Linköping Studies in Science and Technology
Dissertations, No. 1140

Advanced MRI Data Processing

Joakim Rydell

Linköping University
INSTITUTE OF TECHNOLOGY

Department of Biomedical Engineering
Linköpings universitet, SE-581 85 Linköping, Sweden
http://www.imt.liu.se/

Linköping, November 2007

Linköping Studies in Science and Technology
Dissertations, No. 1140

Advanced MRI Data Processing

Joakim Rydell

Department of Biomedical Engineering
Linköpings universitet
SE-581 85 Linköping
Sweden

ISBN 978-91-85895-59-5 ISSN 0345-7524

Printed in Linköping, Sweden by UniTryck AB 2007

Abstract

Magnetic resonance imaging (MRI) is a very versatile imaging modality which can be used to acquire several different types of images. Some examples include anatomical images, images showing local brain activation and images depicting different types of pathologies. Brain activation is detected by means of functional magnetic resonance imaging (fMRI). This is useful e.g. in planning of neurosurgical procedures and in neurological research. To find the activated regions, a sequence of images of the brain is collected while a patient or subject alters between resting and performing a task. The variations in image intensity over time are then compared to a model of the variations expected to be found in active parts of the brain. Locations with high correlation between the intensity variations and the model are considered to be activated by the task.

Since the images are very noisy, spatial filtering is needed before the activation can be detected. If adaptive filtering is used, i.e. if the filter at each location is adapted to the local neighborhood, very good detection performance can be obtained. This thesis presents two methods for adaptive spatial filtering of fMRI data. One of these is a modification of a previously proposed method, which at each position maximizes the similarity between the filter response and the model. A novel feature of the presented method is rotational invariance, i.e. equal sensitivity to activated regions in different orientations. The other method is based on bilateral filtering. At each position, this method averages pixels which are located in the same type of brain tissue and have similar intensity variation over time.

A method for robust correlation estimation is also presented. This method automatically detects local bursts of noise in a signal and disregards the corresponding signal segments when the correlation is estimated. Hence, the correlation estimate is not affected by the noise bursts. This method is useful not only in analysis of fMRI data, but also in other applications where correlation is used to determine the similarity between signals.

Finally, a method for correcting artifacts in complex MR images is presented. Complex images are used e.g. in the Dixon technique for separate imaging of water and fat. The phase of these images is often affected by artifacts and therefore need correction before the actual water and fat images can be calculated. The presented method for phase correction is based on an image integration technique known as the inverse gradient. The method is shown to provide good results even when applied to images with severe artifacts.

Populärvetenskaplig sammanfattning

Funktionell magnetresonanstomografi (fMRI) är en metod för att detektera och lokalisera de områden i hjärnan som aktiveras när en viss uppgift utförs. Metoden används dels vid planering av kirurgiska ingrepp och dels för att undersöka hjärnans funktion. För att hitta de områden som aktiveras, analyseras sekvenser av MR-bilder som samlats in under tiden uppgiften utförts. Genom att förbättra analysen kan man få högre kvalitet på de kartor som visar aktiva områden. Samtidigt möjliggör en bättre analys kortare undersökningstider. Denna avhandling presenterar tre olika metoder för analys av fMRI-bilder. Gemensamt för de olika metoderna är att de analyserar bilderna *adaptivt*, vilket medför att aktiverade områden kan detekteras med hög noggrannhet.

Magnetresonanstomografi används även för att avbilda anatomiska strukturer. Till exempel kan mängden fett i olika delar av kroppen mätas genom analys av MR-bilder av en viss typ. När dessa bilder samlas in, uppstår ofta fel som måste korrigeras innan det går att skilja mellan fett och vatten. I avhandlingen presenteras en korrigeringsmetod som gör det möjligt att hitta fettstrukturer även när det finns stora störningar i de insamlade bilderna.

Acknowledgements

A number of people have contributed to this dissertation. Some have been involved in the actual work, while others have contributed indirectly by just being there for me during the process. All contributions have been very valuable, and you all deserve my deepest gratitude.

First of all, I would like to thank my supervisors, Associate professor Magnus Borga and Professor Hans Knutsson. You have provided me with a never-ending stream of ideas and insights, and always supported me in my research. Advisors like you make life as a PhD student both fun and enlightening!

Dr. Mats Andersson also deserves special acknowledgement for his valuable input, mostly regarding filtering issues. I would also like to thank Dr. Peter Lundberg for his help and support in all things related to MR.

My fellow PhD students in the Medical imaging group, Anders Brun, Johanna Pettersson and Björn Svensson, all deserve to be mentioned here. Thanks to you, I have had a lot of fun both at work and otherwise.

Furthermore, I want to thank my parents, Lisen and Peter Rydell, for always supporting me in all possible ways. My friends outside the department have also made a highly valuable contribution by keeping my mind occupied with other things than work. Special thanks go to Maria Axelsson for long talks about research and everything else, and to Samuel Flodeby for lending me his brain for the cover illustration and for experiments.

Finally, I would like to acknowledge the Swedish Research Council, SMIV and the MOVIII project, whose financial support has made this work possible.

Thank you!

Table of Contents

1
Introduction

Around 300 B.C. Greek physicians Herophilos and Erasistratus performed the first anatomical research based on dissections of human cadavers. As a result of their research they obtained a more informed view of the human body than was earlier available. Dissection remained the only method for gaining knowledge about the human anatomy until the end of the 19th century, when the first X-ray photograph, depicting a human hand, was taken. Just a few years later, X-ray was used for medical purposes. Since then, the X-ray technology has undergone tremendous development, and today it is used in computed tomography (CT) to acquire three-dimensional images of e.g. the human body. Alternative techniques which do not utilize the potentially dangerous ionizing radiation in X-rays are also in widespread use. The most prevalent of these is magnetic resonance imaging (MRI), which uses magnetic fields and radio frequent radiation to acquire images of the body. CT and MRI each have their advantages and disadvantages. For example, while CT provides images of higher resolution, MRI is better suited to distinguish between different types of soft tissue, i.e. fat, water and gray and white matter in the brain.

MRI is a very versatile modality and can be used to obtain information about a large number of different anatomical and physiological parameters. Images showing pathologies, different tissue properties, and even brain activity and blood flow velocity can be obtained. Obviously some of these properties can not be directly measured. Instead, they are obtained by post-processing and analyzing images or image sequences acquired using MRI. The subject of this thesis is processing and analysis of MR images for detection of brain activity and for separate imaging of fat and water.

1.1 Detection of brain activity

Even earlier than the first dissections, already about 400 B.C, Hippocrates stated that the brain is the seat of intelligence. Since then, the function of the human

brain has attracted much interest. A major breakthrough occurred when tools such as positron emission tomography (PET), single photon emission computed tomography (SPECT) and functional magnetic resonance imaging (fMRI) made it possible to create images of the functional regions of the brain. Using these methods, it is possible to see what parts of the brain are activated when a certain task is carried out. Earlier, this could only be studied by observing patients with brain lesions, drawing conclusions from the damaged regions of their brains and the neural functions that seemed to be affected. The new possibilities provided by functional mapping techniques have facilitated many aspects of neurological research, and have shed light on several issues within the field. Another important application of functional imaging is surgical planning, i.e. preoperative planning of brain surgery, for example when removing brain tumors. By knowing where important neural functions are located, it is possible to minimize the damage to the patient's brain during surgery.

The most important method of those mentioned above is fMRI, since it provides functional maps with high resolution without the need for possibly exogenous contrast agents. In contrast, SPECT and PET rely on the use of radioactive substances, which are inhaled, ingested or injected into the bloodstream of the patient to be examined. This makes fMRI the most viable alternative for repeated examinations and for examinations of healthy volunteers.

An fMRI experiment generates a large amount of data, which needs to be analyzed in order to obtain a map of the activated regions of the brain. Several analysis methods have been proposed. Most of these are based on searching for regions where the data matches a template which depends on the experiment. However, the result of the experiment is heavily dependent on how the data is analyzed. The aim of the fMRI-related work described in this thesis is to develop an analysis method that finds active regions with the best possible accuracy.

1.2 Separation of fat and water

In order to visualize active brain regions, they are usually overlaid on an image showing the anatomical structure of the brain. This provides a means of localization, thus aiding the interpretation of the results. When acquiring the anatomical images, the scanner parameters are usually selected to provide contrast between different types of tissue in the brain, such as gray matter, white matter and cerebrospinal fluid. In other types of examinations, high contrast is instead desired between other tissue types. In e.g. obesity studies, MR imaging is used to distinguish between water and fat.

Obesity is becoming an increasingly large problem, commonly causing cardiovascular problems, diabetes and osteoarthritis. It is generally considered that abdominal fat accumulation is particularly dangerous. By measuring the abdominal fat,

it is possible to predict the risk of these and other diseases. Several MRI protocols for measuring intra-abdominal fat exist. The simplest methods classify each pixel as either belonging to fatty or water-based tissue. This, however, has the disadvantage that the fat content is almost always underestimated. A more advanced method, known as the Dixon technique, utilizes that the MR signals from fat and water have slightly different frequencies. This difference can be used to distinguish between the two types of tissue. However, the signal difference of interest is very small, and artifacts from the acquisition process make it difficult to detect. These artifacts need to be corrected before an accurate estimate of the fat accumulation can be obtained. A possible solution to this problem is presented in chapter 8, where phase sensitive reconstruction of MR images is discussed.

1.3 How to read this thesis

The thesis is made up of two main tracks: functional MRI and phase sensitive reconstruction. Depending on the reader, one or both of these tracks may be of interest. The next two chapters provide the background of magnetic resonance (chapter 2) and magnetic resonance imaging (chapter 3) and are relevant for both tracks. Chapter 4 describes the background of functional MRI and the following chapters (5, 6 and 7) each describes one method for fMRI analysis. Chapter 8 outlines the background of fat/water segmentation and phase sensitive reconstruction and describes a method for analysis of such data. Chapter 9 contains a review of included papers, which are available in the second part of the thesis. Finally chapter 10 presents a discussion and some pointers to possible further work.

1.4 Contributions

Naturally, not everything discussed in this thesis is the result of my own work. The background chapters, for example, present a brief summary of results from countless man-years of research in the fields of magnetic resonance and anatomical and functional imaging. Chapters 5 – 8, however, are largely based on my research.

In chapter 5, a method for analysis of fMRI data is presented and its invariance to the orientation of active regions is discussed. A modified method is introduced and shown to provide better invariance properties. The original analysis method is the result of work presented in Friman et al. (2003), while the discussion of rotational invariance and the modified algorithm are based on my research.

Chapter 6 presents another method for fMRI data analysis, based on bilateral filtering. While bilateral filtering per se is well known in the signal processing community, the presented application to functional MRI data is a result of my research.

Chapter 7 presents a method for robust estimation of correlation coefficients. The discussed statistical methods (the general linear model and canonical correlation analysis) are well-known, while the robust estimation technique is based on my own work.

Finally, chapter 8 presents an algorithm for phase reconstruction in complex MR images. The method is presented in the context of Dixon imaging (Dixon, 1984), and utilizes the inverse gradient operation (Farnebäck et al., 2007). The presented method for phase reconstruction based on iterative application of the inverse gradient is a result of my research, while Dixon imaging and the inverse gradient itself are not.

1.5 Publications

Four papers are included in the second part of the thesis. Chapter 9 describes these in more detail.

- Joakim Rydell, Hans Knutsson and Magnus Borga. *On rotational invariance in adaptive spatial filtering of fMRI data.* NeuroImage 2006;30(1):144-150.
- Magnus Borga and Joakim Rydell. *Signal and anatomical constraints in adaptive filtering of fMRI data.* IEEE International Symposium on Biomedical Imaging (ISBI). Arlington, Virginia, USA. 2007.
- Joakim Rydell, Magnus Borga and Hans Knutsson. *Robust correlation analysis with an application to functional MRI.* Submitted manuscript.
- Joakim Rydell, Hans Knutsson, Johanna Pettersson, Andreas Johansson, Gunnar Farnebäck, Olof Dahlqvist, Peter Lundberg, Fredrik Nyström and Magnus Borga. *Phase sensitive reconstruction for water/fat separation in MR imaging using inverse gradient.* International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). Brisbane, Australia. 2007.

The following papers are also related to my research, but are not included in the thesis:

- Gunnar Farnebäck, Joakim Rydell, Tino Ebbers, Mats Andersson and Hans Knutsson. *Efficient computation of the inverse gradient on irregular domains.* IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA). Rio de Janeiro, Brazil. 2007.
- Joakim Rydell, Hans Knutsson and Magnus Borga. *Adaptive fMRI data filtering based on tissue and signal similarities.* Joint Annual Meeting ISMRM-ESMRMB. Berlin, Germany. 2007.

- Joakim Rydell, Hans Knutsson and Magnus Borga. *Rotational invariance in adaptive fMRI data analysis.* IEEE International Conference on Image Processing (ICIP). Atlanta, Georgia, USA. 2006.
- Joakim Rydell, Hans Knutsson and Magnus Borga. *Tissue-selective adaptive filtering of fMRI data.* ESMRMB. Warsaw, Poland. 2006.
- Joakim Rydell, Hans Knutsson and Magnus Borga. *Adaptive filtering of fMRI data based on correlation and BOLD response similarity.* IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Toulouse, France. 2006.
- Joakim Rydell, Hans Knutsson and Magnus Borga. *Correlation controlled adaptive filtering for fMRI data analysis.* Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC). Umeå, Sweden. 2005.
- Joakim Rydell, Magnus Borga and Hans Knutsson. *Correlation controlled bilateral filtering of fMRI data.* ISMRM. Miami, USA. 2005.
- Joakim Rydell, Magnus Borga, Peter Lundberg and Hans Knutsson. *Dimensionality and degrees of freedom in fMRI data analysis – a comparative study.* IEEE International Symposium on Biomedical Imaging (ISBI). Arlington, Virginia, USA. 2004.

1.6 Notation

Inevitably, some mathematics will appear in a text on this topic. The notation conventions used in this thesis are summarized below.

s	Scalar
$\mathbf{v}$	Vector
$\|\mathbf{v}\|$	Norm of vector
$\hat{\mathbf{v}}$	Vector of unit norm
$\mathbf{M}$	Matrix
$\mathbf{M}^T$	Transpose of vector or matrix
$x * y$	Convolution of the signals x and y
$E[x]$	Expectation value of a random variable
$Var[x]$	Variance of a random variable
$Cov[x, y]$	Covariance between two random variables
$Corr[x, y]$	Correlation between two random variables

1.7 Abbreviations

This table lists some of the abbreviations used in this thesis along with their meanings.

MR	Magnetic resonance
MRI	Magnetic resonance imaging
fMRI	Functional magnetic resonance imaging
BOLD	Blood oxygen level dependent
SNR	Signal to noise ratio
GLM	General linear model
PCA	Principal component analysis
CCA	Canonical correlation analysis
RCCA	Restricted canonical correlation analysis

2

Magnetic resonance

This chapter describes how the chemical properties of a sample can be measured using nuclear magnetic resonance (NMR[1]). While this background is not strictly necessary to understand the signal processing topics in later chapters, knowledge of the signal and how it is acquired may provide some additional insight. NMR is a very complex topic, and a complete description of all details requires extensive knowledge of quantum mechanics. Therefore, this presentation describes a simplified model of the actual process. This model is, however, sufficient to understand the remainder of this thesis. For further information, the reader is directed to the very thorough explanation of spin physics available in Levitt (2001).

2.1 Properties of elementary particles

An atom consists of three types of elementary particles: protons, neutrons and electrons. The protons and neutrons are located in the nucleus of the atom, and are thus sometimes collectively referred to as nucleons, while the electrons surround the nucleus. As is familiar to most people, two properties of the nucleus are mass and electric charge. Two less familiar properties are magnetism and spin. These latter properties have much smaller effect on the behavior of an atom, but they can be utilized to obtain information about the chemical environment in which the atom is located. This is done by means of nuclear magnetic resonance. All four properties depend on what type of atoms we are dealing with. Since the atoms of interest in a medical MR examination are located in a human body, a very large part of them are hydrogen (^{1}H), carbon (^{12}C) or oxygen (^{16}O) atoms. Among these, only hydrogen has non-zero spin and magnetic moment, and thus only hydrogen nuclei contribute to the NMR signal. While most of what is stated here is valid also for other nuclei, this text is focused on hydrogen nuclei (single protons).

[1] Nowadays NMR is usually referred to as simply “MR” or “magnetic resonance”, avoiding the negative connotations to the word “nuclear”.

2.2 Magnetism and spin

The property of magnetism can be thought of as a small bar magnet located inside the atom nucleus (see figure 2.1). Like all magnets, it has a north and a south pole, and it points in a certain direction.

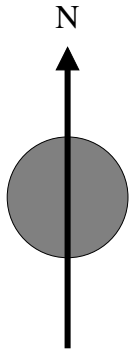


Figure 2.1: The magnetic moment of a nucleus.

The spin property is less intuitive. The nucleus is not spinning in the way that things do when they are rotating. However, it is sometimes convenient to think of the spin as an actual rotating motion. Then, the spin angular momentum can be visualized as in figure 2.2, where the direction of the arrow is the axis of rotation and the length of the arrow is proportional to the angular velocity. The spin vector is always parallel or anti-parallel to the magnetic moment of the nucleus.

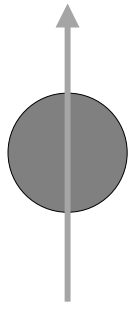
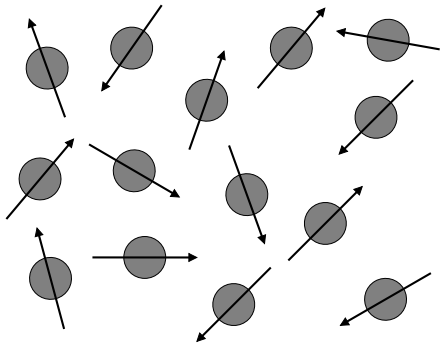


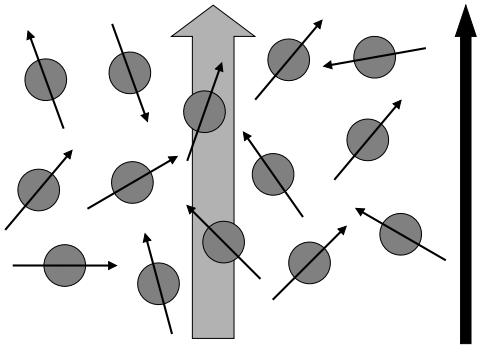
Figure 2.2: The spin of a nucleus.

In the absence of an external magnetic field, the magnetic moments of all nuclei are free to point in any direction. Since all directions are equally likely, the sum of all magnetization vectors is very close to zero. If an external magnetic field, often referred to as the B_0 field, is applied, more of the magnetic moments of the nuclei will be aligned with the field than against it. Then, a net magnetization in the

direction of the external field will be obtained. The stronger the external magnetic field, the more nuclei will align with it, producing a stronger net magnetization. Figure 2.3 shows possible orientations of the individual nuclei with and without an external magnetic field. The external field is illustrated by the black arrow, and the net magnetic moment is illustrated by the gray arrow. The alignment of the individual magnetic moments to the external field is greatly exaggerated in this figure, even though the external field in a typical medical MR scanner is very strong; 1.5 T or approximately 30000 times the Earth's magnetic field.



(a) Without external field.

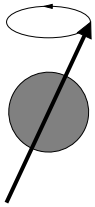


(b) With external field.

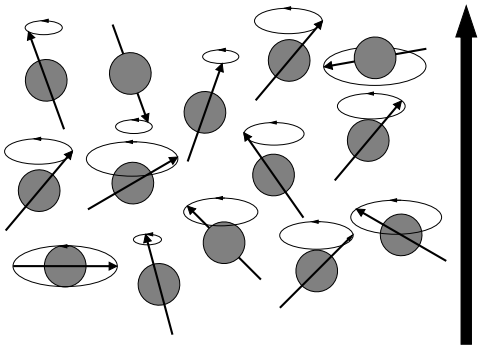
Figure 2.3: Orientation of nuclei with and without an external magnetic field (illustrated by the black arrow). In the presence of an external field, there is a net magnetic moment (illustrated by the gray arrow) pointing in the same direction.

In the presence of an external magnetic field, the spin causes the magnetic moments of the nuclei to precess around the field, as shown in figure 2.4. The frequency of the precession is proportional to the strength of the magnetic field and is referred to as the *Larmor frequency*. If we consider one individual magnetic moment, we essentially have a tiny precessing magnet. This is shown in figure 2.4(a). If a receiver coil was placed near the magnet, a small current would be induced by the precessing magnetic moment. Thus, it would be possible to measure

the precession, and thereby gain knowledge about the nucleus and its interaction with the surrounding sample. However, since there are countless little magnets, all precessing at the same frequency but with random phase, the net magnetism is constant. Thus, no current will be induced in the receiver coil. This is shown in figure 2.4(b).



(a) An individual spin.



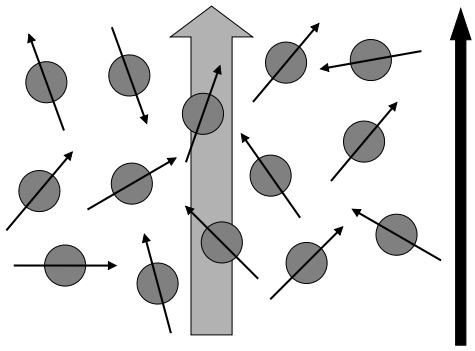
(b) A collection of spins.

Figure 2.4: Magnetic moments precessing around an external magnetic field.

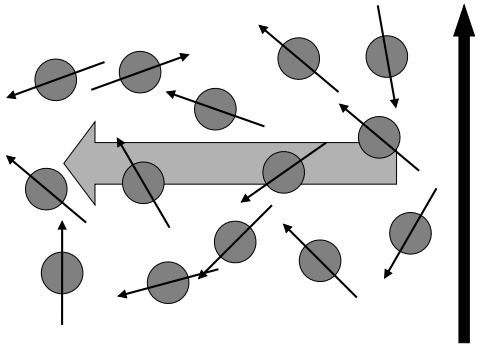
2.3 Radio-frequency pulses

By applying a radio-frequent (RF) pulse at the frequency of the precession of the spins, it is possible to change the orientation of all magnetic moments in a sample. Applying a RF pulse is often referred to as *exciting* the sample. If we, for example, rotate all the magnetic moments 90 degrees, as shown in figure 2.5, the net magnetization will change from being aligned with the external magnetic field (pointing up) to pointing to the left. However, it will still precess around the external field. Thus, after the application of an RF pulse, the precessing magnetic moments will be in phase and there will be a precessing net magnetization. This can be measured using a receiver coil, in which an alternating current at the Larmor frequency will be induced. By measuring this current, much information

about the magnetic and chemical properties of the sample can be revealed. For example, since all spins contribute with the same magnetism and precess at the same frequency, the current induced in the receiver coil is proportional to the number of spins in the sample.



(a) Before RF pulse.



(b) After RF pulse.

Figure 2.5: Magnetic moments before and after a 90 degree RF pulse. The net magnetization changes from pointing up to pointing to the left.

So why does the application of an RF pulse at a certain frequency change the orientation of the magnetic moments? This is most easily understood by considering the RF pulse as a rapidly changing magnetic field which is added to the strong B_0 field. This magnetic field rotates at the same frequency as the precession of the magnetic moments, and thus the latter are continuously nudged away from their axis of precession, as illustrated in figure 2.6. If the frequency of the RF pulse is not the same as the precession frequency, the magnetic moments will not be affected.

An RF pulse is characterized by two important properties: its *flip angle* and its *phase*. The flip angle determines how far from the precession axis the magnetic moment is flipped and thereby also how large the current induced in the receiver

coil will be. This is controlled by the amplitude and length of the pulse. The phase determines in which direction the magnetic moments are flipped, and affects the phase of the current induced in the receiver coil.

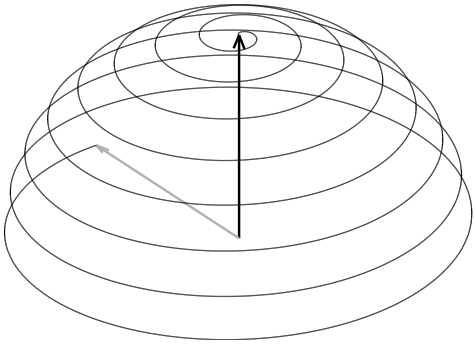


Figure 2.6: Motion of a magnetic moment as a result of an RF pulse with a flip angle of 90° at the precession frequency of the spin. The black arrow indicates the original direction of the magnetic moment, the gray arrow points in its final direction and the thin line traces its trajectory.

2.4 Relaxation

The current that is induced in the receiver coil decays with time. This process is called relaxation, and has two primary causes. One is that the thermal motion at the molecular level brings the net magnetic moment back to its original orientation along the external magnetic field. This is called T_1 relaxation. The other, referred to as T_2 relaxation, is that random interactions between the nuclei cause the precession of their magnetic moments to lose coherence, thus reducing the net magnetization. When the magnetic moments are entirely out of phase, no net magnetization remains. Since the precession frequency of the spins is proportional to the strength of the magnetic field, inhomogeneities in the field also cause dephasing of the magnetic moments. The combined effects of random nuclei interactions and inhomogeneity-induced dephasing is known as T_2^* relaxation. The rate of relaxation differs between different types of tissue and is typically expressed as a relaxation time after which a certain percentage of the signal remains.

By using cleverly designed sequences of RF pulses, relaxation times and other properties can be measured, and conclusions can be drawn about the chemical composition of the sample. In fMRI, T_2^* relaxation is measured to detect the difference between oxygenated and deoxygenated blood. This is possible since the iron in deoxygenated blood create small inhomogeneities in the magnetic field, causing the spins nearby to dephase more quickly than those in the vicinity of oxygenated blood.

2.5 Magnetic resonance spectroscopy

As stated in the previous section, inhomogeneities in the magnetic field cause different precession frequencies in different parts of a sample. Apart from this, differences in the chemical environment of the hydrogen nuclei also contribute to the different precession frequencies. For example, hydrogen bound in fat precesses at a frequency approximately 3.5 ppm lower than hydrogen in water. This difference is known as *chemical shift* and is used in MR spectroscopy for e.g. differentiating between different substances. This also forms the basis for the Dixon technique for fat/water segmentation, where the different resonance frequencies of hydrogen in fat and water are used to form two images from which the water and fat content can be determined. Dixon imaging is explained in greater detail in chapter 8.

3

Magnetic resonance imaging

In a homogeneous sample, for example a liquid solution, the chemical properties are the same everywhere. That is obviously not true in the human body, which is of interest in the context of this thesis. Otherwise, it would not be meaningful to discuss activated regions or separation of different types of tissue. While a certain property of a homogeneous sample can be described by just one value, an entire volume where each voxel represents a measurement is needed to describe something inhomogeneous, such as a part of the body. Hence a means to distinguish between signals from different locations in a sample is needed.

3.1 Spatial encoding

The process of acquiring spatially resolved magnetic resonance signals is referred to as magnetic resonance imaging (MRI). To encode the spatial position of the MR signals, the fact that the precession frequency of the spins is proportional to the strength of the magnetic field is utilized. Imaging can be performed in one, two or three dimensions. To illustrate the spatial encoding, a simple one-dimensional example is shown here. In figure 3.1, the distribution of hydrogen nuclei in a hypothetical one-dimensional sample is shown. The horizontal axis represents the spatial location while the vertical axis represents the amount of hydrogen at each position. To further simplify the example, all of the hydrogen is assumed to be located at three distinct positions in the sample.

As explained in the previous chapter, if the sample is exposed to a magnetic field, the magnetic moments of the nuclei will start precessing around the field with a frequency proportional to the field strength. By applying an RF pulse, the net magnetic moment can be tilted from the direction of the magnetic field, which will cause the precession to induce a current in a receiver coil. Thus, a signal whose amplitude depends on the amount of hydrogen, and whose frequency is proportional to the magnetic field strength, can be measured. This, however, does not provide us with any information about the spatial distribution of the hydrogen.

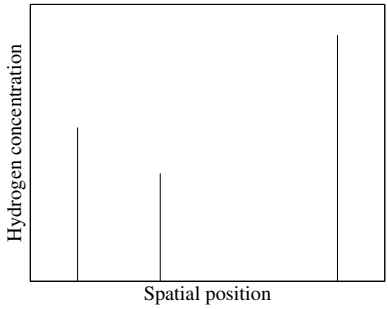


Figure 3.1: Spatial distribution of hydrogen in a one-dimensional sample.

To obtain this extra information, another magnetic field is used. This field is aligned with the main field of the MR scanner, but unlike the main field it is not spatially constant. Instead, the strength of this field varies linearly with the position along the horizontal axis, according to the figure below:

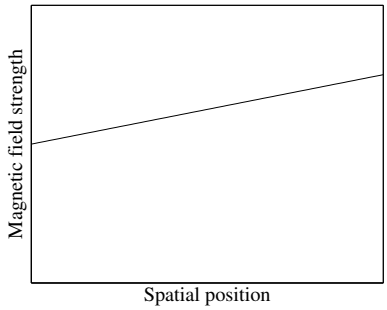
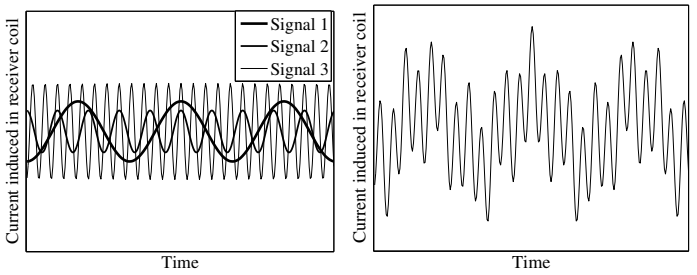


Figure 3.2: Spatially varying magnetic field.

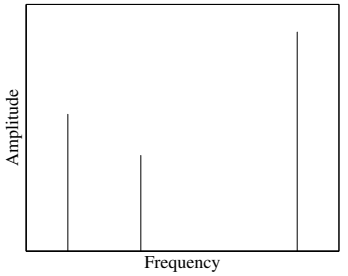
This magnetic field is referred to as a *gradient field*. By activating this field after applying the RF pulse, spins at different spatial positions are made to precess at different frequencies. Thus, the signal induced in the receiver coil will be the sum of several different frequency components, each of which originates from spins at a certain position in the sample. In this example, hydrogen nuclei far to the right will precess faster and generate high frequency signals, while nuclei in the left part of the sample will precess slower and generate signals of lower frequency. By calculating the Fourier transform of the measured signal, its frequency distribution and thus also the spatial distribution of the hydrogen is obtained. Figure 3.3 shows the signal induced in the receiver coil and the absolute value of the Fourier transform of the signal. Clearly, the Fourier transform is identical to the actual spatial distribution shown above.

This example illustrates the basic idea and shows that the acquired signal is ac-



(a) Signals from hydrogen at different locations.

(b) Measured signal (sum of the individual signals).



(c) Fourier transform of the measured signal.

Figure 3.3: Signal induced in the receiver coil and its Fourier transform.

tually the Fourier transform of the image. However, the example is simplified in order to convey the general idea behind MR imaging. The remainder of this section provides a more detailed description of how the acquisition is performed.

As stated above, when no gradients are used, all spins precess at the same frequency, determined by the strength of the B_0 field. Disregarding any relaxation effects, all spins will thus be in phase and this coherent precession will induce a sinusoidal current in the receiver coil. Since the spins are in phase they will all make an identical contribution to the induced signal. Hence the amplitude of the signal indicates the number of spins in the sample. This number, measured before the gradient is activated, is the DC component of the image (the zero frequency in the Fourier transform and the mean value of the image).

After the gradient is activated, spins in different parts of the sample will precess at slightly different frequencies. Thus the spins will lose coherence, becoming increasingly out of phase as time progresses. Figure 3.4 shows the phase of spins in different parts of the sample after different amounts of time.

Hence the current induced in the receiver coil will no longer be the sum of identical contributions from all spins, but instead the spins will contribute with signals of different phase. This may not be obvious at a first glance, but since the magnetic

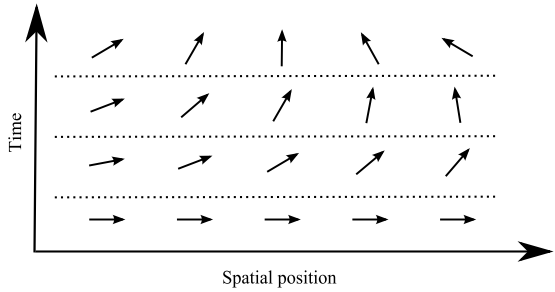


Figure 3.4: Spins at different spatial positions gradually lose coherence while a gradient is active. Because of the magnetic field gradient, spins to the right precess faster than spins to the left.

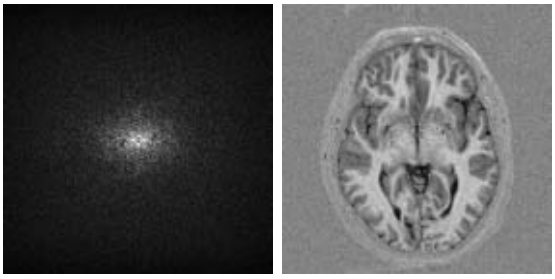
field and therefore also the phase of the individual spins vary linearly with spatial position, each sample of the induced signal corresponds to one frequency component of the desired image. Which frequency component is obtained depends on the phase difference between spins at different positions, which in turn depends on what gradient fields the sample has been exposed to since it was excited by the RF pulse. If no gradients have been applied, the DC component is obtained. If, on the other hand, a constant gradient has been applied for a long time, high frequency components are obtained. If a constant gradient has been applied for some time, and its polarity is reversed, spins with low precession frequencies will begin to precess faster and vice versa. Hence the spins will eventually rephase, and the signal will again correspond to the DC component of the image.

While the above description explains how an image can be acquired one frequency component at a time, it is convenient not having to think about the phase and frequency of each spin. A common alternative is to use the so called *k-space* formalism. The space where the image is acquired is then referred to as the k-space, which is similar to the Fourier space in that each coordinate corresponds to one spatial frequency. The advantage of this formalism is that a very simple relationship is defined between the coordinates of this space and the gradients used during acquisition: the position in k-space is determined by the integral of the gradients. Thus a stronger gradient moves faster through k-space, gradients with opposite signs move in opposite directions and a gradient which is active during a longer time moves further. Defining k as the coordinate in a one-dimensional k-space and $G(t)$ as the gradient strength at time t,

$$k(T) = \int_0^T G(t)dt, \tag{3.1}$$

where T is the time since excitation. Any frequency component can now be acquired by using the gradients to move to the corresponding position in k-space, where the signal is sampled.

In two-dimensional imaging, the procedure is similar but involves two gradient fields, oriented perpendicular to each other. The k-space is also two-dimensional. By using the gradients, it is possible to scan through k-space, after which the Fourier transform can be used to obtain a two-dimensional image of the object in the scanner. In three-dimensional imaging, yet another gradient is used to scan a three-dimensional k-space before a Fourier transform is used to obtain a three-dimensional image. Figure 3.5 shows a typical two-dimensional k-space and the corresponding image. Since the actual k-space samples are complex numbers, the magnitude is shown.



(a) k-space (b) Image space

Figure 3.5: Two-dimensional k-space and corresponding image.

3.2 Slice selection

A typical two-dimensional MR image, such as the one shown in figure 3.5b, depicts a slice through the imaged object. The thickness of this slice is usually just a few millimeters. However, if the sample is excited by an RF pulse as described in the previous chapter and two-dimensional spatial encoding is used, the image will show a "slice" with a thickness of several decimeters. Since the entire part of the body which is located inside the magnet bore is excited by the RF pulse, signals from this entire volume will contribute to the acquired image. Also, since two-dimensional spatial encoding does not provide any localization in the third direction, the signals from different positions in this dimension will effectively be averaged.

In order to receive signals only from a slice at a specific position and with a specific thickness, a *slice selection* gradient is used. As stated in section 2.3, spins are only excited if they experience a RF pulse at their precession frequency. Hence a slice can be selected by activating a magnetic field gradient perpendicular to the desired slice while applying the RF pulse. By choosing the strength of this gradient and the center frequency and bandwidth of the RF pulse, only spins located at certain positions will be excited by the pulse. Figure 3.6 shows a slice selection gradient and the relationship between the frequency and bandwidth of the RF

pulse and the position and thickness of the excited slice.

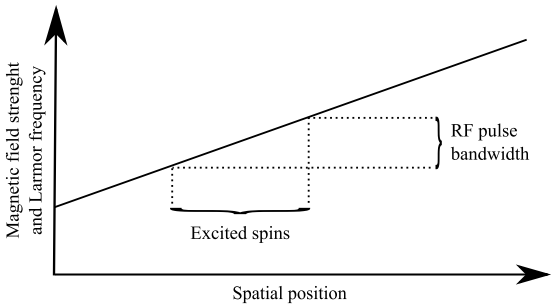


Figure 3.6: A slice selection gradient and the relationship between the RF pulse frequency and bandwidth and the position and thickness of the excited slice.

In three-dimensional imaging a thick slab is selected in the same way, but three-dimensional spatial encoding is used to separate signals from different positions in the slab.

3.3 Pulse sequences

In order to excite a sample and acquire signals from different parts of k-space, a sequence of RF pulses, slice selection gradients and spatial encoding gradients is needed. These sequences are typically called pulse sequences and depicted as diagrams showing RF pulse activity, amplitude and duration of gradients, and readout times (periods during which the signal is acquired, denoted Acq in the figures below). Figure 3.7 shows a common type of pulse sequence called a *gradient echo*. The name is due to the use of the gradient G_x to dephase and then rephase the signal. The dephase-rephase process can be seen in the figure as gradient waveforms with reversing polarity. Since the k-space position is the integral of the gradient waveforms, the initial dip in e.g. the x gradient corresponds to a translation to the edge of k-space. Samples acquired immediately after this correspond to horizontal high frequency image components.

The gradient in the y-direction (G_y) is here drawn as a set of different amplitudes. This indicates that the sequence is repeated a number of times, each time with a different value of G_y. The effect of this is that the k-space data is acquired one line (one vertical frequency) at a time, and that the entire slice is excited once for each line. The repeated excitation has the advantage that the readout period is kept short. Thus it is possible to keep relaxation effects, noise and frequency drifts at very low levels. This improves the quality of the acquired images. It is, however, also possible to acquire the entire slice in one excitation. This technique is called EPI (echo planar imaging; only one echo or excitation is used to acquire an entire image plane). A typical EPI pulse sequence is shown in figure 3.8. Here G_x

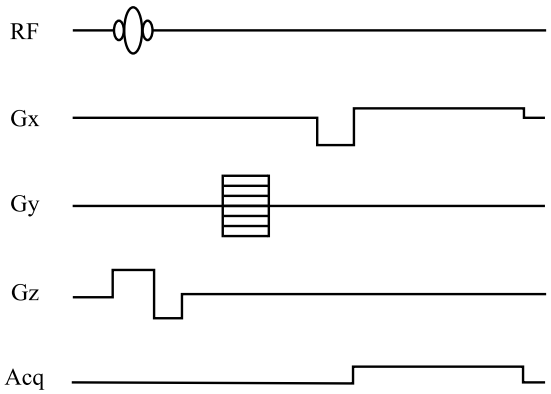


Figure 3.7: A gradient echo pulse sequence.

and G_y are first used to move to one corner of k-space. Then G_x alters between different polarities, i.e. scans back and forth along the x direction of k-space, while G_y is periodically activated for a short time in order to jump to the next vertical frequency. Thus, every second k-space line is scanned left-to-right while the other lines are scanned right-to-left. Because of the longer readout period and increased gradient activity, EPI images often have lower signal to noise ratio than other MR images.

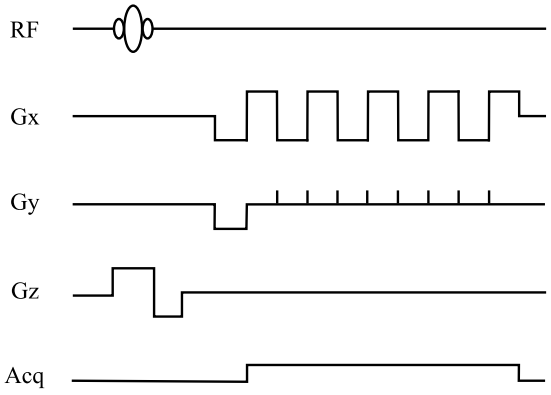
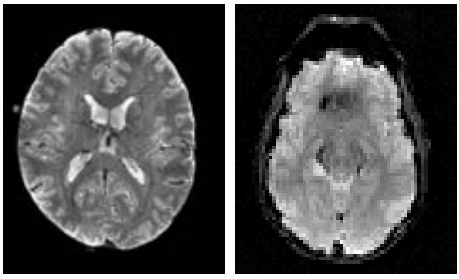


Figure 3.8: An echo planar gradient echo pulse sequence.

When EPI is used, an image may be acquired in as little as a few tens of milliseconds, while ordinary imaging may take several seconds. Thus EPI is the obvious choice if imaging speed is important, for example when capturing dynamic events. An example of this is fMRI, where a sequence of images is used to detect changing oxygenation levels. Examples of images acquired using ordinary imaging and using EPI are shown in figure 3.9. The difference in image quality is huge, and



(a) Ordinary imaging. (b) Echo planar imaging.

Figure 3.9: Two images acquired using MRI. One is optimized for high image quality and the other for high temporal resolution.

definitely motivates the use of ordinary imaging when speed is not essential.

It is also possible to acquire images an entire volume at a time (echo volumar imaging or EVI), but because of the very long readout period this technique is only useful for volumes of very low resolution (e.g. $8 \times 8 \times 8$ voxels).

3.4 Image contrast

The above discussion has primarily focused on measuring the amount of hydrogen present in different parts of a sample. Images showing hydrogen content are often referred to as being proton density (PD) weighted. As mentioned in the previous chapter, other properties such as relaxation times can also be measured using MR. By careful design of pulse sequences, images where each voxel represent for example the local relaxation time can be obtained. Such images are called T_1 weighted, T_2 weighted or T_2^* weighted depending on which type of relaxation affects the image to the greatest extent. The weighting of an MR image is often called the *image contrast*.

For several different reasons, the actual pixel values of a typical MR image have no absolute meaning. A proton density weighted image does not provide information about the absolute concentration of protons, and a T_1 weighted image does not provide the relaxation time in milliseconds. Rather, the pixel values are relative measures, from which conclusions such as "this region has a higher density than that region" or "the relaxation time is shorter in this region" can be drawn. However, special techniques can be used to measure the absolute values of these properties. This is called *quantitative imaging* and is very useful in e.g. fat accumulation studies and for detection of some types of plaque. More information about quantitative imaging can be found in Warntjes et al. (2007).

Images weighted to show proton density and relaxation times are, by the very nature of density and time, real valued. With the exception of some types of

relaxation weighted images, they are also positive. This is advantageous since small errors in the acquisition process tend to disturb the complex phase of the acquired images. When the signal is expected to be real and positive, the phase can easily be disregarded by using the magnitude of the images. Some imaging modes, however, utilize the possibility to encode information in the image phase. One example of this is velocity encoded imaging, which is used to measure e.g. local blood flow. Very briefly, this is achieved by using alternating magnetic field gradients to induce phase differences between spins moving at different velocities. Real images encompassing both positive and negative numbers are also useful for some measurements, for example in Dixon imaging for fat/water separation. In typical Dixon imaging, two images showing water minus fat and water plus fat, respectively, are acquired. Since there is more fat than water in some pixels, the first image contains both positive and negative numbers. Therefore the phase can not be discarded but instead needs to be corrected. This subject will be explored in more detail in chapter 8.

3.5 Further reading

Like NMR, magnetic resonance imaging is a huge topic, and there is no way to provide a complete description within the scope of this thesis. Apart from those mentioned here, there are countless parameters that control different aspects of the data acquisition. Some of them are trade-offs between imaging speed and image quality, some control the amount of radio frequent radiation and rapidly changing magnetic fields that the patient is exposed to, and others control which chemical properties are actually imaged. The interested reader is encouraged to find more information in the large literature on the subject, for example in Liang and Lauterbur (2000), Haacke et al. (1999) or Schmitt et al. (1998).

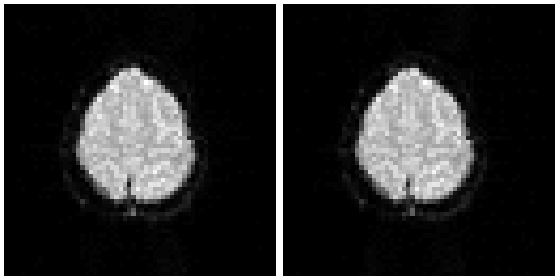
4

Functional magnetic resonance imaging

In functional magnetic resonance imaging (fMRI), the difference between the magnetic properties of oxygenated and deoxygenated blood is used to distinguish between active and inactive brain regions. This chapter describes how neural activity is linked to local changes in blood oxygenation and explains how images showing the oxygenation level can be obtained using MRI. It also describes how the choice of analysis method affects the precision with which activated regions are detected. Evaluation of analysis methods is discussed, and commonly needed pre-processing steps are described.

4.1 Neural activity and blood oxygenation

When neural activity increases in a part of the brain, the activated neurons start consuming an increased amount of oxygen. Since the oxygen supply in the brain is regulated, more oxygen will be transported to the activated region through the blood. However, the regulation overcompensates for the increased demand, and thus more oxygen than needed will be supplied. After a short time, the oxygenation level in the capillaries close to activated neurons will therefore be higher than normal. As mentioned in section 2.4, this affects the T_2^* relaxation by decreasing the local magnetic field inhomogeneity. By acquiring T_2^* weighted images, i.e. images where the voxel values reflect the local T_2^* relaxation rates, it is possible to distinguish between active and inactive regions. However, the difference between an image acquired during rest and one acquired during activity is very small. Figure 4.1 shows a resting-state image and an image acquired during activity. It is clear that the difference between these images is mostly due to noise, and that they are not sufficient to determine which regions are activated.

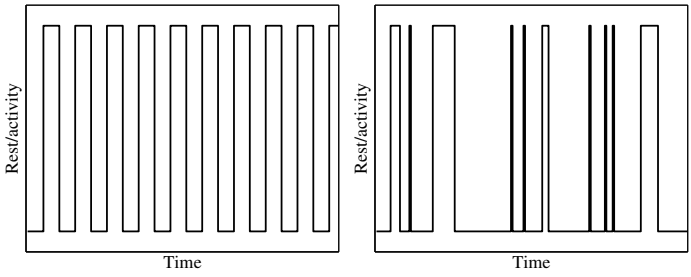


(a) Rest (b) Activity

Figure 4.1: T_2^* weighted images acquired during rest and activity.

4.2 Paradigms

To determine which regions are activated, something more than just two images from different states is obviously needed. The solution is to collect a sequence of images, usually between 100 and 200 over a period of 5 to 10 minutes, while the patient or subject alters between resting and performing some activity. The pattern of rest and activity is referred to as the *paradigm* of the experiment. In the simplest case a block paradigm, with equal periods of rest and activity, is used. For some tasks, however, it may be impossible to design a block paradigm. This is the case when the active condition can not be easily controlled, or does not span a sufficiently long period of time, such as e.g. answering a question or reacting to an instantaneous stimulus. In those cases, event-related paradigms are used instead. In an event-related paradigm, the active condition may appear as instantaneous spikes or blocks of varying length. Examples of block and event-related paradigms are shown in figure 4.2.



(a) Block paradigm. (b) Event-related paradigm.

Figure 4.2: Examples of block and event-related paradigms.

Disregarding the noise, the intensity time-series of a voxel in an activated region can be expected to vary according to the paradigm. This response to the activation paradigm is called the blood oxygen level dependent (BOLD) signal. The inten-

sity of an inactive voxel, on the other hand, will be constant. By finding voxels whose time-series resemble the paradigm, activated regions can be detected. This can be done in several different ways, some of which are presented in this and the following chapters.

4.3 BOLD models

The regulation of the oxygen supply is not immediate, but delayed by a short time. Thus, the BOLD signal from an activated region is not identical to the paradigm. Instead it is a delayed, smoother signal. During the first few hundred milliseconds, there is even a small decrease in blood oxygenation because of the neurons' increased consumption of oxygen. With some knowledge of the regulation process, it is possible to model the expected BOLD response. BOLD models can for example be calculated using convolution of the paradigm with a suitable response function (Boynton et al., 1996; Rajapakse et al., 1998; Friston et al., 1998). Another option is to use Buxton's balloon model (Buxton et al., 1998), which models blood flow, volume and oxygenation in the vicinity of activated neurons. After this, active regions can be recognized as regions where the time-series are similar to the BOLD model. Figure 4.3 shows a paradigm and a possible BOLD model.

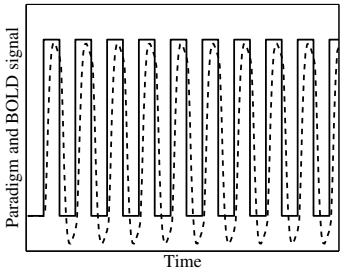
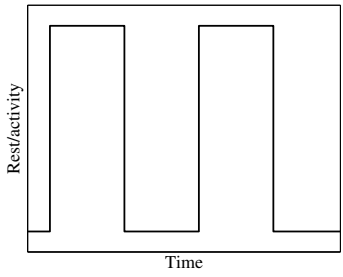


Figure 4.3: A block paradigm and a plausible model of the BOLD response.

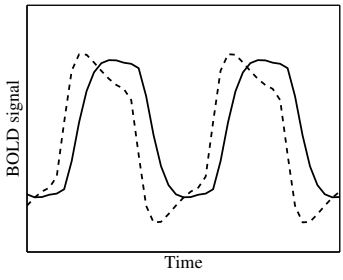
Using a fixed model of the oxygenation regulation is often insufficient. This is due to the large variations between different subjects, different scanning sessions and even between different regions of the brain in the same session. To be able to handle these variations, so called *subspace models* are often used. A subspace model is a model consisting of two or more *basis signals*. These signals are assumed to span the linear subspace of biologically plausible BOLD responses, i.e. it is assumed that any reasonable BOLD response can be expressed as a linear combination of the basis signals.

The simplest subspace models only take into account the variation in delay between the paradigm and the BOLD response. Then, one of the basis functions is defined as the expected response, calculated using for instance Buxton's model, and the other basis function is the temporal derivative of the first one. According

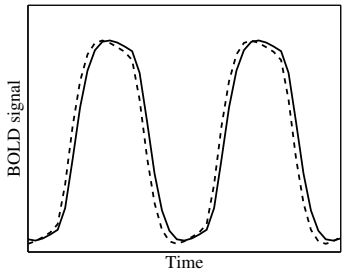
to Taylor's theorem, linear combinations of these two basis signals can approximate any small variation of the delay of the BOLD model with high precision. A more advanced type of subspace model can be generated using principal component analysis (PCA) (Jolliffe, 1986). A large number of different plausible BOLD responses are generated using for example the balloon model. The mean response is used as one of the basis signals, while the other basis signal is chosen to capture as much of the variation between the generated signals as possible. While still only utilizing two basis functions, this subspace model can adapt to different kinds of variations in the BOLD response (Friman et al., 2003). Figure 4.4 shows a part of a paradigm, two BOLD model basis signals generated using PCA and two possible linear combinations of them. These linear combinations of the basis signals represent BOLD responses with slightly different delay and other properties.



(a) Paradigm.



(b) Basis signals.



(c) Linear combinations of the basis signals.

Figure 4.4: A part of a paradigm, the basis signals of a subspace model of the BOLD response and two linear combinations of the basis signals.

4.4 Detecting BOLD-like signals

Several similarity measures can be used to compare the acquired signals to the BOLD model. One that is commonly used is correlation, defined as

$$Corr[x, y] = \frac{Cov[x, y]}{\sqrt{Var[x]Var[y]}}, \tag{4.1}$$

where x is a time-series and y is the model of the BOLD response. While this measure is only applicable to BOLD models with only one basis function, it can easily be extended to the more general case of subspace models. Then, the correlation between a time-series and the BOLD model is defined as the maximum correlation between the time-series and any linear combination of the basis signals. This generalization will be discussed in more detail in chapters 5 and 7.

Unfortunately, the acquired images are very noisy. Hence the activation map obtained by simply comparing the time-series of each voxel to the BOLD model does not reflect the actual activation very well. The simplest solution to this problem, applied in for example SPM (a widespread software package for fMRI analysis) is to average the time-series of nearby voxels, i.e. to low-pass filter the T_2^* weighted images before they are compared to the BOLD model (Worsley and Friston, 1995). While this solution reduces the noise, it also has a major drawback. Because of the low-pass filtering, the map of the activated regions is blurred, which makes it difficult to know the exact size or shape of an activated region. Also, and perhaps worse, small activated regions may disappear completely due to the blurring. The analysis methods described in the next chapters suggest two approaches to *adaptive filtering*, i.e. filtering which only performs averaging where it does not cause unnecessary blurring of the activation map.

4.5 Sensitivity and specificity

In analysis of fMRI data, the final result is usually not a measure of similarity to the BOLD model in each voxel. Rather, a classification of the voxel as activated or not activated is desired. This classification is obtained by thresholding the similarity values. Naturally, this raises the question of how large the similarity should be for a voxel to be declared as activated. A too low threshold will cause a number of voxels to be classified as activated when in fact they are not, while a too high threshold will cause the opposite problem. This can be described in terms of sensitivity (the fraction of truly activated voxels declared as active by the analysis) and specificity (the fraction of non-activated voxels that are declared as inactive). Ideally, the threshold should be chosen such that it maximizes both sensitivity and specificity. This is of course only possible if the similarity measures of all active voxels are higher than those of all inactive voxels. Because of the high noise levels in the data, and the fact that the time-series consist of a relatively low number of

samples, this places impossible requirements on the design of the analysis method and the model of the BOLD response. If the model accepts too much variation, the noise in the time-series from some inactive voxels will almost certainly be considered as possible BOLD responses. On the other hand, if the BOLD model is too restrictive, it will be unable to adapt to the variations between different plausible BOLD responses. As is explained in more detail in the next chapters, other properties of an analysis method also affect the sensitivity and specificity.

Usually, the threshold is chosen such that a certain probability of mistakenly declaring inactive voxels as active is obtained. This probability is referred to as the *significance* or p*-value*, and may either be defined for single voxels (the probability that a specific inactive voxel is declared active), for the entire data set (the probability that *any* inactive voxel in the data set is declared as active), or for smaller groups of voxels. To determine what threshold corresponds to a certain p-value, the distribution of similarity measures in a data set with no activated voxels is calculated, using either parametric or non-parametric statistics (Worsley, 1994; Friman and Westin, 2005; Locascio et al., 1997).

4.6 Exploratory analysis

Up to this point, this chapter has focused on so called *confirmatory analysis*, where a known signal (the BOLD model) is sought in the acquired data. There are, however, also techniques for *exploratory analysis*, which aim at detecting unknown but interesting signals in a data set (McKeown et al., 1998). It is of course not obvious what constitutes an interesting signal, but properties such as slow temporal variation and presence in several voxels may be used to define the set of desired signals. A number of methods for finding interesting signals are based on maximizing statistical independence, signal variance or auto-correlation. These are described in some detail and compared to each other in an fMRI context in Friman et al. (2002b).

In most fMRI experiments, the paradigm is known, and thus confirmatory analysis is the natural choice to detect active regions. However, there are cases where the BOLD signal expected in active voxels is unknown. One example is examinations of complex tasks such as driving, where no objective measure of action or rest is available (Calhoun et al., 2002). Exploratory analysis has also been used to investigate connections between different parts of the brain. By acquiring resting-state data, i.e. data from a scanning session without a paradigm or stimulus, and using exploratory techniques to search for regions with similar signals, conclusions can be drawn about which parts of the brain are connected (Liangsuo et al., 2007).

4.7 Evaluation of analysis methods

Since no method for analysis of fMRI data classifies every voxel correctly, even if provided with an optimal threshold, a means to compare the activation detection performance of different methods is needed. One solution is to use receiver operating characteristics (ROC) curves to evaluate how the sensitivity and specificity of an analysis method varies with the value of the threshold. ROC curves are described in more detail in section 4.7.1. Before an evaluation can be performed, however, a data set needs to be analyzed by the method to be evaluated. This data set can be of any of a few different types. These are real data, pseudo-real data and synthetic data. Each of these types has its own advantages and disadvantages for evaluating an analysis method.

Real data, of course, is data from a real fMRI examination. Since this is exactly the type of data for which the analysis methods are used in the real world, it may seem like the obvious choice. Unfortunately this type of data has one huge disadvantage: the true pattern of activation is unknown. This makes it impossible to simply compare the classifications from the analysis to the correct answer in order to obtain a measurement of the performance of the analysis method. However, as is briefly described in the next section, some aspects of the performance of a detection method can still be evaluated using real data.

Another disadvantage is that the activated regions in real data can not be chosen arbitrarily. Therefore it is often not possible to design a data set with specific properties in order to investigate how it is handled by a certain analysis method. For instance, it is difficult to choose the strength of the BOLD response and the shape of the activated regions when real data is used. Finally, there are large variations between real data sets from different experiments. Hence, the detection performance obtained when one data set is analyzed is not always a good prediction of the performance obtained when analyzing data from another experiment.

Pseudo-real data can be created from real data by mixing resting-state data, i.e. data from a scanning session where no task has been performed, with active time-series from another data set. The active time-series are taken from voxels with very high correlation to the BOLD model, i.e. voxels which are certainly activated (Nandy and Cordes, 2004). If the time-series in a large region of the resting-state data is replaced by time-series from an activated region, most of the properties of the pseudo-real data will be similar to those of real data. The advantage of using pseudo-real data is that the pattern of activation is designed and thus known. The most important disadvantage is that the edges between active and inactive regions are not necessarily similar to those encountered in real data. Since a large portion of the misclassified voxels typically reside close to the edges, this may in some cases be a big problem.

Synthetic data can be divided into two sub-categories. The first is created by inserting time-series similar to the BOLD model into resting-state data, while the

second is created by inserting BOLD-like time-series into synthetic noise. The disadvantage of this data type is, of course, that it is not necessarily similar to real data. The advantage is that the patterns of activation and the signal-to-noise ratio of the BOLD responses can be chosen entirely freely. This makes synthetic data well suited for testing specific properties of an analysis method, such as how well it detects activated regions of different shapes.

4.7.1 Receiver operating characteristics curves

When a pseudo-real or synthetic data set has been created and analyzed, both the results of the analysis and the correct answer in each voxel are known. Then, the detection performance of the analysis method can be presented as a receiver operating characteristics (ROC) curve. This curve shows how the sensitivity and specificity of the detection method vary with the value of the threshold. An example of a ROC curve is shown in figure 4.5. Note that the curve by convention is plotted with one minus the specificity on the x axis and the sensitivity on the y axis. Also note that the x axis is logarithmic, to show the performance for high specificities with greater accuracy.

Unfortunately, there is a problem with the ROC approach. If the data set does not have sharp edges between active and inactive voxels, the signal to noise ratio drops gradually from its maximum value to zero. This corresponds to voxels containing both active and inactive neurons. While some of these voxels are almost completely inside the active region and therefore have a high signal to noise ratio, others mostly contain inactive brain tissue. It is not quite clear whether the latter voxels should be considered as active in the ROC analysis. Because of this effect, it is not always straight-forward to evaluate a detection method using ROC curves. This problem can be diminished by calculating weighted ROC curves, where voxels which mostly contain inactive tissue are not allowed to affect the performance estimate as much as entirely activated voxels.

When real data is used, and the true classification of the voxels is therefore not known, the sensitivity and specificity can not be calculated. Thus, an ordinary ROC curve can not be used to evaluate the analysis method. An alternative is to use a so called modified ROC curve (Nandy and Cordes, 2003). Instead of sensitivity and specificity, the modified ROC curve is based on the number of voxels detected as active in data from a real experiment and the number of voxels detected as active in a resting-state data set. Since the number of voxels classified as active is used to determine the performance of an analysis method, this approach may overestimate the performance of methods which tend to leak strong BOLD signals into surrounding inactive areas.

Comparing two ROC curves is not always trivial, since one curve may be better in one interval while the other curve is better in another interval. To obtain a scalar measure of the detection performance, the integral of the curves can be

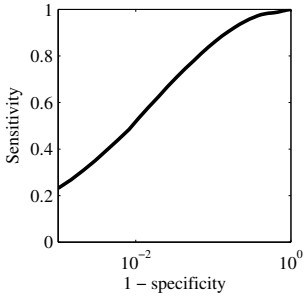


Figure 4.5: An example of a ROC curve, showing the sensitivity and specificity of an analysis method.

calculated. It is important, however, to consider which parts of the ROC curve should be included in such a calculation. For example, the sensitivity obtained where the specificity is very low is typically not of interest in an actual analysis.

Finally, it is important to notice that the choice of data (both the type of data, the pattern of activated regions and the signal to noise ratio) influences the detection performance. This is particularly important when comparing two analysis methods which have been optimized for different types of data sets. When choosing an analysis method, care should be taken to evaluate the candidate methods on data with properties similar to those of the data that will be analyzed. Depending on the actual experiment, certain properties of the analysis method may be more or less important. For example, if the experiment aims at determining whether functions related to language processing reside in the left or right hemisphere, the exact edge of the activated region may not be very important. On the other hand, an experiment mapping the relation between the visual field and the primary visual cortex needs very exact localization of the activated regions.

4.8 Pre-processing of data

4.8.1 Registration

Before data can be analyzed using either confirmatory or exploratory methods, some pre-processing is usually necessary. In the 5 - 10 minutes during which images are typically acquired, the subject or patient often moves a few millimeters in the scanner. Because of this motion, a certain voxel does not contain the same tissue during an entire experiment. Thus the time-series of intensity values recorded in a voxel does not only reflect the oxygenation variation in a specific part of the brain, but the combined result of motion-induced signals and oxygenation variations. This problem can be solved by the use of *image registration*, which aligns the images acquired at different times to each other. The registration algorithms

commonly used to accomplish this range from simple methods based on minimizing the square difference of the registered images (Hajnal et al., 1995) to more elaborate schemes based on local image phase (Knutsson and Andersson, 2005) or maximization of mutual information (Pluim et al., 2003). When registering the images, care must be taken in order to avoid detecting and introducing false motion in the image sequence. The variations in image intensity which are due to blood oxygenation may otherwise be picked up by the registration algorithm and "compensated" by moving the images, thus introducing motion instead of suppressing it. This problem has been shown to be particularly large when using simple registration methods such as least square minimization (Orchard et al., 2003) and is of course larger if very strong BOLD signals are present in the data. It can, however, be alleviated by the use of more advanced registration techniques, or by iterative registration and activity detection.

4.8.2 Detrending

Because of imperfections in the scanner hardware and physiological processes such as blood pulsation, respiratory movement etc, the noise present in both active and inactive voxels is not white but colored, or auto-correlated. The result of these effects sometimes correlates with the paradigm and BOLD model, thereby degrading the performance of the signal detection (Friman et al., 2004). While these noise components are not exactly known, they can be modeled and partly removed from the time-series prior to activity detection. This process is called *detrending*. There are two main classes of detrending approaches. One option is to use known basis functions, for example polynomials or discrete cosine transform bases, to estimate and remove trends in the data. The other alternative is to use exploratory techniques similar to those mentioned in section 4.6 to find trends typical for a particular data set and remove these from the time-series prior to signal detection.

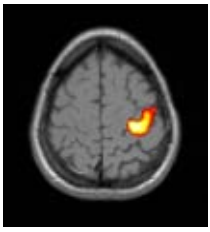


Figure 4.6: An example showing typical visualization of fMRI data. The image shows activation in the primary motor cortex, induced by finger tapping.

4.9 Visualization

The result of the analysis process is either a measure of correlation or some other similarity measure in each voxel, or a classification of each voxel as active or inactive. This, however, is very difficult to interpret unless anatomical information is also provided. Hence, activation is often visualized by overlaying an activation map on a proton density or T_1 weighted image which shows the anatomy of the brain. A typical example is shown in figure 4.6.

5

Analysis based on canonical correlation analysis

As described in the previous chapter, correlation is often used for measuring the similarity between fMRI time-series and the model of the BOLD response. Canonical correlation analysis (CCA) provides a means to measure this correlation while at the same time performing adaptive low-pass filtering of the data in order to avoid unnecessary blurring of the activation map. In this chapter, a previously suggested method for CCA-based analysis of fMRI data (Friman et al., 2001, 2003) is described. A certain disadvantage of the method is identified, and a modified algorithm is presented. Although the description is focused on two-dimensional analysis, an extension to three-dimensional data processing is straight-forward.

5.1 Canonical correlation analysis

Canonical correlation analysis was introduced by Hotelling (1936) as a method for finding the maximum correlation between linear combinations of two sets of variables. That is, given two multivariate variables $\mathbf{x} = (x_1, x_2, ..., x_m)^T$ and $\mathbf{y} = (y_1, y_2, ..., y_n)^T$, CCA finds two projection vectors $\mathbf{w}_x$ and $\mathbf{w}_y$ such that the correlation between the projections $x = \mathbf{w}_x^T\mathbf{x}$ and $y = \mathbf{w}_y^T\mathbf{y}$ is maximized. This maximum correlation is called the *canonical correlation*. The projection directions $\mathbf{w}_x$ and $\mathbf{w}_y$ can also be thought of as weights for linear combinations of the data in $\mathbf{x}$ and $\mathbf{y}$. A more detailed description of canonical correlation is available in chapter 7.

5.2 CCA in fMRI data analysis

When CCA is used for fMRI data analysis, $\mathbf{x}$ represents the output from several different spatial filters and $\mathbf{y}$ represents the basis signals for the subspace of BOLD signals considered to match the paradigm. Because of linearity, filtering the data spatially and then combining the filter outputs (as done by CCA) is equivalent to

first combining the filters and then using the resulting filter on the data. That is,

$$\Big(\sum_{k=1}^{N} a_k \mathbf{f}_k\Big) * I_t = \sum_{k=1}^{N} (a_k \mathbf{f}_k * I_t), \tag{5.1}$$

where $\mathbf{f}_k$ is the k:th filter kernel and I_t is the (two- or three-dimensional) T_2^* weighted image of the brain at time t. CCA is used once for each pixel to be analyzed, each time with $\mathbf{x}$ being the time-series obtained as filter responses from a neighborhood around the pixel under consideration. Thus, in each neighborhood the spatial filters are combined such that the correlation between the BOLD model and the output from the resulting filter is maximized. The correlation, denoted ρ, is assigned to the pixel in the center of the region.

5.2.1 Filter kernels

Several different sets of spatial filters have been proposed for use in the CCA framework for activation detection. The simplest filter set consists of nine different filters, each of which uses the information from one time-series in the 3 by 3 region surrounding the pixel examined for activation (Friman et al., 2001). These filter kernels are shown in figure 5.1.

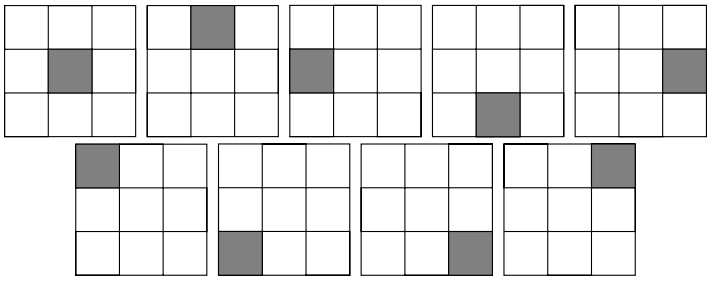


Figure 5.1: Simple filter kernels for use in the CCA framework.

When this filter set is used, the final filter can either use one pixel time-series by itself, or average several neighboring time-series, since CCA can choose any linear combination of the individual filters. Since averaging is only used where it makes detection of activated pixels easier, single activated pixels are no longer lost due to unnecessary blurring. This makes CCA, along with this filter set, very good at finding activated pixels. However, the filter set is not quite that good at correctly classifying inactive pixels. To understand this, imagine a single activated pixel in an otherwise inactive region. When examining this pixel for activation, the correlation with the BOLD model will be maximized by choosing $\mathbf{w}_x$ so that only the filter corresponding to the center pixel is used. However, when examining any of the eight neighbors of the pixel, exactly the same filter response can be obtained by choosing $\mathbf{w}_x$ so that one of the other filters is the only one included.

Thus, a single active pixel will be detected as a group of 3 by 3 activated pixels. In general, this filter set causes a growing of the detected active regions. To alleviate this problem, the symmetric filter set shown in figure 5.2 was proposed in Friman et al. (2002a). When this filter set is used instead, pixels outside an activated region are not as easily declared as active.

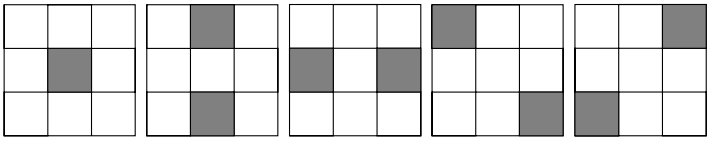


Figure 5.2: Symmetric filter kernels for use in the CCA framework.

To further improve the detection performance, a set of three anisotropic filters ($\mathbf{f}_{o_k}$, $k = 1...3$) and one smaller, isotropic filter ($\mathbf{f}_{i_1}$) was suggested in Friman et al. (2003). The oriented filters were defined as

$$\mathbf{f}_{o_k}(\mathbf{x}) = h(|\mathbf{x}|) \cdot (1 - g_{iso}(|\mathbf{x}|))\big((\hat{\mathbf{n}}_k^T \hat{\mathbf{x}})^2 - \frac{1}{4}\big) \tag{5.2}$$

and the isotropic filter was defined as

$$\mathbf{f}_{i_1}(\mathbf{x}) = h(|\mathbf{x}|) g_{iso}(|\mathbf{x}|). \tag{5.3}$$

$h(|\mathbf{x}|)$ is an isotropic low-pass filter and $g_{iso}(|\mathbf{x}|)$ is the weighting function used for creating the small isotropic filter $\mathbf{f}_{i_1}$ from $h(|\mathbf{x}|)$. h and g_{iso} can for example be gaussian functions, h being wider than g_{iso}. $\hat{\mathbf{n}}_k$ is the orientation of the k:th anisotropic filter. The filter directions are equally spaced, with an angular distance of 60 degrees. These filters have the appealing property that they can be interpolated (linearly combined) to form an oriented filter in any direction. If they are combined with equal weights, a large isotropic filter is obtained. The subtraction of $1/4$ in the directional weighting makes the filters narrower while still maintaining the interpolation property. The filter kernels are shown in figure 5.3. To illustrate the filters more clearly, the filters in the figure are larger than those actually used in the analysis.

This filter set has primarily two advantages compared to the symmetric set shown above. One is scalability, i.e. the possibility to easily create equivalent filters of different sizes (for example, it is not quite clear how to design filters equivalent to those shown in figure 5.2 in a 5 by 5 grid). The other advantage is that the set consists of a relatively low number of filters. This makes it more difficult for CCA to find a filter combination giving a high correlation in an inactive region. Thus, the specificity of the analysis is improved. The anisotropic filters are known as *steerable filters*, and were originally introduced in Knutsson et al. (1983).

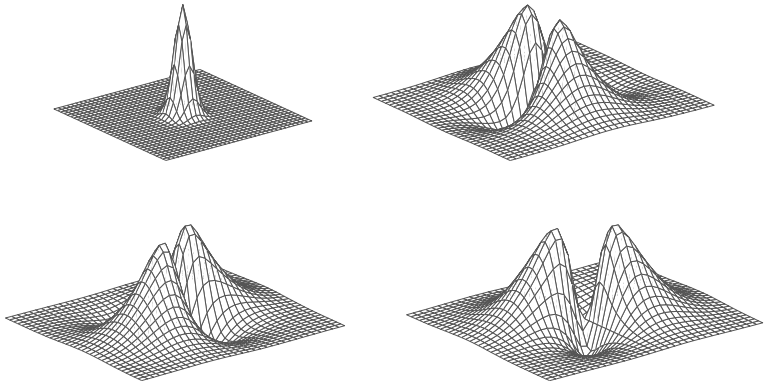


Figure 5.3: Steerable filters for use in the CCA framework.

5.2.2 Restricted CCA

There is still one problem, however, which is not related to the filter sets but rather to an intrinsic property of canonical correlation analysis. CCA can be thought of as an optimization method, which in the fMRI context is used to find an optimal spatial filter and an optimal BOLD model. The drawback is that CCA can only perform unconstrained optimizations, i.e. that there is no way to specify possible ranges for the weights in $\mathbf{w}_x$ and $\mathbf{w}_y$. This makes it possible for CCA to find large correlations where there is no activation by using implausible combinations of the filters. For example, if a negative weight is used for one basis filter while a positive weight is used for another basis filter, the resulting combination is not a low-pass (averaging) filter. One example of such an undesired filter is shown in figure 5.4. The same problem exists for the combinations of BOLD model basis functions. If, for example, two basis functions have been found to span the subspace of BOLD models, not all combinations of these two functions represent valid BOLD responses. Rather, the valid models may be between the two basis functions, such that $\mathbf{b} = \alpha\mathbf{b}_1 + (1-\alpha)\mathbf{b}_2$ is a plausible model of the BOLD response if and only if $0 \leq \alpha \leq 1$. Figure 5.5(a) shows two such BOLD model basis functions and 5.5(b) shows the model obtained when $\alpha = -1$. A time-series which is similar to the signal shown in figure (b) should not necessarily be considered to be activated. For examples of valid BOLD models constructed as linear combinations of the same basis signals, refer to figure 4.4(c) in chapter 4.

Both of these problems can be solved by using *restricted CCA* (RCCA) introduced in Das and Sen (1994), which as the name implies is a method to impose restrictions, or constraints, on the weights obtained by using CCA. The only type of restriction available is a positivity constraint, i.e. the constraint that some or all coordinates in $\mathbf{w}_x$ and $\mathbf{w}_y$ should be positive. It is, however, possible to obtain some other constraints by performing a linear change of basis before the actual

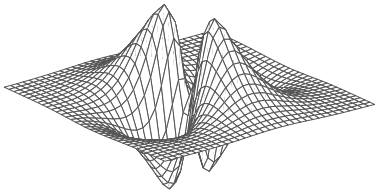
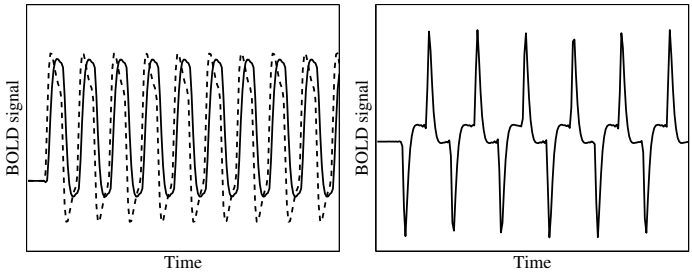


Figure 5.4: An unwanted spatial filter obtained by using a combination of positive and negative weights.



(a) BOLD model basis functions. (b) Invalid BOLD model function.

Figure 5.5: Two BOLD model basis functions and a possible but implausible linear combination.

RCCA calculation. An example of this is shown in the next section. In short, RCCA is based on repeatedly using ordinary CCA with one or more of the filters and/or BOLD model basis functions excluded, until a solution with only positive weights is found.

5.2.3 Analysis

By using the filter set described above together with restricted CCA and a model of the BOLD response, we now have everything that is needed to adaptively analyze fMRI data. The analysis is divided into two steps. In the first step, RCCA is used with the responses to the three anisotropic filters as one input ($\mathbf{x}$) and the BOLD model basis functions as the other input ($\mathbf{y}$). In this step, all weights are constrained to positive values. The resulting oriented filter $\mathbf{f}_o$ can then be obtained as

$$\mathbf{f}_o = \sum_k w_{x_k} \mathbf{f}_{o_k}. \tag{5.4}$$

In the second step, the response to $\mathbf{f}_o$ is combined with the response to the small isotropic filter $\mathbf{f}_{i_1}$. This time, however, simply constraining the weights to non-negative values is not enough. If the weight of the small isotropic filter is not large enough in relation to the weight of the oriented filter, the center pixel will have too

little effect on the filter output. This increases the risk of misclassifying inactive pixels located close to an activated region. To alleviate this problem, a change of variables is performed before the restricted CCA is calculated:

$$\mathbf{f}_1 = \mathbf{f}_{i_1} \quad \text{and} \quad \mathbf{f}_2 = \gamma \mathbf{f}_{i_1} + \mathbf{f}_o. \tag{5.5}$$

After this change of variables, the two filters to be combined by RCCA (again with positivity constraints on all variables) are the isotropic filter by itself ($\mathbf{f}_1$), and the weighted sum of the isotropic and the oriented filters ($\mathbf{f}_2$). Thus, the positivity constraints make certain that the isotropic filter is included. The parameter γ is used to control the minimum weight of the isotropic filter.

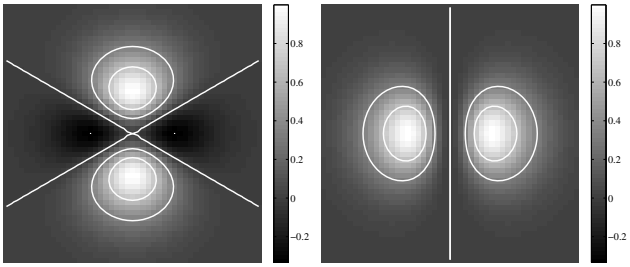
5.3 Rotational invariance

A desired property of any activation detection method is invariance to the shape of the activated region. That is, an activated region should be detected with the same probability regardless of its size, orientation and other properties. Some of these invariances are almost impossible to obtain. It is, for instance, always easier to detect a large activated region than it is to detect a single activated pixel. Invariance to orientation, however, is obtainable, but when restricted CCA is used in the first step of the analysis as described above, it is not achieved.

The reason for the lack of rotational invariance is that the filter shapes which can be constructed from the basis filters, using positive weights, are different depending on the desired filter orientation. If, for example, the desired filter is aligned with one of the basis filters, the smallest possible filter is, of course, that very basis filter. On the other hand, if the desired filter orientation is centered between two basis filter orientations, the smallest possible filter is the sum of these two basis filters. This filter is more isotropic than the individual basis filters, and hence it is not as well adapted to thin, anisotropic activated regions. Figure 5.6 shows the smallest possible filters in two perpendicular orientations. The filter in figure (a) is aligned with a basis filter, while figure (b) shows a filter which is not aligned with any of the basis filters. The aligned filter is clearly narrower than the non-aligned filter.

5.3.1 Isotropic filtering

The simplest method for solving the rotational invariance issue is, of course, to make all filters rotationally symmetric (isotropic). That reduces the complexity of the data analysis at the expense of adaptivity. The only remaining degree of freedom is scale, i.e. the filters always have the same shape, but their size may vary. Isotropic filtering is implemented by skipping the first step of the analysis, where the anisotropic filter is generated as a combination of the oriented basis filters. Instead, two isotropic filters of different size are combined in the second



(a) Aligned with basis filter. (b) Not aligned with basis filter.

Figure 5.6: The smallest possible anisotropic filters in two different orientations.

step. The drawback is of course that these filters can not adapt to anisotropic activated regions.

5.3.2 Anisotropic filtering

As described above, the previous method for anisotropic adaptive filtering using CCA lacks rotational invariance. By modifying the first step of that method, where the oriented filters are combined, a new method which resolves this issue has been developed.

Filter kernels

In the new method, a dynamically generated anisotropic filter $\mathbf{f}_d$ is used instead of a linear combination of the three oriented basis filters. However, the three basis filters are still used to determine the *orientation* of $\mathbf{f}_d$. In addition to the dynamically generated filter, a larger isotropic filter

$$\mathbf{f}_{i_2}(\mathbf{x}) = (1 - g_{iso}(|\mathbf{x}|))h(|\mathbf{x}|) \tag{5.6}$$

is also used. The large isotropic filter $\mathbf{f}_{i_2}$ is shown in figure 5.7, and an example of the dynamically generated anisotropic filter $\mathbf{f}_d$ is shown in figure 5.8. If $\mathbf{f}_{i_2}$ is combined with the small isotropic filter $\mathbf{f}_{i_1}$, the isotropic low-pass filter $h(|\mathbf{x}|)$ is obtained. Instead combining $\mathbf{f}_{i_1}$ with $\mathbf{f}_d$ yields an anisotropic low-pass filter. $\mathbf{f}_{i_2}$ is necessary to handle large isotropic regions of activation, since $\mathbf{f}_d$ is always an anisotropic filter. In the previous method, this case was handled implicitly since $\mathbf{f}_o$ is isotropic when all of the oriented basis filters are weighted equally.

Analysis

The analysis can now be divided into three steps, which are carried out for each pixel. The first two of these steps replace the first step in the previous method.

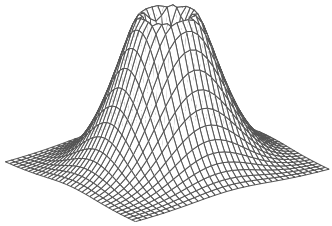


Figure 5.7: Large isotropic filter used in the last step of the analysis.

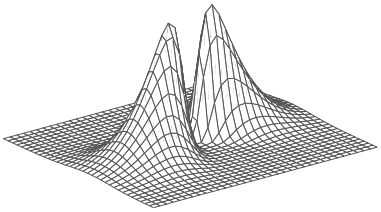


Figure 5.8: Anisotropic filter used in the last step of the analysis.

First, *partly* restricted CCA is used to find $\hat{\mathbf{w}}_x$, the weights for each of the oriented filters. This step is partly restricted in the sense that the filter weights are unconstrained, but the weights for the BOLD basis functions are constrained to positive values. This is important in order to avoid implausible BOLD responses (see figure 5.5). Since the weights for the spatial filters are unconstrained in this step, there is no guarantee that the resulting filter

$$\mathbf{f} = \sum_k w_{x_k} \mathbf{f}_{o_k} \tag{5.7}$$

is anywhere near a low-pass filter, which is what we want. It is, however, possible to create an anisotropic low-pass filter in the dominant direction of $\mathbf{f}$, if that direction is known. This is done in the second step of the analysis. In order to find the dominant direction, a tensor $\mathbf{T}$ is calculated as

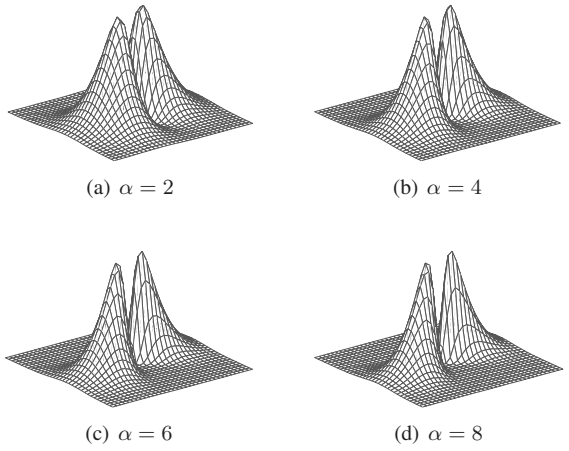
$$\mathbf{T} = \sum_k w_{x_k} \hat{\mathbf{n}}_k \hat{\mathbf{n}}_k^T. \tag{5.8}$$

The first eigenvector $\hat{\mathbf{e}}_1$ of $\mathbf{T}$ describes the dominant orientation of the filter $\mathbf{f}$. This makes it possible to generate the anisotropic low-pass filter $\mathbf{f}_d$ in the direction of $\hat{\mathbf{e}}_1$ according to

$$\mathbf{f}_d(\mathbf{x}) = h(|\mathbf{x}|)(1 - g_{iso}(|\mathbf{x}|))|\hat{\mathbf{e}}_1^T \hat{\mathbf{x}}|^\alpha, \tag{5.9}$$

where $h(|\mathbf{x}|)$ and $g_{iso}(|\mathbf{x}|)$ are defined as in section 5.2.1. Since $\mathbf{f}_d$ is only used to filter the data directly, and is not combined with other oriented filters, it does

not need the interpolation properties of the oriented basis filters $\mathbf{f}_{o_k}$. Thus, the subtraction of $1/4$ is excluded from the generation of this filter, and instead a parameter α is introduced to control the width of the filter. This has the advantage that no negative coefficients appear in the filter. A higher value of α gives a narrower filter, and experiments have shown that $\alpha = 6$ is a reasonable value. Anisotropic filters obtained with a few different values of α are shown in figure 5.9. Note that this kind of parameter could not be used in the old approach, where the interpolation property was needed.



(a) $\alpha = 2$ (b) $\alpha = 4$

(c) $\alpha = 6$ (d) $\alpha = 8$

Figure 5.9: Examples of dynamically generated oriented filters with different values of α.

The resulting filter $\mathbf{f}_d$ is applied to the data in the current neighborhood, as are $\mathbf{f}_{i_1}$ and $\mathbf{f}_{i_2}$. In the final step, these filter responses are combined using RCCA with positivity constraints on all variables. As in the old method, however, it is vital that the weight of the small isotropic filter is large enough, since the pixel in the center of the final filter may otherwise not be included in the averaging. Again, this is solved by performing a change of variables.

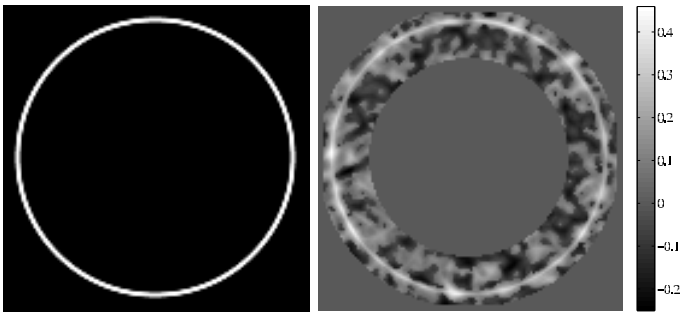
Filter design considerations

To obtain the rotational invariance, sampling effects have to be taken into account when the oriented basis filters $\mathbf{f}_{o_k}$ are designed. Otherwise, the filters will not be good approximations of their respective continuous functions and their interpolation properties will be lost. This affects the relative weights of the filters in the first step of the analysis, which in turn affects the orientation estimates. One method for alleviating this problem is to create large filter kernels and down-sampling those, instead of calculating the final filter coefficients directly from equation 5.2.

5.3.3 Experiments

To evaluate the rotational invariance, both the old and the new method have been tested on a data set with embedded simulated activation. The pattern of activation in the data set has not been selected to mimic real, or even realistic, neural activity. Rather, the purpose is to demonstrate the performance of the different methods when applied to highly anisotropic activation in different orientations. Also, to demonstrate the rotational invariance of the new method, the artificial data set has been designed to contain equal amounts of activation in all orientations. The signal to noise ratio in the artificial data is approximately 5 % in most parts of the activated areas.

Figure 5.10 shows the data set and the activation detected by the new method. Since there is very little visible difference between the correlation maps from the different methods, the results from the previous method are not shown. In figure 5.11, the difference between the correlation maps from the new method and the previous CCA-based method is shown. This figure is based on an average of the correlation maps from 200 data sets. It is clear that in certain orientations, there is a difference between the detection sensitivities of the two methods, while they are equally sensitive in other orientations. In figure 5.12, the angular variations are illustrated by the difference between the correlation maps and the same maps rotated 90 degrees. This figure shows that the differences in figure 5.11 are actually caused by variations in the result from the previous method, while the new method does not show any such variations.



(a) Activated regions. (b) Correlation map.

Figure 5.10: Locations of activated regions in anisotropic test data and correlation map from the new method.

Figure 5.13 shows histograms of filter orientations in the test data. The actual orientations are shown in figure 5.14. The difference between the two methods is particularly visible in the top of the activated region, where the activation is horizontal. The new method aligns the spatial filters with the activation, while the old method mostly alters between the two closest basis filter orientations. The orientations of the basis filters are -30, 30 and 90 degrees ($-\pi/6$, $\pi/6$ and $\pi/2$

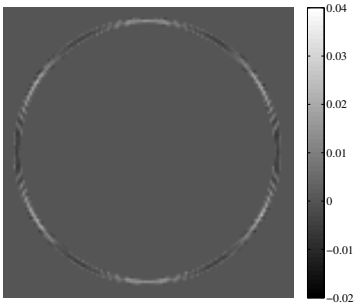
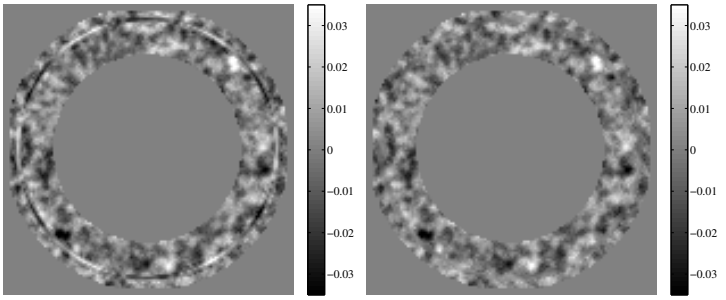


Figure 5.11: Average correlation difference between the new and the old method. The difference is largest where the activation is not aligned with any of the basis filters.

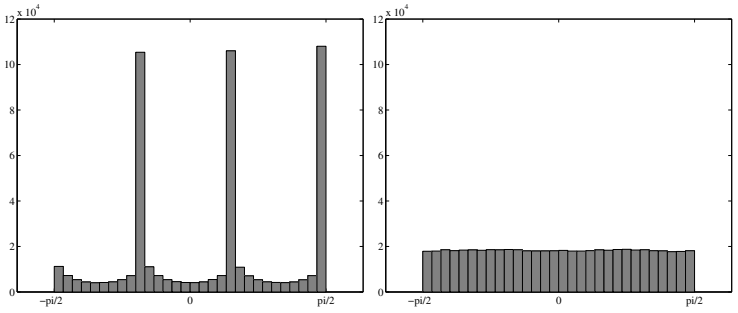


(a) Old method. (b) New method.

Figure 5.12: Average difference between correlation maps and rotated correlation maps. The new method does not show any of the structured variation present in the correlation map from the old method.

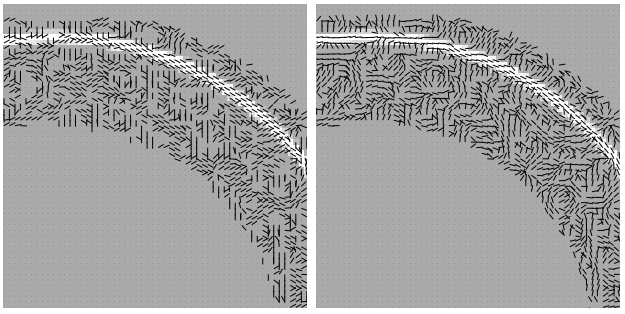
radians), respectively.

Since the activated region is very thin, the old method finds larger correlation coefficients by choosing the smallest possible filters (the basis filters), than by aligning the filters with the activation. Therefore, the detection performance varies with the angle between the closest basis filter and the activated region. The modified method, however, can choose identical filter shapes in all orientations. Hence, the filters are aligned with the activated region, and constant detection performance is obtained. Further experiments and results showing the performance of these methods are available in paper I, in the second part of the thesis.



(a) Old method. (b) New method.

Figure 5.13: Histograms of filter orientations. The overrepresented orientations in the histogram from the old method are those of the basis filters.



(a) Old method. (b) New method.

Figure 5.14: Orientations of the spatial filters when analyzing the anisotropic data. The new method aligns the filters with the activation, while the old method is clearly biased to the basis filters. This is particularly visible in the horizontal activation (in the top of the figures).

6

Analysis based on bilateral filtering

6.1 Introduction

In the previous chapter, it is described how CCA can be used for adaptive spatial filtering of fMRI data. Then, the size and shape of the local averaging region, i.e. the resulting adaptive filter, is chosen such that the correlation between the averaged time series and the model of the BOLD response is maximized. This makes the method very sensitive. Unfortunately, this sensitivity comes at the cost of specificity. Restrictions have to be imposed on the number and range of the parameters in the adaptive filter in order to maintain a reasonable selectivity. If given too much freedom, the method will find false signals in the noise since the filter is optimized to make the filter output as similar to the BOLD repose model as possible. Another problem caused by the correlation maximization principle is that when the filter is centered in a non-activated voxel but close to an activated region, the filter will try to ”reach in” to the activated region in order to pick up as much activation as possible. This will make the resulting regions labeled as activated become larger than they should be.

Instead of maximizing correlation, the filter can be designed to average over *consistent* regions (Sole et al., 2001; Tabelow et al., 2006). A consistent region only contains voxels that, in some aspect, are *similar* to each other. The question is then, of course, how to define similarity. This chapter presents a method for adaptive filtering based on averaging of consistent regions. Two different similarity measures are discussed, and it is shown that a combination of both measures can be used to detect activation with high accuracy.

6.2 Bilateral filtering

The simplest example of a similarity measure in this context is the difference between intensity values in the acquired images. This is a known as *bilateral filtering* (Godtliebsen et al., 1997; Tomasi and Manduchi, 1998). Bilateral filtering

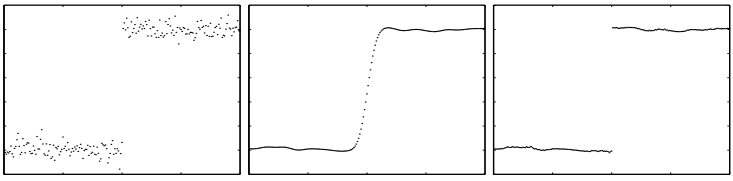
is a non-linear filtering technique for removing noise without degrading important structures such as edges. Ordinary linear low-pass filtering removes noise by averaging pixels that are close to each other. Such averaging, however, destroys sharp edges. Bilateral filtering solves this problem by also taking into account that the pixels to be averaged should have *values* close to each other. The bilateral filter kernel in each neighborhood can be expressed as a product of two filter kernels: the spatial filter and the range filter. The spatial filter is based on spatial distance while the range filter is based on the difference in image intensity. That is, given an image $I(\mathbf{x})$, were $\mathbf{x}$ is the spatial coordinate vector, the bilateral filter kernel $F(\mathbf{x}, \Delta\mathbf{x})$ at image coordinates $\mathbf{x}$ can be written

$$F(\mathbf{x}, \Delta\mathbf{x}) = F_s(\Delta\mathbf{x})F_r(\mathbf{x}, \Delta\mathbf{x}), \tag{6.1}$$

where $F_s(\Delta\mathbf{x})$ is an ordinary spatial filter kernel $g(\Delta\mathbf{x})$ and the range filter is defined as

$$F_r(\mathbf{x}, \Delta\mathbf{x}) = h(I(\mathbf{x} + \Delta\mathbf{x}) - I(\mathbf{x})). \tag{6.2}$$

A common choice of the filter kernels g and h are gaussian functions. When this filter is positioned over a relatively flat region it acts basically as an ordinary low-pass filter. Close to an edge, however, where an ordinary low-pass filter would be strongly affected by the values on the other side of the edge, these values will have a small effect on the result of the bilateral filter. The difference between low-pass and bilateral filtering is illustrated in figure 6.1. A one-dimensional signal is shown along with results from low-pass and bilateral filtering. The signal is a step function with additive gaussian noise. The bilateral filter can be seen as making a robust estimate of the local average; values that are "too far away" are not included.



(a) Noisy data (b) After low-pass filtering (c) After bilateral filtering

Figure 6.1: A noisy one-dimensional signal and the result of low-pass and bilateral filtering.

6.3 Method

6.3.1 Measuring time sequence similarities

In the previous section, the range filter measured the similarities between individual, one-dimensional samples. In fMRI, each voxel contains a whole sequence

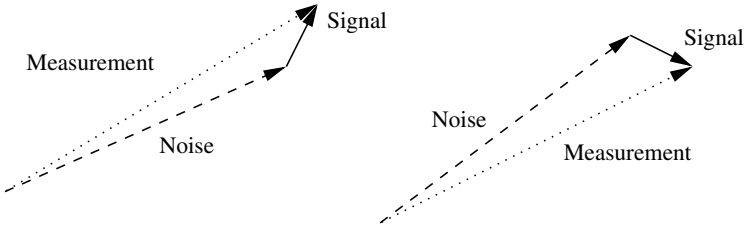


Figure 6.2: Different signals (solid) added to correlated noise (dashed) with low SNR will look similar (dotted) even if the signals are orthogonal.

of values describing the measured time signal at that position. Therefore, a similarity measure for vectors is needed in order to be able to create the range filters. Such a similarity measure is proposed for filtering of fMRI data in Godtliebsen et al. (2001). This is, however, a rather coarse measure giving the similarity one if the Euclidean distance between two time-series is below a certain threshold and zero otherwise. However, the authors recognized that other, more sophisticated similarity measures would perhaps be better.

A straight-forward choice of similarity measure would be the scalar product between the vectors themselves or between the normalized vectors. The latter would correspond to the correlation coefficient between the time-series. One problem with using such a simple similarity measure is that this only combines voxels with high SNR, i.e. the voxels where averaging is least needed. Voxels with low SNR or without any signal would have smaller chances to be combined with their neighbors. This would not solve one of the common problems in fMRI: that spurious correlations are often found in the noise because of the high noise level and the limited number of samples in time. To limit this problem, we would like to average not only in active regions, but also in inactive regions. The optimal similarity measure would indicate a high similarity between two active voxels and between two inactive voxels, but a low similarity between an active and an inactive voxel. This is not straight-forward since two inactive voxel time-series in general are very different.

Another potential problem with a simple similarity measure, such as the scalar product or correlation between time-series, is that all pairs of neighboring time-series would give more or less the same similarity measure because of the spatially auto-correlated noise. This is especially true when the SNR is very low, which is usually the case in fMRI. This is illustrated in figure 6.2 where two orthogonal signal vectors give very similar measurements because of the strong correlated noise vectors added to the signal.

One solution to this problem is to measure the similarities after the measured signal vectors have been projected into a signal subspace, e.g. the subspace spanned the BOLD model basis functions. After this projection, much of the noise has

Figure 6.3: Illustration of the signal subspace. Different colors represent different signals and the saturation represents the strength of the signal.

been removed, leaving only true signal and the part of the noise that fits the signal model. For a signal in a non-activated voxel, the resulting vector will in general have a small magnitude and a random orientation. Signals in an activated region will have larger magnitudes and consistent directions. Figure 6.3 illustrates the space of vectors. Different colors represent different signals and the saturation represents the strength of the signal. The gray region in the center corresponds to noise and very little signal. In this space, a euclidean distance between the vectors can be used to separate different signals from each other as well as signal from noise, while giving a small distance between two noise vectors. Hence the signal range filter can be written as

$$F_{rs}(\mathbf{x}, \Delta\mathbf{x}) = e^{\frac{-\|\mathbf{B}^T\hat{\mathbf{s}}(\mathbf{x}+\Delta\mathbf{x})-\mathbf{B}^T\hat{\mathbf{s}}(\mathbf{x})\|^2}{2\sigma_{rs}^2}}, \tag{6.3}$$

where $\hat{\mathbf{s}}(\mathbf{x})$ is the normalized signal vector in the voxel at location $\mathbf{x}$, $\mathbf{B}$ is a matrix with the temporal basis functions as its columns and σ_{rs} is the width of the signal range filter. The normalization of the signal has erroneously been omitted from equation 2 in paper II.

6.3.2 Anatomical similarities

Another variation of this theme is to average signals if they are located close to each other and they reside in the same type of tissue (fat, cerebrospinal fluid (CSF), or gray or white matter). Neural activation occurs in gray matter and has also been observed in white matter at high magnetic field strengths (Maldjian et al., 1999; Tettamanti et al., 2002). No BOLD signal can be expected in fat or cerebrospinal fluid, however. By constraining the filter to only combine signals from the same type of tissue, activation will not "leak" into parts of the brain where it is biologically implausible. Equally important, noise from regions in other tissue types will not degrade the signal from an activated part of the brain. Hence it seems advantageous to utilize information about anatomical structures in the brain when analyzing fMRI data. Different variations based on this idea have been proposed in for example Kiebel et al. (2000); Walker et al. (2006); Ou and Golland (2005).

This approach can be incorporated in the bilateral filtering framework by using an anatomical range filter F_{ra}, based on e.g. intensity differences in a T_1 weighted image. Since T_1 values reflect different tissue types, this range filter causes signals to be averaged only if they reside in the same type of tissue. In the experiments presented here, we have used a gaussian kernel directly on the T_1 values:

$$F_{ra}(\mathbf{x}, \Delta\mathbf{x}) = e^{\frac{-\|I(\mathbf{x}+\Delta\mathbf{x})-I(\mathbf{x})\|^2}{2\sigma_{ra}^2}} \tag{6.4}$$

where $I(\mathbf{x})$ is the T_1 value. However, a non-linear transformation of the T_1 values might be beneficial in order to improve the discrimination between different tissue types.

6.3.3 Combining signal and anatomical similarities

Each of the variations of bilateral filtering described above has its own advantages and limitations. If only anatomical information is used to constrain the filters, activation may be smeared into adjacent voxels in the same type of tissue. If, on the other hand, only the similarity between different time-series is considered, the available anatomical information is not utilized to improve the detection accuracy. By multiplying the signal range filter F_{rs} and the anatomical range filter F_{ra}, a range filter $F_r = F_{rs}F_{ra}$ which takes both kinds of similarities into account is obtained. This filter is constrained by signal similarities as well as by anatomical similarities, thereby being more selective than either of the filters F_{rs} or F_{ra} alone.

6.4 Experiments

The analysis method has been evaluated on both synthetic and real fMRI data. Some of the experimental results are provided here, while others are available in paper II in the second part of the thesis. The method is applicable to both two- and three-dimensional filtering. Results from both types of analysis are presented.

6.4.1 Two-dimensional data analysis

The synthetic data consist of an image describing the anatomy and a sequence of noisy images with biologically plausible BOLD signals embedded at known locations. The activated locations are chosen such that most of the activation is located within “gray matter”, while some activation is present in “white matter”. Neither the anatomy nor the activated locations have been designed to mimic real fMRI data, but rather to demonstrate the detection algorithm. The anatomy and the activated regions are shown in figure 6.4, where white illustrates white matter, gray is gray matter, black is CSF and the red regions are activated.

Figure 6.5 shows activation maps obtained by a number of different methods on this two-dimensional data. Panels (a)-(c) show the result when using ordinary

spatial low-pass filtering with σ_s = 0.5, 1.0, and 2.0 respectively. Panels (d)-(f) show the activation maps obtained using only anatomical constraints ($F = F_s F_{ra}$), only constraints based on time-series similarity ($F = F_s F_{rs}$) and the proposed combination of constraints ($F = F_s F_{rs} F_{ra}$) respectively. In all three adaptive methods the width of the spatial filter is σ_s = 2.0.

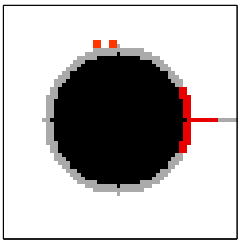
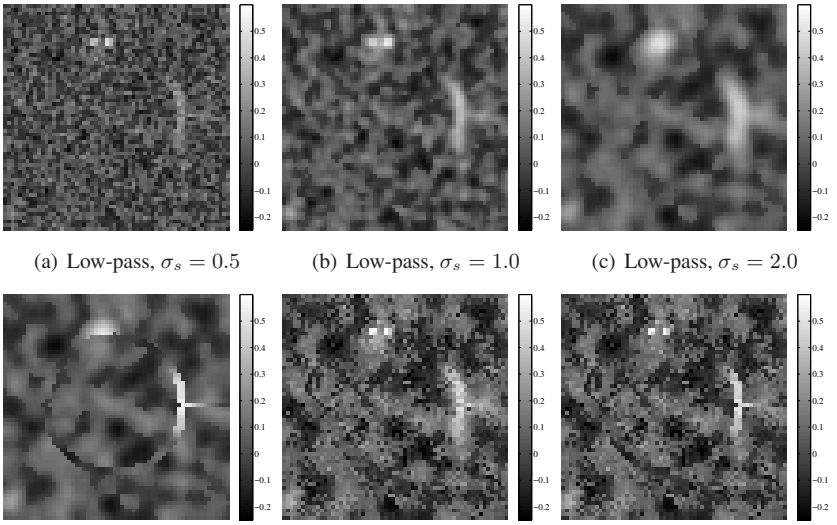


Figure 6.4: Synthetic test image. Red indicates active pixels.

The adaptive methods all provide enhanced detection performance compared to the fixed low-pass filtering. When only anatomical information is used to constrain the filters, however, the activation leaks into surrounding areas with the same type of tissue (figure 6.5(d)). On the other hand, when only the similarity between the time-series themselves is used to constrain the filters, some boundaries between active and inactive regions are slightly blurred (figure 6.5(e)). The proposed method, where both types of constraints are used, provides an activation map with clear boundaries between active and inactive voxels (figure 6.5(f)). These visual results are confirmed by the ROC curves in figure 6.6. It is clear that the method using both signal and anatomical priors provides the best overall performance, followed by the two other bilateral methods, and finally the low-pass filtering. The best performance of the non-adaptive method is obtained when a spatial filter with $\sigma_s = 1.0$ is used.

6.4.2 Three-dimensional data analysis

In three-dimensional data, the correlation between neighboring slices can be used to further improve the signal to noise ratio in activated regions. To demonstrate this, three-dimensional analysis has been carried out on a volume with characteristics similar to the two-dimensional image above. The detection performance is shown in the ROC curves in figure 6.7, and is obviously better than for two-dimensional analysis. The relative performance of the different methods is similar to the two-dimensional case, except for the low-pass filtering with $\sigma = 1$. In three dimensions, this filter kernel averages over a too large region.



(a) Low-pass, $\sigma_s = 0.5$ (b) Low-pass, $\sigma_s = 1.0$ (c) Low-pass, $\sigma_s = 2.0$

(d) Adaptive, anatomical (e) Adaptive signal (f) Adaptive, anatomical and signal

Figure 6.5: Results on simulated 2D data.

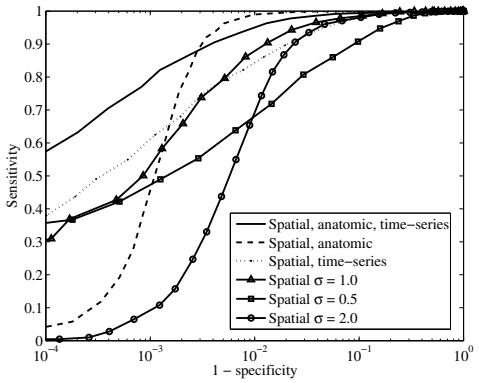


Figure 6.6: ROC curves for the simulated 2D data.

6.4.3 Data without sharp edges

Section 6.2 mentions that bilateral filtering is particularly good at preserving edges in data. It might be argued that fMRI data typically contain very few edges, since each voxel represents the average value in a region and it is highly unlikely that the boundary of an activated region is located exactly at the edge between two voxels. Therefore, the synthetic data shown above may not properly represent the properties of real data. The analysis has also been applied to another two-

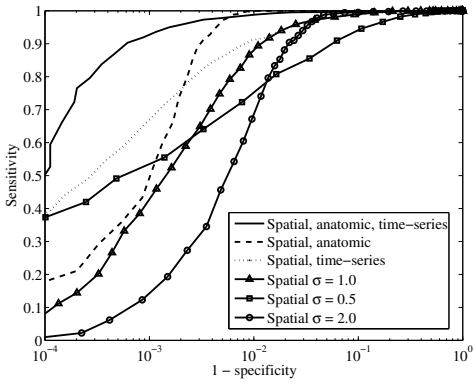


Figure 6.7: ROC curves for the simulated 3D data.

dimensional data set, shown in figure 6.8. This data set is similar to the one shown above, except that edges between active and inactive regions, and between different types of tissue, are smoother.

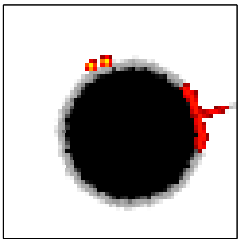
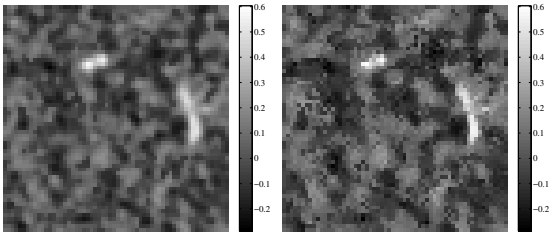


Figure 6.8: Synthetic test image without sharp edges between active and inactive regions. The color scale from red to yellow indicates active pixels with different signal to noise ratio.

As shown in the above section on two-dimensional analysis, $\sigma_s = 1.0$ yields the best detection performance for the fixed low-pass filtering. The best adaptive filter was obtained by combining the signal and anatomical similarity measures. Figure 6.9 shows correlation maps obtained using these two filters. The adaptive filtering appears to preserve the shape of the activated regions with higher accuracy. ROC curves are shown in figure 6.10. According to these curves, the difference between the methods is smaller than for the data set shown above. This has two explanations. First, the similarity measures used to define the range filters F_{rs} and F_{ra} are based on the difference between signals and T_1 values in neighboring voxels. In smooth data, these differences are smaller, which makes it more difficult to adapt the range filters to the data. Second, as mentioned in chapter 4, the ROC analysis requires voxels to be classified as either active or inactive. Some voxels at the edge of an active region contain very little active tissue and hence have a very low



(a) Low-pass, $\sigma_s = 1.0$ (b) Adaptive, anatomical and signal

Figure 6.9: Results on smooth two-dimensional data.

signal level, but are nevertheless considered as active in the ROC analysis. When low-pass filtering is used, the blurring of stronger signals from neighboring voxels causes many of these edge voxels to be classified as active. The adaptive method, however, often do not detect them. It may be argued that a certain percentage of active tissue should be required for a voxel to be considered active, and thus, this effect may be considered either an advantage or a disadvantage of the proposed method. It does, however, degrade the detection performance as measured by the ROC analysis. Still, the ROC curves indicate that the adaptive method performs better than the low-pass filter for most specificities of interest.

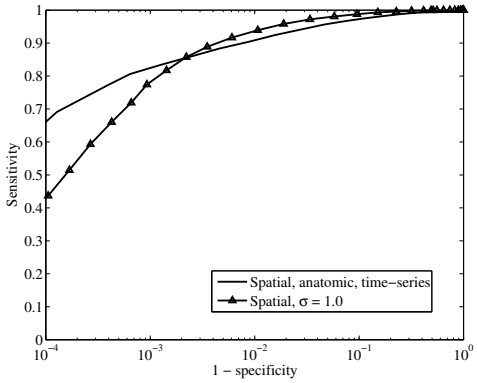


Figure 6.10: ROC curves for the smooth 2D data.

6.5 Discussion

By combining local signal and anatomical similarities, better performance can be obtained compared to when only *one* of the similarity measures is utilized. This can be seen qualitatively in the correlation maps as well as quantitatively in

the ROC curves in figures 6.6 and 6.7. In bilateral filtering the effective filter is a product between a spatial kernel and the range filter. When the similarity measures are combined, the range filter itself is a product between a signal range filter and an anatomical range filter. These filters can be seen as constraints on the spatial filter that prevent it from including other kinds of signals or signals from other kinds of tissues. Utilizing a priori information to constrain the solution space is in general a good idea. This explains the superior performance. The caveat is that it is very important that the a priori information is correct. For example, the T_1 and T_2^* volumes need to be properly aligned. A significant misalignment will make the filters see different regions and, hence, impose wrong constraints. For the same reason it is also important to avoid or correct for geometrical distortions of the T_2^* images.

As was shown in the ROC curves, particularly in figure 6.6, for some choices of specificity the best detection performance is obtained by only using constraints based on anatomical similarity. This is caused by the signal similarity constraint limiting the effective filter size when a spurious signal-like time-series is encountered in an inactive voxel. It should be noted, however, that the specificities for which only anatomical constraints provides the best results are rather low, and that the combination of both types of constraints provides better performance for most interesting specificities.

Compared to linear low-pass filtering, the spatial kernel in bilateral filtering can be rather large, since the effective kernel is restricted by the range kernel if necessary. This makes it possible for bilateral filtering to gain very much from an extension from two- to three-dimensional analysis, since the possibilities to find support for a large kernel becomes greater in high-dimensional data.

In the filtering method described here, the time-series of one single voxel is compared to time-series in surrounding individual voxels. The individual voxels can be thought of as basis functions for the filtering process, similar to the isotropic and anisotropic basis filters described in the previous chapter. Instead of the single voxel bases, isotropic and oriented filters could be used as basis functions also in bilateral filtering. Then the similarity would be measured between the filter response to a small isotropic filter and the responses to a set of other basis filters. Each filter would be included in the averaging to an extent depending on the similarity values. This would cause a slight increase in blurring, but may improve the adaptation to smooth active regions. The exact choice of basis filters remains an open question, though.

In contrast to the CCA-based analysis method presented in the previous chapter, this method does not maximize the similarity between the filter response and the BOLD model. Hence, it does not have the same bias towards sensitivity.

7

Robust correlation estimation

Correlation is used to measure the similarity between signals in many different applications. Examples include speech recognition (Hunt, 1999), fingerprint recognition (Angle et al., 2005), image registration algorithms (Penney et al., 1998) and, as described in previous chapters, analysis of fMRI data. In many of these cases, the correlation between a set of known reference signals and a newly recorded signal is calculated in order to determine whether the new signal belongs to the same group as the reference signals. In the biometric applications, the correlation may be used to determine whether for example a speaker is a certain person or not, while in the fMRI context it is used to determine if a certain part of the brain is activated. Depending on the application, the reference signals may either be recorded in advance or calculated from a model. In either case, it is often possible to ensure a certain quality of these signals since they are recorded or calculated under controlled circumstances. The situation is worse for the new signals, which are compared to the references. These entire signals may have a low signal-to-noise ratio, or sometimes most of the signals are well represented by the recording, while some segments are corrupted by for example background noise or radio frequent interference. Since these signals are recorded while a system is being used, it is often not acceptable to re-record until signals of good quality are obtained. In a voice recognition system, re-recording translates to making users repeat spoken commands, and in the fMRI application it translates to repeating an entire experiment. The latter is both troublesome for the patient and expensive. Hence, a way to measure correlations robustly, i.e. in such a way that corrupted signal recordings have minimal influence on the correlation estimates, is desirable.

This chapter presents an estimator of correlation, which is robust to corrupted segments in the signals. The chapter is organized as follows: in the next section, correlation and its generalizations to multidimensional signals (the general linear model and canonical correlation analysis) are reviewed. In section 7.2 the proposed method is explained. Section 7.3 presents how the method can be applied to fMRI data analysis, while section 7.4 demonstrates detection of partially occluded objects in images. A discussion of the method is presented in section

7.5.

7.1 Theory

7.1.1 Correlation, GLM and CCA

The ordinary Pearson correlation ρ between two one-dimensional signals x and y is defined as the covariance of the signals divided by the geometric mean of their respective variances, i.e.

$$\rho = \text{Corr}(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x)\text{Var}(y)}}. \tag{7.1}$$

Assuming that both x and y are sampled signals of length N with zero mean, an estimate of the correlation, $\tilde{\rho}$, can be calculated as

$$\tilde{\rho} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i^2}}. \tag{7.2}$$

If the variables are not zero-mean, their respective mean values are simply subtracted prior to this calculation.

The concept of correlation can be extended to multidimensional variables. The simplest extension is named the general linear model (GLM) and handles correlations between a one-dimensional and a multidimensional variable. The correlation is then defined as the maximum correlation (disregarding the sign) between the one-dimensional variable and any one-dimensional projection of the multidimensional variable, i.e.

$$|\rho| = \max_{\mathbf{w}} |\text{Corr}(x, \mathbf{w}^T \mathbf{y})|, \tag{7.3}$$

where the vector $\mathbf{w}$ defines the projection of $\mathbf{y}$ which maximizes the correlation with x.

As explained in chapter 5, a further generalization is termed canonical correlation analysis (CCA) (Hotelling, 1936). CCA handles correlations between two multidimensional variables. The canonical correlation is defined as the maximum correlation (still disregarding the sign) between any projections of the two variables, i.e.

$$|\rho| = \max_{\mathbf{w}_x, \mathbf{w}_y} |\text{Corr}(\mathbf{w}_x^T \mathbf{x}, \mathbf{w}_y^T \mathbf{y})|. \tag{7.4}$$

In CCA, the estimated correlation $\tilde{\rho}$ can be written as

$$\tilde{\rho} = \frac{\mathbf{w}_x^T \left(\sum_{i=1}^N \mathbf{x}_i \mathbf{y}_i^T \right) \mathbf{w}_y}{\sqrt{\mathbf{w}_x^T \left(\sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^T \right) \mathbf{w}_x \mathbf{w}_y^T \left(\sum_{i=1}^N \mathbf{y}_i \mathbf{y}_i^T \right) \mathbf{w}_y}} = \\ = \frac{\mathbf{w}_x^T \mathbf{C}_{xy} \mathbf{w}_y}{\sqrt{\mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y}}, \tag{7.5}$$

where $\mathbf{C}_{xx}$ is the within-set covariance matrix of the multidimensional variable $\mathbf{x}$, $\mathbf{C}_{xy}$ is the between-sets covariance matrix of $\mathbf{x}$ and $\mathbf{y}$ etc. It can be shown (Borga, 1998) that the canonical correlation can be calculated by solving the eigenvalue problem

$$\begin{pmatrix} \mathbf{C}_{xx} & 0 \\ 0 & \mathbf{C}_{yy} \end{pmatrix}^{-1} \begin{pmatrix} 0 & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{w}_x \\ \mathbf{w}_y \end{pmatrix} = \rho \begin{pmatrix} \mathbf{w}_x \\ \mathbf{w}_y \end{pmatrix}. \tag{7.6}$$

The first eigenvalue is the estimate of the canonical correlation, while the first eigenvector is the concatenation of the corresponding projection directions $\mathbf{w}_x$ and $\mathbf{w}_y$.

Obviously, GLM is obtained as a special case of CCA when either $\mathbf{x}$ or $\mathbf{y}$ is one-dimensional. Ordinary correlation is the special case obtained when both variables are one-dimensional.

7.1.2 Weighted correlation

A weight can be assigned to each sample, allowing different samples to affect the correlation estimate to a different extent. This is useful if each sample is accompanied by a certainty value. Less certain samples can then affect the correlation estimate less than more certain ones. Assuming that c_i is the certainty (or weight) associated with the i:th sample, the weighted equivalent of equation 7.2 becomes

$$\tilde{\rho} = \frac{\sum_{i=1}^N c_i x_i y_i}{\sqrt{\sum_{i=1}^N c_i x_i^2 \sum_{i=1}^N c_i y_i^2}}. \tag{7.7}$$

It is important, however, to consider the mean values of the signals. Before introducing weights, we could simply subtract the averages to obtain zero mean signals. Since samples with low weights should not affect the correlation estimates to a large extent, this is no longer the case. Instead, the weighted average must be subtracted before applying equation 7.7. That is, the correlation should be estimated as

$$\tilde{\rho} = \frac{\sum_{i=1}^N c_i (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum_{i=1}^N c_i (x_i - \mu_x)^2 \sum_{i=1}^N c_i (y_i - \mu_y)^2}}, \tag{7.8}$$

where

$$\mu_x = \frac{\sum_{i=1}^{N} c_i x_i}{\sum_{i=1}^{N} c_i} \tag{7.9}$$

and μ_y is calculated equivalently.

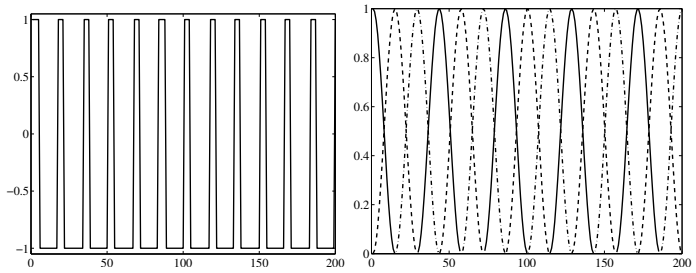
GLM and CCA are adapted in a similar fashion to accommodate weighted calculation of correlation coefficients.

7.2 Method

In order to use weighted correlation, the weights of all samples must be known. In some cases the weights may be available beforehand, for example if there is a natural way to measure the certainty of each sample. This is the case e.g. if a large number of reference signals is available, since the standard deviation of each sample may be calculated and its inverse used as certainty. More often, however, no natural certainty estimate is available. Fortunately, it is possible to find suitable weights using only the two signals whose correlation is to be calculated, by automatically finding outliers in the signals. This is accomplished by dividing both signals into a number of segments. The correlation between the signals is then calculated in each segment using either Pearson correlation, GLM or CCA depending on the dimensionality of the signals. By the use of weighted correlation the segments need not be disjoint subsets of the signals, but can instead be defined by partly overlapping smooth windows. This is preferable to disjoint subsets since it makes the algorithm less sensitive to the exact locations of the boundaries between different segments. Naturally, it is important to make sure that the sum of the weights for all windows is constant for all samples. Figure 7.1 shows a one-dimensional signal and a set of functions defining weights for each segment of the signal. In this example a truncated $\cos^2$ function is used to define each window, and consecutive windows overlap by 50%. The number of windows to use and the amount of overlap depend on the specific application.

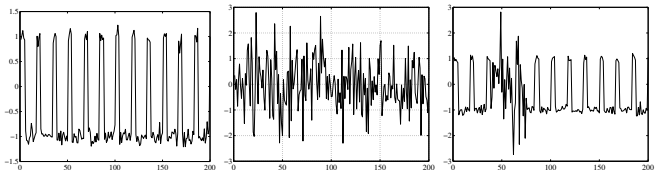
If we calculate the weighted correlation estimates between this signal and each of the signals shown in figure 7.2(a)-(c), using the windows shown above, a number of correlation coefficients are obtained. These are remapped to the similarity values shown in figure 7.2(d)-(f). The remapping is described below. Obviously, the similarity is approximately the same in all windows for the slightly noisy signal shown in figure 7.2(a), and approximately zero in all windows for the pure noise signal in figure 7.2(b). For the signal in figure 7.2(c), however, the similarity is significantly lower in the noisy segment than in the other segments. The ordinary, unweighted, correlation coefficients for these entire signals are 0.99, 0.04 and 0.73, respectively.

Let us now focus on the third signal. We have a sequence of similarities, one for each window. These values can easily be split into two groups: all the relatively

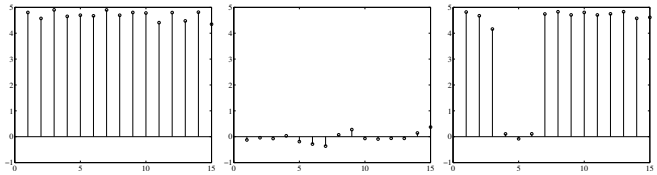


(a) Example signal (b) Window functions

Figure 7.1: An example signal and a set of possible weighting functions. The windows are defined by truncated $\cos^2$ functions and overlap by 50%.



(a) Signal 1 (b) Signal 2 (c) Signal 3



(d) Similarity 1 (e) Similarity 2 (f) Similarity 3

Figure 7.2: Three signals and their similarities (in each window) to the signal shown in figure 7.1.

high values and the three low values. Since there are far more high values than there are low ones, it is intuitively obvious that the low values could be considered as outliers and that the corresponding segments of the signal should therefore have lower weights. The algorithm arrives at the same conclusion by calculating the pairwise differences between the similarity values in every signal segment. The weight of each segment is then defined as a decreasing function of the average distance to the similarity values of all other segments. More formally w_k, the weight of the k:th segment, is calculated as

$$w_k = e^{-d_k^2/(2\sigma^2)}, \tag{7.10}$$

where

$$d_k = \frac{1}{L-1} \sum_{j \neq k} |\Lambda(\tilde{\rho}_k) - \Lambda(\tilde{\rho}_j)| \tag{7.11}$$

and

$$\Lambda(\rho) = \frac{\text{sign}(\rho)}{1 - \rho^2 + r}. \tag{7.12}$$

L is the number of segments. The function $\Lambda(\rho)$, which is similar to Wilks' lambda, can be used to map the estimated correlation coefficients to similarity values in order to obtain a more linear scale. That is, Λ compensates for the fact that a small increase of the resemblance between two signals with low original correlation increases the correlation more than a small increase of the resemblance between two signals with high original correlation. r is a regularization parameter which controls the behavior of Λ when ρ approaches 1.

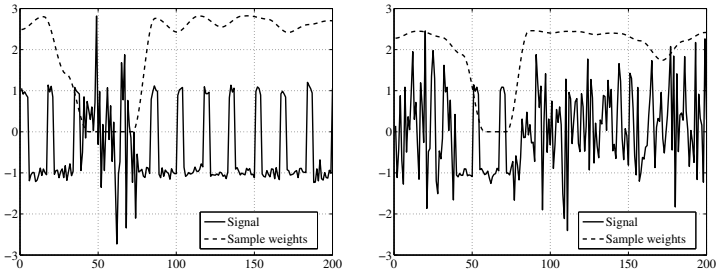
Finally, the weight of each sample can be calculated as the sum of all window functions at the sample position, multiplied by the weight for the respective segments of the signal. That is, the weight c_i of the i:th sample is calculated as

$$c_i = \sum_{k=1}^{L} w_k f_k(i), \tag{7.13}$$

where $f_k(i)$ is the value of the k:th window function at the position of the i:th sample. Thus for the signal in figure 7.2(c), we end up with the weighting shown in figure 7.3(a). The weighted correlation is 0.99, which is very close to the (unweighted) correlation between the first signal above and the reference signal. The automatically calculated weights cause the weighted correlation estimate to disregard the outliers. If the method instead is applied to any of the two other signals, an almost constant weighting is obtained. Thus, in those cases the weighted correlation estimate is the same as the ordinary correlation, which is desired for signals whose correlation to the reference signal is constant.

Figure 7.3(b) shows a fourth example of a signal whose correlation to the reference signal in figure 7.1(a) is to be calculated. This signal consists of only noise, except for a short segment where it is similar to the reference signal. The weighting obtained is also shown in the figure. The robust correlation between this signal and the reference is 0.03, which is very close to the correlation between the reference and the pure noise signal above. Together, the third and fourth examples illustrate how the underrepresented parts of a signal are disregarded, and how this may cause the robust correlation coefficient to be either higher or lower than the ordinary correlation.

The differences between the correlations in different segments can be calculated according to equation 7.11 only when the signals are one-dimensional. When the signals are multidimensional and GLM or CCA is used instead of the Pearson correlation, the projection directions $\mathbf{w}_x$ and $\mathbf{w}_y$ should also be taken into account. If the correlation coefficients in two segments are similar but $\mathbf{w}_x$ or $\mathbf{w}_y$ is different, the signals are related to each other in different ways and the two segments should



(a) Final weights for the signal shown in figure 7.2c.

(b) Another signal and its weighting.

Figure 7.3: Final weights for the signal shown in figure 7.2c and for another signal.

be treated accordingly. To handle this, the distance calculation is replaced by

$$d_k = \frac{1}{L-1} \sum_{j \neq k} \| \Lambda(\tilde{\rho}_k) \mathbf{w}_{x_k} \mathbf{w}_{y_k}^T - \Lambda(\tilde{\rho}_j) \mathbf{w}_{x_j} \mathbf{w}_{y_j}^T \|, \tag{7.14}$$

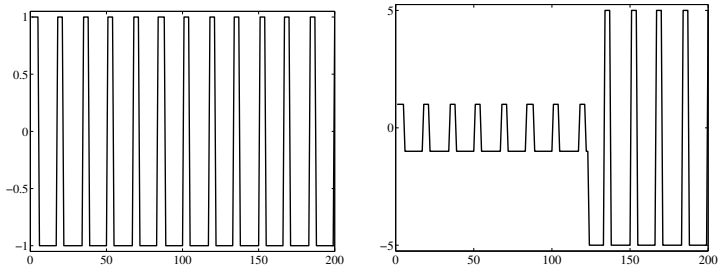
i.e. the difference between outer products of the projection directions (multiplied by Λ) is used to determine the distance between multidimensional correlations. The Frobenius norm is used. In the one-dimensional case, both $\mathbf{w}_x$ and $\mathbf{w}_y$ are 1, and equation 7.11 is thus a special case of equation 7.14.

If desired, additional parameters can be introduced to control how the similarity values in different segments affect the weighting. For example, the distances can be remapped such that small distances become even smaller, while large distances become larger. This has the effect that a small number of segments with almost identical similarity values will receive higher weights, while a larger group of segments with larger differences between their similarity values will receive lower weights.

Apart from similar correlation coefficients and projection directions, it might also be argued that the *ratio of variance* between the signals x and y should be similar. A case where this is important is illustrated in figure 7.4, which shows two one-dimensional signals whose variance ratio is significantly different in different parts of the signals. If small windows are used to measure local correlations between these signals, the correlation coefficients will be similar in all windows except those spanning the part of the signal where the variance changes. Hence an almost constant weighting will be obtained, and the robust correlation coefficient will therefore be very low. If the variance ratio between the two signals is also allowed to affect the weighting, this is avoided. This is achieved by, in each window k, calculating the logarithm of the variance ratio:

$$v_k = \log \frac{V_{x_k}}{V_{y_k}}, \tag{7.15}$$

where V_{x_k} and V_{y_k} are variance estimates of x and y weighted with the k:th window function. Distances between v_k in different windows are then calculated analogously to the distances between the similarity values, and affect the weighting in the same way. The logarithm is used to obtain symmetry; a change of variance ratio from 1 to 0.1 should affect the weighting to the same extent as a change from 1 to 10.



(a) x (b) y

Figure 7.4: Two signals with similar local correlations but different variance ratios.

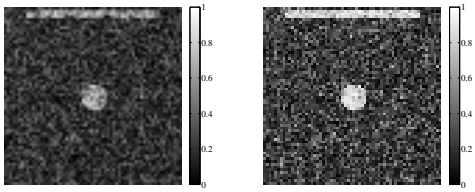
7.3 Application to analysis of functional MRI data

As mentioned in chapter 4, the signal to noise ratio of fMRI data is typically rather low. When the SNR is constant over time, the robust estimator of correlation does not provide any improvement of the activation detection accuracy. If, however, the signals from some or all voxels are corrupted by sudden bursts of noise, for example induced by instruments used for delivering the stimuli to the patient or subject, the proposed method is able to automatically identify and disregard the corrupted parts of the signals.

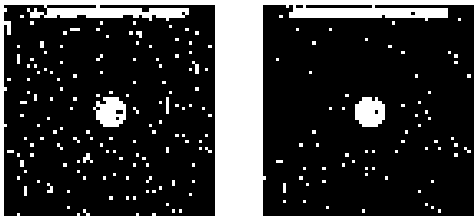
To demonstrate and evaluate the method, synthetic fMRI data was created by embedding BOLD-like signals in noise. The activated regions are a rectangle and a circle. Extra noise was added to 7 consecutive samples out of a total of 40 samples. Most often, fMRI signals are longer (approximately 100 - 200 samples), but short signals are of interest in some applications, e.g. where sliding window approaches are used (Gembris et al., 2000). One example is real-time analysis, which is necessary when fMRI is used in a feedback loop, i.e. when the stimuli delivered to the patient or subject depends on previous brain activity. Two examples of such scenarios are presented in Yoo et al. (2004) and deCharms et al. (2005).

In order to keep the example as simple as possible, ordinary low-pass filtering was applied to the data before the correlation between the signal from each pixel and the BOLD model was calculated. Figures 7.5(a) and (b) show correlation

maps obtained using ordinary correlation and the proposed method, respectively. Thresholded correlation maps are also shown. Obviously, better contrast between active and inactive pixels and fewer misclassified pixels are obtained when using the proposed estimator. A more quantitative evaluation is presented in figure 7.6, which shows the distributions of correlation coefficients when using ordinary correlation (a) and the proposed estimator (b). When the proposed estimator is used, the overlap between the correlation coefficients from active and inactive pixels is smaller than when ordinary correlation is used. This indicates that a lower number of misclassified pixels is obtained when the proposed estimator is used.

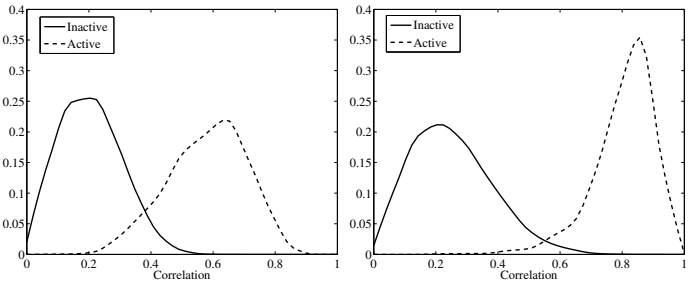


(a) Ordinary correlation (b) Proposed estimator



(c) Thresholded, ordinary correlation (d) Thresholded, proposed estimator

Figure 7.5: Correlation maps (original and thresholded) calculated using ordinary correlation and the proposed method, respectively.



(a) Ordinary correlation (b) Proposed estimator

Figure 7.6: Distributions of correlation coefficients for active and inactive pixels.

7.4 Application to detection of partially occluded objects

Correlation can also be used detect and localize objects in images. This is achieved by, at each position in an image, calculating the correlation between a small image patch and an image of the sought object. In practice, this operation is similar to convolving the image with the sought object, except that the convolution kernel is not mirrored. At the location of the object, a high correlation coefficient will be obtained. However, if the object is partly occluded, the estimated correlation will be lower and the object will be more difficult to detect. By using robust correlation estimation, this problem can be alleviated.

Figure 7.7(a) shows a photograph of an office with a desk and some computer equipment. There is also a book, with a post-it note on its cover, in the image. Figure 7.7(b) shows a close-up of the book without the post-it note. Both ordinary and robust correlation has been used to localize the book in the office image. To make the problem slightly more difficult, only the top half of the book was used in the correlation analysis. The resulting correlation maps are shown in figure 7.8. The positions where the highest correlation coefficients are found are marked in both images. When ordinary correlation is used, the maximum value (0.70) is obtained near the right edge of the image, far from the book. When robust correlation is used, the highest correlation coefficient (0.93) is obtained in the region occupied by the book.

(a) Image showing desk and computer

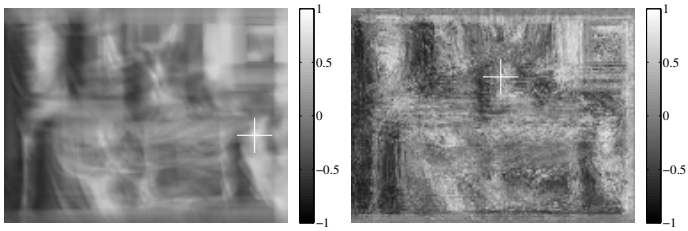
(b) Close-up of book

Figure 7.7: An image showing a typical office environment and a close-up of the book which is to be localized.

7.5 Discussion

Robust methods for correlation estimation have previously been described in e.g. Dehon et al. (2000). Those methods, however, do not consider the locality of groups of outliers. The estimator described here, on the other hand, controls the influence of each segment of the signal based on local estimates of the correlation. Therefore it is particularly suited for signals with local bursts of noise.

For the method to work well, the correlation in the outlier segments need to be



(a) Ordinary correlation (b) Robust correlation

Figure 7.8: Correlation coefficients obtained when searching for the book. The cross marks the location of the highest correlation coefficient.

sufficiently different from the correlation in the other parts of the signal. Otherwise the outliers will not be properly detected and the weights of all segments will be similar. In such cases, the proposed estimator will act like ordinary correlation analysis. The outlier detection is of course influenced by the choice of σ (see equation 7.10), but the estimator also depends on the number of segments and their amount of overlap. If the windows are too narrow, the local correlation estimates will vary greatly. On the other hand, if they are too wide, short bursts of noise may not affect the local correlation enough to be detected as an outlier. The variance ratio measurements are, of course, also sensitive to these parameters. Because of this parameter dependence, the algorithm works best in applications where the properties of the signal and the noise bursts are well-known and the parameters can be fine-tuned to match the specific situation.

8

Phase sensitive image reconstruction

This chapter continues the fat/water segmentation track last visited in chapter 3. First a brief overview of separate water and fat imaging is presented. This is followed by an explanation of why this imaging mode requires phase sensitive reconstruction and correction. Finally a novel method for solving this problem is introduced and described, along with a previously proposed method.

8.1 Obesity, a big problem

There is a tremendous prevalence in obesity worldwide. Studies indicate that more than 50 % of adults in the United States suffered from abdominal obesity during $2003 - 2004$ (Kullberg, 2007). It is well known that obesity, particularly male abdominal fat accumulation, is associated with high risk of developing common diseases such as diabetes, high blood pressure and cholesterol disturbances. Different measures of body composition have been used to predict the risk of these and other conditions. Examples range from very simple measures such as weight, body mass index (BMI) and waist circumference, to more advanced ones based on imaging of fat accumulation in the body using computed tomography (CT) or MRI.

8.2 Imaging of adipose tissue

8.2.1 T_1 weighted imaging

One of the most common MRI-based methods for measuring fat accumulation is simple T_1 weighted imaging. Fat is typically stored in the body as *adipose tissue*, which consists of approximately 80 % fat and 20 % water, protein and minerals. The T_1 relaxation time of fat is very short compared to that of most other types of tissue (according to de Bazelaire et al. (2004), at 1.5 T, the relaxation times for subcutaneous fat, liver and spleen are approximately 343 ms, 586 ms and 1057 ms, respectively). Therefore, adipose tissue is clearly visible as bright regions in

T_1 weighted images. Hence, the images can be segmented in order to determine the amount of adipose tissue. However, fat measurements based on T_1 images sometimes underestimate the fat content (Donnelly et al., 2003). This is likely due to partial volume effects at interfaces between adipose and other tissue, which cause voxels containing fat to be misclassified by the segmentation. Nevertheless, this type of images has been used to assess body composition in e.g. Kullberg (2007), and been shown to correlate well with fat measurements based on CT. However, Dixon imaging (Dixon, 1984) is considered to provide better results.

8.2.2 Dixon imaging

As mentioned in chapter 2, the Larmor frequencies of hydrogen in fat and water are somewhat different. The difference is 3.5 ppm, i.e. 220 Hz at a B_0 field strength of 1.5 T. Since the Larmor frequency of hydrogen in water is approximately 63.9 MHz, this may not seem very important. However, while signals from water and fat are of course in phase immediately after the RF excitation, only 2.3 ms later the signals are entirely out of phase! Another 2.3 ms later, i.e. 4.6 ms after the excitation, the signals are in phase again, and so on. This effect is utilized in Dixon imaging in order to obtain two images: w, showing the water content of each pixel, and f, showing the fat content of each pixel. In typical Dixon imaging, two images are acquired[1]: one image where the signals are out of phase ($I_1 = w - f$) and one image where they are in phase ($I_2 = w + f$). Ideally, the water and fat images w and f can be obtained as

$$w = \frac{I_1 + I_2}{2} \tag{8.1}$$

and

$$f = \frac{I_2 - I_1}{2} \tag{8.2}$$

respectively. By combining this technique with methods for absolute quantification, a true measure of the amount of fat in each pixel can be obtained. Since no segmentation is needed, an unbiased measure of the fat content of any region can then be obtained.

There is one problem, however. Because of experimental factors such as magnetic field inhomogeneity and excitation and receiver coil issues, the complex phase varies across the acquired images or volumes according to

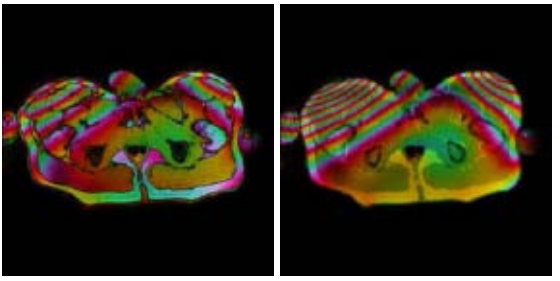
$$\tilde{I}_1 = I_1 e^{i\phi_1} \tag{8.3}$$

$$\tilde{I}_2 = I_2 e^{i\phi_2}, \tag{8.4}$$

where $\tilde{I}_1$ and $\tilde{I}_2$ are the acquired images and ϕ_1 and ϕ_2 are slowly varying, smooth phase fields. Figure 8.1(a) shows an example of an out of phase image with the

[1] Variations of the technique, where more than two images are used, also exist. This discussion, however, will focus on the so called two-point Dixon method.

additional phase variation caused by acquisition imperfections. The image intensity corresponds to the magnitude of the image, while the colors correspond to the complex phase. As is clearly visible in this image, there is a continuous phase variation over the image, in addition to the phase jumps of π at the borders between water and fat. Figure 8.1(b) shows an image where fat and water are in phase. This image does not show any phase jumps between water and fat, but the phase variations are twice as fast as in the out of phase image. The latter is caused by the additional time between excitation and image readout. An important thing to notice in these images is that the out of phase image has very low magnitude near borders between water and fat, since the signals from approximately equal amounts of water and fat in those regions cancel. In the in phase image the signals do not cancel, and the magnitude is unaffected.



(a) Out of phase (b) In phase

Figure 8.1: Example of acquired images with phase variations due to experimental effects.

Because of the phase variations shown in the images above, the water and fat images can not be calculated directly from the acquired images $\tilde{I}_1$ and $\tilde{I}_2$. Instead estimates $\hat{I}_1$ and $\hat{I}_2$ of the desired images I_1 and I_2 need to be reconstructed by correcting the phase errors in the acquired images. These estimates can then be used in the equations above to find w and f. $\tilde{I}_2$ is by far the easiest image to correct. For obvious reasons, w and f are both positive images, and thus $w + f$ is also positive. Hence all pixels in I_2 should have zero phase, and thus $\hat{I}_2 = |\tilde{I}_2|$, i.e. the magnitude of $\tilde{I}_2$. Figure 8.2 shows the in phase image displayed above after correction.

Unfortunately, finding $\hat{I}_1$ is more difficult. The reason for the higher complexity is that $w - f$ is not necessarily positive since some pixels contain more fat than water. Hence the phase of $\tilde{I}_1$ is the combined result of phase shifts due to tissue properties *and* the phase field ϕ_1. The phase shifts due to the sign of $w - f$ need to be preserved in order to retain the possibility to calculate w and f, and therefore ϕ_1 needs to be estimated in order to correct the image. After that, $\hat{I}_1$ is obtained as

$$\hat{I}_1 = \tilde{I}_1 e^{-i\hat{\phi}_1}, \tag{8.5}$$

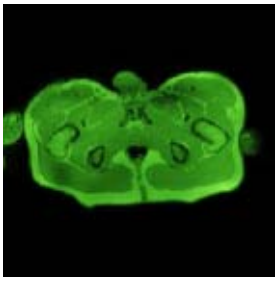


Figure 8.2: In phase image $\tilde{I}_2$ after correction.

where $\hat{\phi}_1$ is the estimate of the phase field ϕ_1.

The problem encountered here also appears in other types of MR imaging where complex or non-positive images are acquired, e.g. in velocity encoded imaging and inversion recovery imaging. There, too, the problem can be solved by estimating the phase field. Hence, an estimation procedure for finding $\hat{\phi}_1$ is not only useful in water/fat segmentation but also in other applications of MRI. The remainder of this chapter is devoted to the estimation of ϕ_1.

8.3 Phase estimation and correction

Separation of the phase of $\tilde{I}_1$ into the sign of $w - f$ and the additional phase field ϕ_1 requires some knowledge about these components. Without this information, any property of the phase variation of $\tilde{I}_1$ may equally well be attributed to either of them. Fortunately, there is a well-known difference between the behaviors of these phase variations: while ϕ_1 is smooth and varies relatively slowly, the phase of $w - f$ only takes on the values 0 and π and is thus piecewise constant with leaps of π at tissue boundaries. By clever use of this information, it is possible to design algorithms which find the two individual contributions of the phase of $\tilde{I}_1$.

8.3.1 Phase estimation using region growing

In Ma (2004) one such algorithm for phase estimation is proposed. Similar to a number of other methods such as those presented in Szumowski et al. (1994); Akkerman and Maas (1995), the method by Ma is based on region growing, i.e. starts from a seed point and grows the solution into neighboring pixels. The method utilizes the smoothness of ϕ_1 by visiting pixels in an order of phase difference, such that the region grows in the direction of the smallest phase gradient. If the phase difference between a pixel and its surrounding points is too large, the sign of that pixel is flipped. When all pixels have been visited, phase variations which are due to the sign of $w - f$ have been removed and thus the estimate $\hat{\phi}_1$ is obtained. To keep track of the phase gradient, the method uses a number of pixel

stacks; one stack for each phase difference interval of 10° from 0° to 90°. First a random pixel is selected as starting point for the region growing and put on one of the pixel stacks, and then the following three steps are repeated until all pixels have been visited:

1. Select a pixel from the lowest non-empty pixel stack.
2. Visit the four nearest neighbors, if not already visited, and place them onto the pixel stacks according to their phase difference in x or y direction, depending on the direction from which the pixel is visited. If the value is in the $0^\circ - 10^\circ$ interval the pixel is placed on the first pixel stack, if the value is in the $10^\circ - 20^\circ$ interval it is placed on the second stack, etc.
3. Determine the phase value of the selected pixel by inspecting the neighboring pixels that have already been checked. If the phase difference between the seed pixel and a summation of already visited pixels within a defined boxcar region exceeds 90°, the sign of the seed pixel is flipped.

The algorithm is rather straight-forward, but the quantization of the phase difference may appear to be a bit ad-hoc. For most images the method works rather well. In some cases, however, regions may be misclassified. This is particularly common in regions where ϕ_1 varies relatively fast, and where the region growing is restricted to occur along thin structures, e.g. in the vicinity of air cavities where the signal level is very low. Two examples, one where the method works well and one where problems are encountered, are shown in figure 8.3. Images before and after correction are shown. It may be difficult to see, but in the right image a region near the left arm is misclassified (some of the subcutaneous fat is red instead of green). This is shown in greater detail in paper IV.

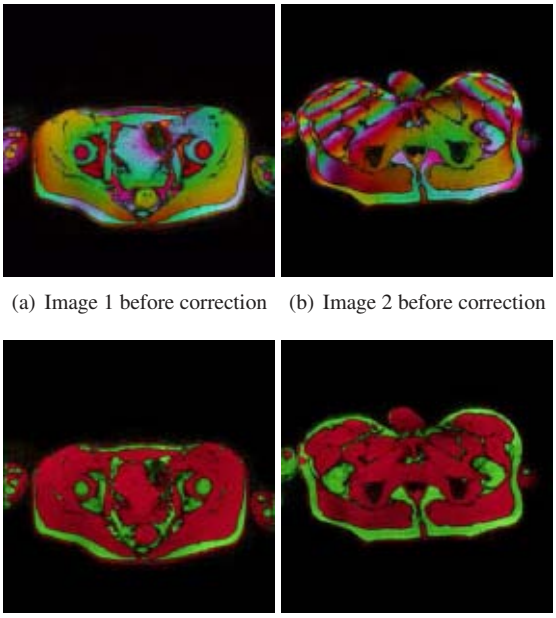
8.3.2 Phase estimation using the inverse gradient

It is a trivial and well-known fact that every image has an associated gradient field, which describes how the pixel values change with the position in the image. In the case of a two-dimensional image I, each point in the gradient field is a two-dimensional vector pointing in the direction of greatest local increase of pixel value, whose norm is proportional to the derivative of the image in that direction. The gradient $\mathbf{g}$ of the image I is typically denoted

$$\mathbf{g} = \nabla I. \tag{8.6}$$

A less well-known fact is that it is also possible to "go the other way" and find the image I associated with a certain gradient field $\mathbf{g}$. This operation may be referred to as the *inverse gradient* and denoted

$$I = \nabla^{-1}\mathbf{g} + C, \tag{8.7}$$



(a) Image 1 before correction (b) Image 2 before correction

(c) Image 1 after correction (d) Image 2 after correction

Figure 8.3: Results of the region growing method for phase correction. The method obviously works well for the left image, while a region close to the left arm is misclassified in the right image.

where C is an unknown additive constant. This is of course obvious for one-dimensional signals, where the gradient is actually the scalar derivative g and I is the integral of g:

$$I = \int g + C, \tag{8.8}$$

but for multidimensional signals such as images or volumes the process is less straight-forward. It can, however, be shown (Song et al., 1995; Horn and Brooks, 1986) that by solving a Poisson equation with inhomogeneous Neumann boundary conditions:

$$\begin{cases} \Delta I(\mathbf{x}) = \nabla \cdot \mathbf{g}(\mathbf{x}), & \text{all } \mathbf{x} \in \Omega, \\ \frac{\partial I}{\partial n}(\mathbf{x}) = \mathbf{n} \cdot \mathbf{g}(\mathbf{x}), & \text{all } \mathbf{x} \in \partial\Omega, \end{cases} \tag{8.9}$$

the image I which has a gradient field as close as possible to $\mathbf{g}$ (in a least squares sense) is obtained. Ω is the region in which the equation is solved, and $\mathbf{n}$ is a normalized, outwards directed, normal vector to the boundary of this region. By solving this equation, the I which minimizes

$$\int_{\Omega} |\nabla I(\mathbf{x}) - \mathbf{g}(\mathbf{x})|^2 \mathbf{dx} \tag{8.10}$$

is obtained. That is, the sum of squared euclidean distances between $\mathbf{g}(\mathbf{x})$ and the gradient of I, for all $\mathbf{x}$ in Ω, is minimized. The equation needs to be solved numerically. If the solver is designed carefully, Ω can have arbitrary shape as long as it is connected. A more detailed description of the inverse gradient and its properties is available in Farnebäck et al. (2007).

So why is the inverse gradient important in the context of phase estimation and correction? Going back to the differences between the two contributions to the total phase of $\tilde{I}_1$ (the sign of $w - f$ and the phase field ϕ_1), a possible way to eliminate one of them presents itself. If the phase of $\tilde{I}_1$ is doubled, we obtain

$$\tilde{I}_1^* = |\tilde{I}_1| e^{i2 \arg \tilde{I}_1} = |w - f| e^{i2\phi_1}. \tag{8.11}$$

Since the phase is doubled, it is no longer affected by the sign of $w - f$, and thus only ϕ_1 contributes to the phase of $\tilde{I}_1^*$! While this is definitely a step in the right direction, it does not completely solve the problem of obtaining $\hat{\phi}_1$. Since the phase is only known modulo 2π, it is not possible to directly obtain $\hat{\phi}_1$ as half the phase of I_1^*. However, if the phase varies sufficiently slowly (less than π per pixel), it is possible to estimate the phase gradient of I_1^*:

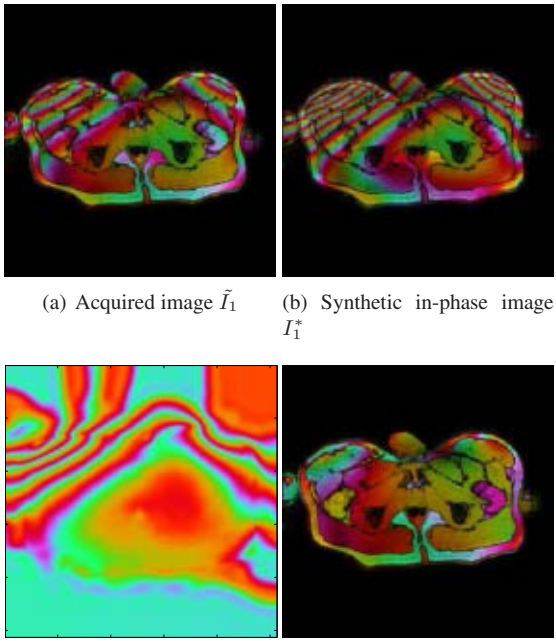
$$\mathbf{g} = \nabla \arg I_1^*, \tag{8.12}$$

Since the gradient $\mathbf{g}$ is determined locally, it does not contain the wraparounds present in the actual phase field $\arg I_1^*$. Hence it is possible to divide $\mathbf{g}$ by 2 without any uncertainty about the interpretation of the obtained value. By finally applying the inverse gradient operator to $\mathbf{g}/2$, a wrap-free phase field $\hat{\phi}_1$ is obtained. This field is an estimate of ϕ_1 and can be plugged into equation 8.5 to obtain $\hat{I}_1$, which is used in conjunction with $\hat{I}_2$ to calculate w and f.

In image regions where the signal strength is very low, e.g. in air cavities or close to the border between water and fat, where the signals cancel, the phase is very sensitive to noise and thus should not affect the gradient $\mathbf{g}$ to the same extent as more certain phase values where the signal magnitude is large. This is solved by a technique known as *normalized convolution* (Knutsson and Westin, 1993), which is used to interpolate $\mathbf{g}$ in low-certainty regions.

The estimation procedure is illustrated in figure 8.4, where different steps in the algorithm are shown. Panel (a) shows the acquired out of phase image $\tilde{I}_1$, panel (b) shows the synthetic in phase image I_1^*, panel (c) shows the estimated phase field $\hat{\phi}_1$ and panel (d) shows a partly corrected out of phase image.

As is clearly visible in figure 8.4(d), the result of this procedure is not perfect; the image is not entirely real valued after correction. This is explained by artifacts in $\mathbf{g}$ caused by noise in $\tilde{I}_1$ and by the interpolation of the gradient field. To overcome this problem, the correction is iterated a number of times, progressively converging to a solution with only two phase values, π apart, corresponding to water and fat. However, these phase values are not necessarily 0 and π. Since



(a) Acquired image $\tilde{I}_1$ (b) Synthetic in-phase image I_1^*

(c) Estimated phase field $\hat{\phi}_1$ (d) Partly corrected image

Figure 8.4: Out of phase image and different steps in the phase estimation procedure.

the image should be real valued, these are the only allowed phase values, but like all integration procedures the inverse gradient is only able to restore the phase up to an unknown additive constant. This problem is easily solved since only two possible mappings between the phase values present in the image and the allowed values exist; one of the existing phase values should be mapped to 0 and the other to π. For typical images it is not difficult to choose one of these mappings. The contrast of the magnitude of the acquired images is largely due to T_1 relaxation, and as was stated above the strongest signals obtained in such images are typically from adipose tissue. Thus, the phase value present where the signal magnitude is highest corresponds to fat, while the other phase value corresponds to water.

Figure 8.5 shows the corrected image $\hat{I}_1$, obtained after a number of iterations. This image is real valued and its phase has discontinuities at borders between different types of tissue. The fat is shown in green while the water is shown in red.

Finally, figure 8.6 shows reconstructed water and fat images. These are obtained from the corrected out of phase and in phase images $\hat{I}_1$ and $\hat{I}_2$ by using equations 8.1 and 8.2 in section 8.2.2.

An advantage of this method is that can easily be generalized to three-dimensional

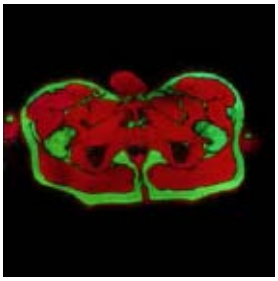
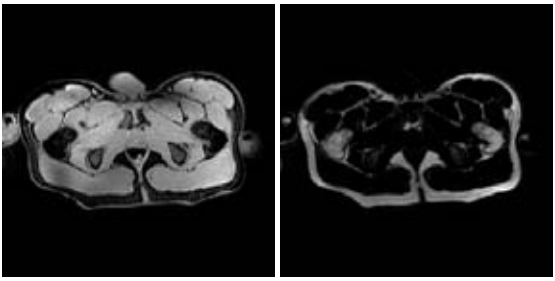


Figure 8.5: Out of phase image after correction.



(a) Water (b) Fat

Figure 8.6: Reconstructed water and fat images obtained from I_1 and I_2 after phase correction.

images. A similar generalization is possible also for methods based on region growing, but that is usually more complicated. Three-dimensional phase estimation and correction is important since the phase in thin sheets of fat may sometimes be very difficult to estimate from a two-dimensional image, while the same procedure is trivial in three dimensions.

9

Review of papers

In this chapter, the papers included in the second part of the thesis are introduced. Each of the papers covers approximately the same material as the corresponding chapter (5 – 8). In some parts, however, the descriptions in the papers are more or less detailed than those in the preceding chapters. Some of the papers also provide more experimental results. Some figures in the included papers may be difficult to interpret since they are reproduced in grayscale. These figures are available in color in the corresponding chapter.

9.1 Paper I: On rotational invariance in adaptive spatial filtering of fMRI data

This paper presents a modification of a previously proposed method for adaptive analysis of fMRI data. The previous method is based on restricted canonical correlation analysis, which enables simultaneous spatial filtering and activation detection. It is demonstrated that the constraints imposed on the spatial filters precludes rotationally invariant detection performance, i.e. that the sensitivity of the method depends on the orientation of the activated regions. In the modified method, another set of constraints are imposed on the filters. The modified method is shown to be invariant to the orientation of activated regions, thereby providing increased detection performance. The experimental results shown in chapter 5 are included in this paper, along with more results showing the increased detection performance. The paper has been published in NeuroImage.

9.2 Paper II: Signal and anatomical constraints in adaptive filtering of fMRI data

This paper describes another method for adaptive filtering of fMRI data. The method is based on bilateral filtering and only averages signals if they are similar and reside in the same type of brain tissue. This filtering strategy is compared to

ordinary low-pass filtering and to bilateral filtering where only signal *or* anatomical similarity is used to constrain the filters. Results show that the combined approach provides superior detection performance. The analysis algorithm is also described in chapter 6, although the paper contains some experimental results which are not presented there. The paper was presented at IEEE International Symposium on Biomedical Imaging (ISBI) 2007.

9.3 Paper III: Robust correlation analysis with an application to functional MRI

A method for robust estimation of correlation is presented along with generalizations to multivariate correlation analysis (the general linear model and canonical correlation analysis). The method is based on identifying outlier segments in the signals, and disregarding those segments when the correlation is calculated. Since correlation is used in a wide range of applications, the presented method can be applied in a large number of different fields. In the paper, the method is demonstrated on synthetic fMRI data. Results show improved detection performance when the data is corrupted by temporally localized bursts of noise. Chapter 7 is based on the material presented in this paper.

9.4 Paper IV: Phase sensitive reconstruction for water/fat separation in MR imaging using inverse gradient

Unlike the other papers, this one is not related to functional MRI. The focus is instead on Dixon imaging, which is used to obtain separate images of water and fat. While most MR images are real and positive, the images acquired in Dixon imaging are complex. The phase is often affected by artifacts and needs to be corrected before the desired water and fat images can be calculated. This paper presents a phase correction method based on integration of phase gradients. The method is compared to a correction method based on region growing, and qualitative results indicate that the new method provides more reliable phase correction in the presence of severe artifacts. A slightly more thorough description of the method is available in chapter 8. The paper has been presented at International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2007.

10

Discussion

Chapters 5 – 8 each introduce one method for MRI data processing. Some of these chapters are concluded by a short discussion. This chapter complements these discussions, and also gives a few pointers to further research.

10.1 Adaptive spatial filtering of fMRI data

Two methods for noise reduction by spatial filtering have been presented. While these algorithms are based on very different strategies, they share the basic idea that the analysis should be adaptive, i.e. that the filtering performed at each location should depend on local patterns of activation in the data to be analyzed. This idea is also shared with other analysis methods, which have been proposed in for example Shafie et al. (1998), Shafie et al. (2003) and Siegmund and Worsley (1995). Still, despite the growing number of adaptive methods providing good detection performance, most people working with fMRI are still using simple, non-adaptive low-pass filtering to analyze their data. Most likely, one important reason for this is that low-pass filtering is a very well-known method for noise reduction. Another probable reason is related to the accessibility of the different algorithms. Low-pass filtering is implemented in the software package SPM, while most of the adaptive filtering schemes are only available as scientific papers and, in some cases, reference implementations which are difficult to use for most people. Thus, in order to increase the general interest for adaptive methods for analysis of fMRI data, the algorithms need to be made available to the people who might actually use them, either clinically or for research. Yet another reason may be that parametric statistics can be used for thresholding the correlation maps when fixed low-pass filtering is used, while adaptive filtering often requires non-parametric statistics based on resampling. Since resampling is very time-consuming, this is a disadvantage of adaptive methods.

It may be argued that low-pass filtering does, in many experiments, detect activated brain regions with sufficient accuracy. However, as mentioned in chapter

4, the properties of real data vary between different experiments. In some experiments, the BOLD response is strong and the activated regions are large. Then, very good results can often be obtained by using simple low-pass filtering. In other cases, however, the benefit of using adaptive filtering may be much larger. Naturally, this depends both on the performed task, the available scanner equipment and the required accuracy.

Different methods sometimes provide equivalent average detection accuracy, but make different types of mistakes. For instance, at a given level of specificity, the CCA-based method is biased towards enlarging the activated regions, while the method based on bilateral filtering instead finds spurious activation in other parts of the brain. This makes it difficult to perform a generally valid evaluation of the different methods, and hence, no attempt to compare them to each other has been made in this thesis. Instead, specific properties of the methods, such as rotational invariance or preservation of edges, have been emphasized in the descriptions of the algorithms.

The rotationally invariant CCA-based method is a modification of a previously proposed algorithm, and has been shown to provide improved detection performance. While the angular sensitivity variations are not very large even in the previous method, this improvement may still be important in some types of experiments, particularly where highly anisotropic activation can be expected. One example of such an experiment is mapping of the visual cortex, where visual stimuli in different orientations generate thin lines of activation (Kim et al., 2000).

Since the CCA-based method optimizes the filters to maximize the correlation to the BOLD model, there is a risk of misclassifying inactive pixels or voxels near the edge of an active region. This may be a problem in some types of experiments, and further research should aim at reducing this risk. It may, for example, be beneficial to impose additional constraints on the adaptive filters. This could be done e.g. by including anatomical priors or by regularizing the filter optimization by requiring that similar filters are used in adjacent pixels or voxels.

The method based on bilateral filtering does not attempt to maximize the correlation coefficient. Hence, it typically does not spread BOLD-like signals beyond active regions. Rather, one of the most important strengths of this method is that it averages over consistent regions. However, when there are no clear edges between active and inactive parts of the brain, consistent regions are hard to find. As shown in chapter 6, the benefit of using bilateral filtering is smaller in such cases. A possible solution to this problem is described in section 6.5: By using smooth basis filters instead of individual voxels, better adaptation to smooth active regions should be obtained.

10.2 Robust correlation estimation

Unlike the methods discussed above, the last fMRI-related method works in the temporal domain. By identifying segments of a signal where the correlation to the BOLD model deviates from its typical value, this method can disregard signal segments with atypical behavior. This improves the detection performance in the presence of bursts of noise. Such bursts may e.g. originate from equipment used to deliver stimuli to the patient or subject. As demonstrated in chapter 7, the method is also useful in other applications where correlation is used to measure similarity between signals.

A number of parameters control how outlier segments are detected, and these parameters need to be properly tuned for the method to work well. If the characteristics of the signal and the noise bursts are well known, this is not a problem. Otherwise, however, the parameter dependence may preclude the use of this method. Further research should aim at development of methods for automatic parameter selection, as this would be a very important improvement.

10.3 Phase sensitive image reconstruction

The last method presented in this thesis is used to correct the phase of complex MR images. In chapter 8, this method is applied to Dixon imaging for separation of water and fat. A qualitative comparison to another method is presented, and visual results indicate that the presented method is more robust to severe artifacts. Unlike the fMRI methods, however, this method is relatively straight-forward to evaluate quantitatively. Unfortunately, no such results are available at this time.

The type of phase artifact which is present in Dixon imaging also affect inversion recovery images. Without modification, this method should be applicable also to correction of such images.

Bibliography

Akkerman, E., Maas, M., 1995. A region-growing algorithm to simultaneously remove dephasing influences and separate fat and water in two-point Dixon imaging. In: Proceedings of the ISMRM Annual Meeting. ISMRM, Nice, France, p. 649.

Angle, S., Bhagtani, R., Chheda, H., 2005. Biometrics: A further echelon of security. In: UAE International Conference on Biological and Medical Physics.

Borga, M., 1998. Learning multidimensional signal processing. Ph.D. thesis, Linköping University, Sweden, SE-581 83 Linköping, Sweden, dissertation No 531, ISBN 91-7219-202-X, http://www.imt.liu.se/∼magnus/.

Boynton, G., Engel, S., Glover, G., Heeger, D., 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. The Journal of Neuroscience 16 (13), 4207–4221.

Buxton, R., Wong, E., Frank, L., 1998. Dynamics of blood flow and oxygenation changes during brain activation: the Balloon model. Magnetic Resonance in Medicine 39 (6), 855–864.

Calhoun, V., Pekar, J., McGinty, V., Adali, T., Watson, T., 2002. Different activation dynamics in multiple neural systems during simulated driving. Human Brain Mapping 16 (3), 158–167.

Das, S., Sen, P., 1994. Restricted canonical correlations. Linear Algebra and its Applications 210, 29–47.

de Bazelaire, C. M. J., Duhamel, G. D., Rofsky, N. M., Alsop, D. C., 2004. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: Preliminary results. Radiology 230 (3), 652–659.

deCharms, R. C., Maeda, F., Glover, G. H., Ludlow, D., Pauly, J. M., Soneji, D., Gabrieli, J. D., Mackey, S. C., Dec 2005. Control over brain activation and pain learned by using real-time functional MRI. Proceedings of the National Academy of Sciences of the United States of America 102 (51), 18626–18631.

Dehon, C., Filzmoser, P., Croux, C., 2000. Robust methods for canonical correlation analysis. Data Analysis, Classification and Related Methods, 321–326.

Dixon, W., 1984. Simple proton spectroscopic imaging. Radiology 153(1), 189–194.

Donnelly, L., O'Brien, K., Dardzinski, B., 2003. Using a phantom to compare MR techniques for determining the ratio of intraabdominal to subcutaneous adipose tissue. American journal of roentgenology 180(4), 993–998.

Farnebäck, G., Rydell, J., Ebbers, T., Andersson, M., Knutsson, H., October 2007. Efficient computation of the inverse gradient on irregular domains. In: IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'07). Rio de Janeiro, Brasil, accepted.

Friman, O., Borga, M., Lundberg, P., Knutsson, H., February 2002a. Detection of neural activity in fMRI using maximum correlation modeling. NeuroImage 15 (2), 386–395.

Friman, O., Borga, M., Lundberg, P., Knutsson, H., June 2002b. Exploratory fMRI analysis by autocorrelation maximization. NeuroImage 16 (2), 454–464.

Friman, O., Borga, M., Lundberg, P., Knutsson, H., 2003. Adaptive analysis of fMRI data. NeuroImage 19 (3), 837–845.

Friman, O., Borga, M., Lundberg, P., Knutsson, H., June 2004. Detection and detrending in fMRI data analysis. NeuroImage 22 (2), 645–655.

Friman, O., Carlsson, J., Lundberg, P., Borga, M., Knutsson, H., February 2001. Detection of neural activity in functional MRI using canonical correlation analysis. Magnetic Resonance in Medicine 45 (2), 323–330.

Friman, O., Westin, C.-F., 2005. Resampling fMRI time series. NeuroImage 25 (3), 859–867.

Friston, K., Fletcher, P., Josephs, O., Holmes, A., Rugg, M., Turner, R., 1998. Event-related fMRI: Characterizing differential responses. NeuroImage 7 (1), 30–40.

Gembris, D., Taylor, J., Schor, S., Frings, W., Suter, D., Posse, S., 2000. Functional magnetic resonance imaging in real time (FIRE): Sliding-window correlation analysis and reference-vector optimization. Magnetic Resonance in Medicine 43 (2), 259–268.

Godtliebsen, F., Chu, C.-K., Sørbye, S. H., Torheim, G., 2001. An estimator for functional data with application to MRI. IEEE Transactions on Medical Imaging 20 (1), 36–44.

Godtliebsen, F., Spjøtvoll, E., Marron, J. S., 1997. A nonlinear Gaussian filter applied to images with discontinuities. Nonparametric Statistics 8, 21–43.

Haacke, M. E., Brown, R. W., Thompson, M. R., Venkatesan, R., Haacke, M. E., Brown, R. W., Thompson, M. R., Venkatesan, R., June 1999. Magnetic Resonance Imaging: Physical Principles and Sequence Design. Wiley-Liss.

Hajnal, J., Saeed, N., Soar, E., Oatridge, A., Young, I., Bydder, G., 1995. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. Journal of Computer Assisted Tomography 19 (2), 289–296.

Horn, B. K. P., Brooks, M. J., February 1986. The variational approach to shape from shading. Computer Vision, Graphics, and Image Processing 33, 174–208.

Hotelling, H., 1936. Relations between two sets of variates. Biometrika 28, 321–377.

Hunt, M. J., 1999. Spectral signal processing for ASR. In: Proc. IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU).

Jolliffe, I. T., 1986. Principal Component Analysis. Springer-Verlag, New York.

Kiebel, S., Goebel, R., Friston, K., 2000. Anatomically informed basis functions. NeuroImage 11 (6), 656–667.

Kim, D.-S., Duong, T. Q., Kim, S.-G., 2000. High-resolution mapping of iso-orientation columns by fMRI. Nature Neuroscience 3 (2), 164–169.

Knutsson, H., Andersson, M., June 2005. Morphons: Paint on priors and elastic canvas for segmentation and registration. In: Proceedings of the Scandinavian Conference on Image Analysis (SCIA). Joensuu.

Knutsson, H., Westin, C.-F., June 1993. Normalized and differential convolution: Methods for interpolation and filtering of incomplete and uncertain data. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 515–523.

Knutsson, H., Wilson, R., Granlund, G. H., March 1983. Anisotropic non-stationary image estimation and its applications — Part I: Restoration of noisy images. IEEE Transactions on Communications 31 (3), 388–397.

Kullberg, J., 2007. Assessment of body composition using magnetic resonance imaging. Ph.D. thesis, Uppsala University, Sweden, iSBN 978-91-554-6828-6.

Levitt, M. H., 2001. Spin Dynamics, Basics of Nuclear Magnetic Resonance. Wiley.

Liang, Z.-P., Lauterbur, P. C., 2000. Principles of Magnetic Resonance Imaging, A signal Processing Perspective. SPIE Optical Engineering Press.

Liangsuo, M., Wang, B., Chen, X., Xiong, J., 2007. Detecting functional connectivity in the resting brain: a comparison between ICA and CCA. Magnetic Resonance Imaging 252 (1), 47–56.

Locascio, J., Jennings, P., Moore, C., Corkin, S., 1997. Time series analysis in the time domain and resampling methods for studies of functional magnetic resonance brain imaging. Human Brain Mapping 5 (3), 168–193.

Ma, J., 2004. Breath-hold water and fat imaging using a dual-echo two-point Dixon technique with an efficient and robust phase-correction algorithm. Magnetic Resonance in Medicine 52 (2), 415–419.

Maldjian, J. A., Gottschalk, A., Detre, J. A., Alsop, D., May 1999. Basal ganglia and white matter activation using fMRI at 4T. In: Proceedings of the ISMRM Annual Meeting. ISMRM.

McKeown, M., Makeig, S., Brown, G., Jung, T., Kindermann, S., Bell, A., Sejnowski, T., 1998. Analysis of fMRI data by blind separation into independent spatial components. Human Brain Mapping 6 (3), 160–188.

Nandy, R., Cordes, D., 2003. Novel ROC-type method for testing the efficiency of multivariate statistical methods in fMRI. Magnetic Resonance in Medicine 49 (6), 1152–1162.

Nandy, R., Cordes, D., 2004. Improving the spatial specificity of canonical correlation analysis in fMRI. Magnetic Resonance in Medicine 52, 947–952.

Orchard, J., Greif, C., Bjornson, G. G. B., Atkins, M., 2003. Simultaneous registration and activation detection for fMRI. IEEE Transactions on Medical Imaging 22 (11), 1427–1435.

Ou, W., Golland, P., 2005. From spatial regularization to anatomical priors in fMRI analysis. Information Processing in Medical Imaging 19, 88–100.

Penney, G. P., Weese, J., Little, J. A., Desmedt, P., Hill, D. L. G., Hawkes, D. J., 1998. A comparison of similarity measures for use in 2-D-3-D medical image registration. IEEE Transactions on Medical Imaging 17, 586 – 595.

Pluim, J. P. W., Maintz, J. B. A., Viergever, M. A., 2003. Mutual-information-based registration of medical images: A survey. IEEE Transactions on Medical Imaging 22 (8), 986–1004.

Rajapakse, J., Kruggel, F., Maisog, J., Cramon, D., 1998. Modeling hemodynamic response for analysis of functional MRI time-series. Human Brain Mapping 6 (4), 283–300.

Schmitt, F., Stehling, M. K., Turner, R., May 1998. Echo-Planar Imaging: Theory, Technique and Application. Springer.

Shafie, K., Sigal, B., Siegmund, D., Worsley, K., 2003. Rotation space random fields with an application to fMRI data. The Annals of Statistics 31.

Shafie, K., Worsley, K., Wolforth, M., Evans, A., 1998. Rotation space: Detecting functional activation by searching over rotated and scaled filters. NeuroImage 7, S755.

Siegmund, D. O., Worsley, K. J., 1995. Testing for a signal with unknown location

and scale in a stationary gaussian random field. The Annals of Statistics 23 (2), 608–639.

Sole, A., Ngan, S.-C., Sapiro, G., Hu, X., Lopez, A., February 2001. Anisotropic 2-D and 3-D averaging of fMRI signals. IEEE Transactions on Medical Imaging 20 (2), 86–93.

Song, S. M.-H., Napel, S., Pelc, N. J., Glover, G. H., May 1995. Phase unwrapping of MR phase images using Poisson equation. IEEE Transaction on Image Processing 4 (5), 667–676.

Szumowski, J., Coshow, W., Li, F., Quinn, S., 1994. Phase unwrapping in the three-point Dixon method for fat suppression MR imaging. Radiology 192, 555–561.

Tabelow, K., Polzehl, J., Voss, H. U., Spokoiny, V., 2006. Analyzing fMRI experiments with structural adaptive smoothing procedures. NeuroImage 33, 55–62.

Tettamanti, M., Paulesu, E., Scifo, P., Maravita, A., Fazio, F., Perani, D., Marzi, C. A., 2002. Interhemispheric transmission of visuomotor information in humans: fMRI evidence. Journal of Neurophysiology 88, 1051–1058.

Tomasi, C., Manduchi, R., January 1998. Bilateral filtering for gray and color images. In: IEEE International Conference on Computer Vision 98. IEEE, Bombay, India, pp. 839–846.

Walker, S. A., Miller, D., Tanabe, J., 2006. Bilateral spatial filtering: Refining methods for localizing brain activation in the presence of parenchymal abnormalities,. NeuroImage.

Warntjes, J., Dahlqvist, O., Lundberg, P., 2007. Novel method for rapid, simultaneous T_1, T_2^*, and proton density quantification. Magnetic Resonance in Medicine 57 (3), 528–537.

Worsley, K., 1994. Local maxima and the expected Euler characteristic of excursion sets of χ^2, F and t fields. Advances in Applied Probability 26, 13–42.

Worsley, K., Friston, K., 1995. Analysis of fMRI time-series revisited-again. NeuroImage 2 (3), 173–181.

Yoo, S. S., Fairneny, T., Chen, N. K., Choo, S. E., Panych, L. P., Park, H., Lee, S. Y., Jolesz, F. A., July 2004. Brain-computer interface using fMRI: spatial navigation by thoughts. Neuroreport 15 (10), 1591–1595.