AALBORG UNIVERSITY
STUDENT REPORT

Something about preprocessing methods and nice results for IDP study

3nd semester Masters, Biomedical Engineering & Infomatics - Fall 2018

Project group: 19gr9411

Christian Korfitz Mortensen, Martin Alexander Garenfeld

ii

Preface

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Contents

1 Background **2**
1.1 MRI physics 2
1.2 MR image reconstruction 4
1.3 Functional MRI 5
1.4 General MRI Pre-processing 7
1.5 Independent Component Analysis 11
1.6 Principal Component Analysis 13

2 Methods **16**
2.1 FSL FIX 16

WORKSHEETS

1 | Background

1.1 MRI physics

Write top

Magnetic resonance imaging (MRI) is founded on the principle of nuclear magnetic resonance (NMR), which exploits the magnetic properties of the hydrogen nucleus that contains a single proton. The proton is not static, but rotates around its own axis. As the proton is positively charged it creates a magnetic moment in the direction described by the thumb rule, and can interact with an external magnetic field. The human body consists of approximately 10% hydrogen atoms, but as the hydrogen nuclei spins are randomly orientated, the net magnetic moment equals zero, as the nuclei cancel each other out. Placing the body in a strong magnetic field will align the nuclei. A property of the hydrogen nucleus is its quantum spin rate, which can either be $\frac{1}{2}$ or $-\frac{1}{2}$ - either in the direction or the opposite direction of the main magnetic field. Most will align in the direction of the magnetic field, while some align in the opposite direction, possibly as a result of heat radiation absorbed by the nuclei. The direction of the nucleus is determined by its energy level, leaving the former in a low energy state and the latter in a high energy state. The nuclei do not simply point in the in the direction or opposite the direction of the magnetic field, but precess. [1] The rate of precession can be calculated by the Lamour frequency:

$$f = y * B_0 \tag{1.1}$$

, where f is precession frequency, y is gyroscopic ratio and B_0 is magnetic field strength. The equation states that the precession frequency is proportional to the strength of the magnetic field. After canceling out all opposing precessing nuclei, the net magnetization, or longitudinal magnetization, will point in the direction of the external magnetic field. However, the longitudinal magnetization can not be detected directly as it points in the direction of the strong external magnetic field. Additional techniques are therefore used in NMR, to facilitate a detectable signal. [1] A depiction of how the nucleus precessing can either align along or opposite to the magnetic field depending on its energy state, can be found in figure 1.1.

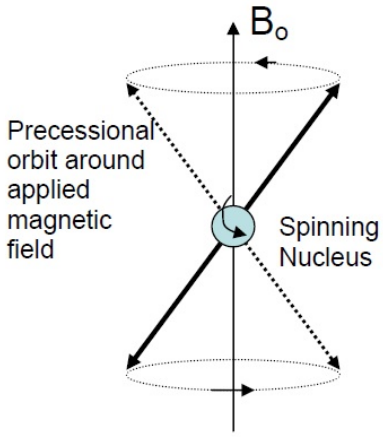


Figure 1.1: The figure illustrates how the nucleus precess and spin in relation to the applied magnetic field B_0 surrounding it. The vectors, indicating the precessing, can go opposite or along the magnetic field depending on the nucleus energy state. [2]

A radio frequency pulse (RF pulse) tuned to the precession of the nuclei is transmitted in the vicinity of the nuclei. The RF pulse is absorbed by the nuclei and more, favorably half of the targeted nuclei population, will enter the high energy state, leaving the longitudinal magnetization to equal zero. The number of nuclei that flip is determined by the amount of energy the RF pulse injects, and the nuclei only exchange energy efficiently if the frequency of the energy from the RF pulse matches the precession rate. The RF pulse furthermore shifts the precession of the nuclei into same phase angle, which creates resonance, and a net magnetization pointing 90° to the longitudinal magnetization. This magnetization is called the transverse magnetization. The coherent nuclei produce a radio signal, or free induction decay signal (FID signal), that can be detected by a radio antenna. After the RF pulse is removed, the nuclei will relax into baseline state. Firstly, the spins of the nuclei will repel each other, as they are positively charged, and thus shift phase. The net magnetization will return to zero. This relaxation is called T_2 or "spin-spin" relaxation, as the energy exchange between the nucleus spins is causing the relaxation. A second relaxation appears as the high energy nuclei returns to the low energy state. The energy that was previously absorbed by the nuclei is dissipated in to the surrounding lattice in the form of heat. During this relaxation the longitudinal magnetization is regrown. This relaxation is called T_1 or "spin-lattice" relaxation, as the spins transfer energy to the surrounding lattice. [1]

(Insert illustration)

The hydrogen nuclei are located in different local environments in the body. Some are for instance associated with free-floating water molecules, while others are associated with structural and storage molecules such as proteins and lipids, and thus more fixed in position. The nuclei have different T_1 and T_2 relaxation characteristics, depending on the local environment or tissue they are associated with. This can be accentuated and measured in NMR. [1]

The chosen pulse sequence is key to how the tissue will be portrayed in an image, and is described by the T_{echo}, time before the FID signal is measured, and T_{rep}, time before a new RF pulse is applied. In a case of nuclei associated with lipids and water molecules, the nuclei in

lipids are fixed and will have a fast T_1 relaxation after exposure to a RF pulse . Meanwhile the nuclei in the water molecules will maintain being in a synchronized phase. At T_{echo}, the nuclei associated with the lipids will have a low amplitude FID signal, as the transverse magnetization is weak, and the nuclei associated with the water molecules will have a high amplitude FID signal, as the transverse magnetization is strong. The water molecules will be assigned a white color on a greyscale image and the lipids as dark grey/black. In this case there is a long T_{echo} and a long T_{rep}, and is referred to as T_2-weighted MRI. [1]
In case of T_1-weighted MRI, the T_{echo} and T_{rep} are short. As in T_2-weighted MRI a RF pulse is applied and the nuclei associated with lipids will quickly return to baseline state and the water molecule nuclei will remain a strong transverse magnetization. At this time point a second RF pulse will be induced, referring to the short T_{rep}. Now the lipid nuclei will return to a strong transverse magnetization state and excite a high FID signal. More low energy state nuclei of the water molecules will absorb the RF pulse and shift to a high energy state, leaving a majority of nuclei in a high energy state. The water molecule nuclei now has a weak transverse magnetization and 180 degrees longitudinal magnetization, thus producing a low-amplitude FID signal. A short T_{echo} after the second RF pulse then shows lipids as white and the water molecules as dark grey/black in a greyscale image. [1]

1.2 MR image reconstruction

Different pulse sequences and certain physiological properties that can be exploited in certain pulse sequences have been laid out in the previous sections. This section aims to describe how the corresponding echo signals are reconstructed as a MR image.

Following the Lamour frequency equation (1.1), the main magnetic field causes all hydrogen nuclei to precess with the same frequency. Without any specification of spatial localization a MRI of a human body would consist of a single number. To prevent this, separate coils in the x, y and z directions are introduced. These coils can be adjusted in position, and thus produce gradient magnetic fields with a varying strength depending on position. According to the Lamour frequency the nuclei will precess with different frequencies when in a magnetic field with varying strength. The gradients can be turned on in combination to create any direction in space. These varying frequencies can be exploited to separate parts of the anatomy and ultimately illustrate a desired area. [1]
Insert Illustration As mentioned, the nuclei only exchange energy efficiently if the frequency of energy, or RF pulse, matches the precession rate. Thus, by altering the magnetic field along the body in one direction, z-direction for the sake of the example, the nuclei will have slightly different precession rates, and the RF pulse will only efficiently affect a desired slice of the nuclei. The nuclei of that slice now precess at the same rate. To get an image with a spatial resolution, the voxels that make out the image needs to be discriminated between. By turning on the gradient of the x-direction the lines in the y-direction are now encoded with a particular frequency. This gradient functions as a frequency encoding gradient. [1]
Turning the y-gradient on and quickly off, will de-phase the nuclei while still remaining the same frequency as before. This gradient functions as a phase encoding gradient. When comparing two locations approximately one voxel apart in the x-direction, then based on the amount

of gradient strength difference, there will be a certain amount of change in phase between the spins spread across that distance. The farther away from isocenter, where the magnetic field strength is B_0, the higher the change in phase will be. This notion is used to assign the correct spatial location of each voxel, when reconstructing the FID signals into an image. This phase encoding procedure is done in different gradient strengths in iterations to assign unique phases to the nuclei in the both directions. One iteration of a certain strength of the phase encoding gradient followed by a measurement is performed at a time. The only change per iteration is the phase encoding gradient strength. These iterations are then series of measurements acquired at different points in time, where each entry of the slice then represent a certain signal intensity. This time domain measurement is referred to at the raw data. [1]
The next step is to Fourier Transform (FT) the raw data, which will yield frequency information to the acquired signal intensities. This step gives a summation of the signal intensities at the different frequencies produced by the frequency encoding gradient. This is called the k-space, as the k-numbers of a signal describes its relative orientation and frequency. The k-space image contains the contrast in the center and the resolution in the periphery, as there is low or no phase encoding at the center and increasing towards the periphery, giving more brightness in the center and dimmer tones in the periphery. To allocate the voxels in correct spatial localization an inverse 2D discrete FT is performed on the k-space image. This provides the desired image of the anatomy slice. [1] Figure 1.2 shows the acquired signal in k-space and the reconstructed image after the Fourier Transform.

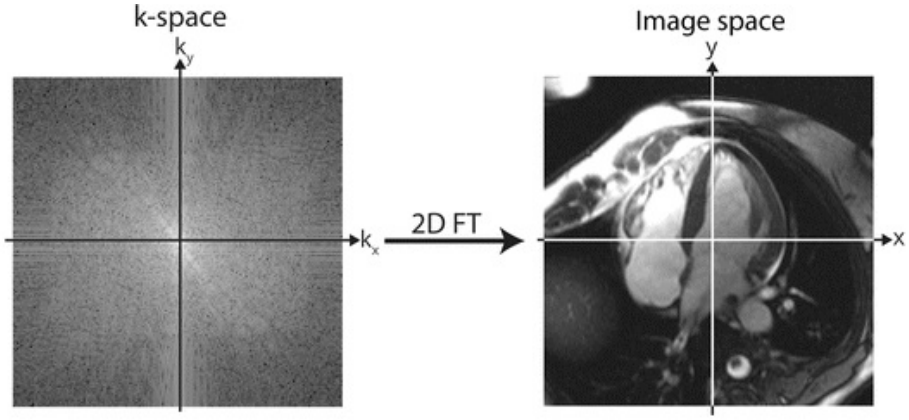


Figure 1.2: A depiction of the acquired signal i represented in k-space and how the reconstructed image looks after the 2D Fourier Transform [3].

1.3 Functional MRI

Different techniques exist to accentuate different tissues, physiological phenomena etc. using MRI. of which the Blood Oxygen Level Dependent (BOLD) functional Magnetic Resonance Imaging (fMRI) will be described in this section. The focus on the BOLD method is due to[1]

Functional magnetic resonance imaging measures the metabolic changes associated with dif-

[1] FiXme Note: find kilde om hvorfor BOLD er nizzle

ferent neurological tasks in different brain areas. fMRI offers advantages which predominantly are, high temporal and spatial resolution, low cost, and most importantly being non-invasive, which has made it a exceedingly popular method for imaging brain activity. The versatility of fMRI has made it a very important tool by being a biomarker for diseases and to study the efficacy of pharmaceuticals. The method offers high resolution of anatomical structures and localization and visualization of vessels. [4]

Multiple steps in forming and transmitting a neurological signal requires Adenosine triphosphate (ATP) consumption e.g. reception and reformation of an action potential. When activating a brain area in e.g. finger tapping, the ATP starts to be processed, leading to a decrease in oxygen concentration and increase in waste. Thereby the metabolic need for oxygen increases. As the movement is planned and executed, factors, which are present in the local tissue of the corresponding brain area, activate a vasodilation, increasing the blood flow to that area and reestablish the local homeostasis. During this regulation a special and not fully understood phenomenon occurs during this process. More oxygenated blood than needed to compensate for the offset is delivered and a overshoot occurs. The overall increase in neural activity in that specific area following the need for metabolic regulation thereby permits two conditions, which can be assessed through fMRI: being the cerebral blood flow and blood oxygen level dependent contrast. An example illustrating the measurable hemodynamic response can se found in figure 1.3. [4, 5]

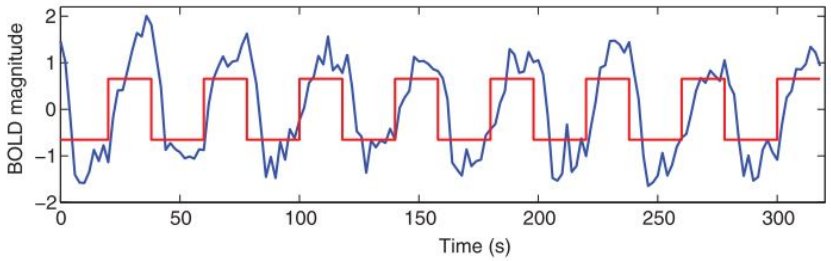


Figure 1.3: Figure showing an induced series of stimuli (red) and the hemodynamic response to the neural activity measured using BOLD (blue). It is shown that the measurable hemodynamic response is delayed compared to the stimuli. [5]

As established in the above section the BOLD signal is effected by the neural activity producing changes in the local blood flow, blood volume and blood oxygenation. The crucial part to why MRI can detect this natural contrast is that oxygenated hemoglobin (HbO_2) is diamagnetic, and deoxygenated hemoglobin (Hb) has four unpaired electrons thereby making it highly paramagnetic. Thus, the more oxygenated blood in an area, the larger the contrast compared to other brain regions would be visible in an acquired image. [4, 5, 6] Figure 1.4 illustrates how the contrast i dependent on the amount of oxygenated hemoglobin. [4]

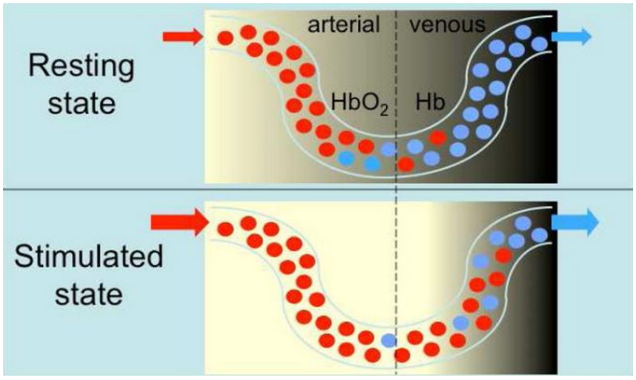


Figure 1.4: Illustration of how the difference in oxygen concentration in the hemoglobin change the magnetic properties, resulting in a higher measurable contrast [4].

This change in local magnetic properties increases the magnetic susceptibility leading to a greater MRI signal when acquiring an T_2^*- weighted sequence.

1.4 General MRI Pre-processing

Maybe do a top saying something about the program we choose to use. Derved lave en afgrænsning der

There are multiple steps to preprocess fMR images depending on the apparent application and outcome intended. However, there is a standard set of methods that is usually used across all applications. The following section seek to elucidate some of the most commonly used correction methods, including those considered for the standard preprocessing method used in this project. An example of some of the general processing steps for MR imaging can be seen in figure 1.5. [5]

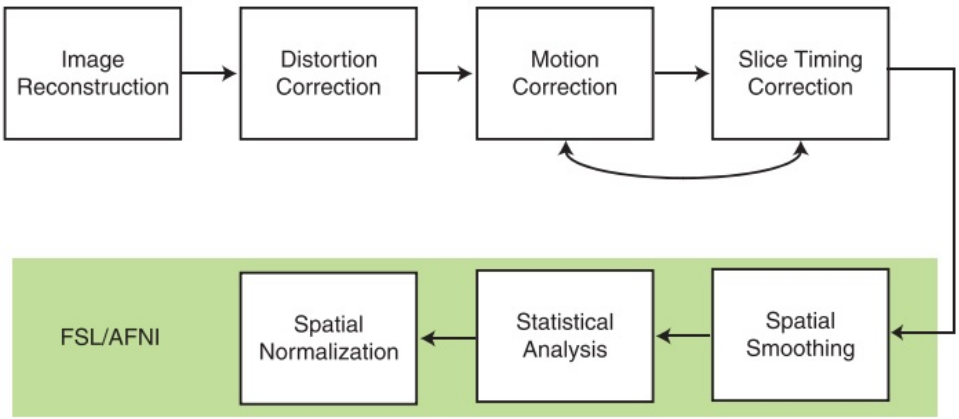


Figure 1.5: The general pipeline for MRI processing done in either FSL or AFNI, showing the different processing steps considered before final statistical analysis. Modified from [5].

1.4.1 Quality control

Conducting a continuous quality control is highly recommended after each performed correction step. Various scanner artifacts can occur while acquiring an MRI series. Before performing any common correction steps, one should consider to look for spike or ghosting artifacts. Spike artifacts are seen as a regular pattern of change in brightness across the entire image. This problem can occur due to instability inside the scanner deriving from e.g. electrical discharges. The artifact called ghosting occur mainly due to two reasons. One being an offset in phase between different lines in K-space and the other due to periodic motion as in heartbeat and respiration. Ghosting can be seen as light copies of the object appearing to either side of the object. Both types of artifacts can corrupt the information contained in the images. However, artifacts of this kind rarely present themselves in newer scanners, nevertheless it is still recommended to perform a quality control of the scan. [5]

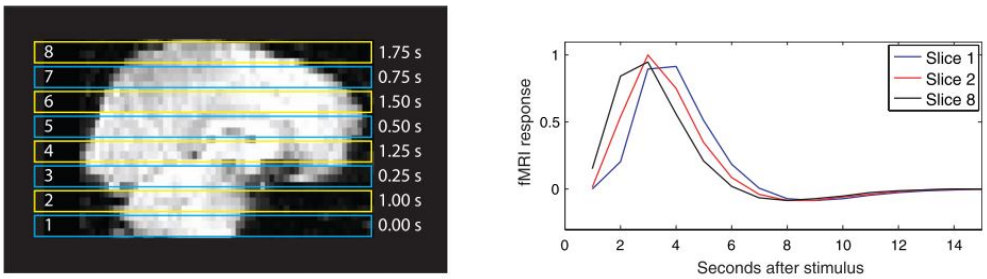
1.4.2 Distortion correction

Some fMRI acquisition methods suffers from artifacts at regions where air and tissue meet. The ear canal and sinuses are areas especially vulnerable. Inhomogeneity in the magnetic field in these areas can cause two types of artifacts: dropout and geometric distortion. A dropout will result in a reduced signal intensity in regions close to the air to tissue transition. When a dropout during an acquisition occurs, the lost signal cannot be restored and the damage is permanent. Therefore it is wise to consider the appropriate acquisition method taking the area of interest into addition. Air to tissue passages can also be subject to spatial distortion due to inhomogeneity created in the magnetic field. This will lead to structures not being located correctly in the captured image. This distortion makes is difficult to align two different scans, as done when aligning fMRI images with structural images. The spatial distortion can partially be corrected by employing field maps. In order to do a field map, the pulse sequence from the scan needs to be known. The process involves acquiring images at two different echo times. This results in images with two different phases, which can be used to compute the field inhomogeneity. Thereby it becomes possible to calculate the relative distance each voxel has shifted. This makes up a map for the distance shift for each voxel, and by inverting the map the original image can be restored. [5]

1.4.3 Slice timing correction

Acquiring fMRI scans is nearly always done in two-dimensions, where the slices are taken one by one. This can either be in an ascending, descending or interleaved order. Interleaved order [2] is sequentially skipping every either odd or even slice and then afterwards acquiring the skipped slices. Regardless of which order the slices are acquired, a difference in effect in each slice to the same hemodynamic response will be present due to the time difference in the slices. The method and result of interleaved MRI acquisition order can be seen figure 1.6. The difference in time between slices can range up to a couple of seconds depending on the acquisition protocol.

[2] FiXme Note: think it is done to avoid crosstalk



(a) (b)

Figure 1.6: Figure (a) illustrates an example of an MRI acquisition using interleaved order, where initially every odd slice is acquired followed by every even. The information about the hemodynamic response in each slice, and thereby also the difference at each time-point is shown in figure b. Figure modified from [5].

The difference in slice timing constitutes a problem when analyzing the data. The data is formed into statistical model, but since this model assumes that all slices are acquired at the same time point, the actual signal and the statistical model creates a mismatch. To counter this problem slice timing correction has been introduced. The common approach of this method is to choose a reference slice. Usually the slice acquired at T/2, where T is the total scan time, is used to interpolate the others. Linear interpolation can be used for simplicity, but most often sinc interpolation is used as it imposes less smoothing to the signal. [5]

1.4.4 Motion correction

Correcting for motion artifacts when doing MRI is inevitable, since even the best subjects will not be able to hold still. Even subtle movements as swallowing will be visible in the acquired image. [5]

Multiple internal and external factors can cause a subject to move. Internal factors are non-avoidable physiologic motion. The heartbeat causes a pulsating movement, which makes the brain move. Additionally, motion created during respiration can cause small changes in the magnetic field around the head. External factors like imposed stimulus might also cause the subject to make sudden movements. Often when doing fMRI the brain activation is measured while the subject is subjected to some kind stimulus. The stimulus would make the patient move, while some brain center might also show activation associated with stimulus. Therefore it is easy to mistake brain activation with stimulus correlated movement when analyzing the data, resulting in a weaker or even false statistical analysis. [5]

Motion during image acquisition can result in two primary artifact effects: bulk motion and spin history. Bulk motion refers to the movement of the head as a whole and requires standard correction methods, e.g. the images throughout the series to be realigned to a reference image. The effect of bulk motion can be visual in the entire image of the brain, but the effect will be most predominant at the edges of the brain. Here the artifact will be noticeable as either a drop or increase in intensity as a voxel would switch from containing brain tissue to suddenly not.

Spin history is head movement interfering with the MRI signal itself. The interference occur during acquisition when a voxel of excited protons is moved in to a neighboring slice. The scanner will thereby receive a different signal than expected which do not correctly represent the actual local properties. This results in an image where the intensities change in a striped pattern, visible when acquiring slices in interleaved order. The standard motion correction methods cannot cope with this type of artifact, but Independent Component Analysis (ICA) might be able to correct for this artifact. [5]

As mentioned earlier motion correction is to realign the series of images to a reference image trying to minimize cost in an introduced cost function. The reference image is usually the one taken midway into the series, justified by it being the closets to the average as well as the scanner at that time would have achieved stable contrast, as the magnetization would have reached steady state. The images are thereafter realigned utilizing an image registration method as it register the brain in each image. The general methods for motion correction treats the brain as a rigid objects, thus only performing rigid body transformations. Subsequently, the brain can either translate or rotate along the three axis, but the shape of the head cannot change. This method is therefore only applicable for bulk motion. [5]

1.4.5 Spatial smoothing

In some cases, introducing spatial smoothing in the preprocessing pipeline proves to be beneficial. Spatial smoothing allows the possibility of gaining a higher signal to noise ratio within the image, though with the consequence of a decrease in spatial resolution as the image gets blurred and smaller areas of activation gets smeared together. The operation can be justified by the closely neighboring voxel being correlated in effect to the hemodynamic response. Spatial smoothing removes the higher-frequency information. This might wash out some of the less significant features in the image, but this is favorable if the signal is increased for the more significant features. Especially when acquiring small voxels spatial smoothing is favorable, as it reduces the overall noise. Smoothing can also be applied to lessen the anatomical variability in images when doing studies with multiple subjects. [5]

Smoothing is done by applying a kernel, also called filter, to the image. The three dimensional image is convoluted with a three dimensional filter. The most commonly used is a Gaussian filter, where the extent of smoothing is controlled by the width of the distribution. The filter works such that for each voxel a new value is calculated based on a weighted average of the neighboring pixels, where the ones closest contribute the most and those further away contribute the least. The amount of smoothing needed to be implemented highly depends on the application and purpose. When smoothing fMRI signals for noise, the width of the filter distribution should not be bigger than the area of activation signals of interest. The effect of smoothing is shown on figure 1.7, where it is seen that as width increases smaller activation areas get removed and bigger areas of activation gets smeared together. Using smoothing to reduce the effect of anatomical variability, the optimal distribution width depends on the amount of variability in the subject population. [5]

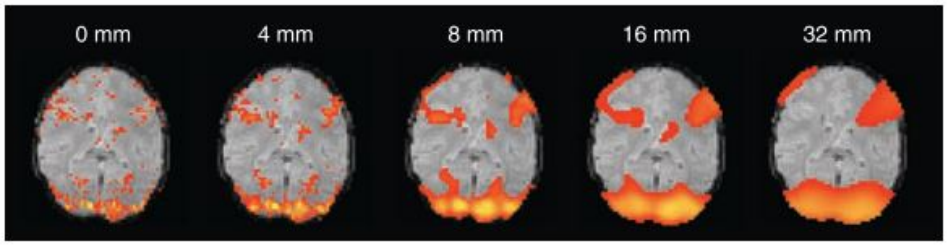


Figure 1.7: Illustration showing the impact of using different distribution width on the activated areas. An increase in width results in greater areas of activation smearing together and the removal of smaller. [5]

1.4.6 Temporal filtering

A characteristic noise which represents itself during fMRI data is the presence of a low-frequency drift. The drift is characterized as a slow increasing trend in the BOLD magnitude, when assessing the signal in the time domain. Doing a Fourier Transform, to analyze the signal in the power spectrum, would reveal low frequency contributers influencing the output. The reason for this type of noise contamination has been heavily investigated, and conclusions state that the noise originates from MRI scanner instability.[3] As this low-frequency drift will always be present, it very crucial to consider the frequency of which tasks or stimuli are performed to avoid output being present in the noise range of 0 to 0.015 Hz. Therefore stimuli or tasks should be performed within intervals of 70 s or less. [5]
A two step approach is used to remove the low-frequency contribute from the signal. Firstly a high-pass filter is used to remove the drift trend in the data. Introducing the high-pass filter impose a correlation of the data making the time series correlated over time. This violated some of the assumptions made in the General Linear Model (GLM)[4] of the data being not temporally autocorrelated and the variance to be constant over observations, as presented in section **??**. Not attending to this problem might cause an elevation in false positive rate. Thus, the second step is to estimate the autocorrelation and undo the correlation structure of the data. This is typically done by pre-whitening the data.
There are multiple ways to implement a high-pass filter, but due to the choice of utilizing the software from FSL, which use a locally weighted scatterplot smoothing (LOWESS), no other methods will be considered and presented. [5] The LOWESS approach will be explained in the methods section in section **??**.

1.5 Independent Component Analysis

ICA has proven to be a very successful tool in separating noise from wanted signal in BOLD-signals, making it a very desirable tool in fMRI [7]. The basic idea about Independent Component Analysis (ICA) is to recover m signal sources, which are mixed in n observed signals. The observed signal is given by:

$$\mathbf{x} = \mathbf{As} \tag{1.2}$$

[3] FiXme Note: heating of coils and such i believe
[4] FiXme Note: make section about data analysis and ref

where $\mathbf{x}$ is a observed signal vector containing the mixed signal elements $x_1, x_2, ..., x_n$, $\mathbf{s}$ is the source signal vector with the elements $s_1, s_2, ..., s_m$ and $\mathbf{A}$ is a mixing matrix with the dimension $n \times m$. Note that the dimensions of the mixing matrix can be equal to each other, meaning that it is required to have at least the number of observed signals as the source signals. The observed signal is assumed to be a linear mix of independent source signals. When performing ICA the goal is to find an inverted mixing matrix A^{-1} that recovers the source signal:

$$\mathbf{s} = \mathbf{A^{-1}x} \tag{1.3}$$

This can easily be achieved if the mixing matrix is known. However, this is rarely the case, as both the mixing matrix and source signals are unknown. There is no reliable way, to fully determine $\mathbf{s}$, thus a set of assumptions are required, which the ICA method is based on:

- The independent components (IC)'s are assumed to be statistically independent.
- The IC's must be non-Gaussian distributions.

Regarding the assumption of independence, a set of variables $y_1, y_2, ..., y_n$ is not allowed to share mutual information so that $i \neq j$. This can be expressed as the joint probability of the variables is equal the product of each marginal probability of the variables:

$$p(y_1, y_2, \ldots, y_n) = p(y_1) \cdot p(y_2) \cdot p(y_n) \tag{1.4}$$

Satisfying this condition assures independency of the variables. Note that uncorrelatedness does not equal independency. Concerning the second assumption on non-Gaussianity, the joint distribution of uncorrelated Gaussian distributions are not necessarily independent, and the distribution will be symmetrical and no information on the direction of the distribution can be derived. Thus, no transformation that allows independency can be performed, and the mixing matrix $\mathbf{A}$ can not be estimated from the mixtures. This can be shown in a two dimensional example with two sources s_1 and s_2 and the mixing matrix $\mathbf{A}$, the joint probability density function (pdf) of Gaussian distributions is calculated as:

$$p(s_1, s_2) = \frac{1}{2\pi} exp(-\frac{s_1^2 + s_2^2}{2}) \tag{1.5}$$

$$= \frac{1}{2\pi} exp(-\frac{\| \mathbf{s} \|^2}{2}) \tag{1.6}$$

For an orthogonal mixing matrix $\mathbf{A}$ the inverse can be written as $A^{-1} = A^T$, and then $s = \mathbf{A^T x}$. The joint pdf can then be rewritten as:

$$p(x_1, x_2) = \frac{1}{2\pi} exp(-\frac{\| \mathbf{A^T x} \|^2}{2}) \mid Det(\mathbf{A^T}) \mid \tag{1.7}$$

Because $\frac{1}{2\pi} exp(-\frac{\|\mathbf{s}\|^2}{2}) = \| \mathbf{x} \|^2$ and $\mid Det(\mathbf{A^T}) \mid = 1$, hence the orthogonality of $\mathbf{A}$, the joint pdf is:

$$p(x_1, x_2) = \frac{1}{2\pi} exp(-\frac{\| \mathbf{s} \|^2}{2}) \tag{1.8}$$

The joint pdf is not changed when choosing an orthogonal mixing matrix as well as the property of independence. No further information about the mixing matrix can thus be revealed when the source signals are from a Gaussian distribution.

1.5.1 ICA approaches

As the mixing matrix and source signal most often are unknown, the IC's must be approximated in an iterative process. There are two main ICA iteration approaches: minimizing mutual information and maximizing non-Gaussianity. The former approach seeks to minimize mutual information by maximizing independence between components. In practice this is done by minimizing the difference between the joint density distribution and the product of the marginal density functions, the left-hand side and right-hand side of equation (1.5). This can for instance be done through a Kullback-Leibler divergence between the $p(y_1, y_2)$ and $p(y_1) * p(y_2)$ in a two dimensional case. The second approach seeks to maximize non-Gaussianity of the components. The central limit theorem states that if sources are mixed, the mix tend to get more Gaussian than the individual sources. The strategy here is to find the directions in the data that is as far away from Gaussian as possible through a linear transformation. That direction will most likely be independent components. To find these directions different ICA approaches use fourth order moments and negentropy of the data.

1.6 Principal Component Analysis

Principal Component Analysis (PCA) is a well renowned and widely used analysis tool, capable of finding the most defining variables in a dataset. This facilitates finding the components that are the most saying for the dataset and makes it possible introduce dimensionality reduction, lowering the amount of redundant information. The PCA is used to transform a set possibly correlated variables into a set of uncorrelated components, called principle components. Each principal component (PC) is orthogonal on the former and are uncorrelated and have zero covariance. They each define the largest variance in an axis, such that PC 1 describes the direction of the maximum variance of the dataset. Each following PC describes the next highest variance of the dataset, with the constraint that is orthogonal and has zero covariance with any of the former PCs. PCA is the orthogonal projection of data onto a lower dimension linear space. A PC is found by minimizing the variance by projecting the feature values (blue dots) onto the line (now red dots) describing the highest variance in the data set (black line) as seen on figure 1.8. The PC is found by minimizing the mean square distance between the data points. [8]

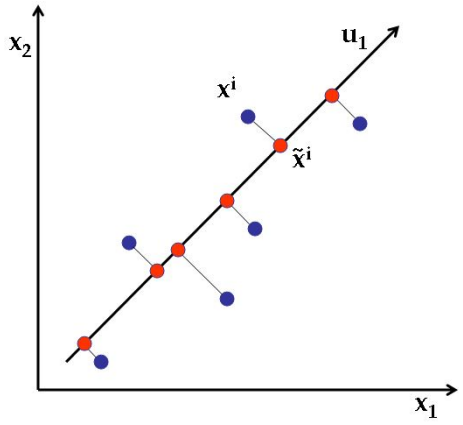


Figure 1.8: Two-dimensional example of projection of data variables (blue dots) onto PC axes (black). (u_1) indicates the direction of the eigenvector. [9]

The algebraic method of calculating the PCs can be done by using Singular Value Decomposition (SVD). The first step is to compute the squared cross product matrix of variances and covariances among every pair of the variables in the data set, where the diagonals are the variances and the off-diagonals are the covariances, as done in the following equation:

$$S = X'X \tag{1.9}$$

Where S is the cross product and X is the dataset matrix. When finding the PCs it includes an eigen-analysis of S. The eigenvalues of are solutions to the following equation:

$$|S - \lambda I| = 0 \tag{1.10}$$

Where λ is the variances of each PC and I is the identity matrix. After solving for λ the eigenvectors can be solved through the following equation:

$$det|S - \lambda I|b_i = 0 \tag{1.11}$$

Where b_i is used to calculate the eigenvectors as in:

$$u_i = \frac{b_i}{\sqrt{b_i' b_i}} \tag{1.12}$$

Where u_i is the i number of eigenvectors that contain a contribution to the principal components. The SVD orders the eigenvalues by size $\lambda_1 > \lambda_2 \ldots > \lambda_i$. The scores for each PC is equal to the corresponding eigenvalue for that exact axis. The eigenvalues describe how much of the variance is accounted for by the associated PC. Summation of all eigenvalues accounts for the total variance of the data set; this is called the trace. To find how much the each PC accounts for, the eigenvalue of that PC is divided by the total variance: $\% \; of \; total \; variance = \frac{\lambda_i}{Trace}$. This can be used for deciding how many components are significant and by how much the dataset can be reduced. [8]

1.6.1 PCA in images

The PCA can also be implemented on images, where the dimensionality reduction principle is mostly used for image compression. Here images can be reconstructed using only very few principle components without much information loss. Similarly PCA can be used to de-noise images, as noise would be presented in some of the less saying components, and by removing these, noise would be removed from the image.
The PCA can be run on the entire image, but methods introducing local PCA in smaller window of the image, has also been introduced for noise removal. [10]

2 | Methods

2.1 FSL FIX

Top saying something about why this methods needs to be tried. The automated de-noising system FSL FIX (FMRIB's ICA-based X-noisefier), made by the Oxford university center for Functional MRI of the brain (FMRIB), provides newly proposed data driven method for de-noising fMRI scans. The method applies the use of Independent Component Analysis (ICA) and numerous classifiers, which will be explained throughout this section.[7]

Bibliography

[1] A. A. Bharath. *Introductory Medical Imaging*. Vol. 3. 1. 2008, pp. 1–186.

[2] John C Edwards. “Principles of NMR”. In: ().

[3] Mushabbar A. Syed, Subha V. Ramen, and Orlando P. Simonetti. *Basic Principles of Cardiovascular MRI*. 2015, pp. 1–338.

[4] Gary H Glover. “Overview of functional magnetic resonance imaging”. In: *Neurosurg Clin N Am* 22.2 (2011), pp. 133–139.

[5] Thomas Poldrack A, Russell; Mumford A, Jeanette; Nichols E. *Handbook of functional MRI data analysis*. 2011.

[6] Nishanth Khanna et al. “Functional neuroimaging: fundamental principles and clinical applications”. In: *Neuroradiology Journal* 28.2 (2015), pp. 87–96.

[7] Gholamreza Salimi-Khorshidi et al. “Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers”. In: *NeuroImage* 90 (2014), pp. 449–468.

[8] John L Semmlow. *Biosignal and Biomedical Image Processing*. 2004.

[9] CIS 520. *PCA*. 2018.

[10] Y Murali Mohan Babu. “PCA based image denoising”. In: *Signal & Image Processing : An International Journal* 3.2 (2012), pp. 236–244.