Perception, 2010, volume 39, pages 1086 – 1093

doi:10.1068/p5990

Perspective-based illusory movement in a flat billboard—an explanation

Thomas V Papathomas
Department of Biomedical Engineering and Laboratory of Vision Research, Rutgers University, Busch Campus, Piscataway, NJ 08854, USA; e-mail: papathom@rci.rutgers.edu;
website: http://ruccs.rutgers.edu/~papathom/index.html

Zoe Kourtzi, Andrew E Welchman
School of Psychology, University of Birmingham, Birmingham, UK
Received 25 April 2008, in revised form 5 May 2010

Abstract. We describe a compelling motion illusion elicited by a huge billboard placed along a street, depicting a building that contains strong perspective cues. When observers move fast along the opposite sidewalk, they perceive the depicted building as rotating in their direction of travel. This is a special case of the 'following', or 'pointing out of the picture', illusion that elicits a strong illusory motion percept. Here we discuss the cause of the illusory motion and suggest that the brain relies on the depicted perspective cues to infer a 3-D shape and a concomitant motion that is incompatible with the physical pictorial surface.

This paper deals with a special case of what is known as the 'following' (Kubovy 1986, page 62) or 'pointing out of the picture' (Koenderink et al 2004) illusion: painted scenes and portraits appear to rotate for viewers who move in front of them. Here we report on a large flat painting that elicits a strong 'following' illusion, especially for viewers who move fast past it. Figure 1 is a photograph of a large two-dimensional (2-D) billboard depicting a building in perspective. The flat advertisement billboard was erected along the sidewalk, in front of a building under construction on Bath Row in Birmingham, UK.

Figure 1. The flat advertisement billboard that elicits the illusory motion for viewers who move fast past it.

It measured about 8.5 m in width and 6.5 m in height at its highest point. Pedestrians walking along the sidewalk on the opposite side of the street experienced a compelling illusory motion: they obtained a vivid impression that the depicted building moved—rotating in their direction of travel. The illusion was much stronger for a fast moving viewer, ie for a runner or for a passenger riding in a vehicle that moved along the road. It was particularly compelling in that the presence of other static objects (eg a lamppost, the building behind) in the scene provided the moving observer with evidence that the viewed object was painted on a flat billboard. Moreover, the 2-D billboard surface was cut along the top edges of the depicted building, creating motion parallax signals for the top edges relative to the scenery behind them, consistent with a 2-D surface. Thus, Pirenne's (1970) condition that "the spectator is unable to see the painted surface, *qua* surface" is not satisfied. As a result, because of the strong perspective cues, the 3-D shape of the depicted building is perceived to be invariant under changes in the vantage point (Kubovy 1986, page 56). In fact, Vishwanath et al (2005) expanded Pirenne's findings by providing evidence that the visual system achieves this shape invariance through an estimate of the local orientation of the painted surface. Furthermore, because the corner of the building 'faces' the viewer by pointing out of the picture, it appears to 'follow' the viewer by rotating in physical space, as do the eyes of a head-on portrait, a pointing finger, or a road leading from the center foreground straight to the horizon (Gombrich 1972; Goldstein 1979; Kubovy 1986, page 84; Koenderink et al 2004).

Observations and experiments on this type of 'following' illusion have been reported extensively in the literature for perspective scenes and full-face portraits (eg Gombrich 1972; Wallach 1976; Goldstein 1979, 1988; Kubovy 1986; Koenderink et al 2004), but they deal with smaller-sized paintings; also, they typically deal with images viewed statically from different angles, rather than experimenting with moving viewers and examining the percept of a figure that appears to rotate. A notable exception is the work of Wallach et al (1974), who examined how accurate humans are in judging, while they walk, whether objects are stationary or moving; however, they worked with physical volumetric, not flat pictorial, objects.

What might give rise to the large-scale illusory motion? Here we discuss a possible explanation, based on inferential theories of perception (Gregory 1975, 1980, 1997; Rock 1983), specifically under self-motion (Wallach 1985, 1987; Gogel 1990; Wertheim 1994). In particular, consider figure 2a that depicts the front view of a simplified model of the building in the billboard of figure 1; it consists of trapezoids ABCD and ABEF that contain linear perspective cues. Figure 2b is the top, or plan, view of the same 2-D stimulus.

The frontoparallel plane on which the picture is drawn is shown as a solid straight line in figure 2b, with points A, D, and F shown in correspondence with those of figure 2a. The viewer's self-motion is indicated by the eye, Y, that moves from position Y_1 to Y_2 over time, as shown by the thin arrow. The (thin) lines of sight for point A and the (thick) lines for the perceived illusory 3-D object for a viewer in position Y_1 are shown by dotted lines. The illusory motion depends on viewers inferring that the painted billboard they are viewing is a 3-D object despite evidence to the contrary from motion parallax, binocular disparity, and occlusion. For example, apex A appears to be located along the line of sight connecting the physical apex A to the eye; it is perceived at position A_1 in front of points D and F, because this is the 3-D shape representation of the building, due to perspective. When the viewer moves to Y_2 while maintaining this 3-D representation of the object (thick dashed lines), apex A is still perceived in front of D and F, along the (new) line of sight connecting A to Y_2 (thin dashed line); the physical position of apex A is stationary, anchored on the 2-D surface of the billboard. As a result, A appears to move to position A_2 as shown in figure 2b.

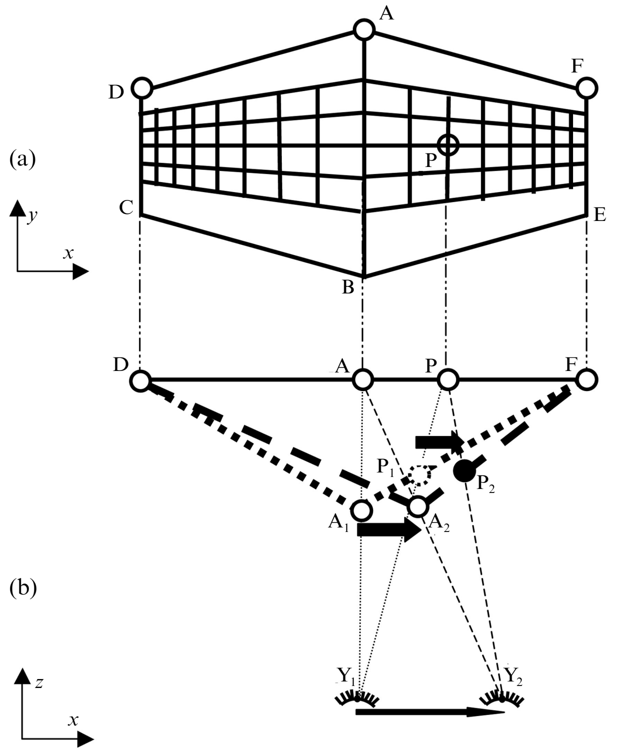

Figure 2. (a) Front view of a simplified stimulus with linear perspective cues. (b) Top (plan) view: the stimulus of figure 2a is shown by a single thick line that contains points A, D, F, and P. The viewer's eye is initially at Y_1. We assume that the viewer perceives apex A at A_1, closer than it actually is, because of perspective. More generally, the viewer perceives the depicted building as protruding (thick dotted lines); lines of sight for eye position Y_1 are shown by thin dotted lines. The viewer's eye moves to Y_2 (thin arrow); lines of sight are shown by thin dashed lines. If the viewer maintains the 3-D percept (thick dashed lines), the perceived location of A will move to A_2; any other point, such as P, will appear to move. Thus the entire object will appear to move (thick arrows).

Of course, any other point, or feature, on the painted building—other than points lying on edges CD and EF—also appears to move, as shown for a generic point P in figure 2. It appears to move from P_1 for eye position Y_1 to P_2 for eye position Y_2. As a result of each feature appearing to move, the surfaces of both trapezoids ABCD and ABEF undergo the kind of transformation that causes the building to be perceived to move by a moving viewer. What is perhaps surprising is that observers do not veto the information provided by the perspective cue in the face of other signals that specify that the viewed surface is flat.

There are several points to be made here: first, this concept of a perceptual, as opposed to abstract, mental 3-D representation is akin to the construct developed by Tyler (1974) to explain illusory motion in stereograms and to Gogel's (1990) theory of phenomenal geometry. Similar perceptual top–down influences have been proposed for biological motion (Shiffrar and Freyd 1993), where the interpretation of the movement depends on knowledge about the probable structure of the environment. Second, observe that points whose 3-D representation protrudes further from the wall appear to move more; for example, in figure 2, point A moves with a larger amplitude than point P because the representation A_1 protrudes more than P_1. Also, notice that the perceived building undergoes a non-Euclidean transformation, in agreement with the findings of Koenderink et al (2004). Third, even though we chose to show the perceived edges CD and EF anchored to the painted surface, this is not necessarily the case; edge CD, for example, can be perceived anywhere in depth along the line of sight that connects the eye to the physical location of edge CD; ditto for edge EF. Further experiments are required to investigate this issue and, more generally, the nature of the 3-D representation of the depicted building. Finally, figure 2b suggests that, as the observer moves from Y_1 (near-center position) to Y_2 (to the right of center), the width of trapezoid ABCD appears larger than that of ABEF. This is in agreement with the findings of Papathomas et al (2004), who obtained estimates of this effect for an obliquely viewed perspective scene.

The logic of this interpretation rests on the same assumptions as those made by Papathomas (2007) based on Gogel's (1990) theory of phenomenal geometry that uses perceived direction, depth, and self-motion. In our case, the assignment of 3-D depth elicited by the flat stimulus appears to be the root cause of the illusory motion experienced by a moving viewer. This explanation is similar to the one proposed to account for the illusory motion of related stimuli (Papathomas 2007), such as hollow masks (Gregory 1975; Hill and Johnston 2007), reverspectives (Wade and Hughes 1999; Cook et al 2002; Hayashi et al 2007), and stereograms (Shimono et al 2002). In particular, our model is consistent with empirical observations that the *magnitude* of the perceived illusory motion is larger for a reverse-perspective painting than for a flat-perspective painting of comparable dimensions. Appendix A contains a detailed analysis of such a comparison. However, our model does not explain why, *qualitatively*, illusory motion is much more robust in reverse-perspective than in flat-perspective paintings. One reason may be that our model does not explicitly take into account motion-parallax signals.

Alternative explanations have been proposed for the perception of pictorial objects, in general, and the illusory 'following' percept, in particular (Hochberg 1971; Gombrich 1972; Wallach 1976; Goldstein 1979, 1988; Kubovy 1986, page 84; Koenderink et al 2004; Todorović 2008; Rogers and Gyani 2010). Rogers and Gyani (2010) advance an explanation that relies less on top–down influences or on comparisons between actual and expected retinal flow patterns. They posit that, for a viewer who moves in front of a 2-D billboard that depicts a 3-D structure and experiences the absence of motion parallax that would have resulted from such a volumetric structure, the only possible explanation is a rotation of the 3-D scene concomitant with the viewer's motion. Their explanation also accounts for the illusory motion of hollow masks and reverse perspectives.

Particularly relevant are theories that explain the illusory motion as a result of changes in the proximal stimulus that run counter to our expectations/predictions (Kubovy 1986; Cook et al 2002). As an anonymous reviewer remarked, our experience-based expectation of seeing more of ABEF and less of ABCD as we move from Y_1 to Y_2 is not realized, thus triggering the percept of illusory motion; in effect, it is as if the viewer's visual system 'reasons' that, since the expected pattern did not materialize, the stimulus must have moved. Nevertheless, this must be occurring at a rather high level of perceptual processing. Vishwanath et al (2005) provide compelling evidence that interpreting pictorial signals requires an estimate of the real 3-D orientation of the projection surface. In our case, sensory information from binocular disparity, motion parallax, and patterns of occlusion are all compatible with viewing a flat billboard surface. Thus, any high-level inferences about movement of the stimulus most likely involve an elaborated representation of the scene that incorporates information from pictorial depth cues. The percept is thus intriguing as it appears to rely on knowing the physical layout of the scene and then 'surprising itself' based on an interpretation of perspective information that is incompatible with other depth cues.

In conclusion, our explanation for the illusion of figure 1 is based on the observation that the 2-D building depicted on the billboard gives rise to a compelling 3-D percept that changes with the observer's point of view (Tyler 2005). The existence of this subjective

3-D percept is supported by evidence that observers viewing 2-D paintings with strong linear perspective cues use vergence eye movements that are governed by the illusory, rather than the physical, depth[(1)] (Enright 1987). The illusory motion in the billboard display offers an interesting test of the depth representations that underlie our routine perceptions. Work on the hollow-mask illusion has combined psychophysics and computational algorithms to simulate human perceptual processes. In particular, a computer model that uses feature tracking to recover 3-D facial geometry from animation sequences 'falls victim' to the hollow-mask illusion and 'perceives' illusory motion because, just like humans, it is endowed with a top – down convex face representation (Kaur et al 2000). Further experiments are necessary to investigate similar effects for painted buildings and scenes via psychophysical and computational approaches.

Acknowledgments. We wish to thank Marty Banks, Michael Kubovy, and Dhanraj Vishwanath for valuable advice during the revision of the manuscript. We also thank two reviewers for providing critical feedback that motivated us to examine the differences between reverse perspectives and flat paintings with perspective cues.

References

Cook N D, Hayashi T, Amemiya T, Suzuki T, Leumann L, 2002 "Effects of visual-field inversions on the reverse-perspective illusion" *Perception* **31** 1147 – 1151

Enright J T, 1987 "Art and the oculomotor system: Perspective illustrations evoke vergence changes" *Perception* **16** 731 – 746

Gogel W C, 1990 "A theory of phenomenal geometry and its applications" *Perception & Psychophysics* **48** 105 – 123

Goldstein E B, 1979 "Rotation of objects in pictures viewed at an angle: evidence for different properties of two types of pictorial space" *Journal of Experimental Psychology: Human Perception and Performance* **5** 78 – 87

Goldstein E B, 1988 "Geometry or not geometry? Perceived orientation and spatial layout in pictures viewed at an angle" *Journal of Experimental Psychology: Human Perception and Performance* **14** 312 – 314

Gombrich E H, 1972 "The 'what and the how': Perspective representation and the phenomenal world", in *Logic and Art: Essays in Honor of Nelson Goodman* Eds R Rudner, I Schemer (New York: Bobbs-Merrill) pp 129 – 149

Gregory R L, 1975 *Eye and Brain* (New York: McGraw-Hill)

Gregory R L, 1980 "Perceptions as hypotheses" *Philosophical Transactions of the Royal Society of London, Section B* **290** 181 – 197

Gregory R L, 1997 "Knowledge in perception and illusion" *Philosophical Transactions of the Royal Society of London, Section B* **352** 1121 – 1128

Hayashi T, Umeda C, Cook N D, 2007 "An fMRI study of the reverse-perspective illusion" *Perception* **1163** 72 – 78

Hill H, Johnston A, 2007 "The hollow-face illusion: object-specific knowledge, general assumptions or properties of the stimulus?" *Perception* **36** 199 – 223

Hochberg J, 1971 "Space and movement", in *Experimental Psychology* Eds J W King, L A Riggs (New York: Holt, Rinehart and Winston) pp 475 – 550

Hoffmann J, Sebald A, 2007 "Eye vergence is susceptible to the hollow-face illusion" *Perception* **36** 461 – 470

Kaur M, Papathomas T V, DeCarlo D, 2000 "Schema- and data-driven influences in the hollow-face illusion: experiments and model" *Investigative Ophthalmology and Visual Science* **41** 224 [abstract]

Koenderink J J, Doorn A J van, Kappers A M L, Todd J T, 2004 "Pointing out of the picture" *Perception* **33** 513 – 530

Kubovy M, 1986 *The Psychology of Perspective and Renaissance Art* (Cambridge: Cambridge University Press)

Papathomas T V, 2007 "Art pieces that 'move' in our minds—An explanation of illusory motion based on depth reversal" *Spatial Vision* **21** 79 – 95

(1) Evidence for similar behavior of vergence eye movements, ie following the illusory rather than the physical depth, has been obtained by Hoffmann and Sebald (2007) for viewing the hollow mask and by Wagner et al (2008) for viewing reverspectives [but see Wade et al (2001) and Wismeijer et al (2008) for different findings that are probably due to differences between saccading and fixating].

Papathomas T V, Vidnyánszky Z, Zhuang X, 2004 "From 2D to 3D and back: Perception of rotated 2D pictorial scenes depends on the 3D surfaces they depict" *Vision Sciences Society Conference* (Novato, CA: Vision Sciences Society Conference Proceedings) page 110 [abstract]

Pirenne M H, 1970 *Optics, Painting & Photography* (Cambridge: Cambridge University Press)

Rock I, 1983 *The Logic of Perception* (Cambridge, MA: MIT Press)

Rogers B J, Gyani A, 2010 "Binocular disparities, motion parallax, and geometric perspective in Patrick Hughes's 'reverspectives': Theoretical analysis and empirical findings" *Perception* **39** 330 – 348

Shiffrar M, Freyd J, 1993 "Timing and apparent motion path choice with human body photographs" *Psychological Science* **4** 379 – 384

Shimono R, Tam W J, Stelmach L, Hildreth E, 2002 "Stereoillusory motion concomitant with lateral head movements" *Perception & Psychophysics* **64** 1218 – 1226

Todorović D, 2008 "Is pictorial perception robust? The effect of the observer vantage point on the perceived depth structure of linear-perspective images" *Perception* **37** 106 – 125

Tyler C W, 1974 "Induced stereomovement" *Vision Research* **14** 609 – 613

Tyler C W, 2005 "A horopter for two-point perspective" *Proceedings of SPIE, Human Vision and Electronic Imaging X* **5666** 306 – 315

Vishwanath D, Girshick A R, Banks M S, 2005 "Why pictures look right when viewed from the wrong place" *Nature Neuroscience* **8** 1401 – 1410

Wade N J, Curthoys I, MacDougall H, Cornell E, 2001 "Fluctuations in perceived depth and convergence" *Perception* **30** Supplement, 28

Wade N J, Hughes P, 1999 "Fooling the eyes: trompe l'oeil and reverse perspective" *Perception* **28** 1115 – 1119

Wagner M, Ehrenstein W H, Papathomas T V, 2008 "Vergence in reverspective: Percept-driven versus data-driven eye movement control" *Neuroscience Letters* **449** 142 – 146

Wallach H, 1976 "The apparent rotation of pictorial scenes", in *Vision and Artifact* Ed. M Henle (New York: Springer) pp 65 – 69

Wallach H, 1985 "Perceiving a stable environment" *Scientific American* **252** (May) 118 – 124

Wallach H, 1987 "Perceiving a stable environment when one moves" *Annual Review of Psychology* **38** 1 – 27

Wallach H, Stanton L, Becker D, 1974 "The compensation for movement-produced changes in object orientation" *Perception & Psychophysics* **15** 339 – 343

Wertheim A H, 1994 "Motion perception during self motion" *Behavioral and Brain Sciences* **17** 293 – 355

Wismeijer D A, Ee R van, Erkelens C J, 2008 "Depth cues, rather than perceived depth, govern vergence" *Experimental Brain Research* **184** 61 – 70

Appendix

Parts (a) and (b) of figure A1 show simplified top views for a viewer moving from position 1 to 2 over a distance M_E in front of a flat and a reverse-perspective painting, respectively. For convenience, we use line DF in (b) to indicate an (imaginary) fronto-parallel surface at the front of the reverse-perspective painting. The front views of the paintings are similar to that of figure 2a. The viewer is at a distance V and moves along a line parallel to the painting's surface in (a), and parallel to the front surface DF in (b). The physical stimuli are shown by solid lines. As a result of the perspective cues, the flat (a) and concave (b) surfaces appear as convex. Thus, the viewer perceives the surfaces shown by dotted lines in position 1, which, due to the viewer's movement, are smoothly transformed to the surfaces shown by dashed lines for position 2.

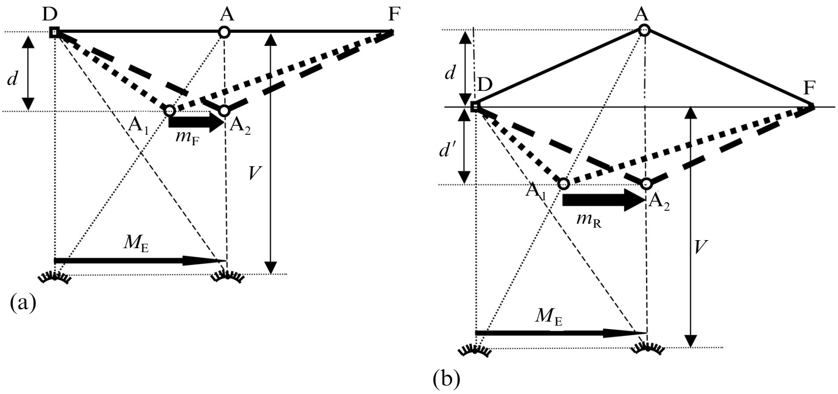

Figure A1. Magnitude of illusory motion for (a) a flat perspective and (b) a reverse perspective.

Specifically, apex A is perceived to be in front of the flat perspective by a depth d. To afford a fair comparison, we consider that the physical depth of the reverse-perspective painting is equal to d. This results in apex A being perceived at a depth d' in front of the front surface DF in (b). Our model assumes that $d' \approx d$; this is generally the case, based on anecdotal observations, although it is not a critical assumption. Apex A appears to move from A_1 to A_2 over a distance m_F for the flat and m_R for the reverse perspective. From similar triangles in (a),

$$\frac{m_F}{M_E} = \frac{d}{V},$$

from which

$$m_F = M_E \frac{d}{V}. \tag{A1}$$

From similar triangles in (b),

$$\frac{m_R}{M_E} = \frac{d+d'}{V+d} \approx \frac{2d}{V+d},$$

from which

$$m_R \approx 2M_E \frac{d}{V+d}. \tag{A2}$$

Equations (A1) and (A2) yield a ratio m_R/m_F [assuming a strict equality in equation (A2)]

$$\frac{m_R}{m_F} = \frac{2V}{V+d}. \tag{A3}$$

Because $0 < d < V$, we get from equation (A3)

$$1 < \frac{m_R}{m_F} < 2\,.$$

In particular, if we assume that d is the fraction of V, $d = cV$, where $c < 1$, then

$$\frac{m_R}{m_F} = \frac{2}{1+c}\,. \qquad \text{(A4)}$$

As a typical example, when $d = 0.1V$, the above equation yields $m_R = 1.82 m_F$. The analysis correctly predicts that the magnitude of m_R is significantly larger than that of m_F.

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 39 2010

www.perceptionweb.com