Acta Psychologica 138 (2011) 359–366

Contents lists available at Sciverse ScienceDirect

Acta Psychologica

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

Do we perceive a flattened world on the monitor screen?

Fulvio Domini [a,b,*], Rajesh Shah [a], Corrado Caudek [c,b]

[a] Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Box 1821, 89 Waterman Street & 229 Waterman Street, Providence, RI 02912, USA
[b] Center for Neuroscience and Cognitive Systems@UniTn, Istituto Italiano di Tecnologia, Via Bettini 31, 38068 Rovereto (TN), Italy
[c] Dipartimento di Psicologia, Università degli Studi di Firenze, Via di San Salvi 12, Complesso di San Salvi, Padiglione 26, 50135 Firenze (FI), Italy

ARTICLE INFO

Article history:
Received 4 March 2011
Received in revised form 27 July 2011
Accepted 29 July 2011
Available online 9 October 2011

PsychInfo classification:
2323
2300

Keywords:
3D shape
Depth perception
Cue combination
Disparity
Natural stimuli

ABSTRACT

The current model of three-dimensional perception hypothesizes that the brain integrates the depth cues in a statistically optimal fashion through a weighted linear combination with weights proportional to the reliabilities obtained for each cue in isolation (Landy, Maloney, Johnston, & Young, 1995). Even though many investigations support such theoretical framework, some recent empirical findings are at odds with this view (*e.g.*, Domini, Caudek, & Tassinari, 2006). Failures of linear cue integration have been attributed to cue-conflict and to unmodelled cues to flatness present in computer-generated displays. We describe two cue-combination experiments designed to test the integration of stereo and motion cues, in the presence of consistent or conflicting blur and accommodation information (*i.e.*, when flatness cues are either absent, with physical stimuli, or present, with computer-generated displays). In both conditions, we replicated the results of Domini et al. (2006): The amount of perceived depth increased as more cues were available, also producing an over-estimation of depth in some conditions. These results can be explained by the Intrinsic Constraint model, but not by linear cue combination.

1. Introduction

From an evolutionary point of view, it is to be expected that the human visual system makes use of all the available information to estimate the three-dimensional (3D) properties of the spatial layout and of the objects within it. In psychophysics, the sources of information indicating depth are often labeled as "cues" (*e.g.*, Cutting & Vishton, 1995). The problem of how information about the 3D structure of objects is combined from various cues is named *depth-cue combination* and has received wide attention (*e.g.* Bruno & Cutting, 1988; Landy et al., 1995; Massaro, 1988). The purpose of the present investigation is to test the hypothesis that some of the distortions of perceived depth that had been reported in the literature can be explained by the presence of flatness cues in computer-generated displays (*e.g.*, Adams & Mamassian, 2004; Enright, 1991; Koenderink, van Doorn, & Kappers, 1994; Rogers, 1995). The role of the flatness cues will be described by making reference to the current depth-cue combination model (Jacobs, 2002; Jacobs & Kruschke, 2011; Knill & Richards, 1996; Körding & Wolpert, 2006; Landy et al., 1995). An alternative model of depth-cue combination will also be discussed (Domini et al., 2006).

1.1. Linear cue integration

Bayesian probability theory provides a normative framework for describing how prior knowledge and information from multiple cues could be combined to make perceptual inferences (Knill & Richards, 1996; Körding, 2007). Landy, Banks and Knill describe a particular instantiation of the Bayesian normative model, which gives rise to the linear models for maximum reliability (Landy et al., 2011). Here, we focus on linear cue integration, because this model has had the largest impact on the literature (*e.g.*, Angelaki, Gu, & DeAngelis, 2009; Arnold, Tear, Schindel, & Roseboom, 2010; Dewing & Ernst, 2006; Helbig & Ernst, 2007; Hillis, Watt, Landy, & Banks, 2004; Knill, 2007; Knill & Saunders, 2003; Körding, 2007; Nardini, Jones, Bedford, & Braddick, 2008; Saunders & Backus, 2006; Todorovic, 2009; Zalevski, Henning, & Hill, 2007). According to linear cue combination, independent estimates of 3D properties are linearly combined for producing the minimum-variance estimate (Cochran, 1937; Ghahramani, Wolpert, & Jordan, 1997). Suppose that we are interested in estimating the depth $\mathbf{z}$ from the image signals $\mathbf{d}$ (disparity) and $\mathbf{v}$ (motion).[1] When both cues are presented together, the optimal estimate of $\mathbf{z}$ is provided by the mode of the posterior distribution

$$p(\mathbf{z}|\mathbf{d},\mathbf{v}) = \frac{p(\mathbf{d},\mathbf{v}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{d},\mathbf{v})}, \tag{1}$$

[1] The information provided by motion and disparities is described in the Appendix.

* Corresponding author. Tel.: +1 401 863 1356.
E-mail addresses: fulvio.domini@brown.edu (F. Domini), corrado.caudek@unifi.it (C. Caudek).

0001-6918/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.actpsy.2011.07.007

where $p(\mathbf{d},\mathbf{v}|\mathbf{z})$ is the likelihood function, $p(\mathbf{z})$ is the prior, and $p(\mathbf{d},\mathbf{v})$ is a constant term. In order to find the maximum *a posteriori* (MAP), both the likelihood function $p(\mathbf{d},\mathbf{v}|\mathbf{z})$ and the prior $p(\mathbf{z})$ must be specified. As stated by the linear model of cue integration, the likelihood function of Eq. (1) is:

$$p(\mathbf{d},\mathbf{v}\,|\,\mathbf{z}) = p(\mathbf{d}\,|\,\mathbf{z};\mathbf{p}_d)p(\mathbf{v}\,|\,\mathbf{z};\mathbf{p}_v), \tag{2}$$

where $\mathbf{p}_d$ and $\mathbf{p}_v$ are the scene parameters (see Appendix A) and $p(\mathbf{d}|\mathbf{z};\mathbf{p}_d)$ and $p(\mathbf{v}|\mathbf{z};\mathbf{p}_v)$ are Gaussian functions. If the scene parameters $\mathbf{p}_d$ and $\mathbf{p}_v$ are estimated correctly then, through inverse geometry, it is possible to obtain an unbiased estimate of the true depth-map $\mathbf{z}_0$ from $\mathbf{d}$ and $\mathbf{v}$. The likelihood functions provide a model of the noise affecting the measurement of $\mathbf{d}$ and $\mathbf{v}$, and, as a consequence, the estimate of the true depth-map $\mathbf{z}_0$. In Fig. 1, these likelihood functions are shown (in red) for a disparity signal (top panel) and a motion signal (middle panel). In both cases, the likelihood functions are generated by the same distal depth $\mathbf{z}_0 = 100$ mm. When both signals are present, the likelihood function $p(\mathbf{d},\mathbf{v}|\mathbf{z})$ is also peaked at $\mathbf{z}_0 = 100$ mm, but it is characterized by a smaller uncertainty (bottom panel, red curve).

For Gaussian distributions and for flat prior distributions, the minimum-variance MAP estimate is found by a weighted linear combination of the depth estimates computed separately for each cue:

$$z' = \sum_j w_j z'_j \qquad w_j = \frac{\sigma_j^{-2}}{\sum_j \sigma_j^{-2}} \tag{3}$$

where z'_j is the amount of depth specified by cue j, and σ_j is the standard deviation of that cue's estimate. The weights w_j are non-negative and sum to one. If the reliabilities of the depth estimates derived from the motion and disparity signals are defined as the reciprocal variances, $r_v = 1/\sigma_v^2$ and $r_d = 1/\sigma_d^2$, then the reliability of the resulting estimate

$$r_c = \frac{1}{\sigma_c^2} = r_v + r_d \tag{4}$$

will be greater than the reliability of either of the single-cue estimates (*i.e.*, $\sigma_c^2 < \sigma_v^2, \sigma_d^2$). By combining information from several cues, therefore, this combination rule yields an estimate having greater precision than could be obtained by using any one cue alone.

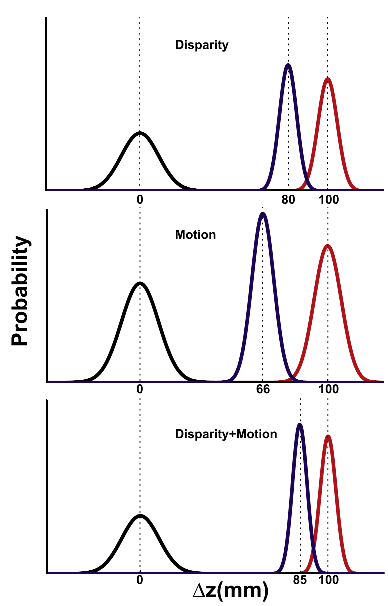

Fig. 1. Maximum likelihood estimates of depth (80, 66, 85 mm) for the combination of disparity information and flatness cues (top panel), motion information and flatness cues (middle panel), and disparity + motion and flatness cues (bottom panel). The red curves represent the likelihood functions for disparity, $p(\mathbf{d}|\mathbf{z};\mathbf{p}_d)$, motion, $p(\mathbf{v}|\mathbf{z};\mathbf{p}_v)$, and disparity + motion information, $p(\mathbf{d},\mathbf{v}|\mathbf{z})$. These functions are assumed to be Gaussian and centered at the true depth value (100 mm). Note that the variance of the disparity + motion likelihood function is smaller than the variances of either of the individual (motion-only and disparity-only) likelihood functions. The black curves, peaked at 0 depth, represent the likelihood functions of the flatness cues, $p(\mathbf{f}|\mathbf{z})$. The blue curves are the likelihood functions resulting from the product of the disparity-only, motion-only, and disparity + motion likelihood functions and the likelihood function of the flatness cues. The peaks of the joint likelihood functions (blue) for the depth cues and the flatness cues provide the MLE estimates that are obtained by combining multiple sources of information.

1.2. Cues to flatness in computer displays

The combination rule proposed by the model of linear cue combination can be used to account for a result often reported in the psychophysical literature, that is, the mis-estimation of depth magnitudes and distances in computer-generated displays and virtual-reality scenes (Allison, Gillam, & Vecellio, 2009; Bradshaw, Parton, & Glennerster, 2000; Braunstein & Tittle, 1988; Caudek & Domini, 1998; Caudek, Domini, & Di Luca, 2002; Caudek & Proffitt, 1993; Creem-Regehr, Willemsen, Gooch, & Thompson, 2005; Di Luca, Domini, & Caudek, 2004; Domini & Caudek, 2003a, 2003b; Domini, Caudek, & Skirko, 2003; Domini, Caudek, Turner, & Favretto, 1998; Domini, Caudek & Proffitt, 1997; Domini, Vuong & Caudek, 2002; Durgin, Proffitt, Olson, & Reinke, 1995; Hibbard & Bradshaw, 2003; Knapp & Loomis, 2004; Thompson et al., 2004). Such perceptual distortions often take the form of an under-estimation of depth (Buckley & Frisby, 1993; Ellis, Smith, Grunwald, & McGreevy, 1991; Frisby, Buckley, & Duke, 1996; Frisby, Buckley, & Horsman, 1995; van Ee, Banks, & Backus, 1999). It has been argued that depth under-estimation may arise from a *cue conflict* among the depth cues manipulated by the experimenter and the depth cues inadvertently provided by the experimental setting (*e.g.*, Buckley & Frisby, 1993; Frisby et al., 1996, 1995; Tittle & Braunstein, 1993; Todd, Thaler, Dijkstra, Koenderink, & Kappers, 2007; van Ee et al., 1999; Watt, Akeley, Ernst, & Banks, 2005). When participants judge 3D shape properties simulated on a CRT screen, in fact, cues for “flatness” are always present.

Young, Landy, and Maloney (1993) introduced the notion of “cues to flatness” by pointing out that cues manipulated in a computer simulation are always accompanied by “other, extraneous cues (vergence, accommodation, motion parallax if the head is free to move, prior knowledge) all of which may signal that the display is flat (which, in fact, it is)” (Johnston, Cumming, & Landy, 1994, p. 2270). When the cues manipulated by the experimenter are of low quality, “more weight is given to these extraneous cues, resulting in a display which appears flattened” (see also Atkins, Jacobs, & Knill, 2003).

Watt et al. (2005) examined this issue in greater detail and proposed that depth flattening in computer displays is caused by the inconsistent information provided by accommodation and by the retinal blur gradient. In computer-generated displays, in fact, the focal distance is fixed (because the images are presented on the CRT monitor) and, consequently, the blur variation in the retinal image is consistent with the distance of the CRT monitor and not with the distances of the different points of the simulated 3D object. In a real scene, instead, the retinal blur varies because the points in the scene are at different distances with respect to the eye's focal distance: the retinal image is sharpest for objects at the focal distance and blurred for points nearer and farther away. A second cue to flatness is provided by accommodation. Ciliary muscle contraction is responsible for the increased refractive power of the lens during accommodation and it is used to minimize the blur at the fixation point. During the scanning of a real scene, the focal distance must be adjusted depending on the depth variation of the 3D object. In a computer simulation, conversely, the focal and the represented distances do not vary appropriately and this can signal the flatness of the display. In a further investigation, Hoffman, Girshick, Akeley, and Banks (2008) suggested that perceived depth distortions in computer displays are reduced (or disappear) when focus cues are correct. Other cues to flatness are also a visible frame around a 3D scene (Eby & Braunstein, 1995), the monocular movement parallax, the pixellation, or any marks on the screen.[2]

1.3. Cue combination and cues to flatness

Depth flattening is easily accounted for by the linear model of cue combination (*e.g.*, Adams & Mamassian, 2004). Consider a cue-combination experiment with motion **v** and disparity **d** cues. The cues are presented either in isolation or together. In order to properly describe the information presented on a flat monitor, Eq. (2) must also include a term modeling the cues to flatness. According to Adams and Mamassian (2004), cues to flatness (**f**) can be combined in a single Gaussian distribution centered at zero depth: $p(\mathbf{f}|\mathbf{z})$. Cues to flatness can be considered as additional cues that are in conflict with the depth cues. The following equations model optimal cue combination of disparity, motion, and disparity + motion cues, together with the cues to flatness:

$$p(\mathbf{d},\mathbf{f}\,|\,\mathbf{z}) = p(\mathbf{d}\,|\,\mathbf{z};\mathbf{p}_d)p(\mathbf{f}\,|\,\mathbf{z}) \tag{5}$$

$$p(\mathbf{v},\mathbf{f}\,|\,\mathbf{z}) = p(\mathbf{v}\,|\,\mathbf{z};\mathbf{p}_v)p(\mathbf{f}\,|\,\mathbf{z}) \tag{6}$$

$$p(\mathbf{d},\mathbf{v},\mathbf{f}\,|\,\mathbf{z}) = p(\mathbf{d},\mathbf{v}\,|\,\mathbf{z})p(\mathbf{f}\,|\,\mathbf{z}). \tag{7}$$

Consider first the situation in which only disparity and flatness cues are present (Fig. 1, top panel). The likelihood function $p(\mathbf{d},\mathbf{f}|\mathbf{z})$ is centered at a value which lies somewhere in between the depth estimate from disparity and the depth estimate from the flatness cues. The position of the center of this distribution depends on the relative strength of the flatness cues and the disparity cue. In Fig. 1, the spread of $p(\mathbf{f}|\mathbf{z})$ (black curve) is larger than the spread of $p(\mathbf{d}|\mathbf{z})$ (red line), thus the combined likelihood $p(\mathbf{d},\mathbf{f}|\mathbf{z})$ (blue) is centered at a value closer to the true depth (100 mm). The estimate resulting from the combination of the disparity and the flatness cues, however, results in an under-estimation of depth (80 mm). In Fig. 1, the estimate from motion information is less reliable (middle panel, red curve) than that resulting from disparity information. The depth underestimation, therefore, is even larger than before (66 mm). But what happens when both disparity and motion cues are presented together? The spread of $p(\mathbf{d},\mathbf{v}|\mathbf{z})$ is smaller then the spread of either $p(\mathbf{d}|\mathbf{z};\mathbf{p}_d)$ or $p(\mathbf{v}|\mathbf{z};\mathbf{p}_v)$. As a consequence, the disparity + motion estimate is less affected by the flatness cues and it results in a smaller depth underestimation (85 mm; Fig. 1, bottom panel, blue curve). In conclusion, according to linear cue combination, the cues to flatness produce (a) a depth underestimation in both single-cue and multiple-cue displays, and (b) a larger amount of perceived depth in multiple-cue than in single-cue displays.

1.4. Rationale of the experiments

The purpose of the present investigation was to test the hypotheses that (*i*) the depth underestimation found with single-cue stimuli, and (*ii*) the "paradoxical" result reported by Domini et al. (2006) can be both explained, within the linear cue combination framework, by the flatness cues of computer displays. To test these hypotheses, we replicated a traditional cue-combination experiment (which isolates disparity and motion cues) by using physical stimuli. We reasoned as follows. If the same results are found with physical stimuli as well as with computer displays, then one must conclude that the distortions of perceived depth that have been described above cannot be attributed to the flatness cues, because cues to flatness are absent within natural viewing conditions.

2. Experiment 1

We replicated the cue-combination experiment of Domini and Caudek (2011) by using physical stimuli. Two vertical rods were placed on a rotating platform (see Fig. 2) and participants were asked to set the platform's position so that the perceived depth separation (Δz) between the two rods appeared to be equal to their horizontal separation (Δx). The task was performed by providing participants with different depth cues: motion-only, disparity-only, and disparity + motion. In the motion-only condition, participants viewed the stimulus monocularly, while the rods were oscillating about a vertical axis. In the disparity-only condition, participants binocularly viewed the stationary stimulus. In the combined condition, participants binocularly viewed the oscillating rods. Consistent with what was found with virtual stimuli by Domini and Caudek (2011), in absence of cues to flatness we expected that (*i*) both single-cue and combined-cue stimuli are perceived in a non-veridical manner, and (*ii*) combined-cue stimuli elicit a larger amount of perceived depth than the single-cue stimuli.

2.1. Method

2.1.1. Apparatus and stimuli

The stimulus consisted of two 0.1 cm × 2.0 cm vertical lines separated by a lateral distance of 121 mm (see Fig. 2). The stimuli were created by cutting vertical slits into an opaque black paper mask

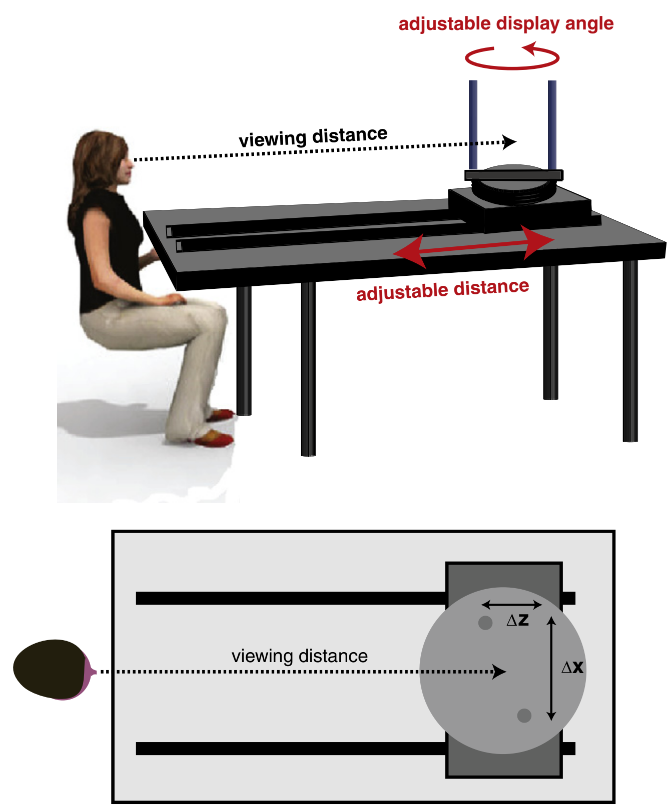

Fig. 2. Schematic representation of the side view (top panel) and of the top view (bottom panel) of the apparatus of Experiment 1. The observer's task was to equate the perceived distances Δx and Δz.

[2] Several methods can be used to weaken the effects of the flatness cues in a 3D scene. This can be done, for example, by minimizing the value of information about distance derived from accommodation through an artificial pupil (Kubovy, 1986), by synoptic or oblique viewing (Koenderink et al., 1994), or by the administration of atropine, a treatment that has the effect of dilating the pupil and paralyzing accommodation.

and mounting a thin diffuse white sheet over the slits. An Ultra Violet Light lamp, with wavelengths no longer than 370 nm, was mounted behind the black mask so that, when viewed from the front, only the two diffuse white slits were illuminated. The black light was used to make the two white strips glow in the dark. While the stimuli were composed of a solid sheet slanted around a vertical axis, the backlighting was covered and the rest of the room dark so that no slanted surface was actually visible. The stimulus rested on a motor driven platform which rotated around the vertical axis.

The rotating stimulus-platform was placed on a track that could be translated along the *z*-axis over a range of 40 to 210 cm. Distances were measured to the axis of rotation which was centered between the two vertical slits. The stimuli were presented within a viewing-box and they were lit by dark light in an otherwise dark room.

The apparatus was endowed with mechanical shutters, which allowed to intermix, across trials, binocular or monocular vision. Informal reports collected at the end of the experiment indicated the participants could not discriminate the trials in which the shutter was closed for one eye, from those in which it was opened.

Stimuli were randomly presented at distances of 60, 100, and 200 cm, under monocular-motion, static-binocular, and combined binocular-motion conditions, with counterbalancing across both positive and negative stimulus angles.[3] Binocular/monocular viewing was automated by shutters attached to HiTech™ servomotors which were controlled by a Dell™ Dimension 8100 PC running a Python-based experimental script. A Phidgets™ USB Servomotor controller and drivers were used to interface the PC to all servomotors. The same hardware was also calibrated to control stimulus motion, which consisted of a triangle wave oscillation at 8°/s around the vertical axis of the stimulus. The initial orientation of the stimuli in each trial was randomly selected in the range of 20° to 60° from frontal–planar.

In the "motion-only" condition, vision was monocular and the 3D stimulus structure oscillated about a vertical axis (see Fig. 2). In the "disparity-only" condition, vision was binocular and the 3D stimulus structure was kept stationary. In the combined "motion-disparity" condition, vision was binocular and the 3D stimulus structure oscillated about a vertical axis (see also Bradshaw et al., 2000).

2.1.2. Participants

Eighteen observers from the Brown University student community participated in the experiment. All observers were naïve to the specific hypothesis in the experiment and unaware of the nature of the stimuli being presented.

2.1.3. Procedure

The method of adjustments was used to set the perceived Δx equal to the perceived Δz (see Fig. 2). Participants indicated with a keypress which stimulus dimension (Δx or Δz) appeared to be larger. After a selection was made by pressing a button on the keyboard, the shutters closed, and the stimulus was rotated in a direction to counter the observer's judgment. While the stimuli were being reoriented, the shutters were closed and participants were not able to observe the stimuli. A mouse key press ended the trial. Participants were seated in a dark room and their heads were stabilized using a chin-rest to reduce head movement. The experimenter was present in the room throughout trials in order to adjust the stimulus platform to the randomized distance.

2.1.4. Design

A 3 (distance: 60, 100, 200 cm) $\times$ 3 (cues: motion-only, disparity-only, motion-disparity) within-subject design was used.

[3] In half of the trials, the vertical slits were coplanar to a solid sheet tilted by a negative angle with respect to the frontal-parallel plane; in the other half of the trials the "stimulus angle" was positive.

2.2. Results

The results are shown in Fig. 3 in terms of the bias of the observers' settings, that is, in terms of the difference between Δx and Δz of the physical stimulus, when observers perceived these dimensions to be equal. Veridical performance (bias equal to zero) corresponds to $\Delta z = \Delta x = 85.60$ mm. A positive bias means depth overestimation, because a smaller amount of physical depth is required to match the perceived amount Δx ($\Delta z < \Delta x$).

Inferential statistics on the observers' responses are based on a Linear Mixed-Effects (LME) model specifying participants as a random factor to control for their associated intraclass correlation.[4] The dependent variable is the amount Δz of the physical stimulus that was set by the observers when they perceived Δz to be equal to Δx (see Fig. 2). The observed mean together with the 95% *C.I.* is reported in Table 1.

The interaction between distance and cues is significant, $\chi^2_2 = 23.073$, $p < .001$. As shown in Fig. 3, the disparity-only and the disparity + motion trials are affected by the viewing distance, whereas the motion-only trials are not, $t_{322} = -0.703$, $p > .05$.

By comparing the 95% *C.I.* of Table 1 with the "true" value of 85.60 mm, we can see that the settings of Δz in the disparity-only trials are compatible with a veridical response at the viewing distances of 60 and 100 cm; at 200 cm, instead, depth from disparity is under-estimated. The settings of Δz in the disparity + motion trials suggest that, at 60 cm, depth is over-estimated whereas, at 100 and 200 cm, the observers' settings are compatible with a veridical response.

In conclusion, the results of Experiment 1 show that the "paradoxical" result of Domini et al. (2006) cannot be attributed to the presence of cues to flatness: with physical stimuli (*i.e.*, in absence of flatness cues), observers report larger amounts of perceived depth when motion is added to disparity information. Interestingly, in our data, such an increase of perceived depth did not necessarily improve the accuracy of the perceptual estimate.

3. Experiment 2

The purpose of Experiment 2 was to ask whether virtual stimuli elicit a smaller amount of depth than an equivalent physical stimulus. In Experiment 2, we replicated the design of Experiment 1 by using virtual stimuli identical in shape and dimensions to those used in Experiment 1; the stimuli were presented with a haploscope, with vergence information always consistent with simulated distance.

4. Method

4.1. Participants

Four observers from the Brown University student community participated in the experiment. All observers were naïve with respect to the hypothesis under investigation.

4.2. Apparatus and stimuli

The stimuli were simulated through a custom C++ program using OpenGL graphic routines and were presented on a custom stereo haploscope. Inter-pupillary distance for each participant was measured and used to adjust the stereoscope viewing geometry so that vergence angle was consistent with the simulated disparity in binocular

[4] We used the `lmer` program (`lme4` package) in the `R` system for statistical computing (queryR Development Core Team, 2010). As indicated by Baayen (2008), *p* values and confidence intervals are generated from the posterior distribution of parameter estimates with Markov Chain Monte Carlo methods, using the `mcmcsamp` program in the `lme4` package with default specifications (*e.g.*, $n = 1000$ samples; locally uniform priors for fixed effects; locally non-informative priors for random effects).

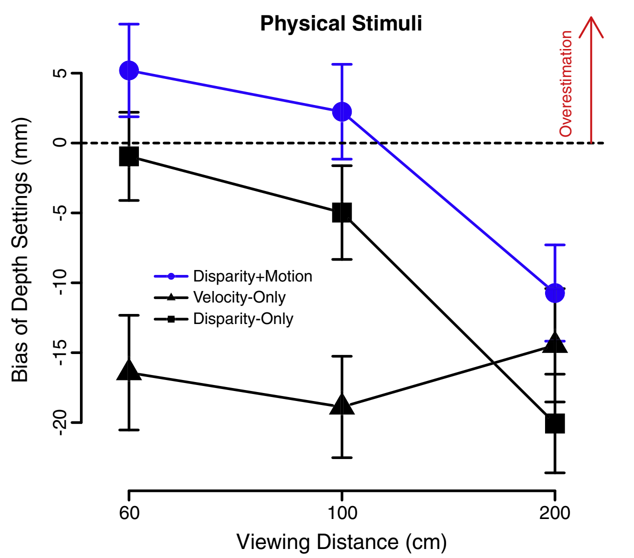

Fig. 3. Experiment 1. Average differences between the Δz and Δx across observers when the two distances are judged to be perceptually equal, as a function of viewing distance and cues. A positive bias indicates depth over-estimation and zero means veridical performance. Vertical bars represent ± 1 SE of the mean. Inferential analysis indicates that, when considering solely the disparity-only and the disparity + motion trials, the distance × cues interaction is not significant, $\chi^2_1 = 0.04$, $p > .05$ For the disparity + motion trials, the Δz setting is smaller, on average, by 3.88 mm in comparison to the disparity-only trials, $p < .001$, 95% *C.I.* [2.28, 5.47]. The main effect of distance is significant, $t_{645} = 8.78$, $p < .001$, indicating that the Δz settings increase with the viewing distance: the same amount of perceived depth requires a larger amount of physical depth. Motion-only trials, instead, are not affected by viewing distance.

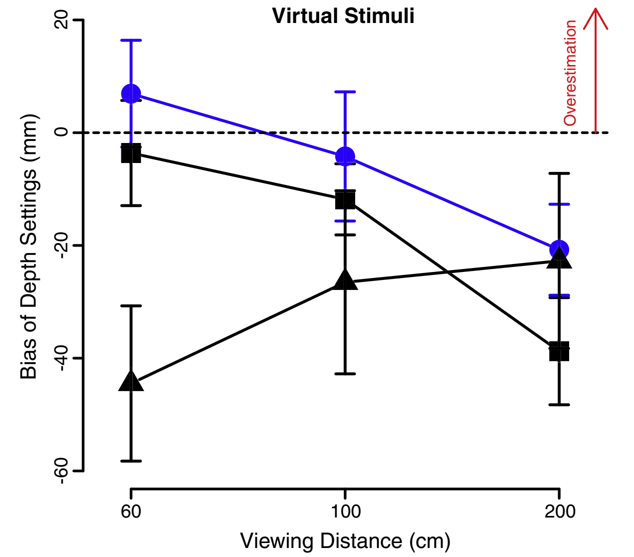

Fig. 4. Experiment 2. Average differences between the Δz and Δx across observers when the two distances are judged to be perceptually equal, as a function of viewing distance and cues. A positive bias indicates depth over-estimation and zero means veridical performance. Vertical bars represent ± 1 SE of the mean. In a LME model with the disparity-only and disparity + motion conditions, the distance × cues interaction is not significant, $\chi^2_1 = 0.18$, $p > .05$. In comparison to the disparity-only trials, Δz for the disparity + motion trials is smaller, on average, by 5.67 mm, $p < .001$, 95% *C.I.* [1.06, 10.08]. The main effect of distance is significant, $t_{45} = 5.67$, $p < .001$, indicating that the Δz settings increase with viewing distance.

conditions. For monocular presentations, observers wore an eye-patch over their left eye.

4.3. Design

A 3 (distance: 60, 100, 200 cm) × 3 (cues: motion-only, disparity-only, motion-disparity) within-subject design was used.

4.4. Procedure

In each trial, participants indicated whether the Δx stimulus dimension appeared to be larger than Δz. An adaptive staircase method was used which incorporated four interleaved staircases. Staircases ended after six reversals and each session ended once the last staircase was completed.

The data were fitted with cumulative Gaussians free to vary in position (PSE) and slope (JND) using the software package psignifit (Wichmann & Hill, 2001). From the fitted psychometric functions, we determined the point of subjective equality (PSE). The PSE corresponds to the Δz magnitude at which the psychometric function reaches 0.5, that is, the PSE is the point at which the Δz stimulus dimensions were equally often judged to be equal to the Δx dimension (see Fig. 2). Eighteen PSEs were computed for each participant.

4.5. Results

In a LME analysis on the data of both experiments, with cues, distance, and experiment as fixed effects, the 3-way interaction is not significant, $\chi^2_5 = 3.65$, $p = .60$, indicating that similar cues × distance interaction in both experiments, $\chi^2_2 = 30.20$, $p < .001$. To perform the task for the physical stimuli, observers set Δz to lower values than for the virtual stimuli. In order for Δz to appear equal to Δx, the simulated depth of the virtual stimuli is 5.74 mm larger on average than the depth of the physical stimuli, $t_{1037} = 2.52$, $p < .02$.

Fig. 4 shows the biases of the observers' settings averaged across participants. The bias is coded as in Fig. 3 of Experiment 1. The observed mean together with the 95% *C.I.* is reported in Table 2.

By using as the dependent variable the amount Δz of the physical stimulus that was set by the observers when they perceived Δz to be equal to Δx, the interaction between distance and cues is significant, $\chi^2_2 = 15.71$, $p < .001$. As in Experiment 1, this interaction indicates that the disparity-only and the disparity + motion trials are affected by the viewing distance, whereas the motion-only trials are not, $t_{70} = -1.38$, $p > .05$.

Table 1
Experiment 1. Average Δz settings (mm), when Δz was perceived to be equal to Δx, as a function of viewing distance and cues (D + M = Disparity + Motion; D = Disparity-Only; M = Motion-only). Veridical performance corresponds to $\Delta z = \Delta x = 85.60$ mm. Lower values indicate depth under-estimation. 95% confidence intervals are reported in parentheses. The confidence intervals are generated from the posterior distribution of parameter estimates with Markov Chain Monte Carlo methods.

Cues	Viewing distance		
	60 cm	100 cm	200 cm
D + M	81.16 [76.30, 84.91]	82.58 [78.72, 87.25]	88.83 [83.94, 93.77]
D	84.42 [79.58, 89.08]	86.20 [82.23, 91.28]	92.95 [87.94, 97.55]
M	90.62 [87.74, 96.49]	92.32 [88.01, 96.54]	89.80 [87.86, 96.86]

Table 2
Experiment 2. Average Δz settings (mm), when Δz was perceived to be equal to Δx, as a function of viewing distance and cues (D + M = Disparity + Motion; D = Disparity-Only; M = Motion-only). Veridical performance corresponds to $\Delta z = \Delta x = 85.60$ mm. Lower values indicate depth under-estimation. 95% confidence intervals are reported in parentheses. The confidence intervals are generated from the posterior distribution of parameter estimates with Markov Chain Monte Carlo methods.

Cues	Viewing distance		
	60 cm	100 cm	200 cm
D + M	81.10 [71.80, 92.11]	86.26 [75.74, 94.57]	94.62 [84.75, 105.95]
D	86.43 [75.77, 96.44]	90.85 [81.19, 100.88]	101.70 [91.37, 112.99]
M	102.67 [89.94, 110.25]	94.87 [88.27, 107.12]	93.57 [82.27, 103.62]

The 95% *C.I.* of Table 2 indicate that, at the viewing distances of 60 and 100 cm, perceived depth is veridical; at 200 cm, depth from disparity is under-estimated. Even though the 95% *C.I.* do not provide evidence of depth over-estimation in the disparity + motion trials, the statistical analyses indicate that, on average, observers reported a larger amount of perceived depth when motion was added to disparity information (see Fig. 4).

5. General discussion

In our previous investigation with virtual stimuli, we found that perceived depth was larger when disparity and motion cues were shown together, with respect to when each cue was shown in isolation (Domini et al., 2006). In Experiment 2 of the present investigation, we replicated this result. As explained in Fig. 1, linear cue combination can explain this result by hypothesizing that the "single-cue" estimates are actually "combined-cue" estimates in which the amount of perceived depth is found by multiplying the likelihood function of the disparity (or motion) cue by the likelihood function of the residual cues associated with using a flat monitor (*e.g.*, accommodation, vergence, and blur cues — see Adams & Mamassian, 2004; Watt, Banks, Ernst, & Zumer, 2002; Watt, Akeley, & Banks, 2003) and by maximizing the resulting likelihood function. In the same way, the disparity + motion condition should be described as disparity + motion + flatness. According to this reasoning, the depth estimate in the disparity + motion + flatness condition is larger than in any of the disparity + flatness or motion + flatness conditions because, when both depth cues are present, the MAP estimate is affected by the likelihood function of the flatness cues in a smaller measure than when only a single depth-cue is present. Linear cue combination, therefore, is consistent with the increase of perceived depth when motion is added to disparity information, but only in the case of computer simulations (when cues to flatness are present), not in the case of physical stimuli (when cues to flatness are absent).

As discussed in the Introduction, the model of linear cue integration can explain this finding by hypothesizing a reduced effect of the residual cues associated with using a flat monitor (*e.g.*, accommodation, vergence, and blur cues — see Adams & Mamassian, 2004; Watt et al., 2002, 2003). To determine whether this is indeed the case, we replicated a classical cue-combination experiment by using a physical object, where accommodation, vergence, and blur cues are always consistent with the depth of the stimulus display. In Experiment 1, we found that perceived depth increased when adding more cues, also with a physical stimulus. This result, therefore, cannot be explained by the reduced effect of the (absent) cues to flatness.[5]

The only possibility for linear cue combination to account for depth *over-estimation* is to abandon the assumption of *unbiased* estimates from single cues. If such assumption is violated, however, linear cue combination becomes meaningless. In the presence of biased estimates, in fact, the criterion of maximum reliability does not guarantee that the least-biased estimate will receive the largest weight. Therefore, we loose any rational justification for choosing maximum reliability as the goal of cue combination (see also Domini & Caudek, 2011).

An alternative explanation of the present result is provided by the cue-combination model proposed by Domini and Caudek (Caudek, Fantoni, & Domini, 2011; Di Luca, Domini, & Caudek, 2007; Di Luca, Domini, & Caudek, 2010; Domini et al., 2006; Domini & Caudek, 2009; Domini & Caudek, 2010; Fantoni, Caudek, & Domini, 2010; Foster, Fantoni, Caudek, & Domini, 2011; Tassinari, Domini, & Caudek, 2008). This approach differs from linear cue combination in an important respect: according to the proposal of Domini and Caudek, the *direct* (*i.e.*, retinal) information about 3D shape provided by different cues is combined *before* a metric interpretation is assigned to each signal in isolation. Specifically, we hypothesized that a *piecewise-affine* estimate of the true depth is derived from each retinal cue in isolation (Caudek & Rubin, 2001; Domini, Caudek, & Richman, 1998). These *piecewise-affine* estimates are then combined into an optimal weighted sum that maximizes the Signal to Noise Ratio (SNR).

There are two aspects that are relevant for the present discussion: (1) the SNR of the combined signal increases as more cues are added; (2) this combined retinal signal is scaled by assigning larger *metric* depth values to larger SNRs (Domini et al., 2006). The approach proposed by Domini and Caudek, therefore, is consistent with the present data because it guarantees larger depth estimates with more cues, without requiring that more information increases veridical perception.

Consistent with previous literature, the comparison of the results of Experiments 1 and 2 indicates that computer-generated displays support smaller amounts of perceived depth than physical stimuli, when equivalent depth information is provided in the two cases. According to linear cue combination, depth under-estimation in computer-displays is produced by (a) the low reliability of the simulated depth cues, and (b) the cue-conflict resulting from the flatness cues. Both these aspects contribute to decrease the amount of perceived depth. The approach proposed by Domini and Caudek explains the different amounts of perceived depth elicited by virtual and physical stimuli in a different manner. According to IC, the flatness cues in a virtual display are ignored, unless they produce non-zero retinal gradients. The different depth magnitudes that are perceived in matched physical and virtual displays are not attributed to a decrease of perceived depth caused by the flatness cues, but rather to an increase in the amount of perceived depth caused by the presence of additional depth cues within natural viewing conditions (*e.g.*, among the others, the blur gradient). If all depth cues are combined together, the SNR of the combined signal will be larger for the physical than for the virtual stimuli (because a larger number of depth cues are combined in the first than in the second case). According to IC, perceived depth depends on the SNR of the combined signal. Therefore, we should expect a larger amount of perceived depth for the physical rather than for the matched virtual displays.

5.1. Conclusions

We found that (a) perceived depth from disparity + motion was larger than perceived depth from either single-cue alone, and (b) the addition of motion to disparity information produced depth over-estimation, even if disparity-only information supported veridical depth perception from physical stimuli. The present results cannot be explained by a cue averaging strategy that attributes depth mis-estimations to the residual flatness cues of the monitor's screen — cues to flatness, in fact, were absent in the physical stimuli of Experiment 1. The Intrinsic Constraint model (Domini et al., 2006) is one of the possible models that are consistent with the results obtained with both the physical (Experiment 1) and the virtual stimuli (Experiment 2) of the present investigation.

Appendix A. Information provided by motion and disparity signals

A.1. Retinal disparities

If the object is small enough, then the relationship between the retinal disparities and the 3D location of the projecting features can be characterized by a simple equation. Let us term d_i the horizontal disparity of the i^{th} feature point. Assume that the object is placed at a

[5] Some researchers have hinted that depth perception might also be affected by a prior for frontal-parallel (Adams & Mamassian, 2004). In terms of Fig. 1, one can imagine that the Gaussian centered at zero might represent not the likelihood function of the cues to flatness, but rather a prior for frontal-parallel. However, given that to our knowledge there are not published studies suggesting the biological plausibility for a prior for frontal-parallel, we don't pursue further such argument here.

fixation distance z_f from the observer and that z_i is the relative depth of the i^{th} feature point with respect to the fixation point F. For an object subtending a small visual angle, Eq. (8) defines the relationship among the *image signals* $\mathbf{d}=(d_1,\ldots,d_n)$, the *depth map* $\mathbf{z}=(z_1,\ldots,z_n)$ and the *scene parameters* $\mathbf{p}_d=(z_f)$:

$$d_i \approx IOD\frac{z_i}{z_f^2} + \varepsilon_{d_i}, \tag{8}$$

where IOD is the observer's interocular distance and ε_{d_i} is a Gaussian random variable representing the noise in the disparity measurements.

A.2. Retinal velocities

Let us term v_i the retinal velocity of the i^{th} feature point. If the observer translates horizontally by T_x while fixating the object, or the object rotates about a vertical axis centered at the fixation point F, then the pattern of retinal velocities is

$$v_i \approx \omega\frac{z_i}{z_f} + \varepsilon_{v_i}. \tag{9}$$

If the observer moves about an otherwise stationary object, then $\omega = \frac{T_x}{z_f}$. ε_{v_i} is a Gaussian random variable representing the noise in the velocity measurements. The *scene parameters* are $\mathbf{p}_v=(z_f,T_x)$.

References

Adams, W. J., & Mamassian, P. (2004). Bayesian combination of ambiguous shape cues. *Journal of Vision*, *4*, 921–929.

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision*, *9*, 1–14.

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology*, *19*, 452–458.

Arnold, D. H., Tear, M., Schindel, R., & Roseboom, W. (2010). Audio–visual speech cue combination. *PloS One*, *5*(4), e10217.

Atkins, J. E., Jacobs, R. A., & Knill, D. C. (2003). Experience-dependent visual cue recalibration based on discrepancies between visual and haptic percepts. *Vision Research*, *43*, 2603–2613.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R.* : Cambridge University Press.

Bradshaw, M. F., Parton, A. D., & Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. *Vision Research*, *40*, 3725–3734.

Braunstein, M. L., & Tittle, J. S. (1988). The observer-relative velocity field as the basis for effective motion parallax. *Journal of Experimental Psychology. Human Perception and Performance*, *14*, 582–590.

Bruno, N., & Cutting, J. E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology. General*, *117*, 161–170.

Buckley, D., & Frisby, J. (1993). Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges. *Vision Research*, *33*, 919–933.

Caudek, C., & Domini, F. (1998). Perceived orientation of axis of rotation in structure-from-motion. *Journal of Experimental Psychology. Human Perception and Performance*, *24*, 609–621.

Caudek, C., Domini, F., & Di Luca, M. (2002). Short-term temporal recruitment in structure from motion. *Vision Research*, *42*, 1213–1223.

Caudek, C., Fantoni, C., & Domini, F. (2011). Bayesian modelling of perceived plane orientation from actively-generated and passively-observed optic flow. *PloS One*, *6*(4), e18731.

Caudek, C., & Proffitt, D. R. (1993). Depth perception in motion parallax and stereokinesis. *Journal of Experimental Psychology. Human Perception and Performance*, *19*, 32–47.

Caudek, C., & Rubin, N. (2001). Segmentation in structure from motion: Modeling and psychophysics. *Vision Research*, *41*, 2715–2732.

Cochran, W. G. (1937). Problems arising in the analysis of a series of similar experiments. *Supplement to the Journal of the Royal Statistical Society*, *4*, 102–118.

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., & Thompson, W. B. (2005). The influence of restricted viewing conditions on egocentric distance perception: implications for real and virtual environments. *Perception*, *34*, 191–204.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In W. Epstein, & S. J. Rogers (Eds.), *Perception of space and motion* (pp. 69–117). San Diego, CA: Academic Press.

Dewing, K., & Ernst, M. O. (2006). Integration of force and position cues to haptic shape. *Cognitive Brain Research*, *1078*, 92–100.

Di Luca, M., Domini, F., & Caudek, C. (2004). Spatial integration in structure from motion. *Vision Research*, *44*, 3001–3013.

Di Luca, M., Domini, F., & Caudek, C. (2007). The relation between dis-parity and velocity signals of rigidly moving objects constrains depth order perception. *Vision Research*, *47*, 1335–1349.

Di Luca, M., Domini, F., & Caudek, C. (2010). Inconsistency of perceived 3D shape. *Vision Research*, *21*, 1519–1531.

Domini, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: a new theory. *Trends in Cognitive Sciences*, *7*, 444–449.

Domini, F., & Caudek, C. (2003). Recovering slant and angular velocity from a linear velocity field: Modeling and psychophysics. *Vision Research*, *43*, 1753–1764.

Domini, F., & Caudek, C. (2009). The intrinsic constraint model and Fechnerian sensory scaling. *Journal of Vision*, *9*(2), 1–15 (25).

Domini, F., & Caudek, C. (2010). Matching perceived depth from disparity and from velocity: Modeling and psychophysics. *Acta Psychologica*, *133*, 81–89.

Domini, F., & Caudek, C. (2011). Combining image signals before 3D reconstruction: The Intrinsic Constraint Model of cue integration. In In J. Trommershäuser, M. S., Landy, & K. Körding (Eds.), *Sensory Cue Integration* (pp. 120–143). New York: Oxford University Press.

Domini, F., Caudek, C., & Proffitt, D. R. (1997). Misperceptions of angular velocities influence the perception of rigidity in the kinetic depth effect. *Journal of Experimental Psychology. Human Perception and Performance*, *23*, 1111–1129.

Domini, F., Caudek, C., & Richman, S. (1998). Distortions of depth-order relations and parallelism in structure from motion. *Perception & Psychophysics*, *60*, 1164–1174.

Domini, F., Caudek, C., & Skirko, P. (2003). Temporal integration of motion and stereo cues to depth. *Perception & Psychophysics*, *65*, 48–57.

Domini, F., Caudek, C., & Tassinari, H. (2006). Stereo and motion information are not independently processed by the visual system. *Vision Research*, *46*, 1707–1723.

Domini, F., Caudek, C., Turner, J., & Favretto, A. (1998). Discriminating constant from variable angular velocities in structure from motion. *Perception & Psychophysics*, *60*, 747–760.

Domini, F., Vuong, Q. C., & Caudek, C. (2002). Temporal integration in structure from motion. *Journal of Experimental Psychology. Human Perception and Performance*, *28*, 816–838.

Durgin, F. H., Proffitt, D. R., Olson, T. J., & Reinke, K. S. (1995). Comparing depth from motion with depth from binocular disparity. *Journal of Experimental Psychology. Human Perception and Performance*, *21*, 679–699.

Eby, D. W., & Braunstein, M. L. (1995). The perceptual flattening of three-dimensional scenes enclosed by a frame. *Perception*, *24*, 981–993.

Ellis, S. R., Smith, S. R., Grunwald, A. J., & McGreevy, M. W. (1991). *Direction judgement error in computer generated displays and actual scenes.* Pictorial communication in virtual and real environments. London: Taylor and Francis.

Enright, J. T. (1991). Paradoxical monocular stereopsis and perspective vergence. In S. R. Ellis (Ed.), *Pictorial communication in virtual and real environments* (pp. 567–576). (2nd ed.). London: Taylor & Francis.

Fantoni, C., Caudek, C., & Domini, F. (2010). Systematic distortions of perceived planar surface motion in active vision. *Journal of Vision*, *10*, 1–20.

Foster, R., Fantoni, C., Caudek, C., & Domini, F. (2011). Integration of disparity and velocity information for haptic and perceptual judgments of object depth. *Acta Psychologica*, *136*, 300–310.

Frisby, J. P., Buckley, D., & Duke, P. A. (1996). Evidence for good recovery of lengths of real objects seen with natural stereo viewing. *Perception*, *25*, 129–154.

Frisby, J., Buckley, D., & Horsman, J. (1995). Integration of stereo, texture, and outline cues during pinhole viewing of real ridge-shaped objects and stereograms of ridges. *Perception*, *24*, 181–198.

Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1997). Computational models of sensorimotor integration. In P. G. Morasso, & V. Sanguineti (Eds.), *Self-organization, computational maps, and motor control* (pp. 117–147). Amsterdam: North-Holland.

Helbig, H. B., & Ernst, M. O. (2007). Optimal integration of shape information from vision and touch. *Experimental Brain Research*, *179*, 595–606.

Hibbard, P. B., & Bradshaw, M. F. (2003). Reaching for virtual objects: Binocular disparity and the control of prehension. *Experimental Brain Research*, *148*, 196–201.

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, *4*, 967–992.

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, *28*, 1–30.

Jacobs, R. A. (2002). What determines visual cue reliability? *Trends in Cognitive Sciences*, *6*, 345–350.

Jacobs, R. A., & Kruschke, J. K. (2011). Bayesian and computational learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 8–21.

Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereo and motion shape cues. *Vision Research*, *34*, 2259–2275.

Knapp, J. M., & Loomis, J. M. (2004). Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments*, *13*, 572–577.

Knill, D. C. (2007). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision*, *7*(7), 1–24 (5).

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference.* New York, NY: Cambridge University Press.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, *43*, 2539–2558.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1994). On so-called paradoxical monocular stereoscopy. *Perception*, *23*, 583–594.

Körding, K. (2007). Decision theory: What "should" the nervous system do? *Science*, *318*, 606–610.

Körding, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, *10*, 320–326.

Kubovy, M. (1986). *The psychology of perspective and Renaissance art.* New York: Cambridge University Press.
Landy, M. S., Maloney, M. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, *35*, 389–412.
Landy, M. S., Banks, M. S., & Knill, D. C. (2011). Ideal-observer models of cue integration. In J. Trommershäuser, M. S., Landy, & K. Körding (Eds.), *Sensory Cue Integration* (pp. 5–29). New York: Oxford University Press.
Massaro, D. W. (1988). Ambiguity in perception and experimentation. *Journal of Experimental Psychology. General*, *117*, 417–421.
Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology*, *18*, 689–693.
Rogers, S. (1995). Perceiving pictorial space. In W. Epsein, & S. Rogers (Eds.), *Perception of space and motion* (pp. 119–163). London: Academic Press.
Saunders, J. A., & Backus, B. T. (2006). The accuracy and reliability of perceived depth from linear perspective as a function of image size. *Journal of Vision*, *6*, 933–954.
Tassinari, H., Domini, F., & Caudek, C. (2008). The intrinsic constraint model for stereo-motion integration. *Perception*, *37*, 79–95.
Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments*, *13*, 560–571.
Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics*, *54*, 157–169.
Todd, J. T., Thaler, L., Dijkstra, T. M. H., Koenderink, J. J., & Kappers, A. M. L. (2007). The effects of viewing angle, camera angle, and sign of surface curvature on the perception of three-dimensional shape from texture. *Journal of Vision*, *7*(12), 1–16 (9).
Todorovic, D. (2009). The effect of the observer vantage point on perceived distortions in linear perspective images. *Attention, Perception, & Psychophysics*, *71*, 183–193.
van Ee, R., Banks, M. S., & Backus, B. T. (1999). An analysis of stereoscopic slant contrast. *Perception*, *28*, 1121–1145.
Watt, S. J., Akeley, K., & Banks, M. S. (2003). Focus cues to display distance affect perceived depth from disparity. *Journal of Vision*, *3*, 66a.
Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision*, *5*, 834–862.
Watt, S. J., Banks, M. S., Ernst, M. O., & Zumer, J. M. (2002). Screen cues to flatness do affect 3d percepts. *Journal of Vision*, *2*, 297a.
Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception & Psychophysics*, *63*(8), 1293–1313.
Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, *33*, 2685–2696.
Zalevski, A. M., Henning, G. B., & Hill, N. J. (2007). Cue combination and the effect of horizontal disparity and perspective on stereoacuity. *Spatial Vision*, *20*, 107–138.