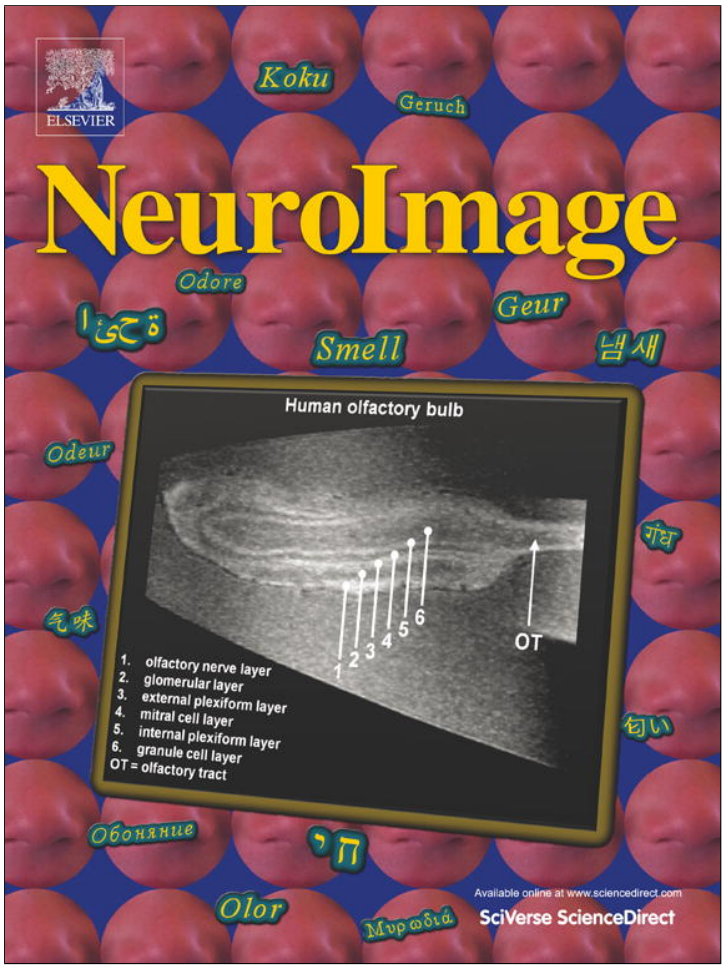

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

NeuroImage 61 (2012) 1143–1152

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Direction-specific fMRI adaptation reveals the visual cortical network underlying the "Rotating Snakes" illusion

Hiroshi Ashida [a,*], Ichiro Kuriki [b], Ikuya Murakami [c], Rumi Hisakata [c], Akiyoshi Kitaoka [d]

[a] Graduate School of Letters, Kyoto University, Sakyo, Kyoto 6068501, Japan
[b] Research Institute of Electrical Communication, Tohoku University, Sendai, Miyagi, Japan
[c] Department of Life Sciences, University of Tokyo, Tokyo, Japan
[d] Department of Psychology, Ritsumeikan University, Kyoto, Japan

ARTICLE INFO

Article history:
Accepted 8 March 2012
Available online 17 March 2012

Keywords:
Motion
Illusion
fMRI
Adaptation
V1
MT+

ABSTRACT

The "Rotating Snakes" figure elicits a clear sense of anomalous motion in stationary repetitive patterns. We used an event-related fMRI adaptation paradigm to investigate cortical mechanisms underlying the illusory motion. Following an adapting stimulus (S1) and a blank period, a probe stimulus (S2) that elicited illusory motion either in the same or in the opposite direction was presented. Attention was controlled by a fixation task, and control experiments precluded explanations in terms of artefacts of local adaptation, afterimages, or involuntary eye movements. Recorded BOLD responses were smaller for S2 in the same direction than S2 in the opposite direction in V1–V4, V3A, and MT+, indicating direction-selective adaptation. Adaptation in MT+ was correlated with adaptation in V1 but not in V4. With possible downstream inheritance of adaptation, it is most likely that adaptation predominantly occurred in V1. The results extend our previous findings of activation in MT+ (I. Kuriki, H. Ashida, I. Murakami, and A. Kitaoka, 2008), revealing the activity of the cortical network for motion processing from V1 towards MT+. This provides evidence for the role of front-end motion detectors, which has been assumed in proposed models of the illusion.

Introduction

Several anomalous motion illusions have been reported in the last decade. They are scientifically intriguing because they can reveal basic characteristics of biological visual motion processing. We are especially interested in cases where motion is perceived in still images, in which there are no motion signals in the images per se. Kitaoka's 'Rotating Snakes' (Kitaoka and Ashida, 2003) is one such illusion that yields a compelling sense of motion in a still image for many people (see Fig. 1). Each disc is perceived as rotating very slowly (roughly around 0.001 r/s in Hisakata and Murakami, 2008), smoothly, and continuously when viewed in the peripheral visual field. The illusion is weakened or abolished by hard fixation or central viewing (Hisakata and Murakami, 2008).

The figure consists of repetitions of micropatterns in which colours are arranged in the order: black–blue–white–yellow. The pattern appears to move in the direction of black to yellow. Spatial asymmetry in the luminance is crucial (Murakami et al., 2006), while colour is not essential, at least not the primal cause, since other combinations of colours including greys can induce a similar illusion (see http://www.lt.ritsumei.ac.jp/~akitaoka/index-e.html for variations).

The neural mechanisms underlying this illusion have not been clearly identified yet. In primates, it is widely accepted that motion signals are generated locally in V1 and are spatially integrated in area MT or later (e.g. Wilson et al., 1992). In accordance with this, direction-specific responses to a variant of 'Rotating Snakes' patterns have been reported in macaque V1 and MT (Conway et al., 2005), but there is not full evidence for humans yet. By using functional magnetic resonance imaging (fMRI), we have shown that human MT+ (possible homologue of macaque MT and MST) is activated by the 'Rotating Snakes' image (Kuriki et al., 2008). However, we did not find illusion-specific activation in V1. While it is possible that MT cells respond to motion directly (Thiele et al., 2004), absence of V1 activity in our previous study can be explained by the design of our experiment. We compared blood-oxygen-level-dependent (BOLD) responses to the illusion stimulus with those to a control stimulus that consisted of the same micropatterns with motion direction reversed between adjacent patterns. While we do not see global illusory motion in the control stimulus, motion detectors at the level of V1 might have responded to the local patterns. So far, there is only a preliminary report that suggests activation of human V1 by a similar image (Beer et al., 2007). Information is also lacking on the role of other visual areas in the occipital cortex.

One possible reason for difficulty in identifying illusion-related activity would be that neurons in the visual areas could respond more strongly to other features such as form and colour in these images,

* Corresponding author. Fax: +81 75 753 2445.
E-mail address: ashida@psy.bun.kyoto-u.ac.jp (H. Ashida).
URL: http://www.bun.kyoto-u.ac.jp/~hashida/ (H. Ashida).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2012.03.033

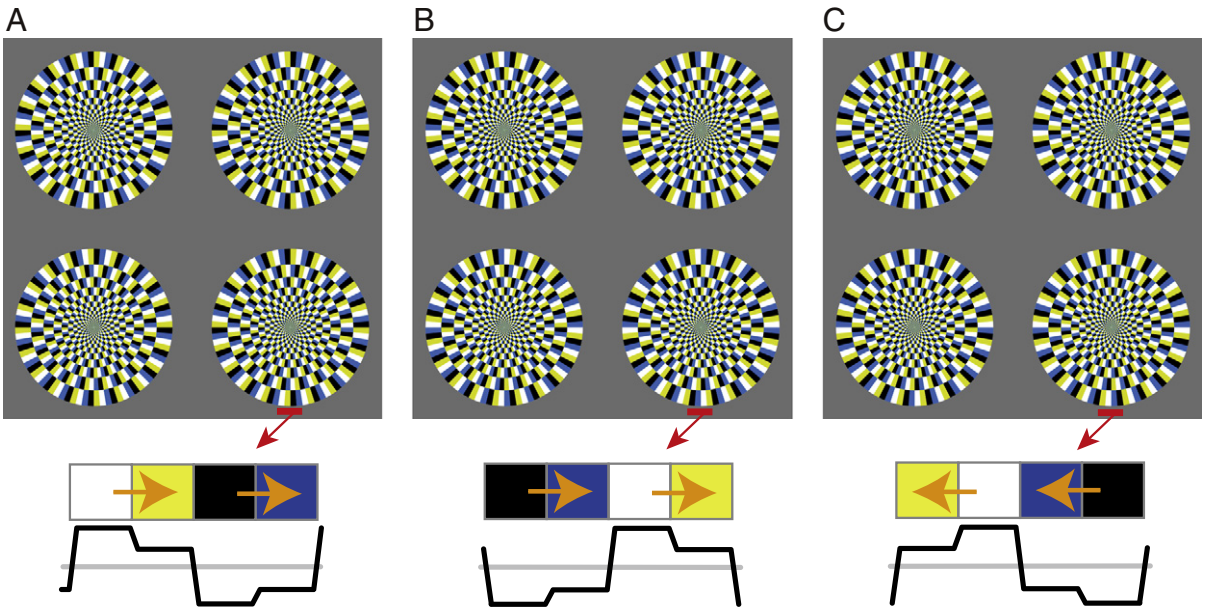

Fig. 1. The images used in the experiment. (A) S1 for adaptation. (B) S2-same, and (C) S2-opposite. A blurred fixation mark was shown schematically in the centre in each panel. Blue arrows in top panels indicate expected directions of perceived motion. The patches below each image show the colour order and the luminance profile. The yellow arrows show expected directions of illusory motion.

making it difficult to find subtle differences between the motion-related responses. We can cope with this problem by assessing adaptation (or repetition suppression) that is specific to motion direction. In a typical fMRI adaptation experiment, two stimuli are sequentially presented and the responses are compared with and without repetition of a certain feature attribute; repetition of the same attribute would reduce BOLD responses if the attribute is relevant for a population of neurons in the area of interest. Despite limitations in interpreting the results of adaptation (Bartels et al., 2008; Sawamura et al., 2006), this paradigm has been successfully applied to reveal properties of the ventral visual areas (see Grill-Spector et al., 2006 for reviews), and also motion related areas (Ashida et al., 2007; Lingnau et al., 2009; Smith and Wall, 2008; Wall et al., 2008).

In this study, therefore, we investigated direction selective adaptation of BOLD responses to the 'Rotating Snakes' image in visual areas of the occipital cortex. The results will be discussed in terms of the cortical network for visual motion processing.

Experiment

Material and methods

Participants

Seven adults took part in the experiment. All had normal or corrected-to-normal acuity and normal colour vision. We pre-screened the participants by showing related illusion figures, both in print and on computer screens, and all of the participants saw vivid motion in the expected direction. One participant was excluded from the analysis because of failure to achieve good performance of the fixation task (see below). The remaining six participants were aged between 20 and 41 (all males), including one of the authors and five naïve volunteers. The experiment was conducted in accordance with the Declaration of Helsinki, and was approved by the ethical and safety committees of ATR Brain Activity Imaging Center. Written informed consent was obtained from each participant, and naïve volunteers were paid for their participation.

Visual stimulation

Visual stimuli were generated with a computer running Windows XP and were back projected onto a screen by reflective liquid crystal projectors[1] (DLA-G150CL/ DLA-HD10KHK, Victor, Japan). The images were created for the resolution of 1024×768 pixels, up-converted by the projectors, and presented to the screen with a refresh rate of 60 Hz. Luminance profiles were measured for each type of sub-pixels (R/G/B), and linear summation of luminance for the three sub-pixels was confirmed for both projectors. The maximum luminance was 90 cd/m^2 or 110 cd/m^2. The screen was viewed through a mirror set above the head coil, with the viewing distance of 1.1 m. Magnifying goggles (MAXTV, Eschenbach, Germany; with metal parts removed) were used to enlarge the image by the factor of two. The goggles also provided acuity correction.

The stimuli in the main experiments were simplified 'Rotating Snakes' patterns with four discs on a grey background (Fig. 1), which were optimised by eye by one of the authors (A. Kitaoka). Here we define the luminance contrast as $c=(l-l_m)/l_m$, where l is the luminance and l_m is the average of the maximum and minimum luminance values. The four colours were black ($c=-0.99$), blue ($c=-0.62$), yellow ($c=0.44$), and white ($c=0.99$). The background was grey ($c=-0.64$). With the magnifying goggles, the image size was 28.2×28.2 deg (766×766 pixels). Each disc was 11.8×11.8 deg (320×320 pixels) in diameter, and the centre of each was 10.4 deg (280 pixels) away from the centre. As a fixation mark, a Gaussian blob (with standard deviation of 8 pixels≅0.2 deg) was presented in the centre of the display. This blurry mark was used in order to relax fixation to some extent, because hard fixation reduces the illusion.

The discs were presented in peripheral locations for stronger effects (Hisakata and Murakami, 2008) and for reducing attention to the change of patterns. The colour orders of the patterns between adjacent discs were arranged so that opposite motion was perceivable. The pattern in Fig. 1 (A) was used as the first stimulus (S1) for adaptation. The patterns in Figs. 1(B) or (C) were used as the second stimulus (S2) for probing adaptation with the same and opposite motion, respectively.

As illustrated in Fig. 1, S2-same had the same colour order, with the spatial phase shifted by a half cycle. The expected motion direction was the same. S2-opposite had the reversed colour order, and

[1] The projector was replaced before scanning the last two participants. The images were adjusted so that the luminance contrast of each image element was preserved.

the expected motion direction was the opposite. The steep edges between black and yellow, and those between white and blue were at the same locations for S2-same and S2-opposite, but the spatial gradients were reversed. In this way, the local colours always changed from S1 to S2 to minimise the effect of local adaptation. We used the same set of stimuli for all participants. It was crucial that S2-same and S2-opposite elicited the same level of responses by themselves. This was confirmed in a control experiment that measured BOLD responses to these two stimuli per se (see Control experiment 1 in Supplementary Document for more details). Perceived strength of illusion does not look different between Figs. 1 (B) and (C), as confirmed by an expert observer (A. Kitaoka, the creator of the original figure).

The regions of interest (ROIs) were defined individually by separate localiser runs. To localise MT+, the areas of the four discs in the main experiment were filled with white random dots, with the whole background being black. Each disc contained 150 round dots (9 min diameter) with the density of 9.4%. In the motion blocks, all dots rotated in one direction, and the overall speed was modulated by a sine function at 0.125 Hz, with the top speed of 5.5 deg/s in terms of visual angle (i.e. radial speed was not constant across radii). The rotation direction was altered every 8 s, which minimised motion aftereffects. Each dot had a lifetime of 1 s, not synchronised across dots, and reappeared at a random new location. Like the illusion images, neighbouring discs rotated in opposite directions. In the control blocks, the dots appeared and disappeared with the same lifetime but did not move.

Other retinotopic areas were defined by a conventional phase encoding technique (Sereno et al., 1995). We used a 45-deg wedge that consisted of random dots in red and green colours and that rotated about the central fixation point. Pilot studies have confirmed that phase maps obtained with this random-dot wedge are consistent, but are generally clearer than those obtained with a conventional checkerboard wedge, especially in MT+. The wedge rotated clockwise six times at a rate of 56 s per round.

Procedure

In the main fMRI adaptation experiment, the procedures generally followed those in earlier works (Lingnau et al., 2009; Wall et al., 2008). In one trial (event), S1 was presented for 3 s, followed by a blank of 0.5 s and S2 of 1.5 s. Both S1 and S2 were presented within trapezoidal temporal windows in order to reduce the effects of abrupt changes and afterimages; the contrast of S1 increased linearly during the initial 0.5 s and decreased during the final 0.5 s, and the contrast of S2 increased linearly during the initial 0.25 s and decreased during the final 0.25 s. With inter-trial intervals of 8 s, trials were initiated every 13 s. With TR of 2 s, this allowed temporal sampling of signal changes every 1 s on average. With the fixation task (see below), participants' expectations of the stimulus onset should be minimised and anyway would not favour a particular type of event.

There were three types of trials: “baseline” trials with S1 and a blank S2, “same” trials with S1 and S2-same, and “opposite” trials with S1 and S2-opposite. In a single scan run, there were six repetitions of each trial type. The first-order transition probabilities were equalised. To complete this, an extra trial was added in the beginning of each run, which was excluded from the analysis. With an initial blank period of 4 s and a final blank period of 11 s, a single run lasted 254 s. Eight runs were conducted for each participant.

The MT+ localiser run consisted of 16-s blocks of moving (M) and flickering control (C) stimuli. A sequence of M–C–C–M blocks was repeated four times. With an initial 4-s blank period, a single run lasted 260 s. The phase encoding run presented six 56-s cycles of the rotating wedge stimulus, lasting 340 s with an initial blank period of 4 s.

Since it is particularly important to avoid an artefact by diverting attention (Larsson and Smith, 2012), gaze and voluntary attention were controlled by a demanding task at fixation. The participants were instructed to count the frequency of a blue fixation mark throughout the run (Ashida et al., 2007; Kuriki et al., 2008; Wall et al., 2008). They were also instructed not to pay attention to the outer discs while performing the fixation task. The colour of the fixation mark changed at 3 Hz in the main adaptation runs. For the phase encoding and localiser runs, where control of attention was less important, the colour changed at 1 Hz.

The whole experimental session took less than 70 min, including an anatomical scan. For some of the participants, the anatomical and phase encoding scans were conducted on separate days.

Data acquisition

MRI images were obtained with a 3-T scanner (MAGNETOM Trio, A Tim system, Siemens, Germany) at ATR Brain Activity Imaging Center (Kyoto, Japan). A whole-head 12-ch array coil was used. T1-weighted anatomical images were obtained by the MP-RAGE sequence (Siemens, Germany), giving 1-mm isotropic voxels in 192 sagittal slices of 256×256 resolution. T2*-weighted functional images were obtained with a gradient-echo, echo-planar sequence (TR = 2 s, 3-mm isotropic voxels, 64×64 in-plane resolution with FOV = 192×192 mm, 30 near-axial slices parallel to the AC-PC line, flip angle = 80°, TE = 30 ms).

Data analysis

The data were analysed by using BrainVoyager QX (version 2.0; Brain Innovation, Inc., Maastricht, The Netherlands). The analyses were conducted in individual brain spaces. The anatomical images were shifted and rotated to a standard position (the axial plane through the anterior and posterior commissures was centred) and were segmented for white and grey matters in order to construct three-dimensional models of the surface between the white and grey matters. The surface model was then inflated to reveal the whole surface. The functional images were corrected for slice timing, corrected for head motion and aligned to a reference volume, by using the first volume of the first run as the reference of all functional volumes within a single scan session, and temporally high-pass filtered at three cycles per run. No spatial smoothing was performed. The images were then co-registered to the anatomical image of each participant and re-sliced.

The MT+ ROI was defined on the inflated cortical surface of each hemisphere by selecting the activated area based on the statistical contrast between responses for motion blocks and control blocks by using the generalised linear model (GLM) analysis. Thresholds for selecting activated voxels were adjusted by the false discovery rate (FDR) correction ($q = 0.05$). In some cases we decreased the threshold in order to keep the ROI within a reasonable area[2] by referring to the anatomical structure (Dumoulin et al., 2000) and the phase map responses. Other retinotopic areas were defined by drawing borders by eye on the basis of field-sign reversals in the polar phase map. Since the wedge stimulus covered a larger area than the stimuli of the main adaptation runs, the data from the main runs were analysed by GLM, using the standard model of hemodynamic response function, and the voxels that showed significant activation for the baseline trials ($p < .05$, FDR correction) were chosen within each visual area. V4 was defined as a full hemifield in the ventral surface (Wandell et al., 2007). V3B was defined following the original definition by (Smith et al., 1998), which corresponds to LO-1 by Larsson, Landy, and Heeger (Larsson et al., 2006).

Event-related averages of time courses were computed from the data of the main adaptation runs in each ROI and for each trial type. The baseline was adjusted for each trial within the time window of −2 to 4 s from the onset of S1. To summarise the effect of adaptation, the signals for the “same” and “opposite” conditions were each

[2] MT+ is usually buried within a sulcus, while it can be in several different sulci (Dumoulin et al., 2000). The criteria for reducing the size were somewhat arbitrary, but we have confirmed that changing the size did not substantially change the pattern of results.

averaged over a time window from 8 to 12 s, within which the response to S2 was prominent, after the baseline time course was subtracted to reveal S2-related responses (a method compatible with Smith and Wall, 2008). Adaptation indices were then calculated as $AI = (\text{“opposite”} - \text{“same”}) / \text{“opposite”}$, following previous work (Smith and Wall, 2008; Wall et al., 2008). A bootstrap procedure (Efron and Tibshirani, 1993) was used to estimate 95% confidence intervals with 10,000 repetitions.

Control experiment with weak motion illusion

There are two possible confounds that may yield spurious adaptation effects. First, if position-invariant order of colours is encoded, this could cause similar adaptation. Second, S2-same resembles the local complementary colours of S1; the preceding afterimage of S1 may reduce responses to S2-same. These possibilities are addressed in a control experiment.

We repeated the main experiment with a modified set of stimuli that elicit much less illusory motion. Yellow and white parts were swapped in the three images of Fig. 1, which resulted in the colour order of black–blue–yellow–white (see Control experiment 2 in Supplementary Document,). Phenomenologically, motion from black to blue competes with motion from white to yellow, and the perceived illusory motion is much weaker if not completely abolished. Three of the original participants were tested with the same procedure.

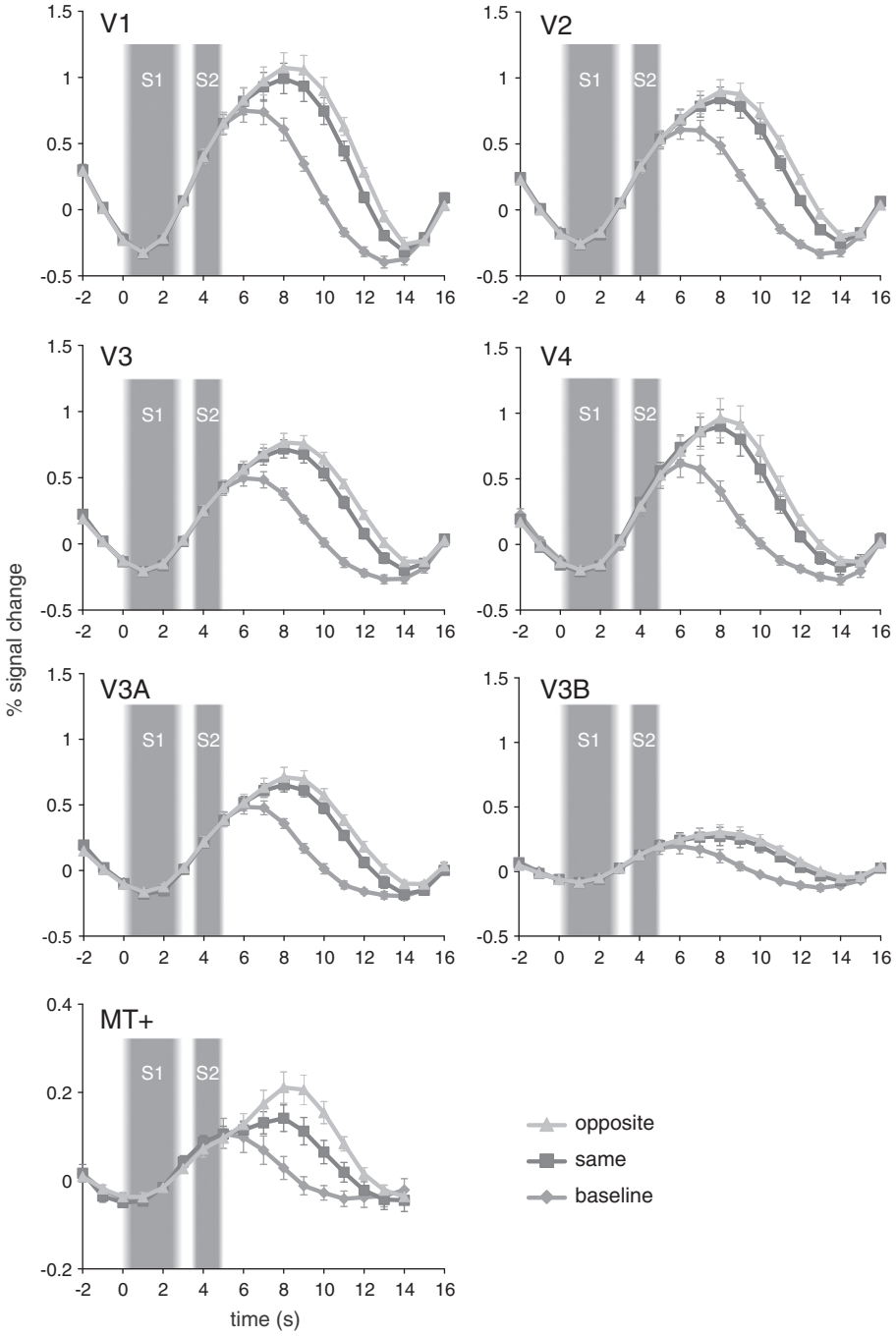

Fig. 2. Event-related averages of BOLD time courses in each visual area. Error bars = ±1 S.E.M. Grey stripes indicate the period of S1 and S2 presentation. Note that the range of the ordinate is different for MT+.

Eye movement recording

Fixation instability (i.e. variation of drift speed) correlates with strength of the illusion (Murakami et al., 2006), and is considered as a major cause of the sense of illusory motion. While the fixation task minimised voluntary eye movements, it remained possible that the S2-opposite image triggered larger involuntary fixational eye movements, which could have resulted in greater BOLD responses to the S2-opposite without neural adaptation.

We therefore conducted offline eye movement recording for the same set of stimuli. We presented images on a CRT monitor (Mitsubishi Electric RDF223H, 1024×768 pixels, refresh rate = 60 Hz, maximum luminance = 95.3 cd/m^2), with the image contrast adjusted to be the same as in the scanner. Stimuli were viewed under dim illumination in a dark room. EyeLink II (SR Research, Canada) was used to track movements of both eyes concurrently at 250 Hz. Analysis of eye movement data followed the method of Murakami (Murakami, 2004, 2010; Murakami et al., 2006). Drift eye movements were analysed along vertical and horizontal axes separately. Instantaneous velocities of drift were computed by differentiating eye position data with the algorithm of three-point differentiation, by excluding those exceeding 10 deg/s as putative microsaccades (Bair and O'Keefe, 1998), and by low-pass filtering (~30 Hz) the velocity within the period of S2 presentation. A histogram of instantaneous velocities was plotted with the bin width of 0.1 deg/s. A Gaussian was fitted and its standard deviation was taken as the index of fixation instability originating from eye drift. Microsaccades were determined as follows. Eye position data were smoothed by a three-point boxcar kernel (window size 16 ms), and two-dimensional instantaneous velocities were computed. Each jumpy movement was counted as a microsaccade if the maximum velocity exceeded 10 deg/s and if the motion directions of both eyes did not differ by more than 60 deg (Rolfs, 2009). Large saccades were scarcely found under the fixation control.

The stimulation sequence was the same as in the main experiment except that the first to-be-discarded trial was omitted because carry-over effects are much less problematic for eye movements than for BOLD responses. One run therefore consisted of 18 trials, which lasted 238 s with an initial blank of 4 s. Two runs were conducted for each participant. Two of the authors, including one who participated in the main fMRI experiment, and four paid volunteers were tested.

Results

The performance of the fixation task was generally good, with error rates below 1% for two participants and below 2% for three. One participant was excluded from the analysis, as noted before, because the error rate reached 18%. Another participant showed relatively poor performance (8.6% errors). Together with the head movement results, the last two runs of this participant were excluded from analysis. We have confirmed that the findings are mostly consistent if this participant is totally excluded. The generally good performance does not mean that the task was too easy. It may be possible just to detect a target without firm fixation, but it was almost impossible to keep track of counting at 3 Hz that requires fast and non-regular updating of working memory without properly attending to the fixation mark.

Results were analysed on the basis of individual hemispheres, because there were some missing hemispheres where significant activation for the baseline stimuli was not found: left V4, right V3A, and left V3B for one participant, and left V3B for another participant. Fig. 2 shows event-related averages of BOLD signal changes in each visual area. Time courses were first computed for each hemisphere and then averaged. Stronger activation was found for the "opposite" condition than the "same" condition in general, indicating direction-specific adaptation to the illusory motion.

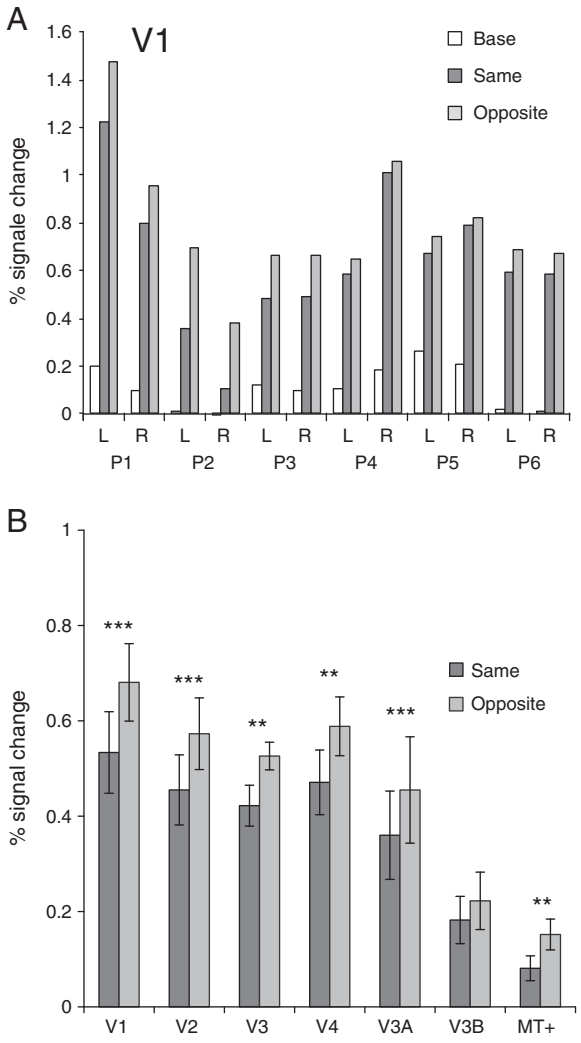

Fig. 3. Averaged responses within a time window between 8 s and 12 s from the onset of S1. (A) Results in V1 for each participant. (B) Averaged results across participants in each visual area, after subtracting the baseline responses. Error bars show ±1 S.E.M. Asterisks show significant differences between "opposite" and "same" conditions (***p<.001, **p<.01).

Fig. 3(A) shows averaged signal changes in V1, within the time window of 8 to 12 s from the onset of S1, for each condition and each participant. Despite variations in the level of responses, all the hemispheres show constant trends of baseline ≪ same < opposite. Fig. 3 (B) plots summaries of "same" and "opposite" responses in each visual area after subtracting the baseline responses. Difference between "same" and "opposite" was assessed by Wilcoxon's signed rank test. The response in the "opposite" condition was larger than that in the "same" condition in V1 ($p = 0.00048$), V2 ($p = 0.00097$), V3 ($p = 0.0024$), MT+ ($p = 0.0034$), V4 ($p = 0.0020$), and V3A ($p = 0.0098$), but not in V3B ($p = 0.19$).

Adaptation indices (AI) shown in Fig. 4 were significantly above zero in all areas except V3B, as shown by the bootstrap 95% confidence intervals, which is consistent with the results in Fig. 3(A). The AI values were all similar, between 0.2 and 0.25, except in MT+ (0.47). These values are similar or slightly larger than those in previous studies (Smith and Wall, 2008; Wall et al., 2008). Table 1 shows a correlation matrix of these AI values across areas, computed over the AI values for each participant (two hemispheres were averaged if both were available). Relatively high correlations suggest substantial inheritance across areas, and variations of levels of correlation may suggest distinct pathways. These points will be assessed in more detail with a path analysis in the Discussion.

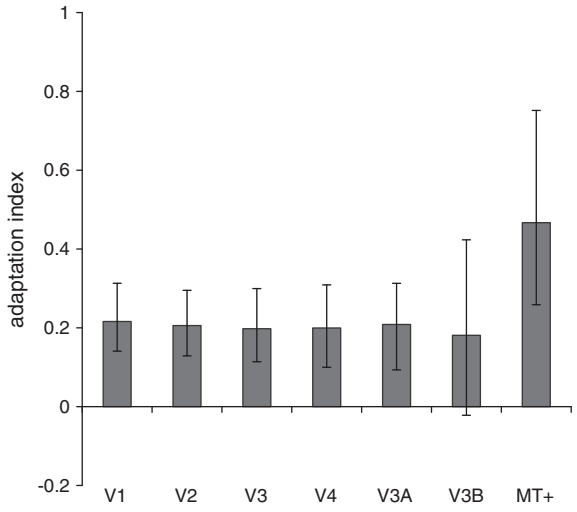

Fig. 4. Adaptation index in each visual area. Error bars show 95% confidence intervals estimated by a bootstrap procedure.

Control results with weak motion illusion

When illusory motion was made weak, just by changing the order of colours, the responses were almost the same for the two S2 conditions (Fig. 5(A)). Wilcoxon's signed rank test (across six hemispheres) revealed a significant difference between the "same" and "opposite" conditions in MT+ ($p=0.031$) and V4 ($p=0.031$) but not in V1 ($p=.15$), V2 ($p=0.21$), V3 ($p=0.15$), V3A ($p=0.21$), or V3B ($p=0.15$). Although the reason for this reversed effect (same>opposite) remains unclear, possible factors are distances between the same colours or spatial phase differences of the fundamental components in luminance between S1 and S2.

In contrast, the pattern of results of the main experiment was robust for the subset of participants (Fig. 5(B)). Wilcoxon's signed rank test (across six hemispheres) revealed a significant difference between the "same" and "opposite" conditions in MT+, V1, V2, and V4 (all $p=0.031$) but not in V3 ($p=0.15$), V3A ($p=0.15$), or V3B ($p=0.15$). Figs. 3(B) and 5(B) show similar tendencies, with significant differences in the most crucial areas of V1 and MT. Lack of adaptation effect therefore cannot be ascribed to the smaller number of participants. It is evident that neither colour order nor afterimage of S1 is important but instead illusory motion is crucial.

Eye movements

We conducted an offline recording of fixational eye movements during observation of our stimuli and compared fixation instabilities between the periods of the S2-same and S2-opposite. Fig. 6 (A) plots the distributions of instantaneous eye velocities for a representative participant who also participated in the main fMRI experiment. The distributions were very similar between the S2-same and S2-opposite conditions. Fig. 6 (B) plots an across-participant scattergram of drift indices for the two stimulus conditions. The slope of the best-fit linear function ($y=ax$) was 0.989 with a bootstrapped (with resampling of 10,000 times) 95% confidence interval (CI95) from 0.956 to 1.022, demonstrating that the data did not diverge from the line of unity. The averaged ratio of drift indices (opposite/same) was 0.990 (CI95: 0.945–1.026) for the horizontal direction and 1.011 (CI95: 0.973–1.057) for the vertical direction, showing no evidence of deviation from 1. The averaged frequencies of microsaccades were 1.28 times/s for the S2-same and 1.05 times/s for the S2-opposite, showing no significant difference ($t(5)=0.73$, $p>0.49$), but also note that it was previously suggested that microsaccades may not be relevant for this illusion in any event (Murakami et al., 2006).

To summarise, there is no evidence that the two stimuli, namely the S2-opposite and S2-same, trigger involuntary eye movement differently. It is highly unlikely that a difference in eye movements is the cause of the observed stronger BOLD signals for the S2-opposite than for the S2-same.

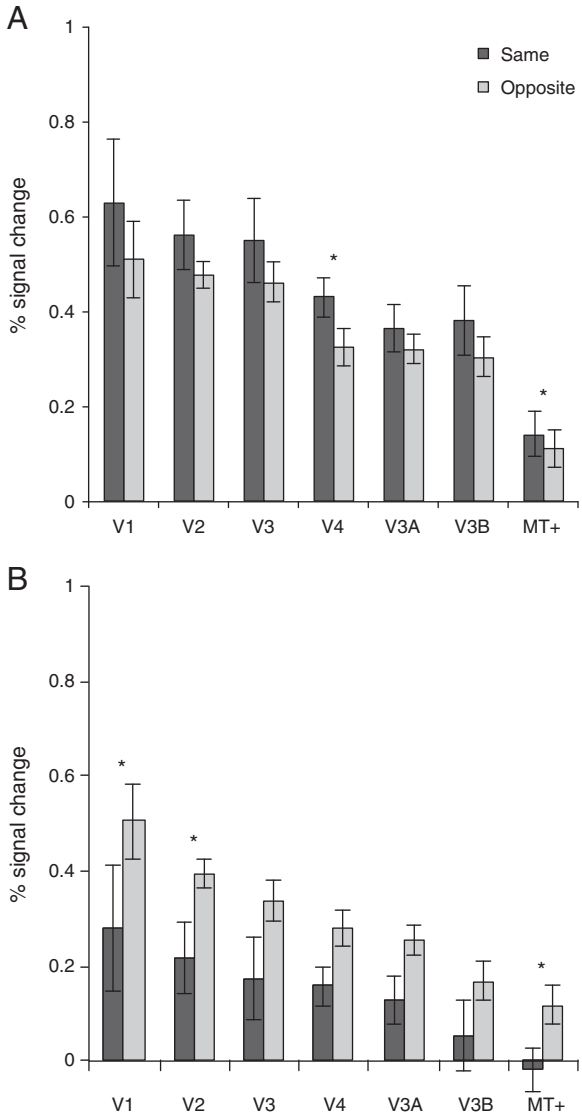

Fig. 5. (A) Results of the control experiment from three participants. Averaged responses between 8 s and 12 s from the onset are plotted for each visual area, after subtracting the baseline responses. Error bars show ±1 S.E.M. (B) Results of the same three participants obtained in the main experiment.

Table 1
Correlation matrix (Pearson's *r*) of adaptation indices.

	V1	V2	V3	V4	V3A	MT+
V1	1.000					
V2	0.961**	1.000				
V3	0.944*	0.993***	1.000			
V4	0.768	0.843*	0.830*	1.000		
V3A	0.847*	0.915*	0.902	0.783	1.000	
MT+	0.928*	0.906*	0.934*	0.654	0.781	1.000

*p<.05, **p<.01, ***p<.001 by *t*-test with false discovery rate correction.

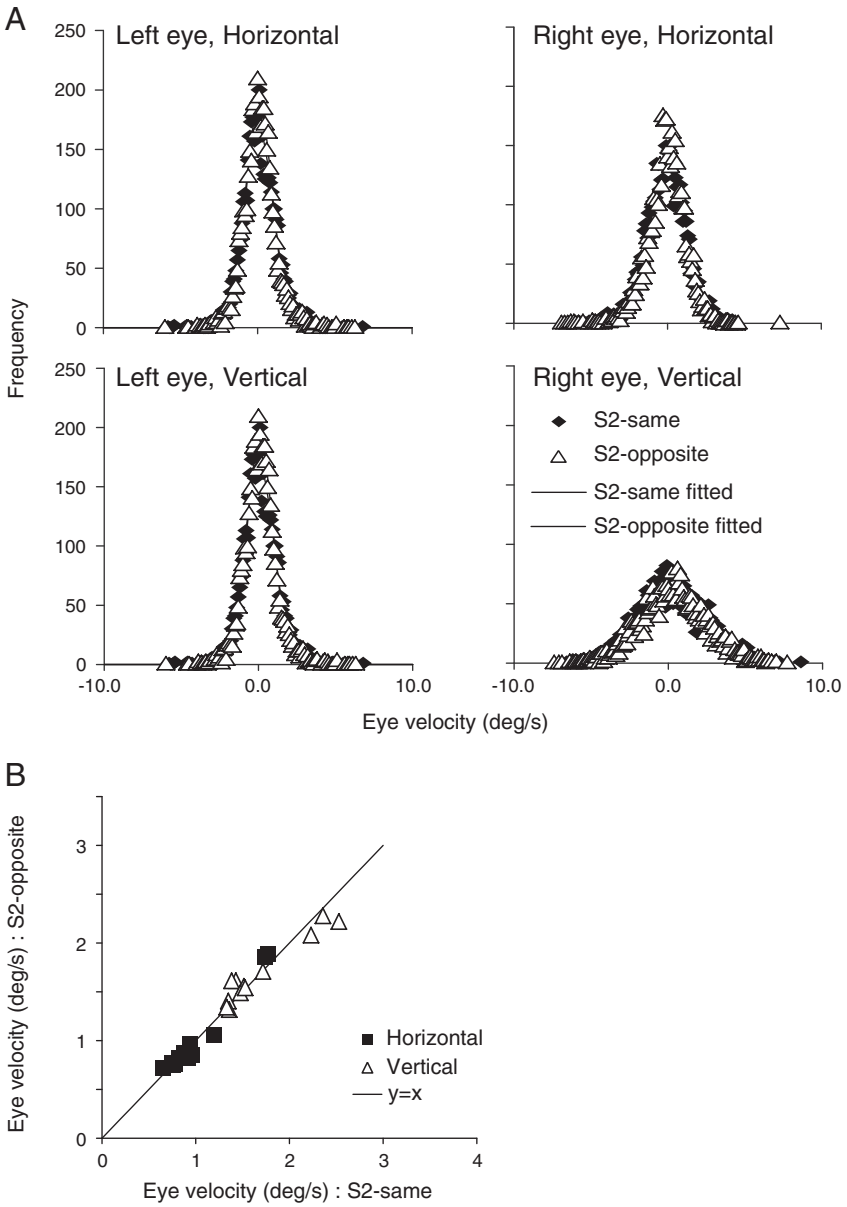

Fig. 6. Summary of the eye-recording results. (A) Histograms of instantaneous eye velocities from a representative participant, separately plotted for the two eyes and two directional axes. Fitted Gaussian curves for the two conditions are overlaid. (B) An across-participant scattergram of eye velocities (SD of fitted Gaussian to the histograms) as indices of drifts, for the S2-same and S2-opposite images. Each marker represents one eye along one directional axis.

Discussion

Adaptation to illusory motion in a static image

We found direction-specific fMRI adaptation to the 'Rotating Snakes' pattern in V1, V2, V3, V4, V3A, and MT+. The stimuli were matched as closely as possible for the two types of S2, while local features did not overlap between S1 and either type of S2. Spurious adaptation due to afterimages of S1 or static orders of colours is unlikely. Artefacts of eye movements are also rejected. It is therefore concluded that these cortical visual areas show direction-selective adaptation to the illusory motion in the 'Rotating Snakes' pattern, suggesting that this pattern elicits motion signals in a very early stage of visual motion processing.

While it is hard to specify the exact site of adaptation, due to downstream inheritance (Bartels et al., 2008), it is most likely that adaptation took place in V1. It is not likely for humans that direction specific adaptation occurs before V1. V2 or V3 is unlikely the main site of adaptation; V2 contains less direction-selective neurons than V1 in macaques (e.g. Orban et al., 1986), and human V3 may be a homologue of macaque V3A that is not strongly sensitive to motion (Singh et al., 2000; Tootell et al., 1997). Macaque physiology supports adaptation in V1 rather than MT because adaptation of MT neurons is observed only locally within the MT receptive field (Kohn and Movshon, 2003). The adaptation index was nearly doubled in MT+ compared to other areas, but this can be explained if only a subset of neurons are direction selective in V1 (e.g. 29% in De Valois et al., 1982; 27% in the central visual field in Orban et al., 1986) while most neurons are direction-selective in MT+ (Albright, 1984; Maunsell and Van-Essen, 1983). If MT+ neurons predominantly receive inputs from adapted direction-specific neurons in V1, overall AI in MT+ can be larger without additional adaptation in MT+. We do not exclude partial effects of adaptation in MT+, because other evidence shows that

MT+ can adapt in a different way from V1 (Kohn and Movshon, 2004; see also Wall et al., 2008).

The cause of adaptation in V4 remains unclear. V4 might adapt to motion per se to some extent. Direction selective adaptation in human V4 has been shown in other fMRI studies (Nishida et al., 2003; Smith and Wall, 2008), and many neurons in macaque V4 that are not directionally selective in a classical sense show direction-specific adaptation (Tolias et al., 2005). Human V3A is sensitive to motion (Singh et al., 2000; Tootell et al., 1997) and it might adapt to illusory motion. It is not surprising that V3B did not show significant motion adaptation, because this area is a part of lateral occipital complex (LOC). Larsson et al. (2006) also showed that this area (LO1) is not very sensitive to motion.

Path analysis on adaptation indices

Overall uniform adaptation or inheritance of adaptation from a single site may underlie the observed adaptation in most of the areas. We attempted to gain insights into this question by a path analysis. Structural equation modelling (SEM) (McIntosh and Gonzalez-Lima, 1994) was applied to the adaptation indices across participants. Note that we do not intend to show dynamic causal or effective connectivity, but to assess how adaptation across visual areas is interrelated. Correlations between areas are already shown in Table 1, but SEM provides more integrated assessment of partial correlations that appear crucial in this case.

Six AI values (one from each participant) were entered into the model for each of the four regions (V1–V3, V4, V3A, MT+). We collapsed the data in V1, V2, and V3 (for each participant) for two reasons. First, correlations among these areas are particularly high (Table 1), which poses a known problem of multicollinearity in SEM. Second, we do not have precise knowledge of anatomical connections between visual areas for humans. Because the sizes of V1, V2, and V3 are different, we re-calculated AIs from raw data with joint regions of interest, but similar results were obtained by simply averaging the AIs in these areas for each participant, or by using the AIs of V1 alone. V3B was not included because it did not show significant adaptation.

SEM analysis was conducted on the covariance matrix by using the "sem" package in R (Fox, 2006). Model paths were drawn assuming that higher areas receive inputs from V1–V3 and correlate with each other, namely, setting V1 as an exogenous variable and connecting other areas by double-headed arrows. We tried to set V3A, V4, and MT as the exogenous variable, since adaptation can originate in any area, but such models were not successful or at least not any better.

We started with a saturated model (with all possible paths), but path values were non-significant between V4 and MT+ ($z=-.451$, $p>.65$ uncorrected), between V4 and V3A ($z=.249$, $p>.8$), and between V3A and MT+ ($z=-0.645$, $p>.51$). We therefore omitted these connections and re-ran the model. As shown in Fig. 7, V1–V3 has significant connection to V3A ($z=6.278$, $p<.001$ with FDR correction), V4 ($z=3.066$, $p<.01$), and MT+ (4.152, $p<.001$). The path values did not change from the saturated model, indicating that the connections among higher areas have very small contribution, if any. This "separate paths" model was fitted fairly well ($\chi^2=0.87$, $df=3$, $p>0.87$; GFI = 0.931, AGFI = 0.772, SRMR = 0.025, Bentler CFI = 1, Tucker-Lewis NNFI = 1.25) and yielded the BIC of −4.66 that was the best among the models we have examined.

The SEM results support the idea that early visual areas (V1–V3) are the origin of adaptation and that adaptation was inherited independently for three pathways towards V4, V3A, and MT+. As we have discussed, V1 is a more likely candidate as a main adaptation site than V2 or V3, although a firm conclusion is open for future studies. Adaptation may have occurred within later areas, but relatively high path coefficients from V1 suggest that inheritance from V1 should have been predominant. It is interesting that MT+ and V3A are adapted independently. V3A may be involved in a different functional pathway from that which includes MT+, while its precise roles in visual motion processing are still open for further investigation. Note that these results do not necessarily exclude inter-connections among higher areas. Adaptation generally shows sharper tuning than neuronal responses (Sawamura et al., 2006), and independence of pathways may well be exaggerated if assessed by adaptation.

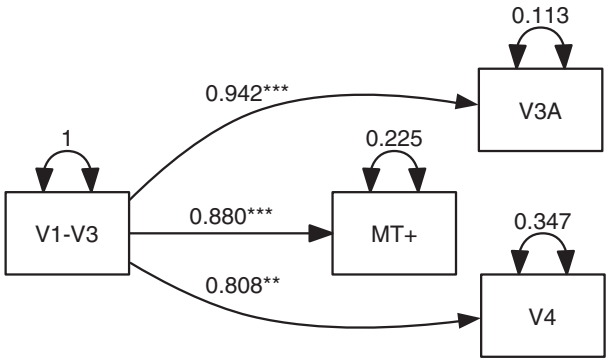

Fig. 7. A path diagram showing the results of SEM analysis. Stars with the path values show statistical significance (***p<.001, **p<.01). Self-directed double arrows show variances (residuals).

Effects of eye movement and attention

While the 'Rotating Snakes' illusion is affected or possibly caused by eye movement (Kuriki et al., 2008; Murakami et al., 2006), eye movement states should have been consistent across conditions in the current study. First, as participants were required to perform a relatively hard fixation task of colour counting at 3 Hz, it is unlikely that they made voluntary eye movement differently for each condition. It is also unlikely that involuntary fixational eye movement was induced differently under different conditions, as this was tested directly and found not to occur in offline measurements. Any change in the cortical activities should therefore be ascribed to neural responses to motion rather than differences in eye movements.

While we did not find significant activation under steady fixation in our previous study (NEM: no-eye-movement condition of Kuriki et al., 2008), we found significant adaptation with fixation in the current study. This apparent discrepancy is reconciled in two ways. First, in this study, the adaptation stimulus (S1) was presented only for 3 s, which corresponds to one interval of fixation in Kuriki et al. (2008) within a 15-s block. S2 was even shorter. Second, Murakami et al. (2006) found that the strength of the 'Rotating Snakes' illusion is not correlated with frequency of microsaccades. It is suggested that saccadic eye movements are not essential for generating motion signals but could help in the disengagement from local adaptation and/or enhance small eye movements. Actually, in the no-eye-movement condition of Kuriki et al. (2008, Fig. 2), the BOLD signals were slightly larger for the illusion stimulus than for the control stimulus in the first half of the block but were obscured later, suggesting adaptation during a block.

Voluntary attention affects BOLD responses especially in MT+ (Huk et al., 2001), which is a possible confound in interpreting the results. In our experiment, however, it is highly unlikely that stronger voluntary attention to S2-opposite than to S2-same stimuli was the cause of the different magnitude of responses. First, the task of colour counting at 3 Hz was quite demanding and it was not easy to attend to S2. Second, the control experiment did not show adaptation effects while the procedure was the same. Third, adaptation effects did not occur uniformly across visual areas, rejecting simple overall modulation of BOLD responses by attention.

Implication for computational models

Two major types of explanations have been proposed for the 'Rotating Snakes' illusion. First, different temporal responses at the onset of the stimulus after saccades and blinks create motion signals between areas of different luminance contrast levels (Backus and Oruç, 2005; Conway et al., 2005). Details vary, but the core assumption is that a high-contrast part of the image is transmitted to the cortex faster than a low-contrast part is, leading to motion from high-contrast to low-contrast regions (Faubert and Herbert, 1999). Second, the retinal images are always slightly moving due to small eye movements, and the illusion arises due to a direction-specific bias in sensing motion due to the spatial asymmetry (Fermüller et al., 2010; Murakami et al., 2006). Smooth perception of motion in one direction might be related to the image stabilisation mechanism that cancels the motion in the other direction (Beer et al., 2008). These two types of explanations are not mutually exclusive and possibly contribute in tandem to induce a greater effect. Both types of explanations assume local motion sensing followed by integration into rotation in the next stage, which is consistent with a widely accepted scheme of local motion detection in V1 and semi-global integration at the level of MT+ (e.g. Wilson et al., 1992).

The present results provide fMRI evidence for two stages of processing underlying the illusory motion. Local illusory motion signals are generated in V1, which pass through the pathway to MT+, possibly via V2 and V3 in part, just as real motion is processed. The motion sensors in V1 could respond to fragments of our illusory motion pattern. The figures of strong illusion are artificial, but such fragments may be more likely found even in natural scenes although they do not yield coherent global motion. In other words, the visual scene can be full of noisy spurious motion signals at the level of V1. Strong constraints are therefore essential at higher levels for understanding consistent visual scenes, such as the rigidity assumption (Ullman, 1979) or invoking a prior slower motion (Weiss et al., 2002). On the other hand, we need a mechanism for visual stabilisation that suppresses not just physical retinal slip by eye movements but all these real and illusory motion signals. This would give a rationale for minimal-motion-based stabilisation (Beer et al., 2008; Murakami and Cavanagh, 1998, 2001).

Conclusions

We demonstrated activation of a cortical motion network from V1 to MT+ underlying the illusory motion perception in the 'Rotating Snakes' figure. This finding supports proposed models of the illusory motion that assume local operation of motion sensors (Backus and Oruç, 2005; Conway et al., 2005; Fermüller et al., 2010; Murakami et al., 2006). Together with previous results, we conclude that local motion signals in response to asymmetric spatial patterns at the level of V1 are integrated and optimised in higher processing stages to create a strong sense of motion in the 'Rotating Snakes' pattern.

Acknowledgment

Supported by JSPS grant-in-aid for scientific research (B20330149 for H. Ashida and A22243044 for A. Kitaoka). Portions of the results were presented at VSS 2010 as an abstract form. We thank Andy Smith for helpful comments and suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2012.03.033.

References

Albright, T.D., 1984. Direction and orientation selectivity of neurons in visual area MT of the macaque. J. Neurophysiol. 52, 1106–1130.
Ashida, H., Lingnau, A., Wall, M.B., Smith, A.T., 2007. FMRI adaptation reveals separate mechanisms for first-order and second-order motion. J. Neurophysiol. 97, 1319–1325.
Backus, B.T., Oruç, İ., 2005. Illusory motion from change over time in the response to contrast and luminance. J. Vis. 5, 1055–1069.
Bair, W., O'Keefe, L.P., 1998. The influence of fixational eye movements on the response of neurons in area MT of the macaque. Vis. Neurosci. 15, 779–786.
Bartels, A., Logothetis, N.K., Moutoussis, K., 2008. fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends Neurosci. 31, 444–453.
Beer, A.L., Heckel, A., Winkler, J., Greenlee, M., 2007. A retinal compensation mechanism for fixation instability as revealed by the peripheral drift illusion. Perception 36, 3–4.
Beer, A.L., Heckel, A.H., Greenlee, M.W., 2008. A motion illusion reveals mechanisms of perceptual stabilization. PLoS One 3, e2741.
Conway, B.R., Kitaoka, A., Yazdanbakhsh, A., Pack, C.C., Livingstone, M.S., 2005. Neural basis for a powerful static motion illusion. J. Neurosci. 25, 5651–5656.
De Valois, R.L., Yund, E.W., Hepler, N., 1982. The orientation and direction selectivity of cells in macaque visual cortex. Vision Res. 22, 531–544.
Dumoulin, S.O., Bittar, R.G., Kabani, N.J., Baker Jr., 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.
Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman and Hall, New York.
Faubert, J., Herbert, A.M., 1999. The peripheral drift illusion: a motion illusion in the visual periphery. Perception 28, 617–621.
Fermüller, C., Ji, H., Kitaoka, A., 2010. Illusory motion due to causal time filtering. Vision Res. 50, 315–329.
Fox, J., 2006. Structural equation modeling with the sem package in R. Struct. Equ. Model. 13, 465–486.
Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23.
Hisakata, R., Murakami, I., 2008. The effects of eccentricity and retinal illuminance on the illusory motion seen in a stationary luminance gradient. Vision Res. 48, 1940–1948.
Huk, A.C., Ress, D., Heeger, D.J., 2001. Neuronal basis of the motion aftereffect reconsidered. Neuron 32, 161–172.
Kitaoka, A., Ashida, H., 2003. Phenomenal characteristics of the peripheral drift illusion. Vision 15, 261–262.
Kohn, A., Movshon, J.A., 2003. Neuronal adaptation to visual motion in area MT of the macaque. Neuron 39, 681–691.
Kohn, A., Movshon, J.A., 2004. Adaptation changes the direction tuning of macaque MT neurons. Nat. Neurosci. 7, 764–772.
Kuriki, I., Ashida, H., Murakami, I., Kitaoka, A., 2008. Functional brain imaging of the Rotating Snakes illusion by fMRI. J. Vis. 8 (16), 11-10.
Larsson, J., Landy, M.S., Heeger, D.J., 2006. Orientation-selective adaptation to first- and second-order patterns in human visual cortex. J. Neurophysiol. 95, 862–881.
Larsson, J., Smith, A.T., 2012. fMRI repetition suppression: neuronal adaptation or stimulus expectation? Cereb. Cortex 22, 567–576.
Lingnau, A., Ashida, H., Wall, M.B., Smith, A.T., 2009. Speed encoding in human visual cortex revealed by fMRI adaptation. J. Vis. 9, 1–14.
Maunsell, J.H., Van-Essen, D.C., 1983. Functional properties of neurons in middle temporal visual area of the macaque monkey: I. Selectivity for stimulus direction, speed, and orientation. J. Neurophysiol. 49, 1127–1147.
McIntosh, A.R., Gonzalez-Lima, F., 1994. Structural equation modeling and its application to network analysis in functional brain imaging. Hum. Brain Mapp. 2, 2–22.
Murakami, I., 2004. Correlations between fixation stability and visual motion sensitivity. Vision Res. 44, 751–761.
Murakami, I., Cavanagh, P., 1998. A jitter after-effect reveals motion-based stabilization of vision. Nature 395, 798–801.
Murakami, I., Cavanagh, P., 2001. Visual jitter: evidence for visual-motion-based compensation of retinal slip due to small eye movements. Vision Res. 41, 173–186.
Murakami, I., Kitaoka, A., Ashida, H., 2006. A positive correlation between fixation instability and the strength of illusory motion in a static display. Vision Res. 46, 2421–2431.
Murakami, I., 2010. Eye movements during fixation as velocity noise in minimum motion detection. Jpn. Psychol. Res. 52, 54–66.
Nishida, S., Sasaki, Y., Murakami, I., Watanabe, T., Tootell, R.B., 2003. Neuroimaging of direction-selective mechanisms for second-order motion. J. Neurophysiol. 90, 3242–3254.
Orban, G.A., Kennedy, H., Bullier, J., 1986. Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: influence of eccentricity. J. Neurophysiol. 56, 462–480.
Rolfs, M., 2009. Microsaccades: small steps on a long way. Vision Res. 49, 2415–2441.
Sawamura, H., Orban, G.A., Vogels, R., 2006. Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the FMRI adaptation paradigm. Neuron 49, 307–318.
Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893.

Singh, K.D., Smith, A.T., Greenlee, M.W., 2000. Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. NeuroImage 12, 550–564.
Smith, A.T., Greenlee, M.W., Singh, K.D., Kraemer, F.M., Hennig, J., 1998. The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). J. Neurosci. 18, 3816–3830.
Smith, A.T., Wall, M.B., 2008. Sensitivity of human visual cortical areas to the stereoscopic depth of a moving stimulus. J. Vis. 8 (1), 1–12.
Thiele, A., Distler, C., Korbmacher, H., Hoffmann, K.P., 2004. Contribution of inhibitory mechanisms to direction selectivity and response normalization in macaque middle temporal area. Proc. Natl. Acad. Sci. U. S. A. 101, 9810–9815.
Tolias, A.S., Keliris, G.A., Smirnakis, S.M., Logothetis, N.K., 2005. Neurons in macaque area V4 acquire directional tuning after adaptation to motion stimuli. Nat. Neurosci. 8, 591–593.
Tootell, R.B., Mendola, J.D., Hadjikhani, N.K., Ledden, P.J., Liu, A.K., Reppas, J.B., Sereno, M.I., Dale, A.M., 1997. Functional analysis of V3A and related areas in human visual cortex. J. Neurosci. 17, 7060–7078.
Ullman, S., 1979. The Interpretation of Visual Motion. MIT Press, Cambridge, MA.
Wall, M.B., Lingnau, A., Ashida, H., Smith, A.T., 2008. Selective visual responses to expansion and rotation in the human MT complex revealed by functional magnetic resonance imaging adaptation. Eur. J. Neurosci. 27, 2747–2757.
Wandell, B.A., Dumoulin, S.O., Brewer, A.A., 2007. Visual field maps in human cortex. Neuron 56, 366–383.
Weiss, Y., Simoncelli, E.P., Adelson, E.H., 2002. Motion illusions as optimal percepts. Nat. Neurosci. 5, 598–604.
Wilson, H.R., Ferrera, V.P., Yo, C., 1992. A psychophysically motivated model for two-dimensional motion perception. Vis. Neurosci. 9, 79–97.