IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 10, 2018, accepted June 26, 2018, date of publication July 9, 2018, date of current version July 30, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2854411

Modification in the SAR Super-Resolution Model Using the Fractal Descriptor LMME in the Term Regularizer

VITOR F. CAMPANA, KLAUS F. CÔCO, EVANDRO OTTONI TEATINI SALLES, AND PATRICK M. CIARELLI

Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Espírito Santo, Vitória 29075-910, Brazil

Corresponding author: Vitor F. Campana (vitor.campana@aluno.ufes.br)

The work of V. F. Campana was supported by IFES—Instituto Federal do Espírito Santo. The work of E. O. T. Salles was supported by FAPES—Fundação de Amparo à Pesquisa e Inovação do Espírito Santo under Grant 244/2016. The work of P. M. Ciarelli was supported by CNPq under Grant 312032/2015-3.

ABSTRACT This paper presents a modified simultaneous auto regressive (SAR) super-resolution model, by adding a fractal operator as a new regularized term. We propose the local morphologic multifractal exponent (LMME) descriptor to determine the hyperparameters estimation of the probability distribution function representing the high resolution (HR) image. Due to the LMME's texture sensibility characteristic, the estimated HR images showed good enhancements in their edge elements and detail regions when compared with the images generated by the original SAR model, at the same time that mitigates the unwanted noise. The proposed modification was also extended to the model that combines the SAR method with ℓ_1-norm. A comparison using peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) metrics is presented in order to measure the quality of the HR images estimated by the proposed methods in relation to other available. The use of the LMME descriptor presented better adjustments of the restored image estimation distributions parameters, reflected on the values of PSNR and SSIM obtained in the experiments using low resolution images with moderate noise levels (SNR of 30 dB and 40 dB). In addition, our proposal converges in fewer iterations.

INDEX TERMS Image processing, super-resolution, SAR, statistical fractal descriptor, regularization.

I. INTRODUCTION

Increasing image resolution can be useful in various fields, such as surveillance, license plate recognition, criminal identification, medical diagnosis, satellite images and computer vision. In such cases, Super-resolution (SR) methods generally retrieve a high resolution (HR) image from one image (singleframe) or a sequence of low resolution (LR) images (multi-frame) [1]. As shown in [14], SR techniques have been applied to the major medical imaging modalities, including magnetic resonance imaging (MRI) and positron emission tomography (PET). In the field of surveillance, license plate recognition and criminal identification, SR can improve the resolution of images generated by security cameras, allowing the identification of people or recognition of characters. In the field of astronomy, SR methods can enhance the resolution of satellites images, and thus can help describing small objects. As it is the case of images of the surface of the Moon.

Recently, several methods of single-frame SR have been proposed to deal with these problems [11]–[13]. The single-frame methods of SR can be divided into example learning-based methods [15], interpolation-based methods [16] and reconstruction-based methods [17]. Example learning-based methods recover a HR image by learning the mapping between HR and LR images [19]. Reconstruction-based methods establish constraints on image reconstruction and prior model to formulate a regularized cost function [27]. Examples of priors applied to reconstruction-based methods: edge prior [20], gradient prior [18], non-local means prior [21] and sparsity prior [22]. Interpolation-based methods [23] show the advantage of simplicity and low complexity. Due to the ill-posed problem of SR reconstruction, recovering the HR image from a single LR image is dificult and the correlation between HR image and LR image is not considered important [8].

Multi-frame SR methods can be divided into two basic types: frequency domain and spatial domain [24]. Although frequency domain methods are eficient, they are not able to incorporate prior knowledge from spatial domain in their

formulation [8]. In order to overcome that disadvantage, many spatial domain methods were developed in the research community, including the non-uniform interpolation-based approaches, iterative back projection (IBP) approaches, projection onto convex sets (POCS) approaches, maximum likelihood (ML) approaches, and maximum a posteriori (MAP) approaches [25].

This article discusses the case of estimating a HR image of a set of observed LR images. SR methods usually include two steps: registration, where the motion between LR images is estimated, and image reconstruction, where the HR image is recovered from the LR images [1]– [4].

In SR multi-frame problems, an image model is used, where the problem is solved iteratively to minimize a cost function [27]. Among the spatial domain approaches, the regularization-based method is one of the most effective multi-frame SR reconstruction approaches [8]. In the SR problem, modeling errors and noise generate unstable solutions and ill-posed nature, making regularization necessary to achieve stable solutions, using a prior knowledge [11], [14]. Models based on prior knowledge depend on parameters that are random variables and they also need to be modeled. These parameters are generally called hyperparameters and their models are called hyperpriors [1]–[4].

An important issue in SR is to find a trade-off between preserving edge information and removing noise. For this, prior regularization models have been proposed, such as Tikhonov regularization, Markov random fields (MRF) regularization, and total variation (TV) regularization [11].

In this article, we propose a modification of an SR method, which is composed by the following steps: registration, hyperparameter estimation and HR image reconstruction, based on a Bayesian perspective, as shown in [2]–[8], [10], [11], [14], and [26]. The Bayesian framework provides uncertainties of the estimates during the restoration process, which helps to prevent error-propagation and improves robustness. Other advantage is that all required algorithmic parameters are estimated along with the HR image and the motion parameters, causing the algorithms not to require user supervision [4].

Considering that statistical fractal descriptor is sensitive to edge elements [33], the objective of this article is to propose a fractal operator in an SR multi-frame problem to reduce unwanted noise and preserve important aspects of the original image.

The main contributions of this paper are the employment of Local Morphologic (LMME) descriptor to determine the hyperparameters in a model based on hierarchical Bayesian inference and to show that in images with high level of detail (edge elements) the application of this new model is more indicated. Another contribution of this proposal is the reduction of the number of necessary iterations for convergence of the algorithm.

This article is organized as follows: Section II presents the SR method according to the Bayesian approach, the concept of the LMME fractal descriptor and a mathematical modeling of the SAR model and the LMME descriptor. In Section III we present our proposal of modifying the SAR model through the mathematical formulation. After showing our model, we present in Section IV the methodology adopted to perform the experiments and the results obtained from such experiments. In this section we chose some images to detail the results and, then, we presented the results as the average value of the experiments performed with 23 images, in order to enable a statistical analysis. Section V provides a better overview of the fields of application of the SR method presented in the article. Finally, in Section VI, we conclude the text presenting our considerations on the results and, also, our proposals for future works.

II. BAYESIAN MODELING FOR SUPER RESOLUTION

Given a set $Y = (Y_1, \ldots, Y_L)$ of L observed LR images, each of size $N = N_h \times N_v$ pixels. The objective is estimate a HR image X of size $M = M_h \times M_v$, where $M_h = P \times N_h$ and $M_v = P \times N_v$, from the set Y of images using the Bayesian formulation. The variable $P > 1$ is the increasing factor on resolution. The images X and Y_k $(k = 1, \ldots, L)$ are represented in lexicographic form. The relationship between the LR images and the original HR image is modeled in SR problems according to

$$Y_k = AH_kQ(S_k)X + N_k = B_kX + N_k, \tag{1}$$

with the matrix $B_k = AH_kQ(S_k)$; A is the $N \times PN$ matrix representing the downsampling operation of the observed LR image acquisition process (Y_k); H_k is the $PN \times PN$ matrix representing the blurring operation, modeled here as an arithmetic mean filter, equivalent to a low-pass filter; $Q(S_k)$ is the $PN \times PN$ warping matrix generated by the motion vector S_k (translation and rotation); X is the HR image to be estimated and N_k is additive white Gaussian noise (AWGN) model [1]–[8], [10], [11].

In this work, as in [4], the blur matrix H_k is considered to be known and the motion model $Q(S_k)$ consisting of translation and rotation, that is $S_k = (\theta_k, c_k, d_k)^t$, where θ_k is the rotation angle, and c_k and d_k are, respectively, the horizontal and vertical translations of the kth HR image with respect to the reference frame X.

For simplicity, it is common to assume that during the observation the image was contaminated by additive white Gaussian noise, whose distribution has mean zero and standard deviation σ_n [36]. In this case,

$$p(Y_k|X) = \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left\{-\frac{\|B_kX - Y_k\|^2}{2\sigma^{2_n}}\right\}, \tag{2}$$

represents the likelihood, where $B_k = AH_kQ(S_k)$.

The estimated image $\hat{X}$ can be expressed by

$$\hat{X} = \arg\min_X \left\{-\frac{\|B_kX - Y_k\|^2}{2\sigma_n^2} - \log p(X)\right\}, \tag{3}$$

where $-\dfrac{\|B_kX - Y_k\|^2}{2\sigma_n^2}$ is called fidelity and $\log p(X)$ is the regularization term.

It is common to use MRF with Gibbs distribution as image priors in Bayesian image restoration and SR problems [4], [5] and [7]. The Gibbs distribution has the following form

$$p(X) \propto \exp\{-\gamma \Gamma(X)\}, \tag{4}$$

where $\Gamma(.)$ is an energy function that defines the relationship among neighbouring image pixels. The constant γ determines how the image prior term affects to the overall maximization problem in (3) [7]. In an SR process via Gaussian Markov Random Fields (GMRF), the $\Gamma(.)$ function represents the square sum of local pixel differences. In the Simultaneous Auto Regressive (SAR) model, the $\Gamma(.)$ function is simplified to a Laplacian operator matrix C as described in Subsection A. The value of γ controls the influence of the regularization on the solution. Setting γ to zero simplifies (3) to the maximum likelihood (ML) solution, where the noise is amplified while the details are preserved. Increasing γ will increase the effect of the GMRF prior, in this case, the differences among neighbouring pixels will be more penalised and the high-frequency content of the reconstructed image will decrease [7].

A. THE SAR MODEL

The SAR statistical model was used in the SR context in [2], [4], and [10]. It is a quadratic model (ℓ_2 norm) that smoothes the border structures of the images and attenuates noise. On such model, the prior is defined as

$$p_1(X|\alpha) \propto \alpha^{\frac{M^2}{2}} \exp\left\{-\frac{\alpha}{2}\|CX\|^2\right\}, \tag{5}$$

where M is the dimension of the HR image, α is the hyperparameter inherent to the model, C is the Laplacian operator, equivalent to a high-pass filter. However, there is a smoothing effect due to the presence of the quadratic term [2], [29].

B. HIERARCHICAL BAYESIAN INFERENCE

From (1), the SR problem can be defined as the estimation of the HR image X from the set Y of LR images using the prior knowledge about $Q(S_k)$, N_k, and X.

For the unknown X, a prior is defined by $p(X|\alpha)$. Prior distributions $p(S_k)$ are assigned to the unknown S_k, for $k = 1, \ldots, L$. The LR images in Y are observations of a random process with the corresponding conditional distribution $p(Y|X, S_k, \beta_k)$. These distributions depend on the additional parameters α and β_k (called hyperparameters), which are modeled by hyperprior distributions.

The hierarchical Bayesian framework treats the SR problem through two stages. The first stage is used to model the acquisition process, the unknown HR image X and the motion vectors S_k. The second stage is used to model the hyperparameters [3]. These stages are iterative processes, that is, they are repeated consecutively until reaching a convergence criterion. Fig. 1 shows a flow diagram of the process.

Once the LR images are acquired, the first step is to perform the interpolation of the reference LR image (usually the first one that was acquired) onto an HR grid. Then, the first estimation of the motion is made using the LR image to get the reference frame. At the first time, the regularization term is also calculated using the interpolated LR image, as shown by arrow 1 in Fig. 1. As in the first iteration there is still no estimated HR image, the interpolated image is used together with the registration and the regularization term to perform the first restoration (estimation) of the HR image, as shown by the arrows marked with the number 2 in Fig. 1.

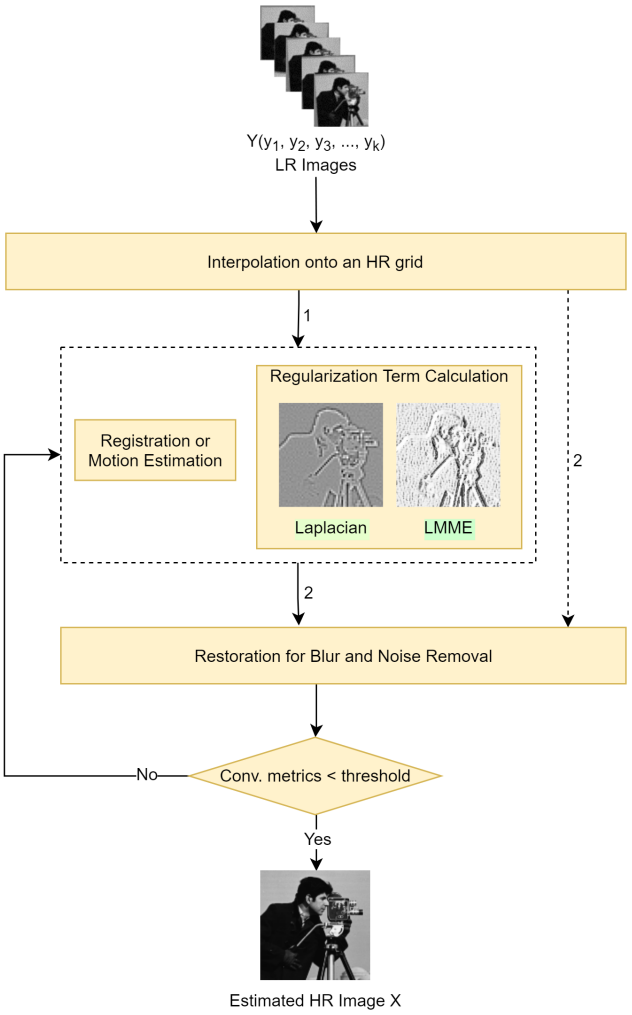


FIGURE 1. Flow diagram of the Super-resolution methodology.

The dashed arrow 2 indicates that the interpolated image is only used in the first estimation of the image. In the second iteration onwards the estimation of the motion and the regularization term are calculated based on the HR image estimated in the previous iteration. When the stopping criterion is reached, the algorithm returns the final estimated HR image.

C. FRACTALS

Fractals are defined by the property of self-similarity over different image scales [33]. According to Turner [41], there are two groups of fractals: geometric and statistical.

Geometric fractals or objects of deterministic self-similarity are composed of distinct features that are returned

by every other identical object at different scales (scale-invariant features).

These objects are modeled by the equation that defines self-similarity, expressed as

$$X(i,j) = \rho^{-d} X(\rho i, \rho j), \tag{6}$$

where ρ is the scaling factor, d is the scaling exponent or self-similarity parameter, and X is the image.

Statistical fractals are those that preserve their statistical properties at different scales.

Statistical fractals are formulated according to

$$pdf(X(i,j)) = \rho^{-r} pdf(X(\rho i, \rho j)), \tag{7}$$

where pdf is the probability density function of X.

In the real world, the most frequent fractal description (self-similar object) is statistical, and it is useful for modeling certain textures [41].

It is important to mention that statistical fractal descriptors are good edge descriptors, perform high-order statistical analysis [33] and interpret a 2-D surface as a 3-D model [43].

Of several definitions of fractal dimension existent, here we use the concept of box-counting [40]. The fractal dimension (D) is a measure for classifying textures [40], given by

$$D = \frac{\log(N_r)}{\log(1/r)}, \tag{8}$$

where r is a scaling reduction factor and

$$N_r = \sum_{i,j} n_r(i,j), \tag{9}$$

where n_r is the number of boxes, which represent gray level, counted in the grid (coordinate) (i,j) and N_r is the quantization of the gray level in the whole partitioned image in grids of dimensions $s \times s$ [40].

D. LMME DESCRIPTOR

An adaptation of the model presented in [40], called Local Multifractal Morphological Exponent (LMME), was implemented in [43] using morphological structuring elements to delimit the local region. In this model, an image $M \times N$ is considered 3-D defined as

$$\{i, j, X(i,j)\}\,; i = 1, \ldots, M, j = 1, \ldots, N. \tag{10}$$

For a scale ϵ, and a structuring element Y_ϵ, we have $\{i_{\epsilon k}, j_{\epsilon k}, \beta_\epsilon\}$, where $k = 1, \ldots, P_\epsilon$, with P_ϵ representing the number of pixels in Y_ϵ and β is the shape factor (height of Y_ϵ) [43].

The X image dilation operation with Y_ϵ generates the X_ϵ image. Each pixel (i,j) of X_ϵ is calculated by

$$X_\epsilon(i,j) = \max_{k=1,2,\ldots,P_\epsilon} \{X(i + i_{\epsilon k}, j + j_{\epsilon k}) + \beta_\epsilon\}. \tag{11}$$

The local natural measure of a pixel (i,j) is defined in a window of size $W \times W$ centered on (i,j) as

$$\mu_\epsilon(i,j) = \frac{|X_\epsilon(i,j) - X(i,j)|}{\sum_{i,j}^{W} |X_\epsilon(i,j) - X(i,j)|}. \tag{12}$$

The intensity measure of order m at scale ϵ represents the intensity of the gray levels of the window and is calculated according to

$$I(m,\epsilon) = \sum_{i,j}^{W} \mu_\epsilon(i,j)^m. \tag{13}$$

In [33] a set of mulifractal texture descriptors was proposed, namely the LMME of the window, Lm, as a modification of the LMME proposed in [43]. In (14) the new formulation is presented. This calculation defines a relation between the intensity measure of order m at scale ϵ of the window [33]. Such relationship is the measure of statistical self-similarity of the window.

$$Lm(i,j) = \frac{1}{1-m} \lim_{\epsilon \to 0} \frac{\ln I(m,\epsilon)}{\ln \frac{W}{\epsilon}}. \tag{14}$$

The grid's LMME (15), also proposed in [33], is called fractal spectrum and defines a relation between the pixel's local natural measure and the grid structuring element's scale of the window. This relationship represents a measure of statistical self-similarity in the grid, as follows

$$Lm(i,j) = \frac{1}{1-m} \lim_{\epsilon \to 0} \frac{\ln \mu_\epsilon(i,j)}{\ln \frac{W}{\epsilon}}. \tag{15}$$

Thus, the LMME is the parameter of statistical self-similarity and represents the linear relationship between the local natural measure and the grid scales in logarithmic space.

The LMME's algorithm is summarized in Algorithm 1.

Algorithm 1 LMME for an image $M \times N$

```
Create positions masks of the Y_ε in the window W × W for each scale ε
for i = 1:M do
    for j = 1:N do
        Create the image window of specified size W × W centered on (i, j)
        for each scale ε do
            Calculate X_ε(i, j) by using (11)
            Calculate μ_ε(i, j) by using (12)
            Calculate the measure of statistical self-similarity Lm(i, j) in the grid using (15)
        end for
    end for
end for
return Lm
```

In order to observe the level of detail preserved by the LMME of the original image (Fig. 2 (a)), the results of the LMME descriptor (Fig. 2 (b)) and the Laplacian operator (Fig. 2 (c)) are shown in Fig. 2.

As we can see, there is a greater level of detail when applying the fractal descriptor, Fig. 2(b), in relation to the Laplacian (linear operator), Fig. 2(c).

(a) (b) (c)

FIGURE 2. (a) Original butterfly image [37], (b) image composed by the LMME descriptors and (c) image obtained with the Laplacian operator.

III. METHODOLOGY

Among the Super-Resolution models, the SAR model stands out as a good noise attenuator in the process of restoring the estimated HR image, although this feature causes the edge elements to be lost. Looking at the prior of such model, presented in (5), we note that α adjusts the influence of high frequency components. In order to obtain a better trade-off between noise attenuation and preservation of edge elements, we propose the use of another edge element detection operator in the hyperparameter α calculation. In this work, we include the LMME descriptor in the formulation because it is sensitive to the edge elements, as can be seen by comparing Figs. 2 (b) and 2 (c).

In [3], [4], [8], and [10], the hyperparameters are modeled using a Gamma distribution as

$$p(\theta) = \Gamma(\theta|a_\theta^0, b_\theta^0) = \frac{(b_\theta^0)^{a_\theta^0}}{\Gamma(a_\theta^0)}\theta^{a_\theta^0-1}\exp[-b_\theta^0\theta], \quad (16)$$

where θ is the random variable; a_θ^0 and b_θ^0 are the parameters of Gamma distribution. To a_θ^0 and b_θ^0, "0" (zero) indicates their initial estimation values and θ is the random variable to which the parameter is associated. The hyperpriors are chosen as Gamma distributions, since they are conjugate priors for the Gaussian distributions [1], [4].

In the original SAR model

$$a_\theta^0 = \frac{PN}{2} + a_\alpha^0 \quad (17)$$

and

$$b_\theta^0 = b_\alpha^0 + \frac{E_x[\|CX\|^2]}{2}, \quad (18)$$

substituting a_θ^0 and b_θ^0 in (16) the following distribution is obtained:

$$q(\alpha) \propto \alpha^{\frac{PN}{2}-1+a_\alpha^0}\exp\left[-\alpha\left(b_\alpha^0 + \frac{E_x[\|CX\|^2]}{2}\right)\right], \quad (19)$$

where a_α^0 and b_α^0 are parameters of hyperprior Gamma and $E_x[X]$ is the expected value of X [3].

In this work, we modify the term of the SR process related to edge detection, which in the SAR model is the Laplacian operator C. Note that the operator C directly affects the estimation of the α value because it is part of the hyperprior $q(\alpha)$, which in turn affects the relevance of the term in the *pdf* used to estimate the value of the image X. The function of the regularization term is to compensate for the lack of prior information about the desirable HR solution. Thus, with the use of the LMME descriptor, we expect to improve the adjustment of regularization term, which in SR is the penalization term, on the fidelity term in order to preserve more details of the estimated image and to mitigate noise.

In the model proposed in this article, the Laplacian operator was replaced by the LMME descriptor in the hyperprior formulation.

In the modified SAR model (SARM) we consider

$$b_\theta^0 = \left(b_\alpha^0 + \frac{E_x[\|Lm(X)\|^2]}{2}\right), \quad (20)$$

and (19) can be rewritten as

$$q(\alpha) \propto \alpha^{\frac{PN}{2}-1+a_\alpha^0}\exp\left[-\alpha\left(b_\alpha^0 + \frac{E_x[\|Lm(X)\|^2]}{2}\right)\right], \quad (21)$$

where $Lm(X)$ is the statistical fractal descriptor applied throughout the image X with predefined structuring element in a given scale.

A. IMAGE ESTIMATOR

The SR problem with the SAR model used in [3], [4], and [10] is solved using Expectation and Maximization (EM), considering the expectation step the estimation of the image X, modeled with the distribution $q(X)$ given as

$$q(X) \propto \exp\left(-\frac{\langle\alpha\rangle}{2}\|CX\|^2 - \frac{1}{2}\sum_{k=1}^{L}\langle\beta_k\rangle\|Y_k - B_kX\|^2\right), \quad (22)$$

where $\langle.\rangle$ represents expected value, α and β are described by a Gamma type *pdf* and $\beta_k = 1/\sigma_n^2$ [3]. The $q(X)$ is a multivariate Gaussian distribution

$$q(X) = \mathcal{N}(X|E_x[X], cov_{q(X)}^{-1}[X]), \quad (23)$$

with

$$E_x[X] = cov_{q(X)}[X]\sum_{k=1}^{L}\langle\beta_k\rangle B_k^T Y_k, \quad (24)$$

and the precision matrix is given by

$$cov_{q(X)}^{-1}[X] = \sum_{k=1}^{L}\langle\beta_k\rangle B_k^T B_k + \langle\alpha\rangle C^T C. \quad (25)$$

Finally, the maximization step involves the estimation of α and β_k.

To estimate the hyperparameter $\langle\alpha\rangle$, in the SAR model presented in [3], [4], and [10] the expected value of the distribution given in (19) is calculated as a Gamma distribution (16) with $E[\theta] = a_\theta^0/b_\theta^0$. Using (17) and (18),

$$\langle\alpha\rangle = \frac{PN + 2a_\alpha^0}{E_x\left[\|CX\|^2\right] + 2b_\alpha^0}. \quad (26)$$

B. HYPERPARAMETER ESTIMATOR

In the proposed model, α is calculated using the expected value of the distribution given in (21), thus

$$\langle\alpha\rangle = \frac{PN + 2a_{\alpha}^{0}}{E_x\left[\|Lm(X)\|^2\right] + 2b_{\alpha}^{0}}. \tag{27}$$

The estimation of β_k is calculated in the same way as in [4]:

$$q(\beta_k) \propto \beta_k^{\frac{PN}{2}-1+a_{\beta_k}^{0}} \times \exp\left[-\beta_k\left(b_{\beta_k}^{0} + \frac{E_x[\|Y_k - B_kX\|^2]}{2}\right)\right]. \tag{28}$$

For the function $\Gamma(.)$, we have that $E[.] = a/b$, therefore

$$\langle\beta_k\rangle = \frac{PN + 2a_{\beta_k}^{0}}{E_x\left[\|Y_k - B_kX\|^2\right] + 2b_{\beta_k}^{0}}. \tag{29}$$

With this new model we expect to get better enhancement of the edge elements of the image and better sharpness at the image restored by the SR process, since the modification is being made in the precision matrix that finds the distribution function that models the estimated image X. As mentioned in [4], the hyperparameters α and β_k are crucial for the performance of the SR algorithm. The LMME fractal descriptor is much more sensitive to edge detector than the Laplacian operator, generating a better sharpness in the HR image estimated by the new model. This will be observed through visual inspection in the tests performed with LR images with high noise contamination, or by the metrics, in cases of LR with no high amount of noise.

C. MODEL BASED ON L1-SAR

Another model also proposed in this work is based on the ℓ_1-SAR presented by [4], wich combines SAR and ℓ_1-norm. In this proposal the term with LMME fractal descriptor will act as a fine adjustment for the detection of edge elements, since ℓ_1-norm also holds this characteristic. The advantage in the resulting HR image quality will be shown through the metrics in the experiments.

The formulation of the ℓ_1-SAR model [4] is given by

$$q_{\ell 1}(X) \propto \Bigg(-\frac{1}{2}\lambda\Bigg[\langle\alpha_2^h\rangle\sum_i\frac{(\Delta_i^h(X))^2 + w_{2i}^h}{\sqrt{w_{2i}^h}} + \langle\alpha_2^v\rangle\sum_i\frac{(\Delta_i^v(X))^2 + w_{2i}^v}{\sqrt{w_{2i}^v}}\Bigg] - \frac{(1-\lambda)}{2}\langle\alpha\rangle\|CX\|^2 - \frac{1}{2}\sum_k\langle\beta_k\rangle\|Y_k - B_k(S_k)X\|^2\Bigg), \tag{30}$$

where λ, value between 0 and 1, adjusts the combination between the two models (ℓ_1-norm and SAR).

The distribution presented in (30) has the following inverse covariance matrix

$$cov^{-1}_{q_{\ell 1}(X)}[X] = \sum_{k=1}^{L}\langle\beta_k\rangle B_k^T B_k + \lambda\left(\langle\alpha_2^h\rangle\Delta^{hT}W(w_2^h)\Delta^h + \langle\alpha_2^v\rangle\Delta^{vT}W(w_2^v)\Delta^v\right) + (1-\lambda)\langle\alpha\rangle C^T C, \tag{31}$$

where, Δ^h and Δ^v represent the $PN \times PN$ convolution matrices associated, respectively, with the horizontal and vertical first order diferences, and $W(w)$, $\forall w \in (R^+)^{PN}$, is a $PN \times PN$ diagonal matrix with elements $W(w)_{ii} = \frac{1}{\sqrt{w_i}}$, for $i = 1, \ldots, PN$.

The estimated image in the ℓ_1-SAR model is given by

$$E_x[X] = cov_{q_{\ell 1}(X)}[X]\sum_{k=1}^{L}\langle\beta_k\rangle B_k^T Y_k. \tag{32}$$

IV. EXPERIMENTS

In this section, we present the tests and results obtained from the experiments. The SARM method, proposed in this work, and the combination of SARM and ℓ_1-norm (ℓ_1-SARM) are compared with other relevant methods, such as TV, SAR, ℓ_1-norm, ℓ_1-SAR and TV-SAR mentioned in [4], similar to what was done in [26]. In the analysis of the results it will be emphasized the performance comparison between the SAR and SARM methods, and between ℓ_1-SAR and ℓ_1-SARM to highlight the gain obtained by the use of the proposed LMME descriptor.

A. METHODOLOGY EMPLOYED IN THE EXPERIMENTS

The images presented in Fig. 3 were used to perform the experiments. Some of these images were chosen because they are commonly used in works published in the area of image processing and others have been included in order to have a greater variation of image type. The resolution of the images are 80×80, 120×90 and 120×120.

For each HR image five LR images were simulated using the same image acquisition model as in [26]. First, the HR image was shifted in the horizontal (c_k) and vertical (d_k) directions and rotated (θ_k) using five different random motion vectors S_k ($-7.74 < \theta_k < 7.74$, $-10 < c_k < 10$ and $-10 < d_k < 10$).

Then, the effect of the point spread function of the camera was simulated by selecting an average kernel of size 3×3, as in [4]. In the next step, the resulting images were downsampled by a factor of 2 in both directions. Finally, the AWGN model with different power values was added. For the first set of experiments, the SNR value was 20 dB; for the second one, it was 30 dB; and for the third one, it was 40 dB.

Regarding the EM algorithm convergence criteria, the first one we used in a subset of experiment is related to the mean square error (MSE). To calculate it we evaluate $\|\hat{X}_j - \hat{X}_{j-1}\|^2/\|\hat{X}_{j-1}\|^2 \leq 10^{-5}$ where $\hat{X}_j$ and $\hat{X}_{j-1}$ are the image estimations at the *jth* and $(j-1)$*th* iterations, respectively.

For another subset of experiments, we used the Structural Similarity Index (SSIM) metric as convergence criterion.

SSIM develops a measure of structural similarity that compares local intensity patterns that are normalized by

FIGURE 3. Images used in the experiments: (a) cameraman (120 × 120), (b) macaw (80 × 80), (c) pepper (80 × 80), (d) butterfly (120 × 90), (e) ring (80 × 80), (f) car (80 × 80), (g) house (80 × 80), (h) armed car (80 × 80), (i) satellite (80 × 80), (j) door (80 × 80), (k) lena (80 × 80), (l) mandril (80 × 80), (m) elaine, (n) hair and dress (80 × 80), (o) card (120 × 120), (p) target (120 × 120), (q) texture D9 [38] (80 × 80), (r) texture D12 [38] (80 × 80), (s) license-plate (88 × 88), (t) MRI image 1 (120 × 120), (u) MRI image 2 (120 × 120), (v) MRI image 3 (120 × 120), (w) newspaper clipping (120 × 98).

luminance and contrast, defined by

$$SSIM(X,\hat{X}) = \frac{(2\mu_X\mu_{\hat{X}} + c_1)(2\sigma_{X\hat{X}} + c_2)}{(\mu_X^2\mu_{\hat{X}}^2 + c_1)(\sigma_X^2 + \sigma_{\hat{X}}^2 + c_2)} \qquad (33)$$

where μ_X and $\mu_{\hat{X}}$ are the mean of X and $\hat{X}$, respectively; σ_X and $\sigma_{\hat{X}}$ represent the variances of X and $\hat{X}$, respectively; $\sigma_{X\hat{X}}$ is the covariance between the two signals and c_1 and c_2 are constants that stabilize the division, in the case of a denominator close to zero [42].

The values of SSIM are in the range $-1 \leq SSIM(X,\hat{X}) \leq 1$. SSIM value aproximates to 1 when X and $\hat{X}$ are closer.

In this case, the EM algorithm stopping when $SSIM(\hat{X}_j, \hat{X}_{j-1}) \geq 0.99$.

For the LMME fractal descriptor, the following values were used for the parameters: window size $W = 5$; shape of structuring element = ''square''; variations of scales $\epsilon = [2, 3]$; shape factor $\beta = 3$ and order $m = 1$. These parameter values were defined in order to keep a compromise between computational cost and edge image improvement.

In this article, the experiments were performed with the 23 images presented in Fig. 3, using the 3 aforementioned SNR values (20 dB, 30 dB and 40 dB). For each SNR value and SR method considered in the analysis the tests were performed 5 times using the MSE convergence criterion and 5 other times using the SSIM convergence criterion. For each executed test the motion vectors S_k were randomly generated.

In the model ℓ_1-SARM the value of λ was chosen to be equal to 0.9. This value was chosen so that the fractal descriptor would not interfere so much in the model, since the ℓ_1-norm model already presents edge detection characteristic.

B. RESULTS

Fig. 4 presents the results of the application of SR methods discussed in this work, where the LR images, derived from Fig. 3 (a), have 62 × 62 pixels. For the simulations showed in Figs. 4 (b), (c) and (d) noise providing SNR of 20 dB, 30 dB and 40 dB was added, respectively. The convergence criterion used was the MSE.

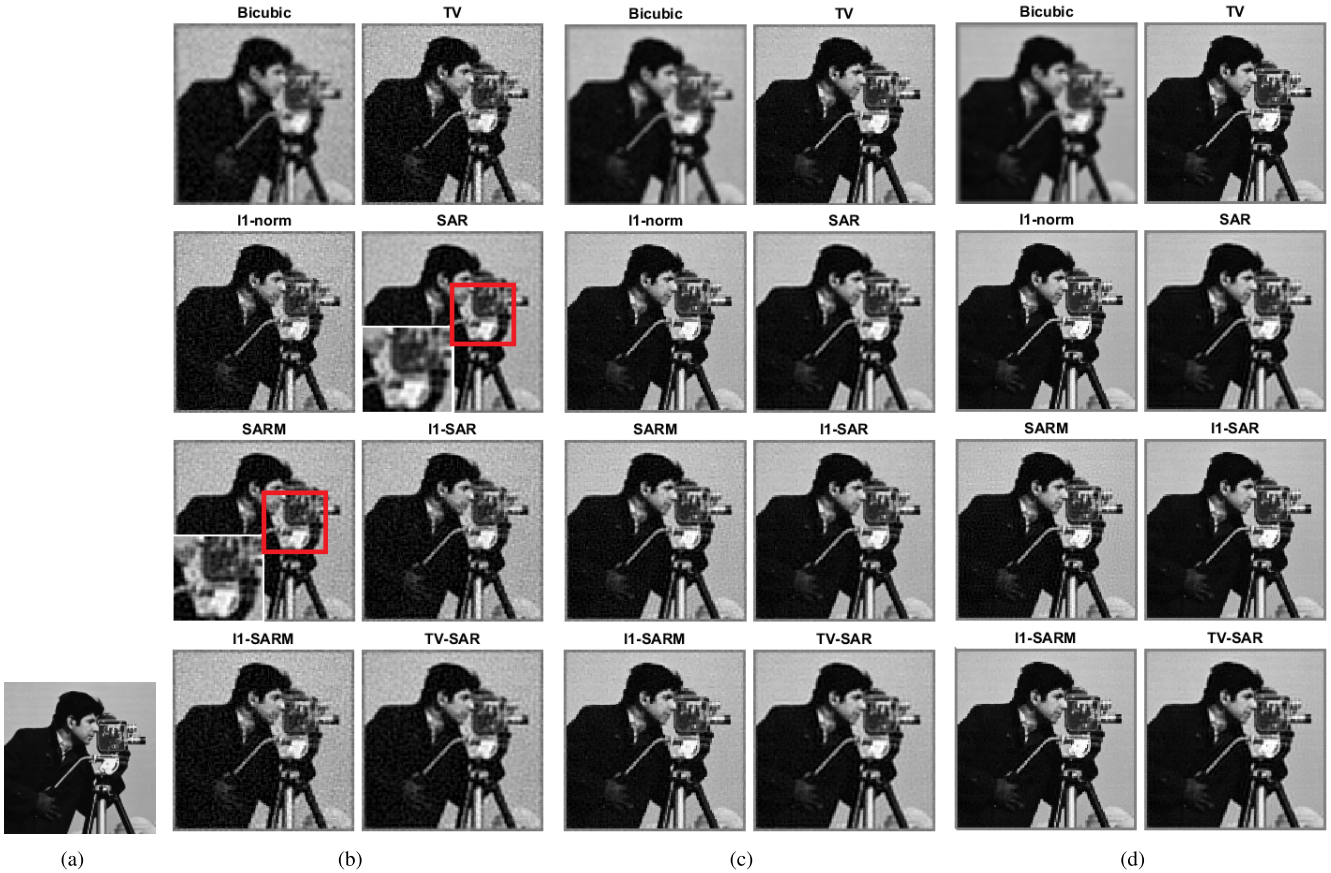


FIGURE 4. Comparative: Fig. 3 (a) (120 × 120) using MSE convergence criterion. (a) Original image, (b) additive noise of SNR = 20 dB, (c) additive noise of SNR = 30 dB and (d) additive noise of SNR = 40 dB.

TABLE 1. Comparison of the PSNR and SSIM values (mean ± standard deviation) among the SR methods applied to Fig. 3 (a). The values in yellow are the cases where the proposed model is better than the original model SAR and l1-SAR. The best values among all the tested methods are in bold.

	TV	l1-norm	SAR	SARM	l1-SAR	l1-SARM	TV-SAR
SNR = 20 dB							
MSE convergence							
PSNR	23.58±2.1E-01	23.57±2.1E-01	23.55±1.7E-01	23.48±1.6E-01	23.74±2.2E-01	23.71±2.0E-01	**23.83±2.0E-01**
SSIM	0.576±1.3E-02	0.572±1.3E-02	**0.704±9.3E-03**	0.590±8.9E-03	0.590±1.3E-02	0.587±1.1E-02	0.647±1.3E-02
SSIM convergence							
PSNR	23.78±2.0E-01	23.82±2.0E-01	23.32±1.4E-01	23.51±1.6E-01	**23.99±2.5E-01**	23.94±2.0E-01	23.86±1.9E-01
SSIM	0.599±1.2E-02	0.598±1.3E-02	0.564±8.7E-03	0.599±1.2E-02	0.622±2.1E-02	0.612±1.1E-02	**0.662±1.1E-02**
SNR = 30 dB							
MSE convergence							
PSNR	28.27±1.3E-01	**28.50±1.4E-01**	25.57±6.8E-02	26.29±5.6E-02	28.07±1.0E-01	28.36±1.2E-01	26.48±8.2E-02
SSIM	0.824±2.8E-03	**0.829±3.0E-03**	0.823±1.7E-03	0.707±3.0E-03	0.827±2.5E-03	0.819±3.0E-03	0.821±1.8E-03
SSIM convergence							
PSNR	27.75±1.5E-01	**27.95±1.1E-01**	25.74±6.3E-02	26.27±6.1E-02	27.62±9.3E-02	27.88±9.9E-02	26.30±7.2E-02
SSIM	0.827±2.8E-03	**0.834±2.2E-03**	0.815±1.7E-03	0.763±2.2E-03	0.832±1.9E-03	0.826±2.2E-03	0.824±1.7E-03
SNR = 40 dB							
MSE convergence							
PSNR	32.30±5.7E-02	**32.62±2.6E-02**	27.30±2.8E-02	29.02±4.7E-02	32.13±5.8E-02	32.57±3.8E-02	29.13±5.2E-02
SSIM	0.928±7.6E-04	**0.934±4.7E-04**	0.875±9.2E-04	0.784±1.1E-03	0.929±6.2E-04	0.930±5.6E-04	0.896±7.3E-04
SSIM convergence							
PSNR	29.00±7.6E-02	**29.76±5.1E-02**	26.02±2.1E-02	26.79±2.7E-02	29.20±4.6E-02	29.74±5.1E-02	27.19±3.0E-02
SSIM	0.923±2.9E-04	**0.932±2.1E-04**	0.876±6.8E-04	0.877±8.1E-04	0.924±7.3E-05	0.930±1.8E-04	0.893±5.3E-04

In Fig. 4, we can observe the edge smoothing characteristic inherent to the SAR method in relation to the sharpening characteristic from the other methods (TV and ℓ_1-norm). Comparing the image restored by SAR model and SARM, in Fig. 4 (b), it is possible to observe the effect of the sensitivity of the LMME fractal descriptor in relation to the Laplacian

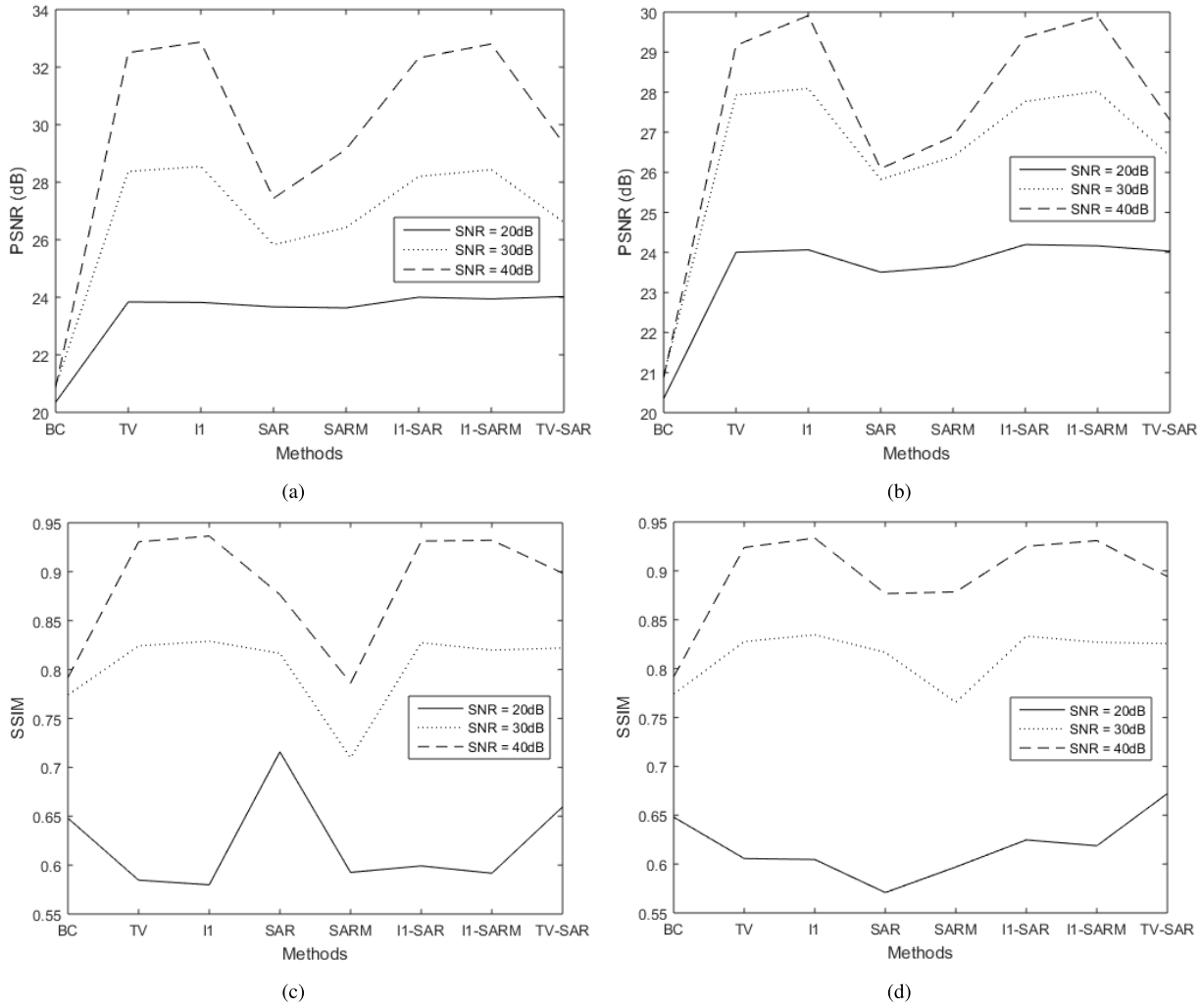


FIGURE 5. Comparing the mean value of PSNR and SSIM metrics versus AWGN (in dB) after performing 5 times the test with Fig. 3 (a). (a) PSNR using MSE convergence criterion, (b) PSNR using SSIM convergence criterion, (c) SSIM using MSE convergence criterion and (d) SSIM using SSIM convergence criterion.

operator. The image restored by the SARM method is sharper and preserves the edge detail of the image compared to the image restored by the SAR method. On the other hand, when comparing models ℓ_1-SAR and ℓ_1-SARM such disparity is not visually noticeable. In this case, it is necessary to analyze the metrics Peak Signal-to-Noise Ratio (PSNR) and SSIM.

Still analyzing Fig. 4, a granulated effect is noticed when the LR image's SNR is 20 dB and decreases when noise reduces, as can be seen in Figs. 4 (c) and (d). This is a side effect of preserving edge elements in the restored image. However, the granulated effect may justify poor performance in the PSNR and SSIM metrics presented in Table 1 for MSE based convergence and SNR equal to 20 dB. We observed that, when the noise is less critical, the values of the metrics also improve.

Due to the high sensitivity of the LMME descriptor on detecting high-frequency elements of the image, for simulations where noise contamination is less critical (40 dB for example), the proposed methods in this article achieved better results. The same occurs with lower frequency for SNR equal to 20 dB because the noise ends up being less penalized by the regularization term.

In Tables 1, 2 and 3 the PSNR and SSIM values of the experiments are shown. The ones in yellow are the results of the cases where the use of the SARM model improves the performance when comparing to SAR results and also the cases where ℓ_1-SARM produces better results than ℓ_1-SAR. The best values obtained, among all SR methods used in the experiments, are marked in bold in each row of the tables. The TV and TV-SAR methods were included in the experiments to give greater comprehensiveness when comparing the proposed method to other methods present in SR research.

Table 1 shows that both PSNR and SSIM values were improved by the proposed models (SARM and ℓ_1-SARM) when low noise levels were added (SNR $=$ 40 dB). Even though they are not the best among all SR methods, in many

TABLE 2. Comparison of the PSNR and SSIM values (mean ± standard deviation) among the SR methods applied to the Fig. 3 (d). The values in yellow are the cases where the proposed model is better than the original model SAR and l1-SAR. The best values among all the tested methods are in bold.

	TV	l1-norm	SAR	SARM	l1-SAR	l1-SARM	TV-SAR
SNR = 20 dB							
MSE convergence							
PSNR	21.63±1.73E-01	21.60±1.72E-01	21.57±1.86E-01	21.59±1.73E-01	21.70±1.76E-01	21.65±1.74E-01	**21.75±1.85E-01**
SSIM	0.667±9.7E-04	0.665±7.4E-04	**0.706±1.7E-03**	0.665±2.0E-03	0.675±8.2E-04	0.668±9.0E-04	0.697±1.3E-03
SSIM convergence							
PSNR	21.72±1.83E-01	21.78±1.83E-01	21.60±1.76E-01	21.64±1.78E-01	**21.81±1.86E-01**	21.75±1.81E-01	21.74±1.86E-01
SSIM	0.680±3.2E-03	0.686±1.0E-03	0.666±1.6E-03	0.673±1.7E-03	0.694±1.1E-03	0.681±9.6E-04	**0.699±1.3E-03**
SNR = 30 dB							
MSE convergence							
PSNR	25.67±1.75E-01	25.77±1.86E-01	24.47±1.37E-01	25.26±1.34E-01	25.67±1.80E-01	**25.78±1.81E-01**	25.00±1.42E-01
SSIM	0.807±4.2E-03	0.810±4.4E-03	0.784±3.5E-03	0.774±4.3E-03	**0.810±4.3E-03**	0.809±4.4E-03	0.797±3.8E-03
SSIM convergence							
PSNR	25.34±1.48E-01	25.54±1.72E-01	24.63±1.12E-01	24.96±1.25E-01	25.42±1.64E-01	**25.55±1.71E-01**	24.89±1.27E-01
SSIM	0.803±3.9E-03	**0.809±4.2E-03**	0.788±3.1E-03	0.791±3.5E-03	0.807±4.1E-03	0.808±4.2E-03	0.795±3.4E-03
SNR = 40 dB							
MSE convergence							
PSNR	28.72±2.76E-01	28.86±2.69E-01	26.02±1.58E-01	28.17±2.19E-01	28.71±2.61E-01	**28.90±2.63E-01**	27.24±2.49E-01
SSIM	0.891±3.7E-03	**0.895±3.8E-03**	0.838±3.9E-03	0.858±4.2E-03	0.893±4.0E-03	**0.895±3.8E-03**	0.868±5.0E-03
SSIM convergence							
PSNR	26.41±1.55E-01	26.85±1.95E-01	25.14±1.02E-01	25.79±1.26E-01	26.65±1.79E-01	**26.90±1.95E-01**	25.76±1.25E-01
SSIM	0.847±3.3E-03	0.859±3.8E-03	0.812±2.7E-03	0.832±3.2E-03	0.854±3.7E-03	**0.860±3.8E-03**	0.831±3.0E-03

TABLE 3. Comparison of the PSNR and SSIM (mean ± standard deviation) between the SR methods applied to all images of Fig. 3, 5 times each image. The values in yellow are the cases where the proposed model is better than the original SAR and l1-SAR. The best values among all the tested methods are in bold.

	TV	l1-norm	SAR	SARM	l1-SAR	l1-SARM	TV-SAR
SNR = 20 dB							
MSE convergence							
PSNR	21.32±4.94E+00	21.33±5.07E+00	21.19±4.16E+00	21.01±4.35E+00	**21.45±5.04E+00**	21.34±5.05E+00	21.44±4.50E+00
SSIM	0.612±1.34E-01	0.613±1.37E-01	**0.651±1.17E-01**	0.601±1.18E-01	0.625±1.27E-01	0.614±1.36E-01	0.640±1.17E-01
SSIM convergence							
PSNR	23.20±4.47E+00	23.33±4.62E+00	22.94±3.97E+00	22.87±4.12E+00	**23.35±4.50E+00**	23.26±4.62E+00	23.21±4.15E+00
SSIM	0.673±1.38E-01	0.679±1.36E-01	0.667±1.39E-01	0.663±1.35E-01	0.686±1.35E-01	0.677±1.40E-01	**0.694±1.29E-01**
SNR = 30 dB							
MSE convergence							
PSNR	25.64±4.60E+00	25.78±4.60E+00	24.23±4.23E+00	24.61±4.10E+00	**25.79±4.65E+00**	25.78±4.62E+00	24.89±4.40E+00
SSIM	0.813±7.03E-02	**0.820±6.78E-02**	0.784±1.02E-01	0.747±8.45E-02	0.819±7.12E-02	0.816±6.98E-02	0.798±8.83E-02
SSIM convergence							
PSNR	25.37±4.68E+00	**25.63±4.73E+00**	24.13±4.23E+00	24.57±4.35E+00	25.48±4.70E+00	25.57±4.75E+00	24.62±4.41E+00
SSIM	0.809±8.65E-02	**0.818±8.26E-02**	0.786±1.02E-01	0.776±9.90E-02	0.814±8.63E-02	0.814±8.57E-02	0.794±9.98E-02
SNR = 40 dB							
MSE convergence							
PSNR	27.40±5.37E+00	27.55±5.45E+00	25.30±4.59E+00	26.15±4.73E+00	**27.60±5.50E+00**	27.56±5.46E+00	26.62±5.01E+00
SSIM	0.899±3.94E-02	**0.904±4.01E-02**	0.828±1.06E-01	0.831±5.98E-02	0.904±4.16E-02	0.903±4.14E-02	0.872±6.81E-02
SSIM convergence							
PSNR	25.90±5.05E+00	**26.29±5.16E+00**	24.17±4.30E+00	24.84±4.54E+00	26.08±5.06E+00	26.23±5.18E+00	24.90±4.57E+00
SSIM	0.861±8.94E-02	**0.875±8.38E-02**	0.812±1.04E-01	0.824±9.47E-02	0.867±8.61E-02	0.871±8.92E-02	0.831±1.01E-01

cases the performance metric values of the SARM and ℓ_1-SARM models were better than the SAR and ℓ_1-SAR models.

From Table 1 it is also shown that the PSNR and SSIM values are interesting for SNR of 30 dB and 40 dB, mainly with SSIM convergence. The TV and TV-SAR methods stand out for experiments with 20 dB in the PSNR metric compared to the other methods, mainly SAR and SARM. This is due to the fact that very noisy images suffer a lot from penalization when using the SAR model, regarding high frequency elements, while the TV preserves these elements, as can be seen in Fig. 4 (b). However, through visual inspection we observed in the images restored by the TV method greater presence of noise, giving the granular appearance throughout the image.

Fig. 5 shows the graphical representation of PSNR and SSIM versus AWGN to facilitate interpretation of the results presented in Table 1.

As previously mentioned, the metric SSIM was also used as the convergence criterion. The motivation for using another convergence criterion to stop the algorithm is because the proposed modification to the new SAR method formulation occurs on the α hyperparameter model. As explained in [4], the hyperparameter models directly affect the performance of the algorithm.

Table 2 presents results of the experiments performed on the image shown in Fig. 3 (d), where the values of the metrics can be compared. Fig. 3 (d) has fractal features, that is, their

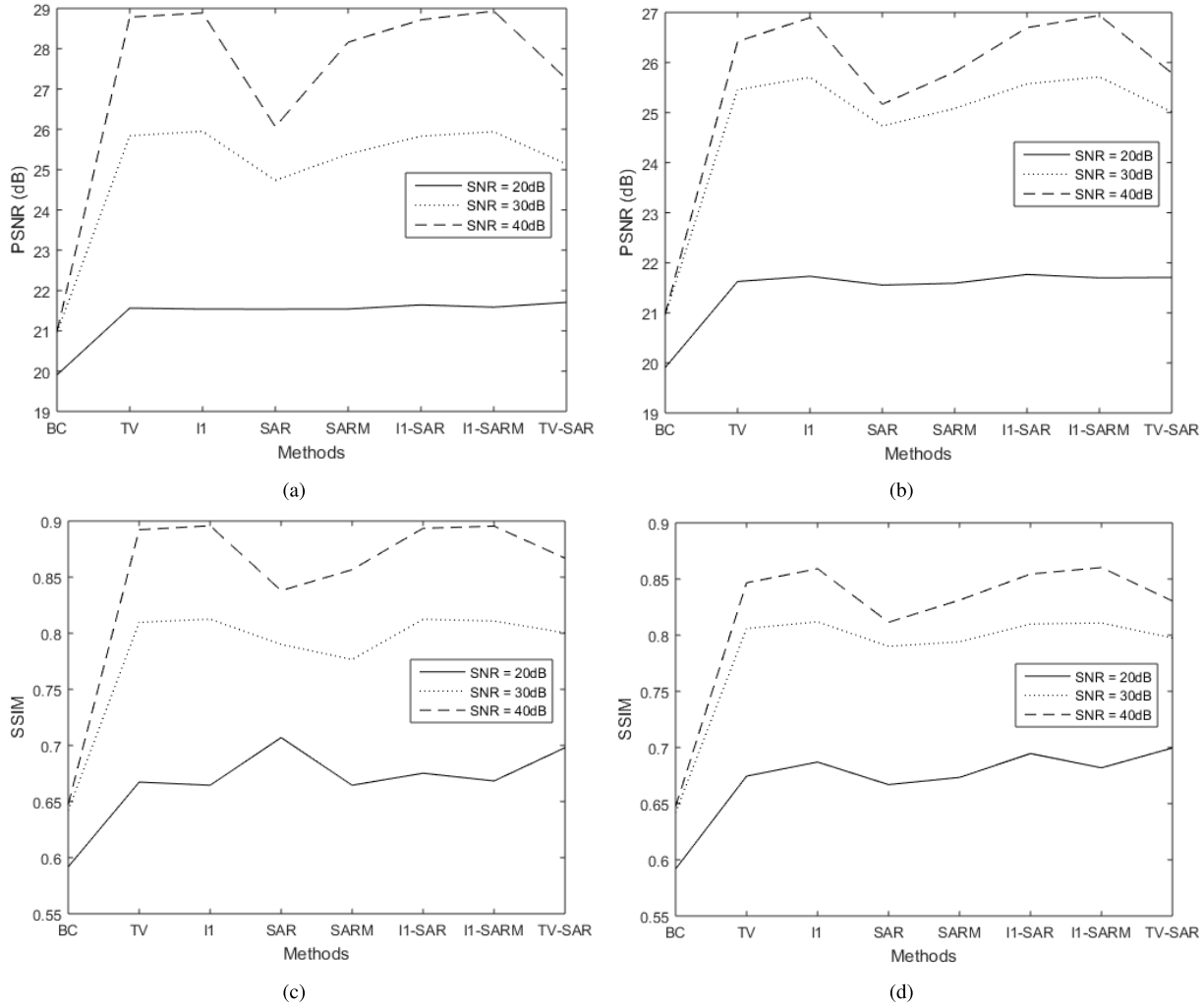


FIGURE 6. Comparing the mean value of PSNR and SSIM metrics versus AWGN (in dB) after performing 5 times the test with Fig. 3 (d). (a) PSNR using MSE convergence criterion, (b) PSNR using SSIM convergence criterion, (c) SSIM using MSE convergence criterion and (d) SSIM using SSIM convergence criterion.

textures are well described by the statistical fractal descriptor, as shown in Fig 2 (b). This explains the good performance of the proposed method in this case. The amount of highlights in yellow shows this in Table 2.

For Fig. 3 (d), the proposed model obtained the best performance of all methods tested in the experiments. This is because the LMME descriptor is sensitive to edge elements.

Table 2 also shows that, for SNR of 30 dB, the application of the proposed model improves the results for Fig. 3 (d). The SSIM convergence makes the results also interesting for SNR = 20 dB, when comparing the SAR and SARM models in this case. But what deserves more attention in this case is that, for no high amount of noise (SNR = 30 dB and 40 dB), the improvement is considerable, making ℓ_1-SARM the best of all methods tested.

Fig. 6 shows the graphical representation of PSNR and SSIM versus AWGN to facilitate interpretation of the results presented in Table 2.

Table 3 contains the results obtained in the experiments performed in all 23 images presented in Fig. 3. In this table, the results were generated by calculating the mean value of the results obtained in each experiment. For convergence based on MSE and SSIM, with noise ranging from 20 dB, 30 dB and 40 dB. In this combined result of all 23 images, we observed that the proposed model (SARM) considerably outperforms the SAR model when the noise decreases to SNR = 40 dB.

Fig. 7 shows the graphical representation of PSNR and SSIM versus AWGN to facilitate interpretation of the results presented in Table 3.

Figs. 8 and 9 show the comparison between the PSNR and SSIM values acquired from the experiments performed on the 23 images presented in Fig. 3. The data used for the graphs were obtained from the experiments performed 5 times for the 23 images, which resulted in 115 samples for each configuration (SNR and metrics). The boxplot graph provides a better

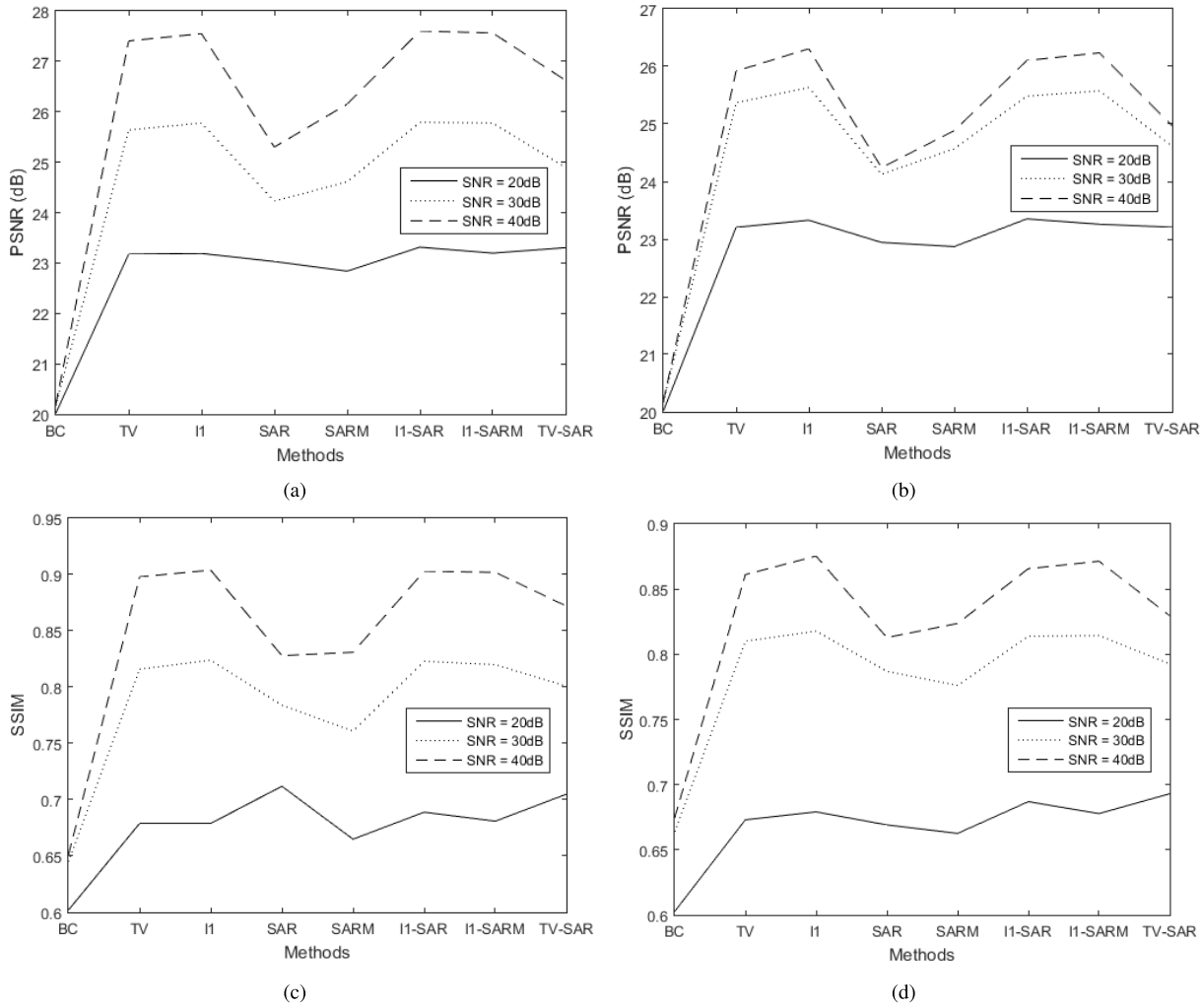


FIGURE 7. **Comparing the mean value of PSNR and SSIM metrics versus AWGN (in dB) performing all images of Fig. 3, 5 times each image. (a) PSNR using MSE convergence criterion, (b) PSNR using SSIM convergence criterion, (c) SSIM using MSE convergence criterion and (d) SSIM using SSIM convergence criterion.**

statistical analysis of the results obtained in the experiments. The boxes in the graphs of Figs. 8 and 9 represent 50% of the most likely values to occur.

Fig. 8 shows values acquired when MSE is used as a convergence criterion. From the graphs shown in Figs. 8 (a), (b) and (c) it can be clearly noticed the performance improvement regarding PSNR metric of the SARM and ℓ_1-SARM methods in relation to the SAR and ℓ_1-SAR methods when the noise level decreases. In Fig. 8 (a), results obtained for SNR $= 20$ dB, the SARM's chart has the median value close to the SAR value, but its boxplot is smalller (results more concentrated around the median).

In Figs. 8 (d), (e) and (f) the values of the SSIM metric can also be analyzed. Comparing SARM and SAR methods, the proposed method does not present better results. Even so, we also see an improvement of performance as the noise decreases.

Fig. 9 presents the boxplot of the experimental results when using SSIM as convergence criterion. In this case, the comparative analysis of PSNR values between the SARM and SAR methods, as well as ℓ_1-SARM and ℓ_1-SAR, are equivalent to those observed for the use of the MSE - based. The Figs. 8 (a), (b) and (c) and the Figs. 9 (a), (b) and (c) show a similar behavior of the values.

For SSIM values using SSIM as convergence criterion, Figs. 9 (d), (e) and (f) show that median and box values of 50% of the most likely values are better for SARM if compared to SAR when the SNR value increases. In case of ℓ_1-SARM this improvement also occurs, but not as intensely.

The use of SSIM as convergence criterion showed that SARM and ℓ_1-SARM obtained better results than when using MSE as convergence criterion. The use of SSIM as stop criterion reduces the number of iterations from the range between 15 and 60 iterations to an average of 5 iterations, according to the experiments performed with the 23 images of Fig. 3.

It is important to mention that the reduction in the number of iterations improves the results of the SARM, showing that the LMME descriptor enhances the hyperparameters

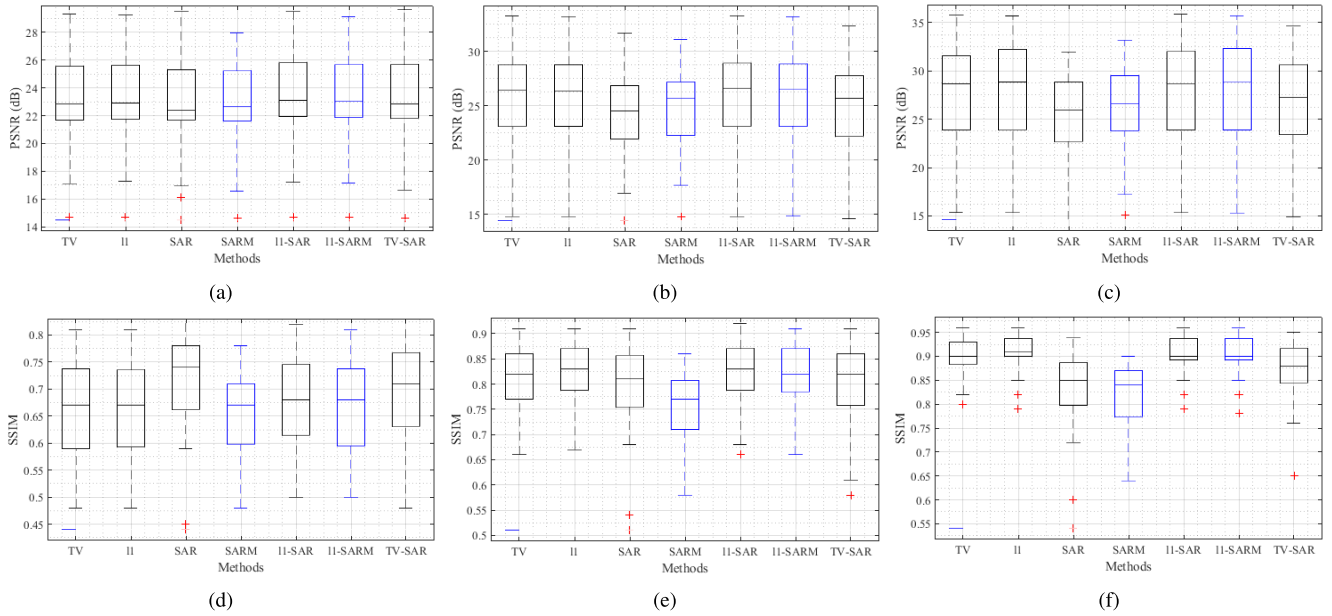


FIGURE 8. Boxplot diagrams showing a comparison between PSNR and SSIM values for different noise levels and MSE convergence criterion. The values in blue are of the proposed methods (SARM e ℓ_1-SARM). The red crosses are the outliers, they were generated mainly by the images of Fig. 3 (q) and (r) due to the accentuated disparity in relation to the other images used in the experiments. (a) SNR = 20 dB. (b) SNR = 30 dB. (c) SNR = 40 dB. (d) SNR = 20 dB. (e) SNR = 30 dB. (f) SNR = 40 dB.

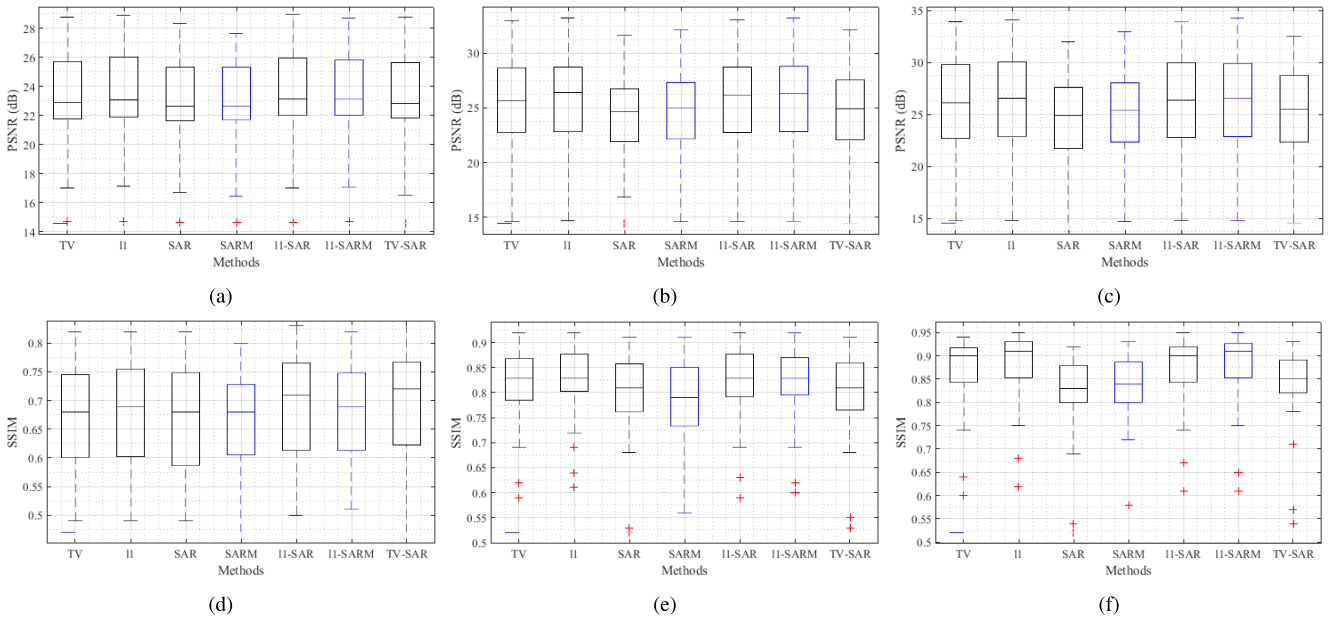


FIGURE 9. Boxplot diagrams showing a comparison between PSNR and SSIM values for different noise levels and SSIM convergence criterion. The values in blue are of the proposed methods (SARM e ℓ_1-SARM). The red crosses are the outliers, they were generated mainly by the images of Fig. 3 (q) and (r) due to the accentuated disparity in relation to the other images used in the experiments. (a) SNR = 20 dB. (b) SNR = 30 dB. (c) SNR = 40 dB. (d) SNR = 20 dB. (e) SNR = 30 dB. (f) SNR = 40 dB.

convergence. This means that modifying the hyperprior model of α has improved the performance of the algorithm.

V. APPLICATIONS

After showing our SR methodology and the results of the experiments performed with a variety of types of images, we will now present some SR applications in the fields of medical diagnosis, surveillance and astronomy. In such contexts, the SR technique can be used as a preprocessing stage for problems related to pattern recognition, since it improves the quality of the image to be treated. The same holds true for segmentation problems, where the image aspects should be more accurately detected.

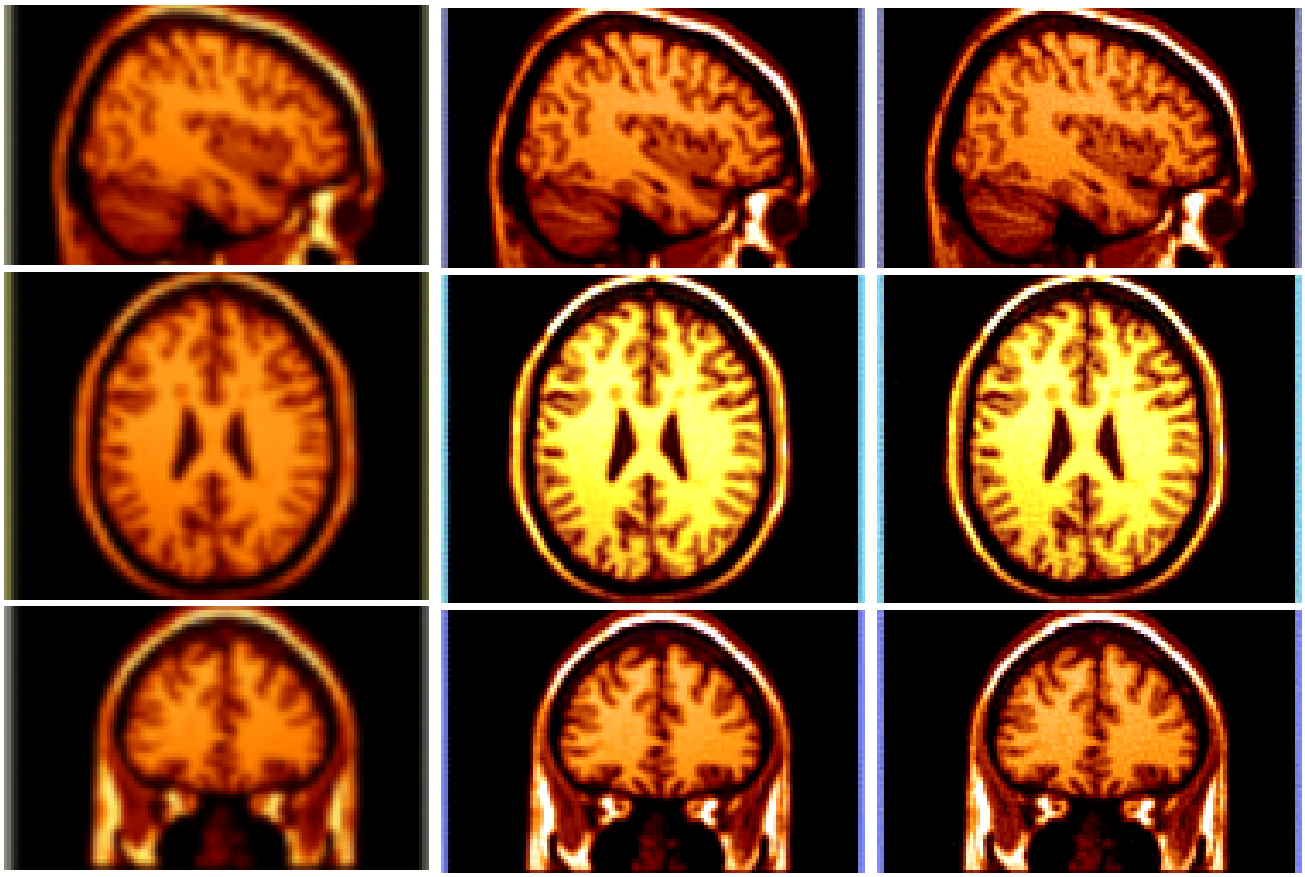


FIGURE 10. SR method applicated in isotropic 3-D medical imaging increasing the resolution to facilitate diagnosis. Comparison of the enhancement of the edge elements of our proposal in relation to the SAR model. LR image adapted from [47]. (a) LR. (b) SAR. (c) SARM.

We will emphasize the comparison between the SAR and SARM methods since the modification of the SAR model is the main focus of this work.

A. MEDICAL DIAGNOSIS

Many types of medical imaging can provide important anatomical information about the structure of the human body and functional information. However, the quality of the images, especially of those that suffer some kind of degradation and therefore offer low resolution, complicates the diagnosis, performed by a specialist.

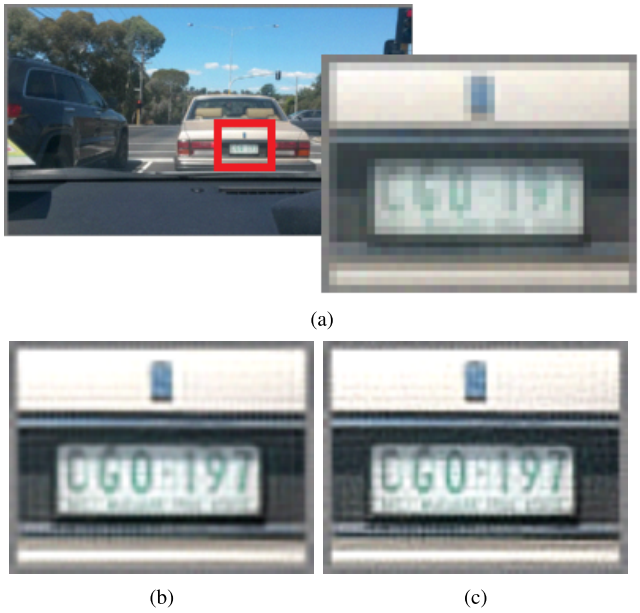
The goal is to increase the resolution of medical images while preserving information from the 3-D input isotropic image, in order to avoid false positive diagnosis. Medical imaging systems can be operated in highly controlled environments and thus a sequence of images which compounds multiple views can be easily acquired [14]. Fig. 10 indicates the results of SR on human brain magnetic resonance data. It is possible to note the sharpness of the images shown in Fig. 10 (c) in relation to Fig. 10 (b), enhancing the preservation of the edge elements of the image. Fig. 10 (a) represents the reference frame of the set of 5 LR images in each dimension.

B. SURVEILLANCE

The multi-frame SR technique can also be applied to license-plate recognition in low-quality surveillance videos, since several video frames can be captured in order to generate the set of LR images. Seibel, Jr., *et al.* [9] use a SR technique as a preprocessing step, enabling character recognition to be made directly by a specialist or by the application of OCR techniques. Fig. 11 shows the application of the SAR and SARM model in increasing the resolution of an LR image to enable recognition of the license plate of a vehicle. In this type of application, the restoration of the image sharpness can be a decisive factor for the correct recognition of the plate characters, especially when performed by humans. From Fig. 11 we can see that in the HR image estimated by the SARM model (Fig. 11 (c)) the characters are sharper than in the HR image estimated by the SAR model (Fig. 11 (b)).

C. APPLICATION FOR RESTORATION OF SCANNED DOCUMENTS

Another application in which the edge enhancement proves to be interesting is for restoration of scanned documents. In Fig. 12 the scanned image of a newspaper was captured 5 times and the region of interest demarcated for execution of



(a) (b) (c)

FIGURE 11. License-plate recognition. LR image adapted from [45]. (a) LR image. (b) SAR. (c) SARM.

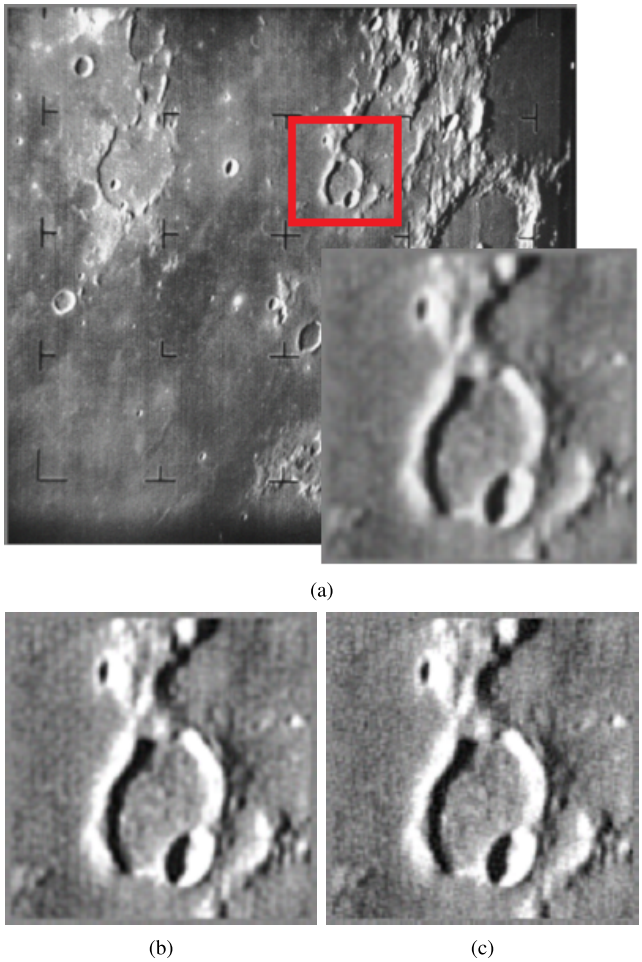
(a) (b) (c)

FIGURE 12. SR method applicated for restoration of information in newspaper texts. Comparison of the enhancement of the edge elements of our proposal in relation to the SAR model. LR image adapted from [46]. (a) LR image. (b) SAR. (c) SARM.

the SR methods. In Fig. 12 (c), the word in the phrase ''Voice from Moon: 'Eagle has Landed'''' are sharper, while in the image of Fig. 12 (b) the text appears more blurry.

D. ASTRONOMY

The application of SR techniques in astronomy images is also mentioned in [14]. In this type of problem increasing the resolution has the objective of improving the sharpness of the image, facilitating the identification of elements of a celestial body under observation.



(a) (b) (c)

FIGURE 13. SR method applicated for restoration of satellite images increasing the resolution. Comparison of the enhancement of the edge elements of our proposal in relation to the SAR model. LR image adapted from [48]. (a) LR image. (b) SAR. (c) SARM.

Fig. 13 presents the application of the SAR and SARM methods in the first image of the Moon's surface provided by the satellite Spacecraft [48], from 4 LR images. Fig. 13 (a) is the reference LR frame. In this case, the increase in sharpness observed in Fig 13 (c) allows better visualization of the contours of the elements on the Moon's surface, while in the image restored by the SAR model, Fig. 13 (b), has some blurry regions.

VI. CONCLUSION

The Bayesian SAR super-resolution model has the characteristic of smoothing the edge elements of the estimated HR images. This work proposes a modification in this model including the LMME fractal descriptor in its modeling in order to improve the preservation of these edge elements, so that the estimated HR images are richer in detail.

The data obtained from the experiments allow to observe that for both SARM and ℓ_1-SARM the regularization term

penalized less the edge elements of the estimated image, as can be seen in Fig. 3 (b), for example.

This means that the proposed method emphasizes the fidelity term by performing a better fit between the values of the hyperparmeters α and β_k. As consequence, the estimated image is sharper, when analysing it visually.

Experiments have shown that the proposed SARM and ℓ_1-SARM methods present better results for the SR process when the input LR images have less additive noise. For the experiments where SNR was 30 dB or 40 dB, we observed that the SARM method is better, on average, than the SAR method used for the simulations. Another relevant result is related to the convergence of the SR algorithm. The use of SSIM as a stopping criterion reduces the number of iterations from the interval between 15 and 60 iterations to an average of 5 iterations, for all the models used in this work.

Using the MSE as a stop criterion in the SR algorithm, the number of iterations is reduced, approximately half, using the SARM model, when compared to the experiments performed with the SAR model.

As future work we intend to use the LMME descriptor for the hyperparameters estimation of models based on standard ℓ_1 norm, such as ℓ_1-norm. A hybrid model using the SAR model and SARM can also be implemented in another work, where in situations of greater noise contamination the SAR model is used and in cases of moderate noise the SARM model is used.

REFERENCES

[1] A. K. Katsaggelos, R. Molina, and J. Mateos, *Super Resolution of Images and Video*. San Rafael, CA, USA: Morgan & Claypool, 2007.

[2] R. Molina, J. Nunez, F. J. Cortijo, and J. Mateos, "Image restoration in astronomy: A Bayesian perspective," *IEEE Signal Process. Mag.*, vol. 18, no. 2, pp. 11–29, Mar. 2001.

[3] S. Villena, M. Vega, D. Babacan, R. Molina, and A. Katsaggelos, "Using the Kullback-Leibler divergence to combine image priors in super-resolution image reconstruction," in *Proc. IEEE Int. Conf. Image Process.*, Hong Kong, Sep. 2010, pp. 893–896.

[4] S. Villena, M. Vega, S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian combination of sparse and non-sparse priors in image super resolution," *Digit. Signal Process.*, vol. 23, no. 2, pp. 530–541, 2013.

[5] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian super resolution," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 984–999, Apr. 2011.

[6] G. Chantas, N. P. Galatsanos, R. Molina, and A. K. Katsaggelos, "Variational Bayesian image restoration with a product of spatially weighted total variation image priors," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 351–362, Feb. 2010.

[7] E. Turgay and G. B. Akar, "Texture and edge preserving multiframe super-resolution," *IET Image Process.*, vol. 8, no. 9, pp. 499–508, Sep. 2014.

[8] S. Huang, J. Sun, Y. Yang, Y. Fang, and P. Lin, "Multi-frame super-resolution reconstruction based on gradient vector flow hybrid field," *IEEE Access*, vol. 5, pp. 21669–21683, 2017.

[9] H. Seibel, Jr., S. Goldenstein, and A. Rocha, "Eyes on the target: Super-resolution and license-plate recognition in low-quality surveillance videos," *IEEE Access*, vol. 5, pp. 20020–20035, 2017.

[10] M. Sonogashira, T. Funatomi, M. Iiyama, and M. Minoh, "Variational Bayesian approach to multiframe image restoration," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2163–2178, May 2017.

[11] S. Zhao, H. Liang, and M. Sarem, "A generalized detail-preserving super-resolution method," *Signal Process.*, vol. 120, pp. 156–173, Mar. 2016.

[12] P. Cheng, Y. Qiu, X. Wang, and K. Zhao, "A new single image super-resolution method based on the infinite mixture model," *IEEE Access*, vol. 5, pp. 2228–2240, 2017.

[13] J. Jiang, X. Ma, C. Chen, T. Lu, Z. Wang, and J. Ma, "Single image super-resolution via locally regularized anchored neighborhood regression and nonlocal means," *IEEE Trans. Multimedia*, vol. 19, no. 1, pp. 15–26, Jan. 2017.

[14] L. Yue, H. Shen, J. Li, Q. Yuan, H. Zhang, and L. Zhang, "Image super-resolution: The techniques, applications, and future," *Signal Process.*, vol. 128, pp. 389–408, Nov. 2016.

[15] R. Timofte, V. De Smet, and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1920–1927.

[16] T. Sigitani, Y. Iiguni, and H. Maeda, "Image interpolation for progressive transmission by using radial basis function networks," *IEEE Trans. Neural Netw.*, vol. 10, no. 2, pp. 381–390, Mar. 1999.

[17] J. Sun, Z. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[18] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, "Gradient profile prior and its applications in image super-resolution and enhancement," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1529–1542, Jun. 2011.

[19] R. Chao, X. He, and T. Q. Nguyen, "Single image super-resolution via adaptive high-dimensional non-local total variation and adaptive geometric feature," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 90–106, Jan. 2016.

[20] Y.-W. Tai, S. Liu, M. S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2400–2407.

[21] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 2272–2279.

[22] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[23] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[24] H. Zhang, L. Zhang, and H. Shen, "A blind super-resolution reconstruction method considering image registration errors," *Int. J. Fuzzy Syst.*, vol. 17, no. 2, pp. 353–364, 2015.

[25] M. K. Ng, H. Shen, E. Y. Lam, and L. Zhang, "A total variation regularization based super-resolution reconstruction algorithm for digital video," *EURASIP J. Adv. Signal Process.*, vol. 2007, pp. 74585-1–74585-16, Dec. 2007.

[26] C. Chen, H. Liang, S. Zhao, Z. Lyu, S. Fang, and X. Pei, "Integrating the missing information estimation into multi-frame super-resolution," *Circuits, Syst., Signal Process.*, vol. 35, no. 4, pp. 1213–1238, Apr. 2016.

[27] Y. Zhang, J. Liu, W. Yang, and Z. Guo, "Image super-resolution based on structure-modulated sparse representation," *IEEE Trans. Hum.-Mach. Syst.*, vol. 24, no. 9, pp. 2797–2810, Sep. 2015.

[28] X. Zhang, J. Jiang, and S. Peng, "Commutability of blur and affine warping in super-resolution with application to joint estimation of triple-coupled variables," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1796–1808, Apr. 2012.

[29] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2008.

[30] M. Woods and A. Katsaggelos, "A Bayesian multi-frame image super-resolution algorithm using the Gaussian information filter," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 2017, pp. 1368–1372.

[31] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1621–1633, Dec. 1997.

[32] C. Deng, J. Xu, K. Zhang, D. Tao, X. Gao, and X. Li, "Similarity constraints-based structured output regression machine: An approach to image super-resolution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 12, pp. 2472–2485, Dec. 2016.

[33] K. F. Côco, E. O. T. Salles, and M. Sarcinelli-Filho, "Topographic independent component analysis based on fractal theory and morphology applied to texture segmentation," *Signal Process.*, vol. 87, no. 8, pp. 1966–1977, 2007.

[34] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.

[35] M. O. Camponez, E. O. T. Salles, and M. Sarcinelli-Filho, "Super-resolution image reconstruction using nonparametric Bayesian INLA approximation," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3491–3501, Aug. 2012.
[36] J. Idier, *Bayesian Approach to Inverse Problems*. Hoboken, NJ, USA: Wiley, 2008.
[37] AbbrA. *Album of Nature—Butterfly*. Accessed: Jan. 3, 2018. [Online]. Available: http://www.abbra.com.br/borboleta2.jpg
[38] *Brodatz Textures*. Accessed: Jan. 3, 2018. [Online]. Available: http://www.ux.uis.no/~tranden/brodatz.html
[39] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
[40] N. Sarkar and B. B. Chaudhuri, "An efficient differential box-counting approach to compute fractal dimension of image," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 1, pp. 115–120, Jan. 1994.
[41] M. J. Turner, J. M. Blackledge, and P. R. Andrews, *Fractal Geometry in Digital Imaging*. San Diego, CA, USA: Academic, 1998.
[42] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[43] Y. Xia, D. D. Feng, and R. Zhao, "Morphology-based multifractal estimation for texture segmentation," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 614–623, Mar. 2006.
[44] X. Zhou, R. Molina, F. Zhou, and A. K. Katsaggelos, "Fast iteratively reweighted least squares for lp regularized image deconvolution and reconstruction," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 1783–1787.
[45] Imgur. (1985). *Rarest Victorian License Plate Slogan 'Victoria—Nuclear Free State'*. Accessed: Mar. 1, 2018. [Online]. Available: https://imgur.com/gallery/eur9S
[46] N. D. Sanders. *Lot #153: Neil Armstrong Signed 21 July 1969 New York Times Newspaper—Men Walk on the Moon*. Accessed: Mar. 1, 2018. [Online]. Available: http://natedsanders.com/neil-armstrong-signed-21-july-1969-new-york-time-lot14613.aspx
[47] BrainWeb. *BrainWeb: Simulated Brain Database*. Accessed: Mar. 1, 2018. [Online]. Available: http://brainweb.bic.mni.mcgill.ca/brainweb/
[48] NASA. *First Image of the Moon Taken by a U.S. Spacecraft*. Accessed: Mar. 1, 2018. [Online]. Available: https://www.nasa.gov

VITOR F. CAMPANA received the degree in computer engineering and the master's degree in informatics from the Federal University of Espírito Santo in 2006 and 2009, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering in image processing. He is currently a Professor of basic, technical, and technological education with the Federal Institute of Education of Espírito Santo.

KLAUS F. CÔCO graduated in computer engineering from the Federal University of Espírito Santo (UFES) in 1998. He received the master's degree and the Ph.D. degree in electrical engineering from UFES in 2003 and 2007, respectively, with a focus on signal processing/images and pattern recognition. It mainly works on the following topics: machine intelligence, digital signal/image processing, segmentation, and identification of patterns and systems. He is currently developing research in the area of identification of patterns in stochastic time series signals. He is currently an Adjunct Professor IV with the Department of Electrical Engineering, UFES.

EVANDRO OTTONI TEATINI SALLES received the B.Sc. and master's degrees in electrical engineering from the Federal University of Espírito Santo, Vitória, Brazil, in 1987 and 1994, respectively, and the Ph.D. degree in electrical engineering from the State University of Campinas, Campinas, Brazil, in 2001. He is currently an Associate Professor with the Department of Electrical Engineering, Federal University of Espírito Santo. His current research interests include digital signal and image processing, pattern recognition, and video processing.

PATRICK M. CIARELLI graduated in electrical engineering from the Federal University of Espírito Santo in 2006, He received the master's and Ph.D. degrees in electrical engineering from the Federal University of Espírito Santo in 2008 and 2012, respectively. He is currently an Assistant Professor I with the Federal University of Espírito Santo. He has experience in the area of electrical engineering, with an emphasis on artificial intelligence, information retrieval, pattern recognition, and image processing.