GIBSON – UNE APPROCHE ECOLOGIQUE DE LA VISION (1950, 1966, 1979)
Principes

"According to cue theory, perception is a consequence of putting together cues of clues available from incoming sensory information, much like a detective in a murder mystery discovers the identity of the murderer by putting together the various clues found at the scene of the crime." (Allard, 2000).

Gibson believed that the cognitive processing described by cue theory was not necessary for human perception. He developed the theory of **direct perception**. It assumes the following:

- all of the information required for perception is provided in the environment,
- perception is immediate and spontaneous, and
- perception and action cannot be separated.

GIBSON – UNE APPROCHE ECOLOGIQUE DE LA VISION (1950, 1966, 1979)

Principes

Approche alternative à celle postulant des traitements partiels par une multitude de modules spécifiques de traitement et de leur combinaison (hiérarchique) à travers des calculs, algorithmes, etc.

- **Le point de départ**
  Non pas l'image rétinienne traitée passivement, mais le *flux optique* ambiant *échantillonné activement* par l'Obs. et dont les caractéristiques spatio-temporelles sont en interaction directe avec le *comportement exploratoire* de l'Obs.

- **Les primitives**
  Le flux optique (et ses déformations) remplace les concepts restreints de points, blobs, barres et toute autre forme pour informer l'Obs. de façon *directe* et non ambiguë sur son environnement (et sa signification).

- **Perception *directe***
  Prise à la lettre, la notion de *perception directe* est analogue à celle de la *boîte noire* des behavioristes à la différence près que pour Gibson la flèche entre Stimulus et Rép. est une double flèche particulière en ceci que l'un ne peut être séparé de l'autre **S**<=>**R**

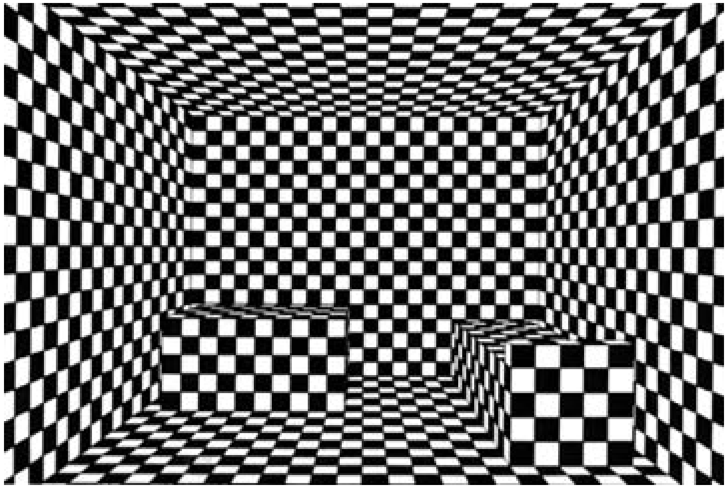
- **Perception et Action**
  Perception et action sont fortement liées et réciproquement contraignantes.

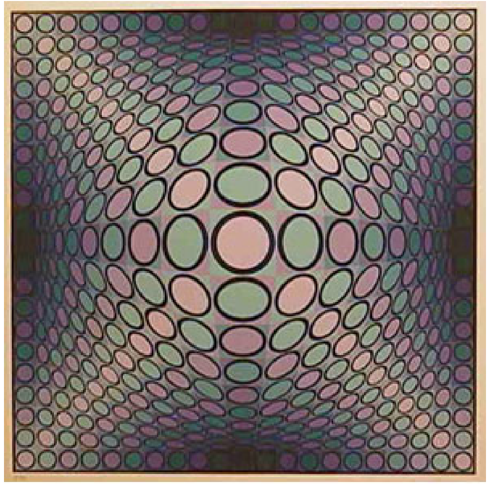
- **Affordance**
  Gibson rejette le concept de perception en tant que *représentation* et le détourne vers la notion de « *affordance* » : c'est à dire le fait de saisir (directement) en quoi un évènement perceptif est *utile* à l'individu (manger, s'asseoir, etc.). L'*affordance* est le concept qui sert de maillon entre perception et action.

GIBSON – UNE APPROCHE ECOLOGIQUE DE LA VISION (1950, 1966, 1979)

Historique

Pendant le guerre **=>** entraînement des aviateurs pour améliorer leur vision de la profondeur et profiter de façon optimale des indices divers de profondeur **=>** peu de réussite **=>** met l'accent sur la perception des *surfaces dans l'environnement* et donc sur les *indices écologiques: le sol* sur lequel ou au dessus duquel l'animal se déplace*, sa texture (*et sa *régularité statistique* – cailloux, gravier, sable, etc.)*, son inclinaison, éclairage,* etc. Il est donc question d'inventer une *géométrie de l'environnement* qui n'inclue pas nécessairement des points, plans, barres, etc.





La structure de l'environnement *structure* à son tour la lumière qui frappe la rétine. Pour évoluer dans l'environnement il faut donc pouvoir *extraire/traiter* cette structure ; une fois extraite, elle contient toute l'information nécessaire **=>** la perception directe. Le flux optique (qui n'est que le reflet de la structure de l'environnement) contiendrait des *invariants* vis-à-vis desquels le système visuel aurait évolué pour les "saisir" directement.

GIBSON – UNE APPROCHE ECOLOGIQUE DE LA VISION (1950, 1966, 1979)
Invariants perceptifs

- ***Exemples d'Invariants*** **:**

1. « *The horizon ratio relation*" **=>** l'horizon croise un objet à une hauteur particulière et tous les objets de même hauteur seront intersectés à la même hauteur quelque soit la distance à l'Obs.);
2. Le *gradient de taille* d'une surface texturée sera toujours le même pour une même inclinaison de cette surface quelque soit sa texture*.
3. Les *gradients de vitesse* dans le flux optique spécifient sans ambiguïté la position instantanée et la direction d'un Obs. en mouvement.
4. Un objet en mouvement sur/au-dessus d'une surface texturée cache progressivement les éléments de la texture dans la direction de son mouvement, les découvre derrière et les partage avec ses bords parallèles à sa direction de mouvement.
5. Les *jonctions en T – occlusion:* le pied du T appartient *toujours* à l'objet caché – et *en L* ou *en – continuité* de l'objet – et *en X – transparence*

*In a remarkable book that was first published over 50 years ago, James Gibson (1950) introduced the concept of texture gradients as a potential source of optical information for the perceptual specification of 3D surface structure. At about the same time, a similar concept was also emerging within a new school of painting called "optical art", especially in works of Victor Vasarely and Bridget Riley. This important new insight had an immediate impact on the study of human perception, and it inspired a succession of theoretical models and empirical investigations that continues today. [James T. Todd, T.J., Oomes, A.H.J., Koenderink, J.J. & Kappers, A.M.L (2004) The Perception of Doubly Curved Surfaces from Anisotropic Textures, *Psychological Sceince*, in press**.**

http://www.psy.ohio-state.edu/faculty/todd/pdf/Todd_Oomes_Koenderink_Kappers_2003.pdf

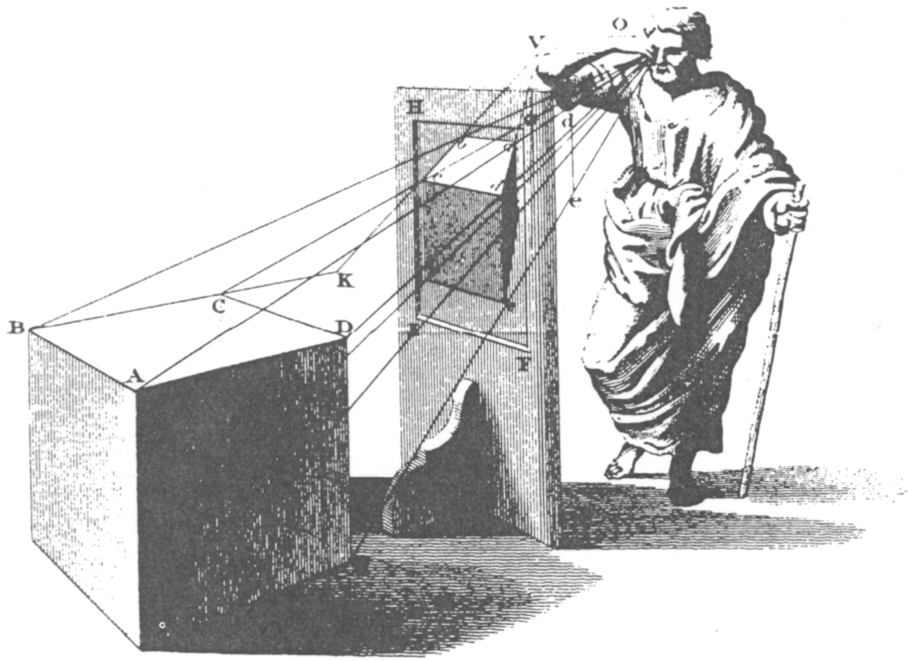


FIGURE 1.2 Principles of perspective, according to which the three-dimensional object with corners ABCDE is rendered in the window EFGH by the points abcde. The latter are located as if taut strings from ABCDE are drawn toward the eye and allowed to intersect the plane of the window. Illustration is from Pirenne (1970), who attributes it to Brook Taylor (1811), **New Principles of Linear Perspective.**

The invariant horizon ratio for terrestrial objects.

The telephone poles in this display are all cut by the horizon in the same ratio. The proportion differs for objects of different heights. The line where the horizon cuts the tree is just as high above the ground as the point of observation, that is, the height of the observer's eye. Hence everyone can see his own *eye-height* on the standing objects of the terrain.

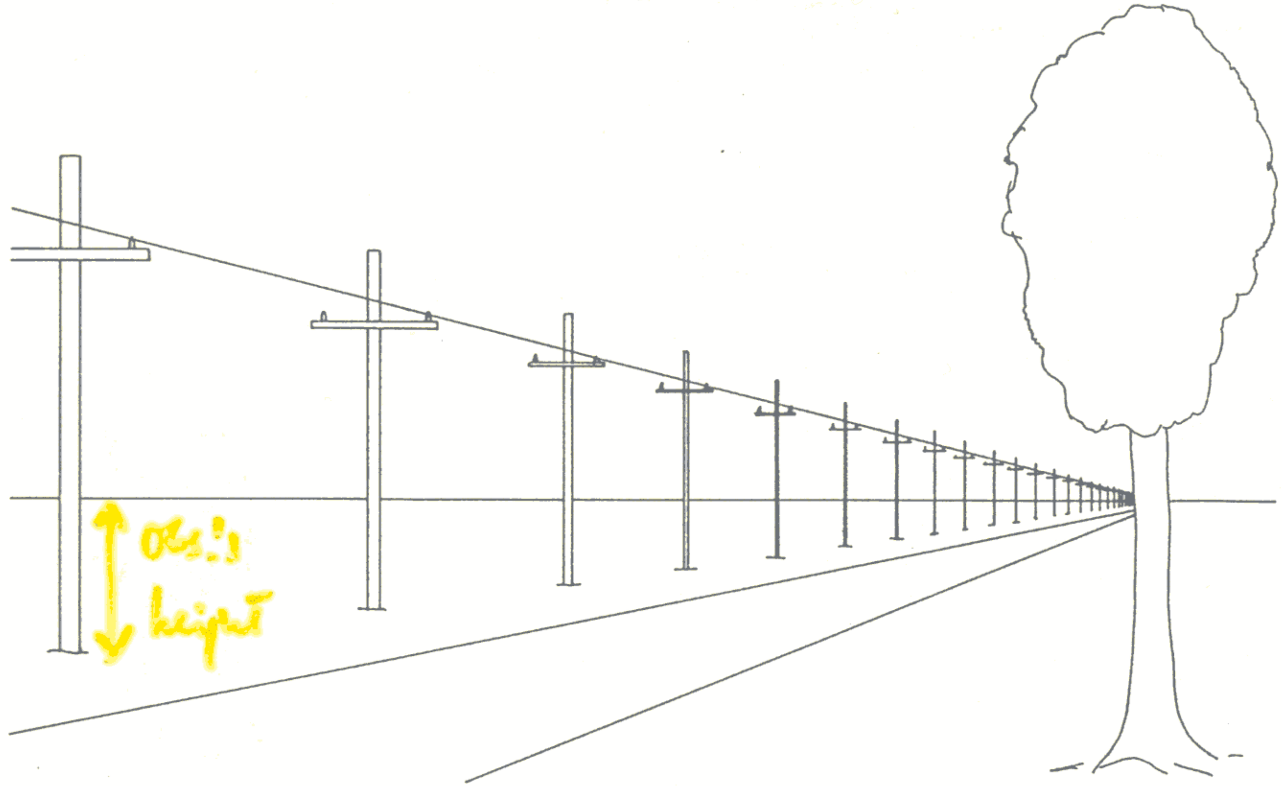
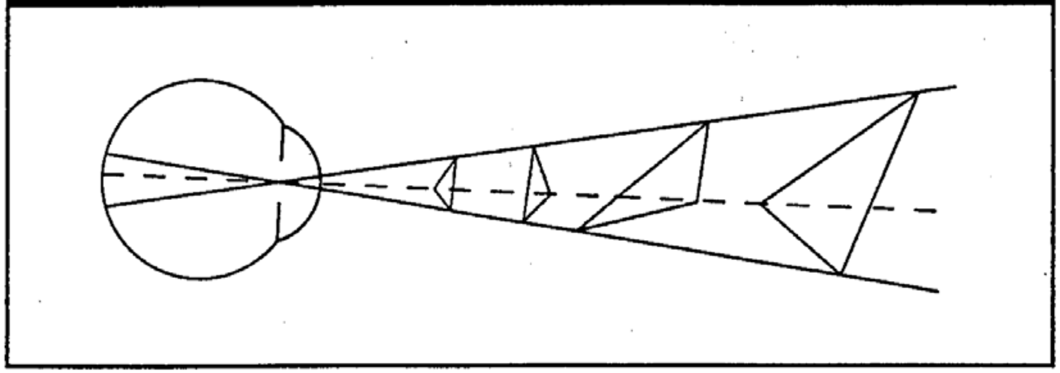
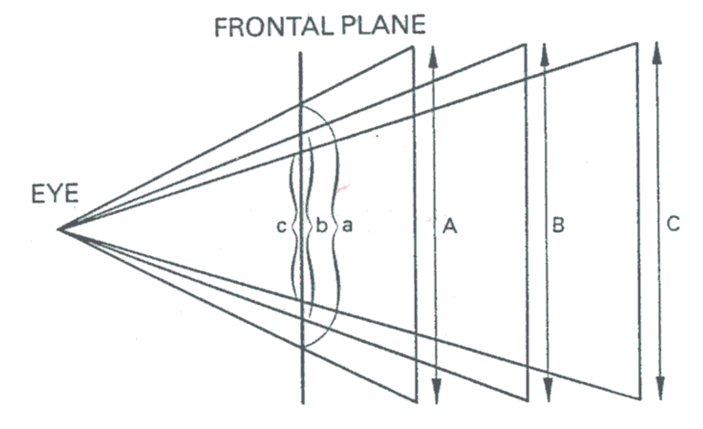


FIGURE 11.3



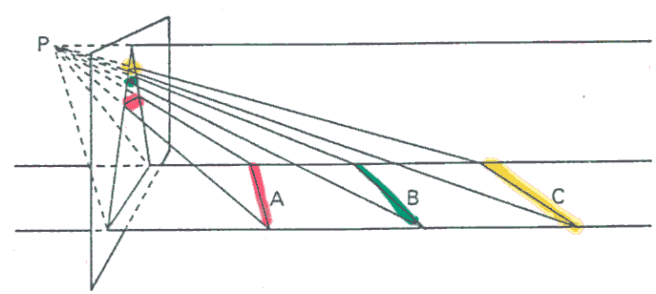
The kind of drawing typically used by students of visual perception to illustrate the ambiguity of the retinal image.

FIGURE 7.13

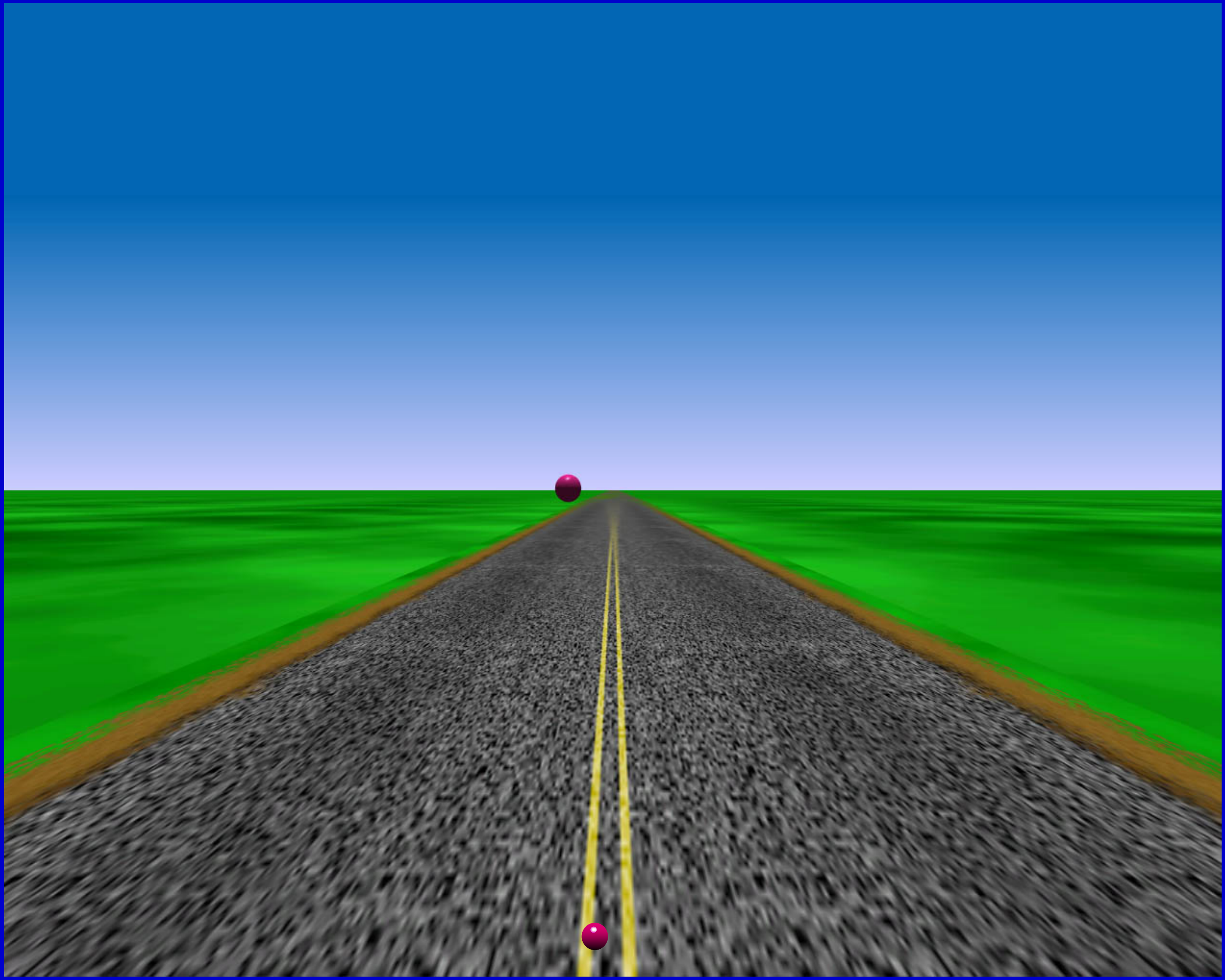


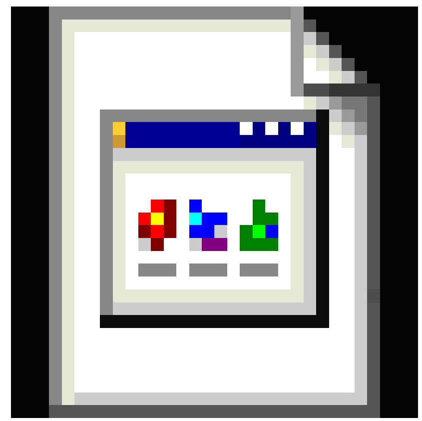
The size of the image cast by an object decreases with increase in distance.

FIGURE 7.14



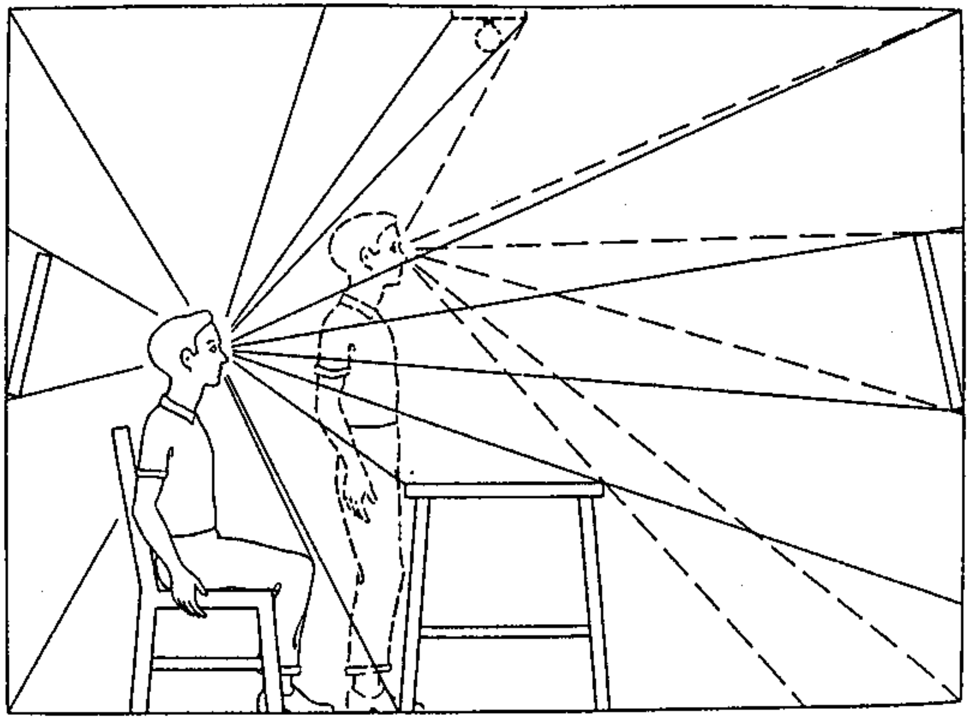
In this diagram, an observer at P looks at a set of square units on the ground (paving slabs, perhaps). The image formed at the frontal plane illustrates how the "cues" of relative size, perspective, and relative height are all simple consequences of the geometry of image formation. Adapted from J.J. Gibson (1950a).





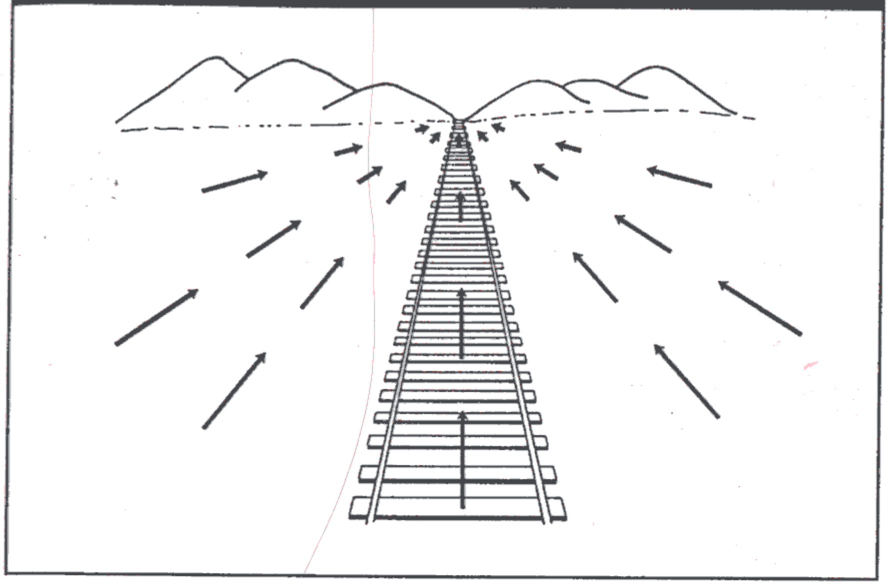
Perspective2.exe

FIGURE 11.4



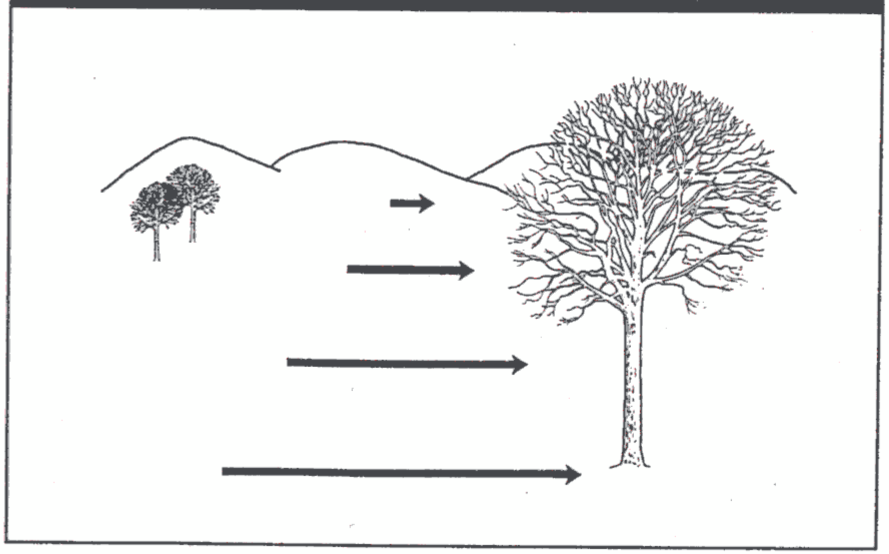
When an observer moves the entire optic array is transformed. From Gibson (1966). Used by permission of Houghton Mifflin Company.

FIGURE 11.7



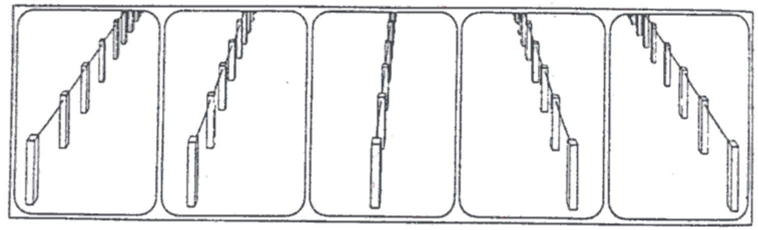
The optic flow field for a person sitting on the roof of a train, facing backwards.

FIGURE 11.8



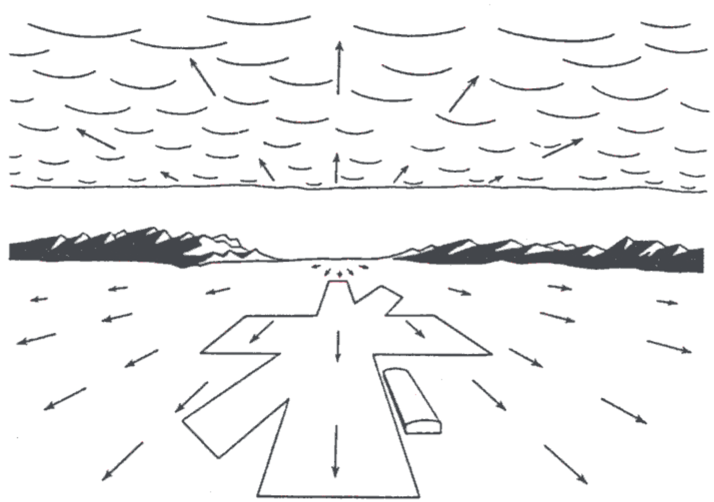
The optic flow field for a person sitting in a train and looking out of the window as they travel from right to left through this terrain.

FIGURE 11.5



Successive views of a row of fence posts as an observer moves past them. The observer travels from right to left between each of the frames from left to right. From Gibson (1950a). Used by permission of Houghton Mifflin Company.

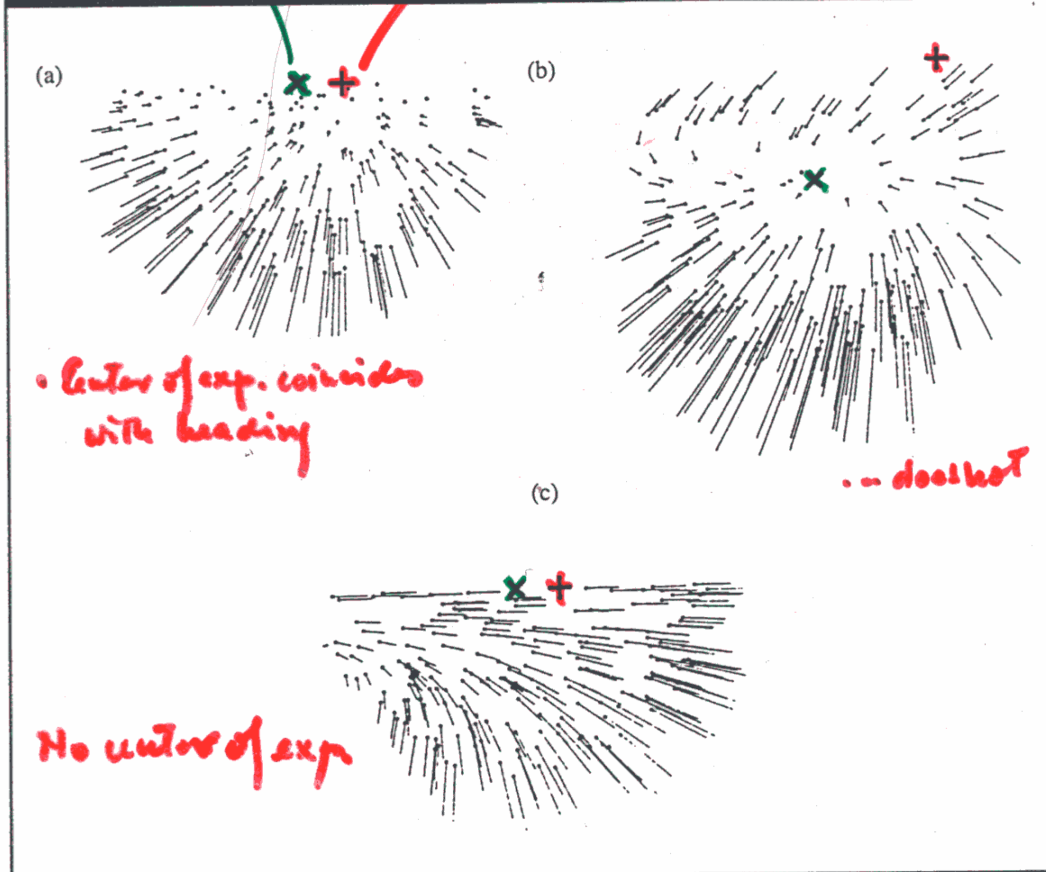
FIGURE 11.6



The optic flow field for a pilot landing an aeroplane. From Gibson (1950a). Used by permission of Houghton Mifflin Company.

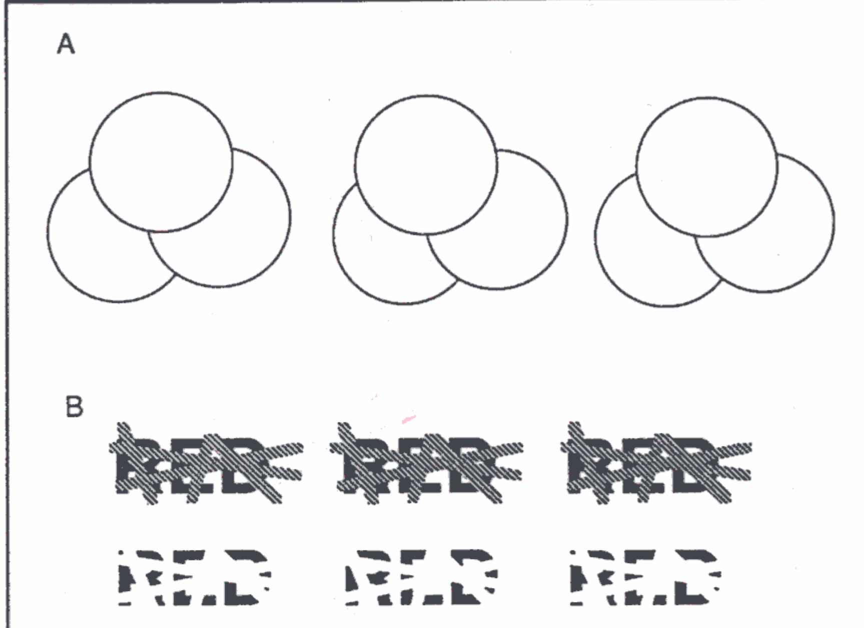
1. Flow of the ambient array specifies locomotion and nonflow specifies stasis.
2. Outflow specifies approach and inflow specifies retreat from.
3. The focus or centre of outflow specifies the direction of locomotion in the environment.
4. A shift of the centre of outflow from one visual solid angle to another specifies a change in the direction of locomotion, a turn, and a remaining of the centre within the same solid angle specifies no change in direction.

FIGURE 13.6



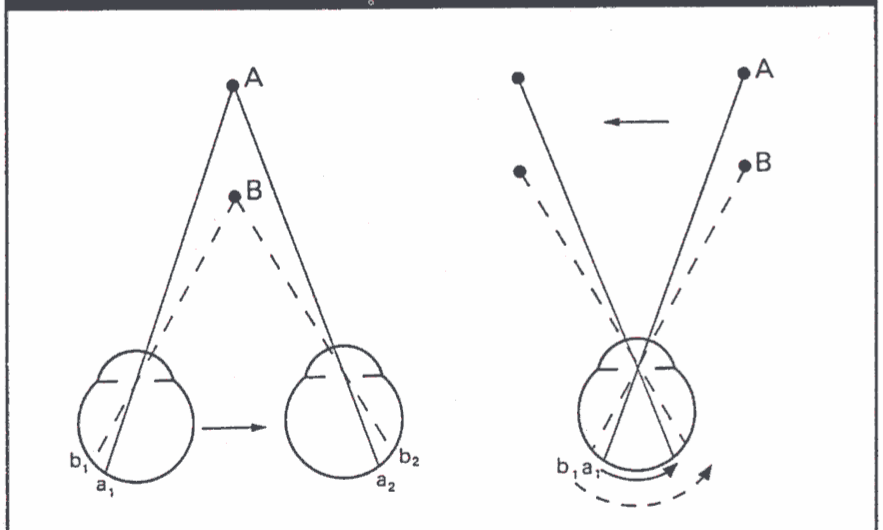
Three examples of retinal flow fields produced during forward movement over a ground plane. The lines attached to each dot show the speed and direction of optic motion at that point in the flow field. In each case, **+** denotes heading and **X** the direction of gaze. (a) No eye movement; there is a centre of expansion that coincides with heading. (b) Eye movement to fixate a point on the ground while moving forwards; there is a centre of expansion, but it does not coincide with heading. (c) Eye movement to track a moving object; there is no centre of expansion. Reprinted with permission from Lappe and Rauschecker (1994), *Nature*, Vol. 369, pp.712–713. Copyright © 1994 Macmillan Magazines Limited.

FIGURE 7.17



Stereo cues can reinforce, or countermand, occlusion cues in depth interpretation. In A and B, with cross-eyed fusion of the centre and right images the cues reinforce each other. Similar fusion of the left and centre images puts stereo depth and occlusion in opposition to each other. Images in B are adapted from Nakayama et al. (1989). Perceptual grouping and recognition of objects fragmented by occlusion are much more effective when occlusion cues are strong.

FIGURE 7.18



Motion parallax. The observer looks at two objects (A and B) at different distances. If the observer moves (as in the left diagram), or the objects move at equal speed (as in the right diagram), the image of the nearer object B moves further across the retina ($b_1 - b_2$) than does the image of A ($a_1 - a_2$).

As an object moves, elements of background texture are progressively wiped out (covered up) by its leading edge, unwiped (revealed) by its trailing edge, and sheared by edges parallel to its direction of movement.

FIGURE 11.9

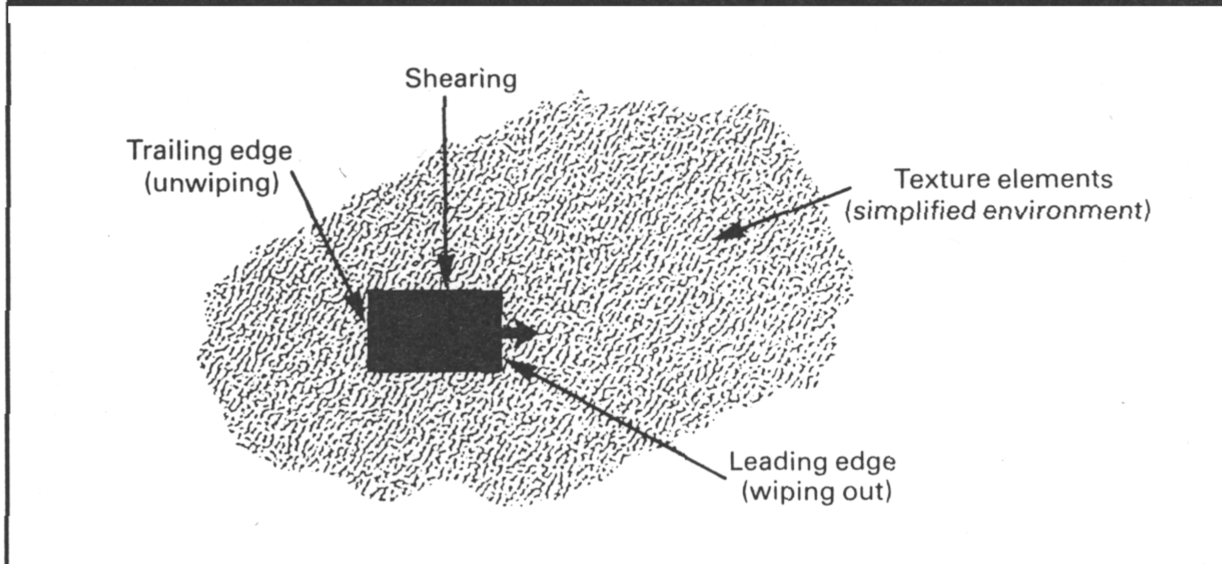
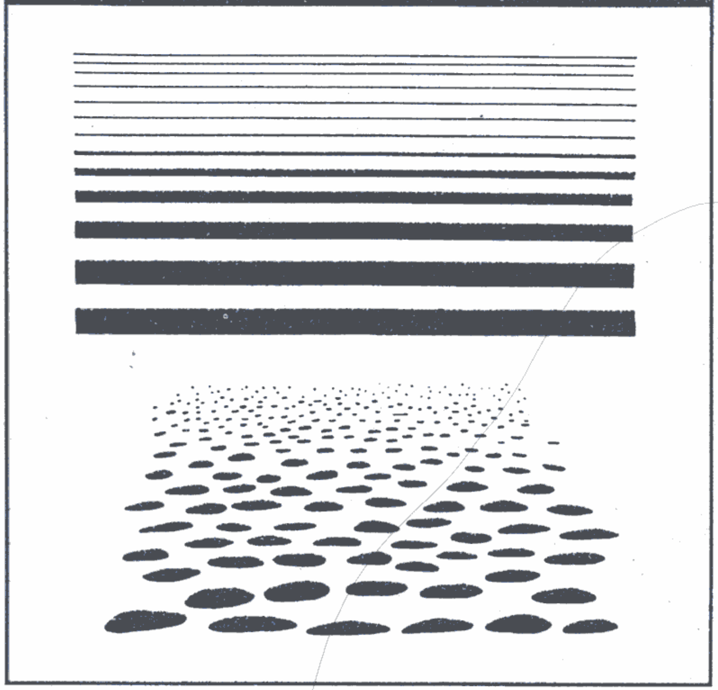


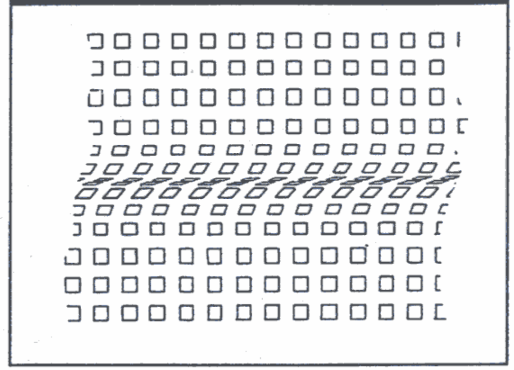
FIGURE 11.1



Examples of texture density gradients.

A given α is nonambiguously related to surface slant

FIGURE 11.2



Surface shape and slant can be revealed by texture.

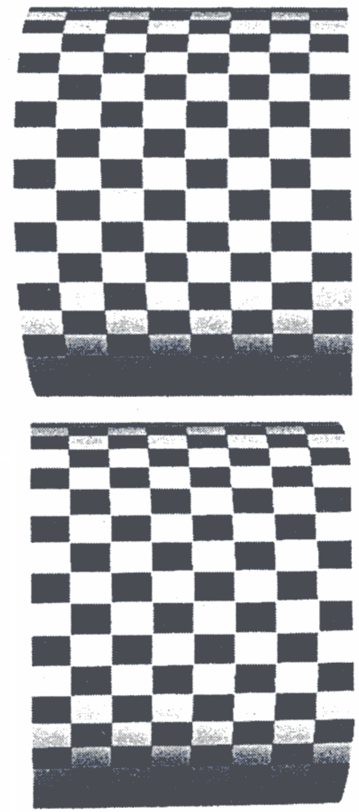


Fig. 22.14

Appearance of the cylinders used to investigate the combination of different cues to 3D shape (stereo, texture, and shading). The photograph is not accurately to scale, so that the pattern of cues in this example does not necessarily correspond in exact geometry to the binocular viewing of a real physical object. By contrast, in the experimental set-up, this condition was fulfilled as accurately as possible.

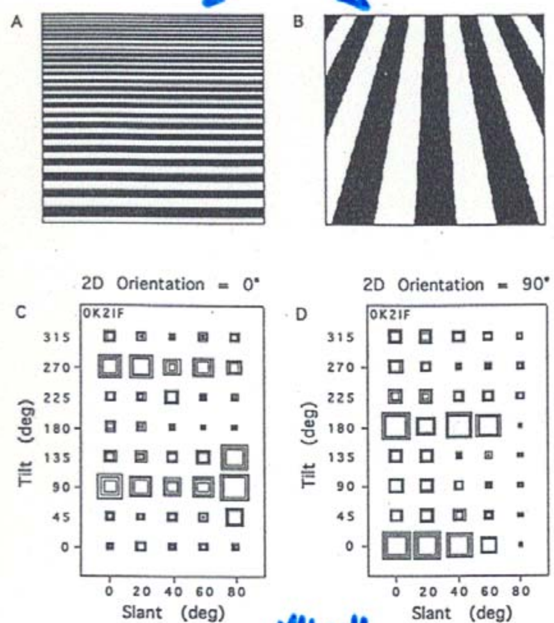


Figure 9.3
(*A*) Example of a 3D frequency modulated (FM) grating with a slant of 80° and a tilt of 90°. The 2D orientation of this grating was selected so that the direction of frequency modulation was parallel to the direction of slant. Gratings used in the experiments were quantized to 250 luminance levels and were displayed for 0.5 sec on a dark background. (*B*) Example of a 3D frequency modulated (FM) grating with a slant of 80° and a tilt of 90°, but whose 2D orientation is rotated 90° from the grating shown in A. The 2D orientation of this grating was selected so that the direction of frequency modulation was orthogonal to the direction of slant. The perspective projection of such a grating into 3D produces a complex frequency modulation. (*C*) Responses of a single V4 cell to 3D FM gratings whose 2D modulation (4 c/deg) was parallel to the direction of slant (as A). Slant is given on the abscissa, and tilt is given on the ordinate. Squares in the two-way plot give normalized relative firing rates and standard errors. This cell responded best to a wide range of slants when the tilt was either 90° or 270°. (*D*) Responses of the same V4 cell shown in C to 3D FM gratings whose 2D orientation was orthogonal to the direction of slant (as in B). The cell responded best to a wide range of slants when the tilt was either 0° or 180°. The pattern of responses for this cell indicates that it was tuned for a specific retinal orientation, but did not distinguish between various slants or tilts.

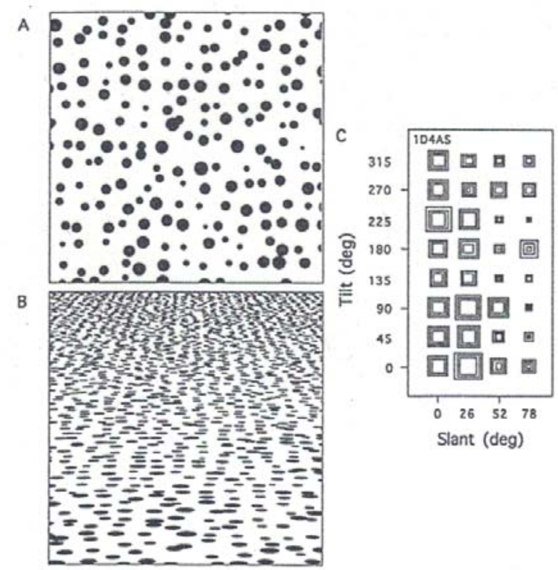


Figure 9.2
(*A*) Example of a texture used in 3D textured surface experiments. Each texture was composed of circular elements whose mean size was optimized individually for each cell. Textures were shown using the optimal color for each cell, and were displayed for 0.5 sec on a dark background. The texture shown here is parallel to the image plane, and so has a slant of 0° and a tilt of 0°. (*B*) Example of a 3D texture surface with slant of 80° and tilt of 90°. The surface appears to be receeding away from the viewer. (*C*) Responses of a single V4 cell to 3D textured surfaces varying in slant and tilt. Surface slant is given on the abscissa, while surface tilt is given on the ordinate. Data were averaged across the three texture sizes used, 0.5, 1.0, and 2.0 times the texture size corresponding to the best grating frequency for the cell. Squares give the normalized firing rates, and standard errors, to each combination of slant and tilt. This cell responded best to low slant values, and gave a fairly uniform response to different tilts.

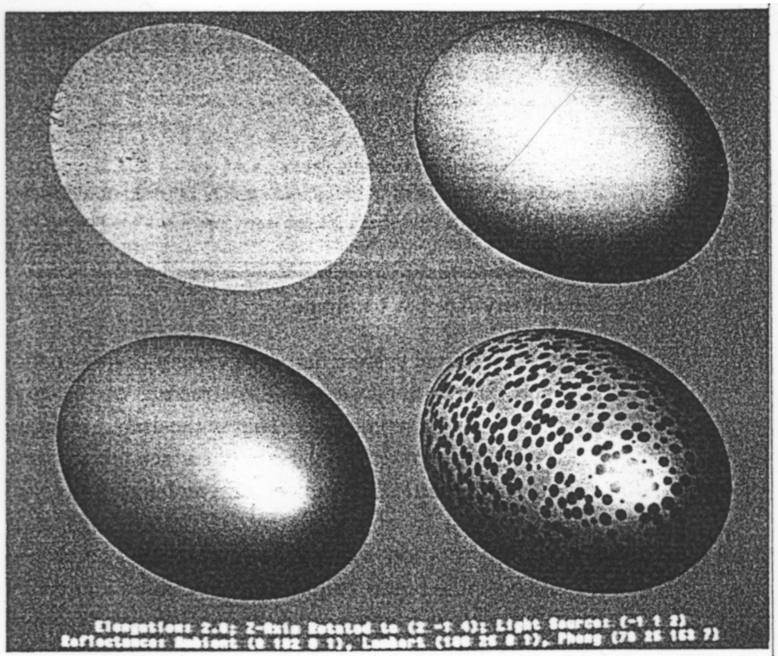


Fig. 20.1
Shape from X. The shape of the four objects looks quite different because the visual system derives different shape information from different shape cues. All four images were generated for the same 3D shape (ellipsoid of rotation) but with different simulated surface properties and under different lighting conditions.

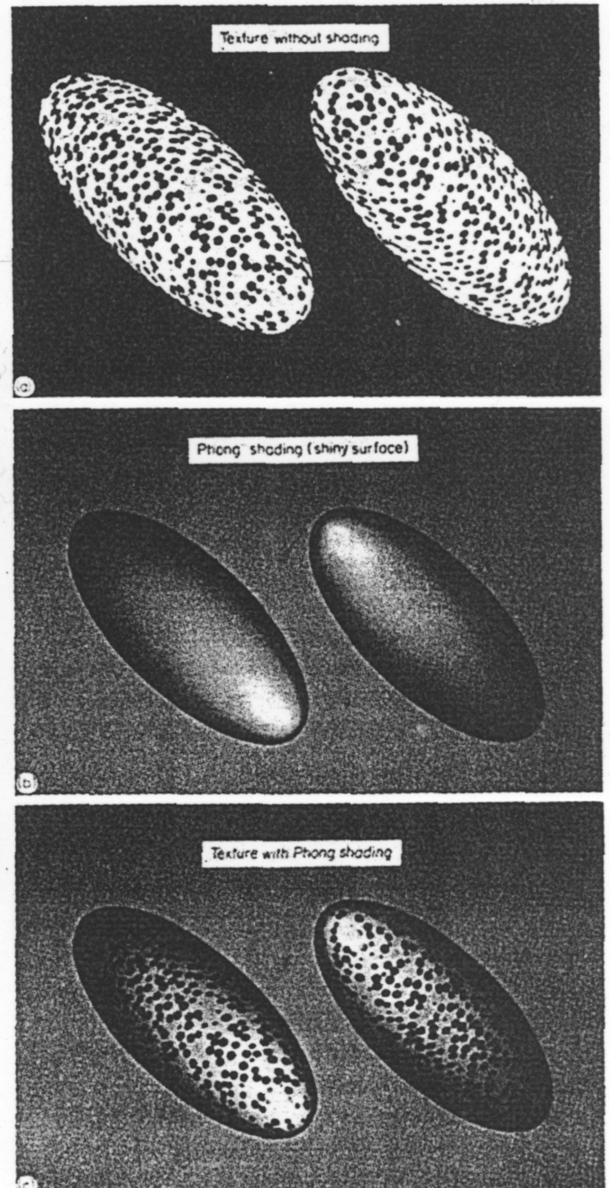
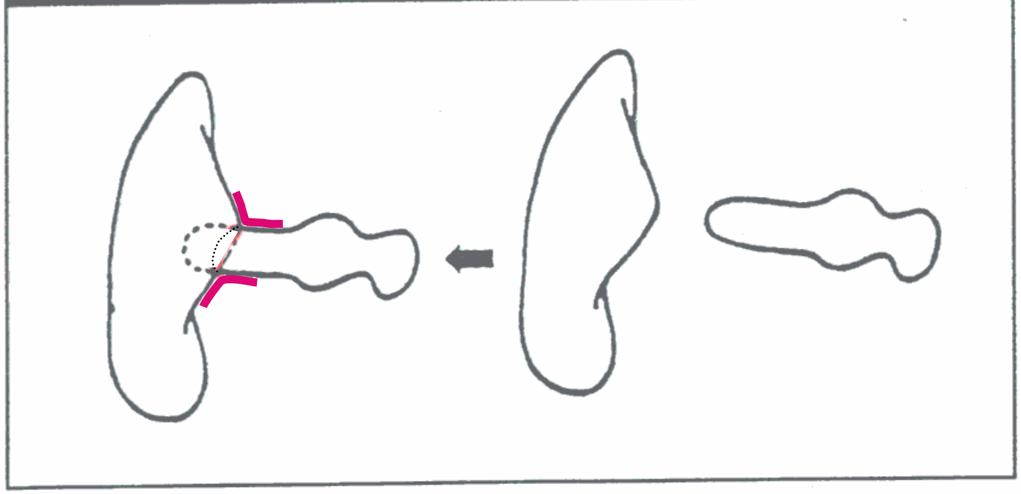


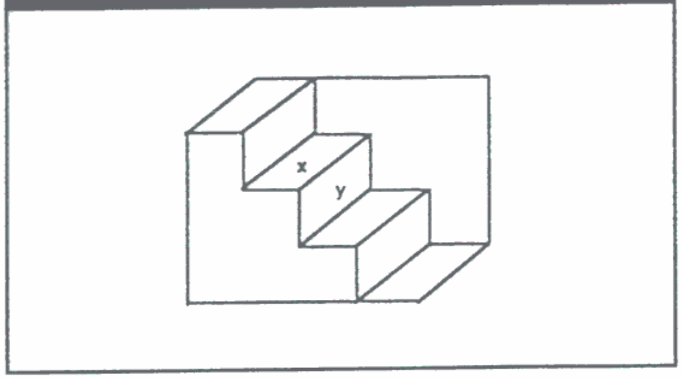
Fig. 20.2
3D descriptors. Different depth cues provide information about different 3D descriptors (e.g., range, shape, orientation). Try to estimate the angle between the long axes of the ellipsoids (for the correct answer, see text). (After Bülthoff & Mallot, 1990.)

FIGURE 6.33



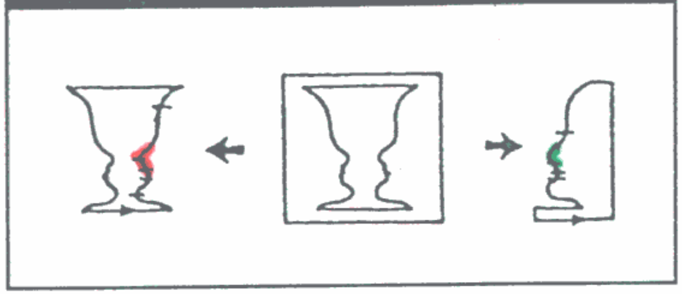
The "transversality regularity": When two surfaces interpenetrate they always meet in concave discontinuities. Reprinted from Hoffman and Richards (1984) with kind permission of Elsevier Science, The Netherlands.

FIGURE 6.34



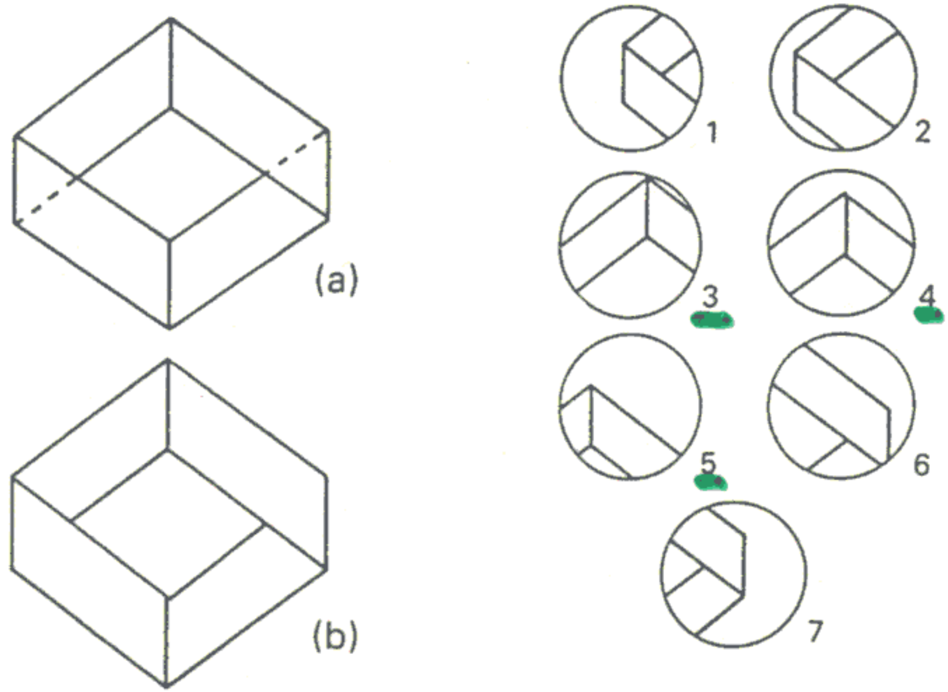
The Schröder staircase shows how part boundaries change as the figure and ground reverse. Reprinted from Hoffman and Richards (1984) with kind permission of Elsevier Science, The Netherlands.

FIGURE 6.35



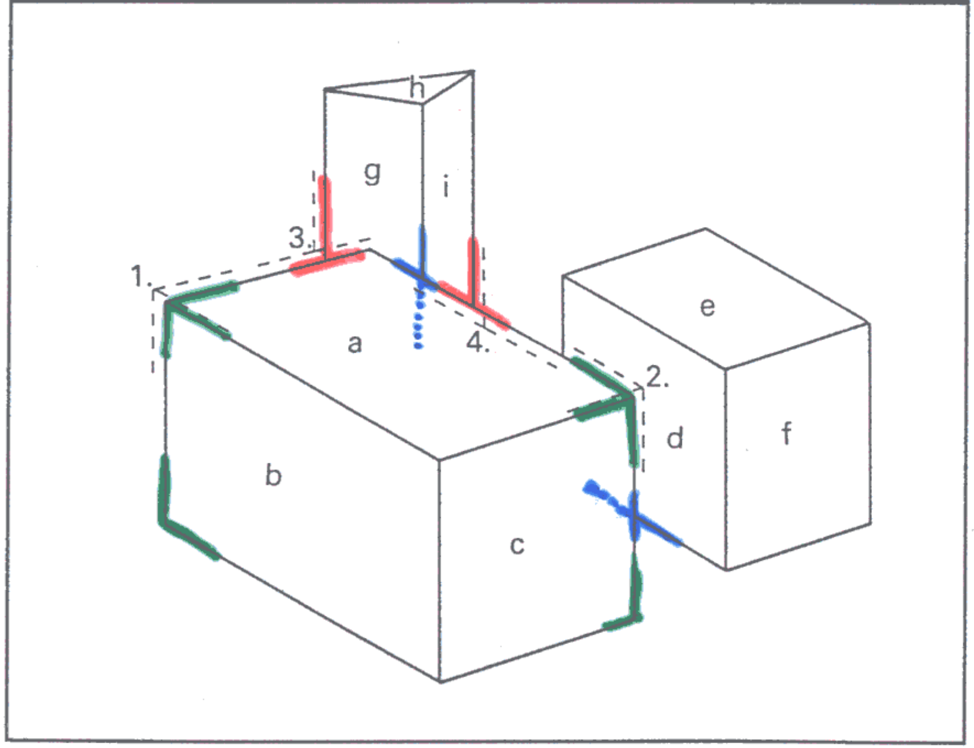
The faces–vase figure. When the vase region is taken as figure, then the concavities (minima of curvature) divide the vase into a base, stem, etc. When the faces regions are taken as figure, the concavities reveal parts corresponding to forehead, nose, etc. Reprinted from Hoffman and Richards (1984) with kind permission of Elsevier Science, The Netherlands.

FIGURE 7.24



(a) shows a figure that is ambiguous in depth and (b) shows an impossible figure. Observers notice ambiguity and impossibility if a figure is viewed piece by piece through an aperture, as illustrated in frames 1–7. Adapted from Hochberg (1968).

FIGURE 6.20



We have no difficulty in seeing that regions a, b, and c belong together (likewise d, e, and f; g, h, and i). Guzman's program SEE interpreted pictures like this by examining the junctions present. Examples of arrow junctions are shown at 1 and 2, and T junctions are shown at 3 and 4. What other kinds of junction are there in this picture?

T - occlusion → le pied du T appartient à l'objet caché

↑ L - continuité de l'objet

X - Transparence

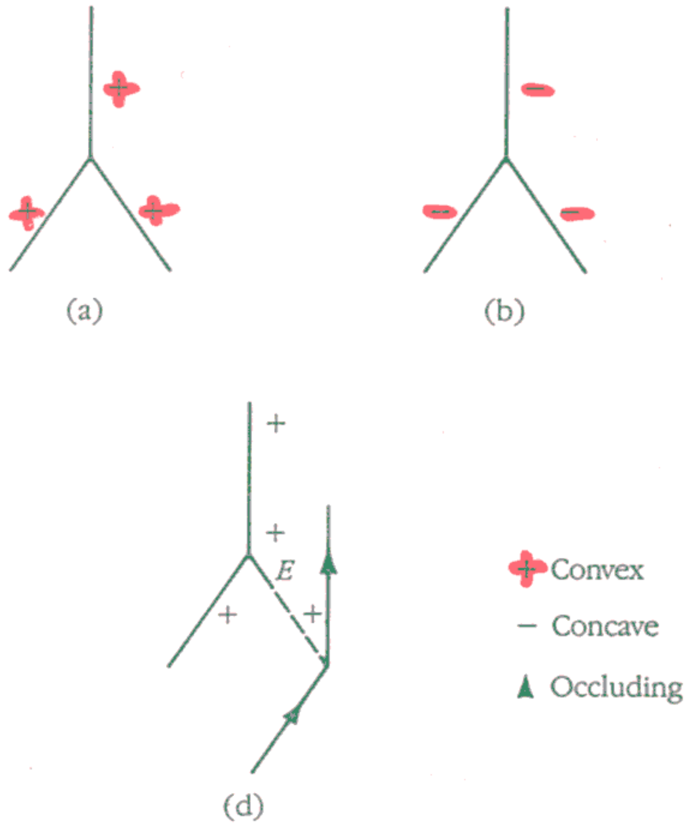
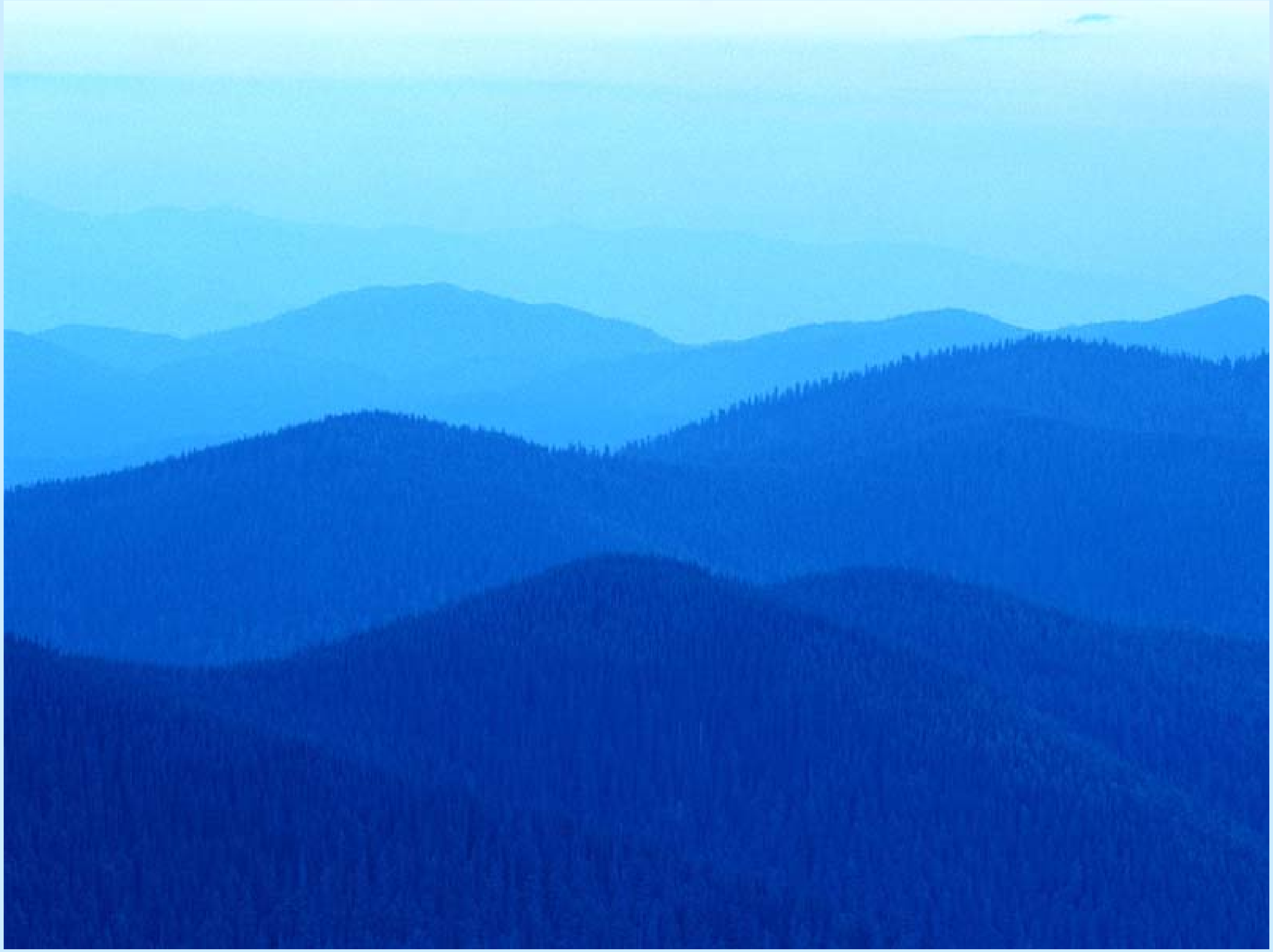
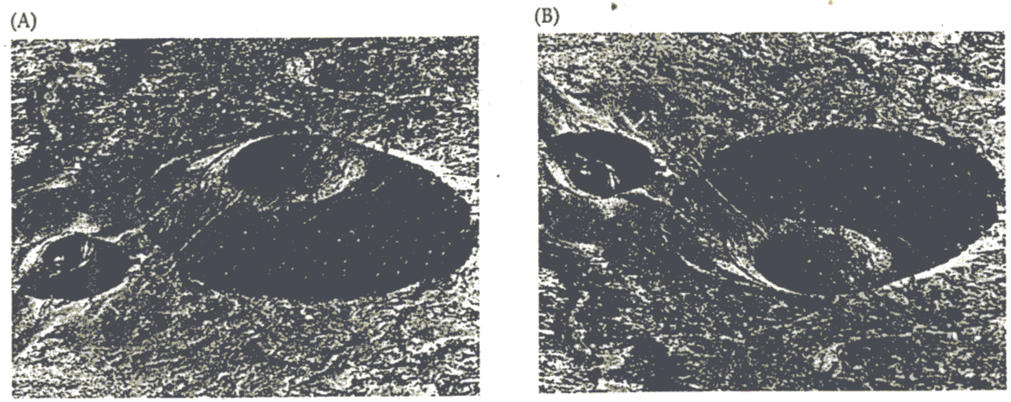


Figure 1–3. Some configurations of edges are physically realizable, and some are not. The trihedral junctions of three convex edges (a) or of three concave edges (b) are realizable, whereas the configuration (c) is impossible. Waltz cataloged all the possible junctions, including shadow edges, for up to four coincident edges. He then found that by using this catalog to implement consistency relations [requiring, for example, that an edge be of the same type all along its length like edge *E* in (d)], the solution to the labeling of a line drawing that included shadows was often uniquely determined.



The image of the near hills is brighter and clearer than the image of the distant ones, from which the light has been scattered by the atmosphere. Photograph by Mike Burton.



11.5 SHADING INFLUENCES SHAPE. The image in (A) is a photograph of two cinder cones in the K'au Desert lava fields of Hawaii. The image in (B) appears to contain craters with mounds in the center. Yet, the two images are the same except for an up-down flip. If you rotate the book 180 degrees, the image containing the mounds will now appear to contain craters, and conversely the image with craters will appear to contain mounds. The spatial relationship between the light and dark regions of the mound/crater is the main source of information defining it as convex or concave. Rotating the image rotates the shading cue and thus changes the shape we infer. After Rittenhouse, 1786.

Light comes from above

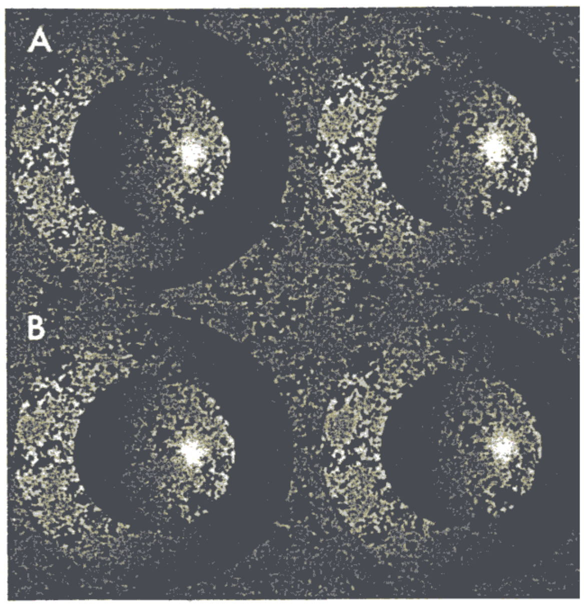
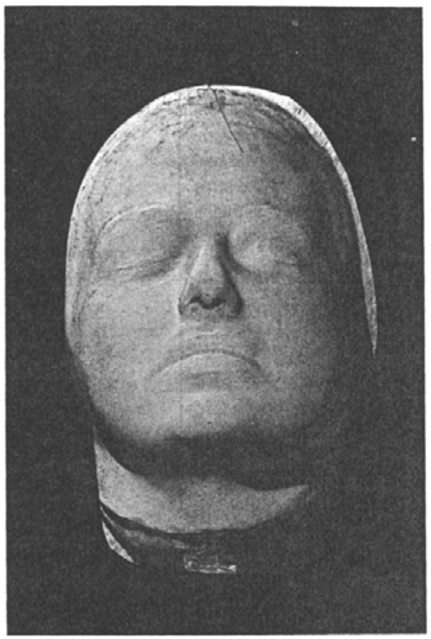


Fig. 20.12
Surface curvature. The perception of surface curvature can change with the position of a specular highlight. In order to demonstrate that the human visual system knows the physics of light reflection we used an image of a surface whose three-dimensional interpretation can flip easily between two states (convex/concave). If a highlight is added to the image the 3D interpretation of the inner part of the surface is biased more towards convex. A stereo pair was made with zero disparity for the textured surface, and then a specularity superimposed either in front of (A) or behind (B) the textured surface (uncrossed view), flipping randomly between the two, with 5 or 10 sec exposures separated by a random-dot masking frame. Subjects made a two-alternative forced choice (2AFC) between convex and concave. After a short training period (20 exposures without feedback) they made more choices that conform to the predictions of the model (C). A control experiment with a white disk of about the size of the specularity that did not look like a highlight at all, did not show any consistent effect between subjects on the perceived curvature. It might be difficult to experience the curvature effect if the images are not displayed on a CRT monitor because of the limited contrast range in the print. In order to get the best effect it is essential that the highlight look like a real reflection of the light source. (Redrawn from Blake & Bülthoff, 1990.)

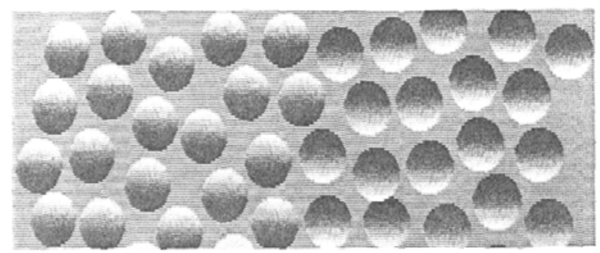
Specularity & Stereo

FIGURE 7.10

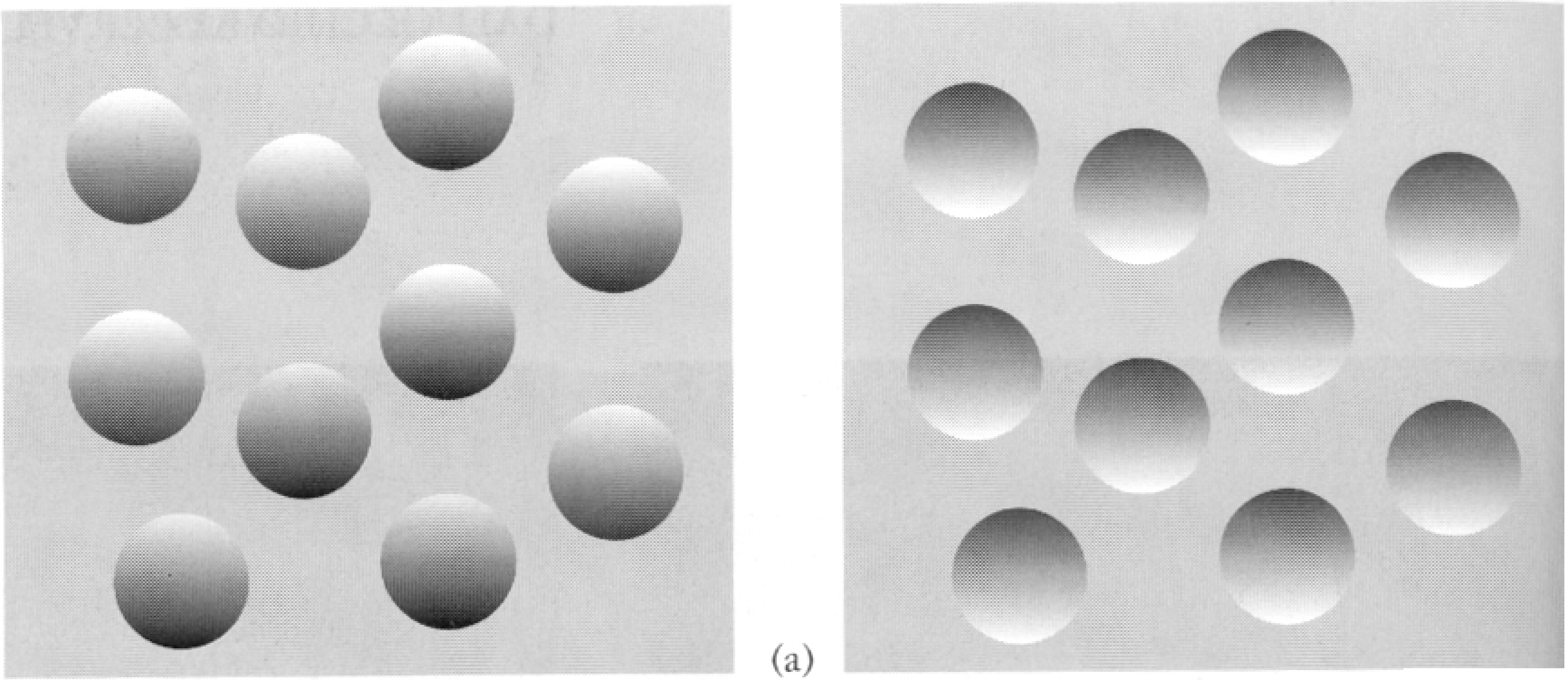


This is a picture of a hollow mask, illuminated from behind. In real life, as in this photograph, we see a normal face, with the tip of the nose nearer to us than the eyelids. Photograph by Sam Grainger.

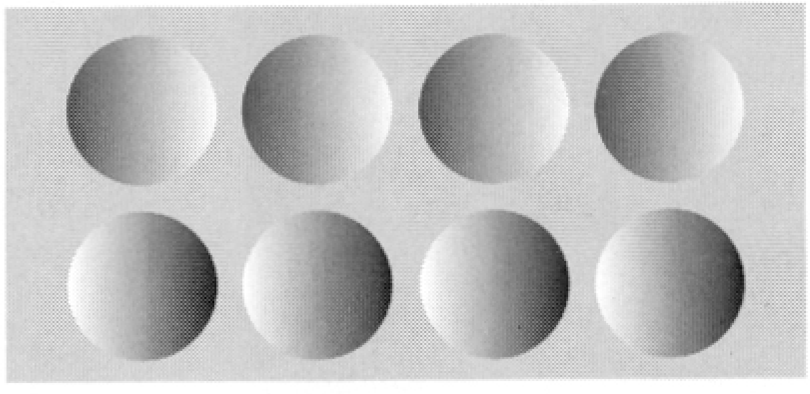
FIGURE 7.15



The shapes on the left are usually seen as bumps, but those on the right are usually seen as dents, consistent with an overhead light souce. From Ramachandran (1988). Reprinted with permission from *Nature* (Vol.331, pp.163–166). © 1988 Macmillan Magazines Limited.

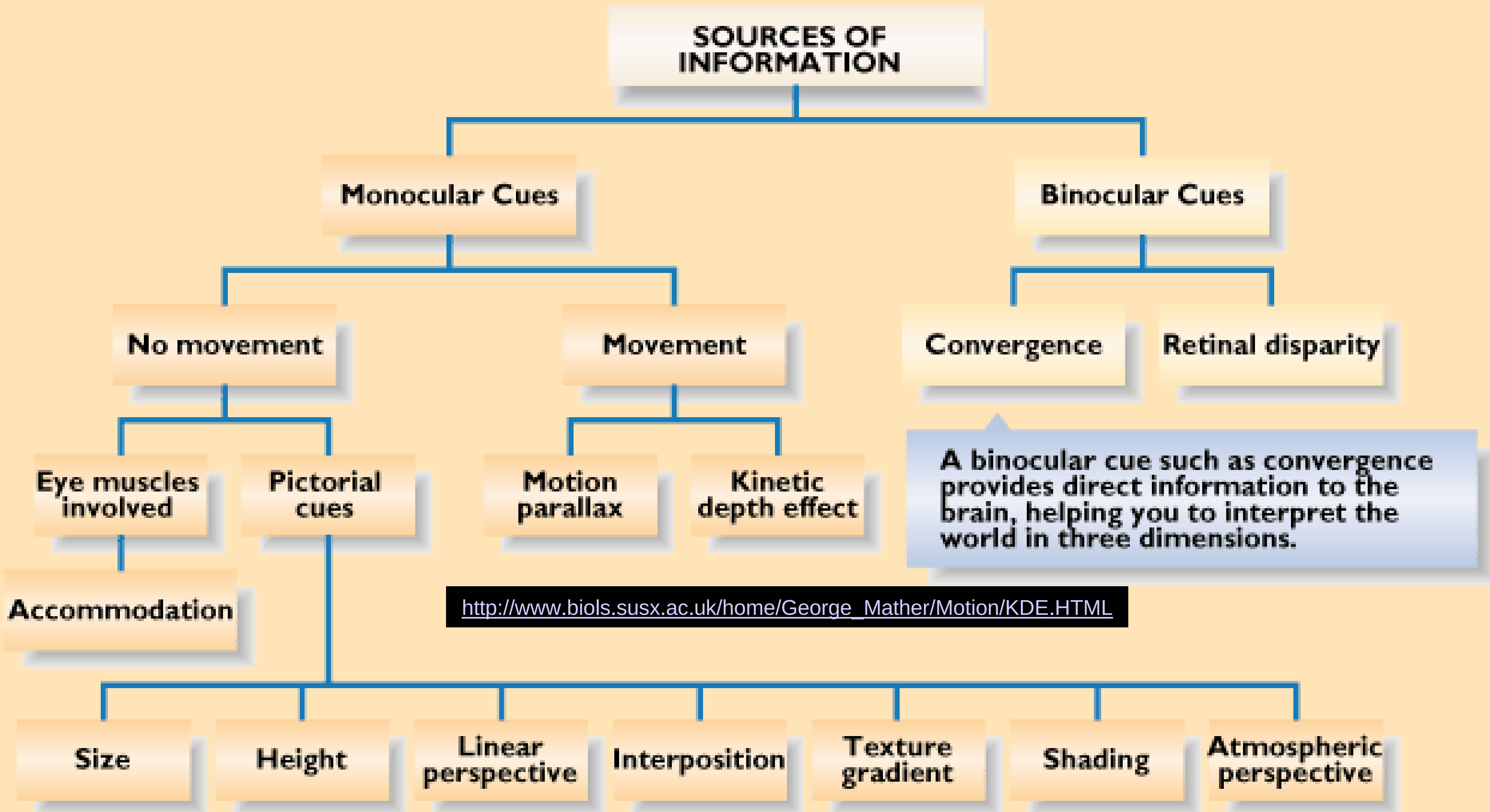


(a)



(b)

Twelve Sources of Information About an Object's Distance



From Andersen G.J. Design of Visual Displays for Driving Simulator Research.
(http://www.uiowa.edu/~neuroerg/Simulator%20Users%20Group/Andersen.HFES.04.vis.ppt)