bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Encoding of motion by spatial and temporal frequency in the human middle temporal cortex measured by 7T fMRI

Anna Gaglianese[1,2,3] , Alessio Fracasso[4], Francisco G. Fernandes[2], Ben Harvey[5], Serge O. Dumoulin[5,6,7], Natalia Petridou[3]

1. *The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, University Hospital Center and University of Lausanne, Lausanne, Switzerland*
2. *Department of Neurosurgery and Neurology, UMC Utrecht Brain Center, University Medical Center Utrecht, Netherlands*
3. *Department of Radiology, Center for Image Sciences, University Medical Center Utrecht, Netherlands*
4. *University of Glasgow, Institute of Neuroscience and Psychology, UK*
5. *Experimental Psychology, Helmholtz Institute, Utrecht University, Netherlands*
6. *Spinoza Center for Neuroimaging, Amsterdam, Netherlands*
7. *Experimental and Applied Psychology, VU University, Amsterdam, Netherlands*

Correspondence should be addressed to Anna Gaglianese, Department of Radiology, University Hospital Center and University of Lausanne, Rue Centrale 7, 1003, Lausanne, Switzerland. Email: anna.gaglianese@chuv.ch, anna.gaglianese@gmail.com and Natalia Petridou, Imaging Division University Medical Center Utrecht, Department of Radiology, Heidelberglaan 100, 3584 CX Utrecht, Netherlands, n.petridou@umcutrecht.nl.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Perception of environmental dynamic scenes results from the evaluation of visual features such as the fundamental spatial and temporal frequencies components of a moving object. The ratio between these two components represents its speed of motion. The human middle temporal cortex hMT+ has a crucial biological role in the direct encoding of object speed. However, the link between hMT+ speed encoding and the spatiotemporal frequency components of a moving object may be more complex than we thought. Both animal studies and recent human electrocorticography data showed that recorded neuronal populations within MT+/V5 change their speed preferences in accordance with the stimulus fundamental spatial frequency. We disentangle whether such mechanism holds for the entire human MT+. We recorded high resolution 7T blood oxygen level-dependent BOLD responses to different visual motion stimuli as a function of their fundamental spatial and temporal frequency components. We fitted each hMT+ BOLD response with a 2D Gaussian model allowing for distinct and independent selectivity for spatial and temporal frequencies of the visual stimuli or tuning for the speed of motion. We show that: 1) hMT+ encodes the speed of motion via independent tuning of the fundamental spatial frequency component of the visual stimuli, 2) the optimal spatial frequency selectivity of hMT+ is tuned for the low spatial frequency of the visual stimuli and is highly reproducible within subjects. Our results show that hMT+ speed preference changes according to the fundamental spatial frequency presented, demonstrating a primary role of the entire hMT+ in the evaluation of the spatial features of the moving visual input. These findings confirm a more complex mechanism involved in the direct perception of speed than initially thought.

Keywords:

7T fmri, motion, gaussian modelling, hMT+, spatial frequency

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Highlights

- hMT+ speed preference changes in accordance with the fundamental spatial frequency
- A small subregion within the complex encodes speed directly
- A reproducible selectivity for the low spatial frequency of the stimuli was found

Introduction

Encoding of visual features from dynamic visual images is essential in humans and nonhuman primates to reconstruct the visual scene and rapidly respond to the ever changing environment. Among visual areas, the human homologue of the macaque middle temporal cortex (hMT+ also known as V5) has been shown to play a functional role in the decoding of features such as the spatial and temporal frequency components of visual motion stimuli. Using electrocorticography, we recently showed that hMT+ neuronal populations separated motion into its spatial and temporal components, with speed preferences changing in accordance with the fundamental spatial frequency of the visual stimuli, rather than being tuned for a particular speed of the attended moving stimuli (Gaglianese et al., 2017). These findings, paired with single cell recording studies in animals, describe hMT+ neurons as spatiotemporal frequency sensors for motion extraction (Lui et al., 2007; Priebe et al., 2006, 2003). However, because animal single-neuron recordings and human electrocorticography do not cover MT+ with closely-spaced recording sites, it remains elusive whether there is a functional organization within the complex for the different mechanisms of speed encoding, i.e. separable responses for spatial and temporal frequencies vs pure speed tuning. The rapid development of ultra-high field (7 Tesla, 7T) functional Magnetic Resonance Imaging (fMRI) allows us to reveal the fine-scale functional organization of the human cortex in vivo (Dumoulin et al., 2018, 2017; Formisano et al., 2003; Fracasso et al., 2016; Petridou and Siero, 2019; Yacoub et al., 2008). Many 7T fMRI studies have been carried out in primary visual cortex V1, although functional

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

studies have been recently extended to reveal the fine-scale functional organization of the human extrastriate cortex and association areas (Dumoulin et al., 2017; Harvey et al., 2020, 2013; Harvey and Dumoulin, 2016; Huber et al., 2020; Klein et al., 2014; Schellekens et al., 2018; Tootell and Nasr, 2017; Zimmermann et al., 2011). A recent high spatial resolution 7T fMRI study in hMT+ in particular, has demonstrated an organization into columnar clusters for preferences for horizontal and vertical motion similar to the columnar organization in monkeys (Schneider et al., 2019). However, human research to date has tended to focus on the organisation of responses to the location and direction of motion in hMT+, rather than the mechanisms involved in the encoding of speed of motion.

Here, we disentangle whether and to what extent the hMT complex is tuned for the speed of motion or is dependent on the fundamental spatiotemporal properties of the visual motion stimuli. We used high resolution 7T fMRI to characterize the organization of hMT+ Blood Oxygenation Level Dependent (BOLD) response amplitudes for different combinations of fundamental spatial and temporal frequency components of visual motion stimuli. By modelling the BOLD responses with a 2D Gaussian model for speed tuning (Gaglianese et al., 2017) we were able to characterize the mechanisms involved in the encoding of speed of motion and to detect reliable spatial frequency selectivity tuning within the complex.

Methods

hMT+ localizer stimulus

Area hMT+ was functionally identified based on responses to moving compared to stationary high contrast black and white dartboards, as conventionally used in literature (Huk et al., 2002; Tootell et al., 1995). We used dartboards in place of standard random dots to match the contrast of the visual stimulus used in the visual motion stimulation experiment. During the motion condition, the full field high contrast black-and-white dartboards were contrast reversed at 5 Hz

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for 10s, interleaved with a stationary period of 10s in which the same static dartboard was presented. The stimuli subtended a visual angle of 30.7x16.1°.

Visual motion stimulation

The visual motion stimulation consisted on five runs of high-contrast square-wave black and white dartboards with different fundamental spatial and temporal frequency combinations (0.33 cycle/deg;1Hz, 0.33 cycle/deg; 3Hz, 0.33 cycle/deg; 5Hz, 0.2 cycle/deg; 3Hz, 1 cycle/deg; 3Hz). Given that the speed of motion of each square-wave dartboard presented is defined by the ratio of temporal to spatial frequencies, speeds of 3deg/sec, 9 deg/sec, and 15deg/sec were presented respectively. Each run is either classified as a fast (15 deg/sec), intermediate (9 deg/sec) or slow (3 deg/sec) moving stimuli, depending on the fundamental spatial and temporal frequency that gives origin to the stimuli speed. The fast and the slow speed respectively were presented twice by using two different spatiotemporal frequency combinations of the moving dartboards (see fig. 1). For each run, we presented only one spatiotemporal frequency combination for a total of 26 trials. The dartboard pattern expanded for 1s alternating with stationary periods of variable length ranging from 6s to 15s, in a pseudo-randomized order. Three additional inter stimulus intervals of 24s were randomly added to allow the BOLD response to return to baseline.

fMRI methods

fMRI data acquisition

MR images were acquired from five participants using a Philips 7T scanner equipped with a volume transmit (Nova Medical, USA) and two high-density16-channel surface coils (Petridou et al., 2013, MR Coils BV, NL). Surface coils covered each lateral occipital pole of the participant to maximize the signal-to-noise (SNR) and BOLD sensitivity in the hMT+ area under examination. For both the localizer and the visual motion stimulation spatiotemporal

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

frequencies experiments a gradient echo EPI sequence was used. Functional images for the localizer were acquired every 1.8s, with an echo time (TE) of 27ms, an isotropic voxel of 1.5mm and 27 coronal slices covering hMT+. For the visual motion stimulation experiment we acquired 15 coronal slices at a fast temporal resolution of 0.849ms, with an isotropic voxel resolution of 1.4mm. For both acquisitions, functional EPIs were acquired with a SENSE factor of 2 in the right-left direction. High-resolution T1-weighted anatomical MRI images were acquired with a 32-channel head coil (Nova Medical, MA, USA) in a different session at a resolution of 0.8x0.8x0.8 mm. Repetition time (TR) was 7 ms, TE was 2.84 ms, and flip angle was 8 degrees.

Data pre-processing

To reduce noise-related signal components, a series of mathematical operations was set and performed. All the pre-processing steps were computed using AFNI (Analysis of Functional NeuroImages, https://afni.nimh.nih.gov/). First, the functional data of the localizer and visual motion stimulation were corrected for motion and aligned to the first image of the first run of each session respectively. Subsequently removal of low frequency signal intensity drift was applied by a quadratic detrending via the 3dDetrend function in AFNI. No spatial smoothing was employed. The visual motion stimulus runs were non linearly co-registered to the localizer via 3WarpDrive function in AFNI. To avoid time series interpolation, we extracted the Regions of Interest (ROIs) on the localizer space and co-registered them to the visual motion stimulus run using the inverse of the obtained transformation matrix.

For visualization purpose the T1-w anatomical images were segmented automatically using the MIPAV software package implemented in CBStool (https://www.nitrc.org/projects/cbs-tools/). White matter and pial surfaces were generated and then imported in SUMA (afni.nimh.nih.gov).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Co-registration of the visual motion stimulation results was performed using first the function 3dAllineate in AFNI, with mutual information as cost function and then non linearly via the 3dWarpdrive function in AFNI.

Localization of hMT+

For each participant, left and right hMT+ areas were functionally defined from the localizer runs by contrasting responses for the high contrast black and white dartboard stimuli that alternated in time between moving and stationary periods. All statistical computations were performed at a single participant level using a general linear model (GLM) with a standard gamma variate hemodynamic response function approach, using the 3dDeconvolve function in AFNI. For each run, outliers due to residual motion were detected via 3dTOutCount function and included in the GLM analysis as regressors of no interest. Voxels that exhibited significant responses for moving vs stationary dartboards ($p<0.001$, Bonferroni corrected) and located within the hMT+ anatomical landmarks (Dumoulin et al., 2000) on the EPI space were selected to define hMT+ ROIs.

Quantification of BOLD responses for the visual motion stimulation

The BOLD responses to each combination of spatial and temporal frequencies presented in the visual motion stimulation were estimated for all voxels using a finite impulse response deconvolution approach (described in Costagli et al., 2014; Dale, 1999; Gardner et al., 2005) implemented in mrTools (available for free download at http://gru.stanford.edu/doku.php/mrTools/overview), a software package running in MATLAB. The response to a given stimulus type was quantified for each voxel by the amplitude of the BOLD response. Amplitude values higher than 6 were discarded from the analysis to avoid contamination from low-specificity signal, dominated by pial vein contributions (Sun et al., 2007). Only amplitude values within each hMT+ ROI were selected

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for further analysis. We computed the significance of the effect of the spatial and temporal frequency components of the visual motion stimulation on the hMT+ BOLD response amplitudes by two-way ANOVA within participants. Furthermore, to investigate how the BOLD response amplitude changes according to the spatiotemporal frequency combination of the stimuli rather than the speed per se, we performed a two-sample t-test between the 3deg1Hz and 1deg 3Hz spatiotemporal frequency combination (both representing 3deg/sec speed of motion of the visual stimuli) and 3deg5Hz and 5deg3Hz (both representing 15deg/sec speed of motion).

Modeling of hMT+ BOLD responses for speed tuning

We ask whether the BOLD responses in hMT+ depend on the fundamental spatiotemporal frequency combination of the presented moving visual stimuli or whether they are based on speed properties. To answer this, we compared the measured BOLD response amplitudes with the predicted BOLD response amplitudes obtained by both a spatiotemporal frequency tuning model and a speed tuning model (Gaglianese et al., 2017; Lui et al., 2007; Miura et al., 2014; Priebe et al., 2006). The first model retains separate and independent responses for the spatial and temporal frequency components of the visual stimuli. The second model describes direct encoding of the speed of motion, resulting in a preference for the same speed at different spatial frequencies, with temporal frequency tuning varying in accordance with the spatial frequency. Both models are represented by a two-dimensional Gaussian function with the addition of an extra parameter Q which allows to characterise the two different types of tuning. A value of Q equal to zero (Q=0) describes separable responses for each spatial and temporal frequency combination of the stimuli. A value of Q equal to 1 (Q=1) describes tuning for particular speeds, i.e., predicts the same optimal speed at different spatial frequencies. Both models are described by the equation below:

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$R(sf,tf) = A \times e^{\left(-\frac{(\log_2(sf)-\log_2(sfopt))^2}{2\sigma sf}\right)} \times e^{\left(-\frac{(\log_2(tf)-\log_2(tfopt(sf)))^2}{2\sigma tf}\right)} \text{ Eq. 1}$$

Where$\log_2(\text{tfopt(sf)})$ is defined as:

$$\log_2(tfopt(sf)) = Q \times (\log_2(sf) - \log_2(sfopt)) + \log_2(tf) \text{ Eq. 2}$$

The term A is the peak of the BOLD response amplitude, sfopt and tfopt are the optimal spatial and temporal frequency, and σsf and σtf are the bandwidths of the spatial and temporal frequencies tuning curves. We fitted the BOLD response amplitudes of each voxel in each hMT+ ROI to both the spatiotemporal frequency and the speed tuned models (setting Q = 0 or Q = 1 respectively). For each voxel and model, three parameters were estimated: a) optimal fundamental spatial frequency b) optimal fundamental temporal frequency c) variance explained. We assigned for each voxel of each hMT+ ROI an optimal tuning model (separable spatiotemporal frequency tuning or speed tuning) based on the best fit (higher variance explained to one of the two models).

Cross validation

We compared each model's goodness of fit in predicting the BOLD response amplitude for each voxel by computing the variance explained in cross-validation (Mante et al., 2005). We evaluated each model's accuracy by fitting the BOLD response amplitudes of one half of the measured data and testing how well the resultant parameters predict the BOLD response amplitudes in the complementary half. For this purpose, two independent halves of the data are needed. Since in our visual motion stimulation paradigm we presented each spatiotemporal condition in a unique run we split each run in two halves according to the incidence of the 24s inter stimulus interval. Hence, to minimize the possible effect of the BOLD response of the last trial of the first half split of the run on the first trial of the second half split. We applied this approach for both the speed tuned model (parameter Q = 1) and the spatiotemporal frequency model (Q=0). Voxels with variance explained below 0.1 in both models were discarded from

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the subsequent analysis. A two-sided paired t-test between the average variance explained of each model within each bilateral hMT ROI+ was computed to define the model that best represented the measured BOLD responses.

Selectivity of hMT+ BOLD responses for spatial frequencies

For each half split and each ROI we obtained a spatial map of optimal spatial and temporal frequencies. We quantified the reproducibility of these maps by computing a Spearman correlation coefficient between the two half split distribution of parameters. Based on our previous observation on neuronal population recordings in humans using the same paradigm (Gaglianese et al., 2017) and neurophysiological findings on animals (Henriksson et al., 2008; Issa et al., 2000; Singh et al., 2000) , we expect the MT+ complex to be more tuned toward low spatial frequency rather than high spatial frequency. To quantify this effect, we classified the distribution of the optimal fundamental spatial frequencies obtained from the model exhibiting the highest variance explained in k clusters. To guide the choice of the number k of clusters to be used to classify the spatial frequencies we computed the within sum of squares accounting for the number of clusters, by using the Bayesian information criterion BIC. We then classified independently the spatial frequency parameters using k-means. Mean and standard deviation of the center of clusters and cluster size for each hMT+ ROI were computed across participants.

Results

hMT+ BOLD responses showed differential responses for each combination of spatial and temporal frequency of the visual motion stimuli

Regions of interest (ROIs) for hMT+ were defined for each hemisphere based on the functional localizer. The grey matter area within each left and right ROIs ranged between 57.7mm^3 and 458.2mm^3, with an average of 203.9mm^3. Deconvolved BOLD response time courses for each fundamental spatiotemporal frequency combination of the visual motion stimulation for a

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

representative hMT+ ROI are shown in Figure 2a. Figure 2b depicts the ROI extent in surface space. Each voxel was assigned a spatiotemporal frequency combination preference in accordance with the maximum BOLD response amplitude across combinations.

The effect of each fundamental spatiotemporal frequency combination on the BOLD response amplitudes was significant across participants (one-way ANOVA, $F = 4.45$, $p = 0.0005$). Four out of five participants exhibited significantly different BOLD response amplitudes ($p<0.05$) for the two pairs of spatiotemporal frequency combinations leading to the same speed of motion of the presented dartboard (3deg1Hz - 1Hz3deg and 3deg5Hz - 5deg3Hz respectively).

hMT+ BOLD response amplitudes were mainly characterized by independent tuning for spatial and temporal frequency

The two-dimensional Gaussian models allowing independent tuning for spatial and temporal frequency (Q=0) or tuning dependent on speed (Q=1), were able to characterize the hMT+ BOLD amplitude responses (Figure 3a). For each model we fitted each voxel's BOLD response amplitude for each combination of the fundamental spatial and temporal frequency on one split half of the data and computed the variance explained by the resulting model in the second complementary half. Overall, for each participant's hMT+ ROIs, in cross validation, the $Q = 0$ model explained significantly more variance than the speed encoding model $Q = 1$ (two-sided t-test, $p<0.001$ for each participant, see Figure 3a). The percentage of grey matter area within each hMT+ ROI of each participant exhibiting higher variance explained for the Q=0 model and the Q=1 respectively is shown in Figure 3b. Figure 4 shows the distributions of optimal fundamental spatial and temporal frequencies and variance explained within hMT+ ROI voxels of each participant for the spatiotemporal frequency tuning model Q=0. Median optimal fundamental spatial and temporal frequency values were consistent across participants.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

hMT+ BOLD response amplitudes were selective for spatial frequency

We further test the reproducibility of the cortical organisation of the optimal spatial and temporal frequency preferences across the two half splits of the data using a Spearman correlation. Only maps of the spatial frequency preferences exhibited a significant correlation in all the participants' hMT+ ROIs (r2 = 0.64, 0.59, 0.79, 0.65, 0.61 respectively, $p<0.0001$). Optimal fundamental spatial frequency values for each participant and each half split were then classified in two clusters respectively using kmeans. K = 2 was based on the optimum value displayed by the BIC score (fig. 5A). Optimal spatial frequency clusters for the complete run and each half split for a representative subject are shown in Figure 6A-C. A parallel-coordinates plot for spatial frequencies is shown in fig. 6D-E in which the starting point of each line on the left side of each plot indicates the spatial location of the voxel and the cluster classification, and the ending point the correspondent classification in the second half. Mean centroids across participants (Fig. 5B) were consistent and centered on 0.20±0.013 cycle/degree (low spatial frequencies cluster) and 0.79±0.035 (high spatial frequencies cluster). The percentage of voxels normalized by the size of each cluster is shown in Figure 5C.

Discussion:

In the current study we investigated the functional organization of hMT+ BOLD responses to different combinations of the fundamental spatial and temporal frequency components of visual motion stimuli. Overall, our results show independent tuning for spatial and temporal frequency within the hMT+ rather than speed tuning, in support of our findings using electrocorticography in human participants (Gaglianese et al., 2017) and previous animal studies using intracellular recordings (Lui et al., 2007; Miura et al., 2014; Priebe et al., 2006, 2003). Especially, our findings extend this tuning mechanism shown for a sample of neuronal populations to the entire hMT+ complex.

Two different mechanisms of speed tuning have been proposed in the literature. The first and more conventional view foresees direct speed tuning, where the same speed preference is

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

maintained for different spatial frequency components of the stimuli. The second proposes independent tuning for the spatial and temporal frequency components of the visual stimuli, where the neuronal population speed preference changes in accordance with the spatial frequency. Here we showed for all participants that the spatial and temporal frequency tuning model explains BOLD responses within the hMT+ significantly better than the speed tuning one. Interestingly, although smaller, we found in all the participants a subregion within the complex exhibiting higher variance explained for the speed tuning model. These findings are in line with neurophysiological recordings in macaque, showing the coexistence of a percentage of MT cells responding to the speed of motion of the presented visual stimuli and a percentage responding selectively to the spatiotemporal properties of moving stimuli (Duijnhouwer et al., 2013; Perrone and Thiele, 2001).

The selectivity for spatial and temporal frequency was further tested by estimating the optimal fundamental spatial and temporal frequency values for each voxel of each participant's hMT+ ROI. Optimal spatial frequencies maps were highly reproducible within participants, in line with the hypothesis that visual areas with specific visual field maps such as MT/V5 exhibit specific responses for the spatial frequency component of the perceived stimuli. Indeed, tuning for spatial frequencies in the occipital cortex has been shown using optical imaging in cat and fMRI in humans (Henriksson et al., 2008; Issa et al., 2000; Singh et al., 2000), by showing a decrease in optimal spatial frequency tuning moving from V1 to V3 and to extrastriate cortex such as MT+. Singh and colleagues (2000) showed, using fMRI in humans, low pass tuning responses for spatial frequency in V5/MT, exhibiting a significant drop in response for spatial frequency above 0.4 cycle/degree. In our dataset we measured the same effect: the optimal spatial frequencies were distributed along two clusters, peaking respectively on low spatial frequencies (0.20 cycle/degree) and on high spatial frequencies (0.79 cycle/degree), where the largest number of voxels in the entire hMT+ complex was tuned for the low spatial frequency cluster. Although the limited spatial frequency sampling of our experiment does not allow us to draw a firm conclusion, we suggest that this effect may reflect the change in eccentricity across the visual field maps, or different responses in the MT and MST (or TO1 and TO2) subdivisions of the complex. A single neuron recording study in the homologue V5 area in monkeys showed different spatial frequency preferences within the area (higher for MT and lower for MST) in accordance with the increase in eccentricity in MST compared to MT (Huk et al., 2002). In humans, the visual field map TO2 has larger pRFs than TO1 (Amano et al., 2009; Kolster et al., 2010). Also, within both of these visual

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

field maps, pRF sizes increase with eccentricity. Spatial frequency preferences typically decrease where pRF sizes increase, at higher eccentricities and in visual field maps with larger pRF sizes (Aghajari et al., 2020).

Finally, optimal temporal frequency tuning was not reproducible within subjects. This can be due to the range of the temporal frequencies used in our experiment (from 1Hz up to 5Hz). It has been shown that the optimal contrast sensitivity of the primate visual system is found at approximately 8 Hz (Hawken M.J., Shapley R.M., 1996; Himmelberg and Wade, 2019; Kastner et al., 2004). A recent fMRI study in humans shows a peak at around 10Hz across visual areas independent of pRF size (Himmelberg and Wade, 2019). Further studies exploring a wider range of spatial and temporal frequencies may help elucidating the spatial organization of the complex with higher detail.

Overall, our findings suggest that hMT+ responses to speed change in accordance with the spatial frequency component of the visual motion stimuli. We speculate that speed tuning properties may emerge from non linear integration of patches within the MT complex preferring the same speed but different spatial frequency. Then, at a later stage, this information is computed in other subregions within the complex as suggested by the presence of small patches showing speed tuning properties rather than separable responses. Moreover, the fact that hMT+ exhibited the same properties as the primary visual cortex V1 in encoding basic features of a visual stimuli, such as the spatial and temporal frequency components, is consistent with previous studies in both humans and primates showing that area MT receives, and is able to process, fundamental properties of the visual input directly from the thalamus, bypassing the role of the primary visual cortex (ffytche et al., 1995; Gaglianese et al., 2012; Sincich et al., 2004) and could explain the absence of deficit in biological motion perception in patients affected by congenital visual deprivation (Bottari et al., 2015). This fundamental low level mechanism of the MT complex in processing visual motion features could explain the multisensory role of this area in encoding motion via other sensory modalities such touch and hearing (Avanzini et al., 2016; Gaglianese et al., 2020; Poirier et al., 2005; Van Kemenade et al., 2014). Indeed, asensory specific areas rely on the process of task information (e.g. motion) responses based on specific low-level properties of the input regardless the sensory modalities in which they are delivered (Amedi et al., 2017; Murray and Wallace, 2002).

Conclusion

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We provided evidence of a functional selectivity for spatial frequency within hMT+, with speed preference changing in accordance with the fundamental spatial component of the presented visual motion stimuli. These measurements suggest that speed encoding in hMT+ is more complex than initially thought and underline the role of this area on computing feature properties of a visual stimuli similar as primary visual cortex.

Acknowledgment

This work was supported by the Netherlands Organization for Scientific Research (NWO), Vidi Grant number 13339 (N.P.) and by the National Institute Of Mental Health of the National Institutes of Health under Award Number R01MH111417. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Aghajari, S., Vinke, L.N., Ling, S., 2020. Population spatial frequency tuning in human early visual cortex. J. Neurophysiol. 123, 773–785. doi:10.1152/jn.00291.2019

Amano, K., Wandell, B.A., Dumoulin, S.O., 2009. Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. J. Neurophysiol. 102, 2704–18. doi:10.1152/jn.00102.2009

Amedi, A., Hofstetter, S., Maidenbaum, S., Heimler, B., 2017. Task Selectivity as a Comprehensive Principle for Brain Organization. Trends Cogn. Sci. 21, 307–310. doi:10.1016/j.tics.2017.03.007

Avanzini, P., Abdollahi, R.O., Sartori, I., Caruana, F., Pelliccia, V., Casaceli, G., Mai, R., Lo Russo, G., Rizzolatti, G., Orban, G.A., 2016. Four-dimensional maps of the human somatosensory system. Proc. Natl. Acad. Sci. 113, E1936–E1943. doi:10.1073/pnas.1601889113

Bottari, D., Troje, N.F., Ley, P., Hense, M., Kekunnaya, R., Röder, B., 2015. The neural development of the biological motion processing system does not rely on early visual

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

input. Cortex 71, 359–367. doi:10.1016/j.cortex.2015.07.029

Costagli, M., Ueno, K., Sun, P., Gardner, J.L., Wan, X., Ricciardi, E., Pietrini, P., Tanaka, K., Cheng, K., 2014. Functional signalers of changes in visual stimuli: cortical responses to increments and decrements in motion coherence. Cereb. Cortex 24, 110–8. doi:10.1093/cercor/bhs294

Dale, A.M., 1999. Optimal experimental design for event-related fMRI. Hum Brain Mapp 8, 109–114.

Duijnhouwer, J., Noest, A.J., Lankheet, M.J.M., van den Berg, A. V, van Wezel, R.J. a, 2013. Speed and direction response profiles of neurons in macaque MT and MST show modest constraint line tuning. Front. Behav. Neurosci. 7, 22. doi:10.3389/fnbeh.2013.00022

Dumoulin, S.O., Bittar, R.G., Kabani, N.J., Baker, C.L., Le Goualher, G., Bruce Pike, G., Evans, A.C., 2000. A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. Cereb. Cortex 10, 454–463. doi:10.1093/cercor/10.5.454

Dumoulin, S.O., Fracasso, A., van der Zwaag, W., Siero, J.C.W., Petridou, N., 2018. Ultra-high field MRI: Advancing systems neuroscience towards mesoscopic human brain function. Neuroimage 168, 345–357. doi:10.1016/j.neuroimage.2017.01.028

Dumoulin, S.O., Harvey, B.M., Fracasso, A., Zuiderbaan, W., Luijten, P.R., Wandell, B.A., Petridou, N., 2017. In vivo evidence of functional and anatomical stripe-based subdivisions in human V2 and V3. Sci. Rep. 7, 733. doi:10.1038/s41598-017-00634-6

ffytche, D.H., Guy, C.N., Zeki, S., 1995. The parallel visual motion inputs into areas V1 and V5 of human cerebral cortex. Brain 118, 1375–1394. doi:10.1093/brain/118.6.1375

Formisano, E., Kim, D.S., Di Salle, F., Van De Moortele, P.F., Ugurbil, K., Goebel, R., 2003.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40, 859–869. doi:10.1016/S0896-6273(03)00669-X

Fracasso, A., Petridou, N., Dumoulin, S.O., 2016. Systematic variation of population receptive field properties across cortical depth in human visual cortex. Neuroimage 139, 427–438. doi:10.1016/j.neuroimage.2016.06.048

Gaglianese, A., Branco, M.P., Groen, I.I.A., Benson, N.C., Vansteensel, M.J., Murray, M.M., Petridou, N., Ramsey, N.F., 2020. Electrocorticography Evidence of Tactile Responses in Visual Cortices. Brain Topogr. 33, 559–570. doi:10.1007/s10548-020-00783-4

Gaglianese, A., Costagli, M., Bernardi, G., Ricciardi, E., Pietrini, P., 2012. Evidence of a direct influence between the thalamus and hMT+ independent of V1 in the human brain as measured by fMRI. Neuroimage 60, 1440–7. doi:10.1016/j.neuroimage.2012.01.093

Gaglianese, A., Harvey, B.M., Vansteensel, M.J., Dumoulin, S.O., Ramsey, N.F., Petridou, N., 2017. Separate spatial and temporal frequency tuning to visual motion in human MT+ measured with ECoG. Hum. Brain Mapp. 38, 293–307. doi:10.1002/hbm.23361

Gardner, J.L., Sun, P., Waggoner, R.A., Ueno, K., Tanaka, K., Cheng, K., 2005. Contrast adaptation and representation in human early visual cortex. Neuron 47, 607–620. doi:10.1016/j.neuron.2005.07.016

Harvey, B.M., Dumoulin, S.O., 2016. Visual motion transforms visual space representations similarly throughout the human visual hierarchy. Neuroimage 127, 173–185. doi:10.1016/j.neuroimage.2015.11.070

Harvey, B.M., Dumoulin, S.O., Fracasso, A., Paul, J.M., 2020. A Network of Topographic Maps in Human Association Cortex Hierarchically Transforms Visual Timing-Selective Responses. Curr. Biol. 30, 1424-1434.e6. doi:10.1016/j.cub.2020.01.090

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Harvey, B.M., Klein, B.P., Petridou, N., Dumoulin, S.O., 2013. Topographic Representation of Numerosity in the Human Parietal Cortex. Science (80-.). 341, 1123–1126. doi:10.1126/science.1239052

Hawken M.J., Shapley R.M., G.D.H., 1996. Temporal-frequency selectivity in monkey visual cortex. Vision Res. 13, 477–492. doi:10.1017/S0952523800008154

Henriksson, L., Nurminen, L., Hyvarinen, A., Vanni, S., 2008. Spatial frequency tuning in human retinotopic visual areas. J. Vis. 8, 5–5. doi:10.1167/8.10.5

Himmelberg, M.M., Wade, A.R., 2019. Eccentricity-dependent temporal contrast tuning in human visual cortex measured with fMRI. Neuroimage 184, 462–474. doi:10.1016/j.neuroimage.2018.09.049

Huber, L., Finn, E.S., Handwerker, D.A., Bönstrup, M., Glen, D.R., Kashyap, S., Ivanov, D., Petridou, N., Marrett, S., Goense, J., Poser, B.A., Bandettini, P.A., 2020. Sub-millimeter fMRI reveals multiple topographical digit representations that form action maps in human motor cortex. Neuroimage 208. doi:10.1016/j.neuroimage.2019.116463

Huk, A.C., Dougherty, R.F., Heeger, D.J., 2002. Retinotopy and functional subdivision of human areas MT and MST. J. Neurosci. 22, 7195–7205. doi:20026661

Issa, N.P., Trepel, C., Stryker, M.P., 2000. Spatial Frequency Maps in Cat Visual Cortex. J. Neurosci. 20, 8504–8514. doi:https://doi.org/10.1523/JNEUROSCI.20-22-08504.2000

Kastner, S., Connor, D.H.O., Fukui, M.M., Fehd, H.M., Herwig, U., Pinsk, M. a, M, H., 2004. Functional Imaging of the Human Lateral Geniculate Nucleus and Pulvinar. J. Neurophysiol. 91, 438–448. doi:10.1152/jn.00553.2003

Klein, B.P., Harvey, B.M., Dumoulin, S.O., 2014. Attraction of position preference by spatial attention throughout human visual cortex. Neuron 84, 227–37.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

doi:10.1016/j.neuron.2014.08.047

Kolster, H., Peeters, R., Orban, G. a, 2010. The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. J. Neurosci. 30, 9801–20. doi:10.1523/JNEUROSCI.2069-10.2010

Lui, L.L., Bourne, J.A., Rosa, M.G.P., 2007. Spatial and temporal frequency selectivity of neurons in the middle temporal visual area of new world monkeys (Callithrix jacchus). Eur. J. Neurosci. 25, 1780–1792. doi:10.1111/j.1460-9568.2007.05453.x

Mante, V., Frazor, R. a, Bonin, V., Geisler, W.S., Carandini, M., 2005. Independence of luminance and contrast in natural scenes and in the early visual system. Nat. Neurosci. 8, 1690–7. doi:10.1038/nn1556

Miura, K., Inaba, N., Aoki, Y., Kawano, K., 2014. Difference in visual motion representation between cortical areas MT and MST during ocular following responses. J. Neurosci. 34, 2160–8. doi:10.1523/JNEUROSCI.3797-13.2014

Murray, M.M., Wallace, M.T., 2002. The Neural Bases of Multisensory Processes.

Perrone, J.A., Thiele, A., 2001. Speed skills: measuring the visual speed analyzing properties of primate MT neurons. Nat. Neurosci. 4, 526–32. doi:10.1038/87480

Petridou, N., Italiaander, M., van de Bank, B.L., Siero, J.C.W., Luijten, P.R., Klomp, D.W.J., 2013. Pushing the limits of high-resolution functional MRI using a simple high-density multi-element coil design. NMR Biomed. 26, 65–73. doi:10.1002/nbm.2820

Petridou, N., Siero, J.C.W., 2019. Laminar fMRI: What can the time domain tell us? Neuroimage 197, 761–771. doi:10.1016/j.neuroimage.2017.07.040

Poirier, C., Collignon, O., DeVolder, A.G., Renier, L., Vanlierde, A., Tranduy, D., Scheiber, C., 2005. Specific activation of the V5 brain area by auditory motion processing: An

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

fMRI study. Cogn. Brain Res. 25, 650–658. doi:10.1016/j.cogbrainres.2005.08.015

Priebe, N.J., Cassanello, C.R., Lisberger, S.G., 2003. The neural representation of speed in macaque area MT/V5. J. Neurosci. 23, 5650–5661. doi:23/13/5650 [pii]

Priebe, N.J., Lisberger, S.G., Movshon, J.A., 2006. Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. J. Neurosci. 26, 2941–50. doi:10.1523/JNEUROSCI.3936-05.2006

Schellekens, W., Petridou, N., Ramsey, N.F., 2018. Detailed somatotopy in primary motor and somatosensory cortex revealed by Gaussian population receptive fields. Neuroimage 179, 337–347. doi:10.1016/j.neuroimage.2018.06.062

Schneider, M., Kemper, V.G., Emmerling, T.C., Martino, F. De, Goebel, R., 2019. Columnar clusters in the human motion complex reflect consciously perceived motion axis. Proc. Natl. Acad. Sci. 116, 5096–5101. doi:10.1073/pnas.1814504116

Sincich, L.C., Park, K.F., Wohlgemuth, M.J., Horton, J.C., 2004. Bypassing V1: a direct geniculate input to area MT. Nat. Neurosci. 7, 1123–8. doi:10.1038/nn1318

Singh, K.D., Smith, A.T., Greenlee, M.W., 2000. Spatiotemporal Frequency and Direction Sensitivities of Human Visual Areas Measured Using fMRI. Neuroimage 564, 550–564. doi:10.1006/nimg.2000.0642

Sun, P., Ueno, K., Waggoner, R.A., Gardner, J.L., 2007. A temporal frequency – dependent functional architecture in human V1 revealed by high-resolution fMRI 10, 1404–1406. doi:10.1038/nn1983

Tootell, R.B.H., Nasr, X.S., 2017. Columnar Segregation of Magnocellular and Parvocellular Streams in Human Extrastriate Cortex 37, 8014–8032. doi:10.1523/JNEUROSCI.0690-17.2017

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tootell, R.B.H., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., Rosen, B.R., Belliveau, J.W., 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J. Neurosci. 15, 3215–3230.

Van Kemenade, B.M., Seymour, K., Wacker, E., Spitzer, B., Blankenburg, F., Sterzer, P., 2014. Tactile and visual motion direction processing in hMT+/V5. Neuroimage 84, 420–427. doi:10.1016/j.neuroimage.2013.09.004

Yacoub, E., Harel, N., Ugurbil, K., 2008. High-field fMRI unveils orientation columns in humans. Proc. Natl. Acad. Sci. U. S. A. 105, 10607–12. doi:10.1073/pnas.0804110105

Zimmermann, J., Goebel, R., De Martino, F., van de Moortele, P.-F., Feinberg, D., Adriany, G., Chaimow, D., Shmuel, A., Uğurbil, K., Yacoub, E., 2011. Mapping the organization of axis of motion selective features in human area MT using high-field fMRI. PLoS One 6, e28716. doi:10.1371/journal.pone.0028716

Zuiderbaan, W., Harvey, B.M., Dumoulin, S.O., 2012. Modeling center – surround confi gurations in population receptive fields using fMRI. J. Vis. 12, 1–15. doi:10.1167/12.3.10.Introduction

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

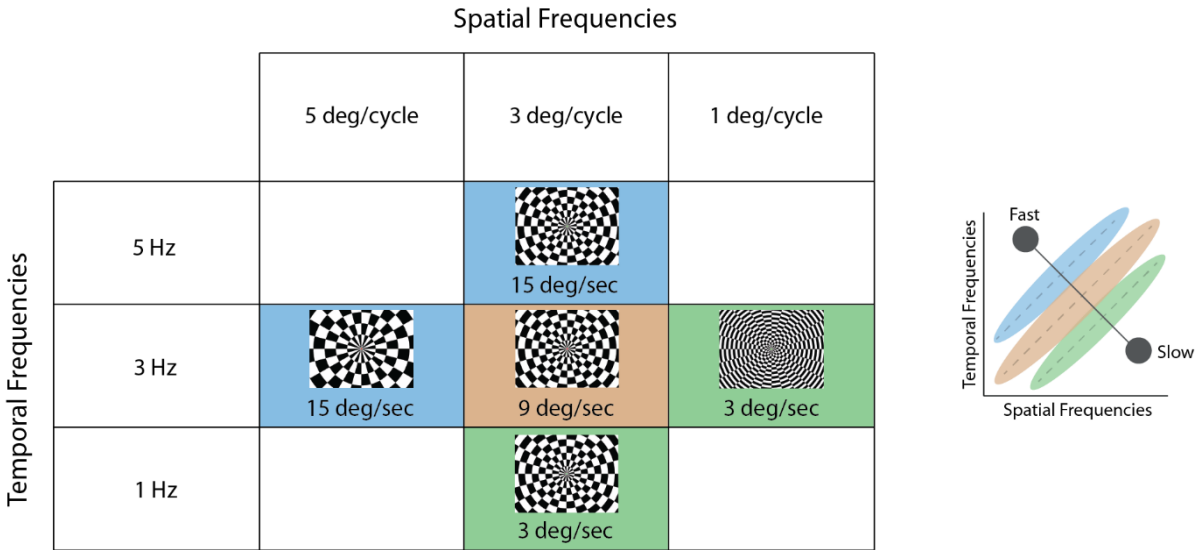

Figure 1: Graphical summary of the visual motion stimuli (high-contrast black-and-white dartboard) presented during the five runs of the study following the localizer run. Each run is either classified as a fast (represented by blue background color), intermediate (represented by orange background colour) or slow (represented by green background colour) moving stimuli, depending on the spatial and temporal frequency that gives origin to the stimulus speed. Graphic on the right provides a simplified depiction of the stimulus space.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

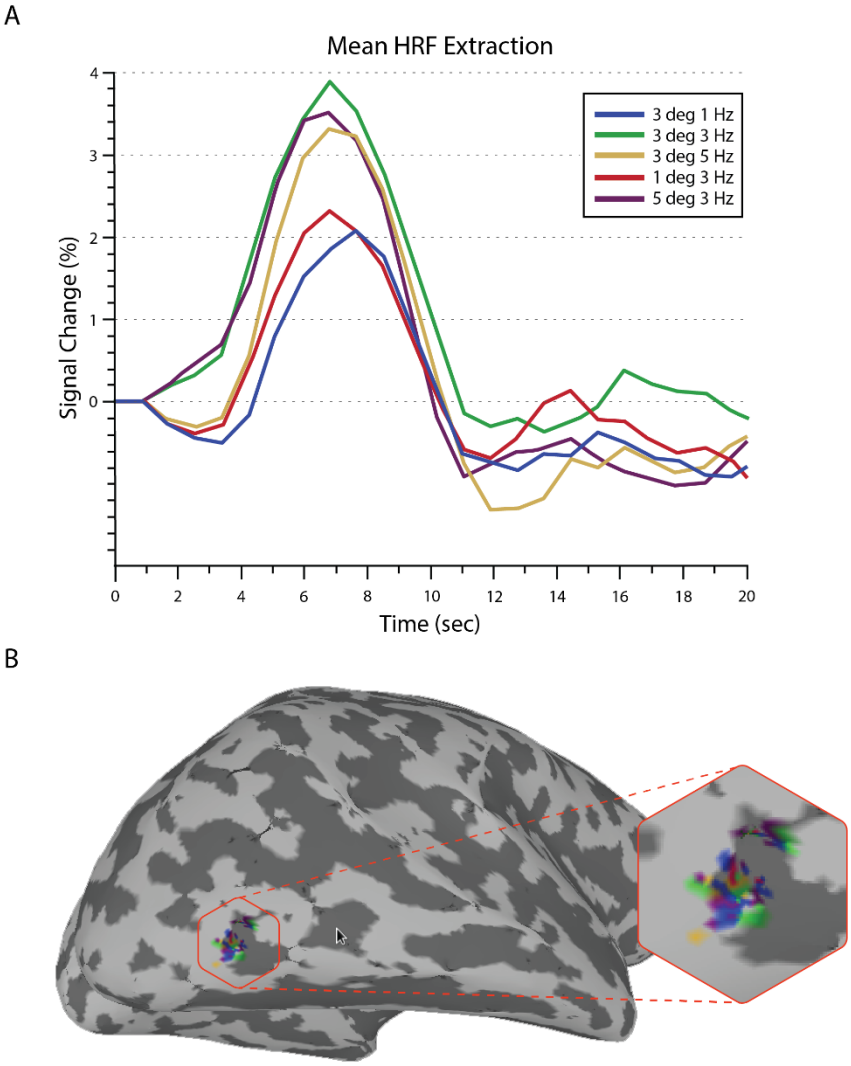

Figure 2: A) Mean BOLD responses for each presented fundamental spatiotemporal frequency of the visual stimulation for a representative hMT+ ROI. B) Spatiotemporal frequency preference within a representative hMT+ ROI plotted on the brain surface. The amplitude with the highest value across all spatiotemporal frequency conditions determined the preferred spatiotemporal frequency combination of each voxel.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

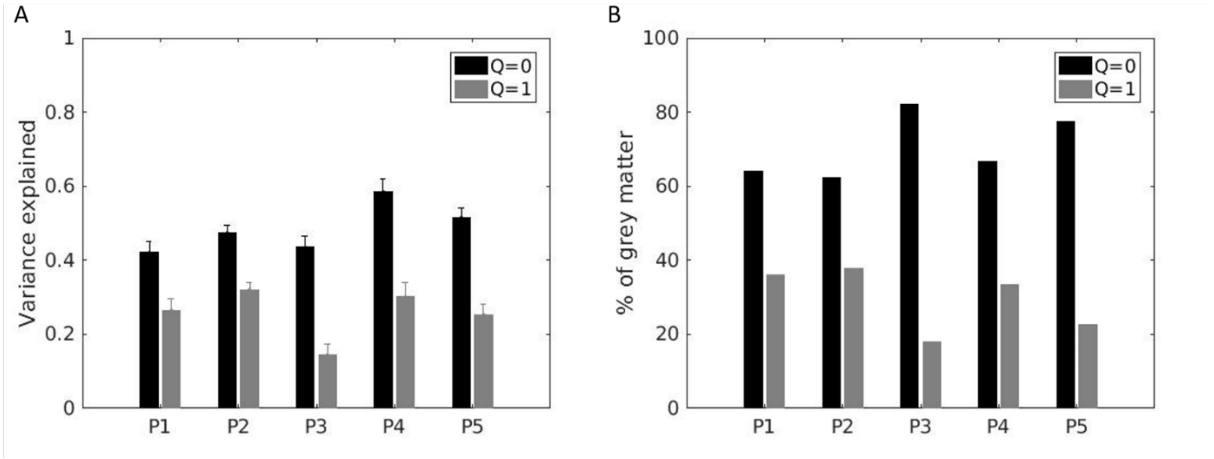

Figure 3: A) Single participant variance explained in split-run cross validation by Gaussian tuning models with independent spatial and temporal frequency tuning (Q=0) and tuning for speed (Q=1). Each bar represents the mean variance explained and standard error for bilateral hMT+ ROI voxels of each participant (P1 to P5, x-axis). B) Single participant percentage of voxels within hMT+ ROIs exhibiting separable spatiotemporal frequency tuning (Q=0, black bars) or speed tuning, i.e. same temporal frequency preference for the different spatial frequency of the moving dartboard (Q=1, grey bars)

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

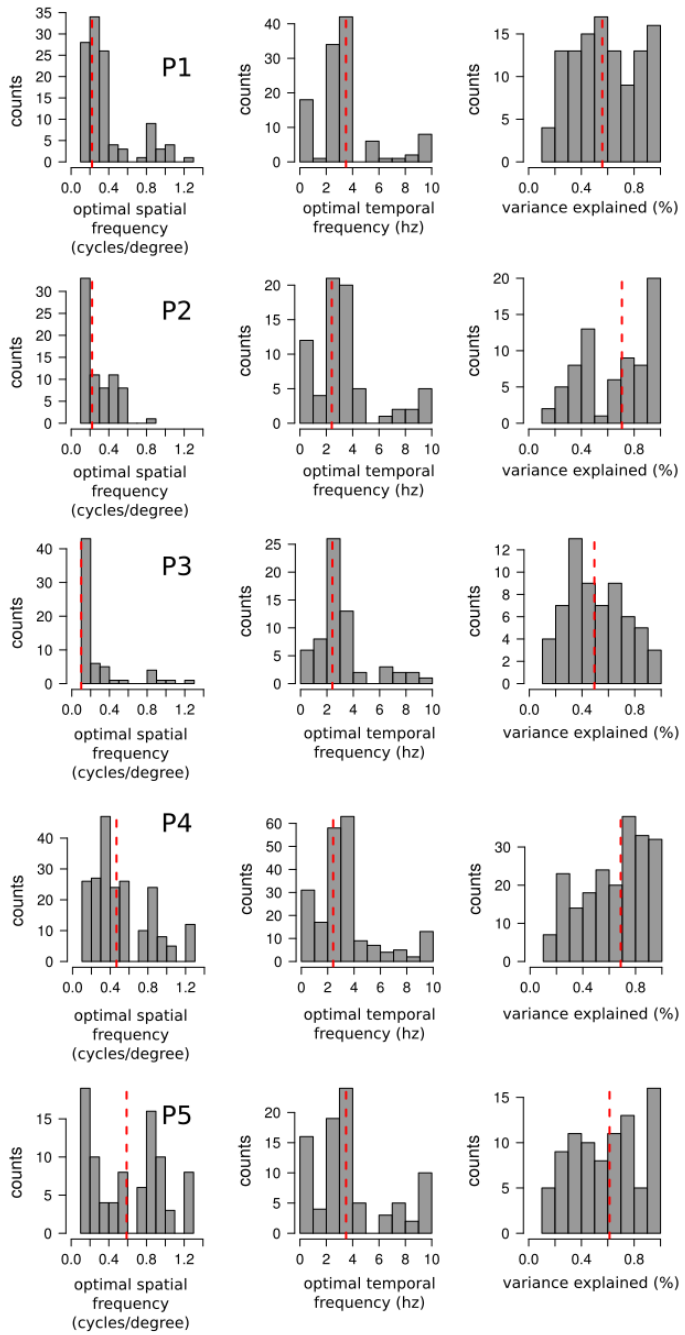

Figure 4: Histograms of estimated optimal temporal and spatial frequencies and variance explained obtained in cross validation for each participant's bilateral hMT+ ROI. Red dashed line shows the median values across voxels.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

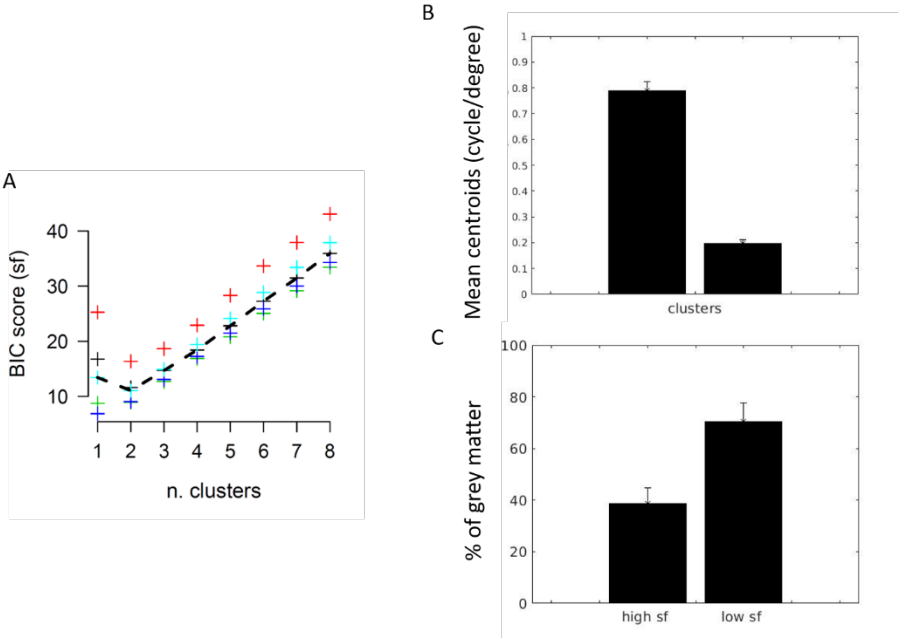

Figure 5: A) Bayesian Information Criterion score as a function of the number k of clusters for the optimal spatial frequency values. Each colour represents a single participant. Dashed line represents the mean score across participants. K = 2 was selected for the optimal number of clusters B) Mean centroids value and standard deviation across participants hMT+ ROIs for each spatial frequency cluster C) Mean percentage of grey matter and standard deviation across participants hMT+ ROIs for each spatial frequency cluster.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.01.429141; this version posted February 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

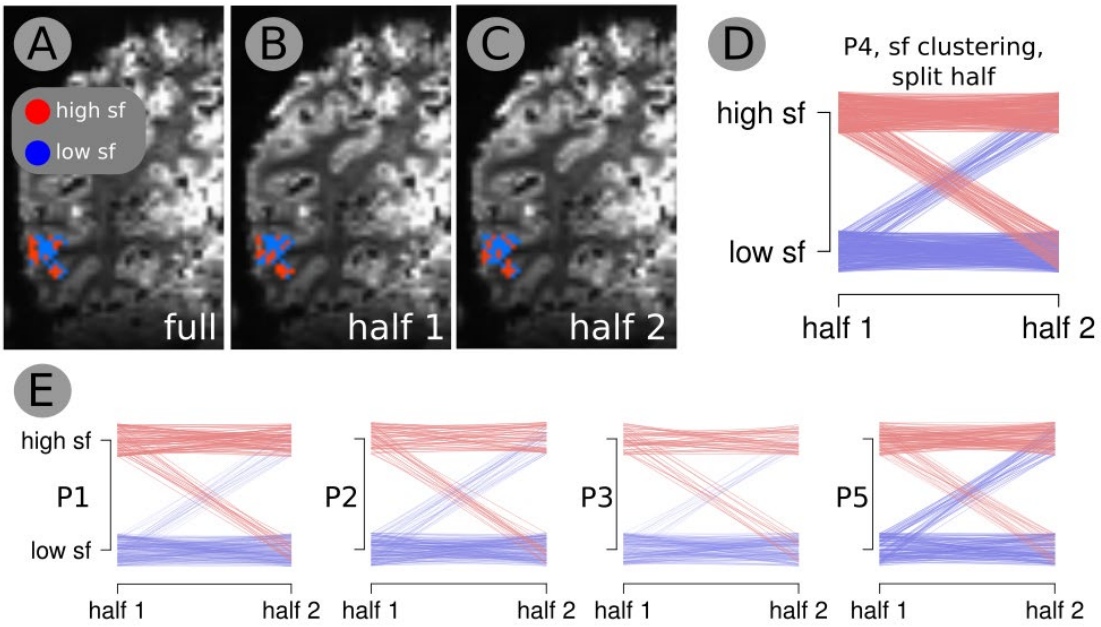

Figure 6: Optimal spatial frequency clusters and their reproducibility for a representative participant. A-B-C Spatial frequency clusters represented on the cortex for the entire run, half 1 and half 2 respectively. D) Parallel-coordinates plot. It represents optimal spatial frequency of each voxel in the first (left side) and second half (right side) run respectively. Y axis maps voxel coordinates in space. E) Same as D for each measured participant