



Research report

Decoding pattern motion information in V1

Bianca M. van Kemenade^{a,b,c,*}, Kiley Seymour^{b,c,d,1},
Thomas B. Christophel^{c,e}, Marcus Rothkirch^{b,c} and Philipp Sterzer^{a,b,c,e}

^aBerlin School of Mind and Brain, Humboldt Universität zu Berlin, Berlin, Germany

^bDepartment of Psychiatry and Psychotherapy, Campus Charite Mitte, Charité-Universitätsmedizin Berlin, Berlin, Germany

^cBerlin Center for Advanced Neuroimaging, Charité – Universitätsmedizin, Berlin, Germany

^dDepartment of Cognitive Science, Macquarie University, Sydney, Australia

^eBernstein Center for Computational Neuroscience, Berlin, Germany

ARTICLE INFO

Article history:

Received 25 September 2013

Reviewed 2 December 2013

Revised 19 February 2014

Accepted 25 April 2014

Action editor Pia Rotshtein

Published online 9 May 2014

Keywords:

Pattern motion

V1

fMRI

ABSTRACT

When two gratings drifting in different directions are superimposed, the resulting stimulus can be perceived as two overlapping component gratings moving in different directions or as a single pattern moving in one direction. Whilst the motion direction of component gratings is already represented in visual area V1, the majority of previous studies have found processing of pattern motion direction only from visual area V2 onwards. Here, we question these findings using multi-voxel pattern analysis (MVPA). In Experiment 1, we presented superimposed sinusoidal gratings with varying angles between the two component motions. These stimuli were perceived as patterns moving in one of two possible directions. We found that linear support vector machines (SVMs) could generalise across stimuli composed of different component motions to successfully discriminate pattern motion direction from brain activity in V1, V3A and hMT+/V5. This demonstrates the representation of pattern motion information present in these visual areas. This conclusion was verified in Experiment 2, where we manipulated similar plaid stimuli to induce the perception of either a single moving pattern or two separate component gratings. While a classifier could again generalise across stimuli composed of different component motions when they were perceived as a single moving pattern, its performance dropped substantially in the case where components were perceived. Our results indicate that pattern motion direction information is present in V1.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Neurons in primary visual cortex (V1), having small receptive fields, suffer an *aperture problem* (Wallach, 1935) for signalling

the speed and direction of a moving line or grating. Specifically, if a grating is viewed through a small aperture, there are a number of motion directions that could correspond to the true motion direction of the grating. The visual system deals with this by combining the responses across many V1

* Corresponding author. Klinik für Psychiatrie und Psychotherapie, Campus Charité Mitte, 10117 Berlin, Germany.
E-mail address: biancavankemenade@gmail.com (B.M. van Kemenade).

¹ Equal contribution.

<http://dx.doi.org/10.1016/j.cortex.2014.04.014>

0010-9452/© 2014 Elsevier Ltd. All rights reserved.

neurons, which signal ambiguous local motion cues, so that an unambiguous global motion signal can be extracted. The mechanism by which these responses are combined has been extensively studied using so called plaid stimuli, created by the superimposition of two gratings that move in different directions. These stimuli can elicit two different percepts; either the two separate gratings are perceived as moving in their respective directions (component motion), or the two gratings are perceived bound together as a plaid, usually moving in the average direction of the two separate gratings (pattern motion). Plaid perception can be influenced by several factors, such as spatial frequency of the components (Smith, 1992), pattern direction (Hupe & Rubin, 2004), component contrast and speed (Smith, 1992), and the luminance of the intersections of the two gratings (Stoner & Albright, 1996).

It has been proposed that V1 only processes component motion, whereas pattern motion is processed at higher levels of the visual hierarchy, particularly in area hMT+/V5, the human motion complex. This theory is based on various electrophysiological studies that have found evidence only for component selective cells in V1, but selectivity for both component and pattern motion in higher visual areas (Gizzi, Katz, Schumer, & Movshon, 1990; Movshon, Adelson, Gizzi, & Newsome, 1985; Movshon & Newsome, 1996; Rodman & Albright, 1989; Rust, Mante, Simoncelli, & Movshon, 2006). Furthermore, several human functional magnetic resonance imaging (fMRI) studies using univariate statistical methods also report evidence for pattern motion processing from V2 onwards, but not in V1 (Castelo-Branco et al., 2002; Huk & Heeger, 2002; Villeneuve, Kupers, Gjedde, Ptito, & Casanova, 2005; Villeneuve, Thompson, Hess, & Casanova, 2012). However, some animal studies have suggested that the aperture problem might be solved already in V1 (Pack, Livingstone, Duffy, & Born, 2003), and there is some further evidence for pattern motion selectivity in V1 (Guo, Benson, & Blakemore, 2004; Schmidt et al., 2006; Tinsley et al., 2003), leaving open the question whether and to what extent V1 might be involved in the processing of pattern motion (van Wezel & van der Smagt, 2003).

In this study, we investigated whether pattern motion is processed in human V1, using fMRI and multi-voxel pattern analysis (MVPA). MVPA has been argued to offer superior sensitivity over standard univariate methods (Haynes & Rees, 2006; Norman, Polyn, Detre, & Haxby, 2006), and has previously been used to discriminate fMRI activity patterns in response to differences in motion direction (Kamitani & Tong, 2006; Seymour, Clifford, Logothetis, & Bartels, 2009). In the MRI scanner, we presented two sets of pattern stimuli moving in one of two opposite pattern directions. The stimuli were composed of two gratings, whose orientations varied across the sets, such that our two pattern motion directions comprised various component motion directions. The key analysis used in our experiments was the training of a classifier on stimuli composed of a certain pair of components, and testing for generalisation across the set, with stimuli composed of different components. Our prediction was that, if decoding in V1 were based on component motion, we should see an increasing drop in decoding accuracy with increasing angular differences in component motion directions between training and test set. If however decoding in V1 were based on the pattern direction, we would not expect to see any

significant differences between the decoding accuracies of any of the cross-classifications, because the pattern directions always remain constant. While our main research question focused on V1, we performed the same analyses also in motion sensitive areas V3A and hMT+/V5, which have previously been reported to show high pattern motion selectivity (Huk & Heeger, 2002), and area V2, based on reports of motion direction tuning in this area (Lu, Chen, Tanigawa, & Roe, 2010).

2. General methods

2.1. fMRI data acquisition and analysis

In two separate experiments, functional MRI data were acquired using a 3 T TIM Trio scanner (Siemens, Erlangen, Germany), equipped with a 12-channel head-coil. A gradient echo EPI sequence was used (TR: 2 sec, TE: 30 msec, flip angle: 78°, voxel size 2.3 × 2.3 × 2.5 mm). Per run, 135 volumes for both Experiment 1 and 2 were obtained, 102 volumes for retinotopic eccentricity mapping, 123 volumes for retinotopic polar angle mapping, and 169 volumes for the functional hMT+/V5 localiser, each containing 29 slices oriented parallel to the calcarine sulcus, acquired in descending order. Initial volumes were not removed, as the scanner sent the first trigger pulse (that started the experiment) only after T1 equilibration had occurred. Anatomical images were obtained using an MPRAGE sequence (TR: 1.9 sec, TE: 2.52 msec, flip angle: 9°).

The functional data of the pattern motion experiments were first realigned to the mean functional image volume and coregistered with the participant's structural image obtained in the retinotopic mapping session using Statistical Parametric Mapping (SPM8, Wellcome Department of Imaging Neuroscience, University College London, UK). No subsequent normalisation to a standard anatomical space or spatial smoothing was performed. During model estimation, the data were high-pass filtered (cut-off = 128 sec). We used an autoregressive model of order 1 [AR(1)] to account for autocorrelation in the data. In Experiment 1, all stimuli with the same pattern direction were modelled by one regressor in a general linear model (GLM), thus pooling the data into two data sets (i.e., pattern directions A and B). For the cross-classifications in Experiments 1 and 2, a GLM was created in which each stimulus type (i.e., different component motion exemplars for A and B) was modelled by a separate regressor. The canonical haemodynamic response function implemented in SPM8 was convolved with a box-car function to model the regressors. Motion parameters were included as regressors of no interest. Our experiments were designed to probe MVPA decoding of pattern motion directions. However, to investigate whether any of our MVPA decoding results could be explained by overall differences in response amplitudes between pattern motion directions, we also performed univariate region-of-interest (ROI) analyses in both experiments (see [Supplementary Material](#)). The only effect relevant to our decoding results was a significant main effect of pattern motion direction in hMT+/V5 in Experiment 2, indicating a potential coarse scale bias for pattern motion direction in this area. No such effect was observed in any of the other ROIs, especially not in V1, which was the area of primary interest in

our study. In Experiment 1, no coarse scale bias for pattern motion direction was found in any of the ROIs.

2.2. ROI definition

To localise hMT+/V5, a standard functional localiser was performed (Huk, Dougherty, & Heeger, 2002; Tootell et al., 1995), in which radially moving dots were presented in block-wise alternation with static dots. Two runs were conducted. The hMT+/V5 localiser data were realigned and coregistered with the participant's structural image obtained in the retinotopic mapping session. As with all the data, we did not normalise to a standard template brain. Thus, ROIs were created in individual subject space. Area hMT+/V5 was defined by the contrast of moving versus static dots while taking anatomical landmarks into account (Dumoulin et al., 2000). For this region's definition only, we also smoothed the hMT+/V5 localiser data (5 mm full-width at half-maximum) to maximise signal from noise in single subject data from this small region. Primary visual cortex (V1), area V2, and motion area V3A were obtained with retinotopic mapping scans using a polar angle protocol, that used a bifield rotating wedges stimulus (Slotnick & Yantis, 2003), and an eccentricity mapping protocol (Engel, Glover, & Wandell, 1997; Sereno et al., 1995). These retinotopy data were only realigned and coregistered to the participant's structural image; no normalisation or smoothing was performed. The data were analysed with *FreerSurfer* (<http://surfer.nmr.mgh.harvard.edu/>) using standard protocols (Engel et al., 1997; Sereno et al., 1995). Voxels were selected using functional data from an independent stimulus localiser scan (see 3.1.3 Procedure). ROI-masks were created by selecting all voxels within each functionally mapped visual area that were responsive to our stimuli in the localiser scan, surviving a liberal threshold of $p < .01$ (uncorrected) using an F-contrast. The average number of voxels in V1, V2, V3A and hMT+/V5 was 459 (standard error of the mean – SEM: 11.8), 527 (SEM: 15.6), 97 (SEM: 11), and 136 (SEM: 4.7) voxels, respectively, in Experiment 1, and 342 (SEM: 13.6), 372 (SEM: 14.8), 77 (SEM: 5), and 141 (SEM: 6.8) voxels, respectively, in Experiment 2.

2.3. Eye movement measurements

In both Experiments 1 and 2, eye movements were recorded with an iView Xtm MRI-LR system [SensoryMotoric Instruments (SMI), Teltow, Germany] using a sampling rate of 50 Hz. A radius of 1.5° from fixation was defined as the fixation area, of which movements beyond this were considered as outliers. Data were detrended and mean-corrected to determine the number of these outliers, and participants with eye movements beyond 1.5° of fixation in more than 5% of all data points were excluded.

A second control analysis was performed to address potential optokinetic biases on classifier performance resulting from our two pattern directions. The direction of movement between each two consecutive data points was calculated from the detrended and mean-corrected eye-movement data and sorted into 30 bins. We subsequently performed statistical comparisons between the distributions of the eye movements performed under the two pattern motion directions using a Kolmogorov–Smirnov test.

3. Experiment 1

3.1. Material and methods

3.1.1. Participants

Fifteen participants with normal or corrected-to-normal vision took part in the experiment. One subject was excluded based on eye movement exclusion criteria (see 3.2.2 Eye Tracking). Thus, the data of 14 subjects are presented (five male, nine female, age 18–30, mean age 25, all right handed). All participants gave written informed consent to participate in the experiment, which was approved by the local ethics committee.

3.1.2. Stimuli

Two sets of plaid stimuli were created, one set moving in pattern direction A (225°), and the other moving in pattern direction B (45°) (Fig. 1). The plaid stimuli were composed of two sinusoidal gratings, but the angle between the gratings' motion vectors (α), reflecting their respective motion direction, could be either 60°, 90°, 120°, or 150°. Thus, for the two opposing pattern motion directions, a total set of eight different types of plaid stimuli were generated, with four possible angles defining the components' motion directions (α) (Fig. 1). The individual component gratings had a spatial frequency of .5 cycles per degree (cpd) and a speed of 1 cycle/sec. Thus, the resulting plaids differed in speed, but this was matched across both set A and B. The plaid stimuli were presented within a centred annulus, with a diameter of 13° of visual angle. The surrounding background had a luminance of 1888 cd/m². The reported luminance values in both experiments are comparably high due to the brightness of the projector at the scanner and the additional gamma correction we performed. The centre of the annulus, in which a fixation cross was presented, had a diameter of 3° to facilitate fixation without optokinetic nystagmus. This was also important for allowing a clear delineation of retinotopic responses by avoiding stimulation of the foveal confluence region.

3.1.3. Procedure

During all fMRI runs, participants were required to fixate and detect a random colour change of a fixation cross in the centre of the screen. Subjects were also required to respond to a speed change of the plaid stimuli, which occurred at random time intervals throughout each block. These tasks were included to improve fixation and enhance subjects' attention to the plaid stimuli, respectively. The scanning session started with a structural scan (MPRAGE), during which participants were presented with two short training runs to become familiar with the stimuli and behavioural tasks. During each run, each plaid stimulus was presented three times, for a duration of 10 sec with an inter-trial interval of 1 sec. The stimuli were presented in a pseudo-randomised order, such that stimuli with the same pattern motion directions were never presented more than twice in a row, reducing possible adaptation effects. After 10 runs, one additional stimulus localiser run was presented in which each plaid stimulus, presented twice in total, was followed by 10 sec of fixation. This allowed us to restrict our classifier analysis to those voxels within each ROI that were generally responsive to our

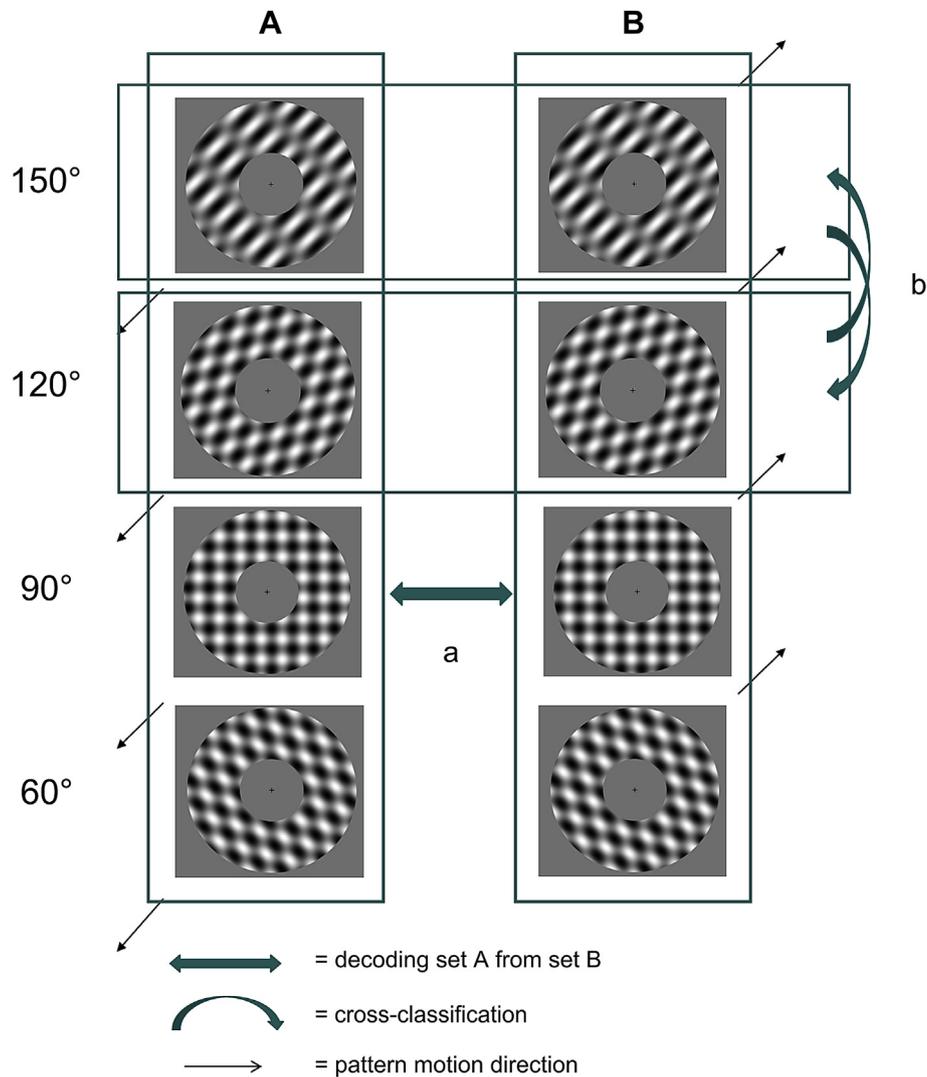


Fig. 1 – Stimuli and analysis of Experiment 1. Plaid stimuli with two pattern directions (A and B), composed of gratings with α ranging from 60° to 150°, were presented in pseudo-randomised order. A classifier was trained on pooled data sets to discriminate between pattern direction A and B (a). Subsequent cross-classifications were performed by training on stimuli with the same α and testing on stimuli with a different α , with all possible combinations of stimulus pairs; (b) shows one of the possible combinations.

stimuli. Standard ROI localisers were run in separate scan sessions. Seven participants were not available for the eccentricity mapping session, so V3A could only be defined in the remaining seven participants. V1 and hMT+/V5 were defined in all 14 participants.

3.1.4. Multi-voxel pattern analysis (MVPA)

MVPA was performed using the Decoding toolbox (Görgen, Hebart, & Haynes, 2012), which implements LibSVM software (<http://www.csie.ntu.edu.tw/~cjlin/libsvm>). In a first step, the data were split into two sets by modelling all stimuli with the same pattern direction with one regressor for each run in a GLM, thus pooling stimuli across different component directions (see 2.1 fMRI Data Acquisition and Analysis). A classifier was then trained on the beta values corresponding to the two pooled data sets to distinguish between the two pattern motion directions (Fig. 1). It is important to note here that by

pooling the responses to stimuli composed of different components, the classifier is forced to ignore component motion information to discriminate the two data sets. Specifically, because the stimuli relating to each data set contained plaids with components moving in various directions, using component motion information to discriminate the two data sets would be uninformative to a classifier. Hence, above-chance decoding of the two data sets would not result from component motion decoding. Classifier performance was tested using a leave-one-run-out cross-validation approach. Training was carried out on all but one run, which served as the test data. This was repeated 10 times until all runs had served as a test run once. The decoding accuracy was averaged across cross-validations and permutation testing was conducted to determine the significance at the group level as described by Stelzer, Chen, and Turner (2013). This method was also used and described in our previous study (van

Kemenade et al., 2014). We provided the classifier with all possible combinations of label assignments for each subject and decoded. Then, we randomly selected one of these decoding accuracies from each subject and calculated the average decoding accuracy. This procedure of random selection and calculation of average decoding accuracy was repeated 10,000 times. A cut-off of 95% was used, based on the resulting distribution of average decoding accuracies, to determine the significance of our results (van Kemenade et al., 2014; Stelzer et al., 2013).

To further probe whether classifier performance was driven by pattern or component motion information, we assessed whether the extent to which a classifier could generalise across different pattern motion stimuli depended on the similarity of their components. We created a new GLM, in which each stimulus condition was modelled by a separate regressor for each run (see 2.1 fMRI Data Acquisition and Analysis). As in the previous analysis, a classifier was trained on the beta values to discriminate the two pattern directions. However, this time we trained the classifier on stimuli with a certain α , for example 60° , and tested on stimuli with another α , for example 90° . If the classifier used pattern motion information to discriminate the two training data sets, it should still be able to decode the pattern direction regardless of the angle between the components in the test data sets. However, if the classifier used component motion information, decoding accuracy should decrease as the difference in α between training and tests becomes larger. We based this hypothesis on the current knowledge of population tuning curves of direction-selective cells; the more the motion direction of a stimulus differs from the preferred direction, the less the cell is responsive to this stimulus (Movshon et al., 1985). Studies measuring tuning curves of direction-selective cells in V1 report a range of tuning widths, but usually no greater than about 90° (Gizzi et al., 1990; Movshon et al., 1985; Movshon & Newsome, 1996). Therefore, cross-classifications between plaid stimuli with $\alpha = 60^\circ$ versus $\alpha = 150^\circ$ should result in the weakest decoding performance if the classifier uses component motion information to make its decision. Furthermore, as the difference in α gets larger between training and test, a graded decline in performance should be observed. The cross-classification procedure was performed on all possible combinations of stimulus pairs. The resulting decoding accuracies from these cross-classifications were then pooled according to

the difference in angle between training and test sets. For instance, performance from training on stimuli with $\alpha = 60^\circ$ and testing on stimuli with $\alpha = 90^\circ$ would be averaged with performance obtained from training on stimuli with $\alpha = 90^\circ$ and testing on stimuli with $\alpha = 120^\circ$, etc. (i.e., a difference of $\Delta = 30^\circ$ between training and test set). A repeated-measures ANOVA was then performed per ROI to investigate differences in decoding accuracy across these angular differences.

3.2. Results

3.2.1. Multivariate pattern analysis

In the first analysis, a classifier was trained and tested using a leave-one-run-out procedure on the pooled data sets, where stimuli in each set contained the same pattern direction, but differed with respect to the angles that separated the component motion directions. The classifier was able to decode the two pattern directions significantly above chance in all ROIs (V1: 78.2%, V2: 77.1%, V3A: 63.6%, hMT+/V5: 66.8%; all $p < .001$; see Fig. 2A).

In subsequent cross-classification analyses, we trained on stimulus pairs with the same α , and tested on stimulus pairs with a different α . We found significant above-chance decoding performance for all Δ s in all ROIs [V1: 71.7% ($\Delta 30^\circ$), 69.4% ($\Delta 60^\circ$), 69.5% ($\Delta 90^\circ$); V2: 73.6% ($\Delta 30^\circ$), 70.2% ($\Delta 60^\circ$), 72.1% ($\Delta 90^\circ$); V3A: 58.6% ($\Delta 30^\circ$), 57.5% ($\Delta 60^\circ$), 62.1% ($\Delta 90^\circ$); hMT+/V5: 57.9% ($\Delta 30^\circ$), 56.4% ($\Delta 60^\circ$), 59.1% ($\Delta 90^\circ$); all $p < .001$; see Fig. 2B]. Furthermore, a repeated-measures ANOVA showed no significant differences in decoding performance across the different cross-classification scenarios in any of the ROIs [V1: $F(2,26) = 1.06$, $p = .36$; V2: $F(2,26) = 3.1$, $p = .06$; V3A: $F(2,12) = 1.59$, $p = .24$; hMT+/V5: $F(2,26) = 2.15$, $p = .14$].

3.2.2. Eye tracking

Due to technical difficulties, no usable eye tracking data were obtained for one participant. For two other participants, eye tracking data were available for seven and nine runs, respectively. One subject was excluded due to frequent eye closure during scanning that could not be attributed to blinking. The fixation analysis showed that no other participants met the exclusion criteria. Our subsequent analysis used a Kolmogorov–Smirnov test of the distributions of the eye movement directions, which revealed no significant difference in eye movements between the pattern motion trial types for any of

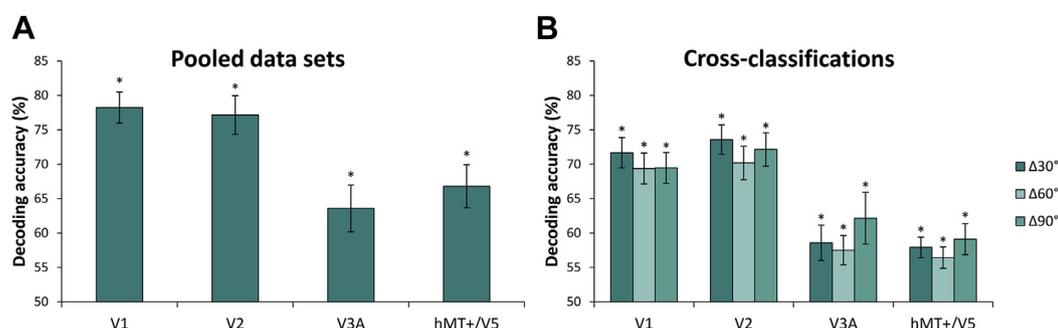


Fig. 2 – Decoding performance Experiment 1. Chance level is 50%. Error bars denote SEM. (A) The two pattern motion directions could be decoded significantly above chance in all ROIs. (B) Cross-classification decoding performance is pooled according to the angular difference between training and test set (Δ). Decoding performance was significantly above chance for all angles in all ROIs. No significant differences were found between angles.

the participants (all $p > .9$). Data from a representative subject are shown in Fig. 3.

4. Experiment 2

In Experiment 1, we hypothesised that cross-classifications between stimuli with different angles between the component motion directions (α) should show a graded drop in decoding accuracy if the classifier used component motion information to classify the test data. However, despite the varying α , all components that induced the same pattern direction in Experiment 1 moved in directions deviating no more than 75° from the pattern direction. This makes it still possible that neurons activated by the components of these separate stimuli had overlapping tuning curves, causing above-chance cross-classification. Hence, to further ensure that decoding performance was not based on component motion information, we conducted an additional control experiment. Here, we manipulated the perception of our plaid stimuli so that either the individual components or the pattern was perceived. Our aim was to show that generalisation of classifier performance across different values of α would be significantly lower if pattern motion was not present in the stimulus. This finding would rule out any dependence of component motion signals on successful pattern motion decoding in V1, as observed in Experiment 1. Our experiment used the two extreme α values from Experiment 1 (60° and 150°) to further avoid an overlap in tuning curves between the component motions in each stimulus set.

4.1. Material and methods

4.1.1. Participants

Thirteen participants with normal or corrected-to-normal vision, of which 10 had also participated in the previous experiment, were invited to take part in this control experiment. Four participants did not perceive the stimuli as intended and were excluded (see 4.1.3 Procedure). Thus, the data of nine subjects are presented (three male, six female, age 19–30, mean age 26, all right handed).

4.1.2. Stimuli

For the purpose of inducing component motion perception (i.e., motion transparency), we closely matched the stimulus properties of the previous experiment but used two square wave gratings, as opposed to sine-wave gratings (Castelo-Branco et al., 2002). We also presented all stimuli dichoptically through a custom made setup (Schurger, 2009) for the purposes detailed below. This meant that stimuli were considerably smaller than those presented in Experiment 1 (annulus of 7.4° of visual angle). In each decoding set (i.e., A vs B), plaids were composed of gratings with $\alpha = 60^\circ$ or $\alpha = 150^\circ$, where their vector average was either motion to the left (180° , case A) or motion to the right (0° , case B). The individual gratings had a spatial frequency of .5 cpd, a duty cycle .3, and a speed of 1 cycle/sec. The term duty cycle refers to the proportion of the width of the darker bars within one cycle of the grating. To enhance the perceptual segmentation of component motion within these plaid stimuli, we also manipulated the luminance of each grating so that they differed (i.e., 1401 cd/m^2 , and 739 cd/m^2). Furthermore, a binocular disparity shift was added such that the components were displaced by $.3^\circ$ across the two eyes, with motion direction at the crossed or uncrossed disparities being counter-balanced equally for each condition. Stimuli that were supposed to be perceived as patterns were presented dichoptically without a disparity shift. Since these stimuli nonetheless elicited bistable perception, instead of pure pattern motion percepts, we brightened the intersections of components (2012 cd/m^2) to enhance pattern motion perception. Thus, for the two opposing ‘pattern’ motion directions (i.e., left vs right, or A vs B), a total set of eight different types of plaid stimuli were generated, which varied in the possible angles defining the components’ motion directions (60° and 150°) and whether they were perceived as patterns or components (Fig. 4).

4.1.3. Procedure

In order to establish whether the perceptual manipulation, i.e., component motion perception versus pattern motion perception, was successful, participants were subjected to a pre- and post-scan behavioural test in which they had to indicate their percept. Each test consisted of 40 trials (five per

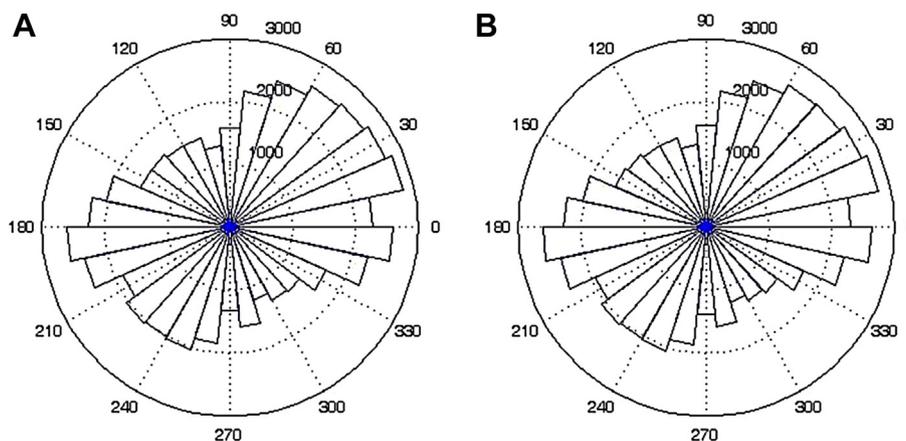


Fig. 3 – Eye movements for each pattern direction. Example from a representative subject for pattern direction A (225°) and B (45°). The directions of all eye movements were sorted into 30 bins for each pattern direction. A Kolmogorov–Smirnov test showed that the distributions for the two pattern directions did not significantly differ.

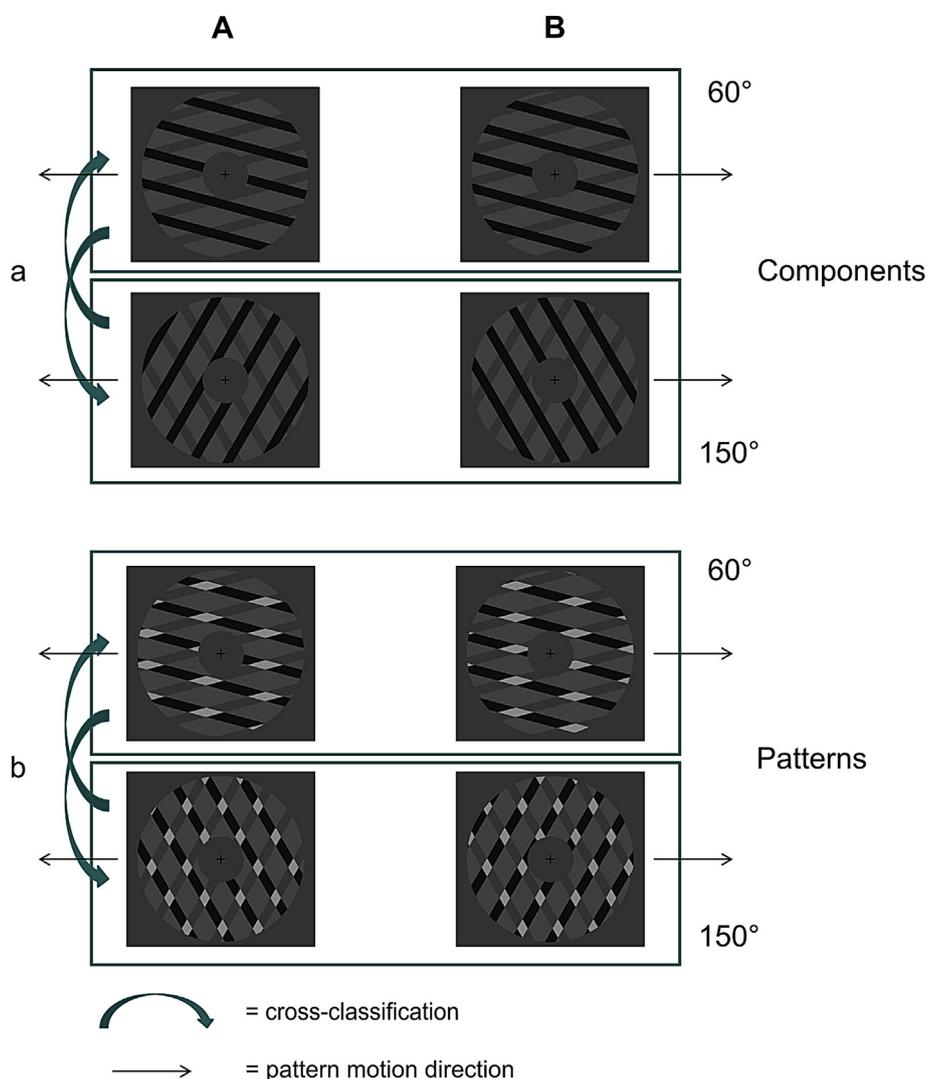


Fig. 4 – Stimuli and analysis of Experiment 2. Plaid stimuli with two pattern directions (A and B), composed of gratings with $\alpha = 60^\circ$ or $\alpha = 150^\circ$, were presented in pseudo-randomised order. Depth cues were introduced to half of the stimuli to induce component motion perception, whereas no depth cues and bright intersections were added to the other half to enhance pattern motion perception. Cross-classifications were performed by training on stimuli with $\alpha = 60^\circ$ and testing on stimuli with $\alpha = 150^\circ$ and vice versa for component (a) and pattern (b) stimuli separately.

stimulus type), which were presented in pseudo-randomised order. After each trial, participants indicated whether they perceived one or two motion directions, or a mixture of these percepts (i.e., pattern motion, component motion, or bistable perception between component and pattern motion). This test was conducted in the scanner bore, before the start and at the end of the scanning session. Participants who perceived more than 12.5% (i.e., 10 trials) of the total 80 trials differently to that intended were excluded. From two participants, the data from the pre-scan test were lost due to a technical issue; however they perceived 13.75% and 25% of the total number of trials differently than intended in the post-scan test alone, which already sufficed for exclusion. Two other participants perceived a total of 17.5% and 15% differently than intended, respectively, leading to a total of four excluded participants.

After the initial pre-scan test, participants completed 10 runs of the main experiment. Each run consisted of 24 trials

(three per stimulus type), presented in pseudo-randomised order. As in Experiment 1, participants had to detect a colour change of the fixation cross and a small speed change of the plaid stimuli, which occurred at random time intervals. Eye tracking was performed to ensure proper fixation. Additional stimulus localisers and retinotopic mapping scans were performed as described above (see [Experiment 1](#)). Two participants were not available for the further eccentricity mapping. Thus V3A could only be defined in seven participants. V1 and hMT+/V5 were defined in all nine participants.

4.1.4. Multivariate pattern analysis

Cross-classifications were performed between the two component stimulus types ($\alpha = 60^\circ$ and $\alpha = 150^\circ$) and between the two pattern stimulus types ($\alpha = 60^\circ$ and $\alpha = 150^\circ$), that is, the classifier was trained on stimuli with $\alpha = 60^\circ$ and tested on stimuli with $\alpha = 150^\circ$, and vice versa, for

component and pattern stimulus types separately (Fig. 4). Since there were no systematic differences in disparity for any pair of conditions that were decoded against each other, the disparity cue was uninformative to the classifier algorithms for differentiating these conditions. To determine whether the cross-classifications were significantly above chance, permutation testing was conducted using the procedure described earlier (3.1.4 Multivariate Pattern Analysis). Using a paired t-test, the cross-classification performance for components was compared with cross-classification performance for patterns.

4.2. Results

4.2.1. Multivariate pattern analysis

Cross-classification of stimuli perceived as single patterns, differing in their component motions, yielded significant above-chance decoding performance in all ROIs (V1: 71.9%, $p < .001$; V2: 69.4%, $p < .001$; V3A: 62.9%, $p < .001$; hMT+/V5: 54.2%, $p = .015$; Fig. 5). Cross-classification of similar stimuli perceived as two separate motion components led to chance-level decoding in V3A and hMT+/V5 (V3A: 49.6%, $p = .6$; hMT+/V5: 49.7%, $p = .53$), and significant above-chance decoding in V1 (55%, $p = .008$) and V2 (55.3%, $p = .004$). Although the cross-classification of components yielded weak but significant above-chance decoding performance in V1 and V2, decoding accuracy was significantly lower compared to the performance in the cross-classification of patterns (both $p = .003$).

4.2.2. Eye tracking

Due to technical difficulties, no usable eye tracking data were obtained for four participants. For another participant, eye tracking data were available for seven runs. The fixation analysis showed that no participant met the exclusion criteria. Our subsequent analysis used a Kolmogorov–Smirnov test of the distributions of the directions of eye movements made during leftward and rightward trials. This test revealed no significant difference in eye movements between the two trial types for any of the participants (all $p > .9$).

5. Discussion

Our results demonstrate the presence of information about pattern motion direction in neural activity in V1. We were able to decode pattern motion from data sets that contained stimuli with the same pattern direction, but different component directions. Subsequent cross-classifications between stimuli with different α also yielded significant above-chance decoding of pattern motion direction, and no graded decline in decoding performance was observed as the difference between α increased over training and test set. Thus, our findings suggest that successful pattern motion decoding generalises across component directions and is largely component-direction invariant. By manipulating perception of the plaid stimuli in Experiment 2 we provided further evidence that our classifier did not use component motion information to decode our stimuli. Here, cross-classification between trials perceived as patterns again yielded significant above-chance decoding performance, whereas decoding accuracies were markedly lower when stimuli were perceived as components. Since a similar drop in accuracy was not observed in the cross-classifications in Experiment 1, our result suggests that successful decoding of pattern motion in V1 is based on pattern direction information and not on overlapping tuning curves for the component motions. Eye tracking analyses also ruled out any contribution of eye movements to these results.

Our findings are in contrast to several studies that did not find evidence for pattern motion processing in V1. While most evidence from electrophysiology demonstrates a lack of pattern-selective cells in primary visual cortex (Gizzi et al., 1990; Movshon et al., 1985; Movshon & Newsome, 1996; Rodman & Albright, 1989; Rust et al., 2006), some studies have provided contradictory evidence: For instance, Guo et al. (2004) observed pattern motion responses in V1 of awake monkeys. Furthermore, Pack et al. (2003) found that the preferred motion direction of end-stopped cells in V1 was independent of the orientation of the presented bar, suggesting that V1 might contribute to solving the aperture problem (van Wezel & van der Smagt, 2003). Tinsley et al.

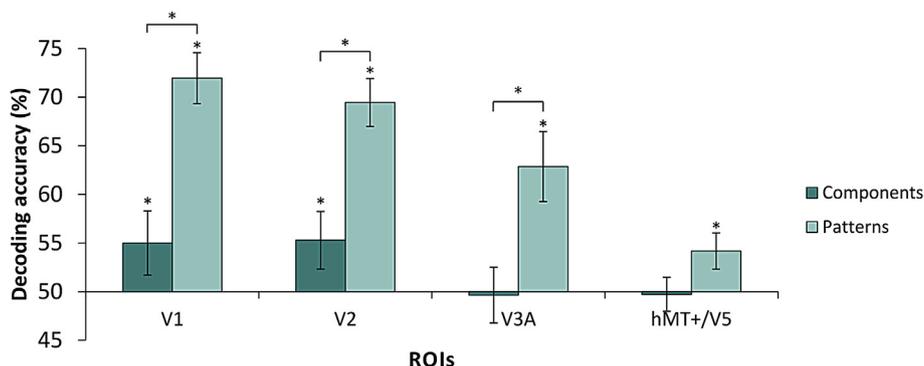


Fig. 5 – Decoding performance Experiment 2. Chance level is 50%. Error bars denote SEM. Cross-classifications across angles on plaids perceived as patterns yielded significant above-chance decoding performance in all ROIs. Cross-classifications across angles on plaids perceived as components resulted in chance-level decoding performance in V3A and hMT+/V5, and above-chance decoding performance in V1 and V2. There was a significant drop in performance when cross-classifying component stimuli compared to pattern stimuli in V1, V2, and V3A.

(2003) also reported that a subset of V1 neurons with short, wide receptive field sizes in the marmoset monkey respond to the pattern direction of a moving plaid stimulus. While these electrophysiological findings are in line with our data, a previous fMRI study in humans using an fMRI adaptation paradigm found pattern motion responses throughout the visual cortex from V2 onwards, but not in V1 (Huk & Heeger, 2002). Similarly, other studies have not found pattern motion activity in V1 but in higher-level visual areas (Castelo-Branco et al., 2002; Villeneuve et al., 2005, 2012). The discrepancy between these findings and our present data could be due to the superior sensitivity of current MVPA methodology over univariate approaches (Epstein & Morgan, 2012; Sapountzis, Schluppeck, Bowtell, & Peirce, 2010). In particular, univariate statistics test one voxel at a time, or rely on averaging across voxels, whereas multi-voxel pattern analysis takes patterns of activity into account (Haynes & Rees, 2006; Norman et al., 2006). Furthermore, MVPA potentially captures other neuronal properties than fMRI adaptation (Epstein & Morgan, 2012; Sapountzis et al., 2010), the method used by Huk and Heeger (2002). MVPA is thus likely to be more sensitive to subtle differences in activity patterns, and may therefore have been able to detect information that was not observed before with univariate methods.

The logic underlying our cross-classification analyses, where a classifier was asked to generalise across varying angular differences between the component motion directions, was based on the current knowledge of tuning curves of direction-selective cells. Direction-selective cells respond maximally to their preferred direction. The more the motion direction of a stimulus deviates from the preferred stimulus, the weaker the neural response to that stimulus will be. Studies measuring tuning curves for direction-selective cells in V1 report varying tuning widths (width of the resultant direction tuning curve at half of its maximum height), which are usually no greater than about 90° (Gizzi et al., 1990; Movshon et al., 1985; Movshon & Newsome, 1996). Thus, a stimulus that moves in a direction 45° away from a cell's preferred direction will elicit a neural response that is about half of the response to the preferred stimulus. In our experiments, α ranged from 60° to 150° , leading to a maximal angular difference between training and test set of 90° . Thus, each component in the stimulus with $\alpha = 150^\circ$ is moving in a direction 45° away from the direction of the components in the stimulus with $\alpha = 60^\circ$. Direction-selective cells responding maximally to the components in the plaid with $\alpha = 60^\circ$ should therefore show a substantially reduced response, namely approximately half of the maximal response, to the components in the plaid with $\alpha = 150^\circ$. Since our classifier could decode the pattern direction regardless of the angle between the components – and especially since there was no drop in decoding accuracy for cross-classifications across increasing angular differences – this suggests that component motion information did not inform the classifier on its decision. Rather, our results indicate that pattern motion signals were the informative feature to allow for the successful decoding of pattern motion direction in our stimuli. Crucially, in our second experiment, a generalisation across α yielded a significantly lower decoding performance when trained and tested on stimuli perceived as components compared to cross-

classifications of stimuli perceived as patterns. This shows that the classifier could not have used component motion information to generalise across α in our first experiment, since there we did not observe a significant drop in decoding accuracy with such cross-classifications.

It should be noted that there are physical stimulus differences between the component and pattern conditions in Experiment 2. The stimuli that were perceived as patterns had bright intersections, whereas these intersections were absent in trials perceived as components. Thus, it is possible that our classifier used neural responses to the motion of these intersections, which have the same direction as the global pattern, to decode pattern direction in Experiment 2. Whilst this is an important point concerning the interpretation of a drop in classifier performance between pattern and component motion conditions, it should be stressed that classifiers in Experiment 1 could successfully generalise across pattern motion stimuli whose components differed by the same values of α (i.e., $\Delta 90^\circ$) without this change in intersection luminance. These results render it unlikely that the high cross-classification performance for pattern stimuli in Experiment 2 was solely based on luminance differences of the intersections. Indeed the possibility that intersections may drive pattern motion responses is an argument commonly discussed in pattern motion studies. However, to date there is no consensus about the role of intersections in pattern motion perception. On one hand, there is evidence suggesting that tracking of local features such as intersections plays a role in pattern motion processing (Alais, Wenderoth, & Burke, 1997; Bowns, 1996; Delicato, Serrano-Pedraza, Suero, & Derrington, 2012; Wenderoth, Alais, Burke, & van der Zwan, 1994). On the other hand, there is evidence suggesting that pattern motion processing is not based on feature tracking mechanisms (Bowns, 2013). Whilst the experiments presented here were not designed to clarify the role of intersections in pattern motion processing, they do suggest that successful cross-classification of pattern motion stimuli in V1 could not be based on the decoding of the separate motion components, since this factor did not change with our stimulus manipulations from Experiment 1 to Experiment 2.

It is known that early visual cortex receives feedback projections from higher visual areas (Felleman & Van Essen, 1991). Since pattern-selective cells have been found mainly in higher visual areas, there is a strong likelihood that the pattern motion information we observed in V1 was a result of feedback. Indeed, an electrophysiological study by Pack, Berezovskii, and Born (2001) supports this possibility. These authors observed a change in neural responses of pattern-selective cells over time, with the initial response resembling component selectivity, before switching to a strong pattern selective response. This integration of component motion information was markedly impaired in anaesthetised animals, suggesting that feedback processes play a role in resolving ambiguous local motion information (Pack et al., 2001). Guo et al. also argued that pattern motion activity in V1 may depend on higher-level information integration mechanisms, such as feedback, that require consciousness, since they observed pattern motion responses in V1 only in awake but not anaesthetised monkeys (Guo et al., 2004). Furthermore, an optical imaging study that temporarily deactivated neural responses in the cat's

homologue of area MT (Schmidt, Lomber, Payne, & Galuske, 2011) showed that neural responses in areas 17 and 18 associated with pattern motion perception were reduced, whereas responses associated with component perception remained unaffected. Furthermore, these responses recovered when deactivation of this area was reversed. Such findings suggest that activity in low-level visual areas that is evoked by pattern motion may depend on feedback from higher-level visual areas (Schmidt et al., 2011). However, it is unclear from our study whether the pattern motion information found in human V1 is due to feedback signals from higher-level areas. Since feedback from higher visual areas to V1 also plays an important role in human visual perception (e.g., Silvanto, Lavie, & Walsh, 2005), this is a realistic scenario. Future studies that investigate the role of feedback signals in the processing of pattern motion in V1 will provide insight.

Taken together, using fMRI and MVPA we present evidence for the presence of direction-specific pattern motion information in V1. This finding suggests that V1 contributes to pattern motion processing in humans.

Acknowledgements

This work was supported by the German Research Foundation (STE-1430/2-1 and STE 1430/6-1) and a grant awarded to K.S. from the Alexander von Humboldt Foundation (100005156). B.V.K. was funded by the Berlin School of Mind and Brain, and by the Stichting dr Hendrik Muller's Vaderlandsch Fonds.

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.cortex.2014.04.014>.

REFERENCES

- Alais, D., Wenderoth, P., & Burke, D. (1997). The size and number of plaid blobs mediate the misperception of type-II plaid direction. *Vision Research*, 37(1), 143–150.
- Bowns, L. (1996). Evidence for a feature tracking explanation of why type II plaids move in the vector sum direction at short durations. *Vision Research*, 36(22), 3685–3694.
- Bowns, L. (2013). An explanation of why component contrast affects perceived pattern motion. *Vision Research*, 86, 1–5.
- Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., et al. (2002). Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21), 13914–13919.
- Delicato, L. S., Serrano-Pedraza, I., Suero, M., & Derrington, A. M. (2012). Two-dimensional pattern motion analysis uses local features. *Vision Research*, 62, 84–92.
- Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Jr., Le Goualher, G., Bruce Pike, G., et al. (2000). A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cerebral Cortex*, 10(5), 454–463.
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, 7(2), 181–192.
- Epstein, R. A., & Morgan, L. K. (2012). Neural responses to visual scenes reveals inconsistencies between fMRI adaptation and multivoxel pattern analysis. *Neuropsychologia*, 50(4), 530–543.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1), 1–47.
- Gizzi, M. S., Katz, E., Schumer, R. A., & Movshon, J. A. (1990). Selectivity for orientation and direction of motion of single neurons in cat striate and extrastriate visual cortex. *Journal of Neurophysiology*, 63(6), 1529–1543.
- Görgen, K., Hebart, M. N., & Haynes, J. D. (2012). The Decoding Toolbox (TDT): a new fMRI analysis package for SPM and MATLAB. In Poster presented at the 18th annual meeting of the organization for human brain mapping.
- Guo, K., Benson, P. J., & Blakemore, C. (2004). Pattern motion is present in V1 of awake but not anaesthetized monkeys. *European Journal of Neuroscience*, 19(4), 1055–1066.
- Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523–534.
- Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, 22(16), 7195–7205.
- Huk, A. C., & Heeger, D. J. (2002). Pattern-motion responses in human visual cortex. *Nature Neuroscience*, 5(1), 72–75.
- Hupe, J. M., & Rubin, N. (2004). The oblique plaid effect. *Vision Research*, 44(5), 489–500.
- Kamitani, Y., & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology*, 16(11), 1096–1102.
- 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.
- Lu, H. D., Chen, G., Tanigawa, H., & Roe, A. W. (2010). A motion direction map in macaque V2. *Neuron*, 68(5), 1002–1013.
- Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pontificiae Academiae Scientiarum Scripta Varia: Vol. 54. Pattern recognition mechanisms* (pp. 117–151). Rome: Vatican Press.
- Movshon, J. A., & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience*, 16(23), 7733–7741.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430.
- Pack, C. C., Berezovskii, V. K., & Born, R. T. (2001). Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature*, 414(6866), 905–908.
- Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. *Neuron*, 39(4), 671–680.
- Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, 75(1), 53–64.
- Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, 9(11), 1421–1431.
- Sapountzis, P., Schluppeck, D., Bowtell, R., & Peirce, J. W. (2010). A comparison of fMRI adaptation and multivariate pattern classification analysis in visual cortex. *NeuroImage*, 49(2), 1632–1640.
- Schmidt, K. E., Castelo-Branco, M., Goebel, R., Payne, B. R., Lomber, S. G., & Galuske, R. A. (2006). Pattern motion selectivity in population responses of area 18. *European Journal of Neuroscience*, 24(8), 2363–2374.

- Schmidt, K. E., Lomber, S. G., Payne, B. R., & Galuske, R. A. (2011). Pattern motion representation in primary visual cortex is mediated by transcortical feedback. *NeuroImage*, 54(1), 474–484.
- Schurger, A. (2009). A very inexpensive MRI-compatible method for dichoptic visual stimulation. *Journal of Neuroscience Methods*, 177(1), 199–202.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268(5212), 889–893.
- Seymour, K., Clifford, C. W., Logothetis, N. K., & Bartels, A. (2009). The coding of color, motion, and their conjunction in the human visual cortex. *Current Biology*, 19(3), 177–183.
- Silvanto, J., Lavie, N., & Walsh, V. (2005). Double dissociation of V1 and V5/MT activity in visual awareness. *Cerebral Cortex*, 15(11), 1736–1741.
- Slotnick, S. D., & Yantis, S. (2003). Efficient acquisition of human retinotopic maps. *Human Brain Mapping*, 18(1), 22–29.
- Smith, A. T. (1992). Coherence of plaids comprising components of disparate spatial frequencies. *Vision Research*, 32(2), 393–397.
- Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. *NeuroImage*, 65, 69–82.
- Stoner, G. R., & Albright, T. D. (1996). The interpretation of visual motion: evidence for surface segmentation mechanisms. *Vision Research*, 36(9), 1291–1310.
- Tinsley, C. J., Webb, B. S., Barraclough, N. E., Vincent, C. J., Parker, A., & Derrington, A. M. (2003). The nature of V1 neural responses to 2D moving patterns depends on receptive-field structure in the marmoset monkey. *Journal of Neurophysiology*, 90(2), 930–937.
- Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, 15(4), 3215–3230.
- Villeneuve, M. Y., Kupers, R., Gjedde, A., Ptito, M., & Casanova, C. (2005). Pattern-motion selectivity in the human pulvinar. *NeuroImage*, 28(2), 474–480.
- Villeneuve, M. Y., Thompson, B., Hess, R. F., & Casanova, C. (2012). Pattern-motion selective responses in MT, MST and the pulvinar of humans. *European Journal of Neuroscience*, 36(6), 2849–2858.
- Wallach, H. (1935). Ueber visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, 20, 325–380.
- Wenderoth, P., Alais, D., Burke, D., & van der Zwan, R. (1994). The role of the blobs in determining the perception of drifting plaids and their motion aftereffects. *Perception*, 23(10), 1163–1169.
- van Wezel, R. J., & van der Smagt, M. J. (2003). Motion processing: how low can you go? *Current Biology*, 13(21), R840–R842.