

Cardiac Image Super-Resolution with Global Correspondence Using Multi-Atlas PatchMatch

Wenzhe Shi¹, Jose Caballero¹, Christian Ledig¹, Xiahai Zhuang², Wenjia Bai¹,
Kanwal Bhatia¹, Antonio M Simoes Monteiro de Marvao³,
Tim Dawes³, Declan O'Regan³, and Daniel Rueckert¹

¹ Biomedical Image Analysis Group, Imperial College London, UK

² Shanghai Advanced Research Institute, Chinese Academy of Sciences, China

³ Institute of Clinical Science, Imperial College London, UK

Abstract. The accurate measurement of 3D cardiac function is an important task in the analysis of cardiac magnetic resonance (MR) images. However, short-axis image acquisitions with thick slices are commonly used in clinical practice due to constraints of acquisition time, signal-to-noise ratio and patient compliance. In this situation, the estimation of a high-resolution image can provide an approximation of the underlying 3D measurements. In this paper, we develop a novel algorithm for the estimation of high-resolution cardiac MR images from single short-axis cardiac MR image stacks. First, we propose to use a novel approximate global search approach to find patch correspondence between the short-axis MR image and a set of atlases. Then, we propose an innovative super-resolution model which does not require explicit motion estimation. Finally, we build an expectation-maximization framework to optimize the model. We validate the proposed approach using images from 19 subjects with 200 atlases and show that the proposed algorithm significantly outperforms conventional interpolation such as linear or B-spline interpolation. In addition, we show that the super-resolved images can be used for the reproducible estimation of 3D cardiac functional indices.

1 Introduction

3D cardiac magnetic resonance (MR) imaging has developed rapidly during the past few years, particularly in the acquisition of 3D cine MR images [1,2]. Near isotropic 3D cardiac MR images allow reliable assessment of complex cardiac morphology. Using 3D images also allows for a more accurate and reproducible estimation of cardiac functional indices [3]. However, 3D cardiac MR imaging is not always available due to several limitations: First, 3D cardiac MR imaging often involves breath-holding for periods that are too long for many patients. In addition, it often has a low signal-to-noise ratio (SNR). Finally, advanced 3D cardiac MR imaging is not yet widely available in clinical practice and still requires substantial specialist expertise.

Image super-resolution is an active field of research in computer vision. Most super-resolution algorithms use an observation model which establishes a relationship between the high-resolution image and the observed low-resolution images. The observed low-resolution images are considered to be warped, blurred,

down-sampled and noisy versions of the original high-resolution image. One of the most common approaches to the super-resolution problem is to use the maximum likelihood (ML) or maximum a posteriori (MAP) estimation [4]. In these approaches, a distance measure between the reconstructed image and the observed images is iteratively reduced. Example-based image super-resolution [5] 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 [6] proposed a distance estimation approach which requires no explicit motion estimation by inverting the position of the patch selection operator and the resampling operator.

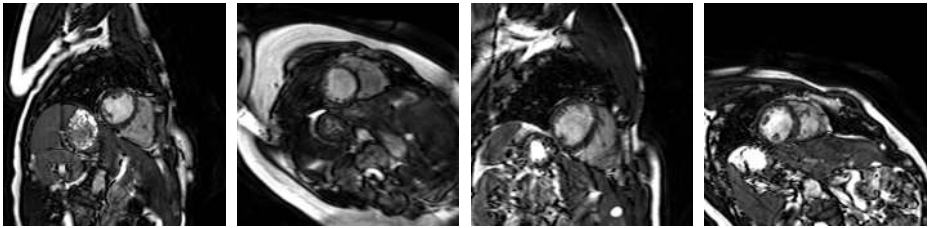


Fig. 1. Variability of heart orientation, position and shape across subjects

The idea of super-resolution has been applied in medical imaging too: Gholipo [7] 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 [8,9]. Rousseau [10,11] 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) 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 [12]. 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 [10] 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 PatchMatch algorithm [13]. Inspired by [10], we assume that information redundancy in similar patches across different subjects can be exploited. Thus, we reformulate the PatchMatch approach to find patch correspondence between a single image and an atlas database. We then use the principles in [6,7] to estimate the super-resolved image using the expectation-maximization (EM) framework.

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 PatchMatch and a closed-form solution in the observation model [6,9]. 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.

2 Methods

2.1 Multi-Atlas PatchMatch

The PatchMatch algorithm proposed by Barnes [13] finds corresponding patches across two images or regions. In contrast with the original PatchMatch algorithm, our multi-atlas PatchMatch (MAPM) finds patch correspondences \mathbf{N} between an image and a database of atlases. Given an image \mathbf{I} and an atlas database \mathbf{A} (individual atlases are denoted as \mathbf{A}_i), we would like to find for each point $\mathbf{x} = (x, y, z)$ in image \mathbf{I} a match in the atlas database \mathbf{A} , $\mathbf{N}(\mathbf{x}) = (\mathbf{p}, i)$ where $\mathbf{p} = (x', y', z')$ is the closest match in atlas \mathbf{A}_i for a given distance function D between patches. The distance function to be used during the search is independent from MAPM and can be customized to different applications.

The MAPM algorithm consists of four different steps which will be described in the following. The reader can find additional figures showing a graphical illustration of the four different steps in the supplementary material¹. The mapping \mathbf{N} can be initialized either by random assignment or by using prior information (Fig.1 in supplementary material). In our case, we assign $\mathbf{N}(\mathbf{x}) = (\mathbf{x}, R(n))$ where $R(n)$ generates a random selection uniformly between \mathbf{A}_1 and \mathbf{A}_n . After initialization, we perform an iterative process of improving the mapping \mathbf{N} using propagation and random search. During the propagation of \mathbf{N} , from point

¹ <https://www.dropbox.com/s/eoeqbviq5kqcdix/MAPdiagram.pdf>

\mathbf{p} neighbouring to point \mathbf{x} (Fig.2 in supplementary material), we attempt to improve $\mathbf{N}(\mathbf{x})$ using the known mapping of $\mathbf{N}(\mathbf{p})$ as in [13]. During the random search step, we attempt to improve $\mathbf{N}(\mathbf{x})$ by testing a sequence of candidate points at an exponentially decreasing distance from $\mathbf{N}(\mathbf{x})$. Different from [13], in our case, the atlas index i can be fixed (Fig.3 in supplementary material) or relaxed (Fig.4 in supplementary material). Each iteration of the algorithm proceeds as follows:

- for each \mathbf{x} propagation from $(x - 1, y, z)$, $(x, y - 1, z)$ and $(x, y, z - 1)$;
- for each \mathbf{x} random search with \mathbf{A}_i fixed then relaxed
- for each \mathbf{x} propagation from $(x + 1, y, z)$, $(x, y + 1, z)$ and $(x, y, z + 1)$;
- for each \mathbf{x} random search with \mathbf{A}_i fixed then relaxed

This process is performed until the sum of all distances in image \mathbf{I} converges.

2.2 Super-Resolution Model with No Explicit Motion Estimation

In the classical observation model, the SR image is reconstructed from a LR training database. The LR images are considered to be degraded versions of the SR image undergoing blurring, downsampling and the addition of noise [6,7,4]. In our case, we aim to reconstruct the SR image from a SR atlas database constrained by a single LR image.

Takeda [6] suggested that the patch selection should be applied before rather than after the downsampling in order to avoid an explicit motion estimation. Gholipour [7] proposed the following formulation designed for MR images:

$$\mathbf{I}_k^L = \mathbf{R}\mathbf{B}_k\mathbf{S}_k\mathbf{M}_k\mathbf{I}^H, \quad (1)$$

Here k denotes a slice, \mathbf{M} denotes motion operator which is no longer needed in our case, \mathbf{S}_k denotes the slice selection operator which can be replaced by patch selection operator \mathbf{P} , \mathbf{B}_k is a blurring kernel representing the point spread function (PSF) of the MR imaging signal acquisition process and \mathbf{R} is the down-sampling operator.

By combining patch redundancy [10] and the formulation proposed in [7], we propose a novel model with two terms $\Phi_{SR} = \Phi_{SR}^1 + \Phi_{SR}^2$ to reconstruct the SR image \mathbf{I} where Ω is the image domain. In this model, the first term constrains \mathbf{N} using the observed LR image so that the selected patches after downsampling operations should be as similar as possible to the LR image:

$$\Phi_{SR}^1 := \sum_{\mathbf{x} \in \Omega} w[\mathbf{x}, \mathbf{N}(\mathbf{x})] \|\mathbf{P}_\mathbf{x}\mathbf{I}^L - \mathbf{R}\mathbf{P}_{\mathbf{N}(\mathbf{x})}\mathbf{B}\mathbf{A}\|_2. \quad (2)$$

The second term constrains \mathbf{I} using \mathbf{N} and \mathbf{A} based on the fact that the reconstructed images should be as similar as possible to the selected patches:

$$\Phi_{SR}^2 := \sum_{\mathbf{x} \in \Omega} w[\mathbf{x}, \mathbf{N}(\mathbf{x})] \|\mathbf{P}_\mathbf{x}\mathbf{I} - \mathbf{P}_{\mathbf{N}(\mathbf{x})}\mathbf{A}\|_2, \quad (3)$$

Here $\mathbf{P}_{\mathbf{x}}$ selects a patch from an image with radius in mm around \mathbf{x} and $\mathbf{P}_{\mathbf{N}(\mathbf{x})}$ with $\mathbf{N}(\mathbf{x}) = (\mathbf{p}, i)$ selects a patch from \mathbf{A}_i with radius in mm around \mathbf{p} . $w[\mathbf{x}, \mathbf{N}(\mathbf{x})]$ is chosen as $\exp\left\{-\frac{D(\mathbf{N}(\mathbf{x}))^2}{2\sigma^2}\right\}$ according to [6] and controls the contribution of the selected patch to the final reconstruction. Finally, we define $D(\mathbf{N}(\mathbf{x})) = \|\mathbf{P}_{\mathbf{x}}\mathbf{I}^L - \mathbf{R}\mathbf{P}_{\mathbf{N}(\mathbf{x})}\mathbf{B}\mathbf{A}\|_2 + \|\mathbf{P}_{\mathbf{x}}\mathbf{I} - \mathbf{P}_{\mathbf{N}(\mathbf{x})}\mathbf{A}\|_2$. We blur the atlases before patch selection to save computation time.

2.3 Expectation Maximization Framework

In this subsection, we construct the whole super-resolution approach within an EM framework: In this context the atlases \mathbf{A} and the LR image \mathbf{I}^L correspond to the observed data, \mathbf{I} is the unobserved data and \mathbf{N} are the parameters. The EM algorithm is initialized by assuming \mathbf{I} to be empty and $\mathbf{N}(\mathbf{x}) = (\mathbf{x}, R(n))$. The distance between an empty patch and any patch is defined as $+\infty$.

In the M-step we optimize \mathbf{N} using the MAPM described in Sec. 2.1. Then, the weighting matrix W is updated according to the distance computed. In the E-step we estimate the SR image \mathbf{I} by optimizing the observation model Φ_{SR} . We can calculate the penalty at each patch $\mathbf{P}_{\mathbf{x}}\mathbf{I}$ independently if \mathbf{N} is fixed similar to the multi-point estimation in [9]:

$$\arg \min_{\mathbf{P}_{\mathbf{x}}\mathbf{I}} \Phi_{SR}(\mathbf{P}_{\mathbf{x}}\mathbf{I}) := \sum_{\mathbf{p} \in \Omega^P} w[\mathbf{x}, \mathbf{N}(\mathbf{p})] \|\mathbf{P}_{\mathbf{x}}\mathbf{I} - \mathbf{P}_{\mathbf{N}(\mathbf{p})}\mathbf{A}\|_2, \quad (4)$$

Here Ω^P is a neighborhood with all patches which contain point \mathbf{x} and centered at point \mathbf{p} and the distance is calculated on overlapping areas as in [9]. This leads to a closed-form solution:

$$\mathbf{P}_{\mathbf{x}}\mathbf{I} = \frac{\sum_{\mathbf{p} \in \Omega^P} w[\mathbf{x}, \mathbf{N}(\mathbf{p})] \mathbf{P}_{\mathbf{N}(\mathbf{p})}\mathbf{A}}{\sum_{\mathbf{p} \in \Omega^P} w[\mathbf{x}, \mathbf{N}(\mathbf{p})]}, \quad (5)$$

3 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 images. Two hundred healthy volunteers were scanned using a 1.5T Philips Achieva 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 $1.25 \times 1.25 \times 2$ 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 ($1.25 \times 1.25 \times 10$ mm) were generated from the 3D images using the operator defined by Eq.1. 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 image is $1.25 \times 1.25 \times 10$ mm. The images were then super-resolved using the previous 200 3D images.

Table 1. The median and interquartile 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) | 19.05 (1.17)* | 19.62 (1.28)* | 19.9 (1.22)* | 20.96 (1.1) |

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 shifting caused by respiratory motion. The inter-slice shifts between SA slices are corrected by registering SA slices to long-axis (LA) slices [14]. Second, a region of interest (ROI) is detected using a Haar feature classifier [15]. Finally, all atlases are intensity normalized [16] to the spatially corrected image. During the experiments we have set our patch size to 14 x 14 x 14 mm.

3.1 Quantitative Evaluation

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 [17] 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.

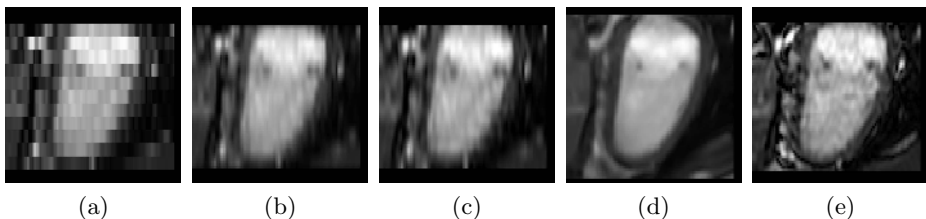


Fig. 2. This figure shows the results of the synthetic evaluation from long-axis view. (a) shows the down-sampled images; (b) shows the linear interpolation; (c) shows the cubic B-spline interpolation; (d) shows the proposed method and (e) shows the original 3D image.

3.2 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

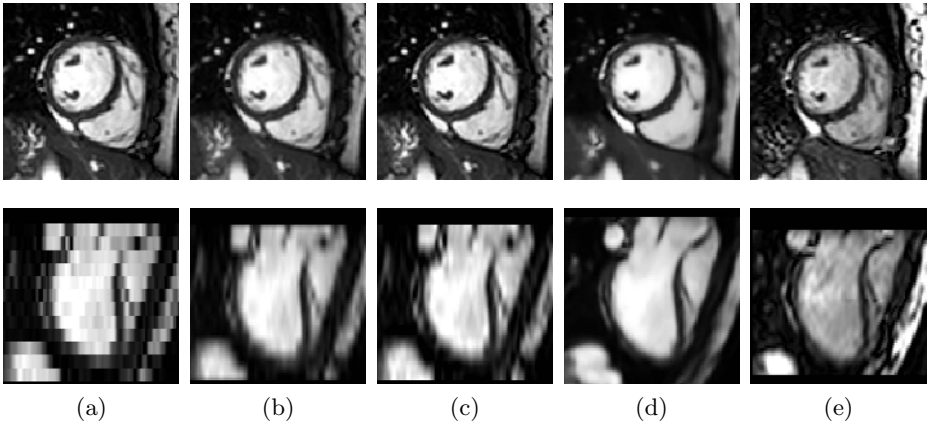


Fig. 3. This figure shows the results from the super-resolution of the SA MR images. (a) original SA image; (b) linear interpolation; (c) cubic B-spline interpolation; (d) proposed and (e) corresponding 3D image of the same subject rigidly align to the SA MR image.

segmentation [8]. We calculate the mean and the standard deviation of absolute differences \mathbf{d} between the left ventricle (LV) volume obtained from two scans of the 19 subjects. The results from SR images ($\mathbf{d}_{SR} : 4.94 \pm 4.36 \text{ ml}$) are more reproducible compare to results from SA images ($\mathbf{d}_{SA} : 6.58 \pm 6.76 \text{ ml}$).

4 Conclusion

In this paper, we developed a MAPM based framework for medical images. We have shown that our framework works well in cases where hundreds of atlases are used as the training database to super-resolve one LR image. In addition, there is no need for any spatial alignment with atlases. The computational time is 2 hours on average per case and does not change with an increasing number of atlases. Finally, the algorithm performs extrapolation as well as interpolation of the images. This is desirable in cardiac images where apical and basal slices may be missing due to limited field of view and thick slices. In the SR image, the original SA slice is a little blurred due to the fusion of multiple patches [6]. This is a trade-off for improved through-plane resolution. Future work will include exploring the possibility to extend MAPM to patch-based segmentation and to exploit neighboring correspondence [18] to preserve the original SA slice and image self similarity [19] to increase the robustness.

References

1. Uribe, S., Muthurangu, V., Boubertakh, R., Schaeffter, T., Razavi, R., Hill, D., Hansen, M.: Whole-heart cine MRI using real-time respiratory self-gating. *Magnetic Resonance in Medicine* 57(3), 606–613 (2007)

2. Davarpanah, A., Chen, Y., Kino, A., Farrelly, C., Keeling, A., Sheehan, J., Ragin, A., Weale, P., Zuehlsdorff, S., Carr, J.: Accelerated two- and three-dimensional cine MR imaging of the heart by using a 32-channel coil. *Radiology* 254(1), 98–108 (2010)
3. Sørensen, T., Körperich, H., Greil, G., Eichhorn, J., Barth, P., Meyer, H., Pedersen, E., Beerbaum, P.: Operator-independent isotropic three-dimensional magnetic resonance imaging for morphology in congenital heart disease. *Circulation* 110(2), 163–169 (2004)
4. Tian, J., Ma, K.: A survey on super-resolution imaging. *Signal, Image and Video Processing* 5(3), 329–342 (2011)
5. Freeman, W., Jones, T., Pasztor, E.: Example-based super-resolution. *IEEE Computer Graphics and Applications* 22(2), 56–65 (2002)
6. Takeda, H., Milanfar, P., Protter, M., Elad, M.: Super-resolution without explicit subpixel motion estimation. *IEEE Transactions on Image Processing* 18(9), 1958–1975 (2009)
7. Gholipour, A., Estroff, J.A., Warfield, S.K.: Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI. *IEEE Transactions on Medical Imaging* 29(10), 1739–1758 (2010)
8. Coupé, P., Manjón, J., Fonov, V., Pruessner, J., Robles, M., Collins, D.: Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. *Neuroimage* 54(2), 940–954 (2011)
9. Rousseau, F., Habas, P.A., Studholme, C.: A supervised patch-based approach for human brain labeling. *IEEE Transactions on Medical Imaging* 30(10), 1852–1862 (2011)
10. Rousseau, F., Kim, K., Studholme, C.: A groupwise super-resolution approach: application to brain MRI. In: *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 860–863 (2010)
11. Rousseau, F.: A non-local approach for image super-resolution using intermodality priors. *Medical Image Analysis* 14(4), 594 (2010)
12. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(9), 1167–1183 (2002)
13. Barnes, C., Shechtman, E., Goldman, D.B., Finkelstein, A.: The generalized patch-match correspondence algorithm. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) *ECCV 2010, Part III*. LNCS, vol. 6313, pp. 29–43. Springer, Heidelberg (2010)
14. Lötjönen, J., Pollari, M., Kivistö, S., Lauerma, K.: Correction of movement artifacts from 4-D cardiac short- and long-axis MR data. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) *MICCAI 2004*. LNCS, vol. 3217, pp. 405–412. Springer, Heidelberg (2004)
15. Viola, P., Jones, M.: Robust real-time object detection. *International Journal of Computer Vision* 57(2), 137–154 (2002)
16. Nyúl, L., Udupa, J., et al.: On standardizing the MR image intensity scale. *Magnetic Resonance in Medicine* 42(6), 1072 (1999)
17. Unser, M., Aldroubi, A., Eden, M.: B-spline signal processing. i. Theory. *IEEE Transactions on Signal Processing* 41(2), 821–833 (1993)
18. Ólafsdóttir, H., Pedersen, H., Hansen, M.S., Larsson, H., Larsen, R.: Improving image registration by correspondence interpolation. In: *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1524–1527. IEEE (2011)
19. Manjón, J.V., Coupé, P., Buades, A., Fonov, V., Louis Collins, D., Robles, M.: Non-local MRI upsampling. *Medical Image Analysis* 14(6), 784–792 (2010)