UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Characterizing spatiotemporal properties of visual motion in the human middle temporal cortex at 7T

Francisco David Guerreiro Fernandes

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil em Radiações em Diagnóstico e Terapia

Dissertação orientada por:
Alexandre Andrade
Anna Gaglianese

2018

Resumo

O ser humano é diariamente exposto a situações onde é necessária a deteção de movimento. Esta capacidade de deteção da velocidade e direção de um objeto em movimento é considerada vital para a perceção, reação e adaptação a todos os eventos dinâmicos com que nos deparamos no meio ambiente próximo.

O *middle temporal cortex* (hMT+) é conhecido como a região central de processamento e deteção de eventos visuais em movimento, cuja velocidade pode ser descrita como resultado de propriedades inerentes: frequência espacial e temporal.

Estudos prévios apresentam resultados conflituosos no que diz respeito ao mecanismo de codificação e resposta a estímulos visuais em movimento. Em primatas foi observada uma predominância de neurónios do MT+ cuja resposta apresenta preferência para combinações especificas de frequências espaciais e temporais. Ou seja, a resposta para estímulos visuais no MT+ é dependente das propriedades da velocidade e não da velocidade em si. Contudo, outros estudos em humanos demonstram exatamente o contrário: a existência de uma resposta preferencial para certos valores de velocidade independentemente das frequências espaciotemporais usadas para gerar tal velocidade.

Este projeto tem como objetivo a caracterização da resposta neuronal a estímulos com diferentes frequências espaciotemporais no hMT+ por uso de uma técnica não invasiva, utilizando ressonância magnética funcional de modo a avaliar padrões de ativação BOLD (Blood-oxygen-level dependent).

Pretende-se inferir se a resposta das populações neuronais depende da velocidade do estímulo em questão ou das suas componentes (frequência temporal e espacial).

As aquisições de resposta BOLD foram feitas num scanner 7 Tesla a voluntários sem patologias durante uma tarefa especifica perante um estímulo visual. Este estímulo consistia num alvo subdividido em porções pretas e brancas bem distintas que se expande a partir de um ponto central durante 1 segundo. O estímulo foi repetido em seis scans, cada scan apresentando combinações diferentes de frequências espaciotemporais.

O estímulo foi criado via script de MATLAB e é apresentado ao voluntário através da sua reflexão num ecrã espelhado presente junto da porção posterior do voluntário.

O primeiro dos seis scans obtido durante a aquisição de scans de ressonância magnética funcional é designado de *localizer* devido à sua utilização na localização funcional do hMT+ de uma maneira independente e não regressiva. Assim, ao realizar o pré-processamento nestes primeiros scans *localizer* é possível obter *clusters* de ativação que representarão a área de estudo dos restantes cinco scans adquiridos. Nos restantes cinco scans de ressonância magnética funcional foram apresentados estímulos com três velocidades diferentes (3 graus/segundo, 9 graus/segundo e 15 graus/segundo), sendo que duas destas velocidades estão codificadas por duas combinações diferentes de frequências espaciais e temporais: 3 graus/segundo é codificado utilizando 1 grau/ciclo e 3 Hz e 3 graus/ciclo e 1 Hz; 15 graus/segundo é codificado utilizando tanto 3 graus/ciclo e 5 Hz como 5 graus/ciclo e 3 Hz. Desta forma, após o pré-processamento da data adquirida e a extração da resposta hemodinâmica e a analise do sinal BOLD podem-se tecer conclusões relativamente ao perfil seguido pelo hMT+.

Inicialmente, após extração da resposta hemodinâmica, a media da amplitude da mesma sobre todos os *clusters* de ativação identificados pelo *scan localizer* foi calculada e comparada entre os vários scans com diferentes características spaciotemporais. Contudo, não foi conclusivo quer a teoria de um perfil no hMT+ baseado na seleção de frequências espaciotemporais, nem um perfil baseado na seleção de velocidades do estímulo. Contudo, foi possível verificar a existência de certas variações entre as várias respostas hemodinâmicas de cada *scan* dentro das regiões de interesse criadas pelos *localizers*. Devido à natureza do projeto (o estudo é efetuado num *scanner* de 7T com a aplicação de uma *coil* de superfície capaz de trazer grande resolução espacial e rácio *signal-to-noise* elevado) foi-nos possível partir para uma análise da amplitude máxima da resposta hemodinâmica em cada voluntário dentro das

regiões de interesse (*clusters*). Ao comparar os máximos da resposta hemodinâmica para cada *run* foram identificadas diferenças significativas na resposta hemodinâmica entre *runs* cuja velocidade do seu estímulo visual em movimento é a mesma, simplesmente codificada com combinações diferentes de frequência espacial e temporal. Ou seja, entre os 10 hemisférios cerebrais presentes no estudo (cada scan pode ser dividido por hemisfério, perfazendo os 10 hemisférios mencionados) foi possível observar uma diferença significativa de atividade entre as mesmas velocidades codificadas com frequências diferentes, em pelo menos 7 (na velocidade 3 graus/segundo) e 5 (na velocidade 15 graus/segundo). Esta diferença de resposta em hemisférios expostos à mesma velocidade de estímulos visuais, ainda que insuficiente para uma conclusão acerca do perfil de codificação presente no hMT+, aponta para uma codificação em torno das componentes da velocidade (frequência espacial e temporal). Isto significaria que o córtex temporal médio apresenta preferência para processar estímulos com certas frequências espaciotemporais.

Com o intuito de investigar o modo como esta preferência por determinadas frequências espaciais e temporais estaria disposta no hMT+ uma análise de voxel por voxel foi de seguida efetuada. Cada *voxel* presente nos clusters de atividade identificados pelo scan *localizer* foi classificado conforme o estímulo visual apresentado que resultaria numa diferença maior de atividade (maior amplitude na resposta hemodinâmica).

Após representação dos *clusters* de atividade em superfícies inflacionadas representativas do córtex individual do voluntario foi possível tecer considerações acerca da presença de um mapa geral para a preferência de frequências espaciais e temporais especificas dentro do hMT+. Deste modo foi-nos possível identificar uma organização espacial dentro do hMT+ com possibilidade de separação do mesmo em subáreas previamente sugeridas noutros estudos mesentéricos do córtex visual: MT (área médio temporal) e MST (área superior temporal).

Estes resultados vão contra as conclusões retiradas de estudos prévios em torno do hMT+ com uso de ressonância magnética funcional: Chawla *et al* e Lingnau *et al*. Contudo, no caso de Chawla *et al*, as diferentes conclusões podem ser justificadas pela natureza dos estímulos utilizados durante o seu estudo (pontos em movimento aleatório). Este tipo de estímulos é conhecido por não permitir selecionar a frequência espacial nas várias velocidades demonstradas nos estímulos, eliminando a possibilidade de fazer conclusões quanto ao perfil de codificação baseado em frequência do hMT+. No caso de Lingnau *et al*. o paradigma para adaptação foi baseado em estímulos de pouco contraste o que leva a que outras populações neuronais tenham sido incitadas ao invés das propostas (hMT+).

Contudo, os resultados apresentados vão de encontro a estudos feitos por Gaglianese, A. com electrocorticografia em pacientes. O que torna este estudo uma extensão dos resultados prévios obtidos em electrocorticografia, mas utilizando agora um método não invasivo (fMRI) numa população saudável.

No entanto, mesmo após resultados promissores, o projeto encontra ainda uma panóplia de desafios que têm prioridade em ser corrigidos de modo a tornar os resultados mais robustos e melhores.

Em primeiro lugar a seleção das regiões de interesse onde a analise do estudo se baseou pode ser de futuro efetuada de maneira menos restrita e com a ajuda de atlas anatómicos. Isto poderia permitir a utilização dos voxeis que foram selecionados de modo a obter esquemas de distribuição de *clusters* de ativação sobre o córtex humano motor mais completos. Para além disso o tempo dispensado no pre-processamento poderia ser estandardizado pelo uso de funções automáticas no que diz respeito à remoção de artefactos fisiológicos.

No que diz respeito à direção do projeto para um futuro próximo: devido aos resultados demonstrados neste relatório de tese de mestrado, existem fortes possibilidades de efetuar estudos completos no sentido a obter clusters suficientemente definidos, que possibilitam a separação do córtex temporal médio em duas subdivisões que podem estar relacionadas com funcionamentos específicos e especificidade no que diz respeito a padrões de codificação de estímulos: MT e MST. Esta divisão poderá permitir tecer novas

conclusões no que diz respeito à importância do hMT+ no que diz respeito a casos em que o V1 (córtex visual primário) se encontra danificado, mas a codificação de movimento ainda é efetuada (*blindsight*).

Palavras-chave: córtex visual associativo, ressonância magnética funcional, resposta neuronal de populações de neurónios, codificação de movimento, codificação de movimento visual.

Abstract

Motion detection comes to humans as an important component in our daily life. Knowing the direction and speed of a moving object helps in understanding, reacting, and adapting to sudden dynamic events in our environment. Among other regions, the middle temporal cortex hMT+ in the human brain is the core region involved in the detection and processing of moving stimuli.

The speed of a moving object depends on the ratio of the change in position in between time samples, or, the spatial and temporal frequencies of a moving object. Therefore, motion can be encoded by speed *per se* or by separate and independent tuning of the specific different spatial and temporal frequencies components.

A recent study using ECoG in humans and complex visual stimuli using square wave gratings at different spatial and temporal frequencies has proven that specific recorded hMT+ neuronal populations exhibited separable selectivity for spatial and temporal frequencies rather than speed tuning. However, due to specific confined localization of the ECoG grid it remains elusive whether this selectivity comprises a spatial organization within the hMT+. Thanks to the advent of new neuroimaging techniques such as ultra-high field MRI at 7T it is now possible to visualize in unprecedent detail the human brain in vivo (0.8mm). Compared to commonly used field strengths, 7T allows for a gain in sensitivity and signal-to-noise, allowing to map for the first time non-invasively, the mesoscopic architecture of brain regions and measuring non-invasively neuronal responses via BOLD.

The aim of this project is to characterize the neural response to stimuli with different temporal and spatial frequencies in the hMT+ in a non-invasive method with the use of 7T fMRI via evaluation of patterns of blood-oxygen-level dependent imaging activation.

We investigated whether the response preferences of the neural populations present in the human middle temporal cortex depend on the stimulus speed or to the independent spatial and temporal frequency components.

The 7 Tesla blood-oxygen-level dependent functional magnetic resonance imaging responses were collected from healthy human volunteers on a Philips 7 Tesla scanner using advanced channels and techniques such as two 16-channel surface coils and gradient echo-planar imaging sequence. This was done during a specific task that consists on a visual stimulus (a high-contrast black-and-white dartboard) that's expanded from the fixation point for one second, presented in six scans (one used to localize the activity region – hence designated localizer – and the remaining five to study said activity), while having a baseline of a homogeneous grey background. Each run consisted on different spatial and temporal frequencies of the dartboard that have been previously used in the mentioned ECoG study.

After computation of the BOLD signal using a deconvolution approach, the results showed that the human middle temporal cortex separates motion into its spatial and temporal components rather than decoding speed directly. Moreover, clusters of activity for specific combinations of spatial and temporal frequencies suggest a spatial organization within the human middle temporal cortex.

Keywords: middle temporal cortex, functional magnetic resonance, neuronal population responses, speed encoding, visual motion.

Acknowledgments

Firstly, I'd like to thank the UMC Utrecht and Nick Ramsey's Lab for the opportunity to go abroad, to grow and mature to the side of a fantastic, talented, intelligent and resilient group of people dedicated to make our planet a better place.

I'm very grateful for the daily guidance, the sympathy and patience that Anna Gaglianese demonstrated throughout this long journey. You are the sweetest supervisor I could ever had hoped for. I know now that Leo is a very lucky young boy! Also, I don't know if I would have been able to handle myself in some of the situations as calmly as you did.

Natalia Petridou and Nick Ramsey, for your belief in my skills (even when I can't see them) and for the secure and confident way of living life you earned my utmost respect and the status of role models.

A big thank you for my internal supervisor, Alexandre Andrade, for the attention, help and time spent along the project and throughout my education in *Faculdade de Ciências da Universidade de Lisboa*.

To my old-time friends, you made me grow up into the person I am today. To everyone in FCUL that shared these years with me here, you were the reason why it was so enjoyable.

To my family, no words can explain the importance you have in my life.

To everyone that I met in the Netherlands…from the ones that said I would go back, the ones that I shared houses and stories with, the ones that made me laugh, the ones that I always spend coffee break with, the ones that travelled half a world with me in an amazing trip, the ones that carried me in their bike. To you, who make me feel like a hero. Kudus to you all. You are bigger and better than I could ever be, and you reminded me something that I had forgotten: to never give up.

And now another battle begins, I just hope I make everyone proud.

Table of contents

List of abbreviations

Consistency of notation throughout the documentation was maintained even in cases where this inevitably led to deviations from the notation of the classic and standard literature or some of the publications cited.

Abbreviation/ Symbol	Meaning
MRI	Magnetic resonance imaging
fMRI	Functional MRI
BOLD	Bold oxygenation level dependent
MEG	Magnetoencephalography
ECoG	Electrocortigography
EEG	Electroencephalography
PET	Positron emission tomography
CBF	Cerebral Blood Flow
CBV	Cerebral Blood Volume
Hb	Hemoglobin
HbO_2	Deoxyhemoglobin
γ	Gyromagnetic ratio; ratio of the magnetic moment to the angular momentum of a particle. This is a constant for a given nucleus.
T_1	Spin-lattice or longitudinal relaxation time; the characteristic time constant for spins to tend to align themselves with the external magnetic field.
T_2	Spin-spin or transverse relaxation time; the characteristic time constant for loss of phase coherence among spins oriented at an angle to the static magnetic field, due to interactions between the spins, with resulting loss of transverse magnetization and MR signal.
T_2*	Transverse relaxation time star; the observed time constant of the FID due to the loss of phase coherence among spins oriented at an angle to the static magnetic field, commonly due to a combination of magnetic field inhomogeneities and spin-spin transverse relaxation with resultant more rapid loss in transverse magnetization and magnetic resonance signal.
TI	Inversion time; time between middle of inverting (180º) RF pulse and middle of the subsequent exciting (90º) pulse to detect amount of longitudinal magnetization.
TR	Repetition time; interval of time between the beginning of a pulse sequence and the beginning of the succeeding (identical) pulse sequence.
FID	Free Induction Decay; the decaying signal originated by the decay towards zero with a time constant T_2 (or T_2*) of the MR signal produced by a transverse magnetization of the spins.
SENSE	SENSitivity Encoding; parallel imaging technique to accelerate the image acquisition.
SNR	Signal to noise ratio.
PI	Parallel Imaging.
R	Acceleration factor; denotes the factor by which the number of samples is reduced to with respect to the full Fourier encoding.
PPA	Partially Parallel Acquisition.

α	Flip Angle; amount of rotation of the macroscopic magnetization vector produced by an RF pulse, with respect to the direction of the static magnetic field (*see* B_0).
FT	Fourier transform; a mathematical procedure to separate out the frequency components of a signal from its amplitudes as a function of time, or vice versa. It is used to generate spectrum from the FID or spin echo in pulse MR techniques which are core to most MR imaging techniques. It can be generalized to multiple dimensions to perform, in MR imaging, connection between an image to its k-space representation.
SE	Spin Echo; the RF pulse sequence where a 90º excitation pulse is followed by a 180º refocusing pulse to eliminate field inhomogeneity and chemical shift effects at the echo. Alternative nomenclature RF spin echo might be more appropriate.
FOV	Field of view; the rectangular region that is superimposed over the human body over which MRI data are acquired. The dimensions are specified in length in each in-plane direction and controlled by the application of frequency and phase encoding gradients
ROI	Region of interest.
hMT+	Human middle temporal cortex
MST	Medial superior temporal area
LNG	Lateral geniculate nucleus
V1	Primary visual cortex (anatomically defined as the striate cortex)
V2	Secondary visual cortex (also called prestriate cortex)
V3	Third visual complex (named visual area V3 in humans)
V4	Visual area V4
V5/MT	Middle temporal visual cortex
V6	Dorsomedial area

List of figures

List of tables

1- Introduction

Even though the perception of speed of a moving object and all the tasks associated with it were identified in several areas of the brain as stated in numerous studies in the past [1]–[6], the general agreement is that the core region responsible for visual motion processing in primates, including humans, is the middle temporal cortex (referred as hMT+ in humans and V5/MT in macaque monkeys and owl monkeys, respectively). Such is demonstrated several times in the several anatomical and functional studies regarding the hMT+ [5], [7], [16], [17], [8]–[15]. While extensive studies have been carried out in the MT due to its retinotopic properties and selectivity for direction of moving stimuli, only recently there was an increase of interest in the mechanism of speed of motion encoding in MT neurons.

This mechanism helps individuals in their spatial navigation, by providing the ability to promptly react to the sudden appearance of potential obstacles or dangers, which in turn leads to the conclusion that any added knowledge behind this complex mechanism holds great potential to help further develop well appropriate solutions for patients who've had any damage on the localized brain region of the middle temporal cortex. Findings report the co-existence of an alternative direct route linking the lateral geniculate nucleus (LGN) to the hMT+, bypassing the V1, which serves as an example of the fundamental role played by the hMT+ , even in the presence of individuals with severe destruction of primary visual cortex (blindsight) [18].

The following ***Subchapter 1.3*** of this chapter will subtly dip into the details on these studies and the several advancements that brought attention to this area of the brain as the responsible for encoding the speed of motion in human beings.

Out of the plethora of studies regarding the MT, it is possible to outline two distinct speed encoding mechanisms: the first suggests that the MT neurons have a separate and independent tuning for the spatial and/or temporal frequencies of visual stimuli; the second proposes direct speed tuning, rather than its separation in the two frequencies components.

The two methods (graphically represented in ***Figure 1.1***) provide different predictions about the tuning profiles for spatial and temporal frequencies in the MT neurons. While the direct speed tuning will predict the same preferred speed, given by all the possible combinations of spatial and temporal frequencies; the separated and independent tuning predicts a preferred representation of the speed components (being it either spatial or temporal) instead of the speed *per se*.

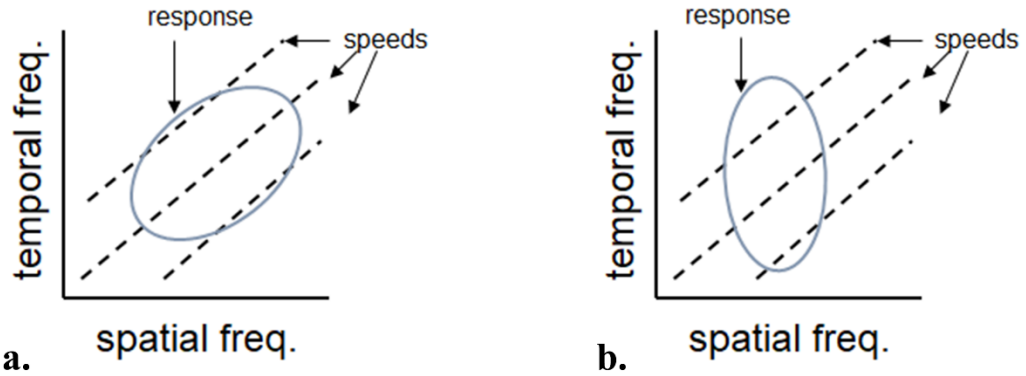

Figure 1.1: Representation of the two paradigms that define speed encoding in the hMT+. a. Several combinations of spatial and temporal frequency, resulting in the same speed, show similar response, indicating a speed focused tuning profile. b. Response is dependent on spatial or temporal frequency, meaning that different combinations of the two parameters result in different responses.

At best, the results that came from countless animal physiology studies can be considered conflicting. While some report that macaque have a more than half of the middle temporal visual area cells tuned for speed [19], with these results being comparable with other reports [20]; others report, for example, in single neuron recordings done in nonhuman primates and also marmosets, to a majority of neurons with distinct responses for spatial and temporal frequencies and only a small percentage displaying a speed tuning profile [21]–[23].

Regarding the human homologue of MT/V5, the studies about responses to speed of motion in the hMT+ come with the same degree of uncertainty present in the animal physiology studies. It is still unsure if the neuronal tuning in human MT+ reflects speed tuning, independent spatial and temporal frequency tuning, or even perhaps a combination of both.

Traditionally, the experiments performed towards a better understanding of speed encoding in the human MT+ were done using functional magnetic resonance imaging (fMRI), magnetoencephalographic (MEG) neuronal responses and intracranial electrocorticography (ECoG). In these, the most widely used motion stimuli were broadband stimuli such as random dots of even spots of light [24]–[26] and its broadband nature doesn't allow for a clear distinction in between speed tuning of independent spatiotemporal frequency tuning. Also, by observing studies that used gratings moving at different speeds to infer about the encoding of speed motion, only an fMRI adaptation experiment using drifting gratings [27] provided evidence of speed tuning of Blood-oxygen-level dependent imaging (BOLD) responses over different temporal frequencies conversely to the results gathered by MEG [28], which showed independent spatial and temporal frequency tuning in response to said gratings in hMT+.

This thesis comes as a follow up of the work done by Gaglianese, A *et al* [29], where ECoG was used to investigate speed tuning properties in hMT+ neuronal populations in response to visual motion stimuli. Its findings pointed to neuronal population response to the spatiotemporal frequency properties and not the speed *per se* of the stimuli, consolidating and extending the findings in macaque and marmoset MT neurons to the now human neuronal populations present in the MT. Here fMRI was found to be the missing link and a valuable addition to corroborate these past results. By combining this non-invasive technique and the extraordinary resolution power of 7T, while extending the findings to a healthy subject dataset, the knowledge of how the hMT+ encodes speed motion would come naturally to light. Also, the proven matching between BOLD and neuronal activity at a scale of mm [30] will also provide proof on the ECoG findings.

The objective is to understand whether the response preferences of the neural populations present in the hMT+ depend on the stimulus speed or the independent spatial and temporal frequency components. More so, the high spatial resolution of the 7T will allow for the division and clustering of neuronal populations inside the MT that exhibit separate and independent spatial and temporal frequency or conversely the stimulus speed directly, with selectivity to multiple combinations of its frequency components.

These findings may point to a division of the human MT+ complex into sub regions [24] that may be identified as homologs to a pair of macaque motion-responsive visual areas, middle temporal visual area (MT) and the medial superior temporal area (MST), with their neuronal populations encoding speed in the two different established methods, explaining the unclearness behind speed encoding in the hMT+.

1.1. Project Specifications and Goals

The vast infrastructures and clinical possibilities present in the Utrecht University Medical Centre (UMCU) research environment allowed for the setup of the experiments that this thesis revolves around. Having vast access to patient data who underwent implantation of ECoG grids for the purpose of epilepsy monitoring gave birth to a novel study that used ECoG, a technique with recent increased interest due to its direct and high temporal and spatial resolution measure of neuronal activity within relatively small neuronal populations in the human brain, as means to investigate the speed tuning properties in response to square-wave dartboard moving with several temporal and spatial frequency combinations. [29] Also the sheer power of the high-field 7T fMRI and the high spatial resolution achieved by the development of high spatial resolution surface coils [31] allowed a signal-to-noise ratio level that allowed the study to be performed. The combination of these last two events make the UMCU one of the few neuroimaging centers in the world that allows the mapping of the visual cortex in a non-invasively way under the millimetre scale. By gathering all conditions, the experiment serves as a bridge in between the several neuron recordings done in non-human primates and non-invasive methods such as MEG, Electroencephalography (EEG) and even fMRI studies in humans.

Part of the fMRI studies on humans were traditionally done with main usage of broadband stimuli [25], [26], [32], like random spots of light and/or random dots, as motion stimuli to investigate speed responses. The nature of these broadband stimuli doesn't allow for distinction between speed tuning or independent spatiotemporal frequency tuning. Although, by using and focusing on adaptation of drifting gratings in an fMRI experiment performed in humans, evidence was provided in the direction of speed tuning of BOLD responses over several temporal frequencies [27].

The final link stands in the methodological use of stimuli that provoke robust responses, like random dot motion patterns and square-wave checker boards, with independent manipulation of temporal and spatial frequencies for a proper analysis of the hypothesis of these components being part of the speed encoding mechanism. Thus, a conclusion was drawn: square-wave checkerboards would be the appropriate stimuli for this type of study. Conversely to the what is expected, checkerboard stimuli hard edges with contrast energy at multiple spatial and temporal frequency harmonics are still regarded as exhibiting a single speed, that being the one of the change in edge position per second. [29]

In combination with the usage of the Philips 7T scanner with the help of two 16-channel high-density multi-element surface coil with minimal electronics developed and present at the UMC and the 16-channel [31], the high contrast potential and accuracy for most of the human tissues brought by this magnitude of B0 strength with the close proximity to the head allowed by the high-density multi-element surface coil allows us to draw some conclusions, contributing to in the definition of a pattern for the speed encoding in the hMT+. Since fMRI is a non-invasive technique and there is a strong positive correlation link between the BOLD amplitude and the ECoG broadband power in some areas of the brain [30] the present study will corroborate or refute some of the conclusions drawn previously in [29], while also expanding them to a healthy population.

1.2.Background

1.2.1. Magnetic Resonance Imaging and functional Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a medical imaging technique rapidly with a galloping increasing interest due to its excellent contrast potential and accuracy provided for most of the tissues present in the human. The technique has already existed for more than thirty years, with a lot of its improvements and developments being achieved as higher resolution, better contrast techniques as well as improved safety mechanisms and potential long-term effects analysis. [33]–[35]

The technique allows for visualization of both anatomical and functional data of the human brain and, broadly speaking, it was birthed in the behavior of the magnetization vector of a specific set of protons precessing after being excited by a controlled radiofrequency pulses. [36]

Functional magnetic resonance imaging (fMRI) has grown to be the most widely used tool to study neuronal activity in the human brain in vivo with its oxygenation level-dependent signal ever since it was discovered twenty years ago with a setup of magnetic resonance imaging that allowed for the reception of level of oxygenation signals [37].

The spatial and temporal resolution provided by state-of-the-art MR technology with its non-invasive character, allowing for multiple studies to be performed in the same subject, grant huge advantages to fMRI over the other functional neuroimaging techniques, like positron emission tomography (PET), which are also based on changes in blood flow and cortical metabolism.

Basic principles and methodology of fMRI are explained during this chapter, as well an extensive description of the BOLD contrast mechanism physiology, after a brief description of the physical fundamentals of MRI at a conceptual level.

1.2.2. Physical Principles of MRI

Physical principles of MRI are shared for anatomical and functional imaging. We can split the operation of MRI in two major themes: excitation and recording of electromagnetic signal (reflecting the properties of the measured object), and construction of two- and three-dimensional images (reflecting on how its measured proprieties vary throughout space).

1.2.2.1. Spin Excitation and Signal Reception

The core of spin excitation resides on the magnetic excitation of body tissue and the reception of returned electromagnetic signals from said tissue. All nuclei with an odd number of protons are considered magnetically excitable and while it can happen within all the magnetically excitable nuclei, MRI focuses on the most common isotope of hydrogen with a nucleus of only one proton. Hydrogen protons (^{1}H) are ideally suited for MRI due to their abundance in human tissue and favorable magnetic properties.

It is the spin that provides the magnetic properties to protons, and due to that, the rotation like a spin top around their own axes, a small directed magnetic field is inducted (***Figure 1.2a)***.

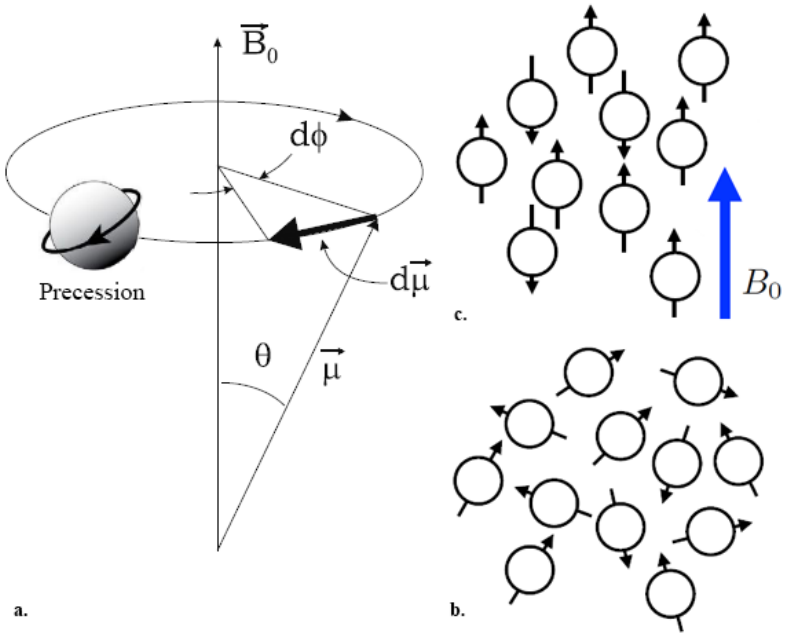

Figure 1.2 a,c. Schematic representation of spinning Hydrogen protons. **a.** Representation of the circular motion of the axis of rotation of a spinning body, in this case hydrogen protons, around another fixed axis, ($\mathbf{B_0}$), produced by a torque applied in the direction of the motion; defined as precession. **b.** Whenever no external magnetic field is present, the direction of the spin is distributed in a random manner. **c.** Upon the presence of an external magnetic field, the affected spins will align either parallelly or antiparallelly. Ultimately this results in a net magnetic field parallel to the external magnetic field due to the predominance of parallelly aligned spins. *Adapted from* [38].

In a normal environment, the collection of all magnetic fields created by the spins in the human body would be oriented randomly, thus canceling each other out (***Figure 1.2b***). However, when placed under the effect of a strong static magnetic field ($\mathbf{B_0}$) of an MRI tomograph, the spins within the subject will orient themselves in line with the present field.

This alignment can either be with the field (parallel to it) or against the field (antiparallel to it) (***Figure 1.2c***). It is to note that the said "alignment" is but an abuse of language since there is no alignment of rotation axis but of movement of rotation around the direction of the scanner static magnetic field: this motion is called precession. The two states of precession differ in the energy their protons carry, with the parallel precession having less energy and thus being more stable. Due to this difference of energy, a higher percentage of parallel protons are found in the volume, causing the body to be magnetized.

1.2.2.2. From Larmor Frequency and Equation to Resonance

The excess number of spins will produce a total excess magnetic field called $\mathbf{M_0}$. The precession frequency of the protons that produce this field depends on the strength of the surrounding magnetic field. More precisely, the precession frequency ω_0 is directly proportional to the strength of the external magnetic field and is defined by the Larmor Equation:

$$\omega_0 = \gamma B_0 \tag{1.1}$$

The symbol of ω_0 is known as the precessional, Larmor or resonance frequency. The symbol γ refers to the gyromagnetic ratio, which is a constant unique to every atom (for hydrogen protons this is 42.56 mHz per Tesla [33]).

When an electromagnetic pulse with the same frequency as the proton's precession frequency is applied to the protons present in the volume, these will absorb the transmitted energy, getting excited in the process. This principle is the basis of the imaging technique and it gives the name to it: Resonance. Since the precession frequency is in the range of the radio frequency (RF) waves, the electromagnetic pulse can be called radio frequency pulse.

As an effect of being excited, all the spins that were in the parallel state (associated with a lower energy level) will then absorb the energy and to go the anti-parallel state (associated with a higher energy level) and will start precessing in phase. This phased precession will cause the magnetization vector M_0 to start moving down towards the x-y plane. This plane is perpendicular to the static magnetic field and, thereby goes by the name of transverse plane. The angle between the magnetization vector and the transverse plane is referred as α and is one of the means of characterization of RF pulse strength and duration.

When α is 90º, the magnetization vector is completely within the transverse plane, meaning that, in terms of proton alignment, the number of protons aligned parallelly and antiparallelly is equal (***Figure 1.3***). The added spins form a net magnetic field called M_{XY} which exists in the transverse plane and can be measured in the receiver coil since it induces a detectable current flow.

Due to the nature of the RF pulse, right after its end, the phasing existent within the precessing protons will also cease to exist, due to the spin-spin interactions of the magnetic fields, which ultimately cause the decay of the transverse magnetization in a few tenths of a millisecond.

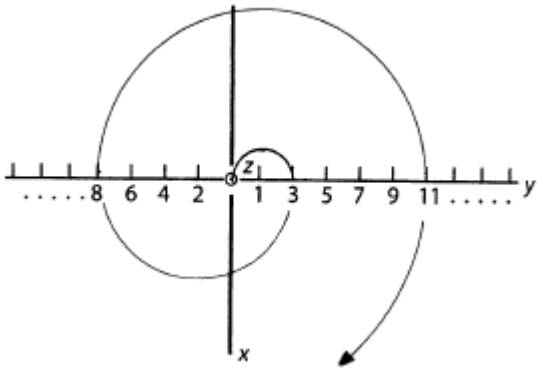

Figure 1.3 Scheme of the magnetization vector viewed from a direction parallel to the static magnetic field after the application of a RF field with frequency of ω_0. *Courtesy of* [39].

The transverse dephasing is a result of the interaction between spins. Such interactions lead to slightly different local magnetic field strengths. In turn, these local magnetic field differences lead to different precession frequencies and thus phase shifts between precessing spins. This process is called dephasing and since it happens in the transverse plane it can also be called transversal relaxation and it evolves over time following an exponential function with time constant of T_2. This relaxation isn't represented by the T_2, but by its successor, T_2*, which accounts for the faster dephasing of the spins due to the differences in the inhomogeneities of the static magnetic field and in the physiological tissue as well. Therefore, the raw signal measured in the receiver coil, also called free induction decay (FID), will decay with the shorter time constant T_2*.

$$M_{XY} = M_0 e^{t/T_{2*}} \tag{1.2}$$

These local field inhomogeneities causing the different precession frequencies that in the end increase the speed of the dephasing allows fMRI to exist as a technique since those field inhomogeneities

have a dependency in the local physiological state, i.e., state of local blood oxygenation and level of local neuronal activity, and the measurements of these field changes will allow the inference of neuronal activity. More details on this and their importance in the fMRI technique will be discussed in the following subchapter 1.1.2.

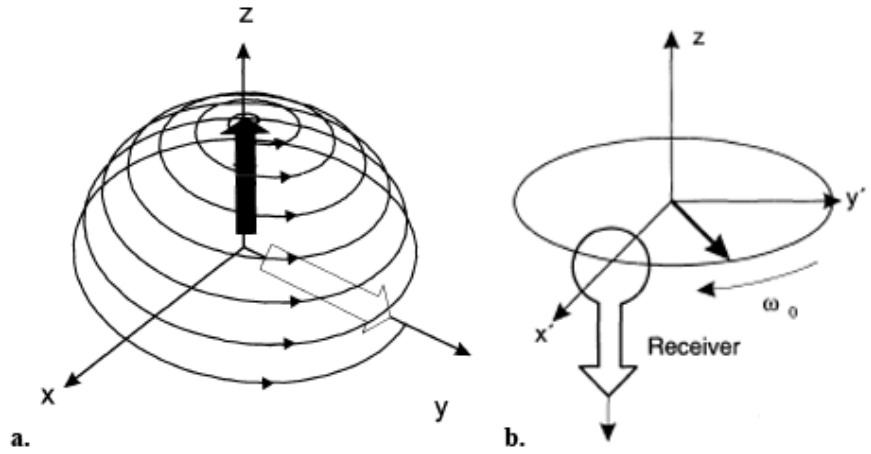

Figure 1.4 a. Excitation of the magnetization when observed in the laboratory frame. **b.** The presence of a current loop (sensitive for transverse magnetization) close to the changing magnetic flux of the rotating magnetization causes an induced current: the MR signal. *Adapted from* [39].

Two effects affect the spin's speed of dephasing: random effects as well as fixed effects caused by magnetic field inhomogeneities. The field inhomogeneities' dephasing effect can be reversed with the application of a 180º RF pulse. By allowing $t = \tau$ to elapse before the application of a reversing RF pulse, the dephased spins will flip on their X' and Y' relative rotation axis, allowing them to reverse their order while maintaining their direction of rotation. Due to the application of the 180º RF pulses, at a time of $t = TE = 2\tau$ the vectors will be again back in phase producing a large signal that goes by the name of spin echo.

This amplitude of the obtained spin echo is smaller than the amplitude of the initial FID. This is due to the spin-spin interactions that inevitable produces irrecuperable loss of signal (during T_2 decay). This goes in a circle by applying consequent 180º RF pulses, since all the spins can start going out of phase the moment that they are back in phase at the echo time. The process can be repeated for as long as enough signal is available, with the interesting part of the process being the possible selection of the T_2 signal amplitude by the correct setting of the timing on the 180º pulses (***Figure 1.5***).

Spins also reorient themselves with the direction of the strong magnetic field that is provided by the scanner while going back to the low energy state after being excited. Longitudinal relaxation is the designation of this spin reorientation with the external magnetic field and it happens to be a slower process than the previous discussed dephasing. The longitudinal component M_z will suffer a recovery dictated by an exponential function with a time constant T_1:

$$M_Z = M_0(1 - e^{t/T_1}) \tag{1.3}$$

It is important to notice that the absorbed energy provided by the RF pulse is released both in a way to be detected outside of the body as RF waves; and some part is released to the nearby surrounding tissue, being designated by the term lattice. Just like interactions between spin-spin determined T_2 and, consequently, T_2*, the interactions spin-lattice will determine the speed of T_1 recovery. What this represents is that T_1 is tissue-specific, and with this value and T_2, MRI allows for a differentiation between tissues when proper MRI sequences are designed to do so.

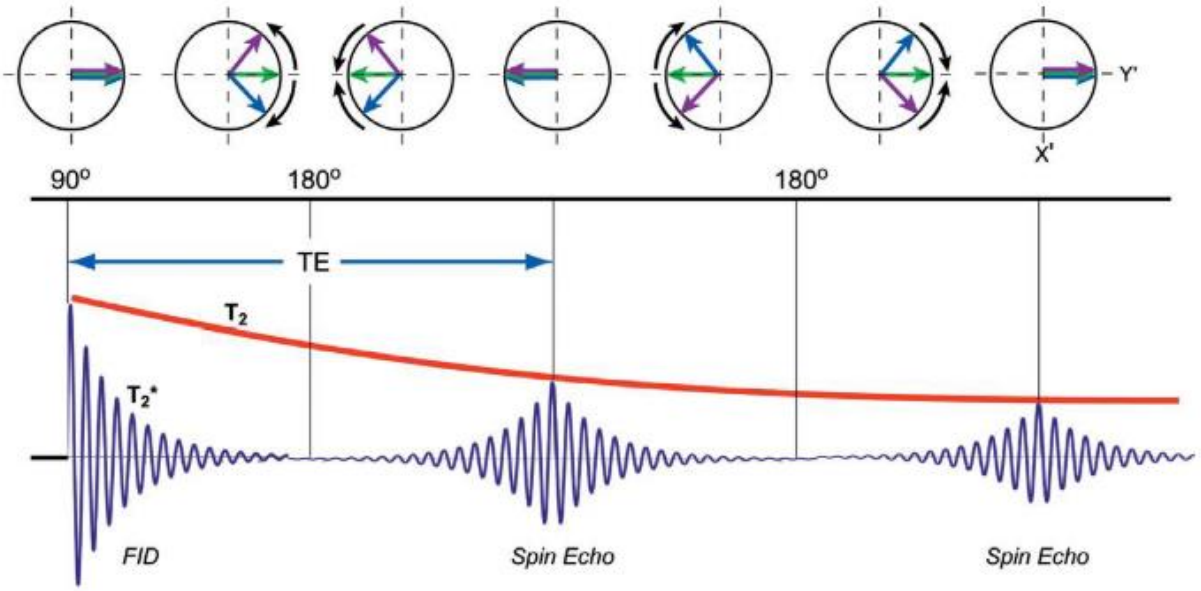

Figure 1.5 Schematic illustration of the process of the effect of the magnetic field inhomogeneities and how they reduce the Free Induction Decay (FID) and how the application of 180º RF allows for a reduction off this decay effect. The three representations of coloured spin vectors precessing at resonance frequency, slightly faster and slightly slower than that (green, violet and blue vectors, respectively). The 180º RF pulse flips the dephased vectors about the X' axis, reversing the order of the spins but not their direction of rotation. This flip culminates in the vectors falling back in phase which induces a large signal called spin echo. The time from the application of the 180 RF pulse and the spin echo is called echo time (TE). *Adapted from* [40].

1.2.2.3. Image reconstruction: Slice selection, Frequency and Phase encoding

The principles of magnetic resonance may provide some insights on the physical effects behind the technique but by themselves they solely are not representative on how it is possible to obtain images of human tissue. To put it into perspective: up to this moment the insight on the magnetic resonance Imaging provided was focused only on the magnetic resonance portion; from now on we discuss the Imaging.

To convert the signals obtained from the excitation of protons into a cohesive image there's a very important step to be taken: assigning components of the signal to its originated positions in space. Not all sequences have the same processes behind them, but the localization of signal source has the combined action of three fundamental techniques: selective excitation of a specific slice; phase and frequency encoding. Three fundamental techniques for three fundamental dimensions, allowing for a complete spatial localization of the received signal.

By analyzing the Larmor equation (1.1), the knowledge that spins exposed to a higher magnetic field will precess faster comes clear to us. This, coupled with the usage of magnetic field gradients created by the gradient coils of the MRI scanner, will add this new magnetic field to the static one existing beforehand. The field strength will then vary linearly with the distance from the magnet's center. In conclusion, spins exposed to the side of higher magnetic field precess faster, while spins exposed to the side of the lower magnetic field will precess slower.

With RF pulses with frequencies within specific bands, a slice of the imaged object will be excited selectively and any spins that fall outside of that range that are precesing at a different speed won't absorb the RF transmitted energy (***Figure 1.6a***).

Slice orientation will always be performed perpendicularly to the axis of the gradient (z-axis), or, by other words, an axial slice will be the result of this slice selection. With the same logic we get to

the conclusion that a gradient along the x-axis produces a sagittal slice selection and along the y-axis a coronal slice selection. Oblique slices fall out of these simple rules, but the general approach is their selection by application of simultaneous gradients combination.

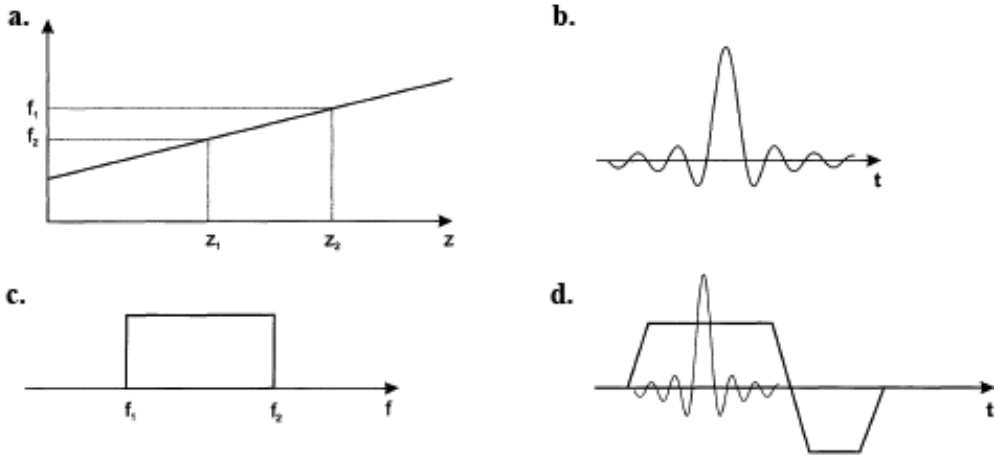

Figure 1.6 a. Selection of a slice with a gradient field defining the frequency band f1-f2 of said selection. **b.** Frequency spectrum of the RF pulse. **c.** Envelope of amplitude of the excitation on time. **d.** Excitation pulse with following refocussing gradient.

The slope of the applied gradient and the frequency band of the applied RF will influence the position and thickness of the selected slice.

After this slice excitation, the echo that is measured is restricted to a compound signal from the protons within said slice and for the rest of the spatial encoding the slice selection gradient is turned off.

Following the slice selection, a second gradient can be applied along one of the two remaining spatial dimensions while receiving the signal from said excited slice. This is the basis of what is called frequency encoding.

By having this second gradient running along one of the other spatial dimensions, the protons within the slide will precess with different frequencies along the respective dimension, which in turn will allow for a differentiation in their spatial position in the received signal.

The gradient applied during frequency encoding is usually designated as the read-out gradient due to being tuned on during reception of the signal coming from the protons. The strength of the signal at each frequency is dependent of the strength of the signal at the specific encoded position.

The sum of all frequency responses will obviously consist of all the measured composite time domain signal. This information allows us to use the Fourier transform (FT) to obtain from the composite signal, the strength of the signal at each specific frequency. By other words, the FT provides the frequency-specific information of the signal that, in the end, merely represents the spatial information required to form an image. The image formed with the frequency-specific information uses a grey level to represent the strength of the signal at each picture element.

Even if, at this point, we're able to recreate an image from all the magnetic information received, one further step is required to be able to separate signal components origination from different positions along the second dimension in the imaging plane. By briefly adding an extra gradient in the static magnetic field with an orientation along the remaining (third) spatial dimension it is possible to phase encode the signal before its reception. The phase encoding merely consists of a slight change in the precession frequency in each row of the slice caused prior to the read-out. The affected protons within each row will precess with a systematic phase shift along the positions on said row after the phase

encoding gradient is turned off. The amount of phase shift will depend on the position of the proton in along the encoded second image dimension.

The phase encoding gradient is turned off a moment before the echo read-out as its existence would cause an ambiguous spatial encoding if coupled with the frequency encoding gradient. The two gradients coexistence on the receiving of the echo would result in two-dimension frequency encoding gradients and thus, ambiguous spatial encoding. In the end would represent two different spatial positions being coded in a similar way, much like different numbers can be added to represent the same number.

Through proper combination of both frequency encoding in one dimension and phase encoding in the remaining dimension, all positions within a 2D image (or slice) can be uniquely coded with a determined resolution.

This is, though, without lack of a flaw: the phase-encoding gradient by its own is not sufficient to encode the second image dimension without being repeated many times for a single slice. The repetitions alter in the strength of the phase-encoding in order to, ultimately, obtain a complete phase and frequency encoding of the selected slice.

1.2.2.4. Two-dimensional κ-space and the formation of the 2D Echo-Planar Images

All the procedures, if applied correctly, will provide a set of data resultant from the series of excitation-recording. This data is then arranged into a two-dimensional space, in frequency space, designated κ-space. Each row of the κ-space corresponds to the data gathered by one cycle of the excitation-recording cycle with different phase-encoding gradient strength. While the slice selection and frequency encoding gradients are constant throughout cycle to cycle, the slope of the phase encoding gradients changes constantly by a set amount from cycle to cycle. This imposed shift for a specific Hydrogen proton depends on the strength of the phase encoding gradient and position of the proton along the second image dimension.

κ-space is then filled in such a way that the second slice dimension also gets frequency encoded. The κ-space contains then two-dimensional frequency encoded information of the slice, which then is transformed into two-dimensional image space by the application of the two-dimensional Fourier transform (2D FT).

This procedure is applied for each slice of a scanned volume. In order to construct a single or multiple 2D images from electromagnetic echoes, a particular sequence of electromagnetic pulses must be applied. These sequences are given the name of MRI pulse sequence and the details of a specific pulse sequence named echo-planar imaging (EPI) will be given in ***Subchapter 1.2.3.4***.

1.2.3. Physiological Principles of functional MRI

The MR images and the techniques used for their acquisition have been perfected throughout the years to describe at its best possible a plethora of biological tissues' properties. Even though MRI is heavily used for its structural view of, for example, the brain matter; to map functional changes in this organ and to track or trace neuronal activity, MRI requires an extension of its approach through time.

This concept is the basis of functional magnetic resonance (fMRI) and the focus of this subchapter. It may seem like fMRI is simply the study of the fluctuation and differences in time of the magnetic properties between arterial blood (oxygenated) and venous blood (deoxygenated), the details and caveats of these types of studies, the processing and interpretation of the data involved are enough to form a distinct field of study.

1.2.3.1. Neurovascular Coupling and The BOLD effect

The existence of neuronal activity is one that requires a constant consumption of energy. Said energy is produced via chemical processes that require the presence of glucose and oxygen which are supplied by the complex vascular network of small and large vessels. Oxygen is transported in the blood via the hemoglobin molecule. Hemoglobin is a molecule that can have two times of terminology associated depending of the presence of oxygen: if the hemoglobin present in the vessels is carrying blood, it is designated oxygenated hemoglobin (HbO_2); if this is not the case it is called deoxygenated hemoglobin (Hb).

When talking about the arterial part of the vascular system, we see a transport of oxygenated blood through an increasingly fine-grained network of blood vessels until its dimensions allow for a chemical transfer of the stored energy (oxygen) to the neurons.

In the brain's resting state only 30 to 40% of the oxygen present in the blood of the capillary bed is extracted. Both the capillary bed and the venous network contain a mixture of oxygenated and deoxygenated hemoglobin. Due to this the venous system, counterpart to the arterial system, performs the transportation of the less oxygenated blood away from the capillary bed to the pulmonary system, thus completing the cycle of blood transportation back to where the blood oxygenation takes place.

There is a local increase of neuronal activity, because of the frequent neuron firings happening in the brain. This event increases the oxygen extraction rate in the capillary bed which, in turn, will increase the relative concentration of deoxygenated hemoglobin.

The vascular system response to the increased energy demand is designated as hemodynamic response. The hemodynamic response is characterized in the local increment of cerebral blood flow (CBF), delayed approximately one to two seconds, and the cerebral blood volume (CBV), due to an increase of the size of vessels. The combined interplay between oxygen extraction, CBV and CBF results in what is classified as the hemodynamic response function (HRF), which is no less than the time course of evoked fMRI signals during an activation of a brain region.

While the HRF has a plethora of physiological events happening in the three phases that it consists of, it is important first to understand the underlying physical properties of the blood and how the MRI becomes sensible to measure these effects due to what is called the BOLD.

Oxygenated hemoglobin has different magnetic properties than deoxygenated hemoglobin, more specifically, while the first is diamagnetic the second is paramagnetic and thus alters the local magnetic susceptibility and inducing magnetic field distortions within and around blood vessels in the capillary bed and venules. During the mentioned oversupply phase of the hemodynamic response, the oxygenated to deoxygenated hemoglobin ratio increases, resulting in a more homogeneous local magnetic field. This homogeneity causes the spins dephasing to happen slower, or in other words, the measured MRI signal in homogeneous magnetic fields is stronger, which leads to the association between the activated brain states and higher measured MRI signals while compared to the resting states. The BOLD effect measures increased neuronal activity indirectly via the local changes in magnetic field (in)homogeneity, caused by oversupply of oxygenated blood following the neuronal activation (***Figure 1.7***). This effect is the most common primary form of fMRI contrast, discovered by Ogawa, S. [37], being detectable with MRI due to the different magnetic properties of oxygenated and deoxygenated hemoglobin, acting as endogenous marker of neural activity.

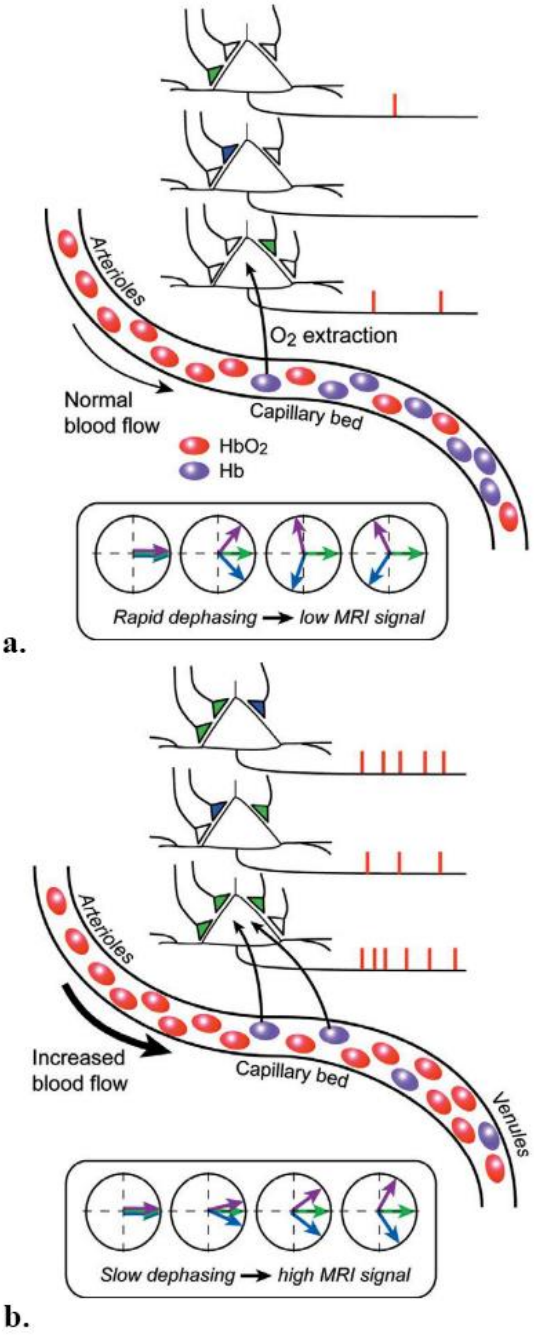

Figure 1.7 a,b. Schematic from neural activity to BOLD MRI response. **a** Represents the case of a cortical region in baseline mode. Here the neural activity is low, the cerebral blood flow (CBF) is at a basal level and the constant oxygen rate that fuels neuronal activity causes the 1:1 ratio of Hb and HbO_2 to distort the magnetic field, due to Hb presence. **b** Represents the case of an activated state in a cortical region. This will lead to an increased oxygen extraction rate, CBF strongly increases delivering more local oxygen beyond local need (overshoot), pushing HbO_2 away from capillary bed. Since HbO_2 does not substantially distort the homogeneity of the magnetic field the MRI signal is enhanced. *Adapted from* [40].

As mentioned before, the hemodynamic response function is divided into several phases: initial dip, main response and undershoot. The main response phase of the HRF is dominated by the quick and strong CBF in response to the slightly increased oxygen extraction rate. Because of that, a local substantial oversupply of oxygenated hemoglobin is originated, and the MRI signal becomes higher due to the homogenization of the magnetic field. While the details of this supply of cerebral blood flow in an overly unnecessary fashion are unclear

and still unexplained, it is known to peak for up to five seconds after the remarkable overshoot. After the stimulus end a post-stimulus undershoot lasting from twelve to fifteen seconds occurs until the signal returns to the baseline value. The duration of the undershoot depends on the time intensity and span of the preceding neuronal activation as well as the brain region activated. Finally, although theoretically expected but not reliably detected in standard human fMRI measurements (for animal studies, see [41]), the initial dip phase, which is characterized by the initial moment of extraction of the oxygenated hemoglobin out of the capillary bed.

An example of an HRF is presented below with all the phases described previously can be seen in the following ***Figure 1.8***.**Figure *1.8*.** Idealized time course of the hemodynamic response following a long stimulus (around twenty seconds). The initial dip is not reliably measured in human fMRI studies. The long stimulus result in a signal that rises initially to a higher value (overshoot) followed by the plateau. Posterior to the stimulus, the signal falls below baseline signal level (undershoot).

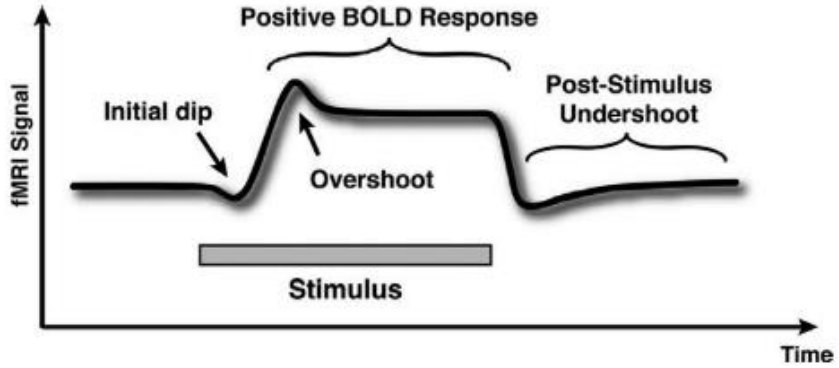

Figure 1.8. Idealized time course of the hemodynamic response following a long stimulus (around twenty seconds). The *initial dip* is not reliably measured in human fMRI studies. The long stimulus result in a signal that rises initially to a higher value (*overshoot*) followed by the plateau. Posterior to the stimulus, the signal falls below baseline signal level (*undershoot*).

1.2.3.2. The search for higher Spatial-Temporal Resolution in fMRI

On the field of fMRI studies, spatial resolution is described as the ability to discriminate between close locations. Since fMRI is a medical imaging technique, the spatial resolution is associated with the size of the voxels, just as in MRI. Unlike standard MRI, fMRI measurements have the goal to localize neural correlation of sensory, motor or cognitive processes. The optimum situation to study brain function would be to be able to individually analyze the columns and layers that form its mesoscopic architecture. Due to these being known to be of the of an approximate size of one millimeter, the desired spatial resolution falls under the one-millimeter mark.

However, there are some issues that arise from this search for the minimum value for the spatial resolution. First of which comprises the fact that the signal originated from a voxel is directly proportional to the volume of said voxel. That and the fact that the signal to noise ratio (SNR) and field inhomogeneities have an approximately linear increase with the B_0 field strength (property associated with higher spatial resolution), creates an overall signal loss when trying to improve the resolution and distortions on the reconstructed images. Lastly, the increase of resolution has a direct effect on the amount of data to sample, resulting in much higher acquisition times.

On the other side, temporal resolution is defined as the smallest period of neural activity reliably separated out from the preceding and succeeding ones. In fMRI, this is however limited by the

hemodynamic response time: the typical BOLD response has a width of ~3 seconds and a peak occurring ~ 5/6 seconds after the onset of a brief neural stimulus, hence the temporal information is heavily blurred as the imaging technique can't keep up with the underlying neural processes occurring. The TR is then crucial in the endeavor of higher temporal resolutions for it can be considered the sampling time associated with the fMRI as a technique, consequently, shorter TR sequences are generally desirable.

Whole-brain coverage on an fMRI study poses an issue, since the increased number in slices to be acquired makes a whole volume TR very difficult to be obtained if standard imaging techniques are used. The following ***Subchapter 1.2.3.4*** is dedicated to the understanding of advanced imaging techniques that overcome some of these challenges.

Furthermore, an fMRI study, due to its longer acquisition times, has inherently more physiological motion effects (such as head/body motion artefacts) that need to be correctly sampled to be effectively removed in post-processing. On the same line than the last issue, the more sampling points there are in time the more statistical power the fMRI measurement will have.

Spatial and temporal frequencies on a scale of the ones present in the functional organization of the human cortex have been achieved [31] by using high density surface coil arrays in close proximity to the head at scanners with a field strength of 7T in combination with acceleration techniques for image reconstruction. These acquisitions have their own set of challenges briefly explained in the next subchapter.

1.2.3.3. High fields technical challenges in acquisition

High magnetic field strengths, such as 7T, in *in vivo* fMRI studies has additional challenges brought by the large B_0 and B_1 inhomogeneity and the increased power of the RF pulse.

Inhomogeneities, as mentioned throughout chapter 1.2 of this thesis report, lead to a severe degradation of the imaging quality.

The reduction of the layer thickness in the MRI images helps reducing the inhomogeneity of the B1 field issue while helping to increase the resolution of said images. Although, by addressing the problem with this solution, there must be either an increase in the measurement time or a reduction in the covered volume.

Solutions like partial parallel imaging (PI) or segmented acquisitions help in further increasing the resolution by reducing the amount of κ-space measured in the acquisition.

The improvement in the imaging is noticeable whenever PI techniques, like sensitivity encoding (SENSE), are applied, with the only caveat being the use of suitable coil arrays during the acquisition.

To solve transmit field (B_1^+) inhomogeneities that arise with the increasing field strength and shorter RF wavelength, several techniques like shimming have been used. The shimming is a local field inhomogeneity correction that is applied before the data acquisition.

In the next subchapter a special attention is given to two acceleration techniques for imaging reconstruction: EPI, as the pulse sequence responsible for fast acquisitions in fMRI studies, and SENSE, as a PI technique used in the acquisition of data for this study.

1.2.3.4. Parallel Imaging and Acceleration Techniques in fMRI

PI have been used in the past to accelerate clinical routine acquisitions in the order of two to three times.

A specific branch of PI techniques, named partially parallel acquisition (PPA), consists of the application of strategies to use the spatial information contained in the component coil to reduce the

amount of special information encoded using gradients. In doing so, acquisition times are reduced as the full field of view of a PPA acquisition is reconstructed from the few lines conventionally encoded.

The most frequent pulse sequence used for rapid imaging in fMRI is echo planar imaging (EPI).

This sequence comprises an excitation pulse preceded by a magnetization preparation, meanwhile a continuous acquisition signal in the form of a gradient echo train is used to rapidly fill in the κ-space within a single TR (***Figure 1.9***).

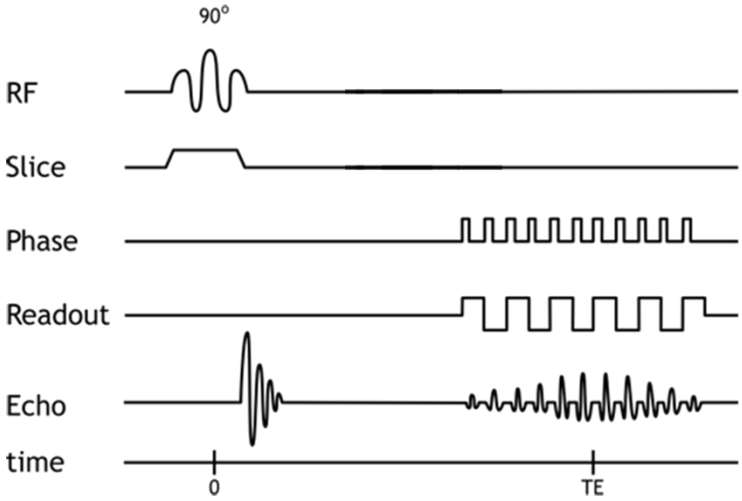

Figure 1.9: Pulse sequence diagram for EPI. The multiple gradient echoes are formed within a spin echo. Additionally, each pulse of the phase encoding gradient begins a new line of the κ-space.

Echo planar imaging can be performed based both on gradient echo and inversion recovery sequences. It is often possible to acquire the entire κ-space matrix in more than one shot, acquiring several chunks over the course of, say, three to four repetitions (***Figure 1.10*** and ***Figure 1.11***).

The readout and phase-encoding gradients are adapted to spatial image encoding, having several possible trajectories to fill the κ-space (constant or intermittent phase encoding gradient and spiral acquisition are a few examples).

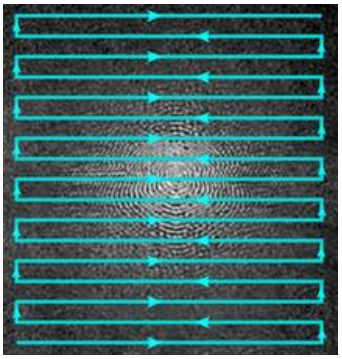

Figure 1.10: Example of a zig-zag traversal of the κ-space in one of the early EPI techniques. Frequency-encoding is along horizontal axis; phase-encoding is long the vertical. *Figure from* [42].

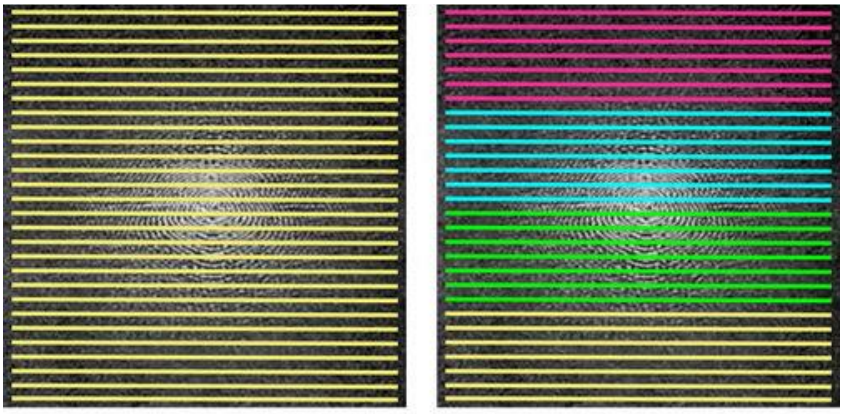

Figure 1.11: Representation of the difference between the single shot EPI sequence that acquires the whole of the κ-space within the TR and the multi-shot EPI sequence that divides the κ-space acquisition between segments (each coloured region represents a region). *Figure from* [42].

In a functional MRI study, this sequence is considered a pulse sequence with good sensitivity regarding signal changes in the cortex with a temporal resolution of two to three seconds and a spatial resolution of about three to four millimeters.

However, to reduce the influence of physiological fluctuation and spatial resolution without drastically increasing acquisition times, partial PI techniques come into play.

SENSE is one of the techniques used to accelerate the acquisition process. The technique is based on the fact that the receiver sensitivity has generally an encoding effect that is complementary to the Fourier preparation done by the linear field gradients. The scan time is then reduced by using multiple receiver coils in parallel during this stage.

In 2D Fourier imaging with common Cartesian sampling of the κ-space, sensitivity encoding represents a reduction in the time spent in the Fourier encoding steps, achieved by increasing the distance in sampling positions in the κ-space while maintaining the maximum κ-values (maintaining the spatial resolution).

The factor by which the reduction of the κ-space is done is referred to as the reduction factor or acceleration factor (R). When describing standard Fourier imaging, the reduction of the sampling density leads to a reduction of the FOV, causing aliasing. The SENSE reconstruction when applied in the Cartesian case is usually performed after the creation of one intermediate aliased image per array element by applying the discrete Fourier transform (DFT) and then combining them all into a full-field of view (FOV) image. The fold-over effect that would be caused by the combining the intermediate images of the reduced FOV, with their signal contributions, into the full FOV needs to be addressed (***Figure 1.12***). By separating the signal contributions given to the full FOV from each pixel of the reduced FOV we avoid the fold-over effect. That is done due to the signal superposition of all the single coil image signals with their respective weights according to the local coil sensitivity.

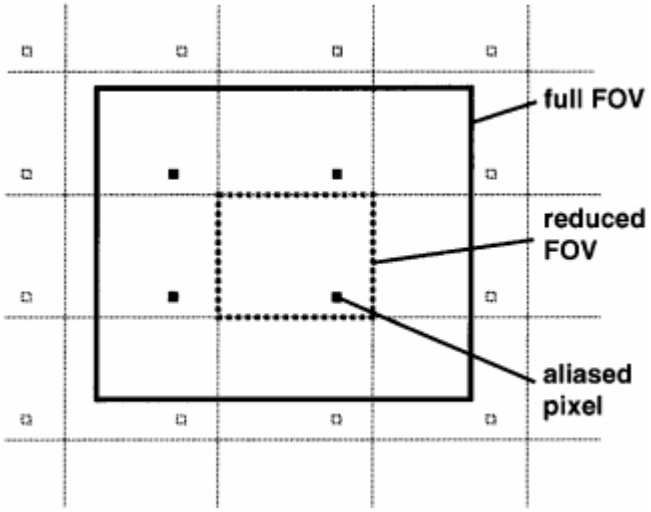

Figure 1.12: Aliasing in 2D Cartesian sampling: the full FOV, represented by the solid box, is reduced both in dimensions. A pixel in the reduced FOV, represented by the dotted box, represents the superposition of pixels forming a Cartesian grid. In this schematic example, four complete pixels are in the full FOV; causing the aliasing degree to be of four. *Adapted from* [43].

If we consider the case in which n_P is the number of superimposed points and n_C is the number of coils used, by analysing one pixel in the reduced FOV and the corresponding set of pixels in the full FOV we can assemble the vector $\vec{a}$, representing the complex image values that the pixel takes in all the intermediate images. The complex coil sensitivities at the n_P superimposed positions form a $n_p \times n_c$ sensitivity matrix **S**.

S can be used to calculate the unfolding matrix **U**, defined as:

$$U = (S^H \Psi^{-1} S)^{-1} S^H \Psi^{-1} \tag{1.4}$$

where **H** indicates the transposed complex conjugate and Ψ is a $n_p \times n_c$ noise matrix, describing the level and correlation of noise.

The signal separation is performed by

$$\boldsymbol{v} = U\boldsymbol{a} \tag{1.5}$$

where **v** is the resulting vector that contains the separated pixel values for the originally superimposed positions. The non-aliased full-FOV image is obtained by repeating this process for each pixel in the reduced POV.

SENSE images can be affected by two distinct types of noise: noise in the sample values and noise in the sensitivity data. The second can be reduced until a negligible level. The noise in SENSE reconstruction is different from channel to channel and there are correlations between samples taken simultaneously. Furthermore, the noise level in SENSE imaging varies from pixel to pixel while also existing noise correlation between pixels.

Due to the mathematical theory behind SENSE imaging, there are two parameters that have a general meaning in other techniques and thus their definition is valuable. The reduction factor R denotes the factor by which the number of samples is reduced with respect to the full Fourier encoding; n_K the number of sampling positions in the κ-space, R can be defined as:

$$R = \frac{n_k^{full}}{n_k^{red}} \tag{1.6}$$

The local geometry factor g describes the capability to separate superimposed pixels by aliasing with the adopted coil configuration:

$$g = \sqrt{[(S^H \Psi^{-1} S)^{-1}](S^H \Psi^{-1} S)} \geq 1 \tag{1.7}$$

and strongly depends on coil geometry.

Since SNR of the reduced FOV (SNR^{red}) can be expressed as

$$SNR^{red} = \frac{SNR^{full}}{g\sqrt{R}} \tag{1.8}$$

g allows an SNR estimate and furnishes a criterion in the coil array design.

The most important advantage of SENSE is its flexibility because reconstructed images are free of specific artefacts, depending on the coil configuration, whereas the main disadvantage presented is the necessity for an accurate coil sensitivity map.

1.3.Brain organization: a visual cortex study

Due to the nature of the project and its objectives, a brief explanation of the brain organization and the visual cortex is assessed in this subchapter.

As complex as the human cortex is, its function is fundamentally organized into columns and layers. While the columns are known to be grouped neurons focused on performing the same function with dimensions inferior to one mm in diameter; the layers are thin sheets of less than one mm of thickness of neurons including the cortex where the neurons are organized with respect to cortical depth.

The visual cortex is the part of the cerebral cortex responsible for the processing of visual information and it's located in the occipital lobe in the back of the head. The visual cortex is divided into six regions: the primary visual cortex (V1), or striate cortex, is responsible for the reception of sensory input from the thalamus; and the remaining five areas, also named extrastriate areas, V2, V3, V4, V5 and V6, are responsible for the processing the visual information coming from the primary visual cortex.

The primary visual cortex is also organized into multiple horizontal layers and vertical columns with a thickness of about two mm. The cells display themselves in a vertical organization according its eye of origin, forming alternating parallel ocular dominance columns. A second set of columns is specific for stimulus orientation and responds in the same direction of a light slit of edge.

On the other hand, the striate cortex has the purpose of combining and analyzing the visual information and transmitting the information present to the remaining areas on the extrastriate cortex, present in the vicinity of the primary visual cortex, responsible and related to visual processing [44].

In ***Figure 1.13*** we can observe a schematic representation of the physiological organization of the visual cortex in columns and layers and its connection to the LGN.

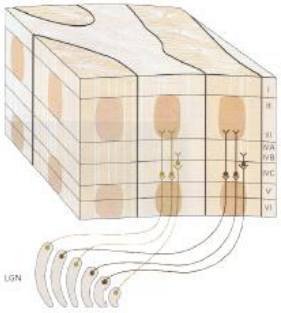

Figure 1.13: Layer and column as core functional parts of the visual cortex. *Adapted from* [45].

2. Methods and Materials

One of the major goals of fMRI measurements is the location of the neuronal correlates of sensory, motor and cognitive processes. Another major goal of fMRI measurement is the characterization of the response profile in specific regions of interest (ROIs) and the plotting of the average signal time courses that arise from different experimental conditions. The close inspection of the shape of said time courses helps to separate signal fluctuations due to measurement of artefacts from stimulus-related hemodynamic responses.

Given the relatively small amplitude of task-related BOLD signal changes of typically 1-5%, and the continuous presence of a plethora of confounding effects, like signal drifts and head motion, the localization and characterization of certain brain regions responding differently to certain conditions of the stimulation protocol, makes this a non-trivial task.

The major analysis steps of functional and associated anatomical data will be described in this chapter right after a small explanation and demonstration of the protocol followed in this study by the participants and the task to be performed.

2.1. Subjects and Procedure

A Philips 7T scanner with two 16-channel surface coils was used to collect the BOLD fMRI responses were from five healthy human volunteers (V4493, V4494, V4528, V5069, V5107 were the identification codes used for the subjects) while the observation of a visual motion stimulus was presented. All five subjects had normal or corrected-to-normal vision. All the anatomical and functional data was acquired with the understanding and written consent from each subject in accordance with the protocol approved by the Medical Research Ethics Committee of the UMC Utrecht, and in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki 2013).

During each scanning session, six scans were acquired (the first scan was acquired in order to independently localize the activity in the hMT+, hence gaining the definition of a localizer scan; while the remaining scans are referred as runs or EPI runs).

The visual stimuli presented consisted of several high-contrast black-and-white dartboards with spatial frequencies of 1, 3 and 5 deg/cycle. For each spatial frequency, the dartboard pattern expanded from the fixation point for 1 second at several different temporal frequencies while having a baseline of a homogeneous grey background. For the dartboard with spatial frequency of 3 deg/cycle the temporal frequencies were 1, 3 and 5 Hz randomly interleaved. As for the 5 and 1 deg/cycle dartboards the temporal frequencies were, on both, 3 Hz. A graphical representation of the combination of stimuli that were presented during the five scans (excluding the localizer scan) is demonstrated in ***Figure 2.1***.

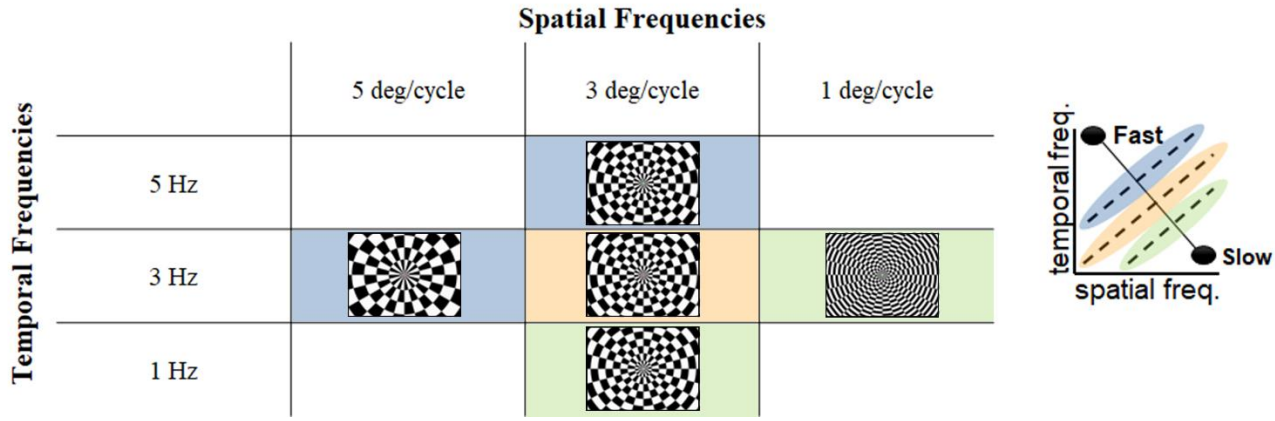

Figure 2.1: Graphical summary of the visual stimuli (high-contrast black-and-white dartboard) presented during the five runs of the study following the localizer scan. Each run is either classified as a fast (represented by blue background color), intermediate (represented by red background colour) or slow (represented by green background colour) depending on the spatial and temporal frequency that gives origin to its speed. Graphic on the right allows a better understanding of this classification.

The choice of the type of stimulus was based on previous in other fMRI studies regarding the encoding of the visual motion [46], [47] and the choice of stimuli speeds was also based on studies regarding visual motion processing, in which it was proved that optimal results in the hMT+ occurred at speeds between 1.5 and 30 deg/s [9], [25]–[28], [46], [48], [49].

The visual stimuli were created via MATLAB R2009a (The Mathworks Inc. Natick, MA, USA) and presented to the subject via reflection of a frosted glass screen that's placed inside the magnet bore behind its head, using a 1024 X 768-pixel LCD projector. These projected visual stimuli were viewed by the subjects via a mirror that was mounted in front of their eyes. The subjects fixated on a dot located on the centre of the screen that was alternation in between colours red and green at random intervals. To maintain accurate fixation and consistent level of arousal throughout the experiment, subjects were instructed to press a button every time that the fixation dot had a switch of its colour.

2.2. 7T MRI Acquisition

2.2.1. Localizer scans

Imaging was performed with a Philips 7T scanner two 16-channel high-density multi-element surface coil with minimal electronics [31]. For the localisers a GE-EPI sequence was used with a voxel size of 1.34 x 1.34 x 1.50 mm^3, TR/TE = 1800/27 ms, $\alpha = 60º$, SENSE factor = 2, FOV = 40.5 x 160 x 171.9 mm^3 and 27 slices (***Figure 2.2***).

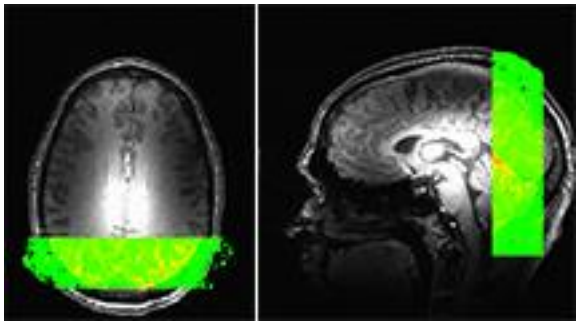

Figure 2.2: Projection of the localizer scan on the anatomical T_1 scan, representing the coverage of the former. Scans are not aligned.

A high resolution T_2*weighted (T_2*w) scan was acquired as an anatomical reference. In this acquisition a 3-D multi-shot GE-EPI was used, with a voxel size of 0.49 x 0.49 x 0.5 mm^3, TR/TE = 54.9/27 ms, $\alpha = 14^o$, SENSE factor = 2, FOV = 40 x 160 x 172 mm^3 (***Figure 2.3***).

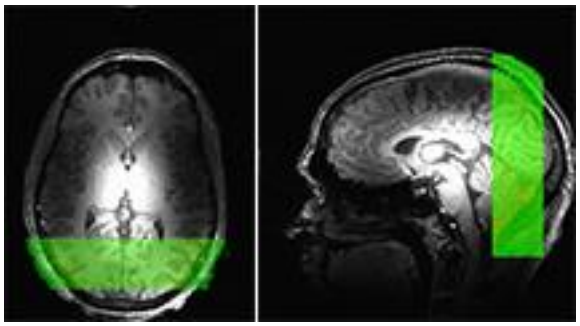

Figure 2.3: Projection of the anatomical T_2*w scan on the anatomical T_1 scan, representing the coverage of the former. Scans are not aligned.

In each the localizer scans, semi-circular dartboard stimuli were presented four times to left and four times to the right visual field, as well as a full circular dartboard stimulus presented four times, in a previous set-up order (full dartboard, left side moving dartboard, right side moving dartboard), with inter-stimulus intervals of 20s and a stimulus duration of 10s (***Figure 2.4***).

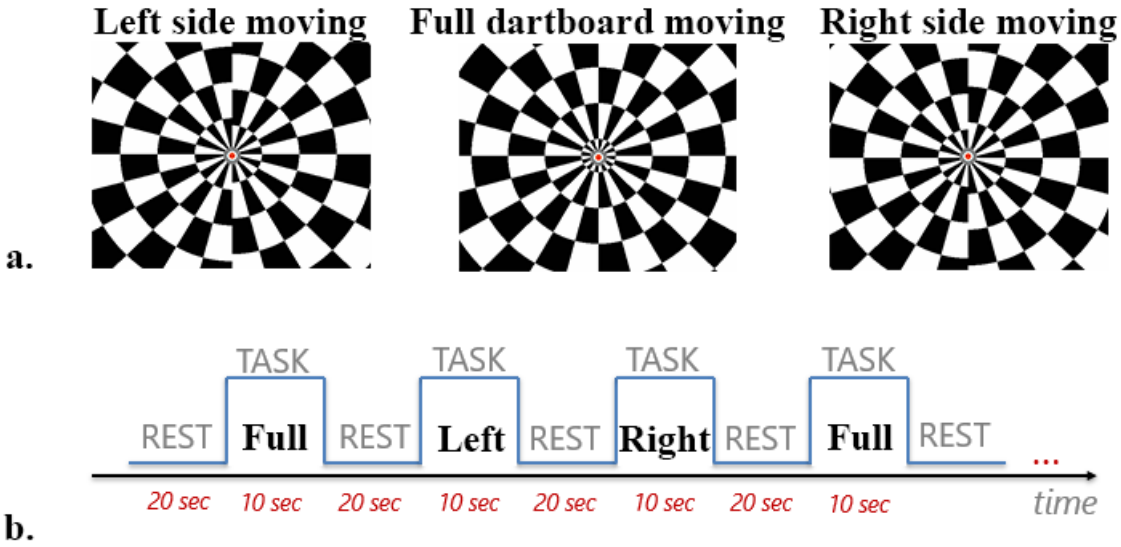

Figure 2.4 a. Representative frames of the localizer task in their several variations: dartboard with moving left side of the visual field, full movement of dartboard, dartboard with moving right side of the visual field. **b.** Schematic representation of the task with 20 second rest periods and 10 seconds of task. The task periods follow a set-up order of full dartboard movement, left-side moving dartboard and right-side moving dartboard in a total of 360 seconds of task overall (cycled is repeated four times).

2.2.2. Run scans

The functional data for the runs was obtained using a GE-EPI technique with a voxel size of 1.43 x 1.43 x 1.50 mm^3, TR/TE = 849/27 ms, $\alpha = 50^o$, SENSE [50] factor = 2, FOV = 22.5 x 160 x 160 mm^3, and 15 slices covering up the primary middle temporal cortex area.

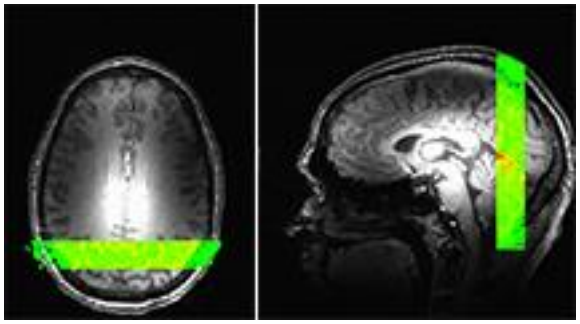

Figure 2.5: Projection of the EPI runs on the anatomical T1 scan, representing the coverage of the former. Scans are not aligned.

Finally, throughout the scans a pressure sensor was placed in the abdomen and a finger pulse oximeter was used to monitor and acquire respiration and heart-beat signals, which were afterwards used in the pre-processing with the purpose of offline removal of physiological artefacts contained in the functional images [51].

In each of the five runs, semi-circular dartboard stimuli were presented 26 times to the complete visual field in a predetermined order, with pseudo-randomized inter-stimulus intervals ranging from 7 to 25s.

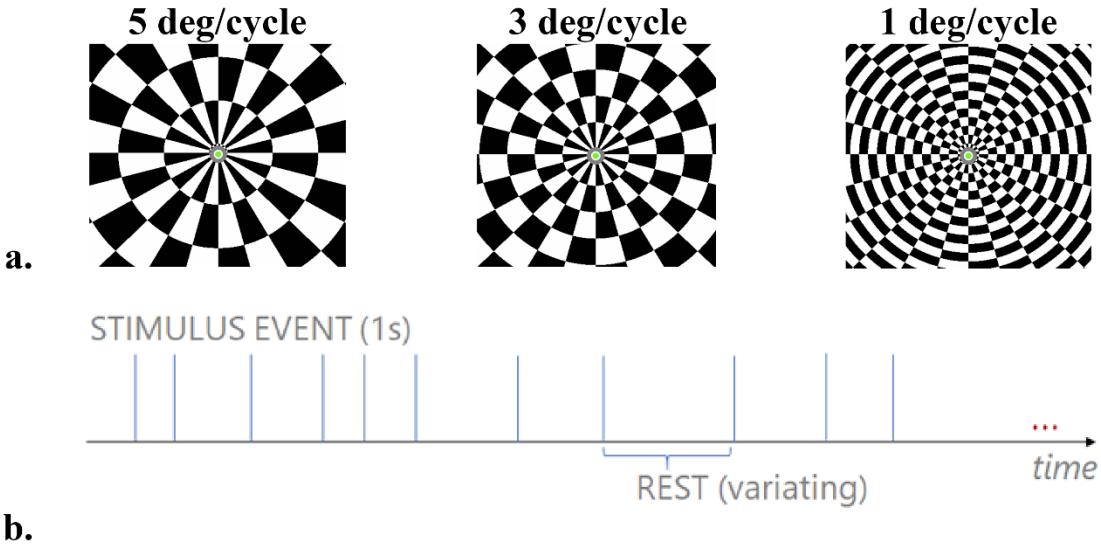

Figure 2.6 a. Representative frames of the several runs present in the task consisting of high-contrast black-and-white dartboards with spatial frequencies of 1, 3 and 5 deg/cycle. For the dartboard with spatial frequency of 3 deg/cycle the temporal frequencies were 1, 3 and 5 Hz randomly interleaved and for the 5 and 1 deg/cycle dartboards the temporal frequencies were, on both, 3 Hz. **b.** Schematic representation of the Event-related design task with stimulus events lasting 1s and with variating resting time intervals, comprising in total 360 seconds.

2.3.Functional BOLD data pre-processing

To reduce artefact and noise-related signal components a complete series of mathematical operations was set and performed prior to the statistical data analysis. The most essential steps of the pre-processing operations in this study include: head motion detection and correction (***Subchapter 2.3.1***), removal of linear and non-linear trends in voxel time courses (***Subchapter 2.3.3***), physiological motion effect corrections (***Subchapter 2.3.4***), general despiking (***Subchapter 2.3.5***) and alignment of functional and high resolution anatomical images (***Subchapter 2.3.6***).

All the mentioned pre-processing steps were performed using AFNI (Analysis of Functional NeuroImages) [52]. Briefly, the functional data were realigned in-plane to correct for motion, and

subsequently de-trended using AFNI and corrected for cardiac and respiratory fluctuations using RETROspective Image CORrection (RETROICOR) [53]. General despiking was executed in all the runs and localizers. All the images were aligned to their respective localizers and, if more than one localizer was present in the subject, to the first localizer taken. The stimuli times and the pre-processed images were used as input functions for a general-linear-model (GLM) analysis. Convolution was performed using a standard gamma variate hemodynamic response function.

The details of each processing step mentioned before are described in further detail in the next ***Subchapters 2.3.1-2.3.6***.

2.3.1. Motion Correction

The quality of fMRI data is strongly hampered by the presence of substantial head movements, so much that some data is even rejected from further analysis upon discovery of head movements equal or higher than 5 mm. Despite possibility of having the head motion correction done in the image space, any displacement of the head will reduce the homogeneity of the magnetic field, which is fine-tuned prior to the functional scan for the then current head position.

Motion correction was operated by selecting one functional volume of a run as a reference to which all the other functional volumes are aligned within the same run. Performing the alignments using Fourier interpolation, the head movements during the runs were described and recorded by six parameters: three translation (displacement) parameters and three rotation parameters. These six parameters are appropriate to characterize the motion of rigid bodies, describing the spatial displacement by translation along the x, y and z-axes and the rotation around these axes. The parameters were obtained by estimating iteratively on how a source volume should be translated and rotated to better align with the given reference volume, which in this case was chosen to be the 5^{th} volume in the acquisition. When no further significant improvements can be achieved, the iterative adjustment stops, and all the steps were applied to the source volume to produce a new volume, replacing the original volume in the motion corrected data set.

The motion parameters were recorded, saved and subsequently used to further remove residual motion artefacts.

Finally, it is important to mention that the head movements from volume to volume were analyzed and considered small, thus allowing the largely valid assumption of rigid body. Even so close inspection during the scanning of the runs was valuable, since the six parameters of a rigid body were incapable of registering within-volume movements and if these were constant and significant, it can lead to the loss of homogeneity during the image acquisition.

2.3.2. Slice Scan Time Correction

All the data obtained under a certain time point within the acquisition protocol is given the name of functional volume. However, the functional volumes usually have a sequence of 2D-EPI measurements distributed within them. Despite the sluggishness of the hemodynamic response (***Figure 1.8***), an impressive specification of the time in the order of a couple of seconds will lead to a suboptimal statistical analysis.

A way to attenuate this loss of statistical power is to preprocess the data in such a manner that the resulting data appears to have all the slices of a functional volume measured in the same moment in time. This temporal shift of the slice time courses requires the resampling the original data acquired at points that fall between measurement time points. The new values were estimated by interpolation of

values from past and future time points (***Figure 2.7***). To minimize the interpolation of the run scans acquired in this study, they were not slice time corrected, leaving only the localizers to be corrected by this procedure. The localizer scans were then corrected using the Fourier method with the slices alternating in the plus direction.

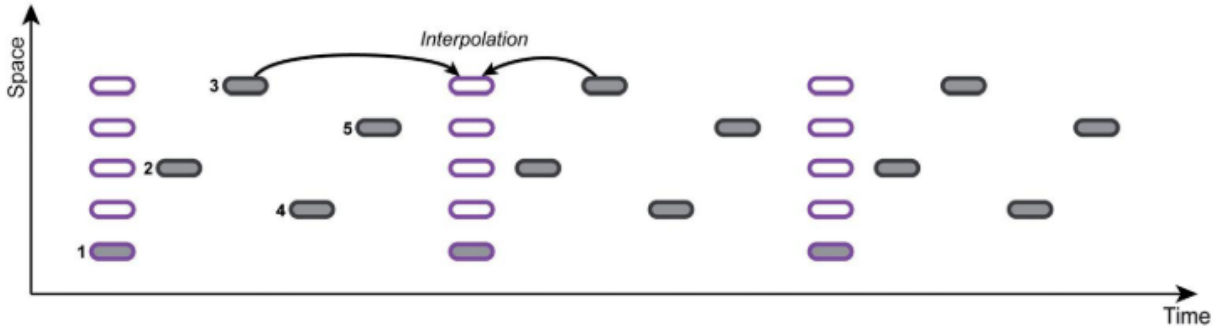

Figure 2.7: Slice time correction of a functional dataset. During the slice scan time correction, slices within each functional volume (the black triangles) are shifted in time resulting in a new resampling time series (violet rectangles). In this new time series, all slices of a functional volume are virtually measured at the same moment in time. To calculate the intensity values of the time points that fall in between measurement time points, past and future values were integrated using a sinc interpolation. Here, five slices are scanned in interleaved order.

2.3.3. Detrending

Due to physical and physiological noise, voxel time courses usually are non-stationary and exhibit signal drifts all along the acquisition period. If not accounted for, these drifts cause pronounced autocorrelation in the residual noise process and complicates signal detection and statistical procedures for threshold finding as proven by Friman, O. [54]. The lack of prior information about the drifts' nature makes it difficult to formulate a model of them. However, two major observations can be addressed regarding the nature of the drifts: due to the slow nature of the change over time that occurs with drifts, their components are well characterized by a high correlation; the drift components are likely to be present in multiple voxels. These two observations serve as background for the development of producing drift models that are adapted to the dataset at hand.

In this study, linear least squares detrending was performed on all datasets (***Figure 2.7***). This meant orthogonalizing the datasets with respect to the basic time series provided, which in this case were two two-degree polynomials (a sine and cosine wave) with a period of 428s and 429s for the runs and 200s and 195s for the localizers. These values were chosen due to their proximity to the TR of the respectful scans.

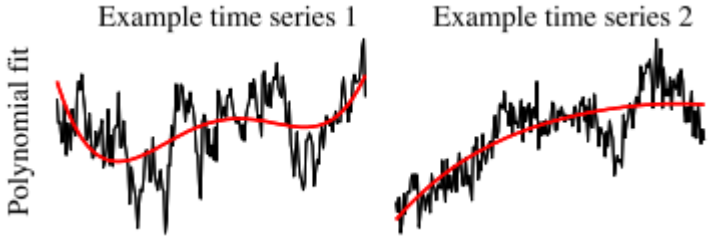

Figure 2.8: Schematic representation of the Least square fits (*in red*) of the detrending method (*polynomial fit*) applied in both the Localizer and Run datasets. *Adapted from* [54].

2.3.4. RETROICOR

Cardiac, respiratory and other low frequency is particularly prominent in the grey matter. This region is the primary focus in many fMRI studies [55], [56]. This physiological noise increases signal

variance, reducing the signal detection power. Also, the structural nature of this noise compromises two of the assumptions had for noise in most fMRI data analysis as described by Lund, T.E. *et al.* [57]: noise is usually considered both independent and identically distributed in fMRI acquisitions.

The pulsatility of blood and cerebrospinal fluid due cardiovascular processes will cause artefacts that tend to be localized spatially near the ventricles, sulci and large vessels as described by G. Glover, H. *et al.* and M. Dagli *et al.* [53], [55]. Respiration may be accompanied by the bulk motion of the head as well as modulation of the magnetic field by thoracic and abdominal movement and the noise is considered to be more spatially global as described by G. Glover, H. *et al.* [53]. These two types of physiological fluctuations were considered a dominant source of noise in BOLD fMRI, particularly at high fields [56] such as the one used in this study (7T).

To correct for these artefacts, a physiological correction technique by the name of RETROICOR [53] was applied. RETROICOR models the cardiac and respiratory fluctuations using a Fourier series defined by the phase relative to the cardiac and respiratory cycles at the time of acquisition. The cardiac and respiratory cycles can be described by the following equations:

$$y_{c/r}(x,t) = \sum_{m=1}^{M} a_{c/r}(x)\cos(m\phi_{c/r}(t)) + b_{c/r}(x)\sin(m\phi_{c/r}(t)) \quad (2.1)$$

$$\phi_c(t) = 2\pi(t - t_1)/(t_2 - t) \quad (2.2)$$

$$\phi_r(t) = \pi \frac{\sum_{m=1}^{100.rnd[\frac{R(t)}{R_{max}}]} H(b)}{\sum_{1}^{100} H(b)} \quad (2.3)$$

Where the phase of the cardiac cycle, $\phi_c(t)$, is defined by the time to the nearest preceding heart beat divided by the time between the heart beats (the cardiac period); the phase of respiratory cycle, $\phi_r(t)$, is defined by the depth of the breath at the time of the image acquisition relative to a histogram (scaled from 1 to 100) of the respiration depth across the entirety of the imaging run [58].

In equations (2.1) and (2.2), $y_{c/r}(x,t)$ is the cardiac or respiratory (c/r) induced signal fluctuation, $\phi_{c/r}(t)$, is the phase of the cardiac or respiratory cycle at the time of the acquisition, M is the Fourier fit order, $a_{c/r}$ and $b_{c/r}$ are the Fourier fit coefficients determined in the regression analysis, t_1, is the time of the preceding heartbeat, and t_2 is time of the following heartbeat. In equation (2.3) $R(t)$ is the respiration depth, R_{max} is the maximum depth of respiration, and $H(b)$ is the histogram of respiration depth over the entire imaging run.

2.3.5. Despiking

Another temporal preprocessing step applied, besides the Detrending detailed in ***Subchapter 2.3.3***, consists in temporal smoothing of voxel time courses. This step removes any high-frequency signal fluctuations considered noise. While despiking the signal leads to an increase the signal-to-ratio, it was not advisable for event-related designs due to the distortion estimation of temporally relevant parameters. This led to the decision of despiking the localizer scans by simply censoring the datapoints on the runs that had abnormal high-frequency fluctuations from the statistical analysis presented on ***Subchapter 2.4***.

2.3.6. Alignment

The alignment of the several scans that comprise the whole study is intimately related to the overall objective of the project: localizer scans were acquired to functionally detect the hMT+ areas in

an independent (non-regressive way). After the processing steps mentioned in the subchapters 2.3.1-2.3.6, the localizers went through statistical analysis and regression, resulting in defined ROIs where the hMT+ would be anatomically defined. Since the aim is to study is to grasp how the hMT+ processes different spatial and temporal frequencies (or speeds), we kept the interpolation to the minimum in the EPIs of the runs. This meant that the alignment of the scans in the study always happens towards the EPIs of the runs and thus, the hMT+ ROIs obtained from the localizers were put on the same image space as the runs to proceed with the statistical analysis on the later.

Finally, to align the functional EPIs of the runs to the anatomical information, the alignment matrix of the alignment between localizer scans (EPI) and T_2* (EPI) was obtained. Here a cost function that minimizes the interpolation, the Local Pearson Absolutes (LPA) cost function, was used. Because the T2* shares some anatomical information, it's recommended to align it to the T_1 anatomical. By multiplying the alignment matrixes of Localizer (EPI) > T_2* and T_2*>T_1 and inverting the resulting matrix we can align the T_1 anatomical image to the EPI of the runs, since the ROIs obtained with the localizers were already in the runs image space.

The acquisition of the ROIs and its application in the statistical analysis of the functional data in the EPIs of the runs is explained in the next chapter.

2.4.Statistical Analysis of functional data

Statistical data analysis aims to identify the brain regions that exhibit increased or decreased responses when compared to other conditions within the experimental conditions applied during a study. Due to the presence of physiological and physical noise fluctuations, observed differences between conditions may occur because of mere chance. The statistical analysis was required to explicitly assess the effect of measurement variability on estimated condition differences. In standard fMRI this independently performed for each voxel time course and the collection of the obtained statistical values (one per voxel) form a three-dimensional statistical map.

2.4.1. The General Linear Model

The general linear model (GLM) aims to predict and explain the variation of a dependent variable in terms of a linear combination (or weighted sum) of a plethora of reference functions. The mentioned dependable variable corresponds to the observed fMRI time course of a voxel and the reference functions to the time courses of excepted fMRI responses for the specific conditions of the experimental paradigm followed in the study. The reference functions are called predictors or regressors and when collected in a set they form the design matrix, also called the model. A predictor time course is typically obtained by convolution of a condition box-car time course with the standard hemodynamic response function, which in all the analysis in this study was assumed to be a standard gamma variate hemodynamic response function. Each predictor time course X gets an associated coefficient or beta weight b, quantifying its potential contribution in explaining the voxel time course *y*. The voxel time course y is modeled as the sum of all the dined predictors multiplied by their associated beta weight *b*. Since the model can't explain the noise fluctuations, an error value *e* is added to the GML system of equations. Consecutively, the GLM is comprised of a collection of equations with *n* data points and *p* predictors as follows:

$$y_1 = b_0 + b_1 X_{11} + \cdots \cdots \cdots + b_{\mathrm{p}} X_{1\mathrm{p}} + e_1)$$
$$y_2 = b_0 + b_1 X_{21} + \cdots \cdots \cdots + b_{\mathrm{p}} X_{2\mathrm{p}} + e_2)$$

$$y_3 = b_0 + b_1 X_{31} + \cdots \cdots \cdots + b_{\mathrm{p}} X_{3\mathrm{p}} + e_3) \tag{2.4}$$

$$y_n = b_0 + b_1 X_{\mathrm{n}1} + \cdots \cdots \cdots + b_{\mathrm{p}} X_{\mathrm{np}} + e_{\mathrm{n}})$$

The *y* variable corresponds to the measured time course of a single voxel and with each line we advance in time (y_1 corresponds to the measured value at time point 1 and so on). This means that when applying the GLM we estimate the measured signal as a combination of the estimates (*b* values), their extraction depending on all the details present in the design matrix, plus the residuals, representing all the predictors not included in the model and reference functions not related to experimental conditions (such like the saved estimated motion parameters mentioned in chapter 2.3.1). All beta values together characterize a voxel's preference for one of more of the experimental conditions. Also, the weights of the estimates indicate how strongly does the brain is active during the modeled experimental condition (or task) relative to the baseline (or rest).

By changing the notation presented in Equations (2.4), and representing the matrix and vectors with single letters we get:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{e} \tag{2.5}$$

where the matrix **X** represents the design matrix containing the predictor time courses as column vectors. The beta values now appear in a separate vector **b**, with the term **Xb** representing the matrix vector multiplication. Given the data **y**, the GLM fitting procedure tries to find the set of beta values explaining the data as best possible.

To draw a palpable measurement of how good the fitting done by the GLM is, the least squares estimated method was applied and this method allows us to decompose the variance of the measured time course into the sum of the variance of the predicted values (model-related variance), plus the variance of the residuals, as shown in Equation 2.6:

$$\mathrm{Var}(\mathbf{y}) = \mathrm{Var}(\hat{\mathbf{y}}) + \mathrm{Var}(\mathbf{e}) \tag{2.6}$$

By trying to find the minimum error variance by least squares, we also find the maximum variance of the values explained with the model, because the variance of the voxel time course was fixed. The square of the multiple correlation coefficient R provides then a measure of the proportion of the variance of the data which is explained by the model:

$$R^2 = \frac{\mathrm{Var}(\hat{\mathbf{y}})}{\mathrm{Var}(\mathrm{y})} = \frac{\mathrm{Var}(\hat{\mathbf{y}})}{\mathrm{Var}(\hat{\mathbf{y}}) + \mathrm{Var}(\mathbf{e})} \tag{2.7}$$

In order to test whether a specified model significantly explains variance in a voxel time course, the R value is then transformed into an F statistic with the *p-1* degrees of freedom in the numerator and *n-p* degrees of freedom in the denominator:

$$F_{n-1,n-p} = \frac{\mathrm{R}^2(n-p)}{(1-R^2)(p-1)} \tag{2.8}$$

By knowing the degrees of freedom, this F test can be converted in an error probability value *p*. A high F value (and low *p* value) indicates that one or more experimental conditions substantially modulate the data time course. While the overall statistic F may express whether the specified model significantly explains a voxel's time course, it can't provide an assessment of which individual conditions differ significantly from each other. To do so, we need to formulate linear combinations of

beta values corresponding to null hypothesis in a t-test. These are called contrasts and allow for comparison between conditions.

All these statistical steps are describing a single voxel's time course since standard statistical methods are performed independently for each voxel. However, if the statistical test results for all individual voxels are integrated in a 3D data set we can get a visualization of the statistical map. Here the original voxels showing anatomical information are replaced with the intensity of the obtained statistical values of, say, t values for a contrast. Applying a threshold on a statistical map allows for the visualization of significant signal modulations on specific regions of the cortex.

In this study, the GLM analysis was done with the stimuli times, the estimates of motion correction and the pre-processed images were used as input functions. A threshold of 0.05, Bonferroni's corrected, was utilized when visualizing the activation on the localizer scans. The Bonferroni correction method is a simple multiple comparison correction approach adjusting the single-voxel threshold in such a way that the error probability of, in this case, 0.05 was retained at the global level. This was done to restrict the supposed area of the hMT+ in order to create ROIs to be used in the runs. Furthermore, the HRF extraction was realized within the projected ROIs in all the run scans.

2.4.2. HRF Models and Deconvolution

The hemodynamic response to an infinitely short stimulus is not an infinitely short increase in the BOLD signal. The signal variation following a sensory stimulation has a particular shape and delay. These temporal characteristics must be considered in the GLM in order to obtain a better fit of the model with the data present. This is done by convolving the design matrix with a model of the HRF to obtain a new model that reflects the variation of the fMRI rather than the variation of the experimental design. Three models are currently implemented: double Gamma function (the HDR is modelled after the difference of two gamma distribution functions); basis set (the HDR is modelled using a number of independent components that span a plausible space of shapes for the HRF) or by using deconvolution. In this last special case there is no attempt to explicitly model the HRF. The deconvolution was the methodology used when extracting the fMRI signal from the different runs. The deconvolution process tries to estimate the shape of the HRF by the use of a identity matrix whose size defines on how many TRs the HRF is being estimated on.

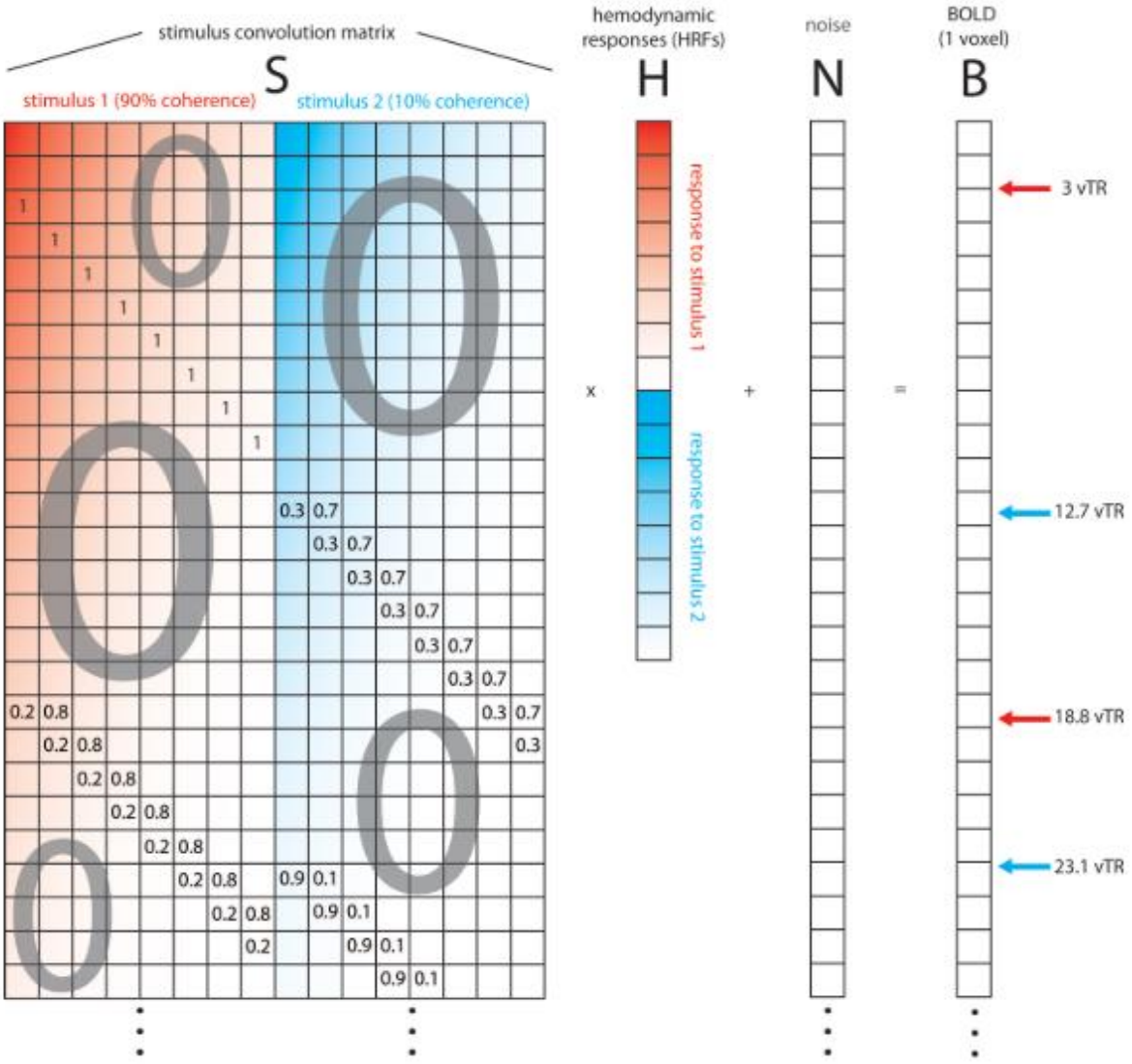

Figure 2.9: Example of the matricial model for estimating hemodynamic responses by the deconvolution approach used in the runs. For simplicity, each estimated HRF in this figure is an 8×1 vector. For example, the stimulus presentation times are indicated on the right and expressed in volume-TRs (vTR). The stimulus convolution matrix S on the left is constructed based on stimulus presentation times. The HRFs are estimated by solving $H = (S^T S)^{-1} S^T B$. *Courtesy of* [59].

3. Results

To be able to characterize and profile the spatiotemporal properties of visual motion in the human middle temporal cortex by using an fMRI paradigm and acquisitions in 7T, a plethora of methodologies described in the former chapter, 2. Methods and Materials, were applied.

In this chapter we focus on demonstrating any relevant data, even if intermediate in the grand scheme of the project, to allow for a complete, careful and logical observation of the steps taken during the study to answer the research questions proposed in the beginning of this thesis report. Any discussions or justifications on the choices made or the data presented will be present in the following chapter, 4. Discussion. Also, any repeated methodologies applied in multiple scan acquisitions during the study will be represented symbolically as to avoid cluttering the Results chapter.

All five subjects were submitted to the localizer task and the run task. The localizer and the run scans were corrected for motion (***Subchapter 2.3.1***) using as base the fifth volume of each series ***(Figure 3.1***). Fourier interpolation was used in the motion correction. In the case of subjects with multiple localizer scans, the second localizer was motion corrected to the first localizer as a base and then concatenated to provide stronger statistical power to the functional detection of the hMT+ areas.

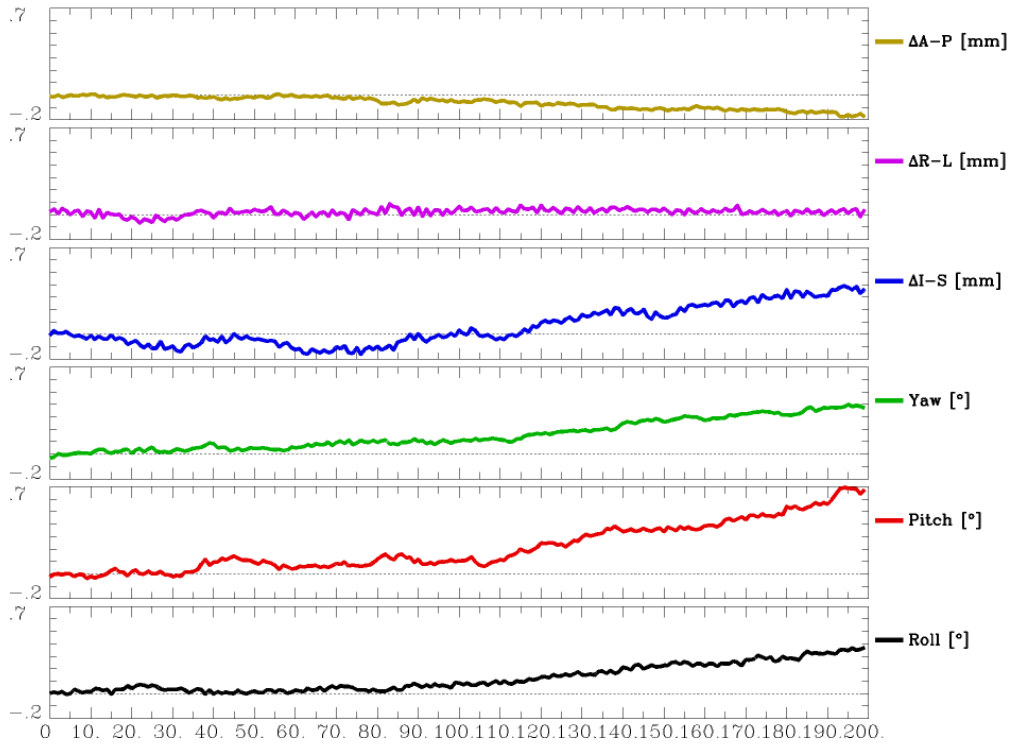

Figure 3.1: Representation of the registration parameters (*from top to bottom*: difference in the anterior-posterior, right-left, inferior-superior directions and yaw, pitch, roll angles) upon performing a motion correction using Fourier interpolation on one of the localizer scans. These parameters were saved to be used later during the GLM Analysis. All scans, both localizer and run scans, were aligned to their fifth volume of the series.

Subsequently, the physiological artefacts (cardiac time series and respiration rate) were removed by application of RETROICOR (2.3.4) in both localizer and run scans (***Figure 3.2***).

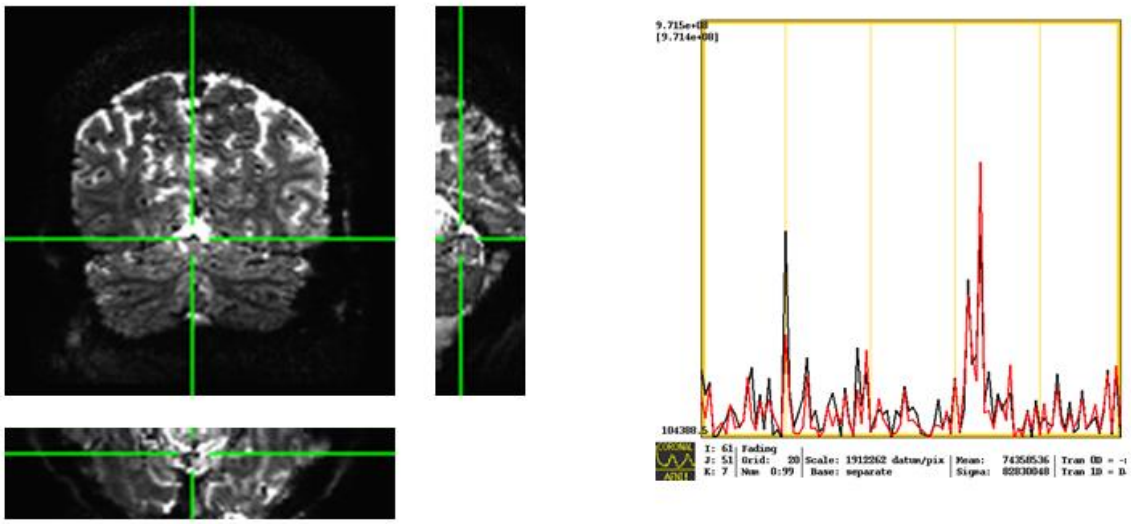

Figure 3.2: Example of the removal of physiological artefacts by using RETROICOR. In the chosen voxel, indicated by the green selection axis, the time series were corrected from their original state (in black) to physiological effect free ones (in red).

The RETROICOR corrected localizer scans were then time shifted to have all the separate slices aligned to the same temporal origin. The slice wise shifting information was retrieved from the header of the scans and done in an alternating manner starting from the first slice of the scan. All the scans, both localizer and run scans, were then promptly detrended (***Subchapter 2.3.3***) using a linear least squares detrending. Localizers were orthogonalized with respect to two two-degree polynomials (a sine and cosine wave) with a period of 200 and 195, while the runs were orthogonalized with respect to a sine and cosine wave with a period of 428 and 429 (***Figure 3.3a***). Posterior to the detrending, voxel-by-voxel arithmetic addition of the mean of the scans was added to the detrended scans (***Figure 3.3c***).

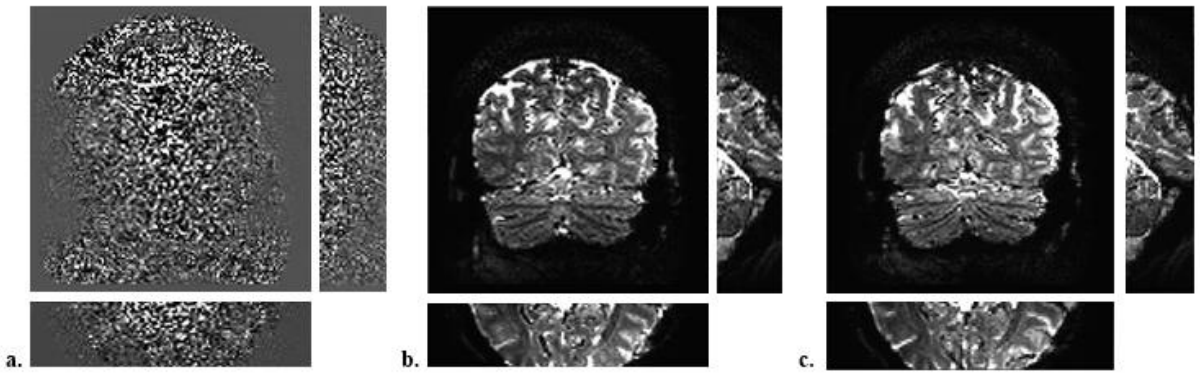

Figure 3.3 a c: Detrending of motion corrected (and time slice corrected localizer) scans by a linear least squares detrend. Top left represents a coronal slice, top right a sagittal slice and bottom an axial slice. **a.** The scan was orthogonalized with respect to two-degree polynomial functions (a sine and cosine wave) with variable period depending on the type of scan: 200 and 195, respectfully, for run scans and 428 and 429, respectfully, for Localizer scans. **b.** Before the detrending, all the mean intensity values per voxel maps of every scan were generated. **c.** The mean intensity value per voxel maps calculated in b. were combined with the detrended images to continue the pre-processing analysis.

Before being submitted as input to the GLM (***Subchapter 2.4.1***), the localizers scans went through an outlier count and censor processing step. Here, each voxel time series was detrended with Legendre polynomials of the third order prior to outlier estimation, and whenever the outlier count was of several orders of magnitude higher, they would be excluded from the GLM regression process (***Figure 3.4***).

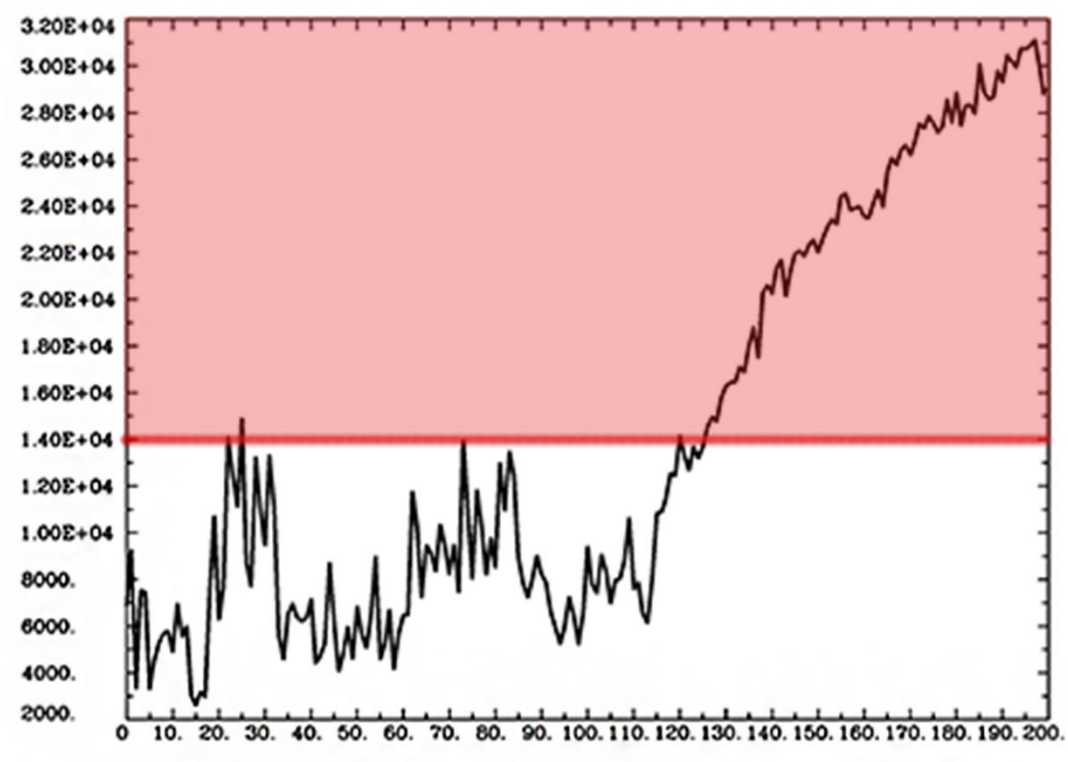

Figure 3.4: Outlier count of 200 volumes in a localizer scan. The volumes where the outlier count went beyond the threshold (in red) of several orders of magnitude (1.4×10^4) were removed from the proceeding GLM analysis.

The GLM analysis was firstly performed with the localizer scans to, in a non-regressive way, functionally localize the hMT+ in the runs. For this a GLM regression was performed with the pre-processed scans, the conditions and stimuli times of the task protocol, the motion correction parameters and the outlier censors as input (***Figure 3.5***). The convolution was performed using a standard gamma variate HRF. The t-value activation maps were visualized with a t-value threshold $P < 0.05$, Bonferroni's corrected, and clustered to form ROIs (about 20 voxels) of the hMT+ activation during the visual motor excitation task (***Figure 3.6***).

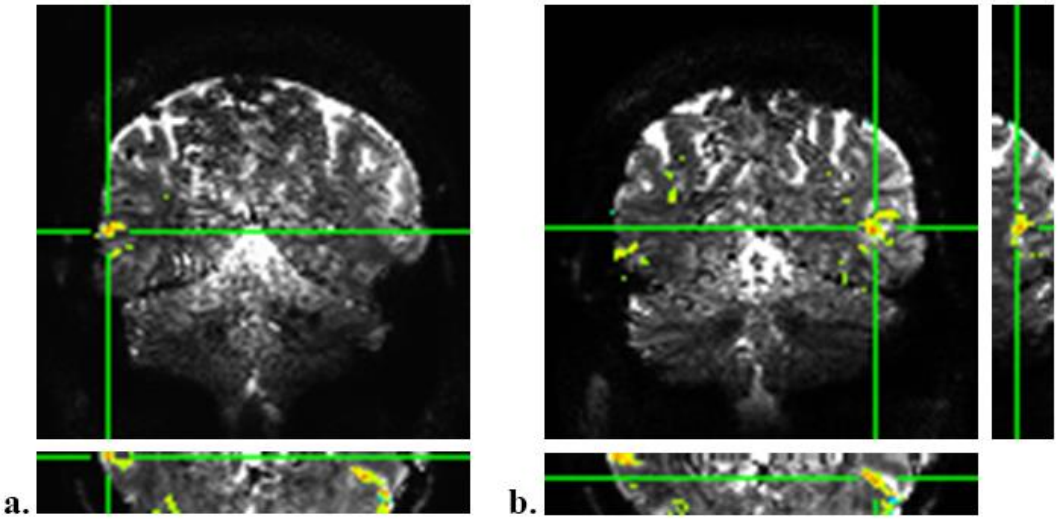

Figure 3.5 a b. Thresholded t-value activation maps of the localizer task visualized with a Bonferroni corrected threshold p-value of 0.05. These maps are being overlayed in coronal (top), axial (bottom) and sagittal (right) slices of the acquisition. **a.** Green guiding lines indicate the general position of the activation region on the left human middle temporal cortex. **b.** Green guiding lines indicate the general position of the activation region on the right human middle temporal cortex.

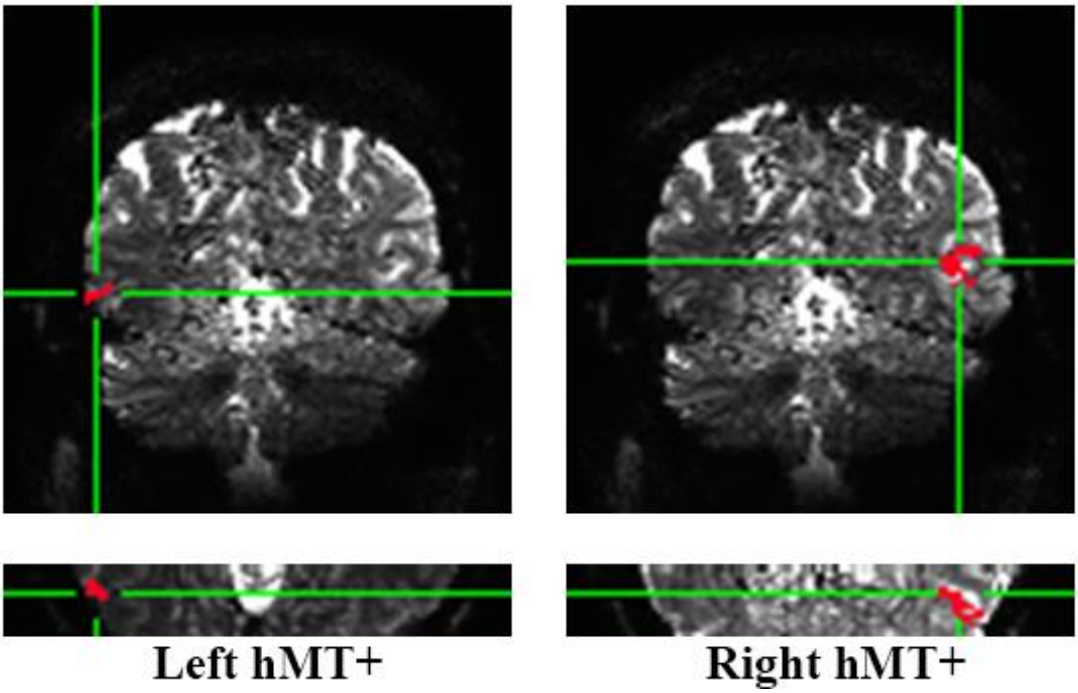

Figure 3.6: Representation of clusters of at least 20 voxels of the activation maps ($P < 0.05$, Bonferroni corrected) from the localizer scan acquisition overlayed in coronal (top) and axial (bottom) slices of the acquisition. These serve as left and right hMT+ activity ROIs for further BOLD analysis.

After the ROIs creation on the localizer image space, we aligned the localizer scans and run scans so that the study analysis could continue, now on the run image space. This was done by aligning the localizer scans to the runs, obtaining the transformation matrix required to transform one image space to the other (***Figure 3.7***). The inverted transformation matrix was then applied on the ROIs clusters obtained (***Figure 3.6***) to proceed to the HRF extraction on the several run acquisitions.

Figure 3.7: Superimposition of the acquired sagittal (left) and axial (right) run scans on the localizer scans. From this alignment the transformation matrix required to go from localizer space to run space was obtained allowing the generated ROIs to be put into run space for further analysis without any interpolation of the later images.

Having the ROIs on the run space, the HRF extraction was performed using a software package, mrTools [60], by taking a blind approach by using the event related analysis. It was decided not to have

any hypothesis on the BOLD findings within the ROI and only the stimuli starting times and the duration of the activation were used as inputs for the analysis. In this way we were able to obtain the activation maps of the runs and the highest levels of activation were remarkably all within the defined ROI (***Figure 3.8***).

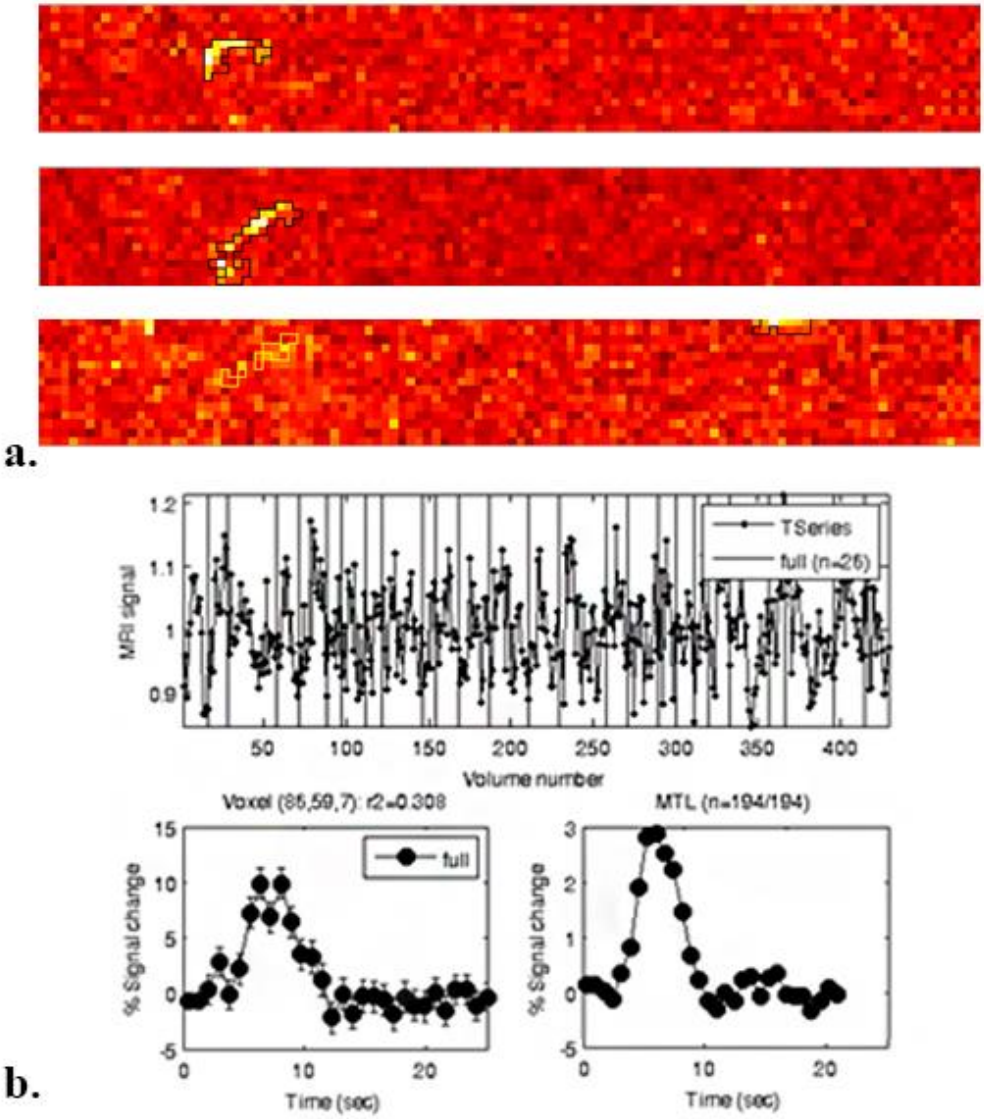

Figure 3.8 a b. HRF extraction of one of the runs acquired from one of the subjects. **a.** These three non thresholded t-value run activation maps are overlayed on axial slices of the acquisition. These demonstrate that the ROI generated by the localizer scan of this subject (*black lines*) falls over the region of the activation maps with the highest activation probability (darkest colours represent near to no activation while bright yellow and white represent highest activation). **b.** The time series of the acquisition (top), the extracted HRF curve of a selected voxel (bottom left) and the mean extracted HRF curve of the whole ROI (bottom right) were plotted for control purposes.

The HRF extraction was performed for the five different runs that were stated in the study protocol. Each run consists of an acquisition with visual stimuli of a high-contrast black-and-white dartboard of a single spatiotemporal frequency combination. The spatial frequencies were of 3 deg/cycle, 5 deg/cycle and 1 deg/cycle and while the board with spatial frequency of 3 deg/cycle was present in three different stimuli dartboards (one with 1 Hz, another with 3 Hz and the third with 5 Hz), the 5 deg/cycle and 1 deg/cycle dartboards had both a temporal frequency of 3 Hz (***Table 3.1***).

Table 3.1: Summary of the visual stimuli (high-contrast black-and-white dartboard) presented during the several runs of the study. Each run gains a designation in respect to the spatiotemporal frequency of the dartboard presented as stimuli on said run.

Spatial Frequency (deg/cycle)	Temporal Frequency (Hz)	Speed (deg/s)	Designation of the run
3	1	3	Slow Motion
1	3	3	Slow Motion
3	3	9	Intermediate Motion
1	5	15	Fast Motion
5	1	15	Fast Motion

An amplitude analysis in the ROI was conducted after the HRF reconstruction, per subject, of every run. All the mean extracted HRF of each run were compared to understand if different spatiotemporal properties on the stimuli would result in different activation of the hMT+ in both left (***Figure 3.9***) and right (***Figure 3.10***) hemispheres. All significant ($P < 0.05$, Bonferroni's corrected) voxels within the ROI created with the localizer scans were used in the HRF extraction.

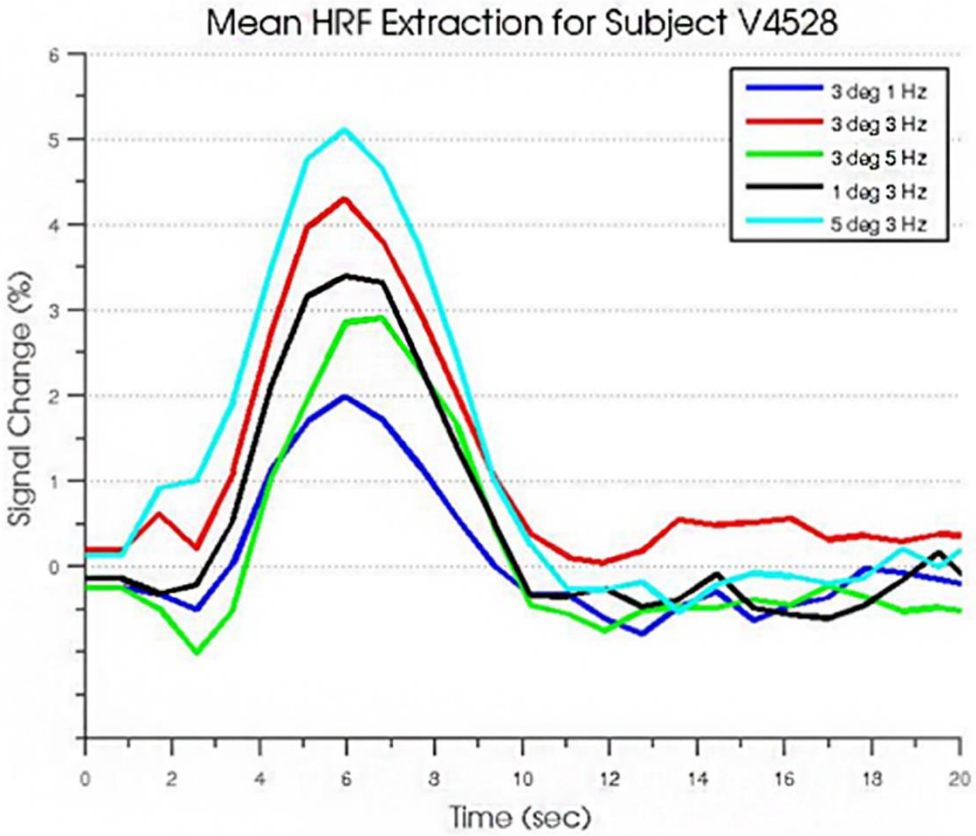

Figure 3.9: Mean HFR extraction for one of the subjects of the study (V4528) of the left hMT+. All the runs are colour coded (blue for 3 deg/cycle, 1 Hz; red for 3 deg/cycle, 3 Hz; green for 3 deg/cycle, 5 Hz; black for 1 deg/cycle, 3 Hz and light blue for 3 deg/cycle, 1 Hz) and the hemodynamic response signal change in percentage can be seen for about 20 seconds.

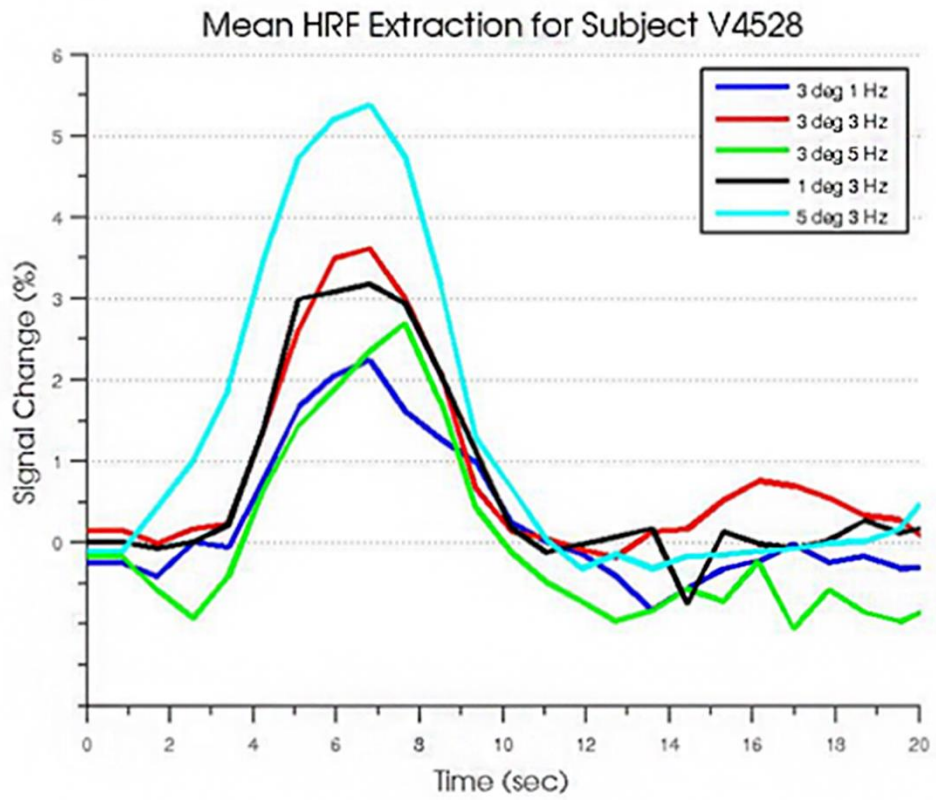

Figure 3.10: Mean HFR extraction for one of the subjects of the study (V4528) of the right hMT+. All the runs are colour coded (blue for 3 deg/cycle, 1 Hz; red for 3 deg/cycle, 3 Hz; green for 3 deg/cycle, 5 Hz; black for 1 deg/cycle, 3 Hz and light blue for 3 deg/cycle, 1 Hz) and the hemodynamic response signal change in percentage can be seen for about 20 seconds.

The maximum amplitude of the HRF responses of all the runs were plotted for both hemispheres per subject to research about the possibility of a preferred spatial or temporal frequency (***Figure 3.11***).

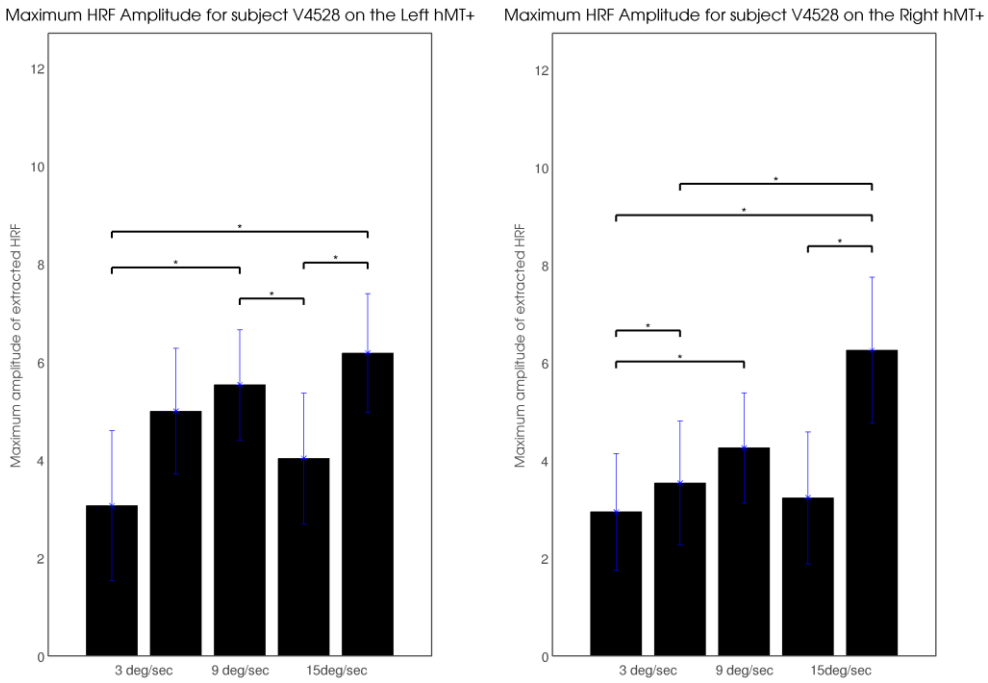

Figure 3.11: Maximum amplitude of the HRF response of both left (*to the left*) and right (*to the right*) hemispheres of one of the subjects in the study. All the differences are indicated with the curly brackets and are of at least one standard deviation.

It was then possible to assess the number of differences in the maximum amplitude of the HRF between runs that have the same speed encoded by difference combinations of spatiotemporal frequencies (3 deg/sec can be encoded by the following combinations 3 deg/cycle, 1 Hz and 1 deg/cycle, 3 Hz; and 15 deg/sec can be encoded by the following combinations 5 deg/cycle, 1 Hz and 1 deg/cycle, 5 Hz). A significant difference in the maximum amplitude of the HRF will allow us to assess the type of encoding present in the hMT+. Since five subjects were submitted to the study it is possible to calculate the differences in responses of 10 hemispheres. In the 3 deg/sec speed runs 7 of the hemispheres show significant differences in their response (***Table 3.2***), while in 15 deg/sec speed runs, only 5 of the hemispheres show significant differences (***Table 3.3***).

Table 3.2: Table with the recollection of significant differences in the maximum amplitude of the HRF responses of the runs with the dartboard speed of 3 deg/sec, generated with different spatiotemporal frequency combinations (3 deg/cycle, 1 Hz and 1 deg/cycle, 3 Hz). The differences were seen in both hemispheres with a total difference of responses for the same speed in seven of the 10 brain hemispheres present in the study.

Left hemisphere	Right hemisphere	Total
3	4	7

Table 3.3: Table with the recollection of significant differences in the maximum amplitude of the HRF responses of the runs with the dartboard speed of 15 deg/sec, generated with different spatiotemporal frequency combinations (5 deg/cycle, 1 Hz and 1 deg/cycle, 5 Hz). The differences were seen in both hemispheres with a total difference of responses for the same speed in five of the 10 brain hemispheres present in the study.

Left hemisphere	Right hemisphere	Total
3	2	5

The maximum amplitude analysis also allows to check the influence of speed in the responses to the visual stimuli. When the maximum HRF amplitude in both hemispheres for all subjects was compared no significant differences were seen (***Figure 3.12*** and ***Figure 3.12***). Also, a clear preference for a specific spatiotemporal frequency combination was also not observable.

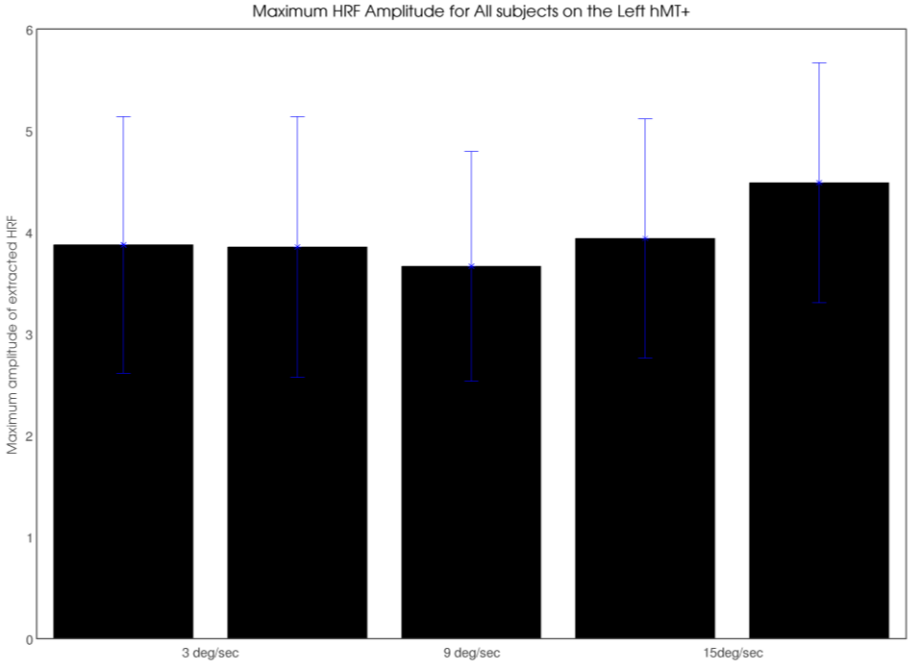

Figure 3.12: Maximum amplitude of the HRF response of the left hemisphere's hMT+ for all the subjects in the study.

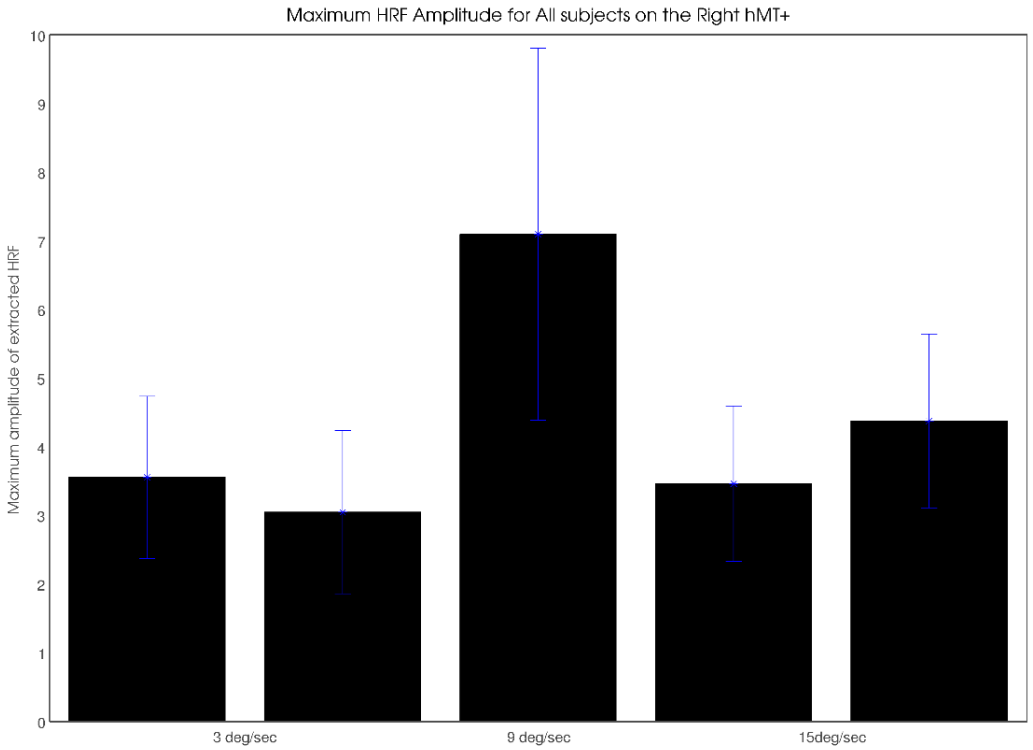

Figure 3.13: Maximum amplitude of the HRF response of the right hemisphere's hMT+ for all the subjects in the study.

To further inspect the influence of the spatial and temporal frequencies of the visual stimuli in the hMT+ a voxel-by-voxel analysis was performed. In this analysis, each voxel within the ROI was given the value of the speed (with both spatiotemporal frequencies being mentioned) in which run the voxel presented the highest maximum amplitude of its HRF response. Both hemispheres were considered and separately analysed and plotted. A count of the number of voxels with highest HRF amplitude was performed and each run was colour coded and subsequently represented on a surface representation of the subject's brain (***Figure 3.14*** and ***Figure 3.15***). The surface created had to be inflated due to the use of a surface coil [31] during the study scan acquisition which, in case, influenced the surface representation on the T_1 scan. Several surfaces with different depths were generated, but we focus on the one roughly at the middle of cortical depth.

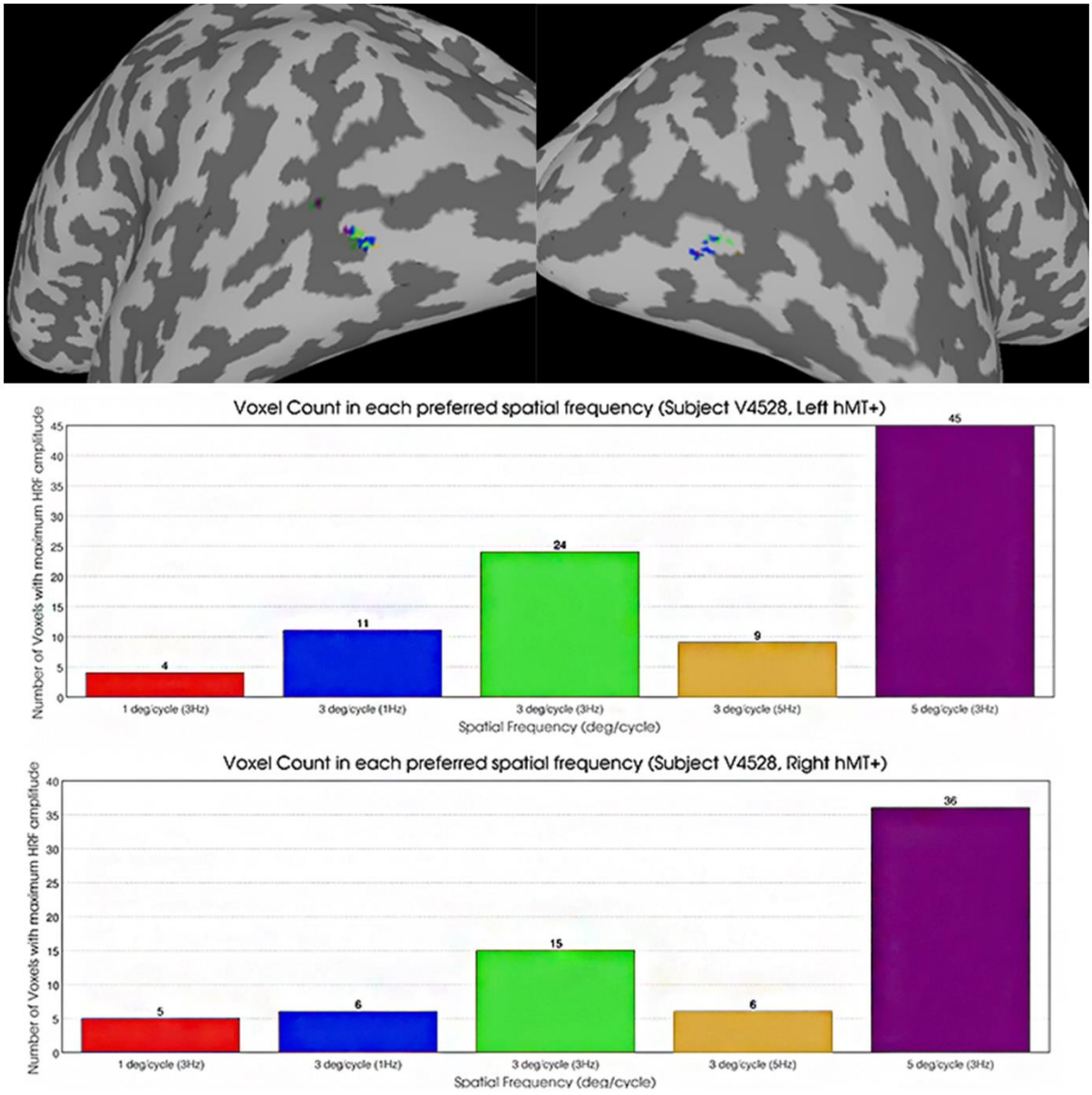

Figure 3.14: Voxel-by-voxel analysis of subject V4528. Voxels with the highest amplitude of HFR on a specific run are coloured with the colour code of said run and represented on an inflated brain surface generated by their T_1 scan (*on top*). Both left (top left) and right (top right) hemispheres are represented. A voxel count was performed to infer on the tuning of the hMT+ for speed components on both the left hemisphere (*in the middle*) and right hemisphere (*in the bottom*). Here it is also possible to see the colour code for the brain surface representation: red – 1 deg/cycle, 3Hz; blue – 3 deg/cycle, 1 Hz; green – 3 deg/cycle, 3 Hz; yellow – 3 deg/cycle, 5 Hz; purple – 5 deg/cycle, 3 Hz.

In the surface representation of the subject V4528 (***Figure 3.14***), it's possible to observe clusters of voxels that have as preferred spatiotemporal characteristics the 5 deg/cycle and 3 Hz and 3 deg/cycle and 3 Hz. These also have a very big voxel count during the analysis (36 and 15, respectfully). Altough not very well represented in the voxel count with just 6 voxels, the run with 3 deg/cycle and 1 Hz is quite evident on the surfare. This can be due to the particular selection of the layer to generate the surface that ended having all the voxels of this speed represented there. The clusterization present in this subject is a clear example on the spatial organization and spatial frequency tuning of the hMT+.

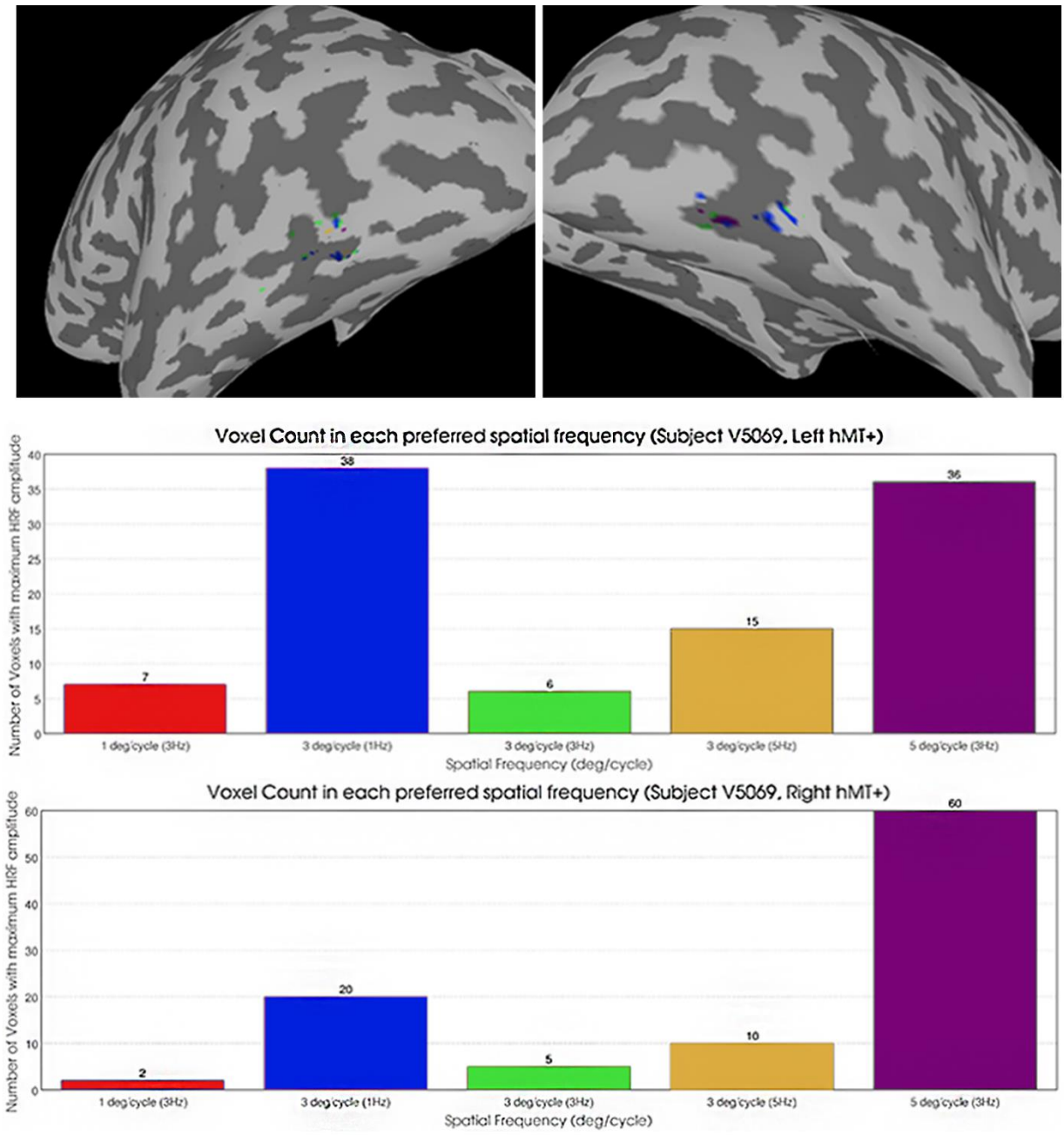

Figure 3.15: Voxel-by-voxel analysis of subject V5069. Voxels with the highest amplitude of HFR on a specific run are coloured with the colour code of said run and represented on an inflated brain surface generated by their T_1 scan (*on top*). Both left (top left) and right (top right) hemispheres are represented. A voxel count was performed to infer on the tuning of the hMT+ for speed components on both the left hemisphere (*in the middle*) and right hemisphere (*in the bottom*). Here it is also possible to see the colour code for the brain surface representation: red – 1 deg/cycle, 3Hz; blue – 3 deg/cycle, 1 Hz; green – 3 deg/cycle, 3 Hz; yellow – 3 deg/cycle, 5 Hz; purple – 5 deg/cycle, 3 Hz.

In the surface representation of the subject V5069 (***Figure 3.15***), it's possible to observe clusters of voxels with the preferred spatiotemporal characteristics of 5 deg/cycle and 3 Hz and 3 deg/cycle and 3 Hz. These also have the highest voxel count during the analysis (60 and 20, respectfully). The rest of the runs have their own representation every though is much more subtle both in the surface representation and in the voxel count. The clusterization present in this subject, just like the one in subject V4528, allows us to accept a spatial frequency tuning profile for the hMT+.

4. Discussion

The aim of this research is to further develop the knowledge of the hMT+, area of the visual cortex responsible for visual motion processing and direction and speed of motion in visual stimuli. This mechanism is invaluable for day to day tasks, but its functional organization is yet to be fully understood. Most studies suggest that the neuron population of the hMT+ falls into two paradigms: the first suggests that the MT neurons have a separate and independent tuning for spatial and/or temporal frequencies of visual stimuli; the second gives the neurons in the MT an encoding of the stimulus speed directly, rather than its separation in the two frequencies components.

To understand the mechanism of speed encoding in the hMT complex neurons, 7T fMRI data was acquired from the visual cortex of 5 healthy subjects while passively viewing a black and white dartboard moving at different spatial and temporal frequencies. The study aims to extend results obtained on single neuronal population in hMT+ [28] to the whole complex and characterize hMT+ responses spatially thanks to the high spatial resolution achieved by the 7T.

4.1. Result analysis

Five subjects that participated in the study were presented with visual motion stimuli consisting of a localizer and five different runs while lying in the 7T fMRI scan. The acquired localizers scans were pre-processed through a pipeline that include steps such as head motion detection and correction (***Figure 3.1***), removal of linear and non-linear trends in voxel time courses (***Figure 3.3***), physiological motion effect corrections (***Figure 3.2***), general despiking (***Figure 3.4***) and alignment of functional and high resolution anatomical images (***Figure 3.7***). The pre-processed localizer scans were, in conjunction with the conditions and stimuli times of the task protocol, the motion correction parameters and the outlier censors, used as input (***Figure 3.5***). The convolution was performed using a standard gamma variate HRF and the activation maps were visualized with a $P < 0.05$, Bonferroni's corrected. The purpose of this GLM with the localizer scans was to generate clustered ROIs in a non-regressive way (independent) of the hMT+ during the chosen visual motor excitation task (***Figure 3.6***).

With the ROI of the hMT+ activation a mean amplitude analysis was performed over all the runs for all the subjects. The purpose of this analysis was to observe if the mean HRF of each run presented significant differences that would allude to a preference of a spatiotemporal frequency combination over other, or even a preference for speed (which would contradict some of the studies that influenced this one). However, the mean HRF analysis of both hemispheres for every subject proved inconclusive. As observed in both ***Figure 3.9*** and ***Figure 3.10*** even though the difference in maximum amplitude is present, the extracted mean HRF of every run is not remarkably different to allows us to ascertain the encoding mechanism behind the hMT+. However, the differences noticeable on every subject's own mean HRF extracted from the runs guided the analysis into single subject single voxel maximum amplitude analysis. The voxels were put under the condition of having to be have significant activity ($P < 0.05$, Bonferroni's corrected) to be included in the HRF extraction. This would mean that, theoretically speaking, there was a change for the ROIs of each run to have variable sizes while still being used to extract HRF that would be compared. Upon closer inspection it was noticeable the difference of voxels used for the HRF extraction from subject to subject, but the number of voxels used in the analysis per subject (within the several runs) was mostly equivalent.

The maximum amplitude of the HRF responses of all the runs were plotted in a bar plot separately per hemispheres for both left and right hemisphere (***Figure 3.11***). Significant differences between activity in the different runs was found. Two main aspects were closely inspected in this

maximum amplitude comparison between runs of the subjects: whether there is significant difference in the response to runs with different spatiotemporal properties coding the same speed; or significant difference in the responses to runs of different speeds.

We assumed the hypothesis that frequency tuning was present in more than half of the hemispheres present in the analysis (10 hemispheres in total) showed significant differences in the HRF response of runs with the same speed encoded by different combinations of spatial and temporal frequencies (3 deg/sec and 15 deg/sec were both encoded in two different ways, as described in ***Table 3.1***). In the 3 deg/sec speed runs 7 of the hemispheres show significant differences in their response to runs of the same speed (***Table 3.2***), while in 15 deg/sec speed runs, 5 of the hemispheres show significant differences (***Table 3.3***). This led to the belief that the hMT+ followed a spatial and temporal frequency tuning profile.

Although this conclusion was drawn, we still cannot formulate an hypothesis on which one of the spatiotemporal frequency combinations has a preferred tuning within the hMT+ on either of the hemispheres (***Figure 3.12*** and ***Figure 3.13***). A single voxel analysis was performed with the purpose of observing the distribution of the significantly active voxels on a personalized brain surface created from its anatomical scan acquisitions, categorizing them under a run (the run which the single voxel presented the highest activation amplitude on its HRF) and identifying any clusters of activity for combinations of spatial and temporal frequencies on the generated surface of the cortex. Due to the usage of a surface coil in the data acquisition for the study, the surface used in this analysis had to be inflated to avoid bad surface representations.

The results of the surface analysis are shown in ***Figure 3.14*** and ***Figure 3.15***. In the images, these two subjects, present clear clusters of preferred spatial and temporal frequencies, specifically on the 15 deg/sec when encoded as 5 deg/cycle and 3 Hz and on the 3 deg/sec speed encoded as 3 deg/cycle and 1 Hz. In the subject V4528 there is also the presence of a cluster with preferred speed of 9 deg/sec (3 deg/cycle and 3 Hz), which reinforces the existence of areas in the hMT+ with frequency specific activity. The presence of these separable and identifiable clusters in almost all the subjects and hemispheres suggests a spatial organization within the hMT+.

The separable spatiotemporal properties seen in the hMT+ found in this study do not correspond to those of the former fMRI studies in humans focused on investigating speed responses. Chawla *et al.* [25] was able to demonstrate nonlinear correspondence between random moving dots and BOLD activity in the hMT+, while Lingnau *et al.* [27] provided evidence for speed tuning coding of BOLD responses over several temporal frequencies. Since random moving dot patterns with the same size and densities, like the ones used by Chawla *et al.*, are incapable of disentangling the spatial and temporal frequencies presented, the spatial frequency was constant in these stimuli. The speed differences reflect temporal frequency differences only. In the case of Lingnau *et al.*, the paradigm used in this case was a complex one, using drafting gratings. Due to the adaptation paradigm used to study speed tuning properties there may have been potential confounds. It has been proven that the speed tuning functions of neuronal populations tend to shift towards lower speeds with decreasing contrast [61], [62]. By using different contrasts in the adaptation block Lingnau *et al.* may have incited different populations of neurons to be tested.

The findings stated on this report are in line with the ones presented by Gaglianese, A. [29] on patient data. Here the hMT+ neuronal populations also exhibited selective preference for specific spatial frequencies. However, these preferences were linked to the electrode location within the hMT+ complex. MT neuron's spatial frequency preference correlates with their receptive field sizes and MST neurons have larger receptive fields but lower spatial frequency preferences than their counterpart [63]. This and the fact that the MT and MST were functionally subdivided into processing slow and fast speeds, respectively, [46] may explain the findings of separable spatially frequency tuned clusters identified as part of the work presented in this report.

It is important to mention that if the several neuronal populations of the hMT+ shown by the clusters on (***Figure 3.14*** and ***Figure 3.15***) act as spatial frequency filters for motion detection with slight different temporal tuning, then the model representation that states that speed depends on the spatial frequency of the stimulus [22], [23], is supported. This mechanism is inherited from the V1 cells. With this information and the findings of by Gaglianese, A. [18] the hMT+ can provide huge insight on blindsight patients (patients with damaged V1 still able to perceive motion due to MT selectivity).

4.2.Limitations and future recommendations

These findings were not void of limitations though. By analysing them and presenting solutions for the future, we are confident that more conclusive proof of the hypothesis presented in this thesis report. The low statistical power of some of the results presented may appear as a clear limitation. However, the nature of the data present in the study, its rarity due to the nature of the protocol followed, and the fact that the spatial resolution and the signal-to-noise ratio were very high, allows for conclusive single voxel and single subject analysis.

In this study, the pre-processing and analysis of the data took a lengthy amount of time when compared to other fMRI studies. This is justified by the nature of the data, although some improvements could be done in the field of the automation of some processes, in part also to avoid human error: making the outlier censor and the cardiac heartbeat frequency filtering automatically selected would improve both the speed and the accuracy of the pre-processing pipeline.

Finally, if the ROI selection is to be done with the help of an anatomic atlas and in a less restrictive way (other than a $P < 0.05$, Bonferroni's corrected), we would be sure about the inclusion of the complete hMT+ in the generated ROIs. This would increase the accuracy and completeness of our analysis over the activation clusters, allowing for further exploration and mapping of the spatial organization behind the hMT+.

5. Conclusion

The purpose of this project was to characterize the neural response to stimuli of different spatial and temporal frequencies in the human middle temporal cortex, using a non-invasive method such as 7T BOLD fMRI. We investigated whether the response preferences of the neural populations present in the hMT+ depend on the stimulus speed or the independent components that are part of it, like the spatial and temporal frequencies. An in-house build two 16-channel surface coil and a specific task based on high-resolution moving dartboards were both used to obtain the 7T fMRI BOLD signal.

The location of the activity of the hMT+ was done functionally in an independent, non-regressive way, using localizer scans. These allowed to build ROI clusters where the hMT+ was subsequently analysed.

The study presents evidence of a frequency tuning profile in the hMT+, although no significant preferred spatiotemporal frequency profile could be drawn. Additionally, clusters of activity for specific combinations of spatial and temporal frequencies were detected in the hMT+. This suggests a spatial organization within the hMT+ that requires further investigation.

Our findings have a few limitations, mostly the pre-processing time and lack of automation may lead to human errors in some of the analysis also the selection of the ROI where the analysis is done can lead to variability in the conclusions, especially regarding the clustering of spatially frequency tuned activity in the hMT+

With use of the above recommendations, we are convinced that it is possible to replicate the results of this study and improve over them, providing a non-invasive, healthy population, take on a highly studied brain region.

6. References

[1] J. M. Allman and J. H. Kaas, "Representation of the visual field in striate and adjoining cortex of the owl monkey (Aotus trivirgatus)," *Brain Res.*, vol. 35, no. 1, pp. 89–106, 1971.

[2] V. A. F. Lamme and P. R. Roelfsema, "The distinct modes of vision offered by feedforward and recurrent processing," *Trends Neurosci.*, vol. 23, no. 11, pp. 571–579, 2000.

[3] L. C. Sincich and J. C. Horton, "Independent projection streams from macaque striate cortex to the second visual area and middle temporal area.," *J. Neurosci.*, vol. 23, no. 13, pp. 5684–5692, 2003.

[4] B. A. Wandell, S. O. Dumoulin, and A. A. Brewer, "Visual field maps in human cortex," *Neuron.*, vol. 56, no. 2, pp. 366–383, 2007.

[5] D. C. Bradley and M. S. Goyal, "Velocity computation in the primate visual system," *Nat. Rev. Neurosci.*, vol. 9, no. 9, pp. 686–695, 2008.

[6] L. G. Ungerleider and R. Desimone, "Cortical connections of visual area MT in the macaque," *J. Comp. Neurol.*, vol. 248, no. 2, pp. 190–222, 1986.

[7] R. Dubner and S. M. Zeki, "Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey," *Brain Res.*, vol. 35, no. 2, pp. 528–532, 1971.

[8] J. H. Maunsell and D. C. van Essen, "The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey.," *J. Neurosci.*, vol. 3, no. 12, pp. 2563–2586, 1983.

[9] R. R, "Coding of Visual Stimulus Velocity in Area MT of the Macaque.pdf." Princeton, p. 14, 1987.

[10] K. Tanaka, "Analysis of Motion of the Visual Field by Direction , Expansion / Contraction , and Rotation Cells Clustered in the Dorsal Part of the Medial Superior Temporal Area of the Macaque Monkey," *Journalofneurophysiology*, vol. 62, no. 3, 1989.

[11] S. O. Dumoulin, "A New Anatomical Landmark for Reliable Identification of Human Area V5/MT: a Quantitative Analysis of Sulcal Patterning," *Cereb. Cortex*, vol. 10, no. 5, pp. 454–463, 2000.

[12] R. T. Born and D. C. Bradley, "Structure and Function of Visual Area Mt," *Annu. Rev. Neurosci.*, vol. 28, no. 1, pp. 157–189, 2005.

[13] K. Amano, B. A. Wandell, and S. O. Dumoulin, "Visual Field Maps, Population Receptive Field Sizes, and Visual Field Coverage in the Human MT+ Complex," *J. Neurophysiol.*, vol. 102, no. 5, pp. 2704–2718, 2009.

[14] S. M. Zeki, "Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey," *J. Physiol.*, vol. 236, no. 3, pp. 549–573, 1974.

[15] S. Zeki, "Thirty years of a very special visual area, Area V5," *J. Physiol.*, vol. 557, no. 1, pp. 1–2, 2004.

[16] R. B. Tootell, J. B. Reppas, K. K. Kwong, R. Malach, R. T. Born, T. J. Brady, B. R. Rosen, and J. W. Belliveau, "Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging.," *J. Neurosci.*, vol. 15, no. 4, pp. 3215–30, 1995.

[17] T. Albright, R. Desimone, and C. G. Gross, "Columnar Organization of Directionally Selective.pdf," *J. Neurophysiol.*, vol. 51, no. 1, 1984.

[18] A. Gaglianese, M. Costagli, G. Bernardi, E. Ricciardi, and P. Pietrini, "Evidence of a direct influence between the thalamus and hMT+ independent of V1 in the human brain as measured by fMRI," *Neuroimage*, vol. 60, no. 2, pp. 1440–1447, 2012.

[19] J. A. Perrone and A. Thiele, "Speed skills: measuring the visual speed analyzing properties of primate MT neurons.," *Nat. Neurosci.*, vol. 4, no. 5, pp. 526–532, 2001.

[20] J. Duijnhouwer, A. J. Noest, M. J. M. Lankheet, A. V van den Berg, and R. J. a van Wezel, "Speed and direction response profiles of neurons in macaque MT and MST show modest constraint line tuning.," *Front. Behav. Neurosci.*, vol. 7, no. April, p. 22, 2013.

[21] L. L. Lui, J. A. Bourne, and M. G. P. Rosa, "Spatial and temporal frequency selectivity of neurons in the middle temporal visual area of new world monkeys (Callithrix jacchus)," *Eur. J. Neurosci.*, vol. 25, no. 6, pp. 1780–1792, 2007.

[22] N. J. Priebe, C. R. Cassanello, and S. G. Lisberger, "The neural representation of speed in macaque area MT/V5.," *J. Neurosci.*, vol. 23, no. 13, pp. 5650–5661, 2003.

[23] N. J. Priebe, S. Lisberger, and A. Movshon, "Tuning for Spatiotemporal Frequency and Speed in Directionally Selective Neurons of Macaque Striate Cortex," *2Journal Neurosci.*, vol. 26, no. 11, pp. 2941–2950, 2006.

[24] A. C. Huk, R. F. Dougherty, and D. J. Heeger, "Retinotopy and functional subdivision of human areas MT and MST.," *J. Neurosci.*, vol. 22, no. 16, pp. 7195–7205, 2002.

[25] D. Chawla, "Speed-Dependent Motion-Sensitive Responses in V5: An fMRI Study," *Neuroimage*, pp. 86–96, 1998.

[26] O. Kawakami, Y. Kaneoke, K. Maruyama, R. Kakigi, T. Okada, N. Sadato, and Y. Yonekura, "Visual detection of motion speed in humans: Spatiotemporal analysis by fMRI and MEG," *Hum. Brain Mapp.*, vol. 16, no. 2, pp. 104–118, 2002.

[27] A. Lingnau, M. B. Wall, and A. T. Smith, "Speed encoding in human visual cortex revealed by fMRI adaptation Hiroshi Ashida," *J. Vis.*, vol. 9, no. 13, pp. 1–14, 2009.

[28] L. Wang, Y. Kaneoke, and R. Kakigi, "Spatiotemporal separability in the human cortical response to visual motion speed: a magnetoencephalography study," *Neurosci. Res.*, vol. 47, pp. 109–116, 2003.

[29] A. Gaglianese, B. M. Harvey, M. J. Vansteensel, S. O. Dumoulin, N. F. Ramsey, and N. Petridou, "Separate spatial and temporal frequency tuning to visual motion in human MT+ measured with ECoG," *Hum. Brain Mapp.*, vol. 38, no. 1, pp. 293–307, 2017.

[30] J. C. W. Siero, D. Hermes, H. Hoogduin, P. R. Luijten, N. F. Ramsey, and N. Petridou, "BOLD matches neuronal activity at the mm scale: A combined 7T fMRI and ECoG study in human sensorimotor cortex," *Neuroimage*, vol. 101, pp. 177–184, 2014.

[31] N. Petridou, M. Italiaander, B. L. van de Bank, J. C. W. Siero, P. R. Luijten, and D. W. J. Klomp, "Pushing the limits of high-resolution functional MRI using a simple high-density multi-element coil design," *NMR Biomed.*, vol. 26, no. 1, pp. 65–73, 2012.

[32] A. C. Huk and D. J. Heeger, "Pattern-motion responses in human visual cortex," *Nat. Neurosci.*, vol. 5, no. 1, pp. 72–75, 2002.

[33] B. M. Dale, M. A. Brown, and R. C. Semelka, *MRI: Basic Principles and Applications*, 5th ed. Wiley-Blackwel, 2015.

[34] M. Nessaiver, *All you really need to know about MRI physics.pdf*, no. August. University of Maryland, Baltimore, 2004.

[35] H. H. Schild, *MRI made easy (...well almost)*, 1st ed. Nationales Druckhaus Berlin, 1990.

[36] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, 2nd ed. Sinauer Associates, 2008.

[37] S. Ogawa, "Brain magnetic resonance imaging with contrast dependent blood oxygenation," *Proc Natl Acad Sci U S A.*, pp. 9868–9872, 1990.

[38] E. Haacke, *Magnetic Resonance Imaging: Physical Principles and Sequence Design*, 1st ed. New York: Wiley, 1999.

[39] M. T. Vlaardingerbroek, J. A. Boer, T. A. den Boen, and L. Dyke, *Magnetic Resonance Imaging, Theory and Practice (2nd ed.)*, vol. 28, no. 7. 2003.

[40] R. Goebel, "Localization of Brain Activity using Functional Magnetic Resonance Imaging," in *Clinical Functional MRI*, pp. 9–51.

[41] D. Kim, T. Duong, and S. Kim, "High-resolution mapping of iso-orientation columns by fMRI," *Nat. Neurosci.*, vol. 3, no. 164, p. 169, 2000.

[42] "Echo-Planar Imaging (EPI)." [Online]. Available: http://mriquestions.com/echo-planar-imaging.html%0A. [Accessed: 23-Sep-2018].

[43] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, 1999.

[44] Lee Ann Ramington, *Clinical Anatomy and Physiology of The Visual System*, 3rd ed., no. 1. Butterworth-Heinemann, 2014.

[45] E. R. Kandel, J. H. Schwartz, T. M. Jessel, S. A. Siegelbaum, A. J. Hudspeth, and S. Mack, *Principles of Neural Science*, Fifth Edit. 2013.

[46] A. Gaglianese, M. Costagli, K. Ueno, E. Ricciardi, G. Bernardi, P. Pietrini, and K. Cheng, "The direct, not V1-mediated, functional influence between the thalamus and middle temporal complex in the human brain is modulated by the speed of visual motion," *Neuroscience*, vol. 284, pp. 833–844, 2015.

[47] S. Kastner and M. A. Pinsk, "Visual attention as a multilevel selection process," *Cogn Affect Behav Neurosci*, vol. 4, no. 4, pp. 483–500, 2004.

[48] A. Mahbod, C. Wang, M. Chowdhury, and Ö. Smedby, "Brain Segmentation Using Artificial Neural Networks with Shape Context," *Under Prep.*, pp. 1–29, 2016.

[49] Y. Kaneoke, "Magnetoencephalography: In search of neuronal processes for visual motion information," *Prog. Neurobiol.*, pp. 219–240, 2006.

[50] P. Kellman, F. H. Epstein, and E. R. McVeigh, "Adaptive sensitivity encoding incorporating temporal filtering (TSENSE)," *Magn. Reson. Med.*, vol. 45, no. 5, pp. 846–852, 2001.

[51] X. Hu, T. Le, T. Parrish, and P. Erhard, "Retrospective estimation and correction of physiological fluctuation in functional MRI.," *Magn. Reson. Med.*, vol. 34, no. 2, pp. 201–212, 1995.

[52] R. W. Cox, "AFNI: software for analysis and visualization of functional magnetic resonance neurouimages," *Comput Biomed Res.*, vol. 29, no. 3, pp. 162–173, 1996.

[53] G. H. Glover, T. Q. Li, and D. Ress, "Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR," *Magn. Reson. Med.*, vol. 44, no. 1, pp. 162–167, 2000.

[54] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Detection and detrending in fMRI data analysis," *Neuroimage*, vol. 22, no. 2, pp. 645–655, 2004.

[55] M. Dagli, J. Ingeholm, and J. Haxby, "Localization of cardiac-in- duced signal change in

fMRI," *Neuroimage*, vol. 9, pp. 407–415, 1999.

[56] G. Kru and G. H. Glover, "INVESTIGATOR Physiological Noise in Oxygenation-Sensitive Magnetic," *Young*, vol. 2001, no. 2001, pp. 631–637, 2001.

[57] T. E. Lund, K. H. Madsen, K. Sidaros, W. L. Luo, and T. E. Nichols, "Non-white noise in fMRI: Does modelling have an impact?," *Neuroimage*, vol. 29, no. 1, pp. 54–66, 2006.

[58] T. B. Jones, P. A. Bandettini, and R. M. Birn, "Integration of motion correction and physiological noise regression in fMRI," *Neuroimage*, vol. 42, no. 2, pp. 582–590, 2008.

[59] M. Costagli, K. Ueno, P. Sun, J. L. Gardner, X. Wan, E. Ricciardi, P. Pietrini, K. Tanaka, and K. Cheng, "Functional signalers of changes in visual stimuli: Cortical responses to increments and decrements in motion coherence," *Cereb. Cortex*, vol. 24, no. 1, pp. 110–118, 2014.

[60] J. L. Gardner, E. P. Merriam, D. Schluppeck, J. Besle, and D. J. Heeger, "mrTools: Analysis and visualization package for functional magnetic resonance imaging data." 2018.

[61] B. Vintch and J. L. Gardner, "Cortical Correlates of Human Motion Perception Biases," *J. Neurosci.*, vol. 34, no. 7, pp. 2592–2604, 2014.

[62] B. Krekelberg, G. M. Boynton, and R. J. A. van Wezel, "Adaptation: from single cells to BOLD signals," *Trends Neurosci.*, vol. 29, no. 5, pp. 250–256, 2006.

[63] K. Miura, N. Inaba, Y. Aoki, and K. Kawano, "Difference in visual motion representation between cortical areas MT and MST during ocular following responses.," *J. Neurosci.*, vol. 34, no. 6, pp. 2160–8, 2014.