Touch can change visual slant perception

Marc O. Ernst[1], Martin S. Banks[2] and Heinrich H. Bülthoff[1]

[1] *Max-Planck-Institute for Biological Cybernetics, Spemannstr. 38, D-72076 Tübingen, Germany*

[2] *Vision Science Program & School of Optometry, University of California, Berkeley, California 94720-2020, USA*

Correspondence should be addressed to M.O.E. (marc.ernst@tuebingen.mpg.de)

The visual system uses several signals to deduce the three-dimensional structure of the environment, including binocular disparity, texture gradients, shading and motion parallax. Although each of these sources of information is independently insufficient to yield reliable three-dimensional structure from everyday scenes, the visual system combines them by weighting the available information; altering the weights would therefore change the perceived structure. We report that haptic feedback (active touch) increases the weight of a consistent surface-slant signal relative to inconsistent signals. Thus, appearance of a subsequently viewed surface is changed: the surface appears slanted in the direction specified by the haptically reinforced signal.

Combining different sources of three-dimensional information helps the interpretation of ambiguous signals and reduces the effects of measurement noise. The method of combination has been successfully examined by using cue-conflict protocols in which signals are manipulated independently[1–5]. In one study, for example, subjects adjusted the three-dimensional shape of convex surfaces with elliptical cross-sections until they appeared cylindrical (circular cross section)[4]. The shape was specified by conflicting disparity (three-dimensional information provided by the differences in images between the two eyes) and texture gradients (three-dimensional information given by projection of a surface with statistically regular markings onto the retina). Affected by both the disparity and texture signals, the shape settings were well described by a linear weight combination rule:

$$S = w_t S_t + w_d S_d \tag{1}$$

S_t and S_d are the outputs of shape estimators with weights w_t and w_d that use texture and disparity signals, respectively. Each shape estimator may use a variety of input signals. For example, the disparity-based estimator uses inputs of horizontal disparity, vertical disparity and their gradients[1]. One cannot distinguish a change in weight from a change in estimator gain, so for our purposes a 'weight change' will refer to both possibilities. Equation 1 is a maximum-likelihood estimator if the estimators, S_t and S_d, are Gaussian distributed and statistically independent and the weights, w_t and w_d, are equal to the estimators' inverse variances, normalized to add up to one[6–8].

This linear weighting scheme suffices for understanding many phenomena in visual perception[1–5,9–11]. A statistically sensible method for estimation would give high weight to more informative estimators and low weight to uninformative ones[5], because such weighting should yield more stable percepts[5–8,12]. The weights must depend on viewing conditions because, for example, the information content of the disparity signal decreases as a function of distance[4,12]. Furthermore, experiments show that the weights vary from one individual to another for a given viewing situation. For example, some subjects consistently give more weight to disparity, whereas others preferentially weight texture gradients[4,12].

These weights affect the appearance of the visual world and the manner in which we interact with it. Here we asked how the nervous system determines weights applied to different estimators. There are at least three ways in which the weights could be determined. First, weights could be fixed for a given viewing situation and individual and not subject to change through feedback (although they may have been changeable during infancy and childhood to compensate for anatomical changes in the sensory apparatus, for example). Second, weights could be adjusted by comparing a given estimator's output with those of other estimators and with feedback from motor behavior. Third, weights could be determined directly from statistical measures of estimator outputs. For example, if the output of one estimator for a given viewing situation fluctuated less over time than that of another estimator, the former's weight could be increased relative to the latter's. We examined the first two of these possibilities. Specifically, we asked whether the weights assigned to different estimators can be changed by haptic feedback (sensation of touch generated by active, exploratory hand and finger movements) that is consistent with one estimator and not another.

In numerous reported visual–haptic interactions, visual appearance is unaffected by haptic feedback. For example, subjects grasping a square viewed through an optical device that distorts the image to make it appear rectangular see and feel the square as a rectangle[13]. In this case, perception is determined entirely by the visual information; thus it is an example of 'visual capture'[13–17]. Perhaps this finding is not surprising, because the visual stimulus clearly specifies perceived shape. Consistent with this idea, touch can affect appearance of an impoverished visual stimulus[18]. These visual–haptic studies sought concurrent perceptual effects, but we were primarily concerned with persistent effects of haptic feedback on visual perception. To increase the probability of observing a positive influence of haptic feedback, we presented visual stimuli with a range of possible interpretations (**Fig. 1**).

Many studies of visuomotor adaptation show that humans can adapt behaviorally to changes in the mapping between the environment and sensory signals[15,19–21]. Helmholtz first demonstrat-

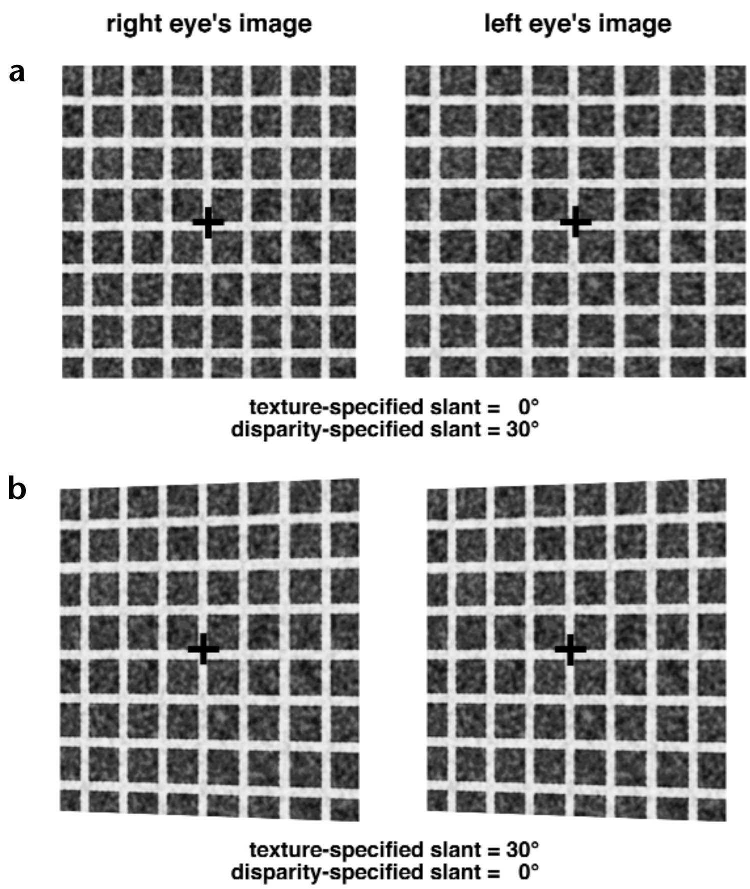

Fig. 1. Two examples of the cue-conflict stimuli. The texture signal to surface slant is the distortion of the regular grid patterns marked on the surfaces. The disparity signal to surface slant gives the right-eye minus left-eye differences in the images. The reader can visualize the stimuli approximately as they appeared by cross-fusing the two panels in each row (directing the left eye to the right half-image and the right eye to the left half-image). The black crosses, not present in the experimental stimuli, were added to help the reader fixate (one fused cross should appear when viewing the stimuli binocularly). (**a**) For a texture-specified slant of 0° and a disparity-specified slant of about 30° (at a viewing distance of 20 cm), most viewers perceive a slant between 0° and 30°, because both signals contribute to the perceived slant. (**b**) For a texture-specified slant of about 30° and a disparity-specified slant of 0° (viewing distance of 20 cm), most people again perceive a slant between 0° and 30°.

ed this plasticity of the visuomotor system by displacing the entire visual field with prismatic spectacles and studying subsequent effects on reaching and other sensorimotor behavior[22]. Initial reaches are strikingly inaccurate, but they improve rapidly. After removing the spectacles, errors are observed in the opposite direction. Thus, adaptation persists after the mapping is restored to its original state. The observed adaptation could theoretically occur in visual perception, in proprioception of body parts or in the translation from visual to motor coordinates. Determining the mechanism for adaptation proves difficult[15,20].

We used a novel approach to find that haptic feedback can alter subsequent visual percepts by changing the weights given to particular sources of visual information. Specifically, we found that haptic feedback to surface orientation subsequently affects the appearance of the surface (when haptic feedback is no longer available). We chose slant estimation because it involves a limited set of well understood environmental and sensory signals and because it allows the use of a simple psychophysical task. As such, the results are easier to interpret than previous work on visual adaptation.

Experiments were conducted in three phases, pre-test, training phase and post-test (**Fig. 2**), using a setup involving an image reflected from a mirror and a force-feedback device (**Fig. 3**). The pre- and post-tests were identical and were used to determine the weights. We presented visual planes with texture- and disparity-specified slants that differed by an angle, α. Subjects indicated the slant of the plane that was perceived as frontoparallel. The subjects' settings were quantified by β, the texture-specified angle at which subjects perceived the surface to be frontoparallel, and were used to estimate the weights assigned to texture and disparity. We looked for differences between pre- and post-test judgments due to haptic feedback provided during the training phase.

There were three experimental conditions, differing only in the training phase. In the texture-feedback condition, haptic feedback was consistent with the slant specified by the texture gradient; the non-reinforced disparity gradient varied randomly from trial to trial. In the disparity-feedback condition, haptic feedback was consistent with the slant specified by disparity, and the texture-specified slant varied randomly. The third condition was a control in which no haptic feedback was given. Subjects viewed a

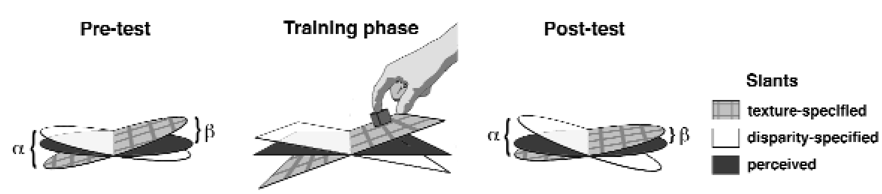

Fig. 2. Experimental design. The pre- and post-tests were purely visual tasks. The visual plane had different texture- and disparity-specified slants. The texture-specified slant is represented by the gray grids and the disparity-specified slant by the light gray planes. The angle between the two specified slants is the conflict angle α. The perceived slant is represented by dark gray planes. β is the slant specified by the texture gradient when the plane (with texture- and disparity-specified slants generally differing) was perceived as frontoparallel. The decrease in β between pre- and post-test indicates an increase in texture weight. During the pre- and post-test, the plane was clipped with a circular window with a diameter of 13°. The projected shape of the window at the cyclopean eye was consistent with the texture-specified slant. Use of the circular window reduced the probability of slant reversals[25]. The haptic training phase occurred between the pre- and post-tests and consisted of visual and haptic stimulation. The cube, surface and targets could all be seen and felt. During the training phase, the boundary of the plane was a 13° × 16° rectangle when it was normal to the line of sight. The shape of the boundary was consistent with the texture-specified slant. (The rectangular boundary increased the salience of the texture signal.)

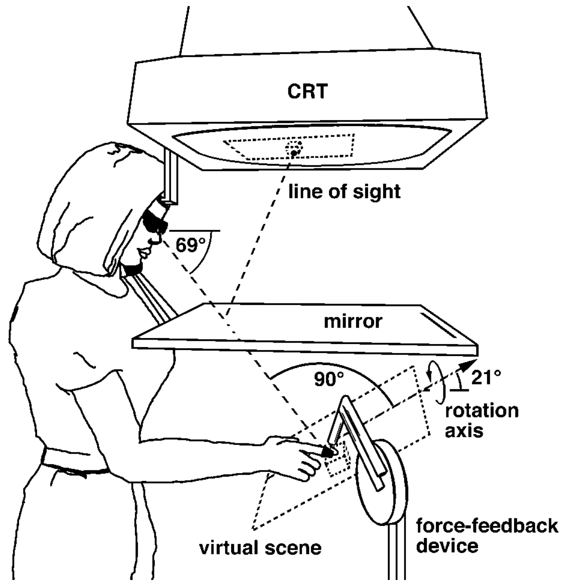

Fig. 3. Experimental setup. The visual stimulus was generated on a cathode ray tube (CRT) and viewed in a mirror that obscured the subject's hand. A dot on the screen indicated finger position. The haptic stimulus was created by a force-feedback device that had six degrees of freedom and could apply force in the three translation directions. The three-dimensional position of the fingertip was monitored, and an appropriate force was applied to the tip when it reached the position of the simulated haptic objects, creating a compelling sensation of touching a solid surface and cube. The haptic stimulus included realistic simulation of gravity's effect on the cube's motion across the plane. The haptic stimulus also simulated friction: it was low between the cube and plane and higher between the fingertip and cube. To view the stimulus, subject's line of sight was pitched 69° downward from earth horizontal. The plane's slant varied about a fixed axis pitched 21° up from horizontal. The line of sight was thus perpendicular to the rotation axis. A chin and forehead rest limited head movements.

recorded sequence of visual images seen in a previous session of the texture-feedback condition. This condition tested whether the weight changes were attributable to the visual experience rather than haptic feedback.

Here we report that haptic feedback can change subsequent visual percepts by changing the weights given to different sources of visual information.

Results

Figure 4 shows the results for the two haptic-feedback conditions and the control. The top row shows the average slant settings in the pre- and post-test. The texture-specified slant at the slant setting (β) is plotted as a function of the conflict angle (α). If settings were based entirely on the texture gradient, the data would lie on a horizontal line through 0°. If they were based entirely on disparity, the data would lie on the diagonal gray line. Settings were actually intermediate, indicating that both signals affected performance. If the visual system used sensorimotor feedback to adjust the weights assigned to different slant estimators, we should observe a change in settings due to the haptic training. Indeed, post-test settings were slightly, but consistently, different from pre-test settings. When haptic feedback was consistent with the texture-specified slant (left), the average weight assigned to the disparity-based estimator decreased from 0.70 to 0.58 (middle left); 9 of the 10 subjects showed a decrease in this weight (lower left). This means that a stimulus that appeared frontoparallel in the pre-test appeared slanted in the direction specified by texture in the post-test.

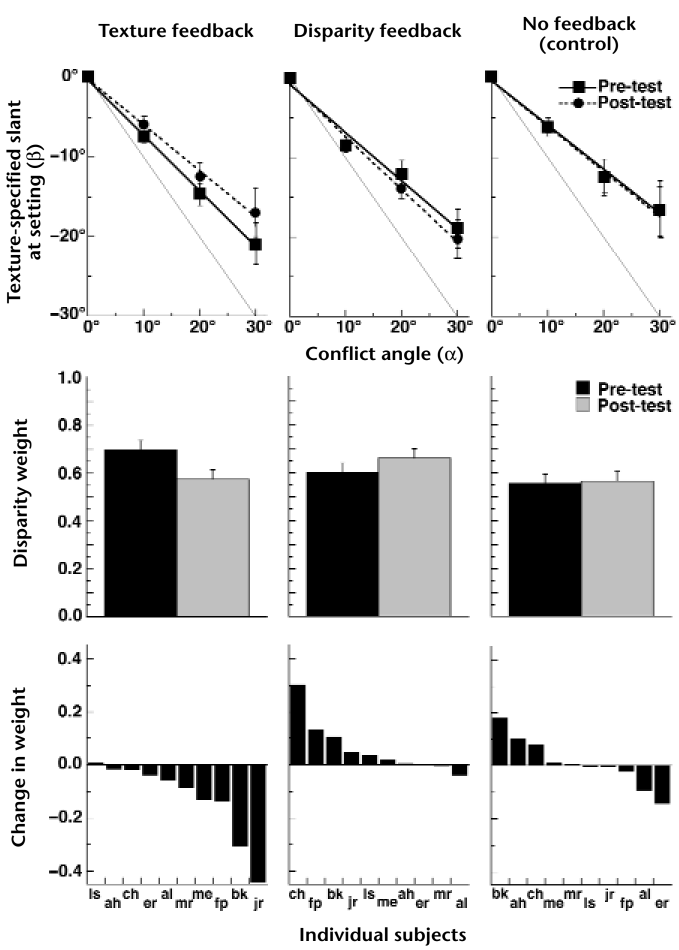

Fig. 4. Results for the haptic-feedback and control conditions. Left, results when the texture-specified slant was reinforced during haptic training. Middle column, results when the disparity-specified slant was reinforced during training. Right, results from the control condition in which no haptic feedback was provided. Top average slant settings in the pre- and post-test. The texture-specified slant at the slant setting (β) is plotted as a function of the conflict angle (α). Individual subjects' slant settings were corrected for constant baseline shifts. If settings were based entirely on the texture gradient, the data would lie on a horizontal line through 0°; if based entirely on binocular disparity, the data would lie on the diagonal gray line. Error bars represent standard errors across subjects. Middle row, average disparity weight for the pre-test (black) and post-test (gray). Weights were calculated from linear regression fits to the data in the upper row. Error bars represent regression errors. Bottom, change in weight for each subject (disparity weight in pre-test minus disparity weight in post-test). Subjects are ranked by magnitude of the effect, with subject's initials below the corresponding bar. The order along the horizontal axis—most positive weight change to most negative—was determined separately for each panel.

When haptic feedback was consistent with the disparity-specified slant (middle column), the average weight assigned to the disparity-based estimator increased from 0.60 to 0.66 (middle panel); 7 of the 10 subjects showed an increase in this weight and 2 showed a small decrease (lower middle).

A statistical test on the pre- and post-test weights revealed a significant interaction: the increase in texture weight was greater after texture-reinforced training than after disparity-reinforced training ($F_{1,9} = 10.597$, $p < 0.01$). These results show that 30–45 minutes of haptic feedback cause an upweighting of the reinforced estimator and a corresponding change in visual appearance. The weight change was small and variable across subjects.

The average weight assigned to the disparity-based estimator in the no-feedback control experiment (**Fig. 4**, right column) revealed an average disparity-based estimator weight of 0.56 in the pre-test and 0.57 in the post-test (middle right), values that are statistically indistinguishable ($F_{1,9} = 0.142$, $p = 0.715$). Thus, haptic feedback is indeed necessary for the weight change to occur.

We conducted a second control experiment (data not shown), in which the texture- and disparity-specified slants were congruent during the training phase. Haptic feedback was consistent with both visual signals, and the weights did not change. Thus, weight changes occur only when haptic feedback is consistent with one signal and not the other.

Discussion

We showed that haptic feedback can change the subsequent visual perception of surface slant. Specifically, when subjects are given haptic stimulation consistent with the texture-specified slant of a visual stimulus, their subsequent visual percepts are closer to the texture slant than they were before training. The opposite occurs when they are given haptic stimulation consistent with the disparity-specified slant. We interpreted the effect as a change in the weights given to different slant estimators, in this case, to texture-based and disparity-based estimators.

The weight changes were small. When the texture-specified slant was reinforced, the texture weight increased from pre- to post-test by an average of 40%. When the disparity slant was reinforced, the texture weight decreased by 15%. There are two obvious reasons for the small magnitudes. First, although we separately manipulated the texture- and disparity-specified slants during training, the values were correlated ($r = 0.59$). Second, calibration of a sensory system is probably best served by slow changes in response to large amounts of data. If brief exposures to new correlations among information sources resulted in large changes in calibration, a sensory system could become an unstable estimator of environmental information because its estimates would be subject to the vagaries of the particular sequence of recent events. Thus, the 15–40% change we observed is a reasonable response to 30–45 minutes of altered experience.

The perceptual effect we observed is persistent: remnants of the effect last at least 24 hours, as subjects' weights had not returned to their initial values by the second day of testing. The evidence for this is the difference in the pre-test weights in the texture- and disparity-feedback conditions (**Fig. 4**). This persistence is striking, given that subjects presumably received substantial haptic feedback from normal interaction with their environment during the intervening period.

It is interesting to consider our observations in light of known visual–haptic interactions. As stated earlier, the majority of this work failed to observe an effect of touch on visual appearance. Why did we succeed where others had not? A sensible answer comes from analyzing information provided by vision and touch for the particular stimulus properties under examination. When the attended property is size, length, curvature or angular separation, vision dominates the percept, even when touch is in conflict[13–17], demonstrating 'visual capture'. The visual system is well designed for making fine discriminations of size, length, curvature and angular separation, so the reliability of visual estimates of those properties is high. The haptic system is not capable of such fine discriminations[23], so its reliability is lower than that of the visual estimates. A sensible estimation method would give greater weight to the more reliable estimate; that is, the final estimate would be dominated by vision and barely affected by conflicting haptic information. However, when the stimulus property under examination is coarseness of texture, vision and touch influence the percept[18]. Visual and haptic discrimination of coarseness (just-discriminable percentage change) are comparable[18], so the reliabilities of visual and haptic estimates are roughly equivalent. Both estimates should, therefore, be able to influence the final estimate[20]. In the experiment reported here, the visual information was by design ambiguous (texture and disparity specified different slants), so it is reasonable to assume that additional haptic information consistent with one of the slants would be used.

In summary, we have shown that haptic feedback provides one means for adjusting the weights given different sources of visual information. An interesting byproduct is that haptic feedback consistent with one source of information can cause a change in subsequent visual appearance: a surface with conflicting disparity and texture signals will look more like the haptically reinforced slant than it did before.

Methods

Nine naive subjects and M.O.E. completed the three conditions of the experiment. All had normal or corrected-to-normal vision and good stereo vision. Informed consent was obtained.

The visual stimulus was a textured plane viewed binocularly. The texture was a regular grid mapped onto the plane and then displayed on a CRT using OpenGL routines. The stimulus was viewed stereoscopically with liquid-crystal shutter glasses. The apparent position of the plane was below the mirror (**Fig. 3**). The surroundings were dark. At the beginning of each stimulus presentation, a bright white field was presented to maintain light adaptation and make the CRT frame less visible.

The stimulus plane's slant was specified by binocular disparity and a texture gradient (**Fig. 1**). In the pre- and post-tests, the disparity- and texture-specified slants differed (conflict angle $\alpha = 0°, 10°, 20°$ or $30°$, tilt $= 0°$). Despite the differing slant specifications, the stimulus always looked like a single plane with a well-defined slant. Each stimulus was presented for 500 ms, and subjects reported whether the plane's left or right side appeared closer. The slant was varied (holding α constant) according to an adaptive staircase procedure (seven reversals) to estimate the surface slant that appeared frontoparallel (β). The estimates were determined from an average of the last four reversals. From these estimates, we calculated the weights assigned to the disparity- and texture-based slant estimators (**Fig. 4**).

The haptic-training phase occurred between the pre- and post-tests. Again, the disparity- and texture-specified slants were manipulated separately. During training, haptic stimulation was provided along with the visual displays. The haptic stimulus was created by a haptic force-feedback device (PHANToM™ Model 1.5, http://www.sensable.com). The virtual haptic stimulus consisted of the plane and a small cube lying flat on the plane. The right index finger was placed in a thimble-like holder attached to the device. To create the haptic sensation, the three-dimensional position of the right index finger was monitored in real time. When the fingertip reached the simulated haptic plane or cube, the device applied an appropriate force on the fingertip, creating a compelling sensation of touching solid objects (the stationary plane and movable cube)[23]. By these means of creating stimuli, we could independently manipulate visual and haptic stimulation.

During haptic training, a subject used the index finger to move the cube along the plane to the target by pressing down on top of the cube and moving the finger in the desired direction as rapidly and accurately as possible. All subjects experienced a convincing sensation of pushing a real cube along a real slanted surface. We chose the task of moving a small cube along the surface because it is visually and haptically demanding, thus requiring more attention than simply touching a surface. The fingertip's three-dimensional position was represented by a small dot. The cube, plane, targets and dot were always visible. A new stimulus with a new cube, slants and target positions appeared after successful target acquisition.

The three experimental conditions differed only in the relationship between haptic feedback and visual stimulation during the training phase. In the texture-feedback condition, subjects received haptic feedback consistent with the plane's texture-specified slant. In other words, the haptic plane had the three-dimensional coordinates of the plane's texture slant, and the haptic cube had the coordinates of a cube lying on that plane. In the disparity-feedback condition, they received haptic feedback consistent with the disparity-specified slant. In the no-feedback condition, subjects viewed the same sequence of visual displays seen in the texture-feedback condition, but received no haptic feedback. To sustain attention on the visually specified slant, they had to indicate the perceived slant and time of target acquisition.

In the two feedback conditions, the slant specified by the reinforced visual signal and the haptic stimulus was ±10° or ±20°. The slant of the non-reinforced signal differed randomly from the reinforced one by ±10°, ±20° or ±30°. There were 240 trials in the training phase, which took 30–45 minutes to complete. The subjects completed all three conditions with a minimum of one day between conditions. The order of the two feedback conditions was counterbalanced, and the no-feedback condition was presented last. No feedback other than the haptic stimulation in the two reinforced conditions was given in the experiment.

Subjects were questioned upon completing the entire experiment. None realized that the texture- and disparity-specified slants differed or that the haptic stimulation was consistent with one and not the other.

Acknowledgements

This research was supported by NSF grant (DBS-9309820) and AFOSR grant (93NL366) to M.S.B. and by the Max-Planck Society. The authors thank Robert Cooper and Mike Landy for comments.

RECEIVED 5 AUGUST; ACCEPTED 18 NOVEMBER 1999

1. Backus, B. T., Banks, M. S., van Ee, R. & Crowell, J. A. Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Res.* **39**, 1143–1170 (1999).
2. Blake, A., Bülthoff, H. H. & Sheinberg, D. Shape from texture: ideal observer and human psychophysics. *Vision Res.* **33**, 1723–1737 (1993).
3. Buckley, D. & Frisby, J. P. Interaction of stereo, texture and outline cues in the shape perception of three-dimensional rigids. *Vision Res.* **33**, 919–933 (1993).
4. Johnston, E. B., Cumming, B. G. & Parker A. J. Integration of depth modules: Stereopsis and texture. *Vision Res.* **33**, 813–826 (1993).
5. Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res.* **35**, 389–412 (1995).
6. Clark, J. J. & Yuille, A. L. *Data Fusion for Sensory Information Processing Systems* (Kluwer, Boston, 1990).
7. Bülthoff, H. H. & Yuille, A. L. in *Attention Performance* XVI*: Information Integration in Perception and Communication* (eds. McClelland, J. & Inui, T.) 3 49–70 (MIT Press, Cambridge, Massachusetts, 1996).
8. Knill, D. C. Discrimination of planar surface slant from texture: human and ideal observers compared. *Vision Res.* **38**, 1683–1697 (1998).
9. van Ee, R., Banks, M. S. & Backus, B. T. An analysis of binocular slant contrast. *Perception* **28**, 1121–1145 (1999).
10. Rogers, B. J. & Bradshaw, M. F. Disparity scaling and the perception of frontoparallel surfaces. *Perception* **24**, 155–179 (1995).
11. Bradshaw, M. F., Glennerster, A. & Rogers, B. J. The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Res.* **36**, 1255–1264 (1996).
12. Backus, B. T. & Banks, M. S. Estimator reliability and distance scaling in stereoscopic slant perception. *Perception* **28**, 217–242 (1999).
13. Rock, I. & Victor, J. Vision and touch: An experimentally created conflict the two senses. *Science* **143**, 595–569 (1964).
14. Gibson, J. J. Adaptation, after-effect, and contrast in the perception of curved lines. *J. Exp. Psychol.* **16**, 1–31 (1933).
15. Rock, I. *The Nature of Perceptual Adaptation* (Basic Books, New York, 1966).
16. Festinger, L., Burnham, C. A., Ono, H. & Bamber, D. in *Journal of Experimental Psychology Monograph* Vol. 74, No. 4 (American Psychological Association, Washington., 1967).
17. Welch, R. B. & Warren, D. H. in *Handbook of Perception and Human Performance* (eds. Boff, K. R., Kaufman, L. & Thomas, J. P.) 25.1–25.36 (Wiley, New York, 1986).
18. Lederman, S. J. & Abbott, S. G. Texture perception: studies of intersensory organization using a discrepancy paradigm, and visual versus tactual psychophysics. *J. Exp. Psychol. Hum. Percept. Perform.* **7**, 902–915 (1981).
19. Harris, C. S. Perceptual adaptation to inverted, reversed, and displaced vision. *Psychol. Rev.* **72**, 419–444 (1965).
20. Welch R. B. *Perceptual Modification: Adapting to Altered Sensory Environments* (Academic, New York, 1978).
21. Stratton, G. M. Vision without inversion of the retinal image. *Psychol. Rev.* **4**, 341–360, 463–481 (1897).
22. von Helmholtz, H. *Handbuch der Physiologischen Optik* Vol. III (Leopold Voss, Leipzig, 1867).
23. Walk, R. D. Intersensory perception and sensory integration. (Plenum, New York, 1981).
24. Massie T. & Salisbury, J. K. in *Dynamic Systems and Control 1994* Vol. 55 (ed. Radcliffe, Clark J.) 295–301 (American Society of Mechanical Engineers, New York, 1994).
25. Gillam, B. Stereoscopic slant reversals: a new kind of 'induced' effect. *Perception* **22,** 1025–1026 (1993)