7. Motion Perception and Motion Imagery: New Evidence of Constructive Brain Processes from Functional Magnetic Resonance Imaging Studies

Rainer Goebel, Lars Muckli and Wolf Singer

7.1 Introduction

Since the pioneering work of the gestalt psychologists, several rules of perceptual organization have been formulated which characterize visual perception as a constructive process of the brain. This view emphasizes that a visual stimulus is not passively mapped onto the brain but is instead actively processed based on innate knowledge and acquired experiences with the visual world. The constructive nature of visual processing can be convincingly demonstrated with perceptual illusions, e.g. the perception of illusory contours. There is accumulating evidence that at the neuronal level the operation of gestalt principles is reflected in synchronous oscillatory discharges of organized cell assemblies (1). Here we report several functional Magnetic Resonance Imaging (fMRI) studies which attempt to localize the underlying neuronal substrate of several constructive brain processes at the system level. We have studied the perceptual illusions of apparent motion, apparent motion of objects defined by illusory contours and the phenomenon of shape-from-motion. Further insights in the constructive nature of vision were obtained by comparing brain activation during imagery of moving objects with activation during imagery of static objects. The influence of slight alterations of the configuration of stimuli on visual perception and the fMRI signal was investigated in detail by means of a transparent motion paradigm.

7.1.1 Two Main Visual Processing Pathways

It has been widely accepted that the mammalian visual system is composed of several parallel channels, each subserving a different aspect of visual function (2). Ungerleider and Mishkin (3) proposed that visual processing in the primate cortex is divided into two main pathways, a ventral stream devoted to the fine analysis of the visual scene including the perception of form and color, and a dorsal stream that codes the spatial characteristics of the visual scene and analyzes motion. The ventral stream which uses mainly foveal signals, processes information relatively slowly but with high spatial resolution. Activity flows from the primary visual cortex (V1) via the inferotemporal cortex (IT) to perirhinal and prefrontal areas. The dorsal stream, which is concerned with spatial localization of objects and the programming of visually controlled motor acts, operates rapidly but with coarse
spatial resolution (4). In the dorsal stream, activity flows from primary visual cortex via several areas in the parietal and superior temporal lobe to premotor and prefrontal areas. In prefrontal cortex, the two streams project to different areas; the dorsal stream in the dorsolateral area, devoted to short-term storage of spatial locations, and the ventral stream in a ventral area specialized for short-term memory of object features.

Most experiments that led to the differentiation of the dorsal and ventral pathways were performed on rhesus monkeys (*Macaca mulatta*). Most neuroscientists agree that the visual system of this monkey approximates very closely the situation in the human brain. With modern brain imaging techniques, the two pathways could also be separated in the normal human brain (5–7).

### 7.1.2 Functional Magnetic Resonance Imaging

Using functional Magnetic Resonance Imaging (fMRI), investigators were able to localize neural activity in human brains during sensory, motor, and cognitive activity. The technique is based on the fact that there are local blood flow and blood oxygenation changes in response to neural activity which are detectable with the technology of Magnetic Resonance Imaging (MRI) by choosing appropriate imaging parameters. The fMRI signal reflects changes in local oxygenation level since oxygenated hemoglobin has a much smaller magnetic susceptibility than deoxygenated hemoglobin (BOLD Effect = Blood Oxygenation Level Dependent Effect). fMRI is rapidly becoming the technology of choice for many functional brain activation studies in humans because it is acceptable for repeated use and it has better spatial and temporal resolution than Positron Emission Tomography (PET) imaging which is based on the same hemodynamic phenomena. The ultimate limitation of spatial resolution of fMRI may be the spatial specificity of the circulatory system and its local changes in response to neural activity. Anatomical and physiological evidence suggests that this circulatory specificity will be less than 1 mm for cerebral cortex. The temporal resolution of fMRI is limited by the vagaries of the hemodynamic responses of the brain’s circulatory system as it changes with neural activation. While some of those changes occur very quickly (on the order of 10–100 ms), most hemodynamic changes that are detectable with current fMRI systems appear only after a delay of about 2000 ms and take another few seconds to reach peak. This means that fMRI is unlikely to replace tools such as electroencephalography (EEG) and magnetoencephalography (MEG). These methods measure the electrical and magnetic neuronal signals, have millisecond temporal resolution but suffer from poor spatial resolution. Current research attempts to combine fMRI/PET and EEG/MEG in order to provide the essential data for the analysis of neuronal processes, their topological distribution and their precise timing.

At the beginning of an fMRI experiment, a fast scan of the subject’s brain is recorded consisting of three orthogonal slices. These are used to select the orientation and position of the slices that will be used in the functional imaging as well as for an anatomical reference scan. The 2D or 3D anatomical reference scan
consists normally of images with higher spatial resolution (e.g., 1.0 mm x 1.0 mm x 1.0 mm) than the functional images and is used to relate the functional findings to known anatomical landmarks in the brain. During functional imaging, the set of specified slices (comprising a volume) is measured at regular intervals while the subject performs different experimental tasks. The changes of neural activity during different conditions are reflected in small changes in the functional images which can be visualized using appropriate data analysis techniques.

7.1.3 Methodological Details of the fMRI Measurements

Fifteen subjects without a previous neurological or psychiatric history participated in one or more experiments. Age ranged from 25 to 33 years and there were six males. Informed consent was obtained from each subject. Visual stimuli were delivered under computer control (Digital DECpc Celebris XL 590) to an LCD display panel (Sharp QA-1000) and a high-luminance overhead projector (Elmo HP-283P) or to a high-luminance LCD projector (EIKI LC-6000). Subjects were in supine position and viewed the screen through an adjustable mirror fixed to the head coil. The image was back-projected onto a frosted screen positioned at the foot end of the scanner. Visual stimuli were generated in real time using the ELSA Winner 2000 Pro/X graphics adapter and the ELSA Powerlib C library.

Each experiment consisted of 128 measurements of 10–15 oblique transversal slices. The slices were positioned either parallel to the calcarine fissure or parallel to the AC-PC line. In each experimental condition, a sequence of eight measurements was recorded lasting 24s. In all conditions a small fixation cross appeared in the center of the screen. This was the only stimulus during fixation and imagery periods (see below). In some studies an objective motion stimulus was used consisting of 400 white dots moving radially outward on a black background (visual field: 38° wide by 23° high, dot size: 0.06° x 0.06°, dot velocity: 3.6°/s – 14.4°/s). This stimulus is known to produce a clear response in the motion-sensitive areas of the dorsal stream (areas MT and mST) without provoking eye movements (8). Additionally, a static stimulus which consisted of 400 stationary dots was used in the first experiments. Eye movements were recorded with a video camera outside the scanner in order to verify whether subjects were able to fixate the cross during apparent motion and motion imagery conditions.

Echo-planar images were collected on a 1.5-T scanner (Siemens Magnetom Vision) using the standard head coil and a gradient echo echoplanar sequence (TR = 3000 ms, TE = 66 ms, flip angle = 90°, FOV = 210 mm x 210 mm, slice thickness = 3 mm, imaging matrix = 64 x 64, voxel size = 3.2 mm x 3.2 mm x 3 mm). In most experiments the Siemens Magnetom gradient overdrive was used allowing functional EPI scans with high spatial resolution (TR = 3000 ms, TE = 69 ms, flip angle = 90°, FOV = 210 mm x 210 mm, slice thickness = 3 mm, imaging matrix = 128 x 128, voxel size = 1.6 mm x 1.6 mm x 3 mm). Before each set of functional scans, we recorded a T1-weighted series of 2D images with the same orientation, slice thickness, and field of view as the functional scans or a fast T1-weighted 3D volume (MPRAGE sequence).
Data analysis was performed using custom software (9; 10). Prior to statistical analysis the time series of functional images was aligned for each slice in order to minimize the effects of head movement. For each slice the third recorded functional image was used as a reference image to which all other images of the slice time series were registered. In order to evaluate statistically the differences between experimental conditions cross-correlation analysis was applied. For the computation of correlation maps, the stimulation protocol served as a reference function reflecting the temporal sequence of experimental and control conditions (experimental condition = 1, control condition = 0). On a pixel-by-pixel basis the signal time course was cross-correlated with the respective reference function (11). Pixels were included into the statistical map if the obtained correlation value was greater than 0.3 given lag values of 1 and 2 (corresponding to a 3–9 secs delay after the beginning of a stimulation condition in order to adapt to the hemodynamic response). Cross-correlation maps were superimposed both on the original functional scans as well as onto the T1-weighted anatomical reference scans. Striate and extrastriate cortical areas V1, V2d, V3, V3a were discriminated based on anatomical location and functional properties (12). The delineation of these areas was validated for each subject using results from separate recordings for generating retinotopic maps (compare 13–14). In order to delineate MT/MST as well as other motion-sensitive areas we defined as regions of interest (ROI) those regions that could be activated with the objective motion stimulus consisting of radially moving dots. In order to assess whether an area is motion-selective we compared the objective motion condition with the static dot stimulus. In experiments 3 and 4 additional ROIs were defined by comparing all imagery conditions with the fixation conditions.

The data for statistical comparisons consisted of the mean time course of all voxels of an analyzed area. Based on these data the mean of the raw fMRI signal for each subject and condition in a given experiment was computed. These mean values were analyzed using ANOVA and post-hoc pairwise comparisons using stimulus condition as a within-group factor. The obtained p-values were corrected for multiple comparisons. Values of percent signal change averaged across subjects were computed on the basis of the difference between the mean values of the fMRI signal in each experimental condition and the mean fMRI signal in the fixation periods for each individual subject.

For 3D visualization and measurement of Talairach coordinates, high resolution T1-weighted 3D data sets (voxel size: 1.0 mm x 1.0 mm x 1.0 mm) were recorded in separate sessions. Statistical maps were transformed into 3D data sets and interpolated to the same resolution as the structural 3D data set. For each subject the structural and functional 3D data sets were transformed into Talairach space (16) which allowed us to compare activated brain regions across different experiments and across different subjects and to determine Talairach coordinates of these regions. Talairach transformation was performed in two steps. The first step consisted of rotating the 3D data set of each subject to be aligned with the stereotaxic axes. For this step the location of the anterior commissure (AC) and the posterior commissure (PC) as well as two rotation parameters for midsagittal alignment had to be specified manually. In the second step the extreme points of the cerebrum
were specified. These points together with the AC and PC coordinates were then used to scale the 3D data sets into the dimensions of the standard brain of the Talairach and Tournaux atlas using a piecewise affine and continuous transformation for each of the 12 defined subvolumes.

7.2 FMRI Experiments

Perceptual illusions and visual imagery are important paradigms for the experimental study of constructive aspects of vision, viz the generation of visual percepts that goes beyond the information contained in the mere physical composition of the stimuli (17). Certain types of illusions arise when the physical properties of a stimulus are supplemented by perceptual features that rely on the assumptions of the brain about what it expects in the outside world. For example, when stimuli separated in space are turned on and off in alternation at appropriate temporal intervals (18), subjects perceive one stimulus moving between the two stimulus positions (apparent motion) rather than two stationary flickering stimuli. The perception of motion of objects that in fact do not change position is an example of a constructive visual process. Visual imagery, on the other hand, may be performed in the complete absence of visual stimuli. Thus, when a subject is requested to imagine a previously seen visual scene, the task usually consists of the construction of a visual image purely from stored information. In order to study constructive aspects of motion perception we gradually reduced the amount of objective motion perception and increased the amount of internally generated motion representations in a series of experiments using objectively moving stimuli, apparent motion stimuli, stimuli inducing apparent motion of figures defined by illusory contours and motion imagery conditions (19, 20).

7.2.1 Apparent Motion

Previous brain imaging studies using either positron emission tomography (PET) or functional magnetic resonance imaging (fMRI) have shown that human cortical areas MT (V5) and mST respond with increased activity to moving stimuli (8; 21; 22) and in certain conditions where stationary stimuli induce illusory motion (17; 23). It remains to be seen to what extent apparent motion stimuli also activate the motion-selective areas MT and mST. Since it is known that these areas respond to some degree to flickering stimuli (8), appropriate control conditions had to be specified which differ from the apparent motion stimuli only in the relative timing of appearing objects. Additionally, form-motion interactions were investigated in a second experiment by inducing apparent motion of squares whose outlines were defined by illusory contours. This paradigm integrates two perceptual illusions and should provide additional insight into the interaction of brain regions responsible for extracting different stimulus properties.
In the first experiment, perception of apparent motion was induced with two concentric rings (diameters: 3° and 10°) that appeared in alternation, separated by a short blank interval of 32 ms (Fig. 7.1A). We opted for this stimulus, which was perceived as a single shrinking and expanding ring by all subjects (n = 10), because it does not elicit eye movements. As a control stimulus devoid of apparent motion cues both concentric rings were turned on and off simultaneously. Thus, this flickering control condition differed from the apparent motion condition only in the relative timing between the large and the small ring. Two flickering control conditions were used (flickering control I and II, see Fig. 7.1A).

Area V1 responded with similar activation levels to all stimulation conditions, with the lowest activation in the two ring conditions. In area V3, the objective motion condition, the apparent motion condition and the flickering control evoked similarly large responses, but in contrast to V1 the stationary stimulus was significantly less effective (p < 0.01, Fig. 7.1B). Area MT/MST responded strongly to both the apparent and the objective motion stimuli but much less to both flickering control conditions (p < 0.01, Fig. 7.1B). The flickering control I condition produced a stronger MT/MST response than the flickering control II condition and was used for statistical comparisons. The results indicate that the MT/
MST complex is the first area in the dorsal pathway that responds selectively to apparent motion. A response pattern very similar to that of MT/MST was found in a region within the intraparietal sulcus.

In a second experiment we used squares whose outlines consisted of subjective contours (24) to generate the perception of apparent motion. These illusory squares were presented in alternation on the left and right side of the fixation spot (Fig. 7.2A). In order to generate the percept of an appearing and disappearing square, the four "pacmen" defining a square were opened and closed simultaneously so that observers (n=7) saw either a Kanizsa square or four filled circles (diameter: 0.9°). Kanizsa squares were centered at an eccentricity of 4.5° and were 2.2° in length. In the control condition without motion cues, the illusory Kanizsa squares appeared and disappeared simultaneously on both sides of the fixation cross whereas in the apparent motion condition they appeared and disappeared in alternation (Fig. 7.2A). Despite this change in relative timing across hemifields, stimulation conditions within each hemifield remained the same in both conditions. In the control condition subjects reported seeing two squares appearing and disappearing in alternation while in the apparent motion condition they reported seeing a single square moving between the two stimulus locations. In two additional no-contour control conditions the pacmen were opened outward and thus did not produce illusory squares (Fig. 7.2A).

Figure 7.2B shows that area V1 responded with similar activation levels to all subjective-contour and no-contour conditions. Area V2 but not V1 and V3 responded significantly more strongly to the two conditions with illusory contours than to the respective control conditions with the outward turned pacmen (p <0.01). The stimulus generating apparent motion of figures defined by subjective contours activated MT/MST more than the flickering squares (p <0.01) and these in turn were more effective than the outward turned pacmen that did not define Kanizsa squares (p <0.01).

Our results suggest that area MT/MST is the first cortical area along the dorsal stream that is capable of extracting apparent motion cues: In contrast to V1, V2 and V3, its responses differed significantly between the apparent motion conditions and the respective flickering controls. Comparison of responses to illusory-contour and no-contour conditions suggests that area MT/MST is also capable of using illusory contour information for computing and representing apparent motion. Area V2 responded more strongly to the two illusory-contour conditions than to the two no-contour conditions suggesting that the illusory contour information is extracted at that level (25-27) and made available as input to area MT/MST.

In the second apparent motion experiment, the difference between the apparent motion and the flickering control condition was smaller than in the first two-ring experiment, although the amount of stationary flicker was reduced. One possible explanation is that the opening and closing of the pacmen evoked a motion experience within the same hemifield. Six subjects did in fact report that they had the impression that in the subjective-contour flickering control condition the two Kanizsa squares were shrinking and expanding at their respective locations.
Fig. 7.2: Main experimental conditions and results of the second apparent motion experiment. The sequence of stimulation conditions consisted of: Fixation, Objective motion, Fixation, Static, Fixation, Subjective-contour apparent motion (SCAM), Fixation, Subjective-contour flickering control (SCFC), Fixation, No-contour apparent motion (NCAM), Fixation, No-contour flickering control (NCFC). (A) In the subjective-contour conditions four circles or four "pacmen" figures appeared on the left and the right side of the fixation cross. By showing the four "pacmen" figures or the four filled circles observers saw either a Kanizsa square or four disks at the respective location. In the subjective-contour apparent motion condition the Kanizsa squares appeared in alternation on each side. In the subjective-contour flickering control the Kanizsa squares appeared and disappeared simultaneously. In the No-contour apparent motion and No-contour flickering control conditions, the "pacmen" figures were turned outward. (B) Statistical comparison of activation levels in V1, V2, V3 and MT/MST (bottom) averaged over 7 subjects. Mean percent signal change ± s.e.m. is displayed for each condition.
7.2.2 Motion Imagery

We further investigated whether the areas activated by apparent motion stimuli are also activated when moving stimuli are solely imagined. Previous brain imaging studies have shown that mental imagery activates the same brain regions engaged in visual perception (28). It is, however, unclear to what extent small circumscribed extrastriate visual areas responsible for the analysis of specific stimulus properties can be selectively activated by imagery. We examined this issue for the motion-selective areas MT and msT as well as other areas by comparing their response patterns during imagery of moving stimuli with imagery of static stimuli.

After subjects (n = 5) had seen one of three motion stimuli for 24 sec (radially moving dots, apparent motion of subjective contours or rotating grating), they were required to fixate a cross for 24 seconds. This long interval was introduced in order for the hemodynamic response to return to baseline level before the start of the following imagery condition. After this fixation period, the fixation cross was turned off for 200 ms, which signaled to the subject that they should now try to imagine the previous motion stimulus as intensely as possible while fixating the cross. After 24 seconds, the cross was turned off again for 200 ms. This signaled to the subject the end of the imagery condition. Motion stimuli were used from previous experiments except the rotating grating (diameter: 16.7°, circular spatial frequency: 15/360°) which rotated 170/sec. The experiment was repeated twice. Activity was measured in twelve slices approximately parallel to the calcarine fissure and in 15 axial slices with a transversal-to-coronal angle of 5–10 degrees with respect to the AC-PC line.

All subjects reported that they were able to evoke mentally a clear motion experience during the imagery periods. At the same time they judged their mental image to contain less figural detail than the percept which was generated when they were seeing moving contours. As expected, the same areas as in the previous experiments responded to the initial presentation of the objectively and apparently moving stimuli (Fig. 7.3B). However, most of these areas also showed a surprisingly high activation during all three conditions of motion imagery. Interestingly, the imagery specific activation (ISA) defined as the ratio (motion imagery – fixation) / (objective motion – fixation) increased with the synaptic distance of an area from V1 along the dorsal processing stream (Fig. 7.3B, Fig. 7.4B). Area MT/MST was activated approximately half as much in the imagery conditions as in the motion conditions (ISA = 0.6). A close inspection of the time course of signal changes at individual voxels revealed that the ISA was different within subparts of the MT/MST complex. During motion imagery the upper anterior part of MT/MST (presumed msT) was activated almost as strongly as during presentation of objective motion stimuli (ISA = 0.8) while the lower posterior part (presumed MT) showed less activation (ISA = 0.4). A weak but significant activation with motion imagery was also found in areas V2 and V3a (p < 0.01), but no significant modulation was detectable in primary visual cortex. In addition to MT/MST several cortical areas in the inferior and superior parietal cortex were strongly and bilaterally activated during motion imagery (Fig. 7.3B, Fig. 7.4). Additional regions activated during motion imagery include the dorsolateral prefrontal cortex.
Fig. 7.3: Results of the first (left) and second (right) motion imagery experiment. The sequence of stimulation conditions in the first experiment consisted of: Fixation, Static, Fixation, Motion I (radially moving dots), Fixation, Motion imagery I, Fixation, Motion II (subjective-contour apparent motion stimulus of experiment II), Fixation, Motion imagery II, Fixation, Motion III (rotating grating), Fixation, Motion imagery III, Fixation. The sequence of stimulation conditions in the second experiment consisted of: Fixation, Static I (stationary dots), Fixation, Static imagery (SI) I, Fixation, Motion I (radially moving dots), Fixation, Motion imagery (MI) I, Fixation, Static II (stationary frame of subjective-contour apparent motion stimulus of experiment 2), Fixation, Static imagery II, Fixation, Motion II (subjective-contour apparent motion of the second apparent motion experiment), Fixation, Motion imagery III, Fixation. (A, C) Signal time course in cortical area MT/MST averaged over five subjects. The sequence of stimulation conditions (triplets of motion-fixation-imagery) was randomized across subjects but is shown in a fixed sequence. (B, D) Statistical comparison of activation levels in areas V1, V2, V3a. MT/MST, area in inferior parietal lobule (IPL), area in superior parietal lobule (SPL), frontal eye field (FEF) and BA 9/46 averaged over five subjects. The three motion conditions and the three motion imagery conditions of the first experiment are pooled together as well as the two static, static imagery, motion and motion imagery conditions in the second experiment. Mean percent signal change ± s.e.m. is displayed for each condition. Activation levels associated with motion...
including Brodmann area (BA) 9/46, the frontal eye field (FEF), the anterior cingulate gyrus and the insular gyrus. Activation of BA 9/46 was found bilaterally but was greater in the right hemisphere (p<0.01). Most of these higher areas were more strongly activated during motion imagery than during the objective and apparent motion conditions.

In a follow-up experiment imagery of moving stimuli was compared with imagery of static stimuli. The experimental procedure was the same as in the previous experiment, except that either a static or a moving stimulus was shown prior to the respective imagery period. One static stimulus consisted of the standard static control condition (400 randomly positioned dots). A stationary frame of the subjective-contour apparent motion stimulus of experiment 2 was used as the second static stimulus. The dynamic versions of these stimuli (radially outward moving dots, subjective-contour apparent motion stimulus) were used in the two motion conditions. Activity was measured in 15 slices approximately parallel to the calcarine fissure and in 15 axial slices with a transversal-to-coronal angle of 1–5 degrees with respect to the AC-PC line.

The results obtained during motion imagery were similar to those in the preceding experiment. A weak but non-significant activation with motion imagery was found in area V1 (Fig. 7.3D). All other analyzed areas were significantly activated during motion imagery (p<0.01). Although several areas of the dorsal pathway were also activated during static imagery, the activation was much weaker than during motion imagery in most areas. Area MT/MST (Fig. 7.4) showed no significant enhancement during static and static imagery conditions but a strong enhancement (p<0.01) during motion and motion imagery conditions (ISA = 0.6). Additional regions activated during both static and motion imagery (not shown in Fig. 7.4) included several small areas in the intermediate and inferior frontal gyri, regions in the insular gyrus as well as the anterior cingulate gyrus. The latter responded with roughly equal strength to both types of imagery but showed no response to all other stimulus conditions.

Our finding that V1 is not significantly active during imagery conditions is in agreement with several PET mental imagery studies (29) but is in conflict with a recent PET study by Kosslyn and colleagues (28) where significant activation of V1 was found when compared to a non-imagery listening control condition. In Kosslyn’s study, however, significant V1 activation was not detectable when compared to the resting baseline condition. Our results suggest that imagery related activation is much less pronounced in V1 than in the extrastriate areas V2 and V3a which were identified as significantly activated in our imagery experiments when matched against the visual fixation control condition.

The motion-sensitive areas MT and MST responded strongly when subjects merely imagined moving stimuli. Since the MT/MST complex was selectively activated during motion imagery but not during static imagery, it is unlikely that the
activated network of areas during imagery periods is merely the consequence of general activation related to vigilance.

Subjects reported that they were able to evoke mentally a clear motion experience during the imagery periods. At the same time they judged their mental images to contain less figural detail when compared to seeing moving contours. We suggest that these subjective reports reflect the lack of detectable V1 activity despite the strong activation of MT/MST during motion imagery and that the activation of area MT/MST during motion imagery mediates an experience of motion that is detached from figural details.

In addition to MT and MST areas at higher levels of the dorsal pathway (30) were strongly activated during motion imagery and during objective and apparent motion conditions. The activated regions in the inferior and superior parietal lobule could be the homologues of monkey VIP, LIP and area 7a. Like the activated region in the inferior parietal lobule (IPL, Fig. 7.3B, D; Fig. 7.4B), monkey area VIP responds to moving stimuli in a similar way as area MT (31). The activation of more dorsal parietal regions might be due to covert spatial attention shifts and spatial representations (see below), which would give them a similar function as...
monkey LIP (32) and 7a (33). Whilst dorsal pathway activity during motion imagery was most pronounced in the areas with the highest synaptic distance to V1, it showed the opposite behavior during objective and apparent motion conditions. Furthermore, FEF and BA 9/46, regions which are known to be involved in spatial attention and visual short-term memory (e.g., 6) were almost exclusively activated during motion imagery. This pattern of activation suggests that attentional mechanisms involved in the generation and maintenance of the mental image generate a top-down information flow toward early visual areas (cf. 34).

The strong activation of the higher areas of the dorsal pathway as well as the prefrontal areas FEF and BA 9/46 during imagined motion, might reflect processes of spatial attention and spatial working memory. It is unlikely that the strong activation of any of these areas is a direct consequence of eye movements. In a PET study of voluntary saccadic eye movements (35) activation at the cortical level was found in the precentral gyrus (presumably FEF), SMA, right insula and left cingulate gyrus; a weak activation was also found in the right fusiform gyrus and the right lingual gyrus. Since this activation pattern does not perfectly match the results found during apparent motion and motion imagery, we consider a strong contribution of eye movement related activity as unlikely.

In order to reduce the occurrence of eye movements, concentric rings were used as the apparent motion stimulus of the first experiment. In the motion imagery experiments the radially outward moving dot stimulus which probably did not provoke eye movements produced similar results as the apparent motion stimulus and the rotating grating. Experiment 3 was also replicated outside of the scanner verifying that subjects were able to fixate the cross during apparent motion and motion imagery (see methods).

From electrophysiological and brain imaging studies it is known that all of the activated higher areas of the dorsal pathway can become active in the absence of eye movements. Cells in the primate posterior parietal cortex are able to maintain activity about inferred (“imagined”) moving stimuli in the absence of either sensory input or motor output (36). Areas in the superior parietal lobule (SPL) are activated in mental rotation tasks (37) and in tasks which require imagined information about spatial relations in extrapersonal space (38) as well as covert spatial attention (39). Likewise, the FEF is activated when attention is directed towards remembered loci in extrapersonal space (40) and it has been suggested that the FEF “keeps track of visuospatial information over time” (41). In the macaque monkey, the FEF is interconnected with the posterior part of the parietal lobe, especially the cortex in the intraparietal sulcus (IPS) which suggests functional coupling of these areas. This is supported by the current results showing that the FEF and areas in the IPS are coactivated during motion imagery. Small regions in the dorsolateral prefrontal cortex including BA 9/46 were strongly activated during motion imagery and to a lesser degree during objective motion, and this activation was consistently lateralized to the right hemisphere. Such a lateralization of BA 46 activation to the right hemisphere was also found in association with a spatial working memory task (42). The activation of the anterior cingulate gyrus might be a correlate of a non-spatial, task-oriented attentional process (43) since it was very strong during all mental imagery conditions but almost absent
during all other stimulus conditions. The coactivation of cortical areas that was found during motion imagery indicates that subjects engaged in processes of spatial attention and spatial working memory during this task.

7.2.3 Perception of Transparent Motion

It has been demonstrated how the human visual system is structured in a hierarchical manner representing apparent motion in the extrastriate area MT/MST (see above), which is located at an intermediate stage of the dorsal stream. In a further study we investigated in detail the influence of slight alterations of the configuration of a transparent motion stimulus on visual perception and whether these changes are faithfully reflected in the fMRI signal.

From a computational point of view, the task of motion detection is relatively easy if there is only a single point-like object moving on a blank background. It is already more complicated for the visual system when an object composed of differently oriented contours is moving in a certain direction, because local motion vectors that are perpendicular to the boundaries of the object are detected in early visual areas. In order to compute the direction of motion of the whole object, the local motion vectors have to be combined and integrated in a second processing stage (aperture problem, e. g., 44). An even more complex situation occurs if there are partial occlusions and translucent surfaces in a scene containing multiple moving objects. In these cases the visual system has to represent more than one motion vector in the same region of visual space. This problem is studied in transparent motion paradigms. As an example, consider a moving animal behind bushes on a windy day or two formations of birds flying "through" each other. In a laboratory demonstration of motion transparency, two sets of randomly placed dots are often used which move in opposite directions within the same area on a computer screen. In this case, subjects report seeing two transparent surfaces, one defined by each set of dots moving continuously and independently across each other.

The visual system also uses contrasts in velocity fields to construct object boundaries that are not visible in a stationary scene. This visual process is termed structure-from-motion or shape-from-motion. A camouflaged reptile is only perceived when it starts to move. A laboratory demonstration of structure-from-motion uses random dots that move differently in a circumscribed area. In those conditions a clear contour is perceived around the area that moves differently relative to the surround. The contour vanishes as soon as all dots stop moving or if they move in the same direction with the same velocity. Area MT/MST representing motion fields and area LO representing constructed contours could be identified as the neuronal substrate of structure-from-motion (19). A visual display that combines motion transparency and structure-from-motion is illustrated in Figure 7.5. Arrows symbolize velocities and directions of dots. This display is usually perceived as four transparent surfaces: Two large surfaces move horizontally across each other and two small ones (left upper quadrant of Fig. 7.5) move vertically across each other. The smaller surfaces are perceived as an object
of rectangular shape even though no such contour exists physically (shape-from-motion). The illustrated display is significant in a further aspect: each moving dot is paired with another dot moving in the opposite direction. This is a special case of motion transparency, because the perception of transparent moving surfaces critically depends on the spatial distance of opponent moving dots. It has been shown that transparency perception vanishes if motion signals in opposite directions are locally well balanced (45). We replicated this finding (46) in a psychological study by varying two spatial parameters of the opponent moving dot display shown in Figure 7.5. The varied parameters — lifetime and dot distance — are schematically shown in a magnification of a dot pair in Figure 7.5 (right lower corner). Dot distance varied between 0.01° and 0.13° and lifetime between 20 and 170 ms. Motion transparency vanished at dot distances smaller than approximately 0.07°. Subjects were asked to report the perceived location of the small surfaces which were presented in one quadrant of the screen. A strong interaction between lifetime and dot distance revealed that performance was mostly impaired if lifetime and dot distances was spatially well balanced (critical ratio of lifetime and dot distance is approximately: 2:1). No motion transparency was perceived at relatively small dot distances but the targets were easy to detect because dot pairs were perceived as smeared lines that were oriented orthogonally between target and surround.

Fig. 7.5: Stimulus used in the transparent motion experiment and topological distribution of responses of a single subject. Arrows indicate motion of a dot in a particular direction. Each dot is paired with another dot moving in the opposite direction. Two parameters of a dot pair, lifetime and dot distance (right lower corner), were varied in the experiment. Subjects had to detect a target consisting of dot pairs moving orthogonally with respect to the background dot pairs (upper left quadrant within the stimulus display). Areas responding to the stimulus include early visual areas, MT/MST as well as areas in the posterior parietal lobule (left).
Extrapolating from human psychophysical experiments and electrophysiological experiments with monkeys Qian and colleagues (47) proposed a two stage model for the processing of transparent moving stimuli. The first stage, located primarily in monkey area V1, detects local motion vectors. Neurons of this stage are equally activated by balanced opponent motion stimuli that are perceived as nontransparent and by locally unbalanced motion perceived as transparent. The second stage, located primarily in monkey visual area MT, produces strong suppression between opposite directions of motion. The suppression for the unbalanced, transparent motion is significantly less than that for the balanced, nontransparent stimuli. Thus, activity at the second stage seems to correlate better with the perception of motion transparency than the activity at the first stage. In order to investigate whether the same suppression effect is observable in human area MT/MST, we presented a transparent motion stimulus and two flickering opponent motion stimuli that were varied in dot distance between 0.4° and 0.1°. This relatively

Fig. 7.6: Behavioral and fMRI responses in the transparent motion experiment with respect to variations of parameter dot distance. Target detectability and confidence ratings are highly correlated and increase with increasing dot distance (upper panel). The fMRI signal change of human visual area MT/MST normalized to the signal strength during fixation periods is shown in the lower panel. With small dot distances (0.1°, 0.2°), which subjects perceive as a flickering dot display, the fMRI signal in MT/MST is significantly reduced as compared to a larger dot distance of 0.4° which subjects perceive as a transparent motion display.
small change in presentation parameters elicited a significant change in target
detectability and confidence ratings (Fig. 7.6, upper panel). The bottom panel of
Figure 7.6 shows the fMRI signal change of human visual area MT/MST normal-
ized to the signal strength during fixation periods. There was a clear decrease in
fMRI signal during flickering opponent dot displays (distance: 0.1°; 0.2°) com-
pared to transparent motion stimulation (distance: 0.4°). Parallel to suppression
of MT activation in the monkey there is a suppression of human V5 activation
during presentation of spatially well balanced opponent motion displays (48).

Targets like the one presented in the upper left quadrant of the visual stimulus
shown in Figure 7.5 are detected consciously only during transparent motion stimu-
lation. Kolb and Braun (49) reported that even though conscious perception of
targets vanishes during the presentation of well balanced opponent motion
stimuli, there is still a good ability to detect targets by guessing. Kolb and Braun
(49; 50) drew an analogy to pathological blindsight which occurs in some patients
with lesions in V1. They proposed that in analogy to the two stage model of Qian
et al. (47) the effect of “blindsight in normal observers” could be represented as a
functional suppression between projections from V1 to MT/MST. After such a
suppression conscious perception was believed to be suppressed but localization
was believed to persist through direct projections from subcortical regions. In a
careful replication study we could not replicate Kolb and Braun’s finding1. Instead
we demonstrated good correspondence of confidence ratings with performance.
Confidence ratings were correlated with localization performance in all conditions
that elicit significant localization (45).

7.3 Conclusions

The finding that areas of the dorsal processing stream as early as MT/MST are
activated by apparent motion of illusory contours and even by motion imagery
nearly as strongly as by objectively moving stimuli is an impressive demonstration
of the constructive nature of visual processing. The differential activation of area
MT/MST reflecting these synthetic operations occurred independently of the
source of stimulation. The motion sensitive area was activated based on bottom-
up information during apparent and transparent motion and it was activated
based on top-down information during motion imagery. The bottom-up activation
originated in the retina flowing via subcortical and early cortical processing stages
to MT/MST. The top-down activation presumably originated in the observed pre-
frontal areas flowing via parietal cortex to MT/MST. The presented findings indi-
cate that the constructive visual processes of apparent motion and motion imagery
are accomplished within the same specialized areas that are held responsible for
the analysis of real motion.

1 Our results are further supported by a study that analyzed the second paradigm of the
Kolb and Braun study which is based on stereoptic rivalry (51). Again the stimulus
paradigm was insufficient to induce blindsight in normal observers.
The results of the transparent motion experiment indicate that the fMRI signal measured in area MT/MST faithfully correlates with the perception of motion since small changes in stimulation conditions produces significant changes both in the fMRI signal as well as in behavioral reports of human subjects. However, a strong correlation between fMRI signal changes on the one hand and changes in motion perception on the other hand does not necessarily imply that activity in these areas corresponds directly to conscious experience of the visual scene. An example of a strong MT/MST activation without accompanying conscious experience is provided by a recent fMRI study of blindsight patients (52).

Acknowledgements

All fMRI studies were performed at the department of Neuroradiology, Johann Wolfgang-Goethe Universität, Frankfurt/Main. We thank F. E. Zanella, H. Lanterman, H. Hacker and C. Goebel for support in performing the fMRI studies and D. E. J. Linden for helpful comments on the manuscript. This research was supported by the Max-Planck-Gesellschaft and the Berlin-Brandenburgische Akademie der Wissenschaften.

References


