Super-Resolution and Self-Similarity in Magnetic Resonance Imaging

Super-Resolutie en Zelf-Similariteit in Magnetische Resonantie Beeldvorming

Proefschrift

ter verkrijging van de graad van doctor aan de Erasmus Universiteit Rotterdam op gezag van de rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 31 januari 2014 om 9.30 uur

door

Esben Plenge

geboren te Aalborg, Denemarken.

Promotiecommissie

Promotor	:	Prof.dr. W.J. Niessen
Overige leden	:	Prof.dr. G.P. Krestin Prof.dr.ir. B.P.F. Lelieveldt Dr. L. van der Weerd
Copromotor	:	Dr.ir. E. Meijering

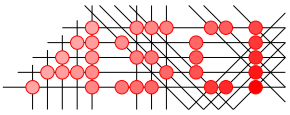
Super-Resolution and Self-Similarity in Magnetic Resonance Imaging

Esben Plenge

ISBN: 978-90-6464-747-5

The work in this thesis was conducted at the Departments of Radiology and Medical Informatics at the Erasmus MC, University Medical Center, Rotterdam, The Netherlands, and at the ASCI graduate school.

ASCI dissertation series number 296.



The publication of this thesis was financially supported by the department of Radiology, Erasmus MC and the ASCI graduate school.

Printed by GVO Printers & Designers B.V.

Layout and cover design: Esben Plenge and his facebook friends

Contents

Chapter 1

Introduction

This thesis is about super-resolution reconstruction (SRR) and self-similarity in MRI. These are two overlapping fields of research and in the studies described here, one has naturally lead to the other. From investigating basic properties of conventional approaches to SRR in MRI and applying these methods to specific research problems, we saw a potential improvement to SRR in MRI by employing the self-similarity of the images. Self-similarity is a versatile methodology, and beside using it for SRR, we have performed a thorough investigation of its application to voxelwise classification in MRI. In this introductory chapter, we will briefly give some background on SRR and self-similarity in MRI and introduce the five studies included in the thesis.

1.1 MRI and image quality

Magnetic resonance imaging (MRI) is one of the main non-invasive imaging modalities used in clinical practice and in pre-clinical studies. Its widespread application is the result of a number of factors: it provides volumetric information, its contrast between soft tissues is unsurpassed, the resolution of MRI is among the highest in the medical imaging field, and the magnetic field strengths used in clinical MRI are not associated with significant health risks, as opposed to for example the X-rays of CT. Furthermore, MRI is a very versatile imaging technique, enabling both the assessment of anatomy and function.

An MRI image is created by placing a subject in a magnetic field [35]. So-called magnetic field gradients are applied along the three axes of the space inside the scanner. The effect of each gradient is a gradual change in the magnetic field strength along the gradient's direction. The protons of the hydrogen atoms in the body are sensitive to the magnetic field and will rotate, or *precess*, at frequencies determined by the field strength. Under the application of a radio frequency (RF) pulse, the precessing protons will generate an electromagnetic signal that can be recorded by RF coils (antennas) in the scanner. The strength of the electromagnetic signal from the subject body is determined by the local proton density (that is, the relative concentration of water molecules). The electromagnetic signal thus encodes the local proton density in its phase and frequency components. This is called Fourier encoding and the process produces a map of the image in the spatial frequency domain. In MRI, this domain is known as the *k-space*.

The acquisition of an MR image thus involves an interplay between physical properties of the precessing protons, the magnetic gradient field strengths, and the way the electromagnetic signal from the protons is sampled. The signal sampling strategy, which is constrained by the physical limitations of the system, determines the acquisition time and basic image properties such as contrast, signal-to-noise ratio (SNR), and resolution. Developing and adopting these strategies is known as pulse sequence design. This is a major sub-field of MRI research and we shall

not delve into the subject here, but just mention one of the fundamental choices in sampling strategy: An MRI of a 3D subject can be acquired either as a so-called *true* 3D image volume, or as a stack of 2D slices. In the former, the signal of the entire 3D volume is sampled simultaneously, in the latter, the signal is sampled one slice at a time. Whether one or the other approach is chosen depends on many factors, including the gradient strength and slew rate, requirements to contrast and SNR, the feasible acquisition time, whether motion is present, and more. In some settings, true 3D imaging will be applied, in others, 2D slice-stacks.

One important disadvantage of 2D imaging is that the slices usually end up being thicker than the in-plane size of the voxels. This effect is known as *anisotropy* and it hampers the localization of structures and events along the slice selection direction, sometimes severely (see Fig. 1.1). It is not uncommon to have anisotropy factors (the ratio between the slice thickness and the in-plane pixel size) larger than 5, owing to requirements to contrast, SNR, and acquisition time.

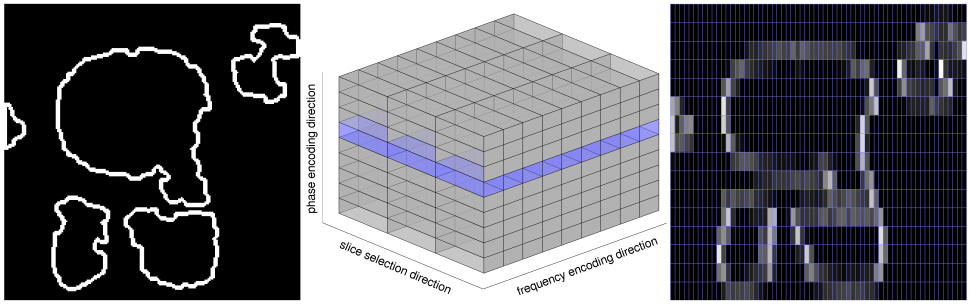


Figure 1.1: *Illustration of the anisotropy of a 2D slice-stack. Left, a high-resolution cross section of some structure is seen. In the center, an MRI slice-stack is illustrated. The purple/dark band represents the cross-section shown on the left. Right, the cross-section is shown after being sampled on an anisotropic grid.*

1.2 Super-resolution in MRI

Resolution is a fundamental property of any image, from any device. It is defined as the smallest distance between two objects at which we can still distinguish them [16], or in more technical terms, as the width of the point spread function (PSF) of the imaging process. Super-resolution reconstruction (SRR) is the general name for a group of methods that aim to recover high-frequency image content that is lost due to wide PSFs. In SRR, a set of low-resolution (LR) images, that each differ with minor shifts and/or rotations in their field of view, are combined into a

high-resolution (HR) image. This is done utilizing a model of the image acquisition process (see Fig. 1.2).

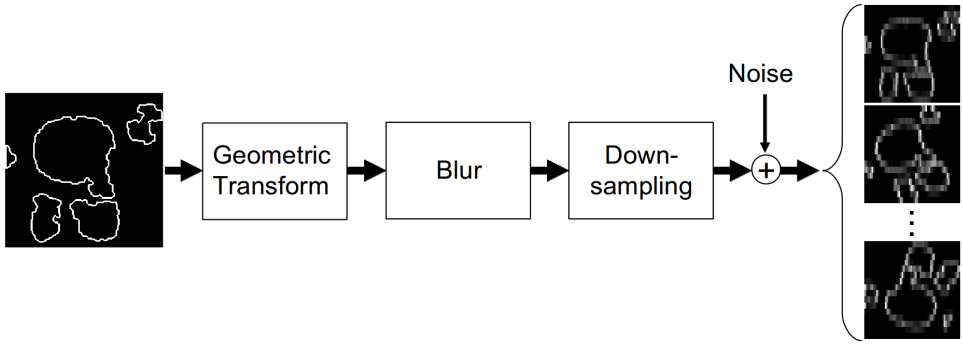


Figure 1.2: *Illustration of the model of the imaging process conventionally applied in SRR. In SRR, the aim is to invert the process, thus going from the LR images to the HR one.*

In MRI, SRR cannot recover high-frequency content in-plane nor in true 3D volumes. This is because the Fourier encoding of the slices/volumes completely excludes any content outside the range of encoding frequencies. In the slice selection direction, however, aliased frequencies may be present and possibly recovered. SRR in MRI is therefore primarily applied to generate isotropic 3D volumes from multiple 2D slice-stacks.

When acquiring MRI slice-stacks for SRR, the aim should be to cover *k*-space as well as possible. A much applied acquisition strategy in the SRR MRI literature simply shifts the acquisitions along the slice selection direction. While this, intuitively, may seem to compensate for partial volume effects in thick slices, it amounts to oversampling the central part of *k*-space in the slice selection direction. The *k*-space can be covered much more effectively by rotating the slice selection direction between each of the acquisitions (see Fig. 1.3). This acquisition strategy was introduced in [83] and is adopted for most SRR acquisitions in the thesis.

1.3 SRR as an inverse problem

The problem of recovering, or *reconstructing*, the HR image, given the LR images and a model of the imaging process, is called an *inverse problem*. In the context of image reconstruction, inverse problems are, in general, ill-posed. This means, that the number of constraints imposed by the LR acquisitions (i.e. the LR voxels) is lower than the number of parameters we aim to recover (i.e. the HR voxels). To make the problem well-posed, regularization is applied to provide additional constraints on

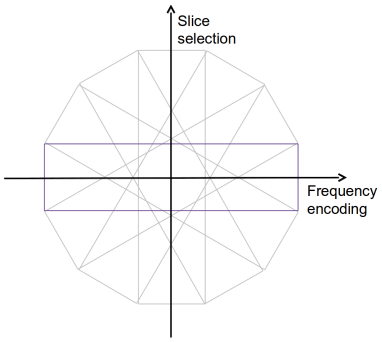


Figure 1.3: *Illustration of how k-space is sampled using a rotated acquisition scheme. Each rectangle covers the central part of k-space of one acquisition.*

the solution. From a Bayesian view-point, regularization corresponds to applying prior knowledge about the probability density function of the HR solution, in order to obtain the maximum-a-posteriori (MAP) solution. By including a properly chosen prior, the solution is thus not only stabilized algebraically but also guided towards an outcome that corresponds to our expectations. In SRR, and many other reconstruction problems, smoothness of the solution has been an immensely popular prior. This is also known as the Tikhonov regularizer. In this thesis, we employ Tikhonov regularization for SRR in the studies described in chapters 2-4.

1.4 Local self-similarity in MRI

More recently, a different type of prior has proven to be very powerful and has been widely applied in image processing and analysis. This prior relies on the observation that local structures tend to repeat themselves throughout an image or within a group of similar images (see Fig. 1.4). Because of this, it has become known as the *self-similarity* prior. This prior is of a statistical nature and is learned from the data itself. By building a dictionary of examples of local structures, in practice consisting of small image patches, a new low-quality patch can be regularized by imposing a fidelity measure between the new patch and the dictionary. This can be exploited for various reconstructive image processing tasks, such as denoising, inpainting, and SRR. The self-similarity prior is discussed in detail in chapter 5 where a novel application of it, specifically aimed at producing isotropic HR volumes from anisotropic MRI slice-stacks, is presented.

The local self-similarity prior has applications beyond reconstructive image processing tasks. It can also be used for classification. Instead of using the dictionary fidelity to regularize the reconstruction of a new image patch, it can be used to mea-

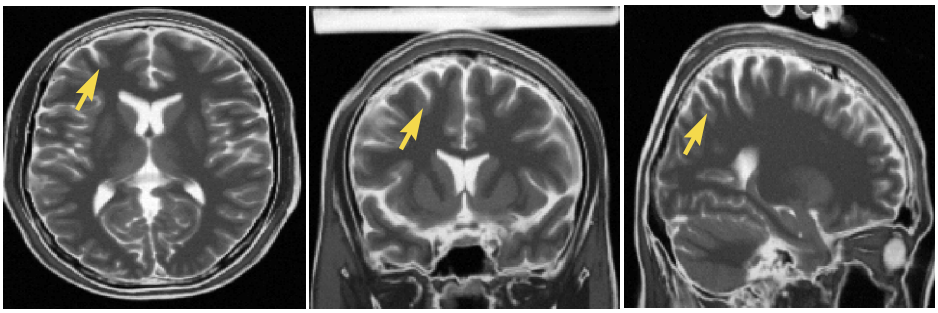


Figure 1.4: *Illustration of self-similarity in brain MRI. At the local level, the image locations pointed to by the arrows have a similar appearance.*

sure how well the new patch is represented using the dictionary. Now, if one dictionary consists of patches from images of trees, and another dictionary of patches from images of human faces, then a patch from a new image of a tree will most likely be represented better by the tree dictionary, especially if we impose the self-similarity prior by requiring the representation to be sparse. This reasoning immediately leads to the concept of *sparse representation classification*, which is explored for several applications, including segmentation of carotid MRI, in chapter 6 of this thesis.

1.5 Thesis contribution and overview

The main contributions of the work presented in this thesis are in the field of SRR in MRI. This topic has been explored from both an applied and a methodological point of view. The first part of the thesis (chapters 2-4) focuses on the former. The presented studies evaluate the feasibility of well-known SRR methods in MRI and applications in small animal research. The aim of these studies was to explore and facilitate practical advantages of SRR in MRI. The second part of the thesis (chapter 5-6) describes two methodological innovations that use the self-similarity of MRI. The first is an SRR method that adapts recent methodological developments in signal processing to perform SRR of MRI slice-stacks. The second is a successful generalization of a state-of-the-art voxelwise classification methodology. Together, they provide background and thorough examination of how self-similarity can be exploited in MRI.

The following paragraphs describe each of the chapters in more detail.

Chapter 2

It remains a major question of the MRI acquisition community whether SRR has any advantage over direct high-resolution image acquisition when SNR and acquisition time are taken into account. In this chapter, we review the theoretical aspects of SRR, briefly describe six representative SRR methods, and present an extensive quantitative and qualitative evaluation of the performance of the six methods in the trade-off space spanned by resolution, SNR, and acquisition time.

Chapter 3

In this chapter, we present an application of SRR for visualization of neuronal activation in manganese-enhanced mouse brain MRI. We show that SRR can be used to reconstruct isotropic high-resolution 3D inversion recovery spin echo (IR-SE) images in a time-frame compatible with *in vivo* experiments. We quantitatively compare the noise, contrast and resolution properties of the SRR IR-SE with those of 3D gradient echo (GE) images conventionally used for manganese-enhanced mouse brain MRI. We qualitatively compare the SRR IR-SE images with the GE images in terms of how well activated mouse brain nuclei can be detected.

Chapter 4

SRR is a computationally expensive method, both in terms of memory requirements and computation time. In this chapter, we present a tool that allows global-to-local interaction with large data-sets. By combining the information from bioluminescence imaging (BLI), CT, and MRI, the tool helps the user identify local regions of interest (ROIs) in a full-body mouse data-set. ROIs can be selected for SRR and high-resolution visualizations be provided in a close to real-time time-frame, on a desktop computer. The tool is qualitatively validated in two case studies involving BLI, CT and MRI of kidney and bone tumors in a mouse model.

Chapter 5

In MRI, 2D slice-stacks yield two native resolutions simultaneously: one at in-plane resolution, and one defined by the slice gap and slice thickness in the slice-selection direction. Local self-similarity of anatomical features occurs both within and across these scales. In this chapter, we present a new method that exploits this cross-scale self-similarity for SRR. The method can be applied to both single and multiple MRI images, thus transcending a traditional methodological division in SRR. We validate the method quantitatively and qualitatively and compare it with basic interpolation and with an existing state-of-the-art SRR method.

Chapter 6

In this chapter we present a new and effective generalization of the basic sparse representation classification (SRC) scheme. Where basic SRC uses the residuals of a single sparse representation, our method fuses multiple sparse representations. The method relies on the joint effect of the accumulation of statistical evidence of class relationship over the multiple representations, and a reduced risk of overfitting in the representations that follow the first one. We validate the approach for three different tasks: texture classification, carotid MRI lumen segmentation, and bifurcation detection in carotid MRI.

Chapter 7

The thesis is concluded with a summary and a general discussion of the results and their implications.

Chapter 2

Super-resolution methods in MRI: Can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time?

Based on:

E. Plenge, D.H.J. 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? *Magn Reson Med 68(6):1983-1993*, 2012.

Abstract

Improving the resolution in magnetic resonance imaging (MRI) comes at the cost of either lower signal-to-noise ratio (SNR), longer acquisition time or both. This study investigates whether so-called super-resolution reconstruction (SRR) methods can increase the resolution in the slice selection direction and, as such, are a viable alternative to direct high-resolution acquisition in terms of the SNR and acquisition time trade-offs. The performance of six SRR methods and direct high-resolution acquisitions was compared with respect to these trade-offs. The methods are based on iterative back-projection, algebraic reconstruction, and regularized least squares. The algorithms were applied to low-resolution data sets within which the images were rotated relative to each other. Quantitative experiments involved a computational phantom and a physical phantom containing structures of known dimensions. To visually validate the quantitative evaluations, qualitative experiments were performed, in which images of three different subjects (a phantom, an ex-vivo rat knee, and a post-mortem mouse) were acquired with different MRI scanners. The results show that SRR can indeed improve the resolution, SNR and acquisition time trade-offs compared with direct high-resolution acquisition.

2.1 Introduction

A fundamental consideration in any MRI experiment is how to optimally balance image resolution, signal-to-noise ratio (SNR), and acquisition time. These three imaging parameters are highly interdependent: higher resolution allows one to observe smaller details, but typically reduces SNR, and/or increases imaging time. At the same time, a certain minimum level of SNR is required to distinguish the signal of interest from system noise, and scan time should be kept low, as MR imaging resources are limited, costly, and long scan times are uncomfortable for the patient, and induce motion artifacts in the images. Together, the three parameters span a space, and their theoretical relations are well known [35]. Techniques such as parallel MRI [37, 73], PROPELLER [67], and compressed sensing [55] affect these relations at the acquisition level. An interesting alternative is to use super-resolution reconstruction (SRR) methods that combine acquisition strategy with post-processing.

SRR is the process of producing a high-resolution (HR) image from a sequence of low-resolution (LR) images, where each LR image transforms and samples the HR scene in a distinct fashion. The idea was first introduced in the 1980s [94] and has since grown into a research field of its own [9, 22, 62]. The first example of SRR applied to MRI was described in a 2001 patent (filed in 1997) [24]. While SRR methods have been studied in MRI for different applications, most work has concentrated on brain MRI [28, 32, 63, 64, 77, 78, 109]. This can be attributed to the fact that SRR is highly dependent on accurate registration of the LR images [54, 76], and in brain MRI a simple global motion model can be applied.

Recently, an SRR scheme based on rotation (rather than translation) of the slice selection direction was extensively investigated in [83]. This scheme, yielding a more effective sampling of the *k*-space, allowed the use of algebraic reconstruction techniques from computed tomography, and the results confirmed the superiority of rotated slice-stacks for SRR. The idea of [83] was extended to slice-stacks of arbitrary orientations and displacements in [71] where they used a regularized reconstruction method and an efficient affine image transform to reduce resampling artifacts.

While SRR in MRI is a developing field, showing its potential in resolution enhancement [31], a major question from the MRI community is whether SRR has any advantage over direct HR acquisition when SNR and acquisition times are taken into account. In this work, we experimentally compare the performance of various SRR methods, and investigate their ability to improve the trade-off between the mentioned imaging parameters compared to direct HR acquisition. After reviewing theoretical aspects of SRR in MRI, we briefly describe the six SRR methods used in our investigation, and present our evaluation framework. The SRR methods have been selected as being representative of the work in SRR in MRI at the time of writing. The evaluation of the methods is based on both phantom data

(allowing quantitative comparisons) and on real image data (allowing qualitative assessment).

2.2 Theory

2.2.1 Super-resolution in MRI

There is consensus [65, 81] that super-resolution in MRI is not achievable in-plane, nor in true 3D acquisitions, since the Fourier encoding scheme excludes aliasing in frequency and phase encoding directions. However, when image contrast requirements yield long repetition times (TR) (as in e.g. T_2-weighted images), it will be significantly faster to acquire 2D slice-stacks than to acquire true 3D volumes, since the acquisition of the slices can be interleaved. In 2D slice-stacks, the individual slices are Fourier encoded, but in the through-plane direction there are no inherent limitations on the frequency spectrum, and aliased frequencies may potentially be recovered. The amount of aliasing depends on the slice profile which ideally is a rect function for non-overlapping slices. However, since the slice profile is determined by the Fourier transform of the finite length slice selection pulse, it will be only an approximation of the rect function. Due to the aliasing present when the object is convolved with the slice profile, SRR is possible.

Several strategies can be adopted when acquiring data for an SRR experiment. We argue, that by acquiring the LR data with rotational increments between the images, as introduced in [83], a more effective sampling of k-space is achieved than by shifting the LR images by sub-pixel distances along the slice selection direction. The latter approach corresponds to increasing the sampling density of the object after convolving it with the slice excitation profile. The signal is thus merely oversampled. If, on the other hand, the slice selection direction is rotated between each LR image, then the narrow slice selection direction bandlimit of each image is oriented in a different direction of the 3D frequency spectrum of the imaged object. In this case, the LR data-set will contain high spatial frequencies in all three dimensions.

In a typical T_2-weighted MR experiment, with an echo time (TE) of 50 ms and a TR of 2500 ms, a maximum of about 40 slices can be acquired within the TR. To reduce the slice thickness for a given volume (which is equivalent to increasing the number of slices), while keeping the contrast (TR and TE) and the in-plane resolution fixed, the only option is to partition the volume into a number of slice packages, N_{SP}, and excite them one at a time. In this way, the acquisition time increases by a factor N_{SP}. However, the slice thickness, and thus N_{SP}, is limited by SNR requirements and the performance of the gradient. The SNR in the final image can be improved by averaging a number of acquisitions, N_{avg}. When an MR experiment is constrained as described above, its total acquisition time T_{acq} is given by

$$T_{\mathrm{acq}} = \mathrm{TR} \cdot N_{\mathrm{PE}} \cdot N_{\mathrm{SP}} \cdot N_{\mathrm{avg}} \ , \tag{2.1}$$

where N_{PE} is the number of phase encoding steps. The SNR depends not only on N_{avg} but also on the slice thickness, h:

$$\text{SNR} \propto h\sqrt{N_{\text{avg}}} \ . \tag{2.2}$$

2.2.2 The imaging model

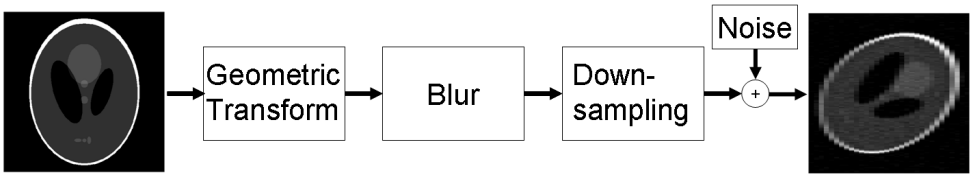


Figure 2.1: *The imaging model assumed in super-resolution reconstruction. Low-resolution images (right) are the result of geometric transformation, blurring, and down-sampling of high-resolution images (left), and the addition of noise.*

Since SRR is an inverse problem, it requires an acquisition model to estimate the HR image underlying the LR images. The imaging model consists of three distinct sub-models: a parameterization of the object space, a model of the physical acquisition process, and a statistical model that appropriately describes the noise. An MRI image can be parameterized by a vector $\mathbf{x}$ of pixels (2D) or voxels (3D). The model of the acquisition process $\mathbf{A} = \mathbf{DBM}$ decomposes into a linear operator $\mathbf{M}$, which describes a geometric transform, a linear space-invariant blurring model $\mathbf{B}$, and a linear downsampling operator $\mathbf{D}$ (Fig. 2.1). According to [34] the noise $\mathbf{n}$ in MRI can be assumed to be additive, white and Gaussian when the SNR > 3. Thus, the acquisition of the k^{th} LR image, $\mathbf{y}_k$, is described in vector form by

$$\mathbf{y}_k = \mathbf{A}_k \mathbf{x} + \mathbf{n}_k. \tag{2.3}$$

Solving the inverse problem corresponds to recovering $\mathbf{x}$ given the $\mathbf{y}_k$ and $\mathbf{A}_k$. The problem can be formulated as an ordinary least squares (LS) problem,

$$\mathbf{x} = \underset{\mathbf{x}}{\text{argmin}} \sum_{k=1}^{N} ||\mathbf{y}_k - \mathbf{A}_k \mathbf{x}||^2, \tag{2.4}$$

which, vertically concatenating matrices $\mathbf{A}_k$ into $\mathbf{A}$, vectors $\mathbf{y}_k$ into $\mathbf{y}$, and assuming a Gaussian noise model, has the closed-form maximum-likelihood solution

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}. \tag{2.5}$$

Using additional prior knowledge about the solution, e.g. that it is smooth, Eqs. (2.4) and (2.5) generalize to

$$\mathbf{x} = \underset{\mathbf{x}}{\operatorname{argmin}} \sum_{k=1}^{N} ||\mathbf{y}_k - \mathbf{A}_k \mathbf{x}||^2 + \lambda \|\mathbf{C}\mathbf{x}\|^2 \tag{2.6}$$

and

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A} + \lambda\, \mathbf{C}^T\mathbf{C})^{-1}\mathbf{A}^T\mathbf{y}, \tag{2.7}$$

respectively, where $\mathbf{C}$ is a regularization operator formalizing the prior knowledge, and λ is a scalar weight. Eq. (2.7) is also known as the *regularized* LS solution.

A direct solution of Eqs. (2.5) and (2.7) is generally infeasible, since $\mathbf{A} \in \mathbb{R}^{n\times m}$, where n and m are the number of voxels in the reconstruction ($\mathbf{x}$) and in all the LR images ($\mathbf{y}$), respectively. Instead, iterative methods are usually applied to approximate the solution. In the following, we briefly review six such methods, including two iterative back-projection methods, two algebraic reconstruction methods, and two regularized LS solvers.

2.3 Methods

2.3.1 Super-resolution reconstruction methods

Iterative back-projection methods: One approach to solving the inverse problem of SRR is to start from an initial guess of the HR image, then simulate N LR images via Eq. (2.3) and use the difference between the simulated and the acquired LR images to correct the HR estimate. The two steps are applied iteratively. Iterative back-projection (IBP) [43] is one such method. It solves the LS problem of Eq. (2.4) by

$$\hat{\mathbf{x}}^{(l+1)} = \hat{\mathbf{x}}^{(l)} + \sum_{k=1}^{N} \mathbf{H}_{\mathrm{BP}}(\mathbf{y}_k - \mathbf{A}_k\hat{\mathbf{x}}^{(l)}), \tag{2.8}$$

where $\mathbf{H}_{\mathrm{BP}}$ is a back-projection kernel of choice, and superscript (l) indicates the iteration number. IBP solves the LS problem and thereby implicitly takes the Gaussian noise model into account. It has been noted [22] that with $\mathbf{H}_{\mathrm{BP}} = \mathbf{A}^T$, IBP is equivalent to the steepest descent method. Our implementation follows [43] and includes the proposed heuristic stabilization and noise reduction schemes.

As a non-regularized LS solver, IBP is sensitive to outliers. An alternative method, termed robust super-resolution (RSR) [110], modifies IBP by exchanging its update term with a pixel-wise median of the errors:

$$\hat{\mathbf{x}}^{(l+1)} = \hat{\mathbf{x}}^{(l)} + N * \mathrm{median}\{\mathbf{H}_{\mathrm{BP}}(\mathbf{y}_k - \mathbf{A}_k\,\hat{\mathbf{x}}^{(l)})\}_{k=1}^{N}\,. \tag{2.9}$$

RSR approximates IBP and thus the LS solution. However, no analytical solution exists for this method.

Algebraic reconstruction methods: The classic algebraic reconstruction technique (ART) is a special case of the projection onto convex sets (POCS) methods [44, 88]. It was introduced in image processing by Gordon et al. [30]. ART starts from an initial guess, $\hat{\mathbf{x}}^{(0)}$, and then iteratively updates $\hat{\mathbf{x}}$ one pixel at a time, via

$$\hat{\mathbf{x}}_j^{(l+1)} = \hat{\mathbf{x}}_j^{(l)} + \lambda^{(l)} \frac{\mathbf{y}_i - \langle \mathbf{a}_i, \hat{\mathbf{x}}^{(l)} \rangle}{\|\mathbf{a}_i\|_2^2} \mathbf{a}_{ij}, \tag{2.10}$$

where $\hat{\mathbf{x}}_j^{(l)}$ denotes the j^{th} element in $\hat{\mathbf{x}}$ in the l^{th} iteration, $\mathbf{a}_i$ is the i^{th} row of $\mathbf{A}$, and $\mathbf{a}_{ij}$ is the $(i,j)^{\text{th}}$ element of $\mathbf{A}$. ART converges to the minimum norm solution [38], satisfying

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} ||\mathbf{x}||_2^2 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}, \tag{2.11}$$

the minimum L_2 norm thus functioning as a regularizer.

An ART implementation is easily extended into POCS by applying constraints to Eq.(2.10) at each projection step l. In our implementation, we applied amplitude and energy constraints as described in [88].

Regularized least squares methods: The term $\mathbf{C}$ in Eq.(2.6) should incorporate prior knowledge about the solution. When it does, the method is known as Tikhonov regularization. The regularization term is usually some norm of an operation (e.g. identity, first and second order derivative, Laplacian) on the HR estimate $\hat{\mathbf{x}}$. Regularization makes the underdetermined problem of SRR less ill-posed, but has the undesirable effect of smoothing sharp features in the image, such as step edges. In our implementation of a Tikhonov-regularized LS solver (TIK), we used the L_2-norm of the second order derivative of the HR reconstruction as our regularization term:

$$\|\mathbf{C}\mathbf{x}\|^2 = \left(\frac{\partial^2 \mathbf{x}}{\partial r_1^2}\right)^2 + \left(\frac{\partial^2 \mathbf{x}}{\partial r_2^2}\right)^2 + \left(\frac{\partial^2 \mathbf{x}}{\partial r_3^2}\right)^2, \tag{2.12}$$

where r_i is the spatial dimension over which the partial derivative is taken. We used an average of the aligned and upscaled LR images as the initial HR estimate. The regularized LS problem of Eq.(2.6) was solved with the conjugate gradient method and the transforms in $\mathbf{A}$ and $\mathbf{A}^T$ were implemented using bilinear interpolation.

A recently proposed alternative SRR method (LASR) [71] also employs Eq.(2.12) for regularization and the conjugate gradient method for solution of Eq. (2.6). However, this method applies $\mathbf{A}$ and $\mathbf{A}^T$ by an affine transformation scheme that minimizes aliasing and spectral distortions.

The experiments with TIK and LASR in this study, were all performed with little regularization ($\lambda = 0.004$), except for the *in vivo* experiments where $\lambda = 0.01$. These values were found by experimenting with a range of λ-values and qualitatively determining the best result.

2.3.2 Evaluation framework

To evaluate whether SRR methods can decrease the slice thickness of MR images at better SNR and timing trade-offs than direct HR acquisition, we devised an evaluation framework, using the relations established above (Eqs. 2.1 and 2.2): 1) for fixed field-of-view (FOV) and N_{avg}, the acquisition time scales inversely with slice thickness, 2) for a fixed N_{avg} the SNR scales linearly with slice thickness, and 3) for a given slice thickness, h, the SNR increases with N_{avg} and thus, using Eq.(2.1), with the acquisition time by $\sqrt{T_{\text{acq}}}$.

Quantitative Experiments: To quantitatively evaluate the performance of the SRR methods, two types of data were used: 1) simulated acquisitions based on the well-known computer generated Shepp-Logan phantom (SLP), and 2) MRI images of a Varian phantom.

Simulations: The SLP experiments were performed to investigate the situation in which the acquisition model is known exactly. In total, 24 anisotropic reformatted LR slices of 128×32 pixels (frequency encoding direction × slice selection direction) were simulated from a 512×512 pixel image, using rotation around the image center (and the virtual phase encoding axis) in increments of 180/24 degrees (see Fig. 2.2). Gaussian noise was added, such that the SNR was around 32 dB, which is within the realistic range for an MRI experiment. From the 24 LR images, eight sub-sets were created containing $n = 1$, 2, 3, 4, 6, 8, 12, and 24 images, respectively, such that the rotation increments within the sub-sets were $180/n$ degrees. From each of these sub-sets, images of 128×128 pixels were produced by each of the six SRR algorithms, and by using a bilinear interpolation scheme (INT). To simulate direct HR acquisition (DAC), six images of 128×128 pixels, covering the same 2D FOV as the LR simulations, were also produced from the 512×512 pixel image, again by applying the imaging model. Averaging these images improves the SNR of the resulting image. According to the inverse relation between slice thickness and acquisition time, six averaged HR images take the same acquisition time as 24 LR images with four times thicker slices, see Fig. 2.2. This allowed us to compare the SRR images with the averaged HR images within the same "acquisition time" range.

MRI phantom acquisitions: A corresponding MRI experiment was performed using the phantom, containing liquid-filled structures (water doped with copper sulphate) yielding homogeneous signal, surrounded by plastic and air, resulting in signal void. LR data were acquired using a T_2-weighted 2D fast spin echo (FSE) sequence on a clinical GE 3T system. TR was 2100 ms, TE was 27.4 ms, with an $N_{\text{avg}} = 1$ and $N_{\text{SP}} = 1$. The slice-stack consisted of 18 slices (4mm thick) with a FOV of 128×128 mm, $N_{\text{PE}} = 128$, and a resolution of 1.0×1.0×4.0 mm. The scan

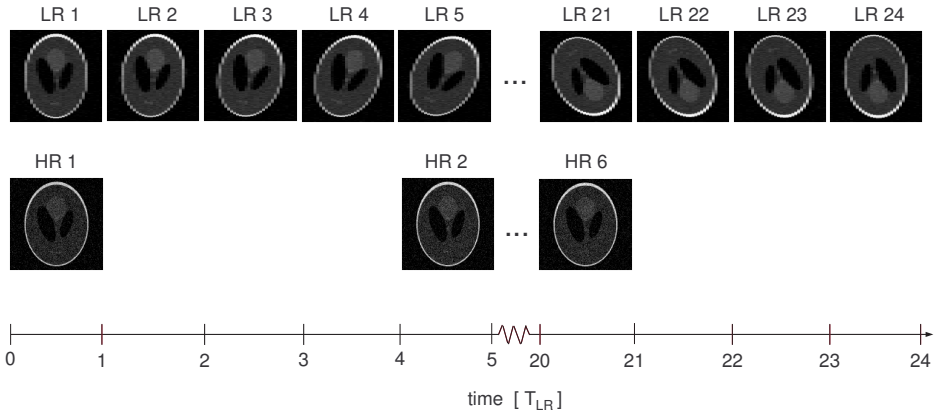


Figure 2.2: *Illustration of the simulation and acquisition scheme of the quantitative experiments. Upper row: 24 LR images are simulated/acquired while rotating the FOV around the phase-encoding axis in increments of 180/24 degrees. Lower row: Six HR images of 4 times thinner slices and 4 times longer acquisition time than the LR images are simulated/acquired, enabling post-acquisition production of multiple averages images. Note that the image planes shown (both LR and HR) are spanned by the frequency encoding and slice selection directions. The unit of the time axis is T_{LR}, the acquisition time of a single LR image.*

time per stack (T_{LR}) was 73 s. The experiment was repeated 24 times, while the slice-stack was rotated around the phase encoding direction in increments of approximately 180/24 degrees.

For analysis, eight sub-sets were picked from the 24 slice-stacks containing $n =$ 1, 2, 3, 4, 6, 8, 12, and 24 stacks, respectively. Each sub-set was selected such that the rotational increments between its slice-stacks were approximately equal. If we denote the acquisition time of one LR image by T_{LR}, each sub-set thus had an acquisition time of $n\,T_{\mathrm{LR}}$. Reconstruction of 1 mm isotropic resolution images from the eight LR sub-sets allowed evaluating resolution and SNR as a function of the acquisition time in the reconstructed images.

For quantitative comparison, isotropic reference images were acquired directly at 1 mm slice thickness. Having four times thinner slices than the LR images but the remaining imaging parameters identical to those of the LR images, these required an acquisition time of $4\,T_{\mathrm{LR}}$ each. For a given imaging sequence and resolution, conventional MRI can improve the SNR only by averaging multiple images. Imaging time increases linearly with the number of averaged images (Eq. (2.1)). This way, the quality of direct acquisitions (DAC) could be evaluated at time points that were multiples of $4\,T_{\mathrm{LR}}$ and spanned the same time range as the reconstructed images,

Fig. 2.2.

Quantitative evaluation measures: We assessed the performance of the six described SRR methods: IBP, RSR, ART, POCS, TIK, and LASR. We compared the SRR methods to simple bilinear interpolation of the data-sets (INT) and to directly acquired HR data-sets (DAC). Resolution (in millimeters) was measured in each image as the average width of 20 edge profiles, see Fig. 2.3. The edge-width of each profile was computed by nonlinear LS fitting of the following sigmoid function [32, 83]:

$$f = a_1 + \frac{a_2}{1 + \exp(-a_3(x - a_4))}, \tag{2.13}$$

where $4.4/a_3$ is the width of the curve's rise length from 0.1 to 0.9 for normalized intensity values ($a_3 = 4.4$ corresponds to a rise length of 1 pixel). The SNR of the evaluation images was computed by dividing the mean of four high-intensity foreground regions with the standard deviation of a background region. The foreground and background regions-of-interest used for the calculation are visualized in Fig. 2.3.

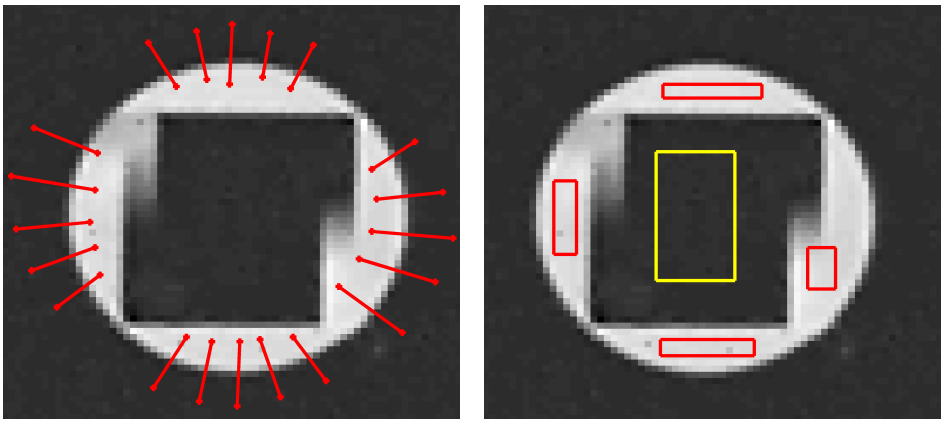


Figure 2.3: *The edge-width measures are based on the intensity profile of 20 edge-crossing strips as illustrated in the left figure. The SNR calculations are based on the high-intensity regions and one background region as illustrated in the right figure.*

Qualitative Experiments: To qualitatively evaluate the capabilities of the SRR methods, four LR data-sets were acquired of four different samples, using three different scanners. In all cases the LR images were acquired just above the scanner systems' lower limits on slice thickness. Reconstructions of HR images were thus aiming at

surpassing these limits. As references, to assess the quality of the reconstructions, data-sets were acquired, where the plane of interest in the reconstructions were captured in-plane.

Phantom: The same phantom as described above was scanned using a T2-weighted 2D fast spin echo (FSE) sequence on a clinical GE 3T system. TR was 2100 ms, TE was 36.5 ms, with $N_{\text{avg}} = 1$ and $N_{\text{SP}} = 1$. The 2D slice-stack consisted of 36 slices (2 mm thick), with a FOV of 128×128 mm, $N_{\text{PE}} = 256$, and a resolution of 0.5×0.5×2.0 mm. The scan time per stack was 278 s. The experiment was repeated 12 times, while the slice-stack was rotated around the phase encoding axis in increments of approximately 180/12 degrees. SRR was performed on an isotropic HR grid of resolution 0.5×0.5×0.5 mm.

Post-mortem mouse: A whole-body scan of a post-mortem mouse was acquired on a 7T Bruker Pharmascan system using a recovery FSE (frFSE) sequence. TR was 6648 ms, TE was 33 ms, with $N_{\text{avg}} = 1$ and $N_{\text{SP}} = 1$. The 2D slice-stack consisted of 64 slices (0.5 mm thick), with a FOV of 50×32 mm, and a resolution of 0.125×0.125×0.5 mm. The scan time per stack was 213 s. The slice-stack was rotated around the phase encoding axis in 24 uniform increments of 180/24 degrees. Due to the scanner's limited FOV, the mouse was scanned in three sections (head, chest, lower abdomen). SRR was performed on each section separately on isotropic HR grids of resolution 0.125×0.125×0.125 mm.

Ex vivo rat knee: An *ex vivo* rat knee, into which cells labeled with super-paramagnetic iron oxide (SPIO) particles were injected, was acquired on a 7T Varian animal system with a fast recovery FSE (frFSE) sequence. TR was 5600 ms, TE was 7.1 ms, with $N_{\text{avg}} = 1$ and $N_{\text{SP}} = 1$. The 2D slice-stack consisted of 24 slices (0.4 mm thick), with a FOV of 26×26 mm, $N_{\text{PE}} = 256$ and a resolution of approximately 0.1×0.1×0.4 mm. The scan time per stack was 252 s. The slice-stack was rotated around the phase encoding axis in 12 increments of approximately 180/12 degrees. SRR was performed on an isotropic HR grid of resolution 0.1×0.1×0.1 mm.

In vivo mouse brain: To gain preliminary insight into the potential of *in vivo* application of SRR in MRI, brain images of a living mouse were acquired on a 7T Bruker Pharmascan system using a RARE T1 weighted spin echo sequence. TR was 680 ms, TE was 8.7 ms, $N_{\text{avg}} = 4$ and $N_{\text{SP}} = 1$. The 2D slice-stack consisted of 21 slices (0.5 mm thick), with a FOV of 150×256 mm, $N_{\text{PE}} = 150$ and a resolution of approximately 0.1×0.1×0.5 mm. The scan time per stack was 210 s. The slice-stack was rotated in 8 increments of approximately 180/8 degrees around the frequency encoding axis. SRR was performed to an isotropic HR grid of resolution 0.1×0.1×0.1 mm. As a reference for the anatomical structures seen in the reconstructions, a RARE 3D T1 weighted gradient echo image was included in the experiment. The resolution

was $0.1 \times 0.1 \times 0.1$ mm, TR was 30 ms, TE was 3.6 ms, and $N_{\text{avg}} = 4$.

2.4 Results

2.4.1 Quantitative results on simulations

The evaluation results of the simulation experiment are summarized in Fig. 2.4, which shows the resolution and SNR of the SLP reconstructions versus "acquisition time". All SRR methods achieve large resolution improvements for acquisition times up to $4\,T_{\text{LR}}$. Especially the regularized least-squares methods (TIK, LASR), and the iterative back-projection methods (IBP, RSR) achieve substantial resolution improvements, around a factor 2; the algebraic reconstruction methods (ART, POCS) in contrast, result in far less resolution gain. Beyond $4\,T_{\text{LR}}$ LASR, IBP, and RSR reconstructions show only modest resolution gains. Only the TIK reconstructions eventually achieve a resolution of 1 mm. This happens around $10\,T_{\text{LR}}$, and at $24\,T_{\text{LR}}$ the resolution is around 0.83 mm. The measured resolution of the directly simulated HR images (DAC) is around 1.18 mm for all time points. It is surpassed by TIK reconstructions around $6\,T_{\text{LR}}$. Standard bilinear interpolation (INT) has virtually no resolution enhancing capabilities. In terms of SNR, both IBP and RSR reconstructions show considerable improvement over the entire time range. TIK, LASR, and ART reconstructions show a modest upward trend, while the SNR improvement of the POCS reconstructions is more pronounced. INT images have a high SNR that increases with increased acquisition time.

2.4.2 Quantitative results on phantom MRIs

The performance of the six SRR methods, bilinear interpolation, and direct acquisition on the Varian phantom data is shown in Fig. 2.5. The general trends are similar to those found with the SLP reconstructions. Only the SNR curves of TIK and LASR reconstructions are different between the SLP and Varian plots. For reference, the graphs also include the theoretical curves (THEO) of isotropic 1.2, 1.3, and 1.4 mm resolution images. It can be observed that the TIK reconstructions surpass the measured resolution of the DAC images around $15\,T_{\text{LR}}$ while its SNR at that point is more than 6 dB higher than that of the DAC images. Furthermore, it can be observed how TIK reconstructions reach a resolution of 1.2 mm already at $5\,T_{\text{LR}}$. For acquisition times beyond $12\,T_{\text{LR}}$, the measured resolution of IBP reconstructions is around 1.2 mm, while the SNR for this method at $12\,T_{\text{LR}}$ is 3.0 dB higher than that of a direct acquisition of 1.2 mm isotropic resolution. At $24\,T_{\text{LR}}$, this difference is even 4.8 dB. From $4\,T_{\text{LR}}$ to $8\,T_{\text{LR}}$, the IBP reconstructions reach a measured resolution of approximately 1.3 mm, while the corresponding SNRs are only slightly (0.5-0.8 dB) higher than that of the theoretical 1.3 mm resolution image.

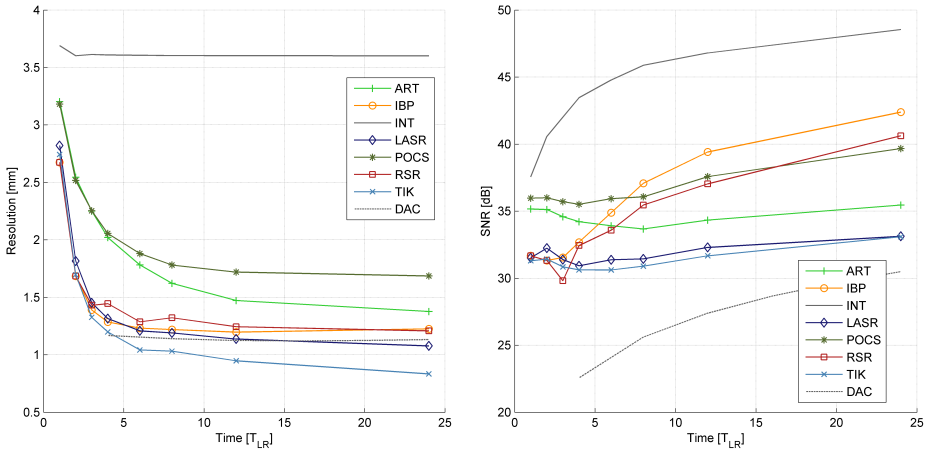


Figure 2.4: *Resolution (left) and SNR (right) of the six SRR methods and bilinear interpolation (solid curves) as a function of acquisition time compared to direct HR imaging (dashed curves) for the computer generated SLP data. The unit of time along the x-axis is T_{LR} (the time required to acquire one LR image).*

2.4.3 Qualitative results

The results of the three qualitative experiments are presented in Figs. 2.6, 2.7, and 2.8. The images in Fig. 2.6 show a detail of radial lines from the Varian phantom and confirm the quantitative results presented in the previous section. The striking improvement in resolution achievable by SRR is clear by comparison with the reference anisotropic LR image, in which the lines can be resolved only when they are perpendicular to the HR dimension of the image. Moving along the radial lines toward the center, at some point the ability to resolve the lines is lost. This point occurs earlier for the reconstructions than for the image where the radial lines are captured in-plane, showing that the latter still has somewhat higher resolution in that plane. The TIK and IBP reconstructions, which had the highest resolution in the quantitative evaluation, are sharper than the other reconstructions, and have good contrast. Over the entire time range, we found that TIK and LASR reconstructions were the most qualitatively pleasing.

Fig. 2.7 shows the results of the experiments (reconstructed with TIK) on a post-mortem mouse. The sample axial slices demonstrate that the quality of the reconstructions clearly improves when an increasing number of LR images is used. In agreement with the results of the quantitative experiments, it can be observed that the quality improvement is substantial going from 2 to 8 LR images, while a smaller effect is seen when going from 12 to 24 LR images. Again, the highest resolution is

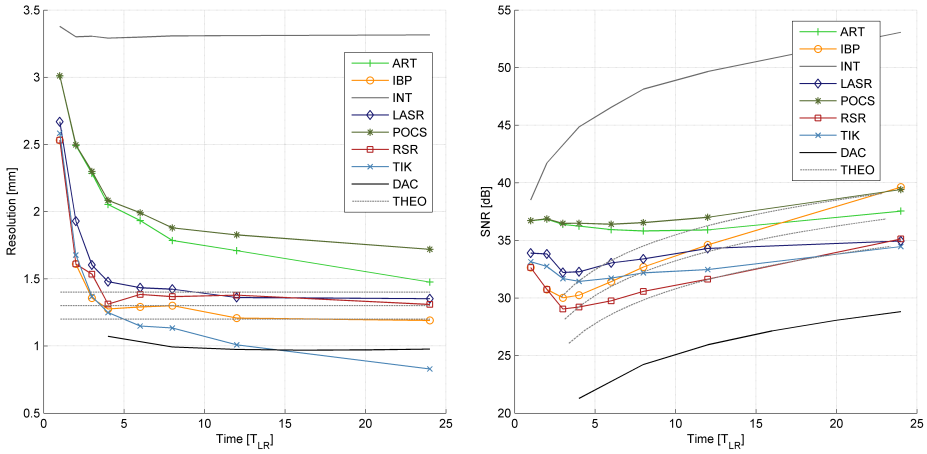


Figure 2.5: *Resolution (left) and SNR (right) of the six SRR methods and bilinear interpolation (solid curves) as a function of acquisition time compared to direct HR imaging at 1 mm (black curve) and 1.2, 1.3, 1.4 mm (from lower to upper dashed curves) for the MRI data of the physical Varian phantom. The unit of time along the horizontal axis is T_{LR} (the time required to acquire one LR image).*

achieved by direct acquisition with the slice of interest in-plane.

Sample results (reconstructed with TIK) of the experiments on an *ex vivo* rat knee with SPIO-labeled stem cells inside are shown in Fig. 2.8. The cell cluster is visible as a darker shadow in the sequence of reformatted LR slices (top row). In the sequence of SRR images (middle row) the cell cluster is considerably better visualized. The bottom row shows an LR acquisition where the slice selection and phase encoding directions were interchanged such that the FOV of the reformatted images in the top row was captured in-plane. In the first of these two slices, the cell cluster is clearly visible. However, the distance between consecutive slices is large, and in the second slice the cell cluster is hardly visible.

Finally, the results of the *in vivo* mouse brain experiments are shown in Fig. 2.9. The reconstructions of all tested methods are included and show characteristics similar to those of the MRI phantom. In all cases, a clear improvement is seen from the LR images to the reconstruction results. TIK and LASR reconstructions are sharper than ART and POCS reconstructions, and less noisy than IBP and RSR reconstructions. The LASR reconstruction is slightly smoother than that of TIK. Comparing with the isotropic 3D acquisition, the reconstructions reproduce the anatomical structures seen in that image with a reasonable degree of accuracy.

2.5 Discussion

In this study we have shown that SRR indeed provides advantages over direct HR acquisition when SNR and acquisition time are taken into account. While the performance of SRR compared with that of direct acquisitions has been studied before [32], no framework has previously been presented in the literature that allows thorough analysis of the relations between resolution, SNR, and acquisition time. So far, the net practical advantage of using SRR has thus remained unclear. To enable a thorough investigation of this, we designed the evaluation framework presented in this work. Image quality evaluation is notoriously tricky, the quality cannot be captured by any one single measure. In our framework we therefore included both quantitative and qualitative comparisons between SRR and direct HR acquisitions.

Our quantitative results suggest that SRR enables surpassing the resolution achieved by direct HR acquisition. This confirms the results of [32]. Quantitative experiments (Fig. 2.5) show that the reconstructions of one SRR method (TIK) surpasses the resolution of the direct HR acquisition, while having substantially higher SNR ($>$ 6 dB) at that time point. However, comparing Fig. 2.6-h with Fig. 2.6-a, this result is not confirmed. The reason for this discrepancy is most likely that the TIK algorithm implicitly applies a sharpening filter and amplifies high-frequency content excessively. This results in overshoot (Gibbs ringing) artifacts at step edges in the reconstructions, causing the sigmoid function to fit to a more steeply rising edge and thus underestimate the edge-width.

The IBP reconstructions do not surpass the resolution of the direct acquisition but approaches and slightly surpasses a resolution of 1.2 mm, while achieving an SNR of 3.0-4.8 dB higher than a theoretical 1.2 mm resolution direct acquisition, in the same total acquisition time. As such, the quantitative results in Figs. 2.4 and 2.5 show a clear improvement in the resolution-SNR-acquisition time trade-off.

The experimental results (in particular those illustrated in Figs. 2.6 and 2.8) highlight the benefits of isotropic resolution. When small structures are to be visualized with MRI, SRR methods can be employed to surpass the limits on slice thickness imposed by the scanner system or protocol. The LR images of all four qualitative experiments have been acquired just above the slice thickness limit and by SRR reconstructed to isotropic resolution below that limit. In this way, the use of SRR alleviates the need for careful orientation of the object or acquisition planes, relaxing experimental conditions.

The ability to compare the performance of SRR methods within the resolution-SNR-acquisition time trade-off space is an additional result of our study. In this work we have compared six SRR methods representative of those found in the literature. The IBP and RSR methods use the data as the only constraint on the solution, while TIK and LASR incorporate prior knowledge as regularization terms. In the latter case, a smoother, lower resolution, solution can be expected. However, using low λ-values, the effect of the regularization is limited. ART and POCS yielded the

overall poorest resolution results. For all six methods, we found that using as few as four LR images for the reconstruction yielded substantial improvements in resolution. This may be related to the anisotropy factor of the LR images, which was also four in our experiments.

The choice of using the phase encoding axis as the axis of rotation of the LR acquisitions was deliberate for the *ex vivo* mouse scans. These were acquired in three parts with a FOV smaller than the whole length of the mouse body. By choosing the orientation of the read-gradient in the head-foot direction, aliasing was avoided. This choice causes chemical shift artifacts to change between the LR images and potentially leads to blurring of the reconstructions. We have, however, not observed this effect in our experiments. By instead choosing the frequency encoding axis as the axis of rotation, the phase encoding axis changes direction in each rotated slice-stack. The related change in direction of the motion blurring artifacts between the slice-stacks, is expected to increase motion blurring when the subject is moving. We tested this setup in our *in vivo* reconstructions but did not observe pronounced blurring.

Good agreement was found between the results of the quantitative experiments on the computer generated phantom (SLP) and the real phantom (Varian) (cf. Figs. 2.4 and 2.5). The results of the qualitative experiments also confirmed the findings of the quantitative experiments (cf. Figs. 2.5 and 2.6). This is in line with a previous observation that the performance of SRR methods on real MRI data can be predicted accurately by measuring the performance on simulated data if an appropriate model of the imaging system is available [97], and suggests that for future SRR algorithm development and testing it is valid to use simulated phantoms in the initial stages and postpone costly MRI experiments to the final stages.

In vivo mouse brain experiments were included in this study to get an idea of the potential of applying SRR in more practical research settings. We deliberately did not include these experiments in the trade-off investigation, since an entirely new set of problems related to motion and blood flow are expected when working with *in vivo* data. The results, however, were encouraging: we did not observe any motion related artifacts in the reconstructions, and the reconstructions reproduced the anatomical structures seen in the reference 3D acquisition with reasonable accuracy. The experiment shows that in principle SRR is possible in live animals. The brain, in particular, seems to lend itself nicely to SRR as pulsatile and breath induced motion is limited compared with in the body, and effective means for animal fixation are available. Our results opens the field for further experiments in this category.

A complete study of the performance of SRR in MRI is an immense task. Extensive quantitative and qualitative experiments have already been performed in [83], that focused on optimizing the acquisition of the LR data, considering the number of slice-stacks, the slice thickness, SNR in the LR images and more. In this study, we have constrained ourselves to studying the resolution-SNR-acquisition

time relationship, and kept all other parameters fixed. These constraints clearly remove much of the flexibility normally available in an MRI experiment but they are necessary to consistently investigate isolated aspects of the imaging process. Future studies of SRR in MRI should include rigorous comparisons of SRR versus specific experimental acquisition schemes (high-resolution 3D and interlaced slice acquisition schemes, different sequence etc.). SRR may prove particularly useful for MRI methods that inherently suffer from long acquisition times, such as fast spin echo techniques (FSE, RARE), and inversion recovery methods (FLAIR, IR-SE), especially when SNR requirements forces the averaging of multiple acquisitions and thereby prolong the acquisition time. Apart from comparing SNR, CNR, resolution and acquisition time, future studies should also look into the experimental factors influencing the real and reconstructed image quality, such as physiological noise, pulsatile motion (heart, blood flow), slice-to-slice interference, etc.

In conclusion, we have shown that SRR is capable of providing better tradeoffs between resolution, SNR and acquisition time than direct HR acquisition, and is particularly useful when limits on slice thickness and the number of slices are imposed by the scanner system or protocol. We have not been able to reproduce earlier results in which SRR reconstructions surpass the resolution of directly acquired high-resolution images, and suspect that these results could be influenced by a bias toward overshoot in the resolution measure. We have demonstrated SRRs potential in application to *in vivo* MRI, and provided directions for future research in SRR in MRI. Finally, by the presented evaluation framework we provide a protocol for quantitatively evaluating and comparing the performance of SRR algorithms in MRI.

2.6 Acknowledgments

The authors thank Mark Halberstadt (Erasmus MC) and Ernst Suidgeest (Leiden University Medical Center) for assistance with MRI acquisitions, Gerben van Buul (Erasmus MC) for the *ex vivo* rat knee sample, and Dana Poole (Leiden University Medical Center) for the *in vivo* mouse brain acquisitions. This research was funded by Medical Delta (HST-Klein project) and by the European Commission in the Seventh Framework Programme (ENCITE project).

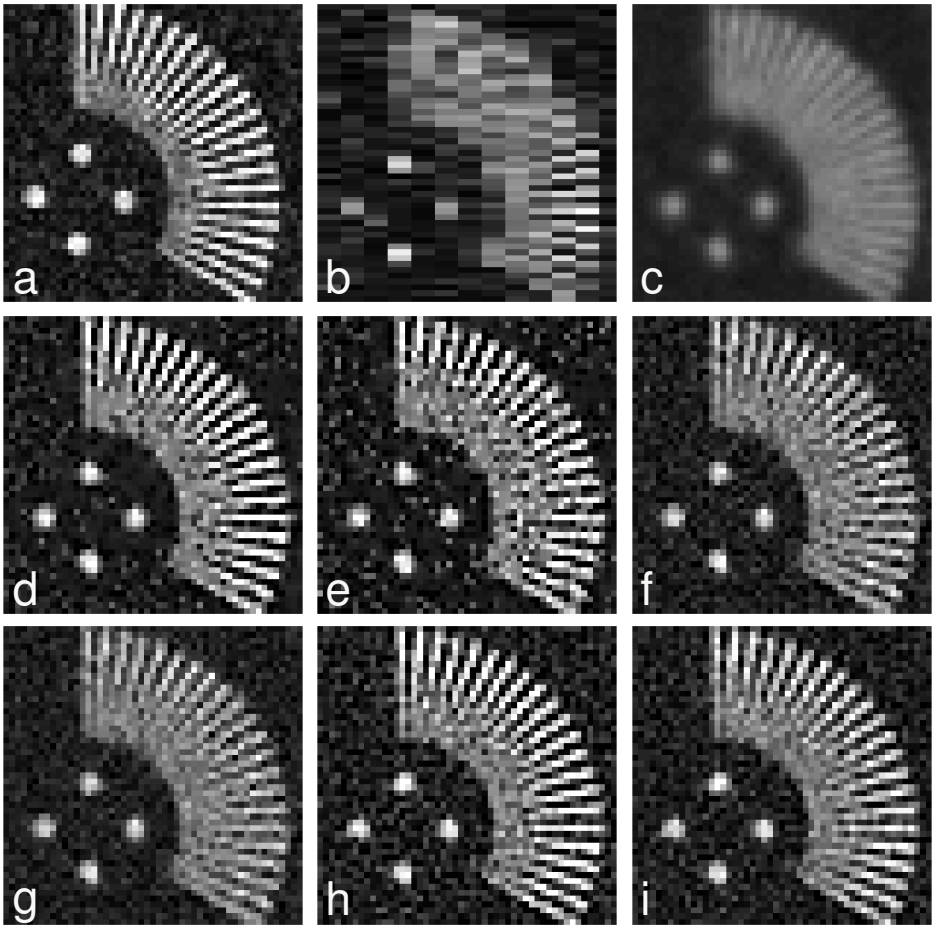


Figure 2.6: *Visual impression of the performance of the SRR methods on the Varian phantom data. The panels show: (a) the structure of interest in the HR plane of a direct acquisition, (b) the structure of interest in the LR plane, and the results of (c) INT, (d) IBP, (e) RSR, (f) ART, (g) POCS, (h) TIK, and (i) LASR, based on 12 LR images. The intensities of each image have been scaled according to the mean value of a homogeneous high-intensity region in the image.*

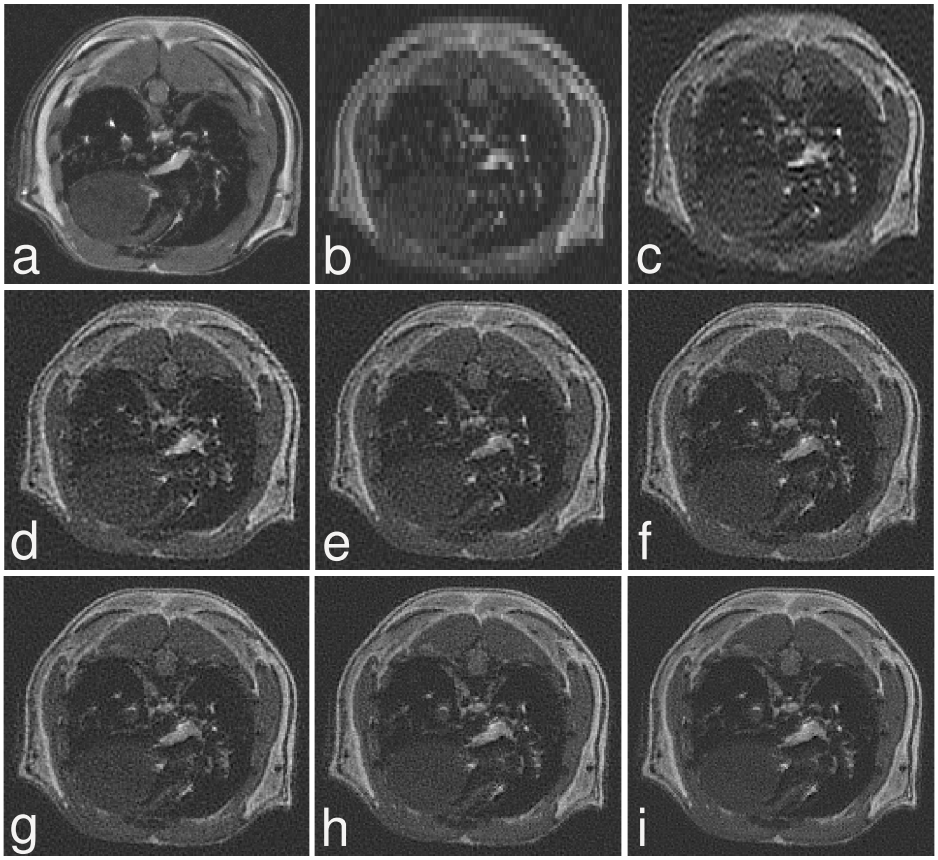


Figure 2.7: *Axial slices of the full-body mouse scan. The images show: (a) a slice from a direct acquisition with thick slices but high resolution in the plane shown, (b) one of the LR acquisitions, and the results of TIK using (c) 2, (d) 3, (e) 4, (f) 6, (g) 8, (h) 12, and (i) 24 LR images.*

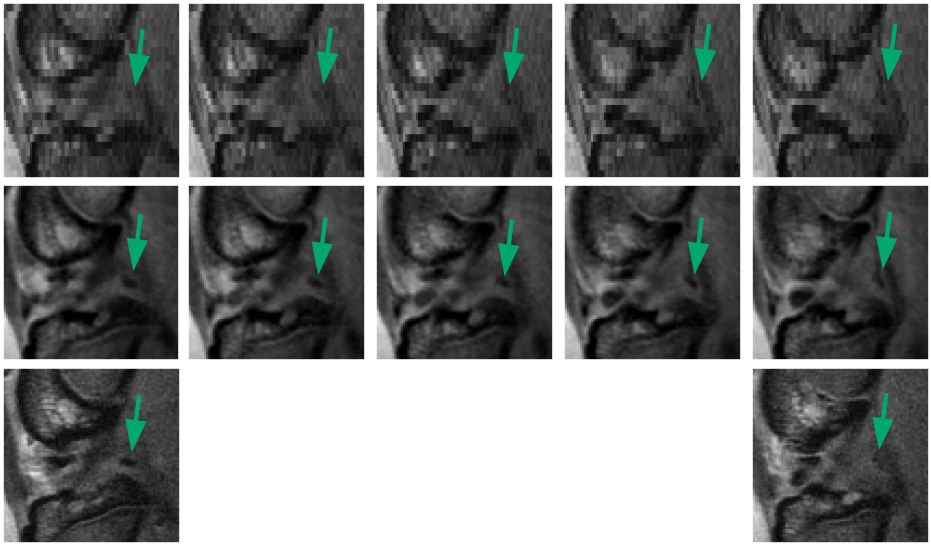


Figure 2.8: *Potential of SRR demonstrated on* ex vivo *data of a rat knee with SPIO-labeled cells (green arrows). Top row: a sequence of reformatted 2D slices of anisotropic LR acquisition. Middle row: corresponding SRR slices, reconstructed using 12 LR images and the TIK method. Bottom row: two slices of anisotropic reference acquisition with target resolution in the plane reconstructed above. The three sequences all cover approximately the same spatial volume.*

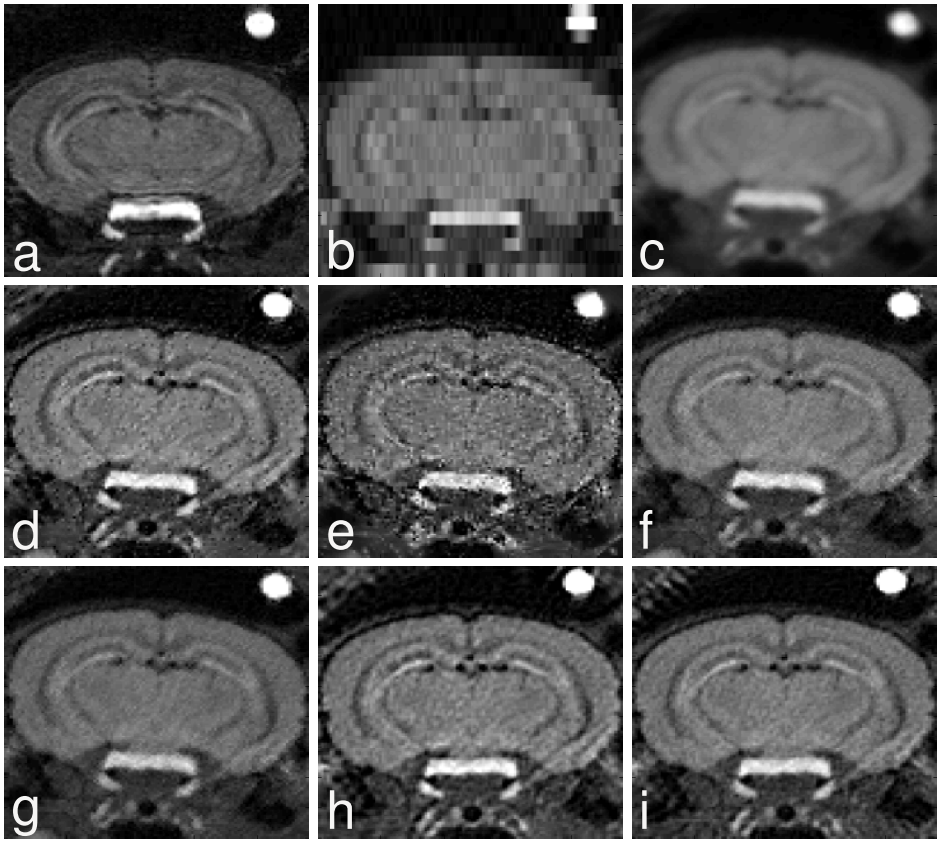


Figure 2.9: *Axial slices of* in vivo *mouse brain. The images show: (a) a direct HR 3D acquisition, (b) one of the eight LR acquisitions, and reconstructions with (c) INT, (d) IBP, (e) RSR, (f) ART, (g) POCS, (h) TIK, and (i) LASR, using 8 LR images.*

Chapter 3

3D inversion recovery spin echo manganese-enhanced MRI of the mouse brain using super-resolution reconstruction

Based on:

E. Plenge*, D. S. Poole*, D.H.J. Poot, E.A.J.F. Lakke, W.J. Niessen, E. Meijering, L. van der Weerd. 3D inversion recovery manganese-enhanced MRI of the mouse brain using super-resolution reconstruction to visualize nuclei involved in higher brain function. *In review*.

Abstract

Visualization of activity in the mouse brain using inversion recovery spin echo (IR-SE) manganese-enhanced MRI (MEMRI) provides unique contrast but suffers from poor resolution in the slice encoding direction. In this study, we investigate whether super-resolution reconstruction (SRR) of IR-SE MEMRI can improve the visualization of mouse brain nuclei involved in higher brain function. Quantitative experiments were performed on a phantom containing compartments of various manganese concentrations. Depending on the inversion time of the IR-SE acquisitions and the manganese concentration, the SRR IR-SE images showed better contrast than the conventional 3D GE acquisition that was used for comparison. *In vivo* experiments were performed using mice that had received manganese using an implanted osmotic pump. The results show that SRR works well as a resolution-enhancing technique in IR-SE MEMRI experiments. The SRR IR-SE image visualizes a number of brain structures more clearly than the 3D GE image, including nuclei involved in specific higher brain function.

* Authors labeled with an asterisk (*) contributed equally to this work.

3.1 Introduction

Manganese-enhanced MRI (MEMRI) is commonly used in rodent studies to visualize neuroarchitecture, neuronal tract tracing, and neuronal activation. Mn^{2+} acts as a Ca^{2+} analog and upon neuronal activation, Mn^{2+} ions enter the neuronal cells through voltage-gated Ca^{2+} channels. Since the efflux of intracellular manganese is slow [89], Mn^{2+} will reside in the cells for days after accumulating in regions of the brain which have been activated. This has the advantage that brain activation performed outside the scanner in awake animals over multiple days can be visualized [7, 33, 41, 60].

In MEMRI studies, T1-weighted 3D gradient-echo (GE) [33, 42, 60, 90, 102, 106] or spin echo (SE) sequences [7, 52, 84, 96, 100] are most commonly used. A few groups have demonstrated that inversion-recovery spin echo (IR-SE) could provide superior image contrast [2, 52, 91]. For IR-SE sequences, however, *in vivo* 3D imaging is not possible within reasonable amounts of time, and instead multi-slice 2D images are acquired. The relatively thick slices of such images result in large partial volume effects which severely hampers visualization of sub-slice-thickness details.

Recently, an image post-processing technique called super-resolution reconstruction (SRR) has been demonstrated to improve visualization of the *in vivo* mouse brain [69]. SRR is the process of producing a high-resolution 3D MRI from a series of 2D slice-stacks, where each slice-stack transforms and samples the image space differently. A major benefit of using SRR in MRI is the possibility of going beyond the resolution limits of the scanner in the slice selection direction. There is consensus [65, 81] that super-resolution in MRI is not achievable in-plane, nor in true 3D acquisitions, since the Fourier encoding scheme excludes aliasing in frequency and phase encoding directions. However, in 2D slice-stacks in the through-plane direction there are no inherent limitations on the frequency spectrum, and aliased frequencies may potentially be recovered.

The aim of this study is to show that isotropic high-resolution 3D IR-SE images can be constructed by SRR using several 2D slice-stacks acquired in a time-frame compatible with *in vivo* experiments. Additionally, we aim to show that the IR-SE SRR images offer better contrast properties than the widely used gradient echo images, and as such constitutes a complementary imaging method.

3.2 Materials and methods

3.2.1 T1 measurements of the mouse brain after manganese administration

To determine the optimal IR parameters, one mouse received a cumulative dose of 300 mg $MnCl_2$/kg body weight (Sigma Aldrich, St Louis, MO, USA) via subcuta-

neously implanted 1007D Alzet osmotic pumps (DURECT Corporation, Cupertino, CA, USA) over 120h, as described in [70]. At the end of this period, T1 and T2 maps were acquired of one representative slice of the mouse brain using a multi-echo multi-TR RARE sequence on a 7T Bruker Pharmascan (Bruker, Ettlingen, Germany) connected to a computer equipped with Paravision 5.1 software (Bruker, Ettlingen, Germany). The following parameters were used: repetition time (TR)/echo time (TE) = 100,300,600,900,2000 ms /5,10,15,20,25,30 ms and a flip angle of 180°. The 5×6 acquired images were processed in Paravision 5.1 to obtain the T1 and T2 map. The region of interest analysis of the map revealed discrete T1 values between 650-750 ms, 900-1000 ms, 700-900 ms, 650-700 ms, 850-900 ms, 600-900 ms in the olfactive bulbs, cortex, hippocampus, superior colliculus, inferior colliculus and cerebellum, respectively.

3.2.2 Phantom experiments

We constructed a 2% agar phantom consisting of several isolated compartments containing manganese concentrations of 0.37 mM, 0.34 mM, 0.27 mM and 0.24 mM $MnCl_2$, with measured T1 values of 650, 720, 900 and 1000 ms, respectively. T1 values were measured with the same sequence as described above. The phantom was based on the findings in the pilot experiment described above, and the T1 values were chosen to cover the whole spectrum of values found in the manganese-enhanced brain. Three sets of 24 IR-SE slice-stacks with varying inversion times (TI) of 650 ms, 850 ms, and 1100 ms were acquired with the rest of the parameters being fixed (TR/TE = 2700/8 ms, 1 average, 30 slices of 1mm thickness, no gap, field of view (FOV) = 40×30 mm, matrix size = 200×150, in-plane isotropic resolution of 0.2mm/pixel, acquisition time (TA) per stack = 6 min). In each of the three sets, the slice-stacks were rotated in 24 increments of 180/24 degrees around the frequency encoding axis, see Fig. 3.1. From sub-sets of these 24 rotated acquisitions, images were reconstructed on an isotropic 0.2 mm resolution grid using SRR. The sub-sets contained 6, 8, 12, and 24 images; for all these sub-sets the rotations for each slice-stack were uniformly distributed.

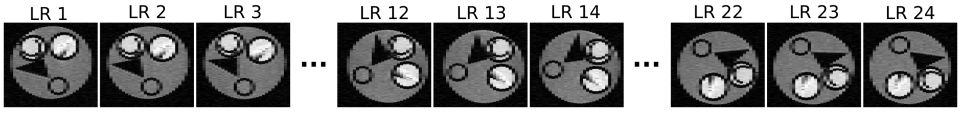


Figure 3.1: *Illustration of the rotated acquisition scheme. 24 anisotropic slices of the LR phantom acquisitions are shown. Rotation is performed around the phase encoding axis.*

For comparison, a 3D GE (FLASH) T1W image was acquired (TR/TE = 30/3.2 ms, flip angle = 30, 4 averages, FOV = $40 \times 30 \times 30$ mm, matrix size = $200 \times 150 \times 150$, isotropic resolution = 0.2 mm, TA = 45 min). All images were obtained with

a 7T Bruker Pharmascan (Bruker, Germany), using a 23 mm birdcage transmit-receive coil.

3.2.3 *In vivo* experiments

Two mice received a cumulative dose of 300 mg $MnCl_2$/kg body weight (supplier $MnCl_2$) via subcutaneously implanted 1007D Alzet osmotic pumps (DURECT Corporation, Cupertino, CA, USA) over 120h, as described in [70]. At the end of this period, the mice were scanned on a 7T Bruker Pharmascan (Bruker, Germany) using MRI protocols described below.

For reconstruction, sets of 2D stacks were acquired using an IR-SE sequence with a TR/TE/TI of 2700/8/670 ms, 1 average, FOV = 20 $\times$ 15 mm, matrix size 200 $\times$ 150, in-plane resolution of 0.1 mm and 30 slices with 0.5 mm slice thickness (no gap), with a TA = 6.75 min. The inversion time was chosen for optimal contrast in the hippocampal formation, based on the phantom experiment described above. 12 slice-stacks were acquired by rotating the field-of-view in uniform increments of 180/12 degrees around the phase encoding direction (inferior-superior).

For comparison, two additional scans have been acquired: a) a 2D IR-SE scan as described above, with 4 averages; b) a 3D T1W GE scan using a 3D FLASH sequence with a TR/TE of 30/3.6 ms, a flip angle of 30°, FOV = 20 $\times$ 15 $\times$ 12 mm, matrix size 200 $\times$ 150 $\times$ 120, an isotropic resolution of 0.1 mm, 5 averages, and a total scan time of 45 minutes. The field of view has been reduced for all the *in vivo* scans in order to acquire more anatomical detail, and this allowed also the addition of an extra average, so that the total scan time remained identical with the phantom scans.

3.2.4 Super-resolution reconstruction

SRR relies on two things: a set of acquisitions of the same scene/object that can be merged into a higher resolution image, and a model of the imaging process that can be used to solve the problem of merging the images in the most optimal way. By acquiring the LR data with rotational increments between the slice-stacks, as introduced in [83], a more effective sampling of *k*-space is achieved than by shifting the LR images by sub-pixel distances along the slice selection direction. In this work the slice-stacks were acquired using a rotated acquisition scheme. The imaging model consists of three distinct submodels: a parameterization of the object space, a model of the acquisition process, and a statistical model that appropriately describes the noise. An MRI image can be parameterized by a vector $\mathbf{x}$ of pixels (2D) or voxels (3D). The model of the acquisition process is given by a linear operator $\mathbf{A} = \mathbf{DBM}$, in which $\mathbf{M}$ describes a geometric transform, $\mathbf{B}$ is a space-invariant blurring model, and $\mathbf{D}$ is a downsampling operator. According to [34] the noise $\mathbf{n}$ in MRI can be assumed to be additive, white and Gaussian when the SNR > 3. Using and, the acquisition of the kth LR image, $\mathbf{y}_k$, can thus be described in vector form by $\mathbf{y}_k = \mathbf{A}_k\mathbf{x} + \mathbf{n}_k$

The objective in SRR is to solve the above equation over all k for $\mathbf{x}$. In general, a direct solution of this problem is not feasible since $\mathbf{A} \in \mathbb{R}^{n \times m}$, where n and m are the number of voxels in the reconstruction ($\mathbf{x}$) and in all the LR images ($\mathbf{y}$), respectively. Instead, iterative methods are usually applied to approximate the solution. Here we apply the low aliasing SRR scheme described in [71]. The SRR method is extended with a bias-field correction step removing inhomogeneities over the each of the images caused by the variations in coil sensitivity.

3.2.5 Quantitative evaluation measures

For quantitative evaluation of the phantom data we have applied a number of measures to the phantom reconstructions and the directly acquired GE image. The measures are based on cuboid sub-regions, Ω, of the image regions A, B, C, and D in Fig. 3.2. These regions contain manganese concentrations corresponding to T1 times of 1000 ms, 900 ms, 720 ms, and 650 ms , respectively. To obtain an accurate measure of the image noise, a large homogeneous sub-region of D was chosen. In the following, std(Ω_X) and mean(Ω_X) denote the standard deviation and mean of the sub-region, Ω_X, respectively, in the image region X, where $X \in A, B, C, D$.

- Noise: std(Ω_D)
- Signal-to-noise ratio (SNR): mean(Ω_A) / std(Ω_D)
- Contrast: mean(Ω_X) $-$ mean(Ω_Y)
- Contrast-to-noise ratio (CNR): (mean(Ω_X) $-$ mean(Ω_Y)) / std(Ω_D)
- Resolution: the average width of 20 edge profiles extracted from a slice in the image. The edge width of each profile was obtained by fitting a sigmoid function [69].

3.2.6 Qualitative evaluation

The SRR image and the GE image were rated by an experienced neuroanatomist (E.L.) for the visibility of the brain nuclei listed in Table 1, on a 3-point scale: not visible ($-$), visible but cannot be delineated ($+-$), or visible and can be delineated ($+$). The neuro-anatomist was requested to list and rate all the visible nuclei in the mouse brain. To avoid a potential bias towards one of the sequences, he was not informed about the aims of this study. Both the coronal and the axial views were used together in order to identify (coronal slices) and confirm (axial slices) the nuclei. In addition, a relative rating was done, specifying which of the two techniques provided the best contrast for each individual structure.

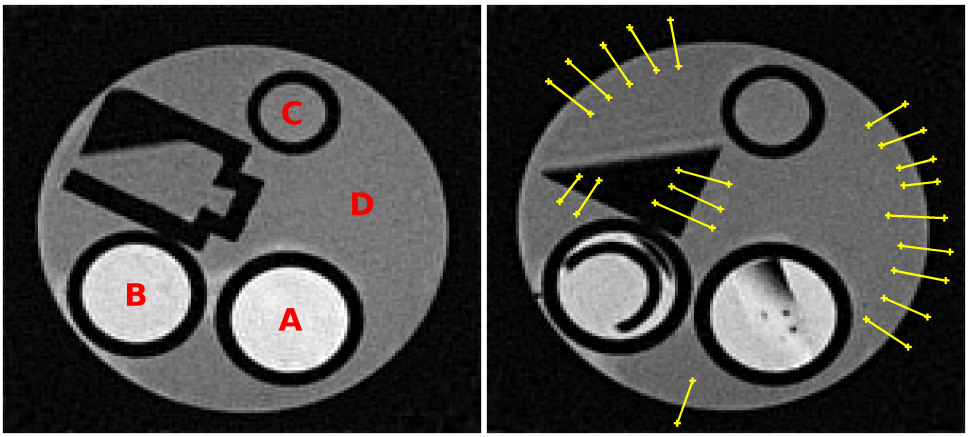


Figure 3.2: *The noise, contrast, SNR and CNR calculations are based on measurements inside the rectangles of the left sub-figure. The resolution measure is based on intensity profiles of 20 edge crossing strips in the right sub-figure.*

3.3 Results

3.3.1 Quantitative results of phantom acquisitions

Fig. 3.3 shows example slices from corresponding locations in the GE acquisition and in SRR reconstruction of IR-SE acquisitions with varying TIs of 650 ms, 850 ms, and 1100 ms using 24 slice-stacks. In the following we shall refer to the IR-SE reconstructions of varying TIs as *TI650*, *TI850* and *TI1100*, respectively. It is seen how the contrast between the regions (*D* and *A* in Fig. 3.3) of low (0.24 mM) and high (0.37 mM) manganese concentrations is considerably higher in the reconstruction of the *TI650* images than in the other IR-SE images and the GE image.

This observation is confirmed by the quantitative results in the middle row of Fig. 3.4. Here, it can be seen that the contrast between regions *A* and *D* is almost two times higher in the *TI650* images than in the GE image. The difference in CNR is not as pronounced because of the higher noise of *TI650* images. Compared to *TI850* and *TI1100*, both contrast and CNR are higher in the *TI650* images.

Contrast and CNR has also been measured between regions of smaller difference in manganese concentrations. In regions *A* and *B* the manganese concentration is 0.37 mM and 0.34 mM, respectively, resulting in T1 values of 650 and 720 ms. In regions *C* and *D* the concentrations are 0.27 mM and 0.24 mM, with T1 values of 900 and 1000, respectively. The results of these measurements are shown in the upper and lower rows of Fig. 3.4. The contrast between the regions with high manganese concentration *A* and *B*, is best in the GE images, while the contrast between

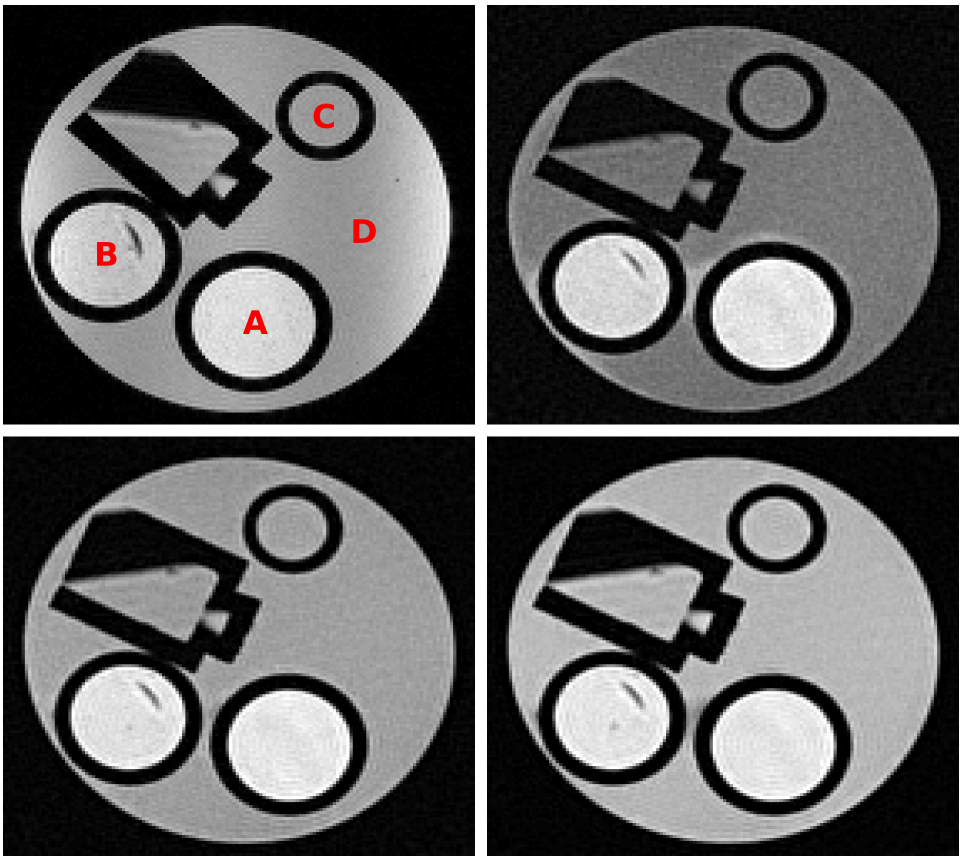


Figure 3.3: *Phantom juxtaposition: From top-left to bottom-right: GE scan, and IR-SE SRR images of TI=650 ms, TI=850 ms, TI=1100 ms. The four compartments, labeled A, B, C, and D, contain manganese concentrations of (A) 0.37mM, (B) 0.34mM, (C) 0.27mM and (D) 0.24mM $MnCl_2$, characterized by T1 values of 650, 720, 900 and 1000 ms, respectively. Each SRR image was reconstructed from 24 rotated LR images. The contrast in each image has been normalized to the range of the minimum and maximum intensities of the image.*

regions of low concentrations, C and D, as well as the contrast between regions of high versus low concentrations, A and D, is best for the *TI650* images. While these rankings hold for the CNR as well, it must be noted that the CNRs between these regions are low, especially between the low concentrations regions where it is < 1, making detection of small structures very difficult (see section 3.4).

Fig. 3.5 groups measurements of noise, SNR, resolution, and acquisition time of the IR-SE images and the GE image. Due to implicit averaging performed by

the SRR algorithm when redundant information is available (when we use more LR images than strictly necessary to cover the *k*-space), the noise in the reconstructions decreases. This explains the corresponding increase in SNR (and CNR in Fig. 3.4). The resolution of the images reconstructed by SRR increases substantially compared with the interpolated images. Even when only 6 LR images are used for the reconstruction, the resolution is substantially increased. Finally, in the lower right sub-figure of Fig. 3.5, the acquisition times are given to allow assessment of the acquisition time/image quality trade-off. Measurements on a single LR image, up-sampled by cubic interpolation to HR grid size, have been included in Figs. 3.4 and 3.5 (1 LR group) for reference.

3.3.2 Qualitative results of *in vivo* acquisitions

Fig. 3.6 shows corresponding coronal and axial slices extracted from the 2D reference scan, the SRR image and the GE image. The resolution enhancement due to SRR is obvious. Though the SRR image has a lower SNR than the GE image, a number of brain structures are more clearly discernible from the surrounding tissues in the SRR image than in the GE image. To exemplify the unique contrast properties of IR-SE MEMRI, arrows have been used in Fig. 3.7 to indicate the location of a number of identified nuclei. For the complete list of identified nuclei, please refer to Table 1.

In the present study, the MRI images were primarily inspected in the coronal direction, as this orientation corresponds to most histological brain atlases and is the most familiar orientation for neuroanatomists. Axial slices have been used to confirm the findings. A total of 41 nuclei were identified in either the GE (32 nuclei) or the SRR IR-SE image (32 nuclei). The visualization of these nuclei were subsequently compared between the images. Out of these 32 nuclei identified in the GE image, 9 are visible only in the GE image, 6 are best visible in the GE image, and 12 are equally visible in the GE and SRR IR-SE images. On the other hand, there are 9 nuclei which are visible only in the SRR IR-SE image, and 5 that are best visible in that image. With the TI chosen in this study, the following regions became visible in the reconstructed IR-SE image: the primary and secondary motor cortex, nucleus accumbens, the hypothalamic paraventricular nucleus, the periaqueductal grey, the superior and the inferior colliculi, substantia nigra, and the nuclei of lateral lemniscus. Also, the visibility of the following nuclei was improved: the nucleus of the lateral olfactory stria, the islands of Calleja in the olfactory tubercle, the habenula, the ventral pallidum and the central grey of the pons. The significance of this modified contrast for scientific neurological investigations will be discussed below. A list of biological functions, based on textbook literature [45], has been included for each mentioned nucleus/region.

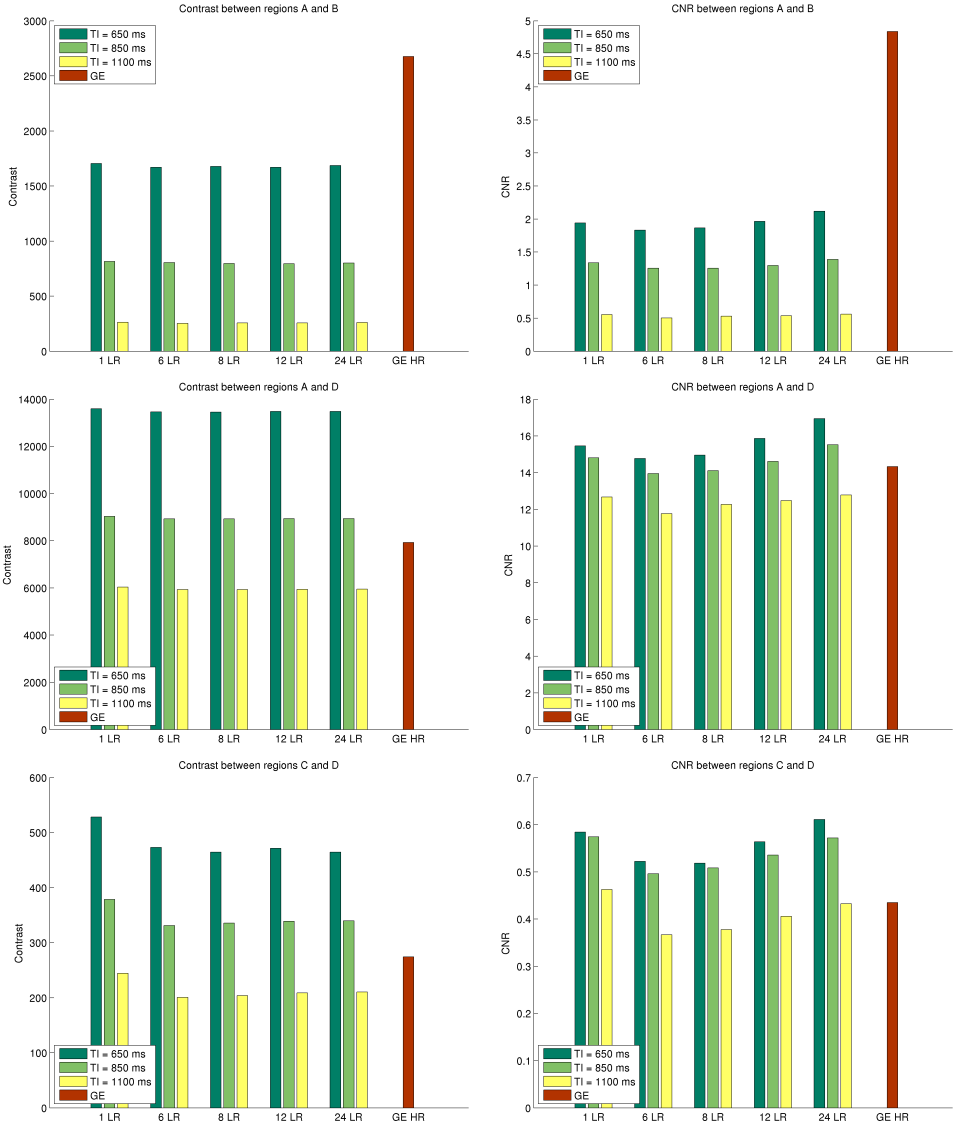


Figure 3.4: *Top row: Contrast and CNR between regions A and B. Middle row: Contrast and CNR between regions A and D. Bottom row: Contrast and CNR between regions C and D.*

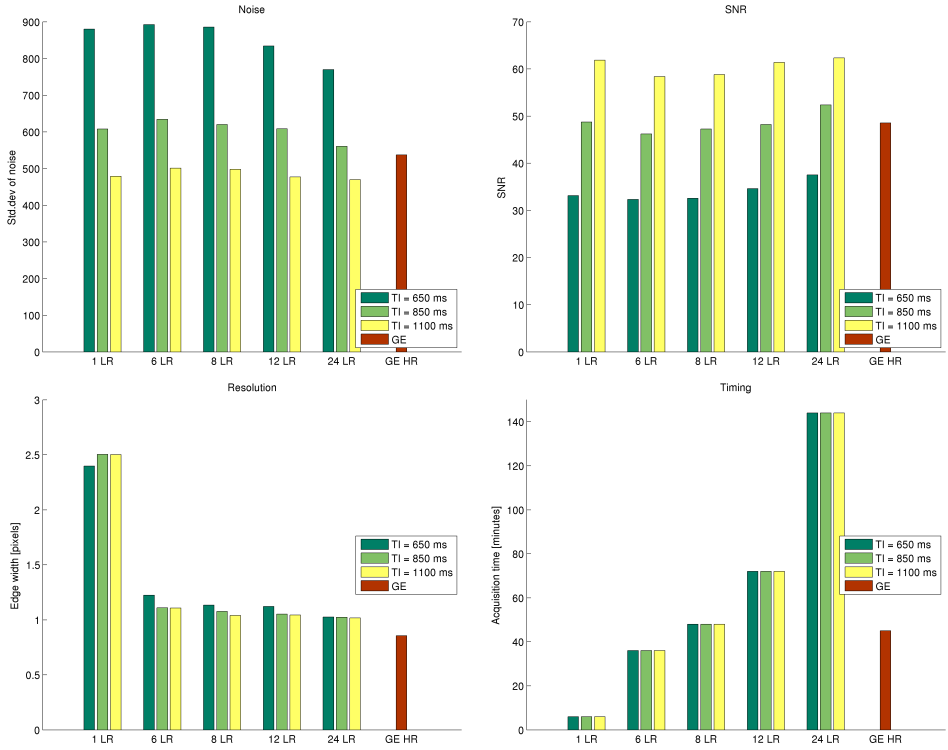


Figure 3.5: *Evaluation of quality and timing of the GE acquisition and SRR IR-SE images reconstructed from varying numbers of LR images. Top row: noise and SNR. Bottom row: resolution and acquisition time of SRR IR-SE images and the GE image.*

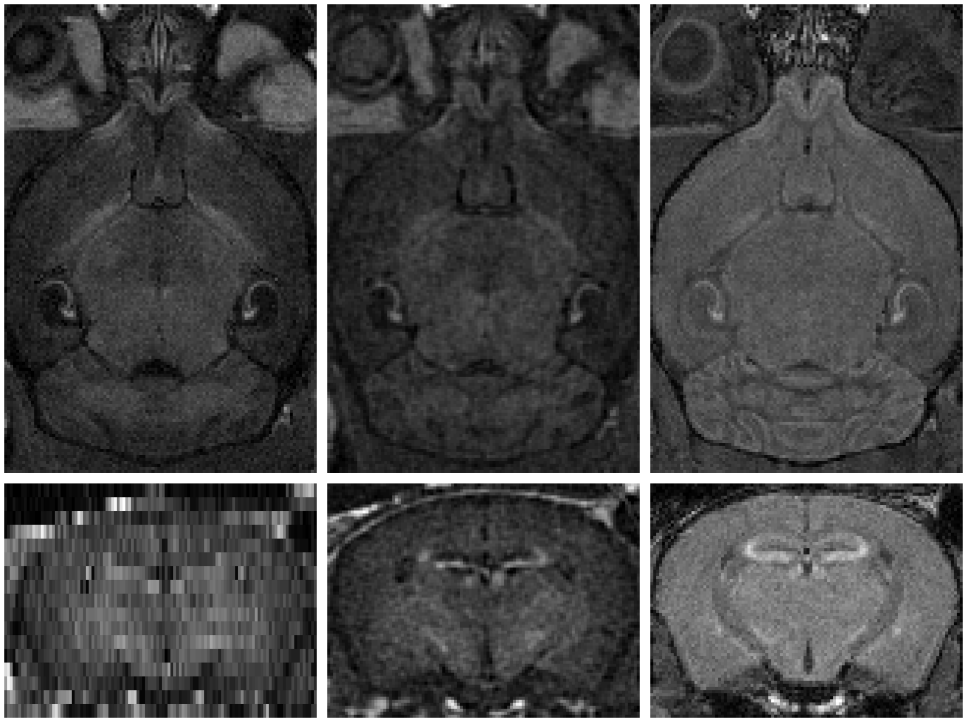


Figure 3.6: *Axial (top row) and coronal (bottom row) slices of the mouse brain.* Left*: Anisotropic LR IR-SE acquisition using 4 averages.* Middle*: SRR IR-SE image using 12 LR images.* Right*: GE image. Note that the LR IR-SE and the SRR IR-SE slices do not correspond exactly due to a small rotational offset around the inferior-superior axis. The contrast in each image has been normalized to the range of the minimum and maximum intensities of the slice.*

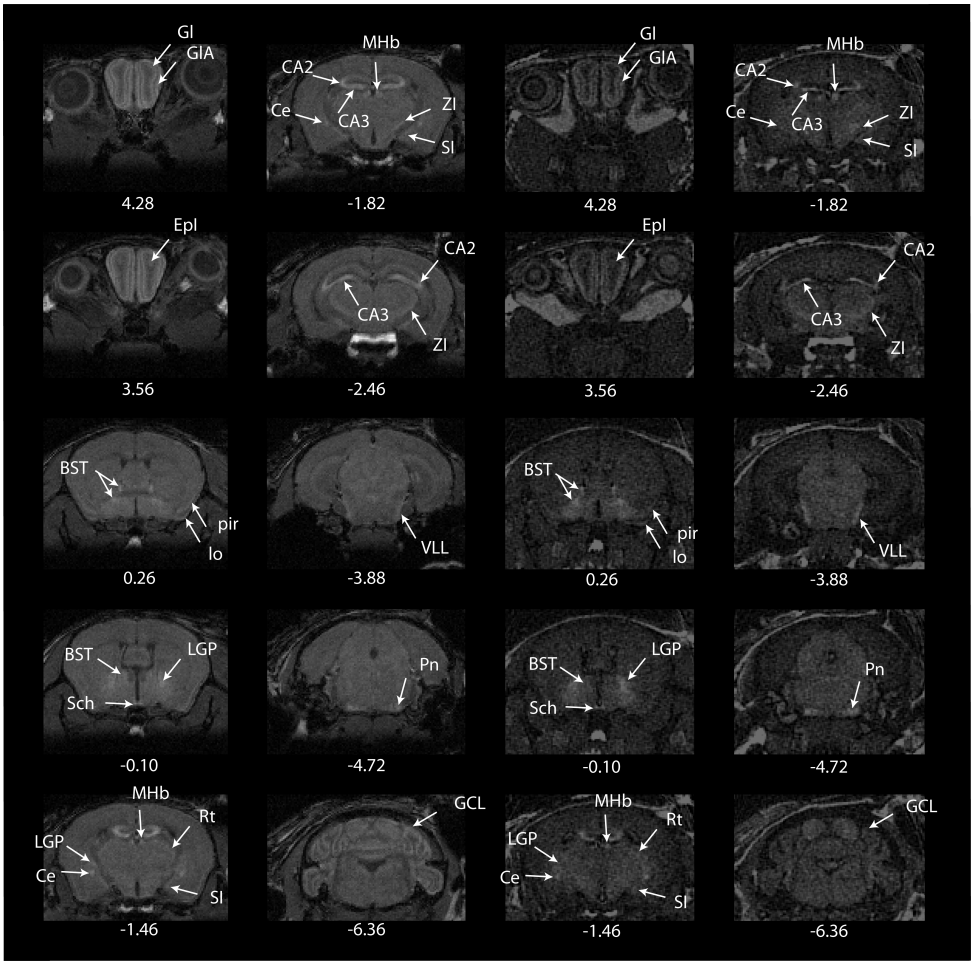


Figure 3.7: *Comparative cranial view of the main brain nuclei (partially) visible in the GE T1W image (two left columns) and the SRR IR-SE image based on 12 rotations (two right columns). The Bregma coordinate is specified under each image. The abbreviations used in the image are specified below. Gl: glomerular layer of olfactory bulb; GlA: glomerular layer of accessory olfactory bulb; Epl: external plexiform layer of the olfactory bulb; BST: bed nucleus of stria terminalis; pir: piriform cortex; lo:lateral olfactory tract; LGP: lateral globus pallidus; Sch: suprachiasmatic nucleus; MHb: medial habenular nucleus; Ce: central amygdaloid nucleus; Rt: reticular nucleus of thalamus; SI: substantia innominata; ZI: zona incerta; CA2: field CA2 of hippocampus; CA3: field CA3 of hippocampus; VLL: ventral nucleus of lateral lemniscus; Pn: pontine nuclei; GCL: granular cell layer of the cerebellum.*

Anatomic description of enhanced region	GE	IR-SE SRR	GE vs IR-SE SRR	Main biological function of enhanced region
glom. layer of olf. bulb	+	+	=	odor decoding
glom. layer of acc. olf. bulb	+	+-	GE	pheromones detection
ext. plexiform layer of olf. bulb	+	+	=	synaptic processing in olfaction
frontal lobe	+-	-	GE	cognition, action planning, movement
primary motor cortex	-	+	SRR	high-level motor control, organization of voluntary movement
supplementary motor area	-	+	SRR	learning sequences of precise movements
piriform cortex	+-	+-	=	olfactory information processing
nucleus of lateral olf. stria	+-	+	SRR	integration of olfactory information
Islands of Calleja	+-	+	SRR	processing of olfactory information
nucleus accumbens	-	+	SRR	regulation of various addictions, pleasant emotions
bed nucl. of stria term. - med. ant. div.	+	+	=	sexual dimorphism
bed nucl. of stria term. - med. div.	+-	-	GE	mediates responses to fear/stress stimuli
bed nucl. of stria term. - lat. div.	+	+	=	diff. regulation of HPA activity
hippocampus - dentate gyrus	+	+	=	formation of memories
hippocampus - CA1 field	+-	-	GE	formation of memories
hippocampus - CA2 field	+	+	=	formation of memories
hippocampus - CA3 field	+	+	=	formation of memories
central amygdaloid nucleus	+	+	=	response to fear/stress stimuli, perception/storage of emotion
reticular nucleus	+	+-	GE	propagation of primary generalized seizures
paratenial nucleus	+-	-	GE	sleep regulation
substantia innominata	+-	-	GE	enhancing cortical response to sensorial stimulation
suprachiasmatic nucleus	+	+	=	control of circadian rhythms
paraventricular nucleus	-	+-	SRR	regulation of autonomic responses
medial preoptic nucleus	+-	-	GE	control of homeostasis, and hormonal and reproductive activity
zona incerta	+	+-	GE	limbic-motor integration
lateral habenula	+	+	=	mediation of dopaminergic activity
medial habenula	+-	+	SRR	mediation of dopaminergic activity
pineal gland	+	-	GE	produces melatonin, regulation of sleep-wakefulness cycles
periaqueductal gray	-	+-	SRR	role in analgesia
superior colliculus	-	+	SRR	major visuo-motor integration region
inferior colliculus	-	+	SRR	sound localization, audio-motor integration
substantia nigra	-	+-	SRR	inhibitory dopaminergic regulation of thalamus and brain stem
ventral pallidum	+-	+	SRR	inhibitory GABA-ergic regulation of thalamus and brain stem
lateral globus pallidus	+	+	=	regulation of voluntary movement
caudate putamen (striatum)	+-	-	GE	movement, reinforcement and implicit learning, neg. emotions
pontine central gray	+-	+	SRR	relays sensorial and motion information (cortex to cerebellum)
nuclei of the lateral lemniscus	-	+	SRR	localization of sound sources
cereb. cortex - gran. cell layer	+	-	GE	motor learning, spatial accuracy, temporal coord., balance
substantia gelatinosa	+	+-	GE	nociception
internal capsule	+	+-	GE	cortico-medullar sensori-motor connection
fornix	+	+-	GE	hippocampal-hypothalamic connection, part of limbic system

Table 3.1: *Overview of positively enhanced brain nuclei identified in the SRR IR-SE and/or GE images. The level of visibility is indicated by (+) for clear positive enhancement (the brain nucleus can be visualized and delineated) (+-) for visible enhancement (the brain nucleus is visible but cannot be delineated) and with (-) for no positive enhancement.*

3.4 Discussion

3.4.1 Phantom experiments

For a well-chosen TI, in our case 650 ms, IR-SE yields better contrast than a GE sequence. Using SRR, the resolution of the IR-SE images can be improved dramatically, making these images a potential alternative to GE imaging in MEMRI. In the quantitative experiments of this study, we have confirmed results from previous studies [91] showing the unique and superior contrast of IR-SE MEMRI. In addition, we have shown that using SRR we can achieve a resolution that is close to that of the GE image, within an acquisition time that is compatible with *in vivo* experiments.

As seen in Figs. 3.4 and 3.5, there seems to be a correlation between the contrast and the noise in the images. The high-contrast *TI650* images, for example, have a high noise-level as well. These images also have a slightly lower resolution than the *TI850* and *TI1100* images. There is no theoretical optimum in the trade-off between contrast, noise and resolution but as a rule of thumb, the well-established Rose criterion [35] requires a minimum "visibility" of around 4 for detection of an object by human observers. An object's "visibility" depends on the CNR and the size of the object in number of voxels, N, and is defined by $\mathrm{CNR}\sqrt{N}$. In 3.4, we see that the CNR between regions A and D is well above 4 for all IR-SE images and the GE image, and it is very likely that single voxel objects ($N = 1$) can be distinguished. As for the regions A and B, the CNR of the IR-SE images is lower than 4, and we cannot expect to be able to separate single-voxel objects. The CNR of *TI650* images is 2 and an object should thus consist of at least 4 voxels for it to be visible according to the Rose criterion. Between the regions C and D the CNR of the IR-SE images is higher than that of the GE image. The largest CNR (*TI650* using 24 slice-stacks) is, however, only slightly larger than 0.6. For this CNR, objects must contain 40-50 voxels to be visible under the Rose criterion.

Based on the quantitative results, we chose to use IR-SE SRR images based on 12 LR images for the qualitative evaluations of the visibility of mouse brain nuclei. While the acquisition time of 12 LR images (1 hour 21 minutes) is longer than that of the GE image (45 minutes), it remains compatible with *in vivo* experiments. If shorter acquisition times are required, a trade-off with resolution and CNR is necessary. In Fig. 3.5 it is seen, that if 8 or 6 LR images are used for SRR reconstruction, the acquisition time is reduced by 25% and 50%, respectively, while the resolution and SNR decrease by much smaller percentages.

3.4.2 *In vivo* experiments

Due to its high resolution, a 3D image enables the identification of small nuclei that are not visible in the relatively thick slices of a 2D image (see Fig. 3.6). As seen in Fig. 3.7, a considerable number of manganese-enhanced brain nuclei can be

identified in 3D MEMRI images acquired with either a GE sequence or produced by SRR IR-SE. Some of these nuclei are equally visible in both types of images, some are better visible in one image type than in the other, and yet others are not visible at all in one of the two types of images. The reason for this resides in the choice of the inversion time.

In the present study, based on the phantom experiments that simulated the experimentally determined T1-values of hippocampus and cortex after manganese administration, TI was chosen to yield maximum contrast between these two regions (TI = 650 ms), while still retaining sufficient background signal to identify the brain. This approach sensitizes the acquisition to nuclei with T1 values in this range. By selecting a different TI, different nuclei can be brought forth. TI can thus be tuned to optimally visualize the nuclei of interest. In contrast, a standard 3D GE sequence results in visualization of manganese-enhancement over a broader range of tissue T1 values. Thus this sequence may be more suitable for exploratory studies.

Structures that were more clearly identified using IR-MEMRI included many that are of interest in neuroscience. In particular, the study of planned, coordinated and oriented motor activity benefits from the improved contrast in the IR MEMRI image. Nuclei of interest include the mouse motor primary and supplementary cortex, which are of central importance for the investigation of the motor control, organization of voluntary movement, as well as learning and precise coordination of complex movements. The pontine central grey is also involved in the motion coordination, by relaying (proprio)sensorial and motion information to the cerebellum, the superior colliculus integrates the visual and motor sensorial information, while the inferior colliculus is responsible with sound localization and audio-motor integration. The lateral lemniscus is also contributing to the localization of sound sources. The spatial orientation of the mouse during motion cannot be considered in the absence of olfactive information - the islands of Calleja and the nucleus of the lateral olfactory stria are involved in the processing and integration hereof.

Other processes of interest involving manganese-enhanced nuclei include the regulation of autonomic responses by the paraventricular nucleus and, together with the hippocampal regions and the amygdala, the response to stress and its related pathologies. The ventral pallidum and substantia nigra are involved in the inhibitory regulation of thalamic and brain stem activity via GABA-ergic and respectively dopaminergic activity, and are crucial for the understanding of reward mechanisms, various addictions and a wide range of positive emotions. The habenula modulates dopaminergic activity and nucleus accumbens is involved as well in the regulation of a wide range of pleasant emotions and addictions. Last but not least, the role of the periaqueductal grey in the maintenance of consciousness, in analgesia, as well as its involvement in various forms of behavior, such as defensive and sexual actions, is indisputable [45].

It is worth mentioning that the regions mentioned above are enhanced due to

the fact that they are all involved in the regular daily activity of the mouse. When disorders are modelled in manganese-administered mice and induce a markedly increased or impaired activity in brain regions other than the ones mentioned above, the optimization of the TI prior to the investigation is a requirement for a good visualization of these regions.

3.4.3 Conclusion

In conclusion, we show in this study that while the classical approach to MEMRI of the mouse brain by means of 3D T1W MRI is appropriate for exploratory studies, super-resolution reconstructed IR-SE 3D images allow a targeted investigation of brain activation patterns and is a valuable complementary tool for manganese-enhanced studies. Moreover, the methods presented here are equally valuable for other applications, such as white matter lesion detection, where high-resolution 3D IR-SE would provide better contrast and resolution.

3.5 Acknowledgements

NWO-VENI (700.56.407), NWO-Athena (700.58.80), Medical Delta.

Chapter 4

Interactive local super-resolution reconstruction of MRI whole-body mouse data: applications to bone and kidney metastases

Based on:

E. Plenge*, A. Khmelinskii*, O. Dzyubachyk*, P. Kok*, T.J.A. Snoeks, D.H.J. Poot, C.W.G.M. Löwik, C.P. Botha, W.J. Niessen, L. van der Weerd, E. Meijering, and B.P.F. Lelieveldt. Interactive local super-resolution reconstruction of MRI whole-body mouse data: applications to bone and kidney metastases. *In review*.

Abstract

A resolution enhancing post-processing technique called super-resolution reconstruction (SRR) has recently been demonstrated to improve visualization and localization of micro-structures in small animal MRI. In such cases, however, the size of the animal under investigation can be very large relative to the size of the structures of interest. Solving the SRR problem thus becomes very expensive, in terms of both computation time and memory requirements. In this chapter we introduce local SRR to overcome these problems. We present a novel method that combines state-of-the-art image processing techniques from the areas of articulated atlas-based segmentation, planar reformation and SRR, and that allows researchers to efficiently capture both global and local scales in whole-body small animal MRI. The approach is validated in two case studies involving CT, BLI and MRI data of bone and kidney tumors in a mouse model. We show that local SRR MRI is a computationally efficient complementary imaging modality for the precise description of tumor metastases, and that it provides a high-resolution alternative to conventional MRI.

* Authors labels with an asterisk (*) contributed equally to this work.

4.1 Introduction

In pre-clinical small animal research on skeletal complications of cancer, imaging modalities like bioluminescence (BLI), CT, and MRI are conventionally used. Such imaging techniques allow non-invasive studies on the metastatic behavior of tumors [87]. BLI gives an indication of metastatic tumor growth anywhere in the body (e.g. bones, liver and lungs), but the spatial resolution is not sufficient to distinguish between lesions located in close proximity to each other and to actually localize all individual metastatic processes in an organ. CT gives excellent contrast in calcified tissue and can be used to study tumor-induced changes in the bone, but due to lack of soft tissue contrast it is less suitable to image organs such as liver and lungs. MRI is the preferred imaging modality for imaging liver and lung metastases as it gives sufficient anatomical detail and good contrast between the organs and tumor masses. So, whereas BLI can be used to indicate the total tumor burden in an organ, MRI will give information on the exact location, size and number or metastatic lesions in that organ. Together, CT, MRI and BLI provide a comprehensive picture of the tumor and metastases development and spread in the entire body.

The sensitivity of MRI for small lesions is, however, relatively low compared to BLI, and the most robust pre-clinical protocols are still 2D MRI experiments, with relatively thick slices. This slice thickness results in a large partial volume effect, making precise detection and localization of tumors difficult, especially for early stage tumors and micro tumors [27]. Recently, a resolution enhancing post-processing technique called super-resolution reconstruction (SRR) has been demonstrated to improve visualization and localization of micro-structures in molecular MRI [69]. In a metastatic disease model, however, the size of the object under investigation (the mouse/rat) relative to the size of the structures of interest (the tumors) can be very large. When attempting to capture both global and local scales in an image, this translates into a large field of view at high image resolution, resulting in images of tens of millions of voxels. In such cases, solving the SRR problem becomes very expensive, in terms of both computation time and memory requirements. Exploring large data-sets in this way calls for conceptual new-thinking.

In this study, we overcome the computational issues of whole-body SRR by the combination of state-of-the-art methods from the areas of articulated atlas-based segmentation of whole-body small animal data [3, 46, 47, 48, 51], planar reformation [50], and SRR in MRI [69, 71]. We integrate these concepts into a novel localized approach to SRR that enables global-to-local exploration of e.g. whole-body mouse MRI data. The idea is similar to that of well-known web-based geographical maps, where it is possible, from a global overview image, to zoom in on a detail of interest. Guided by user interaction or by registration to images of higher sensitivity, such as BLI, local volumes of interest (VOIs) can be identified in the low-resolution MR image. From the global low-resolution overview these VOIs are then enhanced by SRR to show a higher level of detail.

The aim of our study is thus to determine whether SRR-MRI is a feasible method for improving visualization of tumors in small animal imaging. By feasibility we refer to two aspects:

- Image quality: when the number of low-resolution images is constrained by acquisition times compatible with *in vivo* experiments, does our local SRR method improve the visualization of small anatomical details over conventional imaging methods?
- Computational feasibility: can the local SRR computations be handled on a desktop machine in a close-to-real-time time frame?

In the following sections, we first introduce our approach to local SRR in MRI. We briefly describe its components (for details we refer to previously published work in which each of the components has been thoroughly validated) and validate our approach in two case studies with bone and kidney tumor visualization, respectively.

4.2 Materials and methods

4.2.1 Experimental mouse model and imaging

To test the SRR approach, BLI, CT and MRI were acquired in a mouse model of metastasizing breast cancer. One female, *Balb/c nu/nu* mouse of 19.5 g was used. At 7-8 weeks of age, the mouse was injected with 4T1-luc2 [8, 49] breast cancer cells (100 µl, 150,000 cells) into the left heart ventricle under 2% Isoflurane anesthesia.

After 2-3 weeks BLI and CT scans were made *in vivo*. The anesthesia applied was Ketamine:RomPun:PBS (1:1:1), approximately 60 µl/20 g. This was followed by an *ex vivo* MRI scan. The mouse was euthanized to allow flexibility in the MRI experiments and test different acquisition parameters.

The mouse, lying on its belly, was taped to an in-house made PMMA holder that was used in all three scanners. For BLI imaging, the Xenogen^{TM} VIVO Vision IVIS 3D scanner (Alameda, CA, USA) was used. The acquisition was done at a range of wavelengths between 580-680 nm, at 20 nm intervals with an acquisition time of 10 seconds per wavelength. One of the 8 BLI images is presented in Fig. 4.1.

Micro CT data was acquired on a SkyScan^{TM} 1076 *in vivo* X-Ray Microtomograph microCT scanner (Aartselaar, Belgium) at a resolution of 35 µm. The acquisition was performed with a step size of 1.4 degrees over a trajectory of 360 degrees (voltage = 49 kV, current = 200 µA, exposure time = 100 ms, filter: AL. 0.5 mm, frame averaging = 1).

Several strategies can be adopted when acquiring MR data for an SRR experiment. By acquiring the low-resolution slice-stacks with rotational increments around either the frequency or the phase encoding direction, as introduced in [83], a more

effective sampling of *k*-space is achieved than by shifting the low-resolution images by sub-pixel distances along the slice selection direction. In this fashion, a whole-body scan of the *post mortem* mouse was acquired on a 7T Bruker Pharmascan^{TM} system using a fast spin echo (FSE) sequence. TR was 5300 ms, TE was 53.2 ms, with N_{avg} = 4. The 2D slice-stack consisted of 40 slices (0.5 mm thick), with a FOV of 70 $\times$ 45 $\times$ 20 mm, and a resulting resolution of 0.125 $\times$ 0.125 $\times$ 0.5 mm. The scan time per stack was 13 min. The slice-stack was acquired at 4 angles with uniform increments of 180/4 degrees around the phase encoding direction. In this study, we performed SRR on sub-sets of two and four low-resolution images. In the sub-set of two images, the angular increment between them was 180/2 degrees.

To compare the SRR results with a standard fast scan, an additional 2D scan of 5 minutes was performed. A FSE sequence was used, with a slice-stack that consisted of 40 slices (0.5 mm thick), with a FOV of 70 $\times$ 45 $\times$ 20 mm, a resolution of 0.25 $\times$ 0.25 $\times$ 0.5 mm TR = 5300 ms, TE = 55.452 ms, N_{avg} = 4.

Animal experiments were approved by the local committee for animal health, ethics and research of Leiden University Medical Center.

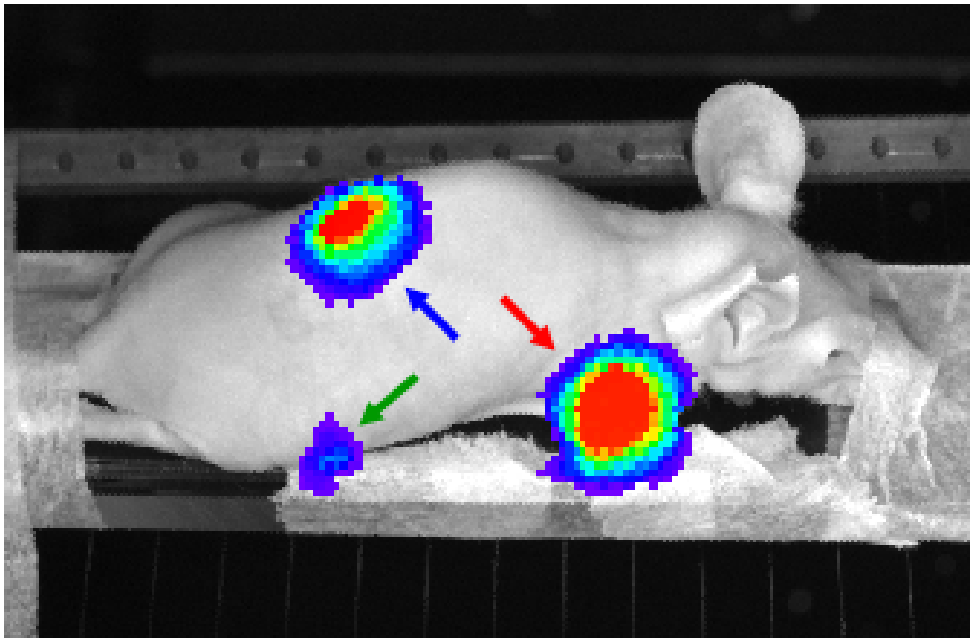


Figure 4.1: *A BLI image of the mouse image to validate the proposed approach. The arrows indicate the different tumor locations: humerus (red), femur (green), kidney (blue).*

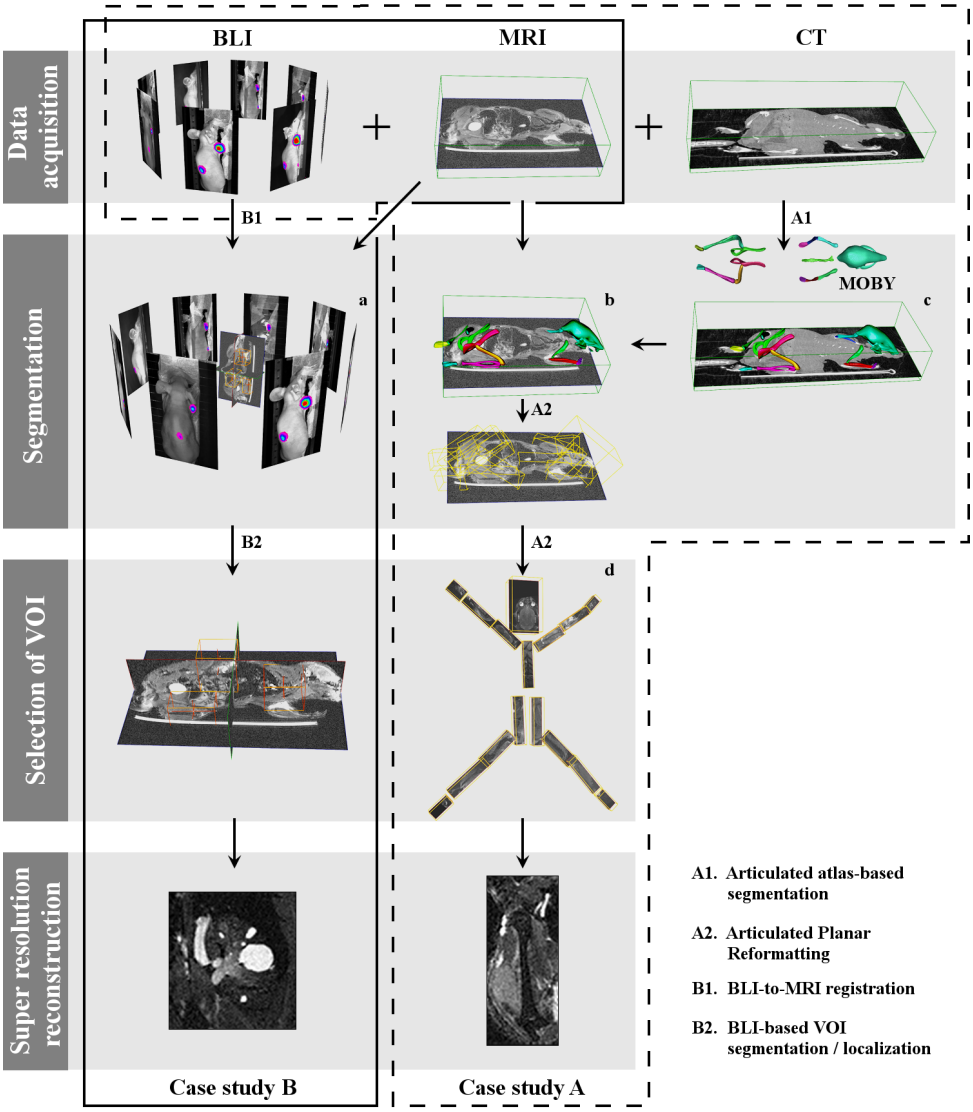


Figure 4.2: *Overview of the interactive local SRR of MRI mouse data applied to two case studies.* A*: After rigidly registering CT to MRI, articulated atlas-based segmentation is performed (*A1*). Subsequently, articulated planar reformation is applied to the segmented MRI and the data is visualized in the standardized atlas space (*A2*). The user can now interactively select any low-resolution bone of interest guided by the BLI images. A high-resolution SRR image of the humerus with a tumor is presented.* B*: BLI+MRI mouse data are first co-registered (*B1*) to define the VOIs (*B2*) using the BLI. A VOI is interactively selected for performing SRR. A high-resolution SRR image of the kidney with metastases is presented.*

4.2.2 Interactive local SRR

The local SRR method integrates a series of processing and analysis steps, which depend on the available complementary data (CT, BLI, etc.) and vary in their level of user interaction. The overview of the presented method can be seen in the flowchart depicted in Fig. 4.2: First, within of a set of low-resolution MRI images, potential VOIs are identified. In our approach this step is either based on user input or it is automated, as described below. Its output is one or more VOIs containing potentially relevant structures. One or more of these VOIs can now be selected for subsequent local SRR.

The methods for segmentation and selection of VOIs are highly specific to the biological problem and to the available complementary data. In the following, we present two situations typical in small animal tumor imaging, in which BLI+CT (case study A) and BLI only (case study B) are used as complementary modalities to MRI (see Fig. 4.2). Each situation presents a different level of automation and requires a different degree of user interaction. The way the relevant information is extracted differs with the choice of the imaging modalities for the study at hand. In case study A, the level of user intervention is minimal. The whole-body mouse is automatically segmented using an articulated atlas. Guided by the BLI, the user can then select the VOIs with tumors for further SRR reconstruction, visualize the results side-by-side with the CT data, and, in case a tumor is present near a bone on one side of the body, compare it to the contralateral side, where most likely there is no tumor. In case study B, user interaction is necessary to co-register the BLI to the MRI data to define the VOIs. After that, the user can select among the VOIs in which the BLI signal indicate the presence of tumors for SRR reconstruction and further high-resolution visualization and analysis.

4.2.3 Case study A: MRI+CT+BLI

This case study was set up to explore the value of SRR-MRI to image bone metastases as a complementary modality to CT, BLI and conventional MRI. In this section, we describe our approach to super-resolution bone MRI.

Articulated atlas-based bone segmentation of CT and MRI mouse data

First, rigid registration of the CT scan to one of the low-resolution MR images was performed [51]. Rigid registration was sufficient in this case because the mouse was fixated in the same animal bed during all imaging procedures and during transport between scanners. The bones were segmented in the CT image using the articulated MOBY mouse atlas [47, 82] (Fig. 4.2c). The fully automated approach presented in [3] was used for this purpose. To deal with the large articulations between bones and/or bone groups, the registration of the atlas to the CT data used a hierarchical model tree. First, a coarse alignment of the MOBY atlas to the CT skeleton was performed. This was followed by the stepwise alignment of the individual atlas bones

to the CT data, using the ICP algorithm [6]: we started with the skull, after which each bone was accurately registered to the correspondent bone in the data. Given the CT-to-MR registration parameters, the transform obtained in the segmentation of the whole-body CT data was propagated to the MR. Figs. 4.2b and 4.2c show the atlas fitted to the CT and MRI data-sets, respectively.

Articulated planar reformation of MRI data

Using the obtained transformations between each bone in the atlas and the low-resolution MR image, articulated planar reformation [9] can be applied to map the labeled data into a standardized atlas space. This method automatically creates a VOI for each bone, which is based on a principal component analysis of the bone shape. By constructing the VOIs in this manner, the final reformatted images are aligned with the principal axes of the bones [50].

Interactive selection of VOI

Upon segmentation and reformatting, the user is presented with a global view of the segmented bones, see Fig. 4.2d. Guided by the BLI signal, the user can now interactively select a bone for SRR reconstruction.

4.2.4 Case study B: MRI+BLI

This case study was set up to explore the value of SRR-MRI as a complementary modality to BLI and conventional MRI when CT data is not available for establishing anatomical correspondence. In practice, this is usually the case for soft tissue tumors, where bone metastases and bone resorption are not expected. In this section, we describe our approach to super-resolution MRI of kidney tumors.

BLI-to-MRI mouse data registration

After acquisition, the BLI images are registered to one of the low-resolution MR images using a landmark-based approach [51]. A minimum of three landmarks is selected. The location of each landmark is indicated in one of the low-resolution MR images and in two separate BLI images at different angulations. Using the known angle between the two BLI images, back projection is used to find the corresponding point in the three-dimensional space. This point is then paired with the point in the MR image and registration is performed. Typical landmarks include the snout and limbs because they are most easily identified in both modalities.

BLI-based VOI localization and segmentation

BLI-based VOIs can be localized by simple thresholding on the raw BLI signal. Once the coordinates of the VOIs in world space are known, the BLI-to-MRI registration transform is used to map the VOIs onto the chosen low-resolution MR

image. The VOIs are then propagated to the remaining low-resolution MRI images using the transform parameters of these acquisitions. Finally, VOIs are extracted and can be used for SRR.

4.2.5 Super-resolution reconstruction

When a VOI has been selected and propagated to all rotated low-resolution MRI images, local SRR can be performed on the volume. SRR is the process of producing a single high-resolution image from a sequence of low-resolution images, where each low-resolution image transforms and samples the high-resolution scene in a distinct fashion. It is an inverse problem in which the acquisition process is modeled as a linear operator on the high-resolution image. When the high-resolution image is vectorized and put into a large vector, the acquisition of the low-resolution image k can be modeled as $\mathbf{y}_k = \mathbf{A}_k\mathbf{x} + \mathbf{n}_k$, where $\mathbf{n}$ is Gaussian noise [34]. The linear operator $\mathbf{A}_k$ models the transform due to the rotation of the field of view of the kth image as well as the point spread function of the acquisition.

The objective in SRR is to find an $\mathbf{x}$ that minimizes the difference between $\mathbf{y}_k$ and $\mathbf{A}_k\mathbf{x}$ for all k simultaneously [69]. In general, a direct solution of this objective is not feasible since it involves many operations with all $\mathbf{A}_k \in R^{n\times m}$, where n and m are the number of voxels in the reconstruction ($\mathbf{x}$) and in a low-resolution image ($\mathbf{y}$), respectively. Instead, the reconstruction is obtained by iterative methods such as the conjugate gradient method. In this study, we apply the method described in [71], which is a Tikhonov-regularized least-squares solver that implements an affine transformation scheme that minimizes aliasing and spectral distortion. The SRR method is extended with a bias-field correction step removing inhomogeneity over the images caused by variations in coil sensitivity.

4.2.6 Software platform

All computational experiments described beneath were implemented in MATLAB R2009b^{TM} and performed on a 2.80 GHz Intel Xeon^{TM} with 12 GB RAM, Windows^{TM} PC.

4.3 Results

4.3.1 Case study A: MRI+CT+BLI (bone tumors)

Local SRR images of the right femur and humerus with metastases were reconstructed at different levels of quality using 2 and 4 low-resolution images, and compared with conventional MRI, BLI and CT. In addition, reconstruction times of individual bones were compared with that of the entire mouse.

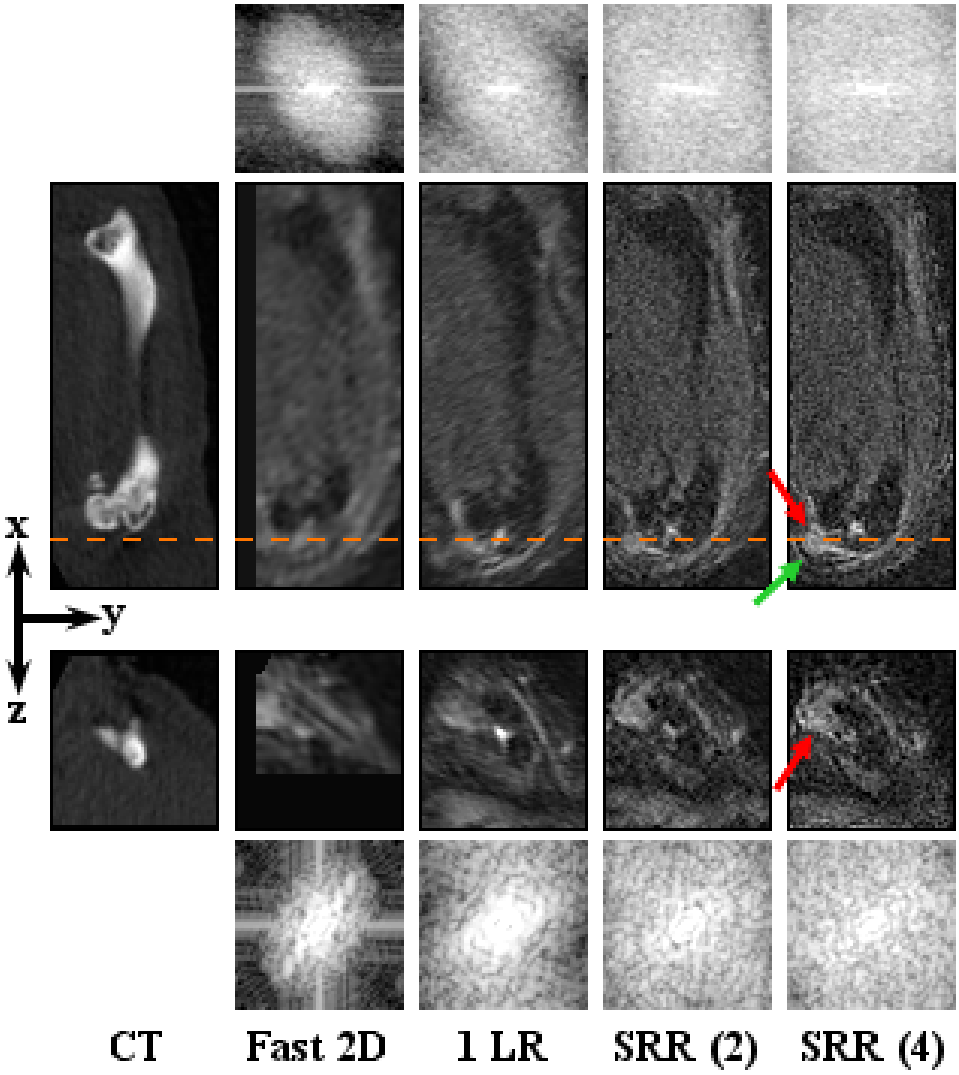


Figure 4.3: *Right femur. From left to right: a CT scan, a fast MRI scan, one low-resolution image, and SRR reconstructions, each based on a different number of low-resolution images. Two orthogonal slices of the same VOI are shown to illustrate the effect of the SRR in a 3D volume. The orange dashed line indicates where the yz-slice (bottom) intersects the xy-slice (top). The red arrows points to the (micro) tumor in the knee. The green arrow points to another location, outside the tumor, at which recovery of the fine details is obvious. The CT and all the MR images are shown in the coordinate system associated with the principal axes of the bone, and Fast 2D and the low-resolution (1 LR) volumes are resampled to isotropic resolution beforehand. Image contrast on the MRI images was increased for visualization purposes. For all MR images the corresponding frequency spectra are shown, above and beneath the xy-slices and the yz-slices, respectively, demonstrating enhanced high-frequency content.*

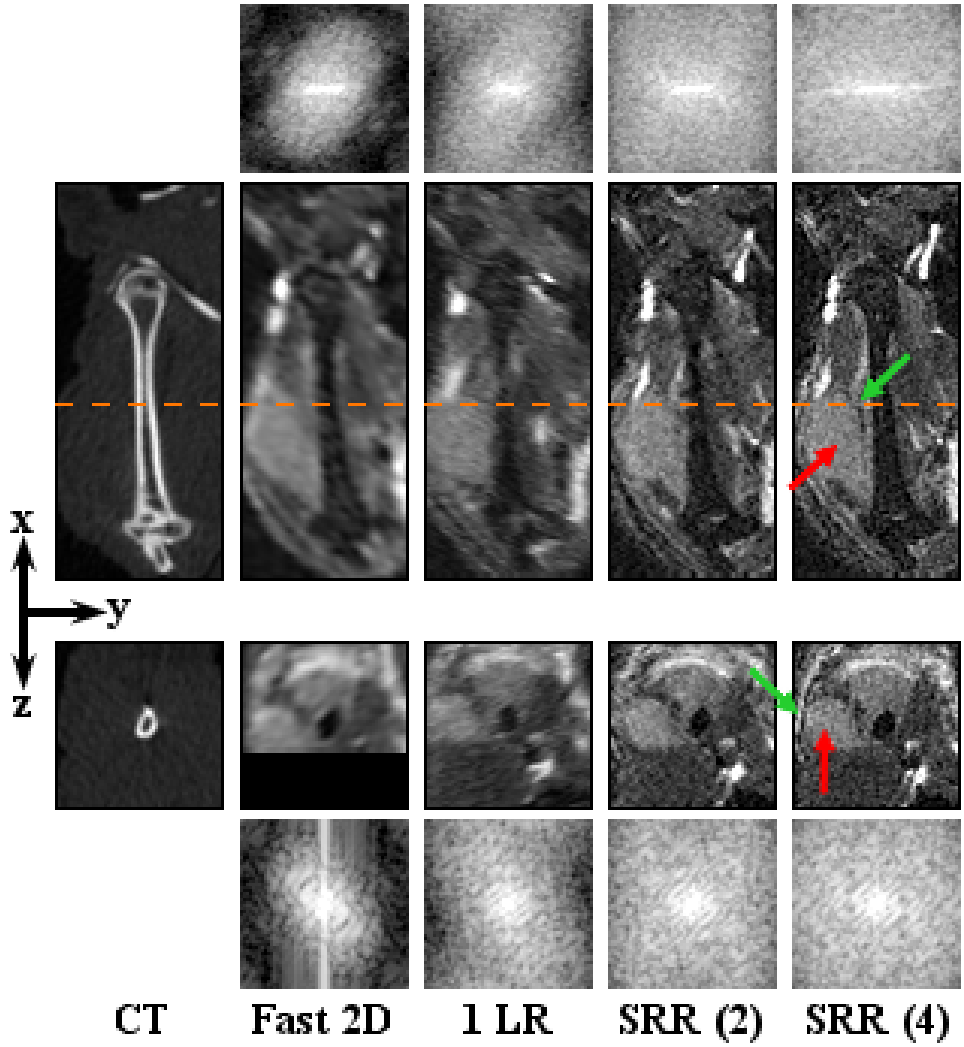


Figure 4.4: *Right humerus. From left to right: a CT scan, a fast MRI scan, one low-resolution image, and SRR reconstructions each based on a different number of low-resolution images. Two orthogonal slices of the same VOI are shown to illustrate the effect of the SRR in a 3D volume. The orange dashed line indicates where the yz-slice (bottom) intersects the xy-slice (top). The red arrows point to the tumor. The green arrows point to some of the locations where recovery of the fine details is the most noticeable. The CT and all the MR images are shown in the coordinate system associated with the principal axes of the bone, and Fast 2D and the low-resolution (1 LR) volumes are resampled to isotropic resolution beforehand. Image contrast on the MRI images was increased for visualization purposes. For all MR images the corresponding frequency spectra are shown, above and beneath the xy-slices and the yz-slices, respectively, demonstrating enhanced high-frequency content.*

In the BLI (Fig. 4.1), three distinct signal areas were observed, the smallest one at the position of the right femur (green arrow). The user therefore manually selected the right femur for SRR of the MRI data, using 2 or 4 low-resolution images for the reconstruction SRR(2), and SRR(4) respectively (see Fig. 4.3). The arrows in the BLI and the SRR(2) and SRR(4) images point to a tumor adjacent to the medial chondyle. This tumor is neither visible in the CT image, nor in the fast 2D MRI or in the raw low resolution image (1 LR). When using 2 low-resolution images for SRR, the image quality increases and the tumor becomes discernible. Using 4 low-resolution images further improves the visibility of the tumor and its margins.

BLI also showed a high intensity area at the location of the right humerus (Fig. 4.1, red arrow). The tumor is not visible on CT. The fast conventional MRI scan does show the tumor, but, due to the relatively thick slices, the tumor margins are blurred, particularly in the transverse plane. As before, the image quality improves when using more low-resolution images, showing a clear delineation of the tumor, with SRR(4) being sharper and less noisy than SRR(2).

To assert that SRR produced an actual resolution enhancement, frequency spectra were produced by applying a windowed Fourier transform to the shown MRI slices. The resolution enhancing effect in Figs. 4.3 and 4.4 is clear: when more images are used for reconstruction, the spectrum of the SRR image contains more high-frequency content than the low-resolution images.

Table 1 shows how the SRR reconstruction times scale approximately linearly with the size of the low-resolution data-set. Since one low-resolution image of the entire mouse contains approximately 20 million voxels, and a typical VOI contains around 250,000 voxels, we accelerate the reconstruction by approximately a factor 80. From the table, it also follows that the SRR times scale approximately linearly with the number of low-resolution images used. While the entire mouse takes more than 40 minutes to reconstruct using 4 low-resolution images, the VOI can be reconstructed within 1-2 minutes. The segmentation and selection of VOIs steps described above, each take less than a minute to perform.

	2 LR images	4 LR images
Femur	56	98
Tibia-Fibula	38	75
Pelvis	79	151
Sternum	31	63
Humerus	48	83
Ulna-Radius	41	78
Whole-body	1282	2479

Table 4.1: *Reconstruction times in seconds for each reconstructed right bone and the whole-body of the mouse, using 2 and 4 low-resolution images.*

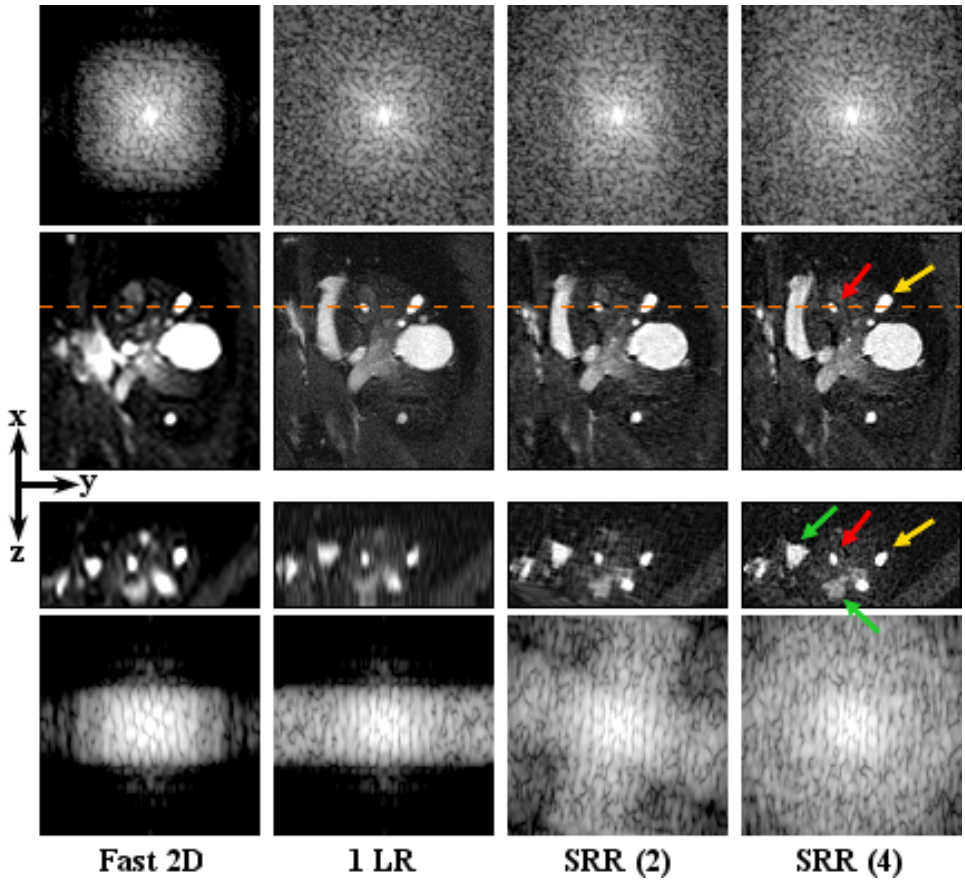


Figure 4.5: *Right kidney. From left to right: a fast MRI scan, one low-resolution image, and SRR reconstructions each based on a different number of low-resolution images, of the right kidney. The red and yellow arrows point to two different tumors. Two orthogonal slices of the same VOI are shown to illustrate the effect of the SRR in a 3D volume. The green arrows point to other locations where the improvement in image quality is particularly noticeable. The orange dashed line indicates where the yz-slice (bottom) intersects the xy-slice (top). In all the MR images, the xy-view is the in-plane direction of the scans. Note that the metastatic lesion seen in the BLI image (Fig. 4.1, blue arrow) actually consists of numerous lesions as shown on MRI scans. For the Fast 2D and the low-resolution (1 LR) the selected views are resampled to isotropic resolution and the image contrast on the MRI images was increased for visualization purposes. Note that for these images the shown slices do not correspond exactly to the SRR views. For all images the corresponding Fourier spectra are shown, above and beneath the xy-slices and the yz-slices, respectively.*

4.3.2 Case study B: MRI+BLI (kidney tumors)

BLI showed a single high signal intensity around the area of the right kidney (Fig. 4.1, blue arrow). Local SRR images of this area were reconstructed at different levels of quality and compared with conventional MRI and BLI. Fig. 4.5 shows orthogonal slices of the kidney for the different image types (fast 2D MRI, one low-resolution image, and SRR on 2, and 4 low-resolution images). On the BLI in Fig. 4.1, the spatial resolution is too low to determine whether multiple tumors are present, but on MRI one large tumor and several small metastases can be detected. Most of these are readily detectable on the fast 2D MRI, and the low resolution (1 LR) image. However, the tumors appear blurred and cannot be clearly delineated. In such images, the smallest metastases will be lost due to partial volume effects, but will be recovered in the SRR(2) or SRR(4) images. The high 3D resolution of the SRR scans also shows that most of these metastases are located in the renal cortex and medulla, whereas the renal pelvis is relatively clean. Again, we asserted the resolution enhancement due to SRR by producing frequency spectra of the shown images.

4.4 Discussion

4.4.1 Relevance to tumor research and other biological applications

Conventionally, bone resorption and metastases in soft tissues (such as kidney, lung and liver) are visualized using BLI+CT and BLI+MRI, respectively. In this study, we have explored the value of adding SRR-MRI to improve soft tissue tumor detection. In two case studies, we have shown how an integrated approach, combining state-of-the-art technologies from the area of image processing with the use of multiple imaging modalities, can be used to detect and study bone and soft tissue metastases with much greater sensitivity than by the conventional methods.

In case study A, we saw how BLI is a sensitive method to visualize luciferase-positive tumors in a living animal. The BLI signal intensity is proportional to the size of a tumor mass, and BLI can thus be used to give a rough estimate of both size and localization of the lesion. In the case of bone metastases, the location and subsequent bone pathology are usually determined using CT [4]. However, in case study A there was no visible bone pathology in the CT scan. When local SRR-MRI was performed, guided by the BLI signal, these images provided the location, size and shape of tumors in the limbs of the animal and confirmed that these metastases were, indeed, soft tissue tumors located outside of the bone. In the conventional fast MRI of the femur, the tumor could not have been identified without the guidance of the BLI images. The SRR-MRI, on the other hand, clearly showed a nodular structure that could be identified as a tumor (Fig. 4.3). In the humerus im-

ages, which contain a large tumor outside of the bone, it can be appreciated how the delineation of the tumor boundary becomes much sharper in the SRR-MRI than in the fast scan and the single low-resolution MR image (see Fig. 4.4; note that the improvement in image quality is especially noticeable when using a high zooming factor). The method thus, has the potential to support detailed quantitative studies of e.g. metastases development and assessment of treatment response.

In case study B of kidney metastases, CT was not used, as this modality gives insufficient soft tissue contrast without the use of contrast agents. BLI indicated the presence of a cancerous lesion in or around the kidney (Fig. 4.1, blue arrow). MRI revealed numerous independent metastases in the kidney (Fig. 4.5), which is not possible with BLI alone due to its limited spatial resolution. Moreover, SRR-MRI allows the researcher to not only distinguish, but also to clearly delineate different tumors in close proximity. This cannot be achieved with conventional MRI, as illustrated in Fig. 4.5. SRR-MRI can thus provide added value in studies where the number of metastases is an important parameter and where experimental treatment is used to intervene with the metastatic process. For instance, a researcher can differentiate between renal, adrenal and peri-renal cancerous lesions with SRR-MRI but not with BLI.

BLI remains the preferred standard measurement for active tumor size as the signal originates only from living cells and not from a necrotic core or cells killed by a certain treatment. Light, however, only has a limited penetration in bone and the bone can thus mask the BLI signal coming from small tumors which grow on the inside the bone, making these tumors appear smaller than they actually are. Having an MRI data-set in which the tumor can be identified and measured clearly helps overcoming these limitations.

An additional point to be made is the possibility to use BLI with SRR-MRI as an alternative for the CT anatomical reference, particularly in longitudinal studies where the repeated exposure to radiation in a CT scan may become a confounding factor or cause adverse effects [40].

Apart from oncology, the presented work flow may be of value in many research areas that requires whole body examination for local (sub-slice-thickness sized) effects. Examples are the homing of labeled stem cells after systemic injection, or imaging of systemic inflammatory diseases.

4.4.2 *Post mortem* to *in vivo* SRR-MRI

In this study, we have applied our approach to *post mortem* image data. However, we have well-founded reasons to believe that our results translate to *in vivo* imaging. The SRR quality depends on the amount of artifacts induced by animal movement. Such artifact are reduced by fast LR acquisitions and accurate subsequent registrations. While accurate non-rigid registration of soft tissue structures, such as liver and kidney may be possible, SRR is expected to be most successful for rela-

tively rigid structures, such as the brain, bone, and tissue surrounding bone: cases in which rigid registration will yield accurate alignment of the low-resolution images. In [69], we showed SRR reconstructions of an *in vivo* mouse brain, and several studies have validated the assumption of accurate motion estimation in applications of SRR in fetal brain MRI [28, 77].

4.4.3 Interactive local SRR

One of the results of this work has been the development of an approach that integrates recent progress in the areas of articulated atlas-based segmentation of whole-body small animal data, planar reformation and SRR in MRI into a novel localized approach to SRR that enables global-to-local exploration of e.g. whole-body mouse MRI data. Together with the preliminary results first published in [48], we have provided a global solution to three possible scenarios that takes into account the availability of complementary data: (i) only MRI is available [48], (ii) MRI+BLI is available, (iii) MRI+CT+BLI is available. From first to last scenario, the proposed approach decreases in the required level of user interaction to segment the data into possible VOIs. Depending on the biological problem, the more complementary data available, the higher the level of automation of the approach and the more data there is for the user to explore, i.e.: in the approach of case study B (MRI+BLI), the user can choose only among VOIs in which BLI signal is present, for a subsequent SRR reconstruction. Alternatively, (if CT is available) the user can select any bone for the SRR reconstruction and thus compare left with right, compare a bone with a tumor, with the same bone without a tumor on the contra-lateral side, etc. Naturally, the more complementary data available in a study, the more information one can extract. Thus, while in (i) only MRI information is available, in (iii) one can fully integrate the information provided by the BLI (which quickly locates tumor growth and indicates tumor burden) together with the anatomical information provided by the CT (used to study tumor induced changes in the bone resorption) and the soft tissue information provided by MRI (which can provide information about the size and the number of metastases).

4.4.4 Image quality vs imaging time

A major constraint when applying SRR in small animal MRI is the limited acquisition time that *in vivo* experiments allow. Each of the low-resolution images takes a certain amount of time to acquire and acquisition of multiple such images may quickly exceed the time in which a mouse can be kept sedated. It was shown in [69], that relatively few images are necessary to achieve significant improvements in the image quality. In this study we have limited the number of low-resolution images to four, with a total acquisition time of 52 minutes, a realistic acquisition time for *in vivo* experiments. If the experimental setting allows it, the number of low-resolution images used can be extended at the expense of additional acquisition

time. This will have some positive effect on the resolution. For an optimal coverage of k-space, the number of low-resolution images should be at least $\pi/2 \times F$, where F is the anisotropy factor, i.e., the slice thickness relative to the in-plane resolution. In our case, that would mean using 7 low-resolution images. Using more than this number of low-resolution images will not have a significant impact on the resolution but will increase the SNR slightly (for an in-depth study of these trade-offs, we refer the interested reader to [69]).

The main advantage of SRR in small-animal MRI is that it enables obtaining isotropic images in scenarios where T2-weighted image contrast is desired, requiring long repetition times and therefore long scan times, particularly for a 3D acquisition. By combining a small number of relatively fast thick-slice acquisitions with SRR, an isotropic resolution close to the original in-plane resolution is obtained. For comparison, direct acquisition of a 3D fast spin echo image with the same resolution and acquisition parameters would take about 28 hours and thus is clearly infeasible.

4.4.5 Reconstruction times

For large data-sets, the SRR method is limited by the memory available on the computer. For the conjugate gradient solver, up to 5 data structures, each the size of the final reconstructed image, and 2 additional data structures, each the size the total low-resolution data, must be kept in memory simultaneously. For large 3D data-sets, this soon becomes impossible, even on a high-performance desktop computer. The interactive approach to locally reconstruct VOIs presented here, allows overcoming the time and memory limitations of the SRR technique. However, as shown in Table 1, the time for the best quality SRR result, i.e., using 4 low-resolution images, is still in the order of minutes (mean time = 91.3 s). The mean time for SRR using 2 low-resolution images is 48.8 s. These results are still far from the real-time target for this approach. Since the results presented here were acquired on a MATLAB^{TM} implemented prototype, the computation times will decrease by re-implementing the algorithm in a C/C++ and GPU programming environment combination.

4.4.6 Conclusions

By combining a number of state-of-the-art image processing techniques, we have enabled a global-to-local exploration of whole-body mouse MRI. We have shown that the SRR-MRI is a valuable complementary modality in studies of tumor metastases. Using only a few low-resolution images, and a total acquisition time compatible with *in vivo* experiments, we have reconstructed SRR MR images from which detailed information about soft tissue metastases, not available in conventional imaging modalities, can be inferred. This cannot be obtained from direct MR acquisition within a feasible acquisition time.

4.5 Acknowledgments

Financial support from Medical Delta is gratefully acknowledged.

Chapter 5

Super-resolution reconstruction using cross-scale self-similarity in multi-slice MRI

Based on:

E. Plenge, D.H.J. Poot, W.J. Niessen, E. Meijering. Super-resolution reconstruction using cross-scale self-similarity in multi-slice MRI. In *Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part III, pages 123-130, Berlin, Heidelberg, 2013. Springer-Verlag.* 2013.

Abstract

In MRI, the relatively thick slices of multi-slice acquisitions often hamper visualization and analysis of the underlying anatomy. A group of post-processing techniques referred to as super-resolution reconstruction (SRR) have been developed to address this issue. In this study, we present a novel approach to SRR in MRI. The approach exploits the high-resolution content usually available in the 2D slices of MRI slice stacks to reconstruct isotropic high-resolution 3D images. Relying on the assumption of local self-similarity of anatomical structures, the method can be applied both to a single slice stack and to the combination of multiple slice stacks that differ in the orientation of their field of view. We evaluate the method quantitatively on synthetic brain MRI and qualitatively on MRI of the lungs. The results show that the method outperforms state-of-the-art MRI super-resolution methods.

5.1 Introduction

In magnetic resonance imaging (MRI) sequences that require long repetition times, conventional 3D imaging is usually not possible due to infeasible acquisition times. In such cases, 2D multi-slice imaging is the common alternative. However, due to hardware induced limitations on gradient strength, requirements on signal-to-noise ratio (SNR), constraints on acquisition time, and other factors, the slices are usually relatively thick compared to the in-plane resolution. Such anisotropy negatively affects visualization and hampers the analysis of underlying processes and anatomy. The isotropy and resolution of images may be improved by super-resolution reconstruction (SRR) methods [28, 69, 71, 98].

SRR methods are often divided into two groups: 1) methods that base the high-resolution (HR) reconstruction on a single image only, and 2) methods that combine the information from multiple low-resolution (LR) images of the same object, acquired under varying fields of views, to reconstruct an isotropic HR image. In both cases, the so-called *inverse problem* of recovering the HR image is ill-posed: the number of constraints imposed by the data (the LR voxels) is, in general, lower than the number of parameters that we aim to recover (the HR voxels). To make the problem well-posed, regularization is applied. By adding prior knowledge of the probability density function of the HR solution, the maximum-a-posteriori (MAP) solution can be found. Prior knowledge could be that the solution is smooth, piecewise smooth, or sparse under a certain transform. A properly chosen prior thus not only stabilizes the process but also guides it towards a plausible outcome.

Recently, a quite different prior has proven to be very powerful for natural images: example-based self-similarity. The self-similarity prior is based on the observation that small-scale structures tend to repeat themselves throughout an image. This concept has been widely applied in image processing tasks such as compression, denoising, inpainting, and SRR [21], and has spawned a number of methodologies including image hallucination [5], sparse coding using learned dictionaries [20, 57], and non-local means (NLM) [10].

In [5, 25, 29], the idea of *cross-scale* self-similarity was investigated. A common feature of these example-based methods is the formulation of a parent/child structure in which the nearest-neighbor of an input LR-patch is sought among the LR-parents of HR-child-patches in either a database [5] or in the image itself [29]. The idea of a correspondence between LR and HR patches for SRR was also applied for regularization of an iterative MAP optimization scheme [21] and later in the context of sparse coding using learned dictionaries [104, 108].

Though the literature on self-similarity-based SRR in natural images is rich, the results are only slowly translating into the field of biomedical image processing. In the context of MRI, the most prominent contributions have been based on the NLM-method. In [77] and [58], isotropic high-resolution T1-weighted images were used to enhance the resolution of T2-weighted slice-stacks. In [59], upsampling was

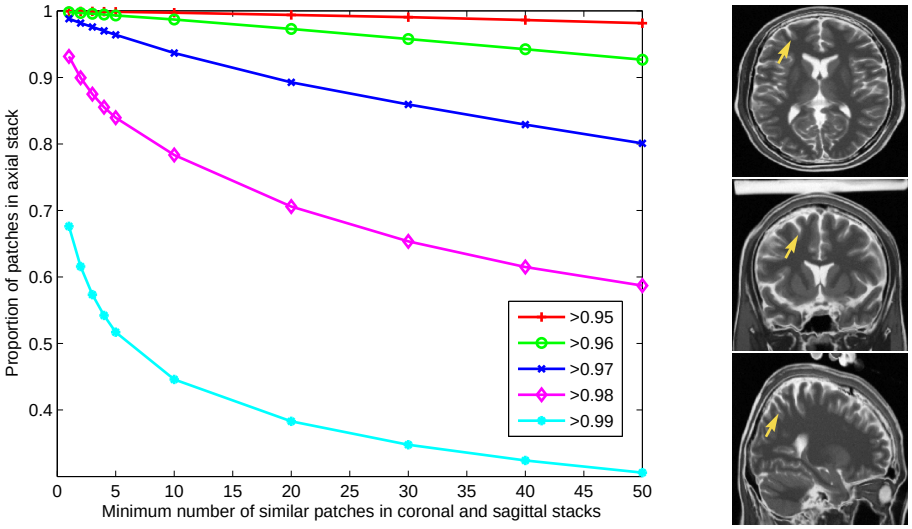


Figure 5.1: *Left: Quantitative check of self-similarity in a T2 weighted brain MRI from BrainWeb [13]. Similarity between the 10% axial 7x7 patches with the highest variance and all coronal and sagittal 7x7 patches was evaluated. For several similarity thresholds, the fraction of axial patches is plotted as function of the number of coronal and sagittal patches exceeding the threshold. Similarity is measured as the inner product between the patches (after vectorization and unit length normalization). Right: Example of local self-similarity in the cortex. The yellow arrows point at a similar cortex structure in the axial (top), coronal (middle), and sagittal (bottom) slices.*

performed on an MRI image using only information from the image itself by iteratively applying NLM-denoising to an interpolated version of the image itself. In [18], an application of self-similarity across scales was introduced as a generalization of the NLM-method.

In this thesis, we propose a novel approach to SRR in multi-slice MRI, based on the concept of cross-scale self-similarity [18]. Multi-slice 2D MRI scanning yields two native scales simultaneously: one at in-plane resolution and one defined by the slice gap and slice thickness in the slice-selection direction. Local self-similarity of anatomical features occurs both within and across these scales (Fig. 5.1), which we exploit to achieve SRR. Our method can be applied to both a single image and to multiple images, thus transcending the methodological division described above. We describe our method (section 5.2), show its performance by comparing it quantitatively and qualitatively with a baseline interpolation scheme and a state-of-the-art SRR algorithm (section 5.3), and we summarize the findings and contributions of this study and discuss limitations and future perspectives (section 5.4).

5.2 Method

5.2.1 From single-scale to cross-scale self-similarity

Our method is developed from one of the most successful applications of image self-similarity: the non-local means (NLM) method [10]. It exploits image self-similarity by locally representing the image as a weighted average of similar patches found elsewhere in the same image. In [72], NLM was derived as a special case of the MAP estimator, in which the posterior, p(HR|LR), is estimated directly from the available HR examples, given the input LR image.

In NLM the ith pixel in image $\mathbf{u}$, defined over the domain Ω, is given by

$$\mathrm{NLM}(\mathbf{u}(i)) = \frac{1}{C(i)} \sum_{j \in \Omega} w(i,j)\mathbf{u}(j) \,, \tag{5.1}$$

where

$$w(i,j) = e^{-\frac{\|\mathbf{u}(\mathcal{N}(i)) - \mathbf{u}(\mathcal{N}(j))\|^2_{2,a}}{h^2}} \quad \text{and} \quad C(i) = \sum_{j \in \Omega} w(i,j) \,. \tag{5.2}$$

$\mathcal{N}(k)$ is a square patch of a fixed size centered on pixel k, and $\|.\|^2_{2,a}$ is the norm, weighted by a Gaussian with standard deviation a, and h is a filtering parameter that controls the decay of the weights as a function of the norm.

In [18], the NLM denoising method is generalized to consider self-similarities across scales. A HR *example image* $\mathbf{v}$ is used to modify the NLM expression in Eq. (5.1), leading to a cross-scale NLM expression:

$$\mathrm{csNLM}(\mathbf{u}(i)) = \frac{1}{C(i)} \sum_{j \in \Omega} w(i,j)\mathbf{v}(j) \,, \tag{5.3}$$

in which

$$w(i,j) = e^{-\frac{\|\mathbf{u}(\mathcal{N}(\mathcal{D}(i))) - \mathcal{H}_z(\mathbf{v}(\mathcal{N}(j)))\|^2_{2,a}}{h^2}} \quad \text{and} \quad C(i) = \sum_{j \in \Omega} w(i,j) \,. \tag{5.4}$$

Here, $\mathcal{D}$ maps $\mathbf{u}$ onto a HR lattice, and $\mathcal{H}_z$ is a linear degradation model (which may include geometrical warp, blur, and more) of a degradation z. Eqs. (5.3)-(5.4) show how HR patches $\mathbf{v}(j)$, whose degraded versions are similar to an up-scaled LR patch $\mathbf{u}(i)$, are used to update the HR estimate $\mathrm{csNLM}(\mathbf{u}(i))$.

Our method is a special case of csNLM. In the limit, when $h \to 0$, the weight of the first nearest neighbor becomes infinitely large. The cross-scale super-resolution expression becomes

$$\mathrm{csSR}(\mathbf{u}(i)) = \frac{1}{w(i,j)} w(i,j)\mathbf{v}(j) = \mathbf{v}(j) \quad \text{with} \quad j \leftarrow \operatorname*{argmax}_j w(i,j). \tag{5.5}$$

Since a search for the first nearest neighbor involves finding the distance to all patches in Ω, there is no large computational gain of using only the first nearest neighbor. In the following we shall, however, constrain ourselves to do so, and later we demonstrate that the first nearest neighbor is indeed sufficient to improve resolution and image quality considerably.

5.2.2 MRI Super-resolution using cross-scale self-similarity

From a 2D viewpoint, MR slice-stacks contain both LR and HR images. The in-plane slices are isotropic HR and the orthogonal slices containing the slice-selection direction are anisotropic LR. Using the knowledge of 2D self-similarity of 3D anatomical structures (Fig. 5.1), we can apply Eq. (5.5) as follows: Let $\mathbf{u}$ denote the 2D LR slices of a reformatted slice-stack, let $\mathcal{D}$ be an interpolation operator that maps the LR slices onto a 2D HR grid, and $\mathcal{H}_z$ a number of blurring kernels applied to the in-plane slices $\mathbf{v}$. We search for the nearest neighbor of $\mathbf{u}(\mathcal{N}(\mathcal{D}(i)))$ over all HR patches j, at all blurring levels z, in $\mathcal{H}_z(\mathbf{v}(\mathcal{N}(j)))$.

A multi-scale dictionary

To create a strict correspondence between HR and LR features, the 2D HR slices are blurred at multiple scales using an approximation of the point-spread function (PSF) of the upsampled LR slice, or of the combination of multiple upsampled LR slices when more LR images are available. The reason multiple blurring levels are applied is to enrich the model and possibly capture cross-scale self-similarity at scales besides the two native ones (relevant for recursive structures such as vessel-trees). The vectorized versions of all HR patches (or a randomly selected subset thereof) are appended column-by-column into a matrix structure. The corresponding vectorized patches of the blurred versions of the HR slices are appended to this matrix, also column-by-column, such that each blurring level is represented by a contiguous block of patch-vectors in the matrix structure. The columns of the matrix are then normalized to have a Euclidean norm of 1. The normalization factors are saved for later rescaling. Adopting the terminology of the sparse coding literature, we shall call this matrix a *dictionary*. It is in this dictionary, that we search for the nearest neighbor of each LR patch.

The reconstruction process

An initial HR estimate is created by interpolating and aligning the available LR images onto a 3D HR grid. The slice-stack is traversed, LR slice by LR slice. That is, for the case when a single slice-stack is available, the direction of traversal should *not* be the original slice selection direction, but any of the two others. When using multiple LR slice-stacks, the direction can be chosen arbitrarily. Around each pixel in each of the traversed slices, a 2D patch (of e.g. 7×7 pixels) is extracted. According to the cross-scale self-similarity assumption, the underlying HR patch of this

LR patch will (to a high degree of approximation) be present among the in-plane patches collected in the dictionary. For each LR patch, its nearest neighbor (NN) is sought in the dictionary. The NN of an input patch is defined as the dictionary column that has the maximum inner product with the vectorized and normalized (to Euclidean unit of 1) input patch. Having found the NN of an LR input patch, the underlying HR estimate can be retrieved from the HR block in the beginning of the dictionary (e.g. by applying a modulus operation on the dictionary index of the NN). The HR patch is scaled back to its original intensity by applying the rescaling factor that was saved when normalizing the dictionary. The HR patches are inserted into the grid and averaged where they overlap.

5.3 Experiments and results

The method was evaluated quantitatively and qualitatively using simulated MRI brain data. Additional qualitative experiments were performed using real MRI data of the lungs. For both brain and lung images, the evaluations considered two applications: SRR of a single slice-stack, and SRR using three orthogonal acquisitions. In this second application, two variants were tested. In the first, the proposed method was initialized from an averaged upsampled image, while the second variant was initialized with the output image of an iterative regularized maximum-likelihood SRR method.

5.3.1 Image data

Simulated brain MRI

The simulated brain MRI was obtained from the publicly available BrainWeb database [13]. To simulate a common case from anatomical neuro-imaging, T2-weighted slice-stacks were generated as follows. Two isotropic HR ($1\times1\times1$ mm) T2-weighted images with intensity non-uniformity of 20%, and noise-levels of 3% and 9%, respectively, were downloaded. From each of the two HR images, slice-stacks were obtained by application of a 1D Gaussian PSF in the slice-selection direction. The full-width-half-maximum of the PSF was equal to the downsampling factor applied in the slice-selection direction. By varying this factor as well as the slice-selection direction, three orthogonal slice-stacks (axial, coronal, sagittal) were generated for downsampling factors of 3, 5, 7, 9. Thus, a total of eight sets, each containing three orthogonal slice-stacks, with resolutions of $1\times1\times3$ mm, $1\times1\times5$ mm, $1\times1\times7$ mm, and $1\times1\times9$ mm, were generated from the two original images with noise-levels of 3% and 9%.

Lung MRI

The lung MRI data was acquired on a 1.5 Tesla GE scanner. Axial, coronal, and sagittal slice-stacks of a test-subject with the lung in expiratory state were acquired during breath hold. A 2D gradient-recalled steady-state sequence (TE = 2.2 s, TR = 0.75 s, flip angle = 35 degrees, number of averages = 1) was used, having a scan time of 15 seconds per slice-stack. The field of view was 400×400 mm, the acquisition matrix was 128×160, yielding an in-plane resolution of 3.125×2.5 mm. The slices were reconstructed by the scanner to a grid of 256×256 with uniform pixel-sizes of 1.56×1.56 mm. 40 slices were acquired per stack with a slice thickness of 8 mm.

5.3.2 Implementation details

Features

For the reconstruction experiments in this study all parameters were set empirically. Raw intensity patches of size 7×7 pixels were used in all SRR experiments. This size was chosen to provide a good balance between the quantitative performance and the visual image quality or the reconstructions.

Dictionaries

The dictionaries were trained on the central 2/3 of the slices of the available slice-stacks to reduce the data load while including the central parts of the brain. Subsequently, a mask was applied, that included the brain region and a rim of the background wide enough to include entire patches. The HR slices were blurred at two scales, empirically set so as to achieve good performance. Patches of HR and blurred slices were extracted at corresponding locations. To further reduce the dictionary size, a random subset of 3/4 of the dictionary atoms were discarded, while maintaining the correspondence between the patches at the three scales. In the single slice-stack experiments, where less slices were available, the dictionary was augmented by rotating all patches by 90 degrees and appending these to the non-rotated patches. For the lung reconstructions, the dictionaries were trained on the central part of the images, containing lung, heart, trachea, stomach and a bit of the surrounding tissue.

5.3.3 Quantitative evaluation

The proposed method, which we will refer to here as patch-based super-resolution (PBSR), was evaluated for the application to 1) a single axial slice-stack (siPBSR), 2) the combination of three orthogonal slice-stacks (3orthoPBSR), and 3) for post-processing of the result of another SRR method using three orthogonal slice-stacks (LASR+PBSR). The PBSR processing was based on initial estimates of the HR image. These were created from the available data by cubic interpolation in the case where a single slice-stack was available, and either the average of three interpolated

and aligned orthogonal slice-stacks(3orthAvg), or another SRR method (LASR) [71], in the case where three slice-stacks were available. The performance of these methods on their own was also measured. All methods were evaluated on the eight sets of simulated brain images.

For quantitative evaluation of the reconstruction quality, we applied two measures, peak-signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [101]. To compute these measures, a noise-free HR T2-weighted image of $1 \times 1 \times 1$ mm was downloaded from the BrainWeb database. PSNR of the reconstructed image $\hat{\mathbf{x}}$ with respect to the *true* HR image $\mathbf{x}$ is defined as

$$\mathrm{PSNR}(\hat{\mathbf{x}}) = 10\log_{10}\left(\frac{L^2}{\mathrm{MSE}(\hat{\mathbf{x}})}\right) \text{ with } \mathrm{MSE}(\hat{\mathbf{x}}) = \frac{1}{|\Omega|}\sum_{i\in\Omega}(\mathbf{x}(i)-\hat{\mathbf{x}}(i))^2 \tag{5.6}$$

and L the dynamic range of $\mathbf{x}$. In SSIM two local patches at the same location i in $\mathbf{x}$ and $\hat{\mathbf{x}}$ are being compared. The mean SSIM between $\mathbf{x}$ and $\hat{\mathbf{x}}$ is

$$\mathrm{mSSIM}(\mathbf{x},\hat{\mathbf{x}}) = \frac{1}{|\Omega|}\sum_{i\in\Omega}\left(\frac{2\mu_{\mathbf{x}(i)}\mu_{\hat{\mathbf{x}}(i)}+C_1}{\mu^2_{\mathbf{x}(i)}+\mu^2_{\hat{\mathbf{x}}(i)}+C_1}\right)\left(\frac{2\sigma_{\mathbf{x}(i)}\sigma_{\hat{\mathbf{x}}(i)}+C_2}{\sigma^2_{\mathbf{x}(i)}+\sigma^2_{\hat{\mathbf{x}}(i)}+C_2}\right) \tag{5.7}$$

where $\mu_{\mathbf{y}(i)}$ and $\sigma_{\mathbf{y}(i)}$ are, respectively, the mean and standard deviation of the local image patch $\mathbf{y}(i)$. C_1 and C_2 are small positive constants to avoid numerical instability, for which we used the default values of [101].

The results of the evaluations are shown in Fig. 5.2. As expected, lower noise and thinner slices yield better reconstructions. Also, according to expectation, lower noise and thinner slices yielded better reconstructions, and in most cases the reconstructions based on three slice stacks were better than the ones based on a single-stack. More interestingly and importantly, it can be seen that applying PBSR on an initial HR estimate, consistently improves the PSNR and SSIM of the resulting images. For example, siPBSR consistently outperforms cubic interpolation. In all cases except one, either 3orthoPBSR or LASR+PBSR ends up on top. Furthermore, the difference in the performance of the two SRR methods, LASR and 3orthoPBSR, is of interest. Both methods find the initial registration parameters of the three orthogonal volumes using a low-aliasing affine transformation model, which is an integral part of the LASR method [71], but besides that, they can be compared as independent, stand-alone SRR methods. It is seen how 3orthoPBSR tends to perform better on small slice-thicknesses, and LASR better on thicker slices.

5.3.4 Qualitative evaluation

Figure 5.3 shows a number of coronal brain MR images reconstructed from either one or three LR slice-stacks generated from the original HR image with a noise level of 3%. Reconstructions using two different slice thicknesses, 3 mm and 7 mm, are

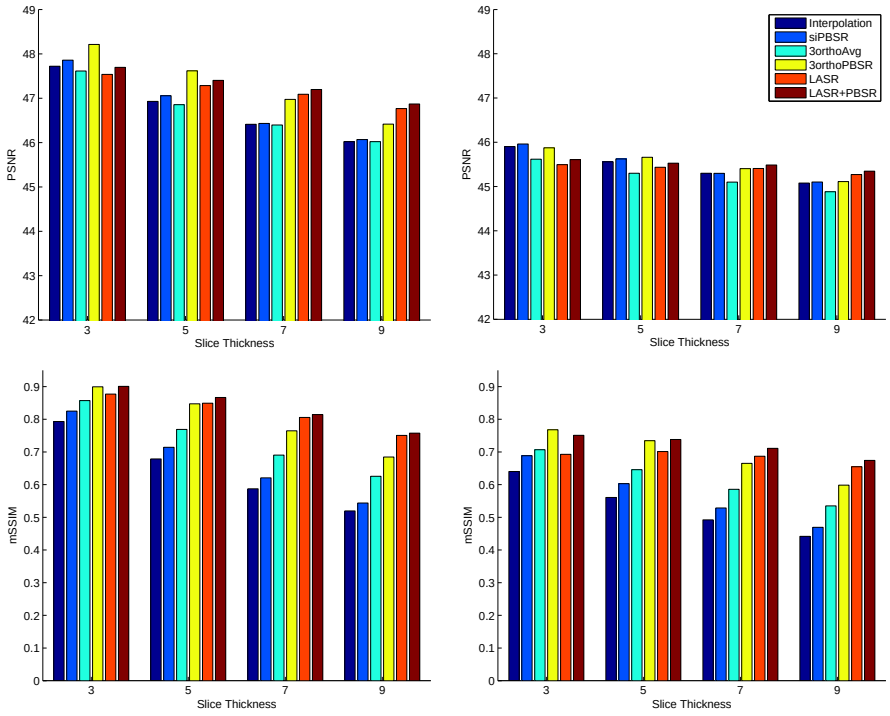


Figure 5.2: *Quantitative evaluations. Top: PSNR of SRR using six different methods. Bottom: mSSIM of SRR by the six methods. Left: Reconstructions from images of the 3% noise group. Right: Reconstructions from images of the 9% noise group.*

shown. It can be seen that the images that have been processed with the proposed PBSR method (images e, f, g, k, l, m) appear both sharper and less noisy than the images that they were created from (images b, c, d, h, i, j). Interpolation artifacts, especially visible at the bright skull-rim in images (b) and (h), are completely removed in image (e), while partially removed in image (k). In image (e), details such as the layers on the top right and left part of the skull become better discernible after PBSR processing. In image (j), reconstructed with the LASR method, an oscillating artifact pattern, due to the thick slices, is visible over the entire image. These artifacts are reduced by the PBSR and edges in image (m) appear sharper and more regular.

In Fig. 5.4, axial reconstructions of the lung MRI are shown. As a general remark, these images nicely display the potential of using multi-image SRR for MRI slice-stacks: comparing image (a), an interpolated coronal slice-stack, with images (b, c, e, f) where three orthogonal slice-stacks are combined using more or less advanced

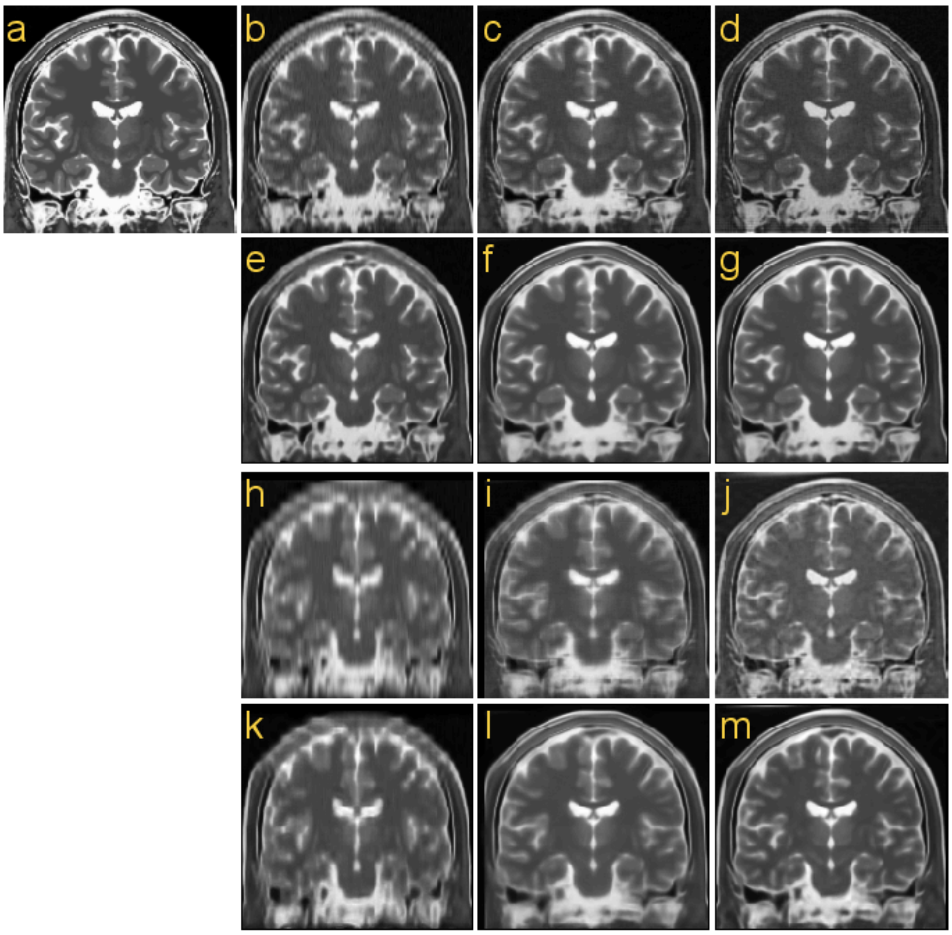


Figure 5.3: *Coronal slice of the simulated brain images, reconstructed using six methods: (a) the noise-free HR reference image, (b) cubic interpolation, (c) 3orthoAvg, (d) LASR, (e) siPBSR, (f) 3orthoPBSR, (g) LASR+PBSR. The reconstruction method of (h)-(m) corresponds to (b)-(g). Refer to text in section 5.3.3 for method nomenclature. Reconstructions (b)-(g) are based on slice-stacks of 3 mm slice thickness and (h)-(m) are based on slice-stacks of 7 mm slice thickness.*

SRR methods, the difference is striking. Comparing the SRR images based on three slice-stacks with each other, again, the PBSR images (e, f) appear sharper and less noisy than the images they are based upon (b, c). PBSR based on a single slice-stack (d) also looks sharper and less noisy than the original interpolated image (a), but in this case it is difficult to assess whether (d) becomes a more accurate

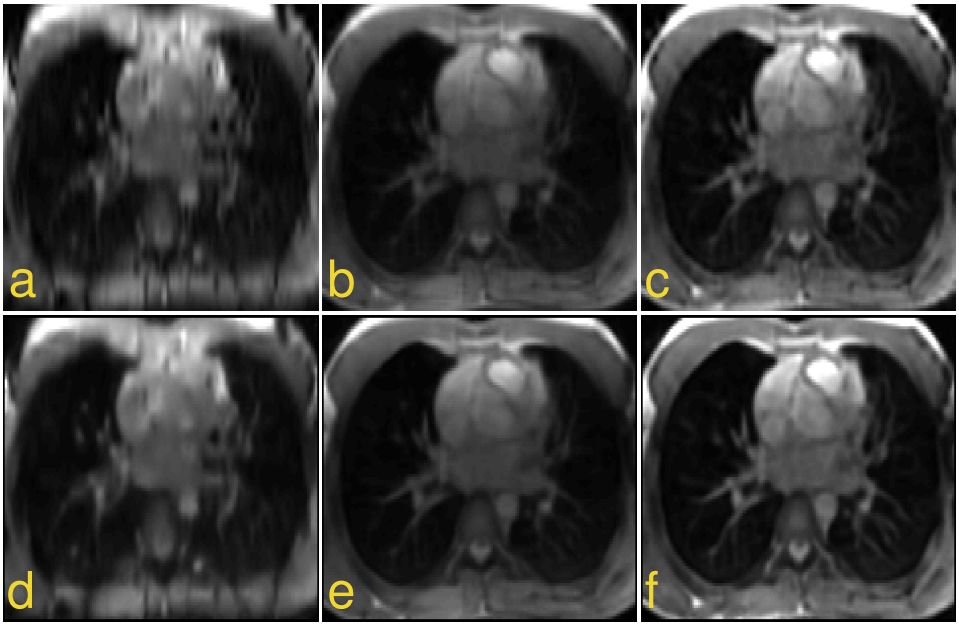


Figure 5.4: *Axial slice of the lung data reconstructed using the six methods. Methods and image order are the same as listed for Fig. 5.3. The lung MRI data are courtesy of P. Ciet and H. Tiddens (Erasmus MC).*

representation of the underlying anatomy. Comparing (c), reconstructed by LASR, and (e), reconstructed by PBSR, we observe that, while reconstruction artifacts are clearly present in the LASR reconstruction (especially along the edges of the image) and not in PBSR, detailed structures in the vessel tree are sometimes better visible in the LASR images.

5.4 Discussion

The main contribution of this chapter is the idea of exploiting redundancy across scales in an MRI slice-stack for SRR. This idea is especially appealing because of the vast amounts of such data that could potentially be enhanced using the proposed method. Cross-scale self-similarity has previously been exploited for SRR in MRI [58, 59, 77], but the present study is the first that takes explicit advantage of the relation between structural features in the HR slices of a slice-stack and in the orthogonal LR slices.

Our quantitative evaluation revealed that applying PBSR on an initial HR esti-

mate consistently improves the PSNR and SSIM of the resulting images. The fact that the PBSR improves the HR estimate of another SRR method (LASR) is worth noting. LASR obtains a regularized maximum likelihood estimated after an iterative procedure that takes into account the acquisition process while maintaining global data consistency. The improved performance when applying PBSR on the LASR estimate indicates that the imaging data contain valuable information that is not exploited by the LASR method but can be exploited using a self-similarity prior. The improvements in siPBSR over the cubic interpolated slice-stacks is small when measured by PSNR. SSIM reports slightly larger improvements, and is known to agree more with the human perception of image quality than PSNR, which corresponds with the observations in the qualitative evaluations (e.g. Fig. 5.3(b,e)), where the siPBSR image looks more similar to the true HR image.

The experiments conducted in this chapter only present a limited evaluation of the proposed method. In future work we plan to study how the patch size and the number of patches in the dictionary affect the performance of our method. In the experiments performed for the present chapter we generally found that smaller patches (e.g. 5×5 pixels) improved the quantitative performance, while degrading in visual quality, and vice versa for larger patches (e.g. 11×11 pixels). The number of blurring levels of the dictionary is another free parameter whose effect on the performance should be thoroughly tested. Such results may, however, not generalize easily, since they will be dependent on the specific anatomy, and on the image quality. Also, we have used only the first NNs for our reconstruction. Using an average of the first n NNs may improve the results. Imposing a structural coherency constraint between overlapping/neighboring patches may also have a positive effect.

The method has high computational complexity ($O(n|\Omega|)$). It would be interesting to see how a dictionary of a smaller size would perform in terms of reconstruction quality. For example, it would be interesting to apply the dictionary learning methods of [104, 108], in which a more compact, while still redundant, representation of the image subspace is achieved.

In this chapter we have presented an SRR method for MRI slice-stacks based on cross-scale self-similarity. The results of both quantitative and qualitative evaluations revealed that it exploits valuable image information missed by another state-of-the-art SSR method. While the preliminary experiments presented here demonstrate that our method improves the quality of MRI slice-stacks, further investigations are needed to study the extent to which this improves diagnostic quality, accuracy of segmentations, and registration quality.

5.5 Acknowledgements

We thank Harm Tiddens, Pierluigi Ciet, and Piotr Wielopolski, Dept. of Radiology, Erasmus MC, for letting us use the lung data.

Chapter 6

Multiple sparse representations for classification in MRI

Based on:

E. Plenge, S. Klein, W.J. Niessen, E. Meijering. Multiple sparse representations for classification in MRI. *To be submitted*

Abstract

In this work we introduce a generalization of the sparse representations classification (SRC) methodology. In conventional SRC, a voxel is classified by representing its features by two or more class-specific dictionaries and assigning the voxel to the class whose dictionary yields the minimum residual. We generalize SRC by, for each voxel, repeating this procedure a number of times, obtaining multiple non-overlapping sparse representations and their residuals with respect to each dictionary. Subsequently, we fuse the residuals into a more statistically informed and robust class assignment. We name this approach *multiple sparse representations classification* (mSRC) and validate it for three applications: pixelwise classification of 2D texture images, lumen segmentation in carotid MRI, and bifurcation point detection in carotid MRI. In all three applications mSRC outperforms basic SRC. We compare our method with K-nearest neighbor and support vector machine classifiers.

6.1 Introduction

The use of learned, overcomplete dictionaries and sparse representations of signals has been shown to yield state-of-the-art performance in reconstructive image processing tasks [20, 57, 58], and in the recent years, the methodology has found application in the image analysis domain as well [56, 103]. In this chapter, we first describe how learned overcomplete dictionaries and sparse representations have been used for voxelwise classification tasks. Then, we introduce a powerful generalization of the conventional so-called *sparse representation classification* (SRC) that allows a more exhaustive exploitation of the redundancy of the overcomplete dictionaries. We begin with an introduction to the concepts of dictionaries and sparse representations.

When a signal can be represented by a number of coefficients that is much smaller than the number of components in the signal itself, it is said to have a sparse representation. More formally, a signal $\mathbf{x} \in \mathbb{R}^n$ is said to be sparse over a dictionary $\mathbf{D} \in \mathbb{R}^{n \times m}$, if there exists a vector $\mathbf{a} \in \mathbb{R}^m$ with low L_0-norm, $||\mathbf{a}||_0$, such that $\mathbf{x} = \mathbf{D}\mathbf{a}$.

In the context of image processing and analysis, the signal $\mathbf{x}$ is often simply a vectorized patch of image intensities. The patch is extracted from the image as a descriptor or *feature* of the voxel in the patch center. The size of the patches is chosen such that they capture the structures that are relevant for the task at hand. Other types of features can be generated, e.g. by applying filters to the images that emphasize particular structures of interest, or dimensionality reducing operations that find more compact feature descriptors.

A learned dictionary is a set of vectors, usually concatenated into a matrix. The dictionary can consist of features that have been extracted directly from a database of images, or it can be derived from this database in a training stage to yield a more compact representation of the training data. The individual vectors of a dictionary are called *atoms*. This terminology is commonly used in the literature, and we will adopt it here as well.

In the context of sparse representations, most studies have focused on overcomplete dictionaries, meaning that $m > n$ for a full rank $\mathbf{D}$. In fact, the overcompleteness of a dictionary is directly related to its ability to represent a signal sparsely. Consider a large dictionary; as this dictionary grows, it becomes increasingly likely that the signal itself is present in the dictionary (up to a scaling factor). As the dictionary shrinks, it becomes necessary to combine an increasing number atoms from $\mathbf{D}$ in order to maintain the expressiveness of the dictionary. In this way, sparsity arises directly from the overcompleteness, and at the same time it provides a natural constraint to the underdetermined system $\mathbf{x} = \mathbf{D}\mathbf{a}$ [12, 53, 61, 85]. We note that example-based regularization methods such as non-local means (NLM) [10] and hallucination [5] are based on similar principles.

It has been shown that learned dictionaries generally outperform analytical

ones [20], and in this study we focus on learned overcomplete dictionaries that incorporate the sparsity prior via the learning process. An excellent overview of the evolution of dictionaries for sparse representation is given in [80].

6.1.1 Sparsity and self-similarity of the data

In the image domain, sparse representations can be applied because small-scale structures tend to repeat themselves throughout an image or a group of similar images. This phenomenon is also known as *self-similarity*. Textures, for example, are by definition highly self-similar (we define a texture as a local structure that repeats itself many times over a contiguous image region) and they often have a very accurate sparse representation by dictionaries trained on similar images (see [66] for an overview of sparse modeling of textures). Natural images are generally rich in texture and other repeating structures and abundant with self-similarity [111]. Think, for example, of the leaves in an image of a tree, the faces in a crowd of people, or the windows of houses in a street view. The bulk of studies on dictionary learning and sparse representations have considered this type of images.

Completely different images are produced in the medical imaging domain (MRI, CT, Ultra-sound). These images tend to be piecewise smooth (at least in the noise-free case) and often do not contain any texture at all. Self-similarity does exist though. Consider, for example, the set of slices of an MRI of a brain or a vessel-tree. Both within and between such slices, much of the content will be repeated and local features can be well approximated by features elsewhere in the image or in other, similar, images. This is illustrated in Fig. 6.1.

Each curve plotted in the figure shows the average correlation between patches sampled from a carotid MRI, and the 100 best *non-overlapping* sparse representations of each patch using a dictionary trained on other carotid MRIs. By non-overlapping we mean, that each atom is used only once in these 100 sparse representations; after it has been used, it is removed from the dictionary. The dictionaries have been trained and tested using varying patch sizes and *sparsity-levels*, i.e. the number of atoms used for sparse representation.

The figure essentially shows, that patches in a carotid MRI can be well represented as a linear combination of atoms in a dictionary trained on patches from other carotid MRIs. A second observation is that this dictionary is sufficiently redundant to accurately represent the new patches *many* times with non-overlapping sets of atoms. This reveals a massive redundancy in the dictionary and in this study we show how this redundancy can be exploited to improve sparse representation classification.

6.1.2 Sparse representation classification

If one thinks of a dictionary as a distribution, then the sparse representation of a new signal, using the dictionary, can be seen as a sample from the distribution. If

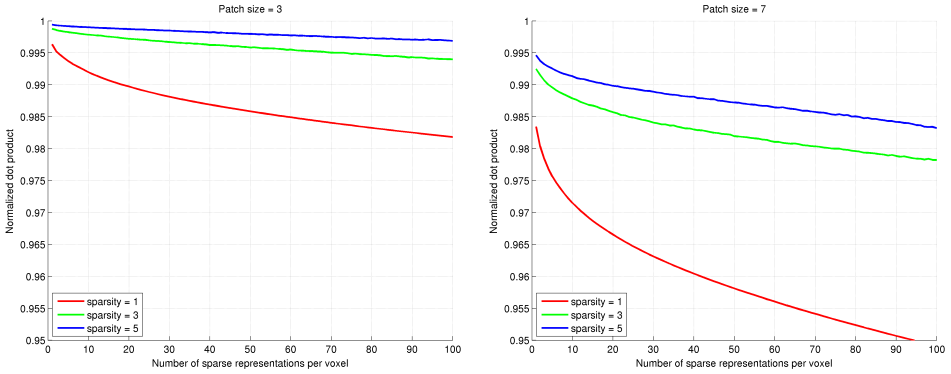


Figure 6.1: *Illustration of the redundancy of dictionaries trained on a carotid MRI data-set. The dictionaries had 1000 atoms and were trained under varying sparsity-levels and patch sizes:* $3 \times 3 \times 3$ *voxels (left) and* $7 \times 7 \times 7$ *voxels (right). Approximately 1000 patches were sampled randomly from a test image (a carotid MRI not included in the training-set) and represented using the dictionaries. The graphs show the average correlation of the sampled patches with their first 100* non-overlapping *sparse representations using the dictionaries.*

the new signal is similar to the training data (say, both are images of brick walls), then it will be well approximated by the dictionary and thus likely under the distribution. If, however, the signal is of a very different type (say, an image of a grass field), then its representation will be less accurate, and the signal less likely under the distribution. This reasoning naturally leads to the use of dictionaries for classification.

Sparse representation classification (SRC) denotes a group of methods that use dictionaries and sparse representations for classification. In the basic SRC scheme, a dictionary is learned for each class, and classification of a new signal is done by finding the dictionary that sparsely represents the signal with the minimum residual.

In this chapter, we present a powerful generalization of SRC that is based on the observations in Fig. 6.1, described above: Instead of extracting a single signal representation from each dictionary, we find multiple *non-overlapping* sparse representations, and for each dictionary we fuse the multiple representation residuals into a soft class assignment. In this way, the statistical robustness of the classifier increases in the same way that flipping an unfair coin multiple times will improve our guess on which face it is more likely to show. We dub the approach *multiple sparse representations classification* (mSRC) and test it on three different applications: texture classification, carotid MRI lumen segmentation, and bifurcation detection in carotid MRI. In all applications, we compare the performance with conventional

SRC, *K*-nearest neighbor (*K*-NN) and support vector machine (SVM) classifiers.

In the rest of this chapter, we first review previous work in the field of SRC. In section 6.3, we provide details about the mSRC method. The experimental design is described in section 6.4, followed by the results in section 6.5, and finally, in section 6.6, we discuss our findings and point out future directions in mSRC research.

6.2 Previous work

Early works in the field of SRC include [86] and [39] that both considered the problem of texture classification. In the method of [86], overcomplete dictionaries are trained for multiple types of textures. Classification of a new sample is performed by associating it with the texture type whose dictionary results in the minimum residual error, essentially implementing the idea described above. Such dictionaries are called reconstructive because they focus on the task of accurate representation, and they lead to a generative classifier.

In [56], the concept of discriminative dictionaries was proposed with the notion that for classification purposes a dictionary should not only represent a certain class of structures well, but also represent structures from other classes poorly. For training of the dictionary, a cost function is formulated over the dictionaries from all classes simultaneously. The cost function includes a discriminative softmax function that penalizes a dictionary's accurate representation of signals from the wrong class. The method was applied to texture classification and object detection.

Another discriminative approach was described in [74]. Here, discrimination is achieved by including the inter-dictionary coherence in the training cost function, promoting dictionaries that are more distinct. The method was applied to textures, object detection and digit recognition.

An entirely different path was taken in [103]. In this method a single dictionary is used for face recognition. The dictionary is not trained but consists of many face images of a number of persons. A new face image is represented as a sparse linear combination of all faces/atoms in the dictionary. Subsequently, the atoms used in the sparse representation are split into their respective classes and the residual with respect to the input image is calculated per class. The input image is associated with the class that yields the minimum residual.

Instead of basing classification on the representation residuals only, in [105] the trained classifier takes the sparse coding coefficients into account. Using a structured dictionary, created by concatenating class-specific sub-dictionaries, the Fisher discrimination criterion is imposed to decrease intra-class scatter and increase inter-class scatter of the coefficients of the training data. Classification is based on the reconstruction error as well as the distance to a mean coefficient vector learned during training. The method was tested on face, digit, and gender recognition.

In the context of MRI, the sparse representation based methodology has mainly

been applied for image processing tasks like denoising and super-resolution [58, 68, 77]. Recently, however, the methodology has been applied to MR image analysis, specifically, segmentation.

In [15], an approach, heavily inspired by the NLM-method [10], was applied to hippocampus and ventricle segmentation. The method is an adaptation of atlas-based segmentation to the local self-similarity paradigm in which images that have been annotated by experts are pre-registered to the target image and classification of a target voxel/patch is performed by weighting the labels of the most similar patches in a local neighborhood in the atlas images. In [14], this method was shown to work well for detection and grading of Alzheimer's disease. In [79], a similar approach was applied to general brain structure segmentation.

In [92], two approaches to hippocampus segmentation in MRI were presented. Both rely on a registered set of atlas images and dictionaries extracted *for each voxel* in the target image. In the first approach, all patches in a neighborhood around the target voxel position are extracted from the labeled training data and used directly as the voxel's dictionary. After a target patch has been coded using the dictionary, it is assigned to the class whose contributing patches yield the minimal reconstruction error, much like in the approach of [103] described above. Sparse coding is performed using elastic nets [112]. In the second approach, discrimination is incorporated by adding a classification error term to the dictionary training objective function, thus training a dictionary and a classifier simultaneously. A target patch is classified by first representing the patch sparsely using the dictionary and then applying the classifier to the sparse coding coefficients.

SRC has also been applied to prostate segmentation in CT images [26]. In a series of steps, involving feature selection among a number of context and appearance features, dictionary training by elastic nets [112], and a linear regression step that trains a soft classifier from the representation residuals, a discriminative classifier is created. Classification of a new sample is done over a number of iterations in which the classification results are refined by updates to the context features.

Through the years, the basic SRC idea, as proposed by [86], has been modified in various ways in order to improve its performance or adapt it to specific classification tasks. While these objectives have been achieved, the complexity of the approaches have increased and involved processing pipelines have been built around the SRC to effectively perform the given tasks. In this work, we focus on how the basic SRC can be improved by a generalization that takes the exploitation of the dictionary redundancy one step further. As such, the idea presented in the following can be adapted to many of the methods described above.

6.3 Methods

6.3.1 Dictionary learning

There are several ways to learn a dictionary (for an overview, see [80]). In this study we adopt the K-SVD algorithm [1], which is a relatively efficient method that incorporates the sparsity prior in the training process. Given a set of L_2-normalized training signals $\mathbf{X} = [\mathbf{x}_1\mathbf{x}_2...\mathbf{x}_M]$, K-SVD attempts to solve the following problem:

$$\min_{\mathbf{D},\mathbf{A}} ||\mathbf{X} - \mathbf{D}\mathbf{A}||_F^2 \quad \text{s.t.} \quad ||\mathbf{a}_i||_0 < T \quad \forall i \ , \tag{6.1}$$

where $\mathbf{A} = [\mathbf{a}_1\mathbf{a}_2...\mathbf{a}_M]$ is the set of corresponding sparse coding vectors, T limits the number of non-zero coefficients in each $\mathbf{a}_i$, and $||\cdot||_F$ is the Frobenius norm. The problem is a combinatorial one and a solution is, in general, intractable. K-SVD, however, approximates the solution by alternating between a greedy sparse coding step using the current dictionary estimate, and a dictionary update step. K-SVD has been shown to converge to a dictionary that is very close to the optimal one [19].

6.3.2 Sparse coding

Given a dictionary, $\mathbf{D}$, the goal in sparse coding is to represent a signal $\mathbf{x}$ as a linear combination $\mathbf{x} = \mathbf{D}\hat{\mathbf{a}}$, where $\hat{\mathbf{a}}$ is sparse. We wish to solve

$$\hat{\mathbf{a}} = \underset{\mathbf{a}}{\operatorname{argmin}} ||\mathbf{x} - \mathbf{D}\mathbf{a}||_2 \quad \text{s.t.} \quad ||\mathbf{a}||_0 < T \ , \tag{6.2}$$

where T is the maximum number of atoms to be used. As in dictionary training, the solution to this system is in general intractable. Instead, a locally optimal solution can be found by a greedy approach, e.g. orthogonal matching pursuit [93].

6.3.3 Sparse representation based classification

In basic SRC, the central voxel of an image patch is classified according to how well the patch is represented by the class-specific dictionaries. After a dictionary $\mathbf{D}_i$ has been trained for each class $i \in \{1...N\}$ of image content, classification of a new patch $\mathbf{x}_{new}$ can be performed by evaluating the representation errors

$$e_i(\mathbf{x}_{new}) = ||\mathbf{x}_{new} - \mathbf{D}_i\hat{\mathbf{a}}_i||_2^2 \quad \forall i \ , \tag{6.3}$$

where $\hat{\mathbf{a}}_i$ has been found using Eq. (6.2). From the class errors e_i a pseudo-probability measure, $p(C_i)$, is computed, leading to the following class assignment rule:

$$i^* = \underset{i}{\operatorname{argmax}}\, p(C_i), \quad \text{where} \quad p(C_i) = 1 - \frac{e_i}{\sum_j e_j} \tag{6.4}$$

6.3.4 Multiple SRC

The underlying idea of mSRC was explored in [23], where the following question was posed for a denoising application: "Suppose we are served with a group of competing sparse representations, each claiming to explain the signal differently. Can those be fused somehow to lead to better results?" The answer came out positive. Though our approach is different than that of [23], the proposed generalization of SRC follows from the similar intuitive idea, that by fusing the residuals of multiple non-overlapping sparse representations for each dictionary, a more robust class estimate of a given signal can be obtained.

The residuals of the non-overlapping sparse representations of a signal $\mathbf{x}_{new}$ are obtained by iteratively solving Eq. (6.3) for all classes i. After each iteration j, each dictionary $\mathbf{D}_i$ is updated by removing the atoms that are non-zero in the current sparse coding vector $\hat{\mathbf{a}}_i$, that is, the atoms that are used in the jth sparse representation of $\mathbf{x}_{new}$. Using L iterations, the voxel is assigned to

$$i^* = \underset{i}{\operatorname{argmax}} \sum_{j}^{L} p_j(C_i) \ , \tag{6.5}$$

where $p_j(C_i)$ is the pseudo-probability of class i found in the jth iteration.

As seen in Fig. 6.1, subsequent representations, $j > 1$, are still highly accurate. It is our hypothesis, that the accumulation of information in the pseudo-probability derived from these representations will improve the performance of our classifier.

6.4 Experimental design

SRC is a generic methodology, and in the computer vision literature it has been shown to work for a number of different applications (texture classification, object detection, face and digit recognition). To validate and explore the performance of our mSRC approach, we have chosen three different application settings. The first, is texture classification using texture images from the Brodatz database [75]. The second, is carotid artery lumen segmentation in MRI, where the centerline of the carotid is given. Finally, the approach is applied to bifurcation point detection in carotid artery MRI. In the following subsections we describe the design of each of these experiments in detail.

Besides testing the effect of fusing multiple sparse representations for classification (mSRC), we perform a series of experiments in which we optimize the mSRC parameters for the given task. The main free parameters of the dictionary training process are the feature (i.e. patch) size, the sparsity-level, and the size of the dictionary. We implement grid search strategies to find the optimal parameter settings using a data-set dedicated to parameter tuning.

In each experiment, we compare mSRC with basic SRC, *K*-nearest neighbors (*K*-NN), and support vector machine (SVM) classifiers. The used SVM applies a radial basis function kernel. This kernel was chosen because it is the most commonly used in the SVM literature. Furthermore, it makes SVM a nonlinear classifier and in that way similar to the other methods. We used the SVM implementation of [11].

In SVM, three parameters were optimized: patch size, the SVM parameter C, which weighs the significance of wrongly classified samples during training, and the parameter γ, which is a parameter of the radial basis function kernel. In [11], it is recommended to test the following ranges: $C \in \{2^{-5}, 2^{-3}, ..., 2^{13}, 2^{15}\}$, and $\gamma \in \{2^{-15}, 2^{-13}, ..., 2^{1}, 2^{3}\}$. The grid search optimization was performed in this range, although in some cases the range was extended to improve performance. In the *K*-NN experiments, the patch size and the values of K were optimized for performance.

In the *K*-NN experiments the training-set was the same as those for mSRC, given below. In the SVM experiments, due to the time complexity of the training process, only 1500 training samples per class were used for each experiment. These samples were picked from random locations in each class-specific part of the training images.

6.4.1 Texture classification

The Brodatz texture data-set has been used in a number of studies of SRC. For this study, three textures (D21, D28, and D77) were selected to accommodate a basic test of our hypothesis that using multiple sparse representations for classification works better than using a single one. D21 and D77 are textures of two types of cloth fabric, while D28 shows a rock surface, see Fig. 6.2. The texture images were split into a training part of size 480×640 pixels and a test part of size 160×640 pixels. Three pairwise classification experiments were performed using hit-rates (the ratio of correctly classified pixels) for evaluation. It is investigated how the performance of mSRC varies with the number of sparse representations used. A dictionary size of 1000 atoms was used, and a grid search optimization was performed on patch sizes of $\{(3 \times 3), (5 \times 5), (7 \times 7), (9 \times 9), (11 \times 11)\}$, and sparsity-levels of $\{1, 3, 5, 7, 9\}$. In each pairwise experiment, the size of the training-set was approximately 300,000 samples per class.

6.4.2 Carotid artery lumen segmentation

Segmentation of the carotid artery lumen is relevant for investigating the state and progression of atherosclerotic disease [36, 99, 107]. In this application, we pose the segmentation of the carotid arteries in MRI as a voxelwise classification problem, in which the carotid centerlines are given to guide the segmentation. Examples of cropped slices of a carotid artery MRI are shown in Fig. 6.3.

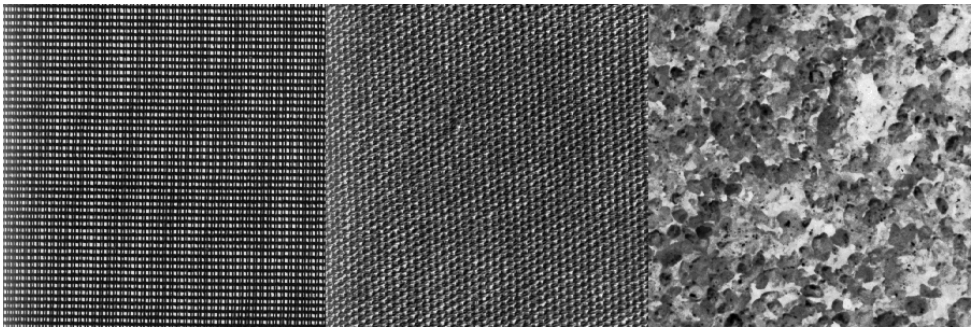


Figure 6.2: *The three Brodatz textures being used for the texture classification experiments. From left to right: D21, D77, D28.*

A data-set consisting of 29 MR images of the head and neck region of 29 patients was used. The MRIs were proton density weighted 2D spin echo images of resolution $0.5 \times 0.5 \times 0.9$ mm, and size $256 \times 256 \times 51$ voxels, acquired with a 1.5 T scanner. For all 29 subjects, the left and right carotid centerlines and the lumens were annotated by experts. The data is a sub-set of the data used in [36] .

The images were bias-corrected with the method of [95], and split into two sets: one was used for parameter tuning, the other was used for the final experiment using the found optimal parameters. The set of images used for parameter tuning consisted of 8 images selected randomly from the full set of 29 images. Parameter tuning was performed by leave-one-out tests. The tuned parameters were feature size, the sparsity-level, and dictionary size. The remaining 21 images were partitioned into a training-set consisting of 11 images and a test-set consisting of the remaining 10 images.

In each experiment, the training data was organized by extracting cubic (in voxel-space) patches centered on all voxels inside the annotated lumen regions of each training image. This resulted in approximately 14,000 lumen samples from each of the 7 training images. A corresponding number of background patches were extracted in each training image, from random locations in a tubular region around the lumen. This region was defined around the annotated centerline with a diameter of 31 voxels. All patches were vectorized and principle component analysis (PCA) was performed to reduce the dimensionality. The components with the lowest 0.1% of energy were discarded. Lumen and background dictionaries were trained on these signals.

Before segmentation, each MRI image was split into a left and a right part, such that only a single carotid was present in each resulting image. A tubular region of

The data can be obtained via http://ergocar.bigr.nl/

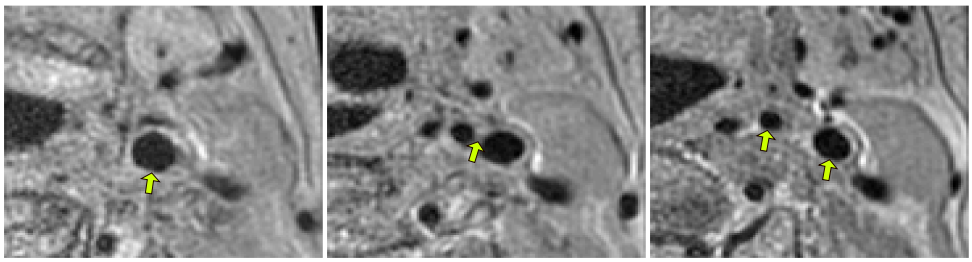


Figure 6.3: *Cropped slices 10, 20, and 30 (left to right) of the left carotid of subject 1 in the carotid MRI data-set. The arrows point to the common carotid artery (left image), the carotid bifurcation point (center image), and the internal and external parts of the carotid after the bifurcation (right image).*

interest (ROI) with a diameter of 31 voxels, was generated around the annotated centerline. At each voxel in the ROI, cubic patches were extracted and, after vectorization, projected into the lower-dimensionality space obtained by PCA during training. After assigning each voxel in the ROI to either foreground (carotid lumen) or background by mSRC, a simple 2D connected component procedure was performed: for each slice in the voxelwise classification image, a logic OR operation was performed with the corresponding slice of an inflated version of the centerline (diameter of 7 voxels). In each slice, only the connected component that included the inflated centerline was considered lumen in the final segmentation.

Using the annotated carotid artery lumen as ground truth, the Dice overlap score [17] was used for evaluation.

A two-stage parameter tuning process was performed for SRC and mSRC. First, using a dictionary of 1000 atoms, a grid search optimization was performed on patch sizes of $\{(3\times3\times3),(5\times5\times5),(7\times7\times7),(9\times9\times9),(11\times11\times11)\}$, and sparsity-levels of $\{1,3,5,7,9\}$. Next, the best performing sub-set of the patch sizes were selected for a 3D grid search in which the additional parameters were the sparsity-levels, with the values listed above, and dictionaries with 200, 500, 1000, 2000, 4000 atoms, respectively.

6.4.3 Carotid artery bifurcation detection

Detection of the bifurcation in carotid artery MRI is useful for carotid centerline extraction, segmentation, and registration (in multimodal imaging or longitudinal studies). For this experiment the MRI data-set of 29 subject was partitioned into the same parameter tuning set and a test-set as described in the previous section, and the same bias-correction method was applied. In all 29 subjects an expert annotated the bifurcation point and lumen of the left and right carotid arteries. An example

of a bifurcation point can be seen at the arrow tip in the central image of Fig. 6.3.

The bifurcation point dictionary was trained on image patches extracted in a region of $7 \times 7 \times 7$ voxels centered on the annotated bifurcation points. To extend and introduce more variability into the training-set, a number of rotations were applied to the training images, and at each rotation, patches were extracted around the annotated (correspondingly rotated) bifurcation points and added to the bifurcation point training-set. This resulted in approximately 4500 samples from each of the 7 training images. For each rotation, and each image, a corresponding number of patches were sampled at random locations in the rest of the image. The set of angular rotations were -30, -20, -10, 0, 10, 20, and 30 degrees, rotations were done around the slice selection direction. Before training of the bifurcation point and background dictionaries, all patches were vectorized and PCA was performed to reduce the dimensionality. The components with the lowest 0.1% of energy were discarded.

As in the carotid artery segmentation experiments, the test images were split into their left and right halves, such that only a single bifurcation was present in each test image. Rectangular left and right ROIs were generated such that the annotated bifurcation points in all 29 images were contained in the ROIs. The sizes of the left and right ROIs were $42 \times 65 \times 39$ and $47 \times 52 \times 39$ voxels, respectively. To reduce computation time, bifurcation point detection was performed inside the ROIs only. A class response image was generated by assigning a pseudo-probability to each voxel of it being the bifurcation point, and the voxel with the highest response was classified as the bifurcation point. While the class assignment in SRC, mSRC, and SVM is done by a continuous range of pseudo-probability values, in K-NN the soft class assignment is discretized into K levels. To decrease the likelihood of having multiple maximum response points, large values of K were used in the tuning, and a Gaussian filter with $\sigma = 0.5$ voxel was applied to the class response image. To make conditions uniform, this filter was applied to the response images of all four methods.

Using the annotated bifurcation points as ground truth, the distance in millimeters between the detected and annotated bifurcation points was used as evaluation measure.

A two-stage parameter optimization was performed for SRC and mSRC. First, using a dictionary of 1000 atoms, a grid search optimization was performed on patch sizes of $\{(3 \times 3 \times 3), (5 \times 5 \times 5), (7 \times 7 \times 7), (9 \times 9 \times 9), (11 \times 11 \times 11)\}$, and sparsity-levels of $\{1, 3, 5, 7\}$. Next, the best performing sub-set of patch sizes were selected for a 3D grid search in which the additional parameters were the sparsity-levels, with the values listed above, and dictionary size, with values of $\{500, 1000, 2000, 4000\}$ atoms.

6.5 Results

6.5.1 Texture classification

The results of the pairwise classifications are shown for patch sizes 3×3 and 5×5 pixels and sparsity-levels of $\{1, 3, 5, 7, 9\}$ in Fig. 6.4. For patch sizes of 9×9 pixels and larger, the results improve to near perfect classification (hit rates > 0.99).

The positive effect of using mSRC is clear from the plots. The hit-rates generally increase with the number of representations used. It can also be seen how the rate of improvement decreases with the number of representations.

An exception is the results obtained with a sparsity-level of 1: after an initial improvement in performance when using 2-4 sparse representations, a small but consistent decay is seen for higher numbers of sparse representations. Another exception to the positive effect of mSRC is seen in the top-left and mid-left sub-plots, for sparsity-levels of 9. Here, a clear negative effect of using multiple representations is found, and the hit-rates move asymptotically towards 0.5, that is, a seemingly random class assignments. In both cases, the hit-rate is quite poor when using just a single sparse representation.

A final observation from Fig. 6.4 is that when only a single sparse representation is used, a sparsity-level of 1 performs best. When using multiple sparse representation, higher sparsity-levels perform better. This effect is discussed in more detail in section 6.6.

In Fig. 6.5 the performance of mSRC is compared with that of basic SRC, K-NN, and SVM for the pairwise texture classifications. For SVM, the patch size, and the parameters C and γ were optimized. For visualization purposes, only the SVM results with the optimal $\gamma = 2^{-19}$ are included in the figure.

For small patch-sizes, the performance of SRC and mSRC is generally inferior to that of K-NN and SVM (in fact, in several cases not even in the displayed hit-rate range) but for larger patch sizes, all methods perform very well.

SRC with a sparsity-level of 1 generally performs almost as good as the highest scoring sparsity-level of mSRC. For sparsity-levels larger than 1, mSRC consistently outperforms SRC.

6.5.2 Carotid artery lumen segmentation

Parameter tuning

In Fig. 6.6, the results of the parameter tuning for carotid artery lumen segmentation are shown. In these experiments the dictionary size is 1000 atoms. It can be seen how the segmentation performance generally improves with the number of sparse representations used, resembling the results of the texture experiments. As in the texture experiments, an exception to the trend occurs (in some of the cases)

for sparsity-levels of 1: after a small increase, the performance decreases slightly when using more sparse representations.

It is clear that the patch size has a significant impact on the segmentation performance, which increases when going from a patch size of $3 \times 3 \times 3$ to $7 \times 7 \times 7$ voxels. The highest Dice scores obtained with patch size $9 \times 9 \times 9$ are slightly better than those of $7 \times 7 \times 7$. Experiments were also performed with patch sizes of $11 \times 11 \times 11$ voxels but the results of these were inferior to those of patch-sizes of both $7 \times 7 \times 7$ and $9 \times 9 \times 9$ voxels and have not been included.

The effect of the sparsity-level is clear for small numbers of sparse representations. When using only a single sparse representation, there is a consistent order to the performance of the sparsity-levels: using 1 atom yields best performance, followed by using 3, 5, 7, and 9 atoms. Interestingly, this general order was also seen in the texture experiments. When more sparse representations are used, sparsity levels of 3, 5, and 7 consistently outperform sparsity-levels of 1 and 9.

Further experiments were performed to examine the influence of the dictionary size on the performance. The results are visualized in Fig. 6.7 for dictionaries of $\{200, 500, 1000, 2000, 4000\}$ atoms using SRC and mSRC with 15 sparse representations. For mSRC, the performance of dictionaries of sizes 2000 and 4000 atoms is very similar and only slightly better than when using a dictionary of 1000 atoms. The overall optimal parameters are a dictionary size of 2000 atoms, a patch size of $9 \times 9 \times 9$ voxels, and a sparsity-level of 7 atoms. These parameters were used to test the method on the test-set.

For the reference methods, the optimal parameters were found by parameter tuning as well. For SRC, the following parameters yielded the best Dice score: dictionary size of 500 atoms, a patch size of $7 \times 7 \times 7$ voxels, and a sparsity-level of 1. For K-NN, patches of size $\{(3 \times 3 \times 3), (5 \times 5 \times 5), (7 \times 7 \times 7), (9 \times 9 \times 9), (11 \times 11 \times 11)\}$ voxels and Ks of $\{1, 11, 21, 31, 41\}$ were tested. The optimal parameters for K-NN were found to be a patch size of $3 \times 3 \times 3$ voxels and a K of 31. For SVM, the same patch sizes were tested and the optimal parameters were a patch size of $3 \times 3 \times 3$, $C = 2$, and $\gamma = 2^{-19}$.

Performance on test-set

The results of the experiments are shown in Fig. 6.8. The best performance, measured by median Dice score, was achieved by mSRC (0.855) followed by K-NN (0.845), and SRC (0.779). SVM performed considerably worse than the other methods with a Dice score of 0.440.

6.5.3 Carotid artery bifurcation detection

Parameter tuning

The results of the parameter tuning in the bifurcation detection experiments are shown in Fig. 6.9. Again, the patch size has a clear effect: for larger patch sizes, the

bifurcation points are detected with smaller median distances to the ground truth points, and the seemingly erratic effect of the sparsity-level is reduced.

The effect of using multiple sparse representations is not clear. For patch sizes of $9 \times 9 \times 9$ and $11 \times 11 \times 11$ voxels, when using more than one sparse representation there is a tendency of decreased distance to the ground truth points, but in many cases this effect reverses when using a high number of representations. From these plots we find that using nine sparse representations yields the best results, but in the lack of a clear trend, no general conclusion can be drawn.

The heatmap visualization in Fig. 6.10 shows the effect of dictionary size and sparsity-level for given patch sizes. Based on the results in Fig. 6.9, nine sparse representations were used here. Again, it is hard to draw general conclusions about the effect of the parameter settings from these plots, but there is a tendency for lower sparsity-levels to perform better. The optimal settings for mSRC are a dictionary size of 500 atoms, a patch size of $9 \times 9 \times 9$ voxels, a sparsity-level of 1, and the nine sparse representations as determined by the previous experiment (Fig. 6.10). These parameters were used for mSRC bifurcation detection on the test-set.

For the reference methods, the optimal parameters were found by parameter tuning as well. For SRC, the same parameters were tested as for mSRC. The optimal set of parameters were found to be a dictionary size of 500, a patch size of $9 \times 9 \times 9$ voxels, and a sparsity-level of 7. For K-NN, patches of size $\{(3 \times 3 \times 3), (5 \times 5 \times 5), (7 \times 7 \times 7), (9 \times 9 \times 9), (11 \times 11 \times 11)\}$ voxels and K-values of $\{100, 200, 300, 400, 500\}$ were tested. The optimal parameters for K-NN were found to be a patch size of $5 \times 5 \times 5$ voxels and a K of 200. For SVM, the same patch sizes were tested and the optimal parameters were a patch size of $3 \times 3 \times 3$, $C = 8192$, and $\gamma = 2^{-19}$.

Performance on test-set

The results on the test-set are shown in Fig. 6.11. The best performance, measured by median distance, of the detected bifurcation points to the ground truth, was achieved by K-NN (1.35 mm), followed by mSRC (1.46 mm), and SRC (2.06 mm). SVM performed considerably worse than the other methods with a median distance of 9.71 mm.

6.6 Discussion

In this work we have presented a generalization of the basic SRC classification scheme, an approach which we have called *multiple* SRC and that coincides with SRC in the case where $L = 1$ in 6.5. The effectiveness of the generalization relies on an extended exploitation of the statistical redundancy of the dictionaries.

6.6.1 mSRC versus SRC (effect of the number of representations)

In Figs. 6.4 and 6.6, the effect of using multiple sparse representation for voxel-wise classification is clear. For sparsity-levels higher than 1, the classification performance increases considerably. The largest increase in performance is seen for the first few additional representations.

In the texture classification experiments there is little gain by using mSRC over SRC. For small patch sizes there is some improvement in performance but for larger patches both methods yield near perfect classification. In fact, all tested methods (depending on the parameter settings) yield near perfect classification in these experiments (see Fig. 6.5), so the problem may be too easy for an informative comparison of the methods.

In the carotid artery lumen segmentation problem, on the other hand, there is a clear difference in the performance between SRC and mSRC. While SRC performs reasonably with a sparsity-level of 1, there are pronounced improvements in performance when using mSRC in combination with higher sparsity-levels and patch sizes larger than $3 \times 3 \times 3$. In the following we shall discuss in detail how these parameters interact.

6.6.2 Effect of patch size

The patch size has a significant impact on the classification performance, not only for (m)SRC, but also for K-NN and SVM. This was to be expected, since it is the patch size that determines whether sufficient local information is captured to enable class discrimination. In the degenerate case of using a single voxel as the feature, no discriminative properties are present (intensity information is discarded in (m)SRC due to the scaling factor in its generative model). The patch can also become too large. If a patch captures more than the essential features of a class, then its distance to other samples of the class increases and its properties do not generalize to discrimination. So, for every application, the patch size should be optimized against the scale of the structural components of interest and the image resolution.

As noted in section 6.1, there is a qualitative difference between the highly textured natural images, and the piecewise smooth medical images. The textures included in this study are highly repetitive, and a good descriptor should capture at least a few periods of such repetitions. In Fig. 6.5 it is seen that a patch size of 5×5 or 7×7 in the texture experiments is sufficient to capture the discriminative structures for both mSRC and the reference methods. Larger patches perform even better but due to a larger overlap they are expected to perform poorly in border regions of adjacent textures.

In the carotid MR images, patch sizes of $7 \times 7 \times 7$ and $9 \times 9 \times 9$ capture the discriminative structures well for (m)SRC, in both the segmentation and bifurcation point detection experiments. Interestingly, for this application K-NN and SVM perform better for smaller patch sizes ($3 \times 3 \times 3$ and $5 \times 5 \times 5$) than for larger ones.

The reason for this may be that both K-NN and SVM include the patch intensity in their class model, and that small patches are sufficient to capture the intensity.

6.6.3 Effect of sparsity-level

Given a dictionary size, the sparsity-level determines the flexibility of the representation. The higher the sparsity-level, the easier it becomes to approximate a given signal, be it from the dictionary's class or not.

One clear result to be observed from Figs. 6.4 and 6.6 is that the flexibility provided by higher sparsity-levels does not help in SRC (that is, mSRC using 1 representation): in all plots, the classification/segmentation performance is ordered by the increasing sparsity-levels, with a sparsity-level of 1 on top.

When a single representation is used, the representation errors are high, and so are the differences between the errors from the various class dictionaries. When the representation power of a dictionary increases by using higher sparsity-levels, the errors decrease, as does the difference in error between the class dictionaries.

This effect was confirmed by an experiment illustrated in Fig. 6.12. The plots show the sum of squared errors (SSE) (normalized by the sum of squares of the corresponding input signals) of the representations in a carotid lumen segmentation. In the plots the noise level is shown as constant line in the bottom of each sub-figure. The noise level is based on the standard deviation of the lumen region σ:

$$noise = \frac{1}{M} \sum_{i=1}^{M} \frac{N\sigma^2}{\sum_{j=1}^{N} x_{ij}^2} \ , \tag{6.6}$$

where M is the number of feature vectors in the entire ROI, and N is the dimensionality of the feature vectors, and x_{ij} denotes the jth element of the ith feature vector of the ROI.

It is seen, that for a sparsity-level of 9 (Fig. 6.12, bottom row), both the correct and the incorrect dictionaries yield representation errors close to the noise level. When the representation error gets lower, it becomes more influenced by the noise and so does the classification. In mSRC, due to the slightly lower accuracy (higher normalized SSE) of the subsequent representations, the effect of overfitting is reduced (see Fig. 6.12, right column). We thus reason, that the improvement of the classification/segmentation results, when using high sparsity-levels and multiple representations, are caused by a *joint effect* of statistical accumulation of evidence of class relationship *and* a reduction of overfitting in subsequent representations.

6.6.4 Effect of dictionary size

The dictionary size also provides flexibility in the representation. The larger/more overcomplete the dictionary, the more sparsely a signal can be represented. The

effect of dictionary size has been tested for the carotid segmentation and bifurcation experiments. The results are displayed in Figs. 6.7 and 6.10. While the latter figure does not show a large impact of the dictionary size (or other parameters for that matter), in Fig. 6.7 we notice at least one effect: in SRC (top row) the size of the dictionary has only a small influence on the performance, and increasing the sparsity-level (as we discussed above) has a small negative effect. In mSRC (bottom row), however, the dictionary size has a pronounced positive impact on the results, although mainly for high sparsity-levels.

The lower performance when using a small dictionary size and high sparsity in mSRC may be easily explained: having high sparsity-levels and small dictionaries, the proportion of atoms that are removed after each representation simply becomes too large: after 15 representations with a sparsity-level of 9, 135 atoms have been removed from the dictionary. The good performance of large dictionaries, on the other hand, may be explained by the increased redundancy of large dictionaries. Using larger dictionaries, we can expect to accurately represent signals a larger number of times than with smaller dictionaries. According to this reasoning, a dictionary of 4000 atoms should perform better than one of 2000 atoms. This is, however, not the case in our experiments. The explanation may be the same as in the case of high sparsity-levels: the redundancy provided by large dictionaries increases the chance of overfitting.

6.6.5 mSRC for classification, segmentation, and point detection

The mSRC method performs well in all the three tested settings and our hypothesis of improved performance over SRC has been confirmed. In the texture classification and lumen segmentation experiments, the effect of fusing multiple representations is nicely reflected by the increasing hit-rates and Dice scores.

The texture classification experiment presented a problem that was easily solved by mSRC as well as SRC and the two reference methods under the right parameter settings. However, mSRC did outperform SRC in the difficult case when a small patch size was used.

The carotid lumen segmentation problem resembles the problems of prostate and hippocampus segmentation in [26] and [92]. The results of these experiments add to the evidence of the applicability of SRC methods in medical image analysis. In this application, we have seen the most clear example of the improvement in mSRC over SRC, and mSRC also clearly outperformed the reference methods K-NN and SVM.

In the bifurcation point detection, the results of mSRC were reasonable (though not as good as K-NN), but the curves and heatmaps in Figs. 6.9 and 6.10 look somewhat erratic. Two tendencies are clear though: First, fusing 3 representations in mSRC consistently works better than SRC. Second, larger patches lead to more stable performance and apparently less dependence on the sparsity-level. It is clear,

that further experiments are needed to investigate the effect of the mSRC parameters in this setting.

One important difference between the classification and segmentation experiments and the point detection task is a huge difference in the statistical basis of the results. In the classification and segmentation settings, the hit-rate and Dice scores are based upon thousands of voxelwise classifications. In the point detection task, a single voxel in a ROI of 100,000 voxels must be found, and the median distance measure shown in the result figures is based on only 16 such detections (20 in the experiment on the test-set). To properly test the effect of mSRC for bifurcation point detection, a much larger test-set would be needed.

6.6.6 Future work

In this study, we have presented the idea of fusing multiple sparse representations into a more statistically informed and robust classifier. We have provided a thorough validation of the method and compared it with two baseline methods. Our approach generalizes SRC and any classification method that uses sparse representation and one or more overcomplete dictionaries should in principle be able to exploit the mSRC idea. In the future, it would be very interesting to see how the mSRC approach works when applied to other SRC methods and for other applications.

A potential drawback of SRC, in the form that we have presented in this study, is that it is independent of intensity information. While this property is advantageous in settings such as lumination invariant object detection or texture classification, it essentially disregards the most informative feature in many types of images, including medical ones. It would be of great interest to re-introduce the intensity information in the classification, for example by including the sparse coding vectors, in which the scaling information is stored.

6.6.7 Conclusion

In summary, we have presented a novel and effective generalization to the basic SRC classification scheme. Where basic SRC uses the residuals of a single sparse representation per class for classification, the presented method fuses multiple sparse representations. The method relies on the joint effect of the statistical accumulation of evidence of class relationship over the multiple representations, and a reduced risk of overfitting in the representations that follow the first one. The resulting classifier is more accurate than the basic SRC.

6.7 Acknowledgments

We thank SURFsara (www.surfsara.nl) for the support in using the Lisa Compute Cluster.

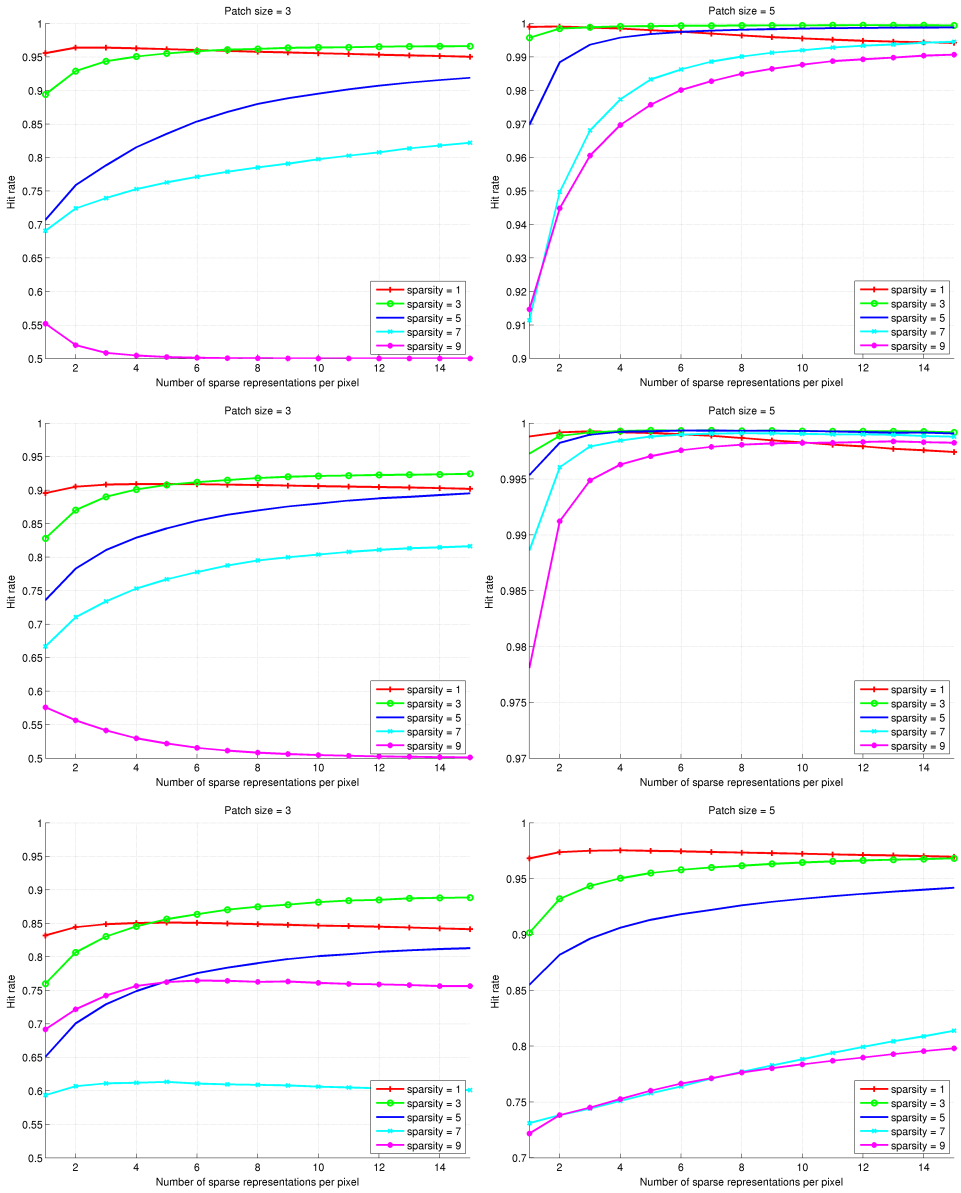


Figure 6.4: *Results of the pairwise texture classification experiments. Top tow: classification of texture D21 and D28. Middle row: classification of texture D21 and D77. Bottom row: classification of texture D28 and D77. Experiments with patch sizes of* $3 \times 3 \times 1$ *and* $5 \times 5 \times 1$ *pixels are shown in left and right columns, respectively. Note that the* y*-axes of the plots are scaled differently to allow appreciation of the effect of using multiple sparse representations.*

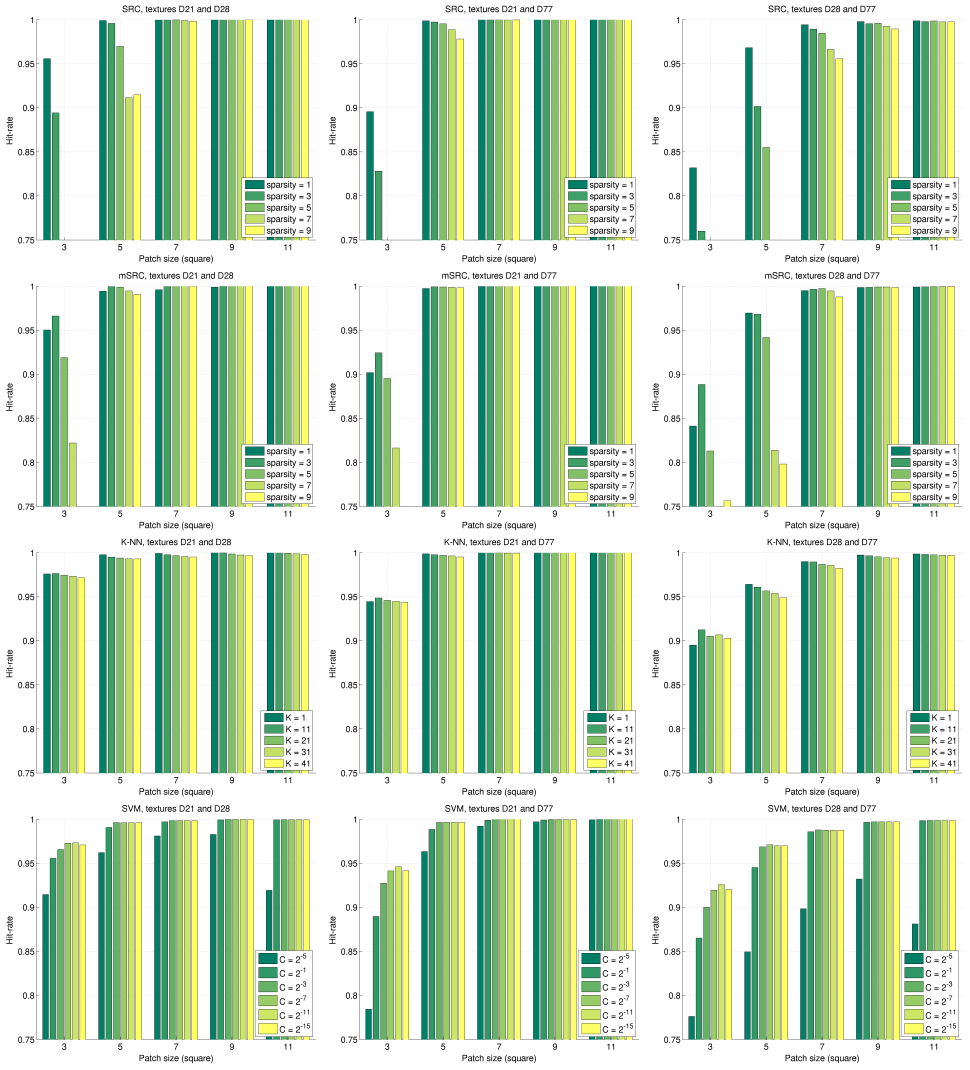


Figure 6.5: *Comparison of the hit-rates of four methods for texture classification. Top to bottom: results of SRC, mSRC, K-NN, and SVM. Left to right: pairwise texture classification of D21 and D28, D21 and D77, and D28 and D77.*

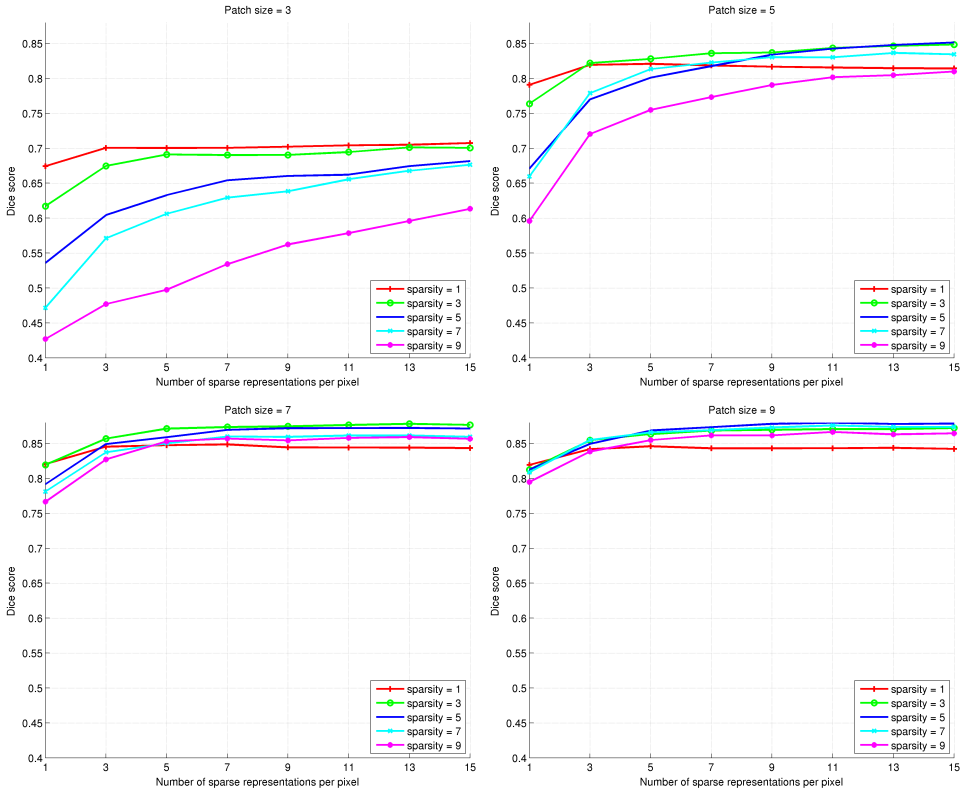


Figure 6.6: *Dice scores in carotid segmentation experiments for parameter tuning. The sub-plots show (in scan-line order) experiments with patch sizes* $\{(3\times3\times3),(5\times5\times5),(7\times7\times7),(9\times9\times9)\}$. *In each sub-plot experiments with sparsity-levels of* $\{1,3,5,7,9\}$ *are included.*

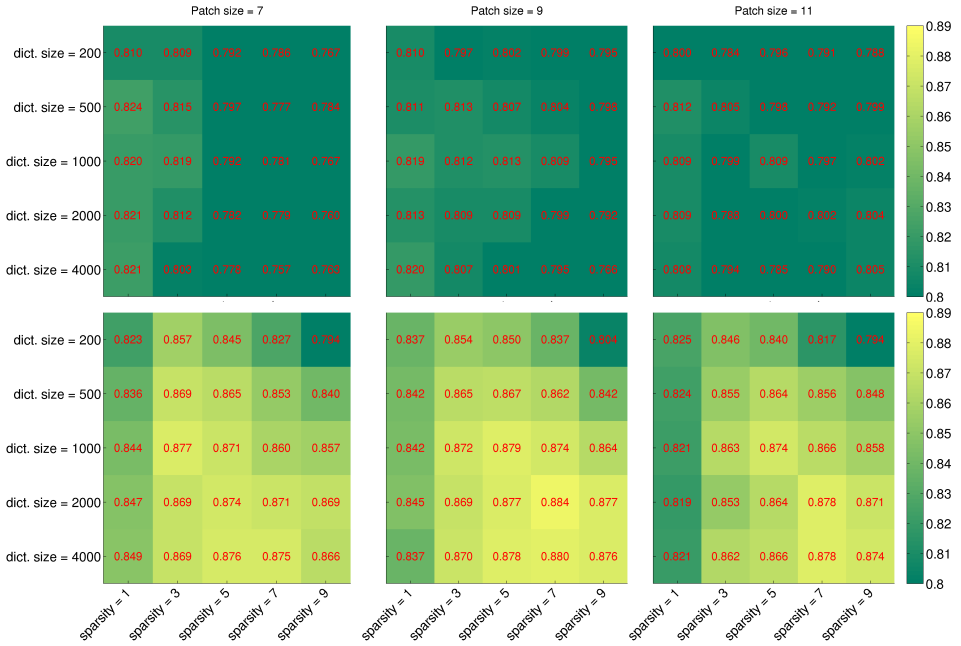


Figure 6.7: *Heatmaps of carotid lumen segmentation results. Numerical and color coded Dice scores are shown for dictionary sizes* $\{200, 500, 1000, 2000, 4000\}$*, patch sizes* $\{(7 \times 7 \times 7), (9 \times 9 \times 9), (11 \times 11 \times 11)\}$*, and sparsity-levels* $\{1, 3, 5, 7, 9\}$*. Top row: a single representation is used for classification. Bottom row: 15 representations are used for classification.*

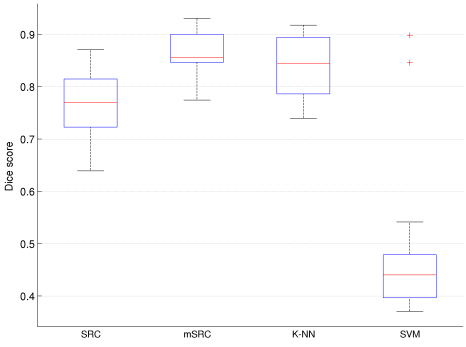


Figure 6.8: *Boxplots of the lumen segmentation performance of SRC, mSRC, K-NN, and SVM on the test-set.*

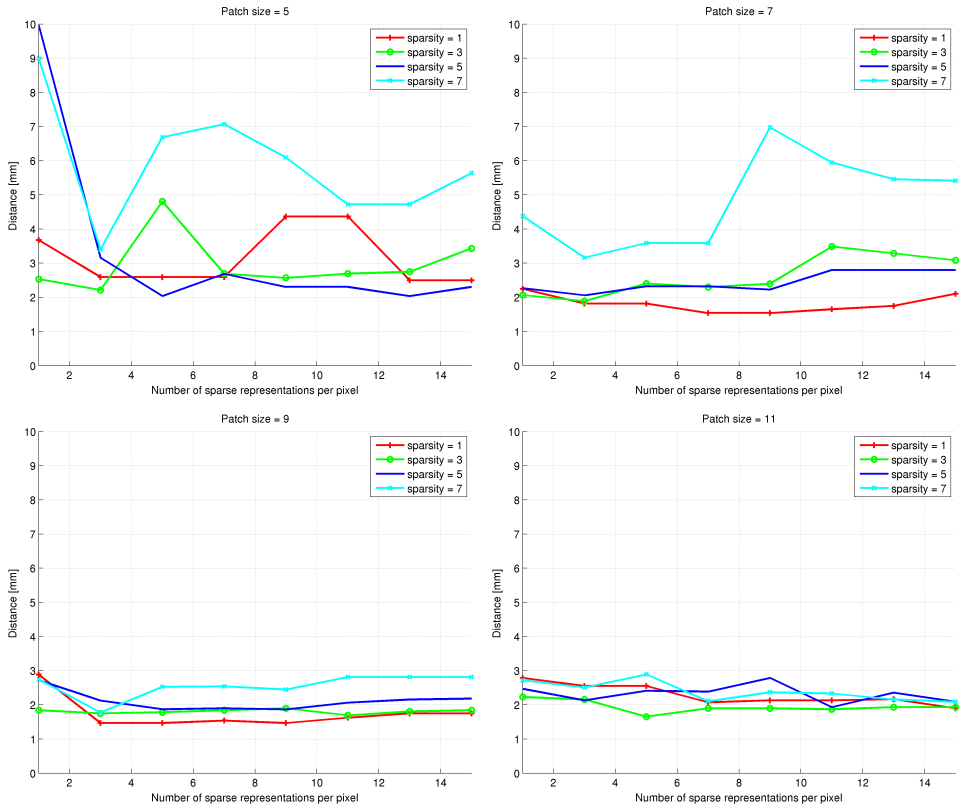


Figure 6.9: *Plot of distances (Euclidean distances in mm) between true bifurcation points and detected ones. Plots are shown for patch sizes of* $\{(5\times5\times5),(7\times7\times7),(9\times9\times9),(11\times11\times11)\}$*, and sparsity levels* $\{1,3,5,7\}$.

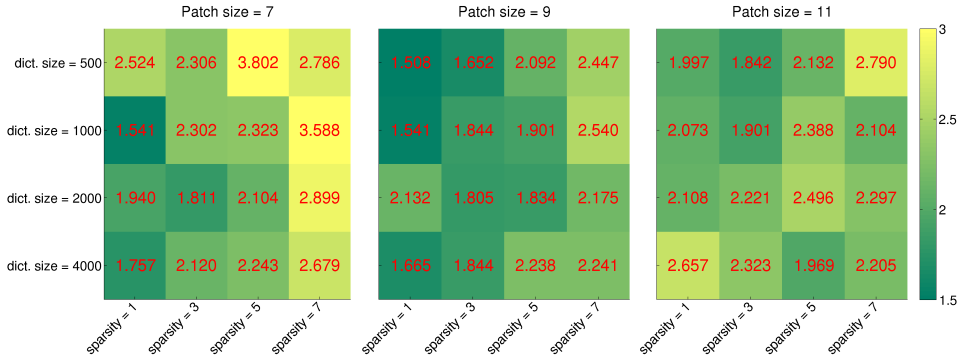


Figure 6.10: *Heatmaps of bifurcation detection results. Numerical and color coded distances are shown for different dictionary sizes* $\{500, 1000, 2000, 4000\}$, *patch sizes* $\{(7 \times 7 \times 7), (9 \times 9 \times 9), (11 \times 11 \times 11)\}$, *and sparsity levels* $\{1, 3, 5, 7\}$.

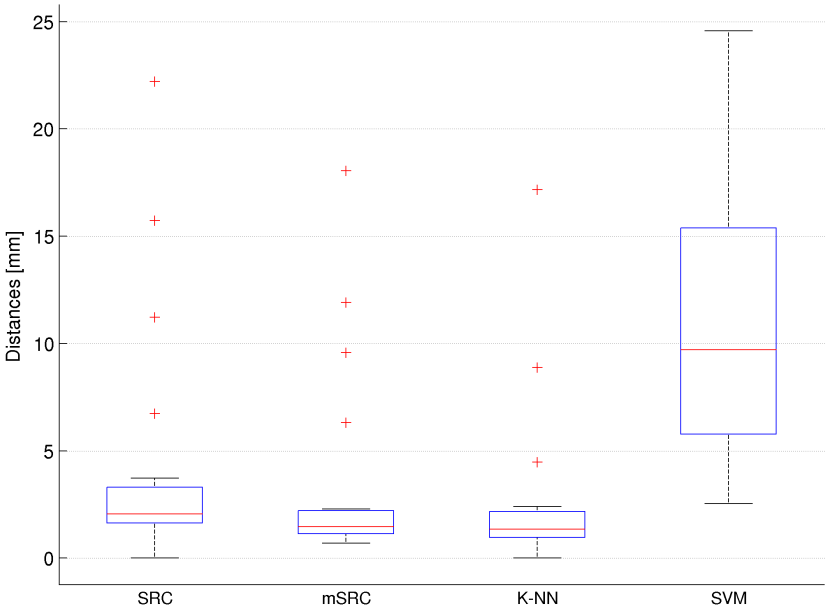


Figure 6.11: *Boxplots of the bifurcation detection performance of SRC, mSRC, K-NN, SVM on the test-set.*

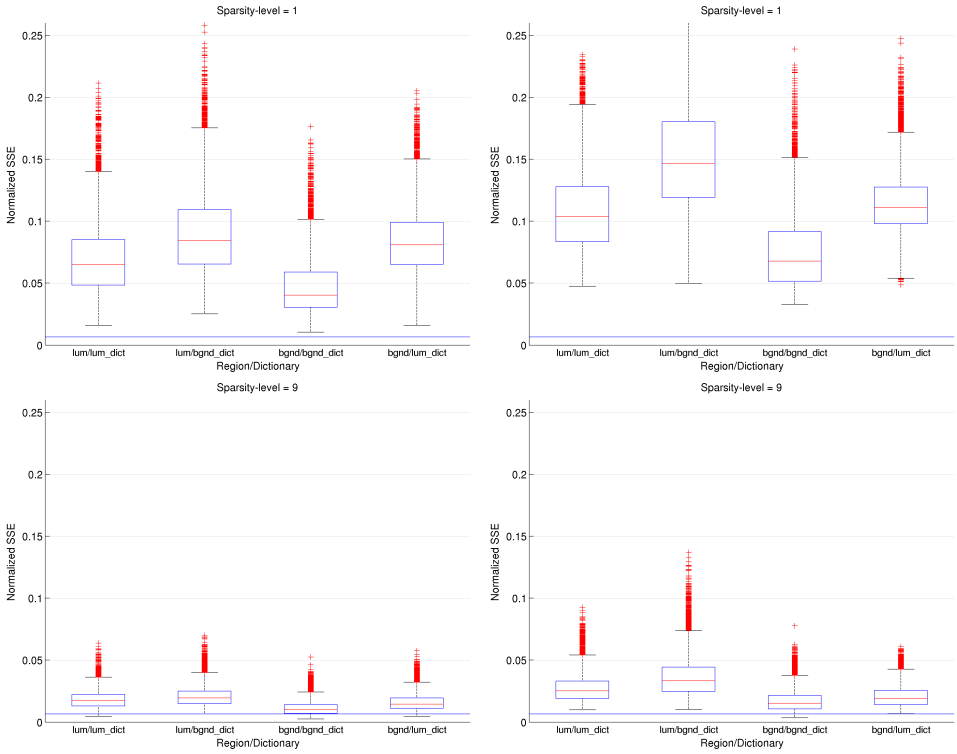


Figure 6.12: *Boxplots of the normalized SSE errors of dictionary representations of sparsity-levels 1 and 9. In the left column, the errors of a single representation are shown. In the right column, the average error of 15 representations are used. The carotid lumen is denoted* lum*, and the background is denoted* bgnd*. The noise level (see Eq. (6.6)) is shown as a constant line in each sub-figure.*

Chapter 7

Summary and Discussion

In this thesis, we have investigated super-resolution reconstruction (SRR) and self-similarity and their application in MRI. We have explored the feasibility of SRR in MRI, both from a general point of view and for several specific applications in small animal imaging. We have developed a new SRR method that exploits multi-scale self-similarity properties that are specific to multi-slice MRI. The exploration of self-similarity in MRI was continued in a study of its application for voxelwise classification. This effort resulted in another new method.

Below, we briefly summarize the contributions of each of the five studies included in this thesis. This will be followed by a discussion of a number of issues encountered in the studies. Finally, perspectives for future research and applications are provided.

7.1 Summary

We began, in chapter 2, by presenting an extensive experimental study of the trade-off between resolution, SNR and acquisition time of SRR in MRI and a comparison of these parameters to direct high-resolution 3D imaging. The main conclusion was that when resolution is limited in the slice-selection direction and the visualization and detection of critical events is hampered, then SRR provides a simple strategy to extend the capabilities of the scanner. In addition, it was clear from the trade-off analysis and experiments that SRR may be a better strategy than direct 3D imaging when SNR requirements make it necessary to average multiple 3D acquisitions.

Armed with the capabilities of visualizing structures smaller than the intrinsic resolution of the MRI scanner, we identified and addressed the specific research problem of mouse brain nuclei visualization in chapter 3. The mouse brain has a diameter of about 15 mm. The study of brain substructures, or *nuclei*, obviously requires high imaging resolution. In manganese-enhanced mouse brain MRI, 3D gradient echo (GE) imaging sequences are usually applied due to their ability to obtain isotropic resolution. It has been shown that using inversion recovery spin echo (IR-SE), a 2D imaging sequence, results in different contrast properties. For studies that involve visualization of brain nuclei, however, the anisotropy of such images severely limits how well they can be analyzed. In the study, we demonstrated how SRR can produce isotropic IR-SE images of approximately $0.1 \times 0.1 \times 0.1$ mm, in which a large number of nuclei could be identified, and more importantly, that a number of these nuclei were better visible in the SRR IR-SE images than in the GE images. In manganese-enhanced mouse brain imaging, SRR IR-SE thus constitutes an imaging method complementary to conventional GE images.

SRR methods can be a computationally expensive. In chapter 4, a practical approach to overcoming this issue was presented. Inspired by the functionality of Google Maps, an interactive image visualization tool was devised that allowed global-to-local exploration of a full-body mouse. The result was a much improved

visualization of small tumors and other potential structures of interest. The key contributions of the study were the facilitation of a near real-time SRR process, and the reproduction of superior visualization by SRR in yet another experimental setting.

In the second part of the thesis, we considered image self-similarity and sparse coding using learned dictionaries, and specifically its application to MRI reconstruction and voxelwise classification. This work focused on methodological innovation and basic validation of the proposed methods.

In chapter 5, we explored the simple idea of exploiting self-similarity across the two native scales in an MRI. The high-resolution information available in-plane of a slice-stack was used as a 2D prior for the reconstruction of the anisotropic planes that includes the slice-selection direction. The method can be applied to both single and multiple slice-stacks. The results indicated that the cross-scale self-similarity prior is capable of exploiting image information that is not used by a conventional state-of-the-art Tikhonov-regularized SRR method.

In chapter 6, other applications of self-similarity in MR imaging were considered. The redundancy of a dictionary trained on a head and neck MRI was shown quantitatively. This was used to motivate a generalization of the basic sparse representation classification (SRC) scheme by using *multiple* sparse representations. Our experiments showed that the generalized method improved classification. In addition, they revealed an effect of reduced susceptibility to over-fitting in the classification process.

7.2 Discussion

The research described in this thesis points to a number of advantages, application areas and future opportunities in the fields of SRR and self-similarity in MRI. In the first part of the thesis, we have shown the capabilities of SRR and presented many arguments in favor of (at least considering) SRR for a range of MRI applications. In the second part, we have shown the potential of exploiting the self-similarity in MR images for a variety of applications: SRR, classification, segmentation and detection. Considering the thesis as a whole, there are a few issues that should be of interest to the research community.

7.2.1 Resolution measures

An obstacle when validating resolution enhancement is the lack of a good quantitative measure. In many studies, the peak signal-to-noise ratio (PSNR) is used. This measure captures the distance between a ground truth noise-free high-resolution image and the reconstructed one. It does, however, not distinguish between resolution enhancement and denoising, while both resolution and noise are affected by the reconstruction process. Another popular way of measuring resolution, is by fitting a sigmoid function to 1D profiles over step-edges in the reconstructed image.

One of the parameters in the sigmoid function controls the width of the transitional part of the function and can thus be used as a measure of the edge-width. The problem with this measure is that the presence of sharp edges does not necessarily equal a high resolution of the reconstructed image. Instead, such edges can be the result of a bias in the reconstruction process, introduced by the regularizer. Consider, for example, Tikhonov and total variation (TV) regularization, which impose smoothness and piecewise smoothness of the solution, respectively. In Tikhonov-regularization, edges as well as homogeneous areas are smoothed. In TV-regularized reconstruction, the smoothness is local to the piecewise constant regions and edges will thus not be smoothed, but rather, enhanced. There is, however, no difference between these two methods that will allow us to observe smaller structures in the TV image than in the Tikhonov image. Edge-width thus becomes a measure of sharpness rather than resolution.

While quantitative measures do provide information about the image quality in terms of SNR and sharpness, resolution is best appreciated qualitatively. In phantoms with structures such as the radial line pattern in the phantom used in chapter 2 the relative resolution of multiple competing images is readily available. We therefore recommend that resolution capabilities of new reconstruction methods are demonstrated and compared to other methods on such phantoms.

7.2.2 Information loss and artifacts from regularization

A risk of any under-sampled reconstruction scheme in which prior knowledge is used for regularization, is that the regularization process removes essential details or introduces artifacts that confuse the interpretation of the reconstructed images. In medical images, where pathologies are often indicated by subtle details, this is an especially sensitive issue. As described above, Tikhonov regularization leads to smooth images, also in locations that are not supposed to be smooth.

The self-similarity prior is susceptible to the same problems of discarding information and introducing artifacts. Instead of guiding the reconstruction towards a realistic solution by imposing analytical constraints, image statistics are used to constrain the solution. The self-similarity prior assumes the presence of similar structures within an image or over a group of images. If a detail, e.g. a pathology such as a small lesion, is *not* present in the training data, then it will be difficult to reconstruct this area accurately, and important information could be lost. This issue was addressed in [77], where high-resolution T1-weighted brain images were used for SRR reconstruction of T2-weighted slice-stacks that contained multiple sclerosis (MS) lesions not visible in the high-resolution T1 images. The recovery of the lesions was successful but the authors noted that the topic needs further investigation.

7.2.3 Towards practical tools

We are well aware that SRR is just one among many approaches to improve the resolution-SNR-acquisition time trade-off in MRI reconstruction. The alternatives include parallel imaging and compressed sensing technologies. One advantage of SRR compared to these other approaches, is that it is a post-acquisition reconstruction method. It relies on conventional 2D sequences and no scanner software modifications or additional hardware are necessary to create the high-resolution images. In that sense, it is an extremely simple and inexpensive way of enhancing the resolution capabilities of a given MRI system.

Planning and performing an SRR-MRI acquisition of an anatomical or pathological structure is simple. The multiple low-resolution images should be acquired in a rotated or orthogonal fashion, such as described in chapters 2-5. The user-interaction in the SRR process is limited: In the conventional Tikhonov regularized method, only the weight of the regularization is set by the user. For the cross-scale self-similarity SRR method, the blurring kernels of the dictionary should be tuned for optimal performance. Both these interactions are done once for a specific application setting. After this, SRR can be performed on new images without any user inputs.

Any sparse representation classification scheme relies on an accurately annotated training set. When such a set is available, tuning of the method parameters should be performed as described in chapter 6. Successfully performing a segmentation, such as the carotid artery segmentation, relies on a strict and standardized imaging protocol, preferably including a image quality check. In the MRI data-set used in chapter 6, there were several images that were blurred or otherwise artifacted. The segmentation performance would improve if such acquisitions had been excluded by visual inspection and repeated.

A standardized imaging protocol also allows simple heuristics to be applied in the image analysis pipeline. In the MRI slice-stacks used in chapter 6, for example, the left and right carotid arteries, including bifurcation points, are present in all acquisitions. Furthermore, the left carotid artery is located in the left half of the slice-stack, and the right carotid artery in the right half. Finally, in all slice-stacks the carotid arteries stretch over all slices, bifurcating somewhere inside the slice-stack. If this information is available a priori, it can help simplify the segmentation task and provide possibilities for sanity checks of the solution.

7.2.4 Decoupling acquisition and reconstruction

We saw in chapter 5 that the self-similarity prior improved the quantitative and qualitative results over the Tikhonov-regularized reference method. This indicates that information was activated by the dictionary and the patch-based reconstruction process that was not readily available in the low-resolution data itself or was removed by the Tikhonov-regularization. This opens the door for a new way of

thinking about image acquisition and reconstruction, in which image statistics are used to decouple the two. There is enough data out there to base the statistics on. In MRI protocols, it is common practice to include multiple acquisitions (often using multiple sequences) of the subject in order to enhance the information. Much of the information content of such a set of images is clearly self-similar, and this may be exploited to yield sparser sampling or, alternatively, improved image quality.

Another interesting possibility, would be to use dictionaries trained on class specific image content to regularize the *scanner side* reconstruction of MRI images, e.g. in a compressed sensing setup. In this way, analysis might be performed *during* reconstruction. This may be a relevant option in studies where mainly quantitative metrics, such as e.g. white or gray matter volume, are of interest.

Successfully exploiting SRR and self-similarity to enhance the quality and analysis of medical images, opens a wealth of opportunities for statistical image processing. It can improve high-end scanning systems, but also boost the performance of current clinical scanners, improving image-based diagnostics at a lower cost. The partial decoupling of sensing and representation provided by the dictionaries could allow data to be shared across centers, and data acquired on low-end scanning systems could be reconstructed or analyzed using libraries of data acquired on high-end scanners in other parts of the world. Such advances would empower remote and low-resource medical facilities to perform improved health care and research.

References

[1] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Sig Proc*, 54(11):4311–4322, Nov. 2006.

[2] I. Aoki, Y.-J. L. Wu, A. C. Silva, R. M. Lynch, and A. P. Koretsky. In vivo detection of neuroarchitecture in the rodent brain using manganese-enhanced MRI. *NeuroImage*, 22(3):1046 – 1059, 2004.

[3] M. Baiker, J. Milles, J. Dijkstra, T. D. Henning, A. W. Weber, I. Que, E. L. Kaijzel, C. W. Löwik, J. H. Reiber, and B. P. Lelieveldt. Atlas-based whole-body segmentation of mice from low-contrast micro-CT data. *Med Image Anal*, 14(6):723 – 737, 2010.

[4] M. Baiker, T. Snoeks, E. Kaijzel, I. Que, J. Dijkstra, B. Lelieveldt, and C. Löwik. Automated bone volume and thickness measurements in small animal whole-body microct data. *Mol Imaging Biol*, 2011.

[5] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Trans Pattern Anal Mach Intell*, 24(9):1167–1183, Sept. 2002.

[6] P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Trans Pattern Anal Mach Intell*, 14(2):239–256, Feb. 1992.

[7] N. A. Bock, F. F. Paiva, and A. C. Silva. Fractionated manganese-enhanced MRI. *NMR Biomed*, 21(5):473–478, 2008.

[8] C. Bolin, C. Sutherland, K. Tawara, J. Moselhy, and C. Jorcyk. Novel mouse mammary cell lines for in vivo bioluminescence imaging (BLI) of bone metastasis. *Biol Proced Online*, 14(1):6, 2012.

[9] S. Borman and R. L. Stevenson. Spatial resolution enhancement of low-resolution image sequences: A comprehensive review with directions for future research. Technical report, Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana, USA, July 1998.

[10] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 60–65, Washington, DC, USA, 2005. IEEE Computer Society.

[11] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM T Intell Sys Tech*, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[12] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Rev*, 43(1):129–159, Jan. 2001.

[13] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans. BrainWeb: Online interface to a 3D MRI simulated brain database. *NeuroImage*, 5:425, 1997.

[14] P. Coupé, S. F. Eskildsen, J. V. Manjón, V. S. Fonov, and D. L. Collins. Simultaneous segmentation and grading of anatomical structures for patient's classification: Application to Alzheimer's disease. *NeuroImage*, 59(4):3736–3747, Feb. 2012.

[15] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins. Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage*, 54(2):940–954, Jan. 2011.

[16] A. den Dekker and A. vd Bos. Resolution: a survey. *J Opt Soc Am (A)*, 14(3):547–557, Mar. 1997.

[17] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, July 1945.

[18] M. Ebrahimi and E. R. Vrscay. Solving the inverse problem of image zooming using "self-examples". In *Proceedings of the International Conference on Image Analysis and Recognition*, volume 4633, pages 117–130. Springer, 2007.

[19] M. Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer Publishing Company, Incorporated, 1st edition, 2010.

[20] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans Image Process*, 15(12):3736–3745, Dec. 2006.

[21] M. Elad and D. Datsenko. Example-based regularization deployed to super-resolution reconstruction of a single image. *Comput J*, 52(1):15–30, Jan. 2009.

[22] M. Elad and A. Feuer. Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Trans Image Process*, 6(12):1646–1658, 1997.

[23] M. Elad and I. Yavneh. A plurality of sparse representations is better than the sparsest one alone. *IEEE Trans Inf Theor*, 55(10):4701–4714, Oct. 2009.

[24] D. Fiat. Method of enhancing an MRI signal, 2001. United States Patent 6294914.

[25] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *IEEE Comput Graph Appl*, 22(2):56–65, Mar. 2002.

[26] Y. Gao, S. Liao, and D. Shen. Prostate segmentation by sparse representation based classification. *Med Phys*, 39(10):6372–6387, Oct. 2012.

[27] K. Gauvain, J. Garbow, S. Song, A. Hirbe, and K. Weilbaecher. MRI detection of early bone metastases in b16 mouse melanoma models. *Clin Exp Metastasis*, 22(5):403–11, 2005.

[28] A. Gholipour, J. Estroff, and S. Warfield. Robust super-resolution volume reconstruction from slice acquisitions: Application to fetal brain MRI. *IEEE Trans Med Imaging*, 29(10):1739–1758, 2010.

[29] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *IEEE 12th International Conference on Computer Vision*, 2009.

[30] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography. *J Theor Biol*, 29(3):471–481, 1970.

[31] H. Greenspan. Super-resolution in medical imaging. *Comput J*, 52(1):43–63, 2009.

[32] H. Greenspan, G. Oz, N. Kiryati, and S. Peled. MRI inter-slice reconstruction using super-resolution. *Magn Reson Imaging*, 20(5):437–446, 2002.

[33] B. Grünecker, S. F. Kaltwasser, Y. Peterse, P. G. Sämann, M. V. Schmidt, C. T. Wotjak, and M. Czisch. Fractionated manganese injections: effects on mri contrast enhancement and physiological measures in c57bl/6 mice. *NMR Biomed*, 23(8):913–921, 2010.

[34] H. Gudbjartsson and S. Patz. The Rician distribution of noisy MRI data. *Magn Reson Med*, 34(6):910–914, 1995.

[35] E. M. Haacke, R. W. Brown, M. R. Thomson, and R. Venkatesan. *Magnetic Resonance Imaging. Physical Principles and Sequence Design*. John Wiley and Sons, New York, 1999.

[36] K. Hameeteman, R. van 't Klooster, M. Selwaness, A. van der Lugt, J. C. Witteman, W. J. Niessen, and S. Klein. Carotid wall volume quantification from magnetic resonance images using deformable model fitting and learning-based correction of systematic errors. *Phys Med Biol*, 58(5):1605–1623, Mar. 2013.

[37] R. M. Heidemann, O. Özsarlak, P. M. Parizel, J. Michiels, B. Kiefer, V. Jellus, M. Müller, F. Breuer, M. Blaimer, M. A. Griswold, and P. M. Jakob. A brief review of parallel magnetic resonance imaging. *Eur Radiol*, 13(10):2323–2337, 2003.

[38] G. T. Herman, A. Lent, and S. W. Rowland. ART: Mathematics and applications. A report on the mathematical foundations and on the applicability to real data of the algebraic reconstruction techniques. *J Theor Biol*, 42(1):1–32, 1973.

[39] J. Herredesvela, K. Engan, T. O. Gulsrud, and K. Skretting. Texture classification using sparse representations by learned compound dictionaries. In *Proceedings of the Workshop on Signal Processing with Adaptative Sparse Structured Representations*, Rennes, France, Nov. 2005.

[40] C. Hindorf, J. Rodrigues, S. Boutaleb, J. Rosseau, A. Govignon, N. Anizan, A. Lisbona, and M. Bardies. Total absorbed dose to a mouse during microPET/CT imaging. In *Eur J Nucl Med Mol I, Volume 37, Issue 2 Supplement*, 2010.

[41] T. Hoch, S. Kreitz, S. Gaffling, M. Pischetsrieder, and A. Hess. Manganese-enhanced magnetic resonance imaging for mapping of whole brain activity patterns associated with the intake of snack food in ad libitum fed rats. *PLoS ONE*, 8(2):e55354, Feb. 2013.

[42] R. J. Immonen, I. Kharatishvili, A. Sierra, C. Einula, A. Pitkanen, and O. H. Grohn. Manganese enhanced MRI detects mossy fiber sprouting rather than neurodegeneration, gliosis or seizure-activity in the epileptic rat hippocampus. *NeuroImage*, 40(4):1718–1730, May 2008.

[43] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graph Models Image Process*, 53(3):231–239, 1991.

[44] S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin de*

l'Académie Polonaise des Sciences et Lettres A, pages 355–357, 1937.

[45] E. R. Kandel, J. H. Schwartz, and T. M. Jessell. *Principles of Neural Science*. McGraw-Hill Medical, 4th edition, July 2000.

[46] A. Khmelinskii, M. Baiker, X. J. Chen, J. H. C. Reiber, R. M. Henkelman, and B. P. F. Lelieveldt. Atlas-based organ bone approximation for ex-vivo MRI mouse data: a pilot study. In *Proceedings of the 7th International Symposium on Biomedical Imaging: from Nano to Macro*, pages 1197–1200, Piscataway, NJ, USA, 2010. IEEE Press.

[47] A. Khmelinskii, M. Baiker, E. Kaijzel, J. Chen, J. Reiber, and B. Lelieveldt. Articulated whole-body atlases for small animal image analysis: construction and applications. *Mol Imaging Biol*, 13(5):898–910, 2011.

[48] A. Khmelinskii, E. Plenge, P. Kok, O. Dzyubachyk, D. H. J. Poot, E. Suidgeest, C. P. Botha, W. J. Niessen, L. van der Weerd, E. Meijering, and B. P. F. Lelieveldt. Super-resolution reconstruction of whole-body mri mouse data: An interactive approach. In *Proceedings of the 9th International Symposium on Biomedical Imaging: from Nano to Macro*, pages 1723–1726. IEEE, 2012.

[49] J.-B. Kim, K. Urban, E. Cochran, S. Lee, A. Ang, B. Rice, A. Bata, K. Campbell, R. Coffee, A. Gorodinsky, Z. Lu, H. Zhou, T. K. Kishimoto, and P. Lassota. Non-invasive detection of a small number of bioluminescent cancer cells in vivo. *PLoS ONE*, 5(2):e9364+, Feb. 2010.

[50] P. Kok, M. Baiker, E. A. Hendriks, F. H. Post, J. Dijkstra, C. W. G. M. Löwik, B. P. F. Lelieveldt, and C. P. Botha. Articulated planar reformation for change visualization in small animal imaging. *IEEE Trans Vis Comput Graph*, 16(6):1396–1404, 2010.

[51] P. Kok, J. Dijkstra, C. P. Botha, F. H. Post, E. Kaijzel, I. Que, C. Löwik, J. Reiber, and B. Lelieveldt. Integrated visualization of multi-angle bioluminescence imaging and micro CT. In K. R. Cleary and M. I. Miga, editors, *Proc. SPIE Medical Imaging 2007*, volume 6509, 2007. 2007.01.01.

[52] J. H. Lee, A. C. Silva, H. Merkle, and A. P. Koretsky. Manganese-enhanced magnetic resonance imaging of mouse brain after systemic administration of MnCl2: Dose-dependent and temporal evolution of T1 contrast. *Magn Reson Med*, 53(3):640–648, 2005.

[53] M. S. Lewicki and T. J. Sejnowski. Learning overcomplete representations. *Neural Comput*, 12(2):337–365, Feb. 2000.

[54] Z. Lin and H.-Y. Shum. Fundamental limits of reconstruction-based superresolution algorithms under local translation. *IEEE Trans Pattern Anal Mach Intell*, 26:83–97, Jan. 2004.

[55] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med*, 58(6):1182–1195, 2007.

[56] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Discriminative learned dictionaries for local image analysis. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[57] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *Proceedings of the 12th International Conference on Computer Vision*, pages 2272–2279. IEEE, 2009.

[58] J. V. Manjón, P. Coupé, A. Buades, D. L. Collins, and M. Robles. MRI superresolution using self-similarity and image priors. *Int J Biomed Imaging*, 2010:1–11, 2010.

[59] J. V. Manjón, P. Coupé, A. Buades, V. Fonov, D. Louis Collins, and M. Robles. Non-local MRI upsampling. *Med Image Anal*, 14(6):784–792, Dec. 2010.

[60] S. Mok, J. Munasinghe, and W. Young. Infusion-based manganese-enhanced MRI: a new imaging technique to visualize the mouse brain. *Brain Struct Funct*, 217(1):107–14, 2012.

[61] B. A. Olshausen and D. J. Field. Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Res*, 37:3311–25, 1997.

[62] S. Park, M. Park, and M. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Process Mag*, 20(3):21–36, May 2003.

[63] R. R. Peeters, P. Kornprobst, M. Nikolova, S. Sunaert, T. Vieville, G. Malandain, R. Deriche, O. Faugeras, M. Ng, and P. V. Hecke. The use of superresolution techniques to reduce slice thickness in functional MRI. *Int J Imag Syst Tech*, 14(3):131–138, 2004.

[64] S. Peled and Y. Yeshurun. Superresolution in MRI: Application to human white matter fiber tract visualization by diffusion tensor imaging. *Magn Reson Med*, 45(1):29–35, 2001.

[65] S. Peled and Y. Yeshurun. Superresolution in MRI - Perhaps sometimes. *Magn Reson Med*, 48(2):409–409, 2002.

[66] G. Peyre. Sparse modeling of textures. *J Math Imaging Vis*, 34(1):17–31, 2009.

[67] J. G. Pipe. Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging. *Magn Reson Med*, 42(5):963–969, 1999.

[68] E. Plenge, D. H. Poot, W. J. Niessen, and E. Meijering. Super-resolution reconstruction using cross-scale self-similarity in multi-slice MRI. In *Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part III*, pages 123–130, Berlin, Heidelberg, 2013. Springer-Verlag.

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

[70] D. S. Poole, N. Doorenweerd, A. Mahfouz, M. J. Reinders, and L. van der Weerd. Continuous infusion of manganese improves contrast and reduces side effects in MEMRI studies. In *Proceedings of 20th annual meeting of International Society of Magnetic Resonance in Medicine 2012*, 2012.

[71] D. H. J. Poot, V. Van Meir, and J. Sijbers. General and efficient super-resolution method for multi-slice MRI. In *Proceedings of the 13th International Conference on Medical Image Computing and Computer-Assisted Intervention: Part I*, pages 615–622, Berlin, Heidelberg,

2010. Springer-Verlag.

[72] M. Protter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the nonlocal-means to super-resolution reconstruction. *IEEE Trans Image Process*, 18(1):36–51, 2009.

[73] K. Pruessmann. Encoding and reconstruction in parallel MRI. *NMR Biomed*, 19(3):288–299, 2006.

[74] I. Ramirez, P. Sprechmann, and G. Sapiro. Classification and clustering via dictionary learning with structured incoherence and shared features. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 3501–3508, Los Alamitos, CA, USA, 2010. IEEE Computer Society.

[75] T. Randen and J. H. Husøy. Filtering for texture classification: A comparative study. *IEEE Trans Pattern Anal Mach Intell*, 21(4):291–310, Apr. 1999.

[76] D. Robinson and P. Milanfar. Statistical performance analysis of super-resolution. *IEEE Trans Image Process*, 15(6):1413–1428, 2006.

[77] F. Rousseau. A non-local approach for image super-resolution using intermodality priors. *Med Image Anal*, 14(4):594–605, Aug. 2010.

[78] F. Rousseau, O. A. Glenn, B. Iordanova, C. Rodriguez-Carranza, D. B. Vigneron, J. A. Barkovich, and C. Studholme. Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images. *Acad Radiol*, 13(9):1072–1081, 2006.

[79] F. Rousseau, P. A. Habas, and C. Studholme. A supervised patch-based approach for human brain labeling. *IEEE Trans Med Imaging*, 30(10):1852–1862, Oct. 2011.

[80] R. Rubinstein, A. M. Bruckstein, and M. Elad. Dictionaries for sparse representation modeling. *Proceedings of the IEEE*, 98(6):1045–1057, June 2010.

[81] K. Scheffler. Superresolution in MRI? *Magn Reson Med*, 48(2):408; author reply 409, 2002.

[82] W. Segars, B. M. Tsui, E. C. Frey, G. Johnson, and S. S. Berr. Development of a 4-D digital mouse phantom for molecular imaging research. *Mol Imaging Biol*, 6(3):149–159, 2004.

[83] R. Z. Shilling, T. Q. Robbie, T. Bailloeul, K. Mewes, R. M. Mersereau, and M. E. Brummer. A super-resolution framework for 3-D high-resolution and high-contrast imaging using 2-D multislice MRI. *IEEE Trans Med Imaging*, 28(5):633–644, 2009.

[84] A. C. Silva, J. H. Lee, C. W. Wu, J. Tucciarone, G. Pelled, I. Aoki, and A. P. Koretsky. Detection of cortical laminar architecture using manganese-enhanced MRI. *J Neurosci Methods*, 167(2):246–257, Jan. 2008.

[85] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger. Shiftable multiscale transforms. *IEEE Trans Inf Theor*, 38(2):587–607, Sept. 2006.

[86] K. Skretting and J. H. Husøy. Texture classification using sparse frame-based representations. *EURASIP J Appl Signal Process*, 2006:102–102, Jan. 2006.

[87] T. Snoeks, A. Khmelinskii, B. Lelieveldt, E. Kaijzel, and C. Löwik. Optical advances in skeletal imaging applied to bone metastases. *Bone*, 48(1):106 – 114, 2011.

[88] H. Stark and P. Oskoui. High-resolution image recovery from image-plane arrays, using convex projections. *J Opt Soc Am A*, 6(11):1715–1726, 1989.

[89] A. Takeda, J. Sawashita, and S. Okada. Biological half-lives of zinc and manganese in rat brain. *Brain Research*, 695(1):53 – 58, 1995.

[90] M. Thuen, M. Berry, T. B. Pedersen, P. E. Goa, M. Summerfield, O. Haraldseth, A. Sandvig, and C. Brekken. Manganese-enhanced MRI of the rat visual pathway: Acute neural toxicity, contrast enhancement, axon resolution, axonal transport, and clearance of Mn2+. *J Magn Reson Im*, 28(4):855–865, 2008.

[91] I. Tindemans, T. Boumans, M. Verhoye, and A. Van der Linden. IR-SE and IR-MEMRI allow in vivo visualization of oscine neuroarchitecture including the main forebrain regions of the song control system. *NMR Biomed*, 19(1):18–29, 2006.

[92] T. Tong, R. Wolz, P. Coupe, J. V. Hajnal, D. Rueckert, and T. A. D. N. Initiative. Segmentation of MR images via discriminative dictionary learning and sparse coding: application to hippocampus labeling. *NeuroImage*, 76:11–23, Aug. 2013.

[93] J. A. Tropp and A. C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans Inf Theor*, 53(12):4655–4666, Dec. 2007.

[94] R. Y. Tsai and T. S. Huang. Multi-frame image restoration and registration. In *Advances in Computer Vision and Image Processing*, volume 1, pages 317–339. JAI Press Inc., 1984.

[95] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*, 29(6):1310–1320, June 2010.

[96] A. Van der Linden, N. Van Camp, P. Ramos-Cabrer, and M. Hoehn. Current status of functional MRI on small animals: application to physiology, pathophysiology, and cognition. *NMR Biomed*, 20(5):522–545, 2007.

[97] A. W. M. van Eekeren, K. Schutte, O. R. Oudegeest, and L. J. van Vliet. Performance evaluation of super-resolution reconstruction methods on real-world data. *EURASIP J Adv Signal Process*, 2007:1–12, 2007.

[98] E. Van Reeth, I. W. K. Tham, C. H. Tan, and C. L. Poh. Super-resolution in magnetic resonance imaging: A review. *Concepts in Magnetic Resonance Part A*, 40A(6):306–325, 2012.

[99] R. van 't Klooster, P. J. de Koning, R. A. Dehnavi, J. T. Tamsma, A. de Roos, J. H. Reiber, and R. J. van der Geest. Automatic lumen and outer wall segmentation of the carotid artery using deformable three-dimensional models in MR angiography and vessel wall images. *J Magn Reson Imaging*, 35(1):156–165, Jan. 2012.

[100] Y. Z. Wadghiri, J. A. Blind, X. Duan, C. Moreno, X. Yu, A. L. Joyner, and D. H. Turnbull. Manganese-enhanced magnetic resonance imaging (MEMRI) of mouse brain development. *NMR Biomed*, 17(8):613–619, 2004.

[101] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process*, 13(4):600–612, 2004.

[102] T. Watanabe, J. Frahm, and T. Michaelis. Functional mapping of neural pathways in rodent brain in vivo using manganese-enhanced three-dimensional magnetic resonance imaging. *NMR Biomed*, 17(8):554–568, 2004.

[103] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell*, 31(2):210–227, Feb. 2009.

[104] J. Yang, J. Wright, T. S. Huang, and Y. Ma. Image super-resolution via sparse representation. *IEEE Trans Img Proc*, 19(11):2861–2873, Nov. 2010.

[105] M. Yang, L. Zhang, X. Feng, and D. Zhang. Fisher discrimination dictionary learning for sparse representation. In *Proceedings of the 13th International Conference on Computer Vision*, pages 543–550, Washington, DC, USA, 2011. IEEE Computer Society.

[106] X. Yu, J. Zou, J. S. Babb, G. Johnson, D. H. Sanes, and D. H. Turnbull. Statistical mapping of sound-evoked activity in the mouse auditory midbrain using Mn-enhanced MRI. *NeuroImage*, 39(1):223 – 230, 2008.

[107] C. Yuan, E. Lin, J. Millard, and J. N. Hwang. Closed contour edge detection of blood vessel lumen and outer wall boundaries in black-blood MR images. *Magn Reson Imaging*, 17(2):257–266, Feb. 1999.

[108] R. Zeyde, M. Elad, and M. Protter. On single image scale-up using sparse-representations. In *Proceedings of the 7th International Conference on Curves and Surfaces*, pages 711–730, Berlin, Heidelberg, 2012. Springer-Verlag.

[109] X. Zhang, E. Y. Lam, E. X. Wu, and K. K. Wong. Application of Tikhonov regularization to super-resolution reconstruction of brain MRI images. In *Medical Imaging and Informatics*, pages 51–56, Berlin, Heidelberg, 2008. Springer-Verlag.

[110] A. Zomet, A. Rav-Acha, and S. Peleg. Robust super-resolution. In *Proceedings of the IEEE Workshop on Applications of Computer Vision*, pages 645–650, 2001.

[111] M. Zontak and M. Irani. Internal statistics of a single natural image. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 977–984, 2011.

[112] H. Zou and T. Hastie. Regularization and variable selection via the Elastic Net. *J Roy Stat Soc B*, 67:301–320, 2005.

Samenvatting

Dit proefschrift beschrijft onderzoek naar super-resolutie reconstructie (SRR) en zelf-similariteit, en de toepassing daarvan in magnetische resonantie imaging (MRI). In dit onderzoek is de mogelijke meerwaarde van SRR in MRI bestudeerd, zowel vanuit algemeen oogpunt, als voor een aantal specifieke toepassingen bij de beeldvorming van kleine proefdieren. Ook is een nieuwe SRR methode ontwikkeld, waarin gebruikgemaakt wordt van zelf-similariteit op verschillende schalen, hetgeen van toepassing is op MRI-technieken waarbij een 3D beeld wordt verkregen door het scannen van meerdere 2D plakken (multi-slice MRI). Zelf-similariteit in MRI is verder verkend voor toepassing in voxel-gebaseerde classificatie. Dit onderzoek heeft geresulteerd in een nieuwe classificatiemethode. De bijdragen van de vijf onderzoeken die in dit proefschrift worden beschreven (hoofdstukken 2 tot en met 6) kunnen als volgt worden samengevat.

Hoofdstuk 2 presenteert de resultaten van een uitgebreid experimenteel onderzoek naar de wisselwerking tussen resolutie, signaal-ruisverhouding, en de opnametijd van SRR in multi-slice 2D MRI, en de vergelijking met directe 3D beeldvorming op hoge resolutie. De voornaamste conclusie is dat wanneer de resolutie beperkt is in de richting loodrecht op de 2D MRI plakken, waardoor de visualisatie en detectie van kritieke structuren wordt belemmerd, SRR een simpele strategie biedt om de mogelijkheden van de scanner uit te breiden. Bovendien kan SRR een betere strategie zijn dan directe 3D beeldvorming, wanneer eisen met betrekking tot minimale signaal-ruisverhouding het noodzakelijk maken het gemiddelde van meerdere 3D opnames te berekenen.

De mogelijkheid die SRR biedt om structuren kleiner dan de intrinsieke resolutie van de MRI scanner te visualiseren wordt in **hoofdstuk 3** toegepast voor de bestudering van substructuren in de hersenen van muizen. Het muizenbrein heeft een diameter van ongeveer 15 mm. Onderzoek naar de substructuren of kernen van de hersenen vereist dus een hoge beeldresolutie. In met mangaan verrijkte muizenhersenen worden bij de MRI opnames gewoonlijk 3D gradiÃńnt-echo (GE) beeldvormingssequenties gebruikt vanwege hun vermogen om isotrope resolutie te verkrijgen. Er is gebleken dat het gebruik van inversion recovery spin echo (IR-SE), een 2D beeldvormingssequentie, resulteert in andere gunstige contrasteigenschappen. De anisotropie van de resulterende afbeeldingen beperkt echter in ernstige mate de analyse van hersenkernen. In dit hoofdstuk wordt aangetoond dat met SRR isotrope IR-SE beelden van ongeveer $0,1 \times 0,1 \times 0,1$ mm kunnen worden verkregen,

waardoor een groot aantal kernen kan worden geÃŕdentificeerd. Een nog belangrijkere bevinding is dat een aantal van deze kernen beter zichtbaar is in de SRR IR-SE afbeeldingen dan in de GE afbeeldingen. In dergelijke studies vormt SRR IR-SE dus een beeldvormende methode die complementair is aan de gebruikelijke GE methode.

SRR methoden kunnen veel rekenkracht vereisen en resulteren in grote databestanden. **Hoofdstuk 4** beschrijft een praktische oplossing voor dit probleem. GeÃŕnspireerd door Google Maps introduceren we een interactieve visualisatiemethode waardoor het hele lichaam van een muis onderzocht kan worden met superieure resolutie. De methode resulteert in veel betere visualisatie van kleine tumoren en andere structuren die van potentieel belang zijn. Samen met de resultaten van het voorgaande hoofdstuk toont dit de veelzijdigheid aan van SRR voor verschillende toepassingen. De belangrijkste bijdrage van het in hoofdstuk 4 beschreven onderzoek is de implementatie van een bijna real-time SRR methode.

In de laatste twee onderzoeken is gekeken naar zelf-similariteit en schaarse codering van beeldinformatie door middel van aangeleerde lexicons van beeldstructuren, en de toepassing daarvan in MRI reconstructie en voxel-gebaseerde classificatie. Het onderzoek was voornamelijk gericht op methodologische innovatie en experimentele verificatie van de mogelijkheden van de voorgestelde methoden.

In **hoofdstuk 5** wordt het idee onderzocht om zelf-similariteit tussen de twee intrinsieke schalen van een multi-slice 2D MRI scan te gebruiken voor SRR. De hoogresolute informatie in de afzonderlijke MRI plakken wordt daarbij ingezet als 2D prior voor SRR in de richting loodrecht op de plakken. De voorgestelde methode kan worden toegepast op zowel een enkele scan als op meerdere scans onderling. De gepresenteerde resultaten geven aan dat de methode in staat is om beeldinformatie te benutten die niet wordt gebruikt door een conventionele doch geavanceerde SRR methode gebaseerd op Tikhonov-regularisatie.

Tenslotte wordt in **hoofdstuk 6** gekeken naar de toepassing van zelf-similariteit op voxel-gebaseerde classificatie van MRI beelden van het hoofd en de hals. Er wordt kwantitatief aangetoond dat aangeleerde lexicons van de relevante beeldstructuren redundante informatie bevatten. Deze redundantie wordt vervolgens gebruikt om een bestaande methode voor voxel-gebaseerde classificatie, de zogeheten schaarse-representatie classificatie (SRC), te veralgemeniseren. In SRC wordt elke relevante beeldregio gerepresenteerd door een eigen lexicon. In de classificatiestap wordt voor elk voxel bepaald door welk lexicon een klein gebiedje rondom dat voxel het beste gerepresenteerd wordt. Het voxel krijgt vervolgens het label van het betreffende lexicon. In SRC wordt de representatie doorgaans bepaald door slechts een paar elementen uit het lexicon. De voorgestelde veralgemenisering van SRC bestaat hierin dat dit proces meerdere keren wordt herhaald, met verschillende elementen uit het lexicon, teneinde meer bewijsmateriaal te verzamelen om te kunnen bepalen tot welke beeldregio het voxel behoort. De gepresenteerde experimenten tonen aan dat de veralgemeniseerde methode de classificatie verbetert.

Acknowledgements

While I'm still in the thesis finalization aftermath and still not sure whether to thank or to scorn myself for having embarked on a PhD study, there are a number of people that deserve credit for having assisted me through the past 4-5 years.

First, I want to thank my promoter. Wiro, thanks for letting me join your group. When I was invited for the interview in Rotterdam I was impressed and excited to see such a large group of potential colleagues. I felt an atmosphere that was at the same time informal and pleasant, and to-the-point and demanding. Though I did not know much about the medical side of image analysis and what I was in for by committing to do a PhD, this feeling made me decide in favour of joining BIGR. From your sharp insights at brain storm sessions to your smooth moves on the Xmas party dance floor, you are a great superior and I feel privileged having worked under your supervision.

Erik, my co-promoter, I owe my stay in BIGR as much to you as I do to Wiro. You found me a fit candidate for your project. Thank you for that! As my daily supervisor, you have given me a lot of responsibility and a lot of freedom. Even the freedom to become lost. Or at least to feel like that. But while I indeed felt quite clueless during the first two years of my PhD study, you provided moral support, and your trust and patience during times of stagnation kept me going. At the same time you have maintained a demanding standard regarding the research quality and the presentation of our studies. I understand that this has been an essential part of my training as a researcher.

Dirk, my faithful co-author. It was a strike of good luck for me when you entered BIGR. Coincidently, you possessed substantial knowledge on the topic that I had embarked on. I didn't leave you alone much after I realized this, and I adopted you as a supervisor on anything MRI related. You maintain high standards on intellectual rigidity and clarity of presentation. Thank you for your input and for setting the bar high.

Louise, you are another faithful co-author, and the main reason why my study of super-resolution in MRI proliferated. After we were introduced, you immediately had a great impact in helping me getting my first paper published in MRM. You saw opportunities in the super-resolution idea and we continued collaborating. During my PhD you were one of my main pillars of support and always provided very qualified an inspired supervision in our projects. Thank you for that!

Stefan, your curiosity and broad knowledge of our field made it natural for me

to approach you for discussions, general input and sanity checks on anything from methods to experimental design. We ended up doing an extensive study together, that became a chapter of this thesis. Your supervision through that process was challenging and inspiring. You made me pursue the unexplained details in the data, and squeeze more insight out of my experiments. Together with Erik, Dirk, and Louise you have really shown and taught me how research is performed. Thank you!

Artem, I have really enjoyed collaborating with you. You took charge of collecting and connecting the pieces of our project, and you were always responsive and communicative. Thanks! Thanks also to Boudewijn for the setting up and providing the main idea for the collaboration project.

Dana, thanks for the collaboration. It took a long, scary and strange path, but you made it, and we made it. Good luck finishing your PhD, and best wishes to you and your family.

Monique, Piotr, Gyula, Gavin, Joost, Gerben, Jamal, Gabi, Sandra, Alessandro, Mark, a.k.a. the Molecular Imaging group, thanks for the discussions and the data! And for the great times of bowling, sailing and dancing!

Miki, my week in your lab in 2010 was one of the best and most intense learning experiences of my life. The inspiration and ideas that I left with are directly reflected in the second half of this thesis. Thank you!

Marcel and Renske! Thanks for taking upon you to be my paranymphs! I address you as a single entity but of course I also know you as the two individuals behind. I have a lot to thank you guys for! Great dinners together, long nights at Nieuwe Binnenweg, concerts, movie nights, moral support, linux support. You guys are a major reason why Rotterdam has become my second home. Thank you for that! I hope you stay there, so I can see you whenever I visit.

Professor Krestin. Thank you for having created such a great space to work and do research. And of course a huge thanks for the fantastic PhD dinners, the xmas parties, and all the Radiology borrels, they make a big difference!

Ihor and Oleh, my first office mates. Thank you for receiving me so well! Though $N = 2$, I am tempted to generalize my experience with you guys and declare Ukranians the friendliest and most helpful people in the world! You are both always ready to help out. I admire your unselfishness and the fact that you both stay so productive and successful despite all the time you spend helping other people.

Valerio, Miro, Pierre, thank you for being quiet! Also, thanks for the football games, the music, and the late-night non-sense.

Michiel, your motivation and curiosity was an big inspiration for me while you were in BIGR. We had a few conversations during my first, difficult years, and they always opened up some perspectives for me. Fedde, it was always interesting to talk with you, whether about regression methods or about the evolution of breakbeat based music. Reinhard, another one of those unselfish souls! You always took the time to help out. Also thanks for a great wedding and for our conversations

about religion. Adria, thanks for your generosity in terms of underwear and thanks for disturbing me in the last months of finalizing my PhD. Marleen and Marco, thank you for your thoughts and guidance. Hope to see you in Copenhagen (or where ever)! Karima, thanks for helping out with the Samenvatting! My apologies for the obscure language of me and my community.

Maart, Marius, Hakim, Henri, thanks for running our cluster and helping out whenever needed! You guys are invaluable. I like you as people as well ;-)

Petra, Desiree, Tineke, Carmen, Fania, Ton, and Linda. I do not dare imagine life as a PhD student without you and the infrastructure that you provide. From my first to my last day in BIGR, I have felt very well taken care of. Thanks you!

Vera, Preci, Sandra, Inge, Leo, Marjolein, Gwen, Aafke, Ingrid, Carmen, Katja, Miriam and all the other IPCI ladies (Daniel, Nico, Gianluca, Seppe, Rene, Osemeke included). Thank you! My life instantly changed to the better when I arrived at 21st floor of the old tower and went from 5 years of studying with geeky guys to having you people around me on a daily basis. Rebecca, Carolina, Rozanna, Anouk, Daniel, and the rest of the Radiology Dept. death metal crew: Cheers! Miss you guys.

LastR but not leastR: Theo, Jifke, Nora, Guillume, Francesco, Andres, Jyotirmoy, Karin, Diego, Arna, Gerardo, Annegreet, Esther, Wyke, Luu, Carolyn, Jean-Marie, Coert, Danijela, Azadeh, Eugene, Emilie, Henk, Ruben, TT, Wei, Gijs, Erwin, Hortense, Rahil and all the rest of (ex-)BIGR, thank you all for making BIGR such a great place!

PhD portfolio

Attended courses

Advanced in vivo NMR (ImagO) 2009
Advanced pattern recognition (ASCI) 2009
Knowledge-driven image segmentation (ASCI) 2009
Front-end vision and multi-scale image analysis (ASCI) 2009
International summer school on molecular and medical image analysis and bioinformatics, Lipary, Italy 2009
Summer school on sparsity in image and signal analysis, Iceland (DIKU) 2010
Biomedical English writing and communication (EMC) 2011
Oral presentation (MI) 2012

Given courses

Introduction to image analysis for medical students 2009-2011

Research visits

Visiting the group of Prof. Michael Elad at Technion, Haifa, Israel 2010

Conference presentations

EMIM - European Molecular Imaging Meeting (poster) 2010
EMIM - European Molecular Imaging Meeting (oral) 2011
SPIE Medical Imaging (poster) 2012
TOPIM - Hot Topics in Molecular Imaging (oral) 2012
TOPIM - Hot Topics in Molecular Imaging (poster) 2012
ISBI - IEEE International Symposium on Biomedical Imaging (poster) 2012
ISMRM - International Society for Magnetic Resonance in Medicine (poster) . 2013
ISMRM Benelux (oral) 2013
MICCAI - Medical Image Computing and Computer-Assisted Intervention (poster) 2013

Other presentations

Dutch medical informatics days, Renesse (oral presentation) 2010
Erasmus MC MI research colloquia (2 oral presentations) 2009-2013
Erasmus MC BIGR seminars (2 oral presentations) 2009-2013
Erasmus MC Radiology research colloquia (oral presentation) 2009-2013

Other conferences and workshops attended
AMIE workshop on animal imaging 2009
ENCITE workshop on molecular imaging 2009
ISBI - IEEE International Symposium on Biomedical Imaging 2010
MICCAI workshop on sparsity techniques in medical imaging 2012
ENCITE workshop on cell imaging and tracking 2012

Awards
SPIE Medical Imaging (poster). Honorable mention 2012
ISMRM Benelux (oral). Best presentation 2013

Travel grants
Erasmus Trustfonds travel grant for research visit to Technion, Haifa, Israel .. 2010

Reviewing activities
IEEE Transactions on Image Processing, IEEE Transactions on Medical Imaging, IEEE Transactions on Biomedical Engineering, Magnetic Resonance in Medicine 2011-2013

Volunteer activities
Board member of Erasmus MC PhD student union, Promeras 2010-2012
Helped organizing Erasmus MC PhD days 2010
Organized the monthly Erasmus MC radiology research drinks 2009-2013

About the author

Esben Plenge studied a Bachelor of Electronic Engineering at Copenhagen University College of Engineering from 2003-2006. During these years he spent one semester as an exchange student at Hong Kong Polytechnic University, and two semesters at the School of Software Engineering of Tongji University in Shanghai. His Bachelor thesis considered efficient implementation of the JPEG2000 image compression algorithm on an embedded platform. From 2006-2008 Esben was a student of the graduate programme in Computer Vision and Graphics at Aalborg University Copenhagen. He obtained the MSc.Eng. degree after writing his thesis on 3D facial motion capture. In 2009 Esben joined the Biomedical Imaging Group Rotterdam (BIGR) and started a Ph.D. study. His research focused on MRI resolution enhancement, in particular it's applications and benefits in small animal imaging. After a visit to the group of Prof. Michael Elad at the Technion in Haifa, Israel, in 2010, he started investigating how the self-similarity of MR images could be exploited for image enhancement and analysis. Since October 2013 Esben has been employed as a post-doctorate fellow in the group of Prof. Michael Elad at the Technion's department of Computer Science.