International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 12
August 2016

Super-Resolution Reconstruction for Cardiac MRI using Coupled Dictionary Learning

Dr. Y. Raghavender Rao

Assoc. Prof & Head of Dept. of ECE,
Jawaharlal Nehru Technological University
Hyderabad College of Engineering Nachupally
(Kondagattu), Karimnagar.

G. Drakshayini

M.TECH Dept. of ECE,
Jawaharlal Nehru Technological University
Hyderabad College of Engineering Nachupally
(Kondagattu), Karimnagar.

Abstract:

Dynamic magnetic resonance imaging (DMRI) requires high spatial and temporal resolutions, which is challenging due to the low imaging speed. To reduce the imaging time, a patch based spatiotemporal dictionary learning (DL) model is proposed for compressed-sensing reconstruction of dynamic images from under sampled data. Specifically, the dynamic image sequence is divided into overlapping patches along both the spatial and temporal directions. These patches are expected to be sparsely represented over a set of temporal dependent spatiotemporal dictionaries. The images are then reconstructed from the under sampled data in (k,t) space under such sparseness constraints [2], where the dictionaries are learned iteratively. Alternating optimization is applied to solve the problem. Simulation results show that the proposed method is capable of preserving details in both spatial and temporal directions. Keywords: Dynamic Magnetic Resonance Imaging (DMRI), Dictionary Learning (DL), Compressed Sensing (CS) [5].

I. INTRODUCTION

Dynamic magnetic resonance imaging (dMRI) plays a key role in many clinical applications, such as cardiac cine imaging, parameter mapping and functional imaging. However, high spatial-temporal resolution dMRI is still challenging due to the long scan time. A lot of work has been done to reduce the amount of acquired data and thus the acquisition time by exploiting the spatial and temporal correlations in the dynamic image sequence. With the emergence of compressed sensing (CS)[5], such correlations are represented as sparseness of the image sequence under a certain scarifying transform. Conventional CS requires prior knowledge about the scarifying transform which is essential for the recovery process. A suitable scarifying transform should have certain structures and fast numerical implementations, such as Fourier and Wavelet transform [1]. However, it is difficult to find a universal transform that can sparsify all signals. The recent research on sparse representation over learned dictionaries indicates great potential on many practical applications. Properly trained or adaptive dictionaries have the ability to offer more flexible sparse representations and provide better performances over the conventional analytical sparsifying transforms. In recent years, the application of dictionary learning (DL) in MRI has attracted a lot of attentions. In a dictionary is learned to sparsely represent images spatially. Although these methods can be directly extended to dMRI by reconstructing images frame by frame, the sparseness is not exploited along the temporal axis.

On the other hand, the concept of DL is used for temporal sparse representation in where the Application concentrates on the parameter mapping. Most recently, a spatiotemporal DL model is proposed for dMRI with an additional low-rank constraint. This method uses all the temporal data points to form the dictionary. It may not capture the local abrupt motion and thus lose some temporal details. Furthermore, the size and the number of atoms increase rapidly with the number of frames. Building such a large dictionary leads to high computational complexity due to the nonlinear operations. In the spatially-based DL techniques, an

image is divided into many overlapping patches, which are used for dictionary learning and sparse representation. Inspired by the work, we extend the patch-based DL model [5] from the spatial domain to the spatial-temporal domain. The image sequence in dMRI is divided into overlapping patches along both spatial and temporal directions. Thus, the atoms in the corresponding dictionaries are of three dimensions. In contrast to the spatiotemporal model, our model adapts to specific local spatial temporal features. The computation time in DL is also reduced due to the smaller size of atoms. The simulation results show that the proposed method is capable of capturing local features, especially the abrupt changes along time.

One of the most common approaches to the super-resolution problem is to use the maximum likelihood (ML) or maximum a posteriori (MAP) [3] estimation. In these approaches, a distance measure between the reconstructed image and the observed images is iteratively reduced. Example-based image super-resolution is another popular approach where correspondences between low- and high- resolution image patches are learned from a database and then applied to a new low-resolution image to recover its most likely high-resolution version. In both approaches, a distance measure between the current estimation and the low-resolution images must be computed. Takeda proposed a distance estimation approach which requires no explicit motion estimation by inverting the position of the patch selection operator and the re-sampling operator.

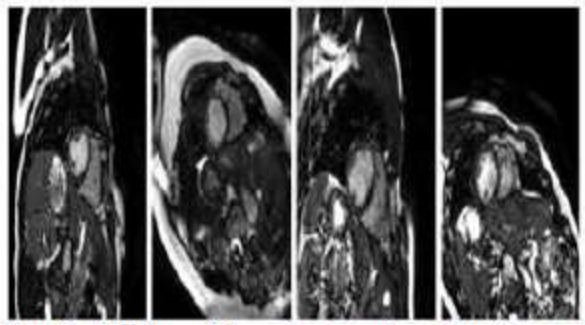


Fig.1.Variability of heart orientation, position and shape across subjects.

The idea of super-resolution has been applied in medical imaging too: reconstructed a high-resolution volume from multiple low resolution (LR) images using image priors based on total variation constraint with MAP estimation. However, this method cannot be directly applied to our problem because it requires multiple instances of low resolution (LR) images from different views. To take advantage of the information redundancy in similar patches across different subjects, patch-based methods have been shown to be highly efficient in applications such as segmentation proposed to combine registration with a patch-based approach to create super-resolution brain MR images from atlases of multiple LR images of different subjects. In this approach the high-resolution image is constructed via non-local fusion of those patches. However, the method requires rough correspondence between images either via explicit motion estimation or other means. This is difficult to guarantee in cardiac MR images due to the large variation in the orientation, position and shape of the heart across subjects (see Fig.1). Moreover, the complexity of these non-local patch-based methods increases with the number of atlases.

In this paper we aim to reconstruct a super-resolution (SR) cardiac MR image from a single short axis (SA) [4] cardiac MR image with a set of 3D atlases available as the training database. Three different aspects are challenging: First, the slice thickness of the SA image is much larger than the slice thickness of the 3D image (approximately 5 times, e.g. 2mm vs. 10mm) while the up-sampling factor of classic super-resolution algorithms is usually around two. Second, the search for the best match of patches in 3D with multiple atlases using conventional approaches is very expensive. Finally, cardiac images exhibit significantly more variability in terms of orientation and anatomy compared to brain images. Local search methods used in brain imaging are thus not suitable. In addition an exhaustive global search for patches is impossible given the computational cost to solve this problem, we propose a framework to combine classic and example- based super-resolution approaches using an approximation graph based search based on the recently proposed Patch Match algorithm. Inspired by, we assume that information redundancy in similar patches across different subjects can be exploited. Thus, we reformulate the Patch Match approach to find patch correspondence between a single image and an atlas database. We then use the principles in to estimate the super-resolved image using the expectation-maximization (EM) framework [2]. The novelty and contributions of this paper are the introduction of a global search strategy as well as an observation model with non-explicit motion estimation that avoids any spatial alignment or registration of the images. Furthermore, the computational cost is kept

low by using Patch Match and a closed-form solution in the observation model. The number of atlases does not influence the computational cost and thus allows full exploitation of a large atlas database. Our results demonstrate that the algorithm can robustly estimate a SR image in the presence of thick slice data and performs both extrapolation and interpolation by recovering missing apical and basal slices.

II. SPARSIFYING TRANSFORMS FOR DYNAMIC MRI

There is a direct relationship between the sparsity level L provided by the transform S and the minimum number of samples necessary for perfect reconstruction. The sparser the model chosen, the higher the achievable acceleration rates will be. In this section, two sparsifying transforms that are well suited to the problem of cardiac cine are described. First, an adaptive patch based transform derived from DL theory [5] is presented for the case of dynamic MR complex data. Then, the temporal gradient transform is proposed as a suitable global sparsity model for cardiac cine that can make reconstructions converge faster and improve performance at high under sampling rates.

A. Dictionary Learning For Dynamic MRI

DL refers to the process of adapting an initial set of basic functions to a specific signal through a training process such that it will provide a sparse representation of that particular signal. Recently, DL has been used for 2D structural MR image reconstruction largely outperforming competing techniques based on fixed basis transforms. The extension to the case of MR sequences was introduced with the use of spatio-temporal 3D dictionaries, but only the reconstruction of synthetic real-valued sequences without a phase component was addressed. This is not feasible in practice since the observed k-space samples always relate to a complex image. The training of complex-valued dictionaries that is suitable for MR data representation is possible defining T and D as complex-valued variables as originally proposed. Instead, we carry out this representation by using a single real valued dictionary which is trained on real and imaginary parts of MR data for their independent coding we look at the differences between these two learning and coding strategies for the processing of MR data. Previous work has already investigated the use of dictionary learning for cardiac cine [1]. In, complex spatio-temporal dictionaries are learnt and used as unique sparse model for patches of the reconstructed data. A key difference with the method proposed here apart from the training strategy is the fact that the dictionary is propagated and updated along the temporal dimension. This was proposed for dictionary learning applications in natural video processing, where the data structure is expected to rapidly change across time and propagating and updating the dictionary adapts it to changing information. However, the variability in cardiac cine data can be assumed to be limited and therefore a single training stage should be able to capture the structural information required, hence avoiding the additional computational load needed for updating the dictionary at each temporal frame. Other variations of learning methods are possible [3]. In, the trained frame consists of temporal functions that sparsely represent the temporal profile of each pixel in the dataset. Although there exist algorithmic differences in the reconstruction process, the concept of this previous work can be thought of as using independent patches that are only temporal and cover the entire dataset. Cardiac cine data is known to be redundant through space as well as through time (although it is usually the case that the temporal dimension is more redundant than the spatial dimension), which is why also exploiting spatial sparsity with a spatio-temporal dictionary could be advantageous.

B. Temporal Gradient Sparsity

An additional sparsity constraint can be imposed on the temporal finite differences (i.e. the first order temporal gradient) of the dataset. Many authors have explored TV for imposing sparsity constraints on a CSMRI reconstruction because it provides sensible sparsity levels, but also because its optimization can be extremely efficient. In many cases, TV is not the main sparsifying transform but rather an auxiliary constraint that can stabilize and correct the solution provided by the main transform. This operation considers an equally weighted combination of the pixel-wise finite differences along space and time, but this is rarely a sensible assumption in cardiac cine because spatial and temporal gradients, which make up the individual dimensions of TV, will usually have different sparsity levels.

III. APPLICATION TO CARDIAC MR IMAGES

The proposed framework was applied to cardiac MR images and evaluated its performance in two scenarios using both simulated and real cardiac MR

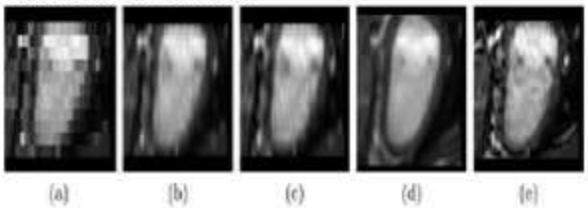
International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 12
August 2016

images. Two hundred healthy volunteers were scanned using a 1.5T Philips Achieve a system with a 32-channel cardiac coil. A single breath-hold 3D balanced steady- state free precession (b- SSFP) sequence is acquired. The final voxel size is 1.25 x 1.25 x 2 mm. The typical breath-hold time is 20 seconds. 11 good quality images were selected and used to build a synthetic data set and the remaining 189 images were used as the atlases. The LR images (1.25 x 1.25 x 10 mm) were generated from the 3D images using the operator defined [6]. In addition, 19 normal volunteers were scanned twice on the same day. A standard acquisition was performed including an axial stack of cine b-SSFP MR images in the left ventricular short axis plane. The voxel sizes for these images are 1.25 x 1.25 x 10 mm. The images were then super-resolved using the previous 200 3D images. There are three pre-processing steps which occur before applying the EM algorithm. First, the SA slices are spatially aligned to remove the inter-slice linear B-spline cubic B-spline shifting caused by respiratory motion. The inter-slice shifts between SA slices are corrected by registering SA slices to long-axis (LA) slices. Second, a region of interest (ROI) is detected using a Haar feature classifier. Finally, all atlases are intensity normalized to the spatially corrected image. During the experiments we have set our patch size to 14 x 14 x 14 mm.

TABLE I: The median and inter quartile range of PSNR for different methods from 11 synthetic cases There is significant difference between the PSNRs of interpolation methods and the proposed method (p-value < 0.05 indicated by *)

	linear	B-spline	cubic B-spline	MAPM
PSNR (dB)	19.65 (1.17)*	19.62 (1.28)*	19.9 (1.22)*	20.96 (1.1)

A. Quantitative Evalution



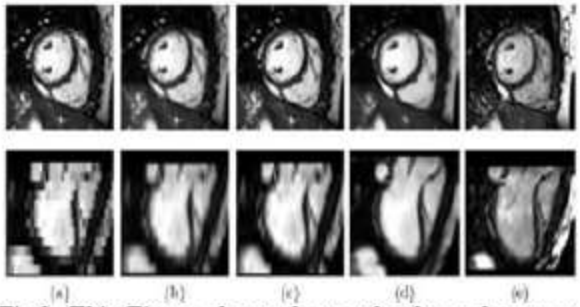
(a) (b) (c) (d) (e)

Fig.2. This Figure shows the results of the synthetic evaluation from long-axis view. (a) shows the down-sampled images; (b) shows the linear interpolation; (c) shows the cubic B-spline interpolation; (d) shows the proposed method and (e) shows the original 3D image.

In this evaluation, we compare the PSNR between the image reconstructed from the synthetic LR image and the original image. We reconstruct the SR image using linear interpolation, spline interpolation and the proposed approach. The result is shown in Tab. 1. During the down-sampling process, part of the apex and base might be missing due to the reduced field of view. This is also a common problem in SA images. It can be seen from Fig.2 that the missing parts of the apical and basal slices can be recovered. This is due to the fact that a patch is copied from the atlases instead of a single voxel. Thus, during the iterative process, the missing topology can be gradually repaired [4].

B. Reproducibility Analysis

In the second experiment, we attempt to super-resolve the SA cardiac MR images using the proposed algorithm (Fig 3). The super-resolved image has better contrast and less noise compared to 3D image of the same subject. In addition, we segment both the SA images and super resolved images using the patch-based segmentation. We calculate the mean and the standard deviation of absolute differences d between the left ventricles (LV) volume obtained from two scans of the 19 subjects. The results from SR images (dSR: 4.94 ± 4.36 ml) are more reproducible compare to results from SA images (dSA: 6.58 ± 6.76 ml).



(a) (b) (c) (d) (e)

Fig.3. This Figure shows the results from the super-resolution of the SA MR images. (a) original SA image; (b) linear interpolation; (c) cubic B-spline interpolation; (d) proposed and (e) corresponding 3D image of the same subject rigidly align to the SA MR image.

IV. EXPERIMENTS

We demonstrate our algorithm on synthetically-generated data and use it to up sample real adult and neonatal datasets. Adult data were acquired on a 1.5T Philips Achieve scanner, using a 32 channel cardiac

coil (TE/TR=1.6/3.2ms, FA=60), giving a reconstructed resolution of 1.25x1.25mm slices, 10 mm thick. SSFP acquisition with retrospective cardiac gating provides 30 frames. For the neonatal data, SSFP acquisition with retrospective ECG gating over 20 cardiac phases was used, with resulting resolution 0.5×0.5×4mm.

A. Results

In the simulation, we take a HR LA sequence of 30 frames, withholding one frame as a test image. This is down sampled by a factor of 8 in the theoretical slice-select direction using a Gaussian blur with FWHM=8. The remaining 29 frames are used for training using the same down sampling operations. We apply our algorithm using overlapping patches of size 16×16 pixels and dictionary size of 400 atoms. The high degree of patch redundancy allows successful reconstruction as shown in Fig.4, with substantial improvement over bi-cubic interpolation. Training and reconstruction took approximately 2 seconds on a 2.4GHz Intel Core i5 machine. Repeating the experiment for similar LA sequences of 15 different subjects, gave a mean squared error (MSE) of 2066 for bi-cubic interpolation compared with 856 for our patch based method. For real MRI, we acquire a SA stack plus a HR LA sequence of an orthogonal view for a given subject. The LA sequence is down sampled and used as training data to reconstruct the corresponding non-planar LR view of the SA stack to isotropic, using an up sampling factor of 8. We present visual results for two markedly different cases: adult and neonatal. In these cases, patches of sizes of 12×24 pixels were used in both training and reconstruction. The results of two example frames from each dataset are shown in Fig. 5. It can be clearly seen that in both cases the patch-based reconstruction algorithm produces sharper images than bi-cubic interpolation.

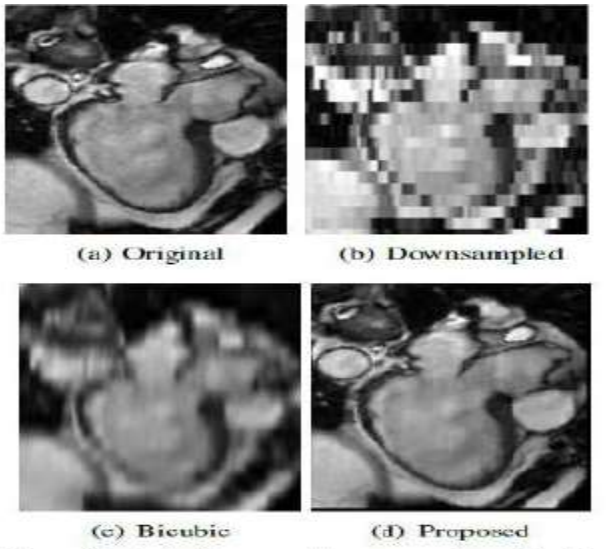


Fig.4. Synthetically-generated example Bi-cubic MSE=2683, proposed MSE=1293.

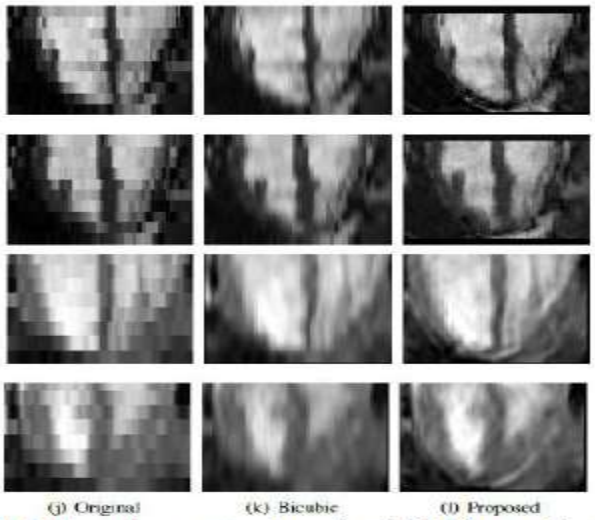


Fig.5. Example reconstruction of real MRI frames. First two rows: adult dataset. Bottom two rows: neonatal dataset.

Comparison Tables

Input	Existing Method	Proposed Method

	PSNR	CSNR	PSNR	CSNR
1	30.5293	35.2030	30.5149	35.2318
2	30.5579	35.1459	30.5137	35.2342
3	30.5621	35.1532	30.3512	35.2156
4	30.5137	35.2342	30.4656	35.3978
5	30.3512	35.3156	30.2398	35.4265

Input	Existing Method		Proposed Method	
	MSE	PSNR	MSE	PSNR
1	0.0756	59.3431	0.0741	59.4343
2	0.3558	52.6186	0.3458	52.7423
3	0.0754	59.3598	0.0759	59.3265
4	5.4797	40.7433	5.6830	40.5850
5	1.5630	46.1911	1.5510	46.2248

V. CONCLUSION

In this paper, we propose a patch-based DL model for dMRI reconstruction. The main concept is to adopt a set of temporal-dependent dictionaries with three-dimensional atoms that are adaptively trained to provide sparse representations for compressed sensing reconstruction. The sparse representation is obtained by dividing the image sequence into three-dimensional overlapping patches and sparsifying these patches using the learned dictionaries. This patch-based DL model is able to capture local structure spatially and temporally. The simulation shows encouraging results that the proposed method can preserve both spatial structures and temporal variations. Finally, the algorithm performs extrapolation as well as interpolation of the images. In the SR image, the original SA slice is a little blurred due to the fusion of multiple patches. Future work will include exploring the possibility to extend MAPM to patch-based segmentation and to exploit neighboring correspondence to preserve the original SA slice and image self-similarity to increase the robustness.

VI. REFERENCES

[1] Kanwal K. Bhatia, Anthony N. Price, Wenzhe Shi, Jo V.Hajnal, Daniel Rueckert, "Super-Resolution Reconstruct-ion of Cardiac MRI Using Coupled Dictionary Learning", IEEE 2014.

[2] J. Caballero, D. Rueckert, and J. V. Hajnal, "Dictionary learning and time sparsity in dynamic MRI," in MICCAI (1), 2012, pp. 256–263.

[3] S. U. Rahman and S. Wesarg, "Combining short-axis and long-axis cardiac MR images by applying a super resolution reconstruction algorithm," in SPIE Med. Im., 2010, p. 7623.

[4] A. Gholipour, J. A. Estroff, and S. K. Warfield, "Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI," IEEE Trans. Med. Imag., vol. 29(10), pp. 1739–1758, 2010.

[5] J. V. Manjon, P. Coupe, A. Buades, D. L. Collins, and M. Robles, "MRI super resolution using self-similarity and image priors," Int J. Biomed. Imaging, 2010.

[6] F. Rousseau, "A non-local approach for image super resolution using intermodality priors," Medical Image Analysis, vol. 14, pp. 594–605, 2010