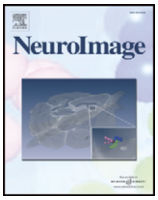
NeuroImage 42 (2008) 1529–1536

Contents lists available at ScienceDirect

NeuroImage

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



The integration of higher order form and motion by the human brain

Pegah Sarkheil [a,c,*], Quoc C. Vuong [a,b], Heinrich H. Bülthoff [a], Uta Noppeney [a]

[a] Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[b] Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK
[c] Department of Psychiatry and Psychotherapy, RWTH Aachen University, Aachen, Germany

ARTICLE INFO

Article history:
Received 17 February 2008
Revised 23 April 2008
Accepted 25 April 2008
Available online 13 May 2008

Keywords:
fMRI adaptation
Motion and form processing
Cue integration
Priming
Object recognition
Dorsal and ventral streams

ABSTRACT

Our experience with a dynamic environment has tuned our visual system to use form and motion as complementary sources of information for object recognition. To identify the neural systems involved in integrating form and motion information during dynamic object processing, we used an fMRI adaptation paradigm which factorially manipulated form and motion repetition. Observers were sequentially presented with pairs of rotating novel objects in which the form or rotation direction in depth could be repeated. They were required to discriminate either dimension of the second target object, while the first object served as a form and/or motion prime. At the behavioural level, observers were faster to recognize the target or discriminate its direction when primed by the same form. Importantly, this form priming effect was enhanced when prime and target objects rotated in the same direction. At the neural level, the two priming effects (i.e., the main effect of form repetition and the interaction between form and motion repetition) were associated with reduced activations in distinct brain regions. Bilateral lateral occipital regions exhibited reduced activation when form was repeated irrespective of rotation direction. In contrast, bilateral anterior fusiform and posterior middle temporal regions (overlapping with hMT+/V5) regions showed an adaptation effect that depended on both form and motion direction. Thus, the current results reveal a visual processing hierarchy with lateral occipito-temporal cortex representing an object's 3D structure, and anterior fusiform and posterior middle temporal regions being involved in spatio-temporal integration of form and motion during dynamic object processing.

Introduction

Form and motion are two natural and complementary sources of information that contribute to visual object recognition (Kelly and Freyd, 1987; Mitsumatsu and Yokosawa, 2003; Pike et al., 1997; Stone, 1998; Stone, 1999; Ullman, 1979; Wallis and Bulthoff, 1999). Motion provides visual cues for object boundaries, local part structures and view-specific features, all of which can facilitate the extraction of form and three-dimensional (3D) structure which are, in turn, important for recognizing objects (Marr and Vaina, 1982). In his ecological approach to perception, Gibson (1979) emphasized that a perceptual system exploits as much of the sensory input for the task at hand. In line with this emphasis, behavioural studies have shown that motion information *per se* is also important for object recognition. For example, the direction of rotation in depth, which preserves 3D shape and view information, affects observers' performance across a range of stimuli and recognition tasks. In this regard, several studies have shown that reversing learned motion patterns impairs recognition performance (Liu and Cooper, 2003; Stone, 1998, 1999; Vuong and Tarr, 2006; Wallis and Bulthoff, 2001). Furthermore, this motion-reversal effect can be modulated by the 3D geometry of objects (Vuong and Tarr, 2004, 2006). Collectively, these findings suggest that high-level form (i.e., 3D structure) and motion (i.e. rotation direction) information are integrated during object recognition. Both 3D form and rotation direction are considered high-level in the sense that they require the integration of low-level form (e.g., contours, colour, shading, etc.) and motion information (e.g., optic flow, contour deformations, etc.).

The complementary nature of form and motion information seems to be reflected in the functional organization of the visual system. Anatomical, physiological and lesion studies in non-human primates suggest that form and motion are processed within two specialized parallel functional streams: a ventral pathway which analyzes progressively more complex form information and a dorsal pathway which analyzes progressively more complex motion information (Ungerleider and Mishkin, 1982). More recently, using functional imaging

---

* Corresponding author. Max Planck Institute for Biological Cybernetics, Spemannstr. 41, 72076 Tübingen, Germany. Fax: +49 7071 601 616.
E-mail address: pegah.sarkheil@tuebingen.mpg.de (P. Sarkheil).

1053-8119/$ – see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.neuroimage.2008.04.265

(fMRI), a similar division has been demonstrated in the human visual system. Regions along ventral occipito-temporal cortex, such as the lateral occipital complex (LOC) and fusiform regions are involved in form and 3D shape processing (Grill-Spector et al., 2001; Kourtzi and Kanwisher, 2001; Malach et al., 1995). By comparison, regions along the human dorsal pathway, such as hMT+/V5 (Tootell et al., 1995; Zeki et al., 1991), V3a (Goebel et al., 1998), KO (Van Oostende et al., 1997) and IPS (Orban et al., 2003; Sunaert et al., 1999), are involved in various functional aspects of motion processing.

However, in line with the reported behavioural interactions between form and motion information, the two functional streams are not completely independent. There are at least two lines of evidence to suggest that interactions can occur at the neural level. First, direction-selective neurons have been found in the ventral pathway (Deyoe and Vanessen, 1985; Zeki, 1978). Conversely, orientation or shape-selective responses are observed in the dorsal pathway (Albright, 1984; Kourtzi et al., 2002). Second, several human fMRI studies investigating structure-from-motion (SFM) processing have identified multiple sites along the ventral and dorsal pathways that are concurrently involved in deriving 3D shape from the motion of random-dots displays (Kriegeskorte et al., 2003; Murray et al., 2003; Paradis et al., 2000; Treue et al., 1991). In SFM processing, however, form perception relies on motion perception. This precondition leaves out many natural circumstances in which form and motion can serve as direct cues to object identity.

To characterize the neural systems involved in integrating form and motion information during dynamic object recognition, we used an fMRI adaptation paradigm (Grill-Spector et al., 2006; Henson, 2003; Henson and Rugg, 2003, see also Geng et al., 2006) which independently manipulated repetition of form and motion information. In our study, observers were presented with pairs of rotating objects in which the form or rotation direction of the second object was the same as or different from the first. They were required to discriminate the form or rotation direction of the second object. Following the rationale of an fMRI adaptation paradigm, we then identified form- and motion-selective regions by testing for a reduced BOLD response for pairs of objects that were the same with respect to form or rotation direction. Importantly, the factorial nature of our paradigm allowed us to identify neural populations in which form and motion information may interact; that is, where the magnitude of form adaptation depends on whether the two objects rotated in the same or different direction or conversely, where the effect of motion adaptation depends on whether or not the two objects had the same form. In other words, by manipulating form and motion factorially, this fMRI adaptation paradigm allowed us to identify neural populations which integrate form and motion information, and relate those regions topographically to brain regions that are selective for either form or motion information. More specifically, we asked whether form and motion information converges in a single higher level visual region (e.g., IPS) or is integrated in a distributed system crossing the ventral and dorsal streams.

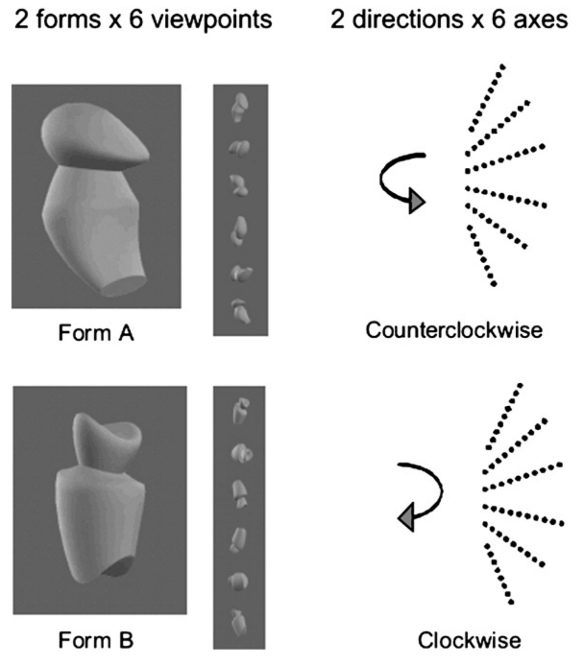


Fig. 1. Two novel 3D objects (form A and B) were animated with a depth rotation of 36° in two directions (clockwise or counter-clockwise). To equate form and 3D motion variability, six axes of rotation and six head-on viewpoints were used.

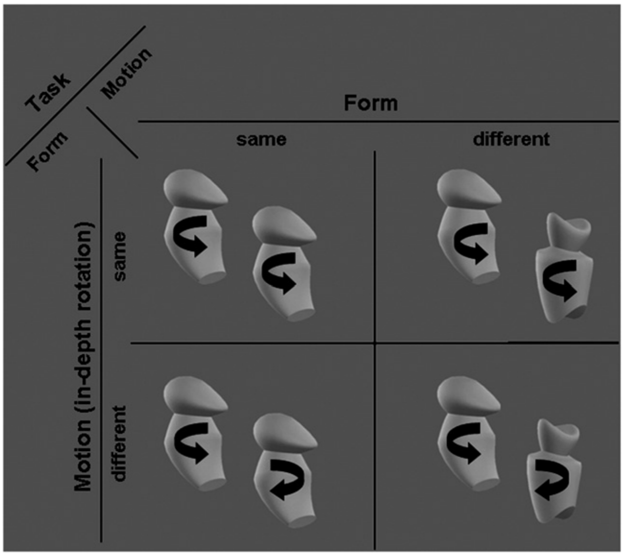


Fig. 2. Study design and example stimuli. The 2×2×2 factorial design with the factors: (i) task (form vs. motion), (ii) form repetition (same vs. different) and (iii) motion repetition (same direction vs. different direction).

Materials and methods

Subjects

19 healthy volunteers (7 females; aged 21 to 32 years, mean 24.8 years) with normal or corrected-to-normal vision gave informed consent to participate in the study. The study was approved of by the joint ethics committee of the Max Planck Institute and University Hospital, Tübingen, Germany.

Stimuli

Fig. 1 illustrates the dynamic stimuli used (see also supplementary videos). Two novel 3D objects (form A and B) were generated in 3D Studio Max version 8.0 (Discreet, Montreal, Canada). These objects were animated with a depth rotation of 36° in two directions (clockwise or counter-clockwise) about a horizontal axis orthogonal to the object's elongation. This axis of rotation was rotated 15°, 45°, 75°, 105°, 135° and 165° in the picture plane (0° is the vertical axis). In addition, each object was presented from six viewpoints from which depth rotation began. A head-on view of the object was arbitrarily designated as the 0° viewpoint and the five other viewpoints were created by rotating the objects in 60° increments about the axis of rotation (see Fig. 1). Importantly, these manipulations allowed us to equate form (i.e. viewpoint) and 3D motion variability (i.e. rotation axis).

Each animation consisted of 10 frames (3.6°/frame) presented at 29 frames/s. The animation was presented at the centre of the screen with a visual angle of approximately 4°.

Design and procedure

A 2×2×2 within-subject factorial design was used which manipulated (1) task, (2) form repetition, and (3) motion repetition (Fig. 2). Observers were presented with pairs of in-depth rotating objects in sequence, in which the first object in the pair was the prime and the second was the target. Both the form and the rotation direction of the target could repeat or change with respect to the prime. In other words, the prime and target objects could have the same or different form (i.e. object A or object B) and rotation direction (i.e. clockwise or counter-clockwise). The observers judged (i) whether the second, i.e., target object was object A or object B (form task) or (ii) whether the target rotated clockwise or counter-clockwise (motion task). They were not asked to perform a task on the prime object to avoid response priming. Observers responded as quickly and as accurately as possible by a two-choice key press.

Fig. 3 illustrates the sequence of trial events. Each trial of the experiment consisted of the presentation of an animation of the prime (350 ms), followed by a visual mask (100 ms), the presentation of an animation of the target (350 ms), and finally 2500 ms response period. Observers were instructed to fixate a cross presented in the centre of the screen. The axis of rotation and the head-on view was constant within a pair, only rotation direction (clockwise or counter-clockwise) or form (object A or object B) varied. Thus, in a trial where form but not motion direction were repeated, identical viewpoints of the object were presented with only the order of the images being reversed.

The trial duration during the experiment was fixed regardless of the subjects' response and the trial onset asynchrony was 3300 ms. No feedback was provided to observers. The visual mask on each trial was a brief 3-frame animation sequence, consisting of two frames of the prime animation which had their Fourier phase component scrambled, separated by a uniform grey image. The masks were used to prevent the percept of any apparent motion continuity between prime and target animations.

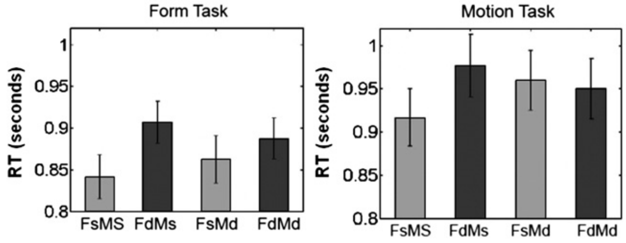


Fig. 4. Bar graphs showing across-volunteer means (+standard errors) of reaction times for form task (left) and motion task (right). n=19. FsMs = Form same Motion same, FsMd = Form same Motion different, FdMs = Form different Motion same, FdMd = Form different Motion different.

Altogether there were 144 animations (2 objects × 2 directions × 6 viewpoints × 6 rotation axes). Each animation was presented four times as primes and four times as targets. The stimuli and order of conditions were randomized with respect to the form and motion repetition factors. In alternating blocks of eight trials, the observers' task was to judge whether the target was object A or object B (form task) or judge whether the target rotated clockwise or counter-clockwise (motion task). The order of the task was counterbalanced within and between observers. The blocks were preceded by a 2 s visual instruction indicating the task to be performed on that block.

1–3 days prior to the scanning session, observers were trained on the task by means of trial-based feedback. The training session was otherwise identical to the test session in terms of paradigm and duration.

Data acquisition

A 3T whole-body scanner (Magnetom Trio, Siemens, Erlangen, Germany) was used to acquire both T1 anatomical volume images and T2*-weighted axial echoplanar images with blood oxygenation level-dependent (BOLD) contrast (TE = 40 ms, TR = 3 s, 37 axial slices, acquired sequentially in ascending direction, matrix 64 × 64 pixels, spatial resolution 3 × 3 × 3 mm^3 voxels, interslice gap 0.5 mm, slice thickness 2.5 mm). Data were acquired in four sessions with a total of 188 volume images per session. The first four volumes were discarded to allow for T1 equilibration effect.

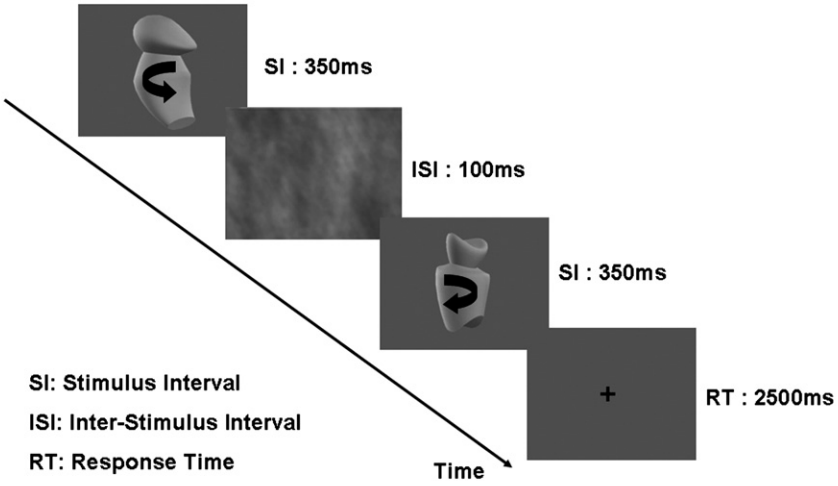


Fig. 3. Example run and timing of one trial. The experimental trial consisted of successive presentations of prime, mask and target stimuli, followed by the response interval.

Table 1
Across-volunteer means (+standard errors) of performance accuracy (%) for the form and motion task

	FsMs	FdMs	FsMd	FdMd
Form task	97.28 (0.8)	95.83 (1.1)	96.58 (0.9)	95.83 (0.9)
Motion task	89.85 (1.7)	92.82 (1.3)	90.28 (1.6)	90.22 (1.9)

$n = 19$.
FsMs = Form same Motion same; FdMs = Form different Motion same; FsMd = Form same Motion different; FdMd = Form different Motion different.

Data analysis

Data were analyzed with statistical parametric mapping (using SPM2 software from the Wellcome Department of Imaging Neuroscience, London, UK; (Friston et al., 1995)). Scans of each observer were realigned using the first as a reference, spatially normalized into MNI standard space (Evans et al., 1992), resampled to $3 \times 3 \times 3$ mm^3 voxels and spatially smoothed with a Gaussian kernel of 8 mm full width at half maximum (FWHM). The time-series in each voxel was high-pass filtered to 1/128 Hz and globally normalized with proportional scaling.

The fMRI experiment was modeled in an event related fashion using regressors obtained by convolving each unit impulse synchronized to the onset of the target with a canonical hemodynamic response function and its first temporal derivative. In addition to modeling the eight conditions in our $2 \times 2 \times 2$ factorial design (only correct trials included), the statistical

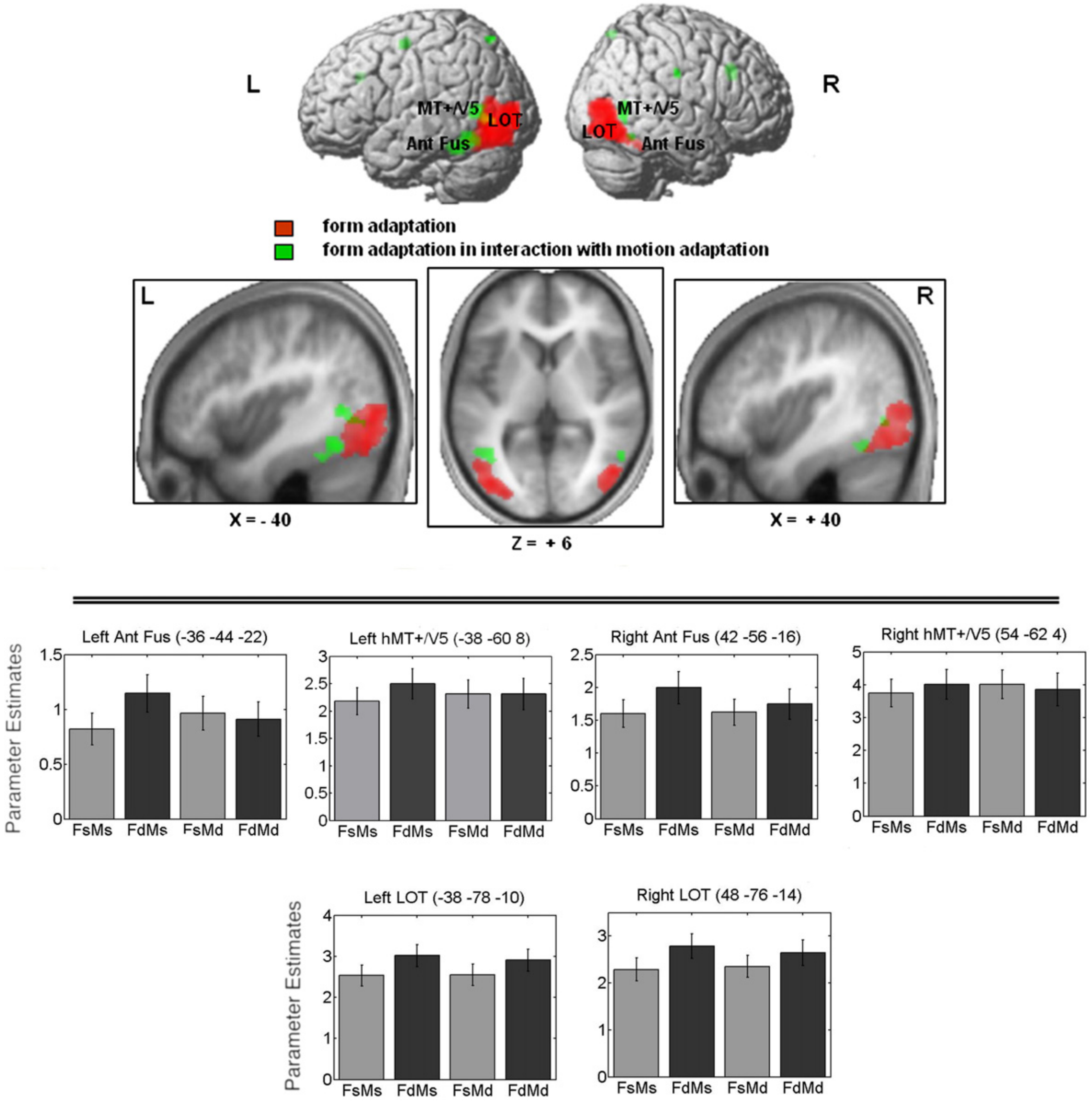


Fig. 5. Top: Activations pertaining to the main effect of form repetition (red) and the interaction of form and motion repetition (green) are shown on sagittal and transverse slices of a mean structural image created by averaging the subjects' normalized structural images. $p < 0.001$ (uncorrected for illustration purposes). Extent threshold >0. Bottom: Parameter estimates (+/− standard errors) averaged across all voxels within a significant cluster for same (grey) and different form (black) trials relative to fixation in (i) the lateral occipital regions showing the main effect of form repetition and (ii) the anterior fusiform and middle temporal regions showing an interaction of form and motion repetition. The bar graphs represent the size of the effect in non-dimensional units (corresponding to % whole brain mean). These effects are activations pooled (i.e. summed) over form and motion tasks. LOT: lateral occipito-temporal region, Ant Fus: anterior fusiform gyrus, FsMs: Form same Motion same, FsMd: Form same Motion different, FdMs: Form different Motion same, FdMd: Form different Motion different. The coordinates refer to the location of the peak activation within a significant cluster.

model included errors and instructions presented at the beginning of the block. Covariates of no interest included the realignment parameters (to account for residual motion artifacts). In a second additional GLM analysis, reaction times were modeled as a covariate of no interest. This second analysis allows us to investigate whether any priming-induced activation decreases persist even when differences in task difficulty (as indexed by reaction times) are modeled as confounds. For both analyses, condition-specific effects for each observer were estimated according to the general linear model and passed to a second-level analysis as contrasts. This involved creating contrast images testing (1) the main effect of form repetition, (2) the main effect of motion repetition and (3) the interaction between form and motion repetition. For completeness, we also tested for (4) the main effect of task and the interactions of task with (5) form repetition or (6) motion repetition. To test for consistent effects across observers, the contrast images were entered into second level one-sample *t*-tests. Inferences were made at the second level to allow a random effects analysis and inferences at the population level (Friston et al., 1999).

Unless otherwise stated, we report activations at $p<0.05$ corrected for the entire brain at the cluster level using an auxiliary uncorrected voxel threshold of $p<0.001$.

Results

Behavioural data

For performance accuracy (Table 1), a three-way analysis of variance (ANOVA) with the factors, task (form vs. motion discrimination), form repetition (same vs. different) and motion repetition (same vs. different), showed a significant main effect of task ($F(1,18)=13.28$; $p<0.05$) and motion repetition ($F(1,18)=9.68$; $p<0.05$). In addition, there was a significant interaction effect between task and form repetition ($F(1,18)=14.29$; $p<0.05$).

An ANOVA of reaction times (RT) limited to correct trials revealed a significant main effect of task ($F(1,18)=10{,}0$; $p<0.05$) and form repetition ($F(1,18)=62.76$; $p<0.05$), but no significant main effect of motion repetition. Observers responded more quickly in the form task than the motion task, and they responded more quickly when the prime and target had the same form than when they had different forms. Importantly, there was a significant interaction between form repetition and motion repetition ($F(1,18)=58.42$; $p<0.05$). As evident in Fig. 4, when the prime and target objects had the same form, observers were faster at making either discrimination (form or motion) when both objects also rotated in the same relative to different direction.

Table 2
Main effect of form repetition

Region	MNI Coordinates	*Z*-score	*p* value (FWE-corr)
	x, y, z	Voxel level	Cluster level
L. lateral occipito-temporal region	−38 −78 −10	4.78 4.12 (F) 4.17 (M)	0.00
R. lateral occipito-temporal region	48 −76 −14	5.30 5.13 (F) 4.13 (M)	0.00

F = form task; M = motion task.

Table 3
Interaction between form and motion repetition

Region	MNI Coordinates	*Z*-score	*p* value (FWE-corr)
	x, y, z	Voxel level	Cluster level
L. fusiform g.	−36 −44 −22	5.13 3.29 (F) 3.16 (M)	0.00
L. post. middle temporal g.	−38 −60 8	5.05 3.32 (F) 2.43 (M)	0.03
R. fusiform g.	42 −56 −16	3.45	>0.05
R. post. middle temporal g.	54 −62 4	4.11	>0.05

F = form task; M = motion task

fMRI data

The two analyses without and with modeling reaction times as covariates of no interest (see Materials and methods) gave nearly identical results. We report the results of the first.

Main effect of form repetition suppression

The bilateral ventral occipito-temporal cortices showed decreased responses for same relative to different objects (Fig. 5, Table 2). To further evaluate whether this effect was due to response priming during the form task, the adaptation of form processing was tested separately for the form and motion task. During both tasks, we observed a significant effect of form repetition. The form priming effect persisted even when reaction times were modeled as covariates of no interest suggesting that it cannot be fully accounted for by differences in task difficulty between primed and unprimed conditions.

Main effect of motion repetition suppression

No brain regions were identified that showed decreased activation for same relative to different motion. Even at an uncorrected threshold of $p<0.05$, we did not observe any significant effect within our regions of interest, i.e. bilateral superior parietal cortex and bilateral posterior middle temporal gyrus as defined by AAL ROI library (MarsBaR Toolbox for SPM). In other words, we did not observe a motion repetition effect that generalizes over multiple objects. This is consistent with the ANOVA of reaction times in the behavioural data, which did not show a significant main effect of motion repetition. This null result in our experiment was not quite expected as neurophysiological studies in non-human primates demonstrate that areas MT, MST and parietal cortices responded to translation and rotation repetition (Priebe and Lisberger, 2002b; Sakata et al., 1986, 1994; Sugihara et al., 2002; Tolias et al., 2001; Van Wezel and Britten, 2002). Apart from obvious methodological differences (i.e. neurophysiology in monkeys vs. fMRI in humans), these inconsistencies across studies may be explained by differences in timing such as the duration of the first adapting stimulus or the prime-target interval. For instance a recent psychophysics study has dissociated three distinct periods of adaptation using sine-wave luminance gratings as stimuli (Kanai and Verstraten, 2005). Similarly, previous fMRI studies have investigated motion priming or aftereffects using block designs rather than randomized immediate repetition priming (Huk et al., 2001; Huk and Heeger, 2002).

Interaction between motion and form repetition

The left anterior fusiform and posterior middle temporal gyri (overlapping with hMT+/V5 based on anatomical landmarks (Dumoulin et al., 2000; Tootell et al., 1995; Watson et al., 1993)) exhibited a significant interaction between form and motion repetition. A significant interaction precludes the possibility that distinct populations within this region independently encode form and motion features. Instead, an interaction (rather than additive effects) suggests that both form and motion information are integrated within a region (Calvert and Lewis, 2004).

An interaction can be interpreted from two complementary perspectives. From one perspective, it can be viewed as a motion repetition effect that is selectively observed when prime and target objects have the same form. Alternatively, it suggests that the form priming effect in these two areas depended on motion information and was primarily observed when two successive objects rotated in the same direction. This latter perspective seems to be more relevant, as the fusiform gyrus play primarily a role in form rather than motion processing. Thus, we are inclined to favour the alternative that a predominantly form-selective region is modulated by the motion direction of an object.

The interaction effect was also observed when reaction times were modeled as covariates of no interest to account for unspecific differences in task difficulty across trials. At an uncorrected threshold, we also observed an interaction effect in the mirror regions of the right hemisphere (Fig. 5, Table 3).

Main effect of task and interactions between task and motion or form repetition

For completeness, we also tested for the main effect of task and the modulatory effect of task context on the form and motion repetition effects (i.e. the interaction between task and motion or form repetition). Comparing the motion and form task revealed a dorso-ventral dissociation: the fusiform gyrus bilaterally showed increased activation for the form relative to the motion task. By comparison, the bilateral inferior parietal lobes, the right intra-parietal sulcus and the left ventral pre-motor cortex exhibited increased activation for the motion relative to the form task (Table 4). Only a small region in the left central sulcus showed a significant interaction between task context and form priming. No significant interactions were observed between task context and motion priming.

Table 4
Main effect of task: form vs motion

Region	MNI Coordinates	*Z*-score	*p* value (FWE-corr)
	x, y, z	Voxel level	cluster level
Form>motion task			
R. fusiform g.	34 −58 −16	4.60	0.00
R. mid. occipital g.	42 −88 10	3.95	0.01
R. insula	46 2 −2	3.91	0.07
L. fusiform g.	−32 −60 −18	3.72	0.02
Motion>form task			
R. sup. occipital g.	42 −78 34	4.58	0.04
R. inferior parietal lobule (IPL)	54 −28 44	4.48	0.00
R. intra-parietal sulcus (IPS)	12 −70 52	4.05	0.00
L. inferior parietal lobule (IPL)	−42 −40 42	3.76	0.05
L. ventral pre-motor (PMv)	−20 −2 60	3.66	0.03

Discussion

This study used fMRI adaptation to identify the neural systems underlying the integration of form and motion information during dynamic object processing. At the behavioural level, observers were faster to recognize an object and discriminate its rotation direction when presented with that object as a prime. Importantly, this priming effect was enhanced when prime and target objects rotated in the same direction relative to when they rotated in different directions. At the neural level, our fMRI adaptation paradigm demonstrated that these two types of priming effects (i.e., the main effect of form repetition and the interaction between repetition of form and rotation direction) were associated with reduced activations in distinct brain regions in the ventral and dorsal streams of the human visual system. More specifically, bilateral lateral occipital regions exhibited a reduced BOLD signal when the form of an object was repeated irrespective of its rotation direction. In contrast, bilateral anterior fusiform and posterior middle temporal regions (close to or even overlapping with hMT+/V5) showed an adaptation effect that depended on both form and rotation direction. In this case, the adaptation for object form was modulated by the consistency of the rotation direction of the two objects. Specifically, adaptation to form was enhanced when prime and target objects rotated in the same direction. Thus, our experiment reveals a hierarchy of regions along ventral and dorsal pathways. In this processing hierarchy, lateral occipito-temporal cortex represents an object's atemporal, i.e., rigid 3D structure, whereas anterior fusiform and posterior middle temporal regions are involved in the integration of form and motion during dynamic object processing.

Furthermore, the adaptation effect in the lateral occipito-temporal regions for form is consistent with previous human fMRI studies which have implicated this region (also referred to as LOC, lateral occipital complex) in the analysis of 3D shape. Collectively, these studies demonstrated that LOC goes beyond pure image-based representations and is involved in processing an object's rigid 3D structure. Increased LOC activation has been reported for objects relative to scrambled images irrespective of the cues (e.g., colour, motion, etc.) that define the object (Grill-Spector et al., 2001; Kourtzi and Kanwisher, 2001; Malach et al., 1995). While these results can be attributed to attention to shapes (Murray et al., 2002), recent fMRI adaptation studies that have largely controlled for attention-related confounds have confirmed the role of LOC in shape processing. For instance, LOC has been shown to adapt to identical shapes across changes in cues, e.g., colour and motion (Self and Zeki, 2005); size and position (Grill-Spector et al., 1999; Kourtzi et al., 2003); viewpoint, e.g., rotation (Kourtzi et al., 2003); and other low-level image properties such as contrast (Avidan et al., 2002). In light of these previous results, it is not surprising that in our study the adaptation effect in the lateral occipito-temporal cortex was not significantly influenced by the rotation direction of the prime. Furthermore, the adaptation effect was observed during both motion and form tasks. This is important as a form priming effect observed during the form task alone could emerge at the level of response selection rather than object processing. However, in our experiment even when the task (i.e. discrimination of motion direction) and repetition (i.e. form) were orthogonal and hence a correct response could not be based on the prime stimulus, a form priming effect was observed indicating that it emerged at the level of object processing rather than response selection.

In contrast to the motion-invariant adaptation in the lateral occipito-temporal cortex, more anterior regions along both ventral and dorsal streams exhibited a form adaptation effect that was modulated by rotation direction. In particular, middle fusiform and posterior middle temporal (hMT+/V5) regions of both hemispheres showed an enhanced form adaptation effect when prime and target objects rotated in the same relative to different directions. These interactions between form and rotation direction can be explained by multiple neural mechanisms.

First, the direction-selective adaptation effect may be inherited from direction-selective V1 neurons with small receptive field sizes (Priebe and Lisberger, 2002). In this case, objects with different shapes would activate partially non-overlapping V1 populations resulting in a smaller direction-selective adaptation effect. This mechanism can be formally tested by presenting prime and target objects in a spatially segregated fashion, so that the two stimuli activate different populations of V1 neurons irrespective of whether the stimuli have identical or different shapes. However, the interaction effects are observed only in higher level visual association areas rendering a mechanism purely limited to V1 rather unlikely.

Second, the interaction of form and rotation direction found at the behavioural and neural levels may be due to a serial-position effect. For instance, if the first frame within the in-depth rotation sequence is more deeply processed than the remaining frames, a stronger “form” adaptation effect would be induced for prime and target sequences that rotated in the same direction and thus had the same initial frames (i.e., a primacy effect, (Henson et al., 1996)). This view-based mechanism is unlikely given that view-based fMRI adaptation effects have been associated consistently with more posterior lateral occipito-temporal regions rather than with hMT+/V5 and fusiform regions (Grill-Spector et al., 1999). In addition, behavioural studies which have randomized the order of frames to remove spatio-temporal continuity have shown that serial-position effects *per se* cannot account for the effect of rotation direction on object recognition (Vuong and Tarr, 2004).

Overall, the results of the present study suggest that high-level form and motion information are integrated in anterior ventral and dorsal (hMT+/V5) occipito-temporal regions. These results converge with and extend previous studies showing that more anterior ventral occipito-temporal regions are involved in integrating different visual cues (Kristjánsson et al., 2007; Self and Zeki, 2005) or information from multiple senses (Amedi et al., 2001). For instance, it has been shown that colour and motion cues that define simple shapes are integrated in an anterior LOC region (Self and Zeki, 2005). In our study, the objects' 3D form and its 3D rotation trajectory were defined by multiple low-level cues such as contour, shading patterns and motion parallax, and their dynamic changes in time. Furthermore, we varied the viewpoint and the axis of rotation on a trial by trial basis making it difficult for the visual system to perform the task based on low-level features such as a specific shading pattern or the retinal velocity field. Instead, the experimental paradigm and task required observers to integrate multiple cues into higher level form and motion representations. Therefore, it is not surprising, yet noteworthy, that our experiment revealed integration sites in a distributed system that cuts across the dorsal and ventral visual processing stream (Sereno et al., 2002; Ungerleider and Mishkin, 1982): An interaction between form and rotation adaptation was observed in the middle fusiform along the ventral stream as well as anterior hMT+/V5 along the dorsal stream. All of these regions have previously been implicated in 3D-shape and structure-from-motion processing (Kriegeskorte et al., 2003). However, the precise functional contributions of dorsal and ventral areas during dynamic object recognition remain to be determined. For instance, one may speculate whether motion parallax and contour-related changes due to in-depth rotation may be related to the interaction effects in the dorsal and ventral regions respectively. Future studies that compare in-depth rotation with and without contour changes may enable us to further disentangle their functional contributions.

To conclude, the present experiment demonstrates that bilateral lateral occipito-temporal regions represent 3D structure of objects irrespective of their dynamics. In contrast, anterior hMT+/V5 and middle fusiform regions show a form adaptation effect that is enhanced when the prime's rotation direction matched the target's rotation direction. Thus, our data reveal a visual processing hierarchy with more anterior regions integrating 3D structure and rotation direction as complementary sources of information during dynamic object processing.

Acknowledgments

This study was supported by the DFG and the Max Planck Society. We thank Andreas Bartels for very helpful discussions and Johannes Tünnerhoff for help with preparing the figures.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2008.04.265.

References

Albright, T.D., 1984. Direction and orientation selectivity of neurons in visual area MT of the macaque. J. Neurophysiol. 52, 1106–1130.

Amedi, A., Malach, R., Hendler, T., Peled, S., Zohary, E., 2001. Visuo-haptic object-related activation in the ventral visual pathway. Nat. Neurosci. 4, 324–330.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., Malach, R., 2002. Contrast sensitivity in human visual areas and its relationship to object recognition. J. Neurophysiol. 87, 3102–3116.

Calvert, G.A., Lewis, J.W., 2004. Hemodynamic studies of audio-visual interactions. The Handbook of Multi-sensory Processes. MIT press, Cambridge, pp. 483–502.

Deyoe, E.A., Vanessen, D.C., 1985. Segregation of efferent connections and receptive-field properties in visual area V2 of the macaque. Nature 317, 58–61.

Dumoulin, S.O., Bittar, R.G., Kabani, N.J., Baker, C.L., Le Goualher, G., Bruce Pike, G., Evans, A.C., 2000. A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. Cereb. Cortex 10 (5), 454–463.

Evans, A.C., Collins, D.L., Milner, B., 1992. An MRI-based stereotactic atlas from 250 young normal subjects. Soc. Neurosci. Abstr. 18, 408.

Friston, K.J., Holmes, A., Worsley, K.J., Poline, J.B., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 189–210.

Friston, K.J., Holmes, A.P., Price, C.J., Buchel, C., Worsley, K.J., 1999. Multisubject fMRI studies and conjunction analyses. Neuroimage 10, 385–396.

Geng, J.J., Egar, E., Ruff, C.C., Kristjansson, A., Rothstein, P., Driver, J., 2006. On-line attentional selection from competing stimuli in opposite visual fields: effects on human visual cortex and control processes. J. Neurophysiol. 96, 2601–2612.

Gibson, J.J., 1979. The Ecological Approach to Visual Perception. Houghton Mifflin Company, Boston, MA.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., Singer, W., 1998. The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. Eur. J. Neurosci. 10, 1563–1573.

Grill-Spector, K., Edelman, S., Kushnir, T., Itzchak, Y., Malach, R., 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. Investig. Ophthalmol. Vis. Sci. 40, S399.

Grill-Spector, K., Kourtzi, Z., Kanwisher, N., 2001. The lateral occipital complex and its role in object recognition. Vis. Res. 41, 1409–1422.

Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23.

Henson, R.N.A., 2003. Neuroimaging studies of priming. Progr. Neurobiol. 70, 53–81.

Henson, R.N.A., Rugg, M.D., 2003. Neural response suppression, haemodynamic repetition effects, and behavioural priming. Neuropsychologia 41, 263–270.

Henson, R.N.A., Norris, D.G., Page, M.P.A., Baddeley, A.D., 1996. Unchained memory: error patterns rule out chaining models of immediate serial recall. Quart. J. Exper. Psychol. Sect. A-Hum. Exper. Psychol. 49, 80–115.

Huk, A.C., Heeger, D.J., 2002. Pattern-motion responses in human visual cortex. Nat. Neurosci. 5 (1), 72–75.

Huk, A.C., Ress, D., Heeger, D.J., 2001. Neuronal basis of the motion aftereffect reconsidered. Neuron 32, 161–172.

Kanai, R., Verstraten, F.A.J., 2005. Perceptual manifestations of fast neural plasticity: motion priming, rapid motion aftereffect and perceptual sensitization. Vis. Res. 45, 3109–3116.

Kelly, M.H., Freyd, J.J., 1987. Explorations of representational momentum. Cognit. Psychol. 19, 369–401.

Kourtzi, Z., Kanwisher, N., 2001. Representation of perceived object shape by the human lateral occipital complex. Science 293, 1506–1509.

Kourtzi, Z., Bulthoff, H.H., Erb, M., Grodd, W., 2002. Object-selective responses in the human motion area MT/MST. Nat. Neurosci. 5, 17–18.

Kourtzi, Z., Erb, M., Grodd, W., Bulthoff, H.H., 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. Cereb. Cortex 13, 911–920.

Kriegeskorte, N., Sorger, B., Naumer, M., Schwarzbach, J., van den, B.E., Hussy, W., Goebel, R., 2003. Human cortical object recognition from a visual motion flowfield. J. Neurosci. 23, 1451–1463.

Kristjánsson, A., Vuilleumier, P., Schwartz, S., Macaluso, E., Driver, J., 2007. Neural basis for priming of pop-out revealed with fMRI. Cereb. Cortex 17, 1612–1624.

Liu, T.S., Cooper, L.A., 2003. Explicit and implicit memory for rotating objects. J. Exper. Psychol.-Learn. Mem. Cognit. 29, 554–562.

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B.H., 1995. Object-related activity revealed by functional magnetic-resonance-imaging in human occipital cortex. Proceedings of the National Academy of Sciences of the United States of America, vol. 92, pp. 8135–8139.

Marr, D., Vaina, L., 1982. Representation and recognition of the movements of shapes. Proceedings of the Royal Society of London Series B—Biological Sciences, vol. 214, pp. 501–524.

Mitsumatsu, H., Yokosawa, K., 2003. Efficient extrapolation of the view with a dynamic and predictive stimulus. Perception 32, 969–983.

Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P., Woods, D.L., 2002. Shape perception reduces activity in human primary visual cortex. Proceedings of the National Academy of Sciences of the United States of America, vol. 99, pp. 15164–15169.

Murray, S.O., Olshausen, B.A., Woods, D.L., 2003. Processing shape, motion and three-dimensional shape-from-motion in the human cortex. Cereb. Cortex 13, 508–516.

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

Paradis, A.L., Cornilleau-Peres, V., Droulez, J., Van de Moortele, P.F., Lobel, E., Berthoz, A., Le Bihan, D., Poline, J.B., 2000. Visual perception of motion and 3-D structure from motion: an fMRI study. Cereb. Cortex 10, 772–783.

Pike, G.E., Kemp, R.I., Towell, N.A., Phillips, K.C., 1997. Recognizing moving faces: the relative contribution of motion and perspective view information. Vis. Cogn. 4, 409–437.

Priebe, N.J., Lisberger, S.G., 2002. Constraints on the source of short-term motion adaptation in macaque area MT. II. Tuning of neural circuit mechanisms. J. Neurophysiol. 88, 370–382.

Sakata, H., Shibutani, H., Ito, Y., Tsurugai, K., 1986. Parietal cortical-neurons responding to rotary movement of visual stimulus in space. Exp. Brain Res. 61, 658–663.

Sakata, H., Shibutani, H., Ito, Y., Tsurugai, K., Mine, S., Kusunoki, M., 1994. Functional-properties of rotation-sensitive neurons in the posterior parietal association cortex of the monkey. Exp. Brain Res. 101, 183–202.

Self, M.W., Zeki, S., 2005. The integration of colour and motion by the human visual brain. Cereb. Cortex 15, 1270–1279.

Sereno, M.E., Trinath, T., Augath, M., Logothetis, N.K., 2002. Three-dimensional shape representation in monkey cortex. Neuron 33, 635–652.

Stone, J.V., 1998. Object recognition using spatiotemporal signatures. Vis. Res. 38, 947–951.

Stone, J.V., 1999. Object recognition: view-specificity and motion-specificity. Vis. Res. 39, 4032–4044.

Sugihara, H., Murakami, I., Shenoy, K.V., Andersen, R.A., Komatsu, H., 2002. Response of MSTd neurons to simulated 3D orientation of rotating planes. J. Neurophysiol. 87, 273–285.

Sunaert, S., Van Hecke, P., Marchal, G., Orban, G.A., 1999. Motion-responsive regions of the human brain. Exp. Brain Res. 127, 355–370.

Tolias, A.S., Smirnakis, S.M., Augath, M.A., Trinath, T., Logothetis, N.K., 2001. Motion processing in the macaque: revisited with functional magnetic resonance imaging. J. Neurosci. 21, 8594–8601.

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

Treue, S., Husain, M., Andersen, R.A., 1991. Human perception of structure from motion. Vis. Res. 31, 59–75.

Ullman, S., 1979. The Interpretation of Visual Motion. MIT press, Cambridge, MA.

Ungerleider, L., Mishkin, M., 1982. Two cortical visual systems. The Analysis of Visual Behaviour. MIT Press, Cambridge, MA, pp. 549–586.

Van Wezel, R.J.A., Britten, K.H., 2002. Motion adaptation in area MT. J. Neurophysiol. 88, 3469–3476.

VanOostende, S., Sunaert, S., VanHecke, P., Marchal, G., Orban, G.A., 1997. The kinetic occipital (KO) region in man: an fMRI study. Cereb. Cortex 7, 690–701.

Vuong, Q.C., Tarr, M.J., 2004. Rotation direction affects object recognition. Vis. Res. 44, 1717–1730.

Vuong, Q.C., Tarr, M.J., 2006. Structural similarity and spatiotemporal noise effects on learning dynamic novel objects. Perception 35, 497–510.

Wallis, G., Bulthoff, H., 1999. Learning to recognize objects. Trends Cogn. Sci. 3, 22–31.

Wallis, G., Bulthoff, H.H., 2001. Effects of temporal association on recognition memory. Proceedings of the National Academy of Sciences of the United States of America, vol. 98, pp. 4800–4804.

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

Zeki, S.M., 1978. Uniformity and diversity of structure and function in rhesus-monkey prestriate visual-cortex. J. Physiol.-London 277, 273–290.

Zeki, S., Watson, J.D.G., Lueck, C.J., Friston, K.J., Kennard, C., Frackowiak, R.S.J., 1991. A direct demonstration of functional specialization in human visual-cortex. J. Neurosci. 11, 641–649.