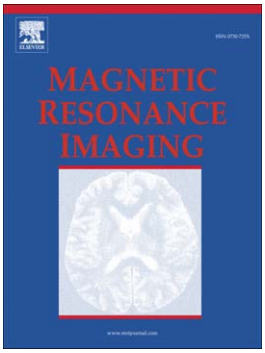


A Sparse Bayesian Representation for Super-Resolution of Cardiac MR Images

Nelson F. Velasco, Andrea Rueda, Cristina Santa Marta, Eduardo Romero

PII: S0730-725X(16)30156-4
DOI: doi: 10.1016/j.mri.2016.10.009
Reference: MRI 8631

To appear in: *Magnetic Resonance Imaging*

Received date: 21 April 2015
Revised date: 19 September 2016
Accepted date: 5 October 2016

Please cite this article as: Velasco Nelson F., Rueda Andrea, Marta Cristina Santa, Romero Eduardo, A Sparse Bayesian Representation for Super-Resolution of Cardiac MR Images, *Magnetic Resonance Imaging* (2016), doi: 10.1016/j.mri.2016.10.009

A Sparse Bayesian Representation for Super-Resolution of Cardiac MR Images

Tue Oct 11 11:34:15 2016

Nelson F. Velasco[a,b]; Andrea Rueda[a,c]; Cristina Santa Marta[d]; Eduardo Romero[a,1]

[a]Computer Imaging and Medical Applications Laboratory – CIM@LAB, Universidad Nacional de Colombia, Bogotá, Colombia

[b]Universidad Militar Nueva Granada, Bogotá, Colombia

[c]Departamento de Ingeniería de Sistemas, Pontificia Universidad Javeriana, Bogotá, Colombia

[d]Departamento de Física Matemática y de Fluidos, Universidad Nacional de Educación a Distancia, Madrid, Spain

Abstract

High-quality Cardiac Magnetic Resonance (CMR) images can be hardly obtained when intrinsic noise sources are present, namely heart and breathing movements. Yet heart images may be acquired in real time, the image quality is really limited and most sequences use ECG gating to capture images at each stage of the cardiac cycle during several heart beats. This paper presents a novel super-resolution algorithm that improves the cardiac image quality using a sparse Bayesian approach. The high-resolution version of the cardiac image is constructed by combining the information of the low-resolution series –observations from different non-orthogonal series composed of anisotropic voxels– with a prior distribution of the high-resolution local coefficients that enforces sparsity. In addition, a global prior, extracted from the observed data, regularizes the solution. Quantitative and qualitative validations were performed in synthetic and real images w.r.t to a baseline, showing an average increment between 2.8 and $3.2\ dB$ in the Peak Signal-to-Noise Ratio (PSNR), between 1.8% and 2.6% in the Structural Similarity Index (SSIM) and $2\ \%$ to $4\ \%$ in quality assessment (IL-NIQE). The obtained results demonstrated that the proposed method is able to accurately reconstruct a cardiac image, recovering the original shape with less artifacts and low noise.

Keywords: Magnetic resonance; Super-resolution; Sparse representation

1 Introduction

Magnetic Resonance Imaging (MRI) has been broadly applied in diagnosis, surgical planning and research of several structural or functional heart disorders. Usually, the limited MRI acquisition time results in highly anisotropic voxels, i.e., 1-2 mm in-plane resolution up to 8 mm inter-plane separation, a distance that may be even larger when both heart movements and apnea

[1] Corresponding author
Email address:
edromero@unal.edu.co

periods are exacerbated. In practice, most cardiac MRI sequences extract slices in arbitrary directions, basically three non orthogonal acquisition planes: 4-chamber long axis, 2-chamber long axis and 2-chamber short axis. In general, these standard views are enough to examine the heart structure, but at a high acquisition time cost. Accordingly, less time demanding methods have been developed, either as protocols that radially acquire images or customized sequences that capture more rapidly a volume. Nevertheless, these techniques produce highly anisotropic voxels and hence very limited image quality [1].

Several advances in cardiovascular Magnetic Resonance (CMR) have brought out this technique as the imaging gold standard for many cardiac diseases [2], [3]. However, for the CMR becomes integrated with the everyday clinical practice, it is crucial the improvement of the image quality and the introduction of supplementary education for the specialists. These imaging educational programs require reliable 3D reconstructions that facilitate efficient data interaction. It might be also useful for patients to be informed about some procedures and in this case the data visualization quality is important [4], [5]. Additionally, a higher resolution image may facilitate post-processing and/or segmentation so that accurate volume calculation of the cardiac chambers or valves may be possible. A great challenge of such reconstruction is that most acquisition techniques result in very large anisotropic voxels and therefore very blurred versions of the 3D object. Essentially, this problem is equivalent to a non-uniform sampling in a regular grid. Traditional interpolation methods fail since they overfit a curve or a surface model that by no means reconstructs discontinuities of some structures such as the interventricular septum or the valve motion.

In general, super-resolution (SR) methods reconstruct a high-resolution image from low-resolution images that have been captured with small displacements between them [6], [7], [8], [9]. Typical SR methods will surely fail in case of CMR images since acquisition planes are non-orthogonal and the voxel anisotropy is much larger, i.e., between 5*mm* and 8*mm* slice thickness. Likewise, heart and chest motions highly increase the uncertainty of any capture.

This paper presents a novel super-resolution method that reconstructs a high resolution version of the CMR image, even in case of non-orthogonal anisotropic voxels and noisy captures. The method is based on a Bayesian representation which may deal with the uncertainty from multiple noise sources, i.e., the non-orthogonal capturing procedure that introduces different degrees of noise from each of the orientations. The acquisition noise is not modeled but approximated by a Bayesian representation of the whole volume. A sparsity hypothesis forces the method to have a minimal neighborhood, a local vicinity where relations are assumed Gaussian and the noise sources independent. In such a case, this local information is driven by a Gibbs probability distribution which in addition is continous. The approach was widely evaluated in both phantom and real data, demonstrating a substantial gain with respect to the base line.

This article is divided into four sections: next section presents a brief comparison with related works. Section 3 shows the general outline of the proposed method, while Section 4 describes the experiments performed to evaluate the accuracy of the proposal and the quality of the obtained images. Finally Section 5 discusses the results and presents some conclusions.

2 Related works

Greenspan et al. [10] introduced the use of SR algorithms in MRI images, using an iterative method that minimizes the mean square distance between the given low-resolution image and low-resolution versions of the high-resolution image. This high-resolution image is obtained by

simulating the imaging process as a geometric transformation of the high-resolution image, which is blurred, downsampled and perturbed by a Gaussian noise. Results were evaluated in a 3D isotropic grid with 4.5 mm inter-plane and 1.5 mm in-plane resolutions, respectively, reporting an improvement of 3% in the Signal-to-Noise Ratio (SNR). This resolution nevertheless is far from what is observed in CMR, typically a slice thickness of 8 *mm*. Plenge et al. [11] demonstrated the advantages of the SR methods to improve the MRI image quality and reduce the acquisition time. This investigation compared six SR algorithms, grouped into three different approaches: iterative back-projection, algebraic reconstruction and regularized least squares, concluding that the SR algorithms remarkably improve the SNR.

Other type of approximations have used different capture protocols that improve the image quality. Peeters et al. [12] acquired shifted versions of brain functional images, whereby simple interpolations would increase the data quality, but at a higher acquisition time. From the obtained results, these authors claim that the detectability of small activated areas is increased and the SNR improved. Poot et al. [13] presented a multi-slice MRI SR method, reconstructing the high-resolution image as a squared minimum problem, regularized by the second order derivatives in three non-orthogonal directions of the estimated image. Experiments were performed using MRI images of brain birds, but the presented results were purely qualitative.

Bai et al. [9] increased the quality of brain MRI images by applying a Bayesian approach to exclusively orthogonal series. In such approach, the local prior is a Markov Random Field approximated as a discretization of a Laplacian operator within a fully regular grid. Rahman et al. [14] estimated a SR image by computing a Maximum a Posteriori (MAP) from a set of orthogonal series of CMR images, but only orthogonal views are considered. The prior term was also a Markov Random field which basically defined conditional independence between the members of the local neighborhood, a statement hardly applicable to images with very similar textures. In addition, their evaluation was purely qualitative on real images, with no numerical or quantitative measures. Gholipur et al. [15] estimated the high-resolution image by maximizing a likelihood term, defined as the difference between the observations and a pseudo-inverse of the generative model of the image. The method was validated with phantoms and brain imaged newborn subjects, reporting a Peak Signal-to-Noise Ratio (PSNR) improvement of 2dB to 3dB in images with a slice thickness of 6 *mm* for phantoms and 3 to 4 *mm* for newborn subjects. This method however has not prior at all.

3 Methods

The proposed method takes as input N low-resolution images, corresponding to series of 2D slices acquired at non-orthogonal and different orientations, and combine them into a single high-resolution volume. The method is divided into two steps, the first or pre-processing aims to correct or eliminate differences in the grayscale intensity and to spatially transform the images to a common reference system. The second step is the super-resolution method, formulated as a regression problem in terms of a set of bases or regressors whose weights are obtained from a Bayesian approach. The method looks for the sparsest solution of a particular prediction at both the local level, determined by the capture procedure, and the global information that ensures fidelity to the particular heart shape. Figure (1) shows the method overview.

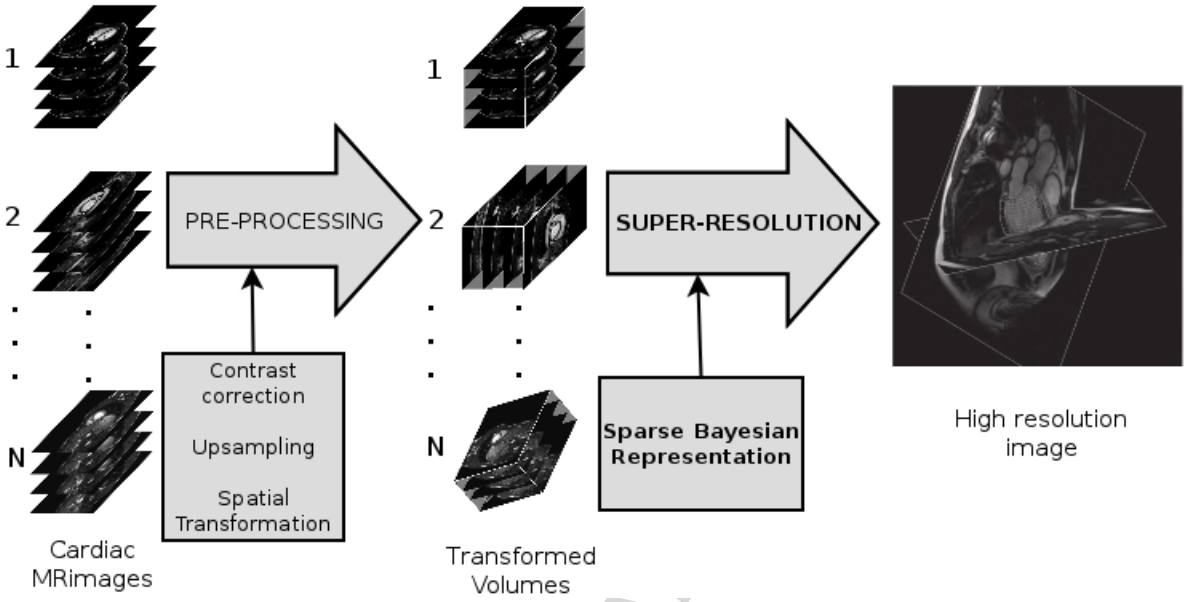


Figure 1: Flowchart of the proposed method

3.1 Image pre-processing

The purpose of this stage is twofold, 1) to homogenize the grayscale low-resolution images and 2) to obtain a common spatial representation. In the present investigation, only those series that contained the whole organ were selected and three different operations were performed on them: a contrast correction, an upsampling (interpolation) and a spatial transformation.

3.1.1 Contrast correction

The image contrast correction is a non-parametric intensity transformation using a reference image. Following the method described in [16], the reference image used in this study is that one with the larger dynamic range, and a mapping transformation is performed by matching the histogram bin coincidences, as follows

$$p = \max_k [T_{m-1} < H_k \leq T_m] \quad (1)$$

Where p is the gray level to be placed in the $k-th$ bin of the histogram to be modified, T_m represents the cumulative histogram of the image to be modified, H_k is the cumulative histogram of the reference image and the subscripts k and m are the gray levels of the reference and modified images, respectively.

3.1.2 Upsampling

As a second step, a high-resolution image is constructed by combining high-resolution versions of the non-orthogonal low-resolution series. Instead of dividing the contribution of the low-resolution Point Spread Function (PSF) between the corresponding four high-resolution

pixels, as usual, an isotropic image is constructed by linearly combining non-orthogonal series that previously have been converted to isotropic. For doing so, either a statistical filtering[2] or an interpolation (linear and bi-cubic) is firstly carried out for each of the anisotropic series, converting the series into isotropic.

3.1.3 Spatial mapping

Finally, to bring all images of each series to a single reference system, the DICOM header information is used to extract the rotation and translation matrices of the mapping, locating all series in the origin of the spatial reference system of the machine. As stated in [18], a 2D image (a slice) is represented by assigning a null vector to the third column of the transformation matrix. In the present investigation, the third column corresponds to the vector $\left[Z_x Z_y Z_z\right]^T \Delta_z$, where Δ_z is the height in millimeters of the low-resolution voxels and $\left[Z_x Z_y Z_z\right]^T = \left[X_x X_y X_z\right]^T \times \left[Y_x Y_y Y_z\right]^T$, and $\left[X_x X_y X_z\right]^T$ and $\left[Y_x Y_y Y_z\right]^T$ are the row direction cosine values (X and Y respectively) of the patient orientation and $\left[P_x P_y P_z\right]^T$ are the voxel coordinates, calculated as

$$\begin{bmatrix} P_x \\ P_y \\ P_z \\ 1 \end{bmatrix} = \begin{bmatrix} X_x\Delta_x & Y_x\Delta_y & Z_x\Delta_z & S_x \\ X_y\Delta_x & Y_y\Delta_y & Z_y\Delta_z & S_y \\ X_z\Delta_x & Y_z\Delta_y & Z_z\Delta_z & S_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i \\ j \\ k \\ 1 \end{bmatrix} \tag{2}$$

Where Δ_x, Δ_y and Δ_z correspond to the voxel size at each dimension, S_x, S_y and S_z are the patient location w.r.t. the scanner, $[ijk]^T$ is a voxel position in the volume before transformation.

3.2 The Super-Resolution algorithm

Given a data set of the form $\{\mathbf{x}_n, \mathbf{I}_n\}_{n=1}^N$, where $\mathbf{x}$ is a vector of known inputs, $\mathbf{I}$ is a desired output or target and N is the number of corresponding pairs of $\mathbf{I}$ and $\mathbf{x}$, a regression model is capable of making accurate predictions of $\mathbf{I}$ for new examples of $\mathbf{x}$. In case of CMR, the observed samples $\{\mathbf{I}_l\}$ stand for the low-resolution images, from which a high-resolution version will be obtained. In the SR context, $\mathbf{x}$ is associated with locations in a sampling grid, where $\mathbf{x}_l$ corresponds to the acquisition spatial locations (low-resolution grid) and $\mathbf{x}_h$ stands for the high-resolution grid ($\mathbf{x}_l \subset \mathbf{x}_h$). Given $\{\mathbf{x}_l, \mathbf{I}_l\}$ and $\{\mathbf{x}_h\}$, the SR reconstruction can be written as

$$I = f(\mathbf{x}) + \varepsilon \tag{3}$$

Where ε is an additive noise from sensors and/or acquisition devices and f is a continuous function representing the real object. The SR problem can then be stated as an optimal search of continuous information from a discrete non-regular grid of samples, a classical regression problem. Under such regression formulation, an approximation of the continuous function f can be stated as

[2] The statistical filtering is a 3D variant of the Non-Local Means filter which has been classically used as a a noise reduction strategy, under the constraint that the downsampled version of the reconstructed image must be as similar as possible to the low-resolution image [17].

$$\hat{f} = \sum_{j=1}^{J} I_{h,j}\delta(\mathbf{x}_{h,j}) = \mathbf{w}\cdot\mathbf{\Phi} \tag{4}$$

Where, $\hat{f}$ is a discrete version of f in the high-resolution locations $(\mathbf{x}_h)$, $\mathbf{w}$ is a vector of weights representing the intensities to be learned and $\mathbf{\Phi}$ is a discrete basis that exists only for the spatial locations $(\mathbf{x}_h)$.

3.2.1 The Sparse Bayesian approach

Provided that a main difficulty of this SR reconstruction is the uncertainty of the capturing grid, a sparse Bayesian framework was herein adapted to deal with it. From a sparse point of view, the SR problem can be stated as the estimation of a minimal number of weights ω_i such that the regression $\omega_i\phi_i$ approximates the observation $\mathbf{I}_j$,.i.e., $\mathbf{I} \approx \Phi^T\mathbf{w}$ with $\mathbf{I} \equiv \left[I_1, \dots, I_j\right]^T$ and $\mathbf{\Phi} \equiv \left[\phi_1, \dots, \phi_j\right]^T$.

Unlike previous approaches [13,14,15], the presented method imposes the sparsity condition by assuming the weights are set within a pre-defined neighborhood, trying to maximize the covariance between weights that follow a Gibbs distribution. The prior probability forces local information to highly depend on the central voxel, provided $\mathbf{\Phi}$ is square and built from a positive kernel function.

The regression bases correspond then to unitary Kronecker pulses placed over the predefined high-resolution grid and the weights correspond to a kernel that models the PSF as a Gibbs distribution. This kernel modulates the likelihood function, approximated herein by a Gaussian function of the difference between the observations and the linear model.

Given a regressor basis $\mathbf{\Phi}$, the full Bayesian treatment of (3) leads to find the predictive distribution $p(\mathbf{w}|\mathbf{I}_l, \mathbf{\Phi})$, which typically is calculated as a product between two probabilities (the likelihood and prior terms):

$$p(\mathbf{w}|\mathbf{I}_l, \mathbf{\Phi}) = p(\mathbf{I}_l|\mathbf{w}, \mathbf{\Phi})p(\mathbf{w}, \mathbf{\Phi}) \tag{5}$$

The *Likelihood term* ($p(\mathbf{I}_l|\mathbf{w}, \mathbf{\Phi})$) is a measure of the ability to generate the observed data $\mathbf{I}_l$ from possible regressors $\mathbf{\Phi}$. Assuming Gaussian noise, it follows

$$\varepsilon = \mathbf{I}_l - \mathbf{w}\cdot\mathbf{\Phi} \tag{6}$$

and

$$p(\mathbf{I}_l|\mathbf{w}, \mathbf{\Phi}) = \frac{1}{(2\pi)^{\frac{M}{2}}\sigma^M}\exp\left\{-\frac{1}{2\sigma^2}(\mathbf{I}_l - g_l(\mathbf{w}\cdot\mathbf{\Phi}))^T(\mathbf{I}_l - g_l(\mathbf{w}\cdot\mathbf{\Phi}))\right\} \tag{7}$$

Where σ^2 represents the image noise variance and g_l is the subsampling and blurring effect that maps the high-resolution to the low-resolution image.

On the other hand, the *Prior term* is estimated by assuming that a sparse solution should neglect nearly every basis except the one belonging to a pre-defined neighborhood, determined by the acquisition procedure. A prior model, following a simple Gibbs distribution, has the advantage of providing a convex cost function with a global minimum, which is also easily differentiable with respect to the estimated $\mathbf{w}$. This kind of model can be written as

$$p(\mathbf{w}|\mathbf{\Phi}) = \frac{1}{(2\pi)^M|\mathbf{k}|^{12}} \exp\left\{-\frac{1}{2}(\mathbf{w}\cdot\mathbf{\Phi})^T\mathbf{k}^T(\mathbf{w}\cdot\mathbf{\Phi})\right\} \quad (8)$$

Where $\mathbf{k}$ is the $M \times M$ kernel to be learned. The exponential term can be factored as a sum of products, where $\mathbf{k} = \mathbf{k}_i\mathbf{k}_i^T$ is the searched kernel, $\mathbf{k}_i = \left[k_{i,1}, k_{i,2}, \ldots, k_{i,M}\right]^T$ is a vector of coefficients. The vectors of coefficients $\mathbf{k}_i$ are the set of prior assumptions about the local relationships between the values of $\mathbf{w}$. This kernel was learned using a Gibbs distribution and a pre-defined neighborhood as

$$\hat{\mathbf{k}} = * \, arg\, max_{\mathbf{k}} p(\mathbf{I}_l|\mathbf{k}) = * \, arg\, max_{\mathbf{k}} \int p(\mathbf{I}_h, \mathbf{I}_l|\mathbf{k})\, \partial\mathbf{I}_h \quad (9)$$

This Maximum a Posteriori (MAP) formulation can be approached by an Expectation-Maximization (EM) framework that includes the high-resolution images $\mathbf{I}_h$ as a hidden variable and onto which the problem is marginalized and treated as an usual EM problem. The E-step estimates a high-resolution image and its covariance using the current kernel, and the M-step determines the best kernel using the high-resolution image from the E-step. The algorithm reads as follows:

1. E-step: considering $p(\mathbf{I}_h) = p(\mathbf{I}_h|\mathbf{I}_l, \mathbf{k})$, an averaged estimation and covariance of the high-resolution image, $\mathbf{I}_h$ is computed
2. M-step: $\mathbf{k}$ is obtained by minimizing $E_{\mathbf{I}_h}(||\mathbf{k}\mathbf{I}_h - \mathbf{I}_\mathbf{l}||^2)$

These local relationships were herein well approximated with the Kronecker function as the canonical base ϕ and within a small pre-defined $3 \times 3 \times 3$ neighborhood, imposed by the acquisition procedure, i.e., the neighborhood was set to a size equivalent to the ventricular septum thickness, a 3×3 neighborhood in the low-resolution image. The weights were set to 1 at the center and -0.0384 at any other location within the neighborhood, after maximizing the covariance of weights of the Gibbs distribution from a set of 1000 randomly selected neighborhoods, as described above.

$$\mathbf{k}_{i,j} = \begin{cases} 1, & for i = j \\ -0.0384 & for j: \phi_j is a neighbor of \phi_i \end{cases} \quad (10)$$

The above approach only takes into account local characteristics of the image. The prior was then complemented to include global information from the low-resolution image by centering the data $\mathbf{w}$ around an estimated of the smooth variations of the low-resolution image $\bar{\mathbf{w}}$. For so doing, $\bar{\mathbf{w}}$ is divided into tiles of a predetermined size and the tile averages are calculated for the whole set of available series, that is to say

$$\mathbf{w}\cdot\mathbf{\Phi} = \bar{\mathbf{w}}\cdot\mathbf{\Phi} + \varepsilon \quad (11)$$

Assuming an additive and independent noise, the complementary prior reads as

$$p(\mathbf{w}|\bar{\mathbf{w}}) = \frac{1}{(2\pi)^M\sigma^{12}} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{w}\cdot\mathbf{\Phi} - \bar{\mathbf{w}}\cdot\mathbf{\Phi})^2\right\} \quad (12)$$

The prior term is then a function that captures the uncertainty of the local influence under a

restricted global shape, as formulated in Equations (8) and (12). The global information $\bar{\mathbf{w}}$ is an estimate of the smooth variations of the average image intensity, calculated for the whole set of available series at each of the neighborhoods defined in Equation (8). This term forces any instance to stay close to the prior global shape and preserves not only shapes but relationships between Regions of Interest. In contrast, the local information prior tends to highlight low-level features like homogeneous regions and/or sharper edges, factors which are straightforwardly associated with the particular acquisition protocol.

3.2.2 Gradient descent

The complete Bayesian formulation is solved using a cost function with the likelihood, the local and global prior terms previously defined

$$\hat{\mathbf{w}} = * \, arg\, min_{\mathbf{w}}\{-\ln[p(\mathbf{I}_l|\mathbf{w},\boldsymbol{\Phi})] - \ln[p(\mathbf{w}|\boldsymbol{\Phi})] - \ln[p(\mathbf{w}|\bar{\mathbf{w}})]\} \quad (13)$$

Solving and eliminating terms that do not depend on $\mathbf{w}$, the cost function $L(\mathbf{w})$ is

$$L(\mathbf{w}) = \frac{1}{2\sigma^2}\sum_{n=1}^{N}\left(\mathbf{I}_{l,n} - g_l\sum_{r=1}^{M} w_{n,r}\phi_r\right)^2 + \frac{1}{2}\sum_{i=1}^{M}\left(\sum_{j=1}^{M} k_{i,j}w_j\phi_j\right)^2 + \frac{1}{2\sigma^2}\sum_{i=1}^{M}(w_i\phi_i - \bar{w}_i\phi_i)^2 \quad (14)$$

Values for $\hat{\mathbf{w}}$ are calculated according to

$$\hat{\mathbf{w}}^{n+1} = \hat{\mathbf{w}}^n - \varepsilon\nabla L(\hat{\mathbf{w}}^n) \quad (15)$$

Where

$$\nabla L(\hat{\mathbf{w}}^n) = \left[\frac{\partial L(\hat{\mathbf{w}}^n)}{\partial w_1} \quad \frac{\partial L(\hat{w}^n)}{\partial w_2} \quad \cdots \quad \frac{\partial L(\hat{w}^n)}{\partial w_N}\right]^T \quad (16)$$

4 Experimental results

Qualitative and quantitative evaluations were carried out by calculating the reconstruction accuracy of the proposed SR method. Several quantitative experiments were conducted using artificial (synthetic) and real images. The proposed method was implemented in C++ using ITK library version 3.20 on an Intel Core i7, 12 GB of RAM and Debian-Linux operating system version 6.0. For a single low-resolution slide (256×256 pixels), a full reconstruction took 280 seconds. For a complete 3D reconstruction, the computation time is calculated by $time = 280 \times slices \times series$. For the datasets used in this study, reconstruction times varies between $1\,h$ $30\,min$ and $4\,h$ $25\,min$.

4.1 Artificial (synthetic) images

The Shepp-Logan phantom was used as a synthetic image model. The size of the phantom is $300 \times 300 \times 300$ voxels, and the intensity values vary from 1000 to 9000 (typical values in CMR). The use of these synthetic images evaluates different factors that affect the reconstruction accuracy under controlled conditions. In this work, five scenarios were covered: influence of prior information, upsampling strategy, orthogonality, number of series and comparison with other methods. For these experiments, up to 7 series were extracted from the original volume. Each series consisted of a set of 43 slices (size 256×256 pixels), with an in-plane resolution of 1.6 mm and a slice separation of 8 mm. A Gaussian noise ($\mu = 0$, $\sigma^2 = 1000$) was added to each slice to emulate the acquisition process. Orthogonal series were extracted along the three principal axes of the phantom, while the non-orthogonal series were extracted along randomly-selected angles (values included in Table 0). Three different metrics were used to quantitatively measure the reconstruction quality: the Peak Signal-to-Noise Ratio (PSNR), the Structural Similarity index (SSIM) [19] and IL-NIQE quality assessment [20]. A noiseless version of the phantom was used as the reference image.

IL-NIQE is a blind image quality assessment (BIQA) which models the way humans score quality. This metric requires a multivariate Gaussian model learned from a collection of paradigmatic images. The quality of an image is quantified by using a Bhattacharyya-like distance, i.e., the learned multivariate model is compared with the distribution of particular features extracted from the test image. In the following experiments, results are presented as percentage of the quality difference between the reference and test images.

Table 1: Random values for evaluated angles in radians

	non-orthogonal series							orthogonal series		
xis	1	2	3	4	5	6	7	1	2	3
	0.5	0	0	-0.8	0.2	-0.1	0	$\pi/2$	0	0
	0	0.7	-0.5	0	-0.5	0.4	0	0	$\pi/2$	0

4.1.1 Prior influence on reconstruction

In this section, the contribution of global prior information is evaluated. Two different strategies were tested: using only local prior information and a combination of local and global information. A set of 7 non-orthogonal series were used in this test. Table (1) presents the quantitative reconstruction performance. Numerical results in Table (1) confirm that the reconstruction, using the two prior hypotheses (local and global), shows a better reconstruction accuracy. The use of the two priors results in an improvement of $1.2\ dB$ in the PSNR, 3% in the SSIM and 2.1% in the quality metrics. Hence, only the local + global prior combination was used for the rest of experiments.

Table 2: Evaluation of the reconstruction accuracy using two prior hypothesis

2*Metric	2*local prior	2*local + global prior
PSNR	21.7112	22.9424
SSIM	0.9530	0.9837
Quality %	76.21	78.29

4.1.2 Upsampling method

This section evaluates the influence of the upsampling strategy used in the preprocessing step (Section 3.1.2). Three different upsampling strategies were tested: tri-linear and bi-cubic interpolations and the statistical filtering method [17]. In this test, a set of 7 non-orthogonal series is used. Table (2) shows the obtained results and values with the statistical filtering demonstrate an important improvement of the reconstruction accuracy, i.e., $3\,dB$ and $1\,dB$ for PSNR, 3.8% and 2.8% for SSIM and 3.17% and 1.7% for the quality metrics, when comparing with linear and bi-cubic interpolation, respectively.

Table 3: Reconstruction performance of the proposed method using three upsampling strategies

2*Metric	Linear interp.	bi-cubic interp.	Statistical Filtering
PSNR	20.1069	22.6689	22.9424
SSIM	0.9456	0.9553	0.9837
Quality %	75.12	76.62	78.29

4.1.3 Orthogonality influence

The influence of orthogonality was tested by using two different configurations: three orthogonal series (sagittal, coronal and transversal views) against three non-orthogonal series, extracted following arbitrary angles. Numerical evaluation using the three selected metrics (see Table 3) shows no important differences between both configurations, suggesting that the orthogonality is not a relevant factor for the reconstruction quality.

Table 4: Reconstruction performance using orthogonal and non-orthogonal series.

Metric	orthogonal	non-orthogonal
PSNR	18.3807	18.3787
SSIM	0.8668	0.8624
Quality %	69.09	69.13

4.1.4 Quality of the Reconstruction of the high-resolution image

This section explores the capability of the method to approximate the high-resolution image from a smaller number of slices. For this purpose, one subset of the Shepp-Logan phantom

and two real cases (named subjects A and B) were prepared. Each series contained a different number of non-orthogonal views, chosen from the original volume. The original series are always composed of $8\,mm$-thick slices at a resolution of 256×256, and the experiments were performed by subsampling these sets of series in the inter-slice dimension, resulting in series with $16\,mm$-thick slices. The distribution of series and slices for each dataset is as follows:

- The Shepp-Logan: The full set of data corresponds to 7 series with 43 slices per series. The subset used for the reconstruction consists of 2 series with 22 slices per series.
- Real subject A: the set of images for subject A is composed of 5 series with 10 slices. The subset used for the experiment is composed of 2 series with 5 slices each.
- Real subject B: The set of images for subject B is composed of 6 series with 12 slices. The subset is composed of 2 series with 6 slices each.

Table (4) shows a net PSNR gain of about $2.68\,dB$ and $2.25\,dB$ when compared with the linear and the bi-cubic interpolations, respectively, a SSIM gain of $4.1\,\%$ and $2.9\,\%$, and an increasing of the quality metrics in about 4% and 2%. Likewise, visual comparison of the resulting images suggests that edge irregularities nearly disappeared when using the proposed method.

Table 5: Comparison of a standard interpolation and the proposed method when combining a subset of slices of the original volume.

Metric	linear	bi-cubic	Proposed method
PSNR	24.6384	25.0727	27.3263
SSIM	0.9564	0.9681	0.9975
Quality %	80.47	82.23	84.04

Table (5) shows the PSNR, SSIM and quality metrics for the two different subjects, using the reconstruction obtained with the full set of slices as the reference image. The proposed method outperforms the linear interpolation by improving the reconstruction in $1.77\,dB$ and $4.22\,dB$ for subjects A and B respectively, a SSIM increment of $2.02\,\%$ is observed for subject B and of $0.03\,\%$ for subject A, while the quality metrics shows an increment of 5% and 4% for subjects A and B respectively. Likewise, the net PSNR gain for the bi-cubic interpolation is $0.96\,dB$ and $2\,dB$ for subjects A and B, respectively, $0.01\,\%$ and $0.9\,\%$ for the SSIM metrics and 3% and 2% in the quality metrics. Overall, SSIM figures are larger than 0.9, indicating a reconstructed image very close to the reference.

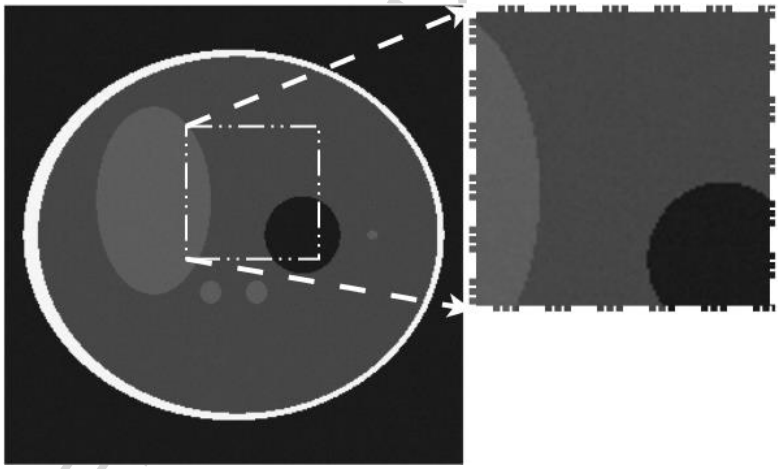
Table 6: Comparison of interpolation strategies and proposed method combining a subset of input cardiac MR series.

Metric	linear	bi-cubic	Prop. method	
PSNR	29.58	30.39	31.35	Subject
SSIM	0.9739	0.9741	0.9742	A
Quality %	83.03	85.35	88.62	
PSNR	25.19	27.41	29.41	Subject

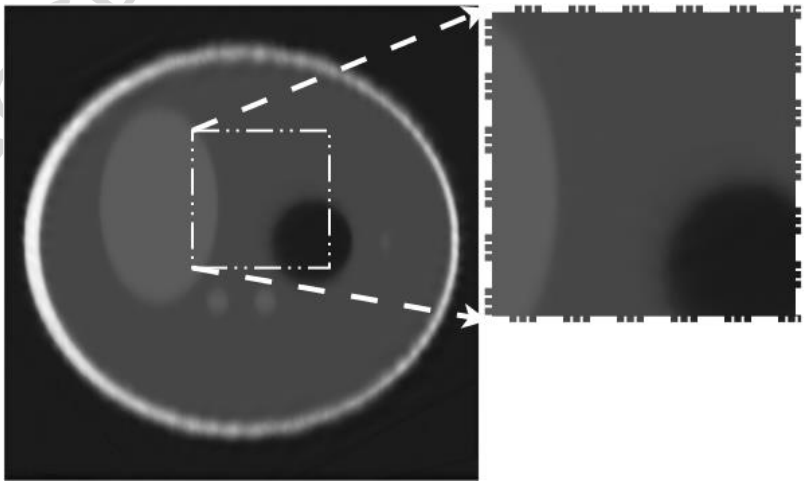
SSIM	0.9381	0.9481	0.9571	B
Quality %	82.76	84.88	86.14	

4.1.5 Baseline comparison

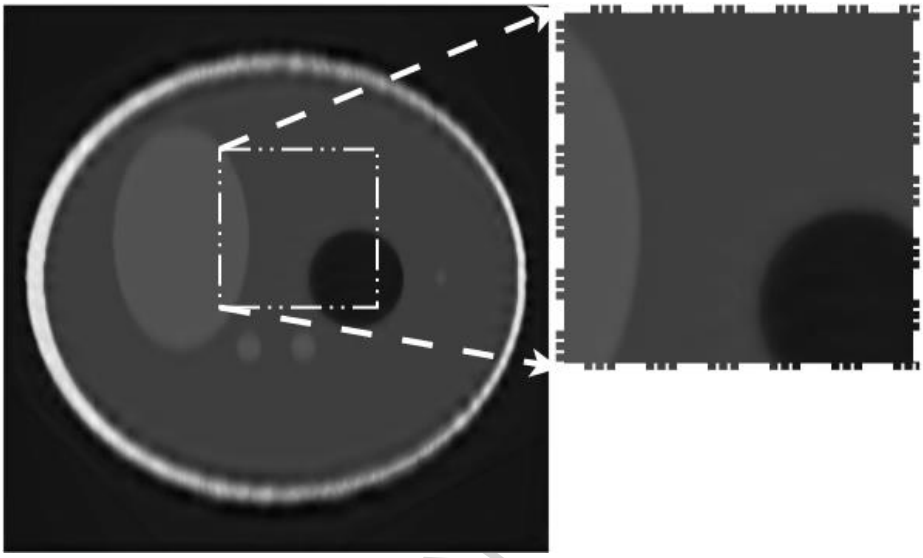
In this section, the performance of the proposed method is compared with a baseline strategy, commonly used for visualization of the CMR images. It consists in a simple average of the voxel intensities after interpolation and spatial normalization of the different series. In this test, the statistical filtering method [17] is used for interpolation, and a set of 7 non-orthogonal series is used.



[a]



[b]



[c]

Figure 2: Comparison with a basic strategy. (2a) original image, (2b) interpolation and averaged image, (2c) result using the proposed method. The difference between the interpolated and the reconstructed images is observed at the zoomed window as a black boundary that is sharper for the reconstructed version.

Fig. (2b) and (2c) show the results of the basic strategy (interpolation and averaging) and the proposed approach, respectively. Zoomed areas illustrate the differences, mainly at the edges, where the proposed method (Fig. (2c)) achieved a reconstruction with edges better defined than the basic strategy. Although edge artifacts are present in both images, they are more visible for the baseline strategy. The quantitative results presented in Table (6) show that the proposed method outperforms the baseline in about 2*dB* in PSNR, 1% for SSIM and 5% in the quality metrics with respect to the interpolation and averaging strategy.

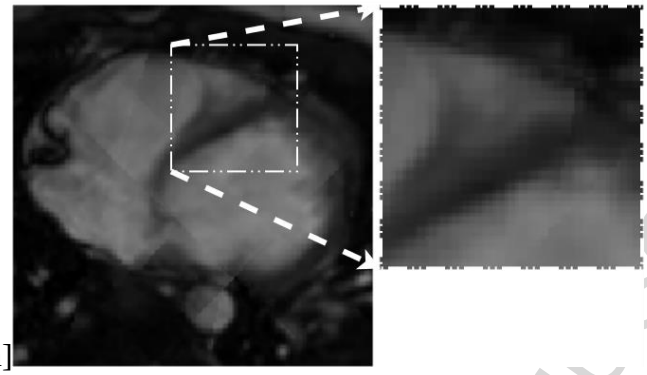
Table 7: Comparison between a basic strategy and the proposed approach.

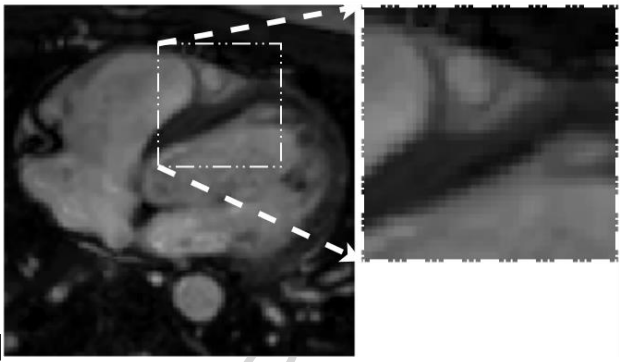
2*Metric	Interpolation and averaging	Proposed approach
PSNR	18.0656	22.9424
SSIM	0.9575	0.9837
Quality %	73.17	78.29

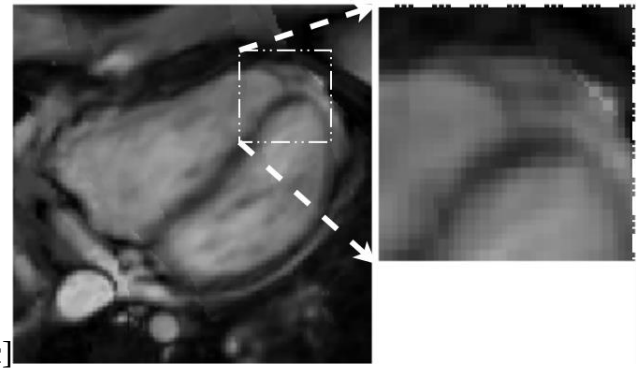
4.2 Real images

For testing the method performance in real images, two different datasets were used. The first was composed of five CMR cases acquired in a 1.5 T scanner from five different subjects. The cases are composed of a variable number of series (between 5 and 11 series) at different non-orthogonal orientations. Each series contains between 2 to 15 slices, an inter-slice separation varying between $8\,mm$ to $11\,mm$, an in-plane resolution of 256 x 256 and a pixel spacing of

1.6 *mm*. In this section, results obtained with the proposed method are qualitatively compared with baseline consisting in a combination of a statistical filtering [17] and a voxel intensity averaging, as previously explained with synthetic images in Section 4.1.5.

[a] 

[b] 

[c] 

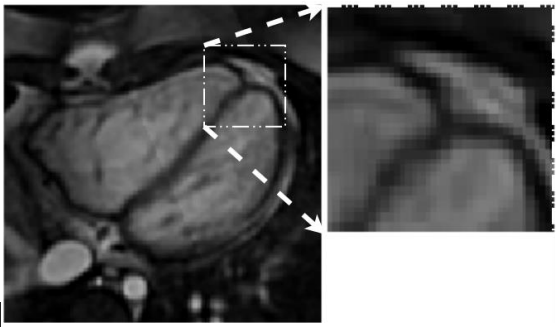
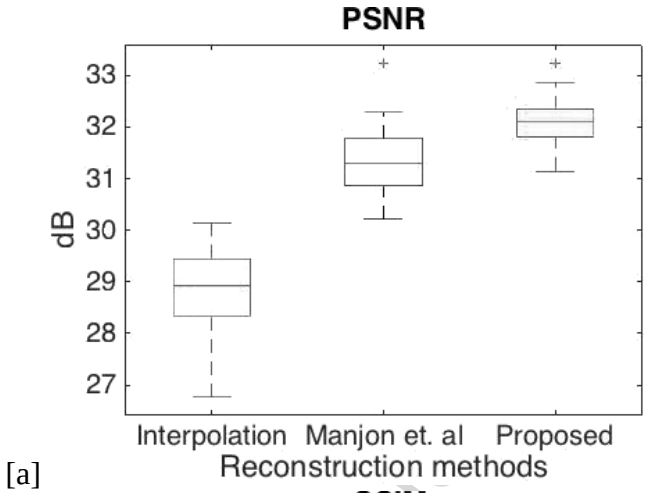
[d] 

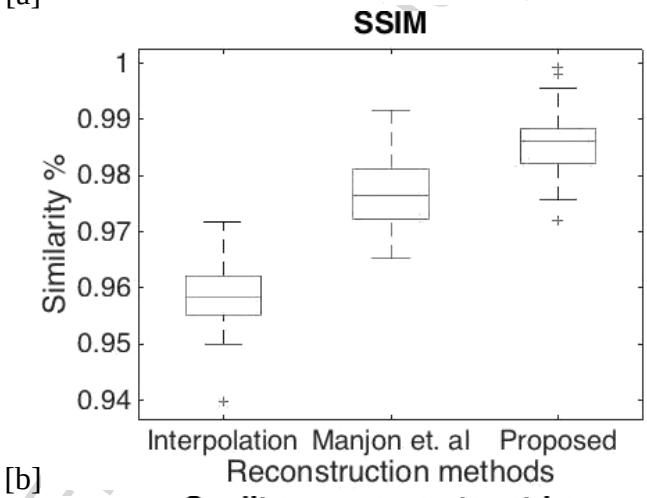
Figure 3: SR reconstruction using real images from two different subjects. Row 1: subject A; Row 2: subject B. (3a) and (3c). The figure illustrates the reconstruction of the interventricular septum at the level of the heart apex. Panels (a) and (c) correspond to the results after interpolating and averaging, the baseline. Likewise, panels (b) and (d) stand for results after the proposed method. Notice how the septum is better defined by the proposed method

Images in panels (3b) and (3d), compared with panels (3a) and (3c), show that the proposed method delivers better structural (shape) details and sharper edges than the averaging-based method, the baseline. Specifically, the reconstructed images (3b) and (3d) have succeeded about improving the details of the ventricular and septum walls, particularly at the level of the heart apex. Note that reconstructed images present some artifacts related with contrast differences. These artifacts are essentially multicausal, associated with the long transient of the sequence which causes the first image to show a different contrast since there is not time to reach the stationary state [21]. A complementary effect is introduced by the differences in the dynamic range of the original low-resolution images. CMR images are reconstructed using captures from different cardiac cycles, each under different non-controlled conditions such as changes in the thorax volume or patient movement. As observed in panels (3b) and (3d), the correction contrast strategy highly reduced the described non-desirable effects. This might also be improved by carefully selecting only those series where the gray value variations can be corrected, but a supervised process would then be required.

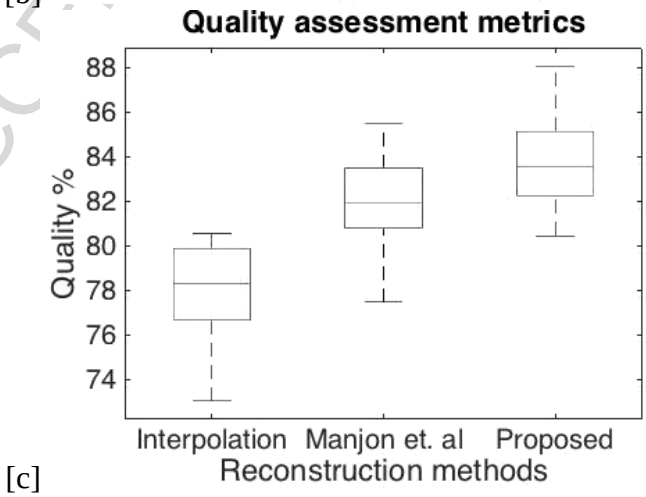
The second dataset allowed us to quantitatively evaluate the reconstruction quality. This dataset, an open project publicly available [22], is composed of 30 real CMR studies captured from different patients. These images were scanned with a GE Genesis Signa MR scanner using the FIESTA scan protocol. Each case is composed of a set of parallel slices that follow the long heart axis, capturing the whole heart and always showing a view with two chambers. The in-plane resolution is 256×256 with a pixel spacing of $1.64\ mm$, along with an interslice separation of $8\ mm$. Series are composed of 10 to 15 slices, depending on the heart size. Before the SR algorithm, a spline interpolation method was applied and isotropic voxels were generated. From this corrected volume, 7 non-orthogonal series were obtained for each of the cases by randomly setting non-orthogonal angles. An additional subsampling process generated anisotropic voxels for the interslice space was set to $8\ mm$. For every available case, three high-resolution images were reconstructed from the anisotropic voxel series, the first using the spline interpolation and averaging strategy, a second reconstruction used the super-resolution method proposed by Manjon et. al and the third was obtained by applying the proposed method. Quantitative comparisons of the reconstruction quality were performed using PSNR, SSIM metrics and IL-NIQE quality metrics. The obtained metrics from the reconstructed images are presented in Figure (4).



[a]



[b]



[c]

Figure 4: Quantitative results of SR reconstruction compared with interpolated and averaged low-resolution data and Manjon et. al method for the three different metrics herein used, namely PSNR, SSIM and the Quality metric.

Numerical results show that our proposal always improves the baseline method (interpolation plus averaging) and SR method (Manjon et.al) using the real CMR dataset in about

two to three dB, as observed in Figure (4a). In terms of the applied PSNR, SSIM and Quality assessment metrics, the reconstruction using the proposed method outperforms in average the baseline strategy, i.e., about 3.2dB in PSNR, 0.26 in SSIM and 6 % in Quality. The PSNR for every SR reconstruction exceeds the 31dB, the SSIM reaches levels of 0.97 and Quality is over 80 %, achieving levels of high-quality reconstruction.

5 Discussion and conclusion

In this paper, a novel SR method for improving the image quality of CMR image series, based on a sparse Bayesian approach, has been described and analyzed. This method can be applied with non-orthogonal series and highly anisotropic voxels, conditions which are associated to a non-uniform sampled grid. Two main stages are proposed, a pre-processing step and a SR algorithm. The pre-processing step is intended to transform the input series into a set of volumes with normalized intensities, isotropic voxels and a common spatial representation. In the SR algorithm, a Bayesian sparse framework includes a likelihood function and a very adapted prior model. Two prior hypotheses were used in the proposal, the former related to preserve local information (defined in Equation (8)), searching for a low-noise estimated image, while the latter attempted to include global features like the organ shape (defined in Equation (12)). The set of hypotheses promotes sparse solutions by defining a prior dependent on the local influences associated to the particular capture protocol and an averaged heart correction. This method generates high-resolution reconstructions of cardiac MR images, much more suitable for accurate heart segmentation and motion analysis.

SR has been introduced for MRI to enhance the quality of medical images [10], [9], [14]. Previous works have considered orthogonal and non-orthogonal views of brain images [15], [13]. In contrast, as far as we know, SR in CMR has been applied only in orthogonal series [14], in which case the SR has been formulated as a multi-resolution problem with a constant sampling factor. This feature allows to manage a high-resolution uniform grid, simplifying the task of estimating the high-resolution image. The complexity generated by the non-orthogonality amounts to a non uniform sampling, or to the ill conditioned problem of a space generated by an uncomplete set of bases. Introduction of specific problem knowledge has facilitated the sub-space regularization, setting thereby an acceptable solution, a way already explored by a precedent work in CMR images [14]. Yet this investigation approached such problem, their conclusions were very limited since the presented results were purely qualitative and the model not general.

In addition, previous works have introduced SR methods for dealing with anisotropic and non-orthogonal multislice images, such as Poot et al. [13], [23] and Plenge et al. [11]. These strategies are dependent on an acquisition process in which the anisotropic sampling is approximated by a linear operator, limiting the capability of successfully dealing with uncertainties or corrupted data coming from the cardiac or thorax motions. In the particular case of Poot et al., the resultant high-resolution intensities are obtained by linear combination of a local model defined as a function of the image tensor, i.e, only first-order relationships are taken into account. Likewise, the PSF model is a kind of linear combination of the low-resolution PSF that is defined with a windowed sinc function along the slice direction. This amounts to ensure that only information contained in two adjacent parallel slices is considered for the reconstruction. With such approach, it is crucial a correct alignment between slices. In contrast, our method presents a more flexible strategy (not dependent on the acquisition model) that establishes more complex relationships, with larger statistical dependencies, i.e., second- or third-order relations, since the

approximation weights can be learned. In case of CMRI, an important uncertainty source is the inter- and intra-series misalignments, coming from the motion effect, and in this scenario a strategy based on a classical acquisition model may crash, because such organ movements are not considered into the model. Finally, provided that main hypotheses of these methods are local, they usually incorporate high-frequency energy minimization terms as regularizers. In contrast, the presented proposal naturally combines different and adapted hypotheses in the regularization strategy, in particular those related with fundamental anatomical relations that may be either local or global.

The proposed method introduces a very adaptable model and has been assessed in several (at least five) non-orthogonal and very anisotropic voxel images (proportion of 5:1). Evaluation of the reconstruction with the proposed method was performed by conducting tests on synthetic and real images. The analysis performed in synthetic images demonstrates the influence of different parameters and conditions related with the input images. Four factors were herein explored: prior influence, upsampling strategy, orthogonality and number of series. The learning phase corresponded herein to the determination of the kernel weights that defined the local neighborhood of the sparse representation by using the covariance metrics, considered as a good estimator under Gaussian support. The chosen local relationship should balance the statistical dependence and the sample distance, a relationship that was herein covered using a $3 \times 3 \times 3$ neighborhood. On the other hand, the upsampling strategy has also shown to directly influence the reconstruction accuracy results. The use of an accurate interpolation method in the pre-processing step, such as the state-of-the-art statistical filtering method [17] for single MRI volumes, results crucial since this step operates as a first estimation of the reconstruction. Furthermore, this preprocessing step can be also considered as a baseline of the present work, case in which our method outperformed this state-of-the-art super-resolution method [17] that fails since it was basically devised to reconstruct parallel series. In addition, the experiment using a variable number of series, to reconstruct the whole volume, has demonstrated that the proposed method obtains more accurate results than a simple approximation (interpolation) or an isotropic super-resolution method [17]. In particular, this method achieves high resolution versions of an image by integer factors, forcing the super-resolution to be performed within a regular grid, an obvious limitation in case of the heart reconstruction. This effect is visually illustrated in Section 4.2, where the blurred image is likely the result of the superposition of different series with intersections at real locations. Finally, the orthogonality acquisition appears to be not important for the reconstruction accuracy. These results suggest that the proposed method is not affected by the orientation of the input series and hence it can be directly applied to any type of protocol acquisition. Figure (3) shows an evident improvement in the right column images, when comparing with images in the left column,

A comparison with a baseline strategy was also performed in both synthetic and real images. The obtained results show that the proposed method always obtains a better high-quality reconstruction than the interpolation or averaging strategies. The main difference between the baseline strategy and the proposed SR method is that the former basically calculates the mean intensity of the available series, while the latter attempts to find the most probable high-resolution image that explains the input series, regularized by a certain prior knowledge of both the heart structure and the capture protocol. The obtained reconstruction results confirm that the inclusion of prior information and a probabilistic framework represents an advantage in SR of CMR images, increasing the reconstruction accuracy. Given the difficulty of the current acquisition process of CMR images, where the heart motion restricts the scanning time and thus the number of slices per volume series, the search of a high-resolution image for diagnosis and surgical planning is still an

open issue. The quality of the high-resolution reconstructions obtained with the proposed method introduce a promising approach for increasing the image quality in tasks such as segmentation, morphometry and heart motion analysis. An initial extension to handle complete 4D CMR image series can be achieved by reconstructing individually each 3D frame volume.

References

[1] H. Childs, L. Ma, M. Ma, J. Clarke, M. Cocker, J. Green, O. Strohm, M. G. Friedrich, Comparison of long and short axis quantification of left ventricular volume parameters by cardiovascular magnetic resonance, with ex-vivo validation, Journal of Cardiovascular Magnetic Resonance 13 (1) (2011) 1.

[2] J. A. Lima, M. Y. Desai, Cardiovascular magnetic resonance imaging: current and emerging applications, Journal of the American College of Cardiology 44 (6) (2004) 1164–1171.

[3] T. D. Karamitsos, J. M. Francis, S. Myerson, J. B. Selvanayagam, S. Neubauer, The role of cardiovascular magnetic resonance imaging in heart failure, Journal of the American College of Cardiology 54 (15) (2009) 1407–1424.

[4] C. E. Goodyer, V. Grau, T. Mansoori, J. E. Schneider, K. W. Brodlie, P. Kohl, 3d visualization of cardiac anatomical mri data with para-cellular resolution, in: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2007, pp. 147–151.

[5] J. P. Earls, V. B. Ho, T. K. Foo, E. Castillo, S. D. Flamm, Cardiac mri: recent progress and continued challenges, Journal of magnetic resonance imaging 16 (2) (2002) 111–127.

[6] E. Carmi, S. Liu, N. Alon, A. Fiat, D. Fiat, Resolution enhancement in mri, Magnetic resonance imaging 24 (2) (2006) 133–154.

[7] R. C. Hardie, K. J. Barnard, E. E. Armstrong, Joint map registration and high-resolution image estimation using a sequence of undersampled images, IEEE Transactions on Image Processing 6 (12) (1997) 1621–1633.

[8] R. C. Hardie, K. J. Barnard, J. G. Bognar, E. E. Armstrong, E. A. Watson, High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system, Optical Engineering 37 (1) (1998) 247–260.

[9] Y. Bai, X. Han, J. L. Prince, Super-resolution reconstruction of mr brain images, in: Proc. of 38th Annual Conference on Information Sciences and Systems (CISS04), 2004, pp. 1358–1363.

[10] H. Greenspan, G. Oz, N. Kiryati, S. Peled, Mri inter-slice reconstruction using super-resolution, Magnetic resonance imaging 20 (5) (2002) 437–446.

[11] E. Plenge, D. H. Poot, M. Bernsen, G. Kotek, G. Houston, P. Wielopolski, L. van der Weerd, W. J. Niessen, E. Meijering, Super-resolution methods in mri: Can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time?, Magnetic resonance in medicine 68 (6) (2012) 1983–1993.

[12] R. R. Peeters, P. Kornprobst, M. Nikolova, S. Sunaert, T. Vieville, G. Malandain, R. Deriche, O. Faugeras, M. Ng, P. Van Hecke, The use of super-resolution techniques to reduce slice thickness in functional mri, International Journal of Imaging Systems and Technology 14 (3) (2004) 131–138.

[13] D. H. Poot, V. Van Meir, J. Sijbers, General and efficient super-resolution method for multi-slice mri, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2010, pp. 615–622.

[14] S. Ur Rahman, S. Wesarg, Combining short-axis and long-axis cardiac mr images by

applying a super-resolution reconstruction algorithm, in: SPIE Medical Imaging, International Society for Optics and Photonics, 2010, pp. 76230I–76230I.

[15] A. Gholipour, J. A. Estroff, S. K. Warfield, Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain mri, IEEE transactions on medical imaging 29 (10) (2010) 1739–1758.

[16] U. E. Ruttimann, R. L. Webber, E. Schmidt, A robust digital method for film contrast correction in subtraction radiography, Journal of Periodontal Research 21 (5) (1986) 486–495.

[17] J. V. Manjón, P. Coupé, A. Buades, V. Fonov, D. L. Collins, M. Robles, Non-local mri upsampling, Medical image analysis 14 (6) (2010) 784–792.

[18] DICOM Working Group 17, Supplement 73: Spat. Reg. Storage SOP Classes, DICOM Standards Comitee (01 2004).

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE transactions on image processing 13 (4) (2004) 600–612.

[20] L. Zhang, L. Zhang, A. C. Bovik, A feature-enriched completely blind image quality evaluator, IEEE Transactions on Image Processing 24 (8) (2015) 2579–2591.

[21] K. Scheffler, On the transient phase of balanced ssfp sequences, Magnetic resonance in medicine 49 (4) (2003) 781–783.

[22] A. Andreopoulos, J. K. Tsotsos, Efficient and generalizable statistical models of shape and appearance for analysis of cardiac mri, Medical Image Analysis 12 (3) (2008) 335–357.

[23] D. H. Poot, B. Jeurissen, Y. Bastiaensen, J. Veraart, W. Van Hecke, P. M. Parizel, J. Sijbers, Super-resolution for multislice diffusion tensor imaging, Magnetic resonance in medicine 69 (1) (2013) 103–113.