



HAL
open science

Spatial and temporal integration of binocular disparity in the primate brain.

Yseult Héjja-Brichard

► **To cite this version:**

Yseult Héjja-Brichard. Spatial and temporal integration of binocular disparity in the primate brain.. Neuroscience. Université Toulouse 3 - Paul Sabatier, 2020. English. NNT: . tel-02956425

HAL Id: tel-02956425

<https://hal.science/tel-02956425>

Submitted on 2 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par
Yseult HEJJA--BRICHARD

Le 10 juin 2020

**Spatial and temporal integration of binocular disparity in the
primate brain**

Ecole doctorale : **CLESCO - Comportement, Langage, Education, Socialisation,
Cognition**

Spécialité : **Neurosciences, comportement et cognition**

Unité de recherche :

CERCO - Centre de Recherche Cerveau et Cognition

Thèse dirigée par
Benoit COTTEREAU

Jury

M. Peter JANSSEN, Rapporteur
Mme Claire WARDAK, Rapporteur
Mme Kristine KRUG, Examinatrice
M. Jean-Marc DEVAUD, Examineur
M. Benoit R. COTTEREAU, Directeur de thèse

SPATIAL AND TEMPORAL INTEGRATION OF
BINOCULAR DISPARITY IN THE PRIMATE BRAIN

Intégration spatiale et temporelle des disparités binoculaires chez le primate

Acknowledgements

Before anything, I would like to express my sincere gratitude to the members of my thesis committee: Prof. Peter Janssen, Dr. Claire Wardak, and Prof. Kristine Krug who, despite the postponing of my defence by seven long months have nonetheless accepted to examine and review my work; and Prof. Jean-Marc Devaud who kindly accepted to join the final version of it and is also symbolically closing a loop, as I arrived in Toulouse to enrol in the master's degree he was co-directing.

I would like to say how grateful I am to Benoit, who supervised me as a master's student and then as his first PhD student. Thank you for trusting me and giving me the right amount of autonomy whilst being always available, for letting me go at my own pace, and for your guidance throughout my PhD studies. Thank you for being so optimistic when I was rather sceptical or a bit demotivated, for pushing me when I needed it, for providing me with the right framework to express my doubts, my hesitations, but also to encourage me to develop my own ideas.

I had a great time at the lab and felt very supported by the Eco-3D team. I am truly grateful for the openness, the kindness, and the support of all of its members and for whom I would like to have a word:

Jean-Baptiste for his priceless help with the analyses, his advice, his availability

Youn for his extensive knowledge of the literature, for his listening

Alexandra for introducing to me the teaching side of academia.

Tushar for the endless conversations we had, for dealing with me and my many questions, for your creativity regarding new experiment ideas, for your enthusiasm

Silvia who I was happy to work with and who helped me a lot for running the human experiments when I had to focus on writing my thesis!

Vanessa who joined the team more lately with her energy and dedication!

Pauline who has picked up the torch to further help develop monkey fMRI in the team.

It was also quite pleasant to see more women joining the team!

I am also thinking of former PhD students and post-docs. Special thanks to Samy who set the path to my PhD and took time to guide me through the different steps, for the many philosophical discussions. Marcello and Volodymyr with whom I shared the office, and for the many debates and discussions that took place in that office; Amirouche as well. Mylène and Emilie who I started this adventure with.

Thank you to my closest toulousain friends and lab mates, Chadlia and Nicolas, with whom I shared my frustrations, my joy, my reliefs; for putting up with me at every moment.

Thank you Manu, Petit Pascal, Jean-Michel for the discussions we had about how to change the world, for sometimes turning words into actions; the PAF students: Sam, Can, Bhavin, Zhaoyang, Milad, Benjamin for the restaurant sessions, and the organisation of events within and outside the lab; to the second floor students, present despite geographical constraints, especially to Mariam, Danaé, Ludovic, and for their involvement in the lab life (which really matters to me, as some may have noticed!), Anna with whom I shared some teaching stress.

Thank you, Marine who I started my journey in Toulouse with, balancing between heavy fun and light seriousness! To our long phone calls across the country and now the ocean!

Thank you, Emilie Rapha for your priceless help and your everyday investment, Camille Lejards for teaching me in the first place how to work with macaques and for believing I could talk with them.

Obviously, I now have to mention Nikita and Tunisie, my two monkey mates, who I worked with during those 5 years. That was not an easy task but it was clearly worth it!

I would like to thank the MRI facility staff, Nathalie, but also Hélène and the 'manip radio' team, Jean-Pierre, Yohan, and Fred.

But also, the several persons in charge of less obvious aspects of research, the IT team for both their technical and moral support: Carmen, Maxime, Damien, Joel; and Claire for dealing with all the financial and some administrative aspects of research.

I also want to thank all the other CerCo members, former and current, that I did not mention but who clearly played a big role into making this place a vibrant place for doing a PhD and for growing up as a future – and hopefully good – researcher. And Simon for giving so much freedom to the students as the director of the lab; this is very valuable and rewarding!

I would like to thank the teaching team of the 'section 69' for welcoming me as a teaching assistant during my first three years.

Merci à InCOGnu and its enthusiastic members who I got to spend time with: Nabila, Lucille, Laura, Joseph, Nawelle, Simon, Mehdi, Alice, Bonnie, Audric, Maxime, Mawa, Amélie, Quentin, Lisa. I had fun during our science popularisation events, it brought me a lot and helped me figure out this is something that really matters to me.

I also want to mention the different persons who paved the way that led me to do this PhD: Natacha Mendès who took me as a trainee research assistant on her postdoc project seven years ago and introduced me to the wonderful world of working with non-human primates; Eugénie Lhommée with whom I did my first clinical internship and who supported me when I chose to study fundamental research instead of clinical neuropsychology; Olivier Pascalis who offered me my first research internship.

A big thought goes to Slowpen who gave me hope into a more human science framework, Amélie, Brice, Lad, and Bertrand, for those countless discussions we had, for those weekends and skype meetings we spent together thinking about how to improve science practices, for teaching me Bayesian philosophy, and for so many other valuable things. I truly value those moments and my friendship with you!

Thanks to my old friends Laureline and Fabien. Even though we don't meet often (and not often enough), you bring me lots and keep me rooted in the 'outside world'.

Endlich, bin ich in Gedanken bei meiner liebe leipzigerinnen Freundinnen: Clo, Moira, Sophie, Anna und Sam!

Thank you to my coolest-roommates-ever, Laura and Valentine, with whom I spent lots of time discussing about how to start a world revolution but who also had to put up with my fluctuating mood. A big hug to Babou the 'cat king' as well. You were like my second family, which means a lot to me!

Thank you to my close family members, and especially my mother, always very proud and supportive, my grandma who makes me constantly look further, and my 'little' brother who finally found his way to do what he always wanted to.

I will close this section by expressing all my gratitude and my heart feelings to the person I am happily walking by every single day, who brings me into a quiet mindset from where I can enjoy the soothing sound of the ebbing tide and the opaque sun light. There is no annoying wind with you, Andrea.

Abstract

The primate visual system strongly relies on the small differences between the two retinal projections to perceive depth. However, it is not fully understood how those binocular disparities are computed and integrated by the nervous system. On the one hand, single-unit recordings in macaque give access to neuronal encoding of disparity at a very local level. On the other hand, functional neuroimaging (fMRI) studies in human shed light on the cortical networks involved in disparity processing at a macroscopic level but with a different species. In this thesis, we propose to use an fMRI approach in macaque to bridge the gap between single-unit and fMRI recordings conducted in the non-human and human primate brain, respectively, by allowing direct comparisons between the two species. More specifically, we focused on the temporal and spatial processing of binocular disparities at the cortical but also at the perceptual level. Investigating cortical activity in response to motion-in-depth, we could show for the first time that 1) there is a dedicated network in macaque that comprises areas beyond the MT cluster and its surroundings and that 2) there are homologies with the human network involved in processing very similar stimuli. In a second study, we tried to establish a link between perceptual biases that reflect statistical regularities in the three-dimensional visual environment and cortical activity, by investigating whether such biases exist and can be related to specific responses at a macroscopic level. We found stronger activity for the stimulus reflecting natural statistics in one subject, demonstrating a potential influence of spatial regularities on the cortical activity. Further work is needed to firmly conclude about such a link. Nonetheless, we robustly confirmed the existence of a vast cortical network responding to correlated disparities in the macaque brain. Finally, we could measure for the first time retinal corresponding points on the vertical meridian of a macaque subject performing a behavioural task (forced-choice procedure) and compare it to the data we also collected in several human observers with the very same protocol. In the discussion sections, we showed how these findings open the door to varied perspectives.

Keywords: fMRI – stereopsis – motion-in-depth – slant perception – horopter – non-human primate

Résumé

Le système visuel du primate s'appuie sur les légères différences entre les deux projections rétiniennes pour percevoir la profondeur. Cependant, on ne sait pas exactement comment ces disparités binoculaires sont traitées et intégrées par le système nerveux. D'un côté, des enregistrements unitaires chez le macaque permettent d'avoir accès au codage neuronal de la disparité à un niveau local. De l'autre côté, la neuroimagerie fonctionnelle (IRMf) chez l'humain met en lumière les réseaux corticaux impliqués dans le traitement de la disparité à un niveau macroscopique mais chez une espèce différente. Dans le cadre de cette thèse, nous proposons d'utiliser la technique de l'IRMf chez le macaque pour permettre de faire le lien entre les enregistrements unitaires chez le macaque et les enregistrements IRMf chez l'humain. Cela, afin de pouvoir faire des comparaisons directes entre les deux espèces. Plus spécifiquement, nous nous sommes intéressés au traitement spatial et temporel des disparités binoculaires au niveau cortical mais aussi au niveau perceptif. En étudiant l'activité corticale en réponse au mouvement tridimensionnel (3D), nous avons pu montrer pour la première fois 1) qu'il existe un réseau dédié chez le macaque qui contient des aires allant au-delà du cluster MT et des aires environnantes et 2) qu'il y a des homologies avec le réseau trouvé chez l'humain en réponse à des stimuli similaires. Dans une deuxième étude, nous avons tenté d'établir un lien entre les biais perceptifs qui reflètent les régularités statistiques 3D dans l'environnement visuel et l'activité corticale. Nous nous sommes demandés si de tels biais existent et peuvent être reliés à des réponses spécifiques au niveau macroscopique. Nous avons trouvé de plus fortes activations pour le stimulus reflétant les statistiques naturelles chez un sujet, démontrant ainsi une possible influence des régularités spatiales sur l'activité corticale. Des analyses supplémentaires sont cependant nécessaires pour conclure de façon définitive. Néanmoins, nous avons pu confirmer de façon robuste l'existence d'un vaste réseau cortical répondant aux disparités corrélées chez le macaque. Pour finir, nous avons pu mesurer pour la première fois les points rétiens correspondants au niveau du méridien vertical chez un sujet macaque qui réalisait une tâche comportementale (procédure à choix forcé). Nous avons pu comparer les résultats obtenus avec des données également collectées chez des participants humains avec le même protocole. Dans les différentes sections de discussion, nous montrons comment nos différents résultats ouvrent la voie à de nouvelles perspectives.

Mots-clés : IRMf – stéréopsie – mouvement 3D – perception des orientations – horoptère – primates non humains

Résumé substantiel en langue française

Présente chez de nombreuses espèces animales, la vision stéréoscopique, ou vision tridimensionnelle, repose sur l'intégration par le système nerveux des différences entre les projections rétiniennes bidimensionnelles d'une scène visuelle. Ces disparités binoculaires, liées au décalage entre nos deux yeux, nous renseignent de façon très précise sur les distances et les profondeurs des éléments constituant une scène visuelle. Avoir accès à ce type d'informations nous permet notamment de localiser et manipuler des objets et donc d'interagir avec notre environnement. Malgré son importance, tous les mécanismes sous-tendant l'extraction et le traitement de l'information tridimensionnelle par notre système visuel ne sont pas bien compris. Les connaissances parcellaires que nous en avons proviennent principalement de l'étude des réponses neuronales à la disparité binoculaire enregistrées dans les aires visuelles de bas niveau et intermédiaires chez le macaque, espèce traditionnellement utilisée comme modèle du système visuel humain. Plus récemment, avec l'avènement des techniques de neuroimagerie, les réseaux impliqués dans le traitement des disparités ont également pu être étudiés chez l'humain. Néanmoins, du fait de la différence de technique d'enregistrement, unitaire d'une part, et macroscopique de l'autre, il reste compliqué de faire des comparaisons inter-espèces. Durant ma thèse, j'ai étudié l'intégration temporelle et spatiale de la disparité binoculaire à un niveau cortical et à un niveau perceptif chez le macaque ainsi que, de façon plus limitée, chez l'humain, en proposant une approche expérimentale qui permette de pallier au besoin d'établir les homologues et différences qu'il pourrait y avoir entre le macaque et l'humain.

Dans un premier chapitre théorique, je propose une introduction générale sur le système visuel qui inclue une explication détaillée du chemin parcouru par un photon de lumière lorsqu'il entre en contact avec la rétine. Je décris ensuite l'anatomie de l'œil ainsi que les différentes cellules qui composent les voies visuelles avant l'entrée de l'information dans le

centre nerveux central, je passe également en revue la phototransduction, phénomène qui permet de transposer le signal électromagnétique du photon à un signal membranaire interprétable par les cellules du système visuel. Je détaille ensuite les deux voies visuelles traditionnellement décrites : la voie magnocellulaire, aussi appelée voie du mouvement ou voie du 'où' et plutôt impliquée dans la vision grossière de par les cellules à grands champs récepteurs qui la composent ; et la voie parvocellulaire, voie des objets, ou voie du 'quoi', constituée d'une multitude de cellules à petits champ récepteur, offrant une vision plus détaillée de la scène visuelle. Je clos ce premier chapitre en introduisant la notion de stéréopsie, ou vision de la profondeur, présente chez de nombreuses espèces du règne animal, et qui constitue le cœur de mes travaux de thèse.

Le chapitre deux, également théorique, offre une revue de la littérature sur l'intégration des disparités binoculaires par le système visuel du primate, en s'appuyant notamment sur des travaux réalisés chez le macaque. Les disparités binoculaires sont un indice de profondeur et de distance utilisées par le système visuel pour extraire l'information tridimensionnelle d'une scène visuelle. Bien qu'il existe de nombreux autres indices, principalement monoculaires (perspectives, ombrages, textures, etc.), les disparités binoculaires permettent au système visuel d'avoir une estimation fine de la structure tridimensionnelle et sous-tendent, au moins chez les primates mais aussi chez d'autres espèces comme le corbeau de Nouvelle Calédonie, l'exécution et le contrôle de gestes précis tels qu'attraper ou manipuler des objets. L'intégration puis le traitement de ces disparités binoculaires auxquelles répondent une multitude d'aires cérébrales représente néanmoins un challenge pour le système visuel de par la nécessité de mettre en correspondance les projections d'une scène visuelle sur chaque rétine. En effet, il n'est pas du tout évident de trouver l'endroit au niveau de chaque rétine sur lequel se projette un point de l'espace visuel. Plusieurs modèles informatiques ont été proposés pour comprendre comment le système visuel résout ce problème de correspondance binoculaire. Parmi eux, le

modèle binoculaire d'énergie, développé par Ohzawa et collaborateurs (1990) et évoqué dans ce deuxième chapitre, propose que les neurones complexes de l'aire visuelle primaire (V1) intègrent l'information apportée par des cellules binoculaires simples arrangées en paires de neurones mutuellement inhibitrices. Cette organisation hiérarchique qui repose sur une addition des différentes réponses émises par plusieurs cellules simples offre une solution, parmi d'autres potentielles, au problème de correspondance et suggère que les cellules complexes agissent comme des détecteurs de corrélation entre les deux projections rétiniennes. Ce modèle contient des limitations qui ne permettent pas de le généraliser à toutes les conditions visuelles testées ni aux aires visuelles situées au-delà de l'aire V1. Par exemple, les réponses du système visuel à des points anti-corrélés, points correspondants dans chaque rétine mais à la polarité inversée et qui ne donnent pas de percept 3D, ne sont pas bien prédites par le modèle. Néanmoins, il a le mérite de fournir un cadre de travail intéressant et dont les hypothèses peuvent être testées et affinées. À la suite de la présentation de ce modèle, je précise qu'il existe trois niveaux d'ordre des disparités binoculaires et que ces différents niveaux de profondeur ne sont pas traités de façon identique ni par les mêmes aires visuelles. Concrètement, les disparités binoculaires peuvent être arrangées spatialement de différentes façons et tantôt représenter une simple surface parallèle au plan de fixation, configuration qui correspond à un niveau de profondeur d'ordre zéro, tantôt être organisées sous forme de gradients et prendre la forme de surface orientées ou inclinées (rencontrées dans les scènes visuelles), configuration de premier ordre, qui peuvent contenir des courbures ou des irrégularités dans la surface (c'est le cas pour les objets qui nous entourent) devenant alors de second ordre. Plusieurs travaux de neuroimagerie (IRM fonctionnelle) réalisés à la fois chez l'humain et chez le macaque ont pu montrer que plus l'organisation des disparités est complexe et plus les aires cérébrales qui seront actives pour ces arrangements de disparités sont avancées dans la hiérarchie du système visuel. Il y a ainsi une spécialisation de certaines aires cérébrales, qui se développe au fur et à

mesure que l'on monte en complexité et qui suit une certaine forme de hiérarchie du système visuel.

Partant ensuite du constat que l'intégration spatiale des disparités binoculaires par le système visuel a été relativement bien documentée chez le primate, j'argumente qu'il y a d'autres questions, complémentaires, qu'il est intéressant de se poser mais qui ont reçu moins d'attention. C'est le cas notamment de l'influence des statistiques naturelles sur le traitement et la perception de ces différentes configurations de disparités. Dans le cadre de ma thèse, je propose d'étudier la façon dont la fréquence d'apparition de différentes orientations et inclinaisons dans les scènes visuelles (premier ordre de profondeur) peut potentiellement impacter leur traitement par le système visuel. En d'autres termes, pourrait-il y avoir une différence au sein des aires cérébrales habituellement impliquées dans le traitement visuel de ces surfaces selon leur fréquence dans les scènes naturelles ? Je suggère également d'étendre cette question à la perception visuelle en étudiant l'influence de ces statistiques naturelles sur la mise en correspondance binoculaire de stimuli tridimensionnels simples.

Enfin, je mets en avant une dimension primordiale dans l'intégration corticale des disparités binoculaires et qui n'a été que trop peu abordée dans la littérature : sa dimension temporelle. En effet, si la vision tridimensionnelle est utilisée pour la manipulation d'objets, elle est aussi extrêmement utile pour la locomotion ou encore la détection et l'évitement d'objets mobiles. Ces dernières nécessitent d'intégrer les disparités binoculaires non seulement à travers l'espace mais aussi à travers le temps. À travers un paradigme de neuroimagerie cérébrale inspiré de deux études s'étant récemment intéressé à ce sujet chez l'humain, je propose d'étudier les aires cérébrales impliquées dans le traitement temporel des disparités en observant les réponses corticales à des stimuli décrivant un mouvement dans la profondeur.

Après une introduction générale sur le système visuel et après avoir passé en revue les principales connaissances accumulées jusqu'à présent sur l'intégration des disparités binoculaire par le système visuel du primate et présenté les objectifs généraux de la thèse, j'introduis l'approche expérimentale principalement utilisée dans le cadre de mes travaux de recherche dans un troisième chapitre. Nous proposons en effet d'utiliser la neuroimagerie fonctionnelle (IRMf) chez le primate non humain afin de pouvoir faire le lien entre les études unitaires pratiquées chez le macaque et les études de neuroimagerie effectuées chez l'humain. Cette approche représente un challenge en soi de par la nécessité de devoir conditionner les animaux à rester immobiles dans le scanner et à effectuer une tâche comportementale, mais aussi de par le besoin de développer des analyses et outils spécifiques pour l'espèce étudiée afin d'avoir une bonne mesure du signal. Dans ce chapitre méthodologique, je mentionne les principes physiques fondamentaux de l'imagerie par résonance magnétique (IRM) et les différents types de séquences qui peuvent être utilisées pour générer des contrastes, qui reflètent différents niveaux d'intensité du signal de résonance magnétique dans les tissus étudiés et parmi eux le contraste BOLD utilisé dans le cas de l'IRMf. Je détaille ensuite la nature du signal mesuré dans nos études en neuroimagerie, le signal BOLD, qui repose sur une comparaison des niveaux d'hémoglobine oxygénée et désoxygénée présentes dans le sang. J'évoque les différences entre le signal BOLD et les autres signaux d'activités cérébrales qui sont traditionnellement mesurés chez le macaque : les potentiels de champs locaux et les potentiels d'action émis par une ou plusieurs cellules nerveuses. Cela me permet ensuite d'enchaîner sur le développement de l'IRMf chez le macaque, des différentes stratégies adoptées par les groupes de recherche pratiquant cette technique chez le singe et de sa pertinence pour faire le lien entre les études d'électrophysiologie réalisées chez le macaque, et qui enregistrent un signal plus précis mais très local (potentiels de champs locaux et potentiels

d'actions neuronales) et les études de neuroimagerie réalisées chez l'humain, qui donnent un aperçu plus global de l'activité cérébrale.

Je souligne le fait que des traitements spécifiques à la fois du signal et statistiques ont dû et sont continuellement en train d'être développés pour rendre possible la neuroimagerie fonctionnelle chez le macaque. Par ailleurs, l'idée de l'IRMf est de considérer le niveau d'oxygénation des aires cérébrales comme marqueur de leur activité, il s'agit donc d'une mesure indirecte et qui présente un décalage temporel spécifique, représentée par une fonction de réponse hémodynamique (HRF). Afin de mieux estimer l'activité cérébrale en réponse à des stimuli, il est nécessaire de déterminer la dynamique de la réponse BOLD et donc la HRF. Je présente la façon dont nous avons procédé pour estimer la HRF dans chacun de nos sujets macaques et comment nous l'implémentons dans nos modèles statistiques. Comme preuve de concept, je développe les résultats de deux études qui ont été réalisées dans l'équipe et auxquelles j'ai en partie participé. La première étude est un travail d'équipe réalisé sur plusieurs années et qui a permis le développement de l'IRM singe au CerCo. Il s'agit d'un travail de recherche portant sur les aires cérébrales impliquées dans le traitement du flux optique. Ce travail, qui a fait l'objet d'une publication en 2017 (*Cerebral Cortex*) et a pu montrer qu'il existe une différence de traitement du flux optique selon que celui-ci est compatible ou non avec la locomotion. À la suite de ces débuts fructueux et dans le cadre de la thèse d'un précédent doctorant, la seconde étude, en cours de révision, a donné de nouveaux outils d'analyse à l'équipe en s'intéressant aux propriétés rétinotopiques des aires visuelles et notamment au sein du cortex pariétal postérieur. J'ai ainsi pu réutiliser les délimitations rétinotopiques des aires visuelles pour réaliser certaines analyses de mes données fonctionnelles.

Le chapitre suivant (chapitre 4) est le premier chapitre expérimental. Il porte sur une forme d'intégration temporelle des disparités binoculaires. Il s'agit d'une étude en

neuroimagerie fonctionnelle, présentée à plusieurs conférences internationales et qui a été acceptée pour publication dans le journal *Cerebral Cortex*. Nous avons adapté deux études réalisées chez l'humain afin de caractériser chez le macaque le réseau cortical impliqué dans le traitement du mouvement tridimensionnel, mouvement défini ici par un changement des valeurs de disparités binoculaires à travers le temps. Nous avons contrasté les réponses BOLD à deux conditions expérimentales pour lesquelles l'information tridimensionnelle n'était disponible que lorsque vue par les deux yeux simultanément. La première condition représentait un mouvement 3D tel qu'émis par un objet se rapprochant ou s'éloignant de l'observateur tandis que la deuxième, la condition contrôle, nous avons mélangé les images vidéo de la première condition avec pour résultat une alternation saccadée de plans dans l'espace. Nos analyses ont permis de montrer qu'au sein de chaque hémisphère pour les deux sujets macaques ayant fait l'expérience trois aires ont été sélectivement plus activées par le mouvement 3D (appelé Cyclopean StereoMotion ou CSM dans notre étude) que par son contrôle. Ces aires étaient localisées sur la berge inférieure du sillon temporal supérieur (CSM_{STS}), sur le gyrus inféro-temporal (CSM_{ITG}), et dans la partie caudale du sillon intrapariétal postérieur (CSM_{PPC}). À l'aide d'analyses rétinotopiques, nous avons pu montrer que l'aire CSM_{STS} ne se trouvait pas dans le cluster MT, un groupe de 4 aires cérébrales associées au traitement du mouvement, mais qu'elle se trouvait à une position plus antérieure. Nous avons également démontré que certaines des aires de ce cluster MT avaient des réponses significativement plus élevées pour le mouvement 3D que pour le stimulus contrôle, et notamment les aires MSTv et FST. Nous n'avons pas trouvé de préférence marquée pour le mouvement 3D dans l'aire MT allant dans le sens des résultats obtenus lors d'enregistrements électrophysiologiques chez le macaque et qui suggèrent une sélectivité au mouvement 3D principalement basée sur un autre indice : les différences de vitesse interoculaires. En utilisant des stimuli révélant la sélectivité au mouvement planaire (ou mouvement bidimensionnel, 2D),

nous avons pu révéler qu'à la fois CSM_{STS} et CSM_{ITG} répondaient au mouvement 2D, mais que CSM_{PPC} semblait montrer une sélectivité unique pour le mouvement 3D.

De façon intéressante, notre aire CSM_{ITG} se trouve à la jonction ou à l'intérieur, selon les hémisphères testés, de l'aire V4 et potentiellement de l'aire V4A, aire documentée comme répondant à la disparité et au mouvement et appartenant à un autre cluster, que nous n'avons malheureusement pas pu définir à partir de nos données. On peut néanmoins souligner un résultat similaire dans une étude en IRMf réalisée chez des sujets humains qui a trouvé que les aires LO-1 et 2, considérées comme potentielles aires homologues de V4A, répondaient aussi au mouvement 3D, appuyant davantage une possible homologie.

Enfin, la troisième aire, CSM_{PPC}, du fait de sa localisation en amont d'aires impliquées dans le traitement des structures 3D et sensibles à la profondeur kinétique, pourrait être une aire complémentaire à l'aire caudale intra-pariétale (CIP) située sur la berge opposée du sillon intra-pariétal postérieur. En effet, l'aire CIP n'a pas montré de sensibilité à la profondeur kinétique, que l'on peut rapprocher du mouvement 3D, mais est connue pour répondre aux orientations spatiales et arrangements d'éléments dans l'espace. L'aire CSM_{PPC} pourrait ainsi être son pendant temporel du traitement des disparités binoculaires.

Plus globalement et considérés dans leur ensemble, nos résultats suggèrent que les aires corticales impliquées dans le traitement du mouvement 3D sont partiellement communes au macaque et à l'humain.

Dans le second chapitre expérimental (chapitre 5), deux études sont présentées, une étude en neuroimagerie chez le macaque et une étude de psychophysique à la fois chez le macaque et chez l'humain. Toutes deux ont vocation à déterminer si le traitement spatial des disparités, déjà assez bien documenté, ne pourrait pas être influencé par les régularités spatiales présentes dans les scènes visuelles. Afin de tester cette hypothèse, nous avons, dans une

première expérience, enregistré le signal BOLD en réponse à différentes surfaces orientées dans la profondeur. Ces surfaces étaient inclinées soit autour de l'axe horizontal ('slants') vers l'avant ou vers l'arrière, soit autour de l'axe vertical ('tilts') vers la gauche ou vers la droite. Nous avons contrasté les réponses à ces orientations avec celles à une condition contrôle pour laquelle il n'y avait pas de percept 3D (condition décorrélée) mais qui était identique aux autres conditions lorsqu'elles étaient vues de façon monoculaire, prenant alors la forme de simples nuages de points. Nous avons dans un premier temps cartographié l'étendue du réseau répondant aux disparités corrélées en opposant nos orientations 3D, que celles-ci soient autour de l'axe horizontal ou de l'axe vertical, à la condition décorrélée. Afin de préciser l'importante étendue des aires significativement activées pour les disparités corrélées, nous avons réalisé une analyse rétinotopique dans les aires visuelles précoces (V1, V2, V3, V4), au sein du cluster MT (MT, MSTv, V4t, FST), au sein du cluster PIP récemment redéfini par un ancien doctorant de l'équipe (PIP1, PIP2, CIP1, CIP2), et au niveau de trois aires plus dorsales (V3A, V6, V6A). De façon inattendue, nous avons trouvé de légères différences selon le type d'orientation avec des activations plus fortes pour les disparités corrélées lorsqu'on considérait les orientations autour de l'axe horizontal ('slant') que lorsqu'on considérait les orientations autour de l'axe vertical ('tilt'). En considérant uniquement les aires répondant de façon systématique aux disparités corrélées pour les deux types d'orientation, nous avons pu trouver un réseau similaire à celui décrit dans la littérature avec notamment les aires MT, MSTv, CIP1, CIP2, PIP2, V3A et V6A qui répondaient plus fortement aux disparités corrélées qu'aux disparités non corrélées. De façon intéressante, l'aire PIP1 ne ressortait jamais dans plus de 2 hémisphères sur 4 comme répondant plus fortement aux disparités corrélées quel que soit le type d'orientation considérée, contrastant ainsi avec les trois autres aires du cluster PIP qui elles répondaient de façon systématique aux disparités corrélées en considérant les deux types d'orientations 3D.

Dans un deuxième temps, nous nous sommes intéressés à l'influence des statistiques naturelles sur les activations cérébrales en réponses aux orientations 3D. Afin de déterminer si les réponses étaient plus fortes pour l'orientation la plus fréquemment rencontrée dans l'environnement, nous avons uniquement considéré les aires précédemment décrites comme répondant de façon systématique aux disparités corrélées sans influence de l'orientation 3D considérée. Afin de mettre en évidence une éventuelle préférence pour ce type d'inclinaison, nous avons comparé deux à deux nos deux types d'orientation ('slants' et 'tilts') : surfaces inclinées vers l'avant versus surfaces inclinées vers l'arrière, et surfaces orientées vers la droite versus surfaces orientées vers la gauche. Nous nous attendions à trouver un biais pour les surfaces inclinées vers l'arrière ('slant') avec des activations plus fortes pour l'inclinaison arrière que pour l'inclinaison avant, cette configuration arrière étant alignée avec l'orientation du sol dans les scènes naturelles et donc plus fréquente dans l'environnement visuel. En revanche, nous n'attendions aucune différence entre les surfaces orientées vers la gauche et vers la droite, ces orientations ne présentant pas de différence dans leurs fréquences d'apparition au sein des scènes visuelles. Étonnamment, lorsque nous avons considéré les cartes d'activations des contrastes BOLD effectués (avant versus arrière et droite versus gauche), nous avons trouvé des activations plus fortes pour l'inclinaison arrière chez un seul de nos sujets macaques, et une absence de différence chez le second sujet. Nous n'avons par contre pas trouvé de différence entre les orientations droite et gauche, et ce pour nos deux sujets, ce qui était attendu. Des analyses rétinotopiques chez le sujet présentant un biais pour les inclinaisons arrière ont révélé que c'était principalement au niveau du cluster PIP que les activations étaient les plus biaisées, c'est-à-dire que c'était au niveau des aires de ce cluster que les activations étaient plus fortes pour les surfaces inclinées vers l'arrière. Ces analyses n'ont pas permis de trouver de résultats similaires pour le deuxième sujet. Nous avons ensuite essayé de considérer l'élévation, en divisant nos stimuli d'inclinaisons 3D ('slants') en trois

parties : une partie supérieure, correspondant au champ visuel supérieur ($>2^\circ$), une partie médiane, correspondant à la zone centrale du champ visuel ($\pm 2^\circ$), et une zone inférieure, correspondant au champ visuel inférieur ($< -2^\circ$). L'idée était d'adresser l'hypothèse que considérer les activations au niveau du stimulus dans sa globalité ne permettait peut-être pas de révéler des biais pour certaines orientations. Nous avons donc essayé de voir si les scores de t obtenus pour chaque inclinaison (arrière et avant) étaient plus élevés pour les disparités dites croisées (donnant un niveau de profondeur en avant du point de fixation) dans le champ visuel inférieur, et plus élevés pour les disparités décroisées (niveau de profondeur en arrière du point de fixation) dans le champ visuel supérieur, ce qui correspond à une configuration de plan incliné vers l'arrière. Si nous avons à nouveau pu confirmer que notre premier sujet macaque montrait un biais pour les surfaces inclinées vers l'arrière, et non un biais général pour les disparités croisées, cette méthode n'a pas non plus permis de montrer que c'était aussi potentiellement le cas chez notre deuxième sujet. Nous ne sommes pas sûrs de l'origine d'une telle différence entre nos sujets. Des analyses supplémentaires seront donc nécessaires, telle que l'utilisation d'une approche multivariée par comparaison de motifs d'activation pour passer outre les limitations des analyses univariées que nous avons effectuées jusqu'à présent. À défaut de pouvoir conclure sur une influence des statistiques naturelles sur l'activité cérébrale, nous avons néanmoins pu décrire de façon robuste le réseau cortical impliqué dans le traitement spatial des disparités binoculaires chez nos deux sujets. Nous avons ainsi trouvé sept aires rétino-topiques qui répondaient plus fortement à nos conditions corrélées (orientations autour de l'axe horizontal et autour de l'axe vertical) qu'à la condition contrôle décorrélée, en accord avec la littérature.

La deuxième étude présentée dans ce chapitre expérimental est une étude de psychophysique réalisée à la fois chez l'humain et chez le macaque. L'objectif était de mesurer la composante verticale de l'horoptère de nos sujets, c'est-à-dire la région de l'espace dans laquelle la

stéréoacuité est la plus fine. Concrètement, il s'agit de déterminer la localisation des points rétiniens correspondants qui sont les points dans l'espace perçus comme superposés au niveau de leur projections rétiniennes et qui permettent de faire correspondre entre elles ces deux projections rétiniennes, donnant naissance au percept 3D. Plusieurs études chez l'humain ont suggéré que la position de ces points correspondants sur l'axe vertical, inclinés vers l'arrière par rapport au plan de fixation, pourrait refléter les statistiques naturelles. Bien que le macaque soit un modèle du système visuel humain, nous ne savons pas si cette inclinaison de l'horoptère vertical est aussi présente chez cette espèce. Nous avons donc adapté un protocole auparavant réalisé chez l'humain et mesuré expérimentalement la localisation des points correspondants chez un sujet macaque et chez huit sujets humains. La tâche expérimentale consistait à fixer un point central pendant qu'apparaissaient successivement et de façon brève deux barres présentées chacune dans un œil. À la suite de cette présentation de barre, les sujets devaient indiquer la direction du mouvement apparent perçu, mouvement résultant de l'apparition successive des deux barres. La distance entre les barres variait au cours du temps, soit de façon constante pour le sujet macaque, soit de façon adaptée via une procédure en staircase (variation de la distance des barres en fonction de la réponse donnée par les sujets) chez les sujets humains. Par ailleurs, les barres étaient présentées à différentes excentricités sur l'axe vertical du champ visuel. L'expérience se terminait lorsque les réponses des sujets pour une distance de barre donnée et pour chaque excentricité testée stagnaient autour du niveau de la chance, reflétant alors des réponses au hasard et donc une absence de perception du mouvement. Ce sont ces distances de barres pour lesquelles un mouvement apparent n'est plus perçu qui nous intéressent. Cela signifie en effet que les barres sont perçues comme superposées par nos sujets et les localisations rétiniennes stimulées par ces barres sont alors dites correspondantes.

La mesure de ces points correspondants à différentes excentricités a permis de confirmer une inclinaison de la composante verticale de l'horoptère chez nos sujets humains, avec des

variabilités inter-individuelles significatives. Surtout, et malgré le fait que toutes les mesures n'aient pas fini d'être collectées cette mesure expérimentale réalisée pour la première fois chez un sujet macaque a révélé qu'une telle inclinaison de l'horoptère était potentiellement aussi présente chez cette espèce. Cette inclinaison est également comparable à celle décrite chez l'humain et retrouvée chez nos huit participants humains. Cette absence de différence entre les deux espèces, attendues si l'on considère les différences de statistiques naturelles auxquelles elles sont exposées, semble favoriser l'idée que la relation entre la distance interoculaire et la hauteur des yeux serait un meilleur facteur explicatif de la forme de l'horoptère vertical. Cette hypothèse permettrait notamment de considérer les mesures indirectes (enregistrements électrophysiologiques des champs récepteurs) de l'horoptère vertical qui avaient été réalisées dans les années 1970 chez le chat et la chouette et qui suggèrent que ces deux espèces ont des inclinaisons similaires mais très différentes de celles trouvées chez l'humain et maintenant chez le macaque. Ces résultats devront évidemment être répliqués chez un autre macaque pour pouvoir réellement conclure sur le rôle de l'expérience visuelle sur la forme empirique de l'horoptère. Néanmoins, ils ouvrent déjà la voie à de nouvelles hypothèses et prédictions.

Pour conclure sur ces différents travaux effectués dans le cadre de ma thèse, le dernier chapitre de ce manuscrit (chapitre 6) comprend la discussion générale dans laquelle les résultats sont remis dans le contexte plus général de l'intégration par le système visuel et au niveau cortical principalement de la disparité binoculaire. La possibilité d'un recouvrement au moins partiel des réseaux corticaux impliqués dans les dimensions spatiale et temporelle de l'intégration binoculaire est discutée, de même que leurs implications au niveau des différentes voies visuelles connues. Finalement, des pistes de réflexions pour de futures études relevant de la psychophysique ou de la neuroimagerie et visant à surmonter les limitations des études menées lors de cette thèse sont proposées. Également, certaines limitations au niveau des connaissances actuelles sont évoquées.

Table of Contents

Acknowledgements.....	i
Abstract.....	v
Résumé.....	vii
Résumé substantiel en langue française.....	ix
Chapter I – General introduction to the visual perception in the primate brain	1
The visual system: from the retina to early visual areas.....	3
From early visual areas to higher visual areas: two pathways.....	15
Three-dimensional vision and binocular disparity.....	17
Chapter II – Context of the thesis: Integration of binocular disparities in the primate brain	23
Disparity processing and binocular integration	25
Absolute versus relative disparity.....	30
Disparity configuration and natural statistics	34
Temporal integration of binocular disparities.....	38
Methodological approach.....	39
Chapter III – Monkey fMRI - Methodology.....	41
The BOLD signal.....	43
Physics of (functional) magnetic resonance imaging	43
Nature of the BOLD signal and how it relates to the underlying neural activity	46
BOLD signal vs. MION-enhanced signal.....	53
Pre-processing steps and general linear model (GLM).....	54
Optimising the HRF estimation and data pre-processing to improve the SNR.....	56
Estimating the individual haemodynamic response function (HRF).....	56
Constant improvement of our methods: Complexity and success	58
Chapter IV – Temporal integration of binocular disparity: The case of motion-in-depth.....	65
General introduction	67
First study: Steremotion processing in the non-human primate (accepted article).....	71

Introduction.....	71
Materials and Methods.....	74
MRI recordings	77
Data analysis	79
Results.....	86
Discussion	99
Conclusion	108
Chapter IV – Spatial integration of binocular disparity: Influence of natural statistics	111
Content of this chapter	113
Second study: Spatial disparity gradients	115
Introduction.....	115
Material and methods.....	117
MRI recordings	120
Data analysis	121
Statistical analyses	124
Orientation biases.....	133
Differential activations with elevation?	138
Discussion.....	142
Conclusion	144
Third study: Measure of the horopter in human and in macaque	147
Introduction.....	147
Experimental setups	152
Results.....	157
Discussion	165
Chapter V – General discussion: What did we learn about the integration of binocular disparities?	169
Summary of the results	171

Overlapping cortical networks between temporal and spatial integration of binocular disparities: the case of the PIP cluster.....	173
Dorsal versus ventral visual pathways	175
Limitations and future directions	178
Linking cortical activation to behavioural responses.....	178
Functional connectivity within the revealed network.....	179
Modularity of the primate brain.....	180
Summing it all up with a nice quotation	183
References.....	185
Appendix I – Consent form for the horopter experiment conducted in humans	iii
Appendix II – Cottureau, B. R., Smith, A. T., Rima, S., Fize, D., Héjja-Brichard, Y., Renaud, L., Lejards, C., Vayssière, N., Trotter, Y., & Durand, J.-B. (2017). Processing of Egomotion-Consistent Optic Flow in the Rhesus Macaque Cortex. <i>Cerebral Cortex</i> , 1–14.	ix

List of figures

Figure 1: Anatomy of the eye.	4
Figure 2: Retina cell layers.	6
Figure 3: Phototransduction.	7
Figure 4. Projections of ganglion cells to the lateral geniculate nucleus (LGN).	12
Figure 5: Visual pathway, from the retina to the primary visual cortex.	13
Figure 6: Hierarchical organisation of the visual system.	18
Figure 7: Stereopsis is present in many different species.	19
Figure 8. A classic random-dot stereogram as described by Julesz (1971).	27
Figure 9. Schematised disparity tuning curves of two neurons in the primary visual cortex. ..	28
Figure 10. The disparity energy model.	30
Figure 11. Absolute and relative disparities.	31
Figure 12. Centre-surround stimuli to assess responses to absolute vs. relative disparities.	32
Figure 13. Responses of three neurons recorded at two different surround disparities.	33
Figure 14. Depth orders. Three orders of depth can be defined and result in different types of disparity-defined stimuli as illustrated here.	35
Figure 15: The BOLD signal reflects variations in oxygen consumption and modification of blood flow.	48
Figure 16: Simultaneous neural and haemodynamic recordings from a cortical site (in the striate cortex) showing transient neural response.	49
Figure 17: Illustration of the different types of cells and synapses that are comprised in a cube that has the equivalent size of one voxel.	50
Figure 18: MRI-compatible chair for macaque.	53
Figure 19. Characterisation of the haemodynamic response function (HRF).	58

Figure 20: A) Statistical parametric maps for the EC versus EI contrast in monkeys M01, M02, and M03. B) Map of overlap between significant activations in the EC versus EI contrast across the 3 monkeys.....60

Figure 21: A) Average sensitivity ratio (%) between the responses to the EC and EI conditions. B) Schematic localization of the 8 areas on the F99 template.....61

Figure 22: Mean PRF results (4 hemispheres) projected on the right inflated cortical surface of the F99 monkey template.....62

Figure 23: Assessing the contribution of the two binocular cues to motion-in-depth.....68

Figure 24. Stimulus design and experimental protocol.76

Figure 25. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) for M01..87

Figure 26. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) for M02.88

Figure 27. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) projected onto individual cortical surfaces and on the F99 template.....90

Figure 28. A) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M01 and delimitation of the MT cluster areas. B) Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) projected on the individual surfaces of M01, for both left and right hemispheres. C) Difference in percent signal change (Δ PSC) between the CSM and TS conditions in retinotopic areas.....93

Figure 29. A) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M02 and delimitation of the MT cluster areas. B) Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) projected on the individual

surfaces of M02, for both left and right hemispheres. C) Difference in signal change (Δ PSC) between the CSM and TS conditions in retinotopic areas	94
Figure 30. Sensitivity to 2D motion in CSM _{STS} , CSM _{ITG} , and CSM _{PPC}	96
Figure 31. Selectivity to 3D and 2D motion within the STS.....	98
Figure 32. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) for M01 and M02.	105
Figure 33. A) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M01 and delimitation of the MT cluster areas. B) Average eccentricities for the four areas of the MT cluster and for CSM _{STS} for both subjects. C) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M01 and delimitation of the MT cluster areas. D) Average pRF sizes for the four areas of the MT cluster and for CSM _{STS} for both subjects	106
Figure 34. Ocular behaviour.	108
Figure 35. Experimental protocol.	119
Figure 36. Activations for the contrast Correlated (GS+nGS) vs. Uncorrelated for M01.....	125
Figure 37. Activations for the contrast Correlated (GS+nGS) vs. Uncorrelated for M02. See Figure 36 for the details of the legend.	126
Figure 38. Activations for the contrast Correlated (RT+LT) vs. Uncorrelated for M01	127
Figure 39. Activations for the contrast Correlated (RT+LT) vs. Uncorrelated for M02.....	128
Figure 40. Difference in signal change (Δ PSC) between correlated (GS+nGS) and uncorrelated conditions in early visual areas and in the MT cluster.....	130
Figure 41. Difference in signal change (Δ PSC) between correlated (GS+nGS) and uncorrelated conditions in the PIP cluster and in V3A, V6, and V6A areas.	131
Figure 42. Activations for the contrast GS > nGS for M01.....	134
Figure 43. Activations for the contrast GS>nGS on the left and for the contrast nGS>GS on the right for M02.....	135

Figure 44. Difference in signal change (Δ PSC) between the ground-aligned slant (GS) and the non-ground aligned slant conditions in the PIP cluster..	136
Figure 45. Activations for the contrasts RT > LT on the left and LT>RT on the right for M01.	137
Figure 46. Activations for the contrasts RT > LT on the left and LT>RT on the right for M02. See Figure 45 for details of the legend.	137
Figure 47. Predictions of t-score patterns as a function of elevation for different disparity preferences.	140
Figure 48. Difference in t-score values between GS and nGS conditions as a function of the elevation for M01.	141
Figure 49. Difference in t-score values between GS and nGS conditions as a function of the elevation for M02.	142
Figure 50. Geometric and empirical horopters.	148
Figure 51. A) Eye-and-scene tracking device and weighted combinations of the four different activities (median horizontal disparities). B) Distributions of preferred horizontal disparity grouped by upper and lower visual field.	150
Figure 52. Apparent motion paradigm for the measurement of corresponding points.	154
Figure 53. Predicted psychometric curves.	157
Figure 54. Data from Observer 1.	158
Figure 55. Location of corresponding points for subjects 2 to 4.	160
Figure 56. Fitted curves for the different eccentricities at which corresponding points were measured.	161
Figure 57. Non-corrected locations of corresponding points for the macaque subject	162
Figure 58. Non-corrected locations of corresponding points for the macaque subject for even (left panel) and odd (right panel) trials.	162

Figure 59. Fitted curves for the different eccentricities at which corresponding points were measured for the even trials.. 163

Figure 60. Fitted curves for the different eccentricities at which corresponding points were measured for the odd trials..... 163

Figure 61. PIP cluster as defined in the wide-field retinotopic study mentioned in Chapter III. (Rima et al., *under review*). 174

Figure 62. Regions sensitive to structural and positional stereoscopic information are projected onto the flattened representations of left and right IPS..... 175

Figure 63. Schematic representation of known functional interactions between different areas from the dorsal and ventral pathways.. 180

Figure 64. Different visual and oculomotor cues can be used by the primate visual system to compute a 3D percept.. 182

List of tables

Table 1: Parameter values of the HRF for each individual.....	57
Table 2. MNI coordinates (in mm) of the local maxima for the 3 regions that were significantly more responsive for the CSM condition than for the TS control in the two hemispheres of the two animals.	89
Table 3. Number of hemispheres that pass the significance threshold for each retinotopic areas and for each contrast..	132
Table 4: Participants characteristics for the measurement of corresponding points.....	155
Table 5. Interspecies comparisons of the optimal shear angles.	165

Chapter I –
General introduction to the visual perception
in the primate brain

From the photon that hits the retina to the visual interpretation of our environment

The visual system: from the retina to early visual areas

Whilst we heavily rely on our vision to interact with our environment and perform everyday life actions, most of the time we do not realise how complex the visual machinery might be. What underlies this ability that most species have to visually perceive the world? How does a nervous system, the brain in our case, manage to make sense of the light information that comes and hits the retina?

These questions are far from being new as already Plato and Galen had their own views on the matter. From the extramission theory, once claiming that the eye emits light to encompass objects, to the intromission theory introduced during the Islamic Golden Age by Ibn al-Haytham (Alhazen) and Ibn Sina (Avicenna) that was finally stating that the eye is using the light to provide visual perception, it was only during the early 17th century with Kepler's work that we approached the idea of a camera obscura, with an image being projected on the retina.

The metaphor of the camera works well to grasp the anatomy of the mammal eye and explain the role of the different elements that compose it.

As illustrated on Figure 1, the light –in the form of an electromagnetic wave – first goes through the cornea, a fixed transparent membrane in charge of about 70% of the light refraction and that covers the iris and the pupil. The muscular action of the iris will allow more or less light to enter the eye by constricting (miosis) or dilating the pupil (mydriasis), a phenomenon also known as the pupillary light reflex. The second step is the aqueous humour, a watery fluid that provides nutrients to the cornea.

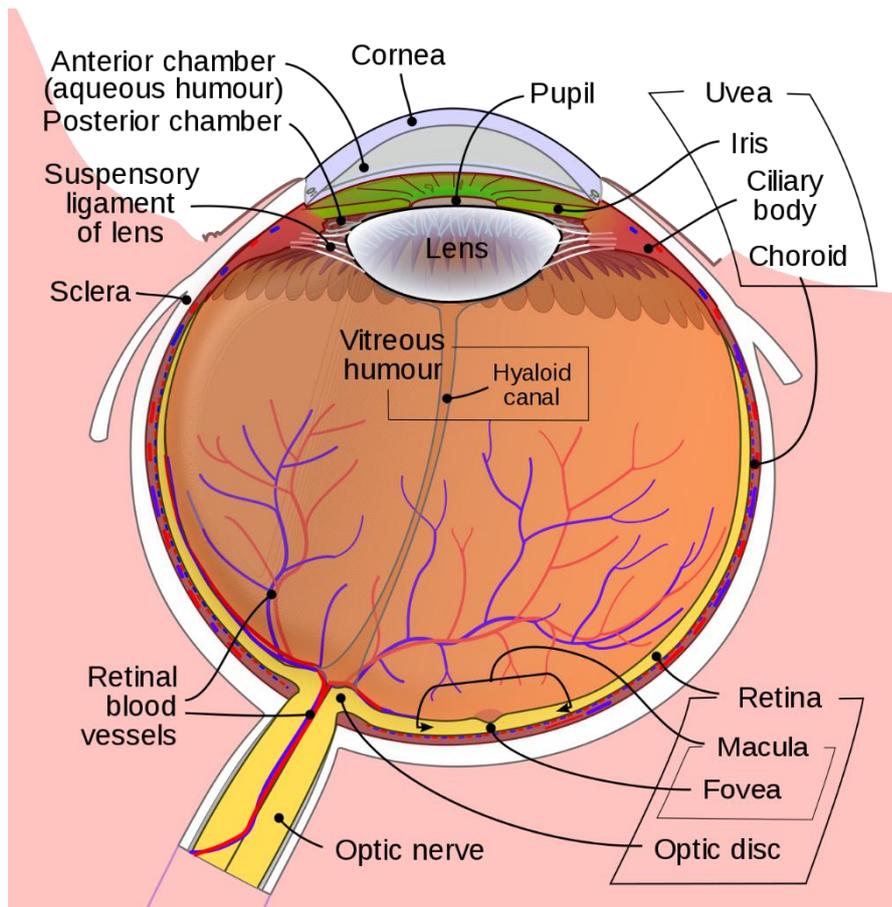


Figure 1: Anatomy of the eye.

The mammal eye is made of different elements that help refract the light to be focused on the retina. The light, in the form of an electromagnetic wave, first goes through the cornea, in charge of about 70% of refraction, crosses the aqueous humour, a watery liquid that provides nutrients to the cornea, then goes through the lens, a biconvex structure that further refracts the light by accommodating, and finally crosses the vitreous humour, a sort of jelly matter that gives its round shape to the eye, before hitting the retina and the different cells that compose it. The most detailed information is projected onto one specific part of the retina: the fovea, where signal processing is finer, due to the asymmetrical retina cell distribution (detailed further in the text). Image retrieved from <https://www.umkelloggeye.org/conditions-treatments/anatomy-eye>

The light will then go through the lens, a biconvex structure composed of different layers and that further helps to refract the light by accommodating under the action of ciliary muscles,

doing the same work as the focus of a camera. Ciliary muscles will contract for a short distance, thus bending and thickening the lens, giving it a strong refraction power, and they will relax for a far distance, then stretching the lens. Finally, the luminous information will cross the vitreous humour, a sort of jelly that gives the eye its round shape before reaching the retina and the cells that it is made of (see e.g. Masland, 2001).

Located in the back of the eye, the retina is in charge of transforming the light information into a signal that will be interpreted by the brain, that is an electric signal. Among the different cell layers retina is made of (see Figure 2), the photoreceptor layer represents the initial step of signal conversion. Photoreceptors are divided into two categories: rods and cones. Rods are mostly responsible for dim light vision and represent the majority of photoreceptors (95% of them, that is about 120 million), whilst cones are more sensitive to bright and coloured light, being themselves divided into three different types: short wave-length (blue colour), medium wave-length (green colour), and long wave-length (red colour) sensitive cones. The distribution of rods and cones varies across the retina, with most cones being in the centre of the retina, at the fovea level, and rods being more highly dense in the periphery surrounding the fovea. A chemical cascade reaction, known as photo transduction (see Figure 3 for more details), is the result of photon detection within the outer part of the photoreceptor cells that contain photoreceptor pigments (rhodopsin in rods and photopsin in cones), leading to an hyperpolarisation of photoreceptor cells that further conveys the now-electric message, in the form of a membrane potential, to the next layer made of bipolar and horizontal cells.

From photoreceptors to bipolar cells, the information is compressed due to a lesser number of bipolar cells than of photoreceptor cells and the resolution varies depending on the type of photoreceptors. High numbers of rods in the periphery of the retina converge onto single bipolar cells, whereas cone information from the retina is less compressed, thus giving rise to a high-resolution vision at the fovea level. This difference will be preserved at the brain level,

since proportionally more optic nerve fibres, made of the axons of ganglion cells, will conduct information from the cones, giving a primary importance to the information projected on the fovea.

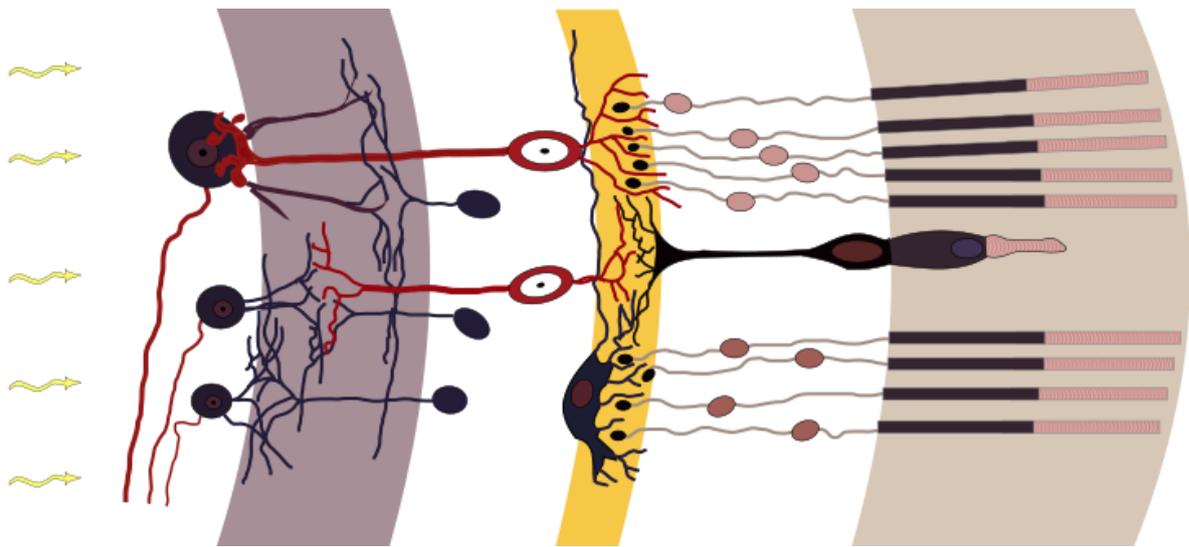


Figure 2: Retina cell layers.

Light (yellow arrows on the left) goes through the different layers in the retina before reaching the photoreceptor layer (beige layer, on the very right). The information encoded by photoreceptors (cones in the fovea and rods in the periphery) is then back propagated to the bipolar and horizontal cells (the outer plexiform layer, in yellow) and further sent to the ganglion cells (the inner plexiform layer, in purple) before leaving the retina through optic nerve fibres made of the axons of ganglion cells. Image retrieved from Wikipedia Commons: <https://commons.wikimedia.org/wiki/File:Retina-diagram.svg> (vectorisation of picture from Ramon y Cajal by [Chris](#)).

Each retina cell encodes a spatial localisation of the visual space, that is one cell's baseline state will be modified in response to the presence of light in one part of the visual field only, leading to a modification of the cell's membrane potential or a generation of action potentials. This cell sensitivity area is its receptive field. The size of those receptive fields varies

accordingly to the visual cell hierarchy, becoming bigger and bigger as one goes up the layers of the visual system. Photoreceptors are their own receptive field, as they will be triggered by luminous information received in their outer segments localised in one physical point of the retina on which visual space is projected. Photoreceptors are connected to bipolar cells and they will then represent the receptive fields of those bipolar cells. Several bipolar cells are also connected to a ganglion cells and they will again compose the receptive field of this ganglion cell. The receptive field of the ganglion cell is thus a collection of the sensory inputs that are received by all the photoreceptors that are synapsing with the bipolar cells that are connected to the ganglion cell. This convergence process will continue up to brain neuron cells, leading to bigger and bigger receptive fields (see e.g. Briggs, 2017).

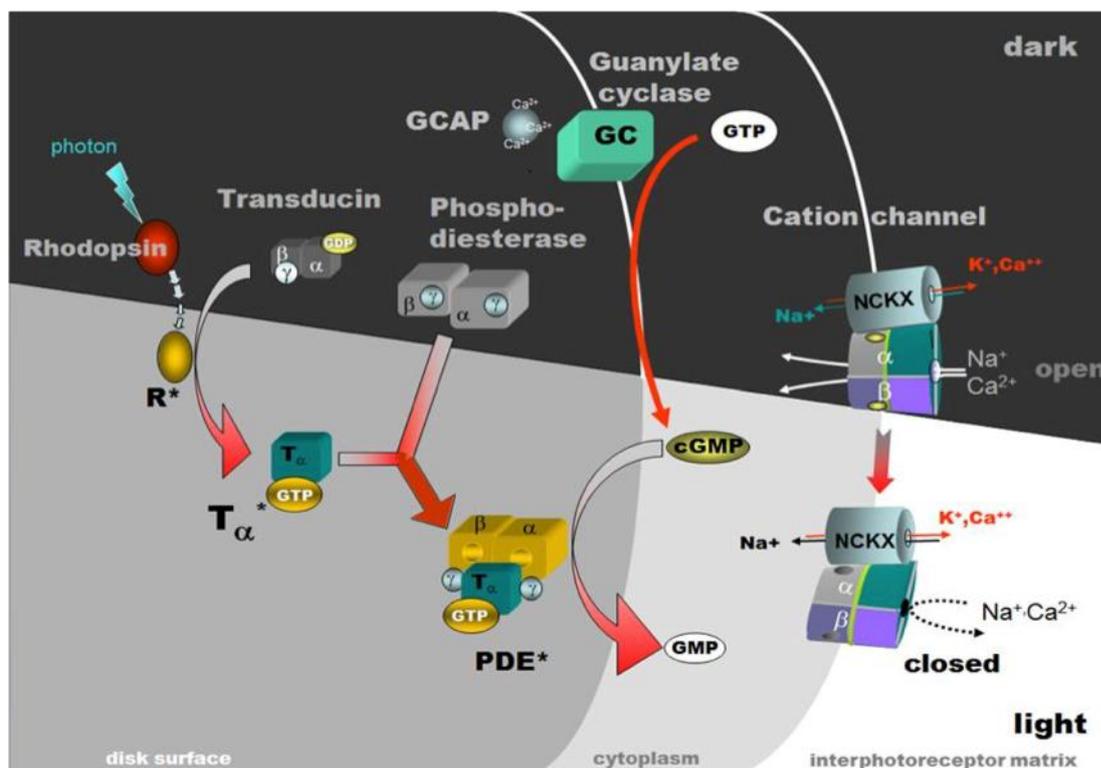


Figure 3: Phototransduction.

Photoreceptors baseline state is within dark conditions, photoreceptors are then depolarised due to high level of cGMP keeping sodium gates open. This allows potassium ions to get out of the cell and sodium to get in, creating what is called a dark current. A neurotransmitter, glutamate, is then released by the photoreceptor. The action of glutamate is to hyperpolarise

ON-centre bipolar cells and to depolarise OFF-centre bipolar cells (for sake of simplicity, we will not mention the horizontal cell pathway). Once a photon is detected by a photopigment present in the disks of the outer segment of a photoreceptor (here, the rhodopsin), it triggers a whole chemical cascade as illustrated on the figure. Briefly, the retinal component (a vitamin A derivative) of the rhodopsin molecule becomes activated (R^*), changing configuration from 11-cis to all-trans. To balance this change, opsin, the other component of rhodopsin, also undergoes a configuration change becoming metarhodopsin II. Metarhodopsin II then activates the transducing protein, causing its dissociation from GDP and its binding to GTP. The alpha subunit of transducing $T\alpha^*$ dissociates from the other subunits but remains attached to the binding GTP. This complex will then activate the cGMP-hydrolysing phosphodiesterase (PDE^*) responsible for the degradation of cGMP in GMP. As a consequence, cGMP concentration within the cell decreases, leading to the closing of cGMP-dependent sodium gates, thus stopping the dark current and causing the hyperpolarisation of the photoreceptor cell. Glutamate is no longer released in big quantities and ON-centre bipolar cells depolarise whilst OFF-centre bipolar cells hyperpolarise. Image copyright: Wolfgang Baehr (<https://pubmed.ncbi.nlm.nih.gov/18193635/>).

Receptive fields are not simply defined as a region getting excited by light. From bipolar cells onwards they also have an ON-OFF organisation, having one or several excitatory parts and one or several inhibitory parts. Their shape varies from being circular at the retina level to being more elongated once at the cortex level, becoming more sensitive to the light orientation.

Bipolar cells for instance have an antagonist centre-surround organisation, with two possible configurations: an ON-centre associated with an OFF-surround or an OFF-centre coupled with an ON-surround (see e.g. Dacey, Crook and Packer, 2017). Located between photoreceptors and bipolar cells, horizontal cells will selectively inhibit the information of some photoreceptors, only allowing strongly emitting photoreceptive cells (i.e. photoreceptors receiving the highest amount of light) to transmit their signal to the bipolar cell they are connected to, thus increasing the signal-to-noise ratio. They are directly involved in the centre-surround organisation of the bipolar cells as they mediate the surround activity of the bipolar

cells. The ON-OFF configuration of bipolar cells and the lateral inhibition property of horizontal cells represent the first step in the detection of edges and contrasts in the visual information received at the level of the retina.

Ganglion and amacrine cells compose the last step before the information leaves the retina through the optic nerve fibres made of the axons of ganglion cells.

Part of the inner plexiform layer like ganglion cells, amacrine cells are interneurons that are involved in integrating and modulating temporal information to transmit it to the ganglion cells. They might also play a similar role as of the horizontal cells in varying the signal from adjacent ganglion cells to enhance signal-to-noise ratio.

Ganglion cells are subdivided into three main categories: magnocellular (M-type), parvocellular (P-type), and koniocellular (K-type) ganglion cells. They all have distinct properties and will follow different pathways. This segregation will remain until the very end of signal integration and processing (see e.g. Briggs, 2017).

Magno-type and parvo-type cells share complementary properties and are usually taught to be part of the “where” pathway with a high temporal resolution and the “what” pathway with a high spatial resolution, respectively.

Parvocellular ganglion cells represent the vast majority of ganglion cells (about 80%). Of small size in term of dendritic tree size, they are also known as midget cells. Mostly connected to cones, they receive visual information in a one-to-one fashion within the fovea and the parafoveal area, that is, one ganglion cell is receiving information from one bipolar cell, which also receives information from a very low number of photoreceptors. Due to this specific configuration, the parvocellular pathway conveys very precise visual information, leading to a great detection of edges and thus, to the later integration of shapes. Their centre-surround

organisation also supports the discrimination of antagonist green-red colours (see e.g. Bowmaker, 1998).

Magnocellular ganglion cells integrate an achromatic visual signal sent by bipolar cells and modulated by amacrine cells faster than parvocellular cells. They are also called parasol cells due to the vast size of their dendritic tree and of their receptive fields, making them great motion detectors. They represent about 10% of the ganglion cells.

Koniocellular ganglion cells are the 10% remaining ganglion cells and represent 8-10% of the cone cell population. S-cone cells are physiologically and anatomically very different from L- and M-cones cells found in the parvocellular pathway, from the outer plexiform layer onwards (bipolar, horizontal, and ganglion cells). Often presented as having heterogeneous properties, koniocellular ganglion cells are the less studied and are believed to be an archaic form of the more recently evolved magno- and parvocellular ganglion cells. They underlie yellow-blue colour discrimination via their antagonist centre-surround organisation (see e.g. Klein et al., 2016; Carvajal et al., 2012).

The axons of the different types of ganglion cells form the optic nerve fibres. Leaving the retina at the level of the optic disc, they mostly carry visual information to the lateral geniculate nuclei (LGN), a 6-layer structure located in the thalamus (see Figure 4). Before reaching the LGN, about 60% of the optic nerve fibres decussate at the level of the optic chiasma. This will have consequences for the processing of the different parts of the visual field. As illustrated on Figure 5, the information received by the nasal retinae will be processed by the contralateral LGN and visual cortex, whereas the ipsilateral visual cortex will process the information received by the temporal retinae. This implies that one visual hemifield is seen by both eyes and will be processed by the contralateral LGN and the contralateral visual cortex. It is also

worth noting that the temporal retinae receive information from the central visual field where information seen by both eyes overlaps, thus supporting binocular vision.

Information processed at the level of the LGN is segregated, based on the type of cells conveying it. Parvocellular ganglion cells will project in the dorsal layers of the LGN (layers 3 to 6), magnocellular ganglion cells project in the ventral layers of the LGN (layers 1 and 2), and koniocellular ganglion cells project between layers. Moreover, layers 2, 3, and 5 will process ipsilateral information, whereas layers 1, 4, and 6 will process contralateral information, thus keeping the information received in each visual hemifield separated.

Axons of LGN relay neurons that transmit retinal information to the primary visual cortex compose part of the optic radiations, forming the link between the retinogeniculate and the geniculocortical pathways. LGN output to the primary visual cortex is quite similar to its retinal input, although precise convergence and divergence of the information still remains unclear. Feedback connections from the primary visual cortex to the LGN form another major portion of optic radiations and most of the input of the LGN and are possibly involved in information filtering.

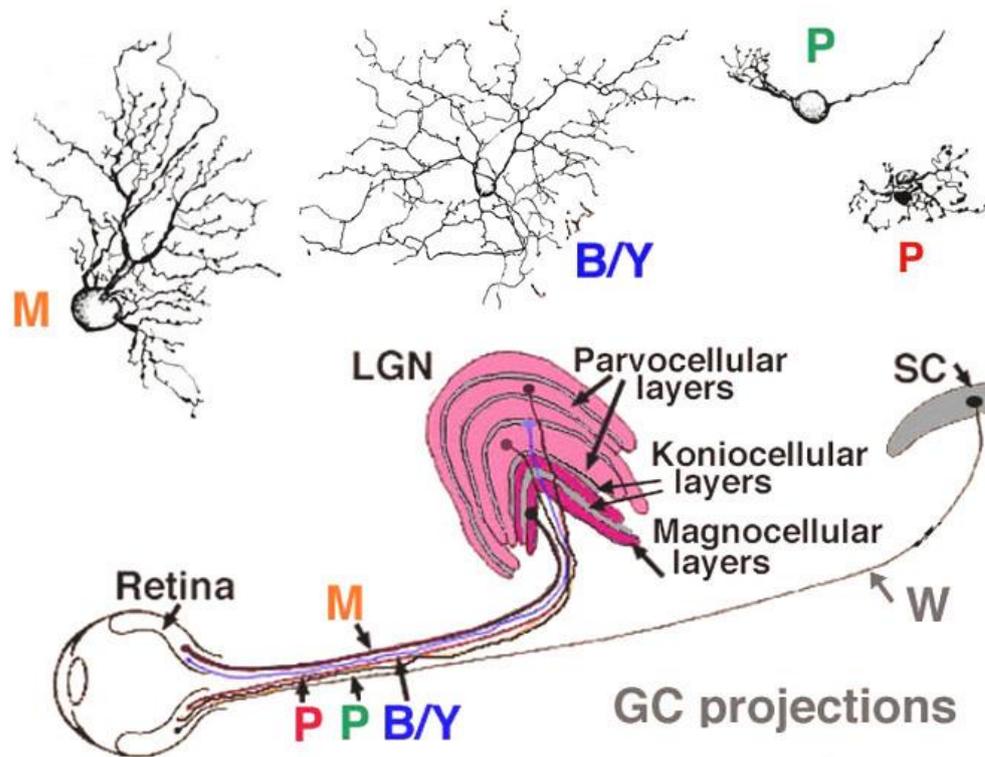


Figure 4. Projections of ganglion cells to the lateral geniculate nucleus (LGN).

Top half: representation of the different types of ganglion cells found in the retina: M-type (magnocellular pathway), P-type (parvocellular pathway), and B/Y or K-type (koniocellular pathway) cells. Dendritic tree differences appear clearly. Bottom half: M-type, P-type, and K-type (B/Y on the figure) ganglion cells project onto different layers in the LGN, keeping the information they convey segregated. Adapted from Dacey and Peterson (1992) & Shapley and Perry (1986). Image retrieved from Webvision (Helga Kolb).

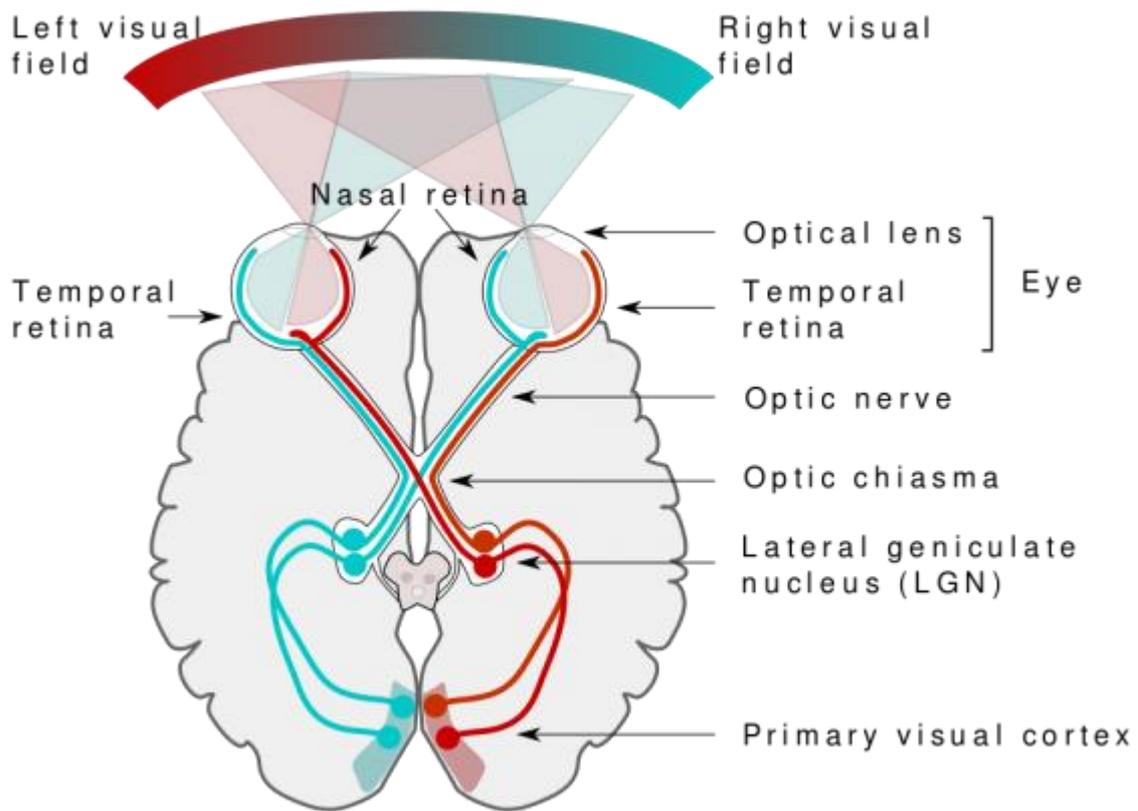


Figure 5: Visual pathway, from the retina to the primary visual cortex.

Hemifields and how the information is conveyed to the primary visual cortex. Image retrieved from Wikipedia Commons (author: [Miquel Perello Nieto](https://commons.wikimedia.org/wiki/File:Human_visual_pathway.svg)) https://commons.wikimedia.org/wiki/File:Human_visual_pathway.svg

The primary visual cortex or V1 is the first cortical structure involved in visual processing. Located in the occipital lobe at the back of the primate brain, it is made of 6 different layers that will be the target of LGN inputs in a differentiate manner: Magnocellular cells first project onto the 4C α layer of V1 and then onto the 4B layer where direct connections to area MT (medial temporal area, known for processing motion, see e.g. Newsome and Pare, 1988) exist; parvocellular cells project onto the 4C β layer of V1 and further on the 2nd and 3rd layers within the blobs and in between blobs; koniocellular cells directly projects into the blobs of the 2nd and 3rd layers. The 6 different layers of V1 are further subdivided in columns. Ipsilateral and

contralateral projections from the LGN will be kept segregated within adjacent columns of layer 4C. Those columns have been called ocular dominance columns (LeVay, Hubel, & Wiesel, 1975).

Contrary to the retinal cells, V1 cells receptive fields are more elongated. Their size and properties vary depending on the complexity of the cell and on the layer where the cell is located. Hubel and Wiesel (1968) distinguished two kinds of cells: simple cells and complex cells. Simple cells are mostly located in layer 4 and have small receptive fields with a monocular input. They have a clear ON-OFF organisation, which can follow a centre-surround organisation or being more lateralised with a spatial offset. Those neurons are selective for a spatial orientation perfectly aligned with their receptive fields. In layers 2, 3, 5, and 6, cells become more complex: their medium sized receptive fields do not have any clear ON-OFF organisation anymore and they display binocular responses, leading the path to binocular visual integration (Cumming and Parker, 2000). Those complex cells are selective for spatial orientation with less constrain than the simple cells: Stimulation can occur whenever in the receptive field; the cell will fire as long as the stimulation remains aligned with its receptive field orientation. Cells selective to a same orientation will be gathered in a single orientation column.

Besides this columnar organisation, the primary visual cortex also has a retinotopic organisation. Each point of the visual field that is projected on the retina has its corresponding processing surface in V1. The ipsilateral hemifield projects on the contralateral primary visual cortex and the superior hemifield on the inferior part of V1. The foveal region that projects on the posterior part of V1 is overrepresented, due to cortical magnification: the cortical representation of the fovea in V1 projects onto a much larger region than the actual size it occupies in the retinal space. As a direct consequence of this cortical magnification, the more peripheral areas are represented on a smaller cortical surface than the surface they actually

occupy in the retinal space. This correspondence between the retina and the primary visual cortex is repeated in each of V1 hypercolumns, which correspond to all the different orientation columns (i.e. a complete 180° sequence) or to each complementary duo of ocular dominance columns (i.e. representing both left and right eyes) (Hubel & Wiesel, 1977). This retinotopic organisation will be also present in extrastriate visual areas (i.e., areas belong the primary visual cortex, also called the striate visual cortex), with a notable difference in the size of receptive fields: the higher in the visual hierarchy, the bigger the size of neuron receptive fields, and the more coarse-grained the retinotopic mapping. This has as a direct consequence a complexification of the visual features being extracted.

As a matter of fact, the information collected at the retina level is not only involved in visual perception per se but is also used for three different functions: the pupillary reflex, with information converging into the pretectum; the circadian rhythm with the suprachiasmatic nucleus as the relay centre; the coordination of head and eyes movements and saccade planning, with the remaining optic nerve projections reaching the superior colliculi before being further processed by cortical structures.

From early visual areas to higher visual areas: two pathways

Visual information entering but also leaving the primary visual cortex is segregated into two parallel pathways¹: the magnocellular (dorsal) and the parvocellular (ventral) pathways. Neuropsychology and brain lesion studies have helped a lot in the understanding of how those different pathways might complement each other and in shedding light on the specialisation of some brain areas. For instance, cortical lesions in areas on the dorsal stream were associated

¹ From now on, the koniocellular pathway will no longer be mentioned, due to a limited knowledge about this pathway and its limited contribution to our research questions.

with various visuo-motoric or visuo-spatial disorders whereas lesions on the ventral pathway led to object recognition deficits or to acquired achromatopsia.

The magnocellular pathway is more generally involved in processing spatial information of visual objects: their movements, the distance they are located at, or their distance relationship; in order to enable interactions with those objects. Also called the “where” pathway, it comprises a network of areas located on the occipito-parietal (dorsal areas) side of the brain. The medial temporal area (MT) and then the medial superior temporal area (MST) are the first cortical areas to receive inputs from V1. Part of the MT complex (the MT cluster in macaques, and the hMT+ cluster in humans²) that also includes the fundus superior temporal area (FST) and V4t, those areas are mainly involved in processing motion information. Higher visual areas of the dorsal pathway, located in the infero-parietal cortex (Shikata et al., 2007), will be involved in visually guided movements, such as reaching and grasping objects (e.g., the anterior infero-parietal area, AIP; Taira et al., 1990; Sakata et al., 1995), or in eye movement planning (e.g., the lateral infero-parietal area, LIP; Andersen et al., 1990).

The parvocellular pathway, or the “what” pathway, involves areas located more ventrally, in the occipito-temporal part of the brain. Those areas, such as V4 (Zeki, 1974) or the infero-temporal area (Conway et al., 2007; Bartels and Zeki, 2000), are involved in the processing of form, shape, and colour perception and, to a further extent, objects identification or recognition. As an example of the specialisation of the occipito-temporal and infero-temporal cortices, some areas showing strong responses to objects (e.g. the lateral occipital complex, LOC), scenes (e.g. the occipital place area, OPA), or faces (e.g. the face fusiform area, FFA) have been extensively studied, establishing a robust selectivity for those objects (Kanwisher, McDermott, and Chun,

² Nomenclature conventions: h = homologous area, p = putative (homologous) area.

1997; Grill-Spector, Kourtzi, and Kanwisher, 2001; Park and Chun, 2009; Dilks, Julian, Paunov, and Kanwisher, 2013).

Not only those two pathways differ in the kind of information they process (Livingstone & Hubel, 1987), they also have different processing speed, with the magnocellular pathway being faster (high conduction velocity and transient responses) than the parvocellular pathway (low conduction velocity and sustained responses). This distinction along with a parallel processing of information underlies a coarse-to-fine processing of the visual information where low spatial frequencies carried by the magnocellular pathway are processed first by the visual system and will then serve as a feedback for the integration of higher spatial frequencies, brought by the parvocellular pathway, resulting in a finer processing of visual information (Marr, 1982; Schyns and Oliva, 1994; Bullier, 2001; Bar, 2004). Now supported by empirical evidence (see e.g., Lu et al., 2018; Petras et al., 2019), this integration theory of the visual signal challenges the traditional views of a purely feedforward signal and sheds light on the dynamics of the visual system and the importance of feedback connexions.

In a nutshell, although some brain areas are very specialised and necessary for the processing of some specific types of information, as for the MT area and its primordial role in motion processing, the brain is generally organised as a hierarchical network with numerous feedforward but also feedback connexions (Figure 6; Van Essen & DeYoe, 1995).

Three-dimensional vision and binocular disparity

The major aim of this thesis being about gaining understanding and knowledge about brain networks involved in two aspects of binocular disparity processing, temporal and spatial

processing of disparity gradients, we shall now dig a bit more into binocular vision and disparity processing.

Stereopsis, or three-dimensional (3D) vision, is the ability to perceive depth by combining information coming from our two eyes, that is, binocular information. Providing a series of evolutionary benefits, it has evolved in a wide range of species in the animal kingdom (see Figure 7 for some examples). Present in most predatory species (i.e., in carnivorous species), stereopsis helps to detect a prey by breaking camouflage and revealing occlusion, via figure-ground segregation; and eventually to catch the prey by providing precise distance estimation. In primates and in one bird species, the New Caledonian crow (Troscianko et al., 2012), stereovision most likely co-evolved with tool manipulation abilities and underlies fine visuo-motoric interactions.

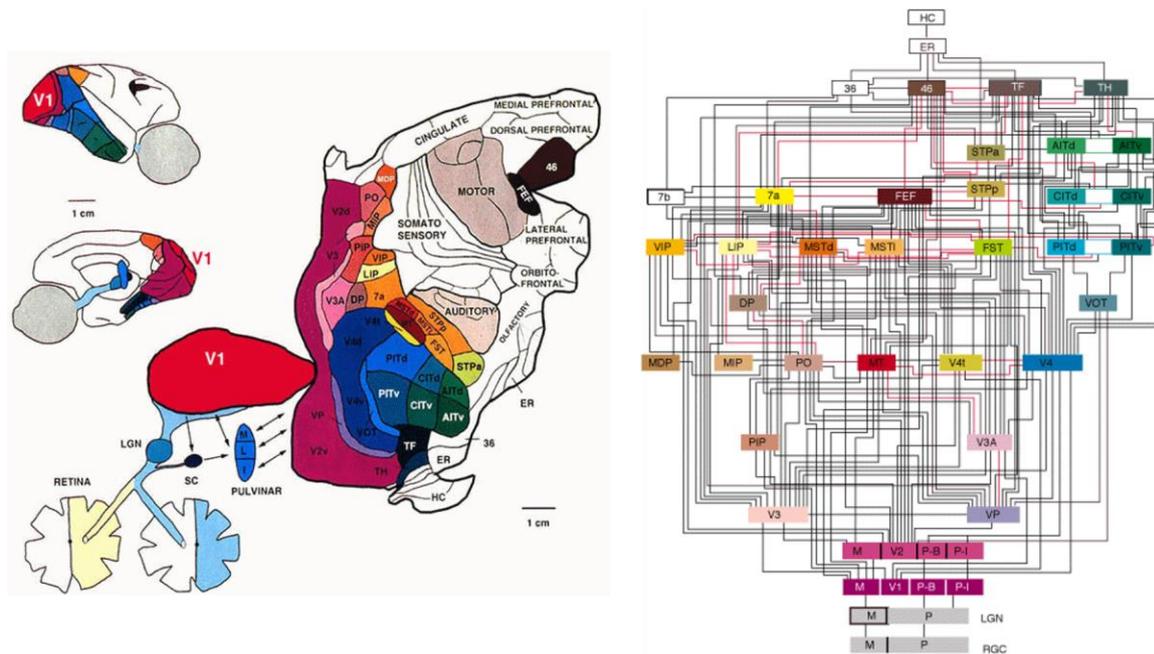


Figure 6: Hierarchical organisation of the visual system, based on Felleman and Van Essen 1991.

For the visual system to be able to integrate binocular information requires an overlap between the visual fields of both eyes, and thus both eyes to be laterally aligned on the horizontal plane. This lateral displacement of the visual field comes with the loss of peripheral vision. One can then easily understand why for prey animal species like rabbits or deer, stereovision was not favoured by evolutionary constraints: their peripheral vision ensures them to detect predators in a wider visual field, giving them the safe opportunity to run away.

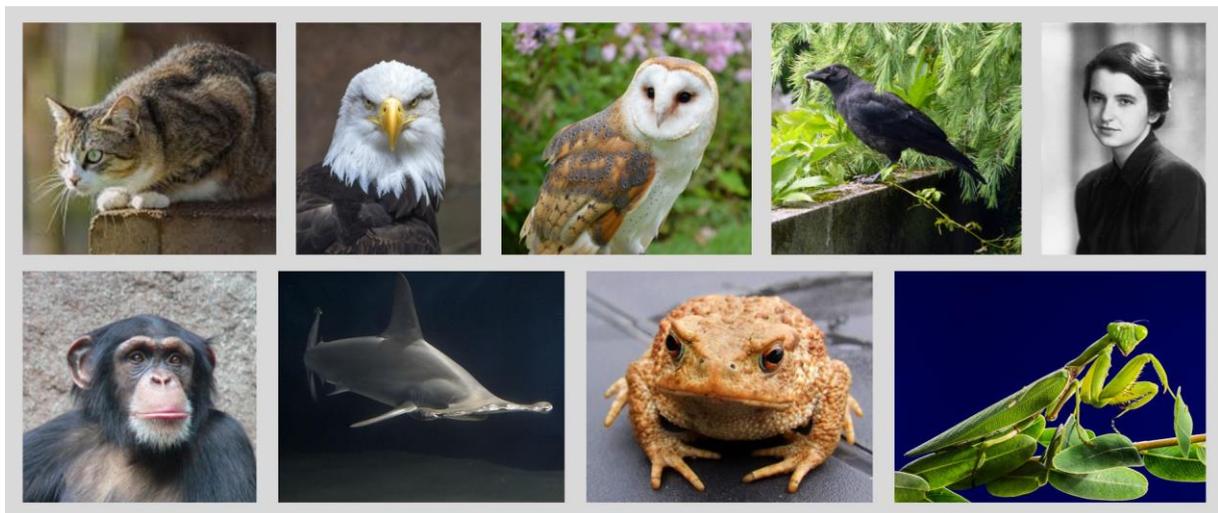


Figure 7: Stereopsis is present in many different species.

Stereopsis has been shown to be present in cats (with the exception of Siamese cats), hawks, owls, new Caledonian crows, humans, chimpanzees, sharks, toads, and more recently in the praying mantis. All those species have a binocular overlap that enables their visual system to compute 3D vision by comparing both eyes' inputs to extract depth information. Home-made composition with Creative commons images.

Because of this lateral shift between both eyes, projections of a visual scene on each retina do not completely overlap, enabling the visual system to use the small image differences, termed binocular disparities, to compute depth perception. Those binocular disparities are mostly horizontal but vertical disparities can also be observed in the parafoveal and in the peripheral space and be used by the visual system (Durand et al., 2002).

Depth perception based on binocular disparities is most accurate at short distances, when disparity values are the biggest. Other visual cues, mostly monocular, can be used in a complementary fashion by the visual system to make sense of object distance and elevation. They include luminance, shading, perspective, occlusion, texture, or motion (via motion parallax). Two additional binocular cues also provide the visual system with depth information: the interocular velocity difference and the vergence angle. The former consists of using phase-opposed movement information, that is the trajectory of a moving object will follow opposed directions in each retina; whereas the latter is based on motor information resulting from converging or diverging the eyes to focus on a point and mainly gives information about object distance. Each of those visual cues has its own limitations and they will be differently combined depending on the visual context.

How is binocular information integrated at the brain level is not yet fully understood and will be at the heart of the next chapter of this thesis ³.

³ Note that not only binocular integration remains to be completely understood, but also the combination of the different cues that can lead to a 3D percept, and which includes monocular cues along binocular ones. Despite being of equal interest, this is out of the scope of this thesis and will, as a consequence, not be addressed in this manuscript.

Chapter II –
Context of the thesis: Integration of binocular
disparities in the primate brain

Disparity processing and binocular integration

“... *the mind perceives an object of three dimensions by means of the two dissimilar pictures projected by it on the two retinae ...*”⁴

Stereopsis, or the ability to compute depth by making use of small retinal disparities, requires more than a binocular overlap of visual hemifields (often given by front-facing eyes) and a good visual acuity in both eyes. Indeed, the visual system needs to be able to triangulate the depth of an object based on its projection on both retinae (i.e. the object should be at the intersection of both lines of sight). Doing so requires solving the correspondence problem, that is to identify corresponding points in both retinae. How the brain solves this problem has been the central question of a fair amount of studies using different approaches and methods: Neurophysiological recording first in the cat, then in the macaque brain; psychophysical studies, mostly conducted in human; neuroimaging recording done in human and more rarely in macaque; or computational modelling of neuron selectivity to disparity. Yet, despite decades of creative and abundant work, we are still far from completely understanding the mechanisms that underlie depth perception. In this contextual introduction, I will go through key studies investigating the processing of binocular disparity to compute depth perception in order to give a comprehensive view of the up-to-date knowledge. I will further develop the aims of this thesis project and the approach we propose to use.

The book that Béla Julesz published on cyclopean perception in 1971 had consequential repercussions on the study of depth perception based on binocular disparity. His popularisation

⁴ As stated by Wheatstone in 1838 in his *Contributions to the Physiology of Vision. – Part the First. On some remarkable, and hitherto unobserved, phenomena of binocular vision*, the very first work dedicated to explain stereopsis.

of the random-dot stereogram (RDS) provided researchers with the ideal tool to study how the visual system processes binocular disparities. RDS is a pair of images made of random dots that correlate between both eyes, with the exception of a subpart of one image that has been shifted horizontally to create disparities (Figure 8). When fused, they give rise to depth perception, revealing a hidden figure from the ground, based on the sole use of binocular disparities. Those stimuli have been widely used to assess the presence of disparity-tuned neurons or more generally of disparity responses in the primate brain. First described in the cat's primary visual area (area 17) in several studies (Barlow et al., 1967; Nikara et al., 1968; Pettigrew et al., 1968), disparity selective neurons (Figure 9) were also found in the macaque brain in many different areas, including the primary visual area V1 (Poggio et al., 1988; Prince et al., 2002) and extra-striate areas V2 (a.k.a. area 18 in Hubel and Wiesel, 1970), V3-V3A (Poggio et al., 1988; Anzai et al., 2011), V4 (Hinkle and Connor, 2001, 2005), MT (Maunsell and Van Essen, 1983; DeAngelis and Uka, 2003), MST (Roy et al., 1992), as well as areas within the infero-temporal cortex (Janssen et al., 2003) and the intra-parietal cortex (Taira et al., 2000). Those different studies showed that disparity processing is happening in the entire visual cortex of the primate brain, starting at a very early stage in the primary visual cortex (for a review, see Parker, 2007) where neurons have binocular receptive fields (i.e. receptive fields with corresponding retinal locations). However, although disparity tuning in neurons has been robustly demonstrated, the mechanisms behind it remain to be understood. Indeed, if the receptive field (RF) properties of simple cells enable a linear spatial summation of the binocular inputs to compute disparity⁵ (Ohzawa and Freeman, 1986)

⁵ Despite their selectivity to binocular disparity, simple cells do not encode those disparities in an independent manner (i.e. they do not have a well-defined preferred disparity) but along with other characteristics of a stimulus such as its orientation or position. This implicates that their responses to disparity are not independent of other stimulus properties but are rather phase dependent and might be affected by stimulus changes such as lateral displacement.

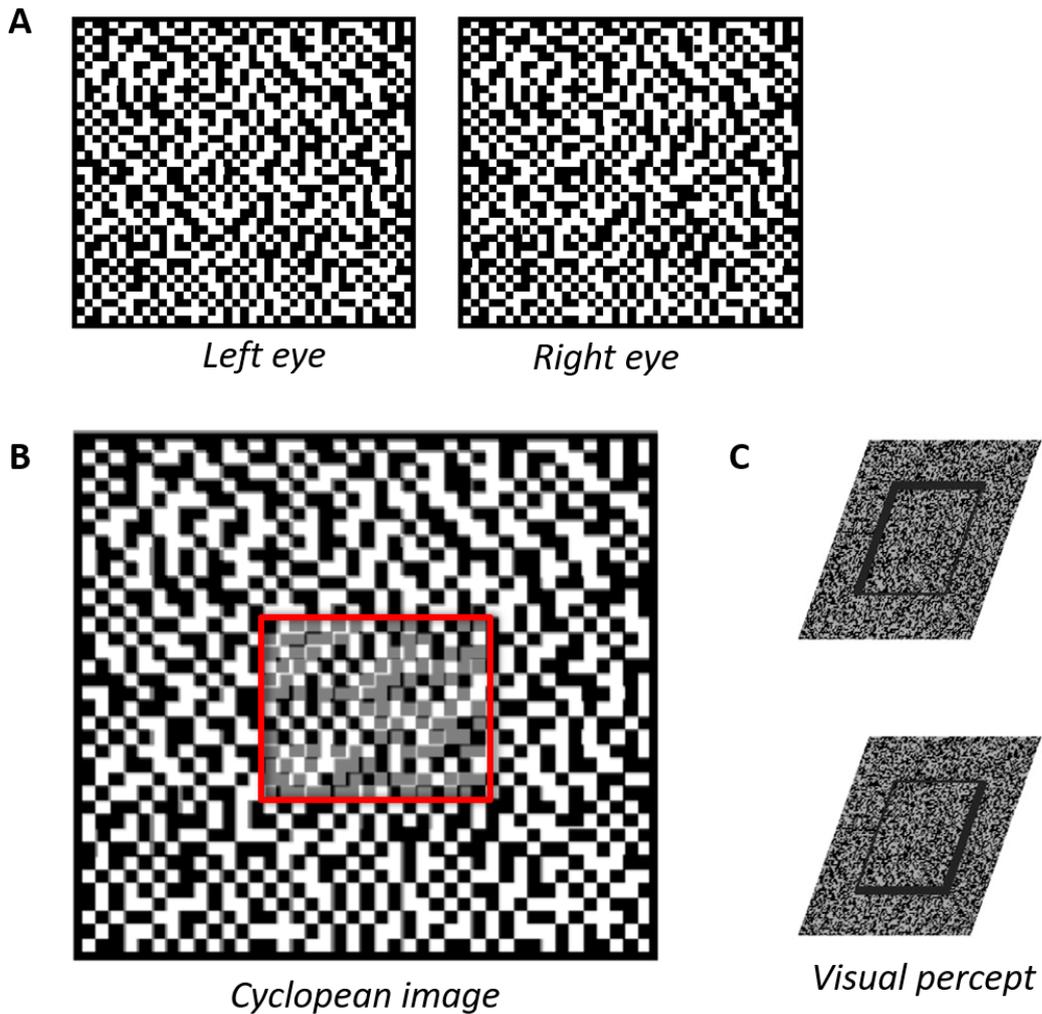


Figure 8. A classic random-dot stereogram as described by Julesz (1971).

A) Example of paired images shown to each eye where no monocular cue can be extracted to aid depth perception. B) When viewed with a stereoscope, images are fused in a cyclopean image. A depth percept emerges because of the small horizontal disparities between the two retinal projections. C) The two images illustrate the possible visual perceptions where a square in the middle of the image detaches itself from the figure background, either coming in front or going behind. Home-made composition made of images retrieved from:

<https://thiscondensedlife.files.wordpress.com/2016/03/random-dot-stereogram.png?w=572&h=276>

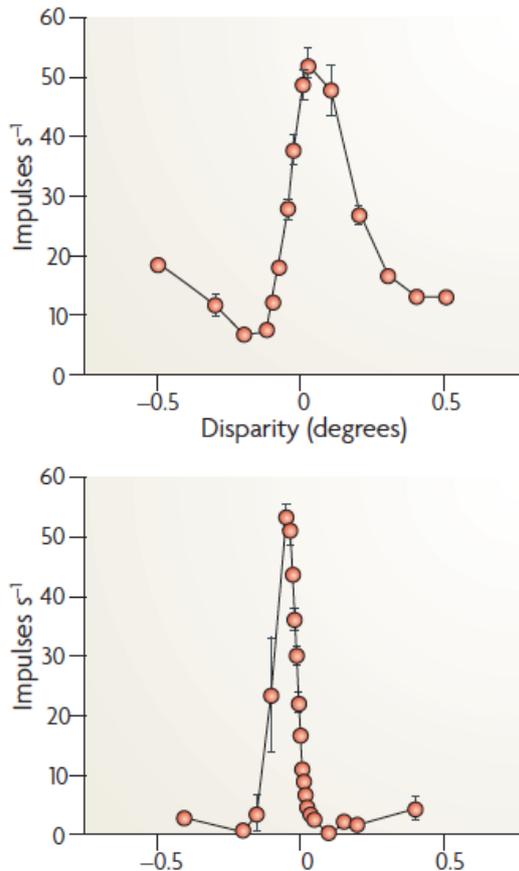


Figure 9. Schematised disparity tuning curves of two neurons in the primary visual cortex.

A tuning curve can be fitted by a Gabor function (i.e. a Gaussian function multiplied by a sine function). The neuron on top is tuned to uncrossed disparities (positive value) around 0.1 degrees, whereas the neuron below is tuned to crossed disparities (negative values) around -0.15 degrees. It is worth noting that despite having a strong preference for one disparity value, those binocular cells will still respond to a broader range of disparity values, hence making it impossible to solve the correspondence problem in a direct manner by relying on neuron tuning values. Adapted from Parker, 2007.

it is less straightforward for complex cells whose larger RFs are not as spatially constrained, making it potentially harder to detect disparity. Yet, single-cell recordings show that complex cells are well suited for fine disparity values detection (Poggio et al., 1985). With their disparity energy model based on previous computational work (see for example Qian, 1994) and on their own empirical data, Ohzawa and collaborators (1990) proposed that complex cells receive their inputs from the squared output of several quadrature pairs of binocular simple cells, that is, two simple cells whose left and right ON/OFF receptive field subregions are shifted by 90° (

Figure 10), acting like correlation detectors. Their model fits well the binocular RF shapes of complex cells recorded in the cat striate cortex (Ohzawa et al., 1990) and in macaque V1 (Cumming and Parker, 1997, but see Read et al., 2002) in response to sinusoidal gratings and dynamic RDS, respectively. Interestingly, this model also explains quite nicely the neuronal responses of V1 cells to anti-correlated dots, which are dots with opposed polarity in each eye. They contain the same disparity information as correlated dots (dots with the same polarity in both eyes) but do not lead to depth perception. The disparity energy model predicts an inverted response with an inversion of the disparity-tuning curve and such patterns could generally be found empirically in V1 (Cumming and Parker, 1997). However, a marked difference was found in higher visual areas, with neurons in V4 showing an attenuated response to anticorrelated RDS compared to correlated RDS (Tanabe et al., 2004), and neurons in the IT cortex not responding at all to anticorrelated dots (Janssen et al., 2003). Yet, it appears that despite its imperfections and limitations – among others, the fact that there is so far no anatomical evidence of the quadrature arrangement, or that this model being based on data recorded in the cat, some differences are expected when extrapolating it to the primate (e.g., a smaller fraction of simple cells in the primate brain than in the cat brain implies that complex cells might receive their inputs from other cells than simple cells) – the disparity energy model provides an interesting framework to explain some important features of V1 complex cell properties such as their RF shape or to some extent (some decrease were found in the neuronal response that are not predicted by the model) their responses to anticorrelated dots.

So far, the aforementioned studies tell us that the visual system relies on cross-correlation to effectively derive disparity, at least in the early processing stage taking place in V1, but that this stage is not sufficient to elicit depth perception, which takes place later in the visual hierarchy.

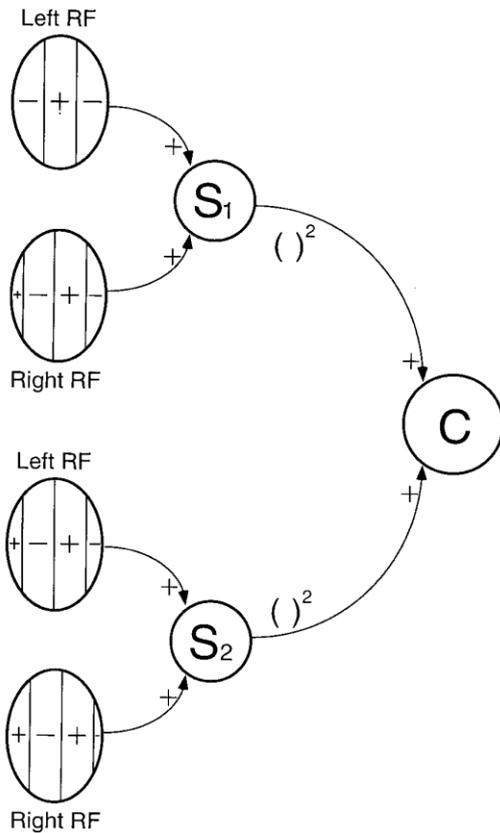


Figure 10. The disparity energy model.

A binocular complex cell C (on the right) receives and sums the squared output of binocular simple cells S arranged in quadrature pairs that are mutually inhibitory. This hierarchical organisation was proposed by Ohwaza and Freeman (1990) as a solution to the correspondence problem. Adapted from Qian, 1997.

Absolute versus relative disparity

Another important aspect of disparity processing in the visual system is the difference between absolute and relative disparity. Indeed, psychophysical studies in human (Westheimer, 1979; Erkelens and Collewijn, 1985; Prince et al., 2000; Chopin et al., 2016) demonstrated that stereoacuity is higher when one feature or one object can be compared to a reference (e.g. a background feature), which defines disparities as relative. However, macaque V1 neurons encode the absolute disparity of a visual object (Cumming and Parker, 1999; Nienborg and Cumming, 2006), which is defined as the difference of the visual angles between the retinal projections of an object on both eyes and the fixation point (Figure 11), or more simply put as the interocular difference of the stimulus retinal projection. Although relative disparities depend on absolute ones, a major difference is that the absolute disparity

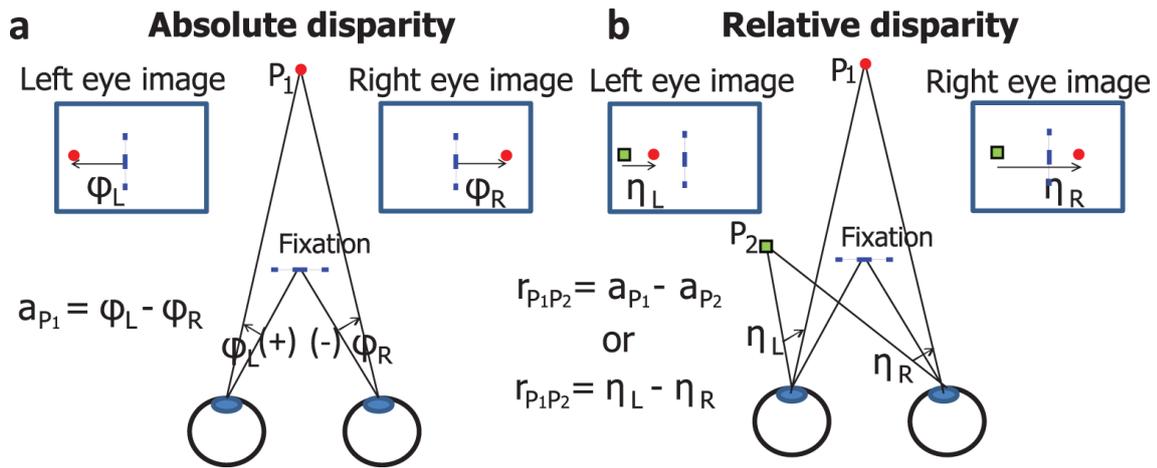


Figure 11. Absolute and relative disparities.

A) The absolute disparity of the object P_1 , a_{P_1} , is defined as the difference between the visual angles ϕ_L and ϕ_R , which are the angles between the fixation point and the projection of the object P_1 on each eye. B) The relative disparity between two visual objects P_1 and P_2 , $r_{P_1P_2}$, can be defined as the difference between the absolute disparities of both objects ($a_{P_1} - a_{P_2}$), or as the difference in each eye of the visual angles between P_2 and P_1 ($\eta_L - \eta_R$). Adapted from Chopin et al., 2016.

of an object changes with the angle of vergence (and therefore with the viewing distance), whereas relative disparities are not sensitive to vergence movement (Cumming and Parker, 1999). The search for neurons encoding relative disparities started in order to understand how absolute disparity was interpreted and turned into relative disparity, and thus, to explain psychophysical data reported in humans. A classical investigation strategy has been to use RDS in the form of a central disk surrounded by an annulus (Figure 12) and to record neuron responses when depth differences between the two parts of the stimulus were introduced. If changing the depth of the surround, either by bringing it in front or by pushing it behind the fixation plane where the central part is located, induces a shift in the tuning curve of a neuron that corresponds to the change of the surround disparity, then it can be concluded that the neuron encodes relative disparity (Thomas et al., 2002). The disparity tuning of that neuron will be independent of the eyes position (vergence movements), and thus of absolute

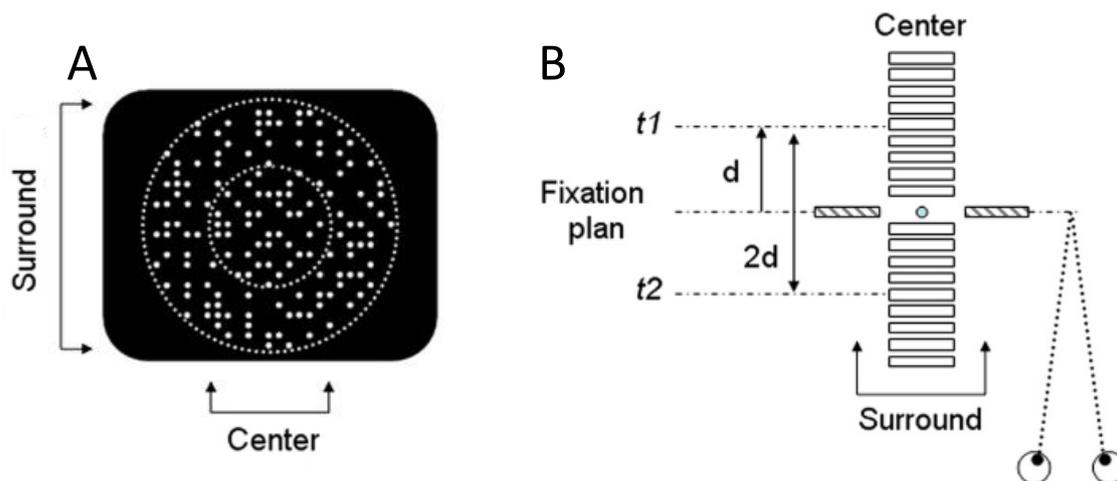


Figure 12. Centre-surround stimuli to assess responses to absolute vs. relative disparities.

A) Front view of the stimulus composed of random dots: The central part of the stimulus is surrounded by an annulus. B) Relative depth differences between the centre and the surround are introduced by moving the surround part back and forth across the fixation plane. This creates the percept that the surround is either in front or behind the centre. Here, the central part has a null disparity: it is on the same plane as the fixation point. Adapted from Cottureau et al., 2011.

disparity. As shown on Figure 13, Thomas, Cumming, and Parker (2002) found neurons in V2 that responded to relative disparity. However, the shift does not completely overlap with the value of the disparity change and most recorded neurons were still selective for absolute disparity. Umeda and colleagues (2007) recorded neurons from the V4 area which receives direct input from V2 and found that the shift of neurons tuned for relative disparity was more closely matching the disparity change of the surround (i.e. the magnitude of their shift was larger than those of V2 neurons), albeit still not ideally aligned with it. Their result suggests that relative disparities are progressively encoded along the visual hierarchy and follows a pathway that involves the areas V2 and V4. To broaden our knowledge on the matter, a more global view on disparity processing would be beneficial but so far still rare due to the limited

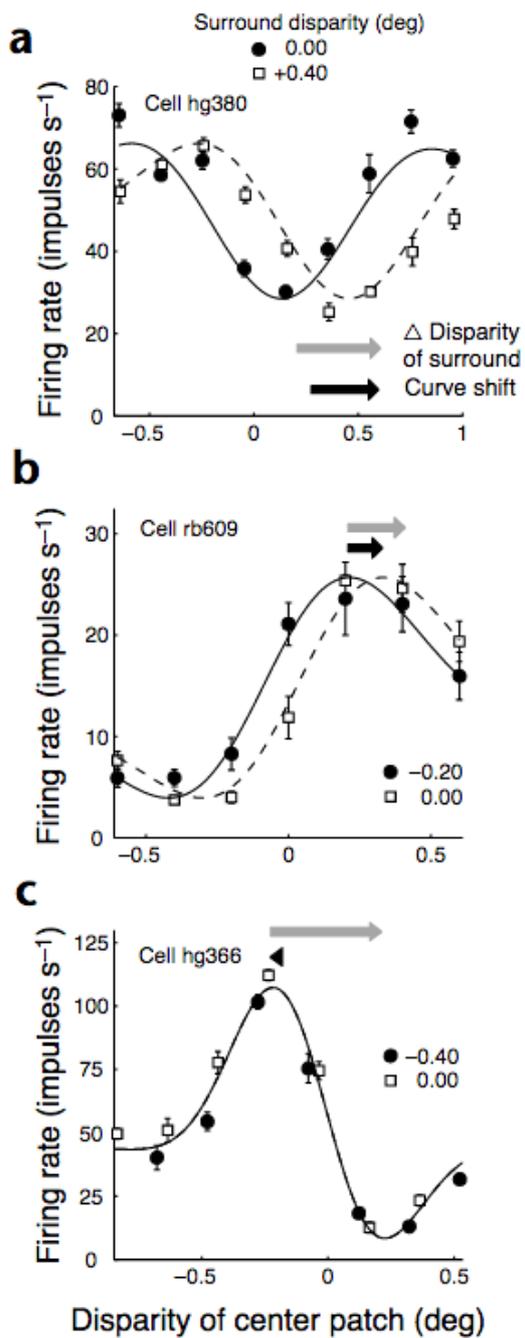


Figure 13. Responses of three neurons recorded at two different surround disparities.

Neurons A and B are tuned to relative disparity: the shift in their tuning curves corresponds more or less to the different surround disparities that were displayed in their receptive field. However, neuron C is tuned for absolute disparity: no shift is present. Adapted from Thomas et al., 2002.

spatial resolution of single-unit recording. However, with the development of new recording techniques, the investigation of disparity processing in human at the whole-brain level was made possible. In agreement with macaque data, neuroimaging studies in human (and in macaque for one of them) have evidenced the implication of several areas in disparity processing (e.g., Backus et al., 2001; Tsao et al., 2003; Neri et al., 2004; Welchman et al., 2005; Preston et al., 2008; Georgieva et al., 2009; Cottureau et al., 2011), demonstrating that it is spread in the whole visual cortex in both species. Two studies further demonstrated similarities with monkeys in response to correlated vs. anticorrelated stimuli. Bridge and Parker (2007) were the first to show an overall reduction of BOLD responses to anticorrelated RDS in higher visual areas. Using a different approach (i.e. they decoded specific activity in targeted brain areas, with a multivariate pattern analysis), Preston et al. (2008) re-established Bridge and Parker's findings (2008), further demonstrating a separation between lower visual areas and higher visual areas in their responses to correlated vs. anticorrelated stimuli, thus also confirming the progressive attenuation of responses to anticorrelated stimuli that were found in macaques (Tanabe et al., 2004; Janssen et al., 2003).

Overall, it seems that, in both species, disparity detection happens at the low level but that it is only in higher visual areas that stereopsis, or actual depth perception, emerges.

Disparity configuration and natural statistics

Several orders of depth can be defined and they have been shown to involve different brain areas (see e.g. Orban, 2011 for a review). The first one, zero-order depth corresponds to a flat surface parallel to the fixation plane that can be located in front (crossed disparities), within (zero disparity), or behind (uncrossed disparities) the fixation plane (Figure 14-A). This kind of stimulus corresponds to an absolute disparity configuration. A disparity gradient or first-

order depth defines a flat surface oriented in depth (i.e. slanted or tilted surfaces, as in Figure 14-B) and a disparity curvature or second-order depth characterises a curved surface (Figure 14-C). These two latter configurations involve the processing of relative disparity to be properly perceived. Importantly, those different depth orders involve different cortical areas. Indeed, whilst most of the visual cortex processes zero-order depth (see the first part of this chapter), only fewer areas are able to process more complex orders (i.e. first- and second-orders). For instance, areas V3A and CIP were found to respond to disparity-defined 3D orientations both in macaque (Shikata et al., 1996; Taira et al., 2000) and in human (Ban & Welchman, 2015). In macaque, disparity-sensitive neurons of areas MT and MSTd were shown to be also sensitive to tilted and slanted surfaces (Nguyenkin and DeAngelis, 2003;

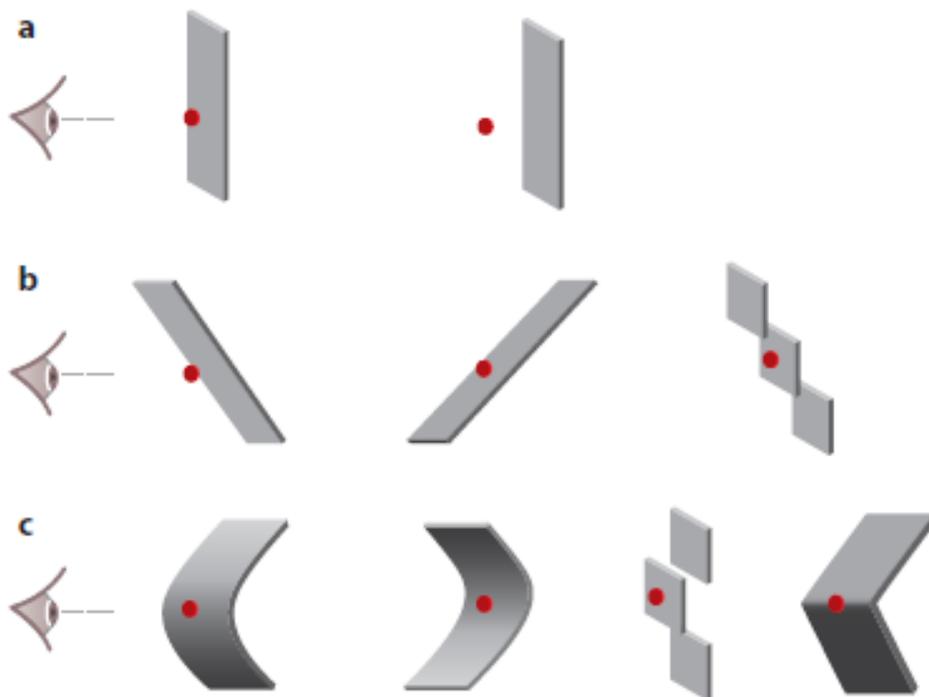


Figure 14. Depth orders. Three orders of depth can be defined and result in different types of disparity-defined stimuli as illustrated here. A) Zero-order depth corresponds to a flat surface parallel to the fixation plane. B) First-order depth takes the shape of a surface oriented in depth or of different planes located at different distances from the fixation point, looking like stairs. C) Second-order depth depicts curved surfaces. Adapted from Orban, 2011.

Sugihara et al., 2002). Regarding second-depth order, it seems that the inferotemporal region, along with the parietal cortex (post-parietal and intraparietal sulci) are largely involved in processing 3D curvatures and 3D shapes, as demonstrated in macaque (Janssen et al., 2001; Durand et al., 2007, 2009; Yamane et al., 2008) and in human (Georgieva et al., 2009).

These different depth orders are defined by a gradient of binocular disparities and rely on a spatial integration of those disparities. Because 3D gradients in the environment are similar across visual scenes, it has been suggested that the spatial integration of binocular disparities might be affected by natural statistics. Given the finite resources of the brain and the computation required to solve the binocular problem, that is to extract precise depth information from the overlap of two retinal inputs, it was proposed that the visual system might optimise the way it processes visual information by allocating most of its resources to more likely visual patterns (Barlow, 1961; 2001). Various theories and models have been made as to how the visual system might efficiently narrow its processing (e.g. efficient coding, predictive coding) but all involve some degree of statistical learning and probabilistic computation.

The link between natural statistics, visual perception, and cortical processing was mostly investigated for monocular visual properties such as luminance and contrast (e.g., Simoncelli and Olshausen, 2001; for a review, see also Geisler, 2008), as estimating those statistics is more direct than for binocular properties, which require dealing with two slightly different views of the world. Nonetheless, a few studies have attempted to measure the natural statistics of depth and later of binocular disparities. For instance, Yang and Purves (2003) documented the distribution of distances in natural scenes and suggested that some incorrect distance estimations might result from probabilistic expectations (i.e. statistical learning of some visual scene properties). Following studies estimated the distribution of binocular disparities by modelling the eye fixation behaviour of observers and revealed that the range of disparities is narrow and centred around 0 degrees (Hibbard, 2007; Liu et al., 2008). More recently, Sprague

and collaborators (2015) developed a device that was simultaneously recording a subject's eye movements and the observed visual scene. They showed that the distribution of binocular disparities varies across the visual field and mostly with elevation. Comparing their results with electrophysiological recordings in macaque V1, they revealed a similar effect of elevation in the disparity preference of neurons that was depending on the location of their receptive field. The human disparity distribution and the macaque disparity preference at the neuronal level follow a disparity gradient: Nearer disparities are more present in the lower visual field and further disparities are more in the upper visual field. More neurons with a receptive field in the lower visual field prefer near disparities and more neurons with a receptive field in the upper visual field prefer far disparities. Interestingly, a high-field fMRI study conducted in human also found such a pattern in areas V2 and V3 at the population level (Nasr and Tootell, 2016).

Altogether these studies reveal the existence of a correlation between the disparity distribution in visual scenes and the processing of those disparities in the early visual system (areas V1 in macaque and V2/V3 in human).

However, it remains unclear whether the cortical processing of binocular disparities in higher brain areas also reflect their spatial regularities and whether depth perception could be biased consequently.

During my doctoral studies, we investigated this first point by focusing on first-order depth and 3D orientations with a neuroimaging approach in macaque. Indeed, continuous disparity gradients make ideal stimuli because they recruit higher brain areas (e.g. areas V3A and CIP) than non-slanted surfaces but are simple enough not to recruit areas that are specialised in object processing as 3D shapes or 3D structures would.

In a second step, we asked whether this cortical bias might have an influence on depth perception, designing a psychophysics experiment for both human and macaque subjects.

Temporal integration of binocular disparities

Beside the spatial component, there is also a temporal dimension to the integration of binocular disparities by the visual system. This temporal integration underlies our ability to interact with the environment, as we need to compute 3D trajectories of moving objects either to catch or grasp them, or to avoid colliding with them. Despite its ecological relevance, this dynamic component has received less attention compared to the spatial one, mainly due to the fact that most studies investigating movement processing focused on planar or bidimensional movement. However, it has been suggested that 3D motion and 2D motion processing involve different cortical pathways, at least in human (Cottareau et al., 2014). Furthermore, while the MT cluster in macaque and the hMT+ complex in human are known to be homologous central hubs for the processing of 2D motion, it is not clear whether 3D motion might recruit areas that are further away in the visual system. For instance, although a couple of neuroimaging studies conducted in human found a consistent network of areas responding to 3D motion, they also revealed some discrepancies, notably regarding an area anterior to the hMT+ complex (Likova and Tyler, 2007; Rokers et al., 2009; Kaestner et al., 2019). In macaque, two studies recorded responses to 3D motion in area MT (Czuba et al., 2014; Sanada and DeAngelis, 2014) and further suggested the involvement of this area in processing 3D motion.

As a way to bridge the link between human fMRI studies and macaque neuronal data, we proposed to identify the cortical network involved in processing 3D motion in macaque, using an fMRI approach and to determine whether there might be an area anterior to the MT cluster that would respond more strongly to 3D motion, thus clarifying the role of both the MT cluster and its surrounding regions in processing 3D motion.

Methodological approach

Be it for spatial or for temporal disparity integration, we could see so far that our knowledge about the primate visual system when it comes to processing binocular disparities is based on both electrophysiological recordings in macaque and functional neuroimaging studies in human, and sometimes also in macaque. These two techniques are complementary: On the one hand, the measured single-unit activity gives a very precise idea about a cortical mechanism in one specific region but does not offer a global perspective. On the other hand, the blood activity of several brain areas is investigated at the macroscopic level but suffers from a limited temporal resolution. However, this major difference in the signal nature also makes it difficult to compare electrophysiological data with neuroimaging recordings collected in macaque and in human, respectively. Thus, it appears necessary to use an approach that can bridge the gap between those different studies. During my doctoral thesis, I focused on the development of functional neuroimaging studies in macaque to investigate and characterise the cortical networks involved in the processing of gradients of disparity. With such an approach we were able to make interspecies comparisons and reveal potential homologies and differences between macaque and human.

In the next chapters, I will first present the methodological aspects that the neuroimaging approach in macaque encompasses (Chapter III). I will then detail two experimental studies where we investigated the cortical networks involved in the processing of both temporal (Chapter IV) and spatial (Chapter V) components of disparity gradients. To bring the topic further, the influence of natural statistics will also be questioned (Chapter V).

Chapter III –
Monkey fMRI - Methodology

The BOLD signal

Physics of (functional) magnetic resonance imaging

Magnetic resonance imaging (MRI) relies on the physical properties of positively charged particles present in the nucleus of an atom, the protons. Some protons, when exposed to an external magnetic field, naturally align with the field direction in a rotating manner, creating a tiny magnetic field at the surface of the proton. This motion is known as their ‘spin’ and is characterised by both a magnetic moment and an angular momentum⁶. The sum of magnetic fields created by all the spins within a system is the net magnetisation vector, which is usually null (due to random motion directions of spins). When submitted to a strong static magnetic field, like the one emitted by MRI scanners, spins will align in a non-coherent way either with the direction of the field or with the opposite direction (Brown et al., 2014). The idea of MR imaging is then to send radio frequency (RF) pulses through electromagnetic coils at a specific frequency, the Larmor frequency, that will disrupt the orientation of the protons’ spins aligned with the scanner field and leads to a transverse magnetisation when spins phase coherence increases. The net magnetisation vector is no longer null and orients transversely to the magnetic field. Some atoms, like hydrogen atoms (^1H) that are present in water molecules and thus in body tissues, will absorb those RF pulses via their protons, modifying their spin phase (excitation). This interaction phenomenon between spins and electromagnetic RF is known as

⁶ The spin motion of a proton about its axis (or its alignment) generates an electrical current on its surface, creating a magnetic field. The magnetic moment corresponds to the strength of this small magnetic source and requires an odd number of protons to be different from zero, whereas the angular momentum represents the relation between the mass of the proton and its angular velocity during the spin motion. Both properties are necessary for a nucleus to be useful for MRI (Brown et al., 2014).

magnetic resonance⁷. Once the RF pulses are no longer emitted, protons of hydrogen atoms will release the accumulated energy in the form of radio waves at the same Larmor frequency and return to the initial alignment created by the scanner-generated magnetic field (equilibrium state) – this is the transversal relaxation. Receiving coils will detect this emitted RF signal (reception) in the form of a changing current. This constitutes the MR signal. Changes in this signal take the form of two types of relaxation: decay and recovery. Decay (or transverse relaxation) represents the loss of phase coherence of the spins, whereas recovery (or longitudinal relaxation) represents the difference between the number of protons still releasing acquired energy (high-energy state) and the number of protons that no longer release energy (low-energy state). Both variables of the MR signal are represented by time constants. T_1 (recovery) describes the time it takes for the system to come back to its initial state (where much more protons are in the low-energy state), and T_2 (decay) describes the decay time of the transverse magnetisation of the net magnetisation vector, with T_2 being shorter than T_1 . On top of spin coherence loss, local inhomogeneities in the magnetic field also play a role in the transverse magnetisation decay. This additive effect is described by the T_2^* (decay) time constant (Brown et al., 2014; Huettel et al., 2009).

⁷ To give a broad idea of how fast/long MR imaging technique developed, it is only in the late 1930s that Isidor Rabi, leading him to be awarded the Nobel Prize in Physics in 1944, demonstrated the effects of nuclear magnetic resonance. The discovery of nuclear magnetic resonance per se is to be attributed to Edward Purcell and Felix Bloch who were awarded a joint Nobel Prize in Physics in 1952 for their work. In 2003 Paul Lauterbur and Peter Mansfield were awarded a joint Nobel Prize in Physiology and Medicine for developing techniques to form images using MRI. Their work was based on the fundamental contribution of Raymond Damadian who suggested a potential use of MRI for biological applications in 1971, after decades of mostly chemical use of MRI (Huettel et al., 2009).

Varying the pulse sequence parameters will allow to generate different contrasts, such as T_1 -weighted or T_2 -weighted static contrasts, commonly used for anatomical acquisitions; or T_2^* -weighted contrasts mostly used for BOLD-contrast fMRI. These contrasts reflect the amplitude of the recovered MR signal, and as a consequence, the intensity difference between different types of tissue, each tissue having its respective properties, as measured by the scanner. Following the RF pulse that targets protons in one 2D-slice⁸ of the brain, a phase-encoding gradient and a frequency-encoding gradient are sequentially applied to provide information about localisation of spins. They will introduce variations in the magnetic field strength in a local manner, giving information about the contribution of different image locations via spin modulation: a voxel position is encoded by its spin frequency. The raw MR signal is encoded in the k -space that contains spatial frequency information from all locations of the imaged object, whilst the interval between a RF pulse and the data acquisition in the k -space is known as the echo time (TE). To reconstruct brain images, an inverse Fourier transform is applied on the k -space. In a spin-echo imaging sequence, one line of imaging data (one line in the k -space) is acquired during one TR. The image acquisition is quite slow, as it requires multiple repetitions of the same pulse sequence to obtain a whole-brain three-dimension image. Nonetheless, this technique allows a detailed view of the brain thanks to its high spatial resolution. In complement to the spin-echo imaging, echo-planar imaging (EPI) is based on gradient echo (GE) sequences and is sensitive to T_2^* -weighted contrast. This faster imaging technique relies on the principle that the whole k -space of an image has to be filled following a single RF pulse (i.e. in one TR), which requires that signal acquisition happens before MR signal has decayed too much, thus in a very short time. Although already proposed by

⁸ Slices are usually acquired in an interleaved manner in order to avoid contamination effects of RF pulse excitation from one targeted slice to the surrounding slices.

Mansfield in the 1970s (Mansfield, 1979), EPI technique has been developed only recently due to the need of a very strong gradient system (larger gradient amplitude) for the fast acquisition of images. After a single RF pulse is sent, a dephasing spatial gradient is applied to form a train of gradient echoes that will alter the spin phase very quickly. Each oscillation corresponds to one line of imaging data and is accompanied by a rapidly alternating frequency gradient that will fill-up the k -space in a zigzag manner: lines are acquired in a back-and-forth trajectory instead of the usual left-to-right direction, requiring the realignment of images after the acquisition is done (Slichter, 2013). EPI allows the acquisition of multiple lines of imaging data in one single TR, making the acquisition of a full brain image much faster and thus less sensitive to motion artefact, at the expense of a high spatial resolution. However, EPI imaging technique is very susceptible to field inhomogeneities that can result in signal loss or image distortions (Weiskopf et al., 2006) and will require dedicated pre-processing steps.

Now that we know more about the creation of a brain image using MRI technique, we shall focus on how to measure brain activity, using functional MRI. In other words, how do we go from MRI acquisition to functional MRI?

Instead of looking at tissue properties, fMRI relies on a marker of brain functioning: the BOLD signal, which is a metabolic correlate of neuronal activity.

Nature of the BOLD signal and how it relates to the underlying neural activity

Blood-oxygen-level-dependent (BOLD) contrast was discovered in the early 1990s by Seiji Ogawa (Ogawa and Lee, 1990; Ogawa et al., 1990). Using strong magnetic fields and GE imaging sequences in rodents, he was able to show that the paramagnetic properties of deoxygenated haemoglobin in blood vessels could be used as a natural contrast agent for fMRI recording, indirectly reflecting brain activity.

Indeed, concentration of deoxygenated haemoglobin, or deoxyhaemoglobin, in blood vessels is an indicator of a brain region activity. Changes in the neuronal activity (e.g. in response to a stimulus) leads to an increase in oxygen consumption (see Figure 15). Oxygen is bound to haemoglobin molecules, or oxyhaemoglobin, and will be used to create adenosine triphosphate (ATP) that carries energy within cells, by oxidising glucose. Oxygenated haemoglobin then becomes deoxygenated. To supply the brain region with energy, cerebral blood flow is increased to bring glucose and oxygen, in the form of oxyhaemoglobin, which will then replace deoxyhaemoglobin. Given their magnetic properties, deoxyhaemoglobin molecules alter the spin of nearby hydrogen nuclei that are present in diffusing water molecules, thus resulting in a reduced intensity of the MR signal. When deoxyhaemoglobin is replaced with oxyhaemoglobin, this results in a local increase of the MR signal (Huettel et al., 2009). These signal differences represent the core of BOLD contrast⁹.

By reflecting the haemodynamic response within a brain region that underwent an increase in its neural activity, the BOLD signal represents the interaction between cerebral blood oxygenation, cerebral blood flow, and energy demand. Most importantly, fMRI relies on the assumption that cerebral blood flow and oxygenation vary as a function of neural activity, expressed by an increase in energy demand.

However, what the BOLD signal tells us about the brain functioning, in the sense of the underlying neural activity it might represent, remains to be fully elucidated.

In a very elegant study, Logothetis and collaborators (Logothetis et al., 2001) simultaneously recorded BOLD signal, local field potentials (LFPs), and single- and multi-unit spiking activity

⁹ These signal differences will be represented on T_2^* -weighted images and will be a function of the ratio between oxygenated and deoxygenated haemoglobin present in the blood flow.

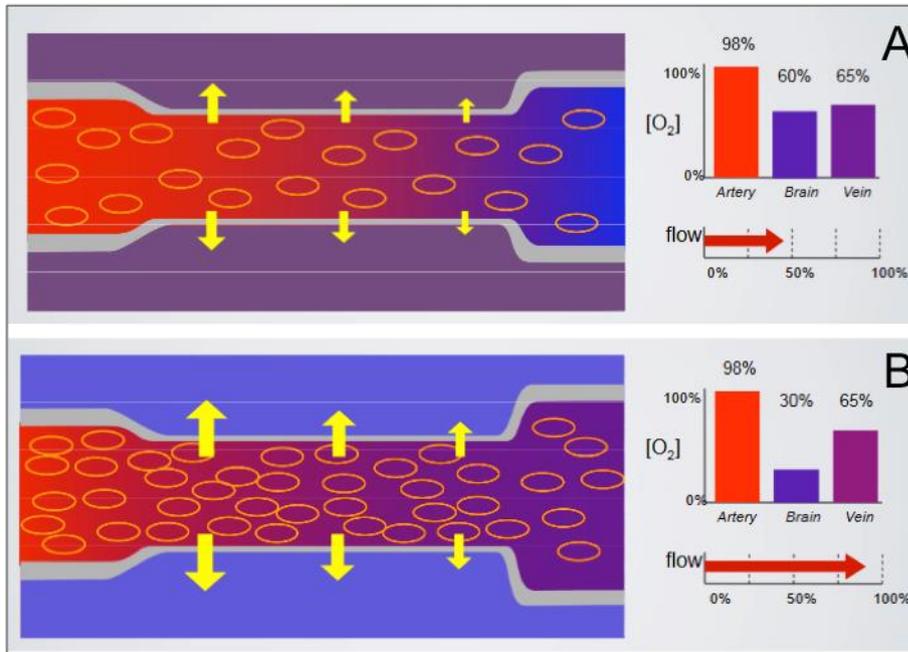


Figure 15: The BOLD signal reflects variations in oxygen consumption and modification of blood flow.

The upper panel represents a baseline condition, whilst the lower panel represents an active condition: dioxygen concentration drops at the brain level and blood flow in the vessels increases to compensate for O₂ decrease, bringing new O₂ molecules to the active brain region.

<http://www.brainmapping.org>

(MUA)¹⁰ of neurons from the brain of anaesthetised macaques (see Figure 16). They found a linear relationship between MUA, LFPs, and BOLD signal for short duration stimulus presentations, as well as a strong correlation between LFPs and BOLD and a slightly weaker but still significant correlation between MUA and BOLD. Their findings indicate that the

¹⁰ LFP tends to reflect the output of a targeted area as well as local intracortical processing, with a bias towards pyramidal cells. Single-unit extracellular recording provides information about neuron spikes, with a bias towards larger neurons. MUA represents the linear sum of extracellular action potential of neurons that are within the range of the electrode measuring them (i.e., 140-300 μ m radius) (Logothetis, 2003).

BOLD signal is more related to the peri-synaptic activity (post- and pre-synaptic activity of multiple neurons)

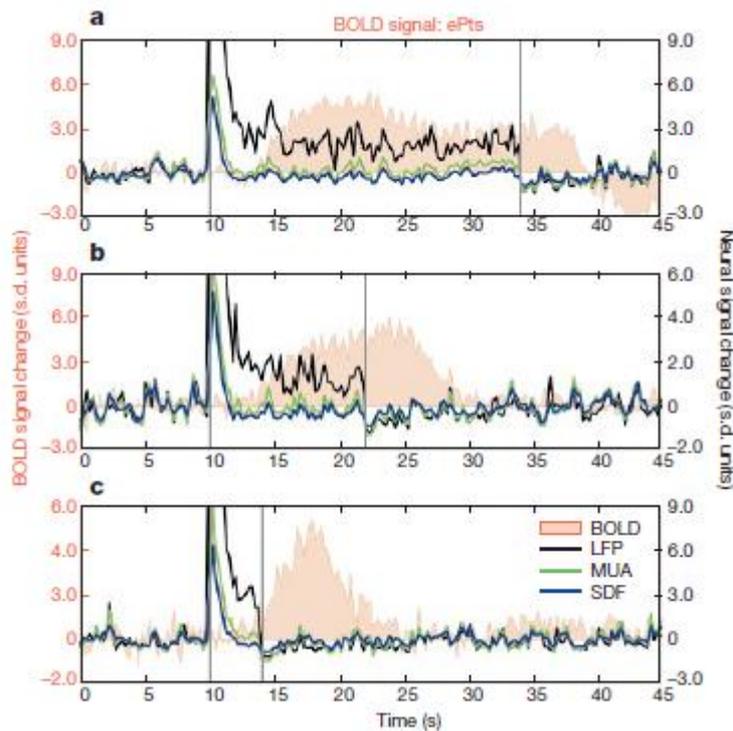


Figure 16: Simultaneous neural and haemodynamic recordings from a cortical site (in the striate cortex) showing transient neural response.

Panels A to C: Responses to a pulse stimulus of 24, 12, and 4 s. Both single- and multi-unit responses adapt a couple of seconds after stimulus onset, with LFP remaining the only signal correlated with the BOLD response. (SDF, spike-density function; ePts, electrode ROI-positive time series.). Adapted from Logothetis et al. (2001).

than to the neuronal output itself (spikes)¹¹. Those correlations were later found in the awake monkey, albeit weaker (Goense and Logothetis, 2008). A study conducted in humans compared single-unit activity and LFPs from the auditory cortex (Heschl's gyrus) of two implanted

¹¹ See Figure 17 for a schematic illustration of the different cells and their synapses that are comprised in a voxel.

patients to the BOLD signal of control subjects and found a strong correlation between the three variables (Mukamel et al., 2005). This latter study, although using a

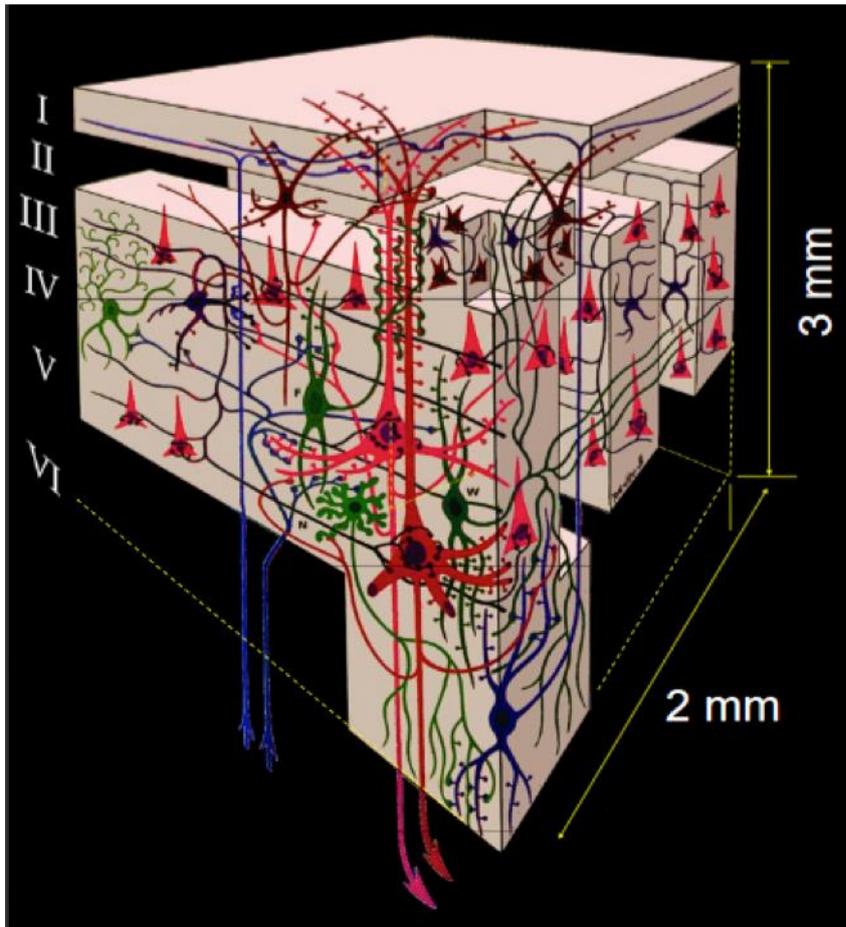


Figure 17: Illustration of the different types of cells and synapses that are comprised in a cube that has the equivalent size of one voxel. Adapted from UCLA Brain Mapping Division

slightly different approach than the one developed by Logothetis and colleagues, confirms the relationship between the BOLD signal, LFPs, and single-unit activity, and evidences similarities between the two primate species. These results, however, are limited to early sensory areas (both visual and auditory) and cannot be extended to the whole cortex. This is especially true, since some studies have found dissociation between the BOLD signal and spike rate and between LFPs and spike rate, suggesting that the correlation between BOLD and spike rates (or MUA) might be due to correlations between LFPs and spike rate and not to a relation

between BOLD and spike rate per se. This dissociation was found in the human hippocampus where spikes are sparser (see e.g. Ekstrom et al., 2007), as well as in the macaque infero-temporal cortex (Kreiman et al., 2006) and favours the BOLD-LFP coupling model (for a review, see Ekstrom, 2010).

Based on those previous studies, it appears that our understanding of what the BOLD signal represents, although strongly correlated with LFPs, is still quite limited. A good way to advance on that question and to address the limitations inherent to any recording technique would be to fill the gap between unitary recordings in macaques, which provide information about LFPs and spikes in a very local manner, and functional neuroimaging in humans, that gives a macroscopic view on the functioning brain. This approach has led to the development of monkey fMRI with the idea that answering the same questions with different techniques (i.e. single-cell and fMRI recordings) in macaque, and comparing results with human fMRI, then using an identical recording technique between both species, can lead to address limitations of each approach, and provide us with the unique opportunity to refine comparisons and reveal homologies and differences between humans and macaques.

Following the aforementioned studies in the anaesthetised monkey, Vanduffel and collaborators in Leuven were among the first to start developing monkey fMRI in the awake, behaving macaque (Vanduffel et al., 2001). Beyond the aim of ‘simply’ understanding the mechanism of fMRI, the idea was to make a link between human neuroimaging findings and monkey electrophysiological studies, investigating functional responses to motion stimuli. Their pioneer work enabled to map the function of higher visual brain areas, in the superior temporal sulcus (MT cluster) and in the intraparietal sulcus (ventral intraparietal area, VIP), and to reveal both homologies and differences with human studies (e.g., functional differences in the area V3A). Overcoming the limits set by the use of anaesthesia under which the highest area reached

was the MT area in a study where a nystagmus was induced (Leopold et al., 2002), their work implies that targeting the function of higher brain areas necessitates an awake state.

However, working with the awake monkey requires a specific conditioning to habituate the animal to remain still and perform a task, as well as the development of a dedicated setup. Conducting their recordings in the same MR scanner as for humans, Vanduffel and colleagues (2001) designed a plastic chair for monkeys to sit in (Figure 18), design that we replicated for our own experiments. In this configuration, the monkey is restrained with a MR-compatible headpost attached to the skull¹² and locked to the chair to limit head movements in the scanner. A rewarding system is attached to deliver water based on fixation performance. Prior to recording sessions, monkeys are trained to get used to the chair in which they sit in a sphinx position, and to have their head fixed. They learn to remain still and to ignore scanner noise, whilst performing a passive fixation task for which they get rewarded. To facilitate this training, a mock scanner can be used, as it has been done in our team. Finally, when the monkeys are ready and behave properly, they can be introduced into the scanner and perform the first MRI acquisitions.

¹² Implanting the monkeys with the headpost is done under anaesthesia. Surgical procedures performed within our team were similar to the ones described in Vanduffel et al. (2001) and can be found in the first fMRI study published by our team, also attached to the appendix section (Cottureau et al., 2017). Briefly, the animals were administered glycopyrrolate prior to the anaesthesia to reduce secretions. They were intubated to induce anaesthesia with isoflurane and their head was placed within a stereotaxic holder. The headpost were attached directly on the skull with ceramic screws that were further cemented with a bone or a dental cement. During recovery, an analgesic and an antibiotic were administered daily to the animals for several days. The animals resumed their behavioural training about two months after surgery.

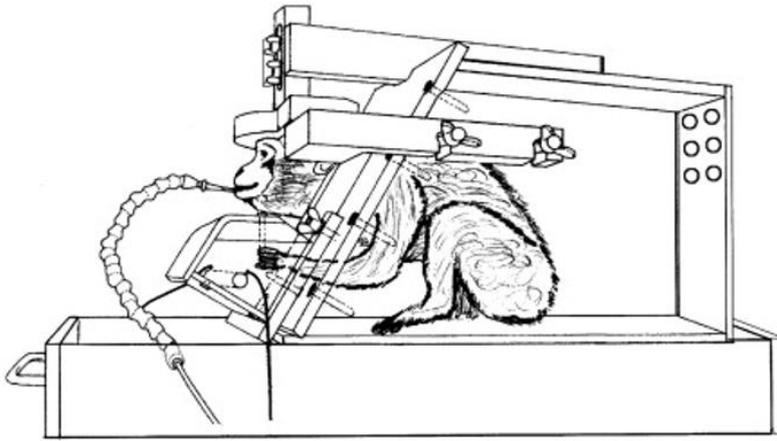


Figure 18: MRI-compatible chair for macaque.

The animal is in a sphinx position, its head fixed to the chair via the headpost attached to its skull. Reward fluid is brought to the mouth via a plastic tube. Adapted from Vanduffel et al. (2001).

BOLD signal vs. MION-enhanced signal

Being able to conduct comparative studies using similar functional neuroimaging protocols is not as straightforward as it might seem. Although macaque and human brains are very similar, macaque brains are much smaller, making it harder to achieve a good spatial resolution with a fair signal-to-noise ratio (SNR). Different strategies have been developed to improve functional sensitivity: high-field scanners specifically designed for macaques (e.g., vertical scanners at 4.7T; Logothetis et al., 2001), dedicated coils adapted to the skull of the animals to optimise signal reception (Cottereau et al., 2017), or the use of a MION (monocrystalline iron oxide nanoparticle) solution, an iron-based contrast agent (Vanduffel et al., 2001).

The latest option is probably the most spread among research groups conducting animal fMRI studies. However, despite improving the SNR (Leite et al., 2002) or being essential for low-field MRI recordings (Vanduffel et al., 2001), the use of MION does have inconveniences. Regarding the MION solution per se, the intravenous injection that it requires makes it stressful to handle for both the monkey and the experimenter, whereas the accumulation of iron in the

monkey organism can be deleterious on the long term even with the complementary use of an iron chelator, thus limiting the amount of recording sessions one subject might go through. On top of those points, the half-life time of MION is difficult to control perfectly and one may be concerned by the impact this might have on data collected across multiple sessions.

Another aspect that should be taken into account is the correspondence between BOLD and MION signals in order to make relevant comparisons between human and macaque fMRI. Differences in the signal sign (the impulse response function with MION is negative compared to BOLD), in the degree of non-linearity (Gautama et al., 2003), in the kinetics (slower onset), or in the vascular mechanisms that are reflected by both signals (Mandeville et al., 1999) require to be handled when comparing BOLD-based and MION-based neuroimaging studies, potentially limiting the direct comparisons. Based on those different considerations and given the scanner configuration (3 Tesla, dedicated to human and non-human animal research), a dedicated 8-channel phased-array coil (RapidBiomed ®) has been purchased within our team, with the main benefit of no longer requiring the injection of MION for our studies. However, although necessary, optimising recording setups is not sufficient. Due to inherent noise in the data, good data processing and data modelling are additionally required.

Pre-processing steps and general linear model (GLM)

The goal of fMRI analysis is to match changes in the BOLD signal with the experimental design in order to estimate the effect of our conditions onto the haemodynamic response. More precisely, the aim is to determine whether variations of intensity in one voxel across time series are correlated with the experimental design convolved with the HRF. Based on the onset time and the duration of the stimuli (i.e. the experimental design), we can make predictions about the BOLD response. The formula of a so-called regression analysis is the following:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

Where y represents the observed data, that is the time series of a voxel, x_i represents the different regressors (i.e. the predictions), β_i the contribution (weight) of those regressors to the observed data y , and ε the residual noise (i.e. error in the measurement).

In a general linear model (GLM), multiple regressions are done for each voxel separately in an independent manner. This kind of analysis applied to fMRI data is known as voxel-wise or univariate analysis and consists in statistically testing the relative contribution of each regressor (β values) to the observed data within a given voxel. Solving the GLM amounts to finding the best β parameters that maximise the explained variance whilst minimising the residual error ε . The statistical significance of the GLM solution can be assessed with a F statistical distribution or a t -test and will result in a statistical parametric map (Friston et al., 1994). However, minimising the residual error when solving the problem is not easy because of the noise contained in the data. Further steps need to be achieved before. First, raw data has to be pre-processed to take into account the order of slice acquisition, as it is often done in an interleaved manner (slice-timing correction), as well as various motion artefacts (registration). Functional images can also be realigned to an anatomical scan (co-registration) and sometimes to a standard brain template (spatial normalisation) to allow between-subject comparisons. Finally, some smoothing can also be applied to increase the SNR. Since motion artefacts in monkeys mainly consist of jaw movements due to the nature of the reward and of body movements, the usual rigid transformation performed onto human data cannot be applied to monkey data. The solution implemented so far consists in performing a principal component analysis (PCA) on voxels located outside the brain mask (e.g., muscles, eyes) to extract noise (i.e. artefact caused by the animal's movements) and use the principal vectors that are necessary to explain 80% of the variance in those voxels as regressors of non-interest in our statistical model (see Farivar and Vanduffel, 2014). Saccade vectors will also benefit the model. Finally, the optimisation of

the haemodynamic response function (HRF) also represents an important advancement in our quest for a high SNR and a minimised residual error.

Optimising the HRF estimation and data pre-processing to improve the SNR

Estimating the individual haemodynamic response function (HRF)

In commonly used general linear models (GLM) the haemodynamic response function (HRF) is convolved with the stimulus design. The idea is to take the delay of the BOLD response (or haemodynamic response) into account to link it to the stimulus paradigm in order to estimate how much of the variance in the time-course of a voxel is explained by the model's predictions.

The BOLD HRF is composed of 4 phases: rise, peak, fall, and undershoot, that represent the use and then the replacement of the paramagnetic deoxyhaemoglobin. In most neuroimaging data processing software, a canonical HRF is convolved to the experimental design, despite the fact that it varies between individuals, regions of the brain, or experimental sessions (Aguirre et al., 1998; Handwerker et al., 2004). Because of interspecies differences in the HRF parameters, we estimated the HRF for each of our monkey subject (N=2), following Dumoulin and Wandell (2008)'s procedure.

Full-field checkerboard (10 Hz) were displayed during 4 seconds, each time followed by a 30-s blank interval. A fixation point was in the centre of the screen where the monkey had to maintain fixation. Only runs for which fixation performance was above 85% were kept, which represents a total of 13 runs for M01 and 11 runs for M02. Data were pre-processed (slice-timing correction, normalisation of the mean functional individual template, no smoothing was applied) and projected on the surface of each individual.

Historically, several analyses have been developed in our team to determine the HRF. We first estimated the HRF based on all the voxels that responded to the visual stimulation, that is

voxels for which the SNR¹³ was above 3. We further restricted those voxels to the posterior half of the brain only, since this is where most visual responses take place. In the latest analysis pipelines, we chose to work directly on the surface by extracting the signal from the nodes that anatomically correspond to the V1 area. This allows the characterisation of the HRF parameters without having to use functional data based on an experimental design (i.e. data that would have to be convolved with an arbitrary HRF).

The time-course of those selected voxels / nodes during one stimulus cycle were averaged to estimate the HRF parameters. The HRF was derived as a response to a 2s-stimulus, which corresponds to the fMRI sampling rate used in our experimental setup. The HRF was parameterised as the linear combination of two gamma functions as described by the following equation:

$$h(t) = A \left(\frac{t^{\alpha_1 - 1} \beta_1^{\alpha_1} e^{-\beta_1 t}}{\Gamma(\alpha_1)} - c \frac{t^{\alpha_2 - 1} \beta_2^{\alpha_2} e^{-\beta_2 t}}{\Gamma(\alpha_2)} \right)$$

Where A is the amplitude, α_1 the response delay relative to onset, α_2 the undershoot delay, β_1 the response dispersion, β_2 the undershoot dispersion, and c the ratio of response to undershoot. The parameter values for each monkey are shown in Table 1 and a graphical representation can be found on Figure 19.

Table 1: Parameter values of the HRF for each individual

<i>Subject</i>	α_1	α_2	β_1	β_2	c
<i>M01</i>	2.8572	29.9973	0.9267	2.6957	10.0000
<i>M02</i>	4.7199	24.8772	1.2660	1.3247	6.3917

¹³ The SNR of a voxel was defined as follows: the signal is the amplitude of the Fourier coefficient at the stimulation frequency ($F = 1/34$), whilst the noise corresponds to the average moduli at the two neighbouring frequencies.

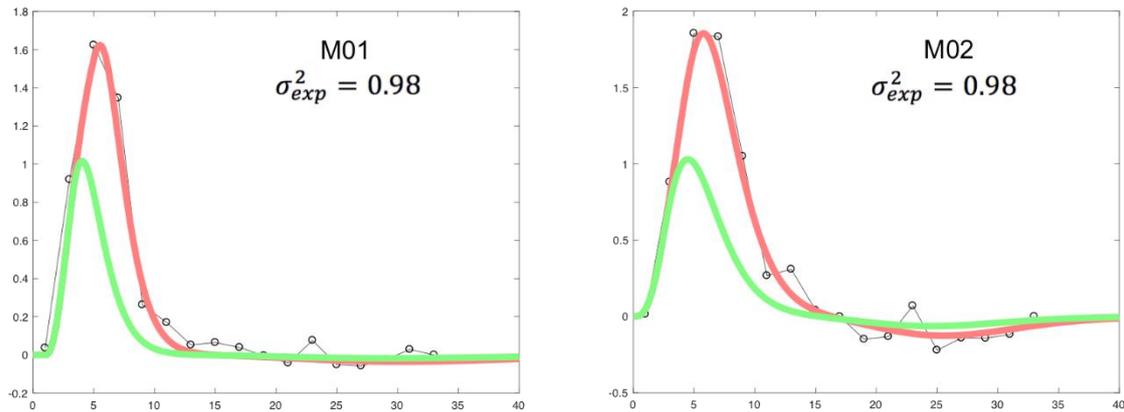


Figure 19. Characterisation of the haemodynamic response function (HRF).

Estimate of the HRF for each of our monkey are displayed on the left panel for M01 and on the right panel for M02. The sigma value indicates the goodness of fit of the model.

In future estimations and notably to further improve the resolution of the GLM, the HRF could be estimated separately for every node or voxels and then averaged to improve the estimation. It could also be interesting to characterise the variation of the HRF parameters in different brain regions, given that vascularity properties vary among different areas (Handwerker et al., 2004), and to compare how different values might affect statistical output.

Constant improvement of our methods: Complexity and success

Developing a platform dedicated to monkey MRI brings its share of challenges, at both technical and practical levels, and the attempts to address them do not always results in satisfying outcomes in a direct manner. However, once they are confirmed or established as valid, results are highly rewarding. The very first fMRI study published by the ‘CerCo monkey team’ (Cottereau et al., 2017)¹⁴ is the result of collaboration between engineers, animal technicians, researchers, and students during several years. Not only was it the proof that after

¹⁴ The paper can be found in the Appendix section: Cottereau et al., 2017

all the investment in terms of both time and money, the MRI platform could host functional neuroimaging research in the awake and behaving non-human primate, it also evidenced the fact that the BOLD signal can be used in macaque monkey without the need of adding a contrast agent such as MION, thus allowing direct comparisons between human and macaque. For instance, this paper sheds light on the network processing optic flow, which is a visual motion arising during locomotion and a primordial monocular cue that helps spatial navigation. Optic flow processing has been investigated in different brain areas of the monkey, with electrophysiological recordings. Those different studies have revealed a role of areas MST (Duffy and Wurtz, 1995), VIP (Zhang and Britten, 2011), VPS (Chen et al., 2011), and FEFsem (Gu et al., 2016). In parallel, neuroimaging studies in human have shown that the network involved in optic flow processing shares homologies with monkeys, with the involvement of areas MST and VIP (Wall and Smith, 2008; Cardin and Smith, 2010). These human studies also showed that other brain areas that had not been explored in macaque, such as the CSv area of the cingulate sulcus were also selective to optic flow. To address the question of whether such an area was also part of the monkey optic flow processing network, we collaborated with Andrew Smith who conducted the two human studies and replicated his findings in 3 macaques, using the same stimuli and experimental design. Two conditions were contrasted, one depicting an optic flow consistent with locomotion (EC for ‘egomotion consistent’) and another one where the optic flow was not consistent with egomotion (EI for ‘egomotion inconsistent’). The local properties (dot number, luminance, etc.) were equivalent between both conditions. The brain areas responding more strongly to the EC condition are shown in yellow to red colours on Figure 20 and the one responding more strongly to the EI condition are shown in blue. A sensitivity ratio (SR) has also been computed (see the paper for more details about the computation) to describe the strengths of BOLD activations between regions of interest for the

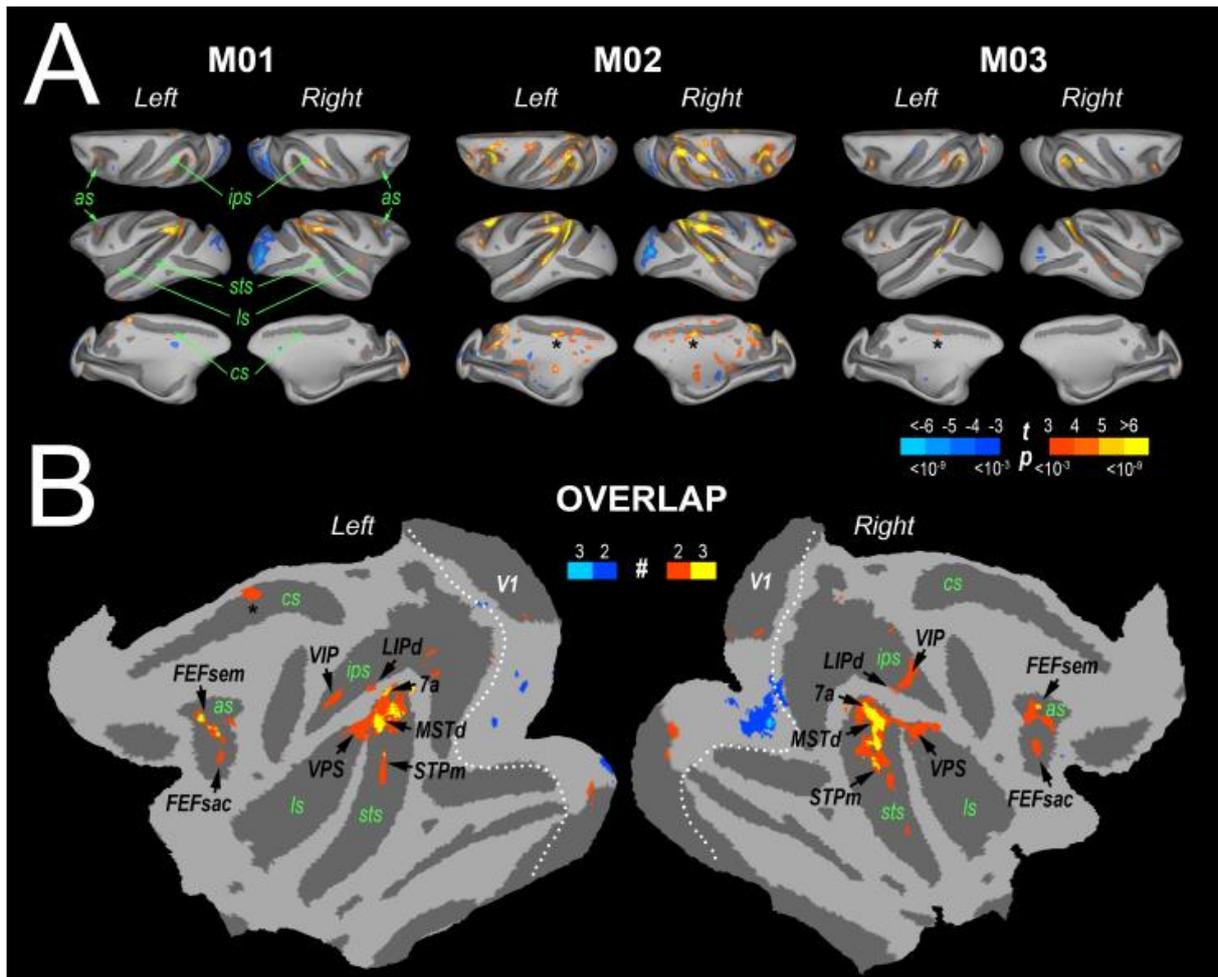


Figure 20: A) Statistical parametric maps for the EC versus EI contrast in monkeys M01, M02, and M03. The colour code reflects the contrast t-values and indicates statistically significant differences between BOLD responses evoked by the EC and EI conditions ($p < 10^{-3}$, uncorrected). Hot and cold colours indicate stronger responses to EC and EI, respectively. B) Map of overlap between significant activations in the EC versus EI contrast across the 3 monkeys. Only activation sites found in at least 2 individuals are shown.

EC condition compared to the EI condition. Figure 21 illustrates the SR for each animal averaged across hemispheres and shows the localisation of the eight brain areas that were found to be more activated for the EC condition.

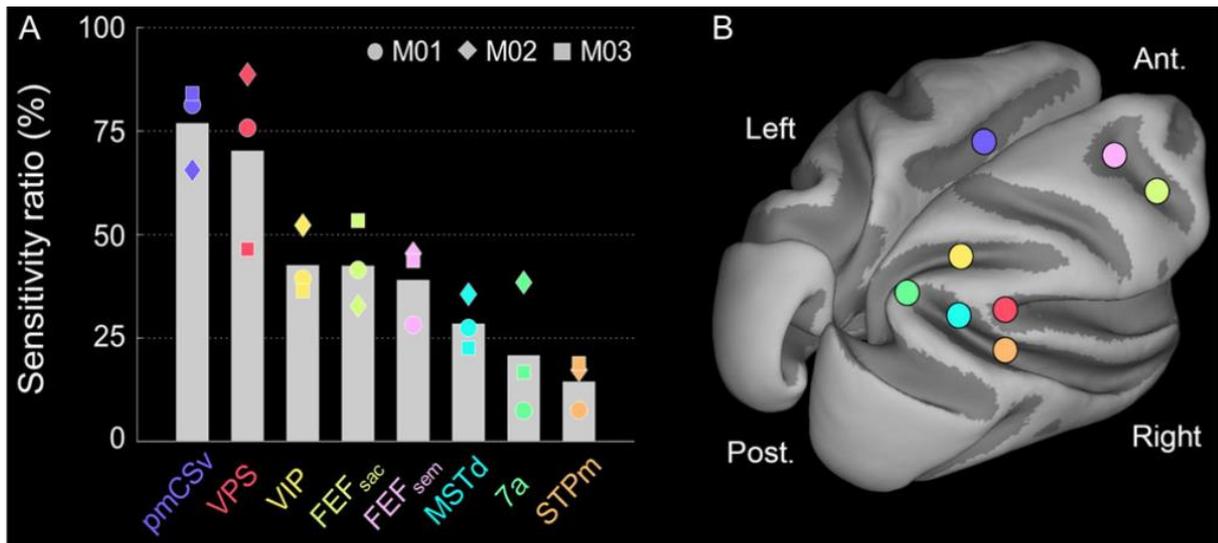


Figure 21: A) Average sensitivity ratio (%) between the responses to the EC and EI conditions. Only areas with significant responses in the 3 animals are shown. Areas were sorted according to their mean sensitivity ratio. Markers provide the individual data corresponding to M01 (circles), M02 (diamonds), and M03 (squares). B) Schematic localization of the 8 areas on the F99 template.

A second study mainly conducted by Samy Rima, the first student who worked on the development of the monkey MRI setup at CerCo, has provided the team with new tools and analyses. This study, currently under review, investigated the retinotopic properties of visual areas in the posterior parietal cortex of two awake, behaving macaques. A wide-field retinotopic presentation of stimuli showing images of a fruit basket moving in different directions was chosen so as to enhance responsiveness of these high visual areas. The population receptive field method (Dumoulin and Wandell, 2008) was applied to the data as it has the main advantage to enable the analysis of areas with large receptive fields. Several features could be determined: the eccentricity, the polar angle, the size for each voxel (population), and the laterality (hemifield differences) of the receptive fields. This approach confirmed the presence of the two retinotopic areas CIP1 and CIP2 that were recently found (Arcaro et al., 2011) and revealed that two other areas, that we labelled PIP1 and PIP2, also have a retinotopic organisation, forming a cluster with the previous CIP1 and CIP2 areas.

Figure 22 shows the results of those analyses: the different parameters of population receptive fields and the localisation of the PIP cluster.

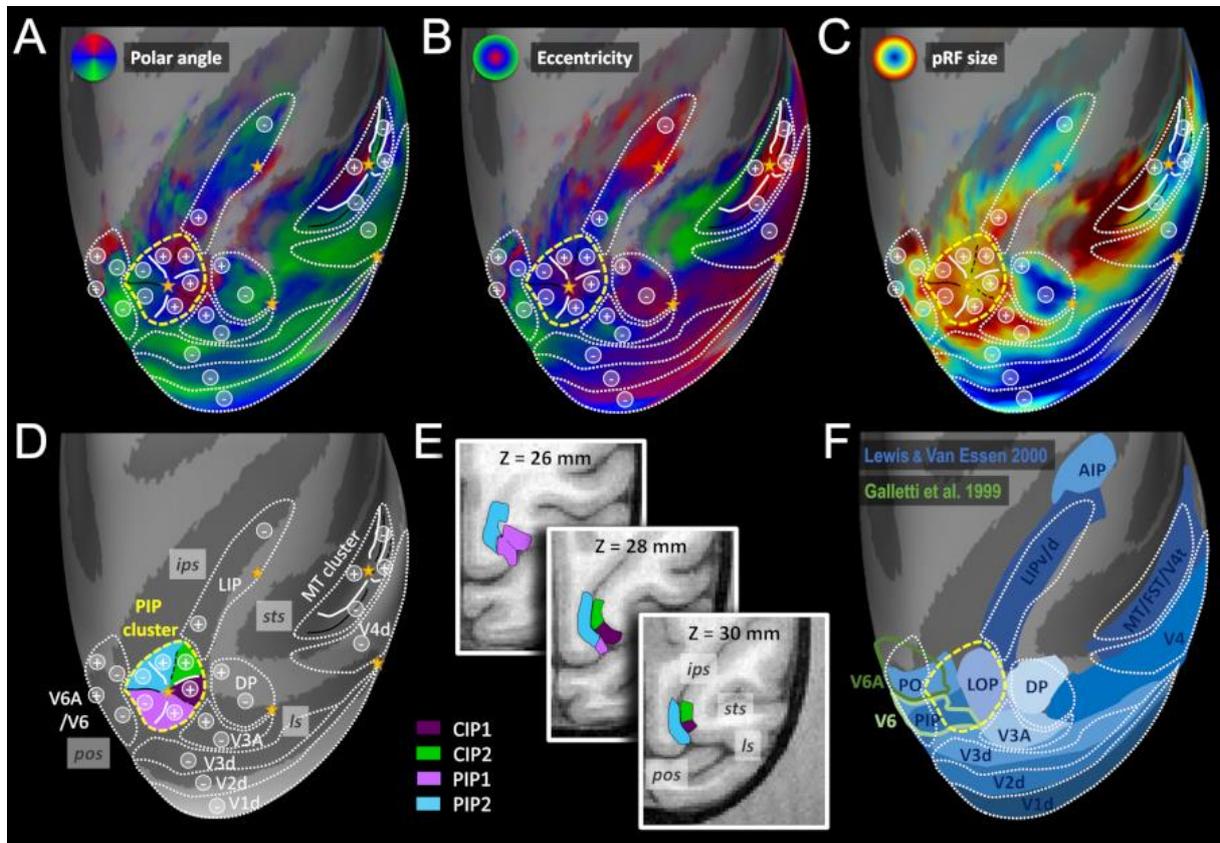


Figure 22: Mean PRF results (4 hemispheres) projected on the right inflated cortical surface of the F99 monkey template.

(A) Mean polar angle map. (B) Mean eccentricity map. (C) Mean population receptive fields' sizes. (D) Outline of the visual areas and visual clusters. (E) Horizontal sections showing the anatomical locations of the PIP cluster's areas. (F) Comparison of our visuotopic parcellation with the F99 locations of various dorsal visual areas (Lewis et Van Essen, 2000) and V6/V6A areas (Galletti et al. 1999). The “+” and “-” bullets indicate representations of the superior and inferior quarter-fields respectively.

These two studies in which I took part showed me how to conduct a study with macaques and taught me about the various analyses that can be done with neuroimaging data. Indeed, I was involved in the animal training and I performed additional analyses. I created retinotopic

masks, projecting the retinotopic data from the nodes of the brain surface back to the volume voxels, thus providing a useful tool to conduct region-of-interest analyses. I also re-estimated the HRF for each animal, extracting the signal from V1 nodes, which led to better estimates of the GLM parameters as explained earlier. This knowledge and know-how have been implemented and further developed in the studies that will be detailed in the next sections and that represent the core of my PhD work.

Chapter IV –
Temporal integration of binocular disparity:
The case of motion-in-depth

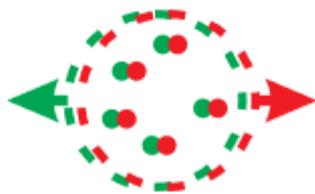
General introduction

Stereomotion or motion-in-depth (MID) is primordial for our interactions with moving objects or to avoid collisions with them, providing information on their trajectories. Yet, despite its ecological relevance, motion-in-depth remains little studied compared to lateral or bidimensional (2D) motion.

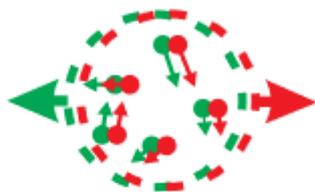
To compute motion-in-depth, the brain can process two types of binocular cues. The first one is changing disparity over time (CDOT). It could be defined as a temporal integration of binocular disparity, i.e. disparity is computed first and then its gradient over time is processed. The second one is about integrating the interocular velocity difference (IOVD). It implies to extract the monocular motion in each retina first and then to compute their difference between the two eyes. If the two mechanisms are very similar, they might nonetheless rely on different neural substrates. In natural situations, the two cues (CDOT and IOVD) are processed together and complement each other (Lages and Heron, 2010). However, it is possible to separate them using specific random dot stimuli to have an insight on the individual contribution of each cue (see Figure 23 for more details). Psychophysical studies rather abundantly investigating the role of each cue have shown mitigate results, some insisting that CDOT was the main signal used by our nervous system for MID perception (e.g., Cumming and Parker, 1994), other that IOVD was more relevant (e.g., Beverley and Regan, 1975; see Regan and Gray 2009 for a review). Regarding cortical activity however, there is only a limited number of studies addressing the question and this has been done from different perspectives. On the one hand, some studies compared cortical responses to planar motion to responses to motion-in-depth without further differentiating MID into its two main components. On the other hand, some papers have investigated neuronal activity differences for CDOT vs. IOVD cues specifically. Cynader and Reagan (1978, 1982) were the first to record neuronal activity in response to MID in the cat visual cortex. They were shortly followed by Poggio and Talbot (1981) who recorded

neurons from the foveal cortex of macaque with similar findings: There are different types of neurons involved in MID processing in the early visual cortex, such as neurons with selectivity to changing disparity, to velocity differences, or to motion direction. These experiments raised more questions: Are the two types of signal (i.e. disparity vs. motion signals, underlying CDOT and IOVD cues respectively) combined and if they are how is it done? Are they processed in parallel? What are the cortical networks involved? Which pathway is followed? Is the pathway for MID perception different from planar motion processing?

Combined condition



CD condition



IOVD condition



Monocular conditions

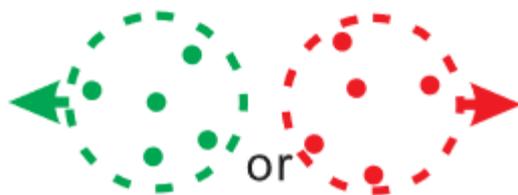


Figure 23: Assessing the contribution of the two binocular cues to motion-in-depth. In the combined condition, both IOVD and CDOT cues are mixed. In the CD condition, the use of dynamic random-dot-stereograms enables the changing of disparity over time (CDOT) whilst randomly relocating the dots allows the elimination of the IOVD cue. In the IOVD condition, uncorrelated random dots between the two eyes are used to disrupt the CDOT cue. Dots in each image are moving in a coherent manner and in opposite direction, thus giving rise to the IOVD cue. The lack of spatial correlation between dots in each image strongly weakens the CDOT cue without completely suppressing it. However, as emphasised by Sanada and DeAngelis (2014), this procedure has been shown to efficiently reduce the use of CDOT cue in those stimuli. Adapted from Sanada and DeAngelis (2014).

To provide more information on the matter, an fMRI-sourced EEG study compared the temporal dynamics of responses to lateral motion and to motion-in-depth in human and suggested that the information might not be processed by the same cortical pathways (Cottereau et al., 2014). Indeed, response dynamics for 2D vs. 3D motion were different, with slower responses to 3D motion, especially within the hMT+ cluster and the parietal cortex. These results are in line with the finding that the MT area in macaque receives motion and depth information from different pathways (Ponce et al., 2008).

Given that the hMT/MT area is a hub for motion processing¹⁵ and responds to binocular disparity (e.g., Maunsell and Van Essen, 1983; Cumming and DeAngelis, 2001), it has received further specific interest. In human, Rokers and colleagues (2009) demonstrated that the hMT+ cluster was strongly responding to both CDOT and IOVD cues, by investigating BOLD signal fluctuations in pre-determined regions of interest. This was somehow not in full agreement with another neuroimaging study (Likova and Tyler, 2007) in which the strongest activations in response to the CDOT cue were located in an area anterior to the hMT+ cluster. However, given the difference in stimuli and analyses, no strict comparisons between both studies can be made. Finally, Joo and colleagues (2016) recently suggested that, at the level of the hMT area, IOVD and CDOT cues are processed independently and might thus arrive through segregated pathways. In macaque it has been shown that there are neurons in MT tuned to MID and that this selectivity is independent of other known selectivity in the MT area (Czuba et al., 2014). However, this selectivity was mostly driven by the IOVD cue with only a small contribution of the CDOT cue (Sanada and DeAngelis, 2014).

¹⁵ Please refer to the introduction section of next part for further details about the motion selectivity in the MT/hMT+ cluster.

Altogether, by limiting themselves to the MT/hMT+ cluster and close-by areas, those studies lacked a bigger picture and, despite their interesting results, their contribution for revealing the cortical network processing motion-in-depth is restricted. In the next part, I detail a neuroimaging study in macaque that was built up on the two aforementioned human fMRI studies so as to overcome their methodological limitations and to shed new lights on how the CDOT cue is processed in macaque, and therefore potentially in human. This work represents the core of my doctorate work.

Currently in press, we collected our data at about the same time as a very recently published study conducted in human by the group of Alex Wade (Kaestner et al., 2019). Whereas we chose to focus on the cortical network involved in processing the CDOT cue only, they were more interested in the differential processing of both CDOT and IOVD cues. However, both studies investigated cortical activity at the whole-brain level and suggest that the parietal cortex might also be involved in processing motion-in-depth, opening the way for further investigations.

First study: Steremotion processing in the non-human primate (accepted article)

Introduction

Motion perception is a fundamental property of the visual system in most animal species. It enables to track over time the position of elements in a scene and thereby facilitates navigation or interactions with moving objects. Historically, numerous studies have characterised planar (i.e. 2D) motion processing in the primate nervous system (see e.g. Maunsell and Newsome, 1987). In macaque, single-cell recordings showed that it is computed at the cortical level within an extended network that begins in the primary visual cortex and includes higher-level visual areas such as area V4 in the ventral pathway (Li et al., 2013) and area V3A (Galetti et al., 1990) or V6 (Pitzalis et al., 2013) in the dorsal pathway. Within the Superior Temporal Sulcus (STS), area MT notably hosts neurons whose responses are highly selective to motion direction (see e.g. Maunsell & Newsome, 1987) and also reflects motion perception (Newsome & Paré, 1988; Britten et al., 1996). In human, neuroimaging studies suggested that 2D motion is also processed within an extended network that includes a putative homologue of area MT: hMT (Huk et al., 2002). Over the last 20 years, the emergence of monkey fMRI (see e.g. Vanduffel et al., 2001) made possible the further characterisation of the correspondence between the cortical areas involved in motion processing in the two species. This comparative approach revealed that 2D motion engaged largely similar networks in macaques and in humans. It notably suggested that motion processing in MT and its satellite areas (V4t, FST, and MSTv) on the one hand and in hMT and its satellites (pV4t, pFST, and pMSTv) on the other hand could be homologous (Kolster et al., 2009; 2010). Yet, functional differences between both species have also been documented, since sensitivity to motion in area V3A and other regions within the intraparietal sulcus (IPS) was found to be more pronounced in human than in macaque (see Orban et al., 2003 for a review). For instance, cortical areas responsive to motion-

defined structures are encountered in the human IPS but not in its monkey counterpart (Vanduffel et al., 2002).

Rather surprisingly, much less is known about the neural mechanisms that process motion along the depth dimension in primates. Motion in depth is nonetheless a very common and important component of motion in everyday life as it can notably signal objects moving towards the head and/or the body. Its estimation derives from two binocular cues: The change of disparity over time (CDOT), which tracks dynamic increase or decrease in horizontal disparity, and the interocular velocity difference (IOVD), based on opponent motion vectors between the two eyes (Nefs et al., 2010). The characterisation of the cortical areas processing each of these cues in both human and macaque would lead to a better understanding of how the 3D motion of objects is integrated in the primate brain and to further establish the similarities but also the differences between motion processing in the two species. Over the last ten years, this important line of research inspired a growing number of studies based on neuroimaging measurements in human and single-cell recordings in macaque.

In human, a pioneer fMRI study (Likova and Tyler, 2007) found that the strongest responses to CDOT arise in a cortical region anterior to the hMT+ complex. This region had reduced responses to 2D motion and was labelled ‘CSM’, for Cyclopean StereoMotion, by the authors. It might therefore be specialised in processing motion-in-depth. From additional analyses within independently defined regions of interest (ROIs), this study also found that several visual areas, such as V3A, V4, and hMT+, were more responsive to CDOT than to static disparity planes. Two years later, Rokers et al., (2009) found specific responses to CDOT and IOVD in the hMT+ complex, but also in area V3A and in lateral occipital regions (LO1/LO2). However, this study did not investigate activations in regions anterior to the hMT+ complex. More recently, Kaestner et al. (2019) confirmed that the hMT+ complex and two groups of ROIs that respectively gathered dorsal (V3A/B and IPS-0) and ventral (V4, LO-1 and LO-2)

areas were more responsive to CDOT than to a control condition where the temporal frames were scrambled. These authors also found significant responses to CDOT in the CSM area but that were not as pronounced as those measured in hMT+. Altogether, these three human studies found consistent patterns of cortical responses to CDOT even though the precise functional role of area CSM, and notably how its responses to both 2D and 3D motion differ from those measured in hMT+, remains to be better understood.

In macaque, explorations of motion in depth selectivity at the single-cell level began with electrophysiological recordings in area MT, notably on the grounds that this area has similar 2D motion responses in human and non-human primates. Two studies found that MT neurons were selective to motion in depth (Czuba et al., 2014; Sanada et DeAngelis, 2014) but that this selectivity was primarily driven by the IOVD cue, with only a small contribution from the CDOT cue (Sanada & DeAngelis, 2014). Based on the human findings, it is possible that stronger responses to CDOT could be found in other satellite regions of the MT cluster (e.g. in FST or MSTv) and/or in more anterior regions of the STS. It is also possible that significant responses to CDOT exist in other regions of the ventral and dorsal visual pathways. In order to clarify which regions of the macaque cortex should be explored to better understand the neural mechanisms underlying motion-in-depth processing in primate, it is essential to first determine whether the areas that were identified in human from neuroimaging measurements could have putative homologues in macaques by using a similar approach.

In the present study, we performed fMRI recordings in awake behaving macaques to identify the cortical regions sensitive to motion-in-depth defined by changing disparity over time (CDOT). To facilitate the comparison with previous human data, we used an experimental protocol that was directly inspired from the neuroimaging studies described above (Likova and Tyler; 2007; Rokers et al., 2009; Kaestner et al., 2019). In order to precisely determine the limits of the MT cluster, we ran an additional retinotopic mapping experiment that permitted

to delineate its four constituting areas V4t, MT, MSTv, and FST (Kolster et al., 2009). This approach allowed us to clarify whether the strongest responses to CDOT in macaque emerge beyond this cluster or within one or several of its components. It also permitted to document CDOT responses in retinotopic area of the early visual cortex (V1, V2, and V3) and along the dorsal (V3A) and ventral (V4) pathways. To determine whether regions activated by CDOT were also responsive to 2D motion, we ran a last experiment where we characterised the responses to moving versus static objects.

Materials and Methods

Subjects

Two female rhesus macaques (age: 15-17 years; weight: 5.35-6.15 kg) were involved in the study. Animal housing, handling, and all experimental protocols (surgery, behavioural training, and MRI recordings) followed the guidelines of the European Union legislation (2010/63/UE) and of the French Ministry of Agriculture (décret 2013–118). All projects were approved by a local ethics committee (CNREEA code: C2EA – 14) and received authorisation from the French Ministry of Research (MP/03/34/10/09). Details about the macaques' surgical preparation and behavioural training are provided elsewhere (Cottureau et al., 2017).

Data availability

Data and analysis code are available on a dedicated platform (OSF: <https://osf.io/yxrsv/>).

Experimental design

Our stimuli were directly derived from those of a previous fMRI study that investigated how cyclopean stereomotion (motion-in-depth based on CDOT) is processed in human (Likova & Tyler, 2007, see also Rokers et al. (2009) or Kaestner et al., 2019). Our aim was to facilitate

the comparison between the cortical networks involved in the two species. We used dynamic random-dot stereograms (dRDS) located within a disk (11 degrees of radius) and refreshed at 30Hz. The dot density was 15%. To manipulate binocular disparity between the two retinal projections, dots were green in one eye and red in the other and stimuli were observed through red-green anaglyphs (stimulus code made available on OSF: <https://osf.io/yxrsv/>). In the ‘cyclopean stereomotion’ (‘CSM’) condition, dots within the upper and lower parts of the disk changed their disparity in opposition of phase, following a triangular function (1Hz) between ± 23.2744 arcmin (see Figure 24-A). This disparity range was chosen so as to maximise the visual system responses to binocular disparities (see e.g. Backus et al., 2001 or Cottreau et al., 2011). The opposition of phase between stereomotion of the dots in the upper and lower parts of the disc led to an average disparity across the visual field of zero at each frame and thereby prevented stimulus-induced change in vergence eye movement. Note that in this condition, motion in depth is defined from the change of disparity over time (CDOT). We chose to use the term cyclopean stereomotion in reference to the original human fMRI study of Likova and Tyler (2007).

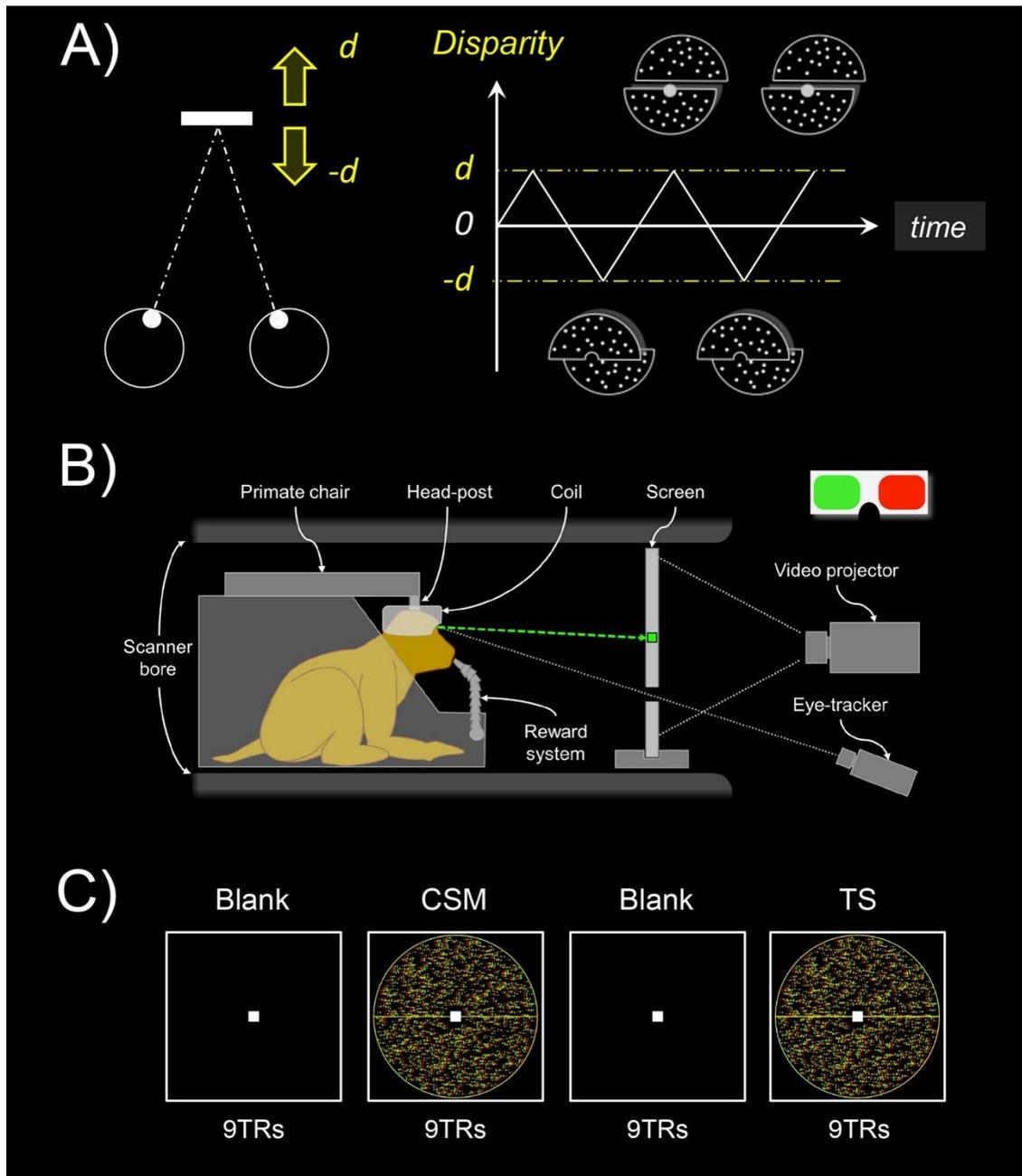


Figure 24. Stimulus design and experimental protocol.

A) In the main condition ('CSM' for cyclopean stereomotion), the motion occurs along the antero-posterior axis (leftward panel). The stimulus consisted of a disk (11° of radius) defined by dynamic random dot stereograms (dRDS) refreshed at 30Hz. Its upper and lower parts moved in depth between $d = \pm 23.3$ arcmin in opposition of phase, following a 1Hz triangular function (rightward panel). In our control condition ('TS' for temporally scrambled), the individual frames of the CSM condition were shuffled in time to disrupt the smooth change of disparity over time. Our two conditions had identical retinal disparity distributions but only the CSM condition conveyed motion-in-depth. B) Schematic representation of the monkey fMRI

setup. The animal was seated in a sphinx position within the primate chair, in the bore of the scanner, with the 8-channel, phased-array coil placed on top of the head. The animal was involved in a passive fixation task while its eye position was monitored by an infrared video-based eye-tracker. Horizontal disparity was introduced through the use of red/green anaglyphs. C) Illustration of the experimental protocol. Recordings were performed using a blocked design, with the alternation of CSM and TS stimuli separated by blank periods. Each run contained 3 repetitions of such blocks plus an additional baseline period of 9TRs (117TRs in total). CSM conditions were shown first in half of the runs and TS conditions appeared first in the other half of the runs.

Our CSM stimulus led to a percept of two planes continuously moving alongside a horizontal axis in opposite directions, one plane being perceived in front of the point of fixation and the other behind the fixation point. The control stimulus consisted of a temporally scrambled version ('TS') of the CSM stimulus. To create this 'temporally scrambled' condition, frames from the 'CSM' stimulus were randomly drawn without replacement. This configuration disrupted the temporal sequence of the stereomotion and thus, the motion in depth. Importantly, both conditions were monocularly identical and contained the same disparity distributions. The average relative disparities between dots in the upper versus lower parts of the disc were also identical between our two conditions.

MRI recordings

Image acquisition: Templates of reference and functional sessions

Whole-brain images were acquired on a 3 Tesla MR scanner (Phillips Achieva) using a custom 8-channel phased-array coil (RapidBiomed) specifically designed to fit the skull of our macaques while preserving their field of view. Four T1-weighted anatomical volumes were acquired prior to the study for each monkey at a high resolution (MPRAGE; repetition time,

TR = 10.3 ms; echo time, TE = 4.6 ms, flip angle = 8°; FOV: 155x155 mm; matrix size: 312x192 mm; voxel size = 0.5 x 0.5 x 0.5mm; 192 sagittal slices acquired in an interleaved order), as well as 300 functional volumes (gradient-echo EPI; TR = 2,000 ms, TE = 30 ms, flip angle = 75°, SENSE factor = 1.6; FOV: 100x100 mm; matrix size: 68x64 mm; voxel size = 1.25 × 1.25 × 1.5mm, 32 axial slices acquired in an interleaved order with a thickness of 1.5 mm and no gap). Those data were recorded in a single session whilst the macaques were slightly anaesthetised (Zoletil 100:10 mg/kg and Domitor: 0.04mg/kg) and their constants monitored with an MR compatible oximeter. Those volumes were then used to create individual anatomical and functional templates of reference.

Our T2*-weighted functional images were acquired with a gradient-echo EPI sequence with interleaved slice acquisition (TR = 2,000 ms, TE = 30 ms, flip angle = 75°, SENSE factor = 1.6; FOV: 100x100 mm; matrix size: 68x64 mm; voxel size = 1.25 × 1.25 × 1.5 mm, 32 axial slices acquired in an interleaved order with a thickness of 1.5 mm and no gap).

Scanning Procedure

During the functional recording sessions, macaques were head-fixed and seated in a sphinx position in a dedicated primate chair (see Figure 24-B). They had to maintain their gaze within a central fixation window (2x2°) during daily sessions of up to 2 hours. Their fixation was monitored with an ASL© infrared video-based eye tracking setup at 60Hz and they were rewarded through a liquid delivery system at intervals whose frequency depended on their fixation performance. Our cyclopean stereomotion stimuli were video-projected using a 23° x 23° field of view (viewing distance = 85 cm). We used a blocked design based on cycles within which our two conditions ('CSM' and 'TS') were interleaved with baseline periods of fixation (see Figure 24-C). Both the condition and baseline blocks lasted 18 seconds (9 TRs) and a cycle was therefore 72-second long (36 TRs). Each run contained 3 repetitions of this cycle plus an

extra baseline that was added at the end for a total duration of 117 TR (234 seconds). We displayed the stimuli, controlled for the delivery of the liquid reward and the fixation performance using the EventIDE software (Okazolab®).

Data analysis

Templates of reference

Anatomical and functional templates of reference were created for each individual with the volumes acquired prior to the current study. The anatomical template was obtained with the four T1-weighted anatomical volumes being realigned, averaged, and then co-registered on the MNI space of the 112RM-SL template (McLaren et al., 2009, 2010). To create the functional template, the 300 functional volumes (GE-EPI) were realigned, averaged, and then co-registered on the anatomical template. Both the T1 and the EPI mean images were segmented separately in order to obtain tissue probability maps for the grey matter, the white matter, and the cerebrospinal fluid (CSF). These probability maps were used to estimate the normalisation parameters from functional (mean EPI) to structural (mean T1) images for each individual.

Pre-processing of the raw functional data

Pre-processing and volume-based analyses were carried with SPM12 in the Matlab environment (MathWorks®). Only runs with central gaze fixation above 85% were kept for further analysis. In total, we kept 43 and 60 runs for both macaques, respectively. The 4 first volumes of each run were discarded (dummy scans) to account for the establishment duration of the BOLD steady-state response. Pre-processing was performed for each volume, run by run. Slice-timing correction was performed first using as a reference the slice acquired in the middle of the acquisition of each TR. Images were then reoriented, co-registered with the EPI template, and transformed to fit the individual T1 template (with the normalisation parameters estimated between the mean EPI and T1 images; see the '*Templates of reference*' section). No

motion correction was applied to the images. Finally, the images were smoothed with a spatial Gaussian kernel (FWHM = 2x2x2 mm).

HRF estimation

Prior to our statistical analyses, we used independent datasets to characterise the BOLD haemodynamic impulse response functions (HRF) separately for each animal. These datasets respectively contained 16 (M01) and 12 (M02) 204s long runs that consisted of 6 cycles of 4s full field counter phasing (10Hz) checkerboards separated by a 30s blank interval (see more details about this procedure in Cottureau et al., 2017). Data were pre-processed using the pipeline described above and projected onto individual surfaces generated with the CARET software (Van Essen et al., 2001). Following Dumoulin and Wandell's procedure (2008), we extracted the BOLD responses from nodes within the anatomically defined V1 of each individual. We only kept visually responsive nodes, that is those whose signal-to-noise ratio (SNR) was greater than 3. This SNR was estimated with a Fourier analysis of the average time courses across runs where the signal corresponded to the Fourier coefficient amplitude at the stimulation frequency F (i.e. $F = 1/34$) and the noise was given by the average moduli at the two neighbouring frequencies (i.e. $F - \delta f$ and $F + \delta f$, where $\delta f = 1/2$ is the resolution of our frequency analysis). We computed the average time course of these nodes during one cycle and used this average time course for estimating the HRF. The HRF was derived as the response to a 2s stimulus (our fMRI sampling rate). Note however that our stimulus duration was 4s rather than 2s because linearity deteriorates at short durations (Boynton et al. 1996; Logothetis and Wandell, 2004) and also because this duration was used in a previous monkey fMRI study that characterised the BOLD HRF in macaque (Leite et al., 2002). For each monkey, the average response to the 4s stimulus was fit as the convolution of the responses to two 2s responses, each of which is the HRF. We parameterised the HRF as the difference of two gamma functions

(Friston et al., 1998). This functional form of the HRF captures the late undershoot of the response better than a single gamma function (Boynton et al., 1996).

General linear model (GLM)

Univariate statistics were performed at the voxel level in SPM12, using a general linear model (GLM). Our visual (CSM and TS) and baseline conditions were implemented as the three main regressors of the GLM. As reported above, we only analysed runs with fixation performance greater than 85%. We used the oculometric data of those runs to define regressors of non-interest that were included in the GLM to exclude the possible contribution of eye movements from our analyses. These regressors were obtained by automatically detecting the presence (1) or absence (0) of saccades in the different volumes of every run. The corresponding saccade regressors were then convolved with the HRF and introduced into the model. To characterise and eliminate noise in our recordings, we also performed a principal component analysis on voxels located outside the brain (see Farivar and Vanduffel, 2014). Indeed, time courses in those voxels should not be influenced by our experimental design but rather reflect artefacts caused by movements of the animal. For each run, we determined the number of principal components that were necessary to explain 80% of the variance in these voxels and used the corresponding principal vectors as regressors of non-interest in our model. This adaptive procedure typically added an average of 13.3 (± 9.3) and 11.3 (± 5.1) additional regressors in the models for Monkey 1 (MO1) and Monkey 2 (MO2), respectively.

We estimated the beta values associated with our GLM using the RobustWLS toolbox (Diedrichsen & Shadmehr, 2005), which is provided as an additional toolbox for SPM12 (<http://www.diedrichsenlab.org/imaging/robustWLS.html>). This approach allows estimating the noise variance for each image in the time series, using the derivative of a maximum likelihood algorithm. Variance parameters are then used to obtain a weighted least square

estimate of the regression parameters of the GLM. It therefore helps to reduce the impact of noisier volumes on beta estimation. Previous studies showed that such a method significantly improved estimations in blocked-design fMRI experiments (see e.g. Takeuchi et al., 2011). The beta weights obtained from the GLM were subsequently used to perform univariate analyses (t -scores) at the whole brain level. These analyses were performed on the pre-processed EPI data and the beta weights were then projected onto the high-resolution volumes of our two animals. They were also projected on the individual cortical surfaces and on the cortical surface of the F99 template using the Caret software (Van Essen et al., 2001).

Localisation of areas selective to motion-in-depth and description of their responses

In order to identify areas with specific responses to motion-in-depth, we examined the statistical parametric map corresponding to the contrast between our two visual conditions ('CSM' > 'TS') and thresholded this map at $p < 10^{-3}$ (uncorrected, t -value > 3.1). All the cortical regions that showed significantly stronger responses to CSM than to TS in both hemispheres and in each animal were considered. We controlled that these areas overlapped when projected on the right cortical surface of the F99 template. To further document the activations in those areas, we identified their local maxima and considered 3x3x3 voxel cubes around their coordinates. We then computed the percentage of signal change (PSC) corresponding to our main condition and its control using the following equations:

$$PSC_{CSM} = (\beta_{CSM} - \beta_{baseline}) / \beta_{constant} \times 100$$

$$PSC_{TS} = (\beta_{TS} - \beta_{baseline}) / \beta_{constant} \times 100$$

These values were extracted within small (3x3x3) voxel cubes rather than within patches determined by anatomical and/or statistical criteria, due to the fact that anatomical borders between areas are difficult to determine precisely and that our contrast led to extended activations that cannot be accurately divided into clusters corresponding to different functional

regions (see Figure 25Figure 26Figure 27). Our approach is more conservative and avoids subjectivity when dealing with borders between areas. Importantly, we reproduced our analyses with betas extracted from smaller (1x1x1) or larger (5x5x5) voxel cubes, and this did not impact our results. Note that here we just document activations around the local maxima of selective areas (notably the relative difference between activations in our main condition and in its control and also the variability across runs) but do not perform additional statistical analyses so as to avoid double dipping (Kriegeskorte et al., 2009).

Definition of retinotopic areas and characterisation of their responses to motion in depth

We also performed a wide-field retinotopic mapping to delineate retinotopic regions that were used for additional ROI-based analyses. Whole-brain images were acquired with an identical setup as for the main experiment. In this case, visual stimuli were displayed using a large field-of-view (80° of visual angle, viewing distance = 25cm) and consisted of videos of a fruit basket that was moving laterally, forward and backward in monocular viewing. Traditional (clockwise/counter clockwise) rotating wedges (radius: 40°, angular extent: 49°) and expanding/contracting rings (eccentricity linearly varying between 0° and 40°) were used as visual apertures through which the fruit basket was displayed. Each run lasted 230s and contained 5 cycles of 44s with the first 10 seconds of a run being discarded (dummy scans) for the signal to reach its baseline. A small green square (0.4° x 0.4°) at the centre of the screen was used to control for fixation during passive viewing. As in our main experiment, only runs with more than 85% of correct fixation (respectively 47 and 48 runs for M01 and M02) were kept for further analyses. A pre-processing pipeline similar to the one described above was performed on the selected runs except that no smoothing was applied to the volumes and a fixed number of components (18 components) was used when performing the PCA, later used as a regressor of non-interest in the GLM. We projected the volume data onto individual surfaces using the Caret software (Van Essen et al., 2001) and a custom reorientation algorithm.

A population receptive field (pRF) analysis (Dumoulin & Wandell, 2008) was performed using the Matlab analyzePRF toolbox developed by Kay et al., (2013). For each surface node, an exhaustive set of theoretical pRF parameters (polar angle, eccentricity and size) was used to generate time courses that were compared to the real recordings. pRF size and position parameters that best predicted the data were selected as the node pRF parameters. With this approach, we obtained polar angle and eccentricity maps from which we characterised retinotopic areas that were described in previous monkey fMRI studies: V1, V2, V3, V4, V3A, as well as the regions within the Superior Temporal Sulcus (STS) (V4t, MT, MSTv, and FST) that form the MT cluster as described by Kolster et al. (2009). Those 8 retinotopically-defined regions were then projected back to the volumetric space to perform a ROI-based analysis of our motion-in-depth data. This was done using the inverse of the transformation between the volumetric and surface spaces mentioned above.

To test whether these retinotopic areas had specific responses to motion in depth, we first estimated their average PSC during the CSM condition and its TS control. We subsequently computed the corresponding difference between PSCs:

$$\Delta_{PSC} = PSC_{CSM} - PSC_{TS},$$

Note that we chose here to use the difference of PSCs because the PSCs for the CSM and TS conditions are paired. In order to estimate whether our observed PSC differences in these retinotopic areas were not due to chance, we computed permutation tests. We randomly attributed a negative sign to our PSC values and computed the mean difference, repeating this procedure 10,000 times. We then calculated a *p*-value defined as the proportion of random differences that were superior to our observed difference.

2D motion localisers

To determine whether the regions that have specific responses to binocular 3D motion are also responsive to 2D motion, we performed a control experiment in which we contrasted responses to static images versus rich 2D motion stimuli. The scanning procedure was identical to the main experiment procedure. Motion localiser stimuli were based on the fruit basket video used for the retinotopic mapping experiment. For the static version, static images were randomly drawn from the video and refreshed at 1Hz. For the moving version, the video was normally played. Stimuli were displayed either centrally ($<3^\circ$ of eccentricity) or peripherally ($>3^\circ$ of eccentricity). As for the retinotopic experiment (see the previous section), these visual stimuli were displayed using a large field-of-view (80° of visual angle) at a viewing distance of 25cm. Each visual condition lasted 6 seconds and was interleaved with a 10-second baseline. The four visual conditions were presented in a pseudo-randomised order and were repeated 3 times within each run. Five extra baseline scans were added at the beginning of a trial for the signal to reach its baseline, thus resulting in a total duration of 202 seconds (101 TRs) for each run. In total 42 and 26 runs with fixation above 85% were kept for our analyses. Selected data was pre-processed as previously described, with an adaptive number of components that were necessary to explain 80% of the variance for each run, adding an average of 12.6 (± 10.1) and 11.9 (± 3.7) additional regressors in the model.

To estimate motion sensitivity in our regions of interest and in our retinotopic areas, we contrasted moving and static conditions, by combining central and peripheral presentations. We then performed a ROI-based analysis, looking at the BOLD activity within our independently defined regions.

Results

The aim of this study was to identify the cortical network that processes disparity-defined motion-in-depth (i.e. cyclopean stereomotion) in two awake, behaving macaques using functional magnetic resonance imaging. Our experimental design was directly derived from previous human fMRI studies (Likova & Tyler, 2007; Rokers et al., 2009; Kaestner et al., 2019) so as to determine the homologies but also the differences between the BOLD activations in the two species (Orban, 2002). Our cyclopean stereomotion ('CSM') condition and its temporal scramble ('TS') control were defined from dynamic random dots stereograms (dRDS). They had identical retinal disparity distributions but differ in their temporal sequences (see the materials and methods section). Only the CSM condition conveyed motion-in-depth. Figure 25 and Figure 26 show the statistical parametric maps obtained for the contrast between 'CSM' and 'TS' on the individual anatomical templates of each subject (M01 on Figure 25 and MO2 on Figure 26). These data are shown for different coronal slices.

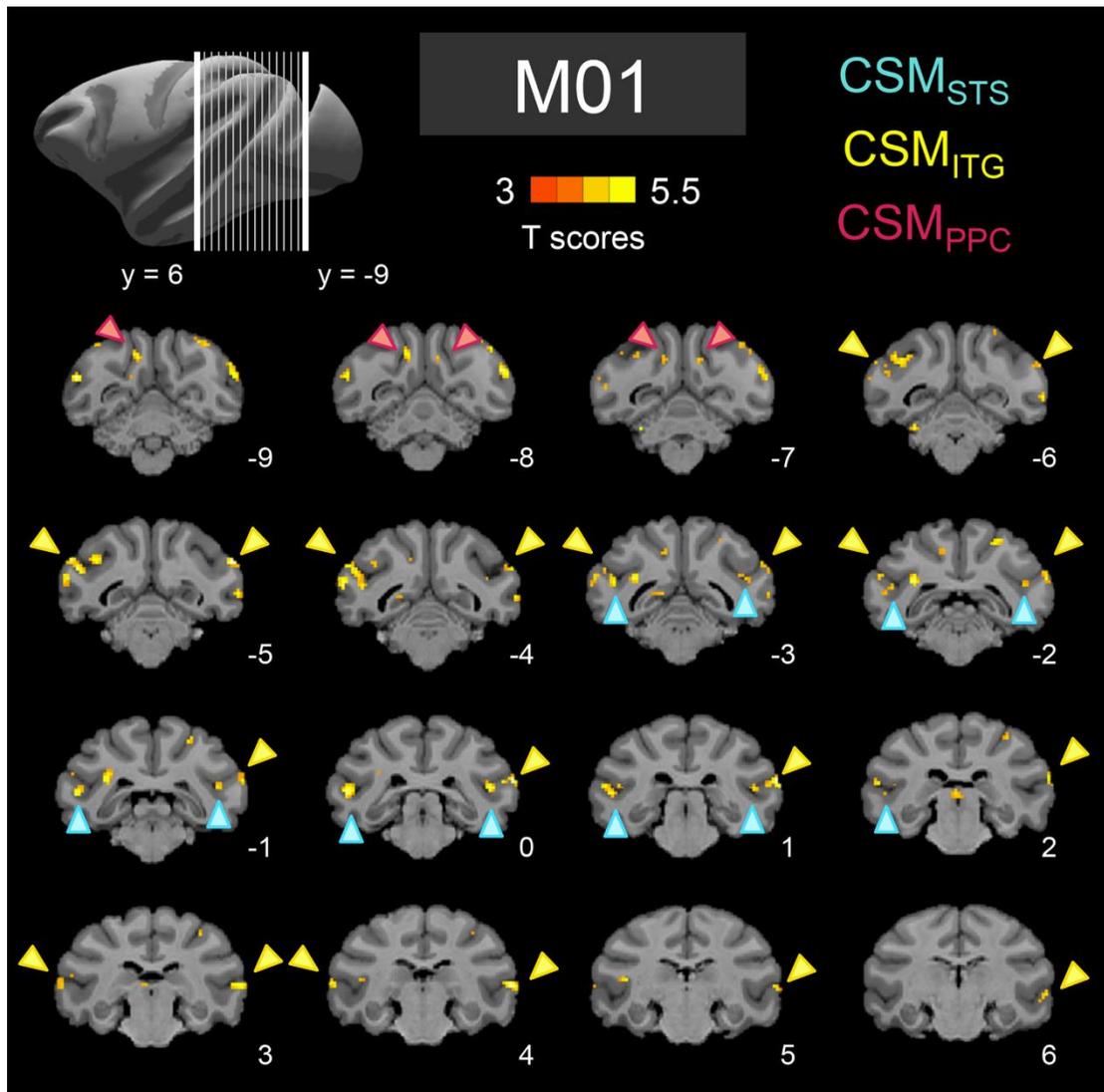


Figure 25. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) for M01. Figure shows activations that were stronger for the CSM condition than for the TS condition. Data are projected on the individual anatomical template of the macaque and are shown for different coronal slices. Coloured arrows indicate the localisation of our three regions of interest: CSM_{STG} (in blue), CSM_{ITG} (in yellow), and CSM_{PPC} (in pink). T-scores were obtained after computing the statistical parametric map for the contrast of interest between CSM and TS.

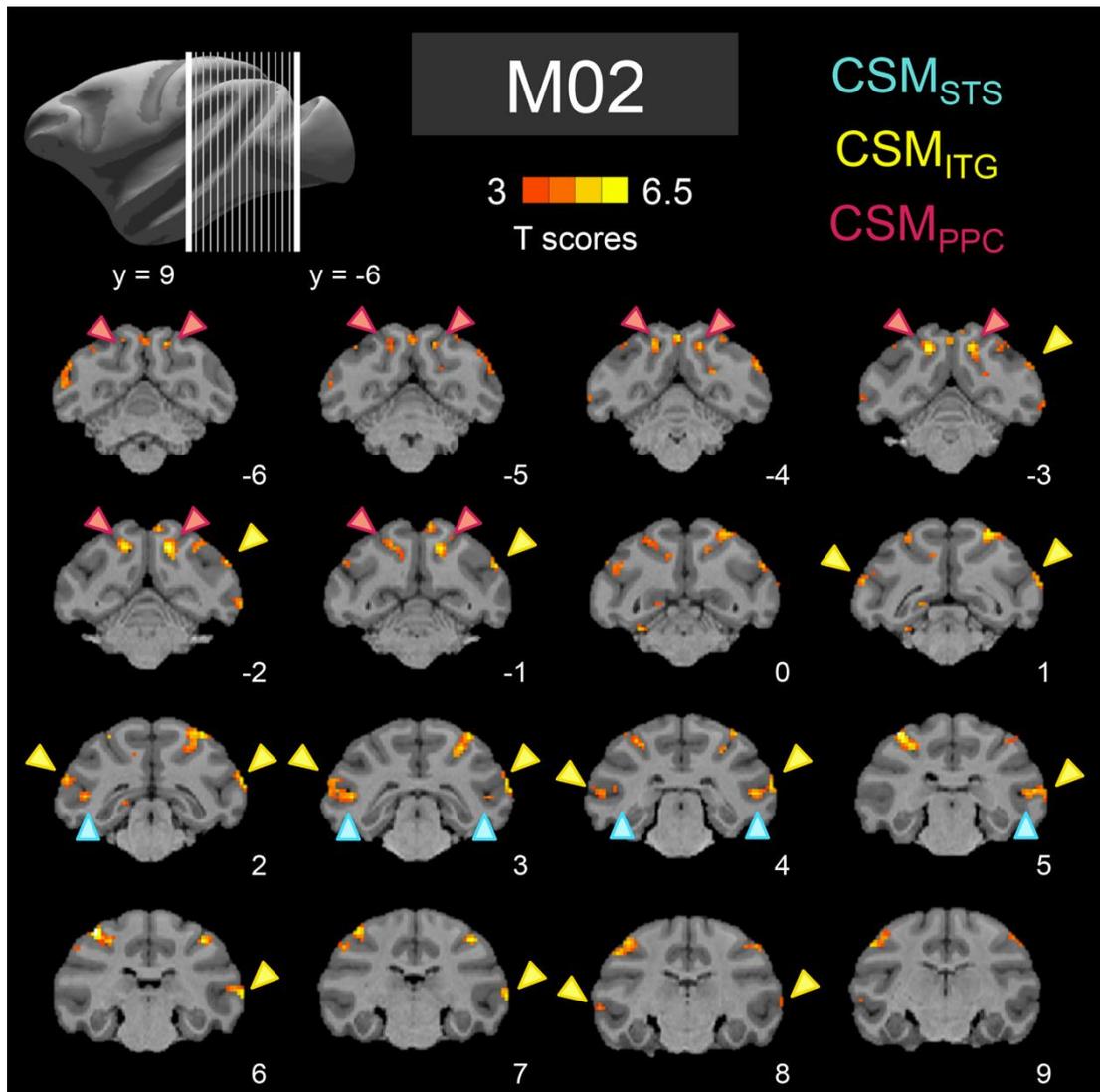


Figure 26. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) for M02. See Figure 25 for the details of the legend.

On Figure 25 and Figure 26, red-to-orange colours indicate significantly stronger BOLD activations for the CSM condition than for the TS condition ($p < 10^{-3}$, uncorrected). Despite differences in the activation patterns observed in the two animals, this analysis reveals a network encompassing the temporal and parietal cortices in both monkeys. Notably, three cortical areas are consistently found in both the left and the right hemispheres of our two macaques. Coloured arrows show these areas. For sake of comparison with previous human neuroimaging studies, we named those areas after Likova and Tyler's denomination (Likova

& Tyler, 2007), that is, CSM for Cyclopean StereoMotion responsive areas. The first area (CSM_{STS}) is located on the posterior bank of the superior temporal sulcus (STS) and extends posteriorly on the infero-temporal gyrus. The second one (CSM_{ITG}) is located on the infero-temporal gyrus, at the intersection between the lunate sulcus, the inferior occipital sulcus (IOS) and the STS. The last area (CSM_{PPC}) is localised in the posterior parietal cortex (PPC), mostly on the medial bank of the intra-parietal sulcus (even though activations can also be observed on its lateral bank in M02). This area might therefore correspond to the posterior intra-parietal (PIP) area. To be sure that the anatomical localisations of these 3 areas are not affected by our projections on the individual anatomical (T1) images, we confirmed their position on the functional (EPI) images in both monkeys (see Figure 32). The MNI coordinates corresponding to the local maxima of these areas in the two animals are provided in Table 2. To demonstrate the consistency of these results across hemispheres, we show on Figure 27 the projections of these activations on the individual cortical surfaces (see panel A).

Table 2. MNI coordinates (in mm) of the local maxima for the 3 regions that were significantly more responsive for the CSM condition than for the TS control in the two hemispheres of the two animals.

ROI	M01			M02		
	x	y	z	x	y	z
CSM_{STS}						
L	-24	0	16	-19	3	16
R	21	0	17	20	4	16
CSM_{ITG}						
L	-29	4	15	-23	3	18
R	28	1	20	26	3	19
CSM_{PPC}						
L	-5	-8	28	-6	-3	30
R	5	-7	27	6	-2	30

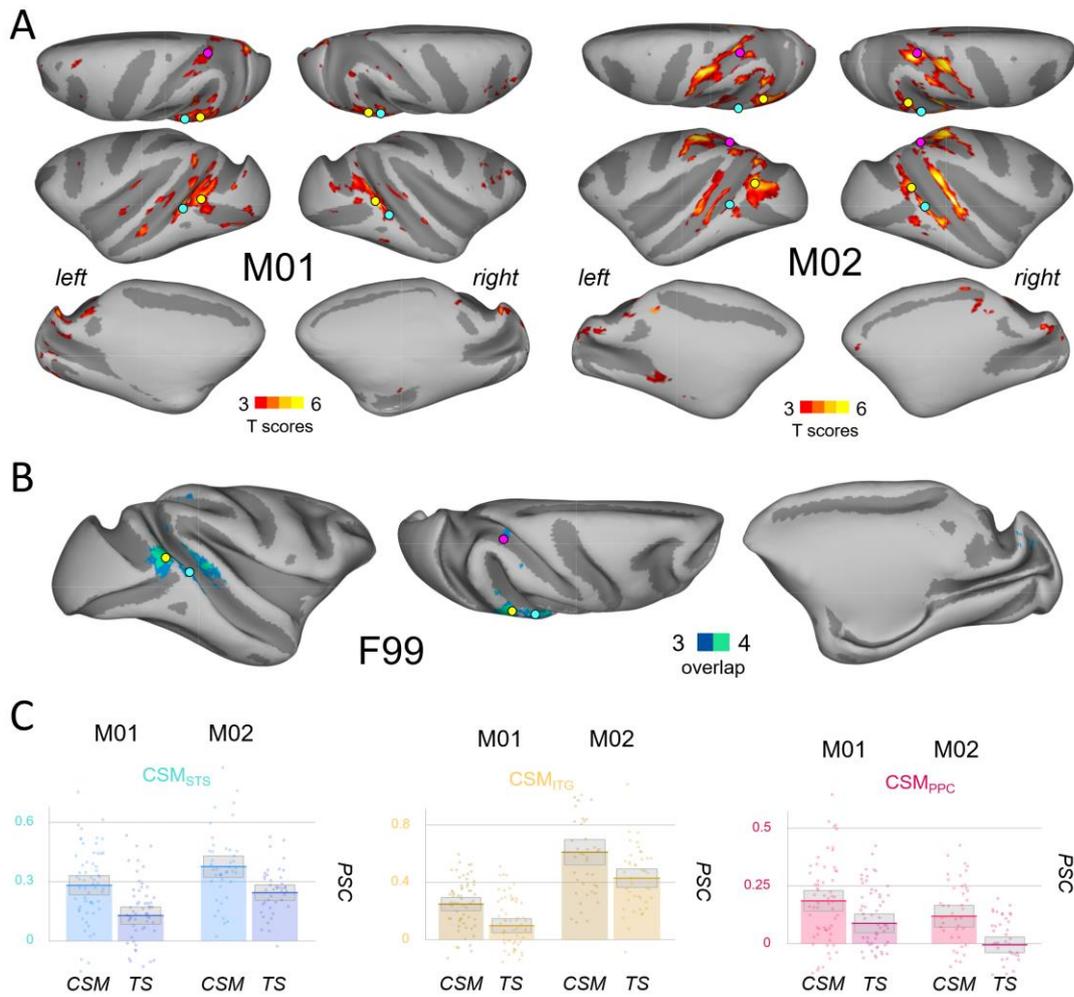


Figure 27. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) projected onto individual cortical surfaces and on the F99 template. A) Activations that were stronger for the CSM condition than for the TS condition. Data were thresholded at $p < 10^{-3}$ (uncorrected) and projected on the individual cortical surfaces of each animal (dorsal, frontal, and medial views). Coloured dots indicate the localisation of our three regions of interest: CSM_{STs} (in blue), CSM_{ITG} (in yellow), and CSM_{PPC} (in pink). B) Degree of overlap between the activations found in the two hemispheres of the two animals. The 4 individual cortical surfaces were morphed onto the right cortical surface of the F99 macaque template for projection of all the thresholded maps (frontal, dorsal, and medial views). Blue colour indicates the overlap of 3/4 hemispheres and green colour of 4/4 hemispheres. C) Percentages of signal change (PSC) for the 2 visual conditions (CSM and TS) with respect to baseline (fixation on a black screen) in our three regions of interest. The boxes give the 95% confidence intervals for the average values. The dots provide the data for each run. A small jitter was introduced to facilitate visibility.

As can be observed, our three regions of interest are found in all the individual surfaces, even though CSM_{PPC} is less visible in the right hemisphere of M01. This was confirmed by our projections of these activations on the right hemisphere of the F99 template. Figure 27 shows that our three regions overlap in at least 3 hemispheres for CSM_{PPC} and in 4 hemispheres for CSM_{STS} and CSM_{ITG}. The bar graphs on Figure 27-C provide the activations elicited by our CSM condition and its TS control relative to baseline (blank screen) around those local maxima (see the material and methods). The thick lines provide the average values and the boxes give the corresponding 95% confidence intervals.

It is worth noting that in monkey M02, significant BOLD activations were also found in more anterior parts of the IPS ($p < 10^{-3}$, uncorrected, see Figure 25, Figure 26, and Figure 27), notably within the ventral and anterior intraparietal areas (VIP and AIP, respectively). VIP has been shown to be involved in egomotion-compatible optic flow processing in both monkey (Cottereau et al., 2017) and human (Wall & Smith, 2008), whereas AIP has been suggested to play a role in 3D object processing and visually guided hand movements in both species as well (Sakata et al., 1997; Durand et al., 2007, 2009; Shikata et al., 2007). Unfortunately, we were not able to find those activations in the other macaque, potentially because of a slightly smaller SNR. In M02, activations were also found on the anterior part of the STS but they reflect responses from the fundus of the STS and/or from its posterior bank that were smoothed by our pre-processing pipeline and/or our transformations from the volume to the individual surfaces. The activations observed in the anterior part of the STS actually belong to clusters centred on the posterior bank.

Finally, it is important to emphasise here that we did not observe significant CSM responses on the cortex medial faces. Neither the anterior bank of the parieto-occipital sulcus, where the motion sensitive area V6 is located (see e.g., Pitzalis et al., 2013), nor the posterior part of the cingulate sulcus, where our group previously identified an area (pmCSV) responsive to

egomotion-compatible optic flow (Cottureau et al., 2017), seem to have strong sensitivity to motion-in-depth.

Retinotopic analysis

Previous studies in human found that the hMT+ complex had significant responses to stereomotion, notably based on changing disparity over time (CDOT) (Rokers et al., 2009; Joo et al., 2016). A single-cell study in macaque also found a weak but significant selectivity to CDOT in area MT (Sanada & DeAngelis, 2014). In order to determine whether the CSM-responsive ROIs we obtained from our univariate analyses overlap with (or correspond to) area MT and/or its neighbour regions, we performed a retinotopic mapping in our two animals (see more details in the Materials and methods section). This method notably allowed us to delineate the areas of the MT cluster: V4t, MT, MSTv and FST (see Kolster et al., 2009), which is not possible with more classical localisers of the MT / hMT+ complex based on thresholded statistical maps (even though some human studies proposed solutions to separate hMT from hMST, see Huk et al., 2002). In Figure 28-A and Figure 29-A, we show the locations of these areas and of our two CSM-responsive regions around the STS and the ITG, CSM_{STS} and CSM_{ITG}, respectively. Figure 28-B and Figure 29-B present the activations for the contrast between cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) projected on the same views.

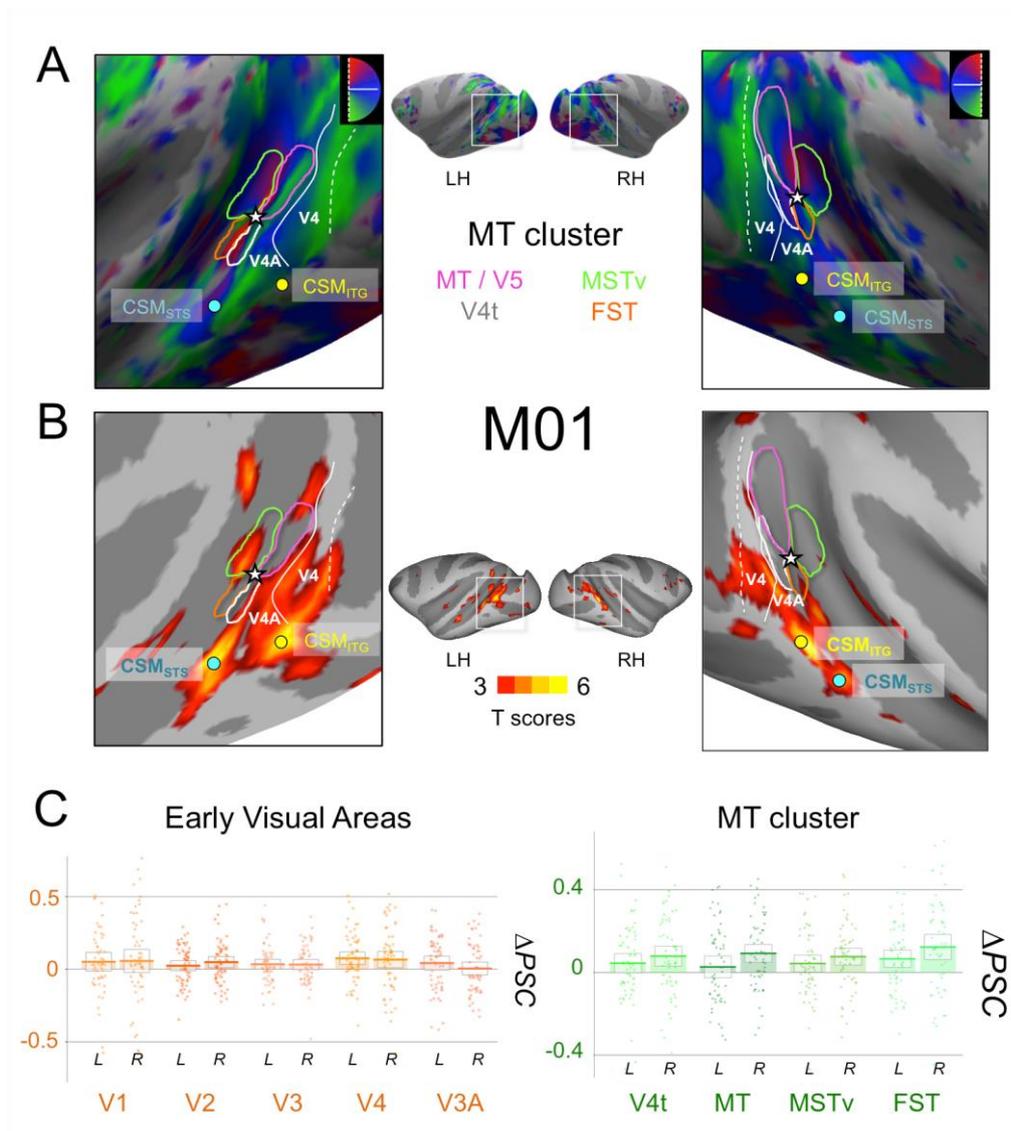


Figure 28. A) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M01 and delimitation of the MT cluster areas: MT (dark blue), V4t (pink), MSTv (orange), and FST (green). We show here the polar angle maps that were used to delineate the borders between these areas. The extent of those areas was obtained from the eccentricity maps. We also show the representations of the vertical and horizontal meridians that delineate the borders between V3 and V4 and between V4 and V4A. Coloured dots indicate the local maxima positions for areas CSM_{STs} (in blue) and CSM_{ITG} (in yellow). As shown on the maps, CSM_{STs} and CSM_{ITG} are in the vicinity of the MT cluster, but clearly exterior to it. B) Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) projected on the individual surfaces of M01, for both left and right hemispheres. C) Difference in percent signal change (Δ PSC) between the CSM and TS conditions in retinotopic areas. On the left, thick lines of the bar graphs provide average values for the left and right hemispheres of early

visual areas: V1, V2, V3, V4, and V3A. On the right, average values are given for the MT cluster areas: V4t, MT, MSTv, and FST. The boxes give the 95% confidence intervals for the average values. The dots provide the data for each run. A small jitter was introduced to facilitate visibility.

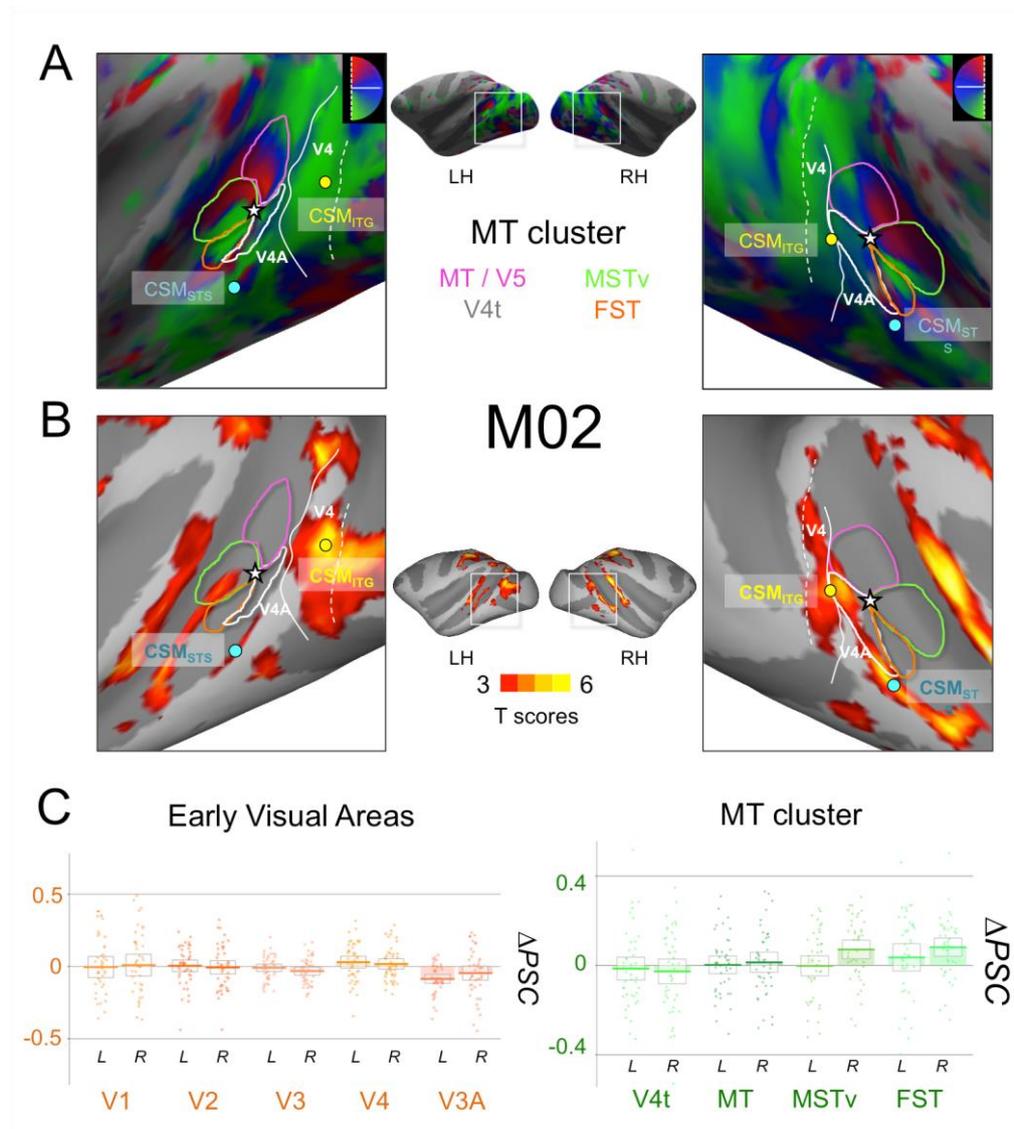


Figure 29. A) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M02 and delineation of the MT cluster areas: MT (dark blue), V4t (pink), MSTv (orange), and FST (green). B) Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) projected on the individual surfaces of M02, for both left and right hemispheres. C) Difference in signal change (Δ PSC) between the CSM and TS conditions in retinotopic areas. Conventions are similar to Figure 28.

We can see that if our two regions are close to the MT cluster, they nonetheless do not overlap with it. Area CSM_{STS} is located more anteriorly along the posterior bank of the STS. Area CSM_{ITG} is located more posteriorly on the ITG, in a position that might correspond to areas V4 and/or V4A (see discussion).

To complete our study and to allow a direct comparison with previous human data (Likova and Tyler, 2007; Rokers et al., 2009; Kaestner et al., 2019), we performed ROI-based analyses within retinotopically-defined areas constituting the MT cluster as well as the early-to-intermediate visual cortex: V1, V2, V3, V4, and V3A. The differences between the percentages of signal change (PSC) for CSM versus TS in our two macaques are shown on Figure 28-B and Figure 29-B for early visual areas (V1, V2, V3, and V4) and the MT cluster (V4t, MT, MSTv, and FST). We can observe that if CSM selectivity in all these areas is not as pronounced as in CSM_{STS}, CSM_{ITG}, and CSM_{PPC}, responses in the MT cluster tend to be stronger than those measured in early visual cortex. Permutation tests demonstrated significant CSM effects in areas MT and V4t (1/4 hemispheres), MST (2/4 hemispheres, right hemispheres only), and FST (3/4 hemispheres). It suggests that selectivity to cyclopean stereomotion exists in these regions. We also found that responses were significantly stronger for motion in depth in area V4 for one animal (2 hemispheres) but not for the other. Responses to CSM were not significantly stronger in V3A.

2D motion analysis

To test whether our three regions are only responsive to motion-in-depth or whether they are activated by motion in general, and notably by 2D motion, we ran an additional motion localiser in our two animals (n = 42 and n = 26 runs in M01 and M02, see more details in the Materials and methods section). We then computed the difference between the percentages of signal change (Δ PSC) corresponding to the 2D motion versus static image conditions. As expected

from such a localiser, this analysis led to significantly stronger responses to motion in most of the retinotopic areas and more specifically within areas of the MT cluster. In particular, permutation tests demonstrated that all 4 regions of the MT cluster had significantly stronger responses to 2D motion in the two animals ($p < 0.05$ except for the left V4t in M02). We show in Figure 30 the results of these analyses in our 3 CSM responsive areas (CSM_{STS}, CSM_{ITG}, and CSM_{PPC}).

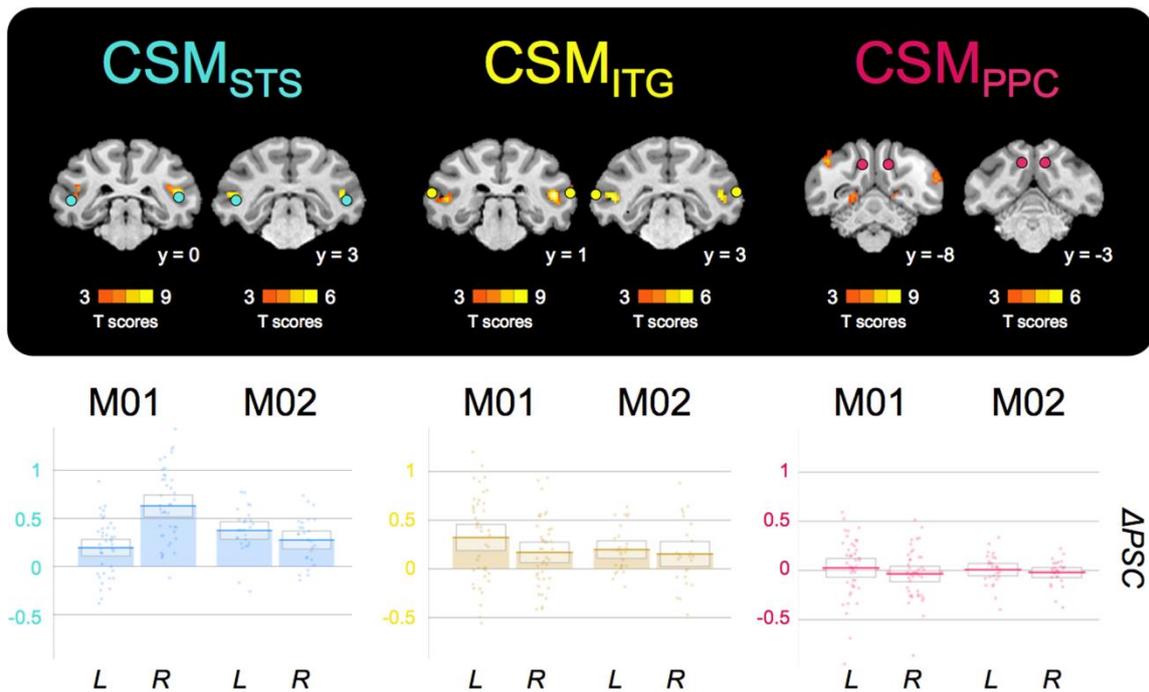


Figure 30. Sensitivity to 2D motion in CSM_{STS}, CSM_{ITG}, and CSM_{PPC}. Strongest responses to 2D motion are shown on coronal slices from the individual anatomical template of each animal (upper panel). The colour dots provide the position of CSM_{STS}, CSM_{ITG}, and CSM_{PPC}. For these 3 regions, PSC difference between responses to 2D motion vs static image are shown on the graphs of the lower panel. The thick lines of the bar graphs provide the average values across runs for the left (L) and right (R) hemispheres of the two monkeys (M01 and M02). The boxes give the 95% confidence intervals for these average values. The dots provide the data for each run. A small jitter was introduced to facilitate visibility.

We can observe that only CSM_{STS} and CSM_{ITG} areas have significant responses to 2D motion (permutation tests, $p < 0.05$), in both hemispheres for CSM_{STS} and in the left hemisphere for CSM_{ITG}, for each monkey. Their motion selectivity (in particular in CSM_{STS}) is therefore not specific to cyclopean stereomotion. On the opposite, responses to 2D motion in area CSM_{PPC} are not different from responses to static patterns (permutation tests, $p > 0.1$). It implies that this region might uniquely respond to cyclopean stereomotion.

Selectivity to 3D versus 2D motion within the lower bank of the STS

To further characterise the selectivity to 3D and 2D motion within the STS, we defined a path running along the posterior bank of this sulcus on the cortical surfaces of each hemisphere of our two animals. Each path departs from MT area and ends in the CSM_{STS} area. For each voxel along this path, we computed the average t-score within its first order neighbourhood (i.e. within a 3x3x3 cube centred on this voxel) for both the stereomotion versus temporal scramble and monocular motion versus static image contrasts. As t-scores for the second contrast were usually higher, we normalised the values along each path by dividing them by the maximum t-scores along the path. This facilitates comparisons between the sensitivity profiles for 3D and 2D motion. As shown in Figure 31, these profiles suggest that selectivity to stereomotion varies within the STS, with lower t-score values within the MT cluster and higher values in the CSM_{STS} area. By contrast, both the area CSM and the MT cluster exhibit strong sensitivity to 2D motion. This observation was confirmed by computing the average percentages of signal changes (Δ PSC) for 2D and 3D motion in those two regions (N.B.: the value for the MT cluster is the average Δ PSC across the path). The corresponding bar graphs are shown in Figure 31-B.).

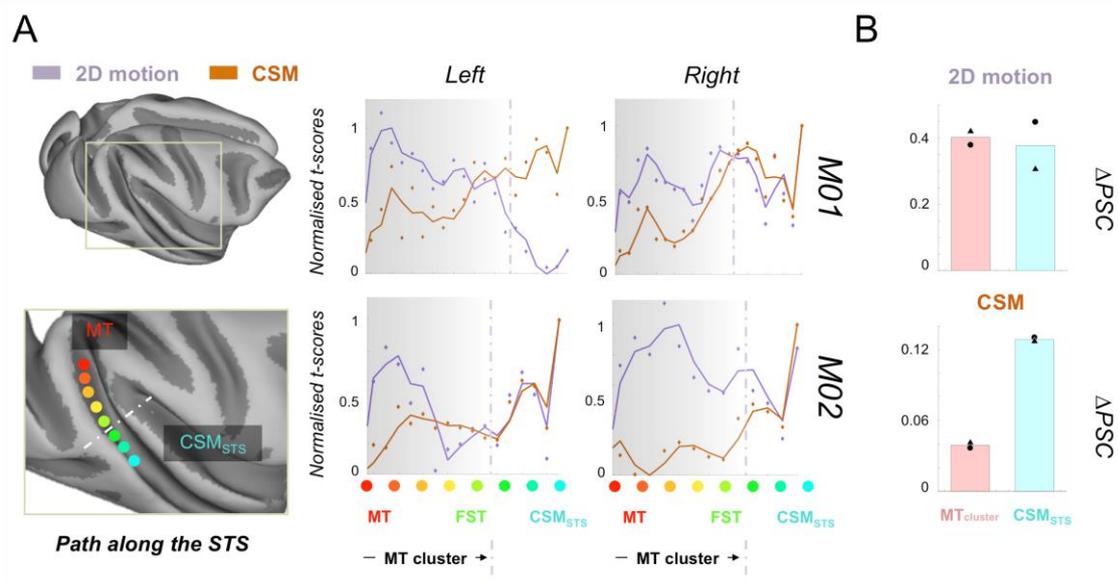


Figure 31. Selectivity to 3D and 2D motion within the STS. A) The left panel is a schematised view of the path drawn along the lower bank of the STS, starting from MT area (red dot) and ending within the CSM_{STS} area (cyan dot) defined from the stereomotion versus temporal scramble contrast. The grey dotted line represents the end of the MT cluster. On the right panel, responses to the 3D motion (i.e. stereomotion versus its temporal scramble control) and 2D motion (i.e. 2D motion versus static images) contrasts are respectively shown in orange and purple for each hemisphere of both macaque subjects. Dots provide the normalised t-score values along the path whilst the curves were obtained from a median filtering of these values. The general trend is an increase of 3D motion selectivity along the STS, with the highest value outside the MT cluster. Selectivity to monocular motion peaks in the MT cluster and tends to decrease along the STS path. B) Average differences in signal change (ΔPSC) values in response to both types of tested motions (2D and CSM) in the four MT cluster areas (red bars) and in the CSM_{STS} area (cyan bars) for both subjects (triangle symbols for M01 and circle symbols for M02).

Discussion

The aim of the present study was to characterise the cortical areas that process cyclopean stereomotion in rhesus macaque. To that end, we adapted the experimental protocol of two previous human fMRI studies that found significant activations in the hMT+ cluster and/or in an area directly anterior to it (Rokers et al., 2009; Likova & Tyler, 2007). Our main condition ('CSM' for cyclopean stereomotion) and its control ('TS' for 'temporally scrambled') shared the same disparity distribution and were monocularly identical (Figure 24) but only the CSM condition conveyed stereomotion, since the temporal sequence was scrambled in the TS condition. We recorded whole-brain BOLD responses from two behaving macaques using a blocked design. Our analyses revealed a network of three areas whose responses to our CSM condition were consistently (i.e. across hemispheres and animals) stronger than those to our control condition (Figure 25 and Figure 26). In reference to the original study of Likova & Tyler (2007), we labelled those regions CSM_{STS}, CSM_{ITG}, and CSM_{PPC}. To complete these analyses, we also documented responses to our CSM condition in visual areas estimated from independent wide-field retinotopic mapping procedures (Figure 28 and Figure 29).

Our CSM_{STS} region is located in the fundus of the superior temporal sulcus (STS) and extends on the infero-temporal gyrus (see Figure 25, Figure 26, and Figure 27), at a location (see Table 2) anterior to the MT area and its satellites (V4t, MSTv, and FST, see Figure 28 and Figure 29). Based on our retinotopic maps, we confirmed that this region is independent from the MT cluster (only a marginal overlap with area FST was found in the right hemisphere of M02). Our additional motion localiser demonstrated that it is also responsive to 2D motion (Figure 30). Previous studies in macaque reported additional motion-sensitive regions in anterior portions of the STS. Using fMRI, Nelissen et al., (2006) notably documented an area in the lower superior temporal sulcus (LST) that responds to opponent motions and to actions. This area was 6-8mm anterior to FST and therefore does not fully coincide with our CSM_{STS} area.

Another monkey fMRI study found several regions of the macaque inferior temporal cortex that had specific responses to disparity-defined stimuli (Verhoef, Bohon, and Conway, 2015). Among them, region ‘Pd’ (for posterior disparity) was localised in the lower bank of the STS, at a position that matches very well with our CSMSTS area. CSMSTS might thus be a distinct motion and disparity-selective area of the STS that notably processes motion in depth. In human, two studies (Likova and Tyler, 2007; Kaestner et al., 2019) found specific responses to cyclopean stereomotion in a cortical region anterior to the hMT+ cluster: CSM (N.B.: Rokers et al., (2009) only performed ROI-based analyses and it is therefore not possible to know whether they also had significant responses to motion in depth in this region). Our CSMSTS area might therefore be its macaque homologue. It is nonetheless important to note here that in the human studies, the delineation of the hMT+ complex (and sometimes its hMT and hMST sub-regions) was based on a contrast between the responses to uniform versus random motion whereas in our case the MT cluster was obtained from retinotopic mapping. To further clarify the potential homology between human CSM and macaque CSMSTS, it would be interesting for future human studies to properly define area MT and its satellites using retinotopic mapping (see Kolster et al., 2010) in order to precisely determine the location of the CSM area with respect to those regions.

Our stereomotion contrast was based on a smooth variation in depth versus its temporally scrambled version (as in Kaestner et al., (2019) and in the ‘TS’ control of the Rokers’ study (2009), see their experiment 2). Although we used dynamic random dot stereograms, it is possible that this temporally scrambled control still evokes some apparent percept of motion. However, it lacked the smooth change of disparity of our main condition. In their experiments, Likova and Tyler (2007) used two planes that alternated between two different depths (thereby generating an apparent motion in depth) in their main condition whereas their corresponding control was a plane at a unique depth. We hypothesise that in both human and macaque, the

CSM_{STS} / CSM area might be activated by different types of motion in depth, and notably by our contrast and the one used by Likova and Tyler (2007). This hypothesis is further supported by a control performed by these authors (see their supplementary materials) that demonstrated that significant activations were also obtained in this area with stimuli smoothly varying in depth (i.e. where binocular disparity was changed according to a sine wave), being therefore closer to those used in our own study.

Although univariate statistics did not show significant responses in the MT area and its satellites, ROI-based analyses demonstrated that for some animal and/or hemispheres, responses in MT (1/4 hemisphere), MSTv (2/4 hemispheres), and FST (3/4 hemispheres) were significantly stronger for our stereomotion condition (Figure 28 and Figure 29). In a pioneer study, Maunsell and Van Essen (1983) concluded from single-cell recordings in area MT of anaesthetized macaques that neurons in this region had no selectivity to motion-in-depth (see also Felleman and Kaas, 1984). More recently, Sanada & DeAngelis (2014) found, using a different method, that MT does host neurons tuned to motion-in-depth (see also Czuba et al., 2014) but that these neurons were mostly driven by the inter-ocular velocity difference (IOVD) between the two eye with only a modest contribution of the change of disparity over time (CDOT): ~10% of their neurons had significant selectivity for CDOT versus ~57% for IOVD. These findings are in line with our study and suggest that if selectivity to stereomotion is observable in area MT, it remains moderate. To our knowledge, selectivity to motion-in-depth was not directly tested in areas MSTv and FST. Based on our data, we hypothesise that a larger proportion of neurons tuned to cyclopean stereomotion could be found there. Altogether, the responses we measured in the STS are consistent with a model where stereomotion would be progressively integrated along a posterior-to-anterior axis with moderate responses in MT, intermediate responses in areas MSTv and FST, and stronger responses in CSM_{STS}. This

hypothesis is supported by our analysis of the responses on a path defined along the STS (see Figure 31), which suggests that selectivity to stereomotion progresses beyond area MT.

Using a ROI-based analysis, all three previous human studies found significant responses to motion in depth in the hMT+ cluster. Likova and Tyler (2007) reported that selectivity in this cluster was weaker than in their CSM region, in agreement with what we found in macaque. In contrast, Kaestner et al. (2019) found that responses in hMT+ (in both hMT and hMST) were stronger than in CSM (see their figure 7). Given their use of a relatively small field of view (i.e. their stimuli had 5° of radius) contrasting with much larger stimuli in our experimental protocol (11° of radius) and in Likova and Tyler's experiment (i.e. their display was a square of 30 x 40°), one possibility would be that neurons in the CSM_{STS} and CSM regions prefer more eccentric (>5°) stimuli. We computed the average population receptive field (pRF) eccentricities and sizes in the MT cluster and in CSM_{STS} (see Figure 33) and showed that in CSM_{STS} these parameters are actually similar to those found in V4t and FST, thus discarding this hypothesis. Further studies, notably in human where retinotopic mapping could be used to better define the position of CSM with respect to area MT and its satellites, will be needed to clarify this point.

Our CSM_{ITG} region is located on the infero-temporal gyrus, at the intersection of the end of the lunate sulcus, the end of the inferior occipital sulcus (IOS), and of the STS (see Figure 25, Figure 26, and Figure 27). It is therefore posterior to the MT cluster (see Figure 28 and Figure 29). This location matches well with area V4A that was previously described using single-cell recordings (Pigarev et al.; 2002) and fMRI (Kolster et al., 2014). CSM_{ITG} also overlaps with area V4, as suggested by our retinotopic analyses for which responses to motion in depth were significantly stronger in this area for one animal (see also panel B of Figure 28 and Figure 29). Unfortunately, signal-to-noise ratios in our retinotopic data were not sufficient to properly map area V4A (notably its anterior border with area OTd) and further studies will be needed to

clarify its responses to Stereomotion relatively to those estimated in V4. Both V4 and V4A are activated by disparity (Watanabe et al., 2002; Verhoef et al., 2015) and motion (Li et al., 2013; Kolster et al., 2014), even though their motion selectivity is not as pronounced as in the MT cluster (Kolster et al., 2014). This in line with our finding that CSM_{ITG} has only moderate responses to 2D motion, notably when compared to motion responses in area CSM_{STS} (Figure 30). Interestingly, two human studies on motion in depth (Likova and Tyler, 2007; Kaestner et al., 2019) reported significant CSM responses in area V4. Kaestner et al. (2019) also found strong stereomotion responses in area LO-1 (and Rokers et al. (2009) in area LO which includes LO-1) which was proposed to be the human homologue of area V4A (Kolster et al., 2014). These results suggest a good correspondence between CSM responses in those regions of the human and macaque brains.

Area CSM_{PPC} is localised in the posterior parietal cortex (PPC), mostly on the medial bank of the intra-parietal sulcus (IPS). Responses in this region were not stronger for 2D motion than for static stimuli, in agreement with previous monkey fMRI studies (see e.g. Vanduffel et al., 2001). Because of its localisation, area CSM_{PPC} might correspond to the posterior intra-parietal area (PIP) (Colby et al., 1988; Markov et al., 2014). Even though further studies will be needed to clarify this point, it is tempting to hypothesise that there might be a functional dissociation for 3D processing between this area and its counterpart on the lateral bank of the IPS, the caudal intra-parietal area (CIP). Indeed, in a previous monkey fMRI study, Durand et al. (2007) revealed sensitivity to kinetic depth in area PIP and AIP (for which we also found activations in M02) but not in area CIP. Area PIP has also been shown to respond to 3D structure (see e.g. Alizadeh et al., 2018). It might therefore play a role in the detection of and interaction with moving objects whereas CIP could be involved in processing 3D orientation and/or 3D features/arrangement of elements (Tsutsui et al, 2002; Durand et al., 2007; Rosenberg et al., 2013). In human, the studies of Likova & Tyler (2007) and of Rokers et al. (2009) did not

explore stereomotion selectivity in the parietal cortex (the latter nonetheless reported significant responses to CDOT in dorsal area V3A). The only study that reported results at the whole-brain level (Kaestner et al., 2019) found strong stereomotion responses in area IPS-0, which is located in the caudal part of the human IPS and therefore constitutes a potential homologue of our CSM_{PPC} region. Further studies will be necessary to clarify this point.

Finally, we did not find CSM specific activations in area V3A although this area is motion selective in macaque (Galetti et al., 1990). Stereomotion significantly activated human area V3A in previous studies (Likova & Tyler, 2007; Rokers et al., 2009; Kaestner et al., 2019). In this area, this is not the only functional difference in motion processing between the two species as responses to structure from motion are known to be much stronger in human V3A than in macaque V3A (Vanduffel et al., 2002, see also Orban et al., 2003).

In order to avoid eye movements, and notably vergence, to contaminate our activations, we took several precautions. As mentioned in the Material and methods section, we only kept runs for which fixation performance was above 85%. Furthermore, our stimuli were designed to avoid driving vergence, with an average disparity value across space that was always equal to zero. Finally, we used the detected saccades as regressors of non-interest in our GLM. It is also worth noting that the activations we observed when contrasting our two conditions of interest (CSM vs. TS) are different from the vergence networks as investigated in terms of vergence tracking and vergence steps by Ward and collaborators (Ward et al. 2015). Additional analyses based on fixation performances and variance of eye position along the x and y axes during the CSM and TS conditions further demonstrated that eye movements did not impact our results (see Figure 34 and the accompanying text).

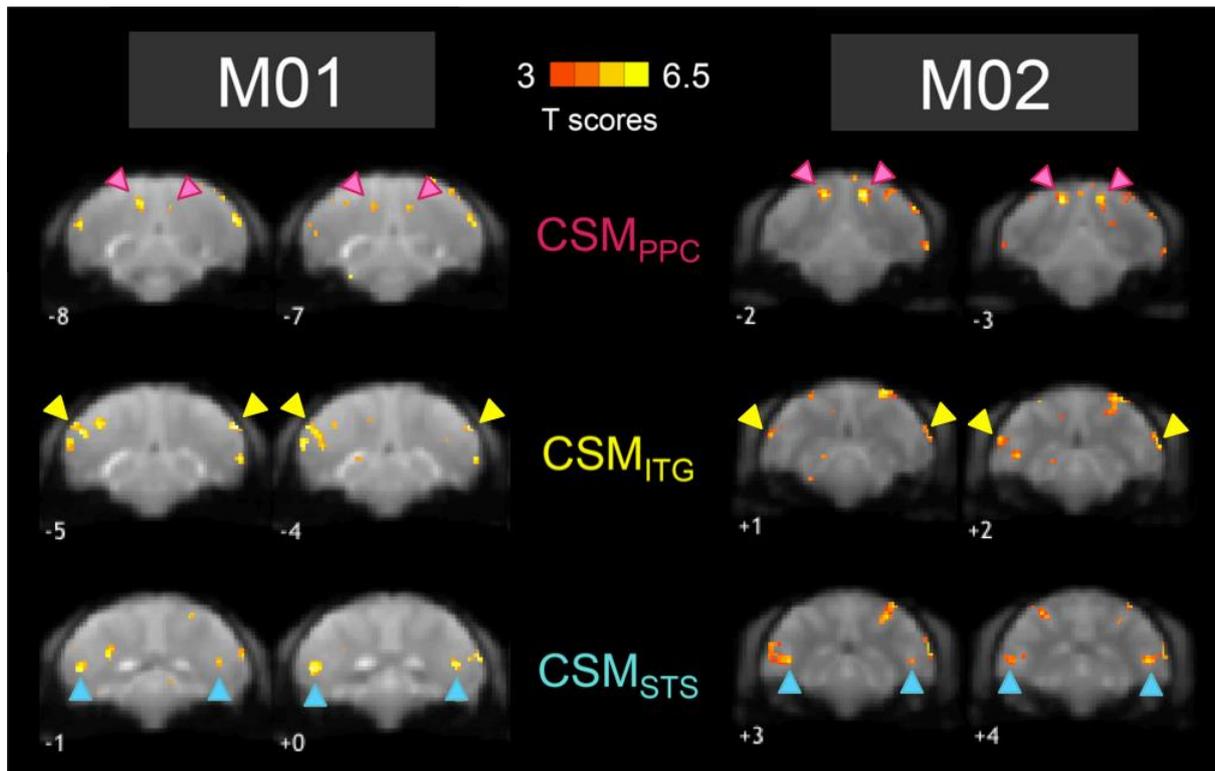


Figure 32. Activations for the contrast between Cyclopean Stereomotion (CSM) and its temporally scrambled version (TS) for M01 and M02. Figure shows activations that were stronger for the CSM condition than for the TS condition. Data are projected on an individual average EPI image of each of the macaque and are shown for different coronal slices. Coloured arrows indicate the localisation of our three regions of interest: CSM_{STS} (in blue), CSM_{ITG} (in yellow), and CSM_{PPC} (in pink). T-scores were obtained after computing the statistical parametric map for the contrast of interest between CSM and TS.

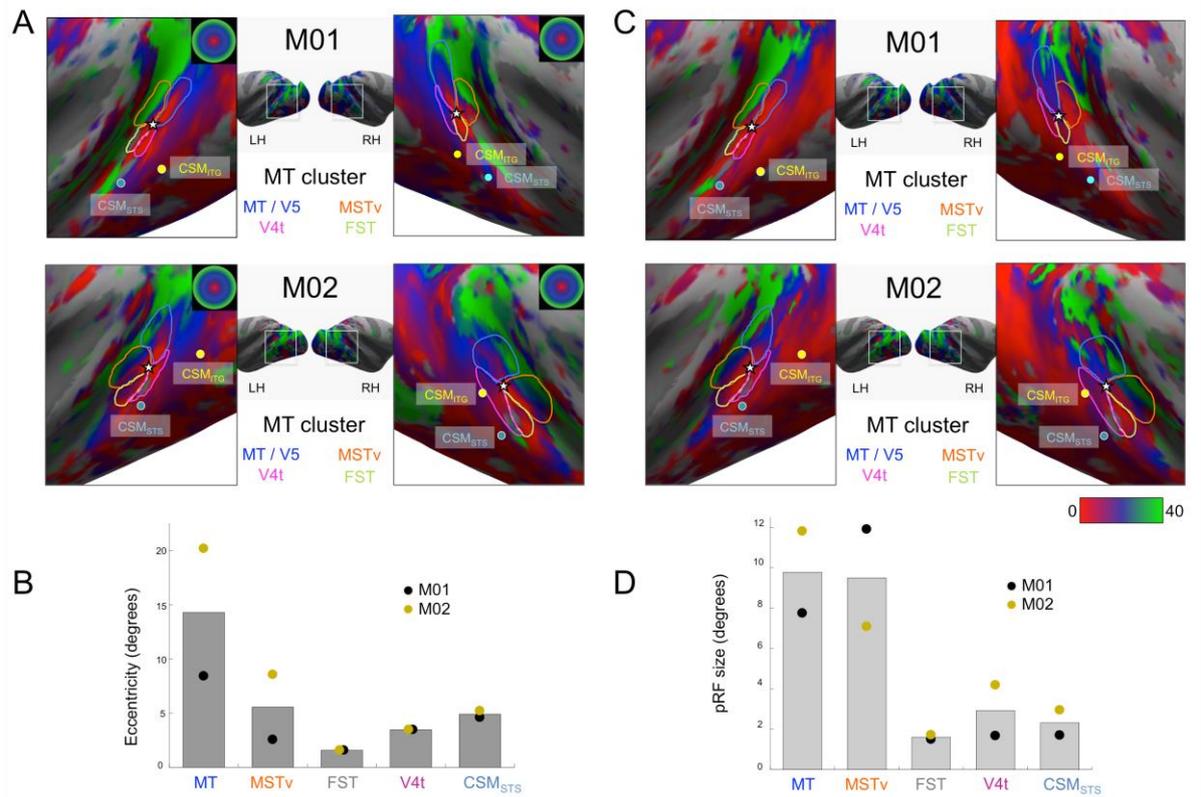


Figure 33. A) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M01 and M02 and delineation of the MT cluster areas: MT (dark blue), V4t (pink), MSTv (orange), and FST (green). We show here the eccentricity maps of both subjects (M01 and M02) and the local maxima positions for areas CSM_{STS} (blue dot) and CSM_{ITG} (yellow dot). B) Average eccentricities for the four areas of the MT cluster (same colours as in A) and for CSM_{STS} for both subjects (M01 in black and M02 in gold). C) Retinotopic mapping of the Superior Temporal Sulcus (STS) for M01 and M02 and delineation of the MT cluster areas: MT (dark blue), V4t (pink), MSTv (orange), and FST (green). We show here a map of the population receptive fields (pRF) size for both subjects (M01 and M02) and the local maxima positions for areas CSM_{STS} (blue dot) and CSM_{ITG} (yellow dot). D) Average pRF sizes for the four areas of the MT cluster (same colours as in C) and for CSM_{STS} for both subjects (M01 in black and M02 in gold).

Influence of eye movements

To confirm that eye movement did not affect our results, we further analysed the oculometric data, by determining whether fixation performances significantly varied between our two conditions of interest (CSM and TS). The distributions of these performances in our two animals for the two conditions are shown in Figure 34. Given that these data were not normally distributed, this was done using permutation tests (see the Materials and Methods for more details). We did not find any significant difference in fixation performances between our two conditions in M02. In M01, fixation scores were slightly higher for the TS condition ($p = 0.0124$). Nonetheless, these differences were not correlated with differences in PSC in our three ROIs (CSM_{PPC}, CSM_{ITG}, and CSM_{STS}). We therefore conclude that our results were not driven by differences of fixations scores between our two conditions. We use the same approach to analyse the variance in eye fixation along the x and y-axes. Variances in X and Y coordinates did not vary between our two conditions in M01. In M02, variance in X and Y were slightly higher for the TS condition but here as well, we did not find correlation between these differences in variance and difference of PSC in our 3 ROIs. It is thus highly unlikely that the ocular behaviour of our subjects had any impact on our results.

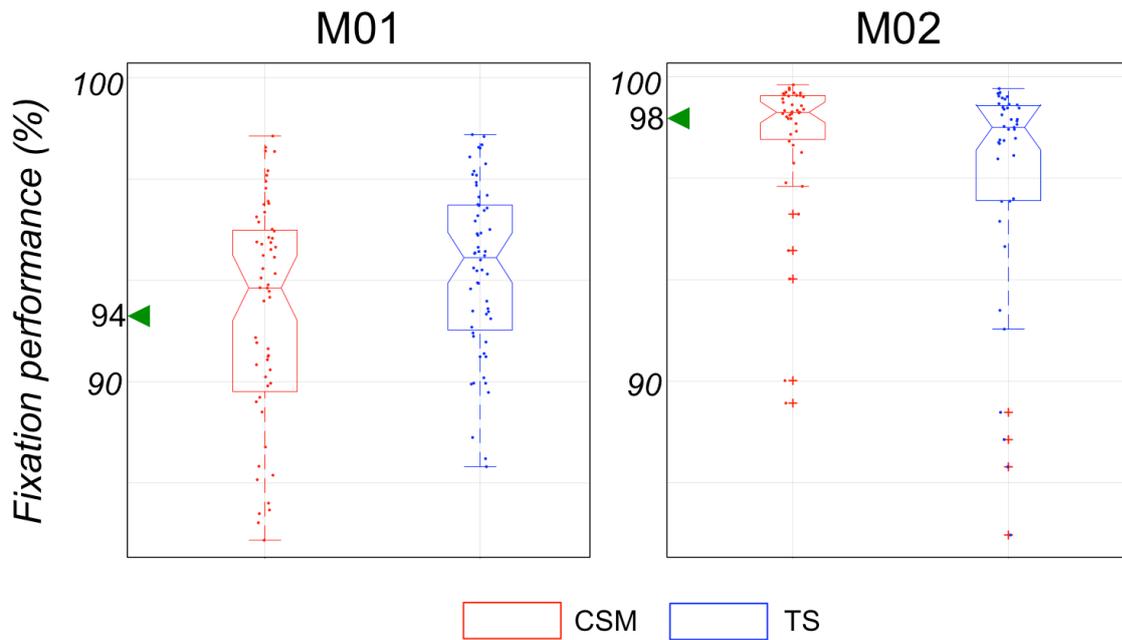


Figure 34. Ocular behaviour. Boxplots represent the distribution of percentages of correct fixation (i.e. less than 2° away from the fixation point) for both CSM and TS conditions and for each monkey (median and interquartile values). Numbers and arrows in green indicate the median values of the percentage of fixation performance during the baseline condition. Jittered dots are the fixation performance values for every run.

Conclusion

Our fMRI recordings in two macaques demonstrated that cyclopean stereomotion is mainly processed by three cortical areas: CSM_{STS} , CSM_{ITG} , and CSM_{PPC} . We also observed a moderate selectivity in areas of the MT cluster, mostly MST_V , and FST. These results are close to those described in human using a similar experimental protocol and therefore suggest that the cortical network processing stereomotion is relatively well preserved between the two primate species.

Chapter IV –
Spatial integration of binocular disparity:
Influence of natural statistics

Content of this chapter

In this chapter, I will present two studies whose shared objective is to better understand the extent to which natural statistics in visual scenes might influence depth processing and perception in human and non-human primates.

The first study is a neuroimaging study conducted in macaque subjects. Its goal was to determine whether cortical areas involved in processing 3D orientations might be biased towards more frequently encountered 3D orientations.

The second study is a psychophysical study aiming to measure the location of retinal corresponding points in both macaque and human subjects and to see whether there are interspecies differences that could be due to different visual environments.

Both studies will be presented separately as independent studies, with their own introduction and discussion. Analyses are still ongoing and the results presented here are thus still preliminary.

Second study: Spatial disparity gradients

Introduction

Although the spatial integration of disparity has been well documented in both human and non-human primates, the extent to which it is influenced by natural statistics remains unclear. As mentioned in the contextual introduction of this thesis, some studies suggested that the way binocular disparities are encoded at the cortical level matches well the distribution of binocular disparities in visual scenes. This was shown for area V1 of the macaque monkey (Sprague et al., 2015), where neurons with their receptive field (RF) in the lower visual field preferred near disparities and those with RF located in the upper visual field preferred far disparities. A similar pattern was found for areas V2 and V3 of the human brain (Nasr and Tootell, 2016) in which disparity-selective columns had different activity as a function of their representation of the visual field: BOLD signal was stronger for near stimuli presented in the lower visual field representation and stronger for far stimuli displayed in the upper visual field representation of those disparity selective columns.

Although this disparity columnar organisation was also evidenced in other visual areas of the macaque brain¹⁶, it is not known whether there is a link between the RF position and their disparity tuning range for those neurons that could reflect an encoding of the disparity distribution found in natural scenes. Furthermore, demonstrating the existence of such a relationship at the neural level required to gather data from five different studies to have a sufficient number of neurons (Sprague et al., 2015). It thus seems that electrophysiology is not an ideal technique given that such an encoding bias towards natural statistics is more likely to

¹⁶ For instance, disparity-selective neurons were found to be grouped together according to the type of disparity they prefer. Near disparity tuned neurons on the one side and far disparity tuned neurons on the other side are segregated in different columns in the macaque visual areas V2 (Cheng et al., 2008), V3 (Adams and Zeki, 2001), and MT (DeAngelis and Newsome, 1999).

be found at the population level. Within the framework of this PhD thesis, we chose to explore statistical biases at the macroscopic scale using functional neuroimaging in macaque, as this approach will give us access to the whole cortex.

In order to target higher visual areas, we used disparity gradients depicting a 3D orientation, such as slanted or tilted surfaces. They are ideal because they are made of both near and far disparities that vary as a function of elevation. Moreover, they are present in natural scenes in a more or less frequent manner, for example the orientation of the ground is a slanted surface, so we can use them to address our question regarding an influence of natural statistics at the cortical level. Finally, we already have a good idea of the cortical areas that respond to 3D orientations. For instance, invasive electrophysiological recordings in the macaque brain notably revealed the major implication of the parietal cortex in processing disparity gradients or first-order depth stimuli (Taira et al., 2000; Tsutsui et al., 2002; 2005). In human, several fMRI studies also showed that slants defined by disparity gradients activate the parietal cortex (Shikata et al., 2007), as well as intermediate areas such as V3A (Ban and Welchman, 2015). A more recent study recorded neurons in the area V3A and in the caudal intraparietal area CIP of macaques (Elmore et al., 2019) and confirmed the sensitivity of both areas to surface orientations (slants and tilts) in macaque. Could there be a specialisation or higher sensitivity for more frequent 3D orientations in those higher visual areas, as the one that has been found in V1 for macaque and in V2 and V3 for human? If the cortical processing of binocular disparities is influenced by or reflects natural statistics, then we should find a differential activity in those 3D orientation selective areas when we display frequently found versus less frequently encountered 3D orientations.

The objectives of this neuroimaging study were thus two folds. First, we aimed to characterise the cortical network responding to correlated disparities in two macaque subjects. Second, we asked whether orientation biases could exist in that network that would reflect natural statistics.

Subjects

Two female rhesus macaques (age: 13-15 years; weight: 5.35-6.15 kg) were involved in the study. Animal housing, handling, and all experimental protocols (surgery, behavioural training, and MRI recordings) followed the guidelines of the European Union legislation (2010/63/UE) and of the French Ministry of Agriculture (décret 2013–118). All projects were approved by a local ethics committee (CNREEA code: C2EA – 14) and received authorisation from the French Ministry of Research (MP/03/34/10/09). Details about the macaques' surgical preparation and behavioural training are provided elsewhere (see for instance Cottureau et al., 2017 in the appendix section).

Experimental design

Our stimuli were defined using dynamic random-dot-stereograms (dRDS) shaped as a disk (23 degrees of radius) and refreshed at 20 Hz with a dot coverage set at 50%. To introduce binocular disparity, dots were green in one eye and red in the other eye, corresponding to two separate retinal projections. Stimuli were observed through red-green anaglyphs. We created different types of conditions by varying some parameters of our stimuli (Figure 35).

In the “GS” (ground-aligned surface) condition, a gradient of binocular disparities was introduced between the left and right monocular images to create a slanted surface. Negative disparity values (crossed disparities) in the lower visual field gradually became positive (uncrossed disparities) in the upper visual field, giving rise to a slant aligned with the ground orientation at infinity: closer in the lower part and further away in the upper part with respect to the fixation point in the fronto-parallel plane.

In the “nGS” (non-ground-aligned surface) condition, the same gradient was reversed: disparities were positive in the lower visual field and negative in the upper visual field. The

slant was then perceived as aligned with the opposite direction: closer in the upper part and further away in the lower part. Importantly, if the ‘GS’ and ‘nGS’ conditions differed by the sign of their disparity gradient, they nonetheless shared the exact same disparity distribution.

In the “LT” (leftward tilt) condition, a similar gradient was used to define a tilted surface, that is, a surface vertically oriented. In this case, crossed disparities (negative) were presented in the left hemifield, making it appear closer than the front plane, whilst uncrossed disparities were in the right visual hemifield, making this part of the stimulus look further away from the fixation plane.

In the “RT” (rightward tilt) condition, a similar tilted surface was shown with the difference that it was rightward tilted: the surface was closer in the left hemisphere (crossed disparities) and further away in the right hemisphere (uncrossed disparities), with disparities values still following a gradient. Here as well, if the ‘LT’ and ‘RT’ conditions differed by the sign of their disparity gradient, they nonetheless shared the exact same disparity distribution.

In the “U” (uncorrelated) condition, dots were not correlated, suppressing any 3D percept.

For all conditions except the uncorrelated one, the angle value of the surface orientation (slant or tilt) was 4.13 degrees. More importantly, when viewed monocularly, our stimuli did not contain any 3D information, (i.e. there was no perception of any slant nor tilt) and were all exactly identical.

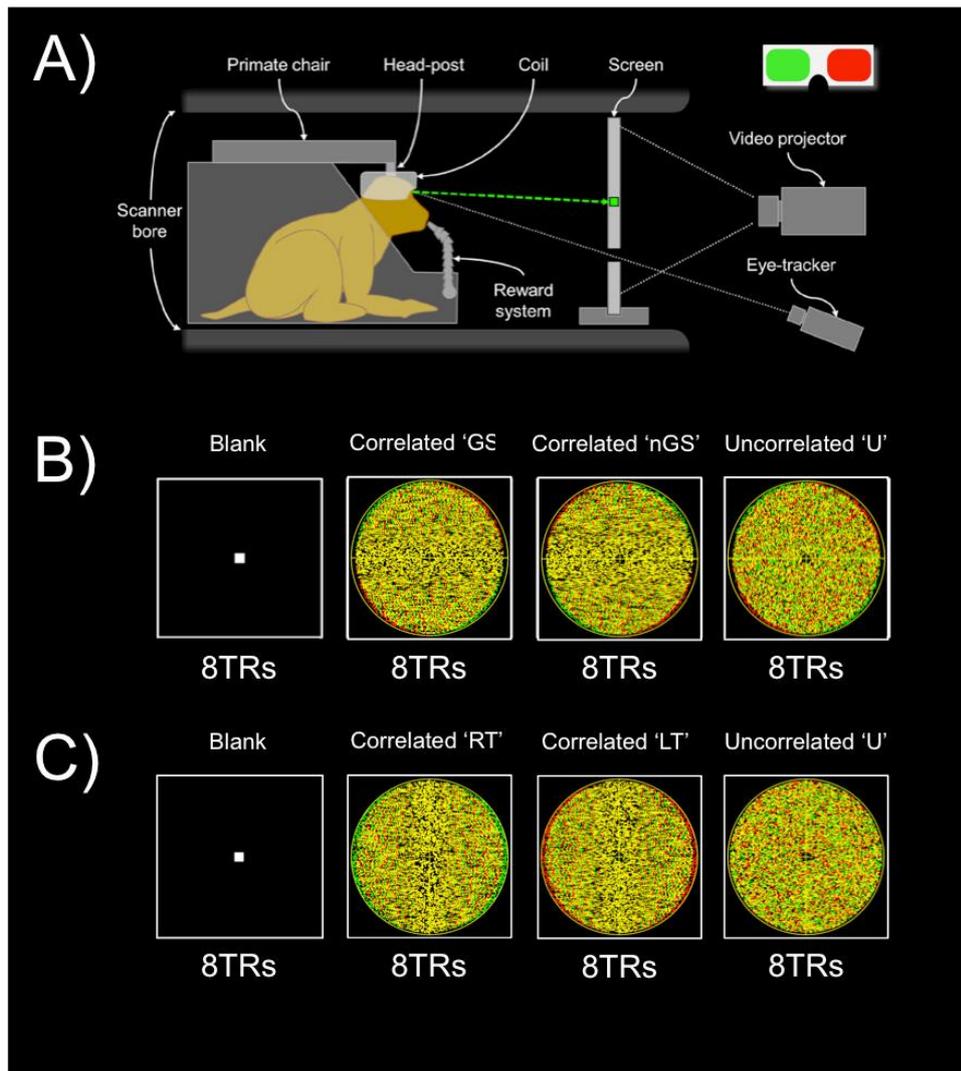


Figure 35. Experimental protocol.

A) Schematic representation of the monkey fMRI setup. The animal *sat* in a sphinx position within the primate chair, in the bore of the scanner, with the 8-channel, phase-array coil placed on top of the head. The animal was involved in a passive fixation task while its eye position was monitored by an infrared video-based eye-tracker. Horizontal disparity was introduced through the use of red-green anaglyphs. B) Illustration of the experimental protocol. Recordings were performed using a block design, with the successive presentation of correlated conditions (here 'GS' and 'nGS'), uncorrelated condition ('U') and a blank screen in the 'slant' runs. Each block contained 3 repetitions of such blocks. A block would always start with a blank screen but the visual stimuli order was pseudorandomised across runs. C) Illustration of the successive presentations of correlated (here 'LT' and 'RT') conditions, uncorrelated conditions, and a blank screen in the 'tilt' runs.

Image acquisition: Templates of reference and functional sessions

Whole-brain images were acquired on a 3 Tesla MR scanner (Phillips Achieva) using a custom 8-channel phased-array coil (RapidBiomed) specifically designed to fit the skull of our macaques while preserving their field of view. Four T1-weighted anatomical volumes were acquired prior to the study for each monkey at a high resolution (MPRAGE; repetition time, TR = 10.3 ms; echo time, TE = 4.6 ms, flip angle = 8°; FOV: 155x155 mm; matrix size: 312x192 mm; voxel size = 0.5 x 0.5 x 0.5mm; 192 sagittal slices acquired in an interleaved order), as well as 300 functional volumes (gradient-echo EPI; TR = 2,000 ms, TE = 30 ms, flip angle = 75°, SENSE factor = 1.6; FOV: 100x100 mm; matrix size: 68x64 mm; voxel size = 1.25 × 1.25 × 1.5mm, 32 axial slices acquired in an interleaved order with a thickness of 1.5 mm and no gap). Those data were recorded in a single session whilst the macaques were slightly anaesthetised (Zoletil 100:10 mg/kg and Domitor: 0.04mg/kg) and their constants monitored with an MR compatible oximeter. Those volumes were then used to create individual anatomical and functional templates of reference.

Our T2*-weighted functional images were acquired with a gradient-echo EPI sequence with interleaved slice acquisition (TR = 2,000 ms, TE = 30 ms, flip angle = 75°, SENSE factor = 1.6; FOV: 100x100 mm; matrix size: 68x64 mm; voxel size = 1.25 × 1.25 × 1.5 mm, 32 axial slices acquired in an interleaved order with a thickness of 1.5 mm and no gap).

Scanning Procedure

During the functional recording sessions, macaques were head-fixed and seated in a sphinx position in a dedicated primate chair (see Figure 35-A). They had to maintain their gaze within a central fixation window (2x2°) during daily sessions of up to 2 hours. Their fixation was monitored with an ASL© infrared video-based eye tracking setup at 60Hz and they were

rewarded through a liquid delivery system at intervals whose frequency depended on their fixation performance. Our stimuli were video-projected on a 768*768-pixel translucent screen (display size: 42.5cm) facing the macaques at a viewing distance of 85cm.

We used a block design based on cycles within which our conditions were displayed during blocks of 16 seconds (8 TRs). Slanted surfaces and tilted surfaces were presented in separate runs: in ‘slant’ runs and in ‘tilt’ runs, respectively. In one run, two correlated conditions (either “GS” and “nGS” or “LT” and “RT”), the uncorrelated condition (“U”), and a simple fixation point (baseline condition) were displayed once during one cycle in a pseudorandomised order (see Figure 35-B and C for an example). Each cycle lasted 64 seconds (32 TR) and was repeated 3 times, resulting in a total duration of 96 TR (192 seconds). We displayed the stimuli, controlled for the delivery of the liquid reward and the fixation performance using the EventIDE software (Okazolab®).

Data analysis

Templates of reference

Anatomical and functional templates of reference were created for each individual with the volumes acquired prior to the current study. The anatomical template was obtained with the four T1-weighted anatomical volumes being realigned, averaged, and then co-registered on the MNI space of the 112RM-SL template (McLaren et al., 2009, 2010). To create the functional template, the 300 functional volumes (GE-EPI) were realigned, averaged, and then co-registered on the anatomical template. Mean images were then segmented in order to obtain tissue probability maps for the grey matter, the white matter, and the cerebrospinal fluid (CSF). Spatial normalisation parameters between both templates could then be estimated.

Pre-processing of the raw functional data

Pre-processing and volume-based analyses were carried with SPM12 in the Matlab® environment. Only runs with central gaze fixation above 85% were kept for further analysis. In total, we kept 49 (26 ‘slants’ and 23 ‘tilts’) and 79 (33 ‘slants’ and 46 ‘tilts’) runs for both macaques, respectively. Pre-processing was performed for each volume, run by run. Slice-timing correction (reference slice: 31st slice – middle one) was performed first. Images were then reoriented, co-registered with the EPI template, and deformed to fit the individual T1 template. No motion correction was applied to the images. Finally, the images were smoothed with a spatial Gaussian kernel (FWHM = 2x2x2 mm).

HRF estimation

Prior to our statistical analyses, we used independent datasets to characterise the BOLD haemodynamic impulse response functions (HRF) separately for each animal. These datasets respectively contained 16 (M01) and 12 (M02) 204s long runs that consisted of 6 cycles of 4s full field counter phasing (10Hz) checkerboards separated by a 30s blank interval (see more details about this procedure in Cottureau et al., 2017 and in the Monkey fMRI chapter). Data were pre-processed using the pipeline described above and projected onto individual surfaces generated with the CARET software (Van Essen et al., 2001). Following Dumoulin and Wandell’s procedure (2008), we extracted the BOLD responses from nodes within the anatomically defined V1 of each individual. We only kept visually responsive nodes, that is those whose signal-to-noise ratio (SNR) was greater than 3. This SNR was estimated with a Fourier analysis of the average time courses across runs where the signal corresponded to the Fourier coefficient amplitude at the stimulation frequency F (i.e. $F = 1/34$) and the noise was given by the average moduli at the two neighbouring frequencies (i.e. $F - \delta f$ and $F + \delta f$, where $\delta f = 1/2$ is the resolution of our frequency analysis). We computed the average time course of

these nodes during one cycle and used this average time course for estimating the HRF. The HRF was derived as the response to a 2s stimulus (our fMRI sampling rate). Note however that our stimulus duration was 4s rather than 2s because linearity deteriorates at short durations (Boynton et al. 1996; Logothetis and Wandell, 2004) and also because this duration was used in a previous monkey fMRI study that characterised the BOLD HRF in macaque (Leite et al., 2002). For each monkey, the average response to the 4s stimulus was fit as the convolution of the responses to two 2s responses, each of which is the HRF. We parameterised the HRF as the difference of two gamma functions (Friston et al., 1998). This functional form of the HRF captures the late undershoot of the response better than a single gamma function (Boynton et al., 1996).

General linear model (GLM)

Univariate statistics were performed at the voxel level, using a general linear model (GLM). Our visual conditions (correlated 1, correlated 2, uncorrelated, baseline) were implemented as the 4 main regressors of the GLM. Slant runs and tilt runs were analysed separately. As reported above, we only analysed runs with fixation performance greater than 85%. To characterise and eliminate noise in our recordings, we also performed a principal component analysis on voxels located outside the brain (see Farivar and Vanduffel, 2014). Indeed, time courses in those voxels should not be influenced by our experimental design but rather reflect artefacts caused by movements of the animal. For each run, we determined the number of principal components that were necessary to explain 80% of the variance in these voxels and used the corresponding principal vectors as regressors of non-interest in our model. This adaptive procedure typically added an average of 18.08 (± 8.91) and 10.12 (± 5.55) additional regressors in the model for the ‘slants’ runs and of 13.61 (± 7.32) and 7.17 (± 1.64) for the ‘tilts’ runs.

We estimated the beta values associated with our GLM using the RobustWLS toolbox (Diedrichsen & Shadmehr, 2005), which is provided as an additional toolbox for SPM12 (<http://www.diedrichsenlab.org/imaging/robustWLS.html>). This approach yields to estimate the noise variance for each image in the time series, using the derivative of a maximum likelihood algorithm. Variance parameters are then used to obtain a weighted least square estimate of the regression parameters of the GLM. It therefore helps reduce the impact of noisier volumes on beta estimation. Previous studies showed that such a method significantly improved estimations in blocked-design fMRI experiments (see e.g. Takeuchi et al., 2011).

Results

We first used SPM12 to compute the statistical parametric maps (SPM maps) corresponding to the contrast between our different conditions. These maps were thresholded at $p < 0.05$ (t -values > 4.8) and corrected voxel-wise for multiple comparisons (FWE correction). Retinotopic analyses were further conducted based on regions that were defined on an independent dataset. Further details about the procedure to obtain those retinotopic areas are given in chapter III.

We first looked at the cortical network involved in processing binocular disparities and thus computed an SPM map for the contrast ‘correlated’ vs. ‘uncorrelated’. Correlated conditions gathered either slanted surfaces (GS and nGS) or tilted surfaces (RT and LT) and were opposed to their respective uncorrelated condition (U). Activation maps are shown on Figure 36 and Figure 37 for the correlated (GS+nGS) vs. uncorrelated conditions and on Figure 38 and Figure 39 for the correlated (RT+LT) vs. uncorrelated conditions.

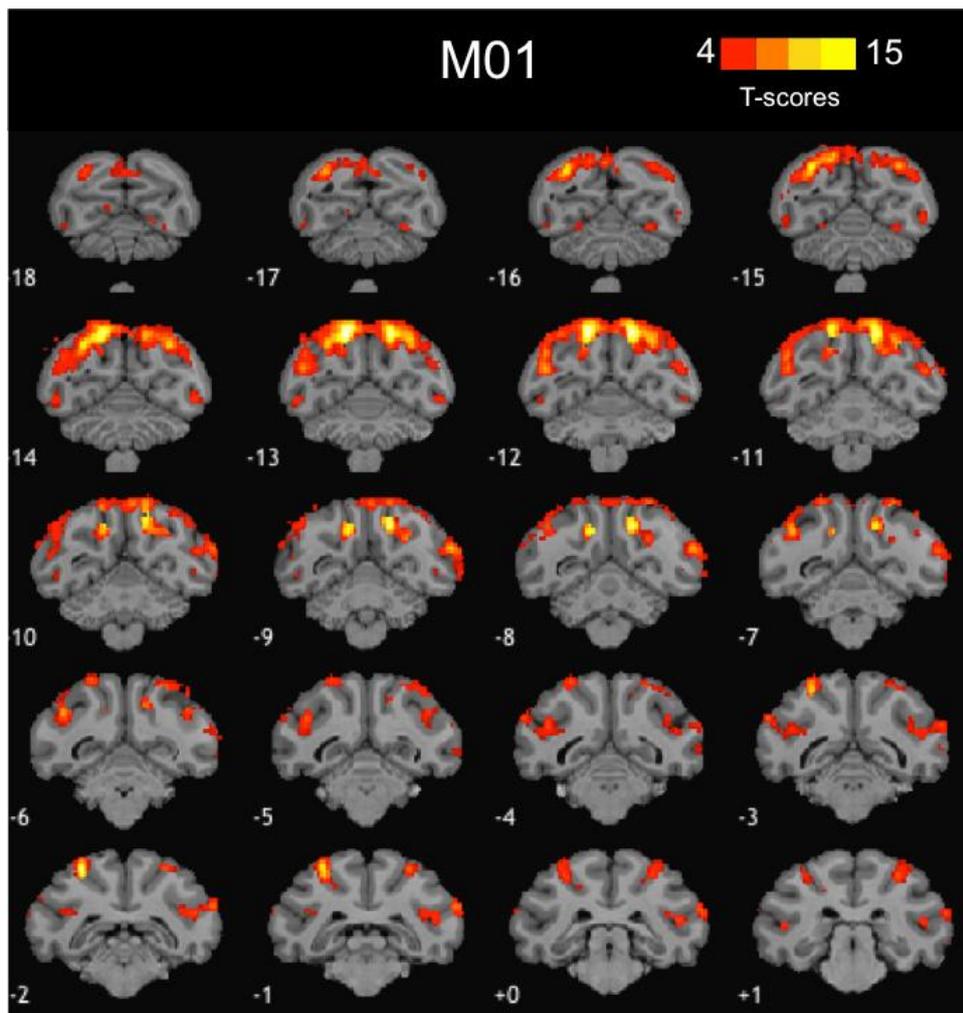


Figure 36. Activations for the contrast Correlated (GS+nGS) vs. Uncorrelated for M01.

Figure shows activations that were stronger for the correlated (GS and nGS together) than for the uncorrelated (U) condition. Data are projected on the individual anatomical template of the macaque and are shown for different coronal slices. T-scores were obtained after computing the statistical parametric map corrected for multiple comparisons (FWE correction) for the contrast of interest between our conditions.

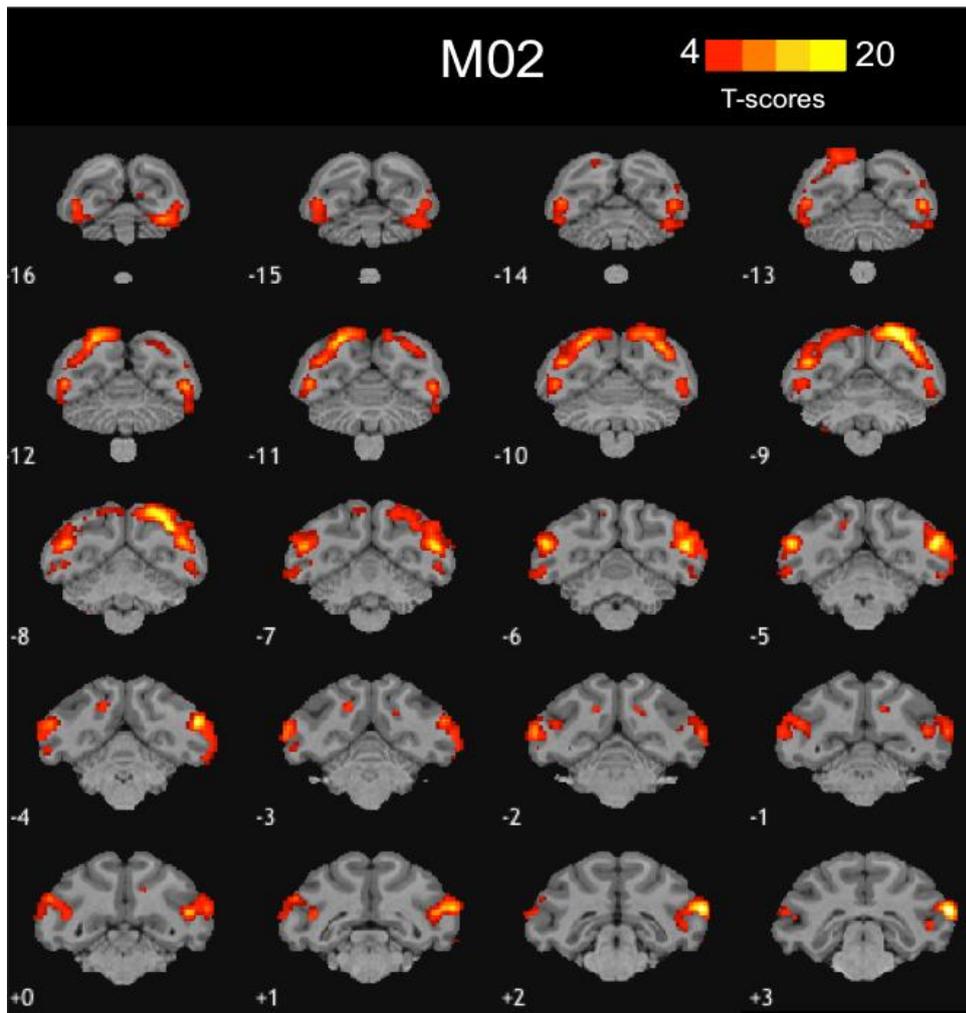


Figure 37. Activations for the contrast Correlated (GS+nGS) vs. Uncorrelated for M02. See Figure 36 for the details of the legend.

On Figure 36 and Figure 37, red-to-yellow colours indicate stronger BOLD activations for the Correlated (GS+nGS) conditions ($p < 0.05$, FWE corrected). This analysis reveals similar bilateral activation patterns between both animals, in particular in the posterior parietal cortex with spread activations that are visible on slices -15 to -11 and -12 to -8 for M01 and M02, respectively (upper part of the brain, mostly). The MT cluster also seems to be more strongly activated for correlated than uncorrelated stimuli as it can be seen on slices -6 to -2 and -6 to +1 for M01 and M02, respectively (middle part of the brain).

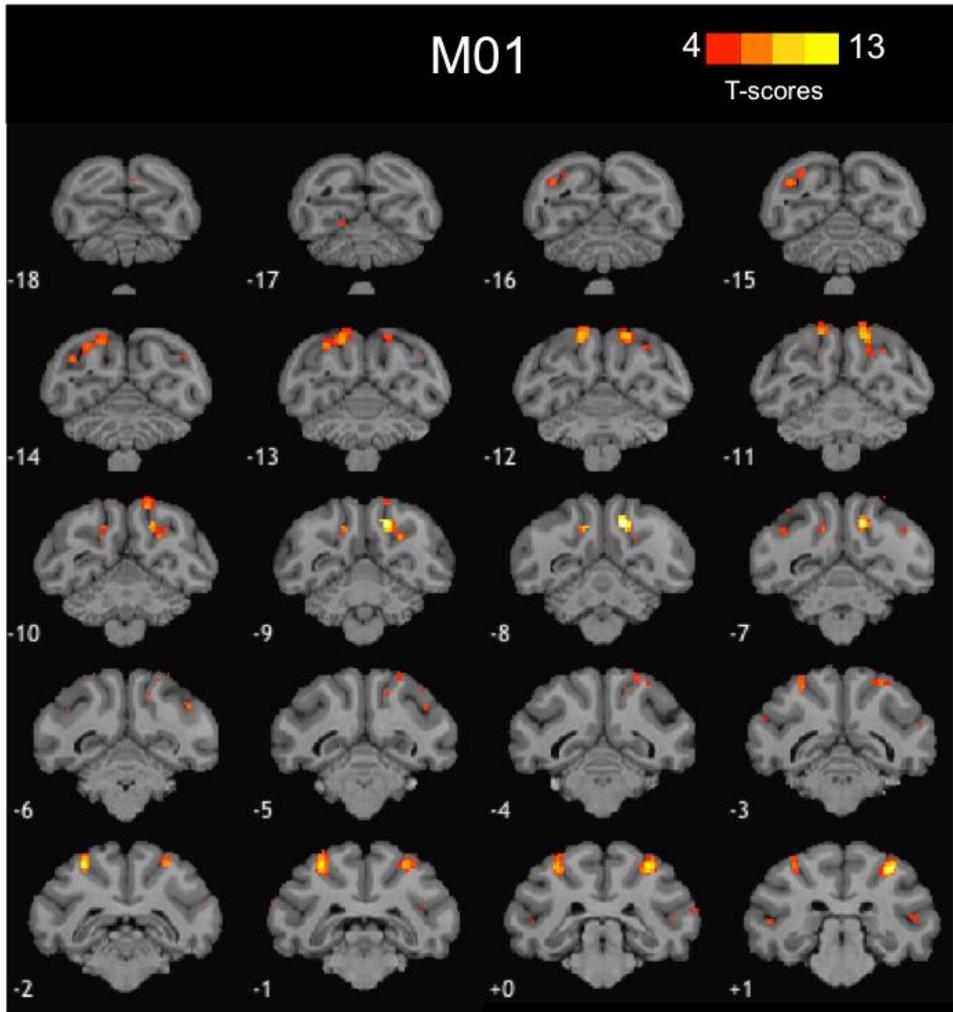


Figure 38. Activations for the contrast Correlated (RT+LT) vs. Uncorrelated for M01. Figure shows activations that were stronger for the correlated conditions (RT and LT together) than for the uncorrelated one (U). Data are projected on the individual anatomical template of the macaque and are shown for different coronal slices. T-scores were obtained after computing the statistical parametric map corrected for multiple comparisons ($p < 0.05$, FWE corrected) for the contrast of interest between our conditions.

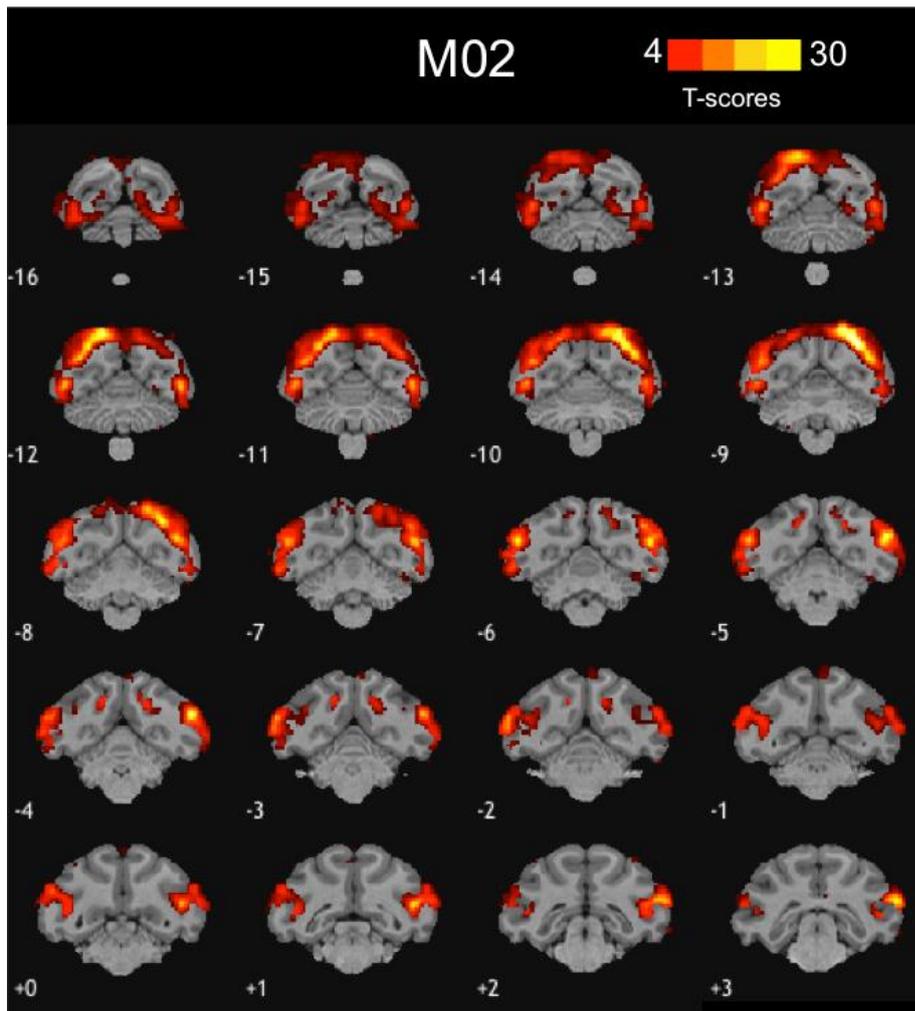


Figure 39. Activations for the contrast Correlated (RT+LT) vs. Uncorrelated for M02.

See Figure 38 for details of the legend.

On Figure 38 and Figure 39, red-to-yellow colours again indicate stronger BOLD activations for the Correlated (RT+LT) conditions ($p < 0.05$, FWE corrected). Although far more visible for M02 than M01 (potentially due to the difference in run number: 46 vs. 23 runs), the activation patterns for correlated vs. uncorrelated disparities are very similar to the patterns obtained for the same contrast but with the ‘slant’ runs.

In order to get a more precise idea of those activations and to be able to compare them to the areas known for being responsive to correlated disparities, we performed a region of interest

(ROI) analysis in 15 independently defined retinotopic areas. We considered voxels within one retinotopically defined area only when they were significantly activated during the visual vs. baseline contrast (voxels whose t -scores were above 3).

ROI analysis in retinotopically defined areas

We computed the difference in the percentage of signal change (PSC) between our conditions, using the following equations:

$$\Delta = PSC_{correlated} - PSC_{uncorrelated}$$

with:

$$PSC_{correlated} = \frac{\frac{\beta_{C1} + \beta_{C2}}{2}}{\beta_{constant}} - \frac{\beta_{baseline}}{\beta_{constant}} \times 100$$

$$PSC_{uncorrelated} = \frac{\beta_{uncorrelated} - \beta_{baseline}}{\beta_{constant}} \times 100$$

where the correlated conditions were averaged to echo the SPM maps representation of our contrast of interest (correlated vs. uncorrelated). However, as mentioned later, control tests were also made independently for each correlated stimulus vs. its respective uncorrelated condition.

To assess the statistical values of the observed PSC differences, we further performed permutation tests. We randomly attributed a negative sign to our PSC values and computed the average difference, repeating this procedure 10,000 times. We could then calculate a

¹⁷ Here C1 stands for the first correlated condition, either ‘GS’ or ‘RT’, and C2 stands for the second correlated condition, either ‘nGS’ or ‘LT’, respectively.

p -value defined as the proportion of random differences that were superior to our observed difference. The significance threshold was set a priori at 0.05 to facilitate data visualisation, we grouped our retinotopic areas as follows: The early visual areas gather V1, V2, V3, and V4; the MT cluster groups V4t, MT, MSTv, and FST; the PIP cluster comprises CIP1, CIP2, PIP1, and PIP2. Finally, the three remaining retinotopic areas we could delineate were grouped together: V3A, V6, and V6A.

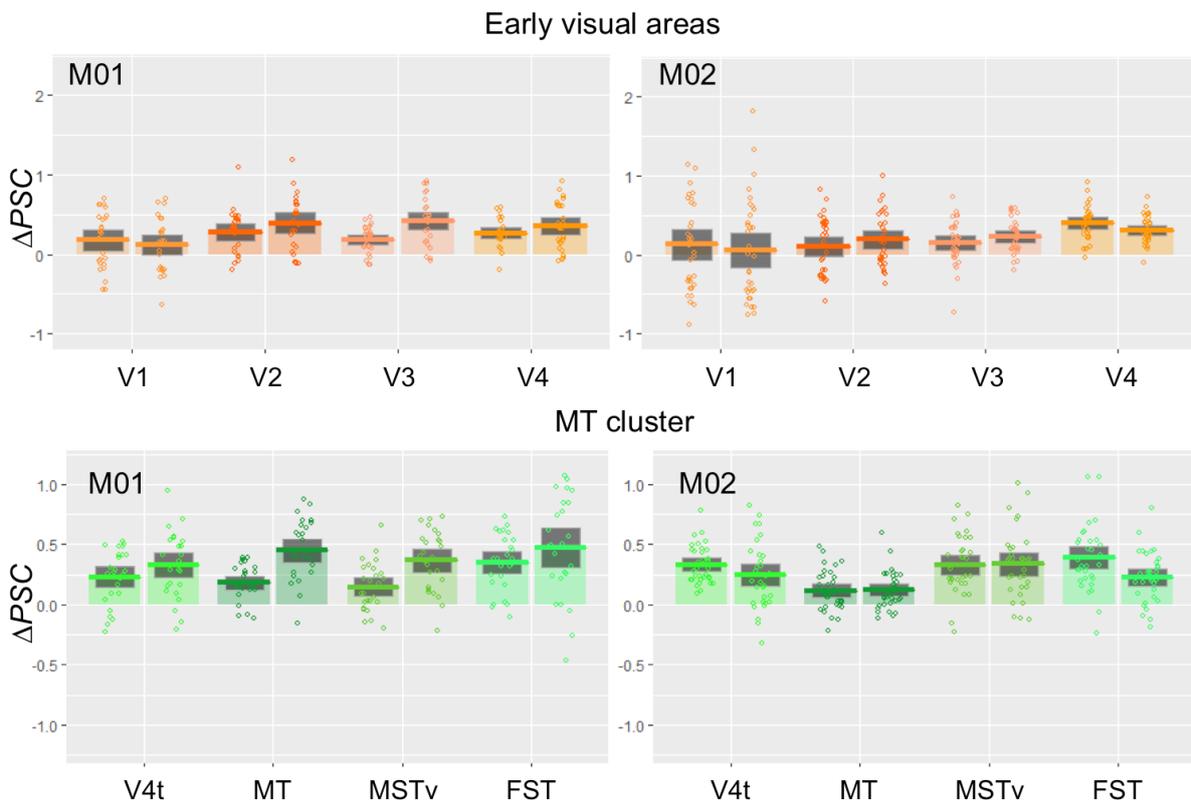


Figure 40. Difference in signal change (ΔPSC) between correlated (GS+nGS) and uncorrelated conditions in early visual areas (in orange) and in the MT cluster (in green).

The thick lines of the bar graphs provide the average values for the left and right hemisphere of the two monkeys (M01 and M02). The box gives the 95% confidence intervals for these average values. The dots provide the data for each run. A small jitter was introduced to facilitate visibility.

As can be seen on the bar graphs shown in Figure 40 and Figure 41, our permutation tests revealed that for most retinotopic areas the average values of the PSC differences are above 0 for both left and right hemispheres ($p < 0.05$), meaning that those areas were more strongly activated for the correlated conditions. More specifically, for M01, both hemispheres of all regions but V1 (only in the left hemisphere) were significantly more strongly activated for the correlated conditions than for the uncorrelated one. For M02, areas of the MT cluster (V4t, MT, MST, and FST), areas V3, V4, V3A, V6A, CIP1, and CIP2 were significantly more strongly activated for the correlated conditions and a tendency was observed for PIP2.

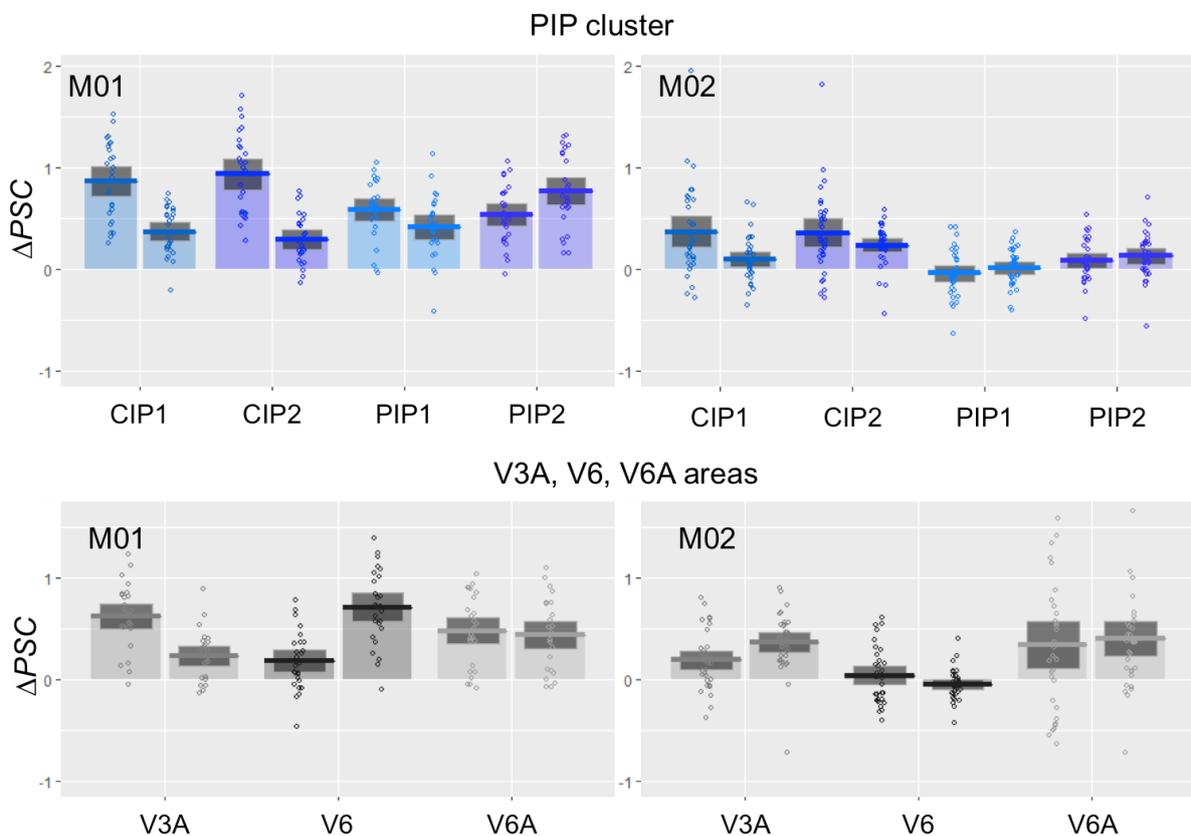


Figure 41. Difference in signal change (ΔPSC) between correlated (GS+nGS) and uncorrelated conditions in the PIP cluster (in blue) and in V3A, V6, and V6A areas (in grey).

See Figure 40 for the details of the legend.

Reproducing this analysis for the ‘tilt’ runs for M02 did not change the areas that were more significantly activated for the correlated conditions than for the uncorrelated one, with the exception of PIP2 area, which was more strongly activated for the correlated conditions in both hemispheres (vs. in the right hemisphere only with the ‘slanted’ runs). However, for M01, results are different since only CIP2 passes the significance threshold in both hemispheres. Asymmetrical results were found for M01 with more areas in the left hemisphere being significantly more activated for correlated conditions, such as PIP1, PIP2, V6A, V3A, MT, and CIP1; and the area MSTv in the right hemisphere only. Despite those differences, it is worth noting that most of the PIP cluster (CIP1, CIP2, and PIP2) and the areas V6A, V3A, MT, and MSTv are significantly more activated for the correlated conditions than for the uncorrelated conditions in at least 3 hemispheres in both ‘slant’ and ‘tilt’ types of runs. Reproducing these analyses by contrasting each single correlated condition to its uncorrelated condition (i.e., GS>U, nGS>U, RT>U, and LT>U) did not change the final pool of areas that were significantly more activated as reported in Table 3.

Table 3. Number of hemispheres that pass the significance threshold for each retinotopic areas and for each contrast. Crosses in the last row indicate the areas that were found in at least $\frac{3}{4}$ hemispheres.

	V1	V2	V3	V4	V4t	MT	MSTv	FST	CIP1	CIP2	PIP1	PIP2	V3A	V6A	V6
C(S) > U	1	1	4	4	4	4	4	4	4	4	2	3	4	4	2
GS > U	1	2	4	4	4	4	4	4	4	4	2	3	4	4	2
nGS > U	1	1	4	4	4	3	3	4	4	4	2	4	4	4	1
C(T) > U	-	-	2	2	2	3	3	2	3	4	2	2	3	3	-
RT > U	-	1	2	2	2	3	3	2	3	4	2	3	3	4	-
LT > U	-	2	3	2	2	2	2	2	3	4	2	3	3	4	1
						x	x		x	x		x	x	x	

Orientation biases

In the former section we described the cortical network involved in binocular disparities processing and we showed that several areas were more strongly activated by correlated disparities than by non-correlated ones: MT, MST_v, CIP1, CIP2, PIP2, V3A, and V6A.

If we cannot explore whether some cortical areas might respond differently to slanted than to tilted surfaces with a traditional univariate ROI analysis, we can address the question of whether cortical activations might reflect a bias for natural statistics. Indeed, since some orientations are more frequent than others, we could expect a specialisation of those cortical networks accordingly. In this section, we will investigate the cortical responses to slants aligned with the ground orientation ('GS' condition), a pattern that is more frequently encountered in natural visual scenes, as opposed to slants that are not aligned with the ground ('nGS' condition). As a comparison, we will look at the cortical responses to our two tilted surfaces: leftward ('LT' condition) vs. rightward tilts ('RT' condition). None of these two tilted orientations is more frequent than the other; we can thus use them as a control since the only parametrical difference between those tilted surfaces and the slanted surfaces is the orientation of their disparity gradients.

As a first step, we computed the SPM maps for our contrasts of interest: GS>nGS, nGS>GS, RT>LT, and LT>RT. Based on the hypothesis that a more frequently encountered orientation influences the cortical activity, we expected to find a difference between the GS and nGS condition, with stronger activations for the GS condition. However, we expected no difference between the RT and LT conditions.

Figure 42 and Figure 43 show the activations for the contrast GS>nGS for M01 and for the contrasts GS>nGS and nGS>GS for M02 projected on individual templates. Those maps reveal marked differences between our two subjects. For M01, as shown on Figure 42, there are

stronger activations ($p < 0.05$, FWE corrected) for the GS condition than for the nGS condition for which nothing passes the statistical threshold even using a more permissive statistical threshold ($p < 0.01$, uncorrected). For M02, activations are very low and pass only a lowered, no longer corrected for multiple comparisons, statistical threshold ($p < 0.01$, uncorrected). There seems to be one spot of activation that is stronger for the nGS condition in the post-parietal cortex.

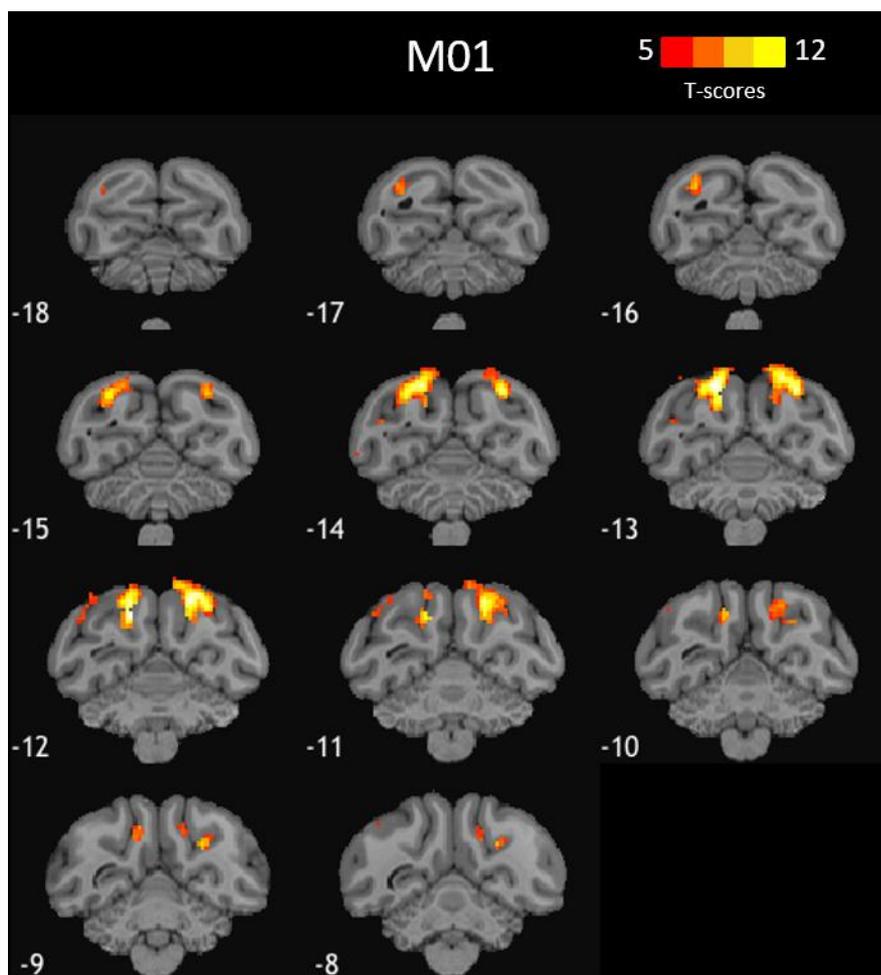


Figure 42. Activations for the contrast GS > nGS for M01. Figure shows activations that were stronger for the ground-aligned slant, ‘GS’, conditions than for the non-ground aligned surface, ‘nGS’, one. Data are projected on the individual anatomical template of the macaque and are shown for different coronal slices. T-scores were obtained after computing the statistical parametric map corrected for multiple comparisons ($p < 0.05$, FWE corrected) for the contrast of interest between our conditions.

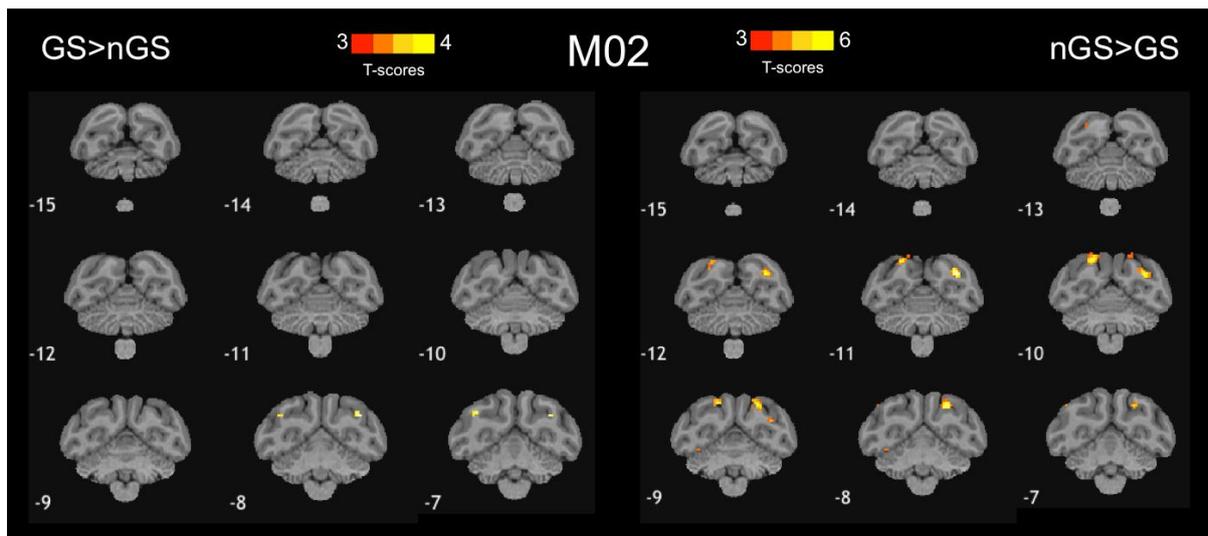


Figure 43. Activations for the contrast GS>nGS on the left and for the contrast nGS>GS on the right for M02. T-scores were obtained after computing the statistical parametric maps for the contrast of interest between our conditions ($p < 10^{-3}$, uncorrected). There seems to be one area that responds more strongly to the nGS condition than to the GS one in the post-parietal cortex.

Retinotopic analyses were performed the same way as previously described. Differences in the signal change between our two conditions of interest, GS and nGS, were computed. Running permutation tests revealed that in M01, three areas in the PIP cluster were significantly more activated for the GS condition than for the nGS condition ($p < 0.05$): PIP1, PIP2, and CIP1, whereas area CIP2 responds significantly more to the GS condition in the left hemisphere only. Bar graphs corresponding to the Δ PSC are displayed for the PIP cluster areas for M01 on Figure 44. For M02, the area V3A was significantly more activated for the nGS condition than in the GS condition but in the left hemisphere only, as visible on the Figure 43.

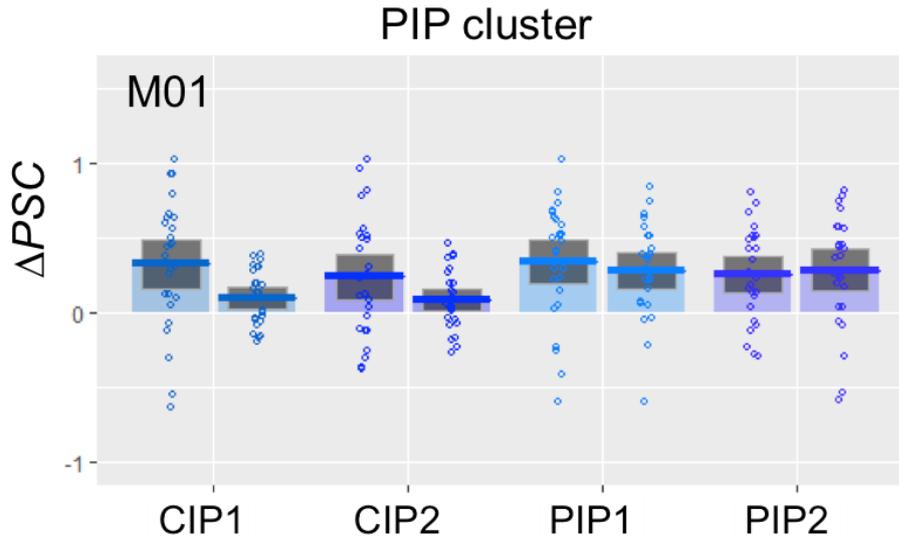


Figure 44. Difference in signal change (ΔPSC) between the ground-aligned slant (GS) and the non-ground aligned slant conditions in the PIP cluster. The thick lines of the bar graphs provide the average values for the left and right hemisphere of M01. The box gives the 95% confidence intervals for these average values. The dots provide the data for each run. A small jitter was introduced to facilitate visibility.

Finally, we computed the statistical maps for our contrasts of interest with the tilted surfaces (RT>LT and LT>RT) as well as the differences in the signal change between the RT and LT conditions for both subjects. The SPM maps are shown on Figure 45 and Figure 46. Again, we found no common pattern between our two animals, with M01 showing stronger activations than M02, despite having much fewer runs (23 and 46 runs for M01 and M02, respectively).

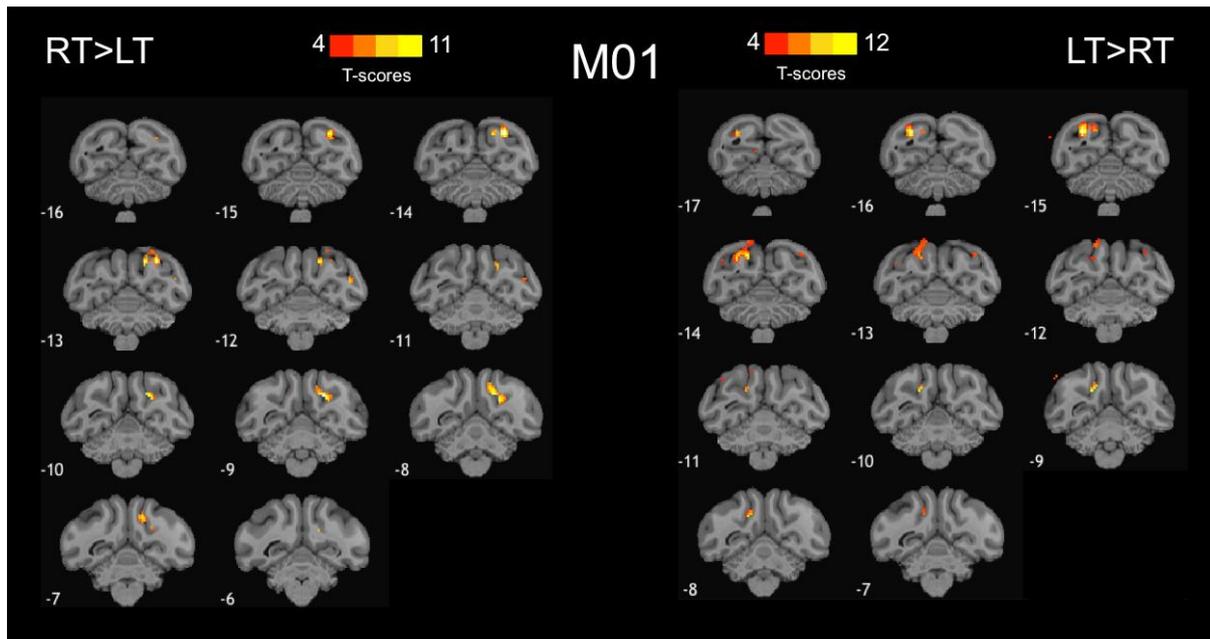


Figure 45. Activations for the contrasts $RT > LT$ on the left and $LT > RT$ on the right for M01. Figure shows activations that were stronger for the rightward tilt on the left and leftward tilt on the right. Data are projected on the individual anatomical template of the macaque and are shown for different coronal slices. T-scores were obtained after computing the statistical parametric map corrected for multiple comparisons ($p < 0.05$, FWE corrected) for the contrast of interest between our conditions.

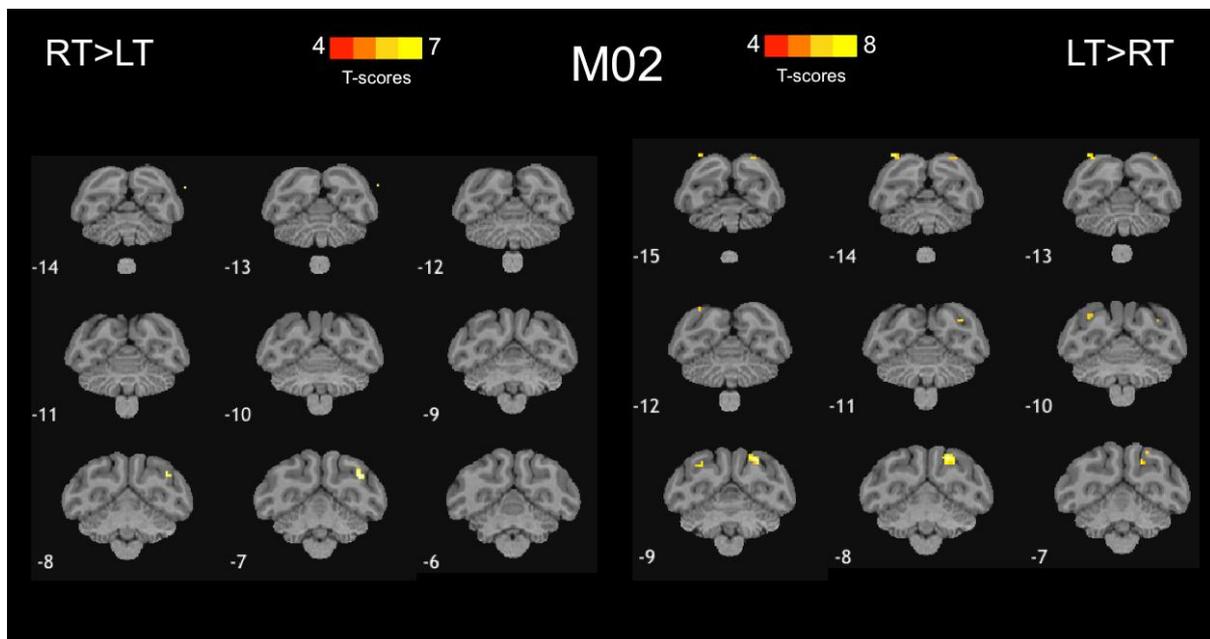


Figure 46. Activations for the contrasts $RT > LT$ on the left and $LT > RT$ on the right for M02. See Figure 45 for details of the legend.

For the ROI analysis in the retinotopic areas, we computed the differences in signal change between RT and LT conditions and performed permutation tests, as described earlier. For both monkeys, we found hemisphere differences but they were not consistent between subjects. For M01, responses in the left hemisphere were stronger for the rightward tilt in V3A, CIP1, and CIP2; and responses in the right hemisphere were stronger for the leftward tilt in V6, V6A, V3, V2, PIP2, and PIP1. For M02, responses were significantly stronger only for the leftward tilt and only in the right hemisphere in V2, CIP1, and CIP2. Given the inconsistency of those results, we cannot conclude much on those activations. However, it can be hypothesised that lateralised activations found in M01 could potentially be driven by crossed disparities present in the contralateral visual hemifield (i.e. a general trend towards stronger activations for crossed disparities in the contralateral hemifield). Indeed, leftward tilts have crossed disparities on the right side of the fixation point whereas rightward tilts have crossed disparities on the left side of the fixation point. Unfortunately, this hypothesis does not seem to hold for M02.

Differential activations with elevation?

In a second attempt to reveal biases towards more frequently encountered 3D orientations, that is stronger activations for the backward slant pattern, we considered the relationship between the elevation and, as a consequence, the disparity values in our stimuli, and the t-scores obtained for our two conditions of interest: the ground-aligned slanted condition (GS) and the not-ground-aligned slanted condition (nGS). We divided our visual field into three parts: A top part, corresponding to the upper part of the visual field ($> 2^\circ$), a middle part corresponding to the central area of the visual field ($\pm 2^\circ$), and a bottom part, corresponding to the lower part of the visual hemifield ($< -2^\circ$). The idea was to address the hypothesis that considering activation differences at the level of the whole visual field was maybe not the best way to reveal potentially very subtle biases towards some 3D orientations. By considering three different sub-regions, we could test whether t-scores obtained for our different slanted conditions would

differ as a function of elevation in the visual field for each of our retinotopically defined visual areas. The elevation could be computed with eccentricity maps obtained from retinotopic measures of the visual field (see Chapter 3, second fMRI study of the team, conducted by Samy Rima).

In order to identify a link between the elevation and t-scores, we determined three different cases that would reveal either a general preference for crossed disparities, or a preference for the ground-aligned slanted surface (GS condition), or a combination of both preferences. The resulting predictions of t-score patterns are summarised in Figure 47, for the three cases that we considered: a preference for crossed disparities, resulting in higher t-scores in the lower visual field for the GS condition and in the upper visual field for the nGS condition; a preference for the GS condition with systematically higher t-scores for the GS condition; a mixed preference pattern with a preference both for crossed disparities and for the GS condition, with higher t-scores for the GS condition in the lower visual field and around-zero t-scores in the upper visual field where the GS and the nGS conditions have opposed disparity values, thus cancelling each other out. Predictions for the central part of the visual field are always the same, with no preference for neither slanted conditions. As a proof of concept, we first investigated the link between elevation and t-score values in the first monkey for which we were confident to find a pattern. For each retinotopic area, we categorised each voxel based on their elevation as belonging to either the upper part, or the lower part, or the central part of the visual field. We then averaged the t-score values of all the voxels comprised in each of the three parts of the visual field. Due to the very small size of areas CIP1 and CIP2, we chose to group them, resulting in a total of 17 retinotopic areas: V1, V2v, V2d, V3v, V3d, V3A, V4v, V4d, MT, MSTv, FST, V4t, CIP, PIP1, PIP2, V6, and V6A.

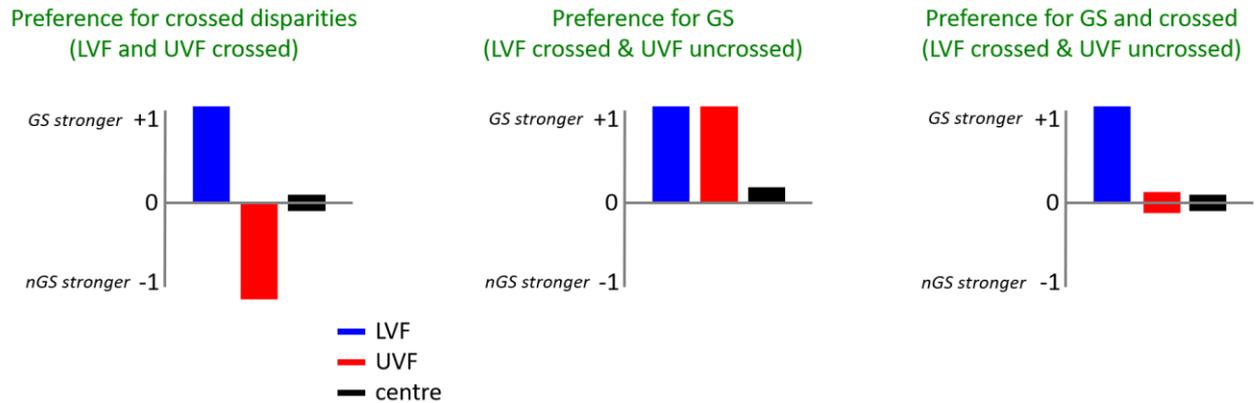


Figure 47. Predictions of t-score patterns as a function of elevation for different disparity preferences.

The graph on the left depicts the known preference for crossed disparities, where t-scores should be higher in the lower visual field of the GS condition and in the upper visual field of the nGS condition, where crossed disparities are present. The central graph represents the case of a preference for ground-aligned surfaces (GS condition) for which case t-score values should be systematically higher for the GS condition regardless of the location in the visual field. The graph on the right illustrates the case of a mixed preference for both the GS condition and crossed disparities, with higher t-scores for the GS condition in the lower visual field and around-zero t-scores in the upper visual field where the GS and the nGS conditions have opposed disparity values, thus cancelling each other out. The lower visual field is represented in blue, the upper visual field in red, and the central part of the visual field in black.

Although statistical analyses are still ongoing (linear mixed models), we found several retinotopic areas with t-score patterns that correspond to a preference for the GS condition in both hemispheres of our first monkey subject. That is the case for V2d, V3d, V4v, V4d, MT, MSTv, V4t, CIP, PIP1, PIP2, V3A, V6A (see Figure 48). Area FST tends to follow a similar pattern in at least one hemisphere, but the very small number of voxels in that area makes those results difficult to interpret. Areas V2v and V6 have opposed hemispheric patterns, whereas both hemispheres of area V1 showed an inverse pattern, suggesting a preference for the nGS condition for this area. It is also worth noting that t-score values in the central area of

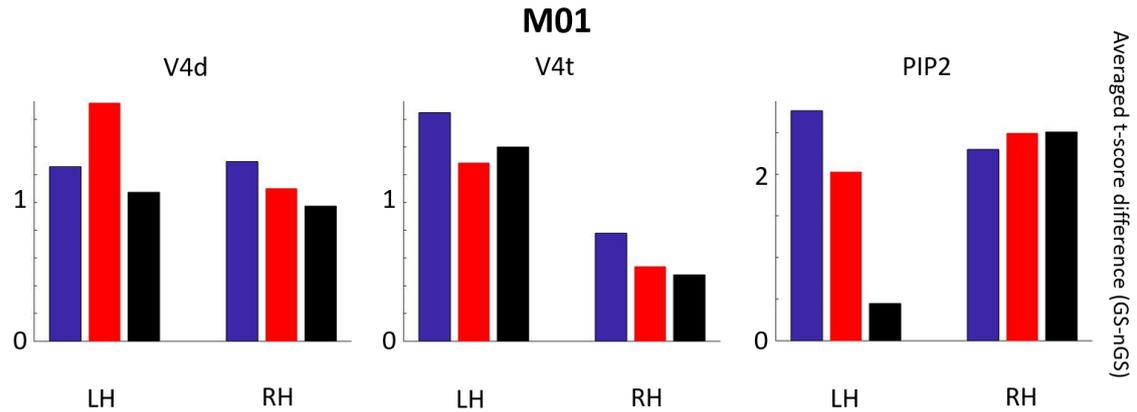


Figure 48. Difference in t-score values between GS and nGS conditions as a function of the elevation for M01. Average t-score values are given for three different parts of the visual field: the lower visual field (blue bars), the upper visual field (red bars), and the central visual field (black bars). Results are shown for three different retinotopic areas: V4d, V4t, and PIP2.

the visual field are higher than predicted and follow the same pattern as t-scores in the lower and upper parts of the visual field.

We applied the same logic to characterise the link between elevation and t-score values for our second macaque subject. For some retinotopic areas, we found a preference for the nGS condition, as it is the case for V1, V2d, V3v, V3d, V4d, MT, PIP2, and V6A (see Figure 49). However, for the other retinotopic areas there is no clear emerging pattern.

These analyses are still very preliminary but may shed some lights about the individual differences we observe at the whole-brain level and at the whole visual field level.

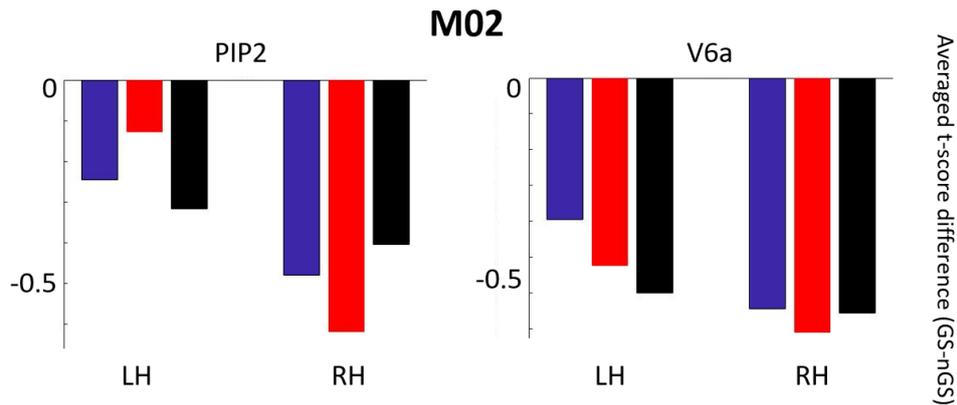


Figure 49. Difference in t-score values between GS and nGS conditions as a function of the elevation for M02. Results are shown for two different retinotopic areas: PIP2 and V6A. See Figure 48 for legend details.

Discussion

The objective of this study was double. First, we aimed at characterising the cortical network responding to correlated disparities. Second, we asked whether orientation biases could exist in that network that would reflect natural statistics. We designed 3D oriented surfaces that were either slanted or tilted and contrasted them with a non-correlated stimulus. All the 3D oriented surfaces shared the same disparity gradients and had the same orientation degree (the angle had the same value). They differed either by their orientation (slant vs. tilt) or by the sign of their orientation, that is, the slanted surfaces were slanted either backward or forward, and the tilted surfaces were tilted either leftward or rightward. More importantly, when viewed monocularly, all correlated conditions did not contain any 3D information and were thus identical to the uncorrelated condition. We recorded BOLD responses at the whole-brain level in two awake, behaving macaque monkeys. We used a block design and kept the slanted and the tilted surfaces apart in different types of run, conducting independent analyses on each type of runs.

We first investigated the cortical network involved in processing correlated disparities. Performing a retinotopic analysis, we found seven areas that were systematically responding more strongly to the correlated stimuli than to the uncorrelated ones: CIP1, CIP2, PIP2, V3A, V6A, MT, and MSTv. Those activations were similar across hemispheres and across run types. These results are in agreement with the literature in macaque. For instance, neurons in the MT area have been found to be selective to correlated disparity (Nguyenkim and DeAngelis, 2003). In a neuroimaging study, Tsao and collaborators (2003) found that V3A and CIP were strongly activated by correlated disparities, whilst PIP area was found to respond to depth structures (Durand et al., 2007). Finally, although not clearly demonstrated, the sensitivity of V6A to binocular disparities has been suggested (Hadjidimitrakis et al., 2011).

Interestingly, areas V3, V4, V4t, and FST were also found to respond more strongly to the correlated stimuli in 4 hemispheres but only in the ‘slant’ type of runs. Those areas are also known to contain disparity-sensitive neurons (Felleman and Van Essen, 1987; Watanabe et al., 2002). The fact that they are responding more to the correlated condition only for the slant runs might be related to a general preference for a vertical gradient of disparities rather than a horizontal one, in other words a stronger preference for slanted surfaces than for tilted surfaces. However, we do not have a proper control to elucidate this eventuality, which deserves further investigation to understand what triggers this activation difference between both types of 3D orientation.

This remark brings us to the second question we were interested in: Is the cortical network involved in processing correlated disparities biased toward more frequently encountered surface orientations? We expected cortical activations for the slanted surfaces to be stronger for the GS condition than for the nGS condition, given that the GS condition reflects spatial regularities (e.g., the orientation of the ground), but no difference between the tilted conditions. We found such a pattern for one subject with three areas in the PIP cluster (PIP1, PIP2, and

CIP1) responding significantly more to the slant aligned with the ground (GS) than the non-aligned one (nGS). However, those results were not replicated in the second subject for whom no clear pattern emerged. Regarding the tilted surfaces, we did not find any consistent pattern of activations between hemispheres or between subjects, although we can note some lateralised activations in one animal that may reflect a bias toward crossed disparities. Overall, the origin of those inter-subject differences is not quite clear and requires further work on the data. Investigating an effect of elevation, we were hoping to find differential activations between the upper and the lower hemifield, reflecting a bias that could have been diminished by averaging BOLD responses across the whole stimulus area. Our preliminary analyses revealed such differences for the monkey subject that was already showing a bias at the whole stimulus level but failed to find similar results in the second monkey subject. It is possible that differences in activations for both slanted surfaces are too subtle to be revealed with traditional univariate analyses. We are now considering conducting multivariate pattern analyses (MVPA) to investigate whether we could differentiate the cortical activity in response to our stimuli. In other words, are the activation patterns in response to both conditions different enough to tell them apart?

Conclusion

This fMRI study confirmed that the cortical network involved in processing correlated disparities is quite spread in macaque and further suggests that area V6A also processes binocular disparity. We could show in one subject that some of the areas of the disparity network are also biased towards more frequently encountered 3D orientation, namely backward slants. We plan to extend our analysis to check whether those biases could also be revealed in the second subject.

Third study: Measure of the horopter in human and in macaque

Introduction

The mechanism underlying stereopsis, matching corresponding points coming from the two retinæ, is found in lots of different species with two aligned eyes on the frontal plane (Nityananda & Read, 2017). It has been of great interest to try to understand how the matching computation was done, given the infinite number of possibilities for having two points in each retinal image correspond to the same point in space and give rise to depth perception. However, some points have a special physiological relationship that facilitates binocular correspondence. For instance, two points, one in each retina, that are simultaneously stimulated by an object, then resulting in the percept that this object comes from the same place in space, will be defined as corresponding points. Those corresponding points form the horopter¹⁸, which is the space area where binocular fusion and depth perception will be the most precise (i.e., where stereoacuity is the highest). This horopter can be determined geometrically (Ibn al-Haytham a.k.a Alhazen, 11th c. and Aguilonius, 17th c. in Howard, 1996) and empirically. As shown on Figure 50, the empirical horopter is slightly different than the geometric one. On the horizontal axis, the empirically found Hering-Hillebrand deviation from the geometrically defined Vieth-Müller circle can be explained in terms of anatomical differences, potentially due to the fact that photoreceptors are not evenly distributed in the retina (Franck, 1905; Fischer, 1924; Halldèn, 1956). On the other hand, the vertical deviation of the horopter, the shear of Helmholtz (Helmholtz, 1910/1962; later confirmed by Nakayama et al. 1977), cannot be explained by anatomical factors. Several developmental or evolutionary hypotheses have been made regarding this vertical tilt.

¹⁸ It is worth noting that the horopter depends on the point of fixation and that different distance of fixation will have a direct effect on the position of the corresponding points, and thus on the shape of the horopter.

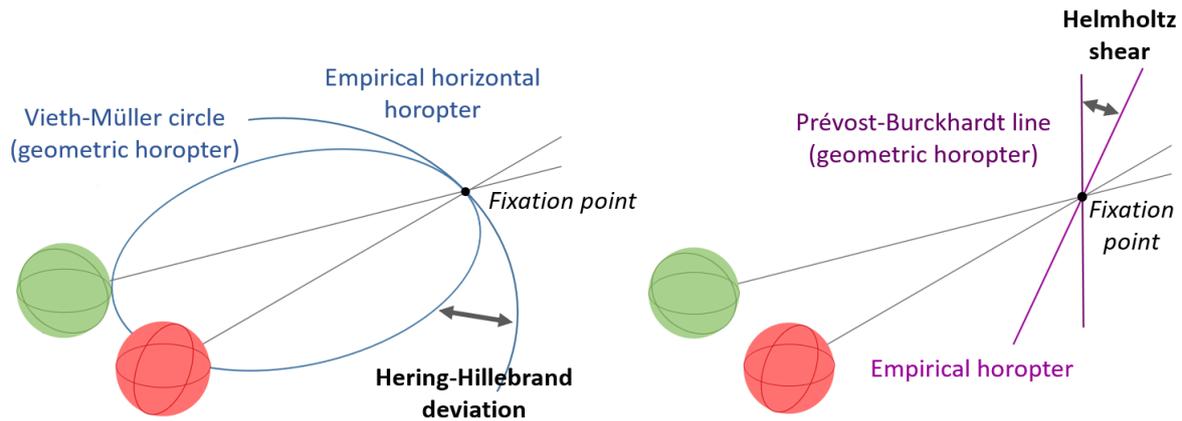


Figure 50. Geometric and empirical horopters.

The horizontal horopters are represented on the right panel: The geometric one or Vieth-Müller circle passes through the two eyes centres (the green and red globes on the picture) and the fixation point whereas the empirical one or Hering-Hillebrand deviation is more ellipsoidal. The vertical horopters are shown on the left panel: The Prévost-Burckhardt line is the geometric horopter and the Helmholtz shear represents the tilted empirical horopter. Importantly, corresponding points that define the horopter are always depending on the fixation point.

Aligned with the ground, the vertical tilt could result from an adaptation to bring the ground in the binocular fusion area, where stereoacuity is the finest (Helmholtz, 1910/1962; Nakayama, 1977; Tyler, 1991). However, Schreiber and collaborators (2008) have shown that the surface of the horopter is more complex than a tilted plane and is not always aligned with the ground but rather changes with fixation distances. It has also been suggested that this tilt varies as a function of the eyes' height and of the interocular distance of the observer (Cooper and Pettigrew, 1979; Cooper et al., 2013). Despite being seemingly convincing in two different species, the burrowing owl and the domestic cat (Cooper and Pettigrew, 1979), the relation between the vertical tilt and the eyes' height and the interocular distance is far less conclusive in human (Cooper et al., 2013): No difference has been found between taller than average and shorter than average subjects (Cooper et al., 2013), implying that the shear could develop at the species level ('adaptive') but not at the individual level ('not adaptable'). A more recent

hypothesis has emerged regarding the origin of the vertical tilt, stating that the shear reflects an adaptation to natural statistics or disparity distributions in visual scenes. For instance, when looking at a fixation point, objects below tend to be closer to the observer and thus to have crossed disparity values, whereas object above the fixation point tend to be further apart, with uncrossed disparity values. This frequently encountered gradient of disparity could explain the shape of the tilted horopter. The team of Martin S. Banks at UC Berkeley has been working on this hypothesis with recent advances. In a very elegant study (Sprague et al., 2015), they looked at the relationship between 3D geometrical properties (e.g. distances) present in a visual scene that participants were freely exploring and the binocular position of their eyes, using an eye-and-scene tracking system. With this setup, they could determine the natural distribution of binocular disparities present in 4 different types of visual scenes and they showed that in the outdoor scene (outside walk) or in the scene involving tool manipulation (making a sandwich), the distribution of natural disparities follows the top-back disparity gradient, and thus the tilt of the horopter, with more crossed disparities in the lower hemifield and more uncrossed disparities in the upper visual field. Interestingly, scenes involving social interactions (ordering a coffee) or walking inside of a building yielded to near-zero disparities aligned on the same plane; no big variations were found across the visual space for those activities. The shear of the vertical horopter could thus originate from the neural encoding of the statistics encountered by the visual system during its development to allocate resources in a more efficient¹⁹ way. A small meta-analysis of neurophysiological recordings in the early visual area V1 of the macaque (Sprague et al., 2015) revealed that more neurons with receptive fields (RF) in the

¹⁹ Here, the term ‘efficient’ refers more to a frequentist learning of visual statistics than to an optimal way of coding those statistics. This coding could also be stated in a Bayesian framework, with more weight (resources) being allocated to more likely stimuli. Both views are not incompatible as likelihoods could be used at a higher level to disambiguate ambiguous stimuli.

lower visual hemifield were tuned to crossed disparities, whereas more neurons with RF in the upper visual hemifield were tuned to uncrossed disparities (Figure 51). The disparity tuning of V1 neurons thus reflects natural statistics and is aligned with the tilt of the vertical horopter. Further supporting this biased encoding idea, a computational study conducted within our team (Chauhan et al., 2018) showed that such results could be mimicked in a neural network using a simple Hebbian learning rule. In agreement with these results, a neuroimaging study conducted in human (Nasr and Tootell, 2016) with a high spatial resolution (7 Tesla) showed that within V2 and V3 areas the BOLD signal in subregions coding for the lower hemifield were stronger when disparity values were crossed than uncrossed, and that the opposite was true for subregions coding for the upper visual field.

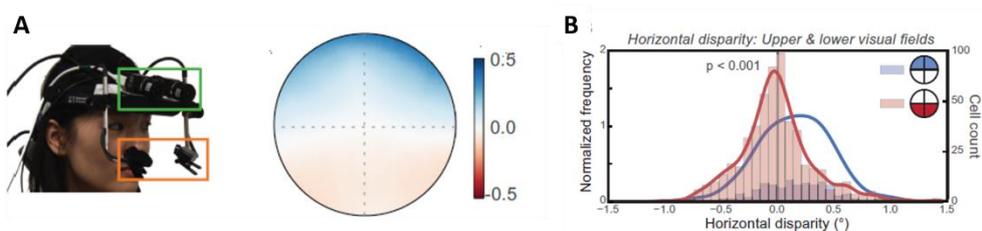


Figure 51. A) Eye-and-scene tracking device and weighted combinations of the four different activities (median horizontal disparities). B) Distributions of preferred horizontal disparity grouped by upper and lower visual field. The histograms represent the number of cells observed with each preferred disparity ($n = 973$). The curves represent those histograms normalized to constant area. The preferred disparities of neurons from the upper visual field (blue) are biased toward uncrossed disparities, whereas the preferences of neurons from the lower visual field (red) are biased toward crossed disparities. Adapted from Sprague et al. (2015).

Although neurophysiological data in different species (in the burrowing owl, the domestic cat, and the macaque monkey) showed that disparity tuning in neurons is either aligned with the

ground²⁰, or more generally biased towards natural statistics (Cooper and Pettigrew, 1979; Sprague et al., 2015), we do not know how natural statistics impact depth perception in those species. Indeed, the empirical horopter has never been measured in an awake, behaving non-human animal. If the shear of Helmholtz truly reflects an encoding of natural statistics, we can expect interspecies differences. For instance, humans and macaques do not evolve in the same natural environment, do not have a similar eye height, and have different visual exploration strategies (Einhäuser et al., 2006). It is thus possible that the natural statistics they are exposed to are different, and this should influence the value of the shear angle. To further investigate this hypothesis, we propose to measure the empirical horopter in macaque, using a similar experimental setup as was used for human subjects (Schreiber et al., 2008; Cooper et al., 2011). Within the framework of a collaboration with Martin S. Banks, we first replicated the experimental setup and recorded the corresponding points on the vertical meridian from eight human participants. In a second time, we adapted the task to efficiently measure the horopter of one macaque subject.

²⁰ In their study, Cooper and Pettigrew (1979) indirectly measured the horopter by mapping the receptive field of binocular neurons in the anaesthetised domestic cat and burrowing owl. They found that the vertical horopter in those two species was tilted away, as in human. The average shear values were 10.0° ($n=4$) and 10.6° ($\pm 4^\circ$) for the domestic cat and the burrowing owl, respectively.

Experimental setups

A- Measurement in human participants

Participants

Eight participants (age: 24-30 years old) took part in the experiment (2 women and 6 men). They had normal-to-corrected vision and the presence of stereopsis was assessed using the online stereoacuity test provided by McGill University (<http://3d.mcgill.ca/>). All participants were informed about the objectives of the task, and agreed on the formalities by signing a written consent form. They received monetary compensation for their participation (30€ in total). The study protocol was submitted to the local ethics committee (Université Fédérale Toulouse Midi-Pyrénées, 2019-160) to make sure it met the European legislation standards

Experimental setup

The experiment took place at the CerCo research centre in a dedicated experimental box. Participants sat in front of a computer screen (screen resolution: 1920*1080 pixels, refresh rate: 60Hz, screen width: 53cm) at a distance of 50 cm. They wore red-green anaglyph filters and were requested to bite a bite bar in order to stabilise their fixation. A keyboard placed in front of participants was used to both initiate a trial (space bar) and give their response (arrow keys). Before starting the experiment, the interocular distance of each participant was measured using a digital pupillometer. Their eyes' height was also measured. Participants' characteristics are indicated in the Table 4.

Stimuli and procedure

In order for participants to maintain their fixation aligned and constant across trials, the fixation cross was specifically designed to help them fuse information coming from both eyes. It consisted of two radial segments (0.15° long) that were presented in each eye and became a cross only when the information coming from each eye was fused. Participants were instructed

to start a trial only when they saw the segments of the cross aligned. Stimuli were two 28 pixel-long bars (0.73°), one green and one red, each presented in one eye through the anaglyph filters. The two bars were successively displayed during 50 ms and separated by a 70ms interstimulus interval. They were presented on the vertical meridian at 6 different eccentricities ($\pm 7^\circ$; $\pm 3^\circ$; $\pm 1^\circ$) to measure corresponding points for one position in space: the fixation point at the centre of the screen (0° , 0°). The rapid succession of the two bars created an apparent motion and participants had to report their perceived direction of motion in a forced choice manner, that is, either leftward or rightward (see Figure 52). The bar presentation order at different eccentricities was randomised to prevent anticipatory eye movements. The distance between bars varied at each trial following a 0.5-up/1-down staircase procedure in order to reach the point of subjective equality, which represents the bar separation for which bars are perceived as superimposed. After reaching 14 reversals at each measured eccentricity, the experimental session stopped. Participants underwent one session per day and each session lasted about 20-30 minutes. From one day to the other, we changed the filter colour in front of the left eye (either green or red) and averaged them to minimise a potential effect of the filter colour. We also estimate crosstalk in an independent session (i.e. the amount of coloured light that would not be properly filtered) by asking participants to report when they could no longer detect a bar of the size of our stimuli by varying the colour intensity (RGB values), for each filter colour and for each eye. We used those values in the main experiment in an individual manner to make sure each bar was perceived by one eye only, thus limiting the influence of crosstalk.

On top of the measure of horizontal disparities, we measured participants' cyclovergence, that is disconjugate rotative movements of the eyes that could alter our measurements. The experimental protocol was very similar to the measurement of the horopter except that the bars were presented horizontally on the horizontal meridian at identical eccentricities and that

participants had to respond whether they perceived the motion direction as upward or downward.

Prior to the recording sessions, all participants underwent 8 training sessions to make sure they understand the task properly and get used to the experimental setup. The number of sessions that followed the training part varied among participants with a minimum of 4 sessions and a maximum of 18, as indicated in Table 4.

The experiment was displayed with the EventIDE software (Okazolab®) and data analysis was carried with Matlab®.

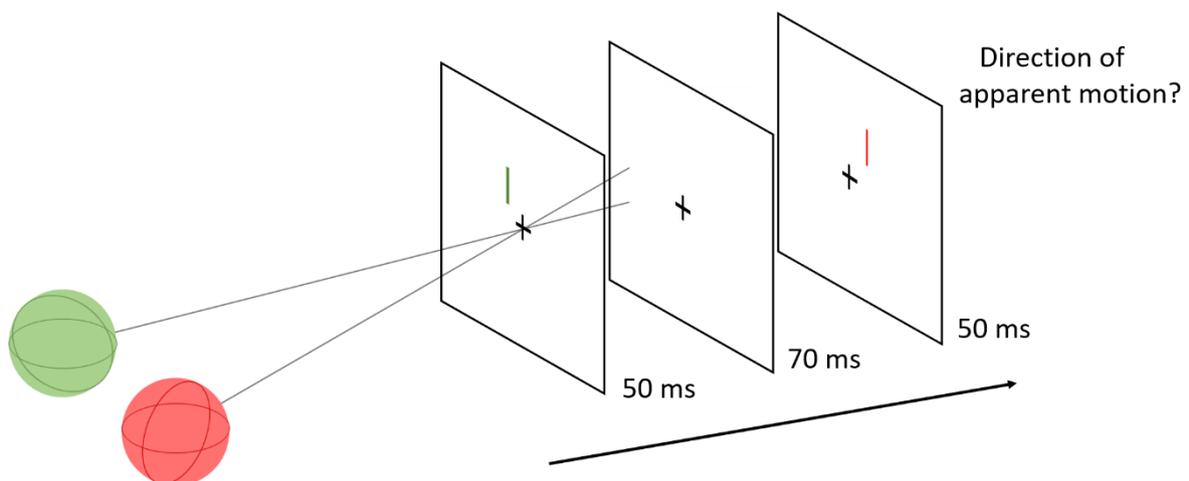


Figure 52. Apparent motion paradigm for the measurement of corresponding points.

Following the rapid succession of two bars of different colour shown each to a different eye (here, the green bar is seen through the green filter, i.e. by the left eye), participants had to report the direction of the apparent motion created by the bar succession. Bars were presented at different eccentricities on the vertical meridian and the bar separation varied across trials following an adaptive staircase procedure. Adapted from Schreiber et al. (2008).

Table 4: Participants characteristics for the measurement of corresponding points.

Subjects	Stereoacuity (arcmin)	IOD (cm)	Eyes' height (cm)	Optimal shear angle (degrees)	Number of sessions ²¹	Shear angle (degrees) <i>4 sessions</i>	Shear angle (degrees) <i>All sessions</i>
S1	≤ 76.6	6.50	153	2.30	18	3.9573	4.1241
S2	≤ 90.2	6.55	166.5	2.25	12	7.2440	7.3698
S3	≤ 84.8	6.30	171.5	2.10	14	4.5809	3.4289
S4	≤ 78.4	6.70	163.5	2.42	12	4.4070	4.4498
S5	≤ 36.7	6.15	157.5	2.24	4	5.7926	-
S6	≤ 65.3	6.90	169.5	2.26	4	4.0396	-
S7	≤ 40.1	6.30	156	2.21	6	2.5355	2.6623
S8	≤ 89.1	6.30	169	2.14	4	5.2537	-

B- Measurement in one Macaque

Adapting the experiment for our macaque subject was not straightforward and required some specific arrangements. Indeed, training a macaque to learn a task necessitates to reward it based on its performance. However, since apparent motion is subjective, a correct or an incorrect response cannot be determined. The training thus consisted of monocular stimuli, with the two bars presented in the same colour to be seen by the same eye, for which an objective response could be defined. To make the task easier to learn, we used constant stimuli instead of an adaptive staircase. Once the macaque had learnt the task (i.e. when training performance was above 90% for 2 consecutive weeks), binocular stimuli were introduced for a small number of trials only (10% of the trials) and were rewarded in a random manner. This method had already

²¹ Here, one session includes two days of recording: one day for recording corresponding points (vertical bars on the vertical meridian) and the second day for recording cyclovergence (horizontal bars on the horizontal meridian). Training sessions were not included in the total number of sessions.

been used in a study involving binocular vision in the macaque (Logothetis and Schall, 1990) and proved to work well.

Like humans, our macaque subject viewed the stimuli through red-green anaglyph filters. Bars were presented at the same 6 eccentricities ($\pm 7^\circ$; $\pm 3^\circ$; $\pm 1^\circ$) on the vertical meridian and were additionally presented around the fixation point (a window of 1° at the centre of the screen). Since the method of constant stimuli was used, six different values for the bar separation (1.16° ; 0.87° ; 0.58° ; 0.17° ; 0.06° ; 0.03°) were pre-determined in order to estimate the point of subjective equality as precisely as possible. Fixation was assessed with an eye-tracker and responses were given in the form of a saccade, either leftward or rightward. To ensure steady fixation, a trial would start only when fixation was maintained during 3 seconds within a circle of 1° around the fixation point. We displayed the stimuli, controlled for the delivery of the liquid reward and the fixation performance using the EventIDE software (Okazolab®).

In total, 108,766 trials were recorded, of which 11,857 were binocular trials. Only those binocular trials were considered for further analyses. Psychometric fitting and statistical analyses were carried using R in the RStudio environment²². Figure 53 shows a prediction of the psychometric curves. We expect the points measured in the lower hemifield to be shifted towards crossed (negative) disparities and the points measured in the upper visual hemifield to be shifted toward uncrossed (positive) disparities.

²² R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
RStudio Team (2019). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.

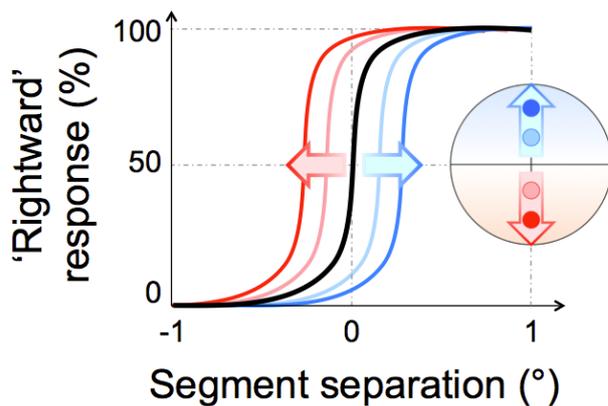


Figure 53. Predicted psychometric curves.

The black curve represents the curve we would obtain if corresponding points have zero disparity or if the stimulus presentation is monocular. Coloured curves represent our predictions for the four points measured in the visual field schematised by a disk on the right. We expect points located in the lower field (red curves) to be shifted toward crossed disparities (negative segment separation values) and points located in the upper field (blue curves) to be shifted toward uncrossed disparities (positive segment separation values).

Results

A- Measurement in human participants

Data were recently collected by a master's student and are still being analysed. Only preliminary results will be presented here.

To estimate the location of corresponding points for our observers, we first averaged the bar separation values from the last 10 reversals of the first 4 test sessions that all participants underwent and plotted them as a function of eccentricity for each eye separately as shown for one participant on the upper panel of Figure 54. Data points were then fitted using a weighted linear regression. The angle between the two regression lines represents our estimate of the shear angle. The same procedure was followed for the cyclovergence data (medium panel of Figure 54) and the obtained angle was subtracted to the shear angle to take any rotational eye

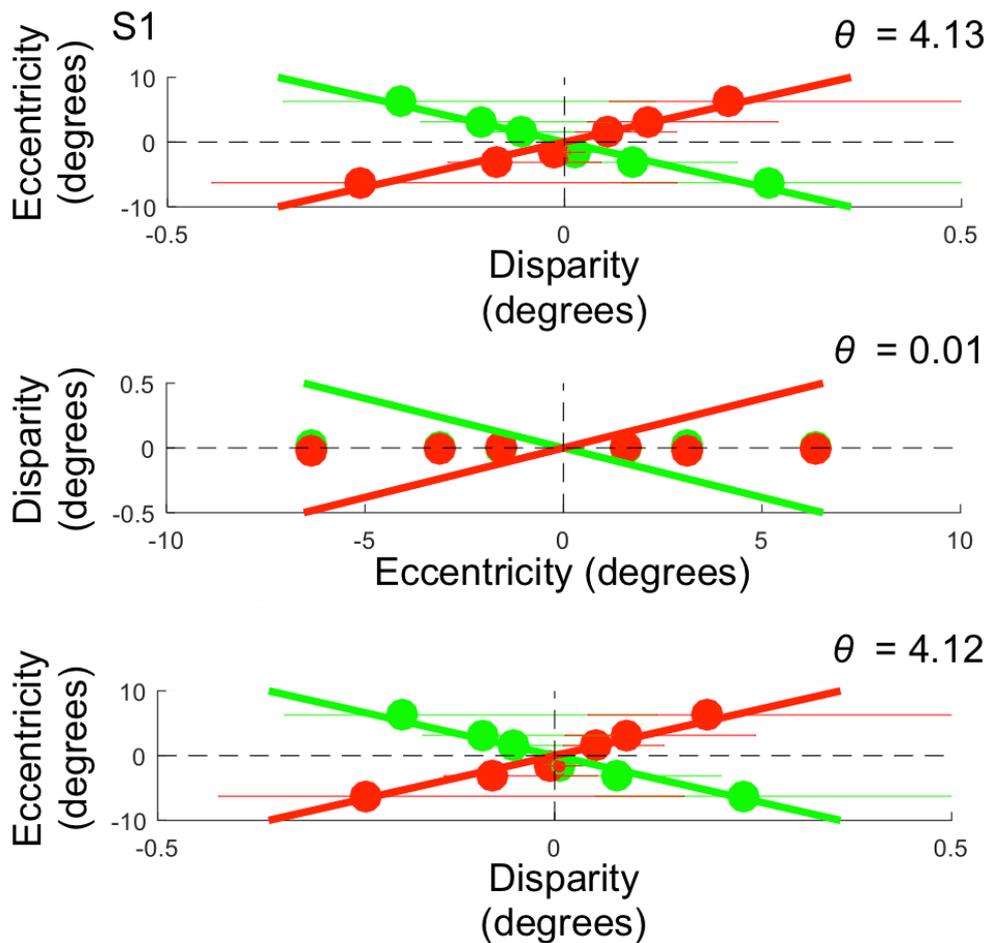


Figure 54. Data from Observer 1.

1) The upper panel shows the data obtained with the measurements of the corresponding points at 6 different eccentricities on the vertical meridian. The abscissa is the bar separation in degrees and the ordinate the eccentricity at which points were presented. Green and red colours indicate the measured location in the left and right eye, respectively. The angle between the regression lines determines the shear angle. 2) The middle panel shows data collected during cyclovergence recordings. The abscissa indicates the eccentricity at which points were presented and the ordinate the bar separation in degrees. 3) The lower panel represents the corrected shear angle after subtracting the cyclovergence data to the shear data.

movement into account. Positions of corresponding points after cyclovergence were re-estimated (lower panel of Figure 54) and are further displayed on Figure 55 for the 3 other participants who were involved in at least 12 sessions. Those results are in good agreement

with the Cooper and colleagues' study (2011) that we replicated: Crossed disparities tend to be in the lower visual hemifield and uncrossed disparities in the upper visual hemifield. Because those participants underwent more sessions, we had a first look at the robustness of the corresponding point measurement. The Table 4 shows the difference in the computed shear angle after correction after 4 and after all the sessions. From a qualitative perspective, the estimated shear angles do not change much when we add more recording sessions. However, the eight training sessions participants performed before were much needed as the value of the shear angle generally stabilised around the 6-7th session (data not shown). Future analyses will aim at addressing robustness from a quantitative point of view.

B- Measurement in one macaque

As indicated earlier, in macaque, only binocular trials (i.e., 11,587 trials) were further analysed and represented 10% of the total trial number. Data at each measured interval (distance between the two bars) was averaged and later fit with a binomial psychometric function at each eccentricity separately. The 95% confidence intervals were also computed. Figure 56 shows the corresponding curves. The points of subjective equality (PSE) are indicated for each eccentricity and were defined as the bar separation value for which responses were at the chance level: 50% of the responses were "rightward". This PSE indicates the value by which corresponding points are shifted, that is the disparity value for which those points are perceived as superimposed. As shown on Figure 56, disparity values of the measured corresponding points below the fixation point, in the lower hemifield are crossed (negative values), which corresponds to objects being closer to the observer. On the opposite, disparity values in the upper visual field, above the fixation point, tend to be uncrossed (positive values) except for

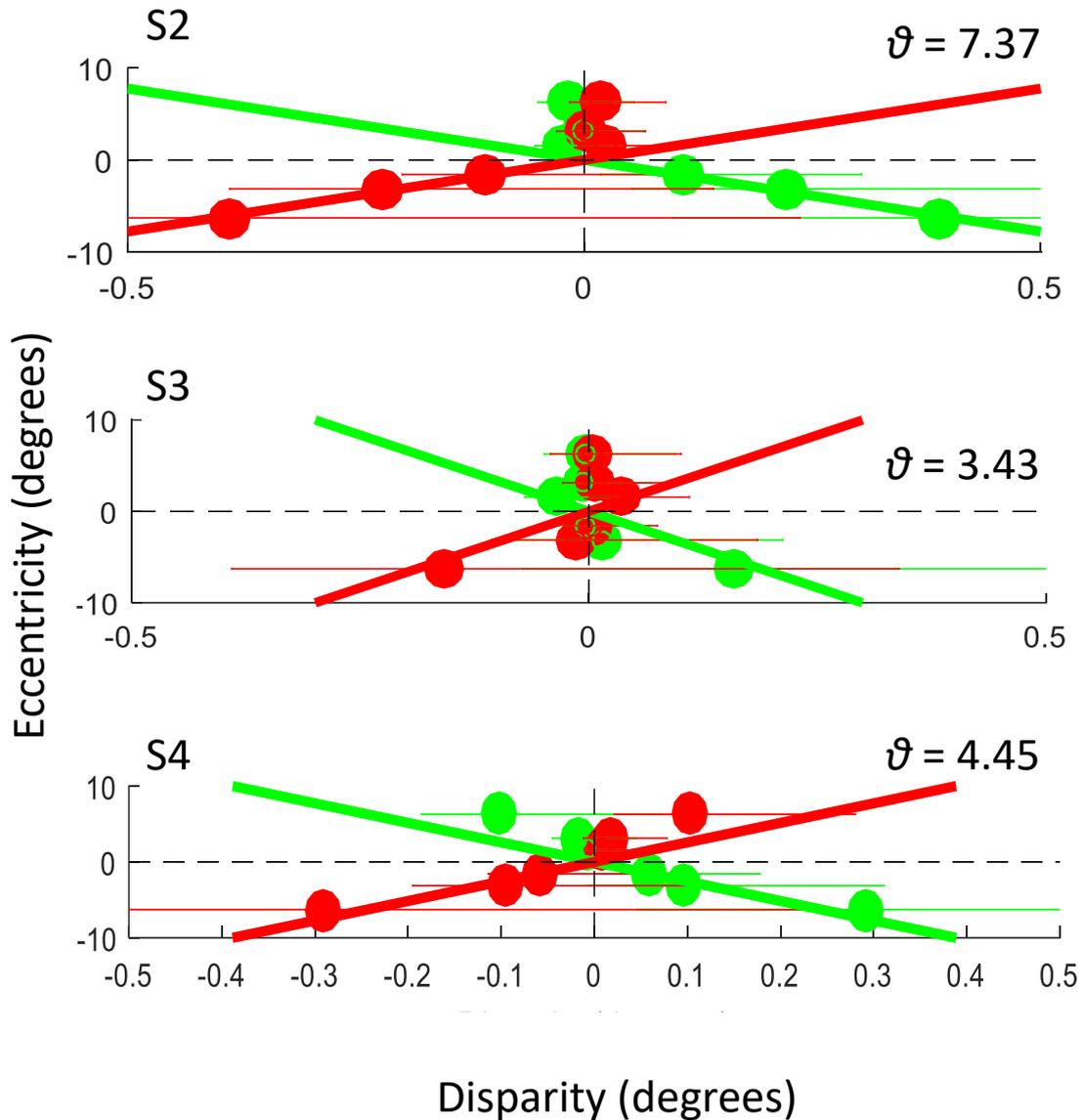


Figure 55. Location of corresponding points for subjects 2 to 4. Each graph represents the corrected shear angle for each of the three observers. For legend details, see Figure 54.

one eccentricity value. This corresponds well to the pattern found in humans (Cooper et al., 2011). To facilitate comparison with human, data was plotted the same way as for humans in the original study as shown on Figure 57.

Unfortunately, given the long duration to record data, the cyclovergence recording is still ongoing and will have to be taken into account later. However, based on the small values

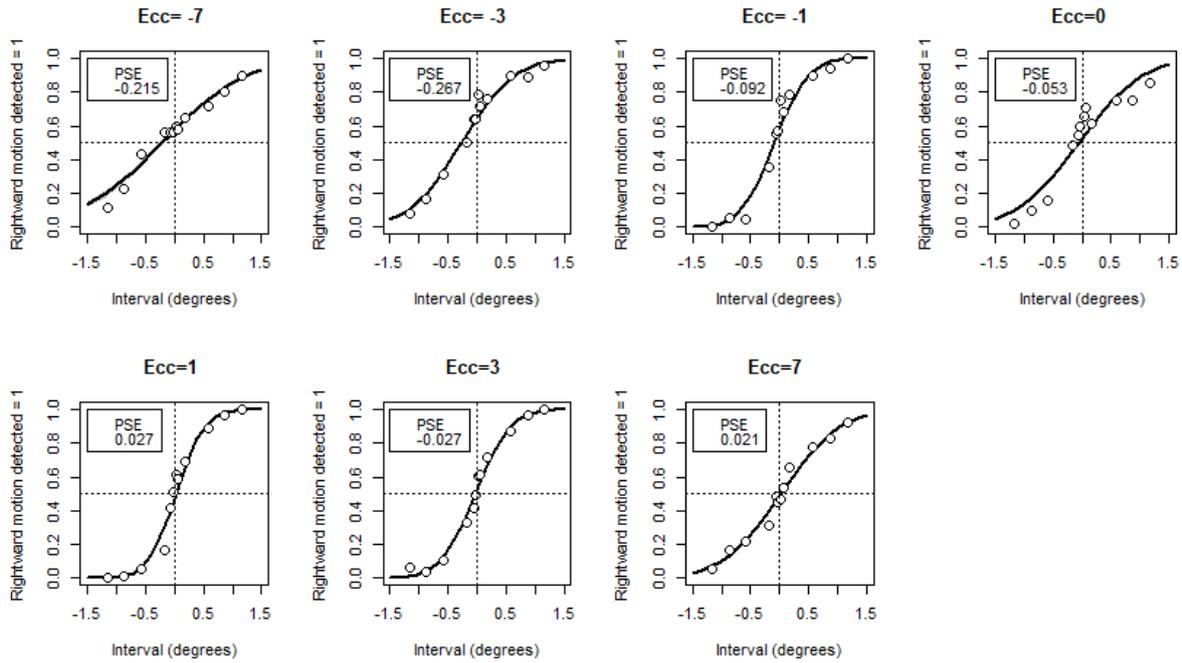


Figure 56. Fitted curves for the different eccentricities at which corresponding points were measured. Abscissa indicate the bar separation or interval in degrees and ordinates indicate the percentage of motion detected as rightward (percentage of rightward responses). Framed values indicate the point of subjective equality which is the bar separation value for which the bars were perceived as superimposed and thus motion was as often perceived as being leftward or rightward, at the chance level. White dots represent the average data point at the measured interval.

usually found in humans (less than 1°) in our own data and in others' (Cooper et al., 2011), we expect the cyclovergence to be rather small and we can consider that, given the small amplitude of cyclovergence effects, our current measurements give a fair idea of the final value of the vertical tilt of our monkey subject.

To further assess the validity of our measurements, data were split into two and the PSE analysis was carried on each subpart separately. The computed shear angles and PSE fit are shown on Figure 58 and on Figure 59 and Figure 60, respectively. They demonstrate that our estimates are robust.

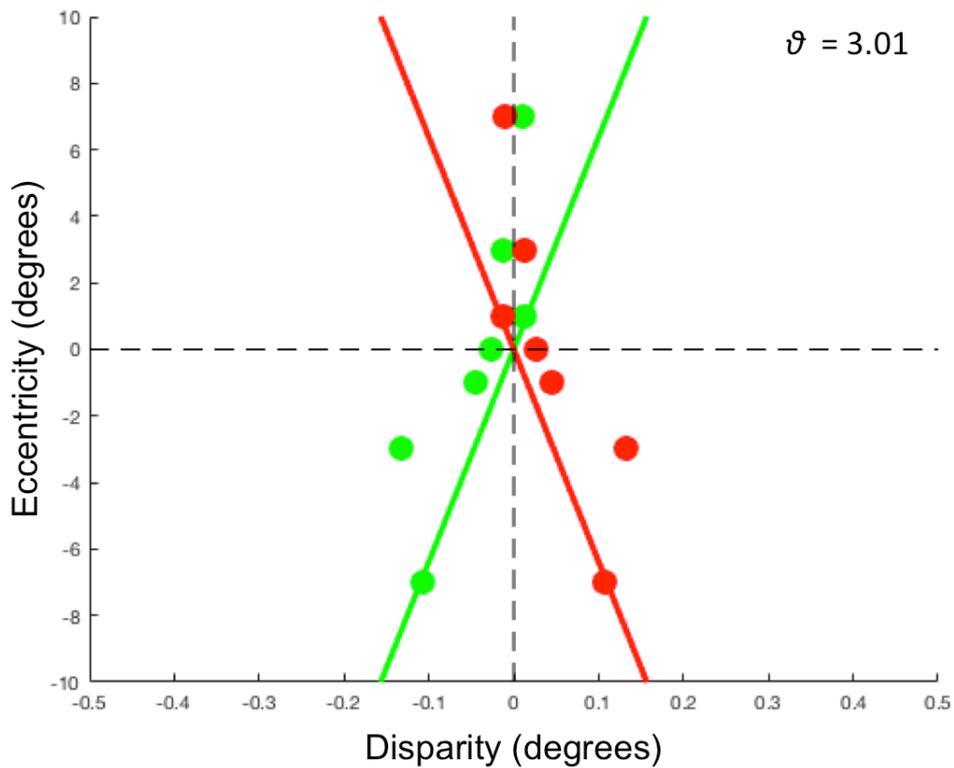


Figure 57. Non-corrected locations of corresponding points for the macaque subject. The angle between regression lines gives the shear angle value.

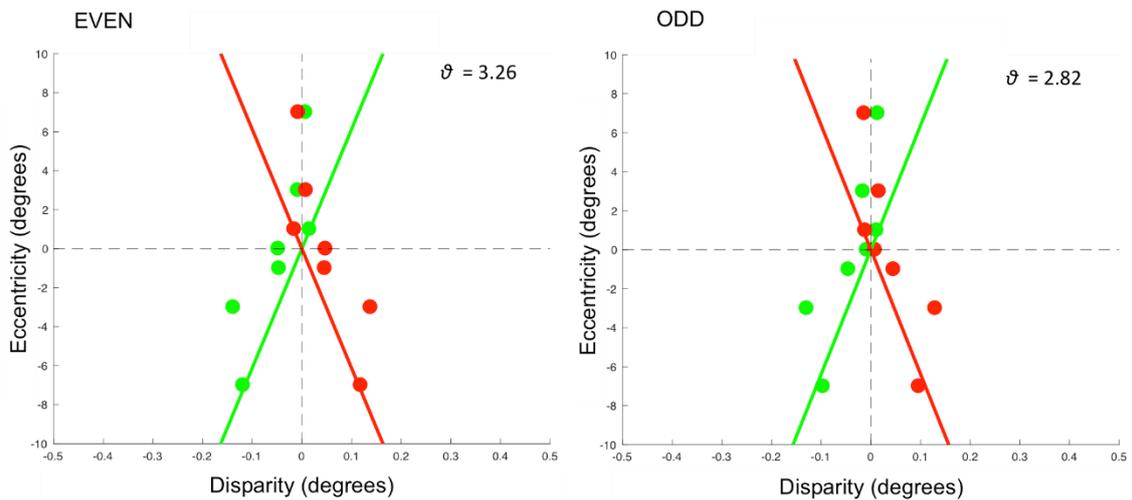


Figure 58. Non-corrected locations of corresponding points for the macaque subject for even (left panel) and odd (right panel) trials. The angle between regression lines gives the shear angle value.

EVEN

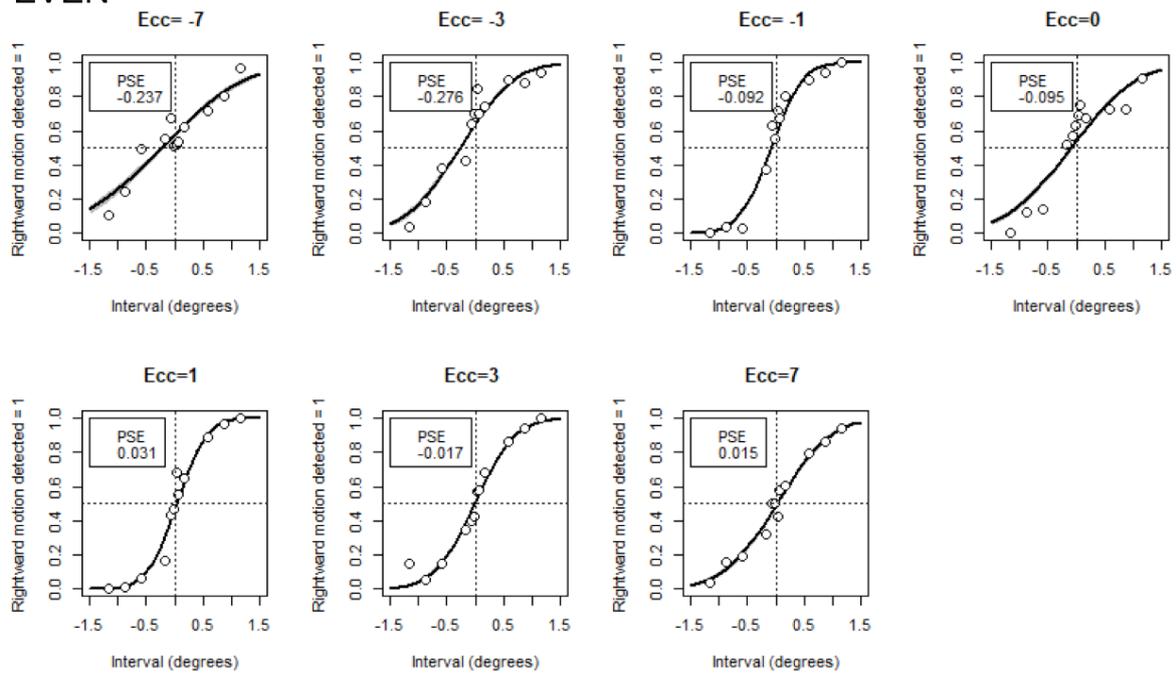


Figure 59. Fitted curves for the different eccentricities at which corresponding points were measured for the even trials. For legend details, see Figure 56.

ODD

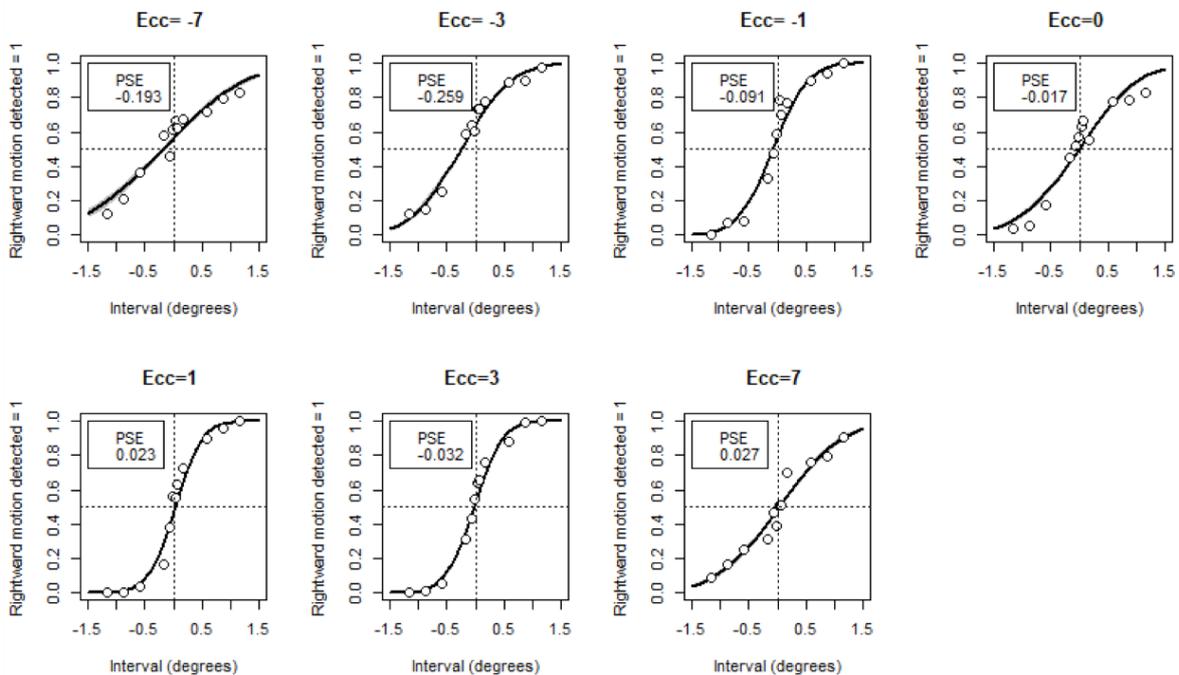


Figure 60. Fitted curves for the different eccentricities at which corresponding points were measured for the odd trials. For legend details, see Figure 56.

C- Interspecies comparisons

It is possible to estimate the optimal shear angle θ for one species by computing the ratio between the average interocular distance (IOD) and the average eyes' height with the following equation:

$$\theta = 2 \tan^{-1} \frac{I}{2h}$$

Where I is the interocular distance of the average population and h is the average eyes' height (Cooper et al., 2011).

Using this equation, we could compute the optimal shear angle for human and for macaque monkey (*Macaca mulatta*). Furthermore, based on the data from Cooper and Pettigrew's study (1979), we could also estimate the optimal shear angle in the domestic cat and the burrowing owl and compare those values with the angle they found by measuring corresponding receptive fields in those two species. The results are reported in Table 5. We found that the ratios computed for the cat and the owl are quite similar to the neurophysiological recordings, and quite similar between the two species (i.e. around 10 degrees). Interestingly, comparing the human ratio with the macaque ratio revealed that the difference between those two species is about 2 degrees. It might be stated that this goes against the hypothesis that they should be more different given that difference in visual environment or physical characteristics could lead to very different natural statistics between the two species.

As a proof of concept, we compared our experimental values collected in human subjects and in our monkey subject to the estimated optimal shear angles. We first averaged the human values to compare it with the single value we have for our macaque subject. The average human shear angle was 4.73° , which is slightly bigger than the shear angle value of the monkey, that is, 3.1° . This result is somehow in contrast with the optimal shear angle values estimated for both species, where the macaque angle should be closer to 4° and the human angle closer to 2° .

Nonetheless, this seems to reflect the high variability found in human observers within and between studies that experimentally measured the shear angle, as reported by Cooper and collaborators (2011).

Table 5. Interspecies comparisons of the optimal shear angles.

<i>Species</i>	<i>IOD</i>	<i>Eyes' height</i>	<i>Optimal shear angle</i>
<i>Homo sapiens</i> (human)	6.5	160	2.33
<i>Macaca mulatta</i> (Macaque monkey)	3.25	42	4.43
<i>Felix catus</i> (domestic cat)	3.8 ⁷	22 ⁷	9.87
<i>Speotyto cunicularia</i> (burrowing owl)	2.8 ⁷	15 ²³	10.66

Discussion

In this study we proposed to measure the corresponding points in a macaque subject to investigate whether its vertical horopter is biased towards natural statistics. Within the framework of collaboration with Marty Banks' team (UC Berkeley), which measured the empirical horopter in human observers, we adapted their experimental setup in our lab. We replicated their measures in several human observers (n=8) and measured the horopter of one macaque subject along the vertical meridian. Although measurements of the shear angle in our macaque subject have not been fully completed (the cyclovergence measurement is still ongoing), we already have a good overview of its value. We showed for the first time that the empirical horopter of a macaque subject deviates from the geometrical horopter and might also

²³ Those values were those reported in the neurophysiological study of Cooper and Pettigrew (1979).

reflect natural statistics. The disparity values of the measured corresponding points were crossed when corresponding points were located below the fixation point, and uncrossed when they were above the fixation cross. We found the same pattern in our human observers as expected from the literature (Schreiber et al., 2008; Cooper et al., 2011).

Comparing the shear angle value obtained from those measurements in our macaque subject with the shear values obtained in our human observers, we found that the averaged shear angle values between both species differed only by 2 degrees. Given the variability in the human shear values we obtained, we cannot consider this difference of 2 degrees as a significant one. We can thus conclude that we did not find the interspecies difference we were expecting. This may go against the hypothesis that, if visual experience in individuals of two species is different, it should be reflected in perceptual biases and confer different values to the shear. To make stronger conclusions, it remains necessary to replicate the horopter measurement in at least a second macaque and to continue the cyclovergence recording. Several questions will need to be answered; such as the role visual experience might play either at the individual level or at the species level. What could further explain the inter-subject variability found in human observers? Would it also be found in macaque subjects? Finally, how can we explain the difference between the similar shear angle values found in the cat and the owl and the similar ones found in human and macaque? Would the ratio between the interocular distance and the eyes' height be a better proxy than visual experience?

Chapter V –
General discussion: What did we learn about the
integration of binocular disparities?

Summary of the results

During my doctorate, I focused on the cortical processing of binocular disparities. More specifically, I investigated the cortical networks involved in the temporal and spatial integrations of disparities, using a functional neuroimaging approach in macaque subjects. I was also interested in determining whether spatial regularities in visual scenes could play a role in shaping responses to disparities at a cortical and at a behavioural level.

The first study I described in this manuscript aimed at revealing the cortical network responding to motion-in-depth (MID) in macaque. Perceiving this motion-in-depth may require at least partially to temporally integrating binocular disparities. We designed stimuli with dynamic random-dot stereograms and contrasted the BOLD response to two conditions: a smooth motion-in-depth and its control, where the frames were shuffled, thus destroying the smooth motion. Our analyses showed that in each hemisphere of both recorded macaques, three areas were selectively more activated by the MID condition than by its control. They were localised in the superior temporal sulcus (CSM_{STS}), in the infero-temporal gyrus (CSM_{ITG}), and in the caudal part of the post-parietal sulcus (CSM_{PPC}). Performing retinotopic analyses, we showed that CSM_{STS} area did not overlap with the MT cluster but was anterior to it. Finally, using a motion localiser, we found that both CSM_{STS} and CSM_{ITG} respond to planar motion, and that on the contrary CSM_{PPC} seems to be uniquely selective to stereomotion. Taken together, those results suggest that cortical network processing MID is partially shared between human and macaque.

In the second part of the manuscript, I detailed how we attempted to determine whether spatial regularities (or natural statistics) in the visual scenes could influence the processing of spatial disparities. In the first study that was conducted in macaque, we recorded the BOLD responses to different surfaces oriented in depth. Those surfaces were either slanted backward or frontward or tilted leftward or rightward. We expected to find a bias for backward oriented

slants, given that they are aligned with the ground orientation in natural scenes. Surprisingly, we found stronger activations for the backward slant in one monkey only, and nothing in the second subject. It is not clear where this difference between our subjects comes from. Further analyses are thus required, such as using a multivariate approach to overcome limitations of univariate analyses. Despite not being able to conclude about the influence of natural statistics at the cortical level, we could nonetheless describe the cortical network involved in processing spatial disparities in two macaque subjects in a robust manner. We thus found seven retinotopically defined areas that were more strongly responding to our correlated conditions (slanted or tilted surfaces) than to the uncorrelated control condition (CIP1, CIP2, PIP2, MT, MSTv, V3A, V6A), in good agreement with previous studies.

In the second section of this part, we described a psychophysical study that we replicated in 8 human observers and adapted for one macaque subject. The aim was to measure the vertical horopter in our subjects. The idea is that corresponding points on the vertical meridian are tilted backward and that this tilt reflects spatial regularities present in our visual environment. For the first time, we experimentally measured the localisation of those corresponding points in a non-human subject. Although measures are not fully completed, we could already show that 1) this backward tilt is also present in macaque and 2) it is comparable to the tilt found in human observers. This seems to be aligned with the idea that the ratio between the interocular distance and the eyes' height might influence the backward tilt angle of the vertical horopter.

Overlapping cortical networks between temporal and spatial integration of binocular disparities: the case of the PIP cluster

As briefly mentioned in Chapter III (Rima et al., *under review*), we investigated the retinotopic properties of the caudal part of the intra-parietal sulcus with wide-field retinotopic stimuli and found a potentially new cluster, the PIP cluster. Given that this paper is currently under review, any conclusions involving the different areas that we showed might compose the PIP cluster (PIP1, PIP2, CIP1, CIP2, see also Figure 61) have to be carefully made. However, those results provide good material to investigate how the areas in this cluster are involved in processing disparity gradients and I wish to develop this point in the next section.

One of the three areas that were strongly responding to motion-in-depth in our study investigating the temporal integration of disparities was located in the post-parietal cortex (CSM_{PPC}). Given its localisation on the medial bank of the intra-parietal sulcus (IPS), this area could correspond to area PIP and was found to overlap with our retinotopically defined PIP2 area in both subjects (data not shown). Furthermore, a retinotopic analysis of the BOLD signal in the PIP cluster showed that PIP2 was the only area of the cluster to be significantly more activated (permutation tests) for the MID condition than for the control condition in 3 hemispheres. In the spatial correlation study presented in the second part of this manuscript, we found three areas in one monkey that responded more strongly to the backward slant (GS condition) than to the forward slant (nGS). Interestingly, those areas all belong to the PIP cluster: PIP1, PIP2, and CIP1 (CIP2 activations were significantly stronger for GS than nGS in the left hemisphere only). Given that the only difference between those two slanted surfaces is their orientation around the horizontal axis, we can assume that the disparity gradient must be processed first in order to discriminate between the two orientations. We suggest that these three areas from the PIP cluster might respond to spatial disparity gradients.

In the discussion of the stereomotion study we suggested a functional dissociation for 3D processing between the CSM_{PPC} area, potentially PIP2, and the caudal intra-parietal area (CIP) located on the lateral bank of the IPS. Indeed, a subpart of PIP area, but not CIP area (see Figure 62), has been shown to respond to kinetic depth (Durand et al., 2007) whereas CIP is known to process 3D surface orientations (Tsutsui et al., 2002; 2005; Rosenberg et al., 2013). The fact that our CIP areas (CIP1 and CIP2) do not respond more strongly to stereomotion than to its control goes in the direction of the finding of Durand and collaborators (2007). However, if we take the results of our two studies together, it appears that the area PIP, and more specifically PIP2, is involved in processing both temporal and spatial disparity gradients, at least in one of our subjects, thus challenging this potential functional dissociation idea between PIP and CIP. Nonetheless, we could use this PIP2 response to stereomotion as a way to differentiate PIP1 from PIP2 and specifically determine whether this functional differentiation of PIP areas overlaps with the retinotopically defined PIP cluster, reinforcing our claim that PIP area should be divided into PIP1 and PIP2 instead.

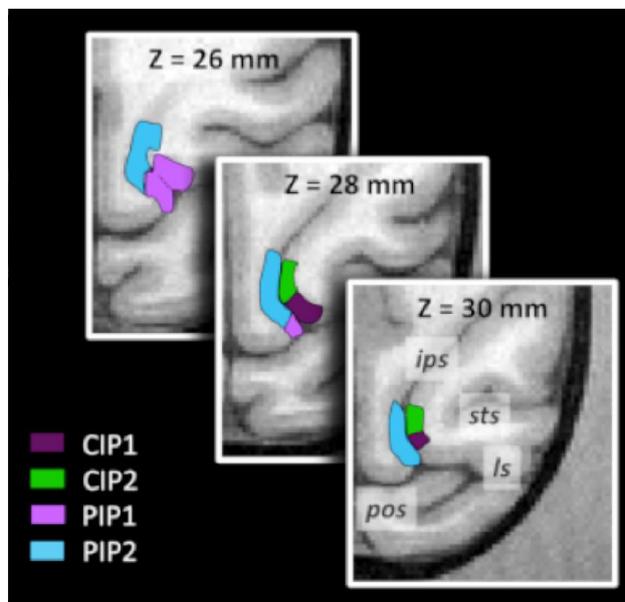


Figure 61. PIP cluster as defined in the wide-field retinotopic study mentioned in Chapter III. (Rima et al., *under review*).

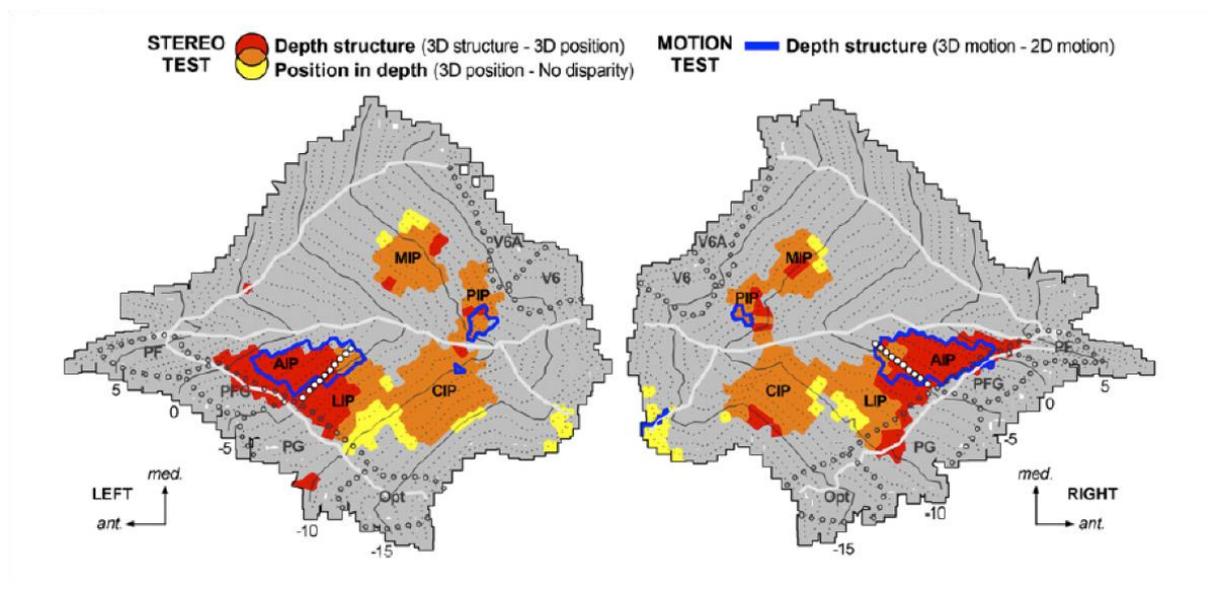


Figure 62. Regions sensitive to structural and positional stereoscopic information are projected onto the flattened representations of left and right IPS, colour coded in red and in yellow, respectively, with mixed sensitivity indicated in orange ($p < 10^{-5}$ uncorrected, masked by all stimulus conditions versus fixation baseline at $p < 10^{-3}$ uncorrected). Blue outlines delineate the regions sensitive to structural depth from motion (same statistical threshold). White dotted lines indicate the AIP/LIP borders derived from the saccade-related activity (same statistical threshold). Grey open symbols demarcate anatomically defined neighbouring areas. Adapted from Durand et al. (2007).

Dorsal versus ventral visual pathways

As mentioned in the introduction (Chapter I), the visual system can be subdivided into two pathways. The dorsal pathway is thought to process spatial information of visual objects, including information about motion and distance. This stream thus supports orientation and navigation in the visual environment (Parker, 2007). It receives input mainly from magnocellular cells, and comprises areas MT, MST (and potentially the whole MT cluster), V3A, and intra-parietal areas (e.g., CIP, AIP, LIP, PIP). On the other hand, the ventral pathway is usually considered to contribute more to colour, form, or shape processing, which underlies

objects identification. It comprises area V4 and infero-temporal areas like TEs. It has been suggested that this segregation of the visual information in two different pathways underlies a differential contribution of each stream for processing binocular disparity (Preston et al., 2008; Srivastava et al., 2009). The main hypothesis summarised by Parker in his review (Parker, 2007) states that different spatial configurations of disparity-defined stimuli are required to elicit respective activations in the dorsal and in the ventral pathway, with a more ‘sophisticated computation’ of disparity in the ventral stream. For instance, zero-order depth and first-order depth (3D oriented surfaces) are at least processed in the dorsal pathway, whereas processing second-order depth (depth structures such as curved surfaces) should be a property of the ventral pathway. It could be shown that TEs neurons, part of the ventral stream, respond to first- and, importantly, to second-order depth (Janssen et al., 2000; 2001). However, several studies have demonstrated that areas in the intra-parietal sulcus, belonging to the dorsal stream, were also significantly involved in processing second-order depth (Durand et al., 2007; Verhoef et al., 2010), and some further demonstrated the causal role of the caudal intra-parietal (CIP) area, and to some extent of the posterior intra-parietal area (PIP), in processing depth structures (Van Dromme et al., 2016). Of important relevance, this latest study also evidenced interactions between the ventral and the dorsal streams, with the reversible inactivation of CIP impacting the activations of the anterior infero-temporal area. This finding emphasises the fact that dorsal and ventral streams are not processing information in a strict independence but rather interact and share information.

We can re-think our results in terms of dorsal and ventral contributions to disparity processing. Indeed, we can consider that temporal integration of disparity is used to detect and interact with moving objects, thus recruiting areas in the dorsal pathway, which is what we mostly found in our results (one area in the superior temporal sulcus, one area in the intra-parietal sulcus, and mild selectivity of the MT cluster). Nevertheless, we also found an area located on the infero-

temporal gyrus, which could belong to the shape sensitive V4A-PITd-PITv-OTd cluster described by Kolster and collaborators (2014), and be potentially more selective to defined-from-motion 3D structures (Vanduffel et al., 2002). This area might equally be involved in the processing of biological motion, a sensitivity that has been shown in neighbouring ventral regions in the STS such as STPa (Oram and Perrett, 1994), and for which the integration of motion-in-depth makes complete sense. Interestingly, the activation of a ventral area in response to motion-in-depth was also found in human in the visual object LO area (Rokers et al., 2009). The LO area has been shown to be indirectly involved in delayed grasping tasks (Cohen et al., 2009; Singhal et al., 2013), further strengthening the interaction idea between the dorsal and the ventral stream. This interaction could be present even for stimuli as simple as stereomotion, which is in favour of the idea that information happens to be shared between both streams at multiple stages and locations either via feedback and recurrent processing or via continuous modulation or both (Cloutman, 2013).

Regarding our study on the spatial integration of disparity however, we found that the areas that were more responsive to our correlated conditions than to the uncorrelated control were almost all part of the dorsal pathway, with the exception of V4. The lack of ventral areas could come from the fact that some ventral areas, despite being selective to correlated tilted or curved surfaces also have neurons that strongly respond to uncorrelated disparity (see e.g. Janssen et al., 2001), making it difficult to distinguish the average area's response to correlated vs. uncorrelated surfaces. This asymmetry between dorsal and ventral areas is thus limiting the interpretation we might give in terms of dorsal and ventral pathways.

Limitations and future directions

Linking cortical activation to behavioural responses

During my PhD work, I attempted to bring together the cortical activity and the visual perception of the environment to have a more integrative view of how the visual system processes information. More specifically, I was interested in addressing the question of whether perceptual biases resulting from statistical regularities (the vertical tilt of the horopter) can be observed in the cortical activity of higher-level brain areas, especially those that are involved in processing surface orientation. As detailed in Chapter III, we found potential cortical biases in one monkey only. One possible explanation for the lack of results in the second subject might come from the angle value we used to design our slanted surfaces. The angle was calculated based on the estimated optimal shear angle of the monkey for which we found stronger activations for more frequent slants (GS condition). Based on psychophysics measures in the second monkey (for which we found no stronger activations), we could see that its vertical tilt angle was about 1° smaller than the angle value we used for the neuroimaging study. Could that value difference be a relevant explanation for the lack of specific response for more frequent orientations? In other words, could it be that the brain areas we found more active in the first monkey are sensitive enough to discriminate between a slant aligned on the true shear angle and a slant aligned on a different-by-one-degree angle value, as the one we used? To get some answers, we could replicate the slant study in our second monkey by using its measured shear angle to design the slants and see whether this makes a difference. Similarly, we could adapt this study to human and record the BOLD signal of our human subjects for which we estimated the vertical shear angle. We could vary the shear angle values, for example taking the values of all the participants, and investigate whether we can discriminate the pattern of cortical activity evoked by the different slant angles. Would we find specific activations for the slanted surface whose angle value is aligned with the measured vertical tilt of the participant?

Finally, another aspect that could be investigated is the relevance of the horopter in visual perception. We started designing a behavioural experiment to determine whether we might be faster or better at detecting or at discriminating slanted surfaces when those surfaces are aligned with the ground and/or when they are aligned with the individual horopter. We will collect data from our human observers that had their horopter empirically measured.

Functional connectivity within the revealed network

Functional connectivity data are being analysed within the team for the same macaque subjects that took part in our neuroimaging experiments. It would be of relevant interest to investigate the connectivity between areas that were responding more strongly to our motion-in-depth stimuli and determine potential interaction points between the dorsal and the ventral streams. We could also have a specific look at the PIP cluster areas that were found to integrate both temporal and spatial disparity gradients and make inference about the areas that might receive their output and from which they might get their input. For instance, it has been suggested that CIP receives part of its input from V3A (Nakamura et al., 2001) and uses it to make 3D orientation discrimination judgement, which does not happen in V3A (Elmore et al., 2019). We also know that areas in the intra-parietal sulcus (IPS) are strongly connected with each other as summarised on Figure 63 (Van Dromme et al., 2016) and that they also follow a hierarchical organisation. Computing the correlation coefficient between brain nodes that encompass our areas of interest might thus reveal broader functional relationship within our network of interest, such as groups of more strongly interconnected areas or hubs that link those different groups to each other, and this could guide future studies.

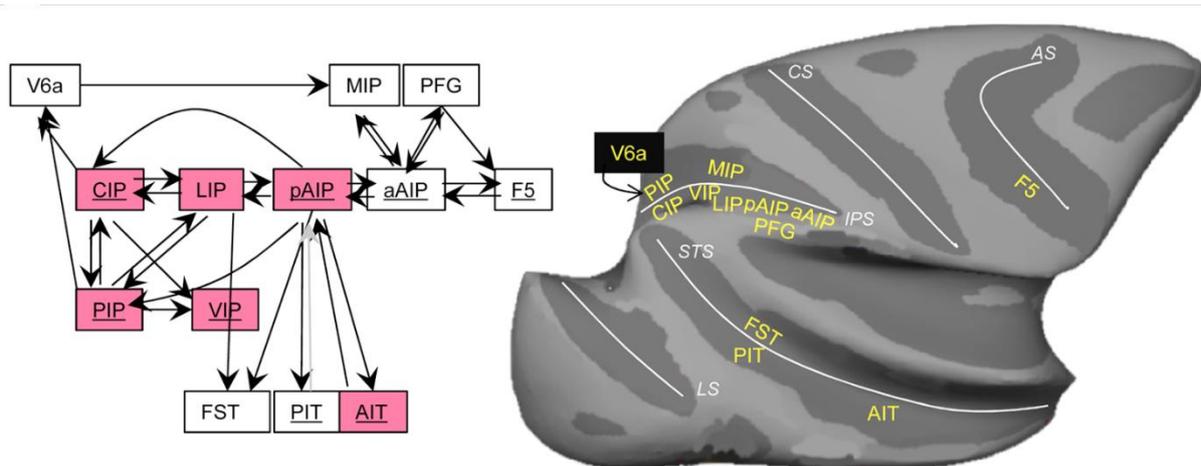


Figure 63. Schematic representation of known functional interactions between different areas from the dorsal and ventral pathways. Red boxes indicate the areas that were affected by the inactivation of the CIP area and underlined area names indicate that they are significantly responding to depth structures. Adapted from Van Dromme et al. (2016).

Modularity of the primate brain

One could argue that one major limitation in our studies is the use of non-ecological visual stimuli or of too simple stimuli, which could limit the exploration of high-level areas. Indeed, it is important to recognise that, despite reflecting encountered visual properties, our stimuli are very artificial, but also that this is needed to have a high level of control regarding the parameters we want to study and regarding the stimuli we will contrast to each other in our neuroimaging paradigms. Equally important is the need to recognise that this approach is based on the idea that the brain is modular and composed of different areas and pathways that integrate complementary information in a hierarchical manner. A more general concern that emerges from these remarks is whether we can understand how the brain works if we break it down into smaller parts. Or whether we should put more efforts into addressing the challenge of designing more ecological protocols instead.

In this work, it was stated that the visual system mostly relies on binocular disparities to compute a depth percept and that the way the correspondence problem is being solved is still not fully understood. Based on electrophysiological data, several models have attempted to explain how neurons might behave like correlation detectors with more or less simple rules (Ohkawa et al., 1990; Read et al., 2002) to partially solve the correspondence problem. However, those computation rules remain limited to the early visual cortex and we are not much more advanced on the matter. We will potentially not be able to solve this problem if we do not take into account the fact that, despite the primordial role of binocular disparities, the visual system also uses other visual cues, such as monocular cues or vergence movements (see Figure 64 for more examples), to build a whole percept. It becomes then important to ask the right kind of questions instead of potentially ill-posed problems that will only lead to partial solutions or explanations. Perceiving depth from bi-dimensional input is not so much about finding the unique solution to the correspondence problem since there is no unique solution but a multitude of them. It is more about using priors, such as learnt statistical arrangements encountered in visual environments, as well as converging and integrating different cues to reduce drastically the amount of potential inter-retinal matches (Jenkin and Harris, 2011). Investigating the interactions between different cues and/or different modalities would allow the development of a more integrative model of the visual system, and of the brain in a general manner (Welchman, 2016, for a review). Obviously, it remains important to also investigate each cue separately to understand the extent to which it might contribute to a whole percept and how this contribution might vary in different situations. However, we should try to keep the bigger picture in mind and put more effort into gathering our current knowledge on the topic.

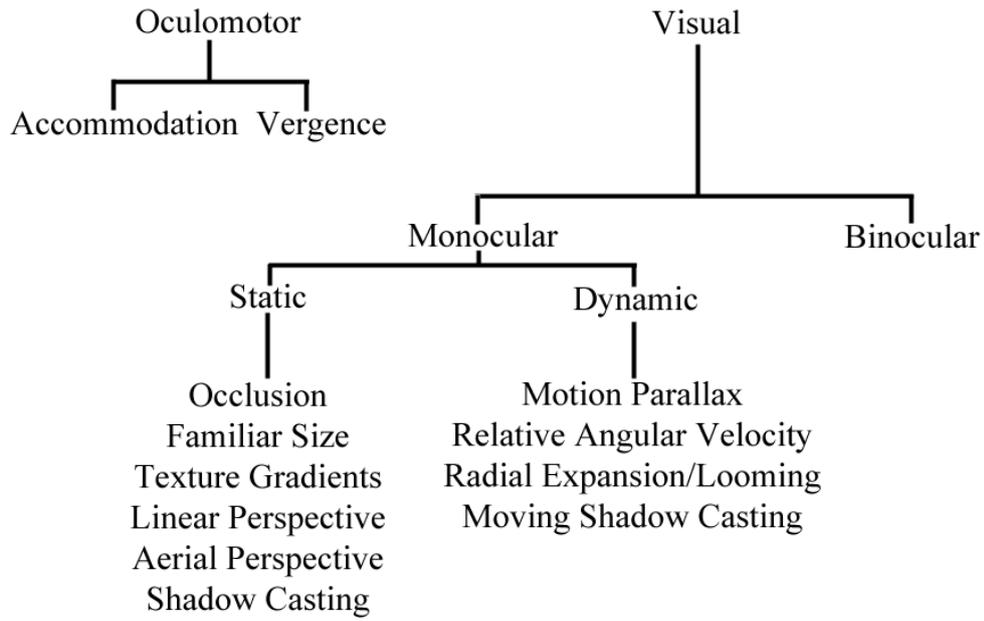


Figure 64. Different visual and oculomotor cues can be used by the primate visual system to compute a 3D percept. Retrieved from <https://slideplayer.com/slide/4713192/>.

Summing it all up with a nice quotation

“Seeing in three dimensions is thus the result of the integration of a large number of different cues, each of which is predicated upon its own set of assumptions and constraints. Many of these cues are not available independently, but rather must be combined in order to be properly interpreted.” (Jenkin and Harris, 2011)

In this work, we proposed to use monkey fMRI as an experimental approach with the aim to bridge the gap between single-cell studies and neuroimaging data in the non-human and human primate, respectively. In our first study, we showed that the cortical network involved in processing motion-in-depth defined by disparity changing over time in macaque is larger than previously thought. In our second study, we asked whether statistical biases found to impact the visual perception of 3D properties could be linked to cortical responses to slanted surfaces that reflect the distribution of disparities in the visual environment. Finally, we showed in a comparative psychophysics study that perceptual biases observed in humans were found in one macaque subject.

Altogether, we believe our findings contribute to a better understanding of binocular disparity processing in the primate brain. The monkey fMRI approach we used allows direct comparisons with human fMRI studies and with monkey electrophysiology, and will help to shed light on interspecies homologies and differences. Furthermore, we demonstrated for the first time that visual perception in macaque is influenced by natural statistics in a similar manner as for human, opening the road to further investigations.

References

- Adams, D. L., & Zeki, S. (2001). Functional Organization of Macaque V3 for Stereoscopic Depth. *Journal of Neurophysiology*, 86(5), 2195–2203. <https://doi.org/10.1152/jn.2001.86.5.2195>
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The Variability of Human, BOLD Hemodynamic Responses. *NeuroImage*, 8(4), 360–369. <https://doi.org/10.1006/nimg.1998.0369>
- Alizadeh, A.-M., Van Dromme, I., Verhoef, B.-E., & Janssen, P. (2018). Caudal Intraparietal Sulcus and three-dimensional vision: A combined functional magnetic resonance imaging and single-cell study. *NeuroImage*, 166, 46–59. <https://doi.org/10.1016/j.neuroimage.2017.10.045>
- Andersen, R. A., Bracewell, R. M., Barash, S., Gnadt, J. W., & Fogassi, L. (1990). Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 10(4), 1176–1196.
- Anzai, A., Chowdhury, S. A., & DeAngelis, G. C. (2011). Coding of Stereoscopic Depth Information in Visual Areas V3 and V3A. *Journal of Neuroscience*, 31(28), 10270–10282. <https://doi.org/10.1523/JNEUROSCI.5956-10.2011>
- Arcaro, M. J., Pinsk, M. A., Li, X., & Kastner, S. (2011). Visuotopic Organization of Macaque Posterior Parietal Cortex: A Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*, 31(6), 2064–2078. <https://doi.org/10.1523/JNEUROSCI.3334-10.2011>
- Backus, B.T., Fleet, D.J., Parker, A.J., & Heeger, D.J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of Neurophysiology*, 86(4), 2054–2068.
- Ban, H., & Welchman, A. E. (2015). fMRI Analysis-by-Synthesis Reveals a Dorsal Hierarchy That Extracts Surface Slant. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 35(27), 9823–9835. <https://doi.org/10.1523/JNEUROSCI.1255-15.2015>

- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5(8), 617.
<https://doi.org/10.1038/nrn1476>
- Barlow, H. (1961). Possible Principles Underlying the Transformations of Sensory Messages. In *Sensory Communication* (pp. 217–234).
<https://doi.org/10.7551/mitpress/9780262518420.003.0013>
- Barlow, H. (2001). The exploitation of regularities in the environment by the brain. *Behavioral and Brain Sciences*, 24(04), 602–607. <https://doi.org/10.1017/S0140525X01000024>
- Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *The Journal of Physiology*, 193(2), 327–342.
<https://doi.org/10.1113/jphysiol.1967.sp008360>
- Bartels, A., Zeki, S. (2000). The architecture of the colour centre in the human visual brain: new results and a review. *European Journal of Neuroscience*, 12, 172–193.
- Beverley, K. I., & Regan, D. (1975). The relation between discrimination and sensitivity in the perception of motion in depth. *The Journal of Physiology*, 249(2), 387–398.
<https://doi.org/10.1113/jphysiol.1975.sp011021>
- Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience*, 16(13), 4207-4221.
- Bridge, H., & Parker, A. J. (2007). Topographical representation of binocular depth in the human visual cortex using fMRI. *Journal of Vision*, 7(14), 15–15.
<https://doi.org/10.1167/7.14.15>
- Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual neuroscience*, 13(1), 87-100.
- Brown, R. W., Cheng, Y.-C. N., Haacke, E. M., Thompson, M. R., & Venkatesan, R. (2014). *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. John Wiley & Sons.

- Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews*, 36(2), 96–107. [https://doi.org/10.1016/S0165-0173\(01\)00085-6](https://doi.org/10.1016/S0165-0173(01)00085-6)
- Cardin, V., & Smith, A. T. (2010). Sensitivity of Human Visual and Vestibular Cortical Regions to Egomotion-Compatible Visual Stimulation. *Cerebral Cortex*, 20(8), 1964–1973. <https://doi.org/10.1093/cercor/bhp268>
- Chauhan, T., Masquelier, T., Montlibert, A., & Cottureau, B. R. (2018). Emergence of Binocular Disparity Selectivity through Hebbian Learning. *The Journal of Neuroscience*, 38(44), 9563–9578. <https://doi.org/10.1523/JNEUROSCI.1259-18.2018>
- Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011). Convergence of Vestibular and Visual Self-Motion Signals in an Area of the Posterior Sylvian Fissure. *Journal of Neuroscience*, 31(32), 11617–11627. <https://doi.org/10.1523/JNEUROSCI.1266-11.2011>
- Chopin, A., Levi, D., Knill, D., & Bavelier, D. (2016). The absolute disparity anomaly and the mechanism of relative disparities. *Journal of Vision*, 16(8), 2–2. <https://doi.org/10.1167/16.8.2>
- Cloutman, L. L. (2013). Interaction between dorsal and ventral processing streams: Where, when and how? *Brain and Language*, 127(2), 251–263. <https://doi.org/10.1016/j.bandl.2012.08.003>
- Cohen, N. R., Cross, E. S., Tunik, E., Grafton, S. T., & Culham, J. C. (2009). Ventral and dorsal stream contributions to the online control of immediate and delayed grasping: A TMS approach. *Neuropsychologia*, 47(6), 1553–1562. <https://doi.org/10.1016/j.neuropsychologia.2008.12.034>
- Colby, C. L., Gattass, R., Olson, C. R., & Gross, C. G. (1988). Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: A dual tracer study. *The Journal of Comparative Neurology*, 269(3), 392–413. <https://doi.org/10.1002/cne.902690307>
- Conway, B.R., Moeller, S., Tsao, D.Y. (2007). Specialized color modules in macaque extrastriate cortex. *Neuron*, 56, 560–573.

- Cooper E.A, Burge, J., Banks, M.S. (2013). The vertical horopter is not adaptable but it might be adaptive. *Journal of Vision*, 11(3).
- Cooper, M.L. and Pettigrew, J.D. (1979). A neurophysiological determination of the vertical horopter in the cat and owl. *Journal of Compared Neurology*, 184, 1-25.
- Cottareau, B. R., Ales, J. M., & Norcia, A. M. (2014). The evolution of a disparity decision in human visual cortex. *NeuroImage*, 92, 193–206. <https://doi.org/10.1016/j.neuroimage.2014.01.055>
- Cottareau, B. R., McKee, S. P., Ales, J. M., & Norcia, A. M. (2011). Disparity-Tuned Population Responses from Human Visual Cortex. *The Journal of Neuroscience*, 31(3), 954–965.
- Cottareau, B. R., Smith, A. T., Rima, S., Fize, D., Héjja-Brichard, Y., Renaud, L., ... Durand, J.-B. (2017). Processing of Egomotion-Consistent Optic Flow in the Rhesus Macaque Cortex. *Cerebral Cortex*, 1–14. <https://doi.org/10.1093/cercor/bhw412>
- Cumming, B. G., & DeAngelis, G. C. (2001). The physiology of stereopsis. *Annual Review of Neuroscience*, 24, 203–238. <https://doi.org/10.1146/annurev.neuro.24.1.203>
- Cumming, B. G., & Parker, A. J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research*, 34(4), 483–495. [https://doi.org/10.1016/0042-6989\(94\)90162-7](https://doi.org/10.1016/0042-6989(94)90162-7)
- Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, 389(6648), 280–283. <https://doi.org/10.1038/38487>
- Cumming, B. G., & Parker, A. J. (1999). Binocular Neurons in V1 of Awake Monkeys Are Selective for Absolute, Not Relative, Disparity. *The Journal of Neuroscience*, 19(13), 5602–5618. <https://doi.org/10.1523/JNEUROSCI.19-13-05602.1999>
- Cumming, B. G., & Parker, A. J. (2000). Local Disparity Not Perceived Depth Is Signaled by Binocular Neurons in Cortical Area V1 of the Macaque. *Journal of Neuroscience*, 20(12), 4758–4767. <https://doi.org/10.1523/JNEUROSCI.20-12-04758.2000>

- Cynader, M., & Regan, D. (1978). Neurons in cat parastriate cortex sensitive to the direction of motion in three-dimensional space. *The Journal of Physiology*, 274, 549–569. <https://doi.org/10.1113/jphysiol.1978.sp012166>
- Cynader, Max, & Regan, D. (1982). Neurons in cat visual cortex tuned to the direction of motion in depth: Effect of positional disparity. *Vision Research*, 22(8), 967–982. [https://doi.org/10.1016/0042-6989\(82\)90033-5](https://doi.org/10.1016/0042-6989(82)90033-5)
- Czuba, T. B., Huk, A. C., Cormack, L. K., & Kohn, A. (2014). Area MT encodes three-dimensional motion. *Journal of Neuroscience*, 34(47), 15522–15533.
- Dacey, D M, & Petersen, M. R. (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Sciences of the United States of America*, 89(20), 9666–9670.
- DeAngelis, G. C., & Newsome, W. T. (1999). Organization of Disparity-Selective Neurons in Macaque Area MT. *Journal of Neuroscience*, 19(4), 1398–1415. <https://doi.org/10.1523/JNEUROSCI.19-04-01398.1999>
- DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology*, 89(2), 1094–1111. <https://doi.org/10.1152/jn.00717.2002>
- Diedrichsen, J., & Shadmehr, R. (2005). Detecting and adjusting for artifacts in fMRI time series data. *Neuroimage*, 27(3), 624–634.
- Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The Occipital Place Area Is Causally and Selectively Involved in Scene Perception. *Journal of Neuroscience*, 33(4), 1331–1336. <https://doi.org/10.1523/JNEUROSCI.4081-12.2013>
- Duffy, C. J., & Wurtz, R. H. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *Journal of Neuroscience*, 15(7), 5192–5208. <https://doi.org/10.1523/JNEUROSCI.15-07-05192.1995>
- Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2), 647–660. <https://doi.org/10.1016/j.neuroimage.2007.09.034>

- Durand, J.-B., Nelissen, K., Joly, O., Wardak, C., Todd, J. T., Norman, J., Janssen, P., Vanduffel, W., & Orban, G. A. (2007). Anterior regions of monkey parietal cortex process visual 3D shape. *Neuron*, 55(3), 493–505. <https://doi.org/10.1016/j.neuron.2007.06.040>
- Durand, J.-B., Peeters, R., Norman, J., Todd, J. T., & Orban, G. A. (2009). Parietal regions processing visual 3D shape extracted from disparity. *NeuroImage*, 46(4), 1114–1126. <https://doi.org/10.1016/j.neuroimage.2009.03.023>
- Durand, J.-B., Zhu, S., Celebrini, S., & Trotter, Y. (2002). Neurons in Parafoveal Areas V1 and V2 Encode Vertical and Horizontal Disparities. *Journal of Neurophysiology*, 88(5), 2874–2879. <https://doi.org/10.1152/jn.00291.2002>
- Einhäuser, W., Kruse, W., Hoffmann, K.-P., & König, P. (2006). Differences of monkey and human overt attention under natural conditions. *Vision Research*, 46(8–9), 1194–1209. <https://doi.org/10.1016/j.visres.2005.08.032>
- Ekstrom, A. (2010). How and when the fMRI BOLD signal relates to underlying neural activity: The danger in dissociation. *Brain Research Reviews*, 62(2), 233–244. <https://doi.org/10.1016/j.brainresrev.2009.12.004>
- Ekstrom, A., Viskontas, I., Kahana, M., Jacobs, J., Upchurch, K., Bookheimer, S., & Fried, I. (2007). Contrasting roles of neural firing rate and local field potentials in human memory. *Hippocampus*, 17(8), 606–617. <https://doi.org/10.1002/hipo.20300>
- Elmore, L. C., Rosenberg, A., DeAngelis, G. C., & Angelaki, D. E. (2019). Choice-Related Activity during Visual Slant Discrimination in Macaque CIP But Not V3A. *ENeuro*, 6(2). <https://doi.org/10.1523/ENEURO.0248-18.2019>
- Erkelens, C. J., & Collewijn, H. (1985). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research*, 25(4), 583–588. [https://doi.org/10.1016/0042-6989\(85\)90164-6](https://doi.org/10.1016/0042-6989(85)90164-6)
- Farivar, R. & Vanduffel, W. (2014). Functional MRI of Awake Behaving Macaques Using Standard Equipment. *Advanced Brain Neuroimaging Topics in Health and Disease - Methods and Applications*. <https://doi.org/10.5772/58281>

- Felleman, D. J., & Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, 1(1), 1–47. <https://doi.org/10.1093/cercor/1.1.1>
- Felleman, D.J. & Kaas, J.H. (1984). Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *Journal of Neurophysiology*, 52, 488–513.
- Fischer, F. (1924). Über Asymmetrien des Gesichtssinnes, speziell des Raumsinnes beider Augen (II.). *Ebenda*, 204, 203-233.
- Franck, M. (1905). Beobachtungen betreffs der Übereinstimmung der Hering - Hillebrand'schen Horopter- abweichung und des Kundt'sehen Teilungs- versuches.
- Friston, K., Holmes, A., Worsley, K., Poline, J.-P., Frith, C., & Frackowiak, R. (1994). Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping*, 2(4), 189–210.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., & Turner, R. (1998). Event-Related fMRI: Characterizing Differential Responses. *Neuroimage*, 7, 30-40.
- Galletti, C., Battaglini, P. P., & Fattori, P. (1990). 'Real-motion' cells in area V3A of macaque visual cortex. *Experimental Brain Research*, 82(1), 67–76. <https://doi.org/10.1007/BF00230838>
- Gautama, T., Mandic, D. P., & Hulle, M. M. V. (2003). Signal nonlinearity in fMRI: A comparison between BOLD and MION. *IEEE Transactions on Medical Imaging*, 22(5), 636–644. <https://doi.org/10.1109/TMI.2003.812248>
- Geisler, W. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, 59, 167–192. <https://doi.org/10.1146/annurev.psych.58.110405.085632>
- Georgieva, S., Peeters, R., Kolster, H., Todd, J. T., & Orban, G. A. (2009). The Processing of Three-Dimensional Shape from Disparity in the Human Brain. *Journal of Neuroscience*, 29(3), 727–742. <https://doi.org/10.1523/JNEUROSCI.4753-08.2009>

- Goense, J. B. M., & Logothetis, N. K. (2008). Neurophysiology of the BOLD fMRI Signal in Awake Monkeys. *Current Biology*, 18(9), 631–640. <https://doi.org/10.1016/j.cub.2008.03.054>
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10), 1409–1422. [https://doi.org/10.1016/S0042-6989\(01\)00073-6](https://doi.org/10.1016/S0042-6989(01)00073-6)
- Gu, Y., Cheng, Z., Yang, L., DeAngelis, G. C., & Angelaki, D. E. (2016). Multisensory Convergence of Visual and Vestibular Heading Cues in the Pursuit Area of the Frontal Eye Field. *Cerebral Cortex*, 26(9), 3785–3801. <https://doi.org/10.1093/cercor/bhv183>
- Hadjidimitrakis, K., Breveglieri, R., Placenti, G., Bosco, A., Sabatini, S. P., & Fattori, P. (2011). Fix Your Eyes in the Space You Could Reach: Neurons in the Macaque Medial Parietal Cortex Prefer Gaze Positions in Peripersonal Space. *PLOS ONE*, 6(8), e23335. <https://doi.org/10.1371/journal.pone.0023335>
- Halldèn, U. (1956). An Optical explanation of Hering-Hillebrand's deviation. *Archives of Ophthalmology*, 55, 830-835.
- Handwerker, D. A., Ollinger, J. M., & D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, 21(4), 1639–1651. <https://doi.org/10.1016/j.neuroimage.2003.11.029>
- Harris, L. R., & Jenkin, M. R. M. (2011). *Vision in 3D Environments*. Cambridge University Press.
- Helmholtz, H. von (1910/1962). *Treatise on Physiological Optics*, Volume 3. New York: Dover (1962); English translation by J. P. C. Southall for the Optical Society of America (1925) from the 3rd German edition of *Handbuch der physiologischen Optik*. Hamburg: Voss (1910).
- Hibbard, P., B. (2007). A statistical model of binocular disparity. *Visual Cognition*, 15(2), 149–165. <https://doi.org/10.1080/13506280600648018>

- Hinkle, D. A., & Connor, C. E. (2001). Disparity tuning in macaque area V4. *Neuroreport*, 12(2), 365–369. <https://doi.org/10.1097/00001756-200102120-00036>
- Hinkle, David A., & Connor, C. E. (2005). Quantitative characterization of disparity tuning in ventral pathway area V4. *Journal of Neurophysiology*, 94(4), 2726–2737. <https://doi.org/10.1152/jn.00341.2005>
- Howard, I.P. (1996). Alhazen's neglected discoveries of visual phenomena, *Perception*, 25(10) 1203 – 1217
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1), 215–243. <https://doi.org/10.1113/jphysiol.1968.sp008455>
- Hubel, D. H., & Wiesel, T. N. (1970). Stereoscopic Vision in Macaque Monkey: Cells sensitive to Binocular Depth in Area 18 of the Macaque Monkey Cortex. *Nature*, 225(5227), 41–42. <https://doi.org/10.1038/225041a0>
- Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture. Functional architecture of a macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological sciences*, 198(1130), 1-59. <https://doi.org/10.1098/rspb.1977.0085>
- Huettel, S. A., Song, A. W., & McCarthy, G. (2009). *Functional Magnetic Resonance Imaging, Second Edition*. Retrieved from <https://book.douban.com/subject/4426184/>
- 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.
- Janssen, P., Vogels, R., & Orban, G. A. (2000). Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science (New York, N.Y.)*, 288(5473), 2054–2056.
- Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2001). Macaque Inferior Temporal Neurons Are Selective for Three-Dimensional Boundaries and Surfaces. *Journal of Neuroscience*, 21(23), 9419–9429. <https://doi.org/10.1523/JNEUROSCI.21-23-09419.2001>

- Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2003). At Least at the Level of Inferior Temporal Cortex, the Stereo Correspondence Problem Is Solved. *Neuron*, *37*(4), 693–701. [https://doi.org/10.1016/S0896-6273\(03\)00023-0](https://doi.org/10.1016/S0896-6273(03)00023-0)
- Joo, S. J., Czuba, T. B., Cormack, L. K., & Huk, A. C. (2016). Separate perceptual and neural processing of velocity-and disparity-based 3D motion signals. *Journal of Neuroscience*, *36*(42), 10791-10802.
- Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: The University of Chicago Press.
- Kaestner, M., Maloney, R. T., Wailes-Newson, K. H., Bloj, M., Harris, J. M., Morland, A. B., & Wade, A. R. (2019). Asymmetries between achromatic and chromatic extraction of 3D motion signals. *Proceedings of the National Academy of Sciences*, *116*(27), 13631–13640. <https://doi.org/10.1073/pnas.1817202116>
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *Journal of Neuroscience*, *17*(11), 4302–4311. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>
- Kay, K.N., Winawer, J., Mezer, A., & Wandell, B.A. (2013). Compressive spatial summation in human visual cortex. *Journal of Neurophysiology*, *110*(2), 481-494.
- Kolster, H., Janssens, T., Orban, G. A., & Vanduffel, W. (2014). The retinotopic organization of macaque occipitotemporal cortex anterior to V4 and caudoventral to the middle temporal (MT) cluster. *Journal of Neuroscience*, *34*(31), 10168-10191.
- Kolster, H., Mandeville, J. B., Arsenault, J. T., Ekstrom, L. B., Wald, L. L., & Vanduffel, W. (2009). Visual field map clusters in macaque extrastriate visual cortex. *Journal of Neuroscience*, *29*(21), 7031-7039.
- Kolster, H., Peeters, R., & Orban, G. A. (2010). The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbours. *Journal of Neuroscience*, *30*(29), 9801-9820.

- Kreiman, G., Hung, C. P., Kraskov, A., Quiroga, R. Q., Poggio, T., & DiCarlo, J. J. (2006). Object Selectivity of Local Field Potentials and Spikes in the Macaque Inferior Temporal Cortex. *Neuron*, *49*(3), 433–445. <https://doi.org/10.1016/j.neuron.2005.12.019>
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience – the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540.
- Lages, M., & Heron, S. (2010). On the Inverse Problem of Binocular 3D Motion Perception. *PLoS Computational Biology*, *6*(11). <https://doi.org/10.1371/journal.pcbi.1000999>
- Leite, F. P., Tsao, D., Vanduffel, W., Fize, D., Sasaki, Y., Wald, L. L., ... Mandeville, J. B. (2002). Repeated fMRI Using Iron Oxide Contrast Agent in Awake, Behaving Macaques at 3 Tesla. *NeuroImage*, *16*(2), 283–294. <https://doi.org/10.1006/nimg.2002.1110>
- Leopold, D. A., Plettenberg, H. K., & Logothetis, N. K. (2002). Visual processing in the ketamine-anesthetized monkey Optokinetic and blood oxygenation level-dependent responses. *Experimental Brain Research*, *143*(3), 359–372. <https://doi.org/10.1007/s00221-001-0998-0>
- LeVay, S., Hubel, D. H., & Wiesel, T. N. (1975). The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. *The Journal of Comparative Neurology*, *159*(4), 559–576. <https://doi.org/10.1002/cne.901590408>
- Li, P., Zhu, S., Chen, M., Han, C., Xu, H., Hu, J., Fang, Y., & Lu, H. D. (2013). A Motion Direction Preference Map in Monkey V4. *Neuron*, *78*(2), 376–388. <https://doi.org/10.1016/j.neuron.2013.02.024>
- Likova, L. T., & Tyler, C. W. (2007). Stereomotion processing in the human occipital cortex. *Neuroimage*, *38*(2), 293–305.
- Liu, Y., Bovik, A. C., & Cormack, L. K. (2008). Disparity statistics in natural scenes. *Journal of Vision*, *8*(11), 19–19. <https://doi.org/10.1167/8.11.19>
- Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*(11), 3416–3468. <https://doi.org/10.1523/JNEUROSCI.07-11-03416.1987>

- Logothetis, N. K. (2003). The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal. *Journal of Neuroscience*, 23(10), 3963–3971. <https://doi.org/10.1523/JNEUROSCI.23-10-03963.2003>
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843), 150. <https://doi.org/10.1038/35084005>
- Logothetis, N.K. & Schall, J.D. (1990). Binocular motion rivalry in macaque monkeys: Eye dominance and tracking eye movements. *Vision Research*, 30(10), 1409-1419.
- Logothetis, N.K. & Wandell, B.A. (2004). Interpreting the BOLD signal. *Annual Review of Physiology*, 66, 735-769.
- Lu, Y., Yin, J., Chen, Z., Gong, H., Liu, Y., Qian, L., ... Wang, W. (2018). Revealing Detail along the Visual Hierarchy: Neural Clustering Preserves Acuity from V1 to V4. *Neuron*, 98(2), 417-428.e3. <https://doi.org/10.1016/j.neuron.2018.03.009>
- Mandeville, J. B., Marota, J. J., Ayata, C., Zaharchuk, G., Moskowitz, M. A., Rosen, B. R., & Weisskoff, R. M. (1999). Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 19(6), 679–689. <https://doi.org/10.1097/00004647-199906000-00012>
- Mansfield, P. (1979). United States Patent No. US4165479A. Retrieved from <https://patents.google.com/patent/US4165479A/en>
- Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., ... Kennedy, H. (2014). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cerebral Cortex*, 24(1), 17–36. <https://doi.org/10.1093/cercor/bhs270>
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco, Cal: W.H. Freeman and Company.
- Maunsell, J. H., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual review of neuroscience*, 10(1), 363-401.

- Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, *49*(5), 1148–1167. <https://doi.org/10.1152/jn.1983.49.5.1148>
- McLaren, D.G., Kosmatka, K.J., Kastman, E.K., Bendlin, B.B., & Johnson, S.C. (2010). Rhesus macaque brain morphometry: A methodological comparison of voxel-wise approaches. *Methods*, *50*, 157–165.
- McLaren, D.G., Kosmatka, K.J., Oakes, T.R., Kroenke, C.D., Kohama, S.G., Matochik, J.A., Ingram, D.K., Johnson, S.C. (2009). A population-average MRI-based atlas collection of the rhesus macaque. *Neuroimage*, *45*, 52–59.
- Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., & Malach, R. (2005). Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex. *Science*, *309*(5736), 951–954. <https://doi.org/10.1126/science.1110913>
- Nakamura, H., Kuroda, T., Wakita, M., Kusunoki, M., Kato, A., Mikami, A., Sakata, H., & Itoh, K. (2001). From Three-Dimensional Space Vision to Prehensile Hand Movements: The Lateral Intraparietal Area Links the Area V3A and the Anterior Intraparietal Area in Macaques. *Journal of Neuroscience*, *21*(20), 8174–8187. <https://doi.org/10.1523/JNEUROSCI.21-20-08174.2001>
- Nakayama, K., Tyler, C. W., & Appleman, J. (1977). A new angle on the vertical horopter. *Supplement to Investigative Ophthalmology & Visual Science*, *82*.
- Nasr, S., & Tootell, R. B. H. (2016). Visual field biases for near and far stimuli in disparity selective columns in human visual cortex. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2016.09.012>
- Nefs, H. T., O'Hare, L., & Harris, J. M. (2010). Two Independent Mechanisms for Motion-In-Depth Perception: Evidence from Individual Differences. *Frontiers in Psychology*, *1*. <https://doi.org/10.3389/fpsyg.2010.00155>
- Nelissen, K., Vanduffel, W., Orban, G.A. (2006) Charting the Lower Superior Temporal Region, a New Motion-Sensitive Region in Monkey Superior Temporal Sulcus. *Journal of Neuroscience*, *26*, 5929– 5947.

- Neri, P., Bridge, H., & Heeger, D. J. (2004). Stereoscopic processing of absolute and relative disparity in human visual cortex. *Journal of Neurophysiology*, 92(3), 1880–1891. <https://doi.org/10.1152/jn.01042.2003>
- Newsome, W. T., & Paré, 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.
- Nguyenkim, J. D., & DeAngelis, G. C. (2003). Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(18), 7117–7128.
- Nienborg, H., & Cumming, B. G. (2006). Macaque V2 Neurons, But Not V1 Neurons, Show Choice-Related Activity. *Journal of Neuroscience*, 26(37), 9567–9578. <https://doi.org/10.1523/JNEUROSCI.2256-06.2006>
- Nikara, T., Bishop, P. O., & Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Experimental Brain Research*, 6(4), 353–372. <https://doi.org/10.1007/BF00233184>
- Nityananda, V., & Read, J. C. A. (2017). Stereopsis in animals: Evolution, function and mechanisms. *Journal of Experimental Biology*, 220(14), 2502–2512. <https://doi.org/10.1242/jeb.143883>
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24), 9868–9872. <https://doi.org/10.1073/pnas.87.24.9868>
- Ogawa, Seiji, & Lee, T.-M. (1990). Magnetic resonance imaging of blood vessels at high fields: In vivo and in vitro measurements and image simulation. *Magnetic Resonance in Medicine*, 16(1), 9–18. <https://doi.org/10.1002/mrm.1910160103>
- Ogawa, Seiji, Lee, T.-M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine*, 14(1), 68–78. <https://doi.org/10.1002/mrm.1910140108>

- Ohzawa, I., & Freeman, R. D. (1986). The binocular organization of simple cells in the cat's visual cortex. *Journal of Neurophysiology*, 56(1), 221–242. <https://doi.org/10.1152/jn.1986.56.1.221>
- Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science (New York, N.Y.)*, 249(4972), 1037–1041. <https://doi.org/10.1126/science.2396096>
- Oram, M. W., & Perrett, D. I. (1994). Responses of Anterior Superior Temporal Polysensory (STPa) Neurons to 'Biological Motion' Stimuli. *Journal of Cognitive Neuroscience*, 6(2), 99–116. <https://doi.org/10.1162/jocn.1994.6.2.99>
- Orban, G. A. (2002). Functional MRI in the awake monkey: The missing link. *Journal of Cognitive Neuroscience*, 14(6), 965-969.
- Orban, G. A. (2011). The extraction of 3D shape in the visual system of human and nonhuman primates. *Annual Review of Neuroscience*, 34, 361–388. <https://doi.org/10.1146/annurev-neuro-061010-113819>
- Orban, G. A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., ... & Vanduffel, W. (2003). Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia*, 41(13), 1757-1768.
- Park, S., & Chun, M. M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *NeuroImage*, 47(4), 1747–1756. <https://doi.org/10.1016/j.neuroimage.2009.04.058>
- Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, 8(5), 379–391. <https://doi.org/10.1038/nrn2131>
- Petras, K., ten Oever, S., Jacobs, C., & Goffaux, V. (2019). Coarse-to-fine information integration in human vision. *NeuroImage*, 186, 103–112. <https://doi.org/10.1016/j.neuroimage.2018.10.086>
- Pettigrew, J. D., Nikara, T., & Bishop, P. O. (1968). Binocular interaction on single units in cat striate cortex: Simultaneous stimulation by single moving slit with receptive fields in

- correspondence. *Experimental Brain Research*, 6(4), 391–410.
<https://doi.org/10.1007/BF00233186>
- Pigarev, I. N., Nothdurft, H.-C., & Kastner, S. (2002). Neurons with radial receptive fields in monkey area V4A: Evidence of a subdivision of prelunate gyrus based on neuronal response properties. *Experimental Brain Research*, 145(2), 199–206.
<https://doi.org/10.1007/s00221-002-1112-y>
- Pitzalis, S., Fattori, P., & Galletti, C. (2013). The functional role of the medial motion area V6. *Frontiers in Behavioral Neuroscience*, 6. <https://doi.org/10.3389/fnbeh.2012.00091>
- Poggio, G. F., & Talbot, W. H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *The Journal of Physiology*, 315, 469–492.
<https://doi.org/10.1113/jphysiol.1981.sp013759>
- Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *Journal of Neuroscience*, 8(12), 4531–4550. <https://doi.org/10.1523/JNEUROSCI.08-12-04531.1988>
- Poggio, G. F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, 25(3), 397–406. [https://doi.org/10.1016/0042-6989\(85\)90065-3](https://doi.org/10.1016/0042-6989(85)90065-3)
- Ponce, C. R., Lomber, S. G., & Born, R. T. (2008). Integrating motion and depth via parallel pathways. *Nature Neuroscience*, 11(2), 216–223. <https://doi.org/10.1038/nn2039>
- Preston, T. J., Li, S., Kourtzi, Z., & Welchman, A. E. (2008). Multivoxel Pattern Selectivity for Perceptually Relevant Binocular Disparities in the Human Brain. *Journal of Neuroscience*, 28(44), 11315–11327. <https://doi.org/10.1523/JNEUROSCI.2728-08.2008>
- Prince, S. J. D., Pointon, A. D., Cumming, B. G., & Parker, A. J. (2002). Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *Journal of Neurophysiology*, 87(1), 191–208.
- Qian, N. (1994). Computing Stereo Disparity and Motion with Known Binocular Cell Properties. *Neural Computation*, 6(3), 390–404.
<https://doi.org/10.1162/neco.1994.6.3.390>

- Read, J. C. A., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Visual Neuroscience*, 19(6), 735–753. <https://doi.org/10.1017/S0952523802196052>
- Regan, D., & Gray, R. (2009). Binocular processing of motion: Some unresolved questions. *Spatial Vision*, 22(1), 1–43. <https://doi.org/10.1163/156856809786618501>
- Rokers, B., Cormack, L. K., & Huk, A. C. (2009). Disparity-and velocity-based signals for three-dimensional motion perception in human MT+. *Nature neuroscience*, 12(8), 1050.
- Rosenberg, A., Cowan, N. J., & Angelaki, D. E. (2013). The Visual Representation of 3D Object Orientation in Parietal Cortex. *Journal of Neuroscience*, 33(49), 19352–19361. <https://doi.org/10.1523/JNEUROSCI.3174-13.2013>
- Roy, J. P., Komatsu, H., & Wurtz, R. H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience*, 12(7), 2478–2492. <https://doi.org/10.1523/JNEUROSCI.12-07-02478.1992>
- Sakata, H., Taira, M., Kusunoki, M., Murata, A., Tanaka, Y. (1997). The TINS lecture; The parietal association cortex in depth perception and visual control of hand action. *Trends in Neuroscience*, 20, 350–357.
- Sakata, H., Taira, M., Murata, A., & Mine, S. (1995). Neural Mechanisms of Visual Guidance of Hand Action in the Parietal Cortex of the Monkey. *Cerebral Cortex*, 5(5), 429–438. <https://doi.org/10.1093/cercor/5.5.429>
- Sanada, T. M., & DeAngelis, G. C. (2014). Neural representation of motion-in-depth in area MT. *Journal of Neuroscience*, 34(47), 15508-15521.
- Schreiber, K.M., Hillis, J.M., Filippini, H.R., Schor, C.M., Banks, M.S. (2008). The surface of the empirical horopter. *Journal of Vision*, 8, 1-20.
- Schyns, P. G., & Oliva, A. (1994). From Blobs to Boundary Edges: Evidence for Time- and Spatial-Scale-Dependent Scene Recognition. *Psychological Science*, 5(4), 195–200. <https://doi.org/10.1111/j.1467-9280.1994.tb00500.x>

- Shapley, R., & Hugh Perry, V. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences*, 9, 229–235. [https://doi.org/10.1016/0166-2236\(86\)90064-0](https://doi.org/10.1016/0166-2236(86)90064-0)
- Shikata, E., McNamara, A., Sprenger, A., Hamzei, F., Glauche, V., Büchel, C., & Binkofski, F. (2007). Localization of human intraparietal areas AIP, CIP, and LIP using surface orientation and saccadic eye movement tasks. *Human Brain Mapping*, 29(4), 411–421. <https://doi.org/10.1002/hbm.20396>
- Shikata, E., Tanaka, Y., Nakamura, H., Taira, M., & Sakata, H. (1996). Selectivity of the parietal visual neurones in 3D orientation of surface of stereoscopic stimuli. *NeuroReport*, 7(14), 2389–2394.
- Simoncelli, E., & Olshausen, B. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24(1), 1193–1216. <https://doi.org/10.1146/annurev.neuro.24.1.1193>
- Singhal, A., Monaco, S., Kaufman, L. D., & Culham, J. C. (2013). Human fMRI Reveals That Delayed Action Re-Recruits Visual Perception. *PLoS ONE*, 8(9). <https://doi.org/10.1371/journal.pone.0073629>
- Slichter, C. P. (2013). Principles of Magnetic Resonance. Springer Science & Business Media.
- Sprague, W. W., Cooper, E. A., Tomic, I., & Banks, M. S. (2015). Stereopsis is adaptive for the natural environment. *Science Advances*, 1(4), e1400254–e1400254. <https://doi.org/10.1126/sciadv.1400254>
- Sugihara, H., Murakami, I., Shenoy, K. V., Andersen, R. A., & Komatsu, H. (2002). Response of MSTd neurons to simulated 3D orientation of rotating planes. *Journal of Neurophysiology*, 87(1), 273–285. <https://doi.org/10.1152/jn.00900.2000>
- Taira, M., Mine, S., Georgopoulos, A. P., Murata, A., & Sakata, H. (1990). Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research*, 83(1), 29–36. <https://doi.org/10.1007/BF00232190>

- Taira, M., Tsutsui, K. I., Jiang, M., Yara, K., & Sakata, H. (2000). Parietal neurons represent surface orientation from the gradient of binocular disparity. *Journal of Neurophysiology*, 83(5), 3140–3146.
- Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R. & Kawashima, R. (2011). Effects of training processing speed on neural systems. *Journal of Neuroscience*, 31(34), 12139-12148.
- Tanabe, S., Umeda, K., & Fujita, I. (2004). Rejection of False Matches for Binocular Correspondence in Macaque Visual Cortical Area V4. *Journal of Neuroscience*, 24(37), 8170–8180. <https://doi.org/10.1523/JNEUROSCI.5292-03.2004>
- Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. *Nature Neuroscience*, 5(5), 472–478. <https://doi.org/10.1038/nn837>
- Troscianko, J., von Bayern, A.M.P., Chappell, J., Rutz, C., and Martin, G.R. (2012). Extreme binocular vision and a straight bill facilitate tool use in New Caledonian crows. *Nature Communication*, 3, 1110. <https://doi.org/10.1038/ncomms2111>
- Tsao, D. Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T. A., Mandeville, J. B., ... Tootell, R. B. H. (2003). Stereopsis Activates V3A and Caudal Intraparietal Areas in Macaques and Humans. *Neuron*, 39(3), 555–568. [https://doi.org/10.1016/S0896-6273\(03\)00459-8](https://doi.org/10.1016/S0896-6273(03)00459-8)
- Tsutsui, K.-I., Sakata, H., Naganuma, T., & Taira, M. (2002). Neural correlates for perception of 3D surface orientation from texture gradient. *Science*, 298(5592), 409–412. <https://doi.org/10.1126/science.1074128>
- Tsutsui, K.-I., Taira, M., & Sakata, H. (2005). Neural mechanisms of three-dimensional vision. *Neuroscience Research*, 51(3), 221–229. <https://doi.org/10.1016/j.neures.2004.11.006>
- Tyler, C. (1991). The horopter and binocular fusion. In *Vision and Visual Disorders. Vol. 9, Binocular Vision* (pp. 19–37). Macmillan Publishers.
- Umeda, K., Tanabe, S., & Fujita, I. (2007). Representation of stereoscopic depth based on relative disparity in macaque area V4. *Journal of Neurophysiology*, 98(1), 241–252. <https://doi.org/10.1152/jn.01336.2006>

- Van Dromme, I. C., Premereur, E., Verhoef, B. E., Vanduffel, W., & Janssen, P. (2016). Posterior parietal cortex drives inferotemporal activations during three-dimensional object vision. *PLoS biology*, *14*(4), e1002445.
- Van Essen, D. C., & DeYoe, E. A. (1995). Concurrent processing in the primate visual cortex. In M. Gazzaniga (Ed.), *The cognitive Neurosciences* (pp. 383-400). Cambridge: Bradford Book.
- Van Essen, D.C., Drury, H.A. Dickson, J., Harwell, J., Hanlon, D., Anderson, C.H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, *8*(5), 443– 459.
- Vanduffel, W., Fize, D., Mandeville, J. B., Nelissen, K., Van Hecke, P., Rosen, B. R., ... Orban, G. A. (2001). Visual Motion Processing Investigated Using Contrast Agent-Enhanced fMRI in Awake Behaving Monkeys. *Neuron*, *32*(4), 565–577. [https://doi.org/10.1016/S0896-6273\(01\)00502-5](https://doi.org/10.1016/S0896-6273(01)00502-5)
- Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J. T., & Orban, G. A. (2002). Extracting 3D from motion: differences in human and monkey intraparietal cortex. *Science*, *298*(5592), 413-415.
- Verhoef, B.-E., Bohon, K. S., & Conway, B. R. (2015). Functional Architecture for Disparity in Macaque Inferior Temporal Cortex and Its Relationship to the Architecture for Faces, Color, Scenes, and Visual Field. *The Journal of Neuroscience*, *35*(17), 6952–6968. <https://doi.org/10.1523/JNEUROSCI.5079-14.2015>
- Wall, M.B. & Smith, A.T. (2008). The representation of egomotion in the human brain. *Current Biology*, *18*(3), 191-194.
- Ward, M. K., Bolding, M. S., Schultz, K. P., & Gamlin, P. D. (2015). Mapping the Macaque Superior Temporal Sulcus: Functional Delineation of Vergence and Version Eye-Movement-Related Activity. *Journal of Neuroscience*, *35*(19), 7428–7442. <https://doi.org/10.1523/JNEUROSCI.4203-14.2015>
- Watanabe, M., Tanaka, H., Uka, T., & Fujita, I. (2002). Disparity-Selective Neurons in Area V4 of Macaque Monkeys. *Journal of Neurophysiology*, *87*(4), 1960–1973. <https://doi.org/10.1152/jn.00780.2000>

- Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. *NeuroImage*, 33(2), 493–504. <https://doi.org/10.1016/j.neuroimage.2006.07.029>
- Welchman, A. E. (2016). The Human Brain in Depth: How We See in 3D. *Annual Review of Vision Science*, 2, 345–376. <https://doi.org/10.1146/annurev-vision-111815-114605>
- Welchman, A. E., Deubelius, A., Conrad, V., Bühlhoff, H. H., & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience*, 8(6), 820–827. <https://doi.org/10.1038/nn1461>
- Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. *Experimental Brain Research*, 36(3), 585–597. <https://doi.org/10.1007/BF00238525>
- Wheatstone Charles. (1838). XVIII. Contributions to the physiology of vision. —Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, 128, 371–394. <https://doi.org/10.1098/rstl.1838.0019>
- Yamane, Y., Carlson, E. T., Bowman, K. C., Wang, Z., & Connor, C. E. (2008). A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nature Neuroscience*, 11(11), 1352–1360. <https://doi.org/10.1038/nn.2202>
- Yang, Z., & Purves, D. (2003). A statistical explanation of visual space. *Nature Neuroscience*, 6(6), 632–640. <https://doi.org/10.1038/nn1059>
- Zeki, S.M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology*, 236, 549–573.
- Zhang, T., & Britten, K. H. (2011). Parietal Area VIP Causally Influences Heading Perception during Pursuit Eye Movements. *Journal of Neuroscience*, 31(7), 2569–2575. <https://doi.org/10.1523/JNEUROSCI.5520-10.2011>

Webvision articles and general resources about the visual system (Chapter I):

Color Vision by Peter Gouras – Webvision. (n.d.). Retrieved 3 June 2019, from <https://webvision.med.utah.edu/book/part-vii-color-vision/color-vision/>

Midget pathways of the primate retina underlie resolution and red green color opponency by Helga Kolb – Webvision. (n.d.). Retrieved 21 May 2019, from <https://webvision.med.utah.edu/book/part-iii-retinal-circuits/midget-pathways-of-the-primate-retina-underly-resolution/>

Bowmaker, J. K. (1998). Evolution of colour vision in vertebrates. *Eye (London, England)*, 12 (Pt 3b), 541–547. <https://doi.org/10.1038/eye.1998.143>

Briggs, F. (2017). Mammalian Visual System Organization. *Oxford Research Encyclopedia of Neuroscience*. <https://doi.org/10.1093/acrefore/9780190264086.013.66>

Carvajal, C., Viéville, T., Alexandre, F. (2012). Konio Pathway: An Instinctive Visual Mechanism for Survival and Decision Making? NeuroComp/KEOpS'12 workshop beyond the retina: from computational models to outcomes in bioengineering. Focus on architecture and dynamics sustaining information flows in the visuomotor system., Oct 2012, Bordeaux, France. hal-00756471

Dacey, D.M., Crook, J.D., & Packer, O.S. (2014). Distinct synaptic mechanisms create parallel S-ON and S-OFF color opponent pathways in the primate retina. *Visual Neuroscience*, 31(2), 139–151. <https://doi.org/10.1017/S0952523813000230>

Howard, I.P. and Rogers, B.J. (2002). Seeing in Depth, Volume 2: Depth Perception. Toronto, Canada: University of Toronto Press.

Klein, C., Evrard, H. C., Shapcott, K. A., Haverkamp, S., Logothetis, N. K., & Schmid, M. C. (2016). Cell-Targeted Optogenetics and Electrical Microstimulation Reveal the Primate Koniocellular Projection to Supra-granular Visual Cortex. *Neuron*, 90(1), 143–151. <https://doi.org/10.1016/j.neuron.2016.02.036>

Masland, R. H. (2001). The fundamental plan of the retina. *Nature Neuroscience*, 4(9), 877. <https://doi.org/10.1038/nn0901-877>

Appendix I –
Consent form for the horopter experiment
conducted in humans

Mesure de l'horoptère et des biais perceptuels qui en découlent

Chercheur titulaire responsable scientifique du projet :

Benoit Cottureau (benoit.cottureau@cnrs.fr), chargé de recherche au CNRS, Centre de Recherche Cerveau et Cognition Pavillon Baudot, CHU Purpan BP 25202 - 31052 Toulouse Cedex. Téléphone : 05 62 74 45 20.

Vous avez été sollicité(e) pour participer à cette étude parce que vous êtes intéressé(e) par ce type de recherche et que vous êtes un adulte bien portant âgé(e) de 18 ans à 50 ans.

Vous êtes libre d'accepter ou de refuser de participer à cette recherche. Afin de pouvoir prendre votre décision en toute connaissance de cause, les modalités de cette recherche sont détaillées ci-dessous.

But du projet de recherche :

Dans cette étude nous cherchons à mesurer votre horoptère, c'est-à-dire la région de l'espace visuel au sein duquel votre vision 3D est la plus fine, et à connaître les conséquences au niveau comportemental de l'existence de biais découlant de notre environnement visuel.

Cette expérience se déroulera en deux phases : une phase de mesure de l'horoptère et une phase comprenant une tâche de discrimination et une tâche de détection qui seront développées autour de la valeur de votre horoptère, afin d'étudier les effets liés aux biais visuels.

Déroulement de la recherche, durée de participation et méthodes :

Dans un premier temps, vous serez recruté par les investigateurs et collaborateurs scientifiques de cette étude. Les objectifs, risques et inconvénients vous seront expliqués ainsi que le déroulement de l'étude.

Après avoir posé toutes vos questions concernant l'étude, vous devrez signer le consentement éclairé pour votre participation à cette étude.

La recherche aura lieu au Centre de Recherche Cerveau et Cognition (CerCo) de l'Institut des Sciences du Cerveau de Toulouse, pavillon Baudot du CHU de l'hôpital Purpan.

Avant de commencer les séances d'entraînement, vous réaliserez un test sur internet afin de vérifier que vous avez une vision stéréoscopique, c'est-à-dire que vous pouvez déterminer la profondeur d'un objet à partir d'indices visuels binoculaires. Ce test n'a pas pour objectif de mesurer votre acuité visuelle mais simplement d'indiquer si vous pourrez percevoir les indices binoculaires utilisés dans nos expériences.

Si vous acceptez de participer à cette étude vous devrez vous rendre au Pavillon Baudot à 10 reprises pour une durée de 30 minutes pour chacune des expériences, et devrez donc pouvoir

être disponible sur une période de trois semaines pour chacune des deux expériences. La durée entre les deux expériences pourra être ajustée en fonction de vos disponibilités et sera au minimum d'une semaine.

Expérience 1 (10 séances de 30minutes) – regroupées sur 3 semaines

Expérience 1 : Mesures répétées de l'horoptère

Dans un premier temps, vous allez réaliser une tâche de détection de mouvement apparent afin de mesurer votre horoptère, c'est-à-dire la région visuelle au sein de laquelle votre vision 3D est la plus fine, de façon précise et robuste.

Expérience 2 (10 séances de 30min) – regroupées sur 3 semaines

Expérience 2 : Tâches de discrimination et de détection

Dans un deuxième temps, vous réaliserez en alternance une tâche de discrimination et une tâche de détection réparties en blocs. Quatre séances d'entraînement à la tâche seront proposées et suivies de six séances de mesures réelles.

Pour chacune des deux expériences, vous effectuerez d'abord quatre séances d'entraînement afin de vous habituer à la tâche, qui seront suivies de six séances de mesures réelles.

Chaque séance durera environ 30 minutes à raison de 4 séances par semaine (rythme modulable en fonction de vos disponibilités mais ne devant pas dépasser 3 semaines).

Il n'y a pas d'intervalle de temps imposé entre les 2 expériences.

Participation à une autre recherche scientifique

Certains protocoles de recherche prévoient une période d'exclusion, c'est-à-dire qu'une fois le consentement signé, vous ne pouvez pas participer à une autre recherche en parallèle. Si vous participez déjà à une autre recherche qui prévoit une période d'exclusion, vous ne pourrez pas participer à la nôtre en même temps, nous devons donc attendre que la période d'exclusion soit terminée avant de vous inclure dans notre protocole. En revanche, notre recherche ne prévoit pas de période d'exclusion. Cela vous permet donc d'être inclus dans un autre protocole que le nôtre.

Vos droits de vous retirer de la recherche en tout temps :

Votre contribution à cette recherche est volontaire. Vous avez le droit de refuser de participer à cette recherche, ou de retirer votre consentement et d'interrompre votre participation **à tout moment**, que ce soit par exemple pendant une expérience ou entre les deux expériences, sans encourir aucune responsabilité, ni aucun préjudice de ce fait.

Vos droits à la confidentialité et au respect de la vie privée :

Toute information vous concernant recueillie pendant cette étude sera traitée de façon confidentielle. Seuls les responsables de l'étude pourront avoir accès à ces données. Les données de recherche non nominatives seront sauvegardées et analysées au sein du laboratoire CerCo. La publication des résultats de l'étude ne comportera aucun résultat

individuel permettant votre identification.

Bénéfices :

Les bénéfices attendus de cette recherche sont d’obtenir une meilleure connaissance de la vision 3D, des biais statistiques existant et de l’effet de ces biais sur des tâches de discrimination et de détection. Une meilleure compréhension du système visuel pourra contribuer à développer des stratégies visuelles pour les personnes qui n’auraient pas une vision 3D efficace.

Risques possibles :

À notre connaissance, cette recherche n’implique aucun risque ou inconfort autre que ceux de la vie quotidienne.

Compensation financière :

Une compensation financière à hauteur de 60€ pour la totalité de l’expérience vous sera remise. Si l’expérience venait à être arrêtée en cours, une contribution calculée au prorata des heures effectuées vous serait versée (5€ par heure entamée).

Diffusion :

Les résultats de cette recherche seront diffusés dans des colloques et publiés dans des actes de colloque et des articles de revue académique.

Vos droits de poser des questions en tout temps :

Vous pouvez poser des questions au sujet de la recherche en tout temps en communiquant avec le responsable scientifique du projet par courrier électronique à benoit.cottureau@cnrs.fr

Consentement à la participation :

En signant le formulaire de consentement, vous certifiez que vous avez lu et compris les renseignements ci-dessus, qu’on a répondu à vos questions de façon satisfaisante et qu’on vous a avisé que vous étiez libre d’annuler votre consentement ou de vous retirer de cette recherche en tout temps, sans préjudice.

A remplir par le ou la participant.e :

J’ai lu et compris les renseignements ci-dessus et j’accepte de plein gré de participer à cette recherche.

Nom, Prénom – Date – Signature

Un exemplaire de ce document vous est remis, un autre exemplaire est conservé dans le dossier.

Appendix II –

Cottereau, B. R., Smith, A. T., Rima, S., Fize, D., Héjja-Brichard, Y., Renaud, L., Lejards, C., Vayssière, N., Trotter, Y., & Durand, J.-B. (2017). Processing of Egomotion-Consistent Optic Flow in the Rhesus Macaque Cortex. *Cerebral Cortex*, 1–14.

ORIGINAL ARTICLE

Processing of Egomotion-Consistent Optic Flow in the Rhesus Macaque Cortex

Benoit R. Cottureau^{1,2}, Andrew T. Smith³, Samy Rima^{1,2}, Denis Fize⁴, Yseult Hégja-Brichard^{1,2}, Luc Renaud^{5,6}, Camille Lejards^{1,2}, Nathalie Vayssière^{1,2}, Yves Trotter^{1,2} and Jean-Baptiste Durand^{1,2}

¹Université de Toulouse, Centre de Recherche Cerveau et Cognition, Toulouse, France, ²Centre National de la Recherche Scientifique, Toulouse, France, ³Department of Psychology, Royal Holloway, University of London, Egham, UK, ⁴Laboratoire d'Anthropologie Moléculaire et Imagerie de Synthèse, CNRS-Université de Toulouse, Toulouse, France, ⁵CNRS, CE2F PRIM UMS3537, Marseille, France and ⁶Aix Marseille Université, Centre d'Exploration Fonctionnelle et de Formation, Marseille, France

Address correspondence to Benoit R. Cottureau and Jean-Baptiste Durand, CNRS CERCO UMR 5549, Pavillon Baudot, CHU Purpan, BP 25202, 31052 Toulouse Cedex, France. Email: cottureau@cerco.ups-tlse.fr (B.R.C.); jbdurand@cnrs.fr (J.-B.D.)

Abstract

The cortical network that processes visual cues to self-motion was characterized with functional magnetic resonance imaging in 3 awake behaving macaques. The experimental protocol was similar to previous human studies in which the responses to a single large optic flow patch were contrasted with responses to an array of 9 similar flow patches. This distinguishes cortical regions where neurons respond to flow in their receptive fields regardless of surrounding motion from those that are sensitive to whether the overall image arises from self-motion. In all 3 animals, significant selectivity for egomotion-consistent flow was found in several areas previously associated with optic flow processing, and notably dorsal middle superior temporal area, ventral intra-parietal area, and VPS. It was also seen in areas 7a (Opt), STPm, FEFsem, FEFsac and in a region of the cingulate sulcus that may be homologous with human area CSv. Selectivity for egomotion-compatible flow was never total but was particularly strong in VPS and putative macaque CSv. Direct comparison of results with the equivalent human studies reveals several commonalities but also some differences.

Key words: egomotion, heading, monkey fMRI, optic flow, vision

Introduction

In macaques, numerous regions of the cerebral cortex contain at least some neurons that are selectively responsive to the direction of motion of a moving visual stimulus. These regions have diverse locations including large parts of the occipital cortex, posterior portions of the temporal cortex, the inferior parietal cortex, and even parts of the frontal cortex. Although the most obvious use of sensitivity to image motion is to specify the

motion of external objects, it is also valuable for monitoring the animal's own movements. Two cortical regions in particular, the dorsal middle superior temporal area (MSTd) and the ventral intra-parietal area (VIP), are associated with the specialized function of encoding visual cues to self-motion. Both contain many neurons that are selectively sensitive to specific components of the optic flow that occurs during self-motion, including direction of heading during locomotion (Tanaka et al. 1989;

Duffy and Wurtz 1991, 1995; Bremmer et al. 2002a). Electrical stimulation of these regions can influence heading judgments (Britten and van Wezel 2002; Zhang and Britten 2011) suggesting that they contribute directly to perceptual awareness of self-motion, although this has recently been questioned in the case of VIP (Chen et al. 2016). Many MSTd and VIP neurons also receive vestibular input (Duffy et al. 1976; Gu et al. 2006; Chen et al. 2011a) and there is evidence that visual and vestibular cues are efficiently integrated by such neurons, with weightings based on cue reliability (Fetsch et al. 2012). Neurons that appear to encode optic flow have also been identified in area 7a of the inferior parietal lobule (Motter and Mountcastle 1981; Steinmetz et al. 1987) and recently in the frontal eye fields (Gu et al. 2015), where again many neurons also respond to vestibular stimuli.

Despite much research, it is not known exactly how visual responses to specific types of optic flow are constructed. The problem has proved challenging and although several sophisticated and biologically plausible models have been proposed (e.g., Perrone and Stone 1994; Grossberg et al. 1999; Yu et al. 2010; Mineault et al. 2012), the computations involved are still debated. However, it is clear that signals encoding motion, which is initially encoded locally, must be spatially integrated in some way. By definition, self-motion generates full-field visual stimulation: When the animal moves, the entire retinal image moves. Receptive fields in MSTd and VIP are large enough (typically 10–50 degs, e.g., Komatsu and Wurtz 1988; Schaafsma and Duysens 1996; Mendoza-Halliday et al. 2014) to integrate local motion signals over a wide area, and in some cases a significant proportion of the visual field, but they are not large enough to integrate signals over the entire visual field. The question therefore arises: Do the responses of flow-selective neurons having different receptive field locations combine to specify the overall optic flow, and if so how? Visual responses are usually studied with a simulated optic flow stimulus positioned such that all key features of the flow are contained within the receptive field. For example, in the case of expansion (forward motion), the center of expansion is typically placed within the receptive field being examined. The implicit assumption is that single neurons are concerned with optic flow only within their receptive fields and can therefore be expected to respond the same way to a given stimulus irrespective of whether the remainder of the visual field is consistent with the same optic flow. However, this has not been tested empirically and consequently we do not know whether neurons in MST and VIP respond 1) whenever what falls in their purview “could be” part of full-field optic flow, or 2) only, or at least more strongly, when signals from other parts of the visual field indicate that it actually is part of full-field flow. Whether MSTd and VIP encode optic flow per se, or localized flow components that can be used to derive overall optic flow, is a key unanswered question.

Surprisingly, perhaps, the above question has been better addressed in humans than in macaques. In the human brain, putative homologs of macaque MST (Dukelow et al. 2001; Huk et al. 2002; Kolster et al. 2010) and VIP (Bremmer et al. 2001) have been identified. These have been shown to be involved in encoding optic flow (Smith et al. 2006; Cardin et al. 2012a). Wall and Smith (2008) addressed the question of whether human MST (hMST) and VIP are active whenever optic flow components are present in the image, or only when full-field flow is present. They presented an array of 9 (3 × 3) optic flow patches. The patches were identical and each contained spiraling flow that would be expected to provide a good stimulus for a macaque MST neuron if presented in its receptive field. However, the array

as a whole was not consistent with self-motion and should not activate neurons that respond selectively to image motion caused by self-motion. Wall and Smith (2008) found that hMST responded almost as strongly to the array as to a single large patch of the same total size, suggesting that it is not strongly sensitive to whether or not image motion reflects self-motion. In putative human VIP, the response was about half that to a single patch, implying stronger selectivity for self-motion. A more extensive study using the same paradigm (Cardin and Smith 2010) confirmed these findings and additionally identified 2 more visually responsive regions that respond at least twice as well to one patch as to an array. One was human V6 (hV6), a region identified in humans only quite recently (Pitzalis et al. 2006) and thought to be the homologue of macaque V6 (Galletti et al. 2001). The other was labeled PIVC (parieto-insular vestibular cortex) but was probably PIC (posterior insular cortex), a visual-vestibular region immediately posterior to human PIVC (Frank et al. 2014) that may be homologous to macaque VPS (visual posterior sylvian area; Chen et al. 2011b). In macaques, both V6 (Fan et al. 2015) and VPS (Chen et al. 2011b) contain neurons that are tuned for visually simulated direction of heading. In both the above human fMRI (functional magnetic resonance imaging) studies (Wall and Smith 2008; Cardin and Smith 2010), the strongest specificity to visual self-motion occurred in a region not previously studied in any detail, the cingulate sulcus visual (CSv) area. Here, a strong response could be elicited by a single optic flow patch but the response was almost completely abolished when an array of optic flow patches was used as substitute. Recent studies (Antal et al. 2008; Fischer et al. 2012) confirm the role of CSv in self-motion processing and an additional piece of evidence implicating CSv in self-motion processing is that it receives vestibular as well as visual input (Smith et al. 2012). Thus, population responses in human visual cortex show a hierarchy of sensitivity to whether the overall visual image is likely to reflect self-motion, from hMST (weakest sensitivity), through hV6 and hVIP (substantial sensitivity), to PIC and CSv (strongest sensitivity).

There have been no single-unit studies in macaques that used either the multi-patch approach or, to our knowledge, any other approach to distinguish responses to true, full-field optic flow from responses to the mere presence of optic flow segments in the receptive field. Before undertaking such studies, it would be valuable to establish with fMRI which macaque visual areas, if any, show such differentiation on a macroscopic scale. This would guide physiological experimentation and also provide a much stronger link with the relevant human fMRI literature. There are numerous important species differences that could make human fMRI studies an unreliable guide to macaque physiology. Not least, area CSv has not been identified in macaques. We have therefore employed the multi-patch paradigm during fMRI in alert fixating macaques with the aim of establishing candidate visual regions for true self-motion specialization, in the sense discussed above.

Materials and Methods

Animal Model

Subjects

Three female rhesus macaques: M01, M02, and M03 (age: 5–7 years; weight: 4.5–6.5 kg) were involved in this study. Animal housing, handling, and all experimental protocols (surgery, behavioral training, and MRI (magnetic resonance imaging) recordings) followed the guidelines of the European Union legislation (2010/63/UE) and of the French Ministry of Agriculture

(décret 2013–118). The project was approved by a local ethics committee (CNREEA code: C2EA – 14) and received authorization from the French Ministry of Research (MP/03/34/10/09). The 3 animals were housed together in a large, enriched enclosure and could thus develop social and foraging behaviors. They returned to their individual cages to be fed twice a day, with standard primate biscuits supplemented with various types of fruits and vegetables. Health inspections were carried out quarterly on these animals. Details about the animals' surgical preparation and behavioral training are provided as Supplementary text 1.

Optic Flow Stimuli

The stimuli were identical to those used in previous human studies (Wall and Smith 2008; Cardin and Smith 2010). They consisted of 800 moving dots arranged in an egomotion-consistent (EC) or egomotion-inconsistent (EI) pattern. The EC condition consisted of a $40^\circ \times 40^\circ$ square field of dots moving in a coherent optic flow pattern containing expansion/contraction and rotation components that varied over time, consistent with self-motion on a varying spiral trajectory (Morrone et al. 2000), displayed at 60 fps. For a given dot with radius r , angle θ , and local speed v , its trajectory was defined by:

$$\frac{dr}{dt} = v \cos \phi$$

$$\frac{d\theta}{dt} = (v \sin \phi)/r$$

Radial and angular velocities are defined by dr/dt and $d\theta/dt$, respectively. The direction of optic flow was defined by ϕ , which varied over time from $-\pi$ to π generating stimuli with radial, circular, and spiral motion. The EI stimulus consisted of a 3×3 array of 9 identical panels, each containing a smaller version of the EC stimulus. Although the individual panels contain optic flow, the overall pattern is not consistent with egomotion because flow induced by observer motion can have only one center of motion. In true optic flow stimuli, the size and speed of motion of the features in the image increase with eccentricity. Because the introduction of these scaling factors would result in different distributions of dot size and speed in our 2 stimuli, and potentially spurious results, we kept the dot size, dot speed, and number of dots in the whole array identical across conditions in order to equate low-level visual characteristics. As a result, our stimulus does not accurately simulate "true" optic flow in terms of the scaling of size and speed with eccentricity typical of motion through a cloud of dots. The use of time-varying flow ensured that all locations were stimulated by all dot directions during the course of the stimulus cycle. It also provides larger responses than continuous expansion because multiple flow-sensitive neurons are stimulated. Finally, it ensures that adaptation at any one local direction is minimal.

MRI Recordings

Images were acquired on a 3 Tesla clinical MR scanner (Phillips Achieva) using a custom 8-channel phased array coil (RapidBiomed) specially designed to fit the skull of macaques while preserving their field of view.

Recordings for Individual Templates

For each individual, anatomical, and functional brain templates were built from acquisitions made in a single session on slightly anaesthetized animals (Zoletil 100:10 mg/kg and

Domitor: 0.04 mg/kg). The animals' constants were monitored during the whole session (about 1 h) with an MR compatible oximeter. During that session, we acquired 4 T1-weighted anatomical volumes magnetization prepared rapid gradient echo (MPRAGE; repetition time [TR] = 10.3 ms; echo time [TE] = 4.6 ms, flip angle = 8° ; voxel size = $0.5 \times 0.5 \times 0.5$ mm; 192 slices), and 300 functional volumes (GE-EPI; TR = 2000 ms, TE = 30 ms, flip angle = 90° , SENSE factor = 1.6; voxel size = $1.25 \times 1.25 \times 1.5$ mm, 32 axial slices).

Recordings for Functional Sessions

The functional scanning sessions were performed on awake behaving animals on a daily basis and lasted for about 1 h (8–12 runs). The animals were head-fixed, seated in a sphynx position within their primate chair (Fig. 1A). They were introduced into the bore of the magnet, facing a translucent screen at a distance of 50 cm. Visual stimuli were rear-projected on the screen by a video projector (Hitachi, CP_X809), at a spatial resolution of 1024×768 pixels and a refresh rate of 60 Hz. The position of one eye was monitored with an infrared video-based eye-tracker at 60 Hz (ASL). Functional images were acquired with the same GE-EPI sequence as that used during the anesthetized sessions. EC and EI stimuli were presented using a block-design. Each run consisted of 224 s (112 TRs) divided into 7 identical cycles of 32 s (16 TRs). In half of the runs, a cycle started with a baseline of 10 s (5 TRs) where only the fixation point was present. It was followed by 6 s (3 TRs) of the EC condition, then by another 10 s of blank and finally by 6 s of the EI condition (Fig. 1B). In the other half of the runs, the EC and EI conditions were reversed within a cycle (i.e., a cycle had 10 s of blank, 6 s of the EI condition, 10 s of blank, and finally 6 s of the EC condition). Video display and reward for correct fixation were controlled using the V-Cortex software.

Data Processing

Anatomical and Functional Templates

Data collected during the anesthetized sessions were used to compute individual functional and anatomical templates. The anatomical template was obtained by realigning and averaging the 4 T1-weighted (MPRAGE) volumes. It was then aligned to the MNI space of the 112RM-SL template (McLaren et al. 2009, 2010). Cortical surface reconstruction were performed using the CARET software (Van Essen et al. 2001). The functional template was obtained by realigning and averaging the 300 functional (GE-EPI) volumes. It was aligned with the anatomical template and spatial normalization parameters (affine and non-rigid) between the functional and anatomical templates were determined based on the gray matter maps of both templates. For group analyses, the same operation was performed to register each individual anatomical template to the F99 template available in the CARET software (Van Essen 2002).

Preprocessing of the Functional Data

In total, 36 runs per animal were kept for further analyses (18 runs with the EC condition first during the blocks and 18 runs with the EI condition first). All those runs were selected based on the quality of fixation (percentage of correct fixation >85%) in order to minimize the influence of eye movements in blood-oxygen-level dependant (BOLD) signal fluctuations (see Supplementary text 2 and Supplementary Fig. 1 for an additional control on the influence of eye position on our results). Within each run, volumes were rigidly realigned with each other on a slice-by-slice basis using a subpixel cross-correlation

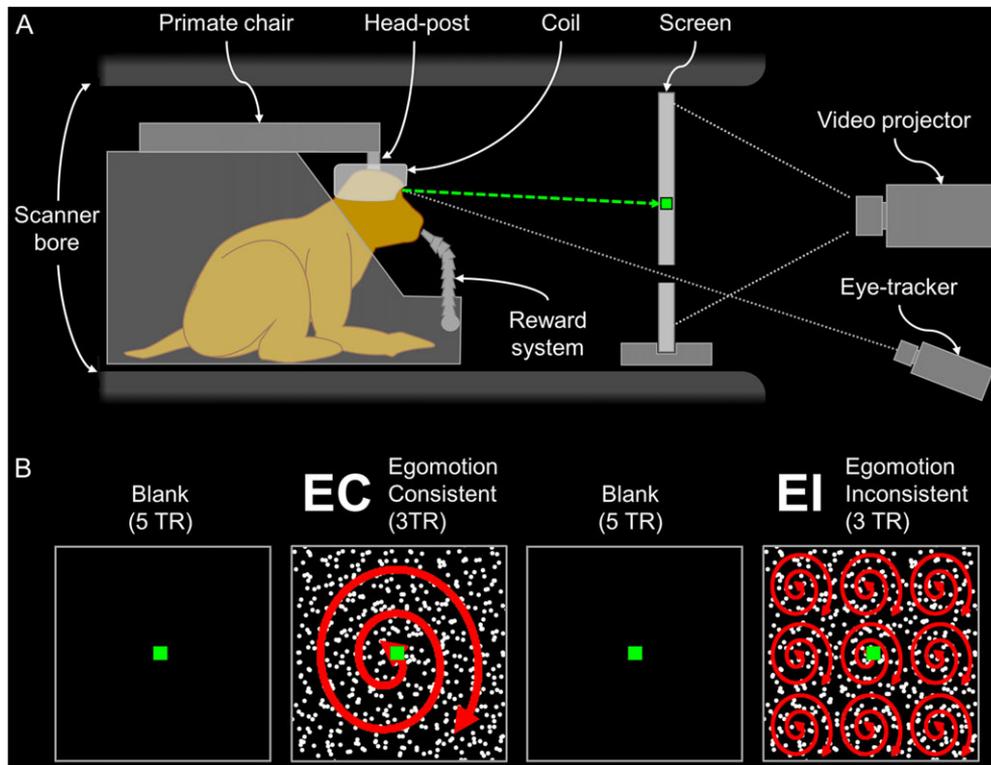


Figure 1. (A) Schematic representation of the monkey fMRI set-up. The animal sits in a sphinx position within the primate chair, in the bore of the scanner, head restrained by the head-post, with the 8 channel, phase array coil located on top of the head. The animal is involved in a passive fixation task, i.e., maintaining the gaze on a green fixation target back-projected on a stimulation screen by a video-projector. Eye position is monitored by an infrared video-based eye-tracker. Correct fixation triggers the delivery of fluid rewards during the runs. (B) Illustration of the stimuli and experimental design. The EC stimulus consisted of a square field of dots moving in a coherent optic flow pattern containing expansion/contraction and rotation components that varied over time, consistent with self-motion on a varying spiral trajectory. The EI stimulus consisted of a 3×3 array of 9 identical panels, each containing a smaller version of the EC stimulus. Recordings were performed using a block-design, with the alternation of EC and EI flow stimuli, separated by blank periods. Each run contained 7 repetitions of such blocks (112 TR in total). EC conditions were shown first in half of the runs and EI conditions appeared first in the other half of the runs.

algorithm (Guizar-Sicairos et al. 2008). This was followed by slice-time correction. A mean image of the functional volumes was then computed for each run and used for normalization on the functional template of the same individual. Those run-dependent normalization parameters were combined to the run-independent parameters linking the functional template to the anatomical one in a single deformation step, during which the functional volumes were resampled at $1 \times 1 \times 1$ mm and slightly smoothed with a spatial Gaussian kernel (FWHM = $1.5 \times 1.5 \times 1.5$ mm).

General Linear Model

Voxel-wise statistics were computed by fitting a general linear model (GLM) to the BOLD signal. The model contained 3 main regressors, representing the 3 experimental conditions: EC, EI, and blank periods (Fig. 1B). Those regressors were convolved with the hemodynamic response function (HRF) estimated from each of the 3 monkeys (HRF estimation from independent datasets is detailed in Supplementary text 3 and Supplementary Fig. 2). In addition, 4 motion regressors were included in the model. For each run, the slice-by-slice rigid realignment yielded 32 vectors of lateral displacements and 32 vectors of antero-posterior (Y) displacements: One for each slice of the functional volume. The principal component analysis was used to derive 2 lateral and 2 antero-posterior motion regressors, which were entered in the model as nuisance regressors (Vanduffel and Farivar 2014). Both the preprocessing steps and GLM analyses were

implemented in Matlab, with the SPM12 software and custom scripts.

Statistics and Results Presentation

As a first step for identifying potential regions of interest (ROIs) (i.e., regions with consistent BOLD response differences between the EC and EI conditions), the volumetric statistical parametric maps obtained for the EC > EI and EI > EC contrasts in the individual GLM analyses were thresholded at $P < 10^{-3}$ uncorrected (t value > 3.1) and spatially normalized for projection onto the cortical surface of the F99 template. Cortical regions showing significant differences in the same direction between the EC and EI conditions in at least 2 individuals were all considered as ROIs. In a second step, 2 new GLM analyses were performed on each individual after splitting the runs into 2 equal parts (18 runs per GLM, 9 of them with the EC condition first). One GLM was used to look for the presence of individual statistical local maxima in the ROIs or in their immediate vicinity. This search was performed at a relaxed statistical threshold of $P < 10^{-2}$ uncorrected. The GLM performed with the other half of the runs was then used to extract the percent BOLD signal changes (PSC) in cubes of $3 \times 3 \times 3$ voxels centered on the local maxima found with the first GLM. This method avoids the “double dipping” that arises when the same data are used both for identifying ROIs and for measuring activity within them. Small cubes were favored over patches determined by anatomical and/or statistical considerations, because anatomical borders between areas are difficult to

determine precisely and, as we will see below, our contrasts (EC > EI but also EC + EI > baseline) lead to extended activations that cannot be accurately divided into clusters corresponding to different functional regions. Our approach is more conservative and avoids subjectivity when dealing with borders between areas. As we will see in the Results section (cf., the MNI coordinates provided in Table 1), our local maxima are separated enough that we were able to associate a single local maximum with each region and with no overlap between the cubes corresponding to the different regions. Importantly, we assessed that the precise size of those cubes (1, 3 or 5 voxels size) did not have significant impact on the extracted results.

For these cubes, we estimated PSC for the EC and EI conditions relative to the Blank condition as follows:

$$PSC_{EC} = 100 \times (\beta_{EC} - \beta_{Blank}) / \beta_{Constant}$$

$$PSC_{EI} = 100 \times (\beta_{EI} - \beta_{Blank}) / \beta_{Constant}$$

where β_{EC} , β_{EI} , β_{Blank} , and $\beta_{Constant}$ represent the regressor coefficients provided by the GLM analyses. For each region of interest, PSC_{EC} and PSC_{EI} were computed for each run independently, and thus expressed as mean \pm standard error across the runs included in the second GLM (18 runs per monkey). Only regions where PSCs were significantly stronger for the EC than for the EI condition (t-tests with $P < 0.05$ and a confidence interval for the difference distribution that does not include 0) were considered for further analysis. Given our number of animals ($n = 3$), there is no statistical test of generalizability. We therefore present the

data on each individual and focus on regions that were consistently found in all the macaques (in at least one hemisphere). In these regions, specificity of the BOLD responses to EC versus EI conditions was quantified by computing a sensitivity ratio of the mean PSC_{EC} and PSC_{EI} with the following formula:

$$\text{Sensitivity ratio (\%)} = 100 \times (PSC_{EC} - PSC_{EI}) / PSC_{EC}.$$

Results

Cortical Network Involved in Processing Optic Flow

In the present study, monkey fMRI techniques are used to characterize the cortical network involved in processing optic flow signals generated by self-motion in nonhuman primate. To that end, 3 macaque monkeys were exposed to optic flow stimuli (moving random dots pattern) either consistent or inconsistent with egomotion (EC and EI conditions, respectively; see Materials and Methods). The experimental design is similar to that developed by Smith et al. in their human fMRI studies (Wall and Smith 2008; Cardin and Smith 2010), allowing a direct comparison of the cortical networks processing optic flow in the 2 primate species (Orban 2002; Orban et al. 2004).

We first assessed the changes in BOLD signal evoked by the visual conditions (EC and EI) relative to baseline (blank screen with fixation point only). These flow stimuli were found to elicit strong statistical increases in BOLD signal across most of the visual cortex in all 3 monkeys (Supplementary Fig. 3A), with a very high degree of overlap between individuals (Supplementary Fig. 3B).

Table 1. Cortical areas activated in at least 3 animals

ROI	M01			M02			M03			Average		
	x	y	z	x	y	z	x	y	z	x	y	z
MSTd												
L	-18	-2	27	-14	-6	29	-13	-4	24	-15	-4	27
R	16	-1	26	14	-4	29	14	-4	26	15	-3	27
FEFsem												
L	-14	23	27	-15	24	29	-15	26	27	-15	24	28
R	13	23	26	12	25	28	-	-	-	13	24	27
VPS												
L	-17	1	28	-19	-3	29	-15	2	27	-17	1	28
R	17	3	27	20	-2	30	-	-	-	19	1	29
7a												
L	-14	-4	30	-15	-7	32	-15	-6	30	-15	-6	31
R	-	-	-	12	-7	34	17	-5	31	15	-6	33
STPm												
L	-	-	-	-20	1	17	-17	2	17	-19	2	17
R	17	2	18	20	0	18	17	1	20	18	1	19
VIP												
L	-8	3	24	-13	3	29	-11	1	27	-11	2	27
R	13	7	27	13	3	29	9	1	24	12	4	27
LIPd												
L	-	-	-	-15	1	33	-15	1	32	-15	1	33
R	-	-	-	14	2	32	16	2	31	15	2	32
FEFsac												
L	-13	23	22	-19	28	21	-15	27	20	-16	26	21
R	-	-	-	17	25	20	-	-	-	17	25	20
pmCSv												
L	-4	15	28	-3	13	27	-3	15	27	-4	14	27
R	-	-	-	2	16	28	-	-	-	2	16	28

Note: Coordinates in MNI space (mm) are those of the statistical local maxima in the left (L) and right (R) hemispheres. Local maxima were determined from the first half of the data and significant selectivity for the EC condition were evaluated from the second half (see details in the text).

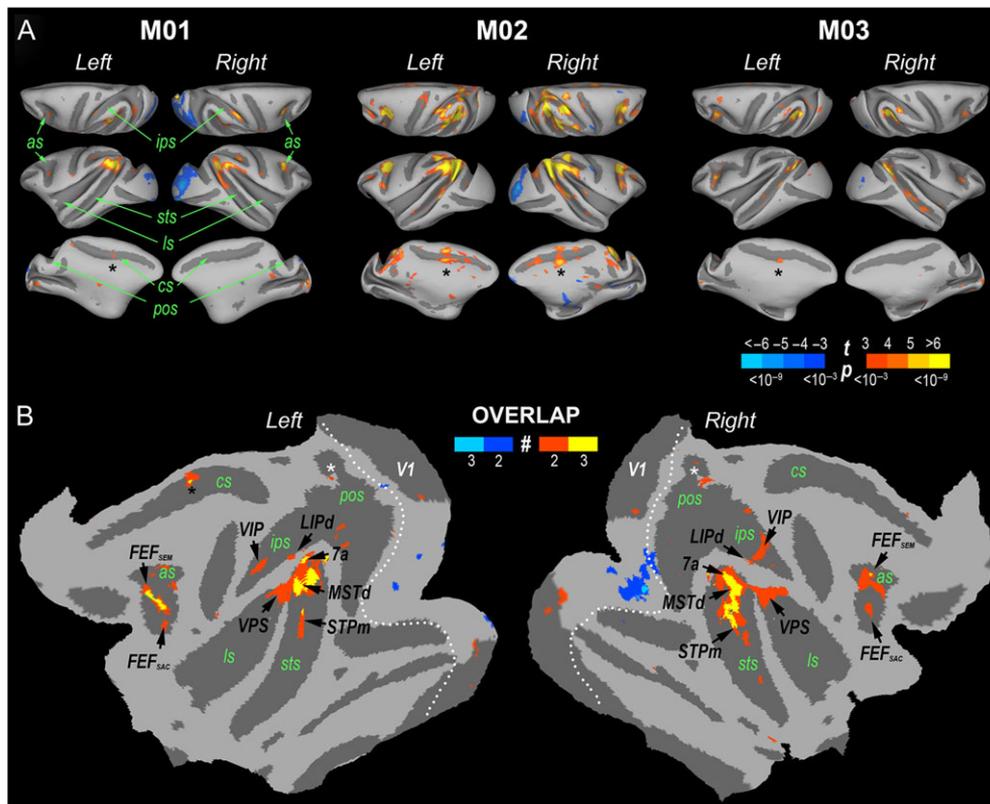


Figure 2. (A) Statistical parametric maps for the EC versus EI contrast in monkeys M01, M02, and M03. Results are projected on dorsal, lateral, and medial views of the left and right hemispheres of the individual cortices. The color code reflects the contrast t-values and indicates statistically significant differences between BOLD responses evoked by the EC and EI conditions ($P < 10^{-3}$ uncorrected). Hot (orange to yellow) and cold (dark to pale blue) colors indicate stronger responses to EC and EI, respectively. (B) Map of overlap between significant activations in the EC versus EI contrast across the 3 monkeys. Only activation sites found in at least 2 individuals are shown. Results are projected on the flattened representations of the left and right hemispheres of the F99 template. Orange and yellow indicate cortical sites significantly more activated by EC than by EI in 2/3 and 3/3 of the subjects, respectively. Dark and pale blue show regions more activated by EI than by EC in 2/3 and 3/3 of the subjects, respectively. Seven cortical were significantly activated in our 3 macaques: MSTd, 7a, STPm, VIP, VPS, FEF_{SEM}, and FEF_{SAC}. Black stars indicate a region of the cingulate sulcus (pmCSv) that was found in the 3 animals in the left hemisphere and in one animal on the right hemisphere. White stars indicate a region of the parieto-occipital sulcus where significant activations were found in 2 animals. Borders of the primary visual area (V1) are shown as white dotted lines (as, arcuate sulcus; cs, cingulate sulcus; ips, intraparietal sulcus; ls, lateral sulcus; sts, superior temporal sulcus).

This first analysis indicates that studying the more specific contrast between the EC and EI conditions is not hampered by a lack of statistical power in any of the individuals. Figure 2A shows the statistical parametric maps (t-values) based on all the available data (36 runs/animal) for the EC versus EI contrast in monkeys M01, M02, and M03, projected on dorsal, lateral, and medial views of the individuals left and right cortical hemispheres (see Materials and Methods). Hot colors (orange to yellow) indicate significantly stronger BOLD responses for EC than for EI condition ($P < 10^{-3}$ uncorrected), while cold colors (dark to pale blue) signal the opposite. Despite differences in the extent of the activation patterns observed in the 3 animals (e.g., monkey M02 is generally more responsive than monkeys M01 and M03), preference for the consistent flow (EC > EI) defines a cortical network encompassing the occipital, parietal, temporal, and frontal lobes in all the monkeys. Many nodes of this network are found consistently across the 3 individuals, as revealed in Figure 2B by overlapping the activations observed in at least 2 of the 3 individuals after normalization on the F99 template (Van Essen 2002). Regions color-coded in orange and yellow are those in which EC evokes significantly stronger activations than EI in 2/3 and 3/3 of the monkeys, respectively. Regions with stronger activations for the EI condition are color-coded in dark and pale blue, depending on whether they are found in 2 or 3 animals.

For each of the anatomical regions showing activation overlap, we looked for the presence of statistical local maxima in the native space of each individual. This search was performed on the volumetric statistical parametric maps obtained with GLMs including only half of the runs (18 runs/animal), at a relaxed threshold of $P < 10^{-2}$ uncorrected (see the “Materials and Methods” section). We chose a relaxed threshold in order to avoid the risk of false negatives while using only half of the full dataset. However, the robustness of those activation sites was tested by running paired t-tests on the BOLD signal change profiles obtained at the same sites in the EC and EI conditions in the other half of the runs.

In the following, we describe the activation sites that were identified with the above method in at least one hemisphere in each of the 3 individuals. Two cortical activation sites were found to prefer EC flow in both hemispheres of all our macaques. The most significant one according to the EC versus EI contrast (i.e., global statistical maximum in all 6 hemispheres) was located in the dorso-caudal portion of the superior temporal sulcus (sts), a location corresponding to the dorsal Medial Superior Temporal (MSTd) area. The second site, located within the intraparietal sulcus (ips), matches the location of the VIP. Both MSTd and VIP areas have been repeatedly shown to play a central role in optic flow processing. Note that for VIP, the

activation found in M03 did not overlap those found in M01 and M02, although the local maxima were very close between the 3 animals (see Table 1).

Four other sites were observed in 5 out of 6 hemispheres. One of them was found in the caudal portion of the lateral sulcus (ls), in a location matching that of the VPS (Chen et al. 2011b). Another site was located dorsally in the arcuate sulcus (as), in a portion of the Frontal Eye Field involved in smooth pursuit eye movements (FEFsem) and also recently shown to house neurons responsive to optic flow stimuli (Gu et al. 2015). The third site, in the postero-ventral portion of the inferior parietal lobule, slightly above MSTd, seems to correspond to posterior area 7a, a region that is known to contain neurons that respond to optic flow (e.g., Siegel and Read 1997). Finally, the fourth site lay in the fundus of the sts, anterior to MSTd. This could be part of the superior temporal polysensory area (STP), which also contains motion-sensitive cells with large receptive fields (e.g., Bruce et al. 1981), although STP occupies primarily the upper bank of the sulcus. We tentatively refer to it as STPm after Nelissen et al. (2006) who report motion-sensitive activity at a similar location with fMRI.

Two further sites were found in all 3 individuals, but less reliably across hemispheres (in 4 out of 6 hemispheres). One of them was located in the as, slightly more anterior and lateral than FEFsem, in a location described as a portion of the Frontal Eye Field involved in saccadic eye movements (FEFsac; Gu et al. 2015). Finally, consistent activations were observed within the postero-ventral lip of the cingulate sulcus (black asterisks in Fig. 2A and B), in a region which had not been documented previously as being involved in optic flow processing in monkeys. However, the location of this region echoes that of the recently discovered CSv area in human, which has been shown to be highly selective for the egomotion-compatible optic flow stimuli used in the present study (Wall and Smith 2008; Cardin and Smith 2010). For that reason, we will refer to this cingulate activation site as putative macaque homologue of CSv (pmCSv; see Fig. 2B). Overlapping activations and corresponding local maxima were also identified in 2 out of 3 animals along the dorsal lip of the intraparietal sulcus (LIPd) and within the parieto-occipital sulcus (white asterisks in Fig. 2B). However, only the LIPd maxima were associated with significant differences between the BOLD signals evoked by the EC and EI conditions in both animals. MNI coordinates of the statistical local maxima for the different areas described above are provided in Table 1.

Note that we also found consistent responses for the EI > EC contrast, but they remained largely restricted to the early visual cortex (see blue patches in Fig. 2B). These activations are not caused by local motion characteristics since they are well matched between the 2 conditions. They might be due to the detection of kinetic boundaries (Reppas et al. 1997).

Quantitative Analysis of Egomotion Selectivity

In the following, we characterize in more detail the BOLD response profiles in the cortical regions enumerated above. For each region, the statistical local maximum was localized based on the GLMs performed in half of the runs for each monkey. The response profiles were estimated with the GLMs performed on the other half of the runs (see Materials and Methods). These response profiles correspond to the average responses within cubes of 27 ($3 \times 3 \times 3$) voxels centered on the local maxima localized on the first GLMs. Figures 3 and 4 show the percentage of BOLD signal changes evoked by the EC and EI conditions with respect to the Blank condition within the

8 areas that were activated in all 3 individuals (areas MSTd, VIP, VPS, FEFsem, 7a, STPm, FEFsac, and pmCSv). Voxel-wise statistical parametric maps obtained for the EC > EI contrast in each individual, superimposed on horizontal sections of the individual anatomical templates, are shown in Figs 3A (areas MSTd, VPS, VIP, pmCSv, and FEFsem), 4A (area 7a), and 4C (areas STPm and FEFsac). Their corresponding BOLD profiles are presented in Figs 3B, 4B and D. The asterisks above the profiles indicate statistically significant differences (paired t-test, $P < 10^{-2}$) between the BOLD responses evoked by the EC (white bars) and EI (gray bars) conditions. The percentages of BOLD signal change in area LIPd are shown in Supplementary Fig. 4 for the 2 animals (M02 and M03) that had significantly stronger responses for the EC than for the EI condition in this area.

In order to characterize the strengths of the BOLD responses elicited by the EC stimulus relative to those evoked by the EI stimulus, we computed a sensitivity ratio (in percentage) between the percentages of signal change obtained for these 2 conditions relative to baseline (see the “Materials and Methods” section). These ratios are shown in Fig. 5A for all the 8 areas reported above and their respective locations are illustrated on the cortical surface of the F99 template in Fig. 5B (the positions of these ROIs on the individual cortical surfaces are provided in Supplementary Fig. 5A). In Fig. 5A, the areas are ranked from the highest to the lowest specificity for the EC condition.

This analysis reveals that 2 regions, pmCSv and VPS, emerge as being clearly the most specific for flow stimuli compatible with egomotion, with a near absence of BOLD responses evoked by the egomotion-incompatible stimuli (see also Fig. 3B). In both regions, the mean ratio (across the 3 animals) was above 70% (77% in pmCSv and 70% in VPS), revealing a nearly 4 times larger response for the egomotion-compatible stimuli. Ratios were much lower but still impressive in areas VIP (43%), FEFsac (43%), and FEFsem (39%), and lower still in MSTd (29%), 7a (21%), and STPm (15%). In LIPd (not shown because it was found in only 2 of the 3 individuals), we observed an intermediate ratio of 47%. Note that these sensitivity ratios are robust to changes in the size of the cubes used to define the ROIs (see Supplementary Fig. 5B).

Finally, we estimated differential sensitivity to egomotion-compatible flow within a number of pre-defined visual ROIs taken from the Caret atlas. This enabled us to cross-check our results for regions such as MSTd and also to check that in visual regions such as V1–V3 and MT, where specificity is not expected, it is not seen. The procedure is described in Supplementary text 4. The results, shown in Supplementary Fig. 6, confirmed that selectivity for EC stimuli is seen in MSTd but not in MT, FST (Fundus of the Superior Temporal sulcus), or V4t and was not seen in V1–V3 or V3A. V6 showed weak selectivity (see Discussion).

Discussion

The aim of present study was to identify, in nonhuman primates, the cortical areas involved in processing visual motion produced by self-displacements, i.e., EC optic flow. To that end, we recorded whole-brain BOLD responses from 3 behaving macaques while they were exposed to optic flow stimuli either consistent or inconsistent with egomotion (Fig. 1). The visual stimuli and experimental design were similar to those used in previous human fMRI studies (Wall and Smith 2008; Cardin and Smith 2010), allowing a direct comparison of the cortical networks between human and nonhuman primates. Our results reveal that in macaque, as in human, many cortical areas are more strongly activated by EC optic flow stimuli. Those regions are broadly

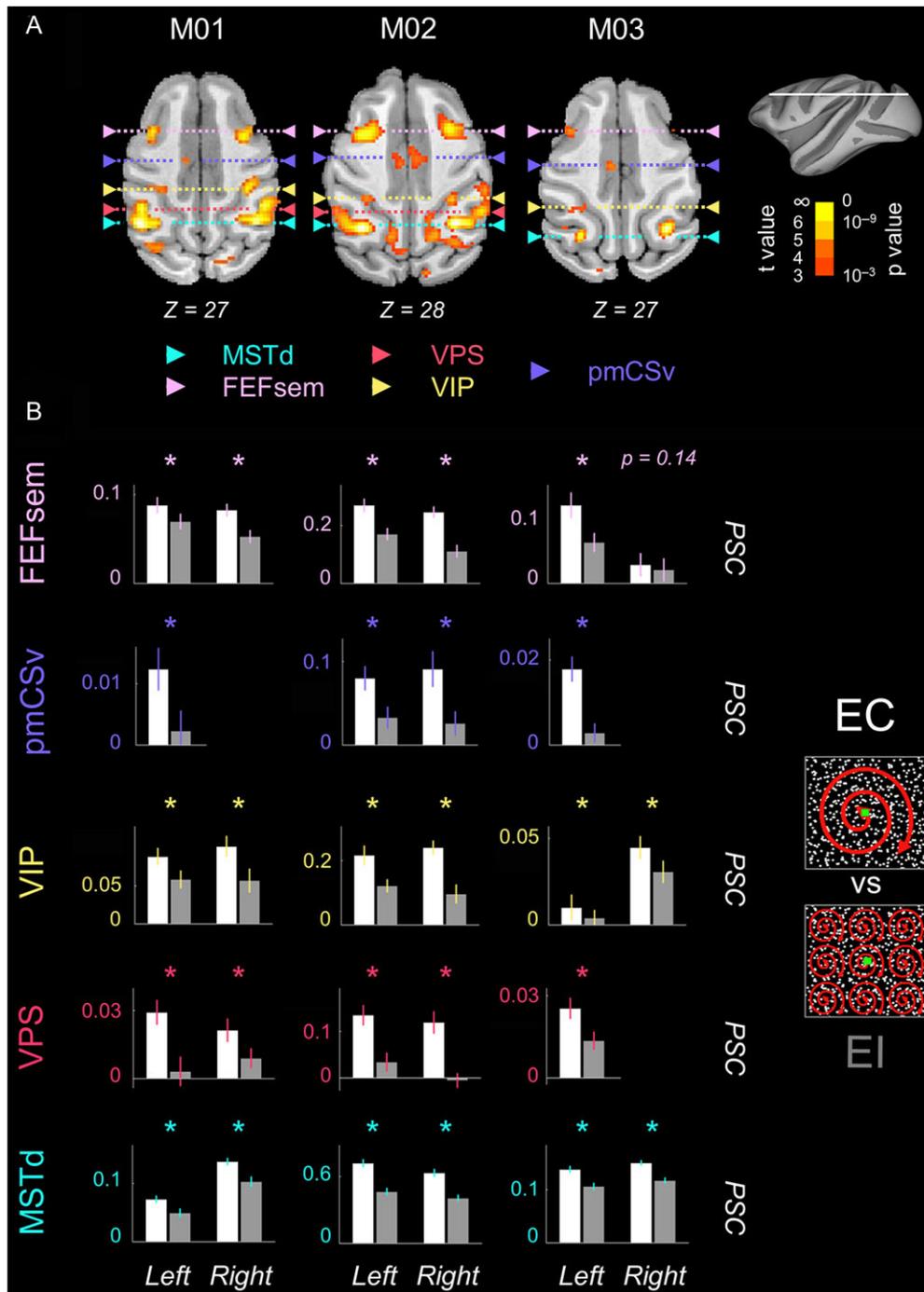


Figure 3. Activity profiles in areas MSTd, VPS, VIP, pmCSv, and FEFsem. (A) Statistical results of the EC > EI contrast shown on axial sections for monkeys M01, M02, and M03 (neurological convention). Areas are indicated by arrows on the 3 monkeys. (B) PSC in these 5 areas for the EC and EI conditions with respect to baseline (blank condition) in both hemispheres of the 3 macaques. The first half of the data was used to define ROIs around the local maxima of these areas and the second half was used to compute PSC (see details in the text). The error bars provide the standard errors across runs ($n = 18$). Stars indicate areas whose PSCs during the EC condition were significantly stronger than during the EI condition (t -tests, $P < 0.05$). P -values of the t -tests are provided for areas that did not pass significance.

distributed, encompassing the temporal, parietal, frontal, and cingulate cortices (Fig. 2). They are now discussed in more detail.

Activations in Temporal Cortex: MSTd and STPm

In all 6 recorded hemispheres, the most statistically significant activations for the contrast between EC and EI stimuli was found in a dorso-caudal portion of the sts, which corresponds

in macaque to area MSTd (Fig. 3). In order to check that these activations were well localized in MSTd and did not overlap with adjacent areas like MT, FST, or V4t, we performed an additional analysis from anatomical atlases provided in the Caret software (see Supplementary text 4 and Supplementary Fig. 6). This analysis confirmed that our activations are specific to MSTd. Numerous electrophysiological studies have shown that MSTd houses neurons selective to optic flow stimuli presented

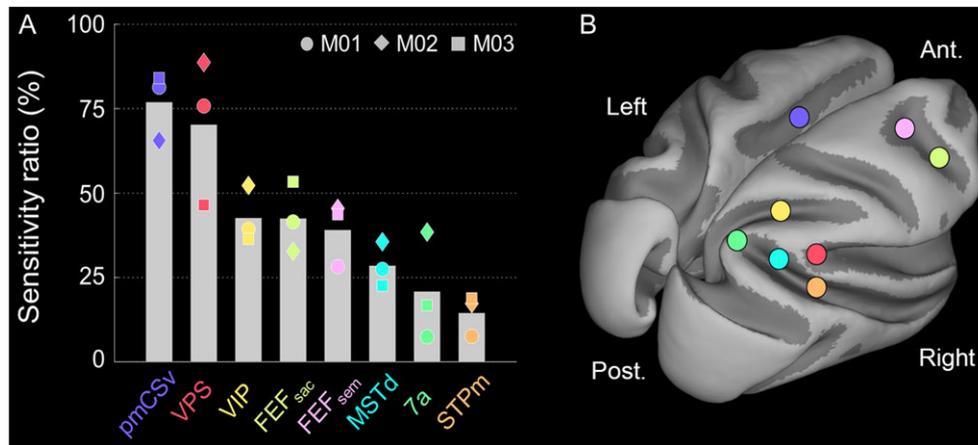


Figure 5. (A) Average sensitivity ratio (%) between the responses to the EC and EI conditions. The ratio (defined in the text) may be thought of as the reduction in response that occurs when an EC stimulus is replaced with EI. As for the PSC, ratios were computed on the second half of the data (see details in the text). Only areas with significant responses in the 3 animals are shown. Areas were sorted according to their mean sensitivity ratio. Markers provide the individual data corresponding to M01 (circles), M02 (diamonds), and M03 (squares). (B) Schematic localization of the 8 areas on the F99 template.

central role for MSTd in processing visual motion produced by self-displacements. Our data also revealed that the sensitivity ratio between responses to consistent and inconsistent flow stimuli is not very high (29%; see Fig. 5), suggesting that flow stimuli that are not consistent with self-displacements can nevertheless evoke strong responses in MSTd neurons. Interestingly, our results are in broad agreement with those obtained in hMST using the same experimental protocol (Wall and Smith 2008). These authors reported about 15% reduction in response to incompatible flow in hMST, compared with about 30% found here in macaque MSTd. This difference is consistent and could reflect a species difference, suggesting a greater specialization in macaque than human. However, it should be remembered that hMST in humans, which is defined simply in terms of the presence of strong ipsilateral drive (absent in hMT), probably does not correspond exactly to MSTd and may include other motion-sensitive regions with large receptive fields. It is therefore unsafe to make a direct comparison of results. Our results leave open the possibility that hMST, or some part of it, is homologous with MSTd for optic flow processing.

Another temporal activation site was observed in all the animals (5/6 hemispheres), situated more anteriorly along the fundus and the dorsal lip of the sts. This site may correspond to a subregion of the STP area, in which neurons selective to optic flow stimuli have been reported (Bruce et al. 1981; Anderson and Siegel 1999), although neurophysiological studies describe STP as located in the upper bank and fundus of sts whereas our cases show activity mainly in the lower bank (Fig. 2A), albeit with overlap across animals mainly in the fundus (Fig. 2B). In a recent monkey fMRI study, Nelissen et al. (2006) confirmed the existence of an optic-flow sensitive region in STP, that they named STPm and whose location is close to that found in the present study (Fig. 4), although again STPm is mainly in the upper bank of the sulcus. Interestingly, Nelissen et al. noted that responses to optic flow stimuli in STPm are similar to those of MSTd, except that the amplitude is lower. This difference is also found in the present study, together with a slightly lower sensitivity ratio (15%) of STPm. In human, a region within the sts and anterior to the hMT+ complex has been recently proposed as a putative homologue of macaque STP (Beauchamp et al. 2004b; Smith et al. 2012). This region was named STSms (STS multi-sensory) because of its multi-sensory

responses (Beauchamp et al. 2004a, 2004b, 2008). Among other modalities, STSms is activated by visual (Beauchamp et al. 2004b) and vestibular (Smith et al. 2012) signals. Therefore, it might be involved in the processing of EC optic flow. However, STSms was not significantly activated by our contrast in 2 human studies based on the same experimental protocol (Wall and Smith 2008; Cardin and Smith 2010). How can this be explained? One possibility might be simply that differential activity was missed in human STSms because of sensitivity limitations (e.g., the human studies used considerably fewer stimulus repetitions). Another is that STSms is not in fact homologous with STPm, or is broadly homologous but differs in its degree of specialization. Further investigations will be needed to clarify this point.

Activations in Parietal Cortex: VIP, 7a, and LIPd

In our 3 animals (6/6 hemispheres), we found statistically significant activation for the EC versus EI contrast in the fundus of the ips, which houses area VIP in macaque (Fig. 3). Together with MSTd, VIP is generally considered as playing a central role in processing heading information provided by both visual and vestibular signals (Bremmer et al. 2002a, 2002b). VIP and MSTd neurons seem to share many characteristics in the way they code both visual and inertial movements (Schaafsma and Duysens 1996; Chen et al. 2011a). However, the mean sensitivity ratio we measured in VIP (43%) was greater than that found in MSTd (Fig. 5). In human, the same contrast significantly activates a region within the anterior part of the ips (see e.g., Wall and Smith 2008) whose coordinates are very close to those of the polysensory motion sensitive area originally described by Bremmer et al. (2001) and proposed as a putative homologue of macaque VIP (hVIP). Wall and Smith (2008) reported a 46% response reduction for EI stimuli, very similar to our result for macaque VIP. Our data are therefore consistent with the hypothesis of a correspondence between these 2 areas, although caution is needed because the ips is organized differently in humans and macaques, with several more areas in humans.

Besides VIP, a consistent site of parietal activation, observed in all our monkeys (5/6 hemispheres), was located within area 7a, which occupies an elongated posterior portion of the inferior parietal lobule (Fig. 4). Area 7a is involved in spatial vision,

through the integration of visual and oculomotor signals (Mishkin et al. 1982; Anderson and Siegel 1999). Both electrophysiological and optical imaging studies have shown that 7a neurons respond selectively to optic flow stimuli (Siegel and Read 1997; Phinney and Siegel 2000; Merchant et al. 2001; Raffi and Siegel 2007). More recently, cytoarchitectonic differences along the inferior parietal lobule have led to its subdivision into 4 sectors: PF, PFG, PG, and Opt (Pandya and Seltzer 1982; Gregoriou et al. 2006). The posterior location of our activation site, together with the fact that it extends ventrally into the dorsal bank of the STS, strongly suggest that it corresponds to the caudal-most region Opt (Gregoriou et al. 2006). Interestingly, tracer injections in Opt (Rozzi et al. 2006) revealed strong connections with the temporal areas that we found to be involved in processing egomotion-compatible optic flow: MSTd and STP, and much weaker connections with the neighboring temporal areas MT and FST. The same study revealed that Opt is also connected to LIP in the ips and to area 23, in the cingulate sulcus, 2 sites that also responded more strongly to consistent than inconsistent optic flow stimuli in a majority of recorded hemispheres, as will be discussed below.

The dorsal part of the lateral intraparietal (LIPd) area was the third site of parietal activation evidenced in the present study (see Fig. 3 and Supplementary Fig. 4). Results were less systematic than those of area 7a/Opt, with 4 out of 6 hemispheres, but closely resembled those found in VIP (with an average sensitivity ratio of 47%). To our knowledge, there is no previous study linking LIPd to the specific processing of egomotion-compatible optic flow. However, the present results, together with the fact that LIPd is connected to 7a/Opt, argue that the possible role of LIPd in optic flow processing deserves further investigation.

Several studies have reported the existence of a possible homologue of macaque area LIP in the human superior parietal cortex (Serenio et al. 2001; Shikata et al. 2008). To our knowledge, the human homologue of area 7a has not been firmly established. In any case, the only robust and reliable activations that were found near ips in human using the same protocol corresponded to area hVIP (see above). Our results therefore suggest that processing of optic flow in 7a (and LIPd, if confirmed) may be specific to macaque.

Activation in Parieto-Insular Cortex: VPS

In our 3 individuals (5/6 hemispheres) stronger BOLD responses for the EC optic flow stimuli were observed in the caudal portion of the sylvian fissure (Fig. 3). The location of this activation site corresponds to the VPS area, which is posterior to the PIVC from which it receives vestibular inputs. VPS is also connected to MSTd (Guldin et al. 1992), which may feed VPS with visual optic flow information. In agreement with this view, VPS neurons have been shown to integrate heading-related information from both visual optic-flow and vestibular signals (Chen et al. 2011b). Importantly, our results revealed that after pmCSv, VPS is actually the cortical region exhibiting the greatest sensitivity ratio for EC stimuli (70%, see Fig. 5), much higher than those found in the temporal and parietal activation sites described so far. In human, the same contrast (Cardin and Smith 2010) also revealed a parieto-insular region, PIC (that was originally mistakenly labeled as PIVC), sharing most of the properties described here for VPS and notably its responsiveness to vestibular inputs (Smith et al. 2012). The sensitivity ratio in PIC is very high (~80%; Cardin and Smith 2010) and close to the one

we found in macaque VPS. Altogether, these results further support the idea that these 2 regions are homologous.

Activations in Frontal Cortex: FEFsem and FEFsac

In 3 animals (5/6 hemispheres), strong activations were found in the dorsal part of the *as* (Fig. 3). This location matches that of FEFsem, a subregion of the frontal eye field involved in the control of smooth pursuit eye movements (Lynch 1987; MacAvoy et al. 1991). The average sensitivity ratio we found in FEFsem (39%) was about the same as those of MSTd, VIP and 7a/Opt, which all share strong recurrent connections with FEFsem (Boussaoud et al. 1990; Maioli et al. 1998; Stanton et al. 2005). Recently, FEFsem neurons have been shown to respond selectively to visual and vestibular signals induced by self-displacements (Gu et al. 2015). Thus, the present results provide further evidence that FEFsem processes heading information.

In 4/6 hemispheres, (3 individuals), a second site of activation was observed within the *as*. Located slightly more anterior and lateral than FEFsem, within the fundus and anterior branch of the *as* (Fig. 4), it nicely fits the anatomical location reported for another subregion of the frontal eye field, FEFsac, which is involved in saccadic eye movements (MacAvoy et al. 1991; Gu et al. 2015). Our analyses of eye movements reveal differences neither in the quality of fixation nor in the number of saccades evoked by EC and EI stimuli, in any of the 3 individuals (see Supplementary text 2 and Supplementary Fig. 1). Thus, the present results argue for a secondary role of FEFsac in the processing of visual motion induced by self-displacements.

In human, studies that used the same experimental protocol (Wall and Smith 2008; Cardin and Smith 2010) did not report any significant activation in or around the frontal eye field region. A re-examination of these data revealed that a few subjects (<20%) actually had significant responses in FEF. This low proportion makes it difficult to determine whether these activations were false positive. Either way, this comparison across species supports the idea that the implication of the FEFsem and FEFsac regions in optic flow processing is at least more pronounced in macaque.

Activation in Cingulate Cortex: pmCSv

In human, a growing number of fMRI studies have described a region within the cingulate sulcus, CSv, which is significantly activated by complex motion patterns (Wall and Smith 2008; Fischer et al. 2012; Pitzalis et al. 2013; Schindler and Bartels 2016). Using the stimuli of the present study, Cardin and Smith (2010) reported that this region had the greatest specificity for EC optic flow, in virtually all the tested subjects. In all our animals (4/6 hemispheres), we measured strong responses in the posterior cingulate sulcus (Fig. 3). Atlas-based comparison in Caret software indicates that this activation site belongs to area 23c (Vogt et al. 2005), which is thought to be involved in spatial vision notably through its connection to area 7a/Opt (Vogt et al. 1992; Rozzi et al. 2006). A striking feature of the activity profiles for the 4 significant hemispheres is the near absence of response to EI stimuli (Fig. 3) which leads to a very high sensitivity ratio (77%, see Fig. 5), as has been found in human CSv (Wall and Smith 2008). Although other studies will be needed to confirm the possible homology between pmCSv and CSv, the identification of pmCSv provides new opportunities to understand activations in human CSv by reference to electrophysiological explorations in macaque.

Homologies with the Human EC-Selective Areas V6, Pc, and 2v?

In human, the contrast used in this study leads to consistent activations in the parieto-occipital cortex (Cardin and Smith 2010), in a site that both retinotopic mapping and response properties point to as the human homologue of macaque area V6 (Pitzalis et al. 2006, 2010; Cardin et al. 2012b). In the present study, our voxel-wise analysis did not reveal any evidence that V6 prefers EC optic flow. However, a suggestive trend was observed in the atlas-based approach (see Supplementary text 4 and Supplementary Fig. 6), which shows a sensitivity ratio for V6 of about 20% that is consistent across the 3 animals. This leaves open the possibility that V6 does possess EC selectivity but that this was not reliable enough to be detected at the voxel level. It is also possible that there is no such selectivity and the trend arises from erroneous inclusion of parts of neighboring visual regions, although the immediate neighbors V2, V3, and V3A do not themselves show strong specificity. Whatever the explanation, our results indicate that selectivity to the EC condition in macaque V6 is not as robust as in hV6. One possible interpretation of this discrepancy might be that human and monkey V6 differ regarding their involvement in heading processing, the involvement being greater in humans. More generally, this observation reinforces the view that the cortical processing of visual motion might differ in several aspects between these 2 primate species (Vanduffel et al. 2002; Orban et al. 2003).

Finally, it should be noted that 2 other human cortical regions have been reported to show selective responses to EC stimuli (Cardin and Smith 2010) that do not appear to do so in macaque. The first is a possible homologue of macaque area 2v in anterior parietal cortex, which has been shown to receive vestibular afferents and may therefore process self-motion. The second is an anterior region of the precuneus, termed Pc by these authors. Since little is known about either region, it is difficult to interpret these differences. At least in Pc, and possibly also in putative human 2v, the difference between EC and EI reflects differences in visual suppression rather than visual responses.

Conclusion

Overall, our results are in excellent agreement with the electrophysiological and anatomical data collected over recent decades in macaque monkey. They demonstrate that a simple contrast between optic flow stimuli that are consistent or inconsistent with self-displacements can reveal the vast majority of cortical areas known to be involved in processing heading information through optic flow, including those also thought to integrate vestibular inputs. An advantage of the monkey fMRI approach is that it allows a direct comparison with results obtained in several human studies based on the same contrast (Wall and Smith 2008; Cardin and Smith 2010). Together, the data collected in the 2 species suggest that although the networks processing optic flow in human and macaque share some properties (i.e., possible homologies between areas MST/hMST, VIP/hVIP, VPS/PIC, and pmCSv/CSv), they nonetheless remain different. On the one hand, some of the significantly activated areas in macaque (7a, STPm, FEFsem and also FEFsac) were not found in human. On the other hand, some areas robustly found in humans (V6, P2v, Pc) did not show significant activations in the present study.

Supplementary Material

Supplementary material are available at *Cerebral Cortex* online.

Funding

CNRS, from the ANR (ANR-13-JSV4-0007-01; ANR-12-BSV4-0005); European Council (Marie Curie grant PIFI-GA-2011-298386 Real-Depth).

Notes

The authors thank A. Sadoun for his help during the surgery and the technical personnel of the Cerco monkey facility for their help with the handling of the animals. We also thank the radiographers of the INSERM U214 TONIC MRI platform for their help with the (f)MRI recordings and M. Furlan for help with generating visual stimuli. *Conflict of Interest:* None declared.

References

- Anderson KC, Siegel RM. 1999. Optic flow selectivity in the anterior superior temporal polysensory area, STPa, of the behaving monkey. *J Neurosci.* 19:2681–2692.
- Antal A, Baudewig J, Paulus W, Dechent P. 2008. The posterior cingulate cortex and planum temporale/parietal operculum are activated by coherent visual motion. *Vis Neurosci.* 25:17–26.
- Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A. 2004a. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci.* 7:1190–1192.
- Beauchamp MS, Lee KE, Argall BD, Martin A. 2004b. Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron.* 41:809–823.
- Beauchamp MS, Yasar NE, Frye RE, Ro T. 2008. Touch, sound and vision in human superior temporal sulcus. *Neuroimage.* 41:1011–1020.
- Boussaoud D, Ungerleider LG, Desimone R. 1990. Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *J Comp Neurol.* 296:462–495.
- Bremmer F, Duhamel J-R, Ben Hamed S, Graf W. 2002a. Heading encoding in the macaque ventral intraparietal area (VIP). *Eur J Neurosci.* 16:1554–1568.
- Bremmer F, Klam F, Duhamel J-R, Ben Hamed S, Graf W. 2002b. Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Eur J Neurosci.* 16:1569–1586.
- Bremmer F, Schlack A, Shah NJ, Zafiris O, Kubischik M, Hoffmann K-P, Zilles K, Fink GR. 2001. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron.* 29:287–296.
- Britten KH, van Wezel RJA. 2002. Area MST and heading perception in macaque monkeys. *Cereb Cortex.* 12:692–701.
- Bruce C, Desimone R, Gross CG. 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J Neurophysiol.* 46:369–384.
- Cardin V, Hemsworth L, Smith AT. 2012a. Adaptation to heading direction dissociates the roles of human MST and V6 in the processing of optic flow. *J Neurophysiol.* 108:794–801.
- Cardin V, Sherrington R, Hemsworth L, Smith AT. 2012b. Human V6: functional characterisation and localisation. *PLoS One.* 7:e47685.
- Cardin V, Smith AT. 2010. Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. *Cereb Cortex.* 20:1964–1973.
- Chen A, DeAngelis GC, Angelaki DE. 2011a. Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *J Neurosci.* 31:12036–12052.

- Chen A, DeAngelis GC, Angelaki DE. 2011b. Convergence of vestibular and visual self-motion signals in an area of the posterior sylvian fissure. *J Neurosci.* 31:11617–11627.
- Chen A, Gu Y, Liu S, DeAngelis GC, Angelaki DE. 2016. Evidence for a causal contribution of macaque vestibular, but not intraparietal, cortex to heading perception. *J Neurosci.* 36:3789–3798.
- Duffy CJ. 1998. MST neurons respond to optic flow and translational movement. *J Neurophysiol.* 80:1816–1827.
- Duffy CJ, Wurtz RH. 1991. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J Neurophysiol.* 65:1329–1345.
- Duffy CJ, Wurtz RH. 1995. Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J Neurosci.* 15:5192–5208.
- Duffy FH, Burchfiel JL, Conway JL. 1976. Bicuculline reversal of deprivation amblyopia in the cat. *Nature.* 260:256–257.
- Dukelow SP, DeSouza JF, Culham JC, van den Berg AV, Menon RS, Vilis T. 2001. Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J Neurophysiol.* 86:1991–2000.
- Fan RH, Liu S, DeAngelis GC, Angelaki DE. 2015. Heading tuning in Macaque Area V6. *J Neurosci.* 35:16303–16314.
- Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE. 2012. Neural correlates of reliability-based cue weighting during multisensory integration. *Nat Neurosci.* 15:146–154.
- Fischer E, Bühlhoff HH, Logothetis NK, Bartels A. 2012. Visual motion responses in the posterior cingulate sulcus: a comparison to V5/MT and MST. *Cereb Cortex.* 22:865–876.
- Frank SM, Baumann O, Mattingley JB, Greenlee MW. 2014. Vestibular and visual responses in human posterior insular cortex. *J Neurophysiol.* 112:2481–2491.
- Galletti C, Gamberini M, Kutz DF, Fattori P, Luppino G, Matelli M. 2001. The cortical connections of area V6: an occipitoparietal network processing visual information. *Eur J Neurosci.* 13:1572–1588.
- Gregoriou GG, Borra E, Matelli M, Luppino G. 2006. Architectonic organization of the inferior parietal convexity of the macaque monkey. *J Comp Neurol.* 496:422–451.
- Grossberg S, Mingolla E, Pack C. 1999. A neural model of motion processing and visual navigation by cortical area MST. *Cereb Cortex.* 9:878–895.
- Gu Y, Cheng Z, Yang L, DeAngelis GC, Angelaki DE. 2015. Multisensory convergence of visual and vestibular heading cues in the pursuit area of the frontal eye field. *Cereb Cortex.* 26:3785–3801.
- Gu Y, DeAngelis GC, Angelaki DE. 2012. Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. *J Neurosci.* 32:2299–2313.
- Gu Y, Watkins PV, Angelaki DE, DeAngelis GC. 2006. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J Neurosci.* 26:73–85.
- Guizar-Sicairos M, Thurman ST, Fienup JR. 2008. Efficient subpixel image registration algorithms. *Opt Lett.* 33:156–158.
- Guldin WO, Akbarian S, Grüsser O-J. 1992. Cortico-cortical connections and cytoarchitectonics of the primate vestibular cortex: a study in squirrel monkeys (*Saimiri sciureus*). *J Comp Neurol.* 326:375–401.
- Huk AC, Dougherty RF, Heeger DJ. 2002. Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci.* 22:7195–7205.
- Kolster H, Peeters R, Orban GA. 2010. The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *J Neurosci.* 30:9801–9820.
- Komatsu H, Wurtz RH. 1988. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J Neurophysiol.* 60:580–603.
- Lynch JC. 1987. Frontal eye field lesions in monkeys disrupt visual pursuit. *Exp Brain Res.* 68:437–441.
- MacAvoy MG, Gottlieb JP, Bruce CJ. 1991. Smooth-pursuit eye movement representation in the primate frontal eye field. *Cereb Cortex.* 1:95–102.
- Maioli MG, Squatrito S, Samolsky-Dekel BG, Riva Sanseverino E. 1998. Corticocortical connections between frontal periarculate regions and visual areas of the superior temporal sulcus and the adjoining inferior parietal lobule in the macaque monkey. *Brain Res.* 789:118–125.
- McLaren DG, Kosmatka KJ, Kastman EK, Bendlin BB, Johnson SC. 2010. Rhesus macaque brain morphometry: a methodological comparison of voxel-wise approaches. *Methods.* 50:157–165.
- McLaren DG, Kosmatka KJ, Oakes TR, Kroenke CD, Kohama SG, Matochik JA, Ingram DK, Johnson SC. 2009. A population-average MRI-based atlas collection of the rhesus macaque. *Neuroimage.* 45:52–59.
- Mendoza-Halliday D, Torres S, Martinez-Trujillo JC. 2014. Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat Neurosci.* 17:1255–1262.
- Merchant H, Battaglia-Mayer A, Georgopoulos AP. 2001. Effects of optic flow in motor cortex and area 7a. *J Neurophysiol.* 86:1937–1954.
- Mineault PJ, Khawaja FA, Butts DA, Pack CC. 2012. Hierarchical processing of complex motion along the primate dorsal visual pathway. *Proc Natl Acad Sci USA.* 109:E972–E980.
- Mishkin M, Lewis ME, Ungerleider LG. 1982. Equivalence of parieto-preoccipital subareas for visuospatial ability in monkeys. *Behav Brain Res.* 6:41–55.
- Morrone MC, Tosetti M, Montanaro D, Fiorentini A, Cioni G, Burr DC. 2000. A cortical area that responds specifically to optic flow, revealed by fMRI. *Nat Neurosci.* 3:1322–1328.
- Motter BC, Mountcastle VB. 1981. The functional properties of the light-sensitive neurons of the posterior parietal cortex studied in waking monkeys: foveal sparing and opponent vector organization. *J Neurosci.* 1:3–26.
- Nelissen K, Vanduffel W, Orban GA. 2006. Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. *J Neurosci.* 26:5929–5947.
- Orban GA. 2002. Functional MRI in the awake monkey: the missing link. *J Cogn Neurosci.* 14:965–969.
- Orban GA, 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.* 41:1757–1768.
- Orban GA, Van Essen D, Vanduffel W. 2004. Comparative mapping of higher visual areas in monkeys and humans. *Trends Cogn Sci.* 8:315–324.
- Pandya DN, Seltzer B. 1982. Intrinsic connections and architectonics of posterior parietal cortex in the rhesus monkey. *J Comp Neurol.* 204:196–210.
- Perrone JA, Stone LS. 1994. A model of self-motion estimation within primate extrastriate visual cortex. *Vision Res.* 34:2917–2938.
- Phinney RE, Siegel RM. 2000. Speed selectivity for optic flow in area 7a of the behaving macaque. *Cereb Cortex.* 10:413–421.
- Pitzalis S, Galletti C, Huang R-S, Patria F, Committeri G, Galati G, Fattori P, Sereno MI. 2006. Wide-field retinotopy defines human cortical visual area V6. *J Neurosci.* 26:7962–7973.

- Pitzalis S, Sdoia S, Bultrini A, Committeri G, Di Russo F, Fattori P, Galletti C, Galati G. 2013. Selectivity to translational egomotion in human brain motion areas. *PLoS One*. 8:e60241.
- Pitzalis S, Sereno MI, Committeri G, Fattori P, Galati G, Patria F, Galletti C. 2010. Human v6: the medial motion area. *Cereb Cortex*. 20:411–424.
- Raffi M, Siegel RM. 2007. A functional architecture of optic flow in the inferior parietal lobule of the behaving monkey. *PLoS One*. 2:e200.
- Reppas JB, Niyogi S, Dale AM, Sereno MI, Tootell RB. 1997. Representation of motion boundaries in retinotopic human visual cortical areas. *Nature*. 388:175–179.
- Rozzi S, Calzavara R, Belmalih A, Borra E, Gregoriou GG, Matelli M, Luppino G. 2006. Cortical connections of the inferior parietal cortical convexity of the macaque monkey. *Cereb Cortex*. 16:1389–1417.
- Schaafsma SJ, Duysens J. 1996. Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. *J Neurophysiol*. 76:4056–4068.
- Schindler A, Bartels A. 2016. Connectivity reveals sources of predictive coding signals in early visual cortex during processing of visual optic flow. *Cereb Cortex*. pii:bhw136 [Epub ahead of print].
- Sereno MI, Pitzalis S, Martinez A. 2001. Mapping of contralateral space in retinotopic coordinates by a parietal cortical area in humans. *Science*. 294:1350–1354.
- Shikata E, McNamara A, Sprenger A, Hamzei F, Glauche V, Büchel C, Binkofski F. 2008. Localization of human intraparietal areas AIP, CIP, and LIP using surface orientation and saccadic eye movement tasks. *Hum Brain Mapp*. 29:411–421.
- Siegel RM, Read HL. 1997. Analysis of optic flow in the monkey parietal area 7a. *Cereb Cortex*. 7:327–346.
- Smith AT, Wall MB, Thilo KV. 2012. Vestibular inputs to human motion-sensitive visual cortex. *Cereb Cortex*. 22:1068–1077.
- Smith AT, Wall MB, Williams AL, Singh KD. 2006. Sensitivity to optic flow in human cortical areas MT and MST. *Eur J Neurosci*. 23:561–569.
- Stanton GB, Friedman HR, Dias EC, Bruce CJ. 2005. Cortical afferents to the smooth-pursuit region of the macaque monkey's frontal eye field. *Exp Brain Res*. 165:179–192.
- Steinmetz MA, Motter BC, Duffy CJ, Mountcastle VB. 1987. Functional properties of parietal visual neurons: radial organization of directionalities within the visual field. *J Neurosci*. 7:177–191.
- Takahashi K, Gu Y, May PJ, Newlands SD, DeAngelis GC, Angelaki DE. 2007. Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. *J Neurosci*. 27:9742–9756.
- Tanaka K, Fukada Y, Saito HA. 1989. Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol*. 62:642–656.
- Van Essen DC. 2002. Surface-based atlases of cerebellar cortex in the human, macaque, and mouse. *Ann NY Acad Sci USA*. 978:468–479.
- Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH. 2001. An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inform Assoc*. 8:443–459.
- Vanduffel W, Farivar R. 2014. Functional MRI of awake behaving macaques using standard equipment. In: Duric D, editor. *Advanced Brain Neuroimaging Topics in Health and Disease – Methods and Applications*. InTech, DOI: 10.5772/58281. Available from: <http://www.intechopen.com/books/advanced-brain-neuroimaging-topics-in-health-and-disease-methods-and-applications/functional-mri-of-awake-behaving-macaques-using-standard-equipment>.
- Vanduffel W, Fize D, Peuskens H, Denys K, Sunaert S, Todd JT, Orban GA. 2002. Extracting 3D from motion: differences in human and monkey intraparietal cortex. *Science*. 298:413–415.
- Vogt BA, Finch DM, Olson CR. 1992. Functional heterogeneity in cingulate cortex: the anterior executive and posterior evaluative regions. *Cereb Cortex*. 2:435–443.
- Vogt BA, Vogt L, Farber NB, Bush G. 2005. Architecture and neurocytology of monkey cingulate gyrus. *J Comp Neurol*. 485:218–239.
- Wall MB, Smith AT. 2008. The representation of egomotion in the human brain. *Curr Biol*. 18:191–194.
- Yu CP, Page WK, Gaboriski R, Duffy CJ. 2010. Receptive field dynamics underlying MST neuronal optic flow selectivity. *J Neurophysiol*. 103:2794–2807.
- Zhang T, Britten KH. 2011. Parietal area VIP causally influences heading perception during pursuit eye movements. *J Neurosci*. 31:2569–2575.