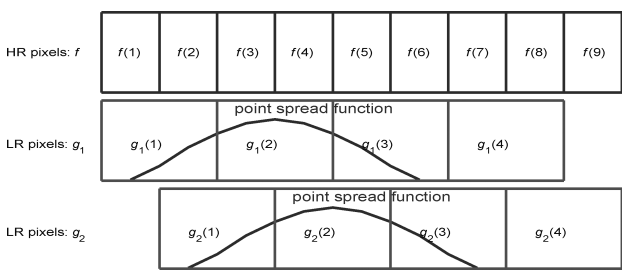


**FIGURE 6.** A third very simple 1D MISO SR model; the relation between LR data $g_i$ and the HR unknown $f$ can be written as $g_k = \mathcal{D}_1\mathcal{B}\mathcal{M}_k f$ or as $g_k = \mathcal{D}_1\mathcal{M}_k\mathcal{B}f$.

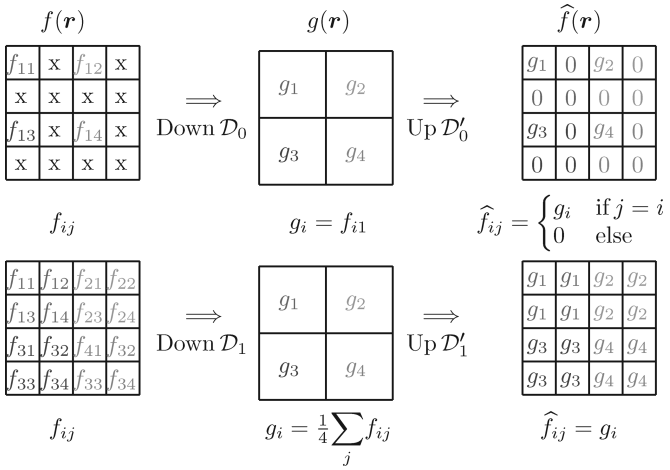


**FIGURE 7.** Adjoint operators of DS operators; top; DS $\mathcal{D}_0$ and its adjoint operator $\mathcal{D}_0'$, bottom: DS $\mathcal{D}_1$ and its adjoint operator $\mathcal{D}_1'$ (we may note that $\mathcal{D}_0'\mathcal{D}_0 \neq \mathcal{I}$ but $\mathcal{D}_0\mathcal{D}_0' = \mathcal{I}$, also $\mathcal{D}_1'\mathcal{D}_1 \neq \mathcal{I}$ but $\mathcal{D}_1\mathcal{D}_1' = \mathcal{I}$ and $\mathcal{D}_0\mathcal{D}_1' = \mathcal{I}$).

### 2.4. General forward-modeling components and associated adjoint operators

A more general realistic forward model is the one that accounts for: (i) Translational movements of the LR images; (ii) Integration of the LR sensor sizes; (iii) Different blurring effects which may include the integration of LR sensor sizes as well as other imaging system defaults and finally (iv) The measurement noise on the observed LR images and all the other unmodeled errors.

In this section, we present a general forward model which partially accounts for these. In this general framework, the relation between LR images $g_i(\boldsymbol{r})$ and the HR image $f(\boldsymbol{r})$ is modeled by

$$g_k(\boldsymbol{r}) = [\mathcal{H}_k f](\boldsymbol{r}) + \epsilon_k(\boldsymbol{r}) \tag{8}$$

where $\epsilon_k(\boldsymbol{r})$ represents the modeling and approximation errors and the operator $\mathcal{H}_k$ is, in general, composed by three main operators:

– A global low-pass filtering $\mathcal{B}$ representing for both real band width limitation of the imaging sensors and the integration over the sensor surface

$$\tilde{f}(\boldsymbol{r}) = [\mathcal{B}f](\boldsymbol{r}) = \int f(\boldsymbol{r}')h(\boldsymbol{r} - \boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \tag{9}$$

where $h(\boldsymbol{r})$ represents the point spread function of the sensor integration and other limiting bandwidth of the imaging system. This PSF may not be known in practice.

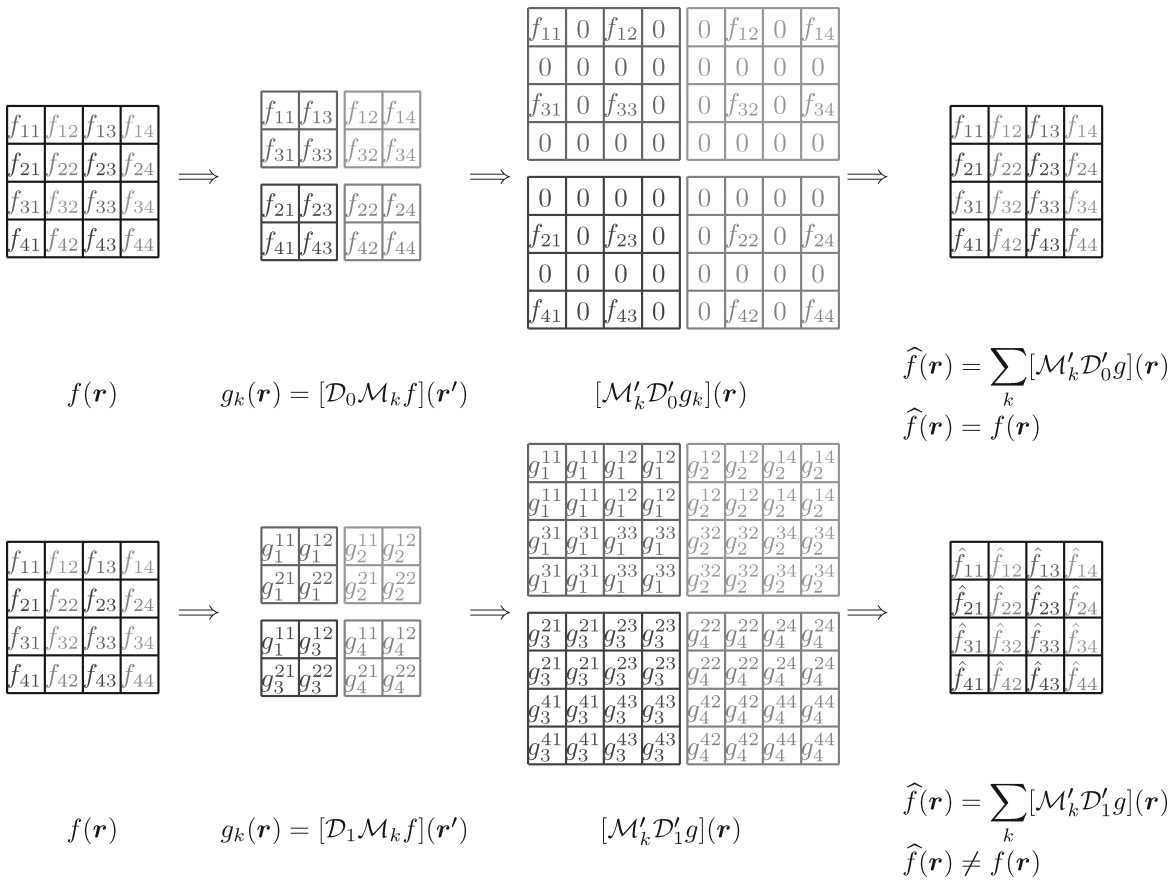


**FIGURE 8.** Adjoint operators of compound movements and DS; top; movements $\mathcal{M}_k$ and DS $\mathcal{D}_0$, bottom: movements $\mathcal{M}_k$ and DS $\mathcal{D}_1$ (we note that $\sum_k \mathcal{M}_k{}'\mathcal{D}_0{}'\mathcal{D}_0\mathcal{D}_0\mathcal{M}_k = \mathcal{I}$ but $\sum_k \mathcal{M}_k{}'\mathcal{D}_1{}'\mathcal{D}_1\mathcal{M}_k \neq \mathcal{I}$).

– A global geometrical transformation $\mathcal{M}_k$, where the simplest model is just a translational movement

$$\tilde{f}_k(\boldsymbol{r}) = [\mathcal{M}_k f](\boldsymbol{r}) = f(\boldsymbol{r} - \boldsymbol{d}_k) \tag{10}$$

where $\boldsymbol{d}_k$ represents the displacement of the coordinates between the two images $\tilde{f}(\boldsymbol{r})$ and $f(\boldsymbol{r})$. More general geometrical transformations taking account for a possible rotation or scaling can also be easily considered. This transformation is then characterized by a few parameters (for example $x$ and $y$ displacement in the case of simple translational movement) which have to be determined. The determination of these parameters is often called *image registration*.

– A DS or decimation $\mathcal{D}$

$$\tilde{f}(\boldsymbol{r}_j) = [\mathcal{D}f](\boldsymbol{r}_j) \tag{11}$$

where $\mathcal{D}$ is a DS with $k$ sampling (or zooming) ratio. In previous section, we emphasized three cases for this operator: A zero order decimation $\mathcal{D}_0$, a first order decimation $\mathcal{D}_1$ and a non-integer $k$ general operator $\mathcal{D}_2$. In the following, when it is not noted, we assume $\mathcal{D} = \mathcal{D}_0$.

The order of these three operators may change. For example, we can change the order of the two operators of $\mathcal{M}_k$ and $\boldsymbol{B}$ which results in two different forward models

$$g_k(\boldsymbol{r}) = [\mathcal{D}\mathcal{M}_k\mathcal{B}f](\boldsymbol{r}) + \boldsymbol{\epsilon}_k(\boldsymbol{r}) \tag{12}$$

and

$$g_k(\boldsymbol{r}) = [\mathcal{D}\mathcal{B}\mathcal{M}_k f](\boldsymbol{r}) + \boldsymbol{\epsilon}_k(\boldsymbol{r}) \tag{13}$$

We may even divide the blur operator $\mathcal{B}$ into an effective imaging system blur $\mathcal{B}_1$ and the LR sensor integration blur $\mathcal{B}_2$ and write

$$g_k(\boldsymbol{r}) = [\mathcal{D}\mathcal{B}_2\mathcal{M}_k\mathcal{B}_1 f](\boldsymbol{r}) + \boldsymbol{\epsilon}_k(\boldsymbol{r}). \tag{14}$$

In the discretized version, if we represent the HR image pixels by $\boldsymbol{f}$, the LR image pixels by $\boldsymbol{g}_k$ and the discretized version of the aforementioned operators by $\boldsymbol{H}_k$, then we can write

$$\boldsymbol{g}_k = \boldsymbol{H}_k\boldsymbol{f} + \boldsymbol{\epsilon}_k \tag{15}$$

or even

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \tag{16}$$

with

$$\boldsymbol{g} = \begin{bmatrix} \boldsymbol{g}_1 \\ \vdots \\ \boldsymbol{g}_k \\ \vdots \\ \boldsymbol{g}_K \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_k \\ \vdots \\ \boldsymbol{H}_K \end{bmatrix} \quad \text{and} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \boldsymbol{\epsilon}_1 \\ \vdots \\ \boldsymbol{\epsilon}_{\mathrm{k}} \\ \vdots \\ \boldsymbol{\epsilon}_{\mathrm{K}} \end{bmatrix} \tag{17}$$

where $\boldsymbol{H}_k = \boldsymbol{D}\boldsymbol{B}\boldsymbol{M}_k$ or $\boldsymbol{H}_k = \boldsymbol{D}\boldsymbol{M}_k\boldsymbol{B}$ or still $\boldsymbol{H}_k = \boldsymbol{D}\boldsymbol{B}_2\boldsymbol{B}_1\boldsymbol{M}_k$ as it is just discussed.

As we will see in the next section, these different models give rise to different intuitive and classical SR methods.

### 2.5. Adjoint operators of forward models and simple MISO SR methods

To each of these three basic operators, we can associate respective adjoint operators. Denote by $\mathcal{B}'$, $\mathcal{M}_k'$, $\mathcal{D}'$ and $\mathcal{H}_k'$ the adjoint operators of $\mathcal{B}$, $\mathcal{M}_k$, $\mathcal{D}$ and $\mathcal{H}_k$. We can then easily see that $\mathcal{B}'$ is also a linear blur operator, $\mathcal{M}_k'$ is a translation operator at the opposite direction of $\mathcal{M}_k$ and $\mathcal{D}'$ is an up-sampling operator. We may note that not all these operators are auto-adjoint. We remember that the operator $\mathcal{H}'$: $\mathcal{Y} \mapsto \mathcal{X}$, the adjoint operator of $\mathcal{H}$: $\mathcal{X} \mapsto \mathcal{Y}$, is such that $\langle x, \mathcal{H}'y\rangle = \langle y, \mathcal{H}x\rangle$ $\forall x \in \mathcal{X}$ and $\forall y \in \mathcal{Y}$, and $\mathcal{H}'$ and $\mathcal{H}$ are auto-adjoint if $\mathcal{H}\mathcal{H}' = \mathcal{H}'\mathcal{H} = \mathcal{I}$, where $\mathcal{I}$ is the identity operator. See details in Figs 7 and 8.

Based on these notations and the definition of the adjoint operators $\mathcal{B}'$, $\mathcal{M}_k'$ and $\mathcal{D}'$ (or equivalently $\boldsymbol{B}'$, $\boldsymbol{M}_k'$ and $\boldsymbol{D}'$), which have particular structures, a very simple scheme for inversion is

$$\hat{\boldsymbol{f}} = \boldsymbol{H}^t\boldsymbol{g} = \sum_k \boldsymbol{H}_k^t\, \boldsymbol{g}_k = \sum_k \boldsymbol{B}'\boldsymbol{M}_k'\boldsymbol{D}'\, \boldsymbol{g}_k \tag{18}$$

or equivalently

$$\hat{\boldsymbol{f}}(\boldsymbol{r}) = \sum_k [\mathcal{H}_k^t\, \mathbf{g_k}](\boldsymbol{r}) = \sum_k [\mathcal{B}'\mathcal{M}_k'\mathcal{D}'\, g_k](\boldsymbol{r}) \tag{19}$$

This corresponds to up-sampling, registration of images in HR grid, filtering and superposition (summing or fusion).

Other methods can be obtained easily by changing the orders of these operators. For example

$$\hat{f} = \sum_k \boldsymbol{B}'\boldsymbol{D}'\boldsymbol{M}_k'\boldsymbol{g}_k \tag{20}$$

corresponds to sub-pixel registration of LR images, up-sampling, filtering and summing.

Many classical methods of SR have been based on these relations. However, even if, in theory, the operators corresponding to the cases of translational movements and DS scheme $\mathcal{D}_0$ are auto-adjoint, in practical applications, these schemes will not give a perfect reconstruction. Between the reasons, we may note the following facts:

- Except for the case of $\mathcal{D}_0$, other operators are not auto-adjoint;
- In general, the movements are not just translational;
- In general, the movements are not an integer factor of HR pixel size;
- We also have to account for the blurring effects and the sensor noise.

We remark, however, that all these schemes are composed of up-sampling (compensation for DS, or more generally interpolation), registration (compensation for the movements) and filtering (compensation for the blurring due to the LR sensor size and other blurring effects). Many classical methods of SR have been based on appropriate combination of these operations. In the next section, we give a brief review of these methods.

## 3. CLASSICAL MISO SR METHODS

As we could see from the different forward-modeling operators and their associated adjoint operators, the MISO SR problem can be summarized as a combination of registration (movement compensation), interpolation (DS compensation) and summation or more generally image fusion. Registration consists in finding some way of bringing together all the input LR images into a coordinate frame that reconstructs an HR output. This corresponds to the adjoint operators $\mathcal{M}_k'$. The combination of the interpolation and fusion is equivalent to the registered LR images to construct an HR image. This corresponds to the combined adjoint operators $\mathcal{B}'\mathcal{D}'$. Then, depending on the order of these operations, we can classify the classical MISO SR methods into two categories of grid mapping and interpolation and interpolation and fusion.

Indeed, due to the Fourier domain properties of sampling and translational movements, there is a great number of SR methods using these relations. In the following, a very short review of these classical methods is presented.

### 3.1. Grid mapping and interpolation

This is the most intuitive SR reconstruction process involving mapping onto a higher-resolution grid (equivalent to $\mathcal{D}_0'\,\mathcal{M}_k'$ or $\mathcal{D}_1'\,\mathcal{M}_k'$) followed by bilinear or higher-order spline interpolation (different approximations or extensions of the operator $\mathcal{B}$). This algorithm is often called *Shift-And-Add*

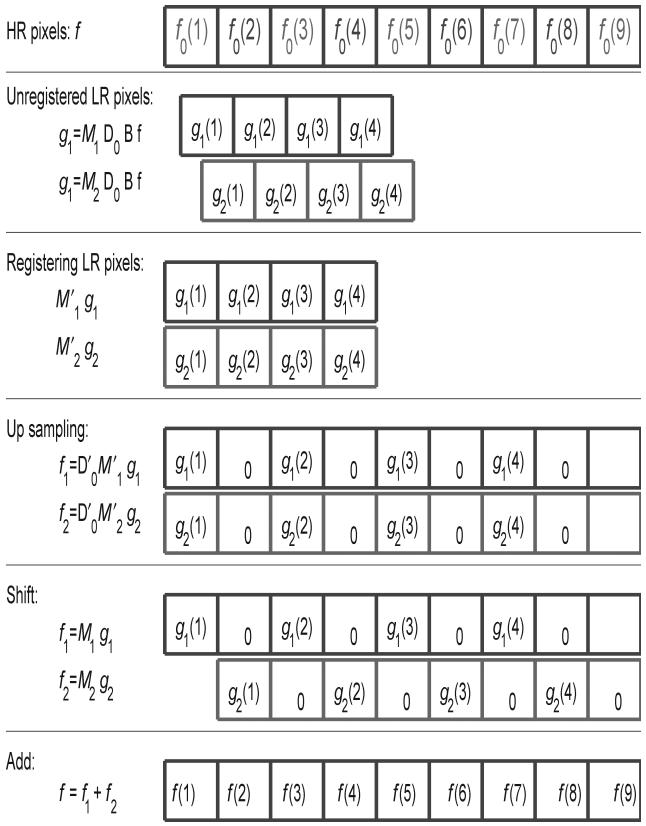


**FIGURE 9.** Sub-pixel LR registration, grid mapping, shift and add operations in 1D.

[29]. This is shown in Fig. 9 for the 1D case and in Fig. 10 for the 2D case.

The main difficult task in this approach is the motion estimation parameters. The main advantage of this method is in its low computational cost making real-time applications possible. On the other hand, only the same blur and noise for all LR frames can be assumed. The missing of some LR frames reduces the overall performance of such algorithms.

### 3.2. Interpolation and fusion

In this method, instead of mapping to an HR grid as in the previous scheme, a linear or nonlinear interpolation method is performed to each LR frame separately to increase their resolution (an approximation or extension of the combined operators $\mathcal{M}_k'\mathcal{D}'\mathcal{B}'$ or $\mathcal{M}_k'\mathcal{B}'\mathcal{D}'$). Then, a fusion between all the resolved frames results in an SR image at the resolution of the interpolated LR frames. This is shown in Fig. 11 for the 1D case and in Fig. 12 for the 2D case.

Depending on the fusion method, not all frames contribute to reconstruct pixels in the SR image. Farsiu *et al.* [4, 30] recommend the median for this purpose. In the particular example of median fusion, only one of the LR frames is used for each reconstructed pixel. For text enhancement in digital video, Li and Orchard [31] use bilinear interpolation followed by averaging of the interpolated frames. Interpolation and fusion is fast and robust to outliers, but it can result in the appearance of some artificial effects in the super-resolved image due to the nature of the fusion process.

### 3.3. Frequency-domain reconstruction

This particular form of SR reconstruction is based on the Fourier transform (FT) properties of sampling, translational motion and rotational movement

$$\begin{array}{lcl} f(\boldsymbol{r}) & \overset{\text{FT}}{\leftrightarrow} & F(\boldsymbol{\omega}) \\ \sum_i f(\boldsymbol{r}_i)\delta(\boldsymbol{r}-\boldsymbol{r}_i) & \leftrightarrow & \sum_i F(\boldsymbol{\omega})\delta(\boldsymbol{\omega}-i2\pi/\boldsymbol{\Delta}) \end{array}$$

with $\boldsymbol{r}_i = i\boldsymbol{\Delta}$

$$\begin{array}{lcl} f(\boldsymbol{r}-\boldsymbol{d}) & \leftrightarrow & \exp\{-j\boldsymbol{\omega}'\boldsymbol{d}\}F(\boldsymbol{\omega}) \\ f(\boldsymbol{R}\boldsymbol{r}) & \leftrightarrow & F(\boldsymbol{R}\boldsymbol{\omega}) \end{array} \tag{21}$$

where $F(\boldsymbol{\omega})$ is the 2D FT of $f(\boldsymbol{r})$, $\boldsymbol{d}$ is the uniform motion parameter and $\boldsymbol{R}$ the rotational operator parameter.

These methods are very often the continuation of frequency-domain motion estimation in the case of pure translational or rotational model assumption. It was first derived by

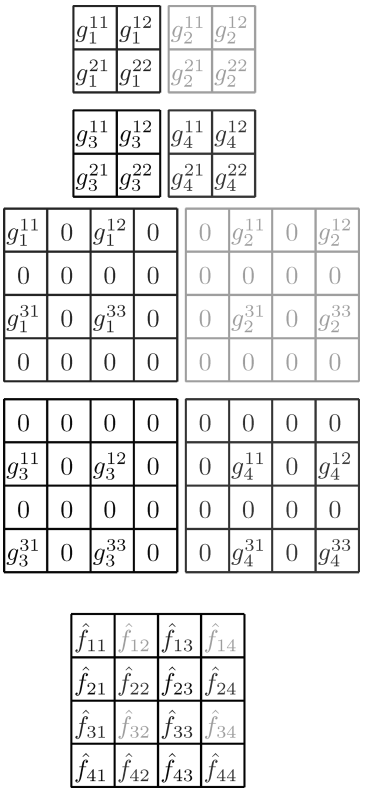

**FIGURE 10.** Grid mapping and addition in 2D.

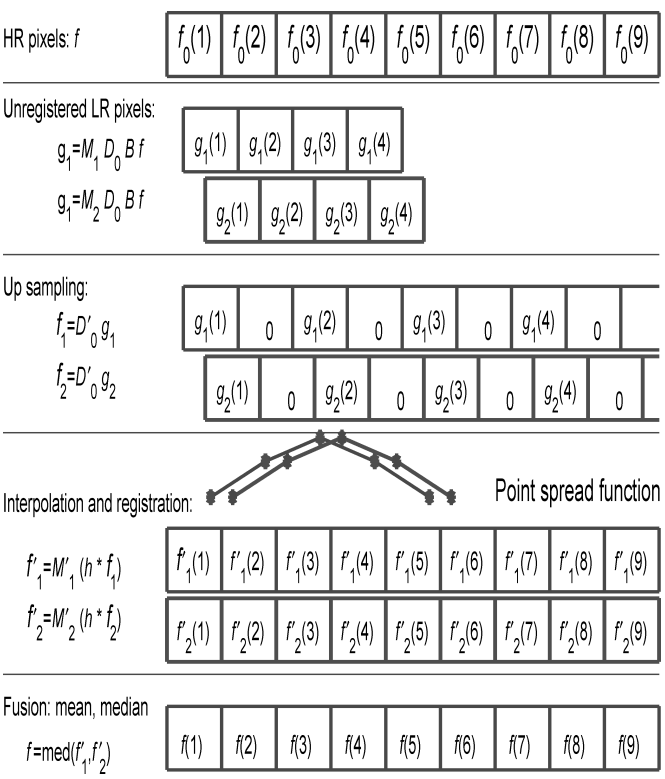


**FIGURE 11.** Interpolation, HR registration and HR fusion in 1D.

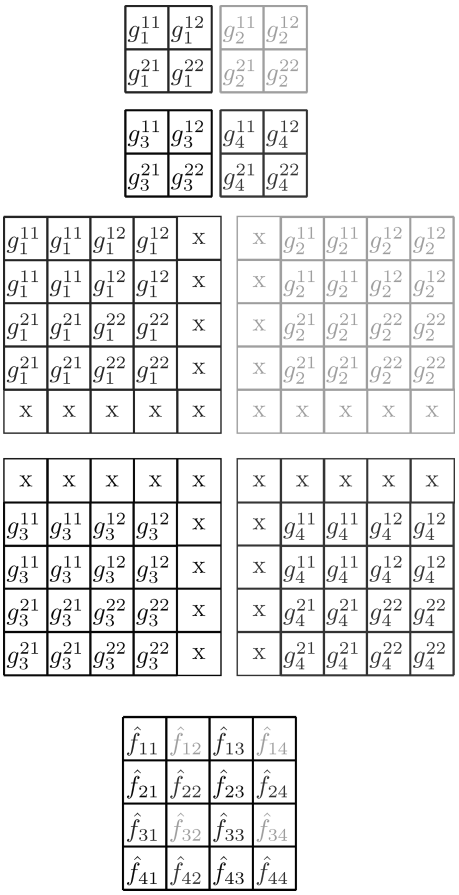

**FIGURE 12.** Interpolation, HR registration and fusion in 2D.

Tsai and Huang [29] and implemented SR reconstruction method, also called alias-removal reconstruction. Assuming that the LR images are under-sampled, the translations between them allow an up-sampled SR image to be built based on the shifting property of the FT and the aliasing relationship between the continuous FT of an original SR image and the discrete FT of observed LR images. Several extensions [32–34] were then proposed to enlarge the initial conditions of Tsai and Huang, which were integer-shift translation only. The major advantage is its simplicity but only global translational models can be considered.

In this paper, we do not detail more these methods. However, we may partially use these methods for the estimation of the uniform motion or rotational motion parameters.

## 4. GENERAL INVERSION METHODS

Based on this forward modeling and assuming, in a first step, that the forward operators $\boldsymbol{H}_k$ are known (which means that the registration parameters and the PSF of the blurring effects are known), the inversion or the estimation of the HR image $f(\boldsymbol{r})$ based on the LR images $g_k(\boldsymbol{r}_i)$, and some prior modeling of

the the HR image $f(\boldsymbol{r})$ can be casted in the following main classes:

- LS methods [23, 29, 35–37],
- Robust estimation (RE) methods [11, 12],
- Regularization-based methods [4, 11, 12, 16], and
- Bayesian estimation methods [1, 7, 27, 38–40].

In the three first classes of methods and in MAP estimation category of the Bayesian estimation methods, the solution is defined as the minimizer of a criterion $J(\boldsymbol{f})$:

$$\hat{\boldsymbol{f}} = \arg \min_{\boldsymbol{f}} \{J(\boldsymbol{f})\} \tag{22}$$

where the expression of $J(\{\boldsymbol{f}\})$ becomes

– LS and RE methods

$$J(\boldsymbol{f}) = \sum_k \|\boldsymbol{g}_k - \boldsymbol{H}_k\boldsymbol{f}\|^\beta = \sum_k \sum_{\boldsymbol{r}\in\mathcal{R}} |g_k(\boldsymbol{r}) - [\mathcal{H}_k f](\boldsymbol{r})|^\beta \tag{23}$$

with $1 \leq \beta \leq 2$ for the general case and $\beta = 2$ for the LS case.

– Regularization methods

$$J(\boldsymbol{f}) = \sum_k \|\boldsymbol{g}_k - \boldsymbol{H}_k\boldsymbol{f}\|^{\beta_1} + \lambda\|\boldsymbol{D}\boldsymbol{f}\|^{\beta_2} \tag{24}$$

where $1 \leq \beta_1, \beta_2 \leq 2$, $\boldsymbol{D}$ represents a high-pass filter operator and $\lambda$ is a regularization parameter.

– Bayesian MAP estimation methods

$$J(\boldsymbol{f}) = -\ln p(\boldsymbol{f}|\underline{\boldsymbol{g}}) = -\ln p(\underline{\boldsymbol{g}}|\boldsymbol{f}) - \ln p(\boldsymbol{f}) + c \tag{25}$$

where: $\underline{\boldsymbol{g}} = \{\boldsymbol{g}_k, k = 1, \ldots, M\}$ represents all the LR images,

$$p(\underline{\boldsymbol{g}} \mid \boldsymbol{f}) \propto \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\sum_k \|\boldsymbol{g}_k - \boldsymbol{H}_k\boldsymbol{f}\|^2\right\}$$

is the likelihood when the noises $\varepsilon_k$ are assumed centred, iid and Gaussian with given variance $\sigma_\varepsilon^2$ and $p(\boldsymbol{f})$ is an a priori model on the HR image.

When a generalized Gauss–Markov prior law is chosen for $p(\boldsymbol{f})$, the MAP criterion (25) becomes equivalent to the regularization criterion (24). However, as we will discuss it in the next section, the Bayesian estimation framework is much richer.

Let us consider here the QR solution, which is the optimizer of the criterion

$$J(\boldsymbol{f}) = \sum_k \|\boldsymbol{g}_k - \boldsymbol{H}_k\boldsymbol{f}\|^2 + \lambda\|\boldsymbol{D}\boldsymbol{f}\|^2$$

This solution can be obtained analytically by differentiating the criterion $J(\boldsymbol{f})$ with respect to $\boldsymbol{f}$ and setting it to zero

$$\nabla J(\hat{\boldsymbol{f}}) = -2\sum_k \mathbf{H}_k'(\boldsymbol{g}_k - \boldsymbol{H}_k\hat{\boldsymbol{f}}) + 2\lambda\boldsymbol{D}'\boldsymbol{D}\hat{\boldsymbol{f}} = \boldsymbol{0} \tag{26}$$

which results to the normal equation

$$\left[\sum_k \boldsymbol{H}_k'\boldsymbol{H}_k + \lambda\boldsymbol{D}'\boldsymbol{D}\right]\hat{\boldsymbol{f}} = \sum_k \boldsymbol{H}_k'\boldsymbol{g}_k \tag{27}$$

and finally to the solution

$$\hat{\boldsymbol{f}} = \left[\sum_k \boldsymbol{H}_k'\boldsymbol{H}_k + \lambda\boldsymbol{D}'\boldsymbol{D}\right]^{-1}\sum_k \boldsymbol{H}_k'\boldsymbol{g}_k \tag{28}$$

This solution which can be compared with (18) is composed of the application of the adjoint operator $\sum_k \boldsymbol{H}_k' = \sum_k \boldsymbol{B}'\,\boldsymbol{M}_k'\,\boldsymbol{D}'$ and a global filtering operator $[\sum_k \boldsymbol{H}_k'\,\boldsymbol{H}_k + \lambda\,\boldsymbol{D}'\boldsymbol{D}]^{-1}$. One then finds the basic operations of up-sampling $\boldsymbol{D}'$, registration or motion compensation $\boldsymbol{M}'$, individual filtering $\mathcal{B}'$ and global filtering or fusion $[\sum_k \boldsymbol{H}_k'\,\boldsymbol{H}_k + \lambda\,\boldsymbol{D}'\boldsymbol{D}]^{-1}$. The case of LS solution corresponds to $\lambda = 0$. In particular cases, we may obtain approximated analytical inversion for the filtering operation.

We may note, however, that even if we have this analytical expression for the QR or LS solution, very often its computation is done via an iterative optimization algorithm such as a gradient ascent one, where at each iteration $(i)$, we adjust the previous result with an increment which needs the computation of the gradient

$$\begin{aligned}&\boldsymbol{f}^{(i+1)} = \boldsymbol{f}^{(i)} - \alpha\nabla J(\boldsymbol{f}^{(i)}) \text{ with}\\ &\nabla J(\boldsymbol{f}) = \sum_k \boldsymbol{H}_k'(\boldsymbol{g}_k - \boldsymbol{H}_k\boldsymbol{f}) + \lambda\boldsymbol{D}'\boldsymbol{D}\boldsymbol{f}\end{aligned} \tag{29}$$

which is again composed of all the basic adjoint operators.

What is more interesting is that, these adjoint operators are the main building blocks of any iterative optimization algorithm trying to optimize any of the aforementioned criteria.

Indeed, the classical methods of iterative back-projection, first introduced by Irani and Peleg [18], have found much use in mainstream SR reconstruction, can actually be considered as a simple gradient-based algorithm for minimizing the LS criterion. However, these methods have no unique solution due to the ill-posed nature of the inverse problem. In fact, minimizing the LS error does not necessarily imply a reasonable solution and a convergent iteration does not necessarily converge to a unique solution.

Note also that, at each step of these iterative algorithms, we need a motion estimation or a registration step.

### 4.1. Methods trying to estimate jointly registration parameters, PSF and the HR image

In the methods of the previous section, the hypothesis was that the PSF of different blurring effects and the registration parameters $\phi$ are known or estimated previously. However, there are also many more advanced methods which try to jointly estimate the PSF and the registration parameters at each iteration of the SR reconstruction. In fact, we can define a criterion $J(\boldsymbol{f}, \boldsymbol{h}, \phi)$ which depends, not only on the HR image $\boldsymbol{f}$, but also on the PSF $\boldsymbol{h}$ and the registration parameters $\phi$. A such typical criterion is

$$J(\boldsymbol{f}, \boldsymbol{h}, \boldsymbol{\phi}) = \sum_k \|\boldsymbol{g}_k - \boldsymbol{H}_k \boldsymbol{f}\|^2 + \lambda_f \|\boldsymbol{D}_f \boldsymbol{f}\|^\beta + \lambda_h \|\boldsymbol{D}_h \boldsymbol{h}\|^2 \quad (30)$$

where $\boldsymbol{H}_k$ depend on $(\boldsymbol{h}, \boldsymbol{\phi})_k$. Then, the main idea is to optimize this criterion, successively, with respect to $\boldsymbol{f}$, $\boldsymbol{h}$ and $\phi_k$, each time keeping the two others fixed at previous iterations. In these methods, appropriate choice of $\boldsymbol{D}_f$ and $\boldsymbol{D}_h$ and the regularization parameters $\lambda_f$ and $\lambda_h$ may be very important for the success of the method. Choosing $\beta = 2$ simplifies the two steps of optimization with respect to $\boldsymbol{f}$ and $\boldsymbol{h}$. However, the optimization with respect to $\boldsymbol{\phi}$ needs great care and the success of the method may depend on this step. Many authors have followed this approach [2, 3, 41].

## 5. MORE ADVANCED PRIOR MODELING AND BAYESIAN ESTIMATION METHODS

In SR problems, as in any inverse problem, the choice of the optimization criterion, and in particular, the regularization terms are very important. These terms, in a Bayesian framework, correspond to the prior laws. In SR, a Gaussian prior for the PSF is often reasonable. This corresponds to the term $\lambda_h ||\boldsymbol{D}_h \boldsymbol{h}||^2$ in (30). More important then is the prior law on the HR image $\boldsymbol{f}$. The Gaussian prior for $\boldsymbol{f}$ leads to fast algorithms. However, it is not appropriate in many imaging applications. This is the reason for choosing $\lambda_f ||\boldsymbol{D}_f \boldsymbol{f}||^\beta$ with $1 < \beta < 2$ in (30), which corresponds to a generalized Gaussian prior which is more appropriate for many applications of imaging systems.

Recently, we developed more sophisticated methods which try to account for the fact that very often the images to be reconstructed are composed of statistically homogeneous regions, and this property can be used to develop methods which still give more accurate reconstruction results [27, 39, 40]. The main idea in these methods is to model the image via a composite Markov model with hidden region labels $z(\boldsymbol{r})$ which takes discrete values $l = 1, \ldots, L$, and is modeled via a Potts Markov field

$$p(\boldsymbol{z}) \propto \exp\left\{\gamma \sum_{\boldsymbol{r} \in \mathcal{R}} \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\} \quad (31)$$

where $\boldsymbol{z} = \{z(\boldsymbol{r}), \boldsymbol{r} \in \mathcal{R}\}$, $\mathcal{V}(\boldsymbol{r})$ is the set of the four nearest neighbors of $\boldsymbol{r}$ and $\mathcal{R}$ is the set of all pixel positions.

This Potts model can also be written as

$$p(z(\boldsymbol{r}) \mid z(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})) \propto \exp\left\{\gamma \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\} \quad (32)$$

where we can see more explicitly the dependency of $z(\boldsymbol{r})$ to its neighbors $\{z(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})\}$.

The image pixels $\boldsymbol{f}_l = \{f(\boldsymbol{r}), \boldsymbol{r} \in \mathcal{R}_l\}$ with a same classification labels $\mathcal{R}_l = \{\boldsymbol{r} : z(\boldsymbol{r}) = l\}$ are then modeled by Gauss–Markov fields

$$p(\boldsymbol{f}_l) = \mathcal{N}(m_l \mathbf{1}_l, \boldsymbol{\Sigma}_l) \quad (33)$$

which can also be written in conditional form

$$p(f(\boldsymbol{r}) \mid z(\boldsymbol{r}) = l) = \mathcal{N}(m_l, \boldsymbol{\sigma}_l^2) \quad (34)$$

where the parameters (means $m_l$, variances $\boldsymbol{\sigma}_l^2$ and covariances $\boldsymbol{\Sigma}_l$) depend on the region labels $l$. By this modeling, naturally the pixels of an image are classified in $L$ independent classes. The pixels having the same class $\boldsymbol{f}_l = \{f(\boldsymbol{r}), \boldsymbol{r} \in \mathcal{R}_l\}$ are naturally grouped in finite set of disjoint regions $\mathcal{R}_{ll'}$ such that:

$$\begin{aligned} &\mathcal{R}_l = \cup_{l'} \mathcal{R}_{ll'} \quad \text{with } \cap_{l'} \mathcal{R}_{ll'} = 0, \quad \cap_l \mathcal{R}_l =, \\ &\cup_l \mathcal{R}_l = \mathcal{R} \end{aligned} \quad (35)$$

See Fig. 13.

The conditional law given in (34) can also be interpreted as a

$$p(f(\boldsymbol{r})) = \sum_{l=1}^{L} \alpha_l \mathcal{N}(m_l, \boldsymbol{\sigma}_l^2) \quad (36)$$

where $\alpha_l = P(z(\boldsymbol{r}) = l)$ with $\sum_{l=1}^{L} \alpha_l = 1$ which is a mixture of Gaussians law.


**FIGURE 13.** Mixture and hidden Markov models for images: Three linked images are intensity image $f(\boldsymbol{r})$, segmentation or classification image $z(\boldsymbol{r})$ and the contour image $q(\boldsymbol{r})$ (note that $f(\boldsymbol{r})$ is gray levels, each color of $z(\boldsymbol{r})$ represents a class and $q(\boldsymbol{r})$ can be obtained from $z(\boldsymbol{r})$ in a deterministic way).

As we assumed the pixels with different labels to be independent, we can write

$$\begin{aligned} p(\boldsymbol{f} \mid z, \boldsymbol{\theta}_1) &= \prod_{l=1}^{L} \mathcal{N}(m_l \mathbf{1}_l, \boldsymbol{\Sigma}_l) \\ &\propto \prod_{l=1}^{L} \prod_{\boldsymbol{r} \in \mathcal{R}_l} \exp\left\{ \frac{-1}{2} \left( \frac{f(\boldsymbol{r}) - m_l}{\sigma_l^2} \right)^2 \right\} \end{aligned} \tag{37}$$

where $\boldsymbol{\theta}_l = \{(m_l, \sigma_l^2), l = 1, \ldots, L\}$.

Now, using the the likelihood

$$p(\underline{\boldsymbol{g}} \mid \boldsymbol{f}, \sigma_\varepsilon^2) \propto \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_k \|\boldsymbol{g}_k - \boldsymbol{H}_k \boldsymbol{f}\|^2 \right\} \tag{38}$$

the prior model (37), the Potts-Markov model of (31) and an appropriate prior for the hyper-parameters $\boldsymbol{\theta} = \{\sigma_\varepsilon^2, \boldsymbol{\theta}_1\} = \{\sigma_\varepsilon^2, (m_l, \sigma_l^2), l = 1, \ldots, L\}$, we can deduce the posterior probability law

$$p(\boldsymbol{f}, z, \boldsymbol{\theta} \mid \underline{\boldsymbol{g}}) \propto p(\underline{\boldsymbol{g}} \mid \boldsymbol{f}, \boldsymbol{\sigma}_\varepsilon^2)\, p(\boldsymbol{f} \mid z, \boldsymbol{\theta}_1)\, p(z)\, p(\boldsymbol{\theta}) \tag{39}$$

from which we can infer on all the unknowns. We have used conjugate priors for all the hyper-parameters $\boldsymbol{\theta}$, i.e. Gaussians for the means, inverse gammas ($\mathcal{IG}$) for the variances and Inverse Wishart ($\mathcal{IW}$) for the covariance matrices. For more details on these, refer to [42].

Finally, from this posterior law MAP or posterior mean estimators are deduced and appropriate numerical optimization or numerical integration (MCMC) algorithms are developed to estimate jointly all the unknowns: the HR image $\hat{\boldsymbol{f}}$, its segmentation $\hat{z}$ and the hyper-parameters $\hat{\boldsymbol{\theta}}$. However, in general, these solutions are obtained in an iterative way either by relaxation or by Gibbs MCMC sampling. The estimation of $\boldsymbol{f}$ given the two other unknowns $z$ and $\boldsymbol{\theta}$ needs optimization of a criterion which looks like the regularization one where the regularization term depends on the segmentation $z$ obtained in previous steps. The estimation of $z$ given previous values of $\boldsymbol{f}$ and $\boldsymbol{\theta}$ can be assimilated to a dynamic segmentation step, and finally, the estimation of the hyper-parameters $\boldsymbol{\theta}$ depends on the data and the previous values of $\boldsymbol{f}$ and $\boldsymbol{\theta}$. At the end of iterations, we obtain, jointly a final solution $\hat{\boldsymbol{f}}$, its corresponding segmentation $\hat{z}$ and statistical properties (means, variances and correlation coefficients) of pixels in each of those homogeneous regions, as well as the statistical properties (mainly the common variance) of the noises.

We may also note that the expression of the likelihood depends on the unknown geometrical translational movements which must also be estimated either once at the beginning using the LR images or during the iterations.

### 5.1. A first algorithm

The main algorithm, in its simplest scheme, can be summarized as follows:

– Initialization:
(i) Estimate the sub-pixel translational movements $\boldsymbol{d}_k$ between the LR images $\boldsymbol{g}_k(\boldsymbol{r})$;
(ii) Estimate a first HR image $\hat{\boldsymbol{f}}(\boldsymbol{r})$ based on LS or QR;

– Iterations:
(i) Estimate a segmentation $\hat{z}(\boldsymbol{r})$ for the HR image $\hat{\boldsymbol{f}}(\boldsymbol{r})$ based on the Potts Markov modeling;
(ii) Estimate the parameters $\hat{\boldsymbol{\theta}}$ of Gaussian mixtures model for the HR image $\hat{\boldsymbol{f}}(\boldsymbol{r})$ using the classification and segmentation result $\hat{z}(\boldsymbol{r})$;
(iii) Update the HR image using the gradient of log posterior which writes here

$$-\ln p(\boldsymbol{f} \mid \hat{z}, \hat{\boldsymbol{\theta}}, \underline{\boldsymbol{g}}) = \frac{1}{2\sigma_\varepsilon^2} \sum_k \|\boldsymbol{g}_k - \boldsymbol{H}_k \boldsymbol{f}\|^2 + \sum_l \|\boldsymbol{f} - m_l \mathbf{1}_l\|_{\boldsymbol{\Sigma}_l}^2 \tag{40}$$

The details of this modeling and more details on this simplest and other MCMC based Bayesian computation algorithms can be found in [28, 27].

### 5.2. New extensions

In aforementioned Bayesian SR image reconstruction algorithm, the translational movement parameters are estimated once from the LR images before using them during the SR algorithm iterations.

In this paper, we propose four extensions:

– Estimation of the translational movements during the iterations using the previously estimated HR image as the reference;
– Extension of geometrical transformation to other projective or rotational transformations;
– Estimation of the PSF during the iterations;

– Implementation in such a way that the LR images can be used successively as they arrive and
– Extension to 3D case.

The new algorithm, in its simplest scheme, can be summarized as follows:

– Initialization:

(i) Estimate a first HR image $\hat{\boldsymbol{f}}(\boldsymbol{r})$ just by interpolating the first LR image;

– Iterations:

(i) Estimate the translational movements $\boldsymbol{d}_k$ between the HR image $\hat{\boldsymbol{f}}(\boldsymbol{r})$ and newly entered LR images $\boldsymbol{g}_k(\boldsymbol{r})$ which is interpolated to the HR dimensions. An extension to a more general geometrical transformation such as rotation and translation can also easily obtained.
(ii) Estimate the blurring PSF. We limited here this step by choosing only a PSF between a finite number of choices. The selected PSF minimizes the LS criterion.
(iii) Estimate a segmentation $\hat{\boldsymbol{z}}(\boldsymbol{r})$ for the HR image $\hat{\boldsymbol{f}}(\boldsymbol{r})$ based on the Potts Markov modeling;
(iv) Estimate the parameters $\hat{\boldsymbol{\theta}}$ of Gaussian mixtures model for the HR image $\hat{\boldsymbol{f}}(\boldsymbol{r})$ using the classification and segmentation result $\hat{\boldsymbol{z}}(\boldsymbol{r})$;
(v) Update the HR image using the gradient of log posterior as in previous case, but also using the segmentation results.

The main advantage of this new algorithm is that it can be applied more easily for the video sequence SR. Also the SR reconstruction result can be available from the first arrival of the LR, and its resolution and precision is increased by the arrival of the new LR images. Its performances become adaptive which is more practical in real video sequence SR applications. Figure 14 shows an example of the results of the different steps of this algorithm.

Figure 15 shows another example of results obtained by two different methods: A robust regularized method and the proposed method. To measure the quality of the results, we used a relative $L1$ distance between the original and the reconstructed results. We limited here this comparison to two examples. As we can see from these results, in both examples, the proposed method gives a better result. However, modeling the image by a Gauss–Markov with Potts prior is more appropriate for the second example, and thus the increase in the performances for this case is still more important.

The 3D extension follows exactly the same steps. However, here it is more difficult to show the results of those different steps.

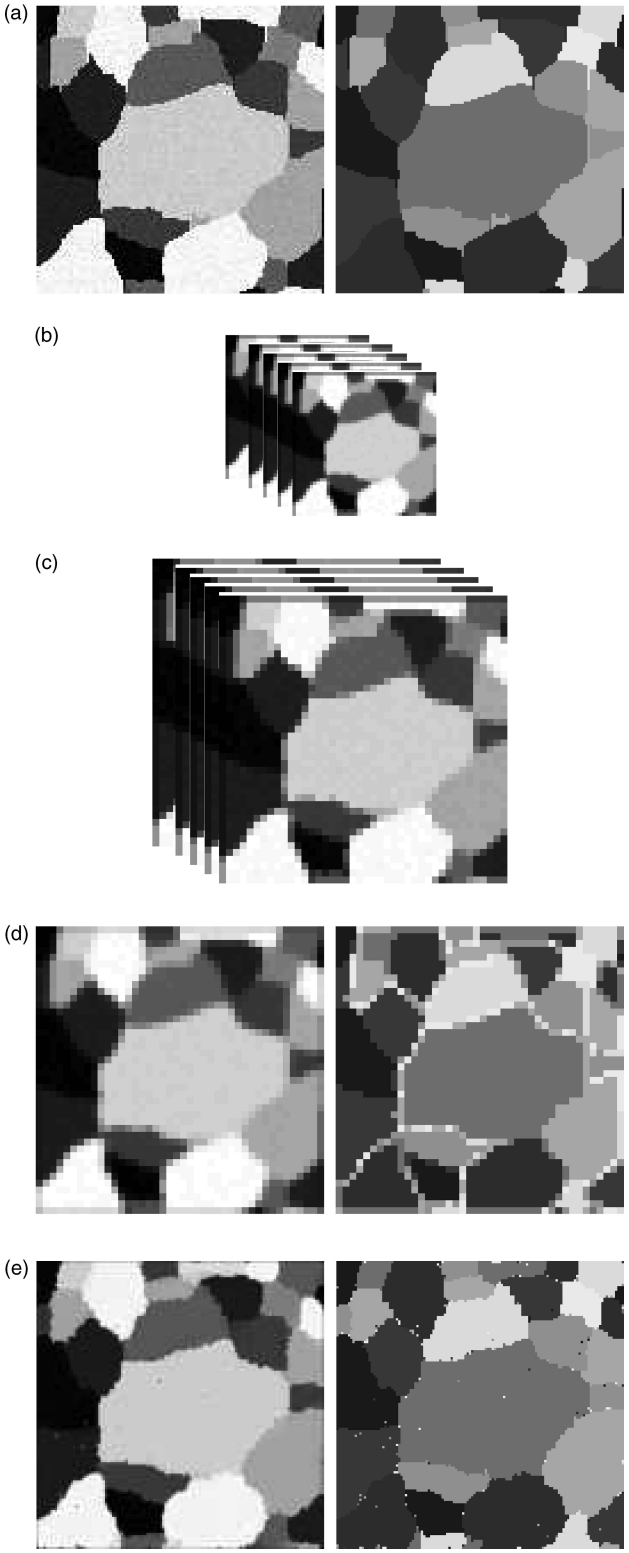


**FIGURE 14.** Different steps of the new SR method: (a) Original HR $f(\boldsymbol{r})$ and its segmentation $z(\boldsymbol{r})$; (b) LR image $g_k(\boldsymbol{r})$; (c) interpolated LR image $\tilde{g}_k(\boldsymbol{r})$; (d) first HR estimate $\hat{f}(\boldsymbol{r})$ and its segmentation $\hat{z}(\boldsymbol{r})$. (e) final HR estimate $\hat{f}(\boldsymbol{r})$ and its segmentation $\hat{z}(\boldsymbol{r})$.

## 6. LIMITATIONS AND CHALLENGES

In all the presented methods, a few steps still need extensions. These needed extensions can be classified into four main

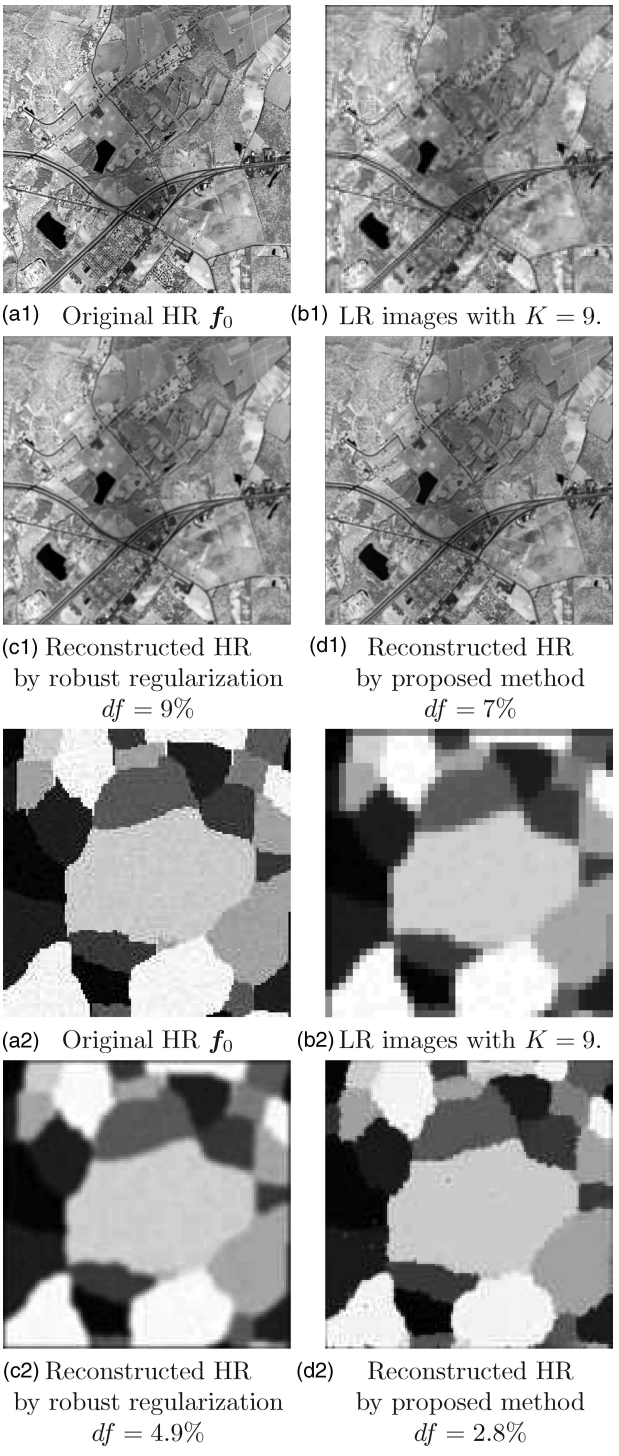


**FIGURE 15.** A comparison of the results obtained by two methods: An RR method and the proposed method; in each group of images, (**a1,2**) is the original HR image $f_0$, (**b1,2**) is one of the $K = 9$ LR images, (**c1,2**) is the reconstructed HR image obtained by an RR method [for $df = 9\%$ and $df = 4.9\%$, respectively, and (**d1,2**) is the reconstructed HR image obtained by the proposed method [for $df = 7\%$ and $df = 2.8\%$, respectively]. (For each of reconstructions the relative $L1$ distances to the original HR is given).

categories: (a) Forward modeling, (b) Prior modeling of the HR image, (c) Estimation or determination of the hyper-parameters and (d) Efficient algorithm implementation. In this section, we give some directions and challenges to them.

### 6.1. Forward modeling

Blurring, noise, geometrical transformation (wrapping), photographic transformation and DS are the main transformations going from the desired HR image to LR images. For blurring, assuming a fixed PSF is good enough in many applications. However, in some applications this model has to be extended to account for spatial variation of the PSF.

Additive, centered, white and Gaussian noise may be too restrictive in some applications. However, it is not too difficult to account for multiplicative or colored or non-Gaussian noise. Indeed, in RE methods, the non-Gaussian noise property is accounted for.

In the presented methods, only integer sub-pixel translational movement and geometrical transformation (wrapping) are considered. This can be extended to non-integer sub-pixel translation and also to other more general geometrical transformations such as rotational, zooming and other projective transformations. However, the computational costs may increase. Indeed, the estimation of the parameters of such transformations may become more sophisticated. This subject is, by itself, a great area of research in image processing which is called image registration. There are different approaches to this problems. A survey can be found in the following references [5, 6, 28, 36, 37, 39, 43, 46].

### 6.2. Prior modeling of the HR image

The Bayesian framework gives many possibilities to model and so to improve the performances of SR methods. The classical Gauss–Markov modeling accounts only for local smoothness or global continuity and regularity. The main advantage of this simple model is that the estimation steps become easy. However, for many images this simple model is not good enough and we need to account for the presence of contours and regions (spatial geometrical properties of the images) as well as the textures. Generalized Gaussian prior modeling of the images has been proposed and used either explicitly or implicitly via RR terms to go farther and mainly trying to preserve the contours during the reconstruction. However, even if these models have given successful results, we can still go farther in prior modeling and would like to model explicitly the contours and regions in images. The compound Markov modeling is an appropriate tool for this. Two main models have been developed. First is the class of *non-Gaussian one-layer Markov* models with

non-quadratic potential and energy functions

$$p(\boldsymbol{f}) \propto \exp\left\{-\gamma \sum_{\boldsymbol{r}} \phi\left(f(\boldsymbol{r}) - \beta \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} f(\boldsymbol{r}')\right)\right\} \tag{41}$$

where $\beta = 1/|\mathcal{V}(\boldsymbol{r})|$ which is in our case $\beta = 1/4$ and where $\phi(t)$ is a positive potential function such as

$$|t|^{\alpha}, \quad \sqrt{1+t^2} - 1, \quad \log(\cosh(t)),$$
$$\begin{cases} t^2 & |t| \leq T \\ 2T|t| - T^2 & |t| > T \end{cases}$$

or

$$\log(1+t^2), \quad \frac{t^2}{1+t^2}, \quad \arctan(t^2), \quad \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases}$$

and many other convex or non-convex functions.

Second is the class of two-or three-layers (double or triple) compound Markovian model with contours or region labels hidden variables. One of the most used models of this class is a Gauss–Markov model with hidden contour variables

$$p(\boldsymbol{f} \mid \boldsymbol{q}) \propto \exp\left\{-\gamma \sum_{\boldsymbol{r}} \left| f(\boldsymbol{r}) - \beta \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} (1 - q(\boldsymbol{r}')) f(\boldsymbol{r}') \right|^2\right\} \tag{42}$$

where $q(\boldsymbol{r}')$ is a binary valued quantity which is equal to 1 if there exist a contour at the position $\boldsymbol{r}'$ and is equal to 0 elsewhere. In this model $\boldsymbol{q}$ represents the image contours and $\beta$ varies as a function of the number of neighbors which do not rely on the contours, i.e $\beta$ takes one of the values {1/4,1/3,1/2,1} depending on the relative position of the pixel $\boldsymbol{r}$ and the contour positions.

To use this prior, we also need to define $p(\boldsymbol{q})$. The simplest model is the Bernoulli model

$$p(\mathbf{q}) \propto \alpha^{\sum_{\boldsymbol{r}} q(\boldsymbol{r})} (1-\alpha)^{\sum_{\boldsymbol{r}} (1 - q(\boldsymbol{r}))} \tag{43}$$

where $\alpha = P(q = 1)$ and $(1 - \alpha) = P(q = 0)$ which does not account for a correlation between the contour elements. There are also more sophisticated models which account for possible correlations between them (Markovian models).

Indeed, there are links between these two classes of models. For example, a truncated quadratic potential function in the first category becomes equivalent to the two-layer (intensity-contours) model in the second.

The compound model we considered is also a compound Gauss–Markov model, but the hidden variables $\boldsymbol{z}$ we considered correspond to region labels. However, there is a deterministic relation between $\boldsymbol{z}$ and $\boldsymbol{q}$ (see Fig. 13). This model can be written in a compact way as follows

$$p(\boldsymbol{f} \mid \boldsymbol{z}) \propto \exp\left\{-\gamma \sum_{\boldsymbol{r}} \left| \tilde{f}(\boldsymbol{r}) - \beta \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} (1 - q(\boldsymbol{r}')) \tilde{f}(\boldsymbol{r}') \right|^2\right\} \tag{44}$$

with

$$\tilde{f}(\boldsymbol{r}) = \frac{f(\boldsymbol{r}) - m(\boldsymbol{r})}{\sigma^2(\boldsymbol{r})}$$

and

$$m(\boldsymbol{r}) = m_l, \quad \forall \boldsymbol{r} \in \mathcal{R}_l$$

and

$$\sigma^2(\boldsymbol{r}) = \sigma_l^2, \quad \forall \boldsymbol{r} \in \mathcal{R}_l.$$

This model is still more appropriate than the compound model in Equation (42), because naturally, it results to the regions with closed contours $q(\boldsymbol{r})$ which can be deduced from the region labels $z(\boldsymbol{r})$ easily

$$q(\boldsymbol{r}) = \begin{cases} 0 & \text{if } z(\boldsymbol{r}) = z(\boldsymbol{r}'), \quad \forall \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r}) \\ 1 & \text{else} \end{cases} \tag{45}$$

With this model, at the end of the estimation process, we have simultaneously $\hat{\boldsymbol{f}}$, $\hat{\boldsymbol{z}}$, $\hat{\boldsymbol{\theta}}$ and, as a by product, $\hat{\boldsymbol{q}}(\boldsymbol{r})$.

A few references where the Markovian models have been used extensively are [7, 9, 14, 27, 31, 38, 47].

### 6.3. Estimation of the hyper-parameters

Practically, all the regularization based methods have, at least, one parameter (the regularization parameter $\lambda$) which has to be fixed either in an ad hoc way or based on cross validation method or still based on some training data sets. It is also the same for the parameters $\beta_1$ and $\beta_2$ of the RR criteria (24). In the Bayesian estimation methods too, in general, there are few hyper-parameters to be fixed, even if, in theory, it is possible to estimate them during the iterations. For example, in the presented Gaussian mixture model with hidden Potts Markov model (31) for region labels, we fixed, in an ad hoc way, the Potts Markov parameter $\gamma$. Another parameter that we fixed is the maximum number $K$ of region label classes.

### 6.4. Efficient algorithm implementation

The forward and prior modeling result to a posterior probability law via the Bayes' rule. Then, by choosing an estimator based on this posterior law, we try to obtain appropriate solutions to an inversion method. More sophisticated are the forward and prior models, in general, more sophisticated will be the inversion and SR reconstruction algorithm and so the cost of proposed algorithm. Finding a good compromise between the complexity, performances and cost is also a challenge. A few references where this compromise has been discussed are [7, 11, 12, 16, 47–51].

## ACKNOWLEDGMENT

The author would like to thank the editors of the journal for their invitation and encouragement to write this paper and the reviewers whose critics, hints and corrections helped to improve considerably, rewrite and extend the first submission of this paper.

## DOWNLOADS

A set of Matlab programs which have been used to create main figures and results in this paper can be downloaded and used for academic use in `http://djafari.free.fr/SR/sr_demo.tar.gz`.

## REFERENCES

[1] Vega, M., Molina, R. and Katsaggelos, A. (2006) A Bayesian super-resolution approach to demosaicing of blurred images. *EURASIP J. Appl. Signal Process.*, Article ID 25072, **2006**, 1–12.

[2] Woods, N., Galatsanos, N. and Katsaggelos, A. (2006) Stochastic methods for joint registration, restoration and interpolation of multiple undersampled images. *IEEE Trans. Image Process.*, **15**, 201–213.

[3] Sroubek, F. and Flusser, J. (2005) Multichannel blind deconvolution of spatially misaligned images. *IEEE Trans. Image Process.*, **14**, 874–883.

[4] Farsiu, S., Robinson, D., Elad, M. and Milanfar, P. (2004) Fast and robust multi-frame super-resolution. *IEEE Trans. Image Process.*, **13**, 1327–1344.

[5] Park, S.C., Segall, C.A., Kang, M.G. and Katsaggelos, A.K. (2004) Spatially adaptive high-resolution image reconstruction of DCT-based compressed images. *IEEE Trans. Image Process.*, **13**, 573–585.

[6] Segall, C.A., Molina, R., Katsaggelos, A.K. and Mateos, J. (2004) Bayesian resolution enhancement of compressed video. *IEEE Trans. Image Process.*, **13**, 898–911.

[7] Molina, R., Vega, M., Abad, J. and Katsaggelos, A.K. (2003) Parameter estimation in Bayesian high-resolution image reconstruction with multisensors. *IEEE Trans. Image Process.*, **12**, 1655–1667.

[8] Segall, C., Molina, R. and Katsaggelos, A. (2003) High-resolution images from low-resolution compressed video. *IEEE Signal Process. Mag.*, **15**, 37–48.

[9] Pickup, L.C., Roberts, S.J. and Zisserman, A. (2003) A sampled texture prior for image super-resolution. *Adv. Neural Inf. Process. Syst.*, **2003**, 1587–1594.

[10] Capel, D. and Zisserman, A. (2003) Computer vision applied to super resolution. *IEEE Signal Process. Mag.*, **20**, 75–86.

[11] Nguyen, N., Milanfar, P. and Golub, G. (2001) A computationally efficient superresolution image reconstruction algorithm. *IEEE Trans. Image Process.*, **10**, 573–583.

[12] Elad, M. and Hel-Or, Y. (2001) A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur. *IEEE Trans. Image Process.*, **10**, 1187–1193.

[13] Tsap, L.V., Goldgof, D.B. and Sarkar, S. (2001) Fusion of physically-based registration and deformation modeling for nonrigid motion analysis. *IEEE Trans. Image Process.*, **10**, 1659–1669.

[14] Patti, A.J. and Altunbasak, Y. (2001) Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *IEEE Trans. Image Process.*, **10**, 179–186.

[15] Lehmann, T., Gonner, C. and Spitzer, K. (1999) Survey: Interpolation methods in medical image processing. *IEEE Trans. Med. Imaging*, **18**, 1049–1075.

[16] Schultz, R.R., Meng, L. and Stevenson, R.L. (1998) Subpixel motion estimation super-resolution image sequence enhancement. *J. Vis. Commun. Image Represent.*, **9**, 38–50.

[17] Irani, M. and Peleg, S. (1993) Motion analysis for image enhancement: Resolution, occlusion and transparency. *J. Vis. Commun. Image Represent.*, **4**, 324–335.

[18] Irani, M. and Peleg, S. (1991) Improving resolution by image registration. *CVGIP Graph. Models Image Process.*, **53**, 231–239.

[19] Aguiar, P.M.Q. and Moura, J.M.F. (2001) Three-dimensional modeling from two-dimensional video. *IEEE Trans. Image Process.*, **10**, 1541–1551.

[20] Sato, Y. and Ikeuchi, K. (1996) Recovering shape and reflectance properties from a sequence of range and color images. *IEEE/SICE/RSJ Int. Conf. Multisensor Fusion and Integration for Intelligent Systems'96*, December, pp. 493–500.

[21] Daum, M. and Dudek, G. (1998) On 3-D surface reconstruction using shape from shadows. *IEEE Proceedings of CVPR 98*, pp. 461–468.

[22] Xu, X. and Narayanan, R.M. (2001) Three-dimensional interferometric ISAR imaging for target scattering diagnosis and modeling. *IEEE Trans. Image Process.*, **10**, 1094–1102.

[23] Shekarforoush, H., Berchod, M. and Zerubia, J. (1996) Subpixel Image Registration by Estimating the Polyphase Decomposition of Cross Power Spectrum. *Proc. 1996 Conf. Computer Vision and Pattern Recognition* (*CVPR 96*), p. 532.

[24] Schatzberg, A. and Devaney, A.J. (1992) Super-resolution in diffraction tomography. *Inverse Problems*, **8**, 149–164.

[25] Nayar, S., Ikeuchi, K. and Kanade, T. (1989) Shape and Reflectance from an Image Sequence Generated Using Extended Light Sources. *Proc. IEEE Int. Conf. Robotics and Automation* (*ICRA '89*), May, pp. 28–35.

[26] Kirihara, S. and Saito, H. (1998) Shape Modeling from Multiple View Images Using Gas, *Asian Conf. on Computer Vision (ACCV'98), Lecture Notes in Computer Science,* Vol. 1352, pp. 448–454.

[27] Humblot, F. and Mohammad-Djafari, A. (2006) Super-resolution using hidden Markov Model and Bayesian detection estimation framework. *EURASIP J. Appl. Signal Process.*, Article ID 36971, 1–16.

[28] Humblot, F. (2005) Detection of small objects using super-resolution image processing techniques. PhD Thesis, Universit de Paris–sud XI, Orsay, France.

[29] Tsai, R.Y. and Huang, T.S. (1984) Multi-Frame Image Restoration and Registration. *Adv. Comput. Vis. Image Process.*, **1**, 317–339.

[30] Farsiu, S., Robinson, D., Elad, M. and Milanfar, P. (2004) Advances and challenges in super-resolution. *Int. J. Imaging Syst. Technol.*, (Special Issue on High Resolution Image Reconstruction), **14**, 47–57.

[31] Li, X. and Orchard, M.T. (2001) New edge-directed interpolation. *IEEE Trans. Image Process.*, **10**, 1521–1527.

[32] Cavicchi, T. (1992) DFT time-domain interpolation. *IEE Proc. F*, **139**, 207–211.

[33] Vandewalle, P., Süsstrunk, S. and Vetterli, M. (2006) A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP J. Appl. Signal Process*. (special issue on Super-resolution), 2006., Article ID 71459, 14 pages.

[34] Vandewalle, P., Sbaiz, L., Süsstrunk, S. and Vetterli, M. (2006) Registration of aliased images for super-resolution imaging. *SPIE/IS&T Visual Communications and Image Processing Conf.*, pp. 13–23.

[35] Gerchberg, R.W. (1974) Superresolution through error energy reduction. *Opt. Acta*, **21**, 709–720.

[36] Foroosh, H., Zerubia, J. and Berthod, M. (2002) Extension of phase correlation to subpixel registration. *IEEE Trans. Image Process.*, **11**, 188–200.

[37] Rochefort, G., Champagnat, F., Besnerais, G.L. and Giovannelli, J.-F. (2006) Super-resolution from a sequence of undersampled image under affine motion. *IEEE Trans. Image Process.*, **15**, 3325–3337.

[38] Molina, R., Mateos, J., Katsaggelos, A.K. and Vega, M. (2003) Bayesian multichannel image restoration using compound Gauss-Markov random fields. *IEEE Trans. Image Process.*, **12**, 1642–1654.

[39] Humblot, F., Collin, B. and Mohammad-Djafari, A. (2005) Evaluation and practical issues of subpixel image registration using phase correlation methods. *Int. Conf. Physics in Signal and Image Process.* (*PSIP'05*), Toulouse, France, January 31, pp. 115–120.

[40] Humblot, F. and Mohammad-Djafari, A. (2005) Super-resolution and Joint Segmentation in Bayesian Framework. *25th Int. Workshop on Bayesian Inference and Maximum Entropy Methods* (*MaxEnt05*). *AIP Conf. Proc.*, San Jose, CA, US, August, pp. 207–214.

[41] Sroubek, F. and Flusser, J. (2003) Multichannel blind iterative image restoration. *IEEE Trans. Image Process.*, **12**, 1094–1106.

[42] Snoussi, H. and Mohammad-Djafari, A. (2002) Information geometry of prior selection. In Williams, C. (ed.), *Bayesian Inference and Maximum Entropy Methods*, *MaxEnt Workshops AIP Conference Proceedings* 570, August. Univ. of Idaho, Moscow, Idaho, USA.

[43] Ballester, C., Bertalmio, M., Caselles, V., Sapiro, G. and Verdera, J. (2001) Filling-in by joint interpolation of vector fields and gray levels. *IEEE Trans. Image Process.*, **10**, 1200–1211.

[44] Argyriou, V. and Vlachos, T. (2003) Sub-Pixel Motion Estimation Using Gradient Cross-Correlation. *The 7th Int. Symp. Signal Processing and its Applications* (*ISSPA*), Paris, 1–4 July.

[45] Alvarez, L.D., Molina, R., Mateos, J. and Katsaggelos, A. (2004) High resolution images from compressed low resolution video: Motion estimation and observable pixels. *Int. J. Imaging Syst. Technol.*, **14**, 58–66.

[46] Shechtman, E., Caspi, Y. and Irani, M. (2005) Spacetime super-resolution. *IEEE Trans. PAMI*, **27**, 531–545.

[47] Tsai, A., Yezzi, A., Jr and Willsky, A.S. (2001) Curve evolution implementation of the Mumford–Shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Trans. Image Process.*, **10**, 1169–1186.

[48] Vrcej, B. and Vaidyanathan, P.P. (2001) Efficient implementation of all-digital interpolation. *IEEE Trans. Image Process.*, **10**, 1639–1646.

[49] Tom, B.C. and Katsaggelos, A.K. (2001) Resolution enhancement of monochrome and color video using motion compensation. *IEEE Trans. Image Process.*, **10**, 278–287.

[50] Baker, S. and Kanade, T. (2002) Limits on super-resolution and how to break them. *IEEE Trans. Pattern Anal. Mach. Intelligence*, **24**, 1167–1184.

[51] Baramidze, V., Lai, M. and Shum, C. (2006) Spherical splines for data interpolation and fitting. *SIAM J. Sci. Comput.*, **28**, 241–259.