Η παρούσα εργασία εκπονήθηκε για την απόκτηση του

Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Βιοϊατρική Τεχνολογία

από την

ΚΑΡΑΒΕΛΑΚΗ ΕΥΘΥΜΙΑ

Πανεπιστήμιο Πατρών, Τμήμα Ιατρικής
Εθνικό Μετσόβιο Πολυτεχνείο, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανολόγων Μηχανικών
Εθνικό Μετσόβιο Πολυτεχνείο, Τμήμα Μηχανολόγων Μηχανικών

Δεκέμβριος 2007
IMPROVEMENT OF PET RESOLUTION WITH SUPER RESOLUTION TECHNIQUES

A Thesis
Submitted for the degree of

MSc in Biomedical Engineering

By
KARAVELAKI EFTHIMIA

University of Patras, Faculty of Medicine
National Technical University of Athens, Faculty of Electrical & Computer Engineering
National Technical University of Athens, Faculty of Mechanical Engineering

December 2007
Acknowledgements

I would like to take this opportunity to express my deep sense of gratitude to Prof. George Kontaxakis and Prof. Andres Santos for their valuable support during the implementation of this thesis. Many thanks to all the people of the Biomedical Image Technology Lab of the Dept. of Electronic Engineering at Universidad Politecnica de Madrid for their friendship and endorsement.
Abstract

Medical imaging is the main tool to extract a 3D modelling of the human body or specific organs within it. In order to accomplish this, various imaging modalities have been developed over the years, such as X-Ray Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). Each one is based on a particular energy source that passes through the body and on specific physical laws, which define the meaning of noise and the sensitivity of the imaging process. In all medical imaging systems the main goal is to increase resolution since higher resolution is a key factor in increased information content, which is critical for increased accuracy in the understanding of the anatomy and in the assessment of size and morphological structure of organs, for early detection of abnormalities, suspected pathologies and more.

In order to overcome the resolution limitations, one promising idea is to use signal processing techniques to enhance the spatial resolution. This approach proposes the acquisition of a high-resolution (HR) image from observed multiple low-resolution (LR) images. This image restoration approach is called super resolution (SR) image reconstruction (or restoration). It is the process of combining multiple low resolution images to form a high resolution image. The basic requirement in order to apply SR restoration techniques is the availability of multiple LR images captured from the same scene, which are sub-sampled (aliased) as well as shifted with subpixel precision. Each observed LR image is expressed as the result of a sequence of operators on the original HR image source, consisting of a geometrical warp, blurring and down-sampling.

The SR image reconstruction method consists of three stages, registration, interpolation and restoration (i.e., inverse procedure). In the registration stage, the relative shifts between LR images, with reference to a certain LR image, are estimated with fractional pixel accuracy. Accurate sub-pixel motion estimation is a very important factor in the success of the SR image reconstruction algorithm. Since the shifts between LR images are arbitrary, the registered HR image will not always match up to a uniformly spaced HR grid. Thus, non-uniform interpolation is necessary, to obtain a uniformly spaced HR image from a non-uniformly spaced composite of LR images. Finally, image restoration is applied to the up-sampled image to remove blurring and noise.

In order to evaluate the performance of SR reconstruction, a 'simulate and correct' approach to reconstruction is selected. First, simulated images of a computer generated phantom are formed and processed in order to comply with the observation model for the LR images. These are used as the images from which the HR image is constructed through the SR method. The iterative back-projection (IBP) algorithm suggested by Irani and Peleg has been chosen to be utilized, which belongs in the spatial domain methods and it is an easily and intuitively understood method. The results of the SR reconstruction are presented separately for the axial and the transaxial case. The evaluation relies on qualitative measures of image enhancement and on objective quantitative measures, such as the resolution (FWHM), the signal-to-noise ratio, the contrast ratio and the contrast-to-noise ratio.

The performed trials demonstrated improvement in both the axial and transaxial resolution. The super-resolution images also provide a significantly improved contrast ratio, which is important for improving sensitivity for detection of small details and features. The improvement in resolution can be achieved without using any hardware changes or any increase in the patient radiation procedure. An important contribution of super-resolution is also the reduction of partial volume effects in the reconstructed image. The loss in SNR, which is a typical characteristic of all resolution enhancement algorithms, was not that considerable to preclude the clinical
application of super-resolution. The overall evaluation demonstrated that the SR reconstruction is a post-processing method, which can provide medical images of higher resolution and better contrast ratio, without increasing the amount of radiation or the duration of the scan.

In the first chapter, the main medical imaging modalities and their applications are outlined. In the second chapter, resolution limitations of the current medical imaging systems are discussed, as well as the physical constraints that determine the resolution and the gain from the increase in resolution for two main modalities, MRI and PET. In the third chapter, an overview of the super resolution techniques is presented, including the main observation models and a short description of the reconstruction algorithms proposed in the literature. In the fourth chapter, the followed experimental procedure is explained and the results are presented and discussed in the fifth chapter. Finally, the conclusions are summarized in the last chapter, along with some directions for future work. In the Appendix, the help lines of the matlab functions which implement the reconstruction method are demonstrated.
List of Figures

Figure 1-1: Medical imaging pipeline example.......................................................... 13
Figure 1-2: (a) MRI scanner; (b) Anatomical image of the human head................. 15
Figure 1-3: MR example images; (a) Angiography; (b) DWI; (c) DTI; (d) fMRI
sequence of images ................................................................................................ 16
Figure 1-4: Basic principle of CT image acquisition .................................................. 19
Figure 1-5: Spiral CT basic concept.......................................................................... 20
Figure 1-6: Coincidence detection in a PET camera.................................................. 22
Figure 1-7 : Types of coincidences in PET ............................................................... 23
Figure 1-8: A block detector consisting of a 7x8 array of crystals coupled to 4 PMT 24
Figure 1-9: Sampling pattern in the transaxial plane for a PET scanner............... 24
Figure 1-10: (a) PET/CT operation schema; (b) PET/CT scanner ......................... 28
Figure 1-11: Example body scans; (a) CT image; (b) PET image; (c) CT and PET
images fused together........................................................................................... 29
Figure 2-1: MRI slice acquisition............................................................................... 31
Figure 2-2: Events affecting the counting statistics................................................... 33
Figure 3-1: Basic requirement for super resolution................................................... 36
Figure 3-2: Observation model relating LR images to HR images............................ 38
Figure 3-3: Example for global translation ............................................................... 39
Figure 3-4: SR reconstruction method ...................................................................... 40
Figure 3-5: Interpolation based reconstruction.......................................................... 42
Figure 3-6: Iterative Back-projection method ............................................................ 43
Figure 4-1: (a) Anterior view of the 4D NCAT phantom. (b) Cardiac and respiratory
motion models. (c) Simulations performed using the phantom. [Ref. NCAT]........ 48
Figure 4-2: (a) Transaxial image. (b) Coronal image. (c) Sagittal image............... 49
Figure 4-3: Sets of images - axial and transaxial – generated in the phase of image
simulation. ............................................................................................................ 50
Figure 4-4: Block diagram of the iterative back-projection method ......................... 52
Figure 4-5: Scheme for interleaving method .............................................................. 55
Figure 5-1: Coronal and sagittal images corresponding to each set of acquisitions . 58
Figure 5-2: Error (difference between input and simulated LR images) for
reconstruction from 2 and 4 acquisitions in the axial case................................. 58
Figure 5-3: (a) Initial coronal image (32 slices). (b) Reconstructed image from 2
acquisitions interleaved. (c) Reconstructed image from 2 acquisitions with super
resolution.............................................................................................................. 59
Figure 5-4: Plot of the pixel values along a column of a coronal image. (a) Selected column of the coronal image. Super resolution result from 2 acquisitions (solid) compared to (b) image without super resolution (dashed) and (c) reconstruction with 2 acquisitions interleaved (dashed). .......................................................... 60

Figure 5-5: (a) Coronal image (16 slices). (b) Reconstructed image from 4 acquisitions interleaved. (c) Reconstructed image from 4 acquisitions with super resolution.......................................................... 61

Figure 5-6: Plot of the pixel values along a column of a coronal image. (a) Selected column of the coronal image. Super resolution result from 4 acquisitions (solid) compared to (b) image without super resolution (dashed) and (c) reconstruction with 4 acquisitions interleaved (dashed). .......................................................... 61

Figure 5-7: (a) Input transaxial image. (b) Rotated transaxial image 4.0856°. (c) Rotated transaxial image 3.8141°. (d) Rotated transaxial image 2.9737° ........ 62

Figure 5-8: Error (difference between input and simulated LR images) for SR reconstruction from (a) Input set 1 and (b) Input set 2.............................................. 62

Figure 5-9: (a) Input transaxial image. (b) SR reconstructed image. ...................... 64

Figure 5-10: Plot of pixel values along a column of the HR image in comparison to the input. (a) Selected column. (b) Super resolution result (solid) compared to image without super resolution (dashed). ......................................................... 64

Figure 5-11: (a) Input transaxial image. (b) SR reconstructed image. ...................... 64

Figure 5-12: Plot of pixel values along a row of the HR image in comparison to the input. (a) Selected column. (b) Super resolution result (solid) compared to image without super resolution (dashed). ......................................................... 65

Figure 5-13: Power spectrum of the (a) input image and (b) SR reconstructed image .......................................................................................................................... 65
List of Tables

Table 5-1: Point spread function FWHM (mm) values (calculated for a Gaussian PSF along the axial direction) ................................................................. 59
Table 5-2: Acquisition time, SNR, contrast ratio and CNR of the SR reconstructed images from 2 acquisitions ................................................................. 59
Table 5-3: Acquisition time, SNR, contrast ratio and CNR of the SR reconstructed images from 4 acquisitions ................................................................. 61
Table 5-4: Point spread function FWHM (mm) values (calculated for a Gaussian PSF) ........................................................................................................ 63
Table 5-5: Acquisition time, SNR, contrast ratio and CNR of the SR reconstructed images for each set of input images .................................................... 63
# Table of Contents

Acknowledgements ........................................................................................................ 4

Abstract ........................................................................................................................ 5

List of Figures ................................................................................................................ 7

List of Tables .................................................................................................................. 9

Table of Contents ......................................................................................................... 10

## Chapter 1  Introduction to Medical Imaging ......................................................... 13

1.1. Introduction ............................................................................................................ 13

1.2. Magnetic Resonance Imaging (MRI) ................................................................. 14
1.2.1 Basic Physics of MRI ......................................................................................... 14
1.2.2 MRI Modalities and Applications .................................................................... 16
1.2.3 Flow Imaging (MR Angiography) .................................................................... 17
1.2.4 Diffusion-Weighted Imaging (DWI) ................................................................. 17
1.2.5 Diffusion Tensor Imaging (DTI) ...................................................................... 17
1.2.6 Functional Magnetic Resonance Imaging (fMRI) ............................................ 18

1.3. X-ray Computed Tomography (CT) ................................................................. 18
1.3.1 Principles of x-ray CT ...................................................................................... 18
1.3.2 Multi-slice Computed Tomography ................................................................. 19

1.4. Nuclear Medicine Imaging .................................................................................. 20
1.4.1 PET and SPECT ............................................................................................... 20
1.4.2 Positron Emission Tomography (PET) .............................................................. 20
1.4.3 Basic Physics of PET ....................................................................................... 21
1.4.4 Coincidence Detection and Electronic Collimation .......................................... 22
1.4.5 Types of Coincidence Events ......................................................................... 22
1.4.5.1 Detectors ...................................................................................................... 23
1.4.6 Acquisition Modes .......................................................................................... 24
1.4.7 Resolution ........................................................................................................ 24
1.4.8 Image Reconstruction ...................................................................................... 25
1.4.9 Applications of PET ........................................................................................ 26
1.4.9.1 PET in Cardiology ...................................................................................... 26
1.4.9.2 PET in Neurology ....................................................................................... 26
1.4.9.3 PET in Oncology ........................................................................................ 26
1.4.10 PET/CT ........................................................................................................... 27

## Chapter 2  Medical Imaging Resolution .............................................................. 30

2.1 Introduction ............................................................................................................. 30

2.2 General Constraints ............................................................................................ 30
2.2.1 Partial Volume Effect (PVE) .......................................................................... 30

2.3 MRI Resolution ................................................................................................... 31
2.3.1 Spatial Resolution .......................................................................................... 31
6.1 Introduction .................................................................................................................. 67
6.2 Application of super-resolution .................................................................................... 67
6.3 Super-resolution performance ..................................................................................... 69
6.4 Additional issues and future challenges ....................................................................... 69
  6.4.1 Spiral CT ............................................................................................................. 69
  6.4.2 Newly emerging hardware ................................................................................... 70
  6.4.3 Registration issues .............................................................................................. 70
  6.4.4 Future work ......................................................................................................... 70

Appendix A ........................................................................................................................ 71

References ......................................................................................................................... 74
Chapter 1  
Introduction to Medical Imaging

1.1. Introduction

Medical imaging is the main tool to extract a 3D modelling of the human body or specific organs within it. In order to accomplish this, various imaging modalities have been developed over the years. Each one is based on a particular energy source that passes through the body. There are many imaging modalities, others providing anatomical information and information about the structure, others providing functional information, e.g. reveal locations of activity within the brain for specific activities and specified tasks.

The medical imaging field is rapidly evolving with the use of increased resolution machines and advanced content-processing tools. The main goal of medical imaging system developers is increased resolution since higher resolution is a key factor in increased information content. Increased information content is critical for increased accuracy in the understanding of the anatomy and in the assessment of size and morphological structure of organs, for early detection of abnormalities, suspected pathologies and more.

Figure 1-1 illustrates a schematic example of the stages involved in medical imaging acquisition and processing. During the acquisition process, there is a transmission or emission process (X-ray transmission and attenuation in CT or RF encoding in MRI or photon emission in PET), through which energy is acquired as it transverses the body (or organ). The energy is captured by an array of detectors which are specific for each imaging modality (RF coils in MRI, crystal detectors in PET, X-ray detectors in CT).

![Figure 1-1: Medical imaging pipeline example](image-url)
The raw data that result from the acquisition process are pre-processed with the use of some a priori knowledge if it is available. This stage includes processes such as noise suppression, contrast enhancement, intensity equalization, outlier elimination, bias compensation, time/space filtering etc. The resulting enhanced data is reconstructed in order to obtain physically meaningful images, volumes and sequences. The followed reconstruction method depends on the physical properties of the device and the principles of the acquisition process. The interpretation stage involves some high-level processing, which includes registration, segmentation, classification, tracking etc, in order to obtain quantitative and qualitative information such as anatomical, functional and metabolic information which have a significant contribution in the decision making process.

In the following of this chapter we present shortly some of the major imaging modalities, the underlying features that make each modality unique along with its clinical relevance.

1.2. Magnetic Resonance Imaging (MRI)

1.2.1 Basic Physics of MRI

Magnetic resonance imaging (MRI) is an imaging technique used primarily in medical settings to produce high quality images of the human body. MRI is based on the absorption and emission of energy in the radio frequency range of the electromagnetic spectrum, producing images based on spatial variations in the frequency of the radio frequency energy being absorbed and emitted by the imaged object.

A basic MRI system comprises a powerful magnetic field created by a resistive or a superconducting magnet. Field strength in these systems varies between 0.2 and 4 Tesla (T) with common commercial systems utilizing field strengths of 1.0 or 1.5T. There is a series of gradient coils within the magnet, which are integral to sequence design, and necessary for the spatial localization of the signal within the patient. Different radio-frequency receiver coils are available depending on which area of the body is being imaged, e.g. shoulder, body or head coil. The body coil is also often used to transmit radio-frequency signals when a peripheral radio-frequency coil (e.g. a shoulder coil) is being used. Advanced computer systems with high-grade hardware and software specifications are necessary. In Figure 1-2 an example MRI scanner is illustrated as well as an example MR image of the human brain.

MRI utilizes principles of nuclear magnetic resonance (NMR), a spectroscopic technique used to obtain microscopic chemical and physical information about molecules. The human body is primarily fat and water. Fat and water have many hydrogen atoms which make the human body approximately 63% hydrogen atoms. Hydrogen nucleus has an NMR signal. For these reasons magnetic resonance imaging primarily images the NMR signal from the hydrogen nuclei. Each water molecule has one oxygen and two hydrogen atoms. Each hydrogen nucleus consists of a single proton which possesses a property called spin. Spin can be thought of as a small magnetic field and will cause the nucleus to produce an NMR signal.

When a patient is placed within a large magnetic field the protons partially align themselves along the main magnetic field, thus producing a net magnetic moment. These protons can be displaced from this north-south alignment with the application of a radio-frequency pulse at a specific frequency. This frequency is defined as the Larmour frequency and is dependent on the strength of the magnetic field. Once
these protons are displaced in the main magnetic field they relax, to align back along the main magnetic field. It is during this process of relaxation that the spinning protons produce a radio-frequency electromagnetic signal, which can be detected by the receiver coil and amplified. This signal is called free induction decay (FID) signal. The magnitude of the signal depends on several factors, including the hydrogen proton concentration within the tissues being imaged.

The relaxation rate is dependant on the interaction of molecules within their local environment and with adjacent molecules, the field strength, and the timing and magnitude of the radio-frequency pulses. Different tissues have different relaxation rates. It is this difference in relaxation between tissues, which enables an image to be created. By using a magnetic field which varies spatially (e.g. changes linearly in all 3 directions), spins in different locations will yield FID signals with different frequencies. Spatial information can be extracted through frequency encoding. In 3 dimensions, a plane can be defined by the process of slice selection in order to reduce spatial encoding to 2 dimensions. In this process, a radio-frequency (RF) pulse of defined bandwidth is applied in the presence of a magnetic field gradient. Spatial encoding can then be applied in 2D after slice selection, or in 3D without slice selection. In either case, a 2D or 3D matrix of spatially-encoded phases is acquired, and these data represent the spatial frequencies of the imaged object. In order to create images from the acquired data, the Discrete Fourier Transform (DFT) is utilized. In conventional 2D multi-slice MR imaging, the RF pulse first selects the 2D slice to be analyzed. A combination of frequency encoding of the signal in one direction and phase encoding in the other direction enables to encode the 2D spatial information within the slice (in-plane). Multiple RF excitations may be required for a single slice to be fully encoded.

The MR image intensities are proportional to the number of nuclei in each voxel. The image contrast is determined by several time constants, each of which represents a certain relaxation process that establishes equilibrium after the RF excitation. During the relaxation and realignment of the high-energy nuclei, the emitted at different rates energy is recorded to provide information about their environment. This realignment of nuclear spins with the magnetic field is termed longitudinal relaxation and the time (typically about 1 sec for tissue water) required for a certain percentage of the tissue nuclei to realign is termed T1 (‘spin-lattice’ relaxation). T2-weighted imaging is based on the local dephasing of spins in the transverse plane. The transverse relaxation time (typically < 100 ms for tissue water) is termed T2 (‘spin-spin’ relaxation). In T2 imaging, a spin echo technique is employed, in which spins are refocused to compensate for local magnetic field inhomogeneities. An important variation of this technique without refocusing is employed in T2* imaging.

Image contrast is created by selecting the image acquisition parameters that weight the signal. The basic types of weighting used in clinical practice are T1, T2,
T2* and proton density weighting (no relaxation time). T1 weighting is good for anatomical detail. T2-weighted images give high signal for tissues with a higher content of free unbound water and are essential in imaging inflammation or neoplastic tissue. Tissues with high free water molecules take longer to relax and, therefore, have high signals on T2 weighting. Proton density weighting depends on the hydrogen proton concentration of tissues and has minimal T1 and T2 contrast-based effects.

Among the reasons for the success of MRI as a diagnostic imaging tool are:

- good spatial resolution
- excellent contrast for all tissue types
- complete non-invasiveness and
- very low risk involved

Disadvantages of the modality are:

- high costs
- long acquisition time
- bulkiness of the equipment
- patient discomfort and
- problems associated with the presence of high magnetic fields

1.2.2 MRI Modalities and Applications

There are several important variants of the MRI process, each utilized for a specific medical application. In the following figure are illustrated some example images of the MRI modalities described below.

![MR example images](image1)

(a) Angiography; (b) DWI; (c) DTI; (d) fMRI sequence of images

**Figure 1-3:** MR example images; (a) Angiography; (b) DWI; (c) DTI; (d) fMRI sequence of images
1.2.3 Flow Imaging (MR Angiography)

Flow imaging is the imaging of flowing blood in the arteries and veins of the body. The intensity in magnetic resonance angiography (MRA) is proportional to the velocity of the flow. Contrast enhanced angiography is based on the difference in the T1 relaxation time of blood and the surrounding tissue when a paramagnetic contrast agent is injected into the blood. This agent reduces the T1 relaxation times of the fluid in the blood vessels relative to surrounding tissues. When the data is collected with a short repetition time (TR) value, the signal from the tissues surrounding the blood vessels is very small due to its long T1 and the short TR. Images of a region of interest are recorded with rapid volume imaging sequences.

1.2.4 Diffusion-Weighted Imaging (DWI)

This modality produces in vivo MR images of biological tissues weighted with the local characteristics of water diffusion. DWI uses very fast scans with an additional series of gradients (diffusion gradients) rapidly turned on and off. Within the brain, water diffuses randomly producing protons that lose phase coherence and thus signal during the application of the diffusion gradients. More precisely, given a spatial direction and a chosen amount of time during which water diffusion takes place, a sophisticated MRI scanner produces a T2 (3D) image attenuated according to the magnitude of the diffusion. The more attenuated the image is at a given position, the more diffusion there is locally. The image intensity varies according to the changes in the spatial direction or the diffusion gradient. Diffusion-weighted images are very useful in the diagnosis of vascular strokes in the brain, in the study of diseases of the white matter or in the effort to infer the connectivity of the brain.

1.2.5 Diffusion Tensor Imaging (DTI)

Diffusion tensor imaging is a variation of DWI in which at least seven images are acquired for every slice, with at least six directions of diffusion weighting. DTI provides a unique tool for visualization of the direction and intactness of white matter fiber tracts in vivo by identifying the preferred direction of diffusion. Axons in the brain are structured in parallel bundles and generally have a myelin sheath that facilitates the diffusion of water molecules along their main direction. Diffusion which is preferentially oriented in one direction is called ‘anisotropic diffusion’. If diffusion gradients are applied (i.e. magnetic field variations in the MRI magnet) in at least 6 directions, it is possible to calculate a tensor (3x3 matrix), for each voxel, that describes this diffusion anisotropy. The fiber direction is indicated by the tensor’s main eigenvector. This vector can be colour-coded, having as a result a cartography of the tracts’ position and direction (red for left-right, blue for superior-inferior, and green for anterior-posterior). The brightness is weighted by the tracts’ anisotropy. High-resolution DTI, combined with algorithms for tracing fibers in 3 directions in tensor fields, has the potential to enable fiber tract mapping of critical functional pathways in the brain. The clinical applications are the tract-specific localization of white matter lesions, the localization of tumours in relation to the white matter tracts and the localization of the main white matter tracts for neurosurgical planning.
1.2.6 Functional Magnetic Resonance Imaging (fMRI)

This method makes it possible to obtain functional information by using haemoglobin as a paramagnetic tracer. fMRI is the use of MRI to measure the haemodynamic response related to neural activity in the brain or spinal cord of humans or other animals. During brain activity there is a rapid momentary increase in the blood flow to the specific thought center in the brain. For example, when moving a finger there is a rapid momentary increase in the circulation of the specific part of the brain controlling the finger movement. The increase in circulation means a decrease in deoxy-haemoglobin which is paramagnetic and affects mainly the T2* of the local brain tissue. The difference in T2* relative to surrounding tissue causes a contrast between the tissues, referred to as Blood Oxygenation Level Dependent (BOLD) contrast. Decreases in the concentration of deoxygenated haemoglobin cause higher BOLD signal intensities since the blood magnetic susceptibility matches more closely the tissue magnetic susceptibility. By using an MRI scanner with parameters sensitive to changes in magnetic susceptibility, changes in BOLD contrast can be assessed. BOLD effects are measured using rapid volumetric acquisition of images (mostly with T2*-weighted acquisition). Such images can be acquired with temporal resolution of 1-4 sec. The voxels in the resulting image typically represent cubes of tissue with size about 2-4 mm on each side in humans.

From fMRI data is produced a time series of samples for each voxel in the scanned volume. There is a variety of methods to correlate these voxel time series with an assigned task in order to produce maps of task-dependent activation. The Statistical Parametric Mapping (SPM) package is a statistical technique to examine the differences in brain activity recorded during functional neuroimaging experiments. These differences are often shown as patches of colour on an MRI brain slice, with the colours representing the location of voxels that have shown statistically significant differences between conditions. The gradient of colour is mapped to statistical values enabling to delineate the relative statistical strength of a given area of activation.

1.3. X-ray Computed Tomography (CT)

1.3.1 Principles of x-ray CT

X-rays were discovered by W.C. Röntgen in 1895 bringing a new prospect for medical diagnosis. They penetrate most biological tissues with little attenuation, and thus provide a comparatively simple means to produce shadow, or projection, images of the human body. The radiographic image represents the distribution of x-ray photons transmitted through the patient. Hence it is a 2D projection of the attenuating properties of tissues along the path of the detected x-rays. The principal interactions causing attenuation are photoelectric absorption and (inelastic) scattering.

However, conventional radiography provides no depth information, as the 3D body structure is projected onto a 2D image. Another limitation is the low soft tissue contrast, which is particularly important in brain imaging, where the soft tissues are enclosed by the highly attenuating skull. In contrast, x-ray computed tomography (CT) imaging produces thin 2D sections of the body and provides good spatial resolution and good discrimination between tissues (better than 1% attenuation change).
In 1972 G.N. Hounsfield first presented a clinical CT scanner at the Annual Congress of the British Institute of Radiology. Since then, the introduction of clinical x-ray computed tomography has revolutionized medical imaging and may be described as the greatest advancement in radiology since the discovery of x-rays. First generation CT systems employed a narrow pencil beam from a collimated source that scans linearly across the patient in order to obtain a parallel projection. The system is then rotated to obtain several such projections.

Since its development in the early 1970s, several generations have been developed and marketed. Current fourth generation CT scanners have a much faster scan speed of about 4-5 sec, enabling several images to be acquired simultaneously. This allows the acquisition of scans with very little motion blur, since the patient can be asked to stop breathing during this period to prevent the presence of image artifacts. A rotating fan beam x-ray source with a continuous ring of about 1000 detectors makes this possible. There are now fifth generation systems available with scan times of only a few milliseconds, thus allowing effective ‘real-time’ imaging.

Tomographic slice images representing attenuation values are reconstructed by inverting the measured projection data. The method most commonly used is called filtered back-projection and employs the following steps:

1. Record projections at different angles around the object to be imaged.
2. Convolve each projection with a filter function (which prevents the occurrence of the ‘star artifact’ in conventional back-projection).
3. Finally, reconstruct the image by back-projecting the filtered projections along their original line-of-sight and summing up the attenuation values.

A CT scanner and the basic concept of CT image acquisition are illustrated in Figure 1-4.

1.3.2 Multi-slice Computed Tomography

Throughout recent years a transition has been made, from slice-by-slice imaging to volume imaging, with the introduction of spiral scan modes. A second transition was made from single-slice to multi-slice scanning. In conventional CT, the slices are acquired sequentially. The x-ray tube rotates around the subject and multiple circular acquisitions are extracted until a slice is acquired. Then, the bed (on which the subject lies) is moved forward incrementally. The size of the step determines the resolution in the longitudinal axis. The spiral (helical) x-ray computed tomography refers to the modern CT scanning technique, in which the rotational movement of the x-ray source is combined with the simultaneous longitudinal movement of the patient’s bed, creating a helical movement of the source around the patient.
The concept of spiral CT is shown in Figure 1-5. Spiral CT advanced conventional CT in speeding up the scanning process as well as in converting it to a 3D modality. With spiral CT, the cross-sectional slices can be reconstructed at any z position, while in the planar CT they were limited to those positions where the circular scans had been performed.

The advantages of multi-slice CT are important to many applications of CT scanning, including survey exams in oncologic or trauma patients and the characterization of focal lung and liver through the creation of thin sections retrospectively. However, the greatest impact has been on CT angiography, cardiac imaging, virtual endoscopy, and high resolution imaging.

1.4. Nuclear Medicine Imaging

1.4.1 PET and SPECT

As the names indicate, positron emission tomography (PET) and single photon emission computed tomography (SPECT) both employ an emitter. The emitter is a molecule tagged with a radio-isotope, administered intravenously or by inhalation. PET and SPECT both also use a detector, a sophisticated camera which can identify at a distance how the emitter has distributed.

While sharing many features, PET and SPECT differ in that SPECT emitters produce single photons, while PET tracers emit a positron, the latter initiating atomic events which make PET a higher resolution modality. In the following sections, the focus will primarily be on PET, which is the most sensitive method for quantitative measurement of physiologic processes in vivo.

1.4.2 Positron Emission Tomography (PET)

Positron Emission Tomography (PET) has enhanced our understanding of the biochemical basis of normal and abnormal functions within the body, and permitted biochemical examination of patients as part of their clinical care. These capabilities are important because:

- The basis of all tissue function is chemical.
- Diseases result from errors introduced into its chemical systems by viruses, bacteria, genetic abnormalities, drugs, environmental factors, aging and behaviour.
- The most selective, specific and appropriate therapy is one chosen from a diagnostic measure of the basic chemical abnormality.
- Detection of chemical abnormalities provides the earliest identification of disease, even in the pre-symptomatic stages before the disease process has exhausted the chemical reserves or overridden the compensatory mechanisms of the brain.
- Assessment of restoration of chemical function provides an objective means for determining the efficacy of therapeutic interventions in the individual patient.
• The best way to judge whether tissue is normal is by determining its biochemical function.

Another principle relates to the value of examining these biochemical processes with an imaging technology. Because in most cases the location and extent of a disease is unknown, the first objective is an efficient means of searching throughout the body to determine its location. Imaging is an extremely efficient process for accomplishing this aim, because data are presented in pictorial form to the most efficient human sensory system—the visual system—for search, identification and interpretation. Recognition depends upon the type of information in the image, both in terms of interpreting what it means and how sensitive it is to identifying the presence of disease.

PET provides the means for imaging the rates of biologic processes in vivo. Imaging is accomplished through the integration of two technologies, the tracer kinetic assay method and Computed Tomography (CT). The tracer kinetic assay method employs a radiolabeled biologically active compound (tracer) and a mathematical model that describes the kinetics of the tracer as it participates in a biological process. The model permits the calculation of the rate of the process. The tissue tracer concentration measurement required by the tracer kinetic model is provided by the PET scanner, with the final result being a three-dimensional (3-D) image of the anatomic distribution of the biological process under study.

Radiolabeled tracers and the tracer kinetic method are employed throughout the biological sciences to measure such processes as blood flow, membrane transport, metabolism, synthesis and ligand-receptor interactions, for mapping axonal projection fields through anterograde and retrograde diffusion, measurement of cell birth dates, marker assays using recombinant DNA techniques, radioimmunoassays, and the study of drug interactions with chemical systems of the body. The tracer technique continues to be one of the most sensitive and widely used methodologies for performing assays of biological systems. PET allows the transfer of the tracer assay methodology to the living subject, particularly humans. PET builds a bridge of communication and investigation between the basic and clinical sciences, based upon a commonality of methods used and problems studied.

The transfer of tracer methods from the basic biological sciences to humans with PET is made possible by the unique nature of the radioisotopes used in PET to label compounds: 11C, 13N, 15O, and 18F. These are the only radioactive forms of the natural elements (18F is used as a substitute for hydrogen) that emit radiation that will pass through the body for external detection. Natural substrates, substrate analogs, and drugs can be labelled with these radio isotopes without altering their chemical or biological properties. This allows the methods, knowledge and interpretation of results from tracer kinetic assays used in the basic biological sciences to be applied to humans by the quantitative measurement abilities of the PET scanner.

1.4.3 Basic Physics of PET

After injection of a tracer compound labelled with a positron emitting radionuclide the subject of a PET study is placed within the field of view (FOV) of a number of detectors capable of registering incident gamma rays. The radionuclide in the radiotracer decays and the resulting positrons subsequently annihilate on contact with electrons after travelling a short distance (~1 mm) within the body. Each annihilation produces two 511 keV photons travelling in opposite directions and these photons may be detected by the detectors surrounding the subject. The detector electronics are linked so that two detection events unambiguously occurring within a certain time window may be called coincident and thus be determined to have come from the same annihilation. These "coincident events" can be stored in arrays corresponding to projections
through the patient and reconstructed using standard tomographic techniques. The resulting images show the tracer distribution throughout the body of the subject.

### 1.4.4 Coincidence Detection and Electronic Collimation

In a PET camera, each detector generates a timed pulse when it registers an incident photon. These pulses are then combined in coincidence circuitry, and if the pulses fall within a short time-window, they are deemed to be coincident. A conceptualised diagram of this process is shown in Figure 1-6. A coincidence event is assigned to a line-of-response (LOR) joining the two relevant detectors. In this way, positional information is gained from the detected radiation without the need for a physical collimator. This is known as electronic collimation. Electronic collimation has two major advantages over physical collimation which is used in SPECT: improved sensitivity and improved uniformity of the point source response function.

When a physical collimator is used, directional information is gained by preventing photons which are not normal or nearly normal to the collimator face from falling on the detector. In electronic collimation, these photons may be detected and used as signal. This results in a significant gain in sensitivity (typically a factor of 10 for 2D mode PET compared with SPECT). This increase in sensitivity means that typical realisable image resolution in PET is around 5-10 mm, whereas in SPECT it is around 15-20 mm.

![Coincidence events in a PET camera](image)

**Figure 1-6:** Coincidence detection in a PET camera

### 1.4.5 Types of Coincidence Events

Coincidence events in PET fall into 4 categories: true, scattered, random and multiple. The first three of these are illustrated in Figure 1-7.

- **True coincidences** occur when both photons from an annihilation event are detected by detectors in coincidence, neither photon undergoes any interaction prior to detection, and no other event is detected within the coincidence time-window.
- A **scattered coincidence** is one in which at least one of the detected photons has undergone at least one scattering event prior to detection. Since the direction of
the photon is changed during the scattering process, it is highly likely that the resulting coincidence event will be assigned to the wrong LOR. Scattered coincidences add a background to the true coincidence distribution which changes slowly with position, decreasing contrast and causing the isotope concentrations to be overestimated.

- Random coincidences occur when two photons not arising from the same annihilation event are incident on the detectors within the coincidence time-window. The number of random coincidences in a given LOR is closely linked to the rate of single events measured by the detectors joined by that LOR and the rate increases roughly with the square of the activity in the FOV. The distribution of random coincidences is fairly uniform across the FOV, and will cause isotope concentrations to be overestimated if not corrected for.

Both scattered and random events add statistical noise to the signal. The number of these false events detected depends on the volume and attenuation characteristics of the object being imaged, and on the geometry of the camera.

- Multiple coincidences occur when more than two photons are detected in different detectors within the coincidence resolving time. In this situation, it is not possible to determine the LOR to which the event should be assigned, and the event is rejected. Multiple coincidences can also cause event mis-positioning.

![Figure 1-7: Types of coincidences in PET](image)

**1.4.5.1 Detectors**

The most critical components of a PET camera are the detectors. In some cases they are large crystals of sodium-iodide coupled to many photo-multiplier tubes (PMTs). A more commonly used configuration is shown in Figure 1-8. In these detectors a rectangular bundle of crystals, a block, is optically coupled to several PMTs. When a photon interacts in the crystal, electrons are moved from the valence band to the conduction band. These electrons return to the valence band at impurities in the crystal, emitting light in the process. Since the impurities usually have metastable excited states, the light output decays exponentially at a rate characteristic of the crystal. The ideal crystal has high density so that a large fraction of incident photons scintillate, high light output for positioning accuracy, fast rise-time for accurate timing, and a short decay time so that high counting rates can be handled. Most current scanners use bismuth-germanate (BGO), which generates approximately 2500 light photons per 511 keV photon, and has a decay time of 300 ns.
One such block, for example, couples a 7x8 array of BGO crystals to four PMTs where each crystal is 3.3 mm wide in the transverse plane, 6.25 mm wide in the axial dimension, and 30 mm deep. The block is fabricated in such a way that the amount of light collected by each PMT varies uniquely depending on the crystal in which the scintillation occurred. Hence integrals of the PMT outputs can be decoded to yield the position of each scintillation. The sum of the integrated PMT outputs is proportional to the energy deposited in the crystal.

1.4.6 Acquisition Modes

Current clinical scanners consist of 18-39 rings of detectors, which are aligned axially. A volumetric PET data set is commonly constructed by collecting a stack of 2D transaxial images perpendicular to the axial (bed) direction. Thus, images along the axial direction (in the coronal or sagittal planes) are generated by re-sampling the volumetric voxel matrix along these planes. Many PET scanners have the option for restricting the LOR for gamma-ray coincident pair detection to the transverse plane perpendicular to the axial direction. This restrictive acquisition mode is called 2D acquisition mode. When no restrictions are applied during the acquisition process the number of events detected is maximized. This mode of operation is termed 3D acquisition mode. In this case, the data is typically re-binned into transverse planes and then reconstructed using a 2D algorithm that generates images from projection data. In either mode, the reconstructed 3D PET data set consists of a stack of 2D transverse images along an axial direction. Coronal or sagittal images are generated by re-sampling the voxel matrix along these planes.

1.4.7 Resolution

If the data are acquired in the 2D mode, the LORs connecting crystals can be binned into sets of parallel projections at evenly spaced angles as illustrated in Figure 1-9. There are two evident characteristics:
1. Samples are unevenly spaced, with finer sampling at the edges of the field-of-view than at the center.
2. The samples along the heavy solid line at angles one and three are offset by one-half of the detector spacing from samples at angle two.

Therefore, adjacent parallel projections can be combined to yield one-half the number of projection angles with a sampling distance of one-half the detector width. A typical block might have 3.3 mm thick crystals, so the resulting sampling distance would be 1.65 mm.

The Nyquist criterion is usually stated in medical imaging applications as requiring that the sampling distance is one-half the spatial resolution expressed as the full-width-at-half-maximum (FWHM). Hence, this block would support a spatial resolution of 3.3 mm. In fact, a scanner with this crystal size has a measured resolution that is somewhat worse, varying from 3.6 mm at the center of the field-of-view to 5.0 mm at 20 cm from the center. This occurs because scintillations usually consist of one or
more Compton interactions followed by photoelectric absorption (assuming the photon is not scattered out of the crystal). Since a 511-keV photon travels on average 7.5 mm in BGO before interacting, the light output is spatially distributed, especially at large radial distances where it is often distributed across two crystals.

The best obtainable resolution is termed the *intrinsic resolution*. This resolution is rarely achieved in practice because unfiltered images are usually very noisy. Although current scanners have intrinsic resolutions of less than 5 mm, the final resolution of the image is usually greater than 8 mm because the reconstruction algorithms trade-off resolution for reduced image variance. This final resolution is called the *reconstructed resolution*. Therefore, the resolution of PET images as they are typically used is not determined by the detectors, but by the degree to which resolution must be degraded to achieve an acceptable image variance. Since the variance is determined by the numbers of counts that can be collected during the scan, the constraints that govern the clinically useful resolution of PET images are the dosage of the radiopharmaceutical, the duration of the scan, the sensitivity of the scanner, and the count-rate capability of the scanner.

### 1.4.8 Image Reconstruction

For 2D reconstruction, the most commonly used algorithm for image reconstruction is the analytical method called Filtered Back-projection (FBP). FBP is straightforward to implement but has the property of amplifying noise in the signal. Considerable interest has been shown in iterative reconstruction schema, such as the Ordered Subsets - Expectation Maximisation (OSEM) algorithm, which has different noise properties than FBP. For 3D reconstruction, the Re-projection and Filtered Back-projection (3D-RP) method has been the most popular, in part because of the significant computational burden of newer 3D iterative reconstruction methods. 3D-RP itself is computationally expensive, and this has led to the development of approximate 3D reconstruction algorithms, such as Fourier Re-binning algorithm, which reduces the 3D problem to a series of 2D problems without significantly distorting the image and results in a significant reduction in the computational burden. The same image reconstruction methods are applied in SPECT imaging as well.
1.4.9 Applications of PET

Currently, FDG is the unique PET tracer with clinical applications. FDG is a glucose analogous that penetrates into the cells using sodium-glucose transporters and specific glucose membrane transporters. Once inside the cell, FDG is phosphorylated by the action of hexokinase to FDG-6-phosphate (FDG-6-P) and remains trapped in tissue, whereas glucose is not. From there, FDG-6-P does not follow the metabolic route of glycolysis or glycogen synthesis, being retained in the cell. The physiological or normal distribution of FDG in the organism corresponds to the cellular and tissue consumption of endogenous glucose. Brain is the organ that shows the major glucose uptake. FDG uptake in heart and skeletal muscles is variable and depends on glucose blood levels, endogenous insulin concentration and muscular contraction. Bowel uptake is also variable, depending on the peristaltism and muscular tone. There is urinary elimination of FDG; kidneys, ureteral tracts and urinary bladder are visualized in PET images.

1.4.9.1 PET in Cardiology

At the moment, FDG uptake measured by PET is considered the gold standard to diagnose myocardial viability. The main application of FDG-PET in cardiovascular pathology is to determine the myocardial viability in patients who will be submitted to coronary surgery. If there is no myocardial viability by PET, patients will be treated medically or included in a transplant waiting list.

PET tracers allow sequential rest-stress studies within a short time frame. The sequential tests help doctors determine the level of stenosis a patient is experiencing. Through PET images can also be detected Coronary Artery Disease by measuring myocardial blood flow and perfusion.

1.4.9.2 PET in Neurology

It is accepted that FDG-PET is the most exact in vivo imaging methodology to evaluate global and regional cerebral metabolism.

- For Alzheimer's testing - PET is the most accurate way to differentiate Alzheimer's from other forms of dementia. Because drugs that can significantly improve quality of life for Alzheimer's patients are now available, early detection of the disease is crucial.
- For movement disorders - PET greatly enhances the ability to diagnose involuntary movement disorders such as Parkinson's Disease, Huntington's Disease, and Tourette's Syndrome. PET makes diagnosis easier by revealing abnormal uptake patterns in the brain's dopamine receptors.

The evaluation of patients with partial epilepsy before the surgery is also recognized as an appropriate use of FDG-PET. PET helps surgeons pinpoint the surgical site for resection, eliminating the need for expensive and invasive methods. PET makes locating the surgical site possible by revealing areas of increased and decreased glucose utilization.

1.4.9.3 PET in Oncology

The main applications of FDG-PET are related to the diagnosis, staging, therapy monitoring, and prognosis of patients with cancer. PET findings have a direct clinical impact on the patient management. FDG-PET can show the pathological increase in glucose consumption that the tumour cells present in vitro. Phosphatase absence in the tumour cells causes an intense metabolic retention of FDG-6-P. A high contrast
between the tumour cell uptake and the healthy cell uptake results in a high detection sensitivity by PET. FDG uptake is related to the high cellularity and the cellular proliferation and, therefore, to the degree of malignancy. The most aggressive tumours require greater glucose consumption to maintain their accelerated growth, whereas those of low degree have less FDG uptake. Scars and already established necrosis and oedema do not show FDG uptake.

PET is a good diagnostic tool in oncology because:
- One can do a whole body study in the same exploratory act.
- It has a great sensitivity for detecting malignant neoplastic tissue, demonstrating tumour infiltration in normal sized lymph nodes and in organs that do not yet present anatomical alterations in CT or MRI.
- The findings are less artefacted by the therapy than those of CT and MRI. Thus, PET can distinguish residual post-therapy fibrosis and/or post surgical anatomical distortion from viable tumour tissue.
- It has a relatively high negative predictive value, since a normal study almost totally discards malignancy.

FDG-PET can have applications in several clinical situations:
- to support a diagnosis of a benign lesion or malignancy in processes detected by other techniques, but with complicated or impossible histological confirmation
- to establish the extension of an already known tumour prior to treatment – staging,
- to differentiate neoplastic tissue from fibrotic residual masses after surgery, chemotherapy and/or radiotherapy
- to locate a tumour recurrence suspected by clinical analysis and/or an increase in tumour markers
- to do a new study of extension after recurrence diagnosis - re-staging
- to evaluate the early therapeutic effects, and
- to look for the primary tumour in a patient with metastatic disease of unknown origin or para-neoplastic syndrome.

In addition, PET may guide needle aspiration or biopsy and allows the definition of the volume of irradiation in radiotherapy planning.

PET is used in cases of brain tumours, head and neck tumours, lung cancer, digestive cancers, lymphomas, melanomas, breast cancer, thyroidal cancer, genitourinary tract tumour as well as in case of unknown origin cancer in order to locate the primitive tumour.

1.4.10 PET/CT

Positron emission tomography facilitates the evaluation of metabolic and molecular characteristics of a wide variety of cancers, but is limited in its ability to visualize anatomical structures. Computed tomography facilitates the evaluation of anatomical structures of cancers, but cannot visualize their metabolic and molecular aspects. Therefore, the combination of PET and CT provides the ability to accurately register metabolic and molecular aspects of disease with anatomical findings, adding further information to the diagnosis and staging of tumours. The recent generation of high performance PET/CT scanners combines a state of the art full-ring 3D PET scanner and a high-end 16-slice (or more) CT scanner.
The combination of PET and CT imaging devices into a single scanner offers several advantages in comparison to PET or CT imaging alone. In combined systems, the CT can be used for the precise anatomical localization of the radiotracer uptake, for the attenuation correction and to reduce the PET examination time. However, the CT-based attenuation correction can lead to artifacts, and thus a review of the uncorrected images may be necessary to differentiate between true radiotracer uptake and tracer activity overestimation caused by artifacts. Only the absence of increased activity in the uncorrected images can truly confirm missing radiotracer activity in the region of the object, preventing “false” interpretations of infection, inflammation, or even malignancy around the object. It is important to take these technical principles into account when interpreting changes qualitatively or quantitatively. The diagnostic CT can be performed for the whole body, or to limit the radiation dose to the patient, centered on the specific region of interest in the body.

Hybrid PET/CT imaging will be very important in oncological applications in the future and possibly for use in cancer screening and cardiac imaging. In Figure 1-10 are shown example CT and PET body scans and the result of their combination.
Figure 1-11: Example body scans; (a) CT image; (b) PET image; (c) CT and PET images fused together
Chapter 2  Medical Imaging Resolution

2.1 Introduction

In all medical imaging systems the main goal is to increase the resolution and to achieve as much as possible true isotropic 3D imaging. Each imaging system has a characteristic resolution which is determined based on physical constraints of the detectors, which are reflected to the signal-to-noise ratio (SNR) and the timing considerations in the system. Within each imaging modality specific physical laws are in control, defining the meaning of noise and the sensitivity of the imaging process. In each case there are signal processing rules, which are applied in the system design in an attempt to achieve an acceptable compromise between resolution and SNR.

In this chapter, resolution limitations of the current medical imaging systems are discussed, as well as the physical constraints that determine the resolution and the gain from the increase in resolution for two main modalities, MRI and PET.

2.2 General Constraints

The main reason that causes the resolution limitations originates from signal-processing issues that need to be considered and in particular the Nyquist criterion. The sampling distance is required to be one-half the spatial resolution, defined as the distance between half-value points of the system impulse response, or the full-width-at-half-maximum (FWHM). Physically, the sampling rate is determined by the detector spacing. If the size (width) of the detectors and the inter-detector distances are reduced the resolution will be increased but there will also be a significant increase in the noise.

2.2.1 Partial Volume Effect (PVE)

One of the outstanding problems that medical image processing deals with, is known as the partial volume effect (PVE), which arises when an interface between two different tissues occur within a single voxel. The PVE is a direct consequence of limited resolution during the acquisition process. In general, PVE blurs the boundary between tissues and adds complexity to tissue characterizations.

In a CT image, each voxel represents the attenuation properties of a specific volume. When more than a single tissue is present within the voxel, the value will be some (non-linear) average of the tissues properties. Increased resolution can help overcome or reduce the problems associated with PVE.

In MRI, the intensity in a particular voxel depends on the entire contents of the corresponding anatomical volume and the sequence that is used. If only a single tissue type is present in the voxel, the signal intensity will be characteristic of that tissue type. However, if more than one tissue type is present, the signal will be a combination of the contributions of the different tissues. This results in an uncertainty in the boundaries between tissues and in the quantitative measurements of structures.
2.3 MRI Resolution

2.3.1 Spatial Resolution

MRI provides intensities for each voxel. These intensities are proportional to the number of nuclei in each voxel and are affected by the nuclear relaxation times and the pulse sequence used. Those effects affect the image contrast. MRI spatial resolution is determined by gradients' intensity, digital imaging filter bandwidth, the number of 'read-out' points and phase encoding steps.

Spatial resolution can be enhanced by:
- Decreasing the field-of-view (FOV).
- Increasing the number of readout points.
- Increasing the number of phase encoding steps.

The FOV is limited by the gradient strength and the subject dimension in the readout direction. The number of readout points is limited by the transverse nuclear relaxation time (T2). Extending the readout period significantly beyond the transverse relaxation time decreases the signal-to-noise ratio (SNR) significantly. The maximum theoretical number of readout points is limited by the local oscillator frequency divided by the exciter/receiver register bit. This limit is impractical since the long acquisition time decreases the SNR to an unfeasible degree. Though T2 decay is usually the only important limit, the amount of available memory for storing the data also comes up occasionally as a limiting factor. The number of phase encoding steps is limited by the acquisition time. Increasing the number of phase encoding steps increases the acquisition time, proportionally.

2.3.2 Slice thickness

In cases in which true 3D image acquisition is not effective or possible, it is common practice to obtain a set of 2D slices. Such is often the case, for example, in T2-weighted imaging, diffusion-weighted imaging and MR angiography. These are all imaging techniques that are important for early medical diagnosis and visualization purposes and usually require coverage of extensive 3D volumes in the imaged object. T2-weighted imaging is difficult to be applied in reasonable times when 3D acquisition methods are used. The problem arises from the need for long signal recovery between excitations to enable operation of the spin-echo mechanism that provides the contrast. Since all the spins are excited by every pulse, the recovery time cannot be utilized and the sequence takes a long time. In diffusion-weighted imaging currently there is no 3D technique for humans. Sequences that acquire raw data referring to the same slice or volume over many excitations cannot be modified in order to provide the contrast because of phase inconsistencies resulting from physiological motion. MR angiography is another application in which 3D acquisition results in better performance.

The problem, as illustrated in Figure 2-1, is that a set of 2D slices does not give a good isotropic 3D image. The slice thickness in MRI is determined by the slice-selection pulse, which is in turn determined by hardware limitations and pulse sequence timing considerations.

Figure 2-1: MRI slice acquisition
As a result, the resolution in-plane (x, y) is high and much reduced in the slice-select (z) direction.

2.3.3 Temporal Resolution and fMRI

In fMRI temporal resolution is also important. Most 3D acquisition procedures cannot reach the required temporal resolution which is necessary for appropriate statistical analyses. Increased spatial resolution would help the visualization of smaller units of neuronal activity. The acquisition of higher resolution images results in a reduction in signal-to-noise ratio, which is proportional to the decrease in the voxel size. In order to restore the SNR, higher magnetic field scanners can be utilized, but this causes increased inhomogeneity and larger distortion artifacts in the images, which are especially evident in fMRI.

2.4 PET Resolution

The increase of resolution of PET images may prove to be beneficial for research and clinical practice. The imaging of small cerebral structures such as the cortical sub layers and nuclei may need PET spatial resolutions of 2 mm or less. Higher PET resolution would also be beneficial for improving sensitivity for detection of small tumours. Cancer lesions need to be of diameters equal or larger than the resolution of the PET scanner to be identified provided they also have a high glucose metabolism. Finally, higher resolution images may show a more differentiated anatomical structure. The increase in anatomical detail may aid in the registration of a PET image with a corresponding anatomical image from another modality, such as CT or MRI.

2.4.1 Physical Constraints

PET is known to have poor spatial resolution. The resolution is limited by physical properties, such as:
- scatter,
- counting statistics,
- positron range,
- patient motion,
- the detectors’ geometry and
- the implemented acquisition protocol.

Current clinical scanners consist of 18-39 rings of detectors, axially aligned. As already mentioned in the previous chapter, a volumetric PET data set is commonly reconstructed by collecting a stack of 2D transaxial images perpendicular to the axial (bed) direction. Images in the coronal or sagittal planes are generated by re-sampling the volumetric voxel matrix along these planes. Accordingly, the spatial resolution in the transaxial plane is mostly limited by the detector width, whereas the resolution along the axial direction is affected by the distance between the detector rings. In practice, the final reconstructed resolution of a PET image is usually poorer than the best obtainable, intrinsic resolution, because reconstruction algorithms typically tradeoff resolution for reduced noise. Typical resolution in clinical scanners is between 4-7mm FWHM.
A solution in order to increase the resolution would be to decrease the width of the detectors. However, detector widths are limited to a certain minimal size, due to SNR considerations. If the width is too small, detection efficiency will be reduced and inter-crystal scatter and penetration will be increased.

Another important factor that needs to be considered is the image variance. In order to achieve an acceptable image variance, resolution in PET scanners is often degraded. The variance is mainly determined by the number of counts (counting statistics) collected during the scan. The noise that affects the counting statistics is comprised of several factors:

- The angular uncertainty of the photons created in the annihilation process. Although the photons emitted in this process should move in a straight line of 180 degrees with respect to each other, there is a small angular divergence.
- The scatter events when one or both of the photons may pass before they reach the detector. This causes mis-estimation of the line where the annihilation process took place.
- The random events that occur simultaneously leaving only one photon detected from a certain annihilation and the other one not. Two events at the same time draw a wrong line between the detector and thus adding wrong information to the reconstructed image.

The above SNR considerations result in an under-sampling of available data. If wider detectors are used, the sampling frequency is lower, while a high SNR ratio is preserved.
Chapter 3  Super Resolution Methods

3.1  Introduction

As the electronic imaging applications are becoming more demanding, images with high resolution (HR) are desired and often required. High resolution means that pixel density within an image is high and therefore an HR image can offer more details that may be critical in various applications. For example, HR medical images offer significant assistance to the doctor to make a correct diagnosis. It may be easy to distinguish an object among similar ones using HR satellite images, and the performance of pattern recognition in computer vision can be improved if an HR image is provided. Although the performance of the existing imaging sensors has improved since their invention, they still suffer from physical limitations and malfunctioning problems, which degrade the quality of captured images. As a result, the current resolution level and the consumer price are most likely that will not satisfy the future demand. The imaging chips and optical components necessary to capture very HR images become prohibitively expensive, especially for scientific applications. Thus, finding a way to increase the current resolution level is needed.

In this chapter we present an overview of super resolution techniques, including the main observation model and a short description of the reconstruction algorithms proposed in the literature.

3.1.1 Limitations of Imaging Sensors

Since the 1970s, systems employing charge-coupled device (CCD) and CMOS image sensors have been widely used to capture digital images for various purposes, such as medical tomography, industrial monitoring system, surveillance system, computer vision system, scientific research applications, broadcasting system, and consumer appliances. Although these sensors are widely used and suitable for most imaging applications, they still have physical limitations. First, an image captured with these imaging sensors may be degraded by noise from various sources that is generated in the imaging sensors. Noise is mainly categorized into photon noise (shot noise), thermal noise, and readout noise. The limited dynamic range is another problem in the image acquisition process with CCD and CMOS imaging sensors. The range of light intensity of natural scenes is very wide, but the dynamic range of an imaging sensor is limited. When capturing a scene with a wide range of light intensity that exceeds the dynamic range of the sensor, there will be a loss of information in the low-intensity areas, the high-intensity areas, or both. Moreover, an imaging sensor is composed of spatial pixels, resulting in limited spatial resolution.

Besides these limitations, CCD and CMOS imaging sensors cannot recognize the colour information of light, but can only generate the electrical signal according to the intensity of the light. Consequently, to reconstruct the colour information, colour filters and multiple imaging sensors are used. Because the use of multiple imaging sensors is very expensive, most digital cameras, camcorders and surveillance cameras use a single sensor to reduce size and cost. A mosaic or grid of colour filters is placed on the face of the imaging sensors to filter light of a specific wavelength. Colours of a whole image are reconstructed based on this filtered light information. Thus, the color
image captured with the single CCD and CMOS imaging sensors do not have sufficient colour resolution.

3.1.2 Hardware Solutions to Increase Resolution

The most direct solution to increase spatial resolution would be in the field of sensor manufacturing. By using special manufacturing techniques, the pixel size can be reduced, so the number of pixels per unit area is increased. However, as the size of the pixel decreases, there is a decrement in the amount of light, which produces shot noise, having as a result a bad quality image. For this reason, in the current image sensor technology there are certain limitations for the reduction of the pixel size, in order to avoid the effects of the noise.

Another solution, again in the field of optics manufacturing, would be to increase the chip size, but this would lead to an increase in capacitance. Because a large capacitance makes it difficult to speed up a charge transfer rate, this approach is considered ineffective.

An important concern in most commercial applications regarding HR imaging is the cost. As high precision optics and image sensors involve high costs, the approaches in the field of hardware are difficult to be effective. Therefore, a software approach is the next logical step to improve image resolution beyond the spatial bandwidth of existing camera systems.

3.2 Signal Processing-Based Approach: Super Resolution

Although researchers of imaging sensors have been endeavouring to overcome the limitations based on device physics and circuit technology, there is an alternate but efficient approach to solve the problems, which is based on signal processing technology. The physical limitations and the resulting artifacts are considered as signals generated from the sensing system. After analyzing the system models, the problems can be solved by mathematical inverse procedure, which is implemented and applied right after the acquisition of the output signal as post-processing. With this signal processing-based approach, image noise can be removed with statistical modelling of the image and the noise and limited dynamic range can be improved through multiple image of the same scene taken with different exposure times. In order to overcome the resolution limitations, one promising idea is to use signal processing techniques to enhance the spatial resolution. This approach proposes the acquisition of an HR image from observed multiple low-resolution (LR) images. This image restoration approach is called Super Resolution (SR) image reconstruction (or restoration). The major advantage of the approach in the signal processing field is that it has lower cost and allows the use of the existing LR imaging systems.

Super Resolution is the process of combining multiple low resolution images to form a high resolution image.

Broadly all signal processing-based approaches to overcome the physical limitations of imaging sensors can be called a super-resolution approach. Super-resolution image reconstruction was originally referring to the methods overcoming the limited spatial resolution of the imaging systems, but the meaning of the term has broadened and now refers to techniques to overcome not only limited spatial resolution, but also all physical limitations of imaging sensors in the signal
processing-based approach. In the following sections we will refer to super resolution reconstruction to enhance the spatial resolution.

### 3.2.1 Applications of Super Resolution

The SR image reconstruction methods are proving to be very useful in practical cases where multiple LR images (frames) of the same scene can be obtained. Application of such restoration methods arises in many areas such as:

1. **Satellite applications**: where several images of the same area are provided and an enhanced resolution image of a target can be obtained.
2. **Video applications**: where typical images obtained with an inexpensive camera are generally of poor quality and unsuitable for printing or frame freeze purposes. Enhancement of a freeze image can be done by using several successive images merged together by a super resolution algorithm. This synthetic zooming of a region of interest is an important application in surveillance, forensic and scientific imaging.
3. **Medical imaging (CT, MRI, PET, etc.)**: where the acquisition of multiple images with limited resolution quality is possible.

### 3.2.2 HR Image from Multiple LR Images

The basic requirement in order to apply SR restoration techniques is the availability of multiple LR images captured from the same scene. These LR images represent different “looks” at the same scene. The LR images are sub-sampled (aliased) as well as shifted with subpixel precision. If the images are shifted by integer units, then there is no new information available to reconstruct an HR image, since each image contains the same information. However, if the subpixel shifts between the LR images are different from each other and if aliasing is present, then each image has different information content and cannot be produced from the others. In this case, the new information contained in each LR image can be utilized to obtain a HR image, as illustrated in Figure 3-1. The different looks at the same scene can result from some relative scene motions from frame to frame by obtaining multiple scenes or video sequences.

![Figure 3-1: Basic requirement for super resolution](image-url)
The different looks at the same scene can result from some relative scene motions from frame to frame by obtaining multiple scenes or video sequences. Multiple scenes can be obtained using one camera with several captures or using multiple cameras located in different positions. The scene motions can be controlled motions in imaging systems, such as in images acquired from orbiting satellites, or uncontrolled motions, such as the movement of local objects or the movement in vibrating imaging systems. For SR image reconstruction to be possible, it is essential that these scene motions are known or can be estimated with sub-pixel accuracy.

As mentioned above, in the process of recording a digital image, there are several factors that degrade the quality of the image:

- Optical distortions (out of focus, diffraction limit, etc.) which cause a natural loss of spatial resolution
- Limited shutter speed, which causes motion blur
- Noise that occurs within the sensor or during transmission
- Insufficient sensor density.

Thus, the recorded image is of low quality as it usually suffers from blur, noise, and aliasing effects. Although the main concern of an SR algorithm is to reconstruct HR images from under-sampled LR images, it covers image restoration techniques that produce high quality images from degraded and aliased images. Therefore, the goal of SR techniques is to restore a single HR image from several blurred, noisy and under-sampled measured LR images.

### 3.2.3 Related Problems

A related problem to super resolution is image restoration, which is a well-established area in image processing applications. The consistent development of computer technology led to a growing interest in image restoration theory. The goal of image restoration is to recover a degraded (e.g. blurred, noisy) image without changing the size of the image. The main directions were non-traditional treatments to the classic problem and looking at new, second-generation restoration problems, allowing for more complicated and more computationally intensive algorithms. Among these second-generation problems is considered to be SR image reconstruction.

Another problem related to SR reconstruction is image interpolation that has been used to increase the size of a single image. Although there has been extensive research in this field, the resulting quality of an image magnified from an aliased LR image is inherently limited even though the ideal sinc basis function is used. Single image interpolation cannot recover the high-frequency components lost or degraded during the LR sampling process. For this reason, image interpolation methods are not considered as SR techniques. The next step to achieve further improvements in this field would be the utilization of multiple data sets in which additional data constraints from several observations of the same scene can be used. The fusion of information from various observations of the same scene allows SR reconstruction of the scene.

### 3.2.4 Observation Model

The key to a comprehensive analysis of the super resolution problem is to formulate the problem and to model it as simply and as effectively as possible. This model will relate the HR image to the observed LR images. A block diagram of the observation model is illustrated in Figure 3-2.
Consider the desired HR image $f$ of size $L_1N_1 \times L_2N_2$, which is the ideal undegraded image that is sampled at or above the Nyquist rate from a continuous scene which is assumed to be bandlimited. The parameters $L_1$ and $L_2$ represent the down-sampling factors in the observation model for the horizontal and vertical directions, respectively. Thus, each observed LR image is of size $N_1 \times N_2$. It is assumed that $f$ remains constant during the acquisition of the multiple LR images, except for any motion and degradation allowed by the model. Therefore, each observed LR image $g_k$ is expressed as the result of a sequence of operators on the original high-resolution image source $f$, consisting of a geometrical warp, blurring and down-sampling. Assuming that each LR image is corrupted by additive noise, the observation model can be represented by the following equation.

$$g_k = (T_k (f) \ast h) \downarrow s + n_k, \quad k = \{1 \ldots K\}$$  \hspace{1cm} (3.1)

where:

- $T_k$ is the geometrical transformation (rotation and/or translation) of the image $f$ to the same reference frame of acquisition for $g_k$,
- $h$ is a blur kernel, often referred to as the point spread function (PSF), defined by the physical properties of the imaging device,
- $n_k$ is additive noise,
- $\downarrow s$ represents the down-sampling of a HR image to a LR image by factors $L_1$ horizontally and $L_2$ vertically, and
- $\ast$ is the convolution operator.

Based on the above model, the aim of the SR image reconstruction is to estimate the HR image $f$ from the LR images $g_k$, for $k=1, \ldots, K$. The index $k$ refers to acquisitions at different points of view.

The motion that occurs during the image acquisition is represented by the geometrical transformation $T_k$, which can be global or local rotation and/or translation. Since this information is generally unknown, there is the need to estimate the scene motion for each frame with reference to one particular frame. The warping process performed on the HR image is actually defined in terms of LR pixel spacing when we estimate it. Thus, this step requires interpolation when the fractional unit of motion is not equal to the HR sensor grid, which is essential for the SR reconstruction.
The necessity of interpolation is shown in Figure 3-3. In this figure, the circle (●) represents the original (reference) HR image, and the triangle (▲) and the diamond (◊) are shifted versions of it. If the down-sampling factor is two for example, a diamond (◊) has (0.5, 0.5) subpixel shift for the horizontal and vertical directions and a triangle (▲) has a shift which is less than (0.5, 0.5). Therefore, a diamond (◊) does not need interpolation, but a triangle (▲) should be interpolated since it is not located on the HR grid.

Although one could use ideal interpolation theoretically, in practice, simple methods such as zero-order hold or bilinear interpolation methods have been used in many literatures.

Blurring may be caused by several factors, such as:

- characteristics of an optical system (e.g., out of focus, diffraction limit, aberration, etc.),
- relative motion between the imaging system and the original scene, and
- the point spread function (PSF) of the LR sensor.

The blurring process can be modelled as a linear space invariant (LSI) or a linear space variant (LSV), and its effects on the HR image are represented in the observation model by the convolution with the blur kernel $h$. In single image restoration applications, is usually considered optical or motion blur. In the SR image reconstruction, however, the physical limitations of the imaging sensors are also considered as a quality degradation factor. Thus, the finiteness of a physical dimension in LR sensors is considered as an important factor of blur. This LR sensor PSF is usually modelled as a spatial averaging operator. In the use of SR restoration methods, the characteristics of the blur are assumed to be known. However, if it is difficult to obtain this information, blur identification should be incorporated into the reconstruction procedure.

The warped and blurred HR image undergoes a down-sampling process, resulting in aliased LR images. Although, the blurring is considered to act more or less as an anti-aliasing filter, in SR image reconstruction, it is assumed that aliasing is always present in LR images.

### 3.2.5 Stages in SR Image Reconstruction

Most of the SR image reconstruction methods proposed in the literature consist of three stages, as illustrated in Figure 3-4:

1. Registration
2. Interpolation
3. Restoration (i.e., inverse procedure)
These steps can be implemented separately or simultaneously according to the reconstruction method adopted. The estimation of motion information is referred to as registration, which is an extensively studied field of image processing.

In the registration stage, the relative shifts between LR images, with reference to a certain LR image, are estimated with fractional pixel accuracy. Accurate subpixel motion estimation is a very important factor in the success of the SR image reconstruction algorithm. Incorrect motion estimation has disastrous implications in the overall SR performance. Since the shifts between LR images are arbitrary, the registered HR image will not always match up to a uniformly spaced HR grid. Thus, non-uniform interpolation is necessary, as mentioned in the previous section, to obtain a uniformly spaced HR image from a non-uniformly spaced composite of LR images. Finally, image restoration is applied to the up-sampled image to remove blurring and noise.

3.3 Overview of Super Resolution Methods

The differences among the several proposed works on super resolution are subject to what type of reconstruction method is employed, which observation model is assumed, in which particular domain (spatial or frequency) the algorithm is applied and what kind of methods is used to capture the LR images. In this section, we provide an overview of the history and the existing super resolution algorithms, categorizing them according to the domain in which they are applied.

3.3.1 Frequency Domain Methods

A major class of SR methods utilizes a frequency domain formulation of the super resolution problem. These techniques utilize the shifting property of the Fourier transform to model global translational scene motion, and take advantage of the sampling theory to enable effect restoration made possible by the availability of multiple observation images. The frequency domain methods include the earliest investigation of the super resolution problem, and although there are significant disadvantages in the frequency domain formulation, work has continued in this area until relatively recently when spatial domain techniques, with their increased flexibility, have become more prominent.

The frequency domain approach is based on the following principles:

i) The shifting property of the Fourier transform
ii) The aliasing relationship between the continuous Fourier transform of an original HR image and the discrete Fourier transform of observed LR images
iii) The assumption that the original HR image is bandlimited
The super resolution restoration idea was first presented by Tsai and Huang [21]. They used the frequency domain approach to demonstrate the ability to reconstruct one improved resolution image from several down-sampled noise-free versions of it, based on the spatial aliasing effect. Other results suggested a simple generalization of the above idea to noisy and blurred images. A frequency domain recursive algorithm for the restoration of super resolution images, which improves the computational efficiency, is suggested in [22]-[24]. These approaches utilize the frequency domain theoretical framework as well as the global translation observation model proposed by Tsai and Huang. However, the formulation is extended to consider observation noise as well as the effects of spatial blurring. These techniques were based on recursive least squares solution methods and later they were extended in order to include a degree of robustness to errors. The technique for achieving this is the method of total least squares [25]-[26].

Ur and Gross [27] propose a SR reconstruction method based on the generalized sampling theorem of Papoulis [28] and a variant thereof, the multi-channel sampling theorem, by Brown [29]. Although the implementation of the reconstruction is achieved in the spatial domain, the technique is fundamentally a frequency domain technique relying on the shift property of the Fourier transform to model the translation of the source image.

### 3.3.1.1 Summary of Frequency Domain Methods

SR reconstruction via the frequency domain approach has significant advantages:

- **Simplicity**: The principles behind the approaches are readily understandable in terms of basic Fourier theory.
- **Computational complexity**: Many of the techniques discussed above are computationally attractive and allow parallel implementation.
- **Intuitive super resolution mechanism**: The techniques are all based on the de-aliasing techniques of [21]. This formulation makes explicit the underlying mechanism of resolution improvement - restoration of frequency components beyond the Nyquist limit of the individual observation image samples.

There are, however, significant disadvantages, which must be considered when dealing with more general SR video restoration problems:

- **Global translation motion model**: All the frequency domain methods discussed utilize a global translational motion model. This is a fundamental limitation of the frequency domain approach.
- **Inflexibility regarding motion models**: The existence of a transformation which is the Fourier domain equivalent of the spatial domain motion model is required. This imposes severe limitations on the range of feasible motion models that can be used in SR applications.
- **Degradation models**: It is difficult to include spatially varying degradation models in the frequency domain reconstruction formulation.
- **Inclusion of spatial domain a-priori knowledge for regularization**: Often the most useful a-priori knowledge used to constrain the solution space to effect regularization is via spatial domain constraints, which are highly convenient and intuitive. Frequency domain methods do not lend themselves well to the inclusion of such constraints.

In the limited case of global translation motion, there is significant benefit in frequency domain approaches to SR restoration, however if we are interested in more general classes of motion, as well as degradations, it is clear that the frequency domain approach is insufficient.
3.3.2 Spatial Domain Methods

The second major division in super resolution literature includes the spatial domain methods. In these methods the super resolution reconstruction is applied in the spatial domain.

3.3.2.1 Non-uniform Interpolation approach

The most intuitive method for SR image reconstruction is the non-uniform interpolation approach. A scheme for this approach is shown in Figure 3-5. The three stages presented in Figure 3-4 are performed successively in this approach:

i) Estimation of relative motion, i.e. registration (if motion information is unknown)
ii) Non-uniform interpolation to produce an improved resolution image
iii) Deblurring process (depending on the observation model)

Though this approach may initially appear attractive, it is, however, overly simplistic as it does not take into consideration the fact that samples of the LR images do not result from ideal sampling but are, in fact, spatial averages. The result is that the reconstructed image does not contain the full range of frequency content that can be reconstructed given the available LR observation data.

Keren, Peleg and Brada [30] describe a spatial domain approach to image registration using a global translation and rotation model, as well as a two stage approach to super-resolution reconstruction. Another interpolation based approach is proposed by Aizawa, Komatsu and Saito [31], who presented a scheme to acquire an improved resolution image by applying the Landweber algorithm from multiple images taken simultaneously with multiple cameras. Tekalp, Ozkan and Sezan [32] propose a two step procedure where the up-sampling of the low resolution images and the restoration are performed sequentially.

The techniques discussed above have the advantage of simplicity but the observation models they use are in general unrealistic as they do not properly account for the effects of optical blurring, motion blurring or noise.

3.3.2.2 Algebraic Filtered Backprojection

An early algebraic tomographic filtered backprojection approach to SR reconstruction is that of Frieden and Aumann [33], in which the authors consider the problem of SR image reconstruction from multiple 1-D scans of a stationary scene by a linear imaging array. In the proposed model, the presence of observation noise is not considered, and since the inverse filtering is highly noise sensitive, this has serious implications in the performance of the method.
3.3.2.3 Simulate and Correct Methods

A large class of SR reconstruction methods consists of algorithms which have in common a ‘simulate and correct’ approach to reconstruction. Given an estimate of the SR reconstruction and a model of the imaging process, the imaging process is simulated using the SR estimate as the input to produce a set of simulated LR observed images. These images are compared with the actual observed images and an error is computed and used to correct the estimate of the SR image. This simulate/correct process is iterated until some stopping condition is met - typically the minimization of some error criterion between the simulated and observed images.

Peleg, Keren and Schweitzer [34] propose a SR reconstruction from a set of globally translated images of an unchanging 2-D scene. Keren, Peleg and Brada in [30] use a more general global translation and rotation model. Irani and Peleg [35]-[37], suggest an approach based on the iterative back-projection (IBP) method adopted from computer aided tomography (CAT). In this approach, the main idea of which is shown in Figure 3-6, the HR image is estimated by back-projecting the error (difference) between synthetically generated LR images and the observed LR images. This process is repeated iteratively to minimize the energy of the error. This method is not limited to specific motion characteristics and allows arbitrary smooth motion flow, although the convergence of the proposed algorithm is proven only for

![Figure 3-6: Iterative Back-projection method](image)
an affine geometric warp between the measured images. The advantage of the IBP is that it is understood intuitively and easily. However, it has no unique solution and it is difficult to apply \textit{a-priori} constraints. We will discuss this method more thoroughly in the following chapter, as it is the method we used to evaluate the SR reconstruction of medical images.

### 3.3.2.4 Probabilistic Methods

Another major research track in SR reconstruction is the probabilistic approach. Since super resolution is an ill-posed inverse problem, techniques which are capable of including \textit{a-priori} constraints are well suited to SR reconstruction. In recent years, Bayesian methods, which inherently include \textit{a-priori} constraints in the form of prior probability density functions, are becoming very popular and are now central to the solution of ill-posed inverse problems in a wide range of applications. The Bayesian approach is synonymous with \textit{Maximum A-Posteriori} (MAP) estimation.

The SR reconstruction from a LR video sequence using the MAP technique is proposed by Schultz and Stevenson [38]. They propose a discontinuity preserving the MAP reconstruction method using the Huber-Markov Gibbs prior model, resulting in a constrained optimization problem with a unique minimum. They also consider independent object motion and inaccurate motion estimates that are modelled by Gaussian noise. A MAP framework for the joint estimation of image registration parameters and the HR image is presented by Hardie, Barnard and Armstrong [39]. The registration parameters, horizontal and vertical shifts in this case, are iteratively updated along with the HR image in a cyclic optimization procedure. Cheeseman, Kanefsky, Kraft and Stutz [40]-[42] have been applying MAP super-resolution reconstruction to Viking, Voyager and more recently Mars Pathfinder imagery. Their formulation assumes Gaussian noise and utilizes a prior which leads to a linear system of equations which are solved using Gauss-Jacobi methods.

Tom and Katsaggelos [43] examine the super-resolution reconstruction problem as composed of three steps which can be performed simultaneously - registration of the low resolution images, restoration of these images followed by an interpolation step which yields the SR reconstruction. This approach is essentially based on a maximum likelihood (ML) formulation of the phases of the SR reconstruction. The proposed ML estimation problem is solved by the expectation-maximization (EM) algorithm. A fundamental problem with this approach is the fact that ML estimation is poorly suited to the solution of ill-posed inverse problems due to high noise sensitivity when the reconstruction problem is under-specified. Thus, MAP estimation is usually used in preference to ML.

Robustness and flexibility in modelling noise characteristics and \textit{a-priori} knowledge are the major advantages of the probabilistic approach. MAP estimation with convex energy functions in the priors ensures the uniqueness of the solution. It is also possible to estimate the motion information and the HR image simultaneously.

### 3.3.2.5 Set Theoretic Methods

One of the prominent approaches to SR reconstruction is based on the method of projection onto convex sets (POCS). It is an alternative iterative approach to incorporating prior knowledge about the solution into the reconstruction process. The POCS formulation was first suggested by Stark and Oskoui [44]. Their method was extended by Tekalp, Ozkan and Sezan [32].

In the POCS formulation, constraint sets are defined which limit the feasible solution space of the SR reconstruction. Constraints are defined as convex sets in the vector space containing all possible SR reconstructions. Sets that represent desirable characteristics of the solution are defined, such as positivity, bounded
energy, fidelity to data, smoothness, etc. The solution space of the SR reconstruction problem is thus the intersection of the convex constraint sets. POCS refers to an iterative procedure which, given any point in the vector space, locates a point which satisfies all the convex constraint sets.

A variant of the POCS based formulation using an ellipsoid to bound the constraint sets has been investigated by Tom and Katsaggelos [45], [46]. This approach takes a form closely related to regularized methods. The observation model used in this work is similar to that proposed by Schultz and Stevenson [38].

The advantage of POCS is that it is simple and it utilizes the powerful spatial domain observation model. It also allows an easy inclusion of a-priori information. However, these methods have no unique solution and require considerable computation and a large number of iterations to achieve convergence.

### 3.3.2.6 Hybrid ML/MAP/POCS Method

Some work has been undertaken on combined ML/MAP/POCS based approaches to SR reconstruction. In particular the desirable characteristics of MAP estimation and those of the very flexible POCS method could be combined in a hybrid optimization.

Earlier efforts for this formulation are found in the work by Schultz and Stevenson [38], where MAP optimization is performed while projections-based constraint is also utilized. Elad and Feuer [47]-[48] propose a hybrid ML/POCS based method which uses the statistical ML formulation to pose SR as a statistical estimation problem, while utilizing projections-based constraints to effect regularization.

The hybrid approach is highly promising as it combines the most favourable characteristics of statistical methods (optimal estimation theoretic solution, mathematical rigor and direct inclusion of a-priori constraints) and POCS based approaches (powerful mechanism for inclusion of linear and nonlinear, set theoretic a-priori constraints).

### 3.3.2.7 Optimal and Adaptive Filtering Methods

Several researchers have proposed inverse filtering approaches to SR reconstruction. Erdem, Sezan and Ozkan [49] have proposed a LMMSE filtering approach, the motion compensated multiframe Wiener filter, for restoration of image sequences degraded by LSI spatial blur and additive noise. A global translation model is assumed, but motion blurring is not incorporated. This approach is notable because a simultaneous multi-frame restoration is undertaken. Elad and Feuer [50] propose a technique based on adaptive filtering applied in time axis. They suggest least squares (LS) estimators for the reconstructed image based on a pseudo-RLS or R-LMS algorithm. The steepest descent (SD) is applied to estimate the HR image at each time iteratively and the LMS algorithm is derived from the SD algorithm. As a result, the HR image is calculated without computational complexity of a direct matrix inversion.

### 3.3.2.8 Summary of Spatial Domain Methods

SR reconstruction via the spatial domain approach addresses has some advantages over the frequency domain approaches:

- **Motion models**: Spatial domain methods are capable of including an almost unlimited range of motion models (local or global).
- **Degradation models**: It is simple to include linear degradations such as motion blurring resulting from a non-zero aperture time, spatially varying or invariant blurs, missing pixels, etc.
• Inclusion of spatial domain \textit{a-priori} knowledge for regularization: Markov random fields as well as the spatial domain POCS formulation provide simple, but very powerful methods to incorporate \textit{a-priori} constraints into the reconstruction process.

• Powerful mechanism for bandwidth extrapolation: The combination of data from multiple images, as well as the use of realistic \textit{a-priori} constraints on the reconstructed image provides spatial domain methods with a powerful mechanism for image bandwidth extrapolation. It is even possible to extrapolate frequency information beyond the diffraction limitations of the optical system.

• Theoretical framework: Probabilistic methods, especially the MAP estimation method, provide a solid mathematical framework within which further theoretical developments can be made.

Spatial domain methods however have some disadvantages:

• Complexity: The optimizations involved in spatial domain methods are more complex than their frequency domain counterparts.

• Computational complexity: The increased flexibility of spatial domain methods tends to come at the cost of much increased computational requirements.

• Non-intuitive SR mechanism

It is obvious that for SR reconstruction of scenes involving anything more than global translational motion, the spatial domain techniques are the preferred approach.
Chapter 4  \textit{Experimental procedure}

4.1 Introduction

As mentioned in the previous chapter, in most imaging applications and specifically in medical imaging applications, resolution enhancement is becoming an essential requirement. Super resolution image reconstruction is one of the most prominent research areas, since it can overcome the inherent resolution limitation of the imaging system and improve the performance of most imaging applications. In medical imaging, the purpose is to improve the resolution as determined by the imaging device detectors, in cases in which the detectors under-sample the original data, while preserving a high signal-to-noise ratio (SNR). SR overcomes the detector sampling limit by over-sampling the data and minimizing the aliasing problem.

An experimental evaluation of the SR reconstruction applied in medical images is performed. SR techniques reconstruct a HR image from a series of LR images, which represent different views of the same scene. The followed experimental process can be divided in three phases:

1. Image simulation
2. Application of SR reconstruction method
3. Performance measures

In the first phase, simulated images of a computer generated phantom are formed and processed in order to comply with the observation model for the LR images, as it was presented in §2.4 of chapter 3. These images are used, in the second phase, as the LR images from which the HR image is constructed through the SR method. There are several SR methodologies proposed, as discussed in chapter 3. The iterative back-projection (IBP) algorithm suggested by Irani and Peleg [37] has been chosen to be utilized, which belongs in the spatial domain methods and it is an easily and intuitively understood method. The selected SR algorithm is applied to the simulated images and certain qualitative and quantitative measures are used to evaluate the performance in the third phase. In this chapter, the details of the experimental procedure are discussed, including the materials and methods utilized and the implementation issues during the process.

4.2 Phase 1: Image Simulation

This phase of the experimental process involves the formation of images which will represent different ‘points of view’ of the same scene, in this case of a computer generated phantom. The images follow the observation model utilized by Irani and Peleg [37].

4.2.1 Medical Imaging Simulation Techniques and Computer Phantoms

Simulation is a powerful tool for characterizing, evaluating, and optimizing medical imaging systems. Simulation involves computer generated phantoms, models of the imaging process, and fast computational methods. Computer phantoms provide a model of the subject’s anatomy and physiology. Provided a model of the physics of the imaging process, acquired data of a computer phantom can be generated using computational methods. A major advantage to using computer-generated phantoms in simulation studies is the knowledge of the exact anatomy and physiological functions of the phantom, providing a gold standard for the evaluation and improvement of medical imaging devices and image processing and reconstruction techniques. Other advantages are that computer phantoms are always available as 'willing participants' and they are easy to be altered in order to model different anatomies and medical situations providing a large population of subjects from which to perform research. It is frequently difficult both ethically and practically to test every combination of parameters on patients under clinical conditions.

The most important aspect of simulation is to have a realistic phantom or model of the subject's anatomy. Without this, the results of the simulation may not be indicative of what would occur in actual patients or animal subjects and would, therefore, have limited practical value.

4.2.2 The 4D NCAT Phantom

The 4D NURBS-based Cardiac-Torso (NCAT) phantom was originally developed to provide a realistic and flexible model of the human anatomy and physiology for use in nuclear medicine research, specifically single-photon emission computed tomography (SPECT) and positron emission tomography (PET) [51]. The phantom is a hybrid between the realism of pixel-based phantoms and the flexibility of geometry-based phantoms. Non-uniform rational b-splines, or NURBS surfaces were used to construct the organ shapes using the 3D Visible Human CT dataset as their basis. By fitting NURBS to actual patient data, the phantom is more realistic than those based on solid geometry. It is also flexible due to the ability of NURBS surfaces to be altered easily via affine and other transformations to model anatomical variations and patient motion.

Figure 4-1: (a) Anterior view of the 4D NCAT phantom. (b) Cardiac and respiratory motion models. (c) Simulations performed using the phantom. [Ref. NCAT]
The NCAT phantom was extended to four dimensions to model common patient motions such as the cardiac and respiratory motions using 4D tagged MRI data and 4D high-resolution respiratory-gated CT data respectively. Both datasets were acquired from normal patient volunteers. With its basis upon human data and the inherent flexibility of the NURBS primitives, the result is a computer-generated phantom that closely resembles the anatomical structures and cardiac and respiratory motions of a normal human subject. Combined with accurate models of the imaging process, the 4D NCAT is capable of simulating imaging data close to that of actual patients. For all these reasons, it is widely used in nuclear medicine imaging research.

Although it is capable of being far more realistic, the 4D NCAT phantom was originally designed for low-resolution nuclear medicine imaging research, and lacks the anatomical detail required for use in higher-resolution imaging modalities such as x-ray CT. At the same time, there is a lack of realistic and flexible computer-based phantoms for use in this area and the NCAT has the advantage, due to its design, that its organ shapes can be changed to realistically model different anatomical variations and patient motion. For this reason, the 4D NCAT phantom was chosen to be utilized to form the simulated CT image sequence used in the experimental process.

### 4.2.3 Formation of Observed LR images

The observation model relating the desired HR image \( f \) with the observed LR image \( g_k \), as proposed by Irani and Peleg [37], is the following:

\[
g_k = (T_k (f) * h) \downarrow s + n_k, \quad k = \{1 \ldots K\} \tag{4.1}
\]

where \( T_k \) is the geometrical transformation (rotation and/or translation) of the image \( f \) to the same reference frame of acquisition for \( g_k \), \( h \) is a blur kernel, often referred to as the point spread function (PSF), defined by the physical properties of the imaging device, \( n_k \) is additive noise, \( \downarrow s \) represents the down-sampling of a HR image to a LR image by a factor \( s \), and the symbol (*) is the convolution operator.

In order to form the observed images, we use a volumetric 3D matrix of size 64x64x64 generated by simulation with the NCAT phantom. The voxel size is 6.25x6.25x6.25mm. We form images in the transaxial plane and images along the axial direction. The transaxial plane runs from the front to the back of the body, dividing it into top and bottom parts, i.e. it is perpendicular to the bed direction. The axial direction is the bed direction or the slice direction if we consider the voxel matrix as a stack of 64 images (slices) of size 64x64. Images along the axial direction, i.e. in the coronal or sagittal planes, are generated by re-sampling the voxel matrix along these planes. The coronal plane divides the body into front and back portions and the sagittal plane into left and right portions.

![Figure 4-2: (a) Transaxial image. (b) Coronal image. (c) Sagittal image.](image-url)
The geometrical transformation in the axial case is implemented as a simple shift of multiples of \(1/s\) of a LR pixel, where \(s\) is the down-sampling factor. In this way, the basic premise for super resolution is achieved, since there is a spatial transformation known to sub-pixel accuracy between successive acquisitions. In the transaxial case, the transformation is a rotation with reference to one transaxial image (slice). In both cases the blur kernel \(h\) is considered as a delta function with width of a pixel (6.25mm) and the noise is ignored.

The sets of images generated are the following:

- **Axial direction**
  - 2 acquisitions 64x64x32
    The initial acquisition of size 64x64x64 is considered as the HR image \(f\) in the observation model (4.1). This acquisition is shifted by 1 pixel along the axial direction and then is down-sampled by 2, in order to achieve a shift of \(\frac{1}{2}\) of a LR axial pixel.
  - 4 acquisitions 64x64x16
    The initial acquisition is shifted by 1, 2 and 3 pixels along the axial direction and then is down-sampled by 4. Thus, we obtain a set of 4 acquisitions each shifted by multiples of \(\frac{1}{4}\) of a LR pixel, relative to the previous one.

**Figure 4-3:** Sets of images - axial and transaxial – generated in the phase of image simulation.
Transaxial direction

- 4 images 64x64

We choose one transaxial image as a reference image and we form 3 more images by performing rotations by 2.9737°, 3.8141° and 4.0856°. This can be performed for any slice. We prefer the slices around the middle of the sequence, which contain more information. The rotation is implemented as an affine transformation determined by two control points and the pixel values are assigned by bilinear interpolation. These control points are set in a default value (=1) in order to be detectable when the transformation will need to be inverted.

In all cases, a Gaussian filter with width of a pixel is applied to the images, corresponding to the blurring process.

The above sets of images represent the LR acquisitions from different points-of-view, which are necessary to apply the SR algorithm.

4.3 Phase 2: Application of Super Resolution

This phase of the experimental process involves the application of the SR method to the images generated in phase 1, including the implementation of the selected algorithm, the definition of the main parameters of the algorithm and the performance evaluation through certain measures.

4.3.1 Iterative Back-Projection Algorithm

As mentioned above, we have chosen the Iterative Back-Projection method, proposed by Irani and Peleg [37], as follows:

- Sets of LR images are available, representing different points-of-view of the same scene (Phase 1).
- These acquisitions are re-sampled in a HR form, mathematically shifted to a common reference frame, and averaged to form a HR initial guess of the desired image.
- This HR guess is shifted back to the reference frames of the initial points-of-view and re-sampled at a LR, to synthetically generate the LR results, provided the HR guess.
- The synthetically generated LR results are compared to the observed LR images of phase 1 and differences are calculated.
- The differences for each point-of-view are re-sampled in a HR form, shifted to a common reference frame, averaged, and used to update the HR guess.
- The process is repeated until the differences are minimized or until a maximum number of iterations is reached.

In Figure 4-4, a flowchart of the iterative back-projection method is illustrated, along with the mathematical formulas utilized, which will be explained briefly further on.
Step 1: Initial guess for the HR image

The initial guess for the HR image is taken as the average of the set of LR acquisitions brought to the same reference frame and up-sampled.

\[ f^{(0)} = \frac{1}{K} \sum_{k=1}^{K} T_k^{-1} (g_k \uparrow s) \]  

(4.2)

- \( f^{(0)} \) is the initial HR guess of the desired image,
- \( g_k \) is the LR observed image,
- \( T_k^{-1} \) is the inverse of the spatial transformation \( T_k \) of the observation model,
- \( K \) is the number of LR acquisitions available, and
- \( \uparrow s \) is the up-sampling operator from LR to HR.

Step 2: Synthetically generated LR images

The \( n \)th synthetically sampled set of LR images \( \{w_k^{(n)}\} \) is obtained from the \( n \)th approximation of the HR image \( f^{(0)} \), ignoring the noise.

\[ w_k^{(n)} = (T_k (f^{(0)})*h) \downarrow s \]

Step 3: Error calculation

\[ e_n = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left( \|g_k - w_k^{(n)}\| \right)^2} \]

is less than a threshold? OR

Number of iterations = \( \text{max} \)?

Step 4: Update the current HR guess

\[ f^{(n+1)} = f^{(n)} + \frac{1}{K} \sum_{k=1}^{K} T_k^{-1} \left( (g_k - w_k^{(n)}) \uparrow s \ast p \right) \]

Figure 4-4: Block diagram of the iterative back-projection method
\[ w_k^{(n)} = (T_k (f^{(n)}) * h) \downarrow s \]  
\[ e_n = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left( \|g_k - w_k^{(n)}\| \right)^2} \]

- \( w_k^{(n)} \) is one of the \( K \) images of the \( n^{th} \) synthetically sampled set of LR images,
- \( f^{(n)} \) is the \( n^{th} \) approximation of the HR image,
- \( T_k \) is the straightforward spatial transformation in the form that it was used in the observation model,
- \( \downarrow s \) is the down-sampling operator from HR to LR,
- \( h \) is the blur kernel, and
- \( (*) \) is the convolution operator.

**Step 3: Error calculation**

The error is calculated as the root mean square of the averaged sum of the Euclidean norm of the difference \((g_k - w_k^{(n)})\) for every set of acquisitions.

**Step 4: Update of current HR guess**

The current HR guess is corrected based on the difference between the two sets of LR images.

\[ f^{(n+1)} = f^{(n)} + \frac{1}{K} \sum_{k=1}^{K} T_k^{-1} \left( \left( g_k - w_k^{(n)} \right) \uparrow s \right) * p \]

- \( f^{(n+1)} \) is the updated HR guess,
- \( f^{(n)} \) is the current HR guess,
- \( T_k^{-1} \) is the inverse of the spatial transformation \( T_k \) of the observation model,
- \( g_k \) is the LR observed image,
- \( w_k^{(n)} \) is one of the \( K \) images of the \( n^{th} \) synthetically sampled set of LR images,
- \( \uparrow s \) is the up-sampling operator from LR to HR,
- \( p \) is the 'sharpening' kernel (ideally the inverse of the blur kernel)
- \( K \) is the number of LR acquisitions available, and
- \( (*) \) is the convolution operator.

### 4.3.2 Implementation Issues

For the implementation in the *axial direction*, we start with an initial guess for the HR image \( f^{(0)} \) by averaging \( K \) consecutive LR slices as given by (4.2). In this case, \( \uparrow s \) is implemented by splitting the axial pixels of the acquisitions \( g_k \) into \( K \) parts. The
spatial transformation $T_k^{-1}$ is the inverse of $T_k$ and represents the shift of the HR form of the acquisitions $(g_k \uparrow s)$ to a common HR reference frame. The spatial transformation $T_k$, as mentioned in §2.3, is implemented as a simple shift of multiples of $1/s$ of a LR pixel along the axial direction. The HR image is updated iteratively by using equations (4.3) and (4.5). In this case, $\downarrow s$ represents the down-sampling by averaging $s$ HR neighbouring pixels along the axial direction. The procedure is stopped when the error $e_n$, as given by (4.4), has reached a minimum, in this case when changes in $e_n$ are approximately 0.1% between two consecutive iterations.

For the implementation in the transaxial direction, we follow the same process as before. The up-sampling $\uparrow s$ represents the splitting of pixels of the acquisitions $g_k$ into $s$ parts, so that 1 pixel of size $m \times m$ is divided into $s$ pixels of size $(m/s) \times (m/s)$. The spatial transformation $T_k$, as mentioned in §4.2.3, is an affine transformation determined by two control points, which have been set to a default value. These points are detected and they are used to determine the inverse transformation $T_k^{-1}$, which represents the shift of the acquisitions $(g_k \uparrow s)$ to a common reference frame. Both the transformations are implemented in the HR reference frame, by rotating the image in-plane and assigning pixel values by bilinear interpolation. In this case, the process is stopped after a maximum number of iterations (=16) or when changes in $e_n$ are approximately 0.5% between two consecutive iterations.

The blur and sharpening kernels, $h$ and $p$, would remove any sharpening or smoothing of the images within the algorithm, but they are included just to keep the model more complete. In both cases, they are implemented as Gaussian filters with width of a pixel.

To aid the comparison between images, some image processing is applied to both sets of images, only for the transaxial case. First, the images are smoothed with a 6.25mm-FWHM Gaussian filter. The contrast is enhanced by trimming the bottom and top 1% pixel magnitudes and stretching the rest of the magnitudes over the available scale. A gamma correction ($\gamma = 1.5$) to the contrast is also applied.

### 4.3.3 Interleaving Method

A traditional alternative to SR reconstruction is interleaving. It is a method to achieve a HR image from a set of shifted LR images by combining the pixels one by one, from alternating LR image inputs, to generate a single large image.

In the axial case, the images reconstructed with SR are compared not only with the initial images without SR, but also with images produced by interleaving the available acquisitions. The voxels are divided into $K$ parts axially. Each part is assigned the value of the corresponding voxel obtained of each acquisition. Thus, each slice of the LR image is replaced by $K$ slices, each of which belongs to one of the LR acquisitions, as shown in Figure 4-5.

### 4.3.4 Super Resolution Parameters

Each SR reconstruction algorithm has certain key parameters to consider for each application scenario. Each parameter needs to be determined to match most closely with the true imaging system characteristics. Two key parameters are the transformation and blur parameters. In estimating the blur $h$, or the point spread function (PSF), the slice excitation profile can be utilized. In medical imaging systems, typical slice profiles are well approximated by Gaussian functions, where the FWHM is the originally selected slice width.
There are two commonly used PSFs:
- The “Box-PSF”, a rectangular pulse PSF, where the box width is taken as the selected slice width, and
- The “Gaussian-PSF”, with FWHM set to the selected slice width.

We consider the blur kernel $h$ as a Gaussian PSF with width of a pixel (slice thickness), since it provides a good approximation of the system characteristics for medical imaging. In the Irani-Peleg algorithm, an additional parameter is $p$, which is ideally the inverse of the blur kernel $h$.

In addition to the blur, another parameter that has an important effect on the performance of the SR algorithm is the transformation parameter. This parameter needs to be accounted for to enable precise image registration, accurate to a small fraction of a pixel, capable of bringing all input images to a common reference frame. The transformation has been implemented as an affine transformation determined by two control points. So it is necessary to have two special points marked in some way, which can be utilized to define the transformation. These points have been chosen to be set to a default value (=1). It is very important to detect these points accurately, so that the transformation $T_k^{-1}$, is determined correctly as the inverse of the initial transformation $T_k$. The relative shifts among the input acquisitions need to be known to subpixel accuracy, so that the acquisitions can be brought to the same reference frame. Incorrect motion estimation has disastrous implications on the performance of the SR reconstruction method.

### 4.3.5 Performance Measures for SR

The most commonly used methods to evaluate the results of image restoration is the Mean Square Error (MSE) and visual inspection by a human observer. The MSE, applicable only when the best obtainable solution - in this case the HR image - is available, is a questionable measure. A small MSE does not always correspond to a
better image quality. Visual inspection, on the other hand, is dependent not only on the viewers but also on the displays and viewing conditions. Thus, it is necessary to utilize a range of objective measures to evaluate the results of the SR reconstruction.

The main goal of SR is to improve the image resolution and to reduce noise and other artifacts. Therefore, the utilized measures need to express the achievement of these requirements and to offer a mean for comparison between images.

Quantitative measures for any image enhancement procedure are a challenge. In general, image enhancement can be expressed as the increase in frequency content as viewed via the image power spectrum. In order to further quantify the image quality, resolution is calculated as the PSF FWHM, which can be computed from an approximate point source. We use a Gaussian PSF with width of a pixel to model the point source. For both axial and transaxial cases, the images of the source are reconstructed as described previously, but no filtering or post-processing is applied. The FWHM values are calculated from the resulting distributions.

When considering a method for resolution improvement, it is important to make sure that the signal-to-noise ratio (SNR) is not compromised. The SNR is measured by computing the mean signal intensity over a certain high-intensity region of interest (ROI) and dividing this by the standard deviation of a region of noise outside the imaged object:

$$SNR = \frac{<\text{high intensity region of interest}>}{\sigma(\text{region outside imaged object})}$$  (4.6)

A variation of the above measure is the contrast ratio, which is defined as the ratio of the average signal over a certain region of interest to the average background:

$$C = \frac{<\text{Sig} >}{<\text{B} >}$$  (4.7)

It is very important to have a high enough contrast to be able to distinguish among different tissues and tissue types, and in particular between healthy and pathological tissue. Another measure of contrast is the contrast-to-noise ratio (CNR), which is defined as the absolute difference between the SNR of a certain region of interest and the SNR of the background:

$$CNR = |SNR(Sig) - SNR(B)|$$  (4.8)

Apart from the resolution and image quality measures, another aspect of the evaluation process is the estimation of the cost in processing time. Especially when we refer to medical imaging acquisitions, time is an important factor that needs to be considered. We calculate the acquisition time of the HR image for each application of the SR reconstruction method.

All the methods are processed in a system with CPU Intel Pentium IV 2.4GHz and 1GB RAM. The algorithms are implemented in Matlab 6.1 environment.
Chapter 5  Results and Interpretations

5.1 Summary of the experimental procedure

Image restoration belongs to the class of inverse problems. It aims to reconstruct the real underlying distribution of a physical quantity - the scene - from a set of measurements. In SR restoration, the scene is restored from a series of degraded LR images, each of which is a projection of the real and continuous scene onto an array of sensor elements. The algorithm uses a forward model which describes how the LR observations are obtained from the HR scene.

The experimental procedure for the evaluation of the SR reconstruction involves two phases: 1) the formation of the LR observations, and 2) the application of the SR algorithm, in order to obtain a HR image from the set of LR observations. For the LR image formation, the observation model proposed by Irani and Peleg [37] is utilized, without taking into consideration the presence of noise. The HR scene is a voxel image produced by using the NCAT phantom. The degradations performed to obtain the LR observations are translation, blurring and under-sampling. The selected SR reconstruction algorithm is the iterative back-projection method, as proposed in [37]. The algorithm is applied to improve resolution in the axial (bed) direction as well as the resolution in the transaxial plane.

This chapter is devoted to an evaluation of the performance of SR restoration. The results of the SR reconstruction are presented separately for the axial and the transaxial case. The evaluation relies on qualitative measures of image enhancement and on objective quantitative measures, such as the resolution (FWHM), the signal-to-noise ratio, the contrast ratio and the contrast-to-noise ratio, as they were defined in the previous chapter. These measurements are presented in tables for each case. To aid the evaluation process, the resulting images along with certain plots are also provided. In the final section of the chapter, the results are discussed for both cases.

5.2 Axial case

The SR reconstruction method is applied on the following two sets of acquisitions, as they have been simulated in the first phase of the experimental procedure, which is described in Chapter 4:

- 2 acquisitions of size 64x64x32
- 4 acquisitions of size 64x64x16

Each of these acquisitions is a stack of 2D images of size 64x64 perpendicular to the axial direction. Thus, images along the axial direction are generated by re-sampling the volumetric voxel matrix along the coronal or sagittal planes. Each simulated voxel matrix is shifted in the axial direction, compared to the previous one, by a fraction of a pixel. In Figure 5-1, coronal and sagittal images corresponding to each set of acquisitions are illustrated.
The SR reconstruction method is applied as described in §4.3.2 of Chapter 4. The process is repeated for 16 and 32 iterations. For the reconstruction from 2 acquisitions, after 16 iterations, the error change between two consecutive iterations is approximately 0.1%, while for the reconstruction from 4 acquisitions this threshold is reached after 32 iterations. In Figure 5-2, a plot of the error values for the reconstruction in the axial case according to the number of iterations, is illustrated.

The SR results are compared with the results of the reconstruction with interleaving. In order to make the comparisons, we resample the voxel matrix along the coronal plane and we calculate the SNR, contrast ratio and CNR over a certain region of interest of the resulting coronal image. To quantify the resolution, we calculate the PSF FWHM values, along the axial direction, for an approximate point source.
Table 5-1: Point spread function FWHM (mm) values (calculated for a Gaussian PSF along the axial direction)

<table>
<thead>
<tr>
<th></th>
<th>64x64x32 voxel matrix</th>
<th>64x64x16 voxel matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Super Resolution</strong></td>
<td>7.7302</td>
<td>7.4851-7.6494</td>
</tr>
<tr>
<td><strong>Interleaving</strong></td>
<td>6.4454</td>
<td>6.2503</td>
</tr>
<tr>
<td><strong>Super Resolution</strong></td>
<td>4.4358</td>
<td>4.4847</td>
</tr>
</tbody>
</table>

**Figure 5-3:** (a) Initial coronal image (32 slices). (b) Reconstructed image from 2 acquisitions interleaved. (c) Reconstructed image from 2 acquisitions with super resolution.

<table>
<thead>
<tr>
<th><strong>Reconstruction from 2 acquisitions</strong></th>
<th><strong>Acquisition time (sec)</strong></th>
<th><strong>SNR (dB)</strong></th>
<th><strong>Contrast ratio</strong></th>
<th><strong>CNR</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial acquisition</td>
<td>35.3419</td>
<td>0.1754</td>
<td>2.1692</td>
<td></td>
</tr>
<tr>
<td>Interleaving</td>
<td>0.0150</td>
<td>35.2606</td>
<td>0.1799</td>
<td>2.0939</td>
</tr>
<tr>
<td>16 iterations</td>
<td>4.9690</td>
<td>33.6251</td>
<td>0.1883</td>
<td>2.0656</td>
</tr>
<tr>
<td>32 iterations</td>
<td>9.8590</td>
<td>33.2580</td>
<td>0.1848</td>
<td>2.0710</td>
</tr>
</tbody>
</table>

**Table 5-2:** Acquisition time, SNR, contrast ratio and CNR of the SR reconstructed images from 2 acquisitions.

The result of the reconstruction is a voxel matrix of size 64x64x64. As shown in Table 5-1, up to 40% improvement of resolution is achieved with SR, which is better than interleaving.

The SR reconstructed image provides better contrast ratio that is up to 7.5% greater than the image without SR and 3.5% greater than the image produced from interleaved data (Table 5-2). The loss in SNR is a result of the trade-off between resolution and noise which is typical characteristic of reconstruction algorithms.
Figure 5-4: Plot of the pixel values along a column of a coronal image. (a) Selected column of the coronal image. Super resolution result from 2 acquisitions (solid) compared to (b) image without super resolution (dashed) and (c) reconstruction with 2 acquisitions interleaved (dashed).

In the case of the reconstruction from 4 acquisitions, the initial image is of low quality and details cannot be resolved. The SR reconstructed image provides a better contrast ratio that is approximately 15% greater than the initial image and almost equal to the image reconstructed from interleaved data (Table 5-3). Again, a loss of SNR can be noted, as a result of the trade-off between resolution and noise.

Comparative plots of pixel values through a selected column in the images are presented in Figures 5-4 and 5-6. The plot of the SR signal has sharper peaks, which indicates that SR is superior in resolving more details compared to the interleaving method. In Figures 5-3 and 5-5 coronal images of the initial and the reconstructed acquisitions are illustrated. It can be seen that the quality of the SR image is better than the reconstructed with interleaving image. The SR images have thinner image planes, $\frac{1}{2}$ of the initial slice thickness in the case of 2 acquisitions and $\frac{1}{4}$ in the case of 4 acquisitions.

The acquisition time for SR is greater than the time for interleaving, which is expected, since SR is an iterative algorithm, while interleaving is a one-step method. However the results of SR are equal to or better than the results of interleaving. In the most important aspect of reconstruction, resolution enhancement, SR provides much better results, which compensates for the cost in time.

Figure 5-5: (a) Coronal image (16 slices). (b) Reconstructed image from 4 acquisitions interleaved. (c) Reconstructed image from 4 acquisitions with super resolution.
<table>
<thead>
<tr>
<th>Reconstruction from 4 acquisitions</th>
<th>Acquisition time (sec)</th>
<th>SNR (dB)</th>
<th>Contrast ratio</th>
<th>CNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial acquisition</td>
<td>34.5507</td>
<td>0.1652</td>
<td>2.3041</td>
<td></td>
</tr>
<tr>
<td>Interleaving</td>
<td>0.0320</td>
<td>32.3403</td>
<td>0.1907</td>
<td>2.0797</td>
</tr>
<tr>
<td>16 iterations</td>
<td>10.0150</td>
<td>32.0241</td>
<td>0.1908</td>
<td>2.0754</td>
</tr>
<tr>
<td>32 iterations</td>
<td>19.9530</td>
<td>30.7190</td>
<td>0.1911</td>
<td>2.0736</td>
</tr>
</tbody>
</table>

**Table 5-3**: Acquisition time, SNR, contrast ratio and CNR of the SR reconstructed images from 4 acquisitions.

![Images](image1.png)

**Figure 5-6**: Plot of the pixel values along a column of a coronal image. (a) Selected column of the coronal image. Super resolution result from 4 acquisitions (solid) compared to (b) image without super resolution (dashed) and (c) reconstruction with 4 acquisitions interleaved (dashed).

### 5.3 Transaxial case

In the transaxial case, the SR reconstruction algorithm is applied on two sets of rotated images of size 64x64 generated as described in Chapter 4. A geometric transformation has been applied on the 2D image of a selected slice, which is determined by certain pixels on the image which have a default value. These pixels need to be correctly detected, so that the transformation can be defined and therefore to be reversed. In Figure 5-7, the images used as an input for the algorithm, are illustrated.
The algorithm is repeated until a maximum number of 16 iterations, when the error change between two consecutive iterations is approximately 6%. The error values for the two sets of inputs are shown in the plots of Figure 5-8. The SR reconstructed image is of size 128x128. The results are compared with the input images, in terms of SNR, contrast ratio and CNR, calculated over a certain region of interest. To quantify the resolution, we calculate the PSF FWHM values, along the horizontal and vertical axis, for an approximate point source, which is simulated by a Gaussian PSF (Table 5-4). The lower and upper values of the FWHM for the case without SR, refer to the cases where the point source is centered within a pixel and when it falls between two pixels. The results presented in the table show that SR provides better resolution in both axes. The improvement in resolution reaches up to 35%.
### Table 5-4: Point spread function FWHM (mm) values (calculated for a Gaussian PSF)

<table>
<thead>
<tr>
<th></th>
<th>FWHM(mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x-axis</td>
</tr>
<tr>
<td>No Super Resolution</td>
<td>8.3053 –11.2976</td>
</tr>
<tr>
<td>Super Resolution</td>
<td>7.4089</td>
</tr>
</tbody>
</table>

### Table 5-5: Acquisition time, SNR, contrast ratio and CNR of the SR reconstructed images for each set of input images.

<table>
<thead>
<tr>
<th></th>
<th>Acquisition time(sec)</th>
<th>SNR(dB)</th>
<th>Contrast Ratio</th>
<th>CNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Super Resolution reconstruction (Input set 1)</td>
<td>Input</td>
<td>36.4694</td>
<td>0.3048</td>
<td>1.6574</td>
</tr>
<tr>
<td></td>
<td>8 iterations</td>
<td>49.9070</td>
<td>32.7264</td>
<td>0.2901</td>
</tr>
<tr>
<td></td>
<td>16 iterations</td>
<td>91.9690</td>
<td>33.1776</td>
<td>0.3264</td>
</tr>
<tr>
<td>Super Resolution reconstruction (Input set 2)</td>
<td>Input</td>
<td>35.5187</td>
<td>0.1763</td>
<td>2.7204</td>
</tr>
<tr>
<td></td>
<td>8 iterations</td>
<td>53.4070</td>
<td>35.4291</td>
<td>0.2061</td>
</tr>
<tr>
<td></td>
<td>16 iterations</td>
<td>94.5000</td>
<td>34.9823</td>
<td>0.1911</td>
</tr>
</tbody>
</table>

As shown in Table 5-5, the SR reconstructed image provides up to 7 - 8% better contrast ratio, in comparison to the input image. In the case of the input set 1, there is also an improvement in the CNR. The loss in SNR, as a typical characteristic of reconstruction methods, is noted in the transaxial case too, but the SNR is maintained in a satisfactory level. The acquisition time for the transaxial case is greater than the corresponding value for the axial case. This results from the additional complexity of the algorithm, referring to the detection and inversion of the geometrical transformations. Because of this cost in time, the maximum number of iterations is set to be 16, since there is no significant improvement in the results for further iterations.

In Figures 5-9 and 5-11 the input and the resulting high-resolution images are illustrated. The comparative plots of Figures 5-10 and 5-12, indicate that SR improves the resolution of details. This is more evident in the plot of Figure 5-12, where the peaks of the SR signal are sharper and some do not exist in the input signal.
Figure 5-9: (a) Input transaxial image. (b) SR reconstructed image.

Figure 5-10: Plot of pixel values along a column of the HR image in comparison to the input. (a) Selected column. (b) Super resolution result (solid) compared to image without super resolution (dashed).

Figure 5-11: (a) Input transaxial image. (b) SR reconstructed image.
Figure 5-12: Plot of pixel values along a row of the HR image in comparison to the input. (a) Selected column. (b) Super resolution result (solid) compared to image without super resolution (dashed).

As mentioned in Chapter 4, image enhancement can be expressed as the increase in frequency content as viewed via the image power spectrum. In Figure 5-11, the power spectrum of the input and the reconstructed image for the input set 1, are illustrated.

Figure 5-13: Power spectrum of the (a) input image and (b) SR reconstructed image

5.4 Discussion

Higher resolution of medical images, especially in the case of PET and MRI images, is of potential value and may have several implications in research and clinical practice. One example is the study of cerebral metabolism. Since brain metabolism is indicative of regional cerebral function, PET imaging provides an excellent map of regional cerebral function. However, PET studies are limited by the spatial resolution of the presently available instrumentation [56]. The imaging of small cerebral structures such as the cortical sub-layers and nuclei may need small spatial resolutions [56].

Cancer lesions should be of diameters equal or larger than the resolution of the PET scanner to be identified provided they also have a high glucose metabolism. Lesions smaller than the PET resolution, may also be detected if the uptake is sufficiently high to overcome the reduction in signal due to partial volume effects. In either case, higher PET resolution would be beneficial for improving sensitivity for
detection of small tumours [57]. Functional PET imaging provides suboptimal
topographic orientation due to the noise inherent to the method. The relatively low
anatomic resolution of PET images may be further degraded in regions with relatively
little differences in uptake between adjacent tissue types [58]. Higher resolution PET
images provided by a super-resolution algorithm may show a more differentiated
anatomical structure. The increase in anatomical detail in the functional PET image
may aid in the registration of the image with a corresponding anatomical image from
another modality such as MRI or CT. This would especially be of significance in
scanners which are not dual-modality.

In MRI, resolution enhancement is important for visualization and early diagnosis.
In certain diagnostic MR imaging applications, isotropic resolution is necessary. So
the main goal is to increase the resolution in the slice-select dimension so as to
achieve high-resolution isotropic, 3D images. In fMRI, the increase in spatial
resolution would help detect and visualize small regions of neuronal activity [37].
Moreover, the activated regions appear sharper and provide better information
regarding the morphological limits and structure.

The super-resolution technique provides a method of approaching these
resolution goals. The performed trials demonstrated an improvement in the axial
resolution, which reached up to 40%, in comparison to the input image. The SR
images can resolve smaller features and more details, since they provide a
significantly improved contrast ratio. In the axial case, the super-resolution method
was compared to the interleaving reconstruction method and was proved to provide
better results. The improvements in the axial resolution can be applied without
hardware changes and without increasing the patient radiation exposure. For the
transaxial rotational and translational shifts, there was no equivalent to interleaving
for reconstruction. In this case, the resolution was improved by 35%. The loss in SNR
which was expected due to the trade-off between resolution and noise was not that
considerable to preclude the clinical application of super-resolution.

Super-resolution techniques require the geometric shifts among the original LR
images to be known with subpixel accuracy. Therefore, brain scans can be more
suitable for the clinical application of super-resolution since there is less likely to be
motion in the head, than other parts of the body since it is usually fixed in a cradle.
Secondly, the typical amount of filtering on clinical brain scan images (2 mm FWHM
Gaussian filter) is considerably less than that applied to whole body scan images, so
the improvements provided by the super-resolution algorithm are more likely to be
preserved.

A critical issue for the transaxial implementation of the super-resolution technique
is the detection of the geometrical shifts among the original low resolution images.
Incorrect motion estimation can have disastrous implications on the overall
performance of the method. This dependency on the accuracy of the motion
detection can be considered as a drawback of the super-resolution method. Even if
the algorithm is not applied in the directions of the transaxial plane, the transaxial
images may also be improved because of the effective thinner slice thickness
generated by the super-resolution. The low-resolution transaxial images can be
conceived as the sum of four “high-resolution” slices, degrading features that are not
common to all four slices (partial volume effects). The thinner effective axial slice
thickness reconstructed by the super-resolution algorithm reduces partial volume
effects induced by this summation in the axial direction, providing improved image
quality.

Improvements of medical images due to the application of super-resolution
methods come at the cost of increased computer processing time. The cost in time is
higher in the transaxial case, due to the need for correct motion estimation. Any
future increases in computer processing speeds will further reduce these time
limitations.
Chapter 6  Summary and Conclusions

6.1 Introduction

Recent works using super-resolution in medical applications have demonstrated that using super-resolution technology enables to effectively break the limits on slice thickness as posed by the physical properties of existing imaging hardware. Each imaging system has a characteristic resolution which is determined based on physical constraints of the detectors, which are reflected to the signal-to-noise ratio (SNR) and the timing considerations in the system. Within each imaging modality specific physical laws are in control, defining the meaning of noise and the sensitivity of the imaging process. In each case there are signal processing rules, which are applied in the system design in an attempt to achieve an acceptable compromise between resolution and SNR.

Higher resolution in the image plane usually means acquisition with a smaller sampling distance, by using a smaller detector (in the case of PET) or by using higher magnetic field scanners and shorter sampling distances (in the case of MRI). Due to physical constraints, high resolution image acquisition results in a lower SNR, i.e. a trade-off exists between resolution and SNR. One of the key features in super-resolution technology is the ability to obtain a high resolution image with almost the same SNR as the original low resolution images from which it is constructed.

In order to improve resolution, researchers of imaging sensors have been endeavouring to overcome the limitations based on device physics and circuit technology. Super-resolution image reconstruction is an alternate but efficient approach, based on signal processing technology. It is one of the most prominent research areas in the field of resolution enhancement, since it can overcome the inherent resolution limitation of the imaging system and improve the performance of most imaging applications.

6.2 Application of super-resolution

Super-resolution is the process of combining multiple low resolution images to form a high resolution image. The major advantage of the approach in the signal processing field is that it has lower cost and allows the use of the existing LR imaging systems. The basic requirement in order to apply super-resolution restoration techniques is the availability of multiple LR images captured from the same scene. These LR images represent different "looks" at the same scene. The LR images are assumed to suffer from blur, noise, and aliasing effects. The observation model which relates the HR image to the observed LR images can be presented by the following equation:

\[ g_k = (T_k f * h) \downarrow s + n_k, \quad k = \{1 \ldots K\} \]  

where:
- \( T_k \) is the geometrical transformation (rotation and/or translation) of the image \( f \) to the same reference frame of acquisition for \( g_k \).
• $h$ is a blur kernel, often referred to as the point spread function (PSF), defined by the physical properties of the imaging device,
• $n_k$ is additive noise,
• $\downarrow$ represents the down-sampling of a HR image to a LR image by factors $L_1$ horizontally and $L_2$ vertically, and
• $(\ast)$ is the convolution operator.

Based on the above model, the aim of the super-resolution image reconstruction is to estimate the HR image $f$ from the LR images $g_k$, for $k=1,\ldots,K$. The index $k$ refers to acquisitions at different points of view.

Most of the super-resolution image reconstruction methods proposed in the literature consists of three stages: registration, interpolation and restoration (i.e., inverse procedure).

We presented a performance evaluation of the super-resolution reconstruction applied in medical images. The method chosen to be utilized belongs to the class of algorithms which have in common a 'simulate and correct' approach to reconstruction. It is an approach based on the iterative back-projection (IBP) method adopted from computer aided tomography (CAT), which was suggested by Irani and Peleg. In this approach, the HR image is estimated by back-projecting the error (difference) between synthetically generated LR images and the observed LR images. This process is repeated iteratively to minimize the energy of the error. This method is not limited to specific motion characteristics and allows arbitrary smooth motion flow. The advantage of the IBP is that it is understood intuitively and easily.

In the first phase of the experimental procedure, simulated images of a computer generated phantom are formed and processed in order to comply with the observation model for the LR images. These images represent different points-of-view of the same scene. Then the IBP algorithm is applied, as follows:

- The LR acquisitions are re-sampled in a HR form, mathematically shifted to a common reference frame, and averaged to form a HR initial guess of the desired image.
- This HR guess is shifted back to the reference frames of the initial points-of-view and re-sampled at a LR, to synthetically generate the LR results, provided the HR guess.
- The synthetically generated LR results are compared to the observed LR images of phase 1 and differences are calculated.
- The differences for each point-of-view are re-sampled in a HR form, shifted to a common reference frame, averaged, and used to update the HR guess.
- The process is repeated until the differences are minimized or until a maximum number of iterations has been reached.

Each super-resolution reconstruction algorithm has certain key parameters to consider for each application scenario. These parameters are critical for the overall performance of the method, so they need to be determined to match most closely with the true imaging system characteristics. Two key parameters are the transformation and blur parameters. We consider the blur kernel as a Gaussian PSF with width of a pixel (slice thickness), since it provides a good approximation of the system characteristics for medical imaging. However, in most true imaging systems the blurring process is hard to be estimated.

In addition to the blur, another parameter that has an important effect on the performance of the super-resolution algorithm is the transformation parameter. This parameter needs to be accounted for to enable precise image registration, accurate to a small fraction of a pixel, capable of bringing all input images to a common reference frame. The transformation has been implemented as an affine transformation determined by two control points. So it is necessary to have two special points marked in some way, which can be utilized to define the transformation. The relative shifts among the input acquisitions need to be known to
sub-pixel accuracy, so that the acquisitions can be brought to the same reference frame. Incorrect motion estimation has disastrous implications on the performance of the super-resolution reconstruction method. This need for accurate motion estimation has also a significant effect on the total processing time of the reconstruction method.

The super-resolution algorithm is applied in corresponding sets of LR images, to improve the resolution in the axial direction as well as the transaxial plane. The evaluation relies on qualitative measures of image enhancement and on objective quantitative measures, such as the resolution (as expressed through FWHM), the signal-to-noise ratio, the contrast ratio and the contrast-to-noise ratio.

6.3 Super-resolution performance

The performed trials demonstrated improvement in both the axial and transaxial resolution. The axial resolution was improved by up to 40% compared to the input image and by 30% compared to the interleaving reconstruction. The super-resolution images also provide a significantly improved contrast ratio. In the case of reconstruction in the transaxial plane, the resolution was improved by 35%. The improvement in resolution can be achieved without using any hardware changes or any increase in the patient radiation procedure. An important contribution of super-resolution is the reduction of partial volume effects in the reconstructed image. In this respect, even if a super-resolution algorithm is applied in a single (axial) dimension, it in effect contributes to the transaxial resolution as well, due to the reduction of partial volume effects. The loss in SNR, which is a typical characteristic of all resolution enhancement algorithms, was not that considerable to preclude the clinical application of super-resolution.

The thinner effective axial slice thickness generated by the super-resolution in the axial direction reduces partial volume effects providing improved image quality, even in the transaxial plane. The increase in contrast ratio is important for improving sensitivity for detection of small details and features, such as small tumours. A higher resolution medical image provided by a super-resolution algorithm may show a more differentiated anatomical structure, which may aid in the registration of the image with a corresponding anatomical image from another modality. This would especially be of significance in PET scanners which are not dual-modality and additional CT scans are required.

The image quality improvement due to the application of super-resolution methods comes at the cost of increased computer processing time. The cost in time is higher in the transaxial case, due to the need for correct motion estimation. The geometrical shifts among the LR initial images need to be known with sub-pixel precision, so there is an additional computational cost in this case.

6.4 Additional issues and future challenges

6.4.1 Spiral CT

An important question to address is the applicability of super-resolution to a given medical imaging system. It is important to note that super-resolution can augment the resolution as acquired by the system detectors, in cases in which the detectors have under-sampled the input data, i.e. high frequencies exist in the signal that reaches
the detectors and the detectors sampling limit leads to aliasing and degradation in the high spatial frequency content, as output in the reconstructed image. Super-resolution reconstructs the aliased high frequency information, thus providing a higher resolution output and minimizing the aliasing problem. In cases in which no frequencies exist that are higher than half of the detectors sampling frequency, no additional improvements in the image resolution can be obtained by super-resolution technology. Such is the case in spiral CT systems, where the main obstacle for achieving increased resolution is the relatively large PSF of the system.

6.4.2 Newly emerging hardware

In the MRI field, new parallel imaging techniques are currently being developed. Such techniques will allow faster acquisition and higher in-plane resolution. In many of the developed techniques the added resolution is at the expense of SNR. The ability to use super-resolution post processing of thick slices may provide the boost needed for the SNR. Novel encoding methodologies, such as Non-Fourier methods (e.g. hadamard wavelets) are starting to emerge in MRI for encoding the third dimension [59]. Such technologies may enable the utilization of true 3D super-resolution techniques.

6.4.3 Registration issues

In the trials presented patient motion was not taken into consideration. For the more general case of a moving subject or organ, accurate image registration methods need to be incorporated.

6.4.4 Future work

Future work can advance the topic in two main directions:
- On the clinical front, by looking for the applications that gain most from the super-resolution technology.
- On the algorithmic front, by extending the investigation into additional medical imaging modalities, as well as comparing between super-resolution algorithms to find the advantages of each in a selected modality.
Appendix A

**Axial case**

- function [varargout] = *model*(img)

  % [im1,im2,...] = model(img)
  % Input: initial HR image
  % Output: LR images shifted by multiples of 1/(#output arguments) of a LR pixel along the axial direction

- function f0 = *HR_guess*(varargin)

  % f0 = HR_guess(im1,im2,...)
  % Calculates a HR initial guess f0 by averaging as many consecutive slices as the number of LR images available (im1,im2,...)

- function [varargout] = *synthetic_LR*(HR)

  % [g1,g2,...] = synthetic_LR(f)
  % Forms LR images by shifting along the axial direction the downsampled HR image
  % Input: HR image (f)
  % Output: LR simulated images (g1,g2,...)

- function updated_HR = *HR_correction*(HR,varargin)

  % f_new = HR_correction(f,im1,im2,...,g1,g2,...)
  % Updates the HR guess (f) according to the difference between the LR original images (im1,im2,...) and the LR simulated images (g1,g2,...)

- function [HR,e,time] = *repeat*(varargin)

  % [f,e,time] = repeat(im1,im2,...,n_iterations)
  % Repeats super resolution algorithm for n_iterations and stops when the error minimizes
  % Input: LR initial images (im1,im2,...), number of iterations
  % Output: HR final image (f), error vector (e), total acquisition time (time)
Transaxial case

- function [varargout] = transform(img,varargin)
  % [initial_im,im1,im2,...] = transform(img,input_points,
  % base_points1,...,marker_value)
  % Produces LR rotated images from the input image(img)
  % Output: rotated images produced by rotation of the input
  % image(img)
  % Input_points: M-by-2 double matrix containing the X and Y
  % coordinates of control points in the image to be transformed
  % Base_points: M-by-2 double matrix containing the X and Y
  % coordinates of control points in the base image
  % Marker_value: value of the pixels corresponding to the
  % control points

- function [varargout] = initial_guess(varargin)
  % f0 = initial_guess(im,im1,...,base_marker,marker1,...)
  % Input: LR images(im,im1,...),markers (control points)
  % detected in the LR images(base_marker,marker1,...)
  % Returns a HR initial guess f0 by upsampling & averaging the
  % available LR images

- function [varargout] = synthLR(f,varargin)
  % [g,g1,...] = synthLR(f,input_points,marker1,...)
  % Input: HR image(f),control points used to determine the
  % rotation of the LR images(input_points,marker1,...)
  % Output: LR simulated set of images (g,g1,...)

- function f_new = updateHR(f,varargin)
  % f_new = updateHR(f,im,im1,...,g,g1,..., input_points,
  % marker1,...)
  % Input: previous HR guess(f),initial LR images(im,iml...),
  % simulated LR images(g,g1,...), control points used to
  % determine the rotation of the LR images (input_points,
  % marker1,...)
  % Output: updated HR guess(f_new) based on the differences
  % between the input and the simulated images

- function [f,e,time] = super_r(varargin)
  % [f,e,time] = super_r(im,im1,...,input_points,ntimes)
  % Repeats the super resolution algorithm for ntimes or until
  % the error is minimized
  % Input: initial LR images(im,iml,...), control points in the
  % base LR image(input_points), number of iterations(ntimes)
% Output: HR guess(f), the error vector(e), and the execution time(time)

> function err = calc_error(varargin)

% err = calc_error(im1,im2,...,g1,g2,...)
% Input: LR initial images (im1,im2,...), LR simulated images (g1,g2,...)
% Output: root mean square of Euclidean norm of the difference between the two sets of images
% Used in both the axial and transaxial case
References


