Research Report

An fMRI study of the reverse perspective illusion

Takefumi Hayashi, Chie Umeda, Norman D. Cook*

Department of Informatics, Kansai University, 2-1-1 Reizenji, Takatsuki, 569-1095 Osaka, Japan

ARTICLE INFO

Article history:
Accepted 8 May 2007
Available online 21 June 2007

Keywords:
fMRI
Reverse perspective
Depth perception
Illusion
Mental rotation
MT

ABSTRACT

"Reverse perspective" is a powerful visual illusion similar to the hollow mask illusion, but more interesting in producing the perception of an illusory motion in a stationary picture. It is caused by conflict between motion parallax and pictorial depth cues in 3D "relief" paintings built with depth inversion. Here we report the measurement of brain activation using fMRI in response to a reverse perspective (RP) object, as well as a normal perspective, 3D-relief object ("shadow-box", SB) and a 2D painting of the same architectural scene. The stimuli were presented to 10 subjects in static and rotating conditions, subtraction of which revealed strong activation of area MT in all three cases. Contrasts between the RP, SB and 2D conditions showed the strongest activation for RP and almost no difference between SB and 2D. The similarity of brain activation between SB and 2D stimuli was interpreted as indicating that observers perceive the illusion of realistic 3D depth in 2D pictures as entirely normal and not qualitatively different from the 3D structure of the shadow-box stimulus. Contrasts between the RP stimulus and either the SB or the 2D stimulus revealed activation of Brodmann Areas 7, 19 and MT (and cerebellar cortex), suggesting the usage of brain regions involved in mental rotation and depth perception in response to the reverse perspective illusion.

© 2007 Elsevier B.V. All rights reserved.

Keywords:
fMRI
Reverse perspective
Depth perception
Illusion
Mental rotation
MT

1. Introduction

The reverse perspective illusion is an illusion of false movement in a stationary picture created by inverting the depth cues in a 3D-relief painting. The familiar "hollow mask illusion", in which a normal face is painted on the inner surface of a mask (Hill and Johnston, 2007), shows related effects, but the "reverspective" artwork of Patrick Hughes (2007; Slyce, 1998) produces a more powerful illusion. In both cases, because of learned expectations about the 3D configuration of familiar (convex) objects, there is a strong tendency to perceive inside-out (concave) objects, such as the hollow-mask, as normal (with protruding nose and chin and receding eye sockets, etc.), until observer or hollow object movement reveals its inverted structure. Such expectations lead the observer to misperceive the static structure of the 3D object, but when the observer moves in relation to the object, a surprising illusion of false movement is typically experienced.

1.1. Structure of the depth inverted objects

The general construction of reverse perspectives with four protrusions from the flat surface of the canvas is shown in Fig. 1A. Perspective cues are drawn on the protruding panels of such structures in order to suggest depth that conflicts with the actual 3D structure (Figs. 1B, C). When viewed at a distance where binocular disparity cues are relatively weak (~2 m), the pictorial visual cues (linear perspective, shadows, etc.) dominate the perception of the picture, which is then seen as a flat picture depicting a 3D scene.
When motion parallax (observer motion) is added to the mixture of visual cues, the visual system is presented with inherently contradictory evidence concerning the 3D structure of the object—a contradiction that produces the illusion of motion in the picture itself. The false motion can be explained by the fact that changes in the retinal image caused by changes in the visual angle are contrary to those expected to occur on the basis of a lifetime of motion parallax experience with normal (convex) solid objects (Cook et al., 2002). Normally, movement to the left in front of a convex object will result in increases in the apparent size (visual angle) of the object’s left side, together with decreases in the right side (Fig. 2, left), and vice versa for movement to the right (Fig. 2, right). Although textbook discussions of motion parallax usually describe the relative motion of near and far objects, static images such as in Fig. 2 suffice to show the before-and-after visual scenes that are obtained due to self-movement. Lateral motion invariably produces visual changes in convex structures that can be described in terms of the relative size of object surfaces. When, however, the depth structure of familiar objects is purposely inverted, the anticipated changes in the size of object surfaces are not realized. On the contrary, movement to the left produces a reduction in the visual angle of the left side of the object and an increase for the right side. Because the anticipated changes in the visual scene are based upon a lifetime of experience of motion parallax when viewing solid objects, observers viewing depth inverted objects consistently report that they perceive a surprising motion in the objects themselves—despite the fact that they are aware that the hollow mask or reverse perspective structure is stationary. Those contradictory cues leave the visual system with only one coherent, if incorrect, interpretation of the nature of the visual stimulus, i.e., that the external object itself has moved. Rather than deny the veracity of the retinal image or dismiss a lifetime of learned expectations concerning motion parallax, observers thus attribute stationary reverse perspective objects with the “magical” property of spontaneous motion. (Subjects also sometimes comment that reverse perspectives induce the queasy feeling of “motion sickness”. This response might be considered as an alternative solution to the perspectival contradiction inherent to the reverse perspective displays, in which the visual changes are perceived as due to unanticipated body movements, rather than due to changes in the visual stimulus.)

### 1.2. Previous experimental work

In a behavioral experiment, we have previously shown that prism goggles (that reverse the normal effects of motion parallax) eliminate the illusory motion in reverse perspective objects only in the direction in which the visual field has been reversed (Cook et al., 2002). Forty subjects first experienced the illusory effect by moving horizontally and vertically in front of large painted objects with inverted depth structure. They were
then fitted with prism goggles that reversed the visual field either vertically or horizontally. Nearly all subjects reported that the characteristic illusory effect then disappeared. The absence of illusory motion was found for both the goggles that produced left–right reversals (when viewing the same objects) and for the goggles that produced up–down reversals (when the same objects were presented upside-down and therefore appeared “normal” with a blue sky above and green grass below). In both cases, the illusion remained intact in the direction unaffected by the prism goggles. Those results clearly demonstrated that the reverse perspective illusion involves the learned expectations of motion parallax: when both the depth structure of familiar objects and motion parallax are inverted, the “double reversal” returns the viewer to a normal state—and no illusion is experienced.

Most previous work on the reverse perspective illusion has been confined to the issues of behavioral phenomenology. In a biography of Patrick Hughes, Slyce (1998) described the nature of the illusion from an artistic perspective. Wade and Hughes (1999) brought the illusion to the attention of psychologists and placed it in its historical context. Papathomas (2002, 2004) measured the distances and angles at which the illusion is experienced in normal subjects. We have recently replicated those findings in two studies where the pictorial depth cues were systematically altered to produce illusions of greater or lesser robustness (Cook et al., in press-a,b). Unfortunately, the illusion cannot be experienced when viewing a 2D image, and is consequently not discussed in most psychology textbooks (an exception being brief mention in Livingstone, 2002, pp. 106–107). Although viewing a 3D reverse perspective object produces a much stronger surprise effect, the visual contradiction can be observed in videos of the illusion in which either the camera or the depth inverted object itself is mobile (Hayashi, 2007).

Because the illusion is highly robust, individual differences are small, and there are no indications of habituation with prolonged viewing (Papathomas, 2002), the visual cues that produce the illusion (Cook et al., in press-a) and the regions of the brain activated during its perception are of considerable interest. In the present experiment, three scenes presenting nearly identical 2D images to the observer when viewed statically were used. In addition to a reverse perspective stimulus (RP), a 3D relief stimulus constructed with normal (not inverted) depth (a “shadow-box”, SB) and a flat 2D image were employed (Fig. 3). Since subject movement within the MRI scanner is of course not feasible, we have designed an experiment in which the stimuli themselves are rotated back-and-forth such that subjects reliably experience the reverse perspective illusion while brain activity is measured.

2. Results

Brain activation for all subjects together showed no significant differences between monocular (left or right) stimulus presentations, so that all data were pooled for a comparison of the three main conditions. Pooling of the data meant that each of the stimuli was presented for a total of 12 min to each subject.

The main results are shown in Figs. 4 and 5, and Table 1. As seen in Fig. 4, brain activations were broadly similar using all three stimuli (RP, SB and 2D). Specifically, by subtracting the activation in response to the stationary object from the activation in response to the rotating object, all three stimuli showed significant activation of visual cortex primarily in the vicinity of area MT, known to be involved in visual motion detection. The strength of the activation, however, differed among the three conditions. The rotation of the 2D image showed the weakest effects, followed by the normal perspective shadow-box, with the strongest effects obtained in response to the reverse perspective object.

Contrasts between the three stimulus conditions (after subtracting out the effects of the static image presentation) provided the most interesting results (Fig. 5). It was found that there was virtually no difference in brain activation between the SB and 2D conditions. In other words, the illusion of depth in the 2D image produced no less (nor more) brain activation than the viewing of the shadow-box which, because of its 3D structure, actually implied more complex changes in the retinal image during rotation. That is, rotation of the SB stimulus presents small, but noticeable changes in the relative widths of the building surfaces as one sees them sequentially from left-of-center, frontally, right-of-center, frontally, and so on (Fig. 2). Viewing of the rotating flat image produces the same sequence of viewer-centered images, but without the small changes in the relative sizes of the building surfaces. It is this lack of (depth-from-motion) motion parallax effects that normally informs us that a visual stimulus is indeed a 2D picture and not a 3D scene. Perhaps because of the ubiquity of 2D visual representations in the modern world, the lack of 3D realism (motion parallax effects) in a 2D picture does not strike us as unusual, and indeed the brain activations to the 2D and SB conditions were virtually negligible...

Fig. 3 – The three stimuli used in the experiment, as viewed from ~30°. (A) is a flat picture, (B) is a normal perspective “shadow-box” and (C) is a reverse perspective object. Note the different widths of the same building sides in the three stimuli.
identical (2D–SB and SB–2D). Future work might include a task in which the subject is requested specifically to detect the 2D or 3D structure of the stimuli, which presumably would lead to differences in brain activation. In the present experiment, however, the subjects were required only to view the stimuli, so that, from the present results, it remains unclear whether the subjects viewed the 2D picture as a 3D structure (the typical illusion of 3D depth induced by pictorial perspective cues) or were unaware of the 3D structure of the shadow-box (and viewed it as a 2D image).

Subtraction of either the shadow-box condition or the 2D image condition from the reverse perspective condition showed similar bilateral activations in several vision-related areas. Specifically, areas BA7 (known to be activated during mental rotation) (Gauthier et al., 2002; Orban et al., 2003; Zacks et al., 2003) and MT (known to be activated during depth-from-motion and optical flow perception) (Cauquil et al., 2006; Peuskens et al., 2001) showed greater activity in the RP than the SB or 2D conditions. There was, moreover, significant activation in the cerebellum in the RP condition (Table 1 and Figs. 4 and 5).

3. Discussion

As is true of all visual illusions, the reverse perspective illusion is of interest primarily because of the light it sheds on normal perception. The main factors that contribute to the illusion are demonstrably motion parallax (observer movement in front of the reverse perspective object) and the pictorial depth cues that are intentionally drawn to indicate 3D depth that is contrary to the actual 3D depth of the relief object (Cook et al., 2002). When normal depth cues are removed from reverse perspective objects, the illusion is weakened and, in the extreme case of a reverse perspective object without perspective lines, shadows or coloring of the buildings, sky or ground (Fig. 1A), depth inversion may or may not be perceived, but the illusion of false movement is absent. Viewers then perceive
either the actual (concave) shape of the "buildings" or the illusion of "normal" (convex) buildings, but normally do not report a sense of "floating" or "twisting" within the pictures. Unlike most visual illusions, reverse perspective involves motion parallax effects—i.e., expectations concerning the changing shape of objects due to self-motion. When those expectations are not consistent with pictorial cues that indicate 3D depth, a situation of conflict between bottom-up motion parallax processing and top-down pictorial depth cue processing is brought about. For this reason, elucidation of the brain regions that are involved in the illusion are likely to contribute to further understanding of the dynamics of perception involving bottom-up and top-down processing.

Previous brain research on the pictorial cues that lead to the perception of 3D depth in 2D images has implicated various parts of the lateral occipital and inferior parietal lobes, notably BA 19 and 7 and areas in and around MT (Cauquil et al., 2006; Welchman et al., 2005). Behavioral study of motion parallax driven by head movements has clarified the relationship between the velocity of head movements and the visual disparity required to perceive depth from motion (Ono and Uijike, 2005). Brain studies of motion parallax employing depth-from-motion have implicated the dorsal inferior parietal area (Rolls and Treves, 1997), but the head and/or body movement that normally produces motion parallax has made it difficult to examine the brain areas involved.

Within the MRI apparatus where subject movement is impossible, motion parallax roughly similar to that due to self-movement can be obtained by rotating the stimuli themselves (consistent with the region of "depth/no motion" in the Ono and Uijike study, 2005), and indeed subjects report seeing the characteristic illusory motion under such conditions. Because the three visual stimuli were shown while rotating around their vertical axes, it is no surprise that, after subtracting the static condition images, brain regions involved in motion detection (notably, MT) were activated for all three stimuli. In the case of the RP stimulus, however, in addition to perception of the rotating motion of the entire structure, a counterintuitive "twisting" and "following" motion of the objects within the scene—as if the objects were moving of their own accord—is also typically reported. Consistent with this perception, contrasts with the SB and 2D conditions showed that the RP condition leads to activation of regions of visual cortex known to be involved in depth perception and mental rotation. Specifically, Brodmann Areas 7 was activated in both the RP and SB conditions and BA19 was activated in all three conditions. They are likely to be a part of the so-called top-down processing of the illusory 3D depth perceived in 2D images. In our study BA40 was significantly activated in the right hemisphere in the RP condition. We conclude, therefore, that areas involved in both the top-down and the bottom-up processing needed to achieve the conflict inherent to the reverse perspective illusion were activated.

Previous work has shown that sufficient top-down pictorial cues must be present in order to over-ride bottom-up (binocular disparity and/or motion parallax) cues and cause misperception of depth inverted objects/faces (Hill and Johnson, 2007). Of particular interest in this regard are the reports of Emrich (1989) and colleagues (Schneider et al., 1996, 2002), indicating that schizophrenics are not susceptible to the depth inversion illusion. They are more likely than normal subjects to perceive the inverted structure veridically as inverted, presumably due to the overpowering influence of bottom-up processing, compared to relatively weak top-down processing. Hyperactivity of BA40 in schizophrenics experiencing delusions of alien control during voluntary movement (Spence et al., 1997) is consistent with this view. Normal subjects engage both the bottom-up processing of motion parallax (BA40 and inferior parietal cortex) and the top-down processing of pictorial perspective cues (BA19 and BA7). As a consequence, they perceive the illusion, whereas, in certain psychopathological states, top-down processing may be dysfunctional, leading paradoxically to the weakening of the illusion.

The bilateral activation of the cerebellum in our study remains to be explained. It may be a result of cerebellar involvement in the sense of "motion sickness" (occasionally reported by subjects viewing reverse perspectives) or related to a loss of balance inherent to the internal contradictions in the RP structures. Further fMRI work on individual differences and various pictorial depth cues is in progress.

### 4. Experimental procedures

The fMRI experiment was undertaken to determine the locations in the brain that are most responsive to the reverse perspective stimulus. Three semi-realistic visual stimuli containing colored objects (~buildings), sky and ground, as well as shading, cast shadows and linear perspective grid lines were used (Fig. 3). When viewed frontally, they presented approximately the same static 2D image to observers at a viewing
distance of 150 cm. In fact, the stimuli differed in their 3D construction: one was a flat 2D image, one was a 3D normal perspective “shadow-box” (SB), and one was a 3D reverse perspective (RP) painting constructed with depth inversion. The pictorial contents of the visual stimuli are immediately recognizable as buildings in an outdoor scene and present the viewer with no paradox concerning the implied 3D structure, when viewed statically. Because the actual 3D structures of the three paintings differed, however, dynamic viewing gives the observer distinctly different impressions of the visual scene. All three images measured 20×50 cm when viewed frontally. The SB and RP stimuli also had a maximal relief depth of 7 cm.

Subjects were 10 undergraduates (6 female, 4 male; ages 19–23) who gave informed consent and were paid for their participation. The entire procedure was undertaken following the ethical guidelines of the Brain Activation Imaging Center of ATR International, Ltd. (Kyoto, Japan). The stimuli were viewed from within the MRI scanner via a mirror directly above the eyes with subjects in a supine position and head immobilized. The stimuli were located 150 cm directly rostral from the subject’s head (Fig. 6A). Because the reverse perspective illusion occurs only with movement, the stimuli themselves were manually rotated back-and-forth (±30°) at approximately ~0.25 Hz around their vertical axis of symmetry. In order to determine whether the perceptual effects of reverse perspective would be possible within the confines of the scanner, we ran pilot experiments involving stationary viewing of rotating reverse perspectives and of dynamic computer graphics. In both cases, subjects report perception of the illusion, although it is typically less robust and/or not as surprising as when self-motion is involved. In the actual experiment, vividness ratings were not collected, but all subjects confirmed that the reverse perspective illusion, i.e., the feeling that the RP object itself was moving in an unusual and counterintuitive way, was apparent while in the scanner. (Lateral motion of the stimuli themselves might also have been employed, but ±30° rotation was chosen for its technical simplicity.) For the rotating RP objects, a rotational velocity of about 0.25 Hz was found to be optimal and monocular viewing produced stronger effects.

The movement condition was contrasted with a static condition in which the subjects viewed the same three stimuli presented frontally without movement. A block design was used: static and dynamic conditions were alternated every 24 s for 5 presentations each of the three stimuli. The 2D, SB and RP conditions were repeated three times each, giving a total scanning time of 6 min per stimulus (1/2 static and 1/2 dynamic), and the entire procedure was repeated for the left eye and the right eye. The sequence of 2D, SB and RP presentations was fixed for each subject, but randomized across subjects. Following calibration with a patch over one eye, the subject viewed the stimuli; recalibration was then done with a patch over the other eye, and the stimuli again viewed. The left/right sequence was balanced among the 10 subjects.

Approximately sinusoidal rotation of the stimuli was timed with a metronome, audible to the experimenter only. Because of the relatively short distance between the subject’s eyes and the stimuli in the MRI scanner room, binocular stereopsis can reveal the actual structure of the RP stimuli, and consequently destroy the illusory effect. At a distance of 1.5 m, the illusion periodically appears and disappears, depending on transient binocular disparity effects, so that all subjects were tested monocularly using an eye patch (left/right order balanced among the subjects).

Brain imaging was performed with a 1.5 Tesla Marconi Magnex Eclipse scanner using an interleaved sequence. For each subject, high-resolution anatomical T2 weighted images were acquired using a fast spin echo sequence. These scans consisted of 50 contiguous axial slices with a 0.75×0.75×3.0 mm voxel resolution covering the cortex and cerebellum. Functional T2* weighted images were then acquired using a gradient echo-planar imaging sequence (echo time, 55 ms; 3.4 ms; 60° flip angle).

Fig. 6 – The experimental set-up (A), the block design (B), and the experimental procedure (C). The stimuli were rotated manually, with the experimenter standing in a position rostral from the subject and out of sight. Stimuli were mounted upside-down so that they were viewed as “normal” scenes with the blue sky above and the ground below. The order of presentation of the stimuli (RP, SB, 2D) and the order of eye patching (left, right) were fixed for each subject, but randomized across subjects.
repetition time, 6000 ms; flip angle, 90°). A total of 50 contiguous axial slices were acquired with a 3.0×3.0×3.0 mm voxel resolution. The field of view included the cortex and cerebellum. Calibration was done three times per subject following eye patch placement or removal.

Images were preprocessed using programs within SPM2 (Wellcome Department of Cognitive Neurology, London UK). Differences in acquisition time between slices were accounted for, movement artifact was removed, and the images were spatially normalized to a standard space using a template EPI image (Bounding Box, x = −90 to 91 mm, y = −126 to 91 mm, z = −72 to 109 mm; voxel size, 3.0×3.0×3.0 mm). Images were smoothed using a 6-mm FWHM Gaussian kernel. Regional brain activity for the various conditions was assessed on a voxel-by-voxel basis using SPM2. A fixed effect model was employed (the data from all subjects were analyzed together). The data were modeled using a box-car function convolved with the hemodynamic response function. In addition, global normalization and grand mean scaling were carried out.

Acknowledgments

This work was supported by a grant from the Japanese Society for the Promotion of Science (JSPS.KAKENHI, Grant No. 16300085) and an Academic Frontiers Project at Kansai University (2003–2007). Preliminary results of this study were reported at the European Conference of Visual Perception in La Coruna, Spain (August, 2005).

References


