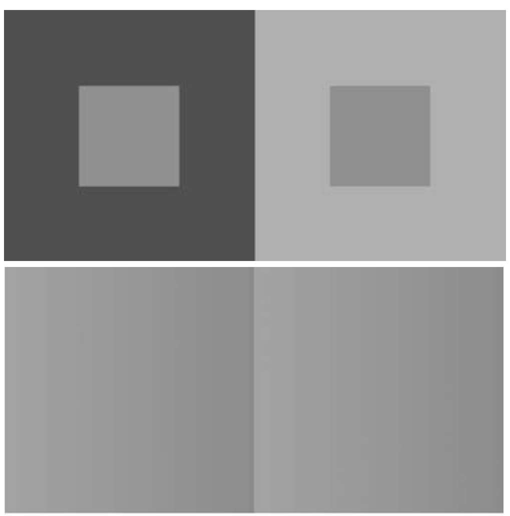
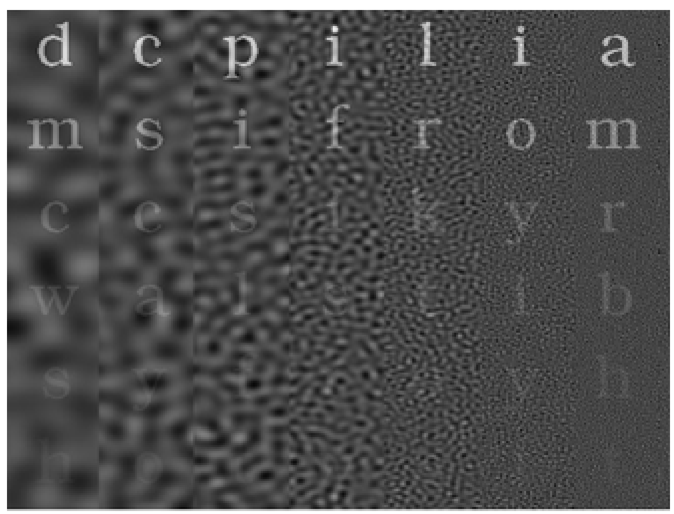
# GESTALT...

# LA **GESTALT** ET LA **GESTALT MODERNE**

L'approche Fouriériste de la perception visuelle a été et reste très contrintuitive, en particulier du fait que « consciemment » nous n'avons pas l'impression de pouvoir « adresser » le contenu spectral d'une image.

A l'origine conçue et utilisée pour rendre compte de phénomènes au seuil (de détection ou discrimination) et aisément interprétable quant à son pouvoir de prédiction de quelques phénomènes perceptifs « simples » (contraste simultanée, bandes de Mach…), elle reste clairement (?) inutilisable pour ce qui est de la reconnaissance et *apparemment* inutilisable pour ce qui est des phénomènes cognitifs de haut niveau. Toutefois, elle s'accommode bien d'un grand nombre de tels phénomènes (voir par ex. la reconnaissance des lettres, l'approche d'Oliva concernant la reconnaissance des scènes naturelles complexes, certaines lois de la Gestalt mais non pas des invariants perceptifs promus par Gibson).





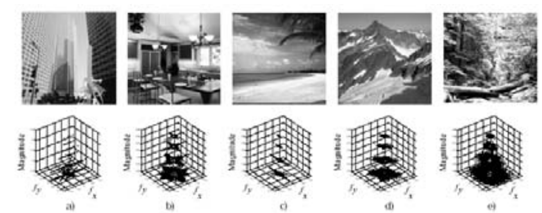




Figure 1: Examples of power spectrum forms for prototypical images (vertical axis is the magnitude in logarithmic scale, horizontal axis are the spatial frequencies $f_x$ and $f_y$). At the bottom, we show sections at several levels of the power spectrum of each image.

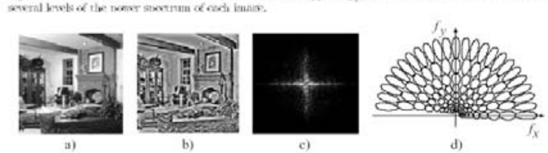



Figure 2: This figure shows the main steps for computing the vector of 100 components used to represent an image. a) Original image. b) Output of the pre-processing stage. The effect of illuminant and shadows have been reduced. c) Power spectrum of the prefiltered image. It is computed as the squared of the magnitude of the Fourier Transform. d) -3dB sections of the set of gabor filters used to sample the power spectrum. The highest frequency is 1/3 cycles/image and the lowest one is 1/72 cycles/image.

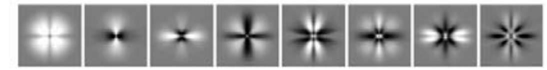


Figure 3: The first 8 Principal Components calculated from the power spectrum of 700 scenes. The horizontal coordinate is $f_x$ and the vertical one is $f_y$. The symmetrical structure of the principal components is due to the mirror transformation applied to the image power spectrum.

Oliva O. et al. (1999). Global semantic classification of scenes using power spectrum templates. In Proceedings of The Challenge of Image Retrieval (CIR99), Springer Verlag BCS Electronic Workshops in Computing series, Newcastle, UK.

# LA **GESTALT** ET LA **GESTALT MODERNE**

La psychologie de la Gestalt a mis l'accent sur l'importance de l'organisation perceptive dont une grande partie des processus sous-jacents pourraient être caractérisée comme étant de type «mid-level». Les concepts clé incluent le groupement, l'appartenance, la bonne continuation, la proximité, etc.

*``The fundamental `formula' of Gestalt theory might be expressed in this way: There are wholes, the behaviour of which is not determined by that of their individual elements, but where the part-processes are themselves determined by the intrinsic nature of the whole. It is the hope of Gestalt theory to determine the nature of such wholes.'' (Max Wertheimer 1938a, page 2)*

# LA **GESTALT** ET LA **GESTALT MODERNE**

❑ ***<u>L'approche classique</u>***

➢ **La loi de la Prägnanz** (Kofka d'après Wertheimer):
*De toutes les organisations possibles sera perçue celle caractérisée par la forme **la meilleure**, **la plus simple** et **la plus stable.***

➢ **Lois relatives au Contexte**

- Proximité
- Similarité
- Sort commun (*common fate*): les objets qui bougent ensemble sont perçus comme groupés
- Bonne continuation
- Fermeture (closure): Les contours fermés sont perceptivement plus prégnants
- Autres "préférences": les objets plus petits ==> Figure; les objets plus grands ==> Fond

Autres concepts: *familiarité* (avec les objets), "*set objectif*", *champ de forces* comme *isomorphes* au stimulus distal. La *Doctrine de l'Isomorphisme* était sensée expliquer ces lois dans la mesure où elles étaient analogues aux lois de la physique des champs de forces.

➢ ***Rappel de quelques "lois" similaires en audition/musique***

- La mélodie quelque soit la hauteur à la quelle elle est jouée
- Les gammes toniques elles-mêmes (dominante, etc.)
- La fission mélodique (i.e. groupement)
- Les « catégories » phonétiques (ba/pa, voyelles, parole)
- etc.

*``The fundamental `formula' of Gestalt theory might be expressed in this way: There are wholes, the behaviour of which is not determined by that of their individual elements, but where the part-processes are themselves determined by the intrinsic nature of the whole. It is the hope of Gestalt theory to determine the nature of such wholes.'' (Max Wertheimer 1938a, page 2)*

Not only was it said that the whole is more than its parts, but the perception of the whole is prior to that of its parts. Gestalt psychology was in the main stream of continental philosophy and used the methods of phenomenology as adumbrated by Goethe, Purkinje, and Hering. Later, Max Wertheimer applied a similar approach to the study of creativity in his Productive Thinking (1945).

In its early years, Gestalt psychology was principally concerned with perception, and a range of robust demonstrations was devised to support its holistic nature. Max Wertheimer (1923, 1938c) described many principles of perceptual organisation, of which proximity, similarity, symmetry, and good continuation were the principal ones. These were illustrated with sets of figures consisting of filled and open dots arranged in patterns which demonstrated the grouping principles (see figure 2). [Nicholas J Wade (2004). Good figures. *Perception*, *33*, 127-134]

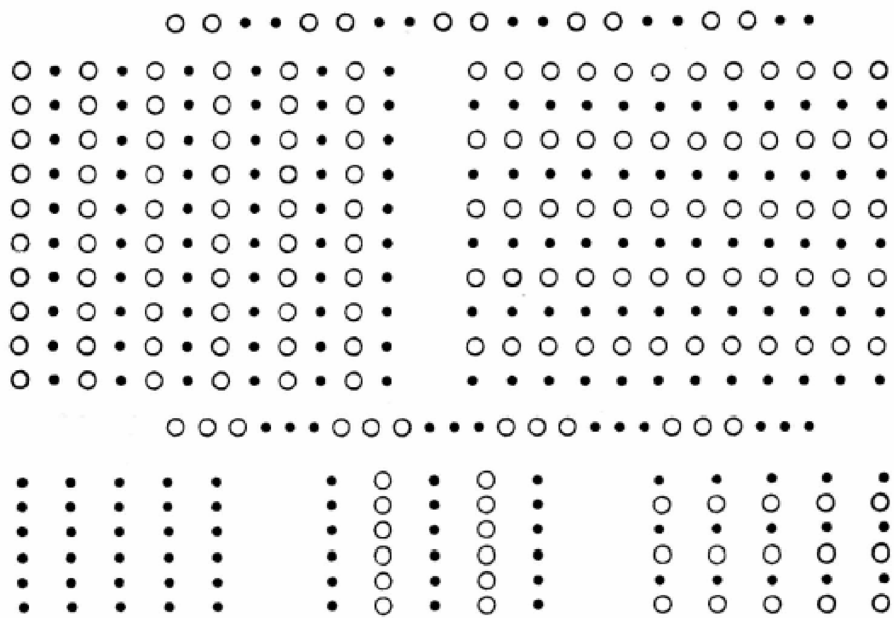


**Figure 2.** Examples of Gestalt grouping by similarity as illustrated by Max Wertheimer (1923). "Other things being equal, if several stimuli are presented together, there is a tendency to see the form in such a way that the similar items are grouped together (*factor of similarity*)." (Max Wertheimer 1958, page 119)

# LA **GESTALT** ET LA **GESTALT MODERNE**

❑ ***<u>L'approche moderne</u>***

- Structuration des champs récepteurs classiques (filtrage passe-bas) et étendus ("binding") ;
- Présuppositions (Bayésiennes: la perception est une interprétation) génétiques et/ou apprises relatives à la structure (physique) du monde afin de contraindre la gamme d'interprétations possibles d'une stimulation – Marr, 1982 ;
- *Similarité*: Parce qu’une même surface absorbe et reflète la lumière de la même façon, il est probable que les régions différentes d'un même objet aient le même aspect (texture, etc.) ;
- *Proximité, bonne continuation & sort commun*: Parce que la matière est cohésive, il est probable que des régions adjacentes appartiennent au même objet et qu'elles subissent le même sort ;
- *Similarité & proximité*: La forme des objets naturels varie en général progressivement (*smoothly*) de sorte que des régions adjacentes tendent à refléter la même quantité de lumière et être dans le même plan de profondeur (<u>algorithmes stéréo</u>; <u>interactions entre CRs</u>); en corollaire, des discontinuitésabruptes signalent probablement une transition entre deux objets (<u>*zero-crossings*</u>) ;
- *Figure-fond*: Un objet solide repose sur une surface donc il est vu d'en haut (<u>cube de Necker</u>) et il tend à être plus petit que cette surface/fond.

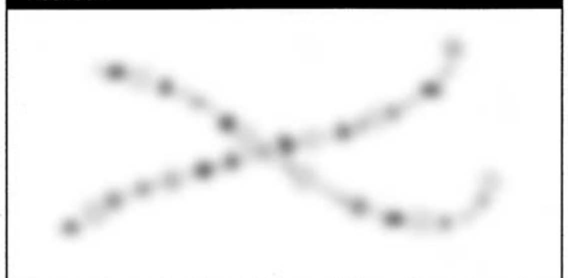
FIGURE 6.8


The dots in (a) form columns because they are closer vertically than horizontally. At (b) we see rows, the dots here are closer horizontally; (c) is ambiguous, the dots are equally spaced in both directions.

FIGURE 6.9

This picture is seen as columns. Similarity in brightness of the dots overrides proximity.

FIGURE 6.11



Quite dissimilar shapes may be grouped together through a combination of proximity and good continuation.

**FILTRAGE PASSE-BAS**

The form at (a) looks like a hexagon, whereas that at (b) looks like a cube. Of course (a) is also a legitimate view of a cube.

Familiarity

FIGURE 6.6

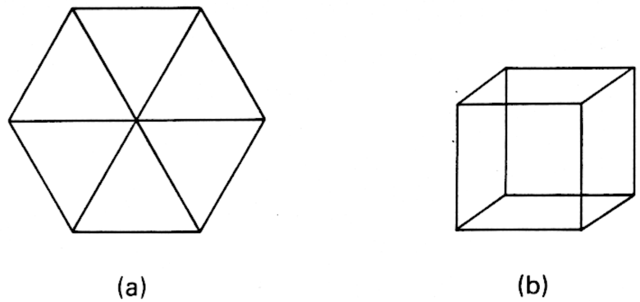
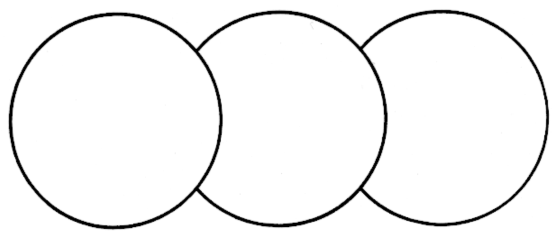
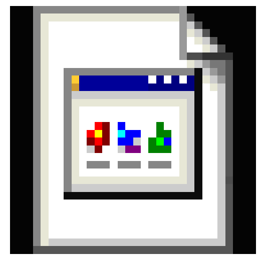



FIGURE 6.7

Closure

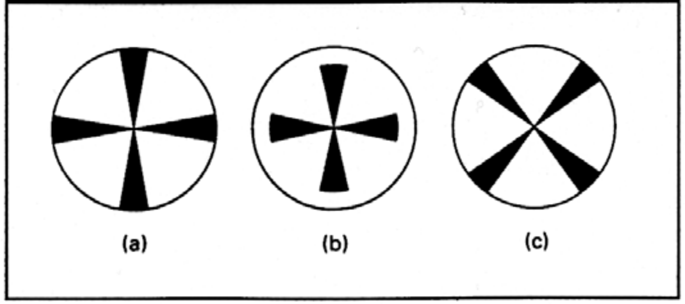


Most people would see this as a set of overlapping circles, although two of the shapes might have "bites" taken out of them.



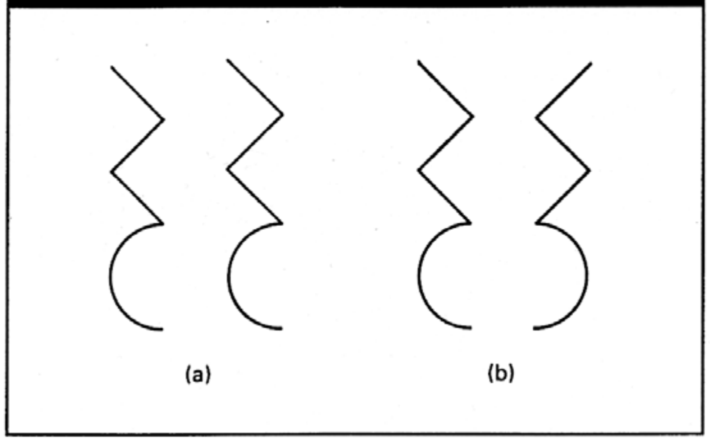
HiddenImage.exe

FIGURE 6.12



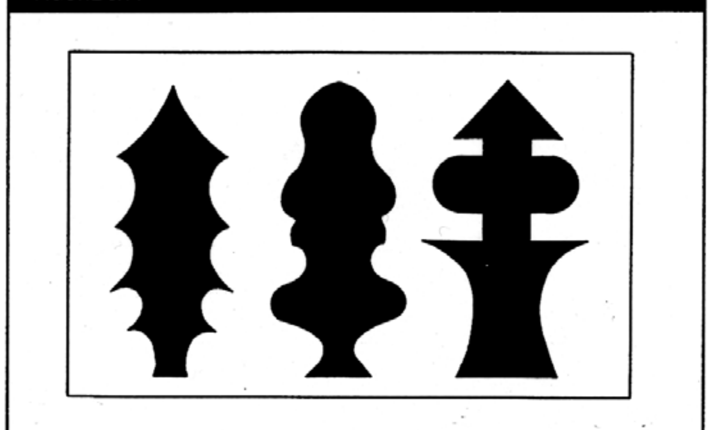

The preferred perception of (a) is a black propeller on a white background. This preference is enhanced if the white area surrounds the black, as at (b). If the orientation of the forms is altered, so that the white area is oriented around the horizontal and vertical axes, as at (c), then it is easier to see the larger white area as a figure.

FIGURE 6.13



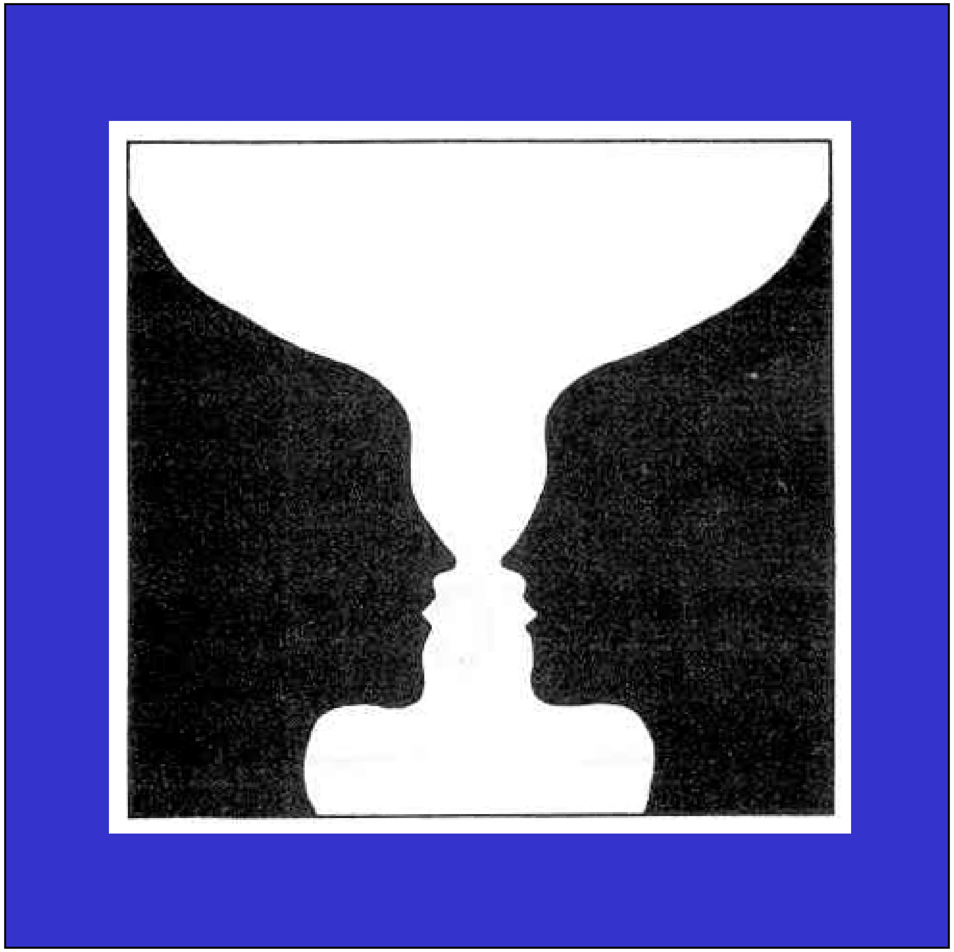

At (a) one form is repeated without reflection around a vertical axis. This arrangement is not as perceptually salient as the arrangement shown at (b), where repetition with reflection around the vertical axis produces bilateral symmetry.

FIGURE 6.14



This picture clearly shows black shapes on a white background. The black shapes are vertically oriented, symmetrical, small (relative to the background), and surrounded by the background.

# Figure-Fond



Rubin, 1915

# Figure-Fond

Fig. 7.2. At first you may see no sign of human occupation in this innocent island scene. However its title is 'St Helena', so look again. If you still have no luck, look for Napoleon in the space between the trees on the left. Once you have seen him, for how long can you look at the picture without him obtruding into your gaze? This illustrates first that mental 'set' is important in what you perceive, for without hints one can easily miss Napoleon altogether. Second, notice how Napoleon and the trees are alternative constructs for that part of the figure; you cannot see both simultaneously, and what you see tends to flip from one interpretation to the other. Gestalt psychologists would say that the trees are either 'figure', in which case you see them as such, or 'ground', in which case the space between them forms the figure. Finally you probably would not interpret the space between the trees as Napoleon unless you had previously seen a picture of him in this posture, so you are being influenced by memories from some time in the past. More factors influence perception than we intuitively believe.

**Figure-Fond**

# Figure-Fond

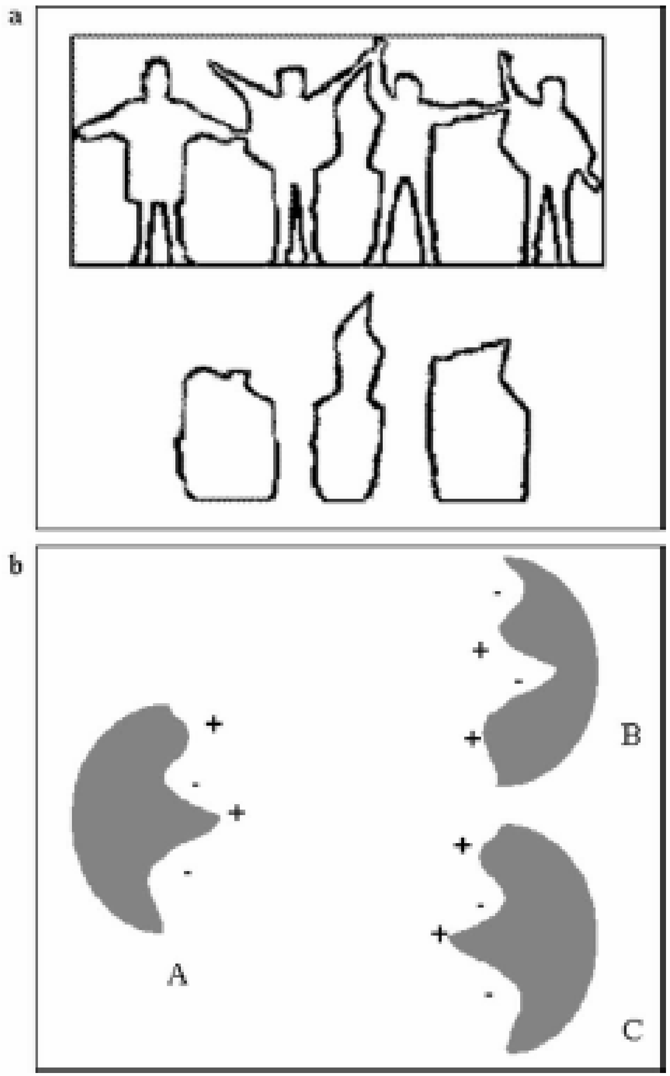
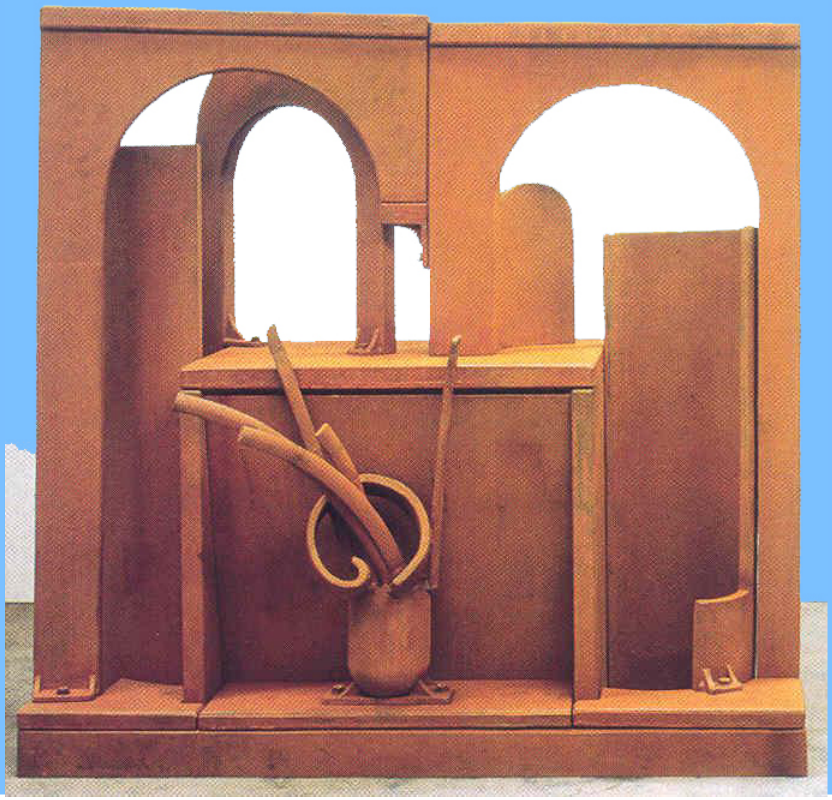



**Fig. 1. (a) Because of how we organize the top scene into figures and background, the ground regions between the figures do not have shape. Notice how the regions of background isolated below are shapes that we do not perceive in the scene above (if this is parsed as a group of human semaphore signals). This demonstration is based on a similar one in Kanizsa (1979). (b) Two halves of a disk have different shapes, even though they share the same bounding contour. This demonstration is based on Attneave (1974). The positive and negative curvature of the contour has been labelled to show that when polarity changes from A to B (because of a change of where the inside of the figure is) the perceived shape is different.**

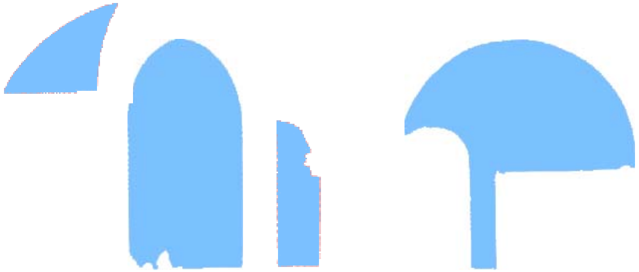
Marco Bertamini M. & Croucher C.J (2003). The shape of holes. *Cognition* **87,** 33–54

**Figure-Fond**

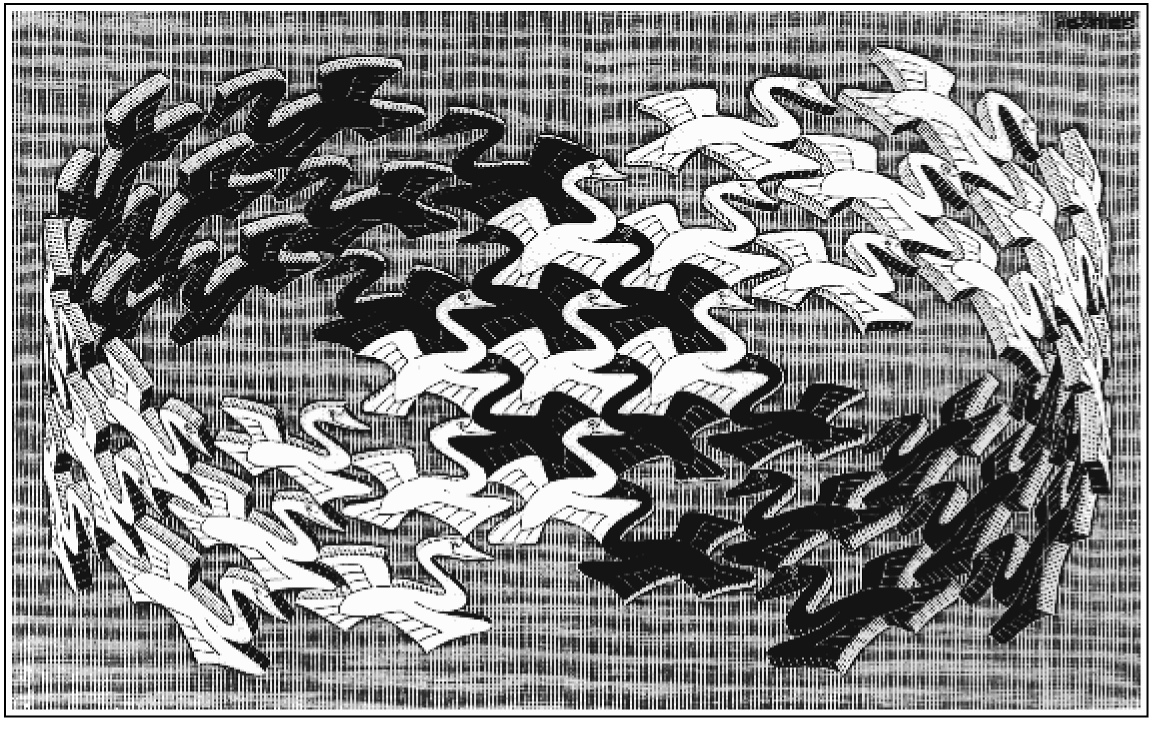
**GROUND**



**FIGURE**

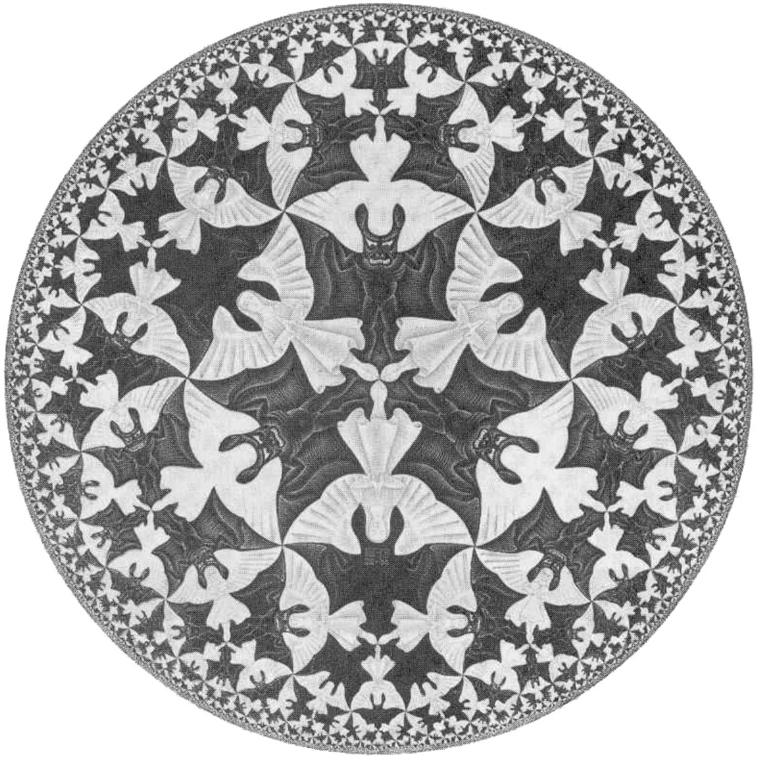


# Figure-Fond



Escher

# Figure-Fond



Escher

# Figure-Fond

Dali

# Figure-Fond

Dali

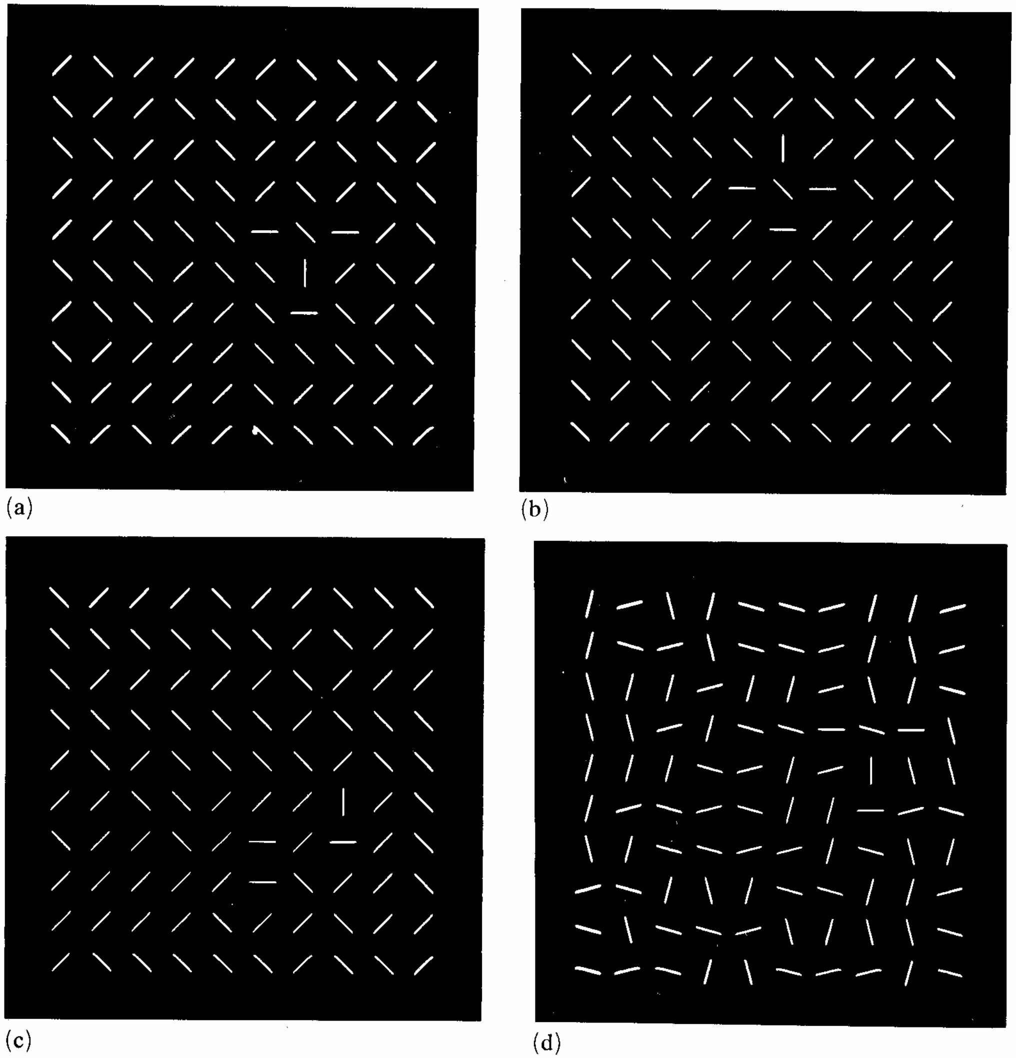
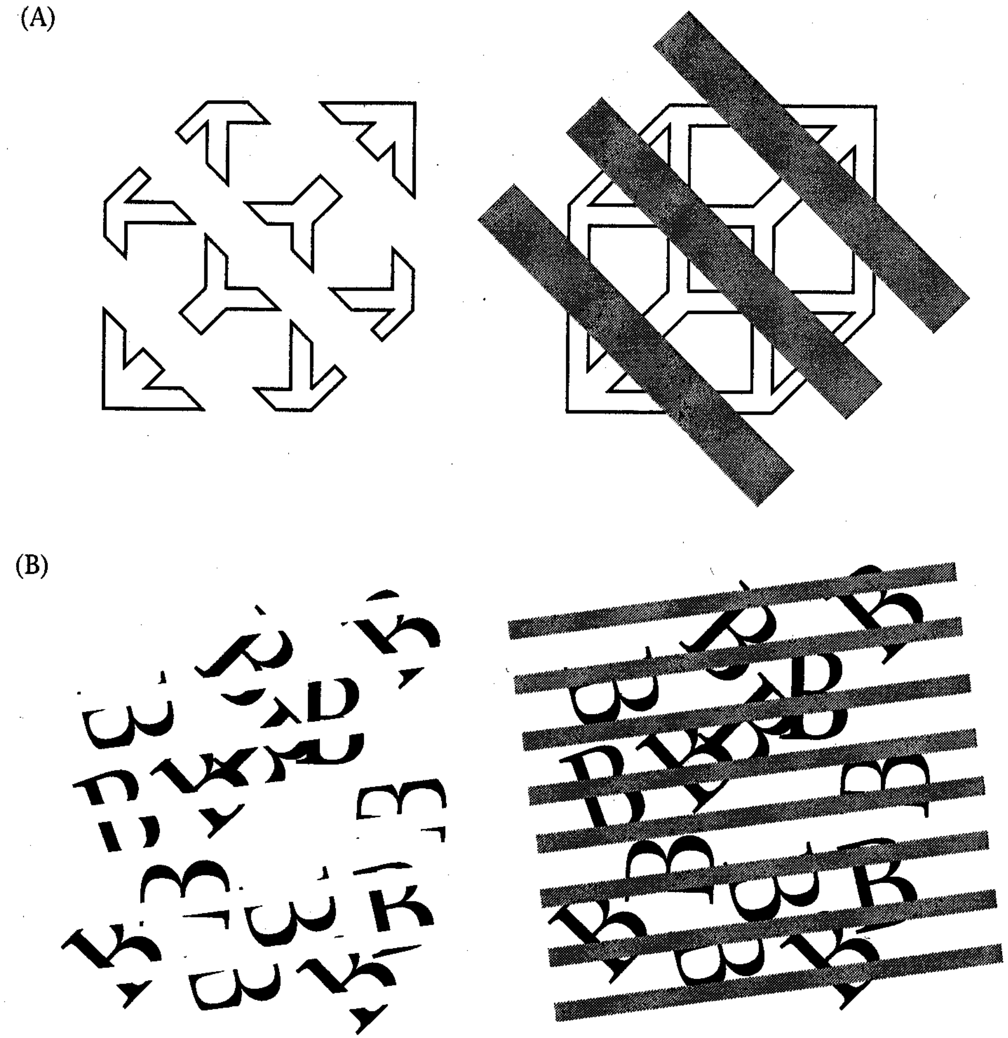



**Figure 1.** Examples of the stimuli used in the experiments. In (a), (b), and (c) target stimuli are embedded among lines oriented at ±45°. These angle differences between target and background elements induced 100% correct detection performances and were not used in the actual experiments. Stimulus (d), where the 'noise' elements are oriented at ±14° and ±76°, was used in the experiments. (a) and (d) show the 'face' stimulus (positioned in the middle-right part of the latter). The symmetrical 'nonface' and the asymmetrical stimuli are illustrated in (b) and (c) respectively.

# L'Entier (visage) détermine (la visibilité de) ses Parties (lignes)

Gorea A. & Julesz B. (1990). Context superiority in a detection task with line-element stimuli: a low-level effect, *Perception*, *19*, 5-16.

## Liage (*binding*) et Contexte




**11.8 OCCLUSION AND OBJECT RECOGNITION.** The presence of a clearly visible occluding surface helps us to integrate otherwise fragmentary image components. (A) When the line segments are seen without an occlusion cue, they appear as a set of uncorrelated two-dimensional patterns. By overlaying occluding boundaries, the pattern is seen as part of an object, namely, a three-dimensional cube. (B) When the pattern on the left is seen on its own, it appears as a jumble of unconnected curves and lines. By placing an occluding object over the white spaces, it is much easier to see that the occluded pattern is a collection of "B's." A after Kanizsa, 1979. B after Bregman, 1981.

**Liage (*binding*)**
**et**
**Contexte**

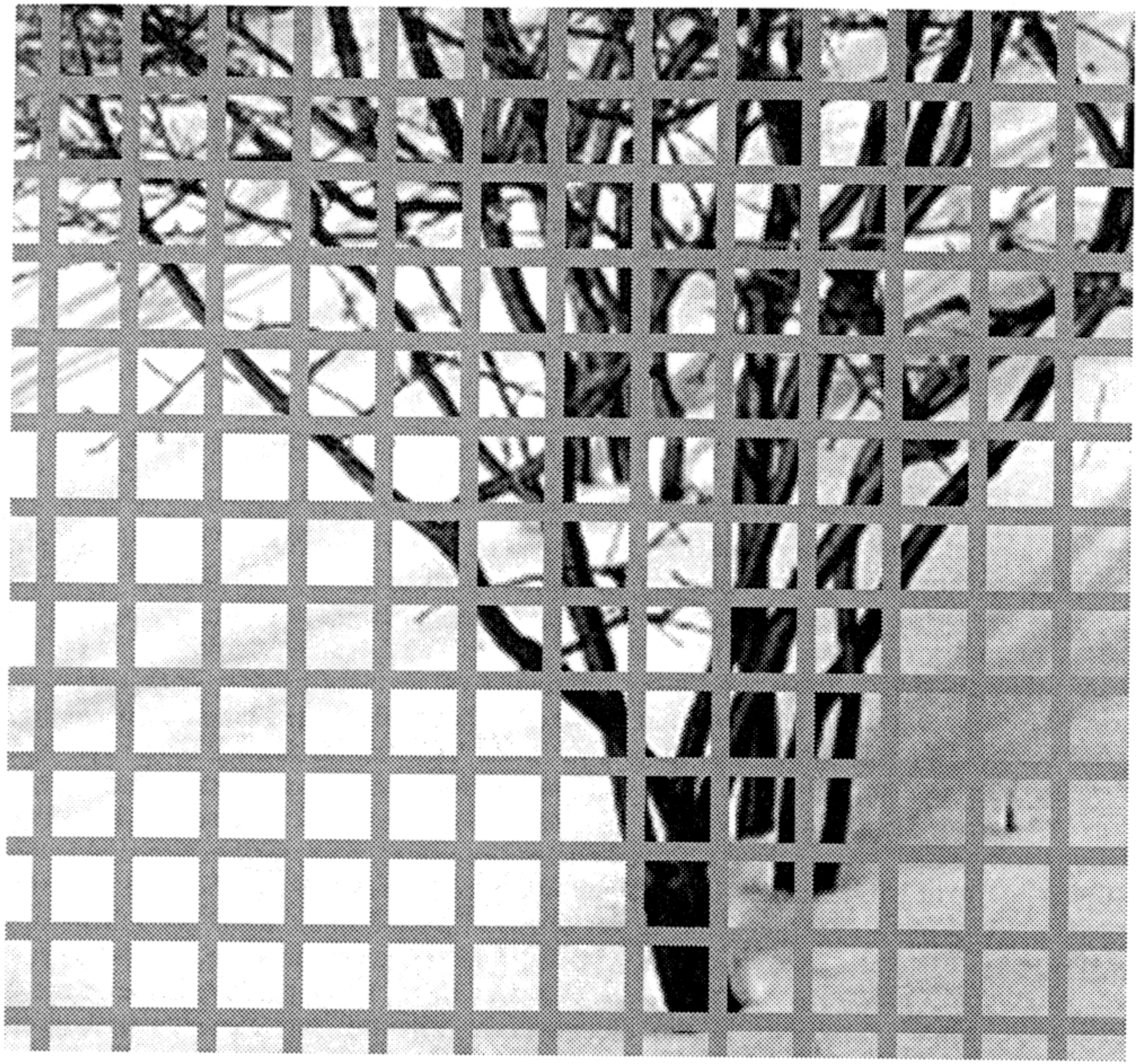


FIGURE 1. A demonstration of the phenomenon of "good continuation". The continuity of the branches is easily identified despite the presence of a disruptive grid.

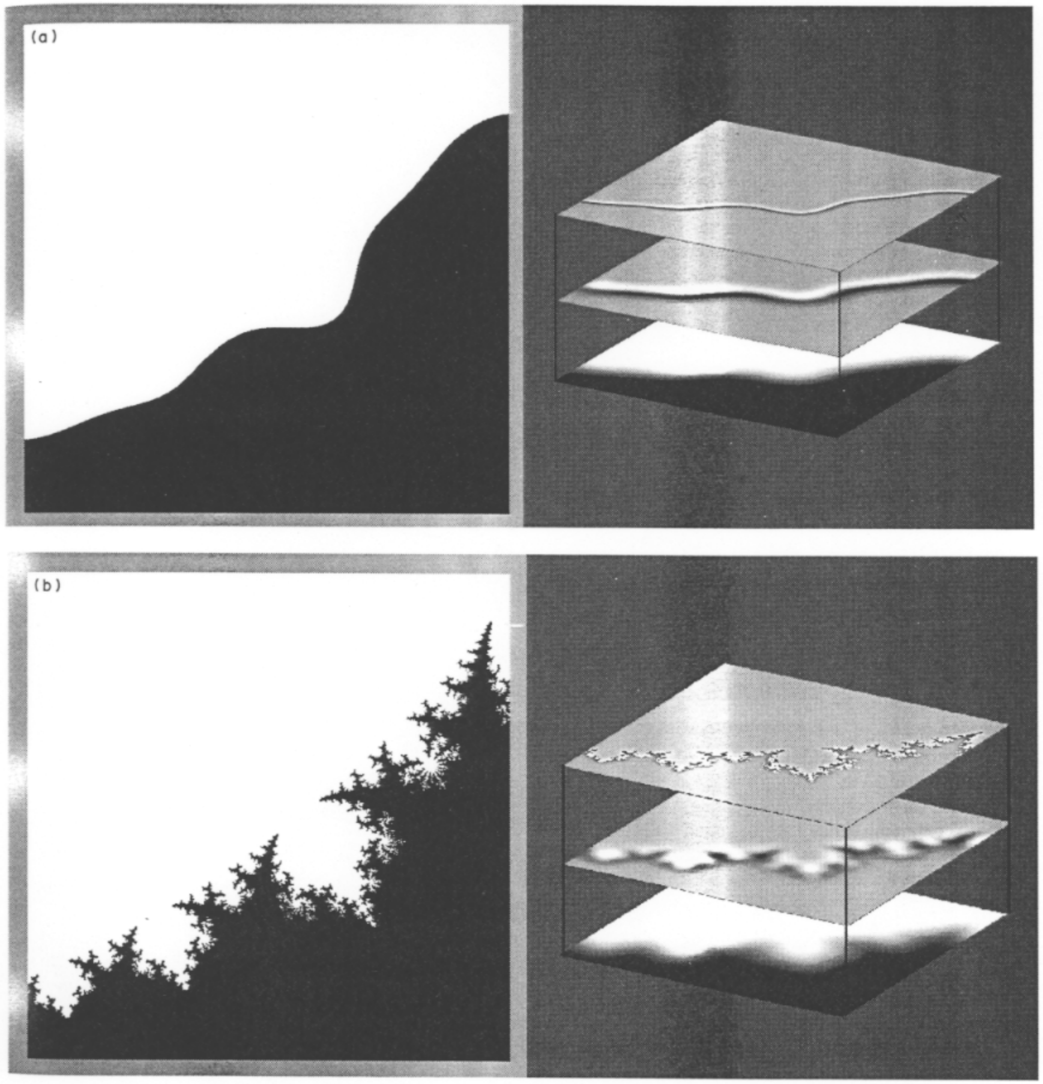



FIGURE 2. Smooth and jagged edges displayed at three spatial scales. (a) Presents a smooth edge (or alternatively a fractal edge with a low fractal dimension) and (b) presents a jagged edge (or alternatively a fractal edge—actually part of the Mandelbrot set—with a high fractal dimension). In the right panel of each figure the three tiers show, from top to bottom, the results of filtering each edge with bandpass filters of high, medium and low peak spatial frequencies. Notice that in (a) the position of the edge is aligned across the three scales. This is not the case for (b).

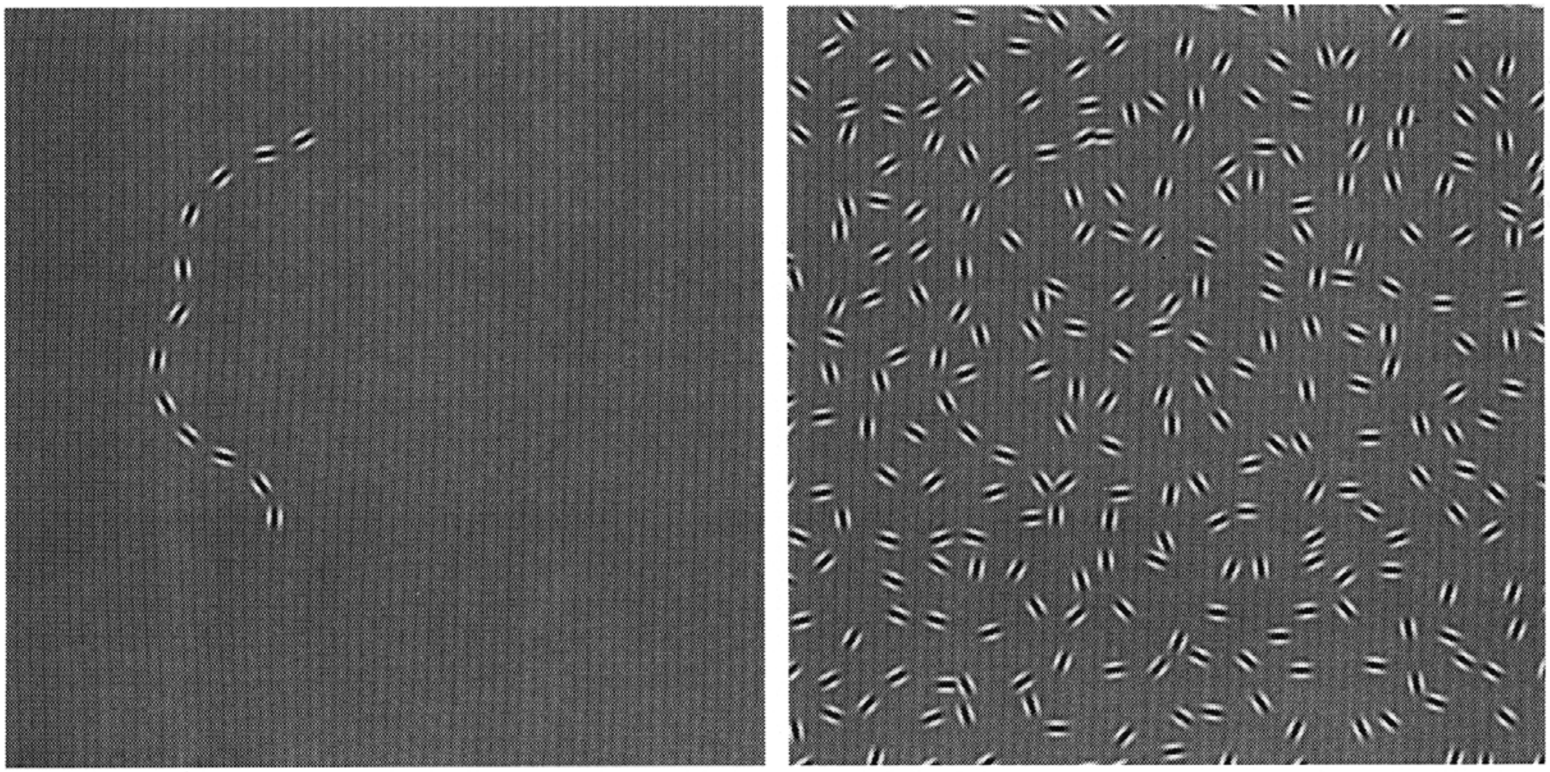


FIGURE 3. An example of a stimulus used in the experiments reported. The left-hand panel shows the path of elements (the stimulus) that the subjects must detect when embedded in an array of randomly oriented elements (the stimulus plus background shown on the right). In all experiments, the stimulus consisted of 12 elements aligned along a path. In this example each successive element differs in orientation by $\pm 30$ deg and for this difference in orientation the string of aligned elements is easily detected.

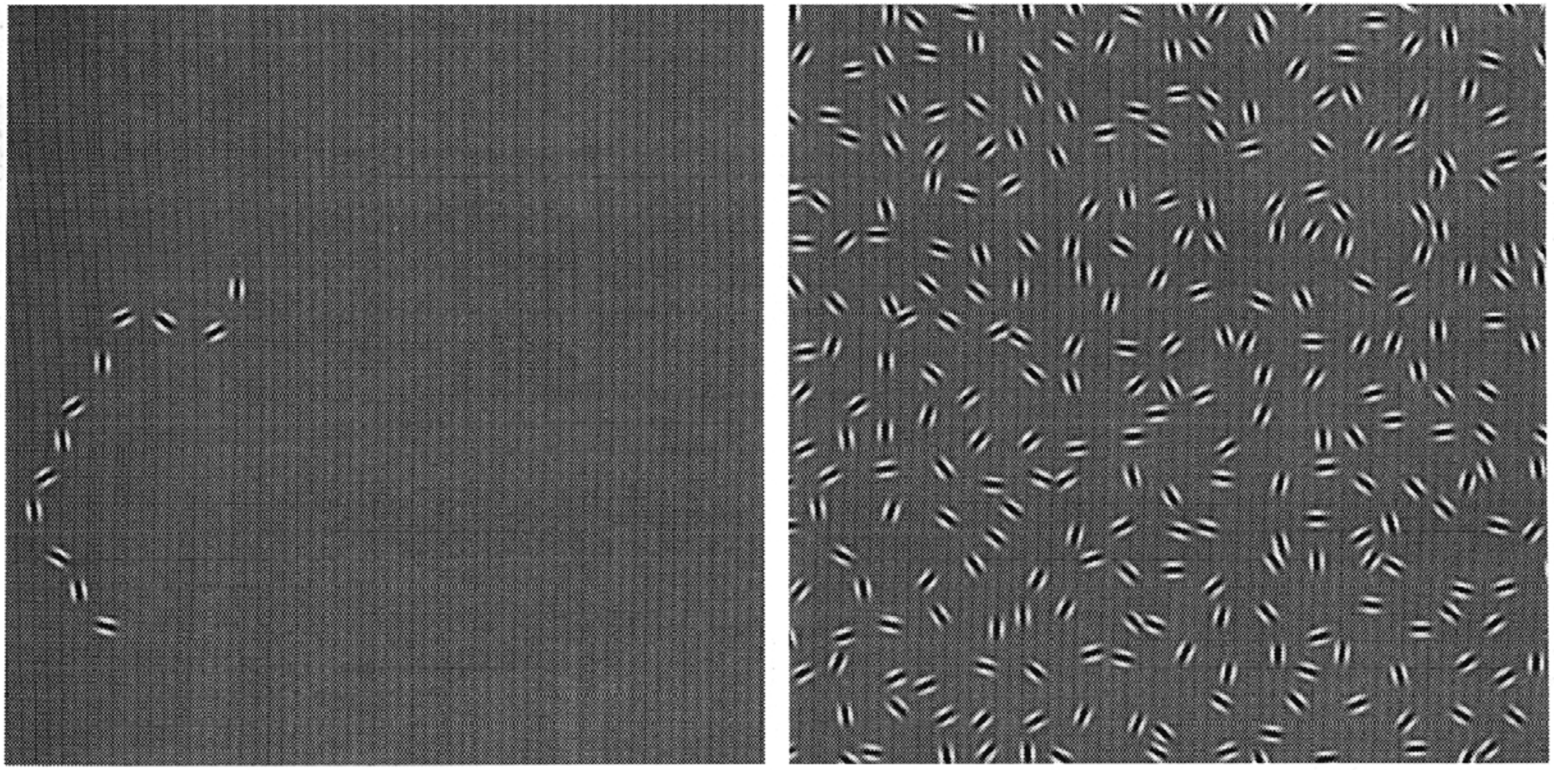


FIGURE 6. The right side of the figure shows an example of a stimulus used in Expt I. In this example, the path orientation variable $\beta$ has the values $\pm 60$ deg. Observers found the task of detecting the path considerably more difficult in this condition than in the condition shown in Fig. 3 ($\beta = \pm 30$ deg).

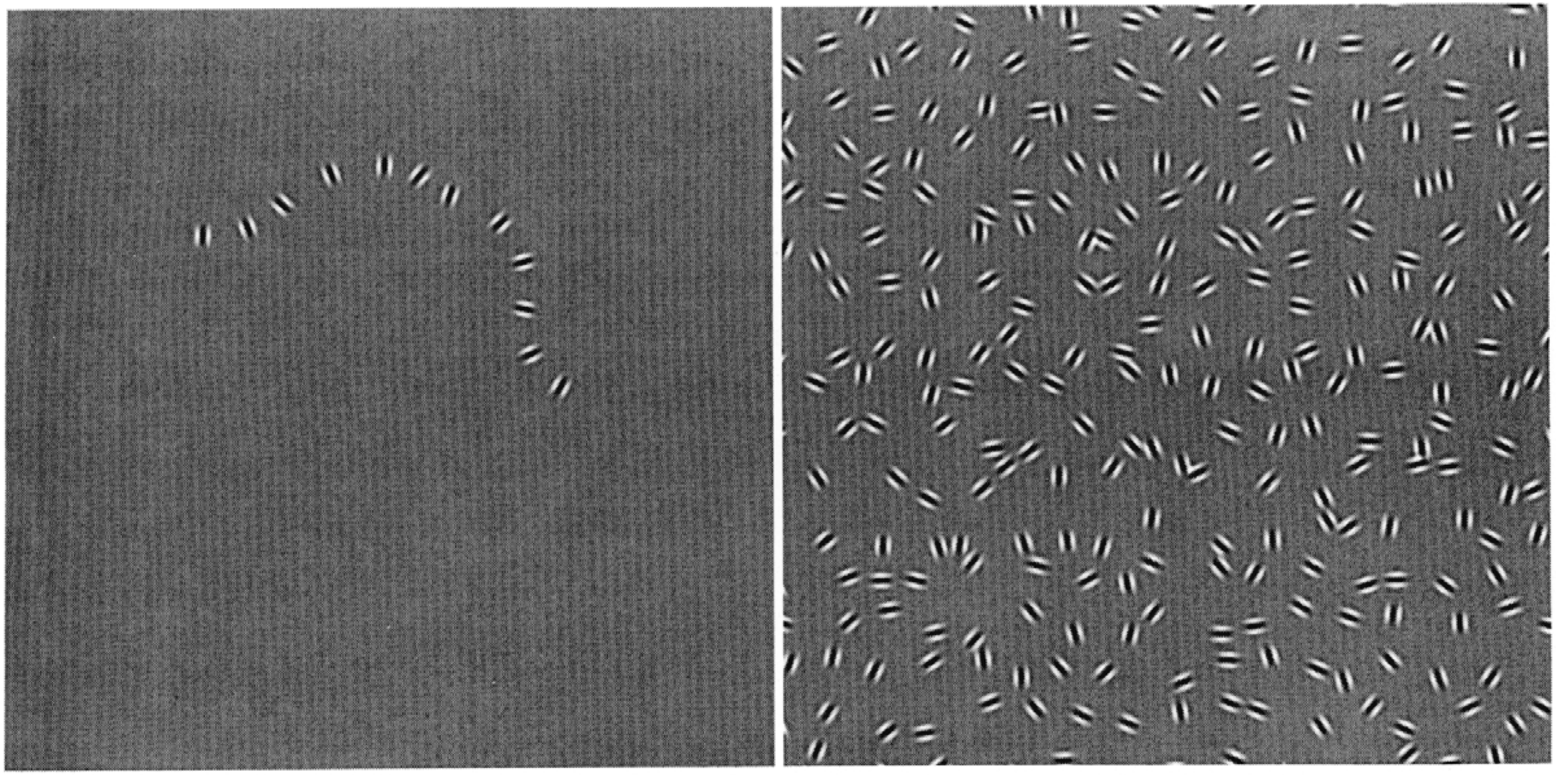
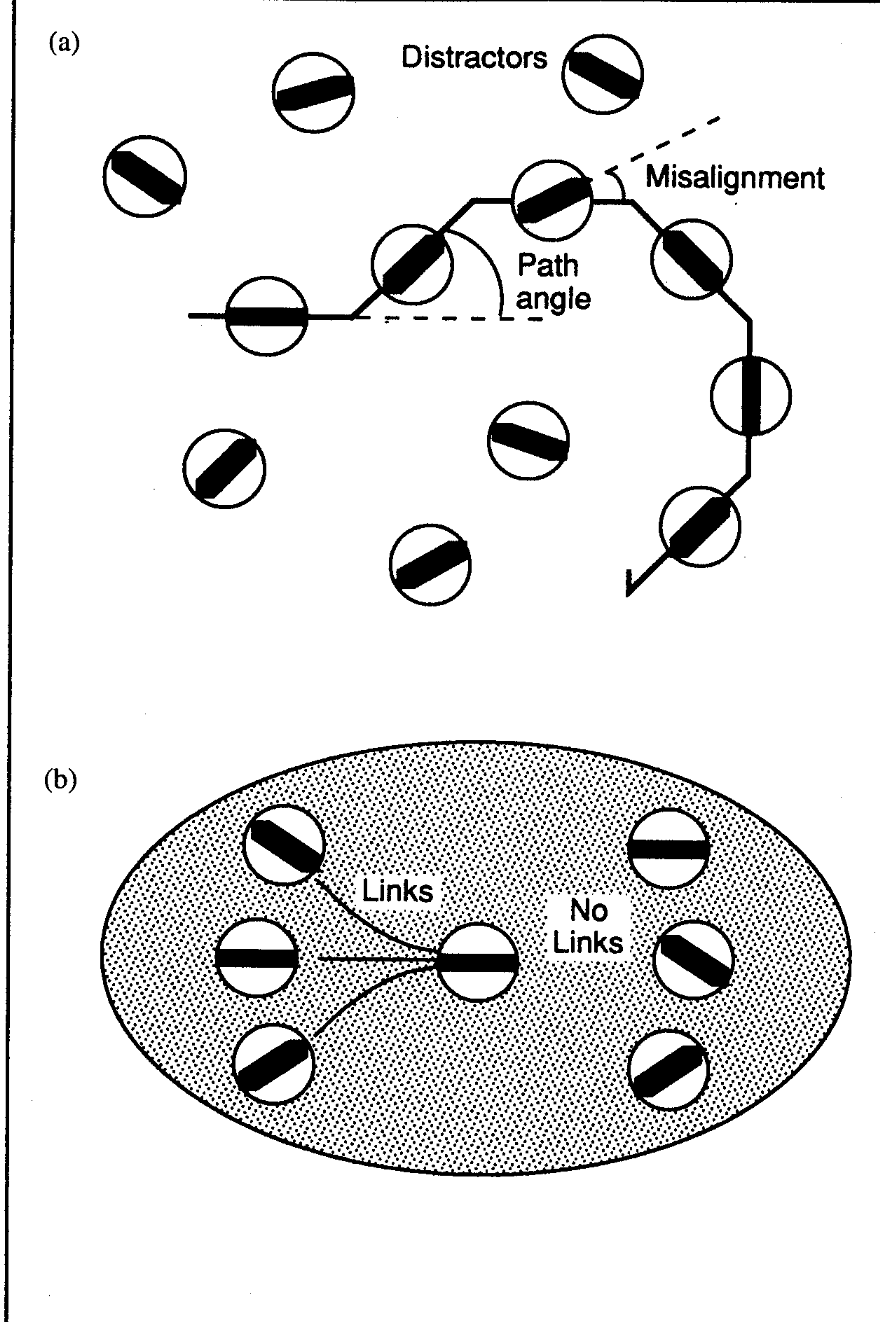


FIGURE 8. An example of a stimulus used in Expt II. In this example the path orientation variable $\beta$ has the values $\pm 30$ deg, and the orientation of the sinusoid of the patch ($\alpha$) is 90 deg with respect to the path. The path has the same parameters as the path depicted in Fig. 3, except the alignment of the Gabor patches is "side-to-side" instead of "end-to-end".

FIGURE 6.32




(a) Display used for research on "good continuation". Subjects had to detect the presence of a continuous path of elements embedded in a field of randomly oriented distractors. The path changed direction by a given angle from step to step, and elements might be misaligned from the local direction of the path. Lines joining the path elements were not present.
(b) Association field proposed to account for the finding that alignment of elements along the path was the crucial factor in detecting its presence. Adjacent elements would be "linked" if a smooth curve could be drawn between them. After Field, Hayes, and Hess (1993).

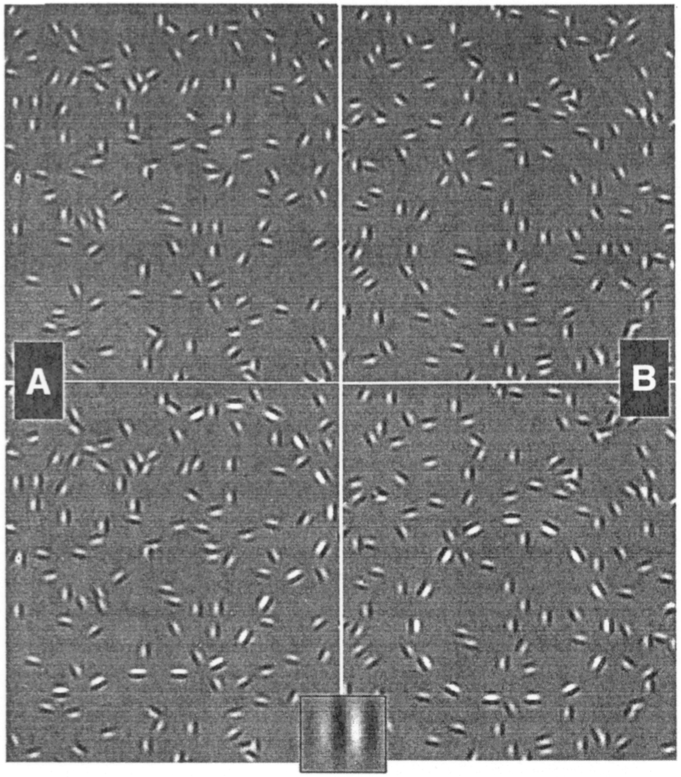



**Figure 2**
Superiority of closed versus open curves. (*Upper*) Two contours embedded in the background of randomly oriented elements. (*Lower*) The same contours are highlighted for didactic reasons. (*A*) A nonclosed contour composed of aligned Gabor patches (GPs) is only barely visible against the background. (*B*) A closed contour with the same angular difference and distance between elements is perceived much more easily. Perception of closed contours is best for brief presentations. For more than 180 msec duration, the observer starts to scrutinize other global structures at the expense of the primordial closed contour. (*Inset*): One GP element, which is a product of a sine wave luminance grating and a circular Gaussian envelope. GP wavelength ($\lambda$) was 0.12 arc deg; Gaussian envelope size was equal to $\lambda$; GP amplitude was 24% of mean luminance (30 cd/m$^2$). (From Kovacs and Julesz, 1993.)

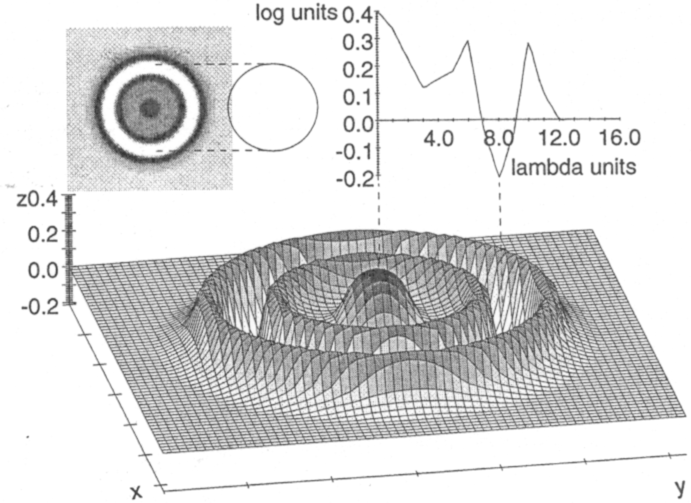



**Figure 3**
Sensitivity change for a single GP as a function of distance from the surrounding line (closed circle). Note the strong threshold enhancement effect inside the circle between 5 and 8 $\lambda$ distance. $\lambda = 8$ represents the center of the circle. (From Kovacs and Julesz, 1993.)

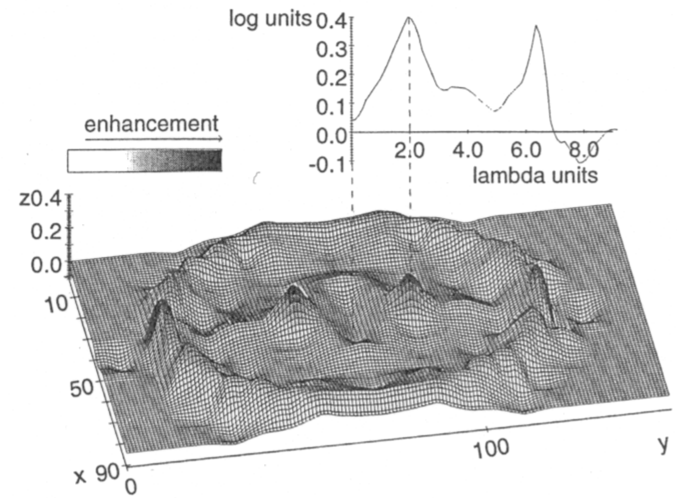



**Figure 4**
Similar to figure 3 but for a closed ellipse with 1.2 aspect ratio. (From Kovacs and Julesz, 1993.)

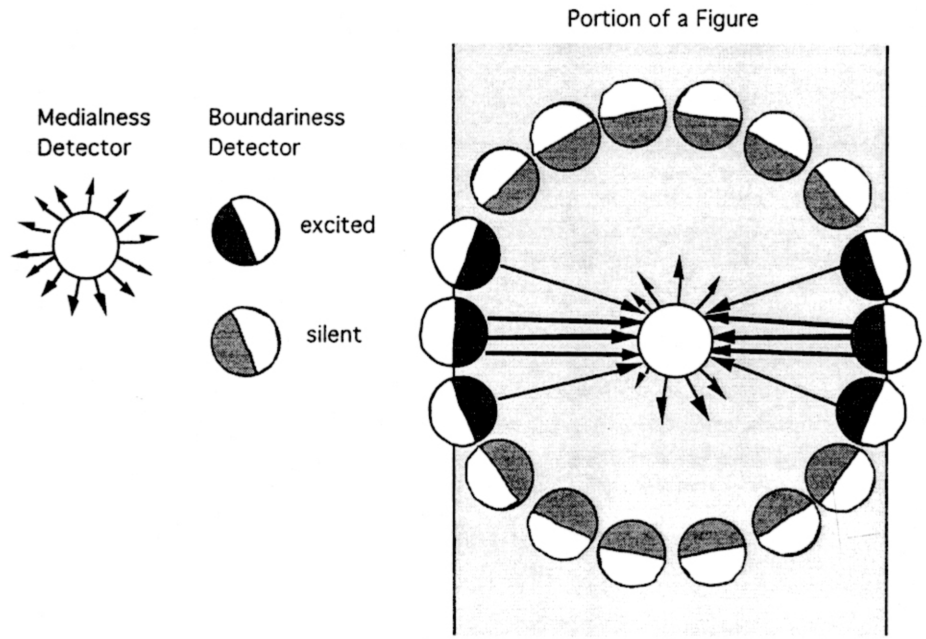



FIGURE 10. Representation of possible excitatory and inhibitory connections that result in enhancement of the core. A medialness detector is postulated to have excitatory connections to all nearby medialness detectors of the same scale (and inhibitory connections to medialness detectors of substantially different scales). The arrows pointing outward from the medialness detector represent the excitatory connections. The length of each outward arrow symbolizes its strength. The other medialness detectors are not shown in this figure for clarity of presentation. All boundariness detectors that can contribute to a given medialness detector are shown by the ring of bipartite circles. Those that are excited vote for the medialness detector, as indicated by the centrally-pointing arrows, and they inhibit the excitatory connections from the medialness detector to other same-scale medialness detectors in the region near which they attach, as shown by the reduction in the arrows pointing outward from the medialness detector. The result is that this medialness detector will excite most strongly those medialness detectors that are of similar scale and that are located in a direction that is consistent with the angle of the boundariness detectors contributing to the excitation of this medialness detector. In this case, this would be medialness detectors immediately above and below the one shown in the drawing.

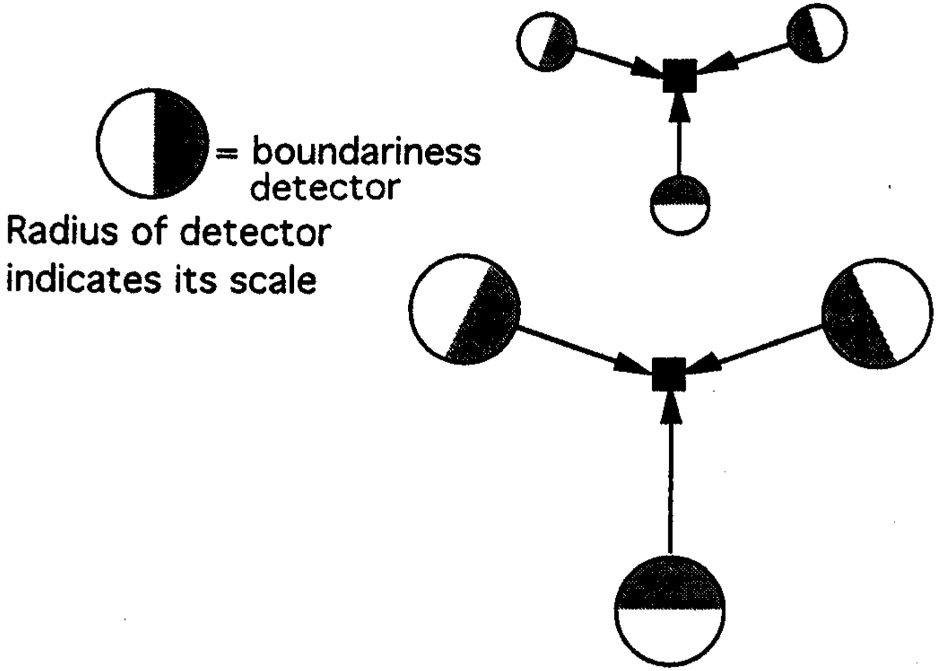



FIGURE 3. Boundariness detectors of the same scale interact at a distance proportional to the scale of the detector and along directions normal to the optimal boundary orientation for each boundariness detector.

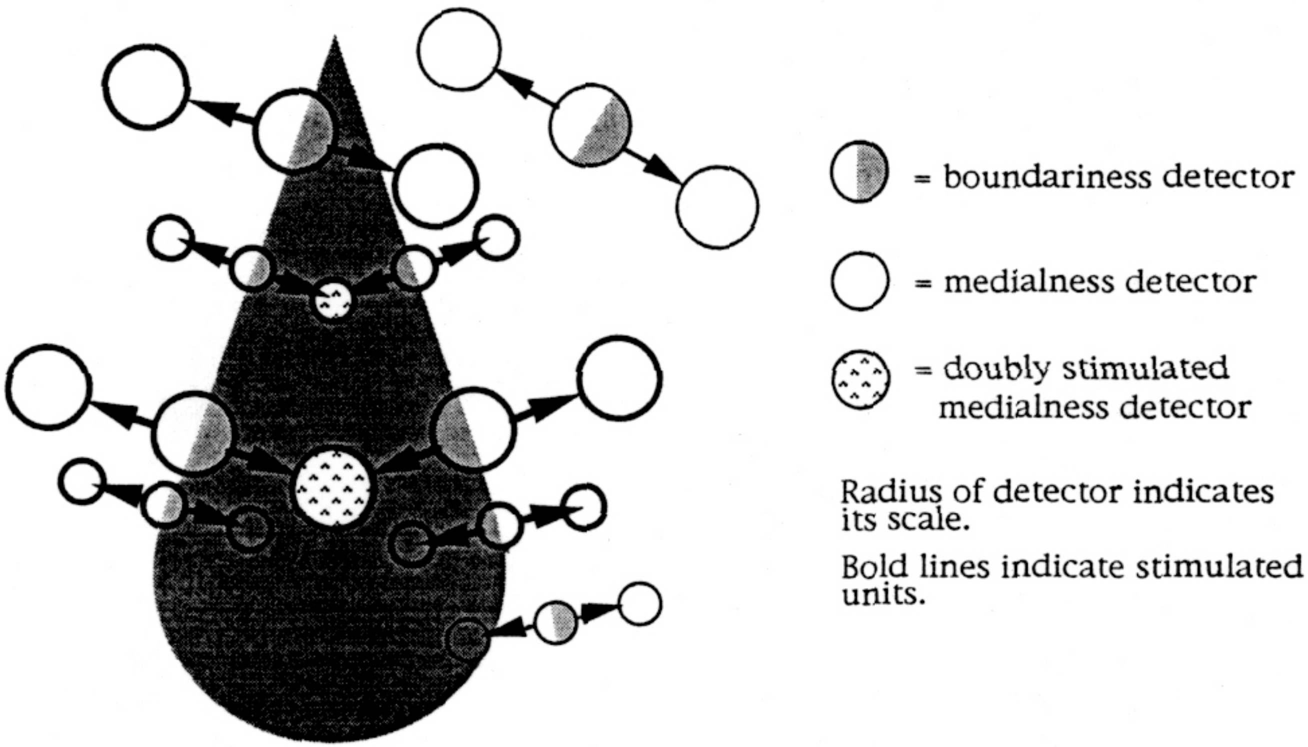



FIGURE 4. Boundariness detectors combining (or failing to combine) to produce strong medialness on a teardrop-shaped region.

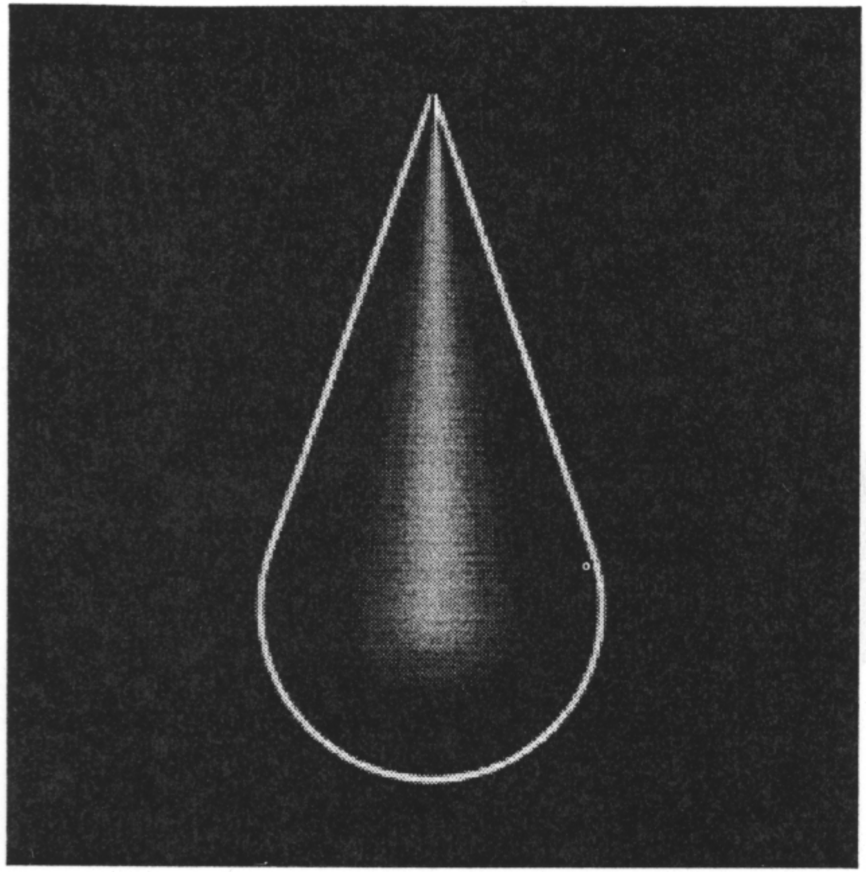
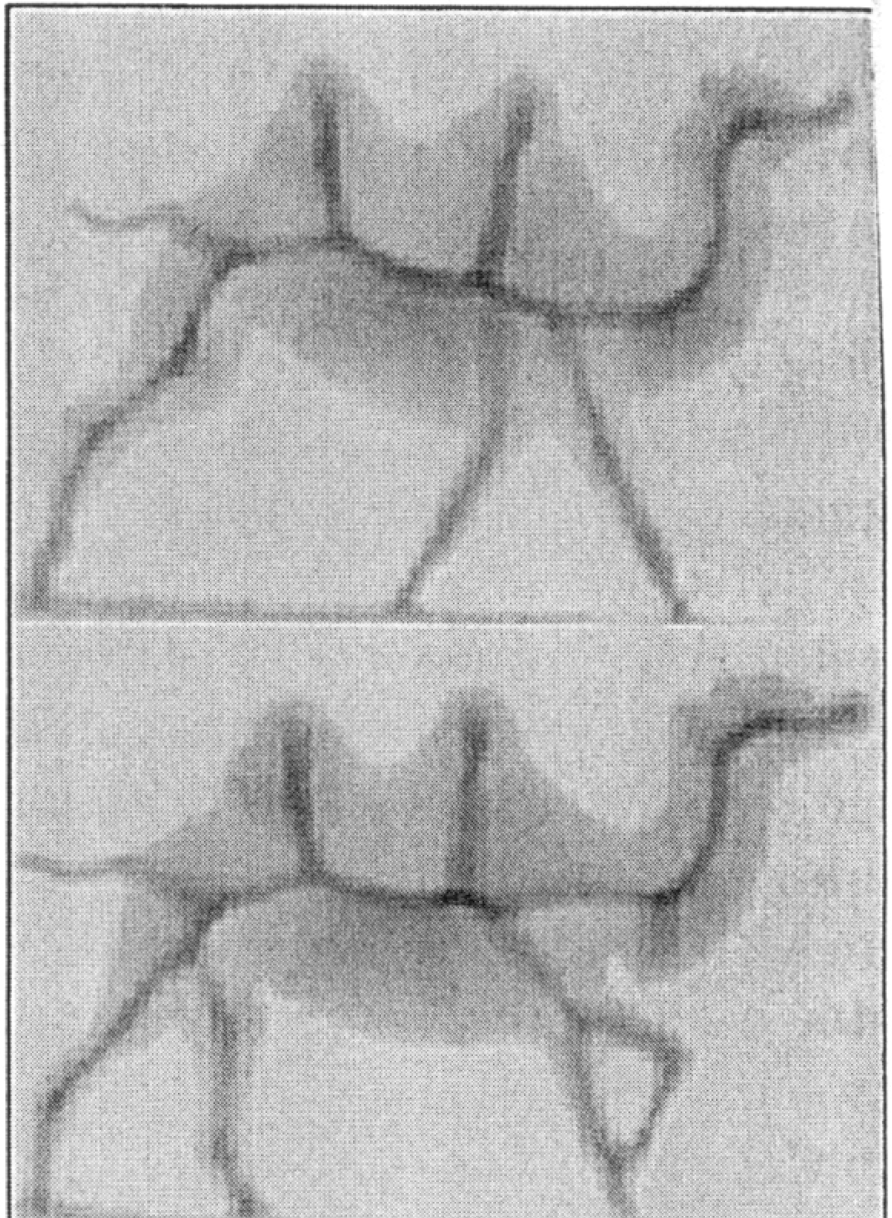
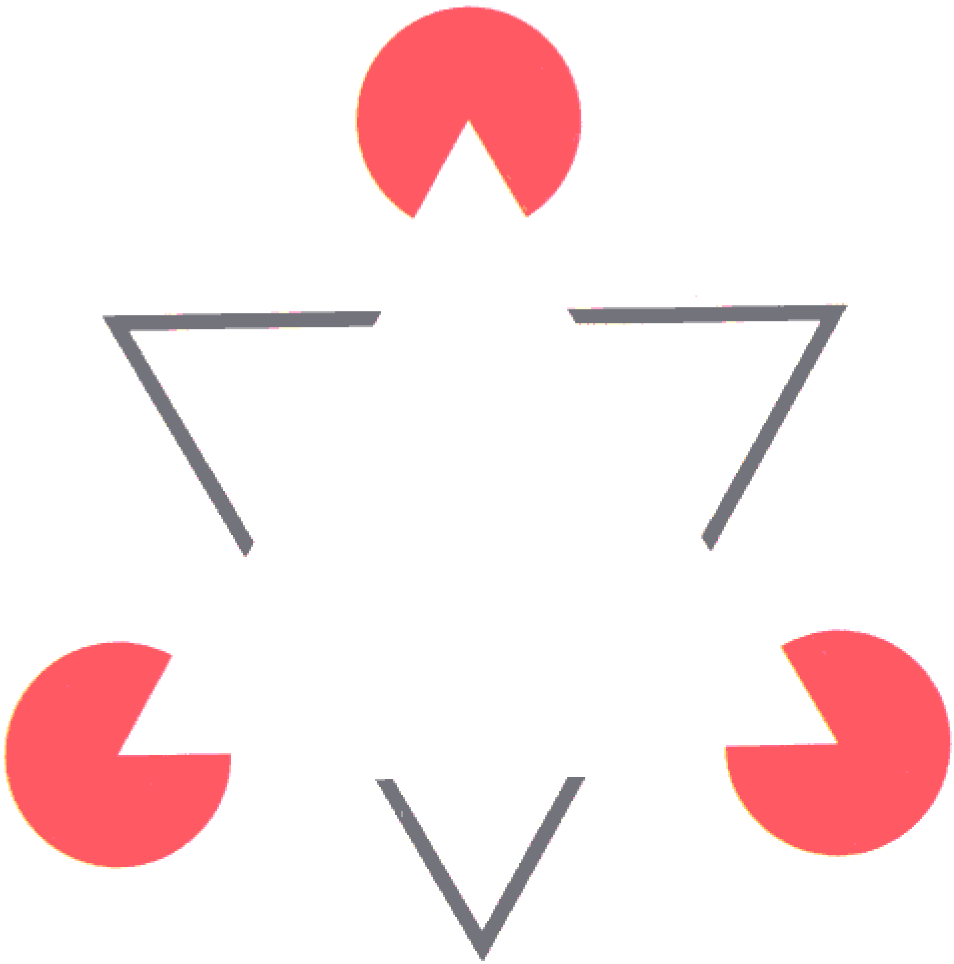
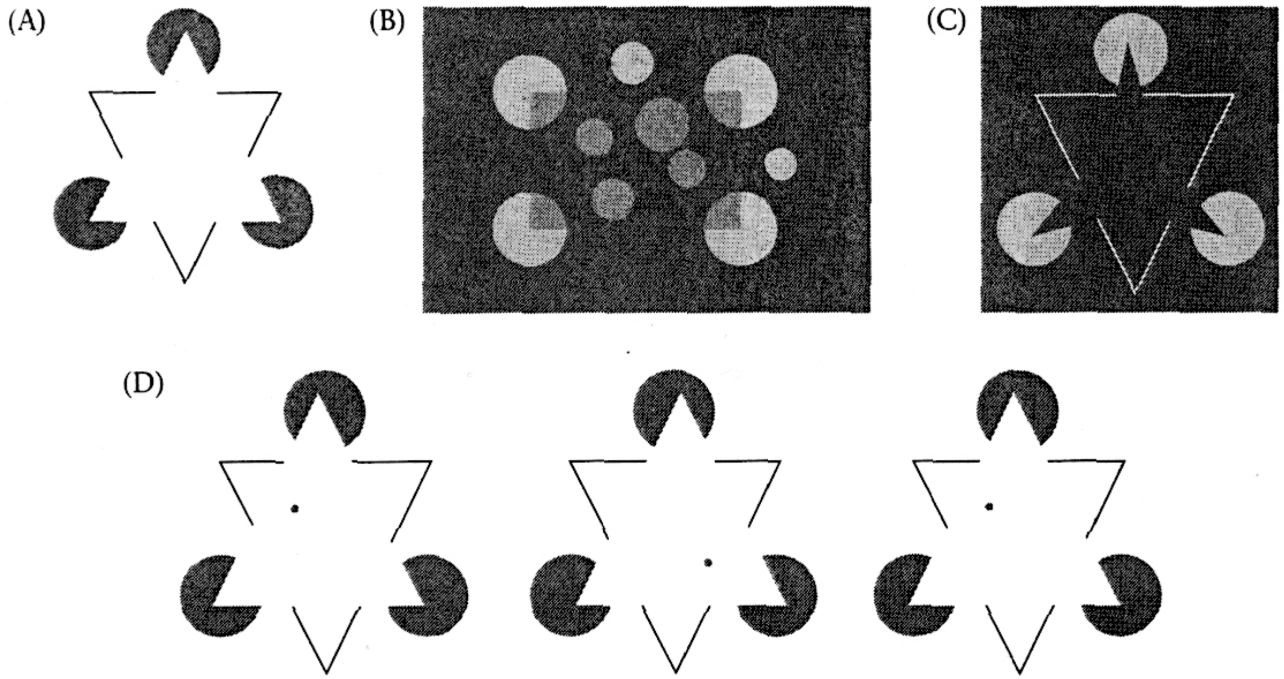


FIGURE 5. A teardrop-shaped figure (shown here by its boundary) and a representation of its core, with core strength being indicated by intensity. Because a core is a locus in 3-space, it is awkward to represent in 2-D. We use two conventions: (1) a fuzzy core in the image plane, where the core's width indicates the figure's width, as in this figure; and (2) a trace in scale space, where the height indicates the width (Fig. 7). We do not yet know what the width of a given core might be; we know only that for this to be an accurate representation of the visual percept, it must be proportional to the figure's width. Thus the exact width (or height) depicted in the figures is arbitrary








**11.7 SUBJECTIVE CONTOURS.** These subjective contours are inferred from occlusion and transparency cues in the images. (A) A triangle is suggested by occlusion, (B) a rectangle is suggested by transparency, and (C) a curved object is suggested by occlusion. (D) Stereo pairs of subjective contours. By diverging your eyes beyond the page, the image pair on the right (left) will fuse and you will see the subjective contours of a triangle in front (behind) of the circles. The subjective contour is somewhat more vivid when the depth cue is added. If you converge your eyes to fuse the images, the depth relationships will reverse. *A–C after Kanizsa, 1976. D after He and Nakayama, 1994b.*

*Figure 3–1.* The interpretation of some images involves more complex factors as well as more straightforward visual skills. This image devised by R. C. James may be one example. Such images are not considered here.