Functional MRI of the visual cortex' response to moving random dots

Qianru Rao

Integrated Program in Neuroscience

McGill University, Montreal

August 2023

A thesis submitted to McGill University in partial fulfillment of

the requirements of the degree of Master of Science

Table of Contents

Acknowledgments i

Contributions of Authors ii

Abstract iii

Résumé v

List of figures and tables viii

List of abbreviations x

1 Introduction 1

2 Background 4

2.1 Visual motion perception 4
2.2 Visual motion perception in the dorsal visual pathway 7
2.3 High-resolution functional magnetic resonance imaging 13
2.4 Multivariate pattern analysis 17
2.4.1 Encoding and decoding of neural activity patterns 17
2.4.2 MVPA 17
2.4.3 SVM 21

3 Methods 22

3.1 Experiment design 22
3.1.1 fMRI experiment at 3T 23
3.1.2 fMRI experiment at 7T 30
3.2 Data analysis 34
3.2.1 Uni-variate analysis 34
3.2.2 Multivariate analysis 40

4 Results of uni-variate analysis 42

4.1 Results of the experiment at 3T 42
4.1.1 Retinotopic mapping of V1 in the lower right quadrant 42
4.1.2 MT+ functional localizer 43
4.1.3 Responses of MT+ and V1 to moving random dots 44
4.2 Results of the experiment at 7T 54
4.2.1 Retinotopic mapping of V1 in the lower right and lower left quadrant 54
4.2.2 MT+ functional localizer 56
4.2.3 Responses of MT+ and V1 to moving random dots 56
4.3 Discussion 66

5 Results of multivariate analysis 69

5.1 Results of the experiment at 3T 69
5.1.1 Decoding of coherence in MT+ and V1 69
5.1.2 Decoding of speed in MT+ and V1 73
5.2 Results of the experiment at 7T 77

5.2.1 Decoding of coherence in MT+ and V1 77
5.2.2 Decoding of speed in MT+ and V1 81
5.3 Discussion 85
6 Conclusions 90
7 Limitations 93
References 95

Acknowledgments

First, I would like to give my heartfelt thanks to everyone who has ever helped me with this study and thesis.

My sincere gratitude goes firstly to my both supervisors, Dr. Amir Shmuel and my co-supervisor, Prof. Junjun Zhang, who supported my study at McGill University and the University of Electronic Science and Technology of China. I am grateful for their guidance, teaching, and help in my master's career, especially in the experimental design and data acquisition in MNI and UESTC. I would also like to express my appreciation to Asa Borzabadi Farahani in Dr. Shmuel's lab for full assistance in data acquisition and suggestions in the data analysis of this study. Besides, I would thank my Advisory Committee, Dr. Janine Mendola and Dr. Aarlenne Khan, for devoting time to reading my thesis proposal and providing many useful suggestions for this study.

In addition, I would like to express my warm thanks to Dr. Edward Ruthazer and Ms. Mengting Hu who offered a wealth of help and efforts in the dual master program so that I could have wonderful experiences in IPN.

Finally, my heartfelt gratitude goes to my family and my friends who gave me much encouragement and support. I also wish to extend my thanks to all professors who read and give me professional and sincere advise.

Contributions of Authors

Qianru Rao completed all chapters of this thesis with the guidance and assistance of her supervisor Dr. Amir Shmuel and her co-supervisor Prof. Junjun Zhang.

The experiments mentioned in this thesis was designed by Qianru Rao, including generating visual stimulus materials and paradigm based on the codes and advisement provided by Dr. Shmuel and modifying the codes for both experiments in MNI 7T MRI center and UESTC 3T MRI center with the help of Prof. Zhang. The recruitment of subjects, data acquisition, and data analysis of the 3T fMRI experiment were completed by Qianru Rao and that of the 7T fMRI experiment were completed by Qianru Rao with the help of Asa Borzabadi Farahani in Dr. Shmuel's lab. All data collected for this thesis were financially supported by Dr. Shmuel (7T MRI data acquisition in MNI) and Prof. Zhang (3T MRI data acquisition in UESTC).

Abstract

Visual motion perception is one of the critical functions of the visual system. Direction and speed are important features of visual motion stimuli, and the processing of these features mainly occurs in the dorsal visual pathway of the visual cortex. The middle temporal complex (MT+) has sensitive direction-selectivity; it plays a key role in visual motion perception. The primary visual cortex (V1) has been widely studied. It exhibits different responses to visual motion stimuli relative to area MT+'s responses.

The majority of previous studies found that V1 responds with higher amplitude to incoherent motion than to coherent motion. As of yet, there is no unified conclusion on how MT+ responds to these two motion patterns. Most studies suggest that MT+ is more sensitive to coherent motion, but some studies have found that human MT+ has a higher response amplitude to incoherent motion, or it shows no difference in response to these stimuli. In addition, V1 is thought to be more sensitive to slow motion, while MT+ shows higher responses as speed increases. These studies have used uni-variate analysis methods and have not clearly explained how visual motion information is represented in the dorsal visual pathway. However, based on activity patterns across voxels, studies that applied multivariate pattern analysis have revealed that object information is represented in the ventral visual pathway. In contrast, it is still unknown how motion information is represented in the dorsal visual pathway. Therefore, multivariate pattern analysis based on support vector machine (SVM) classification was applied in this study with traditional uni-variate analysis as the supplement to investigate how MT+ and V1 encode visual motion coherence and motion speed.

Random dot kinematograms (RDKs) have been widely used in studies of visual motion perception as they do not have specific information on location. Consequently, we used RDKs as stimuli and conducted functional MRI (fMRI) experiments at 3T and 7T to acquire brain imaging data from healthy human subjects with different spatial resolutions. fMRI data at both 3T and 7T were analyzed by SVM decoding and uni-variate analysis.

Our first finding is that SVM classification results at both 3T and 7T showed that MT+ and V1 can discriminate between coherent and incoherent random dot motion. The uni-variate analysis results supported the view that MT+ and V1 show a higher response to coherent motion and incoherent motion, respectively. Second, speed had a slight effect on the encoding of visual motion coherence in MT+ and V1. Based on the results of the MVPA, MT+ distinguished coherent and incoherent motion best when random dots moved at 10 deg/s. According to the uni-variate analysis V1 was most likely to show a difference in responses to coherent and incoherent motion when dots moved at 2.5 deg/s. The results obtained at 3T and 7T were consistent with regard to the performance of MT+ and V1 in encoding visual motion coherence. In contrast, the capacity to decode motion speed was only observed using 7T data. This suggests that 7T fMRI could improve decoding performance when studying visual motion perception based on SVM classification. Finally, the results from both uni-variate and multivariate analyses showed that V1 did not exhibit hemispheric bias in encoding visual motion coherence and speed, while MT+ showed left-sided dominance in most uni-variate results and right-sided dominance in some SVM classification results, suggesting that MT+ may have some hemispheric bias when encoding visual motion coherence and speed.

In summary, this study investigates the encoding of visual motion coherence and speed in MT+ and V1 using traditional uni-variate analysis. More importantly, it is the first study to investigate this issue from a multivariate perspective by providing evidence for the representation of visual motion information in the dorsal visual pathway based on multi-voxel activity patterns. By comparing the SVM classification using 3T and 7T fMRI data, we also demonstrate the advantages of high-resolution fMRI for studying visual motion information encoding.

Keywords: Motion Coherence, Motion Speed, MT+, fMRI, Multivariate Pattern Analysis

Résumé

La perception du mouvement visuel est l'une des fonctions critiques du système visuel. La direction et la vitesse sont des caractéristiques importantes des stimuli visuels de mouvement, et le traitement de ces caractéristiques se produit principalement dans la voie visuelle dorsale du cortex visuel. Le complexe temporel moyen (MT+) a une sélectivité de direction sensible; il joue un rôle clé dans la perception du mouvement visuel. Le cortex visuel primaire (V1) a été largement étudié. Il présente différentes réponses aux stimuli de mouvement visuels par rapport aux réponses de la zone MT +.

La majorité des études précédentes ont révélé que V1 répond avec une amplitude plus élevée au mouvement incohérent qu'au mouvement cohérent. Pour l'instant, il n'y a pas de conclusion unifiée sur la façon dont MT + répond à ces deux modèles de mouvement. La plupart des études suggèrent que mt + est plus sensible au mouvement cohérent, mais certaines études ont constaté que MT + humain a une amplitude de réponse plus élevée au mouvement incohérent, ou il ne montre aucune différence dans la réponse à ces stimuli. En outre, V1 est considéré comme plus sensible au ralenti, tandis que MT + montre des réponses plus élevées à mesure que la vitesse augmente. Ces études ont utilisé des méthodes d'analyse univariées et n'ont pas clairement expliqué comment l'information visuelle sur le mouvement est représentée dans la voie visuelle dorsale. Cependant, sur la base des modèles d'activité à travers des voxels, les études qui ont appliqué l'analyse de modèle multivariée ont indiqué que l'information d'objet est représentée dans la voie visuelle ventrale. En revanche, on ne sait toujours pas comment les informations de mouvement sont représentées dans la voie visuelle dorsale. Par conséquent, l'analyse de modèle multivariée basée sur la classification de la machine à vecteurs de support (SVM) a été appliquée dans cette étude avec l'analyse univariée traditionnelle comme supplément pour étudier comment MT + et V1 codent la cohérence visuelle de mouvement et la vitesse de mouvement.

Les kinématogrammes à points aléatoires (RDKs) ont été largement utilisés dans les études de la perception visuelle du mouvement car ils n ' ont pas d ' informations spécifiques sur l ' emplacement. Par conséquent, nous avons employé des RDKs comme stimulus et avons mené des expériences fonctionnelles d ' IRM (IRMf) à 3T et 7T pour acquérir des données d ' imagerie cérébrale de sujets humains en bonne santé avec différentes résolutions spatiales. Les données d ' IRMf à 3T et 7T ont été analysées par décodage de SVM et analyse univariée.

Notre première conclusion est que les résultats de classification SVM à 3T et 7T ont montré que MT+ et V1 peuvent distinguer entre le mouvement aléatoire cohérent et incohérent de point. Les résultats de l ' analyse univariée ont soutenu l ' opinion selon laquelle MT + et V1 montrent une réponse plus élevée au mouvement cohérent et au mouvement incohérent, respectivement. Deuxièmement, la vitesse a eu un léger effet sur le codage de la cohérence visuelle du mouvement dans MT + et V1. Sur la base des résultats de l ' APMV, MT+ a mieux distingué le mouvement cohérent et incohérent lorsque des points aléatoires se déplaçaient à 10 deg/s. Selon l'analyse univariée, V1 était le plus susceptible de montrer une différence dans les réponses à un mouvement cohérent et incohérent lorsque les points se déplaçaient à 2,5 deg/s. Les résultats obtenus à 3T et 7T étaient compatibles en ce qui concerne la représentation de MT+ et de V1 encodant la cohérence visuelle de mouvement. En revanche, la capacité de décoder la vitesse de mouvement n ' a été observée qu ' à l ' aide de données 7T.Cela suggère que l ' IRMf 7T pourrait améliorer les performances de décodage lors de l'étude de la perception visuelle du mouvement basée sur la classification SVM. Enfin, les résultats des analyses univariées et multivariées ont montré que V1 ne présentait pas de biais hémisphérique dans le codage de la cohérence et de la vitesse visuelles du mouvement, tandis que MT + montrait une dominance gauche dans la plupart des résultats univariés et une dominance du côté droit dans certains résultats de classification SVM, ce qui suggère que MT + peut avoir un certain biais hémisphérique lors du codage de la cohérence et de la vitesse du mouvement visuel.

En résumé, cette étude étudie l'encodage de la cohérence et de la vitesse de mouvement visuel dans MT + et V1 en utilisant l'analyse univariée traditionnelle. Plus important encore, c'est la première étude à étudier cette question d'un point de vue multivarié en fournissant des preuves pour la représentation de l'information visuelle de mouvement dans la voie visuelle dorsale basée sur des modèles d'activité de multi-voxel. En comparant la classification SVM à l'aide de données IRMf 3T et 7T, nous démontrons également les avantages de l'IRMf haute résolution pour l'étude du codage visuel des informations de mouvement.

Mots-clés: Cohérence de mouvement, Vitesse de mouvement, MT +, IRMf, Analyse de motifs multivariés

List of figures and tables

Figure 1.1 Research program. 3
Figure 2.1 Two visual pathways. 9
Figure 2.2 SVM classification 22
Figure 3.1 Stimulus of checkerboard for retinotopic mapping in the 3T Experiment 26
Figure 3.2 Stimuli and paradigm in 3T experiment. 26
Figure 3.3 Stimuli and paradigm in the 7T experiment. 33
Table 3.1 Number of voxels of ROIs (3T, N=12). 38
Table 3.2 Number of voxels of ROIs (7T, N=5). 38
Figure 4.1 Group-level results of retinotopic mapping of V1 in the lower right quadrant (3T). 43
Figure 4.2 Group-level results of MT+ functional localizer by the contrast of 'move-static' (3T). 44
Figure 4.3 Group-level responses to random moving dots(3T). 45
Figure 4.4 Responses to coherent and incoherent random dot motion in MT+ and V1(3T). 48
Figure 4.5 Responses to coherent and incoherent random dot motion in left or right MT+(3T). 48
Figure 4.6 Responses to coherent and incoherent random dot motion in left or right V1 (3T). 49
Figure 4.7 Responses to random dot motion at 5.0 and 10.0 deg/s in MT+ and V1(3T). 49
Figure 4.8 Responses to random dot motion at 5.0 and 10.0 deg/s in left or right MT+(3T). 51
Figure 4.9 Responses to random dot motion at 5.0 and 10.0 deg/s in left or right V1 (3T). 52
Figure 4.10 Interaction of coherence and speed in MT+ (3T). 52
Figure 4.11 Interaction of coherence and speed in V1 (3T). 53
Figure 4.12 Retinotopic mapping of V1 in the lower right (top row) and lower left (bottom row) quadrant at group level (7T). 55
Figure 4.13 MT+ functional localizer at group level (7T). 55
Figure 4.14 Responses to random moving dots (7T). 57
Figure 4.15 Responses to coherent and incoherent random dot motion in MT+ and V1(7T). 60
Figure 4.16 Responses to coherent and incoherent random dot motion in left and right MT+(7T). 60
Figure 4.17 Responses to coherent and incoherent random dot motion in left and right V1 (7T). 61
Figure 4.18 Responses to random moving dots at different speeds (7T). 61
Figure 4.19 Responses to random moving dots at different speeds in left or right MT+ (7T). 62
Figure 4.20 Responses to random moving dots at different speeds in left or right V1 (7T). 62
Figure 4.21 Interaction of coherence and speed in MT+ (7T). 62
Figure 4.22 Interaction of coherence and speed in V1 (7T). 63
Figure 5.1 Decoding performance of coherent and incoherent motion (3T). 72
Figure 5.2 Decoding performance of coherent and incoherent motion in left or right MT+ (3T). 72
Figure 5.3 Decoding performance of coherent and incoherent motion in left or right V1 (3T). 73
Figure 5.4 Decoding performance of speed of random moving dots (3T). 75
Figure 5.5 Decoding performance of speed of random moving dots in left or right MT+ (3T). 76
Figure 5.6 Decoding performance of speed of random moving dots in left or right V1 (3T). 76
Figure 5.7 Decoding performance of coherent and incoherent motion (7T). 80
Figure 5.8 Decoding performance of coherent and incoherent motion in left or right MT+ (7T). 80
Figure 5.9 Decoding performance of coherent and incoherent motion in left or right V1 (7T). 81

Figure 5.10 Decoding performance of speed of random moving dots (7T). 84
Figure 5.11 Decoding performance of speed of random moving dots in left or right MT+ (7T). 84
Figure 5.12 Decoding performance of speed of random moving dots in left or right V1 (7T). 85

List of abbreviations

RDKs: Random Dot Kinemotograms

V1: Visual Area 1 (primary visual cortex)

MT: Middle Temporal Cortex (also called V5)

MST: Media Superior Temporal Cortex

MT+: Middle Temporal Complex

SVM: Support Vector Machine

MVPA: Multivariate Pattern Analysis

T: Tesla

coh: coherent

inc: incoherent

lh: left hemisphere

rh: right hemisphere

lr: left and right hemisphere

1 Introduction

MT+ (Middle Temporal cortex) is a well-recognized brain area highly sensitive to visual motion information. Visual motion perception is essential to daily life and the recognition of motion coherence enables the ability to differentiate coherent motion where the movement towards the same direction from incoherent motion. The majority of previous studies have agreed that the motion coherence in global motion perception originates from the direction selectivity of MT+. However, there are different views on whether the response of MT+ to coherent and incoherent motion is similar. Additionally, speed may also be one of the factors affecting motion coherence perception, and the relationship between speed and motion coherence is still unclear. Therefore, this study focuses on the responses of two brain areas in the dorsal visual pathway, namely the motion-sensitive MT+ (including MT and MST) and visual area V1, in the context of motion coherence and motion speed.

Motion perception is essential in daily life. Area MT+, which is closely related to this cognitive function, has received a lot of attention. Evidence from age-related functional decline, residual motion perception in blind patients (Baker et al., 1991), or auditory motion perception (Jiang et al., 2014) all suggests that MT+ plays a crucial role in visual and other types of motion processes. In this study, based on fMRI experiments and multivariate pattern analysis methods, we study how MT+ in the healthy human brain encodes visual motion information in the dorsal visual pathway. This will support our further understanding of the neural mechanisms of motion processing in the visual cortex.

Currently, there are no reports on human MT+ and V1's capacity to decode visual motion coherence and motion speed information based on multivariate analysis. Previous related studies have used traditional uni-variate analysis methods. Therefore, this study combines uni-variate and multivariate analysis methods to investigate how healthy human MT+ and V1 encode visual motion coherence and speed information and explores

the encoding method of visual motion information in MT+ and V1 from the perspective of multivariate pattern analysis. At the same time, this study combines fMRI data with two different resolutions obtained at 3T and 7T, to explore whether high-resolution fMRI data can enhance the performance of decoding visual motion information (see **Figure 1.1**).

The current popular neural networks such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have computational frameworks that are more geared towards feature recognition mechanisms for static targets in the ventral visual pathway. However, motion processing mainly focuses on the dorsal visual pathway, and the ventral visual pathway only participates in some of the information processing. Some models are based on the processing of visual motion in V1 and MT. Bayerl & Neumann (2007) proposed a feedforward and feedback computational model based on V1 and MT, which involves the early processing stage of both the dorsal and ventral visual pathways. Nishida (Nishida et al., 2018) divided motion into two different stages and proposed a computational model for visual motion processing based on the physiological mechanisms of V1 and MT. Motion information includes different motion features, such as motion speed, motion direction, etc. Diverse motion patterns and scenes pose challenges to the accuracy and effectiveness of machine vision recognition (Snowden et al., 1991).

Here we investigate visual motion processing based on fMRI at both 3T and 7T, using classic random dot kinematograms (RDKs) as visual stimuli. V1 retinotopic mapping and MT+ functional localizer were performed for each subject to obtain the regions of interest (ROIs) for subsequent analysis. This study combines univariate and multivariate analysis methods (using SVM classifiers) to study the fMRI signals of MT+ and V1. By comparing the results of univariate and multivariate analysis, the encoding mechanism of visual motion coherence and speed information in the dorsal visual pathway is explored. The specific research plan is shown in **Figure 1.1**. In addition, we compare the experimental results based on fMRI data at 3T and 7T to investigate whether high-resolution fMRI imaging improves the decoding of visual motion information.

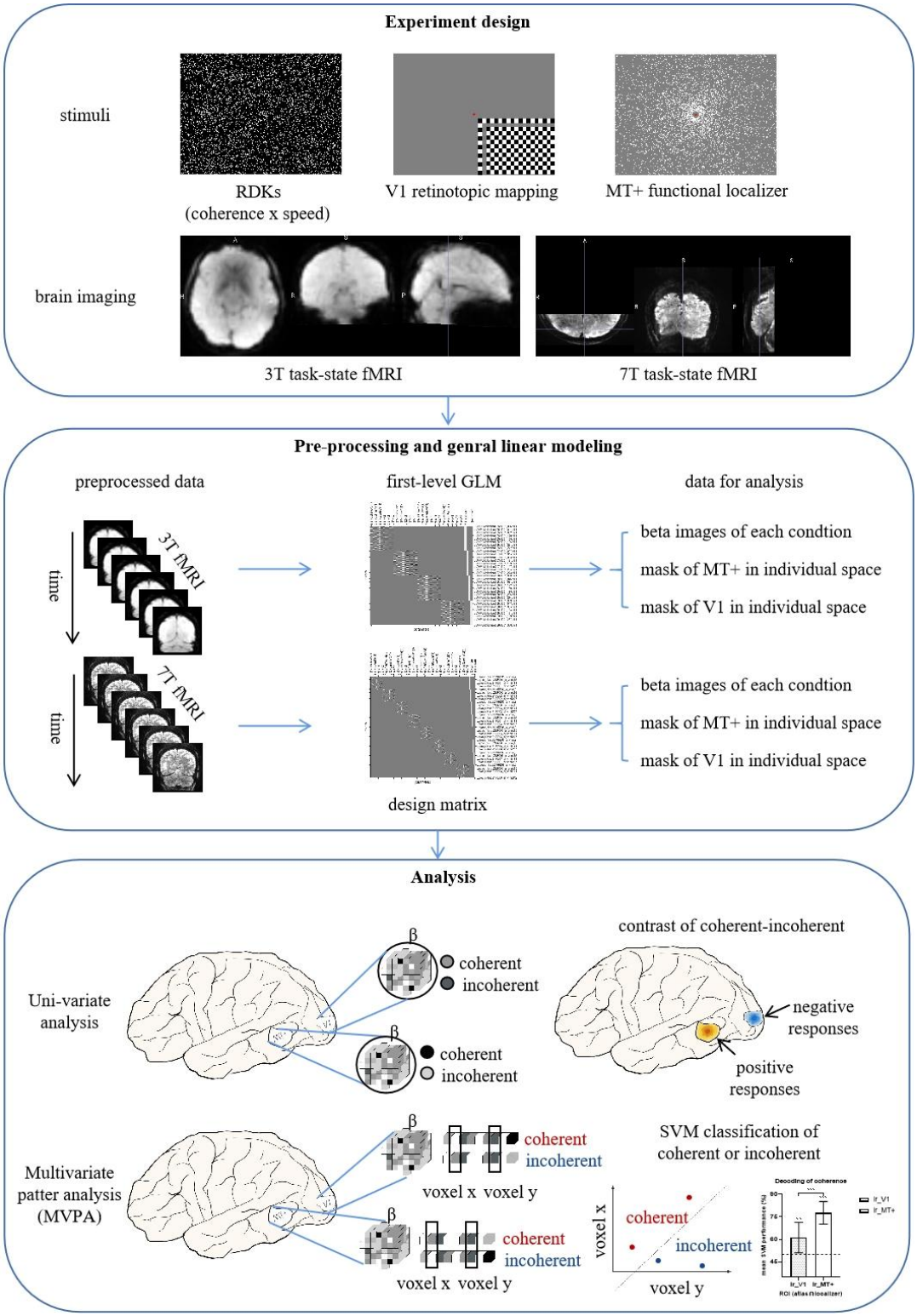


Figure 1.1 Research program. Experiment design, the creation of stimuli, and fMRI are followed by data pre-processed. General linear models are calculated for the univariate and multivariate pattern analysis. This study aims to use functional magnetic resonance imaging (fMRI) technology based on 3T and 7T to investigate the encoding mechanism of visual motion information in the dorsal visual pathway during visual motion perception in healthy human subjects. Only partial brain focusing the visual cortex was imaged by 7T fMRI (more details see MRI data acquisition in 3.1.2).

We defined 3 main research goals:

(1) To explore how MT+ and V1 encode visual motion coherence and speed information in the dorsal visual pathway.

(2) To investigate whether MT+ and V1 show brain lateralization when encoding visual motion coherence and speed information.

(3) To explore whether high-resolution fMRI data is advantageous for decoding visual motion coherence and speed information.

2 Background

2.1 Visual motion perception

Visual perception is one of the most important senses for humans and animals, as it enables us to perceive the world and respond to external stimuli. To understand visual motion perception, it is important to first understand how vision is formed. When external stimuli enter the eye through light, the visual system is responsible for converting the visual stimuli into signals that can be transmitted to the brain. The visual cortex in the brain plays a crucial role in the process of visual perception. Photoreceptor cells on the retina receive light stimuli from the external world and convert them into neuronal signals, which are then transmitted through the optic nerve to the lateral geniculate nucleus. The lateral geniculate nucleus, located between the retina and the visual cortex, receives inputs from different types of retinal ganglion cells. In summary, the visual system processes visual information through different information streams in parallel. Various visual information is transmitted to the brain through the retina and optic nerve, and the processing of these visual signals in relevant brain regions enables us to perceive the real and rich external world.

When the brain processes visual motion information, visual motion perception is formed. In many species such as frogs, rabbits, and flies, the existence of direction-selective neurons can be observed at the retinal level. This detection of external motion may play a role in identifying prey and predators. Animals must recognize moving objects and respond quickly in order to maximize their chances of survival. Therefore, motion

detection is important for animals. In addition to simple translational motion, motion can also be complex and diverse, including contraction, expansion, and rotation. Motion can occur not only in a two-dimensional plane but also in a three-dimensional space. Stereoscopic motion perception is also an important type of visual motion perception (Orban et al., 1999), which enables us to perceive stereoscopic motion and distinguish it from motion in a two-dimensional plane. Without visual motion perception, everything we see would remain stationary, and all actions related to motion would be difficult to complete. Given the importance of visual motion perception to human life, research has to explore the mechanisms by which the human brain processes visual motion information.

Direction selectivity is an extremely important functional characteristic in visual motion perception. Direction-selective neurons in the brain are neurons that are very sensitive to motion stimuli, and they may encode information such as motion direction, speed, and stimulus contrast, enabling the brain to recognize different motion patterns. Motion-sensitive neurons are considered motion detectors. They enable us to not only recognize motion in a specific direction but also easily recognize motion in different directions. Global motion perception benefits from the complex processing of different motion components in the brain, allowing us to recognize a specific direction of motion among many chaotic motion directions, i.e., perceiving the existence of coherent motion as opposed to incoherent motion (Lam et al., 2000). However, when a target moves in a certain direction for a period of time, motion aftereffects occur (Roger BH Tootell et al., 1995), where the neurons responsible for motion detection adapt to the motion direction and weaken neuronal activity, resulting in a perception that the stationary object is moving in the opposite direction, known as the waterfall illusion.

In addition to perceiving real motion, the brain also responds to apparent motion (Tanaka et al., 2007). Real motion refers to an object or target moving from one place to another, while apparent motion occurs when an object is presented successively at the starting and ending positions of motion. When presented at an appropriate speed, the brain may

perceive this as continuous motion. When the speed is too slow, the brain may not perceive the existence of motion, while when the speed is too fast, these stationary targets may be perceived as synchronously presented rather than moving continuously. Visual media such as movies, cartoons, and animations utilize the principle of apparent motion, presenting a series of static images to create videos and animations. When visual images contain inferred motion information even if the images being viewed are static, we can still perceive motion (Fawcett et al., 2007). In addition, motion-sensitive visual areas can also perceive motion transparency, where two targets moving in opposite directions can be perceived as transparent rather than two independent planes (Qian & Andersen, 1994).

In short, it is a complex process for the brain to perceive visual motion, and the realization of visual motion recognition does not depend solely on the change of the physical position of the moving target. An experiment using random dot patterns (random dot patterns) presents thousands of dots at random positions and changes the positions of some of them. At this time, it is difficult to distinguish and find the dots whose positions have been changed. When displayed sequentially in an animated manner, the areas of those points that have been altered and the corresponding motion will be more clearly visible. This kind of visual stimulation material is called Random Dot Kinematograms (RDKs) where subjects are less likely to report the visual motion perception by virtue of recognizing the change in the position of the dots, thus it is widely used in studies of visual motion perception in humans (Lam et al., 2003; Nakamura et al., 2003; Nakayama et al., 1985; Ulbert et al., 2001).

In summary, visual motion perception is one of the most important perceptions of human beings. This perception process involves the complex processing and processing of visual motion information by the brain. More specifically, the brain mainly encodes information such as the direction, speed, and brightness contrast of visual movements. This study will combine the popular RDK random dots motion stimulation to study the

coding mechanism of visual motion coherence and speed in the motion-sensitive brain area of the healthy human visual cortex.

2.2 Visual motion perception in the dorsal visual pathway

Two visual pathways in the visual cortex are responsible for processing different visual information, where the ventral visual pathway is primarily responsible for object recognition and the dorsal visual pathway is mainly responsible for processing motion information (Goodale & Milner, 1992; Ungerleider & Haxby, 1994). (Kaneoke, 2006; Nakamura et al., 2002) suggested that there may be two different connection pathways in the dorsal visual pathway, one related to visual-induced motion that projects from V1 through V6 to the superior parietal lobe (SPL), and the other related to spatial awareness and motion processing. The latter projects from V1, passes through V2 or skips V2 directly to MT/V5+, then enters the superior temporal sulcus (STS) and finally reaches the inferior parietal lobe (IPL). The ventral pathway is generally believed to originate from V1, pass through V2 and V4, and reach the inferior temporal area (IT), which is responsible for object recognition. It is evident that V1 and MT are two crucial brain regions in visual motion processing.

Before visual information reaches the visual cortex, the magnocellular pathway of the lateral geniculate nucleus transmits visual motion information and ultimately converges into the dorsal visual pathway via V1. Livingstone & Hubel (1988) detailed their view on the magnocellular and parvocellular pathways, which they believed were the basis for the functional segregation of the visual system. They proposed that the magnocellular pathway is associated with visual motion perception. When visual signals enter the lateral geniculate nucleus, the signals in the magnocellular layer (M-cell layer) pass through layer 4B of V1. They are then transmitted to V2 before reaching MT, which is primarily responsible for processing motion information. This may be related to the faster conduction speed of the magnocellular pathway, which responds to fast stimuli. Inputs from the parvocellular pathway are divided into two different information streams in layers 2/3 of V1, projecting color and shape information to V2 and then to V4 (Sincich

& Horton, 2005). Research on the magnocellular pathway highlights the crucial role of MT and V1 in perceiving visual motion.

Human subjects can easily tell the direction of movement of external visual stimuli (such as to the left or to the right) suggests that there is an area of the brain responsible for this cognitive function. However, this does not provide much information on where in the brain it is located. Therefore, many studies have explored direction selectivity. Various research methods such as single-cell activity labeling (Clarke & Miklossy, 1990; Tootell & Taylor, 1995), positron emission tomography (Watson et al., 1993; Zeki et al., 1991), functional magnetic resonance imaging (Orban et al., 2003; Roger B Tootell et al., 1995), and magnetoencephalography (Bedny et al., 2010) have been used.

Analyzing the direction and speed of moving objects is considered the main function of the dorsal visual pathway (Nakamura et al., 2002). The key regions associated with this function are in the middle temporal complex (MT+) that belong to the extrastriate visual areas. However, direction selectivity was first discovered in the primary visual cortex (Kamitani & Tong, 2006), where layer 4B cells in V1 have orientation and direction selectivity but are not sensitive to color information (Kamitani & Tong, 2006; Sincich & Horton, 2005). MT (middle temporal area) and MST (media superior temporal area) were first found in the visual cortex of monkeys (Maunsell & Van Essen, 1983), and they are more sensitive to direction and speed than other visual areas (Mikami et al., 1986; Saito et al., 1986). It was later discovered that the human visual cortex has homologous brain regions MT (also called V5) and MST (DeYoe et al., 1996; Huk et al., 2002). Due to their adjacent positions and sensitivity to motion, these two brain regions were not deliberately distinguished in early studies of the human brain. (DeYoe et al., 1996) proposed the naming of MT+, which includes two subregions, MT and MST. Their locations are shown in **Figure 2.1**, with MT being more posterior and MST being more anterior relative to MT. The MT area responds only to stimuli in the contralateral visual field, whereas area MST receives input from both the ipsilateral and contralateral visual fields (Smith et al., 2006). Almost every cell in MT exhibits direction selectivity, and MT's

direction-selective neurons play an important role in judging motion components such as direction and speed. The study of motion after-effects (Roger B Tootell et al., 1995) and opponent motion (Heeger et al., 1999) provides strong evidence for the direction selectivity of human V5/MT.

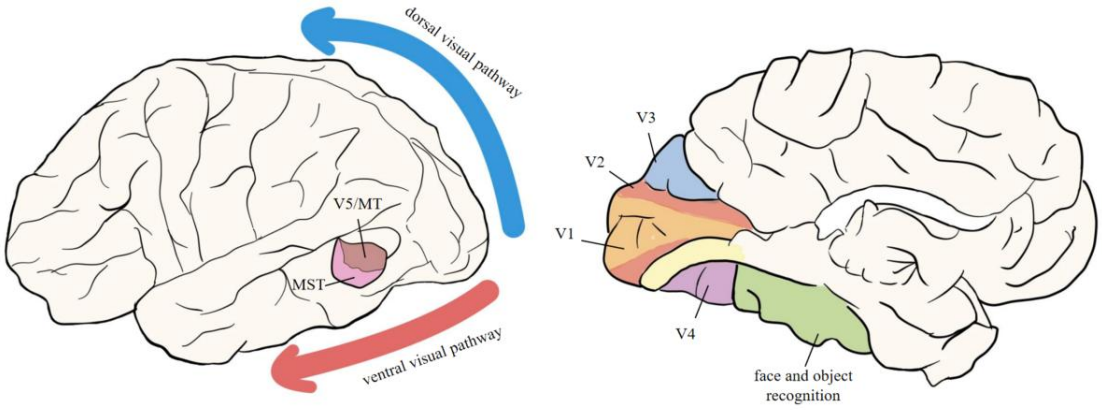


Figure 2.1 Two visual pathways. MT+ consists of MT and MST, and V5/MT is on the posterior side of MST.

Damage to the MT area has been found to affect visual motion perception in both monkeys (Newsome & Pare, 1988) and humans (Baker et al., 1991), and abnormalities in MT and MST activation have also been observed in human migraine patients (Antal et al., 2011). An fMRI study found no significant differences in MT activation intensity between children and adults during visual motion perception (Taylor et al., 2018), while (Ward et al., 2018) used fNIRS to show that older adults have weakened motion coherence perception ability (manifested as an increase in the threshold level of motion coherence).

Damage to the MT area has been found to affect visual motion perception in both monkeys (Newsome & Pare, 1988) and humans (Baker et al., 1991), and abnormalities in MT and MST activation have also been observed in human migraine patients (Antal et al., 2011). An fMRI study found no significant differences in MT activation intensity between children and adults during visual motion perception (Taylor et al., 2018), while (Ward et al., 2018) used fNIRS to show that older adults have weakened motion

coherence perception ability (manifested as an increase in the threshold level of motion coherence).

The integration of local motion information (primarily including motion direction and speed) enables us to perceive global motion, while the selectivity for motion direction helps us judge whether the directions of local motion are coherent. Coherent motion is composed of many different local motion vectors (Howard et al., 1995; Williams & Sekuler, 1984), and the perception of global motion patterns requires the integration of local motion information in visual space (Bex & Dakin, 2002). The brain may process visual motion information by processing two independent scalars (direction and speed), thereby achieving the integration of local motion perception over a larger visual field (Kaneoke & Kakigi, 2006). Therefore, studying the moton-sensitive areas of the MT+ region in the context of perceiving global motion at different speeds may help us understand the mechanism behind processing global motion. Local motion signals are generated in the primary visual cortex and are spatially integrated starting from MT/hMT to form global motion perception. As mentioned earlier, random dot kinematograms (RDKs) have been widely used in visual motion research (Laycock et al., 2007), and discrimination between coherent motion and noise (generally incoherent motion or static random dots) is often applied in studies of global visual motion processing or motion coherence (Van de Grind et al., 1987). The initial visual motion brain imaging studies, (Watson et al., 1993), (Dupont et al., 1994), (Roger BH Tootell et al., 1995) and others, compared motion stimuli to static stimuli, but this method had a problem in that it did not take into account the specificity of motion-sensitive areas for motion direction. Therefore, (Braddick et al., 2001) used a uniformly moving random pixel stimulus, which displayed the same pixels as coherent motion stimuli at the same frame rate. These pixels moved in random directions, thus unifying the spatiotemporal frequency of the two types of motion stimuli. By comparing coherent and random motion noise, not only can high temporal frequency signals be reflected, but brain areas sensitive to motion direction can also be better isolated.

Previous research has shown that both monkeys and humans can detect coherent motion from random dot motion stimuli, even when the proportion of coherently moving dots is low. The previous stimulus's motion speed has a significant impact on the response latency (Lam et al., 2000). The frequency of switching between coherent and incoherent motion stimuli has little effect on the motion-sensitive brain areas (Braddick et al., 2001). Encoding of motion speed information has been found in both MT+ and V1. (Chawla et al., 1998) found that 7-30 deg/s elicits high amplitude responses in human MT based on a 2T fMRI study, which is in agreement with previous electrophysiological results (Cheng et al., 1995). In addition, V1's response decreases linearly with increasing motion speed, and compared to MT+, the motion-sensitive neurons in V1 show direction selectivity for slower motion stimuli. (Kumano & Uka, 2013) indicated that when random dot stimuli contain multiple motion directions, MT neurons are better able to discriminate two-dimensional speed. Therefore, the perception of coherent motion involves the processing of motion direction and speed. Direction-selective neurons have preferred directions, and their response is highest when the target moves in the preferred direction, while their response may be inhibited when the target moves in the opposite direction. In addition to the motion feature information carried by the motion stimulus itself, the receptive fields of visual areas may also affect their encoding ability for information such as motion coherence or motion speed. First, the receptive field of MT is approximately 10 times larger than that of V1, so there are fewer motion stimuli falling within the receptive field of V1 neurons. Even if the motion stimulus falling within the receptive field moves in the direction preferred by the neuron, the resulting neural response in the entire V1 may be very weak. The larger receptive field of MT neurons allows them to process more motion inputs simultaneously. When the stimuli move in a single direction, the group of neurons with a preference for that direction will respond strongly, resulting in strong neural activity or response signals in MT, which means that coherent motion stimuli will elicit strong responses in MT. Secondly, motion opponency may also affect the response of visual areas to motion coherence, and the effect of stimuli presenting random noise with multiple different motion directions (incoherent motion) may be

reduced due to the presence of motion opponency. Studies related to motion opponency generally use stimuli that move in opposite directions simultaneously (usually point-pair stimuli, with each point in a pair moving towards the position of the other point).

Both fMRI and single-unit studies have shown that motion opponency-induced inhibition is much weaker in V1 than in V5 (Heeger et al., 1999). Stimuli presenting random noise or incoherent motion elicit responses from V1 neurons with preferences for multiple direction biases (strongest response to their preferred direction). V1 is subject to inhibition feedback during the perception of coherent motion, resulting in a stronger response to incoherent motion in V1 (Harrison et al., 2007). In contrast, motion stimuli in the surrounding receptive field of MT neurons will inhibit their neural activity, resulting in greater inhibition of MT responses to incoherent motion or random noise. Under these conditions, MT is not completely unresponsive, ultimately showing a weaker response to incoherent motion than to coherent motion (Qian & Andersen, 1994). In addition to motion opponency-related inhibition, speed may also affect the response to coherent motion in V1 and MT. Coherent motion stimuli that are too fast (20 deg/s) may exceed the central visual field's speed range of V1, which may lead to weaker responses to coherent motion in V1 (Braddick et al., 2001). Conversely, motion stimuli in the receptive field surround of MT neurons will inhibit their neural activity. Although there is no significant difference in MT responses to slow and fast random dot motion, Braddick suggested that caution should be exercised when interpreting responses to faster coherent motion, as the observed response may be a response to dynamic noise (incoherent motion). Therefore, the response of V5 to incoherent motion or random noise may be subject to greater inhibition. However, MT is not completely unresponsive in these situations, ultimately showing a lower response intensity (Qian & Andersen, 1994). MEG studies (Aspell et al., 2005; Nakamura et al., 2003) and fMRI studies (Rees et al., 2000) have found that the response of MT increases with the level of motion coherence, but there is no clear conclusion about the response of MT to coherent and incoherent motion. (Cheng et al., 1995) used PET studies to find that both MT and V1 can be activated by coherent and incoherent motion, but no differences were observed between the two

regions in response to the two motion patterns. (McKeefry et al., 1997) reported in fMRI and PET imaging studies that both V1 and MT have higher response amplitudes to incoherent motion, which is in agreement with Manning's results based on VEP studies. However, the authors indicated that fMRI could not reflect MT's sensitivity to motion, and therefore it failed to reveal the fact that MT has a stronger response to coherent motion. (Braddick et al., 2001) hypothesized that this result is due to the use of very low dot density (average 0.19 dots/deg^2) in the experiment, under which there are not enough coherent motion stimuli within any receptive field to effectively activate the corresponding neurons in MT. Therefore, they used a higher dot density and obtained the opposite conclusion using fMRI imaging, namely that MT shows a higher response amplitude to coherent motion, while V1 shows a higher response to random noise (incoherent motion). Maruyama (Maruyama et al., 2002) tested the effects of motion speed and random dot density on motion coherence responses, and found that the response strength and time of MT to both motion patterns are almost unaffected by speed and density, and the response to coherent motion is greater than that to incoherent motion at all speeds or densities. (Becker et al., 2008) also found, based on fMRI studies, that the response of MT+ (including MT and MST) to coherent and incoherent motion is not affected by random dot density. In addition, a combined fMRI and MEG study showed that the response time and intensity of human MT are modulated by motion speed (Kawakami et al., 2002), but there is no difference in the representation of speed in MT between coherent and incoherent motion (Lam et al., 2003).

2.3 High-resolution functional magnetic resonance imaging

Functional Magnetic Resonance Imaging (fMRI) is a technique used in cognitive neuroscience to obtain brain functional images based on the Blood Oxygen Level Dependent (BOLD) effect. Oxygenated hemoglobin and deoxygenated hemoglobin exhibit different magnetic properties in the brain tissue, with oxygenated hemoglobin showing diamagnetic properties while deoxygenated hemoglobin shows paramagnetic properties. When the brain is active, local blood flow changes can cause changes in the

oxygenation ratio of the two hemoglobin species. BOLD-based MRI utilizes this change. The BOLD signal is approximately proportional (with a negative sign) to the content of deoxy-hemoglobin. It measures fMRI signals over time. If a stimulus is presented, it measures the so-called 'BOLD response'. The higher the ratio between the two hemoglobin species, the higher the amplitude the MRI signal shows. The BOLD response is generally believed to be a marker of functional activity in the brain. BOLD fMRI technique (Ogawa et al., 1992) provides a powerful non-invasive tool for studying the human visual cortex.

Functional magnetic resonance imaging can be further divided into resting-state and task-state. Resting-state fMRI (rs-fMRI) refers to brain imaging data collected without any specific stimulation when the brain is in a free activity state. Resting-state brain imaging data is often used for disease research. Task-state fMRI (task-fMRI) requires subjects to perform a series of cognitive tasks while undergoing MRI brain scans, such as watching specific visual stimuli to induce perceptual processes such as visual, memory, emotion, and attention control. Therefore, if significant activation is observed in certain brain regions during a certain period (as evidenced by changes in the BOLD signal under task conditions), it can be considered that the activation of that brain region is related to the task. Multiple studies have shown that visual motion stimuli can affect the fMRI amplitude (Harvey & Dumoulin, 2016). For example, Tootell et al. (1998) studied the functional characteristics of human V1 using fMRI. Therefore, this study explored visual motion perception-related activities of brain regions in the dorsal visual pathway based on fMRI.

Task-state functional magnetic resonance imaging places high demands on experimental designs. Although the BOLD signal recorded by fMRI can reflect neural activity in the brain, there is still a gap between BOLD signals and the actual neural activity. The neural activity begins with the spike firing of neurons, which generally only lasts for a few milliseconds, while local blood flow activity reaches its peak about 5 s after neuronal firing. The analysis of functional magnetic resonance imaging uses hemodynamic

functions to calculate the amplitude of the BOLD signal collected. Therefore, the BOLD response may be delayed compared to actual neural activity, which is also one of the factors that contribute to the low temporal resolution of functional magnetic resonance signals. Experiments in fMRI studies, especially when the time resolution is not high enough, might use block design with each trial lasting several seconds or longer to make efforts that fMRI signal collected corresponds better with the experimental stimuli or tasks before the signal reaches its peak. Setting a baseline between two experimental conditions is also a commonly used method to effectively separate the preceding and following experimental task conditions, ensuring that the collected signal is not affected by the delay in magnetic resonance signal acquisition.

In addition to experimental design, the hardware properties of the magnetic resonance scanner itself may also limit the effectiveness of magnetic resonance imaging. With the development of magnetic resonance imaging technology, the magnetic field strength of magnetic resonance scanners has increased from a maximum of 1.5 T to the widespread use of 3T magnetic resonance scanners, and now ultra-high field magnetic resonance imaging (7T and above) has also been successfully developed. Currently, 1.5T or 3T magnetic resonance scanners are used for clinical diagnosis and treatment, while 3T magnetic resonance scanners are commonly used for scientific research. 7T ultra-high field magnetic resonance scanners are used for human studies, and 9.4T magnetic resonance scanners are used for small animal studies. The increase in field strength means that higher spatial resolution can be achieved, and brain functional imaging under ultra-high magnetic fields can achieve a higher signal-to-noise ratio and spatial resolution (De Martino et al., 2013). For fMRI techniques with poor temporal resolution (in seconds), improving spatial resolution becomes particularly important. The resolution of fMRI imaging based on 3T is usually 2-3 mm^3 per voxel, while fMRI imaging based on 7T can achieve voxel resolutions of 1 mm^3 or even submillimeter levels, greatly improving spatial resolution. Becker (Becker et al., 2008) mentioned that studies of non-human primates have made important contributions to our understanding of global visual motion perception. Limited by research techniques. human studies that measure

brain activity based on the response of a large number of neurons are far from being perfect, far behind what can be achieved with invasive measurements in humans. On the one hand, single-cell recordings allow animal experiments to adjust experimental stimuli according to neuronal preferences, but this technique cannot be used in human studies. On the other hand, what non-invasive functional imaging collects is the collective response of many neurons, which cannot accurately reflect the functional specificity of single neurons, such as the preference for different motion directions, etc. The spatial resolution of imaging technology (the Studies using a 2x2x2 mm^3 voxel size) has been insufficient to functionally separate responses from adjacent cortical regions. However, the emergence of 7T has brought hope for high spatial-resolution imaging (Stephan et al., 2019). Existing studies have explored the encoding mechanism of visual scene information in the primary visual cortex (Muckli et al., 2015) and the spatial attention mechanism (Liu et al., 2021) at the laminar level. In addition, measurements within cortical structures became possible (Lawrence et al., 2019), and (Harvey & Dumoulin, 2016) achieved imaging of the visual cortex toward functional columns by high-resolution fMRI.

Bartels (Bartels et al., 2008) hypothesized that fMRI, as a non-invasive brain imaging technique, should be regarded as a supplement rather than a low-resolution alternative to invasive research techniques such as electrophysiological methods. The emergence of fMRI provides the possibility of non-invasively studying the functional characteristics of the cortex, especially the healthy brain cortex. Combining fMRI techniques makes it possible to compare various cognitive functions between humans and non-human primates. The higher the brain area, the greater the functional differences that may occur between human and monkey brains, so these comparisons mostly occur in the visual system, and homologous visual areas V1-V5 have already been defined in the human brain (Orban et al., 2004).

In summary, high-resolution fMRI imaging techniques provide an opportunity to non-invasively test and study the neural physiological mechanisms of the brain cortex in

humans. This study will study MT+ and V1 in the dorsal visual pathway based on fMRI at both 3T and 7T.

2.4 Multivariate pattern analysis

2.4.1 Encoding and decoding of neural activity patterns

Encoding and decoding computational models are widely used in cognitive neuroscience (Kriegeskorte & Douglas, 2019). The encoding model performs calculations from experimental conditions to brain responses by controlling experimental stimuli to predict and compare neural activity elicited by different experimental stimuli. The decoding model, on the other hand, infers the cognitive events experienced by the subject when generating the neural activity measured by brain imaging techniques (such as what kind of experimental stimuli were presented or what experimental tasks were performed). Simply put, the encoding process involves various brain regions producing specific responses after processing stimuli, while decoding refers to inferring the content of the stimulus solely based on brain activity signals in the absence of knowledge of the stimulus. In the perception of visual motion, encoding can be understood as the subject's brain producing specific responses after receiving different types of visual motion stimuli, from which the brain areas related to each type of motion stimulus can be determined, and thus the functional relevance of these brain areas in processing the motion information can be inferred. Conversely, decoding can be understood as revealing the stimulus information carried by the signals by analyzing the brain response signals collected during the experiment, i.e., which type of stimulus the subject observed during the time corresponding to the signal, which is usually achieved through classification analysis methods.

2.4.2 MVPA

Multivariate Pattern Analysis (MVPA) is one of the most used methods for decoding brain fMRI activity. It shows higher sensitivity than traditional BOLD response analysis. Traditional single voxel fMRI analysis relies entirely on the information contained in the

time series within a single voxel or a group of voxels, while MVPA takes into account the spatial response pattern of all voxels. The conventional time-course analysis of fMRI can lead to the assumption that the same brain region shows equal BOLD activation values for different stimuli, thereby masking pattern information. BOLD response-based activation analysis simply adds up the activation values, while MVPA analysis is a cross-voxel pattern analysis that allows for more detailed analysis on each trial of an individual, meaning specific brain activity patterns can be extracted for each experimental condition (non-average). Therefore, multivariate pattern analysis can extract more valuable information regarding brain function from fMRI data.

MVPA analysis methods commonly include ROI-based classification, searchlight analysis, and Representation Similarity Analysis (RSA). The first method requires prior knowledge of which part of the brain may have better decoding performance. Randomly selecting an ROI for classification may not result in the expected decoding level. Searchlight analysis can perform decoding analysis without prior knowledge. It selects arbitrary clusters of any size for analysis at the whole-brain level, essentially performing classification. Typically, neighboring clusters are selected with each voxel as the center and a radius specified in terms of the number of voxels, or a sphere is drawn in millimeters and all voxels inside the sphere are selected as the ROI. Therefore, ROI-based classification methods are more suitable when a certain brain region is known to be related to a specific functional activity or when there is a clear ROI of interest. On the other hand, searchlight analysis is more suitable for situations where the location of the related brain region is not clear. It helps us find the most sensitive brain regions to stimuli. However, the drawback is that no good standard for the number of voxels in the neighborhood has been established. When the cognitive function being studied requires the involvement of a larger brain region, it may not be possible to select an appropriate size to provide sufficient information for classification. RSA analysis refers to the representation of brain activity signals in high-dimensional space, mapping the neural activity signals of each experimental condition to a high-dimensional space, which is considered as the neural activity pattern corresponding to each experimental condition.

Then, correlation analysis is performed on the mappings of different experimental conditions within the same brain region, usually using Pearson or Kendall correlation. Finally, by constructing a Representative Dissimilarity Matrix (RDM), the correlation under different conditions is presented (what is calculated here is the uncorrelated value, i.e., 1 minus the correlation obtained earlier, which can be simply understood as the distance or dissimilarity between the activity patterns of two conditions). In RSA analysis, by observing the differences in the activity patterns represented by neural signals under different experimental conditions, the brain representation for different stimuli can be compared. Therefore, RSA analysis is more suitable for studying the differences in representation patterns of the same brain region for different experimental tasks or stimuli.

The effectiveness and advantages of MVPA analysis in visual research have been demonstrated in many fMRI studies. (Haxby et al., 2001) used RSA analysis to study the decoding of object recognition by the ventral visual pathway, and the results showed that the functional specificity of the ventral temporal cortex is not limited to the regions that respond most strongly to specific stimuli. The neural activity patterns of brain regions that respond most strongly to certain categories of stimuli may also carry information about other categories of stimuli. For example, the fMRI activity of the cortex that responds most strongly to faces can also effectively discriminate between different types of visual stimuli such as houses, cats, and chairs, indicating that the brain region sensitive to faces also encodes visual information about these other objects. Thus, the results of multivariate pattern analysis may reveal neural encoding mechanisms that cannot be explained by uni-variate activation analysis.

In the dorsal visual pathway, fMRI imaging has successfully identified brain regions with stronger visual motion responses (Heeger et al., 1999; McKeefry et al., 1997; Roger BH Tootell et al., 1995), and has also revealed adaptation responses to repeated motion directions (Huk & Heeger, 2002; Nishida et al., 2003; Seiffert et al., 2003). However, there have been relatively few fMRI decoding studies of motion processing, and the

results of these studies have supported different conclusions regarding the decoding performance of V1 and MT for motion direction. One conclusion suggests that V1 has better motion direction decoding ability than MT. (Kamitani & Tong, 2005) used a decoding classification method to study the discrimination ability of V1 and MT+ (including MT and MST) for eight motion directions, and the results showed that both V1 and MT+ successfully decoded each motion direction, but MT+ did not show better classification accuracy. (Wang et al., 2014) used similar motion stimuli (random dots moving along one of eight motion directions) and found that the fMRI responses of multiple visual areas could successfully decode the direction of coherent motion, and the decoding accuracy improved with an increase in the number of voxels. However, regardless of the number of voxels, the decoding performance of V1-V3 was better than that of MT+ (including MT and MST). (Schwarzkopf et al., 2011) used scattered elements to form coherent motion perception and also found that MT had the poorest decoding performance among multiple visual areas (including V1-V4). However, it should be noted that the coherent motion here is different from the coherent motion mentioned earlier. In this study, the visual stimuli consisted of multiple gratings moving in an arc instead of all stimuli moving in the same direction. Another conclusion supports the view that MT has better motion direction decoding performance. Furlan (Furlan & Smith, 2016) found that the decoding accuracy of MT could reach above 70%, while the fMRI activity of V1 could not decode global motion direction (below chance levels). (Kamitani & Tong, 2006) hypothesized that although MT+ is a brain region highly sensitive to motion, it did not show better motion direction decoding performance than other early visual areas, which may be related to the smaller area and number of voxels in MT+. However, the decoding performance of MT+ for motion direction is still better than that for orientation information, which shows that MT+ is more sensitive to visual motion direction than form information. Thus, the results of MVPA classification may be influenced by the size of the ROI, but they can still reveal functional characteristics of the cortical regions to some extent.

Machine learning-based MVPA analysis imposes higher requirements on data volume. In fMRI-based MVPA classification, higher spatial resolution means more available samples. (Muckli et al., 2005) used SVM classification analysis based on 7T ultra-high-resolution fMRI data and found that the decoding performance of each layer of V1 was different, thanks to the ultra-high voxel resolution of 7T. Although the size of MT is much smaller than that of other early visual areas, high-field fMRI imaging can compensate for the problem of small data volume through higher spatial resolution.

2.4.3 SVM

MVPA classification analysis introduced machine learning (Mitchell et al., 2004), and support vector machines (SVMs) are the most commonly used model. SVM is a commonly used supervised learning method in data analysis and machine learning, which can be used for classification and regression analysis. Specifically, in classification problems, SVM constructs a hyperplane or a series of hyperplanes to divide the space into parts, thus achieving the classification task. In regression problems, SVM finds the range of as many sample points as possible falling on both sides of the hyperplane. SVM is an effective machine learning technique applicable to various situations. It has many advantages over other algorithms such as decision trees and random forests, including good performance in high-dimensional space, the ability to handle non-linear feature relationships, strong generalization ability, reducing the chances of overfitting, and applicability to small and large datasets. However, when used on large datasets, the computational cost of the training process is high, and data preparation steps such as managing missing values are required. To address these issues, some variant algorithms have been developed, such as soft-margin SVM and kernel SVM. SVM can be divided into linearly separable SVM, linear SVM, and nonlinear SVM. Linearly separable SVM is usually called hard-margin SVM, which requires the training data to be completely linearly separable and does not allow any misclassified samples. Its goal is to find a hyperplane that can correctly separate the training data, so that positive and negative

samples are located on both sides of the hyperplane, and the distance between them is maximized.

The basic idea of SVM is to find an optimal hyperplane in high-dimensional space that can separate samples of different classes. Specifically, given a training dataset, the goal of SVM is to find a hyperplane H: $WX + b = 0$ (where W is the normal vector and b is the bias term) that can divide samples of different classes on both sides and satisfy the condition of maximum margin. The margin refers to the distance from the nearest point of each class sample to the hyperplane, including points of the same class as well as different classes, and the points closest to the hyperplane are called support vectors (shown in red circles in **Figure** 2.2), which satisfy the condition $y_i (X_i W + b) = 1$.

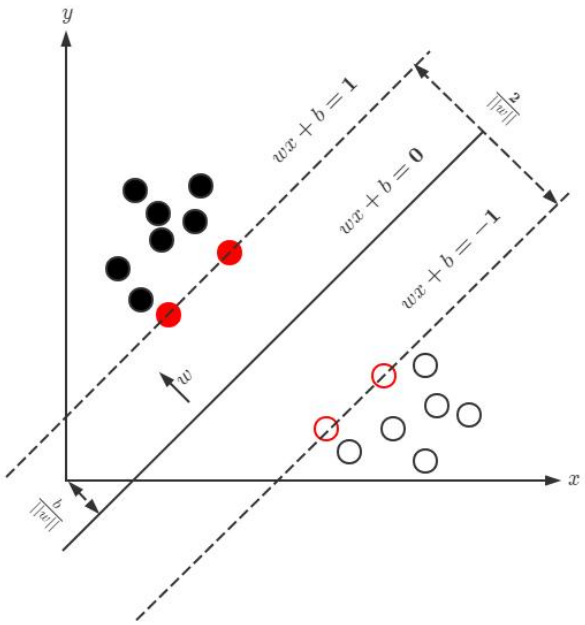


Figure 2.2 SVM classification

In summary, SVM classification can not only be applied to binary classification problems, but also effectively solve multi-classification problems, and the application scenarios of SVM are widely applicable to all types of datasets.

3 Methods

3.1 Experiment design

This study designed two visual motion stimuli experiments based on fMRI imaging. The first experiment was conducted using 3T magnetic resonance scanning (referred to as the "3T experiment"), and the second experiment was conducted using 7T magnetic resonance scanning (referred to as the "7T experiment").

Both experiments used random dot kinematograms (RDKs) as visual stimuli, which have two main advantages. First, the positions of the random dots are not specific, so the perceptual processes of encoding position information and motion information can be distinguished. Second, the motion of the random dots is two-dimensional, distinguishing the visual motion perception evoked by the experimental stimuli from stereoscopic motion perception. RDKs are feasible and efficient for studying the encoding of visual motion information in the visual cortex. In addition, this study used functional localizer runs for localizing MT+ and V1 for each subject using fMRI.

3.1.1 fMRI experiment at 3T

The first experiment was based on 3T magnetic resonance brain imaging to investigate whether stimulus parameters such as dot size, dot density, and motion speed can effectively elicit responses in the regions of interest, MT+ and V1, and to study their encoding of visual motion coherence and speed information. The stimulus parameters for the experiment conducted using high-resolution 7T magnetic resonance imaging were set based on the results of the 3T experiment.

Ethics and participants

The 3T experiment recruited a total of 12 healthy young participants, including 5 females. All participants had normal or corrected-to-normal vision, were right-handed, had no history of neurological or psychiatric disorders or family history of such disorders, had no metal implants in their head or body, and had no claustrophobia. Each participant volunteered to participate in the experiment, read and signed an informed consent form prior to the experiment, and received compensation after the experiment. The experiment was approved by the Ethics and Human Protection Committee of the Magnetic

Resonance Imaging Research Center of the University of Electronic Science and Technology of China.

Stimuli and paradigm

Before presenting any experimental stimuli, the subjects will be sent into the MRI scanner and watch an experimental preparation screen in a dark experimental environment. The subjects were allowed to slightly adjust the position of their heads to bring the entire screen within their field of view so that they could easily fixate on the center of the screen. When the experiment and scanning officially started, the subjects were asked to reduce their head movements as much as possible. All subjects completed the sessions of V1 retinotopic mapping experiment, MT+ functional localizer, and random dots motion experiment. All experimental stimuli and paradigms were generated using PsychToolbox 3.0 in Matlab.

Retinotopic mapping of V1 in the lower right quadrant. The visual cortex has the characteristics of retinotopic mapping. Ganglion cells in the retina are diverse, and the signals sent by these cells to the brain are part of the neural basis for vision (Masland, 2012). V1 has a retinotopic map of the field of view. That is, the signals from the retina are projected to the primary visual cortex in a topographic manner. The entire conduction pathway from the retina to the cortex maintains an anatomical point-to-point relationship. The stimulus from the receptive field is encoded into a secondary dimensional plane and preserves the original spatial organization of visual stimuli (Ibbotson & Jung, 2020). The anatomical boundaries of the V1 region are well-defined and almost all studies agree on this (Tootell et al., 1998).

Non-invasive methods such as fMRI can be used to localize V1 in individual brains by taking advantage of its retinal topological mapping property, which is considered to be one of the most accurate criteria for defining the visual cortex (Felleman & Van Essen, 1991). Checkerboard stimuli are widely used in retinotopic mapping experiments in the primary visual cortex. (Muckli et al., 2015; Muckli et al., 2005) used a checkerboard

flickering stimulus to locate the V1 of the subject, and divided the stimulus into two conditions: targe and surround, so as to isolate the voxels receiving feedback and eliminate the overflow of the feedforward stimulation area (Smith & Muckli, 2010) . Those voxels that only respond to the target region but not the surrounding regions are truly retinotopically specific.

According to the principle of retinotopic mapping, our study refers to the retinotopic mapping experimental stimulus used in (Muckli et al., 2015). In the 3T experiment, the V1 area corresponding to the lower right quadrant of each subject's visual field was retinotopically mapped, and it is regarded as the region of interest used in the subsequent analysis.

In the 3T experiment, visual stimuli were projected by a projector onto a projection board inside the scanner and finally into the subjects' eyes. The maximum viewing angle that the projection device can present is 24°×18° (display resolution is 1024×768), as shown in **Figure 3.1**, the retinotopic mapping in this study uses black and white checkerboard flickering (4Hz) stimulation. The checkerboard stimulus was presented only in the lower right quadrant, and the rest of the visual field was presented with a gray background. There is a red circular fixation point with a diameter of 0.3° in the center of the screen, and the color of the fixation point will switch back and forth between bright red and dark red every 3 to 6 seconds (each change time is randomly generated within this range). The subjects were asked to keep their attention and fixate on the red dot during the whole experiment (including V1 retinotopic mapping, MT functional localizer, random dot motion experiment), and to press the button as soon as possible after its color changed. The retinotopic mapping stimulus included two conditions: target and surround. The checkerboard presented by the target condition is within in the main area of the lower right quadrant (see **Figure 3.1(a)**), and the distribution of checkerboard stimuli in the surrounding area is in the peripheral area of the lower right quadrant (as shown in **Figure 3.1(b)**). That is, the stimulus in the surrounding area is closer to the center of the visual field than the stimulus in the target area. The lower right quadrant (12°×9°) is evenly

divided into 20×14 squares along the horizontal and vertical directions (so the side length of each square is about 0.6°), and the colors are black and white. The target area is composed of 17×11 squares, and the distance between the surrounding area and the target area, and separately, between the surrounding area and the central fixation is one row and one column of squares.

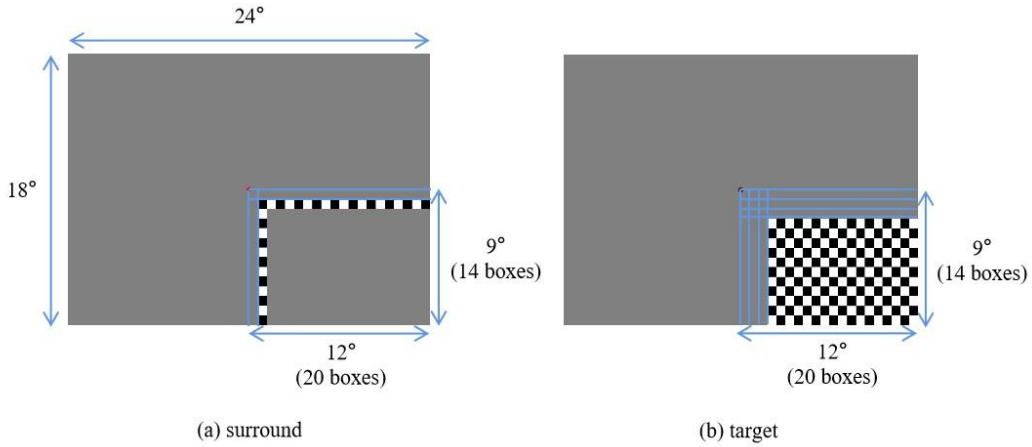


Figure 3.1 Stimulus of checkerboard for retinotopic mapping in the 3T Experiment

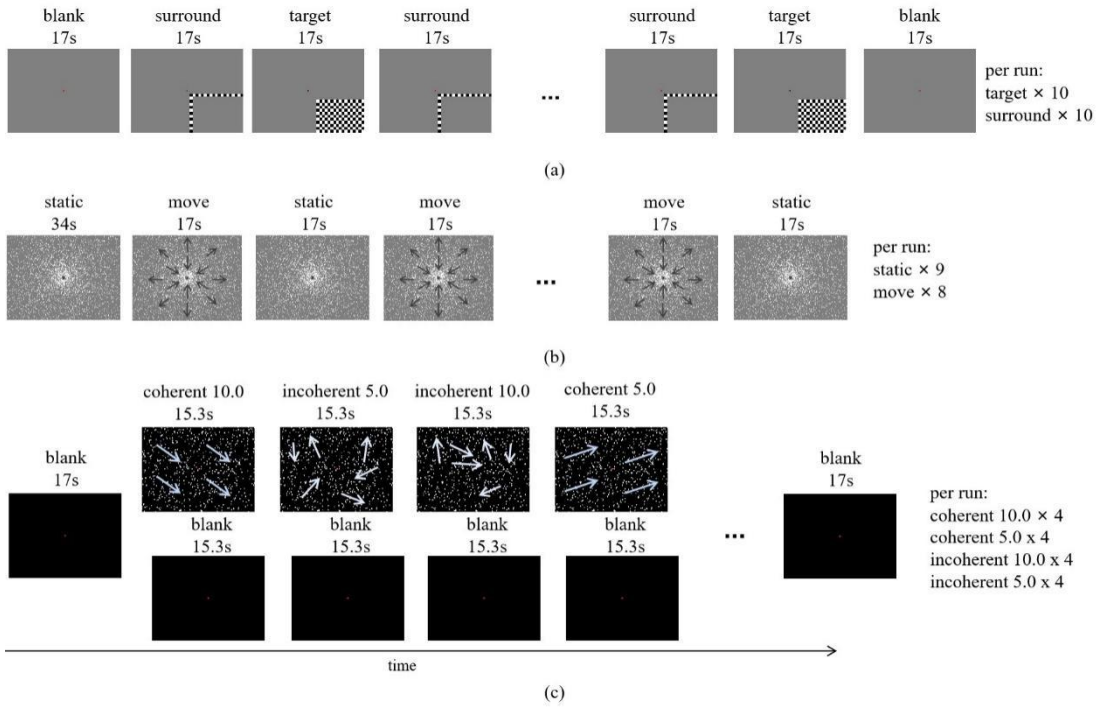


Figure 3.2 Stimuli and paradigm in 3T experiment. (a) retinotopic mapping for V1 in the lower right quadrant. (b) MT+ functional localizer. (c) random moving dots.

In the lower right quadrant V1 retinotopic mapping, a blank screen (only gray background and fixation point) was presented for 17 s at the beginning and end of each run as the baseline condition. The surround conditions and target conditions were alternately presented for the rest of the trials 10 times (17 s for each trial), see **Figure 3.2(a)** for the specific sequence. The above conditions were presented in the same order to all subjects, and the entire run of V1 retinotopic mapping experiment took 374 s (17 s + (17 + 17) s × 10 + 17 s = 374 s).

MT+ functional localizer. (Benson et al., 2014), (Wang et al., 2015), (Huang et al., 2019), and (Rosenke et al., 2021) found inter-individual differences in the position of MT+. For constructing the atlas of the visual cortex, (Huang et al., 2019) carried out a functional localizer of human MT and found that the model began to stabilize when the number of subjects was greater than 200, which means that when the number of subjects is small, there may be large individual differences in MT among subjects. Therefore, our study uses an MT+ functional localizer for each subject. This way, we obtain a more accurate mapping of the MT+ region based on the individual space of the subject. This will support the subsequent SVM analysis performed in individual subjects. Our study focused more on the way the dorsal visual pathway encodes motion coherence and speed information, so the two subregions of MT+ (MT and MST) were not distinguished in detail, and the localized region was referred to as MT+ in this study.

The visual stimuli used in the MT+ functional localizer experiment consisted of white dots and a gray background. White dots have a size of 0.14°, their positions are randomly generated and distributed over the entire gray background with a dot density of 10 dots/deg^2. As shown in **Figure 3.2(b)**, the experiment included two stimulus conditions, with static and moving dots. Under the condition of ‘move’, the white dots shrink toward the center first. Each dot disappears when it arrived at a radius of 2° from the center red dot. Then, all dots expand away from the center, and the dots that arrive at the edge of the screen disappear. Contraction and expansion movements alternated repeatedly every 1.7 s for each stimulus condition. In each MT localizer experiment, all subjects adopted a

fixed sequence of stimulus presentation, that is, the stimulus started and ended with the rest condition, and the rest condition and the stimulus condition were alternately performed. The rest condition was presented 9 times, and the stimulus condition was presented 8 times. The resting conditions at the beginning and end were presented for 34 s each, and the remaining trials were presented for 17 s. Each subject performed one MT functional localizer experiment, a total of 323 seconds (34 s + 17 s × 8 + 17 s × 7 + 34 s = 323 s).

Random dots. In the 3T experiment, the random dot motion stimulus consisted of a black background and white dots. The position of the white dot (size 0.14°, density 10.0 dots/deg) is randomly generated each time (for each subject and each run of experiments the dots are re-generated randomly). A red spot in the center of the visual field is used for the subject to respond to the button (the button task is the same as described above). Random dots may move in different directions. Conditions in which all points move in the same direction are described as ‘random dots moving in unison’. When all points move in random directions, the term we use is ‘incoherent movements of random dots’. MT+ neurons have their preferences for the direction of movement, and different neurons are sensitive to different directions. When the stimulus with a specific movement direction is repeatedly presented, the MT+ neurons will adapt to the direction, and the neuron activity will decrease accordingly. To capture MT+ activities, and to avoid the weakening of neural activity caused by the adaptation direction, the direction of the coherent movement in this experiment was randomly generated from 0-360° and changed continuously (every 1.7 s).

This experiment is a two-factor (2×2) block design, and the experimental factors include coherence and movement speed. Coherence includes two levels, which are coherent and incoherent. Movement speed includes two levels, 10 deg/s and 5.0 deg/s. Therefore, there are four experimental conditions in the random dots motion experiment in the 3T experiment, which are coherent motion with a speed of 5.0 deg/s (hereinafter collectively referred to as "coh 5.0") and incoherent motion (hereinafter collectively referred to as

"inc 5.0"). Two additional stimuli moved at a speed of 10.0 deg/s: coherent motion (hereinafter collectively referred to as "coh 10.0" and incoherent motion (hereinafter collectively referred to as "inc 10.0"). As shown in **Figure 3.2(c)**, each run of the random dots motion experiment consists of 4 blocks. A 17-second blank screen (only a black background and a central fixation point) was presented at the beginning and end of the experiment. Each block consists of four mini-blocks and each mini-block presents one of the above-mentioned conditions. Each experimental condition was performed once (15.3 s each time), and, in order to avoid motion aftereffects, each stimulus was followed by a blank screen lasting 17 s as a baseline, so each experimental condition was presented 4 times in a run. The subjects completed 4 runs of random dot motion experiments, and each run presented a total of 523.6 s (17 s + (15.3 + 15.3) s × 16 + 17 s = 523.6 s).

MRI data acquisition

In the 3T experiment, the random dot motion stimulus consisted of a black background and white dots. The position of the white dot (size 0.14°, density 10.0 dots/deg) is randomly generated each time (for each subject and each run of experiments the dots are re-generated randomly). A red spot in the center of the visual field is used for the subject to respond to the button (the button task is the same as described above). Random dots may move in different directions. Conditions in which all points move in the same direction are described as ‘random dots moving in unison’. When all points move in random directions, the term we use is ‘incoherent movements of random dots’. MT+ neurons have their preferences for the direction of movement, and different neurons are sensitive to different directions. When the stimulus with a specific movement direction is repeatedly presented, the MT+ neurons will adapt to the direction, and the neuron activity will decrease accordingly. To capture MT+ activities, and to avoid the weakening of neural activity caused by the adaptation direction, the direction of the coherent movement in this experiment was randomly generated from 0-360° and changed continuously (every 1.7 s).

This experiment is a two-factor (2×2) block design, and the experimental factors include coherence and movement speed. Coherence includes two levels, which are coherent and incoherent. Movement speed includes two levels, 10 deg/s and 5.0 deg/s. Therefore, there are four experimental conditions in the random dots motion experiment in the 3T experiment, which are coherent motion with a speed of 5.0 deg/s (hereinafter collectively referred to as "coh 5.0") and incoherent motion (hereinafter collectively referred to as "inc 5.0"). Two additional stimuli moved at a speed of 10.0 deg/s: coherent motion (hereinafter collectively referred to as "coh 10.0" and incoherent motion (hereinafter collectively referred to as "inc 10.0"). As shown in **Figure 3.2(c)**, each run of the random dots motion experiment consists of 4 blocks. A 17-second blank screen (only a black background and a central fixation point) was presented at the beginning and end of the experiment. Each block consists of four mini-blocks and each mini-block presents one of the above-mentioned conditions. Each experimental condition was performed once (15.3 s each time), and, in order to avoid motion aftereffects, each stimulus was followed by a blank screen lasting 17 s as a baseline, so each experimental condition was presented 4 times in a run. The subjects completed 4 runs of random dot motion experiments, and each run presented a total of 523.6 s (17 s + (15.3 + 15.3) s × 16 + 17 s = 523.6 s).

3.1.2 fMRI experiment at 7T

In order to obtain higher-resolution brain imaging data, this study also used 7T magnetic resonance scanning functional imaging. The experiment based on 7T is generally coherent with the 3T experiment. The experimental stimulus and paradigm design have been improved according to the results of the 3T experiment. See below for details.

Ethics and participants

A total of 5 healthy young subjects were recruited for the 7T experiment (including 2 females). All subjects had normal or corrected-to-normal vision, no history of mental illness of the subject or family, no metal implants in the brain or body, no claustrophobia, etc. Each subject participated in the experiment voluntarily, read and signed the subject's

consent form before the experiment, and received a certain amount of remuneration after the experiment. The 7T magnetic resonance scanning experiment involved in this study has been reviewed and approved by The McGill University Health Center Research Ethics Board. Before entering the scanning room, all subjects changed their clothes in the waiting room and received experimental instructions. After the main examiner confirmed that the subjects could understand the experimental tasks correctly, the subjects could enter the MRI scanning room to start the experiment. Different from the 3T experiment, the 7T-based MRI experiment was completed in the MRI scanning room of the Montreal Neurological Institute (MNI), Canada, and all subjects were instructed in English.

Stimuli and paradigm

In order to obtain a larger field of view and better experimental results in the 7T experiment, VisuaStim goggles were used instead of a projector to present visual stimuli. The maximum screen range for the subjects to accept visual stimuli was 30 ° × 22.5 °. Since this field of view has changed compared with the 3T field of view, we adjusted the visual stimuli in the 7T experiment accordingly.

Retinotopic mapping of V1 in the lower right and lower left quadrant. In Muckli's study on the ventral visual pathway in 2015, the decoding analysis was performed in V1 corresponding to the lower-right quadrant visual field [60]. However, the reason for choosing this quadrant was not explained in detail in the study. This aroused our curiosity about the decoding performance of V1 corresponding to other quadrant views. Compared with the upper receptive field, the global coherent motion response is stronger in the lower receptive field (Maruyama et al., 2009). To avoid a long experimental time, we only studied the lower-left and lower-right.

As shown in **Figure 3.3 (a) and (b)**, during the 7T experiment checkerboard stimuli in the lower left quadrant and lower right quadrant were used to locate the upper right and left sides of area V1 of each subject. Except for the checkerboard stimulus, the rest of the screen presented a gray background and a central fixation point. To adapt to the change

of screen size, the lower right quadrant or lower left quadrant was evenly divided into 25×18 squares along the horizontal and vertical directions (keep the side length of each square about 0.6° as in the 3T experiment). Likewise, the experiment included both the target area and the surrounding area. The target area is composed of 22×15 squares, and the distance between the surrounding area and the target area, and separately, the surrounding area and the central fixation is one row and one column of squares.

Each subject underwent one run of V1 retinotopic mapping experiments in the lower left quadrant and one run of V1 retinotopic mapping experiments in the lower right quadrant. A blank screen (only gray background and fixation point) was presented at the beginning (20 s) and end (12.5 s) of each run as the baseline, the rest of the trials consisted of 12 alternating presentations of the ambient condition and the target condition for 12.5 s per trial. All subjects received checkerboard stimuli in the same order, and each run of V1 retinotopic mapping experiment lasted 332.5 s (20s + (12.5 + 12.5) s x 12 + 12 s = 332.5 s).

MT+ functional localizer. The stimuli of the MT+ functional localizer experiment in the 7T experiment remained the same as in the 3T experiment, and the experiment duration was adjusted according to the scan time (TR). As shown in **Figure 3.3(c)**, all subjects viewed a fixed stimulus presentation sequence, that is, the static condition and the moving condition were presented alternately; each run of experiments presented 13 static conditions and 12 moving dots conditions. Each run of experiments began with 22.5 s of static dots condition, followed by alternating motion and static dots conditions (15 s each). In the 7T experiment, each subject performed 1 run of MT+ functional localizer experiment, a total of 382.5 s (22.5 s + (15+15) s × 12 = 382.5 s).

Random dots. In order to better study the effect of speed on the coherence of MT+ and V1 coded motion, the 7T experiment adopts a two-factor (2 × 4) block design, and the experimental factors include coherence and speed of motion. Movements are classified into coherent and incoherent movements, each presented at four different speeds (2.5, 5.0,

7.5, 10.0 deg/s). Therefore, there is a total of 8 experimental stimulus conditions in the 7T random dots experiment, including coherent movement (hereinafter collectively referred to as "coh 2.5") and incoherent movement (hereinafter collectively referred to as "inc 2.5") with a speed of 2.5 deg/s, and a speed of 5.0 deg/s. deg/s of coherent motion (hereinafter collectively referred to as "coh 5.0") and incoherent motion (hereinafter collectively referred to as "inc 5.0"), 7.5 deg/s of coherent motion (hereinafter collectively referred to as "coh 7.5") and incoherent motion (hereinafter collectively referred to as " inc 7.5") and both coherent motion (hereinafter collectively referred to as "coh 10.0") and incoherent motion (hereinafter collectively referred to as "inc 10.0") at a speed of 10.0 deg/s. To match the shorter TR (1.25 s) in the 7T experiment, the direction of motion was changed every 1 s in the coherent motion stimulus, and the direction of each presentation was randomly generated from 0-360°.

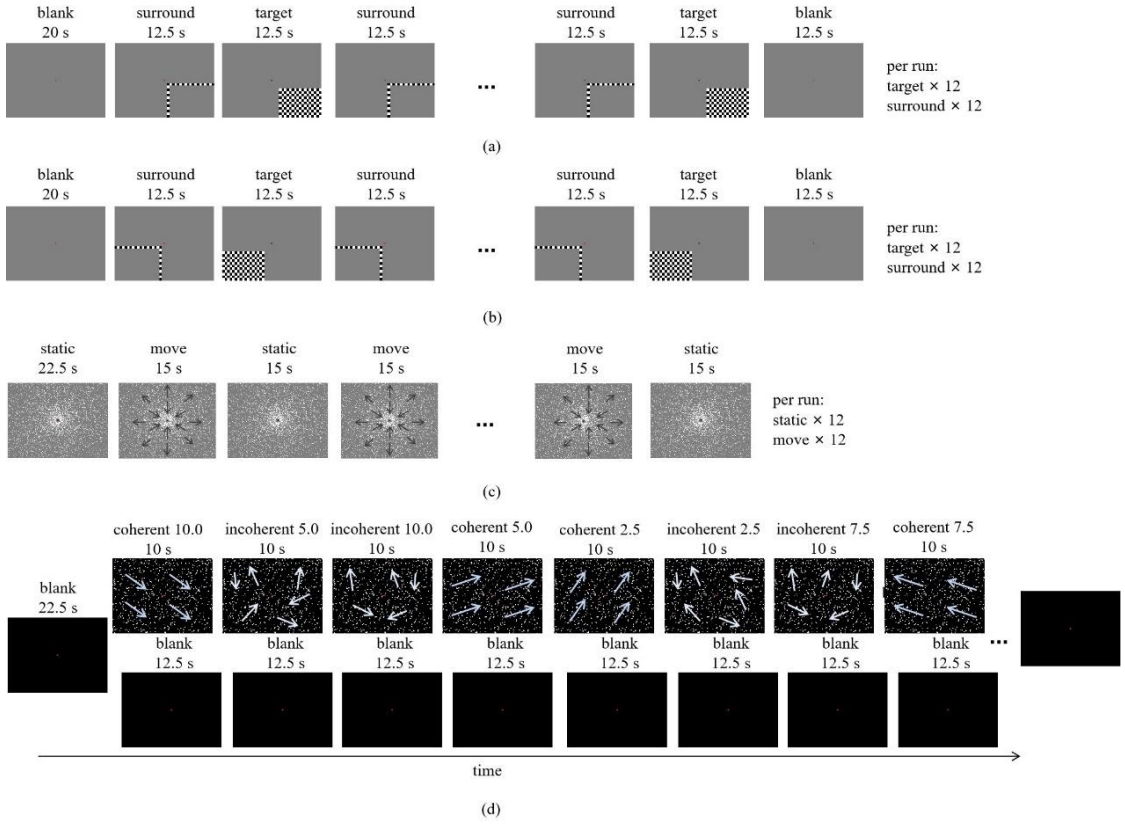


Figure 3.3 Stimuli and paradigm in the 7T experiment. (a) and (b) show the retinotopic mapping of V1 in the lower right and lower left quadrants, individually. (c) shows MT+ functional localizer and (d) shows random moving dots.

In the 7T experiment, as shown in **Figure 3.3 (d)**, each run of random dots experiment consists of two blocks, and a blank screen is presented at the beginning (22.5 s) and end (7.5 s) respectively (only black background and red fixation point). Each block consists of 8 mini-blocks, one of the 8 conditions is presented in each mini-block and each stimulus (10 s) is accompanied by a 12.5-second blank screen as a baseline, so each experimental condition is presented 2 times in each run. Each subject completed 8 runs of this experiment, and each run took 390 s (22.5 s+(10+12.5) ×16+7.5 s=390 s).

MRI data acquisition

The brain imaging acquisition device for this magnetic resonance experiment is the SIEMENS MAGNETOM Investigational scanner (7 Tesla) in the MNI Magnetic Resonance Scanning Laboratory, which uses a 32-channel phased array head coil. Structural images were acquired using a T1-weighted MPRAGE sequence scan with the following parameters: Sagittal imaging, TR = 2600 ms, TE = 2.01 ms, flip angle = 4°, FOV = 256×256 mm2, scan matrix = 256×256, with a voxel size of 1×1×1 mm^3 and a total of 192 slices. 2D-EPI scan sequence was used for task state functional image acquisition, and the parameters were as follows: no spacing scan (no spacing), coronal (Coronal) imaging, TR = 1250 ms, TE = 21 ms, flip angle = 70°, FOV = 240× 240 mm2, field of view phase (FOV phase) 60%, scan matrix = 200×200, voxel size 1.2×1.2×1.2 mm^3, slice thickness = 1.2 mm, slice scanning, 48 slices in total.

3.2 Data analysis

3.2.1 Uni-variate analysis

In this study, fMRI was used to obtain brain functional imaging of different subjects, and all subsequent data analyzes were completed based on these fMRI data. Due to the long of MRI data acquisition, each subject may have different degrees of head motion. All data was pre-processed before analysis. The pre-processing of 3T and 7T data is introduced below.

Pre-processing

Data acquired based on 3T MRI were preprocessed using the SPM12 toolbox (Friston et al., 1994). The acquisition during the initial 10 s at the beginning of each run yields T1w contrast and not T2* contrast because the RF excitation did not reach equilibrium. Therefore, the first 10 s of data in each run were removed before preprocessing (TR=2s, so the data in the first 5 volumes were deleted), and these data were not used in data analysis. Since the functional images are scanned at intervals, considering the difference between slices or the error caused by head motion, slice timing is performed first, then head motion correction is performed, and then the structural image and all functional images are combined. We registered the structural image to the average image of all functional images to complete the alignment between the structural image and functional image within the subject, and the registered functional image and structural image were aligned to the standard MNI space (MNI space). The normalized/resampled functional image resolution is 3×3×3 mm^3, and the structural image resolution is 1×1×1 mm^3. The average head movement of all subjects was less than 3 mm, so they were included in the analysis range, and all analyzes were based on pre-processed data.

The fMRI data collected at the 7T scanner is partial brain imaging, which is different from the fMRI data based on 3T, which includes the whole brain. SPM12 presented problems in the registration between individual structural images and partial functional images, so AFNI (Cox, 1996) was used for preprocessing of 7T data. The TR used in the data acquisition based on 7T is relatively short (1.25 s), and slice acquisition time correction was not adopted to avoid unnecessary calculation errors. The 7T MRI data preprocessing first uses AFNI to correct the head movement of each subject's data, then registers the individual functional images to the structural images, and finally uses SPM12 to standardize the registered functional images and structural images to the MNI space. Subsequent analyzes were performed based on the preprocessed data.

GLM analysis

The preprocessed data were analyzed by General Linear Model (GLM) using SPM12. In the analysis of 3T-based fMRI data, three independent models of V1 retinotopic mapping, MT+ functional localizer, and random moving dots in the lower right quadrant were calculated and obtained. The first two models were used to extract individual V1 and MT+ ROIs (for subsequent SVM classification), and the last model was used for uni-variate analysis, i.e., observing the level of response of V1 and MT to each experimental condition. In the analysis of fMRI data based on 7T, four independent models of V1 retinotopic mapping in the lower left quadrant, V1 retinotopic mapping in the lower right quadrant, MT+ functional localizer, and random dots motion were calculated and obtained. The first three models were used to extract Individual V1 and MT+ ROIs (for subsequent SVM classification), and the last model for uni-variate analysis, i.e., within-subject (first-level) and between-subject (second-level) activation analysis.

For the V1 retinotopic mapping model, the three conditions of the target, surround, and blank screen (baseline) are used as the regression factors of interest. In the MT+ functional localizer model, the two conditions of moving and static dots are used as regression factors of interest. In the random dots motion model constructed based on 3T data, the coherent motion with a speed of 5.0 deg/s (hereinafter referred to as coh 5.0), the coherent motion with a speed of 10.0 deg/s (hereinafter referred to as coh 10.0), the speed of 5.0 deg/s incoherent motion (hereinafter referred to as inc 5.0), incoherent motion with a speed of 10.0 deg/s (hereinafter referred to as inc 10.0), and the blank condition (baseline) were used as the regression factors of interest. The average response was calculated for all trials within the ROI.

In the random dots motion model constructed based on 7T data, the coherent motion at speeds of 2.5 deg/s, 5.0 deg/s, 7.5 deg/s, and 10.0 deg/s and the incoherent motion at these four speeds, and the post-stimulus blank screen (baseline) were used for a regressor of interest for a total of 9 conditions, and the responses corresponding to all trials within

each run for each condition were averaged. In all of the above computational models, the 6 columns of head motion parameters were considered as uninteresting regressors.

In the GLM first-level analysis, activation contrasts between conditions are calculated separately for the above models. In the V1 retinotopic mapping model, in order to obtain the V1 activation area corresponding to the lower left quadrant (in the 7T experiment) and the lower right quadrant (in the 3T and 7T experiments), the target area minus the surrounding area (target-surround), the target area minus baseline (target-baseline), and the surrounding area minus the baseline (surround-baseline) are used to take the intersection of the above three activation results (T target>surround ∩ T target>baseline) to get the corresponding lower left quadrant or right quadrant. The mask of V1 in the lower quadrant is used as the region of interest V1 in subsequent analysis.

In the MT+ functional localizer model, to obtain the MT+ activation area on both sides of the brain, the motion condition minus the static condition is computed (moving-static), and the mask of each subject's MT+ area is obtained according to the comparison of the activation result. The voxel data are used for subsequent analysis of the region of interest MT+. In the random dots motion model, to explore the encoding of visual motion coherence and motion speed information by visual areas, the activation contrast (coherent-incoherent) and the difference between coherent motion and incoherent motion were calculated respectively. We contrasted movement across speeds, as well as activations for coherent and incoherent movements at each speed and differences in activations between individual speeds for the two motion modes. The analysis based on the 3T data calculated activation contrasts at 5.0 deg/s and 10.0 deg/s speeds for coherent and incoherent motion and for both speeds during coherent and incoherent movements. In the analysis based on 7T data, the activation contrasts between coherent and incoherent movements at speeds of 2.5, 5.0, 7.5, and 10.0 deg/s, and the activation contrasts between pairs of four speeds were calculated.

ROI definitions

Based on the uni-variate analysis in this chapter, the individual analysis and group level analysis are first carried out in the whole brain (7T only collects part of the brain, that is, in all voxels of the collected images). Secondly, in order to better study the two brain regions of interest, V1 and MT+, and also test the robust results of this study, three different methods are used to localize them, and values within ROIs are extracted in uni-variated analysis to explore their responses to different experimental stimuli and also decoding performance of each ROI is illustrated in the SVM classification. The ROI size obtained by each method is shown in **Table 3.1** and **Table 3.2**.

ROI	localizer		atlas∩localizer		atlas	
	Mean	SD	Mean	SD	Mean	SD
lh_V1	173	54.76	115	143.69	465	50.47
rh_V1	-	-	-	-	553	43.89
lr_V1	-	-	-	-	1013	84.52
lh_MT+	99	37.61	86	29.91	910	130.13
rh_MT+	110	48.69	94	34.12	891	49.11
lr_MT+	209	81.61	180	59.98	891	49.11

Table 3.1 Number of voxels of ROIs (3T, N=12). lh refers to left hemisphere, rh refers to right hemisphere and lr refers to left and right hemispheres are combined. Note: For V1, only the left side of each subject in 3T experiment was retinotopically mapped.

ROI	localizer		atlas∩localizer		atlas	
	Mean	SD	Mean	SD	Mean	SD
lh_V1	346	271.56	234	150.58	6772	854.41
rh_V1	217	220.74	160	130.10	8802	961.63
lr_V1	551	501.16	390	281.91	15268	1728.50
lh_MT+	296	102.52	287	98.56	10785	1870.31
rh_MT+	166	78.32	159	72.04	10625	1371.89
lr_MT+	456	165.06	440	154.37	21410	3082.25

Table 3.2 Number of voxels of ROIs (7T, N=5). lh refers to left hemisphere, rh refers to right hemisphere and lr refers to left and right hemispheres are combined.

In order to reduce the influence of individual differences, the first method obtains the masks of V1 and MT+ through functional localizer, hereinafter referred to as "localizer". Left hemisphere V1 (hereafter collectively referred to as "lh_V1") or right hemisphere

V1 (hereafter collectively referred to as "rh_V1") corresponding to the lower right quadrant (in 3T and 7T experiments) and lower left quadrant (in 7T experiments only) using the activations localized by the V1 retinotopic mapping in the experiment), and the individual left MT+ (hereinafter collectively referred to as "lh_MT+") and right MT+ (hereinafter collectively referred to as "rh_MT+") obtained through MT+ functional localizer, and through the imcalc tool of SPM12, the left brain and right brain masks are combined to obtain a double side V1 (hereinafter collectively referred to as "lr_V1" or bilateral MT+ (hereinafter collectively referred to as "lr_MT+"). It should be noted that in the 3T V1 retinotopic mapping experiment, a localizer stimulus was presented only to the lower right quadrant, so there is no mask of the right V1 of each subject. In the localizer results, the mask of right brain V1 was obtained based on the visual atlas in the subsequent analysis.

The second method is to obtain the location of the brain region of interest based on the public atlas ("atlas" is used below to indicate that the atlas is used to obtain the ROI). The atlas used was developed based on the cortical (surface) space. Therefore, the structural image that is preprocessed and registered with the individual functional image is used for the cortical reconstruction (surface reconstruction) of the subject. This process uses the recon-all command in FreeSurfer, and then each atlas of the fsaverage space is mapped to the individual space to extract the brain area of interest. The delineation of V1 combines the two atlases of Benson (2014) and Wang (2015), that is, the extraction of V1 of the Benson atlas and the extraction of the Wang atlas V1d, V1v and take the union as the V1 of the subject. The unilateral MT+ of the subject is composed of TO1 and TO2 (i.e., MT and MST as (Amano et al., 2009) reported) of the Benson atlas, the hMT of the Goebel atlas and the hMT+ of the Huang atlas by taking the union of the labels extracted respectively. All the labels obtained according to the atlas are mapped to the volume space of the individual using the mris_surf2vol command of FreeSurfer. The method of extracting the mask through the atlas is called "atlas" in this study.

The third method is to take the intersection of the masks obtained by the above two methods, and the brain area of interest obtained by this method is called "atlas∩localizer". The mask of the brain region of interest obtained by the above three methods are intersected with the brain of the corresponding subject to obtain the voxel positions used to extract the value of β in the SVM classification.

3.2.2 Multivariate analysis

ROI definitions

In the multivariate pattern analysis based on SVM classification, the same three methods of extracting ROI as in the uni-variate analysis were adopted to obtain the brain regions of interest MT+ and V1, that is, based on the functional localizer and the atlases mentioned above. The union (atlas) and the intersection of the two (atlas∩localizer) were used to obtain unilateral and bilateral MT and V1, and extract the β value for SVM classification research.

SVM classification

The multivariate model analysis in this study adopts the SVM classification method. The general linear modeling introduced in the previous chapter is calculated based on the average of all trials within a single run, which not only loses information about the activity pattern caused by each experimental condition but also may lose the activity information of a single trial. Therefore, in order to obtain better samples for training and testing of the SVM classification model, this study first constructs a GLM model of fMRI responses to random dots motion based on a single trial, so as to obtain the activation values of all voxels in the brain region of interest (the β value corresponding to each trial is used). Then each β data is mapped to a vector, and finally a linear SVM classifier is constructed through CosMoMVPA (Oosterhof et al., 2016) and libSVM toolbox for training and testing, and the classification accuracy is calculated. The performance of each brain region of interest in decoding experimental stimuli was viewed. This study uses the leave one-sample out method to cross-validate the SVM

model, that is, each time one of the samples of each condition (a trial β value) is used as the test data, and all the remaining samples are used for training. The average of the results of the model is used as the classification accuracy.

In this study, binary classification calculations were performed on the data of the 3T experiment, including the binary classification of coherent and incoherent motion stimuli (hereinafter referred to as "decoding of coherence") and the binary classification of two motion speeds (5.0 and 10.0 deg/s) (Hereinafter referred to as "decoding of speed"). The former includes the classification of coherence with and without distinction between speeds, and the latter includes the classification of speeds without distinction between coherent and incoherent movements. A similar classification analysis was performed on the data from the 7T experiment, except that there were four speed levels (2.5, 5.0, 7.5, and 10.0 deg/s) in the 7T experiment, so four classifications were used when decoding speed.

When the influence of speed on decoding coherence is not considered, the random dots motion trials of coh 5.0 and coh 10.0 are calculated as samples of coherent motion conditions. Similarly, the trials of inc 5.0 and inc 10.0 are both regarded as samples incoherent motion conditions. Therefore, in the 3T experiment, there are 8 trials for each run of coherent movement and incoherent movement, and each subject completes 4 runs, that is, 32 samples for each of the coherent movement and incoherent movement conditions. In the 7T experiment, coherent motion and incoherent motion are presented in 8 trials in each run, and each subject completes 8 runs, that is, 64 samples of coherent movement or incoherent movement conditions. One sample of each of the two conditions is used for testing, and the remaining trial is used for training the model. After repeating 64 times, the average of the classification accuracy is taken. In addition, coherent and incoherent movements are classified at each speed. In the 3T experiment, the coherent movement and incoherent movement at a speed of 5.0 deg/s and 10.0 deg/s were presented in 4 trials for each run, a total of 4 runs, 16 samples for each condition, repeated 16 times. In the 7T experiment, taking 2.5 deg/s as an example, there are 2 trials

for coherent motion and incoherent motion, and there are 16 samples in 8 runs. Two types of motion stimuli are classified, repeated 16 times and the average value is taken as the accuracy measure. The same procedure was applied 5.0, 7.5, and 10.0 deg/s.

When the influence of coherence on the decoding speed is not considered, in the 3T experiment, the trials of coh 5.0 and inc 5.0 are both regarded as samples with a speed of 5.0 deg/s, and the same is true for 10.0 deg/s; in the 7T experiment, coh 2.5 and inc 2.5 conditions are regarded as 2.5 deg/s samples, and the same is true for the other three speeds. When not distinguishing between coherent motion and incoherent motion in the 3T experiment, the speed was divided into two categories, 5.0 and 10.0 deg/s each with 32 samples (8x4); in the 7T experiment, the speed was divided into four categories, with a total of 32 samples for each speed (4x8). Additionally, speeds were classified separately under the coherent and incoherent motion conditions. In the 3T experiment, 16 samples (4x4) of 5.0 and 10.0 deg/s were used under each of coherent and incoherent motion conditions. The speed was classified into two categories, and the accuracy of the classification speed was obtained by repeating 16 times of cross-validation; the 7T experiment included 16 samples (2x8) for each of the four speeds under each of the coherent and incoherent motion conditions. The speed is classified into four categories, and the SVM classification accuracy is obtained through 16 leave-one-out cross-validation.

4 Results of uni-variate analysis

4.1 Results of the experiment at 3T

4.1.1 Retinotopic mapping of V1 in the lower right quadrant

First level GLM analysis was performed before group level GLM analysis. Aa shown in **Figure 4.1 (a)**, the results of the second-level GLM analysis of all subjects show that the target condition minus the surround condition (target-surround) elicits activation position with a greater response to the target area (expressed as positive activation). Activations

that respond with higher amplitude to peripheral areas (represented as negative activations) are located more anteriorly (parietal direction). Stimulation of the target region further from the center of the visual field elicited more anterior activation. Voxels that respond to the target area (as shown in **Figure 4.1 (b)**) and respond to the target area with higher amplitude than to the surrounding area (T target>surround∩ T target>baseline) are selected as the region of interest V1 for subsequent analysis. The above activation result threshold is p<0.05 (cluster-level FWE correction).

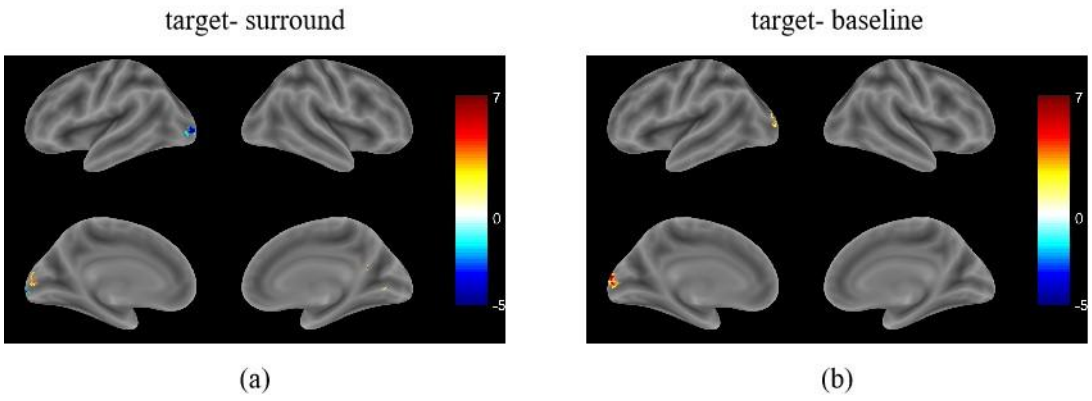


Figure 4.1 Group-level results of retinotopic mapping of V1 in the lower right quadrant (3T). (a) The contrast of 'targer-surrund' and (b) 'target-baseline'. Threshold is p<0.05 (cluster-level FWE corrected).

It should be noted that in the 3T-based V1 retinotopic mapping experiment, the original fMRI data of one of the 12 subjects was irreparably damaged due to an unknown device problem, so in this second-order GLM analysis only the data of the remaining 11 subjects was used as the ROI.

4.1.2 MT+ functional localizer

MT+ functional localizer selects voxels that are more responsive to motion conditions than to static conditions as the region of interest. In the first-level GLM analysis at the subject level, MT+ showed individual differences in location, but in the second-order analysis, MT+ still showed significant positive activation (as shown in **Figure 4.2**, uncorrected p<0.001).

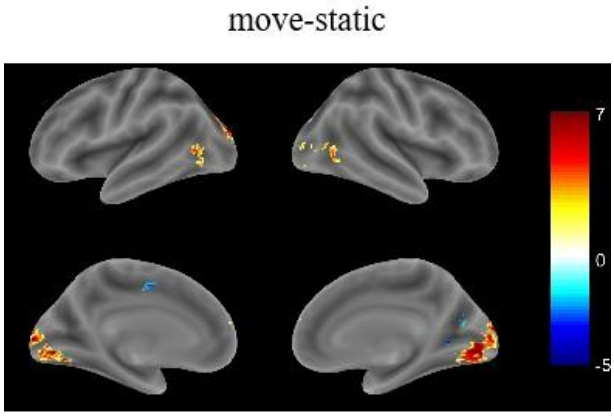


Figure 4.2 Group-level results of MT+ functional localizer by the contrast of 'move-static' (3T). The threshold is p<0.001, uncorrected.

4.1.3 Responses of MT+ and V1 to moving random dots

In the 3T-based random dots stimulus experiment, firstly, the activation comparison map of the whole brain to different experimental stimuli was calculated at the group level (second level, N=12), and then the three methods of extracting ROI described above were used in the brain region of interest. MT+ and V1 were observed for their responses to coherent and incoherent motion and to different motion speeds (β average of all voxels in the ROI).

Without considering the speed as shown in **Figure 4.3 (a)**, the activation comparison results of coherent motion minus incoherent motion (coh-inc) showed significant positive activation for bilateral MT+; for bilateral V1, significant negative activation was observed (p<0.05, corrected for cluster level FWE). The same result was observed at a speed of 10.0 deg/s (coh 10.0-inc 10.0), that is, the bilateral MT+ responded more strongly to coherent movements and the bilateral V1 responded more strongly to incoherent movements (**Figure 4.3 (c)**, p<0.05, cluster level FWE correction). When the speed was 5.0 deg/s, as shown in **Figure 4.3 (d)** (p<0.05, cluster level FWE correction), the bilateral V1 also showed significant negative activation, while only the left side in MT+ showed positive activation. The right side showed no difference in response to coherent and incoherent movements.

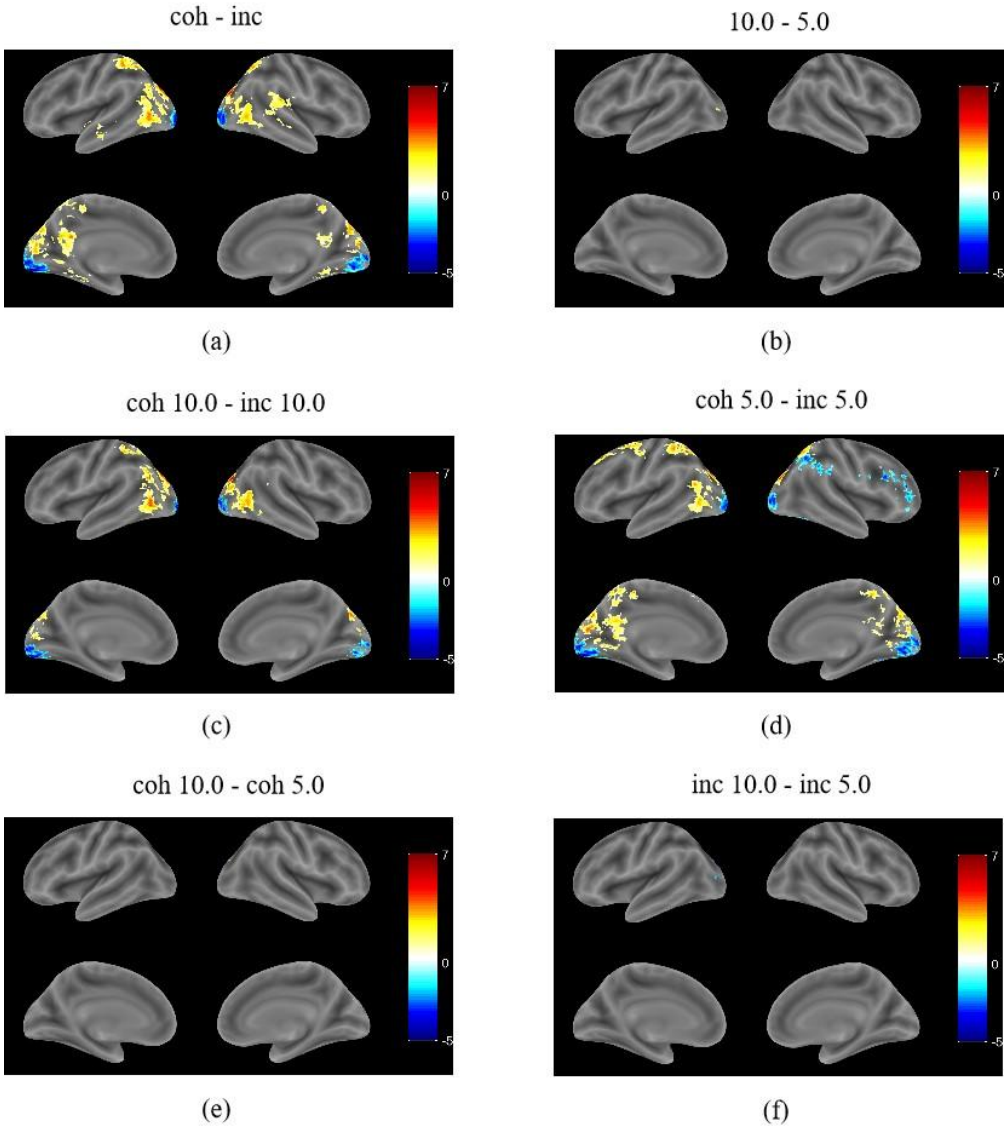


Figure 4.3 Group-level responses to random moving dots(3T). (a) the contrast of 'coh-inc' regardless of speed (p<0.05, cluster-level FWE corrected). (b) the responses to random dots moving at 10.0 deg/s minus at 5.0 deg/s (p<0.001, uncorrected). (c) and (d) show the responses to 'coh-inc' at 10.0 deg/s and 5.0 deg/s (p<0.05, cluster-level FWE corrected). (e) and (f) show the responses to '10.0-5.0' when random dots moving coherently and incoherently (p<0.001, uncorrected).

In the ROI-based analysis, when speed was not differentiated, bilateral MT+ responses to coherent motion were observed to be greater than those of incoherent motion (atlas: t=2.147, p=0.055, atlas∩localizer: t=4.955, p<0.001, see **Figure 4.4(a)**). Since the right V1 was not located in the 3T experiment, in order to better compare the activation levels of the two brain regions, only the activation of the left MT+ on coherent and incoherent movements is shown here (see **Figure 4.4(a)** first column). The response of the left MT+

to the coherent movement was significantly larger than that to the incoherent movement (t=4.714, p=0.001). V1 showed the opposite trend to MT+, i.e., a stronger response to incoherent movements (as shown in **Figure 4.4(b)**). Significantly larger response margins to incoherent motion than to coherent motion (t=2.113, p=0.058) were observed in the left V1 based on localizer extraction, a difference that was not significant in the other two analyses. In the three analyses, the two-factor repeated measures ANOVA (ROIxcoherence) of brain area and coherence identified an interaction effect between brain area and coherence (localizer: F=56.093, p<0.001; atlas: F=8.999, p=0.012; atlas∩localizer: F=37.595, p<0.001).

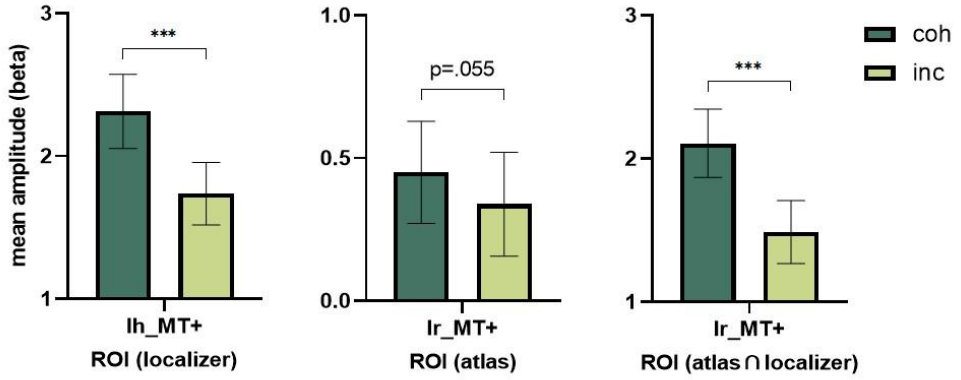
The response of unilateral MT+ to coherent and incoherent movements without distinguishing speed is illustrated in **Figure 4.5**. In the localizer-based analysis, repeated measures ANOVA results of MT+ hemisphere (left or right) and coherence (coherent or incoherent) showed that the main effect of MT+ hemisphere was significant (F=11.227, p=0.006), coherence being the main effect of is significant (F=25.387, p<0.001), and the interaction effect of the two is significant (F=17.612, p=0.001). Left MT+ showed a greater response than right MT+. Left MT+ (t=4.552, p=0.001) and right MT+ (t=5.260, p<0.001) all had a stronger response to coherent movement. In the atlas-based analysis, similarly, the main effect of MT+ hemisphere was significant (F=9.468, p=0.011), the main effect of coherence was marginally significant (F=4.708, p=0.053), and the interaction effect of the two was significant (F=7.870, p=0.017). T-test results showed that the response of left MT+ was significantly greater than that of the right MT+ for both the coherent condition (t=2.763, p=0.018) and incoherent condition (t=3.342, p=0.007). For both left (marginally significant, t=1.860, p=0.007) and right (t=2.406, p=0.035) sides, MT+ always showed a stronger response to coherent condition than to incoherent condition. In the analysis based on atlas∩localizer, the repeated measures ANOVA results of MT+ hemisphere and coherence showed that the main effect of MT+ hemisphere was significant (F=11.676, p=0.006), and the main effect of coherence was significant (F=24.185, p<0.001), the interaction effect between the two is significant (F=13.116, p=0.004). The results of the simple effect t-test showed that the response of

left MT+ was also significantly greater than that of the right MT+ in coherent condition (t=3.152, p=0.009) and incoherent condition (t=3.670, p=0.004), and both left MT+(t=4.442, p=0.001) and right MT+(t=5.260, p<0.001) responded with a higher amplitude to coherent movement than to incoherent movement.

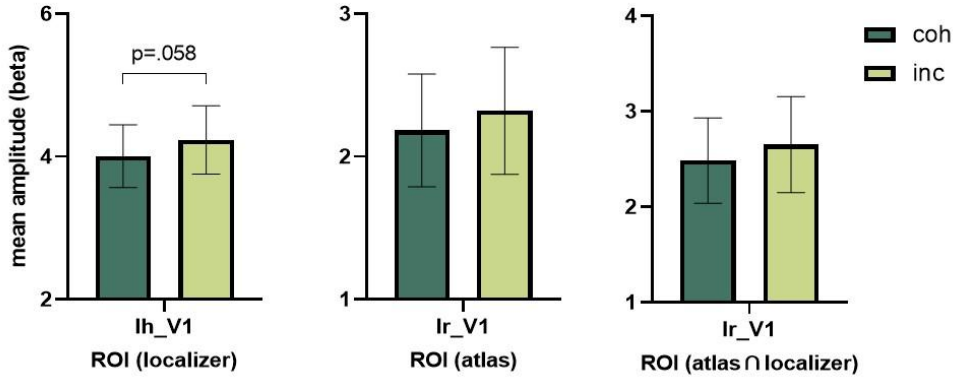
Since the 3T experiment only located the left V1, it is impossible to compare the response of the left and right sides. Here, only the left and right sides of V1 were compared based on the atlas. As shown in **Figure 4.6**, no differences (p>0.05) in response to the two movement modes were observed in either the left V1 or the right V1, and neither in the coherent nor incoherent movement conditions. There were no difference in responses between the left and right sides (p>0.05).

Regardless of coherence, as shown in **Figure 4.3 (b)**, neither MT+ nor V1 showed a significant difference in response to 5.0 and 10.0 deg/s at the group level (p<0.001, uncorrected). There was no significant difference in the responses of MT+ and V1 to random dots motion moving in the two speeds (5.0 and 10.0 deg/s) independent of whether the dots moved coherently (see **Figure 4.3 (e)** for the contrast of 'coh 10.0 - coh 5.0') or incoherently (see **Figure 4.3 (f)** for the contrast of 'inc 10.0-inc 5.0').

In the ROI-based analysis, as shown in **Figure 4.7 (a)**, significant differences in the activation of bilateral MT+ for the two speeds were observed only in the atlas ∩ localizer-based analysis, that is, for 5.0 deg/s. The response to random dots motion at 5.0 deg/s was stronger than at a speed of 10.0 deg/s (t=2.225, p=0.048), while the bilateral V1 responded almost equally to both speeds. In the localizer-based analysis, only the activation of the left MT+ and left V1 is compared. As shown in the first column of **Figure 4.7 (a)**, the paired t-test results indicate that the left MT+ showed a trend of stronger responses to dot movement at 5.0 deg/s than at 10.0 deg/s (t=1.969, p=0.075), while the left V1 showed an opposite activation where the responses were stronger at 10.0 deg/s than at 5.0 deg/s (not significant, t=2.819, p=0.17).



(a)



(b)

Figure 4.4 Responses to coherent and incoherent random dot motion in MT+ and V1(3T). (a) Mean amplitude of MT+. (b) Mean amplitude of V1. Left, middle and right column show the results within ROI generated by localizer, atlas, and atlas ∩ localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. *** indicates pair t-test $p<0.001$. Error bars indicate SEM.

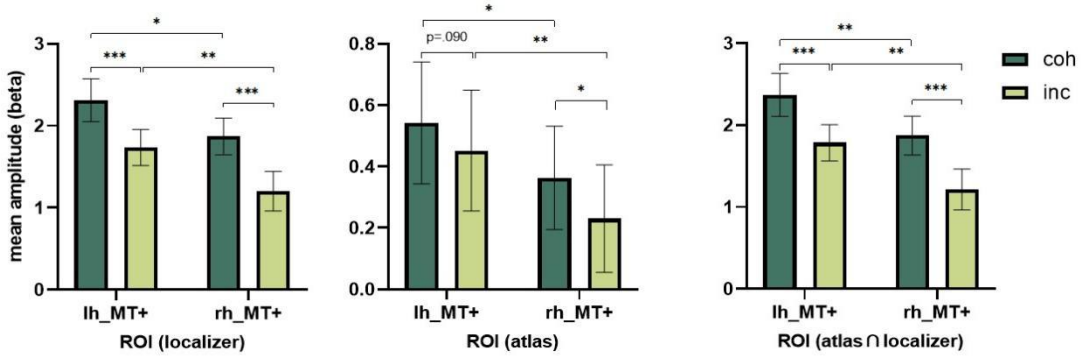


Figure 4.5 Responses to coherent and incoherent random dot motion in left or right MT+(3T). The left, middle, and right columns show the results within ROIs generated by the localizer, atlas, and atlas ∩ localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. * indicates pair t-test $p<0.05$, ** $p<0.01$ and *** $p<0.001$. Error bars indicate SEM.

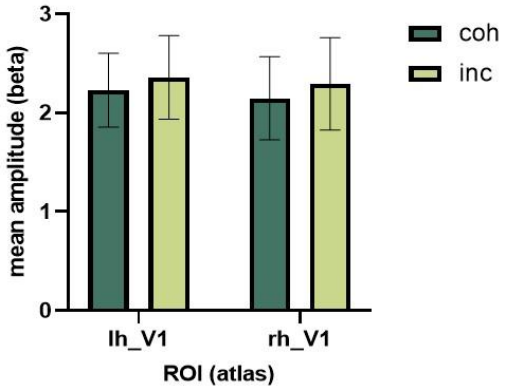


Figure 4.6 Responses to coherent and incoherent random dot motion in left or right V1 (3T). ROI is generated by atlas only. Dark green indicates coherent and light green indicates incoherent motion. Error bars indicate SEM.

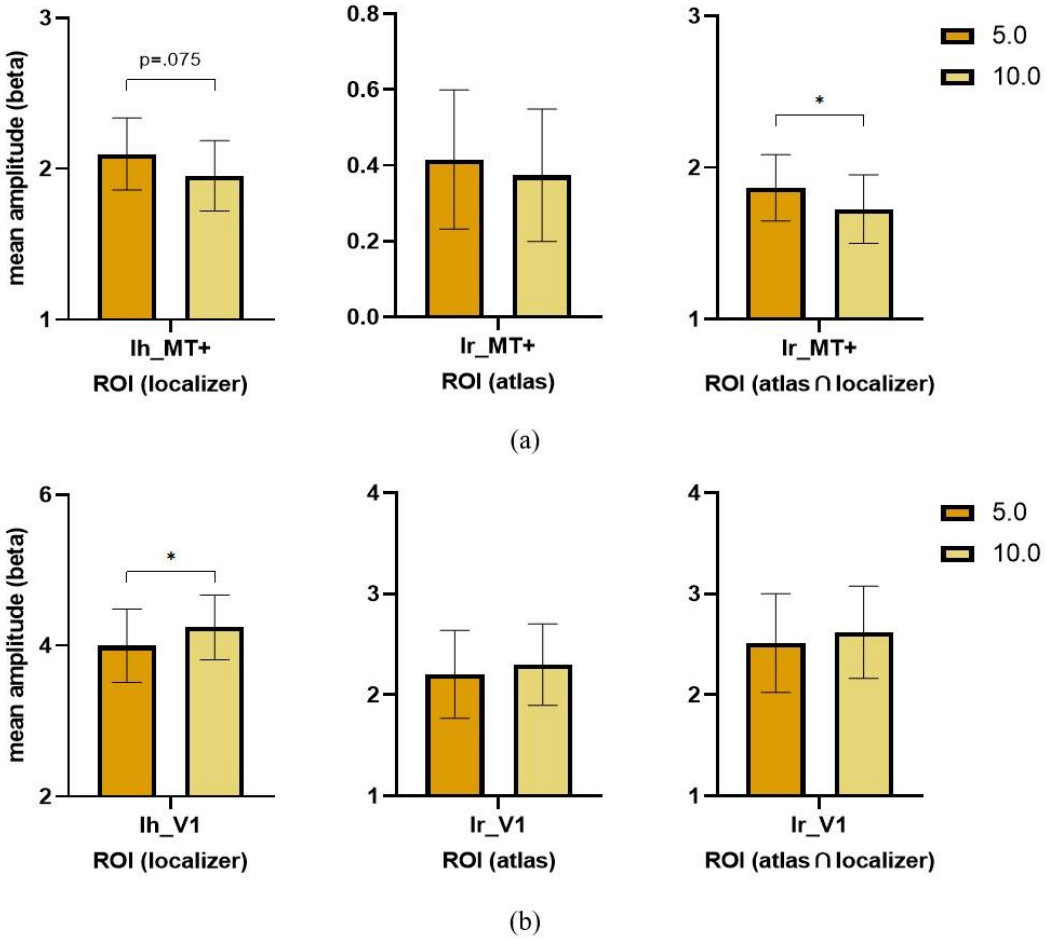


Figure 4.7 Responses to random dot motion at 5.0 and 10.0 deg/s in MT+ and V1 (3T). (a) mean amplitude of MT+. (b) mean amplitude of V1. The left, middle, and right columns show the results within ROI generated by the localizer, atlas, and atlas∩localizer, respectively. Orange indicates 5.0 deg/s and yellow indicates 10.0 deg/s. * indicates pair t-test $p<0.05$. Error bars indicate SEM.

The response of unilateral MT+ to two speeds of movement without distinguishing coherence is shown in **Figure 4.8**. The results based on atlas∩localizer demonstrate that the left MT+ showed a trend (t=1.968, *p*=0.075) for higher responses to moving random dots at 5.0 deg/s than to 10.0 deg/s. The right MT+ showed (t=2.347, *p*=0.039) a statistically significant similar difference. The analysis based on 'localizer' found a similar result, that is, the left MT+ (marginally significant, t=1.969, *p*=0.075) and the right MT+ (significant, t=2.385, *p*=0.036) responded with a higher amplitude to random dot motion of 5.0 deg/s than to 10.0 deg/s. In addition, regardless of the speed of 5.0 or 10.0 deg/s random dots motion stimulation, the response of the left MT+ under the three ROI extraction methods was significantly greater than that of the right MT+. Under the condition of 5.0 deg/s, the difference between the two is significant (localizer: t=3.289, *p*=0.007; atlas: t=3.207, *p*=0.008; atlas∩localizer: t=3.393, *p*=0.006), at 10.0 deg/s, the left MT+ and right MT+ also showed a significant difference in response to random dots motion (localizer: t=3.384, *p*=0.006; atlas: t=2.914, *p*=0.014; atlas∩localizer: t =3.409, *p*=0.006). **Figure 4.9** shows the response of unilateral V1 to two motion speeds without distinguishing coherence. Here, only the unilateral V1 obtained based on atlas is compared. Only the left V1 showed a trend of significant difference (p=0.088) in response to the two speeds of random dots motion stimuli, whereas no such significant difference was observed for the right V1. Additionally, no left-right hemisphere differences were observed in V1 in response to random dots motion responses of either 5.0 (t=0.542, *p*=0.598) or 10.0 (t=0.671, *p*=0.516) deg/s.

Finally, the interaction of coherence and speed was explored (using 2x2 repeated measures ANOVA). **Figure 4.10** demonstrates the bilateral MT+ response to random dots motion. In the analysis results based on the localizer, it was observed that the main effect of coherence was significant (F=25.860, *p*<0.001), the main effect of speed was significant (F=5.047, *p*=0.046), and the interaction between the two was significant (F=52.214, *p*<0.001). At speeds of 5.0 (t=3.465, *p*=0.005) and 10.0 (t=6.514, *p*<0.001) deg/s, bilateral MT+ responses to coherent movements were significantly greater than those to incoherent movements, and the bilateral MT+ response to random dots motion

stimulation at 5.0 deg/s was significantly greater than that at 10.0 deg/s (t=4.554, *p*=0.001). In the atlas-based analysis results, it was observed that the main effect of coherence was marginally significant (F=4.610, *p*=0.055), and the interaction between coherence and speed showed a trend toward significance (F=4.303, *p*=0.062). At a speed of 10.0 deg/s, there was a significant difference in the responses of bilateral MT+ coherent and incoherent movements, with the former being greater than the latter (t=2.474, *p*=0.031). There was a marginally significant difference in bilateral MT+ responses to incoherent movement at different speeds during incoherent movement (t=2.044, *p*=0.066). In the analysis results based on the atlas∩localizer, it was observed that the main effect of coherence was significant (F=24.556, *p*<0.001), the main effect of speed was significant (F=4.951, *p*=0.048), and the interaction between the two was significant (F=45.855, *p*<0.001). At speeds of 5.0 (t=3.419, *p*=0.006) and 10.0 (t=6.250, *p*<0.001) deg/s, both bilateral MT+ showed a significantly greater response to coherent than incoherent movements, and responses to incoherent motion at both speeds also showed a significant difference, i.e., the response to incoherent motion was greater at 5.0 deg/s than at 10.0 deg/s (t=4.392, *p*=0.001).

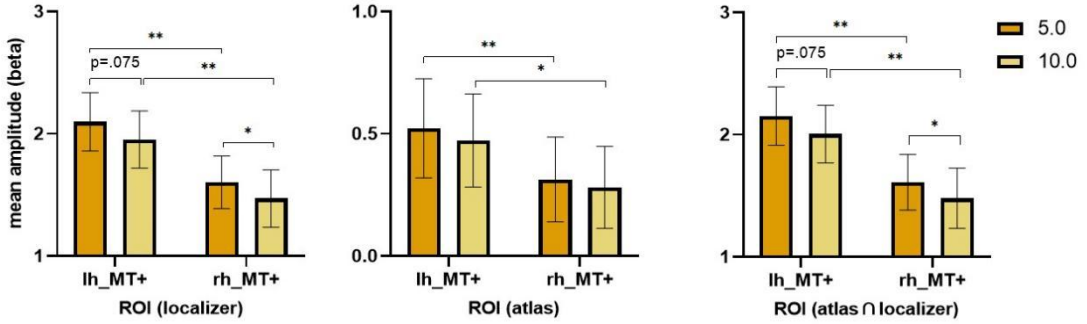


Figure 4.8 Responses to random dot motion at 5.0 and 10.0 deg/s in left or right MT+(3T). The left, middle, and right columns show the results within ROIs generated by the localizer, atlas, and atlas∩ localizer, respectively. Orange indicates 5.0 deg/s and yellow indicates 10.0 deg/s. * indicates pair t-test *p*<0.05, ***p*<0.01. Error bar indicated SEM.

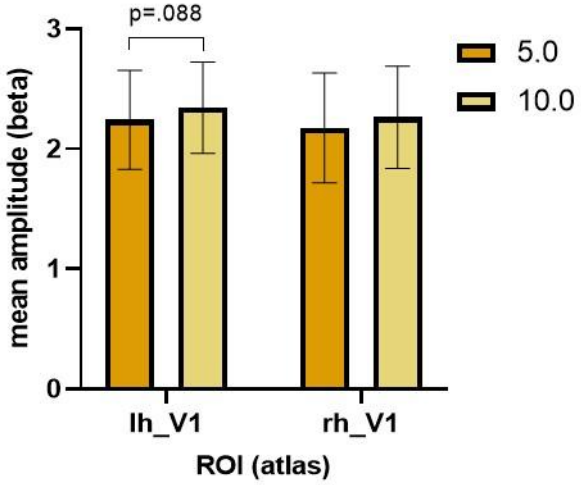


Figure 4.9 Responses to random dot motion at 5.0 and 10.0 deg/s in left or right V1 (3T). ROI is generated by atlas only. Orange indicates 5.0 deg/s and yellow indicates 10.0 deg/s. Error bars indicate SEM.

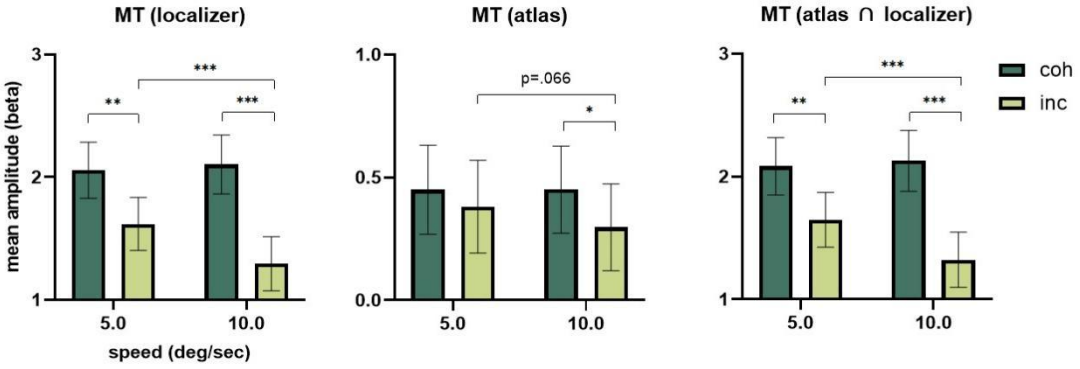


Figure 4.10 Interaction of coherence and speed in MT+ (3T). The left, middle and right columns show the results within ROIs generated by the localizer, atlas, and atlas∩localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. * indicates pair t-test p<0.05, ** p<0.01 and *** p<0.001. Error bars indicate SEM.

The response of bilateral V1 to random dot motion is shown in **Figure 4.11**. Since there is no available data of the right V1 for the 'localizer' analysis, only the results based on the 'atlas' and 'atlas∩localizer' are shown in **Figure 4.11**. Both methods did not find any significant main effect of coherence, main effect of speed or interaction effect of the two factors (p>0.05). Only in the atlas∩localizer-based analysis the observed bilateral V1 response to coherent motion exhibited marginally significant differences at different

speeds, i.e., 10.0 deg/s elicited a stronger bilateral V1 response than 5.0 deg/s during coherent motion time (t=1.840, *p*=0.093).

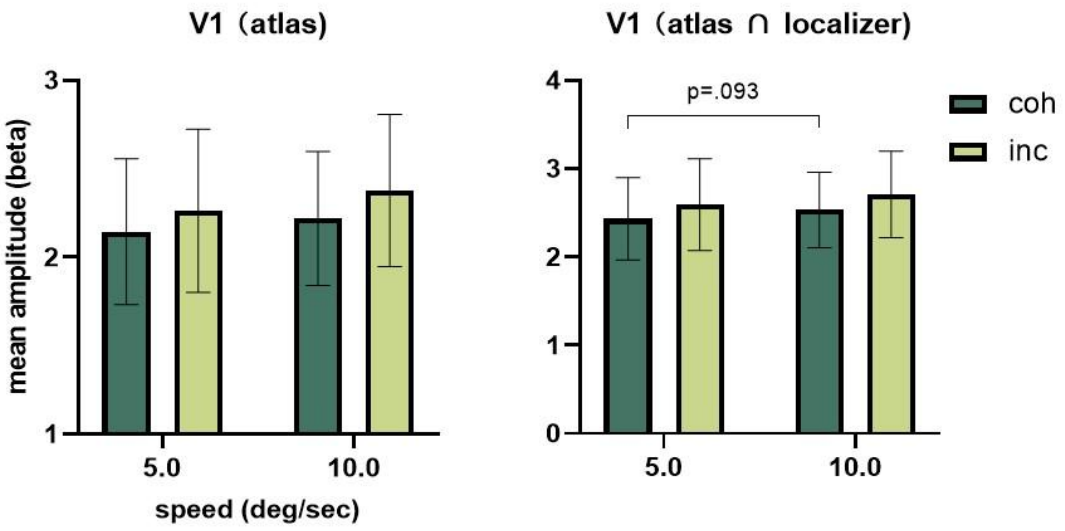


Figure 4.11 Interaction of coherence and speed in V1 (3T). The left, middle and right columns show the results within ROIs generated by the atlas and atlas∩localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. Error bars indicate SEM.

In conclusion, in the 3T experiments, both unilateral (left or right side) and bilateral (left and right sides together) MT+ consistently exhibited greater responses to coherent than incoherent movements regardless of speed, and bilateral MT+ always responded greater to coherent motion at both 5.0 and 10.0 deg/s (except for the results based on 'the atlas' analysis), which is consistent with the results of the contrast maps. V1 showed a trend that the response to incoherent motion was higher than that of coherent motion (the difference was not significant), and the difference was still not significant when analyzing each speed alone, which may be caused by large individual differences. Neither MT+ nor V1 showed a significant difference in the two-speed conditions for coherent movement, while the MT+ response to incoherent movement was higher at 5.0 deg/s than at 10.0 deg/s (this difference was not reflected in the whole brain activation map possibly due to individual differences in the location of MT+). For V1, speed had no significant effect on encoding coherent or incoherent movements, which is consistent with the results shown in the whole-brain activation maps. In addition, the response of

the left MT+ was significantly higher than that of the right MT+ in both coherent and incoherent movement conditions, and V1 did not show such left and right hemisphere differences.

4.2 Results of the experiment at 7T

4.2.1 Retinotopic mapping of V1 in the lower right and lower left quadrant

In the 7T-based V1 retinotopic mapping, the activation comparison between the target area and the surrounding area was first calculated for each subject on an individual level. In the 7T-based experiment, first-order analysis results for all five subjects found areas that responded more to the target area than to the surrounding area. Second-order analysis at the group level (5 subjects) showed that both V1 retinotopic mapping stimuli could localize subjects' V1 corresponding to the lower left and lower right quadrants. In the V1 retinotopic mapping experiment in the lower right quadrant, the response to the target area in the left hemisphere was greater than the baseline level (see **Figure 4.12 (b)**). The brain's region responding with a higher amplitude to the target condition was located in the posteriorily to the region responding with a higher amplitude to the surround condition (see **Figure 4.12 (a)**), which agrees with the principle of V1 retinotopic mapping. In addition, in the V1 retinotopic mapping experiment in the lower left quadrant, the response of the right brain to the target area was also higher than the baseline level (as shown in **Figure 4.12 (d)**), and the activation position was the similar (respectively in different hemispheres) to that in the V1 retinotopic mapping experiment in the lower right quadrant. When calculating the activation comparison between the target area and the surrounding area, positive activation results are obtained at the group level, and the activation position (see **Figure 4.12 (c)**) is also close to the positive activation position in **Figure 4.12 (a)**. The activation result thresholds above are all $p<0.05$ (corrected by cluster-level FWE).

The left V1 (corresponding to the lower right quadrant field of view) and the right V1 (corresponding to the lower left quadrant field of view) obtained by retinotopic mapping

are obtained by taking the intersection method, that is, selecting voxels that respond to the target area and have greater activation than the surrounding area (T target>surround ∩ T target>baseline). The V1 obtained by retinotopic mapping in the subsequent analysis is based on the respective masks of each subject obtained above.

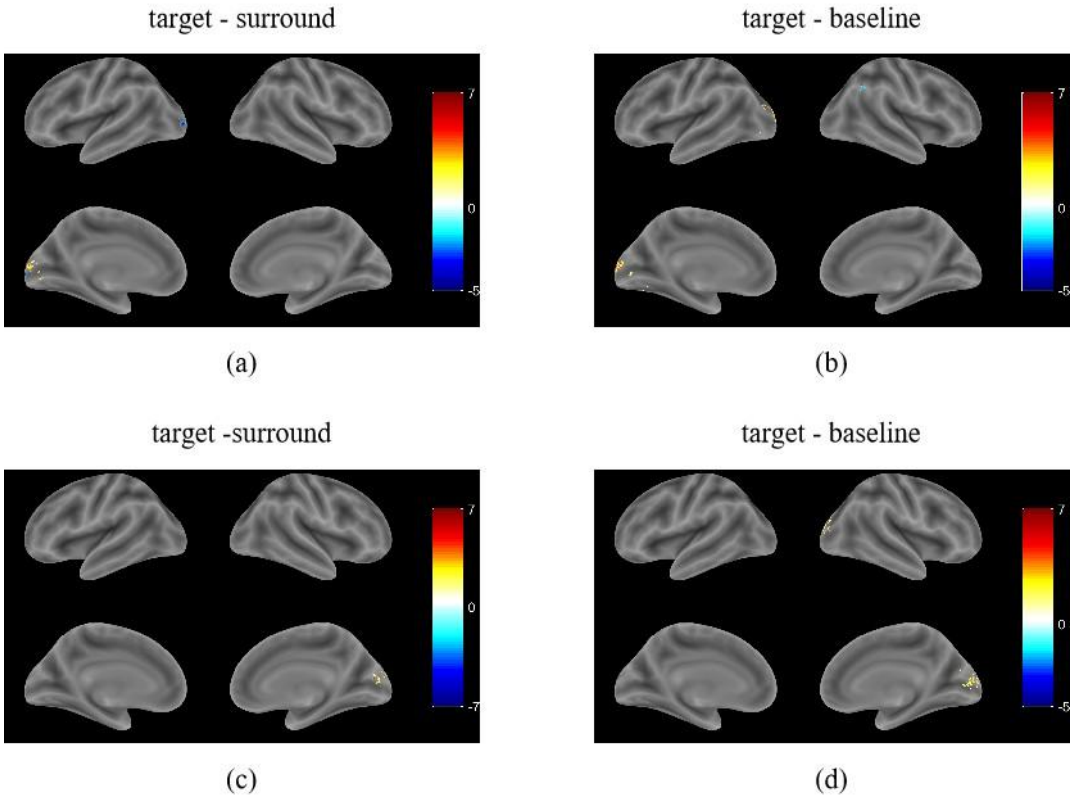


Figure 4.12 Retinotopic mapping of V1 in the lower right (top row) and lower left (bottom row) quadrant at the group level (7T). (a) and (c) show the contrast of 'target-surrun'. (b) and (d) show the responses to the target region that is larger than baseline. Threshold is $p<0.05$, cluster-level FWE corrected.

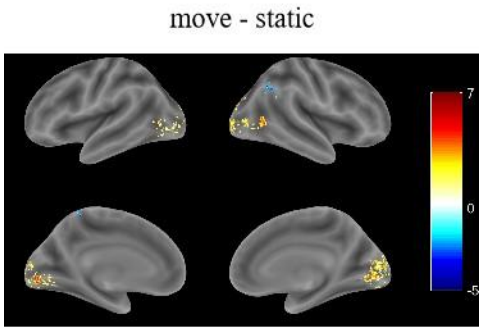


Figure 4.13 MT+ functional localizer at the group level (7T). Threshold is $p<0.05$, cluster-level FWE corrected.

4.2.2 MT+ functional localizer

In the 7T-based MT+ functional localizer experiment, MT+ was defined as the group of voxels that responded with a higher amplitude to the motion condition than to the rest condition, as shown in **Figure 4.13**. By calculating the activation contrast of the motion condition and the rest condition, the group level (5 subjects) bilateral MT+ activation (positive activation), the threshold was p<0.05 (cluster level FWE correction). Voxels activated by each subject in MT+ functional localizer were taken from the left and right hemispheres respectively and used as the mask of their left or right MT+ for subsequent analysis.

4.2.3 Responses of MT+ and V1 to moving random dots

In the random dots exercise experiment based on 7T, the group-level (N=5) whole brain activation contrast maps were calculated for all subjects, and three different ROI extraction methods were used to observe the MT+ and V1 pairs of coherent movements, incoherent movements, and response to different speeds of movement (β).

When the speed is not distinguished, the activation comparison results of coherent movement and incoherent movement (coh-inc, see **Figure 4.14 (a)**) show significant negative activation of bilateral V1 (p<0.05, cluster level FWE correction). When the speed is 10.0 deg/s, as shown in **Figure 4.14 (b)**, a significantly stronger bilateral V1 response to incoherent motion was also observed (coh 10.0-inc 10.0, p<0.05, cluster level FWE corrected). However, no similar negative activation of V1 was observed at 2.5, 5.0 or 7.5 deg/s (**Figure 4.14 (b)-(d)**). In either case, no significant differences in MT+ responses to the two stimuli were observed. In addition, no significant differences were observed in the responses of V1 or MT+ to different speeds, regardless of whether coherence was distinguished or when coherent and incoherent movements were observed separately (**Figure 4.14 (f)** shows the results of the F-test for the effect).

In the ROI-based analysis, as shown in **Figure 4.15 (a)**, it was observed that bilateral MT+ responses to coherent movements were significantly greater than those to incoherent

movements when speed was not differentiated, and bilateral MT+ also showed a higher mean response to coherent movements in atlas-based analyzes but was not significantly different than responses to incoherent movements. As shown in **Figure 4.15 (b)**, under the three methods, the response of bilateral V1 to incoherent motion is always greater than to coherent motion (localizer: marginally significant, t=2.731, p=0.052; atlas: significant, t=3.323, p=0.032; atlas∩localizer: significant, t=3.847, p=0.018). A significant interaction effect of brain region and coherence was observed in all three analyses (localizer: F=23.429, p=0.008; atlas: F=16.093, p=0.016; atlas∩localizer: F=38.358, p=0.003) And the main effect of the brain area was significant (localizer: F=108.855, p<0.001; atlas: F=48.756, p=0.002; atlas∩localizer: F=100.765, p=0.001).

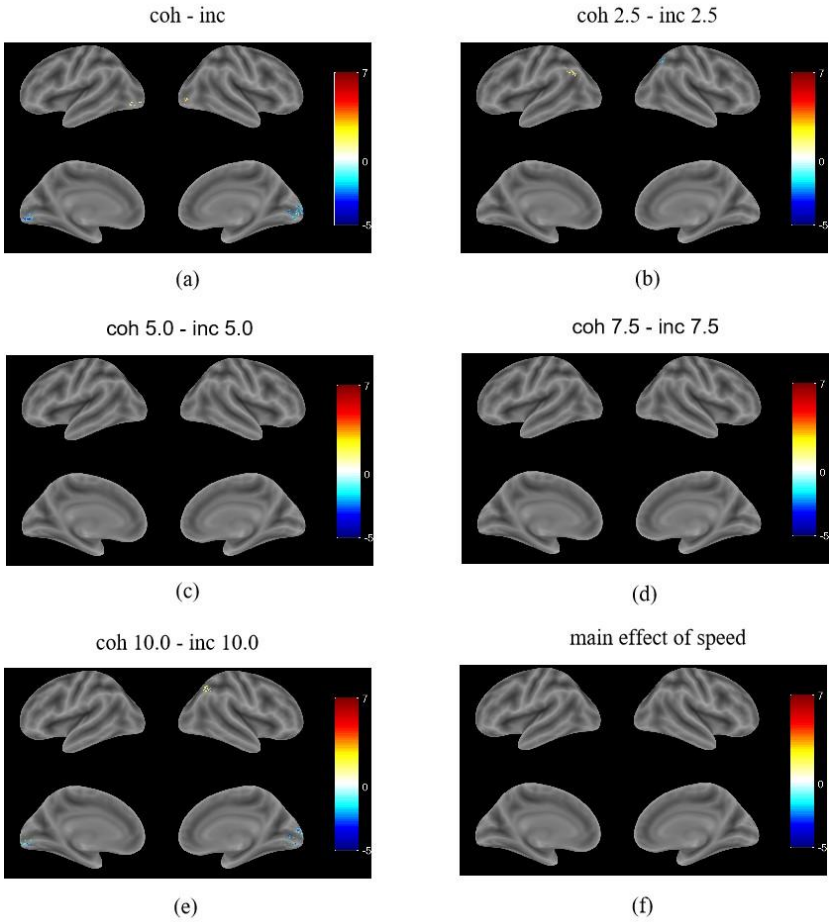


Figure 4.14 Responses to random moving dots (7T). The contrasts of coherent and incoherent motion are calculated (a) regardless of speed, and at (b) 2.5 deg/s, (c) 5.0 deg/s, (d) 7.5 deg/s, and (e) 10.0 deg/s. Thresholds for (a), (b), and (e) are p<0.05 (cluster-level FWE corrected) and p<0.001 (uncorrected) for (c) and (d). (f) shows results of the main effect of speed (p<0.001, uncorrected). Error bars indicate SEM.

Figure 4.16 shows the response of unilateral MT+ to coherent and incoherent motion without distinguishing speed under the three ROI extraction methods. In the localizer-based analysis, both left MT+ (t=8.204, p=0.001) and right MT+ (t=3.881, p=0.018) showed significantly stronger responses to coherent movements. Likewise, left MT+ (t=8.164, p=0.001) and right MT+ (t=3.890, p=0.018) were also observed to respond significantly more to coherent movements than to incoherent movements in the atlas∩localizer-based analysis. In the atlas-based analysis, repeated measures ANOVA results for MT+ hemisphere and concordance showed a significant main effect of MT+ hemisphere (F=53.015, p=0.002). The results of the paired t-test of simple effects showed that, except for the left MT+ based on the atlas extraction showed a significantly higher response than the right (coherent movement, t=8.911, p=0.001; incoherent movement, t=5.863, p=0.004), no significant difference (p>0.05) was observed in the MT+ side responses under the other two methods.

As **Figure 4.17** shows the response of one-sided V1 to coherent and incoherent motion under the three ROI extraction methods without distinguishing the speed. Repeated measures ANOVA for V1 hemisphere location and observed marginally significant or significant main effects of coherence in all three analyses (localizer: F=6.115, p=0.069; atlas: F=10.001, p=0.034; atlas∩localizer: F=15.232, p=0.018). Only the right V1 was significantly more responsive to incoherent than coherent movement observed in the localizer-based analysis, whereas no significant difference was found in the response of the two motion stimuli to the left V1. Both left V1 (t=3.105, p=0.036) and right V1 (t=3.186, p=0.033) had significantly stronger responses to incoherent motion in the atlas-based analysis. In the localizer analysis, it was also observed that the left V1 (t=3.811, p=0.019) and the right V1 (t=3.766, p=0.020) both responded with a higher amplitude to incoherent motion conditions. No significant difference was observed between left V1 and right V1 responses to coherent or incoherent movements in all analyses (p>0.05).

Figure 4.19 shows MT+ and V1 responses to different speeds without distinguishing coherence. The speed main effect of MT+ or V1 in the three analyzes was not significant (p>0.05), and the results of paired t-tests for speed showed that in both-sided MT+ (**Figure 4.18 (a)**) or bilateral V1 (**Figure 4.18 (b)**) there was no difference in response at each speed (p>0.05).

Figure 4.20 demonstrates the response of unilateral MT+ to four speeds of motion without distinguishing coherence. No significant main effect of unilateral MT+ was observed in any of the three analyses (p>0.05), and the paired t-test results for two speeds showed that either the left MT+ or the right MT+ responded equally at each speed. No significant difference was shown (p>0.05). Except for the left MT+ extracted based on the atlas showed a significantly larger response than the right MT+ in each speed condition (2.5: t=8.362, p=0.001; 5.0: t=5.158, p=0.007; 7.5: t=7.752, p=0.001; 10.0: t=9.272, p=0.001), the other two analyses did not show differences in activation between the left and right sides of MT+ for all other speed conditions.

Figure 4.21 demonstrates the response of unilateral V1 to random dots motion at four motion speeds without distinguishing coherence. No significant main effect of V1 hemispheric position was observed in any analysis (p>0.05), and paired t-test results for two speeds showed that there was no significant difference (p>0.05). In addition, no significant differences (p>0.05) were observed in the activation levels of left V1 and right V1 in any case.

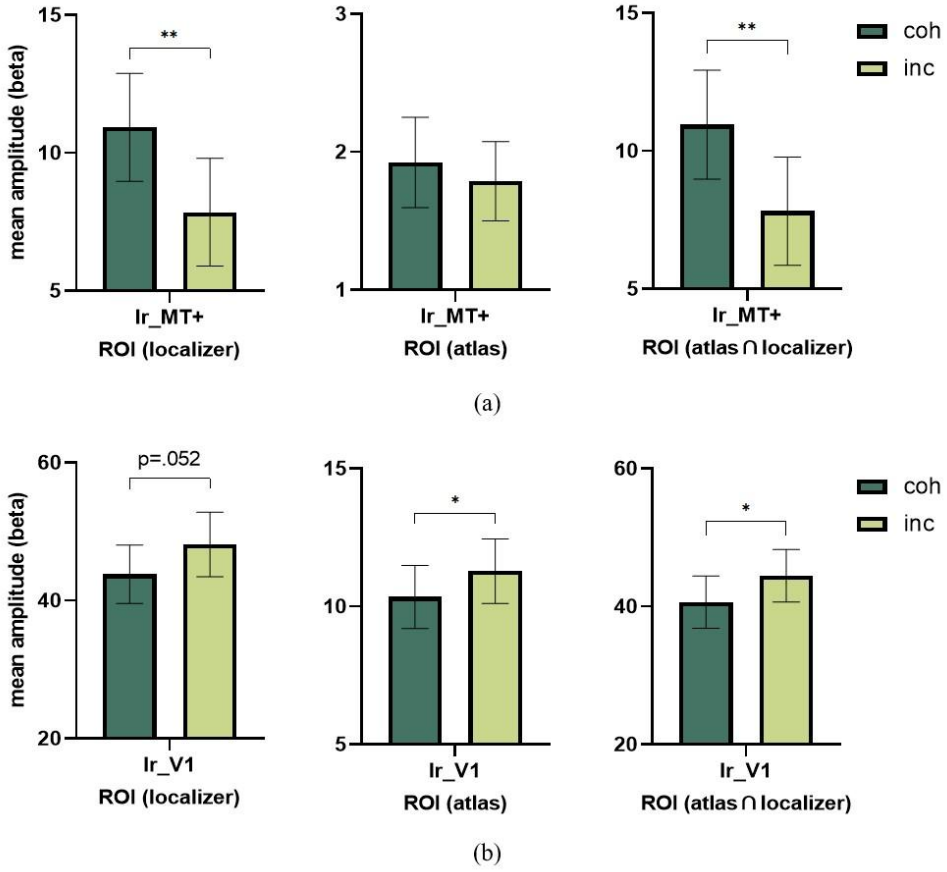


Figure 4.15 Responses to coherent and incoherent random dot motion in MT+ and V1(7T). (a) Mean amplitude of MT+. (b) Mean amplitude of V1. The left, middle, and right columns show the results within ROIs generated by the localizer, atlas, and atlas ∩ localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. * indicates pair t-test $p<0.05$, ** $p<0.01$. Error bar indicated SEM.

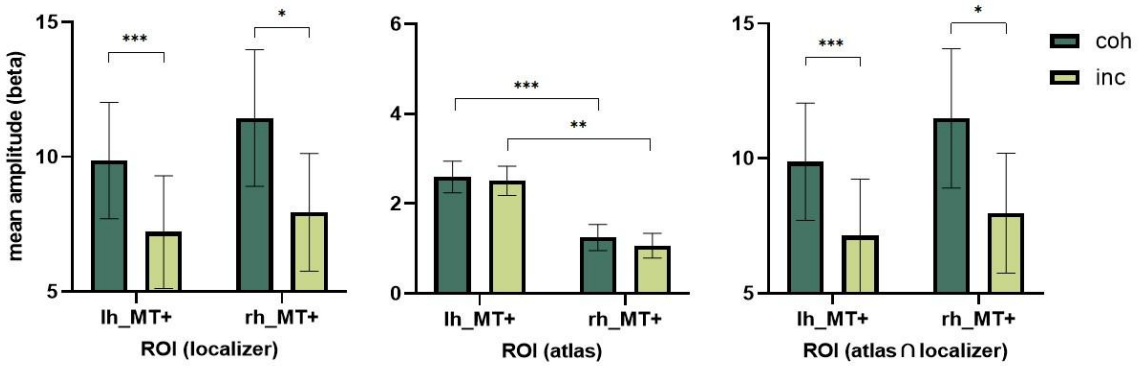


Figure 4.16 Responses to coherent and incoherent random dot motion in left and right MT+ (7T). The left, middle, and right columns show the results within ROIs generated by the localizer, atlas, and atlas ∩ localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. * indicates pair t-test $p<0.05$, ** $p<0.01$, *** $p<0.001$. Error bars indicate SEM.

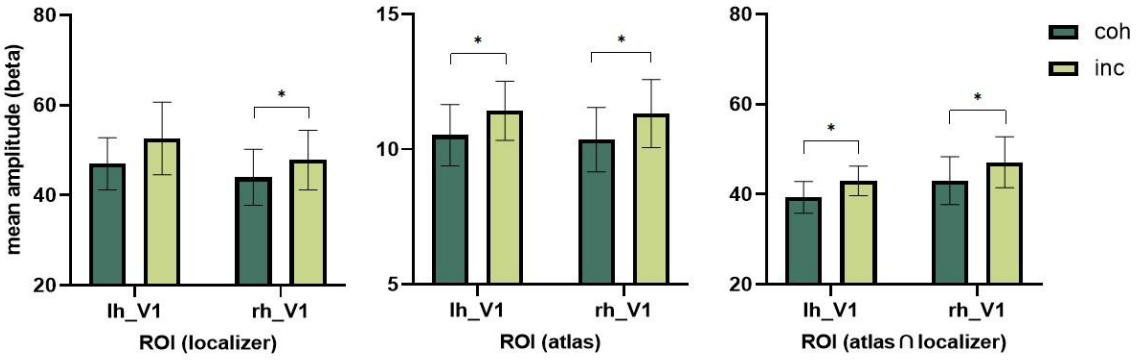
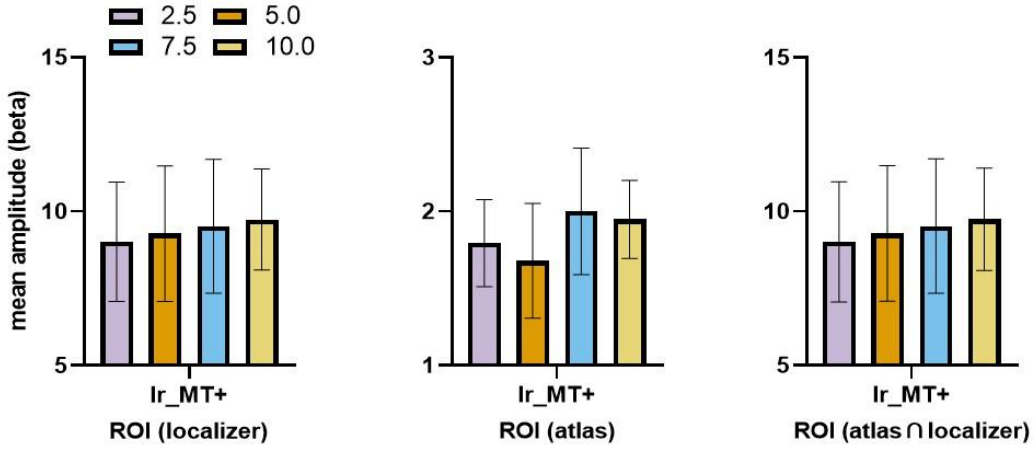
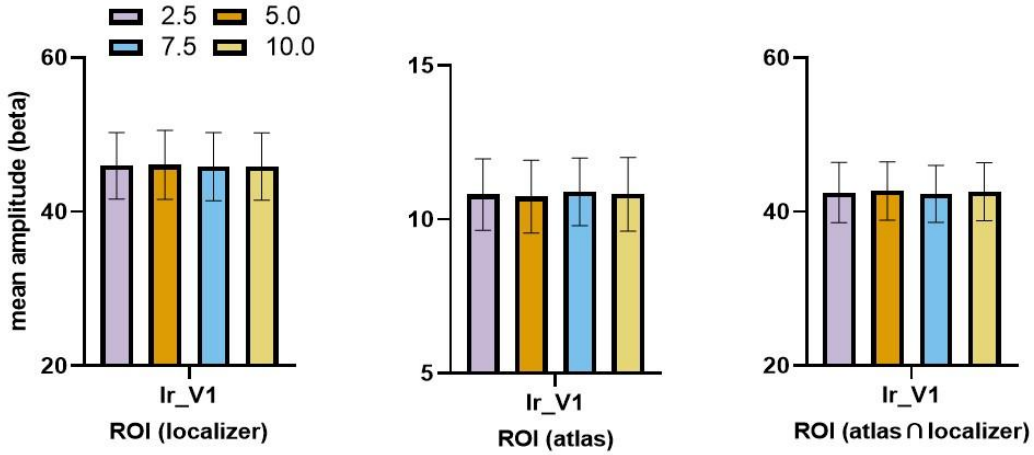


Figure 4.17 Responses to coherent and incoherent random dot motion in left and right V1 (7T). The left, middle, and right columns show the results within ROI generated by the localizer, atlas, and atlas ∩ localizer, individually. Dark green indicates coherent and light green indicates incoherent motion. * indicates pair t-test p<0.05. Error bars indicate SEM.



(a)



(b)

Figure 4.18 Responses to random moving dots at different speeds (7T). (a) Mean amplitude of MT+. (b) Mean amplitude of V1. Left, middle and right column show the results within ROIs generated by the localizer, atlas, and atlas ∩ localizer, individually. Purple, orange, blue, and yellow indicate 2.5, 5.0, 7.5, and 10.0 deg/s, respectively. Error bars indicate SEM.

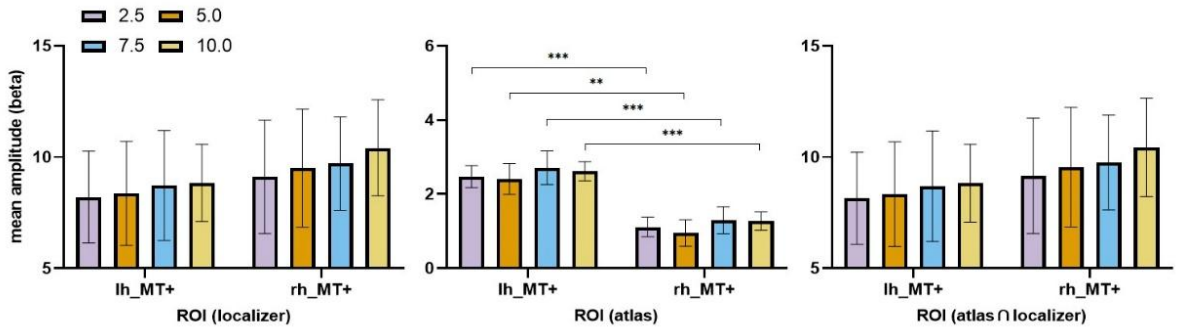


Figure 4.19 Responses to random moving dots at different speeds in left or right MT+ (7T). The left, middle, and right columns show the results within ROI generated by the localizer, atlas, and atlas ∩ localizer, respectively. Purple, orange, blue, and yellow indicate 2.5, 5.0, 7.5 and 10.0 deg/s, respectively. ** indicates pair t-test $p<0.01$ and ***$p<0.001$. Error bar indicated SEM.

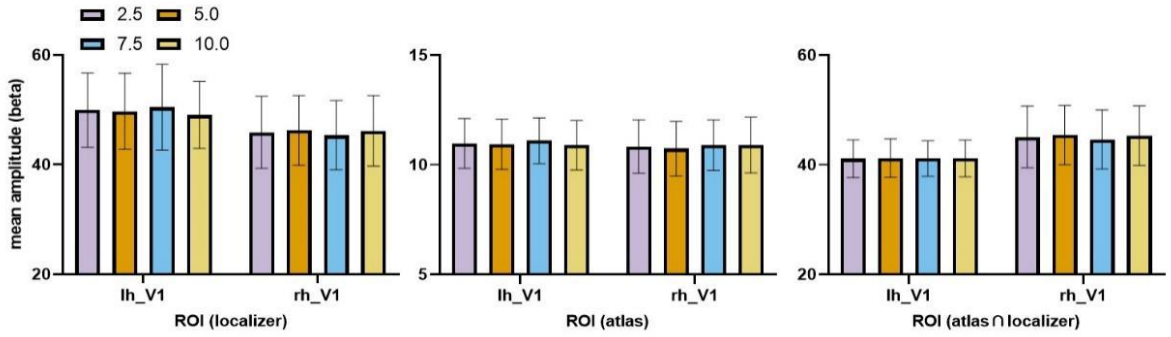


Figure 4.20 Responses to random moving dots at different speeds in left or right V1 (7T). The left, middle, and right columns show the results within ROIs generated by the localizer, atlas, and atlas ∩ localizer, respectively. Purple, orange, blue, and yellow indicate 2.5, 5.0, 7.5, and 10.0 deg/s, respectively. Results did not show statistical significance ($p>0.05$). Error bars indicate SEM.

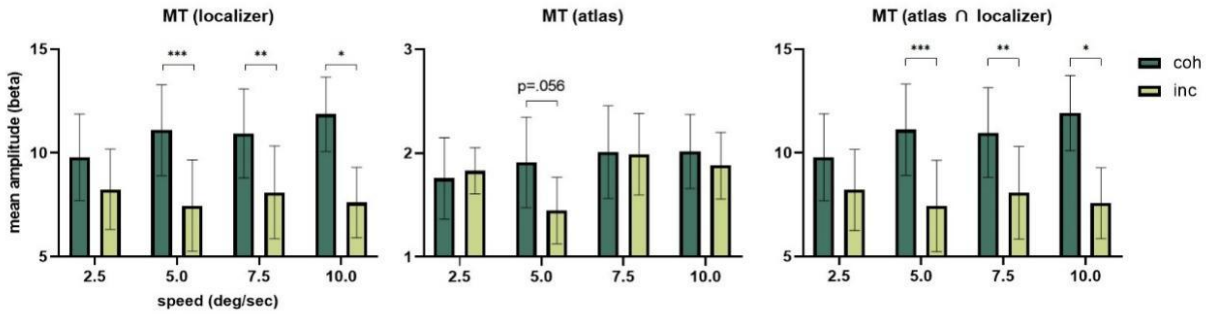


Figure 4.21 Interaction of coherence and speed in MT+ (7T). The left, middle, and right columns show the results within ROIs generated by the localizer, atlas, and the atlas ∩ localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. * indicates pair t-test $p<0.05$. **$p<0.01$, ***$p<0.001$. Error bars indicate SEM.

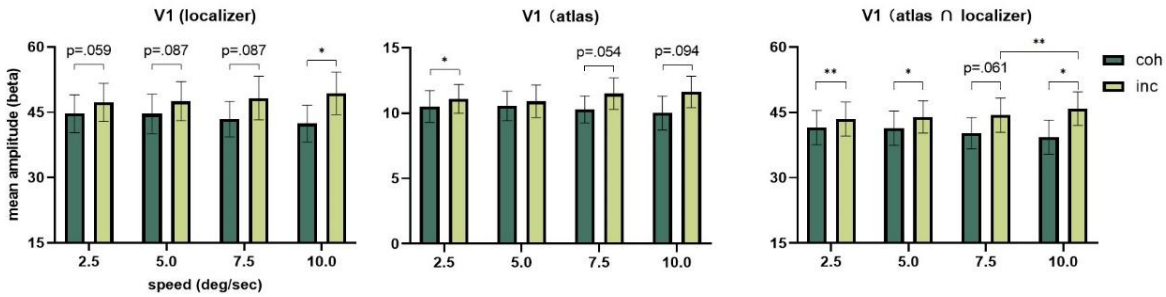


Figure 4.22 Interaction of coherence and speed in V1 (7T). The left, middle, and right columns show the results within ROIs generated by the localizer, atlas, and atlas ∩ localizer, respectively. Dark green indicates coherent and light green indicates incoherent motion. * indicates paired t-test $p<0.05$. ** $p<0.01$. Error bars indicate SEM.

Finally, the interaction of coherence and speed is explored. The repeated measures ANOVA results of coherence and speed showed that the main effect of coherence (F=47.773, p=0.002) and the interaction effect of coherence and speed were significant (F=20.155, p=0.048); the main effect of coherence in the two-sided MT+ extracted based on atlas∩localizer was significant (F=44.531, p=0.003); no significant main effect or interaction effect was observed in the two-sided MT+ extracted based on atlas ($p>0.05$). As shown in **Figure 4.21**, in the localizer-based analysis, the two-sided MT+ is 5.0 (t=12.164, $p<0.001$), 7.5 (t=5.121, p=0.007) and 10.0 (t=3.604, p=0.023) deg/s, the response to coherent motion was significantly larger than the response to incoherent motion; peripherally, in the analysis based on atlas∩localizer, this difference was between 5.0 (t=12.748, $p<0.001$), 7.5 (t=5.150, p=0.007) and 10.0 (t=3.659, p=0.022) deg/s were significant; in the atlas-based analysis, this difference was only marginally significant at 5.0 deg/s (t=2.659, p=0.056) . There were no significant differences in bilateral MT+ responses at any speed, whether during coherent or non-coordinated movements.

In addition, the bilateral V1 was also tested for the interaction effect of coherence and speed. The bilateral V1 extracted based on the localizer showed a marginally significant main effect of coherence (F=7.458, p=0.052), and the interaction of coherence and speed Significant effect (F=4.026, p=0.034); in the analysis based on atlas, the two-sided V1 only showed a significant main effect of coherence (F=10.387, p=0.032); in the analysis

based on atlas∩localizer, the two-sided V1 showed a significant main effect of coherence (F=14.799, p=0.018), a significant main effect of speed (F=26.222, p=0.037) and a significant interaction effect of the two (F=4.529, p=0.024). As shown in **Figure 4.22**, in the localizer-based analysis, the bilateral V1 responded significantly more to incoherent motion at 10.0 deg/s than coherent motion (t=2.889, p=0.045), and at 2.5 (t=2.611 , p=0.059), 5.0 (t=2.253, p=0.087) and 7.5 (t=2.254, p=0.087) deg/s showed marginally significant differences in the responses to the two movements; in the atlas-based analysis, Bilateral V1 showed significantly different responses to the two movements at 2.5 deg/s (t=3.218, p=0.032), this difference was not observed at 5.0 deg/s (p>0.05), whereas at 7.5 deg/s (t=2.704, p=0.054) and 10.0 (t=2.185, p=0.094) deg/s showed marginal significance; in the analysis based on atlas∩localizer, bilateral V1 was at 2.5 (t=7.619, p=0.002), 5.0 (t=2.836, p=0.047) and 10.0 (t=3.640, p=0.022) deg/s showed a significantly stronger response to incoherent motion, and the difference was marginally significant at 7.5 deg/s (t=2.587, p=0.061). Bilateral V1 showed a significantly stronger response to incoherent motion at 10.0 deg/s than at 7.5 deg/s in the atlas∩localizer results (p=0.023, Bonferroni correction). Responses of side V1 to coherent or incoherent movements differed significantly at different speeds.

In conclusion, no significant activation of MT+ was found in the group-level activation comparison results of the 7T random dots exercise experiment, which may be related to the small voxel size of the 7TfMRI data and the large individual differences, due to individual differences in the location and size of MT+ Larger, so no significant activation was observed at the group level. Significant negative activation of V1 was observed bilaterally only when the effect of speed was not considered and at the group level when the speed was 10.0 deg/s, suggesting a stronger V1 response to incoherent movements. The analysis results based on ROI also support the conclusion that V1 has a stronger response to incoherent motion. The response of bilateral V1 to incoherent motion is significantly stronger than the response to coherent motion under the three analyses, and bilateral MT+ is also based on the localizer and atlas∩localizer showed a significantly stronger response to coherent motion, supporting the notion that MT+ was more

responsive to coherent motion. In addition, no significant differences in MT+ or V1 responses to different speeds were observed in group-level activation contrast plots, whereas some degree of interaction between coherence and speed was observed in ROI-based analysis. In the localizer-based and atlas∩localizer-based analyses, there was a significant interaction effect of coherence and speed in bilateral V1, and in the localizer-based analysis, bilateral MT+ showed a significant interaction effect of both. Overall, neither bilateral V1 nor bilateral MT+ showed differences at different speeds, regardless of the response to coherent or incoherent movements, but MT+ responses to coherent movements showed an increasing trend with increasing speed. Bilateral V1 consistently showed main effects significant or marginally significant for coherence, and responses to incoherent motion were significant or marginally significantly larger than coherent motion at all speeds, where, when speeds were 2.5 and 10.0 deg/s The difference is more obvious when, which also shows that V1 is more sensitive to incoherent motion. A significant coherent main effect was observed for bilateral MT+ in all but atlas-based analyses, and consistently showed a significantly higher response to coherent motion at all three speeds except 2.5deg/s, again illustrating the characteristic of MT+ being more sensitive to coherent motion was eliminated. Finally, in addition to the atlas-based analysis observed that either coherent or incoherent movements, or random dot motions at speeds of 5.0 or 10.0 deg/s, resulted in greater activation of the left MT+, whereas this left-right hemisphere difference did not It is not observed in the results based on localizer or atlas∩localizer. Considering that it may be related to the large MT+ acquired by atlas, which leads to large individual differences, it indicates that the method of extracting brain regions of interest based entirely on the intersection of different visual atlases may not be reliable when the number of subjects is not huge. Overall, neither V1 nor MT+ showed significant differences in left versus right activation levels for random dot movements.

4.3 Discussion

The checkerboard stimulus used for V1 retinotopic mapping in this study is limited to the lower left quadrant or the lower right quadrant. There is no research report on how V1 corresponding to a certain quadrant encodes visual movement coherence and speed information. Based on the 3T and The 7T fMRI imaging experiments, we successfully located V1 corresponding to a single quadrant (lower left or lower right quadrant) for each subject, and by distinguishing the target area from the suround area, the activation area that overlaps with other quadrants can be avoided, and the extracted V1 is obtained based on the activation of each subject. This makes it possible to avoid the error caused by individual differences. Thus, obtaining a more accurate individual V1, and using it as the brain area of interest for subsequent SVM activity pattern classification will also make the MVPA analysis results more accurate. In addition, the size of V1 obtained by retinotopic mapping in the single quadrant is about 1/4 of the corresponding V1 in the full field of view. Therefore, the number of voxels in V1 obtained by this method is reduced compared with the complete V1, which might reduce the impact of significant differences in the size of MT+ and V1 size on MVPA analysis to some extent.

Secondly, in the fMRI experiments based on 3T and 7T, MT+ functional localizer experiment was performed by presenting motion and static random dots stimulation, and bilateral MT+ was successfully localized in all subjects, which will be used as the brain region of interest MT+ for subsequent SVM analysis. The individual-based functional localizer reduced the influences of MT+ individual differences on SVM decoding performance.

In the random dot motion experiment, comparing the activation of coherent movement and incoherent movement without considering the influence of speed found that MT+ showed bilateral higher responses to coherent motion, and V1 showed bilateral preference to incoherent motion, which was observed in both the 3T and 7T experiments. In the ROI-based analysis, MT+ showed a significant main effect of coherence except for the 7T atlas-based analysis (probably related to the large ROI extracted by the atlas,

which included some irrelevant voxels). For V1, all analysis at 7T showed the main effect of coherence that was significant or marginally significant. Besides, the localizer-based analysis at 3T also suggested this significant main effect of coherence. The above results support the view that MT+ is more sensitive to coherent motion and V1 is more sensitive to incoherent motion.

Also, speed has some influence on MT+ encoded random dots motion and less on V1. Except for the analysis of 3T (localizer) and 7T (atlas∩localizer) data, neither V1 nor MT+ showed a main effect of speed, which is consistent with the group-level activation maps at 3T and also 7T, while the results of simple effect t-test showed that the differences in responses amplitude (beta value) between different speed conditions were more pronounced during incoherent motion compared to coherent motion, which might indicates that V1 and MT+ are more sensitive to speed of incoherent random dot movement. The resulting activation difference is more manifested in incoherent movement, and it is speculated that V1 and MT+ may be more sensitive to speed discrimination during incoherent movement, while the activation during coherent movement is not affected by speed. In the ROI analysis of 3T and 7T, except for 7T's localizer and atlas∩localizer methods, V1 did not show the interaction effect of coherence and speed, indicating that speed has no obvious effect on the random dots motion encoded by V1. MT+ exhibited a significant interaction of coherence and speed observed in all three ROI analyzes of 3T and was consistently more responsive to coherent motion at a speed of 10.0 deg/s, a difference that was greater than that at 5.0 deg/, which is consistent with the results of group-level activation contrast maps. In the 7T results, this interaction was only observed in the localizer-based analysis, which may be related to an upward trend of the bilateral MT+ response with increasing speed during coherent motion. In general, no difference was observed in the response of bilateral MT+ to the two types of motion at 2.5 deg/s, which indicates that MT+ has better coherence discrimination ability in fast motion, and the faster motion can be considered in follow-up studies.

Finally, unilateral MT+ and unilateral V1 were extracted by three methods to explore their response to random dots motion. In addition to the atlas-based analysis of ROI extraction, it was observed in all results that the left and right MT+ had stronger responses to coherent motion, and the unilateral V1 had a stronger response to incoherent motion, again indicating that MT+ was more sensitive to coherent motion and V1 is more sensitive to incoherent motion. Regardless of coherence, the 3T (localizer and atlas ∩ localizer) results showed that the right side was better than the left side in distinguishing the two speeds of 5.0 and 10.0 deg/s, but both left and right side of MT+ at 3T showed no significant differences in responses to the two speeds; no significant unilateral V1 activation differences were observed for motion speed in both 3T and 7T results. The left and right activation levels were then compared for MT+ and V1. When not distinguishing motion speeds, the results of the ROI analysis at 3T showed that the response of the left MT+ was significantly stronger than that of the right MT+ for both coherent and incoherent motion, and this difference was also observed in the 7T atlas-based results; When not distinguishing motion coherence, 3T ROI analysis results demonstrated that the left MT+ showed a stronger response to any speed (2.5-10.0deg/s), and this result was also observed in the 7T atlas-based analysis. Overall, the left MT+ showed some degree of dominance in encoding motion coherence and speed information, whereas V1 showed no significant lateralization.

The 3T and 7T experiments were inconsistent in the activation comparison results of the second-order analysis, while the two areas showed similar results in the ROI-based analysis, which may be due to the small number of subjects in the 7T experiment and the variability of MT+ positions. On the other hand, the 7T partial brain fMRI images acquired in this study have a higher spatial resolution (isotropic 1.25 mm^3), and the obtained activation is more localized to the gray matter of the subject's brain compared to the 3T data. The results agree with the 3T experiments' results.

5 Results of multivariate analysis

5.1 Results of the experiment at 3T

5.1.1 Decoding of coherence in MT+ and V1

First, the performance of V1 and MT+ in decoding coherence is compared, here only bilateral V1 (lr_V1) and bilateral MT+ (lr_MT+) are analyzed. Second, the effect of speed on decoding coherence is explored; that is, the decoding performance of the coherence condition is compared at different speeds. Similarly, only bilateral V1 and bilateral MT+ are compared here. Finally, whether there is laterality in the decoding performance and consistency regarding laterality between the two brain regions, V1 and MT+, was investigated by comparing the decoding performance of the left V1 to the right V1, and the left MT+ to the right MT+.

Since there is no right V1 in the ROI extracted based on the localizer, we analyze the data based on the left V1 functional localizer; the right V1 results obtained with the atlas∩localizer are based on the right V1 from the atlas.

The SVM classification results of extracting ROIs based on the localizer show that when the influence of speed is not considered, left V1 (59.90%, t=2.896, p=0.015), left MT+ (64.37%, t=3.744, p=0.003), the right MT+ (69.10%, t=11.819, p<0.001), and the bilateral MT+ (77.78%, t=12.168, p<0.001) showed significantly higher decoding rate than the chance level (50%). There was no significant difference in accuracy between left V1 and left MT+ (p=0.348). When the speed was 10.0 deg/s, only the right MT+ (59.38%, t=5.084, p<0.001) and the bilateral MT+ (61.46%, t=5.24, p<0.001) showed significantly higher coherence classification accuracy than chance level. The accuracy rates of left V1 and left MT+ were 51.80% and 53.82%, respectively, and the difference between them was not significant (p=0.555). When the speed was 5.0 deg/s, only the left side V1 (54.36%, t=1.81, p=0.098) and bilateral MT+ (56.69%, t=1.936, p=0.079) showed a trend of higher accuracy than the chance level. The accuracy rate of the left

MT+ (54.95%, p>0.10) was not significantly different from that of the left V1 (p=0.858). No matter whether the influence of speed was considered or not, there was no significant difference in the performance of the left V1 and the left MT+ classification coherence (as shown in the first column of **Figure 5.1 (a) and (b)**). In addition, no significant differences in the performance of classification at different speeds were observed in the left V1, left MT+, and bilateral MT+ (p>0.10). Finally, since the 3T experiment did not include a retinotopic mapping of the right V1, only the classification performance of the left and right sides of MT+ was compared here. As shown in **Figure 5.2**, when the speed is not distinguished, there is no significant difference in the accuracy of the classification between the left MT+ (64.37%) and the right MT+ (69.10%) (t=-1.319, p=0.214).

The classification results obtained using the data from the ROI extracted based on the atlas show that when the speed is not distinguished (as shown in the second column of **Figure 5.1(a)**), bilateral V1 (61.24%, t=3.837, p=0.003) and bilateral MT+ (73.18%, t=10.302, p<0.001) successfully decode coherence. The accuracy of classifying coherent motion and incoherent motion was significantly higher than 50%, and the former was significantly lower than the latter (t=-4.416, p=0.001). When the speed is 5.0 deg/s, the accuracy rate obtained from bilateral V1 (54.95%, t=2.372, p=0.037) and right MT+ (55.12%, t=2.477, p=0.031) coherence decoding is significantly higher than 50 %. The decoding rates with data from the left side of V1 (56.43%, t=2.135, p=0.056) and bilateral MT+ (55.38%, t=2.011, p=0.069) are marginally significantly higher than 50%. When the speed is 10.0 deg/s, the accuracy of V1's coherence classification is close to the chance level, and the performance of MT+ is higher than the chance level, where the left MT+ (57.21%, t=3.101, p=0.010), the right MT+ (58.60%, t=4.983, p<0.001) and bilateral MT+ (58.51%, t=3.769, p=0.003) all show statistical significant decoding rate. As shown in the second column of **Figure 5.1(b)**, bilateral MT+ performs significantly better than bilateral V1 at a speed of 10.0 deg/s (t=3.439, p=0.006). In addition, the accuracies of decoding with data from bilateral V1 and bilateral MT+ in classifying coherent and incoherent motion did not show significant differences with data obtained using other speeds (p>0.10). Finally, when the speed is not distinguished, the

accuracy of the classification of the left MT+ (68.40%, t=6.763, p<0.001) and the right MT+ (74.13%, t=10.221, p<0.001) is significantly higher than 50 % and the performance of the left MT+ is significantly better than that of the right (as shown in the second column of **Figure 5.2**, t=3.455, p=0.005), the left V1 (57.47%, t=2.307, p=0.042) and the right V1 (58.86%, t=3.285, p=0.007) can also effectively classify coherent motion and incoherent motion, but there is no significant difference between the levels obtained in the two hemispheres (as shown in **Figure 5.3**, t=-0.824, p=0.428).

The classification results using ROI extraction based on atlas∩localizer show that when the speed is not distinguished (as shown in the third column of **Figure 5.1(a)**) the accuracy of classification based on data from the bilateral V1 (59.64%, t=3.284, p=0.007) and the bilateral MT+ (77.74%, t=12.69, p<0.001) were significantly higher than the chance level (50%). The difference between the two was significant (t=-6.094, p<0.001). When the speed is 5.0 deg/s, only the two-sided MT+ (58.16%, t=2.462, p=0.032) classification accuracy is significantly higher than 50%. When the speed is 10.0 deg/s, left MT+ (55.21%, t=3.856, p=0.003), right MT+ (57.90%, t=3.610, p=0.004), bilateral MT+ (60.94%, t= 5.632, p<0.001) show classification accuracy significantly higher than 50%, and only the left V1 (54.17%, t=2.073, p=0.062) had a marginally higher accuracy rate than 50%. As shown in the third column of **Figure 5.1(b)**, the accuracy of classification coherence (51.74%, p>0.10) of the bilateral V1 at 10.0 deg/s is close to the chance level and significantly lower than that of bilateral MT+ (t=-3.109, p=0.010). When the speed is not considered, the bilateral MT+ shows better coherence classification than the bilateral V1, but when distinguishing the speed, the difference between these brain areas is only reflected when the speed is 10.0 deg/s. In addition, neither the accuracy of bilateral V1 nor that of bilateral MT+ showed significant differences between the 5.0 and 10.0 deg/s conditions. Finally, we performed a side-by-side comparison of the performance of MT+ and V1 classification. As shown in the third column of **Figure 5.2**, both the left and right MT+ can classify coherent and incoherent motions, and there is no significant difference in accuracy between the two (t=-1.503, p=0.161).

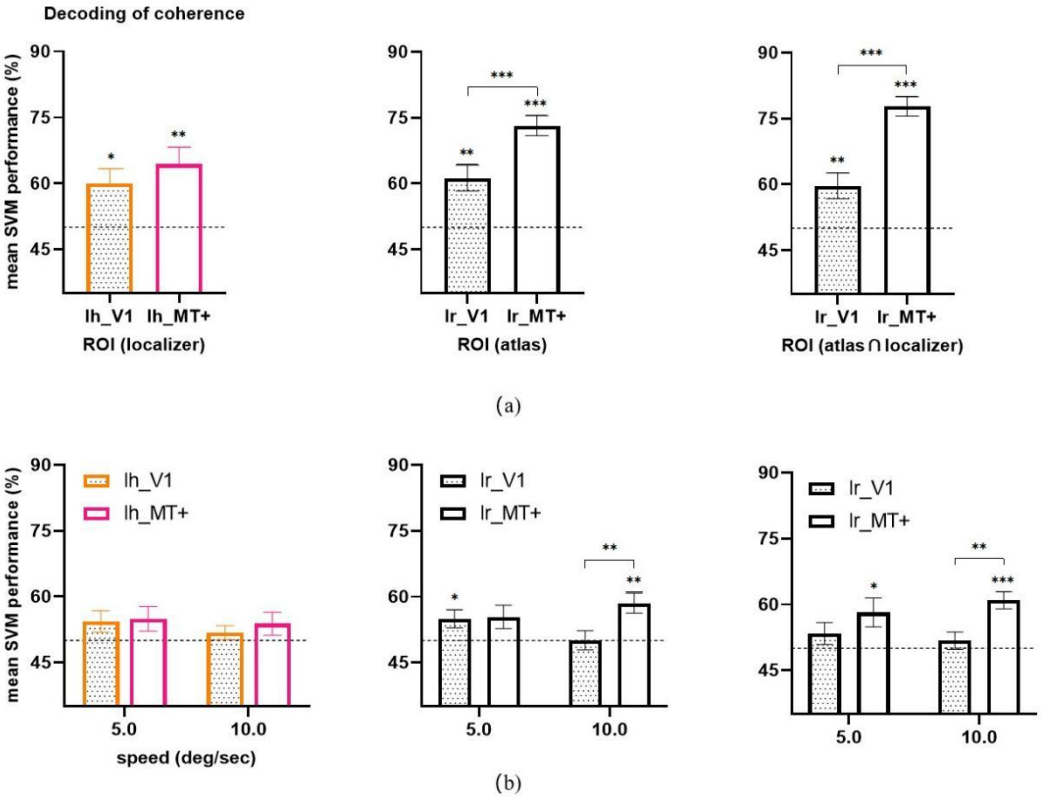


Figure 5.1 Performance of decoding coherent and incoherent motion (3T). (a) Decoding performance in V1 and MT+ regardless of speed. (b) Decoding performance in V1 and MT+ at 5.0 or 10.0 deg/s. Left, middle and right columns show the results in V1 or MT+ generated by localizer, atlas, and atlas ∩ localizer, respectively. In 3T the experiment, V1 is mapped only in the left hemisphere, so SVM decoding is analyzed only in left V1. The first column presents the performance of the left MT+, for comparison. The chance level is 50% (dashed lines). * indicates $p<0.05$, **$p<0.01$, ***$p<0.001$. Error bars indicate SEM.

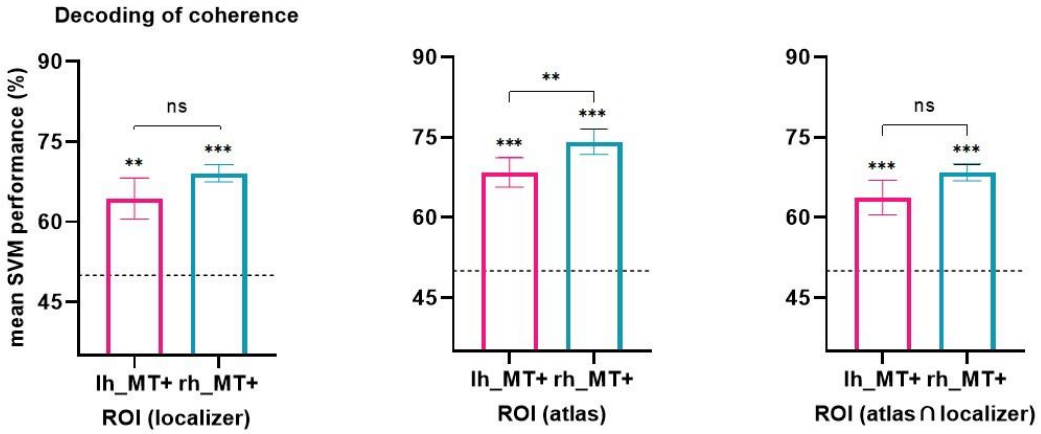


Figure 5.2 Performance of decoding coherent and incoherent motion in left or right MT+ (3T). The left, middle and right columns show the results in a single side of MT+ generated by the localizer, atlas, and the atlas ∩ localizer, respectively. The chance level is 50% (dashed lines). **$p<0.01$, ***$p<0.001$, ns indicates $p>0.05$. Error bars indicate SEM.

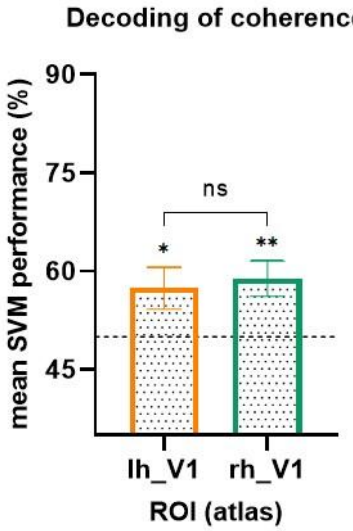


Figure 5.3 Performance of decoding coherent and incoherent motion in left or right V1 (3T). V1 is mapped only in the left hemisphere in this 3T experiment. Therefore, to compare between the left and right V1, only the results generated by the atlas are shown. The chance level is 50% (dashed lines). *p<0.05, **p<0.01, ns indicates p>0.05. Error bars indicate SEM.

5.1.2 Decoding of speed in MT+ and V1

First, the performance of V1 and MT+ in decoding motion speed was compared, here only bilateral V1 (lr_V1) and bilateral MT+ (lr_MT+) were analyzed. Second, the effect of coherence on decoding motion speed was explored, that is, comparing the performance of bilateral V1 and bilateral MT+ in decoding motion speed in the two motion modes of coherent motion and incoherent motion. Finally, whether there is laterality in the decoding of motion speed in two brain regions, V1 and MT+, was investigated by comparing the decoding performance of left V1 and right V1, left MT+ and right MT+ for the two types of motion speed.

The classification results based on the ROI extracted by the localizer showed that when the coherence was not considered, the accuracy of the classification speed of all brain regions of interest was close to or even lower than the chance level (50%). When the random dots move coherently, only the speed classification based on data from bilateral MT+ (53.47%) approaches a success rate higher than 50% (t=2.046, p=0.065), and the performance of other brain regions is close to the chance level. Under incoherent motion conditions, all regions of interest failed to effectively distinguish the two motion speeds, and the accuracy of the left MT+ classification speed was significantly lower than the

chance level (t=-2.343, p=0.039). When comparing the performance of V1 and MT+ in decoding motion speed, considering that the right V1 was not located in the 3T experiment, the decoding performance of left V1 and left MT+ is compared. As shown in the first column of **Figure 5.4(a) and (b)**, there is no significant difference in the accuracy of the speed classification between the left V1 and the left MT+ regardless of whether the influence of coherence is considered. Also, motion coherence had no effect on speed classification for either V1 or MT+ (both were below chance level). Finally, as shown in the first column of **Figure 5.5**, when the coherence is not considered, the classification speed accuracy of the left MT+ (51.09%) and the right MT+ (52.74%) is close to the chance level and the difference is not significant (t=-0.781, p=0.451).

The classification results with data obtained with ROI based on the atlas show that when the coherence is not considered, only the accuracy rate (52.56%) of the bilateral V1 approaches significance (t=2.163, p=0.053), while MT+ shows success rate close to or below chance levels. With coherent motion, only the right MT+ (54.43%, t=3.263, p=0.008) had a classification accuracy significantly higher than 50%, while the right V1 (54.26%, t=2.069, p=0.063) and bilateral MT+ (52.17%, t=1.938, p=0.079) approachd significance. With incoherent motion, the accuracy of unilateral and bilateral V1 and MT+ speed classification was close to or below the chance level. As shown in the second column of **Figure 5.4(a) and (b)**, regardless of whether the effect of motion coherence is considered, the performance of V1 and MT+ classification speed is close to the chance level and there is no significant difference. Under the incoherent motion condition, both bilateral V1 (49.13%) and bilateral MT+ (49.74%) speed classification success rates are lower than the chance level (50%). accuracy. When the coherence is not considered, as shown in **Figure 5.5**, neither the left MT+ (49.92%) nor the right MT+ (49.70%) can effectively distinguish the speed and there is no significant difference in the decoding performance of the two (t=0.137, p=0.893); As shown in **Figure 5.6**, there is no significant difference (t=0.933, p=0.371) in the accuracy of the speed classification of the left V1 (52.26%) and the right V1 (50.52%).

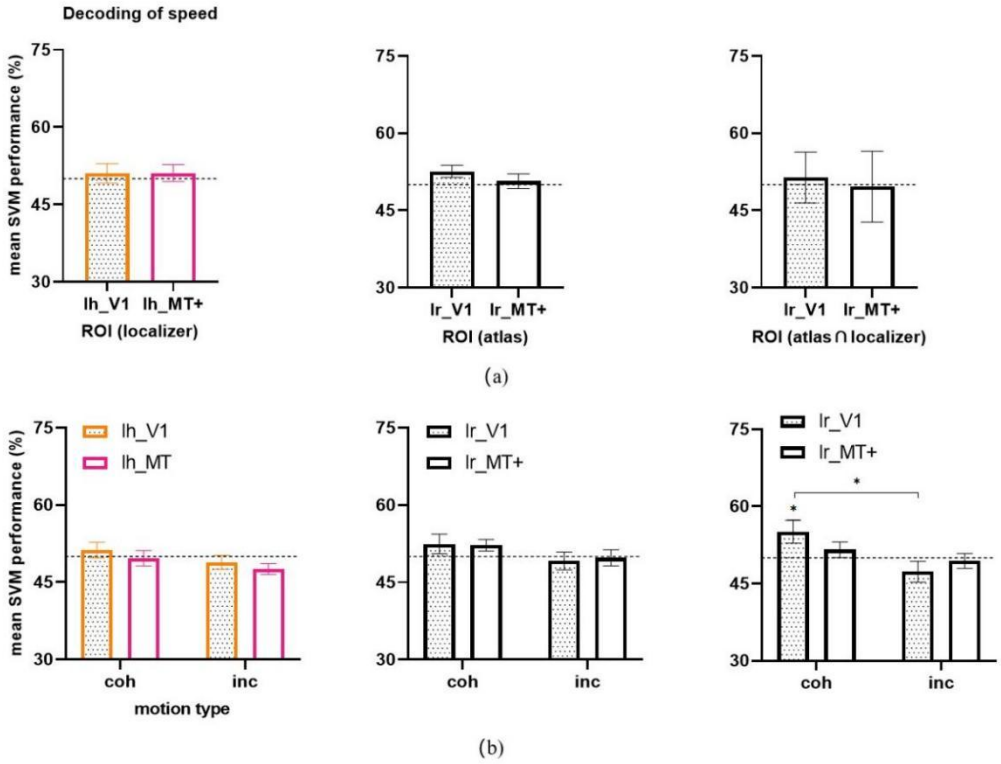


Figure 5.4 Performance of speed decoding of random moving dots (3T). (a) Decoding performance in V1 and MT+ regardless of coherence. (b) Decoding performance in V1 and MT+ when random dots move coherently (coh) or incoherently (inc). The left, middle and right columns show the results in V1 or MT+ generated by the localizer, atlas, and atlas ∩ localizer, respectively. In the 3T experiment, only the left V1 is mapped, so SVM decoding is analyzed only in left V1 and the performance of left MT+ is illustrated in the first column for comaparison. The chance level is 50% (dashed lines). *p<0.05. Error bar indicated SEM.

The classification results based on the ROI extracted by atlas∩localizer show that the accuracy rate of MT+ for decoding motion speed is close to or lower than the chance level, regardless of motion coherence. As shown in the third column of **Figure 5.4(a) and (b)**, there was no significant difference in speed classification performance between bilateral V1 and bilateral MT+ (p>0.1). Compared with incoherent motion, V1 better distinguishes motion speed under coherent motion conditions (t=2.723, p=0.020), and bilateral V1 (55.04%, t=2.257, p=0.045) classification accuracy is significantly higher than 50%. When the influence of coherence is not considered, as shown in the third column of **Figure 5.5**, the accuracy of speed classification of the left MT+ (51.96%) and right MT+ (53.56%) is higher than the chance level but not significant (p>0.10) and there is no significant difference between the two (p=0.418). Similarly, as shown in **Figure 5.6**,

V1 also does not exhibit a significant difference in the accuracies obtained by the left and right V1 (t=0.933, p=0.371).

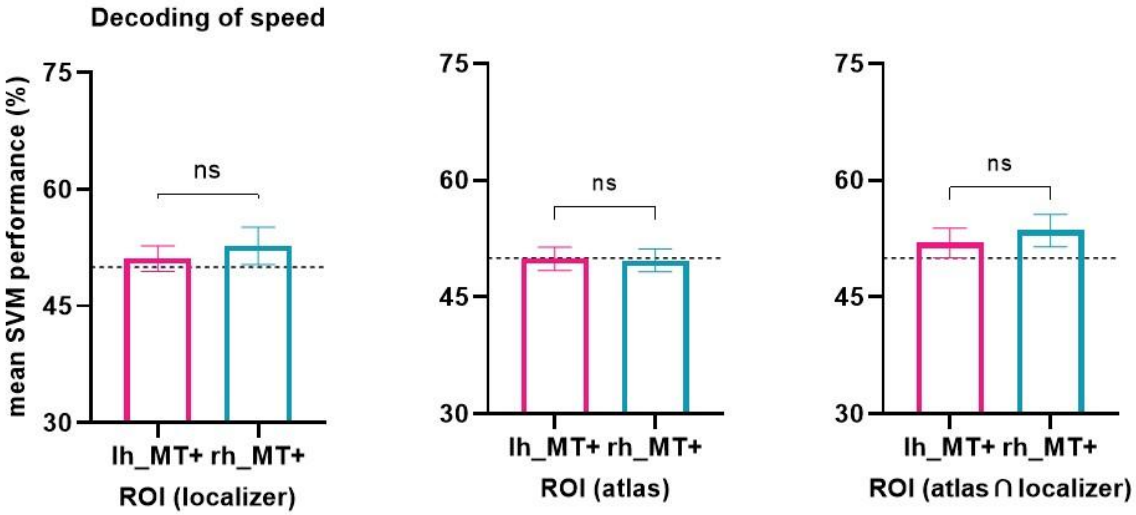


Figure 5.5 Performance of decoding speed of moving random dots in the left or right MT+ (3T). The left, middle and right columns show the results from a single side of MT+ as delineated by the localizer, atlas, and atlas ∩ localizer, respectively. The chance level is 50% (dashed lines). ns indicates p>0.05. Error bars indicate SEM.

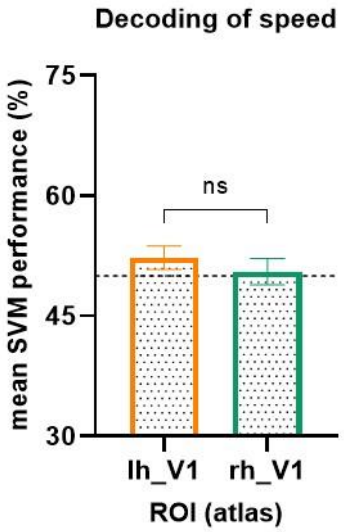


Figure 5.6 Performance of decoding speed of the random moving dots in left or right V1 (3T). V1 In the 3T experiment, V1 is mapped only in the left hemisphere. Therefore, to compare the results obtained in the left and right V1, we only use the data from the two parts of V1 as delineated by the atlas. The chance level is 50% (dashed lines). ns indicates p>0.05. Errors bars indicate SEM.

5.2 Results of the experiment at 7T

5.2.1 Decoding of coherence in MT+ and V1

The steps of SVM classification analysis for 7T data were the same as those for 3T data. First, we explored the performance of V1 and MT+ in decoding motion coherence, that is, whether the accuracy of V1 and MT+ in classifying coherent and incoherent motion is significantly higher than the chance level, regardless of speed, and compare the two-sided V1 (lr_V1) and two-sided decoding performance of MT+ (lr_MT+). Second, we investigated the effect of speed on decoding coherence by comparing the performance of bilateral V1 and bilateral MT+ coherence decoding at different speeds. Finally, we tested whether there is laterality in the performance of decoding coherence in the two brain regions, V1 and MT+, by comparing the left and right sides of V1 and the left and right sides of MT+ for coherent and incoherent movements decoding performance.

The classification results with ROI extracted based on the localizer show that when the influence of speed is not considered, both V1 and MT+ can distinguish coherent motion and incoherent motion. As shown in the first column of **Figure 5.7(a)**, the bilateral V1 (59.24%, t=7.820, p=0.001) and bilateral MT+ (73.61%, t=3.622, p=0.022) classified coherent motion and incoherent motion successfully. The accuracies of coherent motion are both significantly higher than the chance level (50%), but the difference between the two only approaches significance (t=-2.085, p=0.105). As shown in the first column of **Figure 5.7(b)**, bilateral MT+ only exhibited a decoding rate higher than the chance level at a speed of 10.0 deg/s (57.43%, t=3.880, p=0.018). The motion coherence cannot be effectively decoded under the condition of other speeds. Especially when the speed is 7.5 deg/s, it is significantly lower than the chance level (<50%, t=-3.379, p=0.028). Bilateral V1 only showed 54.10% (t=2.211, p=0.092) decoding performance at speed 2.5. It was unable to distinguish between coherent motion and incoherent motion at other speed conditions (5.0, 7.5 and 10.0 deg/s). Overall, V1 and MT+ showed the worst decoding coherence (lowest accuracy) when the speed was 7.5 deg/s, and the two brain regions only showed significant differences in decoding coherence ability when the

speed was 10.0 deg/s (t=-6.489, p=0.003). No significant main effect of speed was observed in either bilateral V1 (F(df=3)=1.964, p=0.173) or bilateral MT+ (F(df=3)=1.483, p=0.269). In other words, the performance of classification accuracy is not affected by speed. When the speed is not considered, as shown in **Figure 5.8**, the left MT+ (71.30%, t=3.952, p=0.017) and the right MT+ (64.62%, t=4.400, p=0.012) can effectively classify coherent motion and incoherent motion, and the former shows better accuracy than the latter (marginally significant, t=2.576, p=0.062). As shown in **Figure 5.9**, left V1 (59.31%, t=8.404, p=0.001) and the right V1 (56.08%, t=4.896, p=0.008) successfully decode motion coherence; there is no significant difference between the the performances of these regions (t=1.931, p=0.126).

The classification results based on the atlas-extracted ROIs show that both V1 and MT+ can classify coherent motion and incoherent motion when speed is not considered. As shown in the second column of **Figure 5.7(a)**, the coherence classifications by bilateral V1 (63.91%, t=10.212, p=0.001) and bilateral MT+ (69.12%, t=5.267, p=0.006) were significantly higher than 50%. There was no significant difference between the two (t=-1.960, p=0.122). As shown in the second column of **Figure 5.7(b)**, the performance of bilateral V1 at speeds of 2.5 (54.24%, t=2.720, p=0.053) and 10.0 (54.38%, t=2.747, p=0.052) deg/s approaches significance. Likewise, bilateral MT+ exhibits accuracy significantly higher than chance level at 2.5 (55.97%, t=4.088, p=0.015) and 10.0 (53.68%, t=3.119, p=0.036) deg/s. At each speed, no significant difference between bilateral V1 and bilateral MT+ decoding coherence performance was observed. In addition, no significant main effect of speed was observed in bilateral V1 and bilateral MT+, however, the performance of coherence decoding in both brain regions was worst at a speed of 7.5 deg/s (about 51.70%). When the speed is not considered, as shown in the second column of **Figure 5.8**, the left MT+ (69.98%, t=6.104, p=0.004) and the right MT+ (67.74%, t=5.872, p=0.004) can both decode motion coherence, with no significant difference between the two (t=1.301, p=0.263). As shown in the second column of **Figure 5.9**, left V1 (62.85%, t=8.564, p=0.001) and right V1 (63.33%,

t=8.944, p=0.001) can also decode coherent information, and similarly, there is no significant difference between the two (t=-0.335, p=0.754).

The classification results based on the ROI extracted by atlas∩localizer show that both V1 and MT+ can classify coherent and incoherent motions when speed is not considered. As shown in the third column of **Figure 5.7(a)**, the classification accuracy rates of bilateral V1 (60.26%, t=7.096, p=0.002) and bilateral MT+ (73.77%, t=3.775, p=0.020) are both statistically significant. The difference between the two is not significant (t=-1.998, p=0.116). As shown in the third column of **Figure 5.7(b)**, the success rate of bilateral MT+ in decoding coherence is significantly higher than 50% (58.68%, t=3.419, p=0.027) only when the speed is 10.0 deg/s. The two-sided V1 showed a decoding accuracy slightly higher than the chance level when the speed was 2.5 deg/s (54.17%, p=0.106, not significant), while its decoding performance was below chance level under other speed conditions. Bilateral V1 exhibited a significant main effect of speed (F=38.777, p=0.025), with significantly higher accuracy for decoding coherence at 2.5 deg/s than at 7.5 deg/s (p=0.005, Bonferroni correction). No significant main effect of speed was observed in bilateral MT+, and there were no significant differences in decoding performance across speeds. When the speed is not distinguished, as shown in the third column of **Figure 5.8**, the left MT+ (71.44%, t=3.969, p=0.017) and the right MT+ (64.90%, t=4.499, p=0.011) can classify coherent motion with statistically significant success rate. The decoding performance of the left MT+ i\shows a trend of being higher relative to that of the right MT+ (t=2.337, p=0.080). As shown in **Figure 5.9**, the left V1 (57.90%, t=5.596 , p=0.005) and the right V1 (56.50%, t=5.473, p=0.005) show no significant difference in decoding coherence (t=1.127, p=0.323).

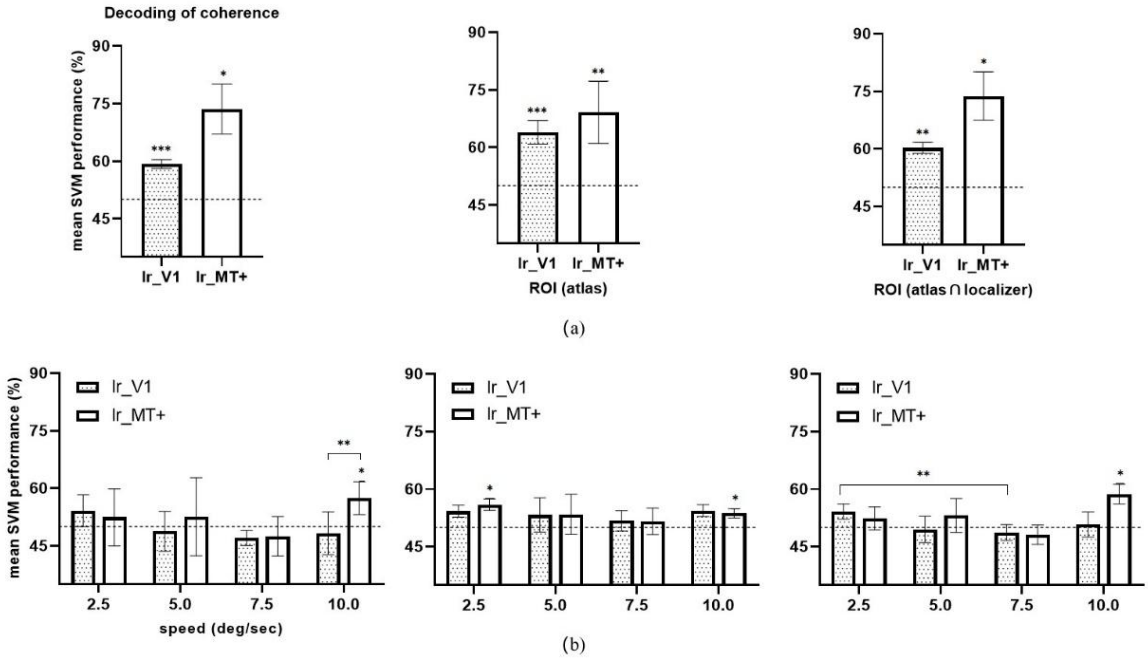


Figure 5.7 Performance of decoding coherent and incoherent motion (7T). (a) Decoding performance in V1 and MT+ regardless of speed. (b) Decoding performance in V1 and MT+ at 2.5, 5.0, 7.5, 10.0 deg/s. The left, middle and right columns show the results in V1 or MT+ generated by the localizer, atlas, and atlas ∩ localizer, respectively. The chance level is 50% (dashed lines). * indicates p<0.05, **p<0.01, ***p<0.001. Error bar indicated SEM.

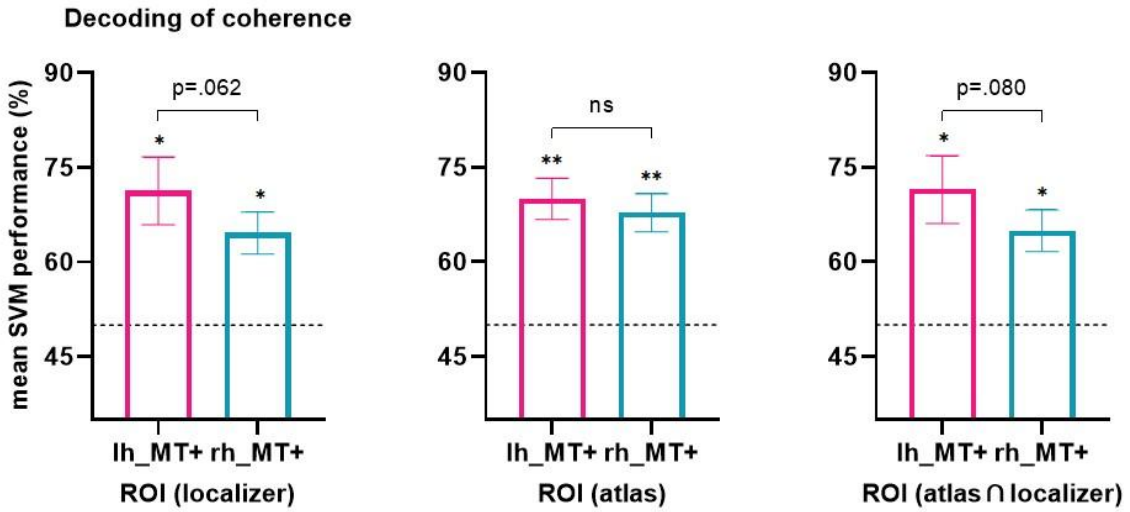


Figure 5.8 Performance of decoding coherent and incoherent motion in left or right MT+ (7T). The left, middle and right columns show the results in a single side of MT+ generated by the localizer, atlas, and atlas ∩ localizer, respectively. The chance level is 50% (dashed lines). * indicates p<0.05, **p<0.01. Error bar indicated SEM.

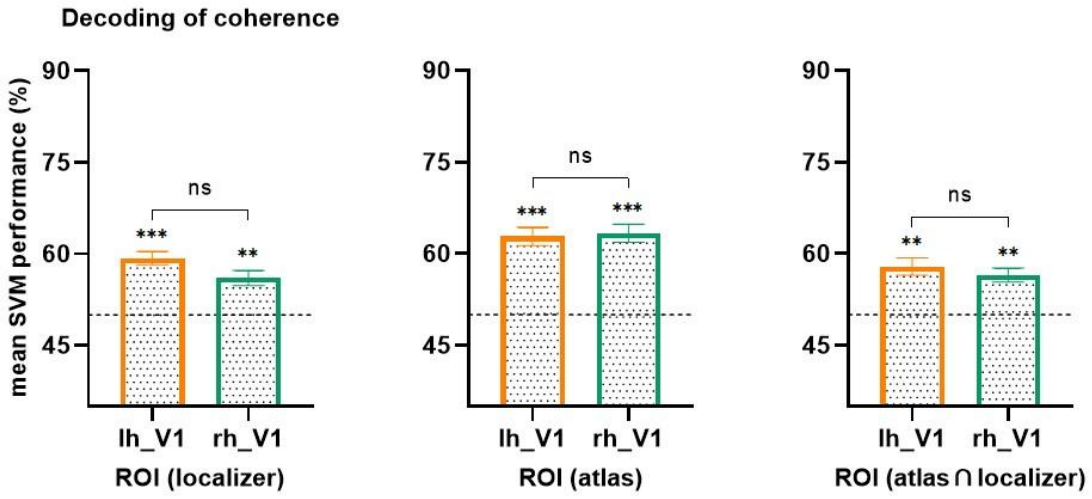


Figure 5.9 Performance of decoding coherent and incoherent motion the in left or right V1 (7T). The left, middle and right columns show the results from a single side of V1 generated by the localizer, atlas, and atlas ∩ localizer, respectively. The chance level is 50% (dashed lines). **p<0.01, ***p<0.001. Error bars indicate SEM.

5.2.2 Decoding of speed in MT+ and V1

This section will illustrates the performance of MT+ and V1 in decoding random dot motion speed. First, the classification performance of MT+ and V1 of four motion speed conditions was explored without considering the influence of motion coherence, and the decoding ability of two brain regions for motion speed was compared. Second, the effect of coherence on decoding motion speed was explored by comparing the speed decoding performance of bilateral MT+ and bilateral V1 under coherent or incoherent motion conditions. Finally, whether there is laterality in the decoding of motion speed in two brain regions, V1 and MT+, was investigated by comparing the left and right sides of MT+ and the left and right sides of V1, respectively.

The classification results based on the ROI extracted by the localizer show that when the motion coherence is not considered, both V1 and MT+ showed the decoding speed ability and no significant difference between them was observed. . The right V1 (29.50%, t=3.699, p=0.021), bilateral V1 (31.13%, t=3.924, p=0.017) accuracy rate of classification speed was significantly greater than the chance level (25%). In the right MT+ (30.92%, t=2.777, p=0.050) and bilateral MT+ (30.14%, t= 2.574, p=0.062) the

decoding performance approached significance. As shown in the first column of **Figure 5.10(a)**, there is no significant difference in the accuracy of speed classification between bilateral V1 and bilateral MT+ (t=0.511, p=0.636). Additionally, V1 and MT+ did not exhibit significant differences in decoding speed performance under either coherent or incoherent motion conditions. Both bilateral V1 and bilateral MT+ showed no significant difference in decoding speed levels in different motion modes, however, as shown in the first column of **Figure 5.10(b)**, bilateral V1 (25.84%) and bilateral MT+ and both lateral MT+ (27.43%) show a trend of classifying speed only in coherent motion (>25%, p>0.10). Especially, the accuracy rate of bilateral V1 in incoherent motion is lower than 25% (p=0.087). When the coherence is not considered, as shown in the first column of **Figure 5.11**, the accuracy of the speed classification of the left MT+ (28.60%) is higher than the chance level (not significant, p=0.133). Although the corresponding accuracy of the right MT+ was significantly higher than 25%, it did not show a significant difference relative to the left (t=-1.428, p=0.227). As shown in the first column of **Figure 5.12**, the speed classification of the left V1 and the right V1 were 28.59% (p>0.10) and 30.92% (p<0.05), respectively, and there was no significant difference between them (t=-0.324, p=0.762).

The classification results based on the ROI extracted from the atlas show that only V1 exhibits decoding speed when coherence is not considered. As shown in the second column of **Figure 5.10(a)**, the accuracy rate of speed classification of bilateral V1 (28.56%, t=3.472, p=0.026) is significantly higher than 25%, and the accuracy rate of bilateral MT+ is 27.28% (not significantly higher than the chance level, p=0.160). The difference between the two is not significant (t=1.306, p=0.262). No significant differences were observed in the performance of bilateral V1 and bilateral MT+ speed decoding even when coherent or incoherent movement conditions were considered separately (**Figure 5.10 (b)** second column). Neither bilateral V1 nor bilateral MT+ showed significant differences in the accuracy of decoding speed in different motion modes. However, none of the decoding accuracies reached 25% during incoherent motion, with bilateral V1 and bilateral MT+ decoding speed accuracies of 26.77% and 25.87%, respectively (not

significantly different from chance levels, p>0.10). No significant difference in the decoding performance between coherent and incoherent motion patterns might be caused by the few numbers of subjects at 7T which is difficult to reach a statistical significance between them. When the coherence is not considered, as shown in the second column of **Figure 5.11**, the accuracy rates of the left MT+ and right MT+ speed classification are 27.01% and 27.74%, respectively. There was no significant difference relative to the chance level (p>0.10) and there was no significant difference between the two (t=1.021, p=0.365). As shown in the second column of **Figure 5.12**, the left V1 (29.32%, t=3.022, p=0.039) and the right V1 (28.42%, t=3.993, p=0.016) distinguished the four speeds, but there was no significant difference in decoding accuracy between the two (t=0.615, p=0.572).

The classification results with data from ROI extraction based on atlas∩localizer show that when coherence is not considered, as shown in the third column of **Figure 5.10(a)**, bilateral V1 (29.29%, t=4.628, p=0.010) and bilateral MT+ (29.50%, t=2.801, p=0.049) show speed classification accuracy higher than 25%; the difference between V1 and MT+ accuracies is not significant (t=-0.178, p=0.867). Neither bilateral V1 nor bilateral MT+ showed significant differences in decoding speed levels between coherent and incoherent motion (p>0.10). As shown in the third column of **Figure 5.10(b)**, both bilateral V1 and bilateral MT+ showed decoding accuracy lower than the chance level during incoherent movements, while the decoding levels of the two brain regions were slightly above the chance level during coherent movements (bilateral V1: 25.66%, bilateral MT+: 26.49%). When coherence is not considered, as shown in the third column of **Figure 5.11**, the accuracy rate of speed classification by the left MT+ is 28.61% (not significantly higher than 25%, p=0.10). The speed classification of the right MT+ (32.15%) is significantly higher than 25% (t=3.835, p=0.019) and significantly higher than the that of the left MT+ (t=3.681, p=0.021). As shown in the third column of **Figure 5.12**, the left V1 (29.56 %, t=2.545, p=0.064) shows a trend approaching significant speed classification accuracy. The right V1 (29.53%, t=3.876, p=0.018) speed

classification accuracy rate is higher than 25%, with no significant difference between the two (t= 0.040, p=0.970).

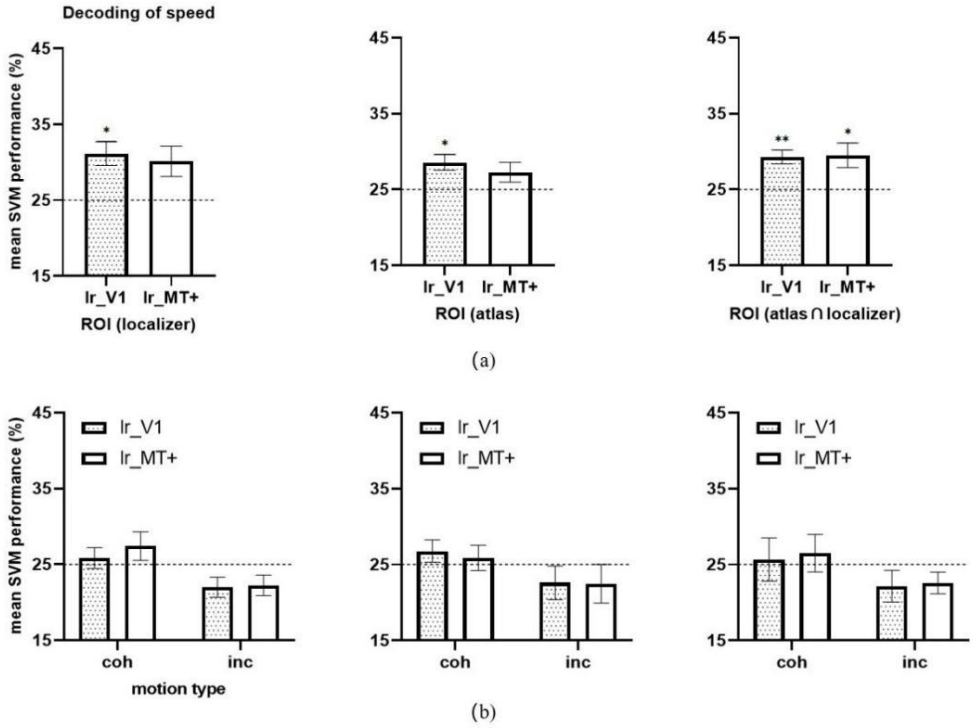


Figure 5.10 Performance od decoding speed of random moving dots (7T). (a) Decoding performance in V1 and MT+ regardless of coherence. (b) Decoding performance in V1 and MT+ when random dots move coherently (coh) or incoherently (inc). The left, middle and right columns show the results in V1 or MT+ generated by the localizer, atlas, and atlas ∩ localizer, respectively. The chance level is 25% (dashed lines). * indicates p<0.05, **p<0.01. Error bars indicate SEM.

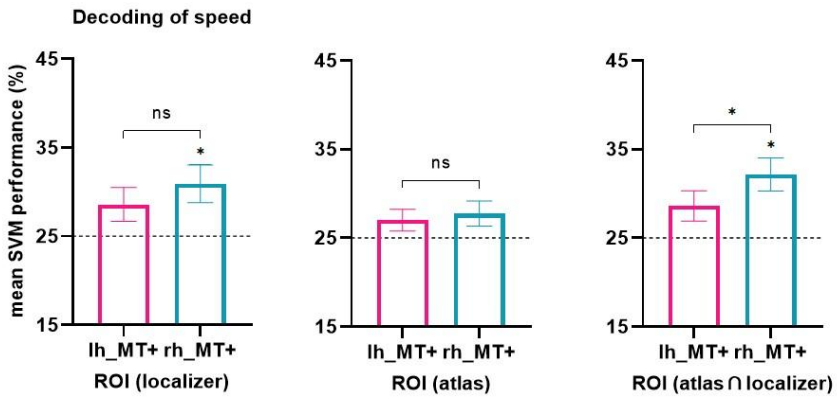


Figure 5.11 Performance of decoding speed of random moving dots in left or right MT+ (7T). The left, middle and right columns show the results in a single side of MT+ generated by the localizer, atlas, and atlas ∩ localizer, respectively. The chance level is 25% (dashed lines). ns indicates p>0.05, *p<0.05. Error bar indicated SEM.

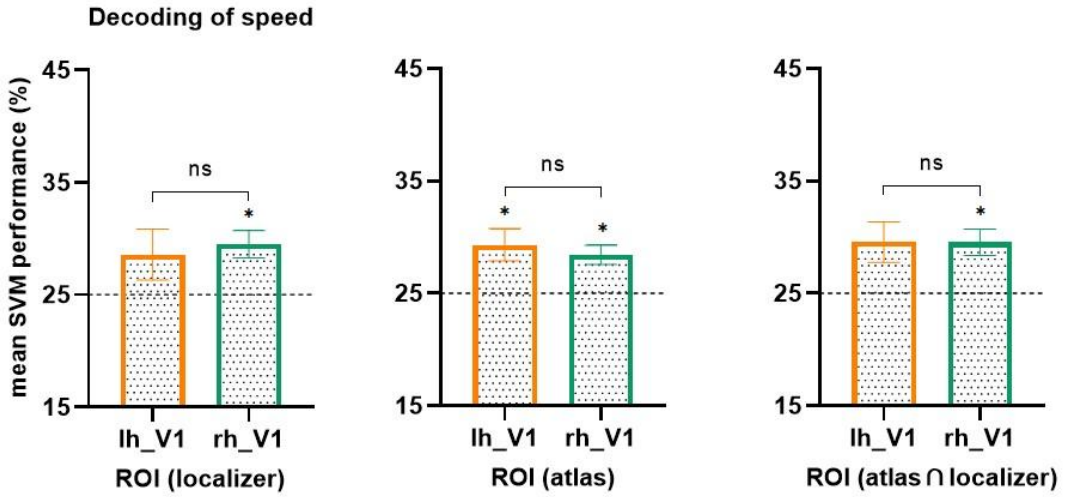


Figure 5.12 Performance of decoding the speed of moving random dots in the left or right V1 (7T). The left, middle and right columns show the results in a single side of V1 generated by the localizer, atlas, and atlas ∩ localizer, respectively. The chance level is 25% (dashed lines). ns indicates p>0.05, *p<0.05. Error bars indicate SEM.

5.3 Discussion

The classification results of SVM based on 3T or 7T fMRI data show that MT+ and V1 can effectively decode random dots motion coherence information, and the decoding performance of MT+ (69%-78%) is always better than that of V1 (approximately 60%) (i.e., higher classification accuracy). In the 3T experiment, the classification results of ROI extraction based on the visual atlas (atlas) and intersection (atlas∩localizer) methods showed that the accuracy of the bilateral MT+ classification of coherence was significantly higher than that of the bilateral V1 when no speed was distinguished (p<0.05). The classification accuracy of the bilateral V1 in the 7T experiment was still lower than that of the bilateral MT+, but the difference between the two was not significant (p>0.05), which may be caused by the small number of subjects.

The decoding of motion speed showed inconsistent results in the 3T experiment compared to the 7T experiments. In the 3T experiment, regardless of whether the ROI was obtained based on the functional localizer (localizer), the visual atlas (atlas) or the intersection of the two, no region of interest (including One-sided V1, two-sided V1,

one-sided MT+, two-sided MT+) successfully decoded the two-speed conditions of 5.0 and 10.0 deg/s (classification accuracy 47%-55%, not significantly different than chance level, p>0.05). However, in the classification results of the 7T experiment, bilateral V1 showed consistent decoding of the speed. Each decoding analysis was performed for a pair of speeds from the four conditions of 2.5, 5.0, 7.5, and 10.0 deg/s. The accuracy was consistently and significantly higher than the chance level by 25%, p<0.05, under the three ROI methods. More importantly, in the classification results of the 7T experiments, bilateral MT+ showed decoding at a rate higher than the chance level. In the analysis results based on atlas∩localizer, the accuracy of bilateral MT+ classification of speed was 29.50% (significantly higher than 25%, p<0.05).

None of the above analyzes considered the impact of different speeds or different coherences on decoding performance. Our discussion explores the mutual influence of the two.

First, speed may affect the performance of MT+ and V1 in decoding motion coherence. In the 3T and 7T results, although MT+ and V1 showed no significant difference in the accuracy of decoding coherence at different speeds, V1 and MT+ still showed different differences in distinguishing coherent and incoherent motion. The results obtained at 7T show that the V1 decoding based on all three ROI extraction methods all show coherence decoding accuracy higher than the chance level for a speed of 2.5deg/s (although the result is not significant). Data from V1 could not decode motion coherence at speeds of 5.0 or 10.0 deg/s. The results of 3T showed that all the subjects showed a classification accuracy higher than the chance level at a speed of 5.0 deg/s (the average classification accuracy of the bilateral V1 obtained based on the atlas was significantly higher than that of the chance level, p<0.05). However, the bilateral V1 responses obtained with 10.0 deg/s yielded decoding accuracy almost equal to the chance level. The results of bilateral V1 obtained at 7T demonstrated that some subjects showed decoding performance higher than the chance level at these two speeds (5.0 and 10.0 deg/s), and the atlas-based results showed that the mean decoding accuracy of bilateral V1 of all 5 subjects was marginally

significantly higher than the chance level when the speed was 10.0 deg/s. The V1 ROI obtained by retinotopic mapping only includes the response area of the lower visual field (lower left quadrant and lower right quadrant). This result indicates that a larger ROI may improve the level of SVM classification. Subsequent analysis should consider the influence of this factor on decoding performance. The results from both 3T and 7T demonstrated that when the speed was 10.0 deg/s, MT+ showed the ability to decode coherence (the classification accuracy of the bilateral MT+ was significantly higher than the chance level in all results, p<0.05), and the performance difference between MT+ and V1 coherence decoding is the largest at this speed (3T results based on atlas and atlas∩localizer and 7T results based on localizer all observe that the decoding level of bilateral MT+ at 10.0 deg/s is significantly better than that of the bilateral V1). The results obtained from the 7T data show that neither MT+ nor V1 can effectively distinguish between coherent and incoherent motion at a speed of 7.5 deg/s. Especially the bilateral MT+ based on localizer extraction shows a classification accuracy significantly lower than the chance level at this speed, and the decoding performance of bilateral V1 based on atlas∩localizer extraction at this speed is significantly worse than that obtained at 2.5 deg/s (p<0.05). Therefore, V1 is capable of distinguishing coherent and incoherent motion at 2.5 deg/s, while the decoding ability of MT+ at 5.0 and 10.0 deg/s is unstable. In follow-up studies, we recommend considering an increase of the time devoted to acquire the MT+ localizer data as well as the number of subjects for further exploration.

Second, motion coherence may also affect the performance of decoding motion speed based on data from MT+ and V1. The results from 3T and 7T did not show a significant difference in the performance of speed decoding with data from the two brain regions under different motion patterns. However, the data from both MT+ and V1 yielded better decoding results in decoding different motion speeds under coherent motion conditions, while speed decoding accuracy with incoherent motion stimuli was lower than chance level for almost all subjects (this was even more evident in the 7T results). During incoherent motion, the accuracy of decoding speed based on the left MT+ delineated by

the localizer in the 3T experiment was significantly lower than the chance level (<50%, p<0.05), and the accuracy of decoding based on the bilateral V1 extracted by localizer in the 7T experiment showed a trend of being lower than chance level (<25%, p=0.087). Therefore, for MT+ and V1, coherent motion may be more conducive to encoding the speed information of random dots, while incoherent motion is not conducive to the encoding of the speed difference of random dots.

In addition, we discuss here the laterality in decoding motion coherence and speed by MT+ and V1. When decoding coherence information regardless of motion speed, the classification results of both 3T and 7T experiments showed no lateralization of V1 in encoding motion coherence (i.e., there was no significant difference in the decoding performance of left and right V1, p>0.05). The 3T and 7T experiments show inconsistent results regarding whether MT+ demonstrated laterality in decoding coherence. In the classification results based on the 3T data, MT+ shows a trend of classification accuracy being higher with data from the right MT+ than that on the left side. Especially for the data obtained from the MT+ extracted based on the atlas, the accuracy obtained based on the data from the right MT+ is significantly higher than that of the left MT+ (p<0.01). The coherence classification results of 7T show an opposite trend. MT+ shows a trend of the accuracy with the data from the left side is higher than that on the right side, but the trend is not significant overall. Based on the two methods of localizer and atlas∩localizer, the difference is marginally significant. Overall, neither V1 nor MT+ showed consistent significant lateralization when decoding coherence information.

When decoding motion speed information regardless of motion coherence, 3T results show that either left or right MT+, left or right V1 decode speed with accuracy close to chance level. Neither V1 or MT+ exhibit laterality in decoding accuracy. The results of imaging at 7T showed that the right MT+ and the right V1 were more likely to distinguish four movement speeds (the classification accuracy was significantly higher than 25%, p<0.05), while the left MT+ showed no statistically significant decoding

accuracy. The left V1 showed decoding accuracy higher than the chance level only when the ROI was based on the atlas. Overall, neither MT+ nor V1 showed significant lateralization when decoding speed information.

Considering the different number of subjects imaged at 3T and 7T, especially the small number of subjects in the 7T experiment, it is currently impossible to draw a definite conclusion on whether MT+ has lateralization when encoding coherent information. However, the contradictory results of the two experiments indicate that, to a certain extent, MT+ possibly shows laterality when encoding visual motion coherence. Future studies should consider increasing the number of subjects. In the 7T experiment, both the lower left and lower right quadrants were stimulated, to make it possible to map the two corresponding dorsal quadrants in V1. We recommend continuing this method of symmetrically imaging V1 in the two hemispheres.

Taken together, the research results of this chapter show that V1 is more sensitive to visual movement speed information, while MT+ is more sensitive to visual movement coherence information, which explains that MT+ has higher sensitivity to direction than V1 from the perspective of multivariate pattern analysis. The results verify the important role of MT+ in the dorsal visual pathway in the process of visual motion perception. On the other hand, the SVM classification results of 7T data not only confirmed the ability of MT+ and V1 to decode motion coherence already observed through the 3T data, but also demonstrated findings that were not revealed by the 3T data. It can be seen that high-resolution (7T) fMRI data can reveal the processing of visual motion coherence and speed information in visual areas MT+ and V1 under the condition of a small number of subjects, and can obtain the same effect as 3T data (where the number of subjects was more than twice that of the 7T experiment). The classification results show similar or even better decoding levels, which indicates that improving the spatial resolution of fMRI improve the decoding performance of visual motion information.

Finally, here we discuss the selection of the samples for thr SVM classification and ROI. In the SVM classification analysis in this study, in some calculations w combined samples from two experimental conditions. For example, when comparing the coherence decoding performance of the V1 and MT+ brain regions, increasing the number of samples may increase the accuracy. In future experiments, how should one coordinate the length of the experiment and the number of times the experimental conditions are presented is still a challenge. In principle, as much data as possible should be obtained for the SVM classification or other multivariate pattern analyses. Also, since V1 contains more voxels than MT+, directly comparing the classification results of these two brain regions is not the most reliable method. However, even though MT+ has a smaller volume than V1, this study still concludes that MT+ is better at decoding coherence information than V1, which further illustrates the advantages and importance of MT+ in the processing of motion coherence in the visual cortex. At the same time, this study also compares V1 based on the lower left quadrant or lower right quadrant with MT+. The size of V1 obtained through retinotopic mapping is only about a quarter of the entire V1, which also makes up for the difference in the size of the two brain regions to a certain extent.

6 Conclusions

In this study, two random dots experiments were designed, and the brain imaging data of healthy human subjects were collected based on 3T and 7T magnetic resonance scans, The data were analyzed by uni-variate analysis and multivariate pattern analysis. By comparing the two methods, we explained how MT+ and V1 play an important role in processing visual motion, more specifically, how they encode visual motion coherence and speed information.

In this study, uni-variate analysis was first used to investigate the activation of MT+ and V1 to random dots motion. The uni-variate analysis results of 3T and 7T showed that MT+ was more sensitive to coherent motion, while V1 was more sensitive to incoherent

motion. Second, this study also explores this issue from the perspective of multivariate pattern analysis, and the SVM-based classification results show that, regardless of using 3T or 7T functional imaging, MT+ and V1 have consistently demonstrated the ability to decode the coherence or incoherence of random dots motion. (i.e., the accuracy of distinguishing random dots as coherent or incoherent motion is significantly greater than the chance level of 50%), and MT+ (>70%) tends to show better decoding performance than V1 (about 60%). Therefore, this study not only verifies the role and difference of these two brain regions in the process of encoding visual motion information from the perspective of uni-variate analysis, but more importantly, the results of this study verify that MT+ encodes visual motion information from the perspective of multivariate pattern analysis.

The results of uni-variate analysis and SVM classification corroborate each other on the problem of encoding the coherence of motion of random dots. However, the results of the multivariate analysis were more sensitive than the uni-variate analysis to the problem of encoding the speed of motion of random dots. In uni-variate analysis, neither MT+ nor V1 constituted a significant difference in response to random-dot motion stimuli at different speeds, while SVM classification results based on 7T fMRI showed that V1 responded to four speeds (2.5, 5.0, 7.5 and 10.0 deg/s) random dots motion stimulation has robust discrimination capacity, and the MT+ obtained in the way of intersection based on atlas and functional localizer also shows a significantly higher speed classification accuracy than the chance level (25%). The above results validate the advantages of multivariate analysis in studying the decoding of visual motion information and suggest that MT+ of the dorsal visual pathway encodes visual motion coherence and speed based on multi-voxel pattern information. At the same time, the results also show that the 7T fMRI data with higher spatial resolution can still achieve similar or even better decoding performance relative to 3T data when the number of subjects is small, that is, the improvement of spatial resolution improves decoding performance for visual motion information.

The study also discusses the interplay of coherence and speed. The uni-variate results showed that V1 was most likely to exhibit significantly different responses to the two movements at 2.5 deg/s, followed by 5.0 and 10.0 deg/s; MT+ responses to coherent movements were stronger and the differences were more likely to be significant at speeds of 5.0 and 10.0 deg/s, whereas at 2.5 deg/s the levels of MT+ responses to coherent and incoherent movements remained similar. In multivariate analysis, the results of 3T showed that V1 was more likely to distinguish between coherent and incoherent random dots motion at 5.0 deg/s, and the results of 7T showed that only when the speed was 2.5 deg/s, the classification accuracy of V1 was slightly higher than the chance level (The result is not significant, but the accuracy rate of all subjects is greater than the chance level, so likely the determining factor is the small number of subjects). Under the conditions of 5.0 and 10.0 deg/s, the data from some subjects can decode the motion coherence, which indicates that V1 may have large individual differences in decoding coherence. Neither V1 nor MT+ showed significant decoding for motion coherence at 7.5 deg/s. Therefore, combining the results of uni-variate and multivariate analysis, speed does not have a very significant effect on the processing of MT+ or V1 encoding motion coherence and speed information, but MT+ still shows a preference for fast motion, while V1 shows preference for slow motion.

Finally, we explore whether MT+ and V1 are lateralized in encoding random dots motion coherence and speed information. V1 showed no significant lateralization in both 3T and 7T uni-variate and multivariate analyses. MT+ showed left-side dominance (i.e., greater response than the right) in some uni-variate analyzes and right-side dominance (i.e., higher decoding accuracy than the left) in some multivariate analyses. Although the left MT+ showed significantly greater responses than the right MT+ in some cases, the right MT+ responded differently to coherent and incoherent motion relative to the left MT+, so the right MT+ showed better decoding performance (distinguishing between coherent and incoherent motion). The levels of MT+ decoding coherence or speed did not show consistent left-right differences in multivariate analysis results. In summary, the results of

this study show that MT+ shows a certain degree of laterality, but the results are not consistent. Follow-up studies can increase the sample size for further exploration.

In conclusion, this study explores the encoding of visual motion information by MT+ and V1 of the dorsal visual pathway by combining uni-variate and multivariate analyses. Our results support the previous view that MT+ is more sensitive to coherent motion and V1 is more sensitive to incoherent motion. MT+ is more likely to distinguish between coherent and incoherent motion during fast motion (10.0 deg/s) while V1 has better decoding power at slow motion (2.5 deg/s). In addition, the experimental results indicate that MT+ and V1 do not have significant laterality in the process of encoding visual motion information. Finally, high-resolution fMRI at 7T shows improved decoding of visual motion information than low-resolution imaging at 3T.

7 Limitations

Based on fMRI at two different resolutions (3T and 7T), this study combined two different analysis methods (uni-variate analysis based on individual voxel activation values and multivariate pattern analysis based on SVM classification) to explore the encoding of visual motion information (mainly coherence and speed of motion) by the MT+ and V1 brain areas in the dorsal visual pathway. At present, preliminary conclusions have been obtained, that is, MT+ is highly sensitive to coherence through uni-variate and multivariate analysis methods, and MT+ and V1 encode coherence information based on complex activity patterns rather than a simple superposition of neural responses. In addition, multivariate analysis methods have been verified to be more sensitive than uni-variate analysis, and high-resolution fMRI imaging has also been shown to be beneficial to mining brain activity pattern information in this decoding analysis. However, although two experiments were designed in this study, the number of subjects in each experiment was small (12 subjects in the 3T experiment and 5 subjects in the 7T experiment), and the two experiments did not recruit the same group of subjects. Therefore, in subsequent experiments, it may be considered to recruit the same group of subjects for 3T and 7T fMRI data acquisition.

In addition, the number of tools that can currently support the processing of high-resolution fMRI data (especially partial brain data) is small. Existing tools such as SPM12 cannot achieve registration for partial brain fMRI data. The 7T data used in this study did not go through distortion correction, which may lead to errors in subsequent registration. Due to the high spatial resolution of 7T fMRI data, it is observed that there are fewer and scattered voxels with significant activation. Not only they cannot be well connected into a cluster, but they may also carry more noise, so valuable voxels may be excluded due to statistical insignificance during calibration. (Kriegeskorte & Bandettini, 2007) report high resolution (≤ 2 mm^3) fMRI imaging may provide inaccurate imaging of neural activity; they require smoothing for increasing the signal-to-noise ratio. However, this will lose the original image resolution. In this study, neither the 3T nor 7T data were smoothed. Although the activation in the 7T data was relatively scattered, the 7T data still showed its advantages in the multivariate analysis method.

References

Amano, K., Wandell, B. A., & Dumoulin, S. O. (2009). Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *Journal of Neurophysiology*, *102*(5), 2704-2718.

Antal, A., Polania, R., Saller, K., Morawetz, C., Schmidt-Samoa, C., Baudewig, J., Paulus, W., & Dechent, P. (2011). Differential activation of the middle-temporal complex to visual stimulation in migraineurs. *Cephalalgia*, *31*(3), 338-345.

Aspell, J., Tanskanen, T., & Hurlbert, A. (2005). Neuromagnetic correlates of visual motion coherence. *European Journal of Neuroscience*, *22*(11), 2937-2945.

Baker, C., Hess, R. F., & Zihl, J. (1991). Residual motion perception in a" motion-blind" patient, assessed with limited-lifetime random dot stimuli. *Journal of Neuroscience*, *11*(2), 454-461.

Bartels, A., Logothetis, N. K., & Moutoussis, K. (2008). fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. *Trends in neurosciences*, *31*(9), 444-453.

Bayerl, P., & Neumann, H. (2007). Disambiguating visual motion by form-motion interaction—a computational model. *International Journal of Computer Vision*, *72*(1), 27-45.

Becker, H. G., Erb, M., & Haarmeier, T. (2008). Differential dependency on motion coherence in subregions of the human MT+ complex. *European Journal of Neuroscience*, *28*(8), 1674-1685.

Bedny, M., Konkle, T., Pelphrey, K., Saxe, R., & Pascual-Leone, A. (2010). Sensitive period for a multimodal response in human visual motion area MT/MST. *Current Biology*, *20*(21), 1900-1906.

Benson, N. C., Butt, O. H., Brainard, D. H., & Aguirre, G. K. (2014). Correction of distortion in flattened representations of the cortical surface allows prediction of V1-V3 functional organization from anatomy. *PLoS computational biology*, *10*(3), e1003538.

Bex, P. J., & Dakin, S. C. (2002). Comparison of the spatial-frequency selectivity of local and global motion detectors. *JOSA A*, *19*(4), 670-677.

Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., & Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception*, *30*(1), 61-72.

Chawla, D., Phillips, J., Buechel, C., Edwards, R., & Friston, K. (1998). Speed-dependent motion-sensitive responses in V5: an fMRI study. *Neuroimage*, *7*(2), 86-96.

Cheng, K., Fujita, H., Kanno, I., Miura, S., & Tanaka, K. (1995). Human cortical regions activated by wide-field visual motion: an H2 (15) O PET study. *Journal of Neurophysiology*, *74*(1), 413-427.

Clarke, S., & Miklossy, J. (1990). Occipital cortex in man: Organization of callosal connections, related myelo-and cytoarchitecture, and putative boundaries of functional visual areas. *Journal of Comparative Neurology*, *298*(2), 188-214.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, *29*(3), 162-173.

De Martino, F., Zimmermann, J., Muckli, L., Ugurbil, K., Yacoub, E., & Goebel, R. (2013). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. *PloS one*, *8*(3), e60514.

DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., Miller, D., & Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the*

National Academy of Sciences, *93*(6), 2382-2386.

Dupont, P., Orban, G., De Bruyn, B., Verbruggen, A., & Mortelmans, L. (1994). Many areas in the human brain respond to visual motion. *Journal of Neurophysiology*, *72*(3), 1420-1424.

Fawcett, I. P., Hillebrand, A., & Singh, K. D. (2007). The temporal sequence of evoked and induced cortical responses to implied-motion processing in human motion area V5/MT+. *European Journal of Neuroscience*, *26*(3), 775-783.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, *1*(1), 1-47.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, *2*(4), 189-210.

Furlan, M., & Smith, A. T. (2016). Global motion processing in human visual cortical areas V2 and V3. *Journal of Neuroscience*, *36*(27), 7314-7324.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20-25.

Harrison, L. M., Stephan, K. E., Rees, G., & Friston, K. J. (2007). Extra-classical receptive field effects measured in striate cortex with fMRI. *Neuroimage*, *34*(3), 1199-1208.

Harvey, B. M., & Dumoulin, S. O. (2016). Visual motion transforms visual space representations peripherally throughout the human visual hierarchy. *Neuroimage*, *127*, 173-185.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425-2430.

Heeger, D. J., Boynton, G. M., Demb, J. B., Seidemann, E., & Newsome, W. T. (1999). Motion opponency in visual cortex. *Journal of Neuroscience*, *19*(16), 7162-7174.

Howard, R., Bullmore, E., Brammer, M., Williams, S., Mellers, J., Woodruff, P., David, A., Andrew, C., Allin, M., & Simmons, A. (1995). Activation of area V5 by visual perception of motion demonstrated with echoplanar MR imaging. *Magnetic resonance imaging*, *13*(6), 907-909.

Huang, T., Chen, X., Jiang, J., Zhen, Z., & Liu, J. (2019). A probabilistic atlas of the human motion complex built from large-scale functional localizer data. *Human brain mapping*, *40*(12), 3475-3487.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*(16), 7195-7205.

Huk, A. C., & Heeger, D. J. (2002). Pattern-motion responses in human visual cortex. *Nature neuroscience*, *5*(1), 72-75.

Ibbotson, M., & Jung, Y. J. (2020). Origins of functional organization in the visual cortex. *Frontiers in Systems Neuroscience*, *14*, 10.

Jiang, F., Stecker, G. C., & Fine, I. (2014). Auditory motion processing after early blindness. *Journal of Vision*, *14*(13), 4-4.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature neuroscience*, *8*(5), 679-685.

Kamitani, Y., & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology*, *16*(11), 1096-1102.

Kaneoke, Y. (2006). Magnetoencephalography: in search of neural processes for visual motion information. *Progress in neurobiology*, *80*(5), 219-240.

Kaneoke, Y., & Kakigi, R. (2006). Spatial integration of visual motion with separate speed and direction information. *Neuroreport*, *17*(18), 1841-1845.

Kawakami, O., Kaneoke, Y., Maruyama, K., Kakigi, R., Okada, T., Sadato, N., & Yonekura, Y. (2002). Visual detection of motion speed in humans: spatiotemporal analysis by fMRI and MEG. *Human brain mapping*, *16*(2), 104-118.

Kriegeskorte, N., & Bandettini, P. (2007). Analyzing for information, not activation, to exploit high-resolution fMRI. *Neuroimage*, *38*(4), 649-662.

Kriegeskorte, N., & Douglas, P. K. (2019). Interpreting encoding and decoding models. *Current opinion in neurobiology*, *55*, 167-179.

Kumano, H., & Uka, T. (2013). Responses to random dot motion reveal prevalence of pattern-motion selectivity in area MT. *Journal of Neuroscience*, *33*(38), 15161-15170.

Lam, K., Kaneoke, Y., Gunji, A., Yamasaki, H., Matsumoto, E., Naito, T., & Kakigi, R. (2000). Magnetic response of human extrastriate cortex in the detection of coherent and incoherent motion. *Neuroscience*, *97*(1), 1-10.

Lam, K., Kaneoke, Y., & Kakigi, R. (2003). Human cortical response to incoherent motion on a backgrun of coherent motion. *Neuroscience letters*, *347*(1), 41-44.

Lawrence, S. J., Formisano, E., Muckli, L., & de Lange, F. P. (2019). Laminar fMRI: Applications for cognitive neuroscience. *Neuroimage*, *197*, 785-791.

Laycock, R., Crewther, D. P., Fitzgerald, P. B., & Crewther, S. G. (2007). Evidence for fast signals and later processing in human V1/V2 and V5/MT+: a TMS study of motion perception. *Journal of Neurophysiology*, *98*(3), 1253-1262.

Liu, C., Guo, F., Qian, C., Zhang, Z., Sun, K., Wang, D. J., He, S., & Zhang, P. (2021). Layer-dependent multiplicative effects of spatial attention on contrast responses in human early visual cortex. *Progress in neurobiology*, *207*, 101897.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*(4853), 740-749.

Maruyama, K., Kaneoke, Y., Watanabe, K., & Kakigi, R. (2002). Human cortical responses to coherent and incoherent motion as measured by magnetoencephalography. *Neuroscience research*, *44*(2), 195-205.

Maruyama, M., Palomo, D. D., & Ioannides, A. A. (2009). Stimulus - contrast - induced biases in activation order reveal interaction between V1/V2 and human MT+. *Human brain mapping*, *30*(1), 147-162.

Masland, R. H. (2012). The neuronal organization of the retina. *Neuron*, *76*(2), 266-280.

Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, *49*(5), 1127-1147.

McKeefry, D., Watson, J., Frackowiak, R., Fong, K., & Zeki, S. (1997). The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. *Neuroimage*, *5*(1), 1-12.

Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology*, *55*(6), 1308-1327.

Mitchell, T. M., Hutchinson, R., Niculescu, R. S., Pereira, F., Wang, X., Just, M., & Newman, S. (2004). Learning to decode cognitive states from brain images. *Machine learning*, *57*(1-2), 145-175.

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., Goebel, R., & Yacoub, E. (2015). Contextual feedback to superficial layers of V1. *Current Biology*, *25*(20), 2690-2695.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS biology*, *3*(8), e265.

Nakamura, H., Kashii, S., Nagamine, T., Matsui, Y., Hashimoto, T., Honda, Y., & Shibasaki, H. (2003). Human V5 demonstrated by magnetoencephalography using random dot kinematograms of different coherence levels. *Neuroscience research*, *46*(4), 423-433.

Nakamura, M., Kaneoke, Y., Watanabe, K., & Kakigi, R. (2002). Visual information process in Williams syndrome: intact motion detection accompanied by typical visuospatial dysfunctions. *European Journal of Neuroscience*, *16*(9), 1810-1818.

Nakayama, K., Silverman, G. H., MacLeod, D. I., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception*, *14*(2), 225-238.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, *8*(6), 2201-2211.

Nishida, S. y., Kawabe, T., Sawayama, M., & Fukiage, T. (2018). Motion perception: From detection to interpretation. *Annual review of vision science*, *4*, 501-523.

Nishida, S. y., Sasaki, Y., Murakami, I., Watanabe, T., & Tootell, R. B. (2003). Neuroimaging of direction-selective mechanisms for second-order motion. *Journal of Neurophysiology*, *90*(5), 3242-3254.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, *89*(13), 5951-5955.

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in neuroinformatics*, *10*, 27.

Orban, G. A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., Todd, J., & Vanduffel, W. (2003). Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia*, *41*(13), 1757-1768.

Orban, G. A., Sunaert, S., Todd, J. T., Van Hecke, P., & Marchal, G. (1999). Human cortical regions involved in extracting depth from motion. *Neuron*, *24*(4), 929-940.

Orban, G. A., Van Essen, D., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in cognitive sciences*, *8*(7), 315-324.

Qian, N., & Andersen, R. A. (1994). Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *Journal of Neuroscience*, *14*(12), 7367-7380.

Rees, G., Friston, K., & Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. *Nature neuroscience*, *3*(7), 716-723.

Rosenke, M., van Hoof, R., van den Hurk, J., Grill-Spector, K., & Goebel, R. (2021). A probabilistic functional atlas of human occipito-temporal visual cortex. *Cerebral Cortex*, *31*(1), 603-619.

Saito, H.-a., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience*, *6*(1), 145-157.

Schwarzkopf, D. S., Sterzer, P., & Rees, G. (2011). Decoding of coherent but not incoherent motion signals in early dorsal visual cortex. *Neuroimage*, *56*(2), 688-698.

Seiffert, A. E., Somers, D. C., Dale, A. M., & Tootell, R. B. (2003). Functional MRI studies of human visual motion perception: texture, luminance, attention and after-effects. *Cerebral Cortex*, *13*(4), 340-349.

Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: integration of color, form, and motion. *Annu. Rev. Neurosci.*, *28*, 303-326.

Smith, A., Wall, M., Williams, A., & Singh, K. D. (2006). Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience*, *23*(2), 561-569.

Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surruning context. *Proceedings of the National Academy of Sciences*, *107*(46), 20099-20103.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience*, *11*(9), 2768-2785.

Stephan, K. E., Petzschner, F. H., Kasper, L., Bayer, J., Wellstein, K. V., Stefanics, G., Pruessmann, K. P., & Heinzle, J. (2019). Laminar fMRI and computational theories of brain function. *Neuroimage*, *197*, 699-706.

Tanaka, E., Noguchi, Y., Kakigi, R., & Kaneoke, Y. (2007). Human cortical response to various apparent motions: a magnetoencephalographic study. *Neuroscience research*, *59*(2), 172-182.

Taylor, C. M., Olulade, O. A., Luetje, M. M., & Eden, G. F. (2018). An fMRI study of coherent visual motion processing in children and adults. *Neuroimage*, *173*, 223-239.

Tootell, R. B., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., & Dale, A. M. (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences*, *95*(3), 811-817.

Tootell, R. B., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., Brady, T. J., & Rosen, B. R. (1995). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature*, *375*, 139-141.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R., & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, *15*(4), 3215-3230.

Tootell, R. B., & Taylor, J. B. (1995). Anatomical evidence for MT and additional cortical visual areas in humans. *Cerebral Cortex*, *5*(1), 39-55.

Ulbert, I., Karmos, G., Heit, G., & Halgren, E. (2001). Early discrimination of coherent versus incoherent motion by multiunit and synaptic activity in human putative MT+. *Human brain mapping*, *13*(4), 226-238.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What'and 'where'in the human brain. *Current opinion in neurobiology*, *4*(2), 157-165.

Van de Grind, W., Koenderink, J., & Van Doorn, A. (1987). Influence of contrast on foveal and peripheral detection of coherent motion in moving random-dot patterns. *JOSA A*, *4*(8), 1643-1652.

Wang, H. X., Merriam, E. P., Freeman, J., & Heeger, D. J. (2014). Motion direction biases and decoding in human visual cortex. *Journal of Neuroscience*, *34*(37), 12601-12615.

Wang, L., Mruczek, R. E., Arcaro, M. J., & Kastner, S. (2015). Probabilistic maps of visual topography in human cortex. *Cerebral Cortex*, *25*(10), 3911-3931.

Ward, L. M., Morison, G., Simmers, A. J., & Shahani, U. (2018). Age-related changes in global motion coherence: Conflicting haemodynamic and perceptual responses. *Scientific Reports*, *8*(1), 10013.

Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S., & Zeki, S. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, *3*(2), 79-94.

Werring, D., Clark, C., Parker, G., Miller, D., Thompson, A., & Barker, G. (1999). A direct demonstration of both structure and function in the visual system: combining diffusion tensor imaging with functional magnetic resonance imaging. *Neuroimage*, *9*(3), 352-361.

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *ACM SIGGRAPH Computer Graphics*, *18*(1), 24-24.

Zeki, S., Watson, J., Lueck, C., Friston, K. J., Kennard, C., & Frackowiak, R. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience*, *11*(3), 641-649.