On Single-Image Super-Resolution in 3D Brain Magnetic Resonance Imaging

Farah Bazzi[1]*, Muriel Mescam*[1]*, Adrian Basarab*[2]*, Denis Kouamé* [2]*,*

[1] University of Toulouse 3 - Paul Sabatier, CerCo, CNRS UMR5549, Toulouse, France
[2] University of Toulouse 3 - Paul Sabatier, IRIT, CNRS UMR 5505, INP-ENSEEIHT, Toulouse, France
{farah.bazzi,muriel.mescam}@cnrs.fr , {basarab,kouame}@irit.fr,

ABSTRACT

The objective of this work is to apply 3D super resolution (SR) techniques to brain magnetic resonance (MR) image restoration. Two 3D SR methods are considered following different trends: one recently proposed tensor-based approach and one inverse problem algorithm based on total variation and low rank regularization. The evaluation of their effectiveness is assessed through the segmentation of brain compartments: gray matter, white matter and cerebrospinal fluid. The two algorithms are qualitatively and quantitatively evaluated on simulated images with ground truth available and on experimental data. The originality of this work is to consider the SR methods as an initial step towards the final segmentation task. The results show the ability of both methods to overcome the loss of spatial resolution and to facilitate the segmentation of brain structures with improved accuracy compared to native low-resolution MR images. Both algorithms achieved almost equivalent results with a highly reduced computational time cost for the tensor-based approach.

***Index Terms*—** Single image super-resolution, structural MRI, segmentation, tensors, total variation regularization, low rank.

1. INTRODUCTION

Cerebral aging is a complex process where severe morphological and structural changes in the brain occur causing functional and network disruptions that can lead to many disorders such as epilepsy, Parkinson or Alzheimer's diseases. The assessment of these changes provides an important tool for following the development or the regression of brain-related diseases [1]. Magnetic resonance imaging (MRI), in particular T1-weighted scans, is well suited for structural studies of brain changes since it provides high soft tissue contrast and allows multiple acquisitions without potential hazards [2]. Due to practical and ethical limitations of using human beings, nonhuman primates such as marmosets (*Callithrix jacchus*), which present more neuroanatomical similarities with the human brain than rodent models and yet a short life expectancy (around 10 years), are used in longitudinal studies of brain aging [3]. However, imaging small brains in a 3T MRI platform dedicated to humans is a challenging task because the spatial resolution and the contrast obtained are insufficient compared to the size of the anatomical structures observed. In the absence of a higher field MRI scanner for generating high quality images, it becomes crucial to develop appropriate post-processing methods that enhance the resolution of preclinical images and allow the analysis of morphological changes.

Imaging beyond the resolution of an imaging system is referred to as super-resolution (SR), which increases the spatial resolution by extracting information from a set of low resolution images that may have been translated, blurred, rotated or scaled. When such multiple frames are not available, another class of methods exists, i.e., single-image SR that restores a high-resolution (HR) image based on an image formation model and a single low-resolution (LR) image [4]. In this study, we aim at evaluating the impact of enhancing the resolution of MR volumes using single-image SR on the robustness of the segmentation of the gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). This segmentation step is crucial towards a reliable analysis of morphometry and highlighting anatomical cerebral aging markers. In this study, we consider two single image SR approaches that solve the SR problem within two different frameworks. On one hand, we evaluate a recent Tensor-Factorization based (TF) method [5], proposed by our group, that has been validated and proven computationally-efficient on computed tomography images. On the other hand, we consider a standard approach for SR by solving the associated inverse problem based on low rank and total variation regularization [6], originally proposed for MR imaging. Both methods use the same image formation model, that relates the HR image to be estimated to the observed LR image through blurring and down sampling operators and additive white Gaussian noise. However, the TF method avoids the unfolding of the volume into 2D matrices which usually results in the loss of information regarding the locality of pixels [5].

The results are evaluated from a brain segmentation perspective, using a state of the art brain segmentation method [7] applied to native LR and estimated HR volumes.

The remainder of the paper is organized as follows. Section 2 summarizes the SR approaches and the segmentation method considered in this study. Section 3 provides the simulated and experimental results. Finally, Section 4 draws the conclusions and perspectives.

2. METHODS

2.1. MR image formation model

MRI is a non-invasive imaging technique used to visualize internal body organs. MRI signal results from the relaxation of hydrogen spins in the body after their excitation with an external magnetic field. Similar to any other imaging modality, the loss of spatial resolution in the acquisition process can be expressed through a linear model that relates the observations (measurements) to the HR image to be estimated. Moreover, in MRI, data acquisition is performed in the k-space (Fourier domain). This requires the acquisition matrix to be incorporated into the image formation model. Finally, the image formation model commonly used in MRI is

$$\mathbf{g} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{x} = [x_1, ..., x_N]^T \in \mathcal{R}^N$ stands for the non-observable HR image, $\mathbf{g} = [g_1, ..., g_M]^T \in \mathcal{C}^M$ are the collected data in the k-space, $\mathbf{A} \in \mathcal{C}^{M \times N}$ is the system matrix and $\mathbf{n} \in \mathcal{C}^M$ is an additive, zero-mean white Gaussian noise. All the images are expressed in

their standard vectorized version in a lexicographic order. As explained in the introduction of this paper, our goal is to evaluate post-processing SR methods in brain MRI. For this reason, the starting point of our study are reconstructed LR MR volumes. Thus, the model in (1) will be further simplified in the following sections by considering only the operators accounting for spatial resolution loss and not the system acquisition geometry.

To invert the direct model in (1), a common way is to express the estimation of $\mathbf{f}$ as the minimization of a cost function composed by a data fidelity term and a regularization term aiming at stabilizing the solution:

$$\min_{\mathbf{x}} \|\mathbf{g} - \mathbf{A}\mathbf{x}\|_2^2 + \beta \mathbf{R}(\mathbf{x}) \tag{2}$$

where β is a hyper parameter weighting the data fidelity and regularization terms and $\mathbf{R}$ is a function incorporating prior knowledge about the HR image. In the two following subsections we provide basic details about the two SR approaches considered, both in terms of forward model and regularizer employed.

2.2. 3D Super-Resolution using Tensor Factorization

The TF method, recently intreduced in [5], is based on the traditional image degradation model assuming that the LR image can be expressed as a noisy, blurred and decimated version of the HR image. Within this algorithm, 3D images are associated to tensors of order 3, resulting into the following degradation model:

$$vec(\boldsymbol{Y}) = DH\, vec(\boldsymbol{X}) + vec(\mathbf{N}) \tag{3}$$

where vec() vectorizes the elements of the 3D tensor in lexicographical order. $\boldsymbol{Y} \in \mathbb{R}^{I/r \times J/r \times K/r}$ and $\boldsymbol{X} \in \mathbb{R}^{I \times J \times K}$ are the LR and HR images respectively, $\mathbf{H} \in \mathbb{R}^{IJK \times IJK}$ is the block-circulant version of the 3D point spread function (PSF), $\mathbf{N}$ is an independent identically distributed (IID) additive white Gaussian noise (AWGN), I, J and K are the 3D volume dimensions and *r* is the decimation rate accounting for the voxel resolution loss. Based on the hypothesis of separable PSF (valid in the present study given the choice of Gaussian PSF), $\mathbf{H}$ can be decomposed in three block circulant matrices with circulant blocks (BCCB), $H_1 \in \mathbb{R}^{I \times I}$, $H_2 \in \mathbb{R}^{J \times J}$ and $H_3 \in \mathbb{R}^{K \times K}$. Similarly, the down sampling operators for the three dimensions can be given as $D_1 \in \mathbb{R}^{I/r \times I}$, $D_2 \in \mathbb{R}^{J/r \times J}$ and $D_3 \in \mathbb{R}^{K/r \times K}$.

Using the canonical polyadic decomposition of $\boldsymbol{X}$ with $\bar{U} = \{U^1 \in \mathbb{R}^{I \times F}, U^2 \in \mathbb{R}^{J \times F}, U^3 \in \mathbb{R}^{K \times F}\}$, where F is the rank of the tensor, (3) can be rewritten as the mode-*n* product of the separated kernels:

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{X} \times_1 D_1H_1 \times_2 D_2H_2 \times_3 D_3H_3 + \mathbf{N} \\ &= [[D_1H_1U^1, D_2H_2U^2, D_3H_3U^3]] + \mathbf{N} \end{aligned} \tag{4}$$

Thus the reconstruction of the HR image from the LR image requires to find the set of matrices $\bar{U}$ by solving the following minimization problem:

$$\min_{\bar{U}} \|\boldsymbol{Y} - [[D_1H_1U^1, D_2H_2U^2, D_3H_3U^3]]\|_F^2 \tag{5}$$

Since the minimization problem posed in (5) is NP-hard, U^1, U^2 and U^3 are minimized sequentially as follows:

$$\begin{aligned} &\min_{U^1} \tfrac{1}{2}\|\mathbf{Y}^{(1)} - D_1H_1U^1(D_3H_3U^3 \odot D_2H_2U^2)^T\|_F^2 \\ &\min_{U^2} \tfrac{1}{2}\|\mathbf{Y}^{(2)} - D_2H_2U^2(D_3H_3U^3 \odot D_1H_1U^1)^T\|_F^2 \\ &\min_{U^3} \tfrac{1}{2}\|\mathbf{Y}^{(3)} - D_3H_3U^3(D_2H_2U^2 \odot D_1H_1U^1)^T\|_F^2 \end{aligned} \tag{6}$$

The solution of the three minimization problems in (6) is obtained using the least-square estimator with a Tikhonov regularization. The solution is further computed using the Moore-Penrose pseudo-inverse$^+$:

$$\begin{aligned} U^1 &= (D_1H_1)^+\mathbf{Y}^{(1)}(D_3H_3U^3 \odot D_2H_2U^2)^{+T} \\ U^2 &= (D_2H_2)^+\mathbf{Y}^{(2)}(D_3H_3U^3 \odot D_1H_1U^1)^{+T} \\ U^3 &= (D_3H_3)^+\mathbf{Y}^{(3)}(D_2H_2U^2 \odot D_1H_1U^1)^{+T} \end{aligned} \tag{7}$$

Unlike the conventional unfolding performed in 3D reconstruction algorithms such as in [6], this method unfolds the tensor sequentially, in each direction, thus preserving the local 3D information.

2.3. 3D super-resolution with low-rank and total variation

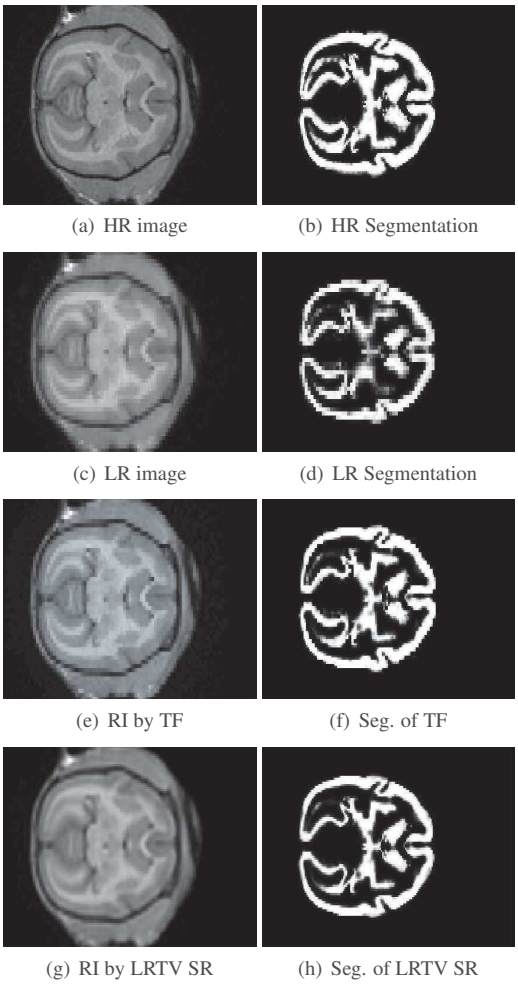
This algorithm was originally proposed in [6] for 3D MRI super-resolution. It exploits the same model as (3). In order to invert this 3D forward model, the combination of two regularization terms was shown to be particulary efficient for MRI SR in [6]. In particular, in addition to 3D total variation that provides local regularization, a low rank assumption was used to account for global prior information in the HR image recovery process. Consequently, the SR problem was expressed as the minimization of the following cost function:

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{y} - \mathbf{D}\mathbf{H}\mathbf{x}\|_2^2 + \lambda_{Rank} Rank(\mathbf{x}) + \lambda_{TV} TV(\mathbf{x}), \tag{8}$$

where λ_{Rank} and λ_{TV} are two hyperparameters tuned manually to their best values. The minimization problem (8) is further divided into three sub problems that are solved iteratively in an alternating direction method of multipliers (ADMM) framework and the total variation (TV) is implemented using gradient descent thus rendering this method computationally expensive [6]. This method will be referred to as LRTV hereafter.

2.4. Brain MR image segmentation

The main objective of this study is to evaluate the segmentation of WM and GM brain regions from MR images, after the application of the SR algorithms discussed above. In our work, segmentation was done using Structural Parametric Mapping (SPM) [8] software that implements the expectation maximization (EM) segmentation methods based on a Bayesian classifier framework. We have previously developed a module that can be integrated into 3D Slicer [9] for the semi-automatic registration of a marmoset brain template on any marmoset brain MR image. The registration method employed relies on a template built from a single fully segmented marmoset brain image, which was transported onto the Karcher mean of 13 adult marmoset brain images using a diffeomorphic strategy that fully preserves the brain topology. MR images are then segmented into GM, WM and CSF compartments using the tissue probability maps resulting from the registration process. Segmentation is done using SPM software based on image intensities and prior information [7].



(a) HR image (b) HR Segmentation

(c) LR image (d) LR Segmentation

(e) RI by TF (f) Seg. of TF

(g) RI by LRTV SR (h) Seg. of LRTV SR

Fig. 1. Simulation results from the data set representing one axial slice and their corresponding GM segmentation results: (a,b) reference HR image, (c,d) LR image obtained by blurring and downsampling the HR image in (a), (e,f) super-resolved image obtained with TF, (g,h) super-resolved image obtained with LRTV.

3. RESULTS AND DISCUSSION

A dataset of T_1-weighted MR images was used in this study. Acquisitions on marmosets [1,2] were done on a 3T MRI platform using a gradient echo sequence with $T_R = 10.5msec$, $T_E = 4.7msec$ and flip angle of 8^o. The size of the dataset was $98 \times 182 \times 113$ with voxel size of $0.35 \times 0.35 \times 0.35\ mm^3$.

Since SR algorithms require the previous knowledge or estimation of the blurring point spread function (PSF) we have followed the conclusion drawn by [10] and [11] by approximating the PSF as a Gaussian function with its full width at half maximum (FWHM) being the selected slice width. The standard deviation is then computed by:

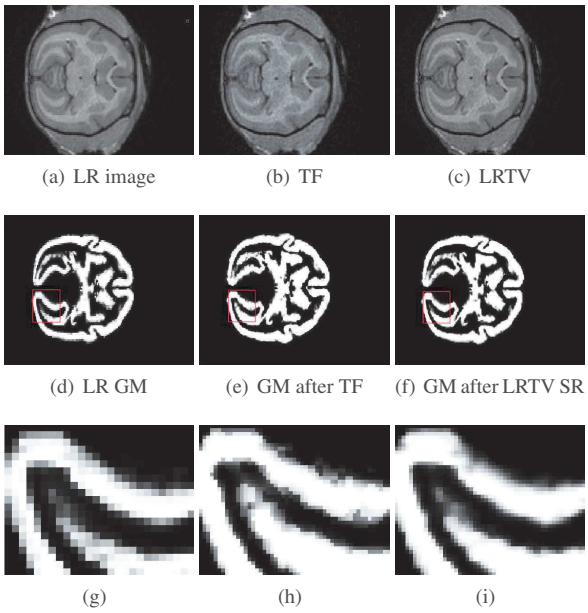
$$\sigma = \frac{\text{FWHM}}{4\sqrt{2ln2}} \tag{9}$$

The parameters used in the LRTV algorithm were kept as suggested by the authors in [6] given that the original paper already addressed an MR application. Note that TF method only requires one hyperparameter, the tensor rank F which was set to its best value, while LRTV requires the right tuning of the weights of each regularization term and the parameters of the ADMM and gradient descent optimizers. The SR algorithms were applied on the dataset following two setups, called simulation (with ground truth available) and experimental hereafter. For the simulation study, the quality of the segmentation results was quantified by two metrics: the structural similarity index (SSIM) and the DICE coefficient. For two images A and B, SSIM and DICE are defined as:

$$\text{SSIM} = \frac{(2\mu_A\mu_B + C_1)(2\sigma_{AB} + C_2)}{(\mu_A^2 + \mu_B^2 + C_1)(\sigma_A^2 + \sigma_B^2 + C_2)}$$

$$\text{DICE} = 2\frac{A \cap B}{A + B}$$

where μ_A, μ_B are the local means, σ_A, σ_B are the local standard deviations, and σ_{AB} is the cross–covariance. SSIM assesses the visual impact of three characteristics of an image: luminance, contrast and structure. DICE coefficient measures the accuracy of the overlapping of two binary images and is given as twice the number of elements common in A and B ($A \cap B$) divided by the sum of the number of elements in each (A+B). For both coefficients, a value of 1 represents a perfect similarity/overlap . Experimental results were



(a) LR image (b) TF (c) LRTV

(d) LR GM (e) GM after TF (f) GM after LRTV SR

(g) (h) (i)

Fig. 2. Experimental results showing one axial slice of the data set: (a) the observed LR image, (b) the super-resolved image using TF and (c) the super-resolved image using LRTV. Note that the images in (b) and (c) have 4 times more pixels than the LR in (a). GM segmentations from (d) LR image, (e) TF image and (f) LRTV image. For a better visualization purpose, (g), (h) and (i) represent zooms from the segmented images in (d), (e) and (f).

[1]Governmental authorization from the MENESR (project #05215.03) was given for the experimental procedures involving animal models described in this paper.

[2]We thank Caroline Fonta and the MRI platform of INSERM TONIC UMR1214 for their help with image acquisition.

only evaluated qualitatively because of non availability of the ground truth (reference) HR image.

3.1. Simulation results

The simulated data was computed by considering the experimental MR scans as HR images. Blurring with a Gaussian kernel and down sampling by a factor of $r = 2$ in each spatial dimension were applied on the MR images resulting into LR volumes that were used as input for the SR methods. The super-resolved images provided by TF and LRTV methods were compared to the reference HR image. Moreover, the effectiveness of SR on MR images was tested by segmenting the brain volumes (GM, WM and CSF) and comparing them to the segmentation of the ground truth HR image. For illustration purposes, we present in Fig. 1 the results of the algorithms for an axial slice of the dataset and the corresponding GM segmentations. Table 1 summarizes the quantitative results, i.e. the average similarity index for 113 slices of the dataset between the ground truth and the resized LR image using cubic interpolation (LR, GT), the recovered images (RI) by TF (GT, TF) and RI by LRTV $(GT, LRTV)$. In addition, it shows the average SSIM for the GM segmentation from GT and recovered TF image (GM_{GT}, GM_{TF}), and the GM segmentation from GT and recovered LRTV image (GM_{GT}, GM_{LRTV}).

Table 1. Average SSIM values for simulated results.

	LR,GT	TF,GT	LRTV,GT	GM_{GT} , GM_{TF}	GM_{GT} , GM_{LRTV}
Avg SSIM	0.15	0.51	0.56	0.95	0.95

These results show that the volumes recovered by TF method and LRTV method provide almost equivelant results in comparison to the ground truth. Morover, SR algorithms show important enhancement in the image when compared to the resized LR image. These results are confimed by the average DICE coefficients of the 113 slices (see Table 2) computed between the GM segmentations obtained from the GT, and from the TF and LRTV images respectively. The numerical results are confirmed by the visual inspection of images in Fig. 1. The effectiveness of the proposed TF method is accompanied by the advantage of being computationally more efficient than LRTV. The computation time to recover 113 slices was roughly 16 minutes for LRTV and 2 minutes for TF method with standard Matlab (2017b) implementation on a desktop computer.

Table 2. DICE coefficients computed from GM segmentations.

	GM_{GT}, GM_{LR}	GM_{GT}, GM_{TF}	GM_{GT}, GM_{LRTV}
DICE Coefficient	0.92	0.93	0.91

3.2. Experimental results

Herein, we applied the SR methods directly to the MR images, thus considered as the LR images. The performance of the TF algorithm for SR was compared to the LRTV algorithm considered as a benchmark. An example of the results is shown in Fig.2 for one axial slice. We performed the segmentation similarly to the previous section. However, in the absence of ground truth HR images, we only analyze the results qualitatively. Qualitative analysis of segmentation results confirm that SR has provided regions with borders that are better defined, with more confident mapping of the tissues and less partial volume effect, compared to the LR segmentation. This may lead to an underestimation of the gray matter, and inversely to an overestimation of the white matter, in this region. The texture of the GM region also appears smoother in Fig.2 (i) compared to Fig.2 (h). The LRTV approach seems more congruent with the original LR image. Indeed, the TF method may overestimate the GM (and underestimate the WM).

4. CONCLUSION

We investigated the effectiveness of single-image SR in MRI. A 3D fast SR approach based on the tensor factorization of the image originally validated on CT images was evaluated and compared to the 3D low-rank SR with TV regularization approach proposed for MR images. As expected, the TF method is up to 8 times faster than the other approach. The visual inspection of a zoomed region of the GM segmentation in the experimental results show that LRTV provides thinner tissue regions and more partial volume effect and that TF approach tends to overestimate the GM and underestimate the WM. However, both SR methods reduce the partial volume effect and thus improve the segmentation. In the future, obtaining manual segmentation by experts can present a ground truth reference to which we compare our results. Moreover, we also prospect to get ground truth images acquired using higher magnetic fields (ex: 7T scanner) allowing us to compare the results of the SR algorithms and hence enhance their performance by for example optimizing the choice of their parameters. This data could also open the path to machine learning SR algorithms in this particular application. The next step after confirming the results will be to study the cerebral aging markers by calculating the cortical thickness and the volumes of the brain structures in a longitudinal study over the marmoset life time or for different marmosets at different ages.

5. REFERENCES

[1] K. A. Jellinger and J. Attems, "Neuropathological approaches to cerebral aging and neuroplasticity," *Dialogues in clinical neuroscience*, vol. 15, no. 1, p. 29, 2013.

[2] R. L. Hendren, I. D. Backer, and G. J. Pandina, "Review of neuroimaging studies of child and adolescent psychiatric disorders from the past 10 years," *Journal of the American Academy of Child Adolescent Psychiatry*, vol. 39, no. 7, pp. 815–828, 2000.

[3] S. D. Tardif, K. G. Mansfield, R. Ratnam, C. N. Ross, and T. E. Ziegler, "The marmoset as a model of aging and age-related diseases," *ILAR journal*, vol. 52, no. 1, pp. 54–65, 2011.

[4] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE signal processing magazine*, vol. 20, no. 3, pp. 21–36, 2003.

[5] J. Hatvani, A. Basarab, J.-Y. Tourneret, M. Gyöngy, and D. Kouamé, "A tensor factorization method for 3d super-resolution with application to dental ct," *IEEE Transactions on Medical Imaging*, 2018.

[6] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "Lrtv: Mr image super-resolution with low-rank and total variation regularizations," *IEEE transactions on medical imaging*, vol. 34, no. 12, pp. 2456–2466, 2015.

[7] D. Mietchen and C. Gaser, "Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution," *Frontiers in neuroinformatics*, vol. 3, p. 25, 2009.

[8] J. Ashbruner and K. J. Friston, "Unified segmentation," *Neuroimage*, vol. 26, no. 3, pp. 839–851, 2005.

[9] K. M. Pohl, J. Fisher, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "A bayesian model for joint segmentation and registration," *Neuroimage*, vol. 31, no. 1, pp. 228–239, 2006.

[10] E. Carmi, S. Liu, N. Alon, A. Fiat, and DanielFiat, "Resolution enhancement in mri," *Magnetic Resonance Imaging*, 2006.

[11] Greenspan, Hayit, G. Oz, N. Kiryati, and S. Peled, "Mri inter-slice reconstruction using super-resolution," *Magnetic Resonance Imaging*, 2002.