

Philips Research Laboratory
Avenue Albert Einstein, 4
B-1348 Louvain-la-Neuve, Belgium

Working Document WD65

A twofold model of edge and feature detection

C. Ronse

September 1990

Abstract. Horn's model of surface reflectance shows that edges in three-dimensional surfaces lead to grey-level edges combining in various ways sharp or rounded steps, lines and roofs. The perceptual analysis of extended edges necessitates the localization not only of step and line edges, but also of roof edges and Mach bands, and more generally of discontinuities and sharp changes in the n -th derivative of the grey-level. Arguments are given which indicate the inadequacy of locating features at zero-crossings of any type of smooth operator applied to the image, and the necessity of orientationally selective operators. The null space of feature detection is defined; it contains in particular all constant signals. Oriented local features are modelled as the linear superposition of a featureless signal (in the null space), an even-symmetric and/or an odd-symmetric feature, measured by convolution with respectively even-symmetric and odd-symmetric functions. Advantages of energy feature detectors are given.

Key words. Edge types, zero-crossings and peaks, orientational selectivity, linear processing, feature symmetry, energy feature detector.

1. OPTICAL CHARACTERIZATION OF EDGES

An image is formed as light coming from certain sources is reflected towards the viewer by surfaces in three-dimensional space. Horn [9] analysed the optical generation of image intensities under some simplifying assumptions. First the reflectance of a surface element is *achromatic*, it does not depend on the frequency spectrum of light. Second, it is a multiplicative factor of incoming light intensity. Third, the reflections of intensities from several light sources are combined additively. Let us call *illumination* the function giving for each orientation the power of light flow by unit solid angle and unit surface perpendicular to that orientation. Then the illumination function of light reflected at a point on a surface will be given in terms of the surface reflectance and illumination by the formula

$$I_r(e) = \int_{S^+} R(i, e) I(i) d\omega,$$

where i denotes the orientation of incoming light, e that of reflected light, I is the illumination function of incoming light, I_r that of reflected light, $R(i, e)$ is the reflectance function for light coming from i and reflected towards e , ω the measure of solid angle, and the integration is made on the half-sphere S^+ of all possible orientations i of incoming light.

The reflectance function is usually a linear combination of two basic ones: *Lambertian* and *specular* [9]. Lambertian reflectance is that of a pure matte surface, which looks equally bright when seen from any direction; moreover its brightness is proportional to the amount of light it receives. Thus we have $R(i, e)$ proportional with $\cos \alpha_i$, where α_i is the angle between an illumination orientation i and the surface normal. Specular reflectance is more or less like that of a mirror: for i fixed, $R(i, e)$ is strongly localized, with its energy concentrated on values of e around i' , the orientation opposite of i w.r.t. the surface normal. A small amount of specular reflectance makes a matte surface glossy.

The illumination of a scene has two components: the *primary* one where light from a source hits directly a surface, and a *secondary* one where light from a source hits a surface only after having been reflected by one or more surface portions. As reflectance is multiplicative, for a fixed scene geometry, the ratio of secondary to primary illumination increases with surface albedo. This is why the image of a black scene under strong light is distinct from that of white scene under weak light [8]. Usually there is a single light source, so that primary illumination is *coherent*, in other words its contribution to the illumination function $I(i)$ is concentrated around a single value of i (the orientation of the light source w.r.t. the surface). In indoor scenes secondary illumination, especially from walls and ceilings, becomes *diffuse*, which means that its contribution to the function $I(i)$ is relatively flat (see [31], Chapter 1). In both outdoor and indoor scenes there is a component of secondary illumination which is not diffuse, namely the *mutual illumination* between *neighbouring surfaces* forming a *concave edge*.

Grey-level features arise mainly from two factors. First the changing patterns of reflectance on a surface, indicating changes in the material composition of that surface; this is for example the case for many texture edges. Second, the interaction of features of the scene geometry with illumination, in relation to the viewpoint: edges and corners in the three-dimensional shape of a surface, occlusion, shadows, etc. (however diffraction of light around

objects is visually negligible, see [2], Chapter XI, Section IV). Note that the two types of features can interact: for instance changes in the scene geometry lead to modifications in the appearance of a texture.

Shadows give a dark grey-level profile whose contour is an *extended edge*, namely a very gradual ramp: see Figure 1.1. This will be discussed further in the next section. Edges arising from the scene geometry are mainly of three types: *convex*, *concave*, and *occluding*. We illustrate them in Figure 1.2. As pointed out in [9], the angle made by a convex or concave edge is not completely sharp, but rounded, although this appears sometimes at a small scale. Note also that an occluding edge depends on the viewpoint, and can be combined with a convex edge.

Under a totally diffuse illumination ($I(i)$ flat), for both Lambertian and specular reflectances the geometry of the scene does not influence the grey-level profile, unless it combines with local reflectance patterns (e.g., texture); hence only edges involving reflectance changes are seen. When the illumination is relatively diffuse, shadows and geometrical features of the scene have dim grey-level profiles. Let us now consider grey-level profiles arising from geometrical edges under coherent illumination. We first assume that the surface is of uniform Lambertian reflectance, and we will examine later the effects of specularity.

Consider a convex edge with two faces A and B , and assume a single light source, as shown in Figure 1.3(a). As the angle between A and B makes a small round, the grey-level profile makes one of the following features, according to the orientation of the light source: a positive line edge (when both faces are equally lit), a slightly rounded step edge, or a linear combination of both (see Figure 1.3(b)).

For a concave edge, the grey-level profile arising from the illumination by the light source is the same as for a convex edge, but with the two sides interverted. It is known that a concavity on a surface illuminated from one side looks nearly the same as the mirror-symmetric convex bump illuminated from the other side. However in a concave edge the mutual illumination between the two sides gives rise to a roof edge in the grey-level profile, whose orientation is opposite to that of the step edge due to primary illumination. See Figure 1.4.

If the surface of an occluding edge does not face the light source, it may project a shadow on the surface behind. Otherwise, the grey-level profile depends on whether the occluding edge is combined with a convex one. For an occluding edge without convex edge, the grey-level profile is a sharp step edge, or an extended round ramp (when the occluding edge and the occluded surface are equally lit): see Figure 1.5. Note that when the occluding edge has a relatively small radius of curvature, we get a sharp step followed by a round profile, as in Figure 1.6. If the occluding edge is combined with a convex edge, then we have a succession of grey-level profiles corresponding to both types of edges, namely a sharp step (or nothing) followed by a linear combination of a step and a positive line, as shown in Figure 1.7. Note that we can obtain in this way a negative spike edge, or its superposition with a step edge (see Figure 1.8).

For a unique punctual light source, specularity leads to positive peaks or line edges in the grey-level profile (see Figure 1.9). If the illumination is more diffuse, we get roof edges

instead.

These differences in grey-level profile may help to distinguish between the different types of three-dimensional edges of a scene from their appearance in the image [9]. In particular a sharp step is more likely to correspond to an occluding edge, while a slightly rounded step or a positive line suggests a convex edge, or a concave edge if it is superposed on a roof.

Let us now give three classifications of grey-level edge types from the literature. In Figures 1.10 to 1.12 we reproduce the models of Herskovits and Binford (from [9]), Nalwa and Binford [24], and Ling [15]. The ones of Herskovits-Binford and Ling were obtained experimentally as the most frequent one-dimensional profiles in real images. Note first that in both Nalwa-Binford and Ling, a line edge is represented as a rectangular bar, although from the above discussion on optics it should rather be a triangular or round peak, as in Herskovits-Binford; we will discuss bar edges in more details in the next section. Second, what Ling calls a convex edge is classified by Nalwa-Binford as a step edge; we prefer to maintain this distinction, because such a round grey-level edge arises through the occlusion by a curved surface, as in Figure 1.6. Finally Ling's valley edge corresponds to the negative spikes of Figures 1.6 and 1.8, due to occlusion. Thus Ling's classification seems the most detailed of all three.

2. PERCEPTUAL INTEGRATION OF EXTENDED EDGES AND MACH BANDS

A sharp edge as the step and line of Figure 2.1(a) is unambiguously located, and most algorithms succeed in this task. However, the edge may be gradual, as in Figure 2.1(b), where the step forms a short ramp, and the line is a bar with a short plateau. In this case it is customary to locate the edge at the point p in the middle of the ramp or plateau. However for an extended edge as in Figure 2.1(c), this localization at p becomes completely arbitrary. Indeed, the signal is linear in a small neighbourhood around p , so that there is no local feature at that point. One might say that p locates an edge at a large scale, that can be obtained by convolving the image with a wide mask. However this smoothing will blur the features and lead to an imprecision in the localization of p . Moreover this localization is sensitive to low frequency noise; in Figure 2.1(d) we show the superposition of two low amplitude wide sinusoids of opposed sign on the edge, and each of them shifts the position of p to one extremity of the edge (p_1 and p_2 respectively). This imprecision in the localization of p can cause a classical edge detector (based on gradient maxima) to oscillate on an extended edge, giving thus an ugly edge map (while the gradient image is visually satisfactory). We conclude then that in Figure 2.1(c) there is no local feature at p , but rather two features at a and b ; in other words the extended step becomes a succession of two ramp edges (sharp steps in the gradient), while the extended bar becomes a succession of two step edges.

Thus we must include as local features not only step edges (where there is a sharp change in grey-level) and line edges (where there is a sharp change in the grey-level integral), but also ramp edges (where there is a sharp change in the grey-level gradient). Note in particular that a triangular delta or rounded peak (cfr. Figures 1.3(b) and 1.10) is a better model of line edge than a rectangular bar (as in Figures 1.11 and 1.12). Indeed, when the

bar is narrow, it is nearly the same as the rounded peak; on the other hand when it becomes wide (as in Figure 2.1(c)), it must be classified as a succession of two step edges.

The feature formed by the junction of a grey-level ramp and plateau is called a *Mach band*; at the extremity of the plateau a narrow line is seen, which is light when the plateau is at the top of the ramp, and dark when the plateau is at the bottom of the ramp. As remarked in [23], such lines are also perceived at positive and negative roofs, for example in a triangular wave. We show in Figure 2.2 some roofs and Mach bands; those in the left column are classified as positive, because they lead to the perception of a light line, while those of the right column are classified as negative, with a dark line perceived. Clearly each class can be obtained as the addition of a symmetric roof in that class and a linear signal.

A common interpretation of Mach bands is that they are visual illusions produced by the mechanism of *lateral inhibition*. As we must recognize the albedo of objects under various illumination intensities, the visual system does not measure the absolute luminance of each point, but luminance contrasts between neighbouring areas. Thus the grey-level of each point is compared to those in its neighbourhood. Then at the location of a positive roof or Mach band, the grey-level is higher than the average in its neighbourhood, leading to the perception of a light line, and conversely for the dark line seen at a negative one. There are several arguments against this explanation. First, experiments by Morrone et al. [23] on periodic vertical gratings (images whose grey-level is constant in the vertical direction and periodic in the horizontal one) have shown that the sharpness of the perceived Mach band is not related to the sharpness of the angle between the ramp and the plateau in the grey-level profile; it rather depends on Fourier phase characteristics of the image. Second, lateral inhibition should also apply to colours, since the spectrum of sunlight varies a great deal between morning and evening, and we can still perceive the colour of objects and distinguish hues which differ by much less than the daily variation of the sunlight spectrum [1]. However in *isoluminant chromatic* images (made with colours having all the same lightness), Mach bands are not seen. This was found by coworkers of Koffka (see [12], pp. 170-171), and has been verified in recent experiments (D. Burr, private communication). Now it is well-known that the perception of the structure of an image (figure and ground, perspective, etc.) depends on luminance changes and vanishes for isoluminant chromatic images (see [12], pp. 126-128, where it is called the "Liebmann effect"). This has been verified in neurophysiological studies of Livingstone and Hubel (see [16,17]). We can thus agree with Koffka that Mach bands are not side-effects of visual mechanisms, but rather a form of perceptual organization of the image.

In order to see what may happen in the grey-level profile of an extended edge, let us consider a shadow. In Figure 2.3 we show a scene where an occluding edge gives rise to a shadow. Below we give the grey-level profile. The edge of the shadow extends from a to b ; neglecting secondary illumination, the grey-level profile will be constant to the left of a and to the right of b . Hence for any $\epsilon > 0$, the grey-level function I is not analytic on the open interval $]a - \epsilon, b + \epsilon[$; then there is at least one point p in the closed interval $[a, b]$ such that for any $\epsilon > 0$, I is not analytic on $]p - \epsilon, p + \epsilon[$. Such a point where analyticity stops will usually be found at a and b . Note however that the grey-level gradient is not necessarily

discontinuous at a and b ; quite to the contrary, the grey-level gradient will in general be continuous everywhere. It is known that a function f is analytic on an open interval $]x_0, x_1[$ if and only if

- (a) for any $k > 0$, the k -th derivative $f^{(k)}$ is defined on all $x \in]x_0, x_1[$; and
- (b) there exist $M, \delta > 0$ such that for any $k \geq 0$, all $x \in]x_0, x_1[$ give $|f^{(k)}(x)| < M k! \delta^k$ (for $k = 0$: $|f(x)| < M$).

Thus there is at least one $p \in [a, b]$ (usually $p = a$ and $p = b$) such that either $I^{(k)}(p)$ is not defined for some $k > 0$, or in the vicinity of p the function $\log(|I^{(k)}(x)|/k!)$ grows faster than linearly in k . Any such a point p must be considered as the location of a feature.

Note that in [14] a feature is located at each point where there is a step discontinuity of the grey-level function I or of its k -th derivative ($k > 0$), and such features are classified according to the value of k and the size of the step in $I^{(k)}$. The criterion of analyticity stop given here is more general.

3. ALGEBRAIC PROPERTIES AND NULL SPACE OF FEATURE DETECTION

An *edge detector* \mathcal{E} associates to a grey-level image I an *edge map*, which is a binary image giving the location of all significant edges. More generally we can assume that \mathcal{E} detects various types of local features (edges, Mach bands, etc.). Note that $\mathcal{E}(I)$ is analogous to a pencil sketch of the image.

We will consider here a few desirable properties of an edge detector which can be expressed in a simple algebraic form. A first one is *idempotence*: the edge map $\mathcal{E}(I)$ of an image contains the same edges as that image I , so that $\mathcal{E}(\mathcal{E}(I)) = \mathcal{E}(I)$. This property was formulated mathematically in [25], and further discussed in [27] from the point of view of human vision and art. We have commented [25] in detail in [26]. As explained there, idempotence can be correctly formulated if the binary edge map $\mathcal{E}(I)$ can be represented as a grey-level image, and the physically most obvious representation of $\mathcal{E}(I)$ is an image made of line edges. From this it follows that in order to be idempotent, an edge detector must detect and locate accurately a line edge. In particular pure step edge detectors (locating an edge at gradient maxima or Laplacian zero-crossings of the smoothed image) are in general not idempotent.

We will now consider algebraic properties of invariance of \mathcal{E} w.r.t. some types of image transformations. This will lead us to the *null space* of \mathcal{E} .

3.1. Spatial symmetries and grey-level transformations

Grey-level images are functions $\mathcal{S} \rightarrow \mathcal{G}$, where \mathcal{S} is the underlying space (\mathbb{R}^2 or \mathbb{Z}^2), and \mathcal{G} the set of grey-levels (\mathbb{R} or \mathbb{Z}). Given an image $I : \mathcal{S} \rightarrow \mathcal{G}$ and a symmetry σ of \mathcal{S} , $\sigma(I)$ is defined by applying σ to the graph of I , in other words

$$\sigma(I)(p) = I(\sigma^{-1}(p)) \quad \text{for } p \in \mathcal{S}. \quad (3.1)$$

In particular if τ_h is the translation by h (where $h \in \mathcal{S}$), we have

$$\tau_h(I)(p) = I(p - h) \quad \text{for } p \in \mathcal{S}. \quad (3.2)$$

We say that the edge detector is σ -invariant if it commutes with σ , that is $\sigma \circ \mathcal{E} = \mathcal{E} \circ \sigma$. Let $\Sigma_{\mathcal{E}}$ be the set of all symmetries of \mathcal{S} which commute with \mathcal{E} . Clearly $\Sigma_{\mathcal{E}}$ is a group (that is, it contains the identity and is closed under composition and inversion).

A translation or rotation of an object in the frontoparallel plane leads to the corresponding translation or rotation of the object image. The edges of the object should then be translated or rotated accordingly. In practice the vertical and horizontal orientations play particular roles in perception, but there are no privileged positions. It is thus normal to require that $\Sigma_{\mathcal{E}}$ contains the group \mathbf{T} of all translations τ_h by elements h of \mathcal{S} . Moreover in the analog case ($\mathcal{S} = \mathbb{R}^2$) $\Sigma_{\mathcal{E}}$ usually contains also the group of rotations and bilateral symmetries. In the digital case ($\mathcal{S} = \mathbb{Z}^2$) rotations are difficult to implement, and so we assume that $\Sigma_{\mathcal{E}}$ is the group of symmetries of \mathbb{Z}^2 , namely the product of \mathbf{T} with the square dihedral group, that is the group of symmetries of the pair of X and Y axes (this group has size 8, and is generated by the bilateral symmetries w.r.t. a diagonal and a vertical axis).

Given an image $I : \mathcal{S} \rightarrow \mathcal{G}$ and a function $\psi : \mathcal{G} \rightarrow \mathcal{G}$ of grey-levels, $\psi(I)$ is defined by applying ψ to each grey-level of I , in other words

$$\psi(I)(p) = \psi(I(p)) \quad \text{for } p \in \mathcal{S}. \quad (3.3)$$

In particular there is the grey-level shift by a constant $g \in \mathcal{G}$ and the grey-level magnification by a constant $k > 0$, that is:

$$\begin{aligned} (I + g)(p) &= I(p) + g \\ (kI)(p) &= kI(p) \end{aligned} \quad \text{for } p \in \mathcal{S}. \quad (3.4)$$

Given $\psi : \mathcal{G} \rightarrow \mathcal{G}$, we say that the edge detector is ψ -invariant if we have $\mathcal{E}(\psi(I)) = \mathcal{E}(I)$ for every image.

As we said in [26], given an image I and a grey-level g (positive or negative), the two images I and $I + g$ should have the same edge map, in other words $\mathcal{E}(I + g) = \mathcal{E}(I)$, because they represent the same scene. Indeed, if the intensity of all light sources illuminating a scene is multiplied by a constant factor K , then (as reflectance is a multiplicative factor in Horn's model) the light intensity reflected by each point towards the viewer will also be multiplied by K ; as both cameras and the human eye have a logarithmic gain correction, this will lead to the addition to each grey-level of a term proportional to $\log K$. We require thus that \mathcal{E} is invariant under grey-level shift. It is also reasonable to postulate that for $k > 0$, I and kI represent the same scene, because this corresponds to a modification of the logarithmic gain correction, and such a grey-level magnification does not affect the appearance of objects in the image. Hence we assume that $\mathcal{E}(kI) = \mathcal{E}(I)$ for $k > 0$. Note however that we do not necessarily postulate $\mathcal{E}(-I) = \mathcal{E}(I)$, because the edge model given in Section 1 is not invariant under grey-level inversion, and it is known that negatives of photographs are not perceived in the same way as their positives (see [27]).

3.2. The null space of edge detection

For $g \in \mathcal{G}$, we write also g for the image with constant grey-level g . Clearly we postulate $\mathcal{E}(0) = \emptyset$, and as $\mathcal{E}(I + g) = \mathcal{E}(I)$, we get $\mathcal{E}(g) = \emptyset$. We can consider whether there are other

signals than constant ones with the same properties. We define the *null space* $\mathcal{N}(\mathcal{E})$ of the edge detector \mathcal{E} as the set of all images $Z : \mathcal{S} \rightarrow \mathcal{G}$ such that for every image $I : \mathcal{S} \rightarrow \mathcal{G}$ we have $\mathcal{E}(I + Z) = \mathcal{E}(I)$. In particular we have then $\mathcal{E}(Z) = \mathcal{E}(0) = \emptyset$. The following facts are easily shown:

- (i) $\mathcal{N}(\mathcal{E})$ is an additive group (that is, $0 \in \mathcal{N}(\mathcal{E})$, and for $Z, Z' \in \mathcal{N}(\mathcal{E})$ we have $Z + Z', Z - Z' \in \mathcal{N}(\mathcal{E})$).
- (ii) If \mathcal{E} is invariant under grey-level magnification (that is, $\mathcal{E}(kI) = \mathcal{E}(I)$ for $k > 0$), then $\mathcal{N}(\mathcal{E})$ is closed under grey-level magnification (that is, for $Z \in \mathcal{N}(\mathcal{E})$ and $k > 0$ we have $kZ \in \mathcal{N}(\mathcal{E})$).
- (iii) $\mathcal{N}(\mathcal{E})$ is closed under $\Sigma_{\mathcal{E}}$ (that is, for $Z \in \mathcal{N}(\mathcal{E})$ and $\sigma \in \Sigma_{\mathcal{E}}$ we have $\sigma(Z) \in \mathcal{N}(\mathcal{E})$).

The null space $\mathcal{N}_{\mathcal{E}}$ must be defined experimentally. We know that $\mathcal{N}(\mathcal{E})$ must contain all constant signals. A good candidate for $\mathcal{N}(\mathcal{E})$ is the set of all linear signals $(x, y) \mapsto ax + b + y + c$. It is known (see in particular [12], p. 169) that an image portion where the grey-level gradient is constant (or changes very slowly) looks uniform; if we cut it by an edge, then we will perceive two uniform regions of distinct luminance on both sides of the edge. Moreover we saw in Section 2 that roofs and Mach bands of the same sign lead to the same perceived local feature (a light or dark line); now it is clear that a Mach band can be obtained by linear superposition of a roof of the same sign and a linear signal.

4. ZERO-CROSSINGS OR PEAKS?

Usually edge detection is achieved in two stages. First we enhance the image by applying to it an operator ψ ; generally ψ is a linear convolution of bandpass type. Second, edges are located at some special places in $\psi(I)$. There are two main choices for them:

- (a) *Zero-crossings* of $\psi(I)$, that is points where $\psi(I)$ changes sign.
- (b) *Peaks* of $\psi(I)$, that is local maxima of $|\psi(I)|$ (generally along some direction).

Often ψ corresponds to a particular type of edge, namely for a fixed size scale, orientation, or shape (e.g., line or step). Then we must apply in parallel to I several operators ψ forming a family Ψ , and select edges in each $\psi(I)$.

The edge detector of Marr and Hildreth [20] locates step edges at zero-crossings of the convolutions of the image with Laplacian of Gaussian functions chosen at three or four size scales; note that these operators are isotropic. On the other hand gradient-type approaches take for each size scale orientationally selective operators ψ_j corresponding to several orientations (say, the 4 main ones, so that $j = 0, \dots, 3$), and define the “gradient” ψ by setting $\psi(p) = \max_j |\psi_j(I)(p)|$ for each point p ; then p is an edge point if for the j with $|\psi_j(p)| = \psi(p)$, ψ has a local maximum at p along the direction associated to ψ_j , which is the direction normal to the edge at p .

We will recall here two known major defects of the approach based on zero-crossings, which apply not only for the Laplacian of Gaussian, but to smooth operators in general. This will imply that the edge detector may not be based on isotropic image operators, but must use orientationally selective ones.

4.1. Defects of zero-crossings

Suppose first that we take an operator ψ which is isotropic, that is commuting with trans-

lations and rotations. Assume that $\psi(I)$ is continuous. Let P be the set of points $p \in \mathcal{S}$ such that $\psi(I)(p) > 0$, and N the set of those for which $\psi(I)(p) < 0$; let Z be the set of zero-crossings of $\psi(I)$. Then Z is the edge separating P from N . Thus it will in practice have the topological characteristics of the edge of a binary image. Any Jordan curve (smooth simple closed arc) will cross Z an even number of times. Usually Z will consist of a union of mutually disjoint Jordan curves, or curves that terminate at the boundary of the image [6,29]. In Figure 4.1 we show two situations of edges which cannot be the separation between two sets (and where a circle can be drawn crossing the edge an *odd* number of times):

- (a) An edge endpoint arising for example from a crack in a region. Here the edge is within a single connected region. Then Z will usually turn around the true edge in order to separate the image into two regions: see Figure 4.1(a).
- (b) A T-junction, generally between three regions. Then Z will usually separate the image into two regions as in Figure 4.1(b).

Situation (b) has been analysed in detail in [6] when ψ is the Laplacian of a Gaussian. Other shortcomings of that particular operator (in particular that zero-crossings turn around corners) have been considered in the literature (see [6] and the references in it). Daugman [5] has shown that human vision can see features which cannot be detected as zero-crossings of a Laplacian of Gaussian. However the above defects apply to zero-crossings in general, even when ψ is not linear.

The second defect of zero-crossings was pointed out by Ling [15]. Let us now consider one-dimensional features, that is in a signal which is locally constant in the Y direction and varies only according to the X axis. Thus what we say here applies both to isotropic or orientationally selective operators. Zero-crossings along a direction are of two types: negative-to-positive and positive-to-negative, see Figure 4.2(a). Assuming again that $\psi(I)$ is continuous, it is clear that between two zero-crossings of one type lies a zero-crossing of the other type. Now consider a staircase made of two step edges having the same sign (see Figure 4.2(b)). Then at the location of each step we normally have a zero-crossing, and both should be of the same type (since the steps have the same sign). Then there will be a spurious zero-crossing of the other type between them. For example if $\psi(I) = I * \partial^2 G_\sigma / \partial x^2$, the convolution with the second derivative of a Gaussian, then the two positive-to-negative zero-crossings at the step edge locations correspond to local maxima of the smoothed gradient $|I * \partial G_\sigma / \partial x|$, while the spurious negative-to-positive zero-crossing between them corresponds to a local minimum of that smoothed gradient. Hence we must distinguish between the two types of zero-crossings along a direction.

4.2. Anisotropic operators

From the inadequacy of zero-crossings we will show that a wide class of isotropic image transformations are insufficient for edge detection. Moreover, our argument will highlight a third defect of the zero-crossing approach for that class of operators. Let ψ be an operator on the space of grey-level images, satisfying the following two conditions:

- (i) ψ commutes with translations and 180° rotations.
- (ii) For an image I , $\psi(-I) = -\psi(I)$.

Condition (i) is weaker than isotropy; it is satisfied by the convolution with any even-symmetric mask. Condition (ii) is satisfied by linear operators, but also by many non-linear operators, such as the median filter. Now let S be a perfect odd-symmetric edge about a point p , and let ρ be the 180° rotation about p . Then $\rho(S) = -S$ and from (i, ii) we get

$$\rho(\psi(S)) = \psi(\rho(S)) = \psi(-S) = -\psi(S),$$

which means that $\psi(S)$ will also be odd-symmetric about p , and in particular $\psi(S)(p) = 0$. Thus with ψ satisfying (i, ii), the step edge in S will be a zero of $\psi(S)$. Thus for any such operator ψ (or any family of such operators), step edges in an image I should be detected as zero-crossings of $\psi(I)$, which leads to the two defects mentioned above.

Moreover, the above argument shows that for any signal S which is odd-symmetric about p we will have $\psi(S)(p) = 0$, in other words p will automatically be labelled as a feature point. For example with the Marr-Hildreth edge-detector [20], the convolution of S with the Laplacians of 3 or 4 Gaussians will have a zero-crossing at each scale, in other words p will be the location of a significant edge. This definition of edge is not restrictive enough.

Thus with property (ii) we will need operators which are not invariant under 180° rotations. In particular, convolutions with even-symmetric masks are inappropriate for detecting step edges. We will take a family Ψ of operators ψ which are translation-invariant (commute with each translation $\tau \in \mathbf{T}$), but will not commute with rotations. In other words they will be *orientationally selective*, each ψ will correspond with a particular orientation. In order to implement the invariance of the edge detector \mathcal{E} under the group $\Sigma_{\mathcal{E}}$, we require that for each $\sigma \in \Sigma_{\mathcal{E}}$ and each $\psi \in \Psi$ there is some $\psi' \in \Psi$ with $\sigma\psi\sigma^{-1} = \psi'$; here the orientation associated to ψ' results from that of ψ by the symmetry σ .

5. LINEAR PROCESSING AND ASSOCIATED REQUIREMENTS

One usually takes a family Ψ of linear translation-invariant operators defined on a certain space of measurable images. Moreover, for each image I the edge map $\mathcal{E}(I)$ must be determined solely from the set of transformed images $\psi(I)$, $\psi \in \Psi$. Given a signal Z such that $\psi(Z) = 0$ for each ψ , then for any image I we will have $\psi(I + Z) = \psi(I)$ for each ψ , so that $\mathcal{E}(I + Z) = \mathcal{E}(I)$, and hence $Z \in \mathcal{N}(\mathcal{E})$, the null space of \mathcal{E} . Conversely, one must choose Ψ in such a way that $\psi(Z) = 0$ for any $Z \in \mathcal{N}(\mathcal{E})$ and $\psi \in \Psi$.

Generally these linear operators are convolutions by masks having certain shapes. Given a measurable function $F : \mathcal{S} \rightarrow \mathbb{R}$ (where we recall that \mathcal{S} is the space \mathbb{R}^2 or \mathbb{Z}^2), the convolution operator $\psi : I \mapsto I * F$ is translation-invariant and satisfies the following properties:

(1°) If F is integrable (in the Lebesgue sense, that is integrable in absolute value), then ψ is stable with respect to the L^p norm ($p \geq 1$) $\| \cdot \|_p$ and the uniform norm $\| \cdot \|_u$ defined by $\|I\|_p = (\int_{\mathcal{S}} |I(x)|^p dx)^{1/p}$ and $\|I\|_u = \sup_{p \in \mathcal{S}} |I(p)|$, because

$$\text{and} \quad \begin{aligned} \|I * F\|_p &\leq \|I\|_p \cdot \|F\|_1 \\ \|I * F\|_u &\leq \|I\|_u \cdot \|F\|_1 \end{aligned}$$

for each image I . Thus convolution by F is a well-posed problem in the sense of [28,29].

- (2°) $\psi(g) = 0$ for a constant signal g if and only if F is integrable and $\int_{\mathcal{S}} F(x) dx = 0$.
- (3°) $\psi(L) = 0$ for a linear signal L if and only if for $X_i : (x_1, x_2) \mapsto x_i$ ($i = 1, 2$), F , $X_1 F$, and $X_2 F$ are integrable $\int_{\mathcal{S}} F(x) dx = \int_{\mathcal{S}} X_1(x) F(x) dx = \int_{\mathcal{S}} X_2(x) F(x) dx = 0$.
- (4°) For $\mathcal{S} = \mathbb{R}^2$, given a linear transformation σ of \mathcal{S} (fixing the origin), then $\sigma(I * F) = |\det \sigma|^{-1} \cdot \sigma(I) * \sigma(F)$; in particular for $\det \sigma = \pm 1$, $\sigma \psi \sigma^{-1}(I) = I * \sigma(F)$. For $\mathcal{S} = \mathbb{Z}^2$, this is true if σ belongs to the square dihedral group of 8 symmetries of the pair of X and Y axes.

Therefore we will choose each operator to be the convolution by an (absolutely) integrable function F with $\int_{\mathcal{S}} F(x) dx = 0$. If we require that $\mathcal{N}(\mathcal{E})$ contains linear signals, we will impose the same condition (integrability and vanishing integral) on $X_1 F$ and $X_2 F$. The invariance of \mathcal{E} under symmetries in $\Sigma_{\mathcal{E}}$ is achieved by taking for each F its rotated versions $\sigma(F)$ by all $\sigma \in \Sigma_{\mathcal{E}}$ fixing the origin.

Since our convolutions are orientationally selective, each F corresponds to a particular type of feature with unit vectors $\vec{1}_n$ in the normal direction and $\vec{1}_t$ in the tangential direction ($\vec{1}_n \perp \vec{1}_t$). Given $x \in \mathcal{S}$, let x_n and x_t be its coordinates in the orthonormal basis $(\vec{1}_n, \vec{1}_t)$. We assume that F is separable as the product of a “detection function” D in the normal direction and a “continuation function” C in the tangential direction: $F(x) = D(x_n) \cdot C(x_t)$. Generally C is smooth, non-negative, and even-symmetric (typically, a Gaussian). Then the condition of integrability and vanishing integral for F becomes: C and D are integrable, with $\int D = 0$. The same condition for $X_1 F$ and $X_2 F$ (if $\mathcal{N}(\mathcal{E})$ contains linear signals) becomes: $x_n D(x_n)$ and $x_t C(x_t)$ are integrable, and $\int x_n D(x_n) dx_n = 0$ (as C is even-symmetric, $\int x_t C(x_t) dx_t = 0$ anyway).

The functions with which we convolve the image vary not only in orientation, but also in size, in order to detect features at several size scales. If F is used in the detection of features of size 1, then in order to detect similar features of size $s > 0$ one generally uses the magnification F_s of F by a factor s : $F_s(x_n, x_t) = F(x_n/s, x_t/s)$. However this is not always the case, for example in [21]; we reproduce some of their masks for two sizes in Figure 5.1.

Finally these functions can also vary in shape, since they must be used to detect different types of edges: lines, steps, roofs, etc. Generally one takes the same continuation function C , but vary the profile of the detection function D . There are mainly two types of detection functions: even-symmetric and odd-symmetric. See Figures 5.1 and 5.2, where we show those used respectively by Morrone and Burr [21] and by Klein and Levi [11]. Although their analytical expressions are very different, the masks proposed in both papers have similar profiles: for the even-symmetric one, a positive lobe in the middle, flanked by two smaller negative lobes; for the odd-symmetric one, a succession of four lobes of alternating sign, two large ones at the middle, and two smaller ones at the extremities. They are typical of smooth functions which are localized in both the spatial and Fourier domains, and give a zero response on a linear signal.

In the visual cortex of cats and primates the response of simple cells is linear, with odd-symmetric and even-symmetric receptive fields [4]. Following the works of Hubel and Wiesel who first investigated them, they are usually called “edge detectors” and “bar detectors”

respectively (see for example [3]), which means that they can be used in the detection of step edges and line edges respectively (as said in Section 2, we prefer to consider triangular or rounded line edges instead of rectangular bars). Convolution of an image I with an odd-symmetric “edge detector” F_o , $|I * F_o|$ will have a local peak at the location of a step edge in I with normal orientation aligned to that of F_o . However convolution of I with an even-symmetric “bar detector” F_e , as F_e has zero response on a linear signal, a local peak of $|I * F_e|$ will be obtained not only at the location of line edges and roofs with normal orientation aligned to that of F_e , but also their superposition on a linear signal, for example a Mach band. This is no wonder if we remark that the profile of the even-symmetric detection functions shown in Figures 5.1 and 5.2 implements “lateral inhibition”. Thus the expression “bar detector” is too restrictive.

6. THE TWOFOLD MODEL

There are many shapes of edges or local features for a fixed size and orientation. We suggest here that they can be decomposed as the sum of a featureless signal in the null space $\mathcal{N}(\mathcal{E})$, an even-symmetric and an odd-symmetric edge feature (with the symmetry w.r.t. the feature location). The associated perception is a mixture of line and step corresponding to the even-symmetric and odd-symmetric edge components, while the featureless signal does not influence the local appearance of the edge. We propose to detect such a mixed type edge and to characterize its type by suitably combining the convolution of the image with masks having even-symmetric and odd-symmetric detection functions respectively. We will see that the results of both convolutions may not be analysed separately, but must be measured together, in particular by a quadratic energy function.

Our analysis is made in the case of one-dimensional signals. The two-dimensional case is dealt with briefly at the end of this section.

6.1. Justification of the model

In [21,23] experiments have been made on the human perception of local features in vertical gratings (images whose grey-level is constant in the vertical direction). Thus the edge model considered there is one-dimensional. Three types of local features are perceived: a step edge, a line edge, or a combination of both. As we have seen in Section 2, in a roof and a Mach band of the same sign, a line is seen, which is dark or light according to that sign; moreover such roof and Mach band of the same sign differ by the addition of a linear signal. As we hinted in Section 3, such a linear signal may belong to the null space, so that a roof or Mach band can be decomposed as the sum of an even-symmetric roof feature and a featureless linear signal in the null space.

Given a one-dimensional signal I and a point p in the line, the *Fourier phase of I at p* is the Fourier phase that I would give if the origin was positioned at p ; in other words it is the phase of $\tau_{-p}(I)$, the translate of I by $-p$ (see [26], Section 3). For every frequency ν , the Fourier phase at p is obtained by adding $2\pi\nu p$ to the ordinary phase. Morrone and coworkers [21,23] found that human observers see a local feature in vertical gratings at points of *maximum phase congruency*, in other words points p such that the Fourier phases

at p for all positive frequencies are maximally concentrated. If the angle around which they concentrate is close to 0 or 180° , a line edge (light or dark respectively) is seen. If that angle is close to 90° or -90° , a step edge (positive-to-negative or negative-to-positive respectively) is seen. For intermediate values a mixture of step and line is seen; in particular when the angle is 45° , the perception is instable, oscillating between line and edge [21].

Clearly a signal with constant Fourier phase 0 or 180° is even-symmetric; moreover an even-symmetric triangular delta has constant Fourier phase 0 or 180° , according to its sign [26]. Similarly a signal with constant Fourier phase 90° or -90° is odd-symmetric; moreover an odd-symmetric step has constant Fourier phase 90° or -90° according to its orientation [26]. Finally a signal with constant Fourier phase φ comprised between 0 and 90° is a linear combination $\cos \varphi R + \sin \varphi S$, where R and S have the same Fourier amplitudes as the original signal, and constant Fourier phases 0 and 90° respectively.

These facts are consistent with a feature model in which a signal I has a feature at p if it can be decomposed as $I = Z + R + S$, where $Z \in \mathcal{N}(\mathcal{E})$, and given the translation τ_{-p} mapping p to the origin, $\tau_{-p}(R)$ is even-symmetric, $\tau_{-p}(S)$ is odd-symmetric, and both have some salient characteristics which make them features. From the experiments of Morrone and coworkers [21,23], such characteristics can be Fourier phase congruency, namely about 0 or 180° for $\tau_{-p}(R)$, and about 90° or -90° for $\tau_{-p}(S)$. The significance of such Fourier phase characteristics for multiscale edge detection will be discussed in the next section.

6.2. Detection and classification of simple features

Let us now see how we can detect and classify simple one-dimensional features like a step or line (or the separate signals R and S of the previous paragraph) by the convolution with even-symmetric and odd-symmetric masks F_e and F_o respectively. As we will see in the next section, if we take into account Fourier phase characteristics of features, we will require that F_e and F_o have constant Fourier phase 0 and 90° respectively, while they have equal Fourier amplitudes.

The grey-level profiles of F_e and F_o will be like in Figures 5.1 and 5.2, and those of relevant features will be somewhat analogous. We will thus consider four types of functions in one variable, two even-symmetric and two odd-symmetric, which will include F_e and F_o as well as the ideal step and line edges. An even-symmetric function F is of type P (for “positive”) if its profile has a significant positive lobe around the origin, with a peak there; it is of type N (for “negative”) if $-F$ is of type P. An odd-symmetric function F is of type NP (for “negative-to-positive”) if its profile has two significant lobes on both sides of the origin, a negative one to the left and a positive one to the right, with slope peaking at the origin; it is of type PN (for “positive-to-negative”) if $-F$ is of type NP. For example the even-symmetric functions of Figures 5.1 and 5.2 are of type P, while the odd-symmetric functions are of type NP in Figure 5.1 and PN in Figure 5.2. Let us also say that P is opposite to N, and NP is opposite to PN. Convolution functions of these four types will usually give a function in one these types according to the following rules:

- Convolution with a function of type P does not change the type. Convolution with a function of type N changes the type into its opposite.

- Convolution of two odd-symmetric functions of the same type gives a function of type N.
- Convolution of two odd-symmetric functions of opposite types gives a function of type P.

Now we associate the numbers $\pm 1, \pm i$ to types as follows:

$$\sigma(F) = \begin{cases} +1 & \text{if } F \text{ is of type P,} \\ -1 & \text{if } F \text{ is of type N,} \\ -i & \text{if } F \text{ is of type NP,} \\ +i & \text{if } F \text{ is of type PN;} \end{cases} \quad (6.1)$$

then the above rules become $\sigma(F * G) = \sigma(F)\sigma(G)$. More than a mnemotechny, $\sigma(F)$ indicates the most likely sign taken by the Fourier transform of F in the range of positive frequencies. This will be explained in the next section, where we will characterize more precisely the four types we have introduced.

Now we suppose that the functions F_e and F_o are of types P and PN respectively (for types N and NP respectively we only have to take $-F_e$ and $-F_o$ instead, and the analysis remains similar). Let the image have a simple feature at p , and let I be its section in the direction normal to the feature; then the local profile of I around p will be similar to the profile around 0 of a function of one of the above four types. Then $I * F_e$ and $I * F_o$ will have a local profile around p which will be similar to the profile around 0 of a function whose type is given by the above convolution rule. This type P, N, NP, or PN manifests itself by a positive or negative peak, a negative-to-positive or positive-to-negative zero-crossing at p respectively:

- For I around p like type P, at p we have a positive peak of $I * F_e$ and a positive-to-negative zero-crossing of $I * F_o$.
- For I around p like type N, at p we have a negative peak of $I * F_e$ and a negative-to-positive zero-crossing of $I * F_o$.
- For I around p like type NP, at p we have a negative-to-positive zero-crossing of $I * F_e$ and a positive peak of $I * F_o$.
- For I around p like type PN, at p we have a positive-to-negative zero-crossing of $I * F_e$ and a negative peak of $I * F_o$.

This straightforward method works also when I is the superposition of a simple feature like above and a featureless signal Z in the null space $\mathcal{N}(\mathcal{E})$, since $Z * F_e$ and $Z * F_o$ will be identically zero.

6.3. Compound or noisy features

We will show here that if in a one-dimensional signal I (the image section in a given orientation) we have a simple feature which does not match perfectly the model (for example, a noisy one), it may be necessary to compare or combine $I * F_e$ and $I * F_o$. If we have a compound feature (a superposition of an even-symmetric one and an odd-symmetric one) a superposition of the form $(I * F_e)^2 + (I * F_o)^2$ becomes mandatory.

A combination of $|I * F_e|$ and $|I * F_o|$ which preserves the peaks of each of them must be a disjunction or superposition. The two disjunctions

$$\text{and} \quad \begin{aligned} & \max(|I * F_e|, |I * F_o|) \\ & \max((I * F_e)^2, (I * F_o)^2) \end{aligned} \quad (6.2)$$

are equivalent (the first is the square root of the second). Two peak-preserving linear superpositions can be envisaged:

$$|I * F_e| + |I * F_o| \quad (6.3)$$

$$\text{and} \quad (I * F_e)^2 + (I * F_o)^2. \quad (6.4)$$

Let us show that (6.3) is inadequate. Suppose that $I = S$, a perfect odd-symmetric step edge located at the origin 0. Then $|S * F_o|$ has a maximum at 0, while $|S * F_e|$ has a minimum there, equal to 0. Assume that $S * F_o$ and $S * F_e$ are derivable. In some neighbourhood of 0, $S * F_o$ has no zero-crossing, while $S * F_e$ has 0 as unique zero-crossing; as can be seen in Figure 6.1, it follows that $|S * F_o|'(0) = \pm(S * F_o)'(0) = 0$, while the left and right derivatives of $|S * F_e|$ at 0 are $-(S * F_e)'(0)$ and $+(S * F_e)'(0)$ respectively, so that $|S * F_e| + |S * F_o|$ will have a local minimum at 0. On the other hand with (6.4) we have the following derivative:

$$[(S * F_e)^2 + (S * F_o)^2]' = 2(S * F_e) \cdot (S * F_e)' + 2(S * F_o) \cdot (S * F_o)'; \quad (6.5)$$

as $(S * F_o)'(0) = (S * F_e)(0) = 0$, the expression (6.5) has value 0 at point 0; in the neighbourhood of 0, $|S * F_o|$ is much stronger than $|S * F_e|$, so that the influence in (6.5) of $(S * F_o)'$ will be enhanced w.r.t. to that of $(S * F_e)'$. In other words, the risk of having a local minimum at 0 is diminished.

Suppose now that $I = N + S$, where S is a perfect odd-symmetric edge located at 0, and N a signal which is not in the null space, but whose convolutions by F_e and F_o have weak values near 0. For example N can contain noise, or a systematic deviation from the ideal step edge model; or else I can have features at other points than 0, and they belong to the signal N . Then $S * F_o$ will have a strong peak (positive or negative) at 0, while $N * F_o$ will be weaker near that point. Thus $I * F_o$ will not be significantly different from $S * F_o$ near 0. But $S * F_e$ will have a zero-crossing at 0, and near that point the values of $N * F_e$ can be stronger. Moreover if N consists mainly of high frequencies near 0, in that region the slope of $N * F_e$ can be higher in absolute value than that of $S * F_e$. Thus we may have small local peaks of $I * F_e$ near 0. These peaks in $I * F_e$ are weaker than the one of $I * F_o$ near 0, and they can be dominated if we take one of the combinations (6.2) or (6.4) of $I * F_o$ and $I * F_e$. The same argument applies if we have $I = N + R$, where R is a perfect even-symmetric line or roof. Hence $I * F_o$ and $I * F_e$ should not be analysed separately, but rather combined.

Consider now a compound feature $I = R + S$, where R and S are perfect edges located at 0, with R even-symmetric and S odd-symmetric. Let us write $R_o = R * F_o$, $R_e = R * F_e$, $S_o = S * F_o$, $S_e = S * F_e$, $I_o = I * F_o$, and $I_e = I * F_e$; thus $I_o = R_o + S_o$ and $I_e = R_e + S_e$. Assume again that R_o , R_e , S_o , and S_e are derivable. Without loss of generality, we can suppose that R and F_e are of type P, while S and F_o are of type NP (the argument remains the same in other cases). By the above convolution rules for the types P, N, NP, and PN, we have the following:

- R_o is of type NP. Thus $R_o(0) = 0$ and $R_o'(0) > 0$.
- S_o is of type N. Thus $S_o(0) < 0$ and $S_o'(0) = 0$.

— R_e is of type P. Thus $R_e(0) > 0$ and $R'_e(0) = 0$.

— S_e is of type NP. Thus $S_e(0) = 0$ and $S'_e(0) > 0$.

Therefore $I_o(0) < 0$, $I'_o(0) > 0$, $I_e(0) > 0$, and $I'_e(0) > 0$, in other words $|I_o|$ has negative slope at 0, while $|I_e|$ has positive slope at 0. In another way, we have:

$$\begin{aligned} (I_o^2)'(0)(I_e^2)'(0) &= 4I_o(0)I'_o(0)I_e(0)I'_e(0) \\ &= 4[R_o(0) + S_o(0)][R'_o(0) + S'_o(0)][R_e(0) + S_e(0)][R'_e(0) + S'_e(0)] \\ &= 4S_o(0)R'_o(0)R_e(0)S'_e(0) < 0, \end{aligned}$$

that is I_o^2 and I_e^2 have slopes of opposite sign at 0. Thus the peaks in $|I_o|$ and $|I_e|$ arising from the superposition of R and S will be located at both sides of the origin. With (6.2) we will get two peaks for the compound edge. This edge duplication is undesirable. Now if we take (6.4), the derivatives at 0 of I_o^2 and I_e^2 will tend to cancel each other, in other words the slope of $I_e^2 + I_o^2$ at 0 will be weaker. This will not lead to an edge duplication, but to an edge displacement, a weaker defect.

To see the perceptual plausibility of such an edge displacement, we show in Figure 6.2 a particular type of vertical grating whose grey-level is horizontally periodic, obtained by adding a constant term (in the null space) and a convex linear combination of a square and a triangular wave (taking the role of S and R respectively in the above argument). The weight of both waves varies linearly in the vertical direction, from a pure square wave to a pure triangular one. In the middle, where both waves are superposed, the compound edge is seen at a location slightly away from the true edge location, where it appears at the top and the bottom.

The signal described in (6.4) is called the squared *energy function* of I , and a feature detector based on an analysis of peaks of such function is called an *energy feature detector* [21,22,25,27,30]. We have seen here that it can be useful for detecting features modelled as above, namely $I = Z + R + S$, where $Z \in \mathcal{N}(\mathcal{E})$, R and S are respectively even-symmetric and odd-symmetric features (with the symmetry w.r.t. the feature location).

We have made a one-dimensional analysis of edges. In a two-dimensional model we take F_e and F_o decomposable as products of a continuation function C in the tangential orientation and detection functions D_e and D_o respectively in the normal direction: $F_e(x) = D_e(x_n)C(x_t)$ and $F_o(x) = D_o(x)C(x_t)$. Here D_e and D_o will have the properties mentioned above for F_e and F_o . On the other hand C will be even-symmetric and non-negative; furthermore if we take into account Morrone's phase congruency model [21,23], we will also require that C has constant zero Fourier phase.

7. CONSIDERATIONS IN THE FOURIER DOMAIN

Some of the statements made in the previous sections will be related to properties of the Fourier transform. We restrict ourselves here to signals in one-dimensional analog space, that is functions $\mathbb{R} \rightarrow \mathbb{R}$ in the spatial domain and $\mathbb{R} \rightarrow \mathbb{C}$ in the Fourier domain.

Let us first recall from [26] some basic facts of the L^1 theory of the Fourier transform. All functions are implicitly assumed to be (Lebesgue) measurable. For any function F , the *reflection* F_ρ of F is given by $F_\rho(x) = F(-x)$; every function F can be decomposed in a

unique way as the sum of an even-symmetric function and an odd-symmetric one, namely $(F + F_\rho)/2$ and $(F - F_\rho)/2$.

Let L^1 be the set of (Lebesgue) integrable functions $\mathbb{R} \rightarrow \mathbb{C}$ (that is, integrable in absolute value); it is a vector space. The *convolution* $F * G$ of two functions F and G is defined by $[F * G](x) = \int_{\mathbb{R}} F(x-t)G(t)dt$ (whenever the integral exists). We have the following:

- (a) Let $F, G \in L^1$; then $F * G \in L^1$. If G is bounded (resp., uniformly continuous), then $F * G$ is also bounded (resp., uniformly continuous).

Given $F \in L^1$, its *Fourier transform* F^\wedge is defined by

$$F^\wedge(\nu) = \int_{\mathbb{R}} F(x) \exp[-2\pi i\nu x] dx.$$

Note that the Fourier transform is linear and preserves even- and odd-symmetry. Write F^\vee for $(F_\rho)^\wedge = (F^\wedge)_\rho$, that is

$$F^\vee(\nu) = \int_{\mathbb{R}} F(x) \exp[2\pi i\nu x] dx.$$

Given $F \in L^1$, F^\wedge has the following properties (among others, see [26]):

- (b) F^\wedge is bounded and uniformly continuous.
(c) If F is continuous and F^\wedge is in L^1 , then $F = F^{\wedge\vee}$. In particular F is bounded and uniformly continuous.
(d) For $G \in L^1$, $[F * G]^\wedge = F^\wedge G^\wedge$.
(e) If $XF : x \mapsto xF(x)$ is in L^1 , then F^\wedge is differentiable, and $F^{\wedge'} = [-2\pi iXF]^\wedge$.

From properties (a, b, c, d) we derive the following:

- (f) Let $F, G \in L^1$ such that F is continuous and $F^\wedge \in L^1$. Then $F * G$ is bounded and uniformly continuous, $[F * G]^\wedge \in L^1$, and $F * G = [F * G]^{\wedge\vee}$.

When F is real-valued, F^\wedge has an even-symmetric real part and an odd-symmetric imaginary part. We write $F^\wedge(\nu) = F^{\mathcal{A}}(\nu) \exp[iF^{\Phi}(\nu)]$, where $F^{\mathcal{A}}(\nu) = |F^\wedge(\nu)|$; here $F^{\mathcal{A}}$ is the *Fourier amplitude* of F and F^{Φ} is the *Fourier phase* of F ; thus $F^{\mathcal{A}}$ is even-symmetric, while F^{Φ} is odd-symmetric. Note that $F^{\Phi}(\nu)$ is defined modulo 2π for $F^{\mathcal{A}}(\nu) > 0$, while it is usually undefined for $F^{\mathcal{A}}(\nu) = 0$. By (d), the convolution of two real-valued functions in L^1 multiplies their amplitudes and adds their phases modulo 2π : $[F * G]^{\mathcal{A}} = F^{\mathcal{A}}G^{\mathcal{A}}$ and $[F * G]^{\Phi} = F^{\Phi} + G^{\Phi} \pmod{2\pi}$.

We will now state a few simple results which will indicate what may happen to the sign of a signal after convolution with certain functions. The following result is contained in Lemma 2.1 of [26]:

LEMMA 1. *Let $F \in L^1$ be even-symmetric with non-negative real values, not almost everywhere equal to zero. Then $F^\wedge(0) > |F^\wedge(t)|$ for all $t \neq 0$.*

Applying Lemma 1 in the Fourier domain and then (c), we deduce the following result, contained in Proposition 2.2 of [26]:

COROLLARY 2. Let $R \in L^1$ and real-valued, not almost everywhere equal to zero, even-symmetric, continuous, with $R^\wedge \in L^1$ and $R^\wedge(\nu) \geq 0$ for all ν . Then $R(0) > |R(t)|$ for all $t \neq 0$.

An even-symmetric (resp., odd-symmetric) function in L^1 has an even-symmetric (resp., odd-symmetric) and real-valued (resp., imaginary-valued) Fourier transform, and its Fourier phase takes values 0 or π (resp., $\pi/2$ or $-\pi/2$). Thanks to (d, f), we obtain thus the following:

COROLLARY 3. Let $I, T \in L^1$ and real-valued, not almost everywhere equal to zero, with T continuous and $T^\wedge \in L^1$. Assume one of the following:

- (i) I and T are even-symmetric, and $I^\Phi(\nu) = T^\Phi(\nu) = 0$ for all ν .
- (ii) I and T are odd-symmetric, and $-I^\Phi(\nu) = T^\Phi(\nu) = \pi/2$ for all $\nu > 0$.

Then $[I * T](0) > |[I * T](t)|$ for all $t \neq 0$.

We will now prove in some way a converse of the above results.

LEMMA 4. Let F be real-valued, bounded, and continuous. For any $\sigma > 0$ let H_σ be defined by $H_\sigma(t) = \exp[-2(\pi\sigma t)^2]$. Let $\theta > 0$ with $F(\theta) \neq 0$. Assume one of the following:

- (i) F is even-symmetric, and for $\sigma > 0$ we define K_σ^θ by $K_\sigma^\theta(t) = H_\sigma(t - \theta) + H_\sigma(t + \theta)$.
- (ii) F is odd-symmetric, and for $\sigma > 0$ we define K_σ^θ by $K_\sigma^\theta(t) = H_\sigma(t - \theta) - H_\sigma(t + \theta)$.

Then for σ sufficiently large, the integral over \mathbb{R} of FK_σ^θ has the same sign as $F(\theta)$.

PROOF. We show only case (ii); the proof of case (i) is similar. We can assume that $F(\theta) > 0$; for $F(\theta) < 0$ the result will follow by interverting F and $-F$.

As F is continuous, there exist $\varepsilon, \eta > 0$ with $\varepsilon < \theta$, such that for $|t - \theta| < \varepsilon$, $F(t) > \eta$. Now $K_\sigma^\theta(t) > 0$ for $t > 0$, so that $F(t)K_\sigma^\theta(t) > \eta|K_\sigma^\theta(t)| > 0$ for $|t - \theta| < \varepsilon$. Let $X =]-\theta - \varepsilon, -\theta + \varepsilon[\cup]\theta - \varepsilon, \theta + \varepsilon[$. Clearly $K_\sigma^\theta(t)$ is odd-symmetric (as F), so that FK_σ^θ is even-symmetric. Thus we have

$$F(t)K_\sigma^\theta(t) > \eta|K_\sigma^\theta(t)| > 0 \quad \text{for } t \in X.$$

Let $Y = \mathbb{R} \setminus X$. There is some $M > 0$ such that $F(t) < M$ for all $t \in \mathbb{R}$. Thus we have

$$|F(t)K_\sigma^\theta(t)| < M|K_\sigma^\theta(t)| \quad \text{for } t \in Y.$$

Setting $A_\sigma = \int_X |K_\sigma^\theta|$ and $B_\sigma = \int_Y |K_\sigma^\theta|$, we get then:

$$\int_{\mathbb{R}} FK_\sigma^\theta = \int_X FK_\sigma^\theta + \int_Y FK_\sigma^\theta \geq \int_X FK_\sigma^\theta - \int_Y |FK_\sigma^\theta| \geq \eta A_\sigma - M B_\sigma$$

As $B_\sigma/A_\sigma \rightarrow 0$ for $\sigma \rightarrow \infty$, the result follows. ■

The Gabor cosine function GC_σ^θ and the Gabor sine function GS_σ^θ (with $\sigma, \theta > 0$) are defined by

$$GC_\sigma^\theta(x) = G_\sigma(x) \cos[2\pi\theta x] \quad \text{and} \quad GS_\sigma^\theta(x) = G_\sigma(x) \sin[2\pi\theta x],$$

where G_σ is the Gaussian given by

$$G_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-x^2}{2\sigma^2}\right].$$

Now the Fourier transform of G_σ is the function H_σ of the above Lemma, so that $2GC_\sigma^{\theta\wedge}$ and $2iGS_\sigma^{\theta\wedge}$ give the two forms of K_σ^θ in items (i) and (ii) respectively. As the integral over \mathbb{R} of FK_σ^θ is equal to $[FK_\sigma^\theta]^\wedge(0)$, we obtain thus the following counterpart of Corollary 3:

COROLLARY 5. Let $I \in L^1$ and real-valued, not almost everywhere equal to zero, and let $\theta > 0$ with $I^\wedge(\theta) \neq 0$. Then:

- (i) If I is even-symmetric, for $\sigma > 0$ sufficiently large, $[I * GC_\sigma^\theta](0)$ has the same sign as $I^\wedge(\theta)$.
- (ii) If I is odd-symmetric, for $\sigma > 0$ sufficiently large, $[I * GS_\sigma^\theta](0)$ has the same sign as $-iI^\wedge(\theta)$.

In fact convolutions by the Gabor cosine and sine functions are of bandpass type, with bands located around θ and $-\theta$, and bandwidth narrowing as σ increases. So if I is even- or odd-symmetric, and I^\wedge has zero-crossings outside 0, then we can have $[I * T](0)$ of arbitrary sign according to the choice of a function T as in Corollary 3. Thus Corollary 5 provides a converse to Corollary 3.

Many convolution masks used in multiscale edge detection (such as Gaussians and their first and second derivatives) are functions $F \in L^1$, continuous, with $F^\wedge \in L^1$, and constant Fourier phase 0 or π in the even-symmetric case, $\pi/2$ or $-\pi/2$ in the odd-symmetric case. We see that a one-dimensional signal I will give a peak of the same sign for the convolutions by masks at all scales, provided that it has constant Fourier phase (0, π or $\pm\pi/2$ according to the symmetry of the masks). This justifies Morrone's characterization of local features by Fourier phase congruency, which we discussed in Subsection 6.1.

Let us now characterize more precisely the four types P, N, NP, and PN of functions that we used in Subsections 6.2 and 6.3. Recall the numbers associated to these four types in (6.1). Given a real or imaginary number x , we define its *sign* $\text{sgn}(x)$ as follows:

$$\text{sgn}(x) = \begin{cases} 0 & \text{if } x = 0; \\ x/|x| & \text{if } x \neq 0. \end{cases}$$

In other words $\text{sgn}(x) = 1$ if $x > 0$, -1 if $x < 0$, i if $-ix > 0$, and $-i$ if $ix < 0$. Then for $F \in L^1$, even- or odd-symmetric, we will say that F is of type P, N, NP, or PN, with corresponding $\sigma(F)$ as in (6.1), if $\text{sgn}(F^\wedge(\nu))$ is generally equal to $\sigma(F)$ for $\nu > 0$. All this depends on what is meant by "generally". In the strict interpretation this means "always", in other words $\text{sgn}(F^\wedge(\nu)) = \sigma(F)$ for all $\nu > 0$ (the Fourier phase of F is constant, multiple of $\pi/2$). More generally we can require that this equality holds for all $\nu > 0$ in a certain frequency band, outside which the Fourier transform of the convolution masks is very weak. Then the convolution of any two functions of these four types results in a function in one of the four types, according to the rule given with (6.1). For the convolution masks F_e and F_o we will be strict and require constant Fourier phase.

In order to formalize the arguments concerning the sign of these types of functions and of their derivatives, we must assume that F_e^\wedge , F_o^\wedge , $X F_e^\wedge$, and $X F_o^\wedge$ are in L^1 (where X is the function $x \mapsto x$). Indeed, for any $F \in L^1$ with $X F \in L^1$, by (e) F^\wedge is derivable and $F^{\wedge'} = [-2\pi i X F]^\wedge$, from which we get $F^{\vee'} = [2\pi i X F]^\vee$. Then by (d, f) for a signal $I \in L^1$ we get $I * F_e = [I^\wedge F_e^\wedge]^\vee$, $(I * F_e)' = [2\pi i X I^\wedge F_e^\wedge]^\vee$, $I * F_o = [I^\wedge F_o^\wedge]^\vee$, and $(I * F_o)' = [2\pi i X I^\wedge F_o^\wedge]^\vee$, so that $(I * F_e)(0) = \int_{\mathbb{R}} I^\wedge F_e^\wedge$, $(I * F_e)'(0) = \int_{\mathbb{R}} 2\pi i X I^\wedge F_e^\wedge$, $(I * F_o)(0) = \int_{\mathbb{R}} I^\wedge F_o^\wedge$, and $(I * F_o)'(0) = \int_{\mathbb{R}} 2\pi i X I^\wedge F_o^\wedge$. The integrated functions having a sign which is "generally" constant, their integrals over \mathbb{R} should normally get that sign. In view of the results proved

above, if we require constant Fourier phase for our functions of types P, N, NP, and PN, everything we said in Subsections 6.2 and 6.3 concerning peaks of these functions or their derivatives at 0 will be strictly true.

If one locates features at maxima of the energy function $(I * F_e)^2 + (I * F_o)^2$, besides requiring constant phases for the masks, say $F_e^{\Phi}(\nu) = 0$ and $F_o^{\Phi}(\nu) = \pi/2$ for all $\nu > 0$, we must also demand equal Fourier amplitudes, namely $F_e^{\mathcal{A}} = F_o^{\mathcal{A}}$. As shown formally in [26] this guarantees that the energy feature detector will effectively locate features at points of maximum phase congruency, but also that it will be invariant under grey-level shift, since we will have $\int_{\mathbb{R}} F_e = \int_{\mathbb{R}} F_o = 0$. We refer the reader to that paper for more mathematical details.

In order to apply this analog model to digital images, the masks will be digitized by the method of Hummel and Lowe [10]. A *sampling* σ of analog convolution masks into digital ones must be considered in relation with a *reconstruction* ρ of digital signals into analog ones, where σ and ρ commute with all digital translations and with the reflection about the origin, in such a way that for a digital signal D and an analog mask F we have $[\rho(D) * F](p) = [D * \sigma(F)](p)$ for every digital point p . As $\rho(D)$ will be in L^1 while F is chosen in L^1 and uniformly continuous, $\rho(D) * F$ will also be uniformly continuous by (a), so that its peaks will appear in the digital convolution $D * \sigma(F)$, provided that the resolution of the sampling is fine enough.

REFERENCES

1. P. Brou, T.R. Sciascia, L. Linden, J.Y. Lettvin: The colors of things. *Scientific American*, Vol. 255, no. 3, pp. 80–77, Sep. 1986.
2. G. Bruhat: *Cours de Physique Générale: Optique*, 5e. édition, Masson & Cie., Paris (1959).
3. D.C. Burr, M.C. Morrone, D. Spinelli: Evidence for edge and bar detectors in human vision. *Vision Research*, Vol. 29, no. 4, pp. 419–431, 1989.
4. J.G. Daugman: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am.*, Ser. A, Vol. 2, no. 7, pp. 1160–1169, Jul. 1985.
5. J.G. Daugman: Pattern and motion vision without Laplacian zero crossing. *J. Opt. Soc. Am.*, Ser. A, Vol. 5, no. 7, pp. 1142–1148, Jul. 1988.
6. E. De Micheli, B. Caprile, P. Ottonello, V. Torre: Localization and noise in edge detection. *IEEE Trans. Pattern Analysis & Machine Intelligence*, Vol. PAMI-11, no. 10, pp. 1106–1117, Oct. 1989.
7. D. Gabor: Theory of communication. *J. IEE*, Vol. 93, pp. 429–457, 1946.
8. A.L. Gilchrist, A. Jacobsen: Perception of lightness and illumination in a world of one reflectance. *Perception*, Vol. 13, no. 1, pp. 5–19, 1984.
9. B.K.P. Horn: Understanding image intensities. *Artificial Intelligence*, Vol. 8, pp. 201–231, 1977.
10. R. Hummel, D. Lowe: Computational considerations in convolution and feature-extraction in images. *From Pixels to Features*, J.C. Simon ed., Elsevier Science Publishers

- (North-Holland), Amsterdam (1989), pp. 91–102.
11. S.A. Klein, D.M. Levi: Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation. *J. Opt. Soc. Am., Ser. A*, Vol. 2, no. 7, pp. 1170–1190, Jul. 1985.
 12. K. Koffka: *Principles of Gestalt Psychology*. Harcourt, Brace & Co., New York (1935).
 13. V. Lacroix: The primary raster: a multiresolution image description. *Proc. 10th. Internat. Conf. Pattern Recognition*, Atlantic City, NJ (1990), Vol. I, pp. 903–907.
 14. D. Lee: Coping with discontinuities in computer vision: their detection, classification, and measurement. *IEEE Trans. Pattern Analysis & Machine Intelligence*, Vol. PAMI-12, no. 4, pp. 321–344, Apr. 1990.
 15. J. Ling: *Analysis of Edge-Detection Algorithms for a Difficult Image Processing Application*. Unpublished Thesis (1989).
 16. M.S. Livingstone: Art, illusion, and the visual system. *Scientific American*, Vol. 258, no. 1, pp. 68–75, Jan. 1988.
 17. M.S. Livingstone, D.H. Hubel: Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, Vol. 240, pp. 740–749, 6 May 1988.
 18. D. Marr: Early processing of visual information. *Phil. Trans. Royal Soc. London, Ser. B*, Vol. 275, pp. 483–519, 1976.
 19. D. Marr: *Vision*. W.H. Freeman & Co., San Francisco, CA (1982).
 20. D. Marr, E. Hildreth: Theory of edge detection. *Proc. Royal Soc. London, Ser. B*, Vol. 207, pp. 187–217, 1980.
 21. M.C. Morrone, D.C. Burr: Feature detection in human vision: a phase-dependent energy model. *Proc. Royal Soc. London, Ser. B*, Vol. 235, pp. 221–245, 1988.
 22. M.C. Morrone, R. Owens: Feature detection from local energy. *Pattern Recognition Letters*, Vol. 6, pp. 303–313, 1987.
 23. M.C. Morrone, J. Ross, D.C. Burr, R. Owens: Mach bands are phase dependent. *Nature*, Vol. 324, pp. 250–253, 20 Nov. 1986.
 24. V.S. Nalwa, T.O. Binford: On detecting edges. *IEEE Trans. Pattern Analysis & Machine Intelligence*, Vol. PAMI-8, no. 6, pp. 699–714, Nov. 1986.
 25. R. Owens, S. Venkatesh, J. Ross: Edge detection is a projection. *Pattern Recognition Letters*, Vol. 9, pp. 233–244, 1989.
 26. C. Ronse: On the idempotence of edge detection and the relevance of energy feature operators, (second version). Submitted.
 27. J. Ross, R. Owens: Sketches: man-made and machine-made. *Leonardo*, to appear.
 28. A.N. Tikhonov, V.Y. Arsenin: *Solutions of Ill-Posed Problems*. V.H. Winston & Sons, Washington (1977).
 29. V. Torre, T.A. Poggio: On edge detection. *IEEE Trans. Pattern Analysis & Machine Intelligence*, Vol. PAMI-8, no. 2, pp. 147–163, Mar. 1986.
 30. S. Venkatesh, R. Owens: On the classification of image features. *Pattern Recognition Letters*, Vol. 11, pp. 339–349, 1990.
 31. R.J. Watt: *Visual Processing : Computational, Psychophysical, and Cognitive Research*. Lawrence Erlbaum Ass., London (1988).



FIGURE 1.1. *Extended grey-level edge; this is a typical intensity profile for the edge of a shadow.*

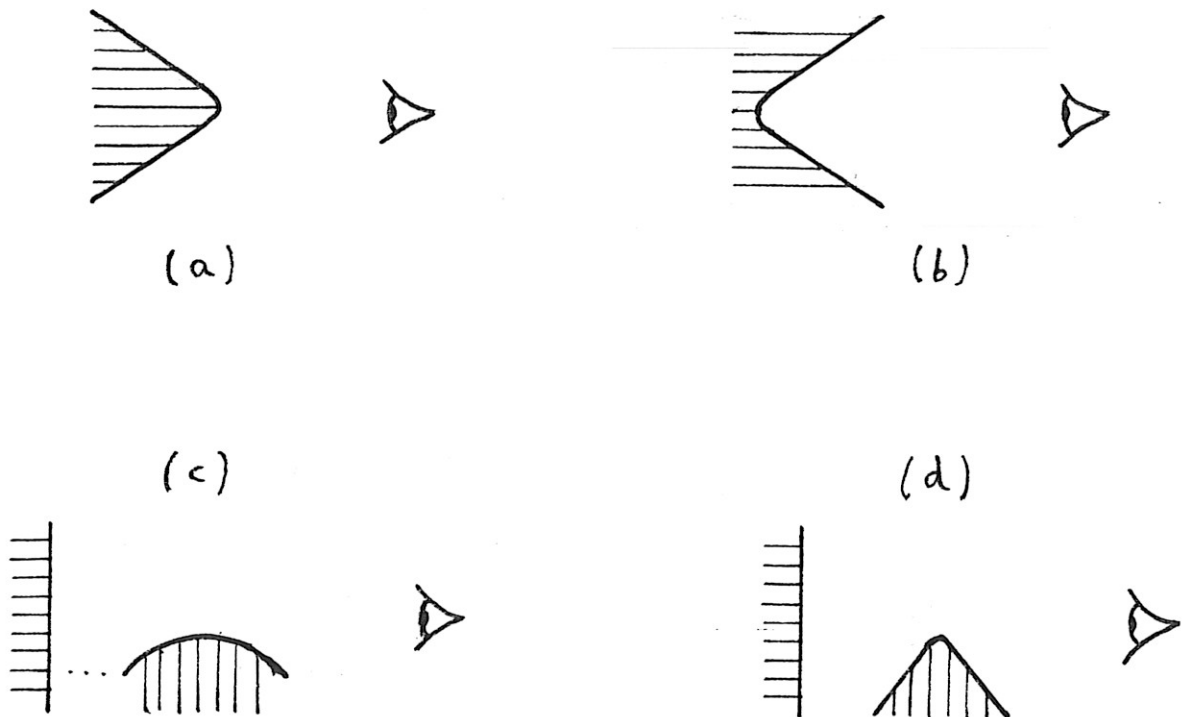


FIGURE 1.2. *Types of geometrical edges in a scene: (a) convex edge; (b) concave edge; (c) occluding edge; (d) combination of occluding and convex edges.*

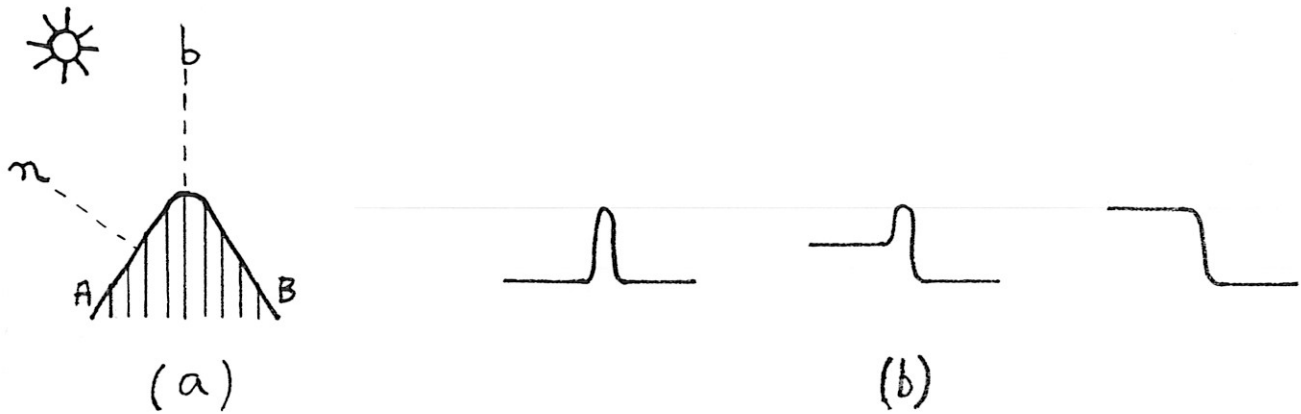


FIGURE 1.3. (a) Convex edge under coherent illumination: n and b are the directions respectively normal to face A and bisecting faces A and B . (b) Resulting grey-level profile according to the direction i of the light source: a line edge when i is at b , a superposition of line and step edges when i is between b and n , and a slightly rounded step edge when i is at n or further.

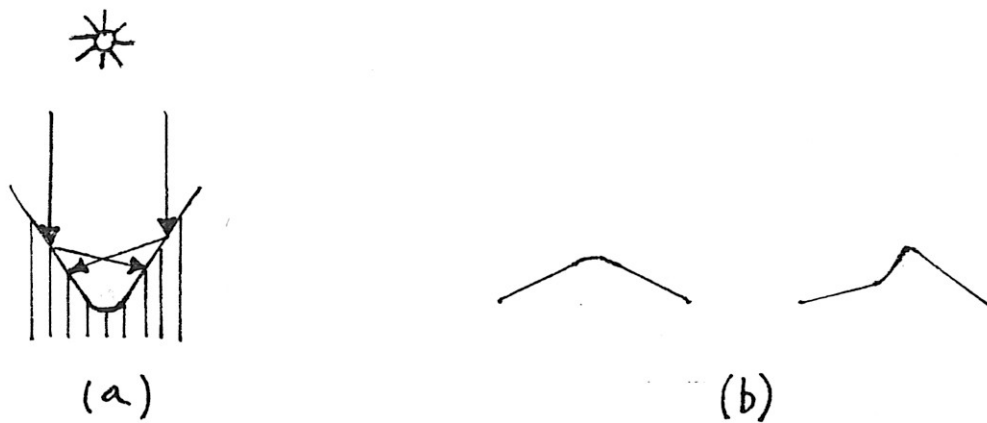


FIGURE 1.4. (a) Mutual illumination of faces in a concave edge. (b) The resulting secondary component of grey-level profile forms a roof, which is symmetric when the direction of the light source bisects the two faces, and otherwise has orientation opposite to that of the step edge due to primary illumination.

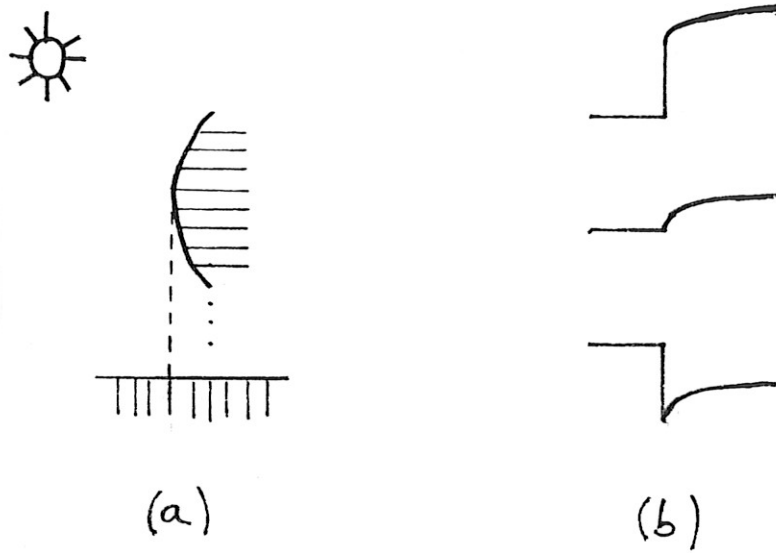


FIGURE 1.5. (a) A non-shadowing occluding edge. (b) The resulting grey-level profile, according to the orientation of the light source.

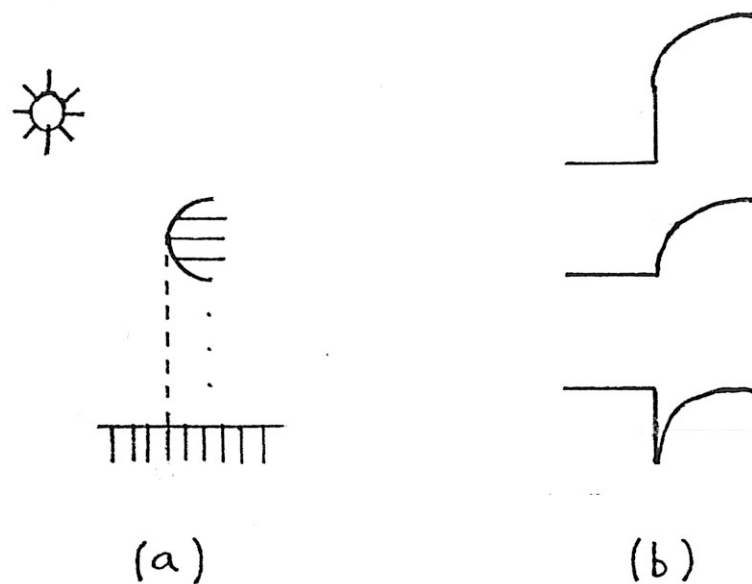


FIGURE 1.6. The same situation as in Figure 1.5, but where the radius of curvature of the occluding edge is small.

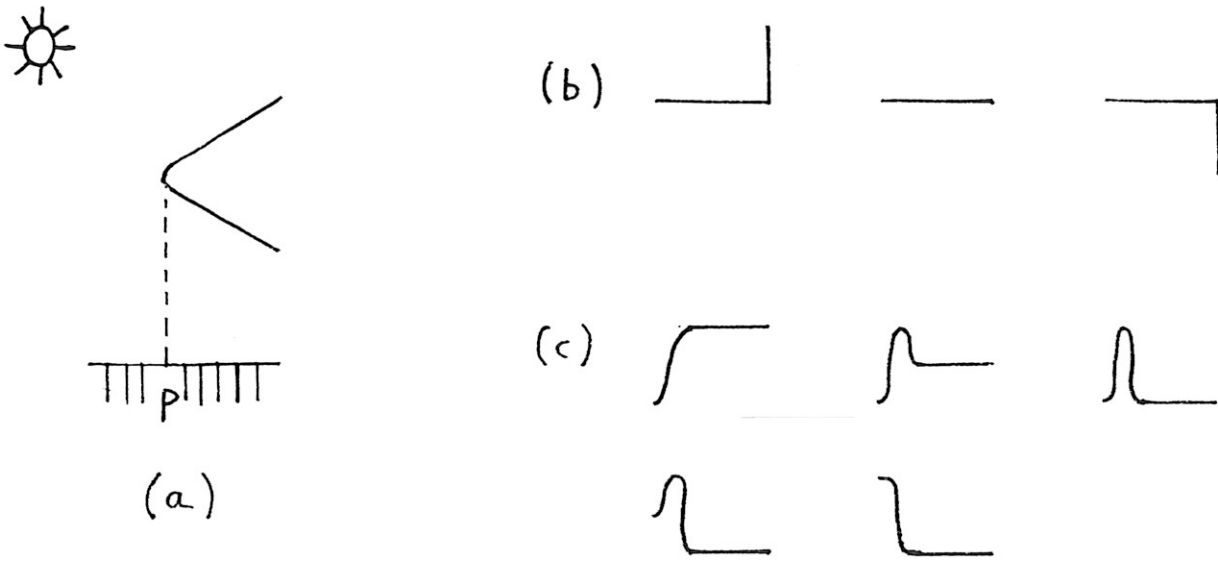


FIGURE 1.7. (a) A non-shadowing occluding edge which is at the same time convex. The resulting grey-level profile is a succession of the grey-level profiles corresponding to both types of edges, namely: (b) to the left of p , a sharp step, and (c) to the right of p , a linear combination of a step and a positive line. Note that not all combinations from (b) and (c) are necessarily found in real situations.

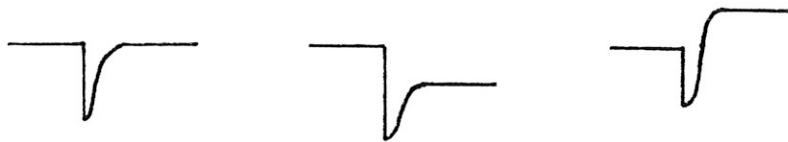


FIGURE 1.8. Some frequent grey-level profiles arising in the situation of Figure 1.7.

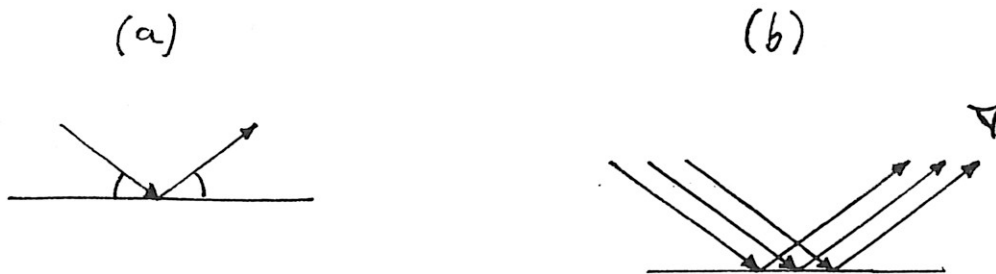


FIGURE 1.9. (a) Specular reflectance. (b) Under coherent illumination, it leads to a positive peak or line edge in the intensity profile

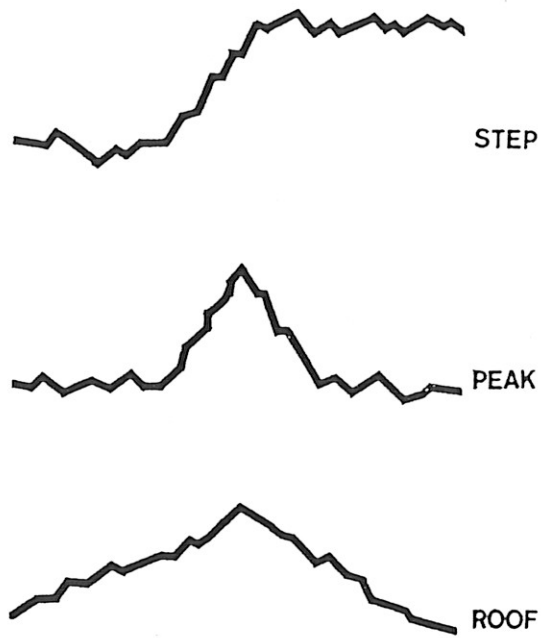


FIGURE 1.10. Classification of edge types according to Herskovits and Binford.

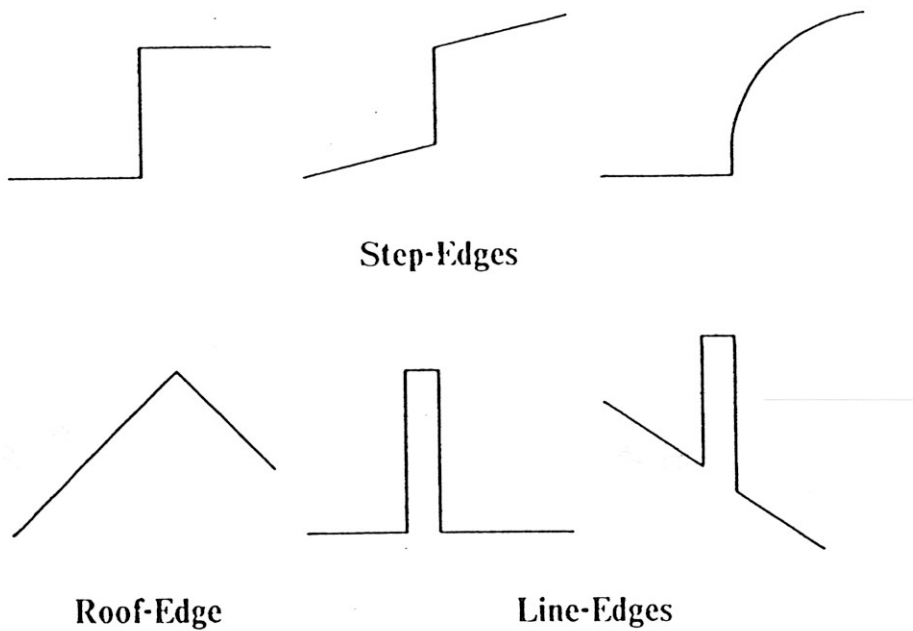
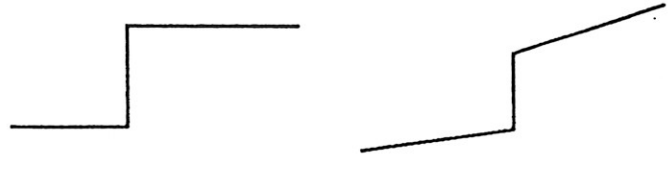


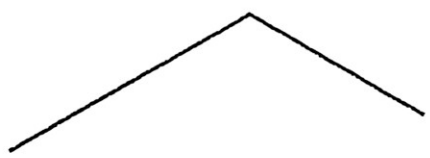
FIGURE 1.11. Classification of edge types according to Nalwa and Binford.



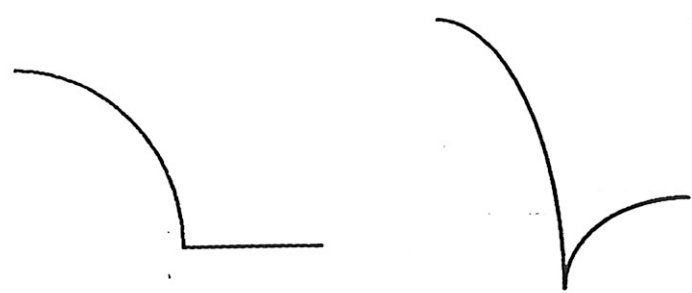
a) step edge



b) line edge



c) roof edge



d) convex edge and valley-edge

FIGURE 1.12. Classification of edge types given by Ling.

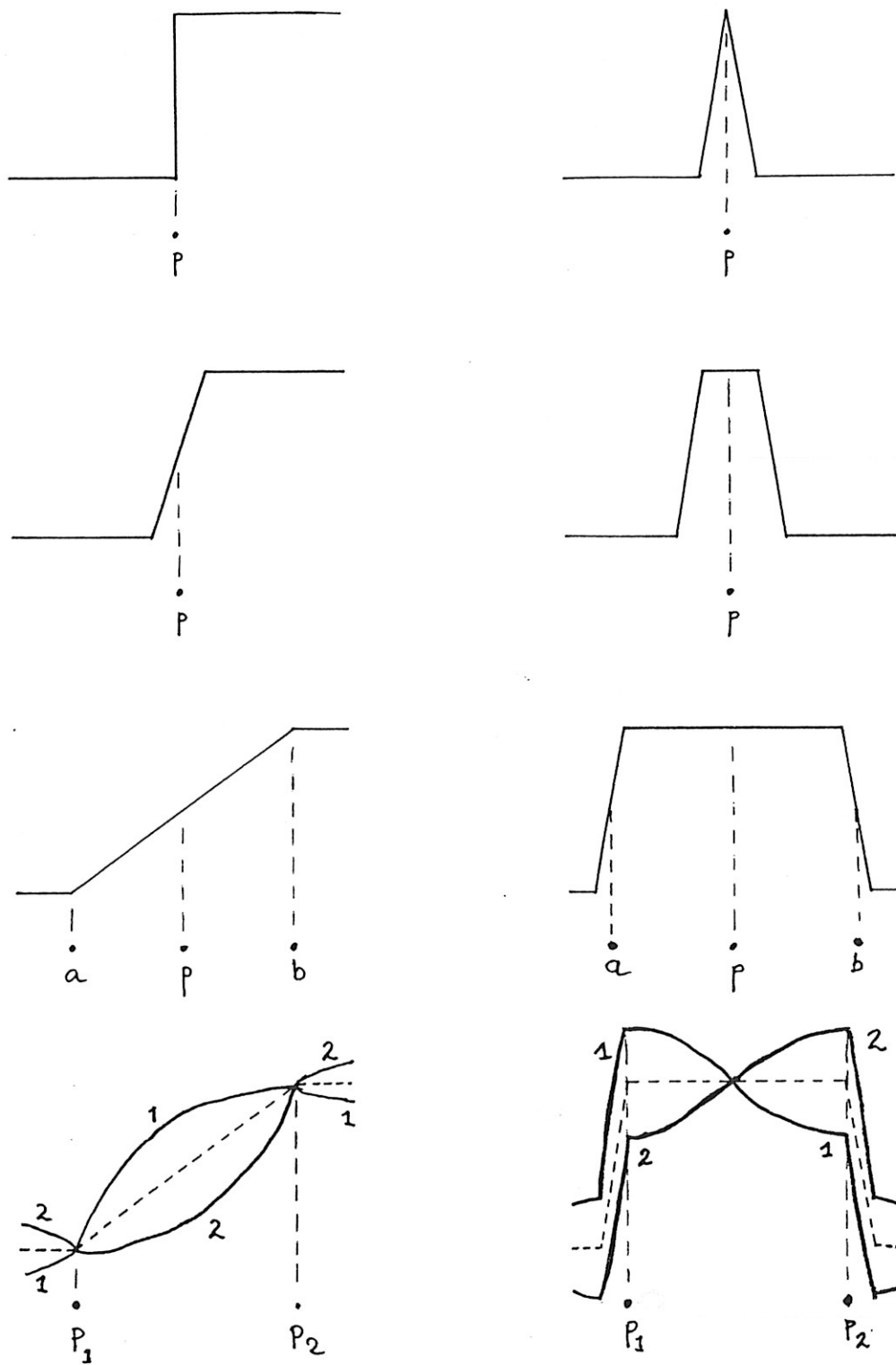


FIGURE 2.1. (a) A sharp step or line edge located at p . (b) A gradual step or line edge is located at the point p in the middle of the ramp or plateau. (c) In an extended edge, the localization at p becomes arbitrary. (d) Adding a low frequency sinusoid (1 or 2) to an extended edge shifts the localization point p (to p_1 and p_2 respectively).

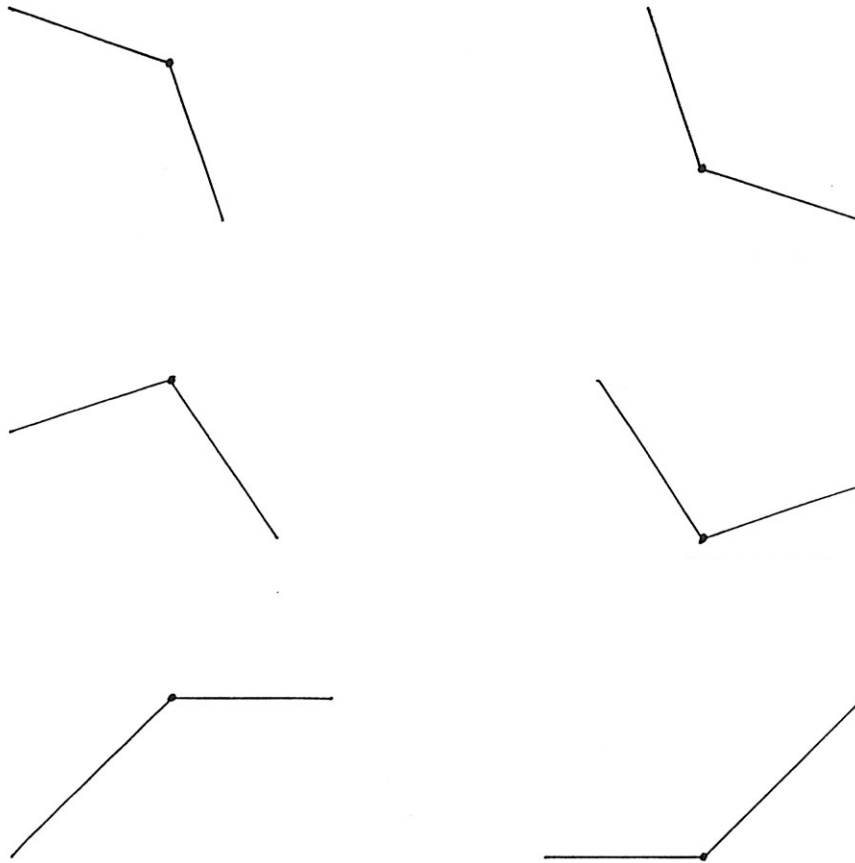


FIGURE 2.2. Some grey-level roofs and Mach bands. Those in the left column are positive, while those in the right column are negative, leading to the perception of respectively light and dark lines.

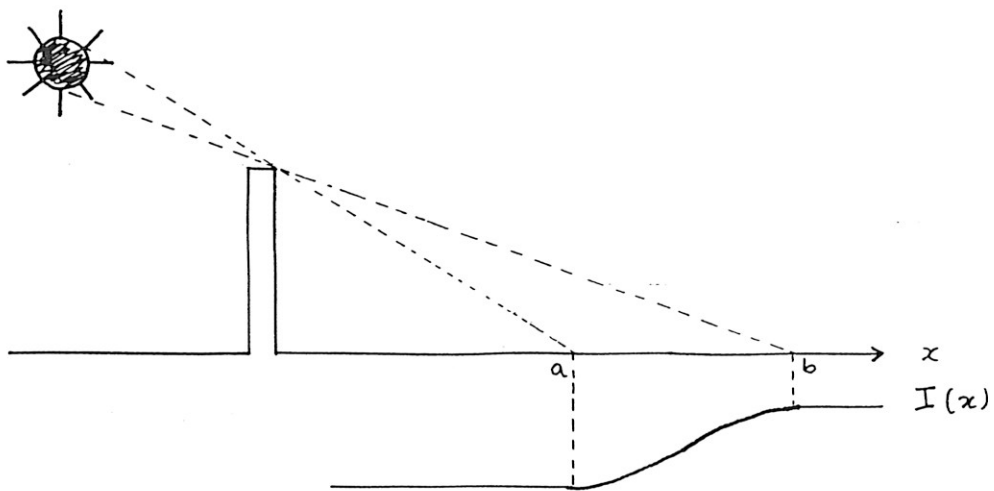


FIGURE 2.3. A scene where a shadow has an extended edge.

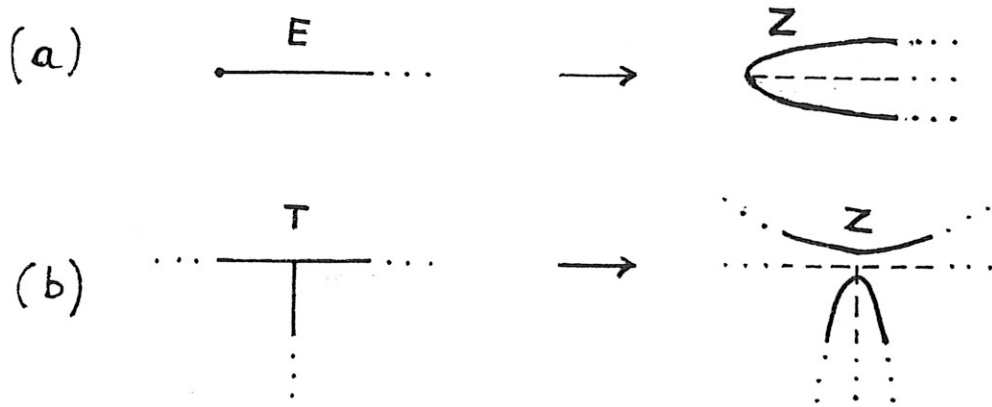


FIGURE 4.1. Two situations where an edge E cannot be the frontier between two regions, and the type of zero-crossing set Z one can obtain then: (a) an edge endpoint; (b) an edge T-junction.

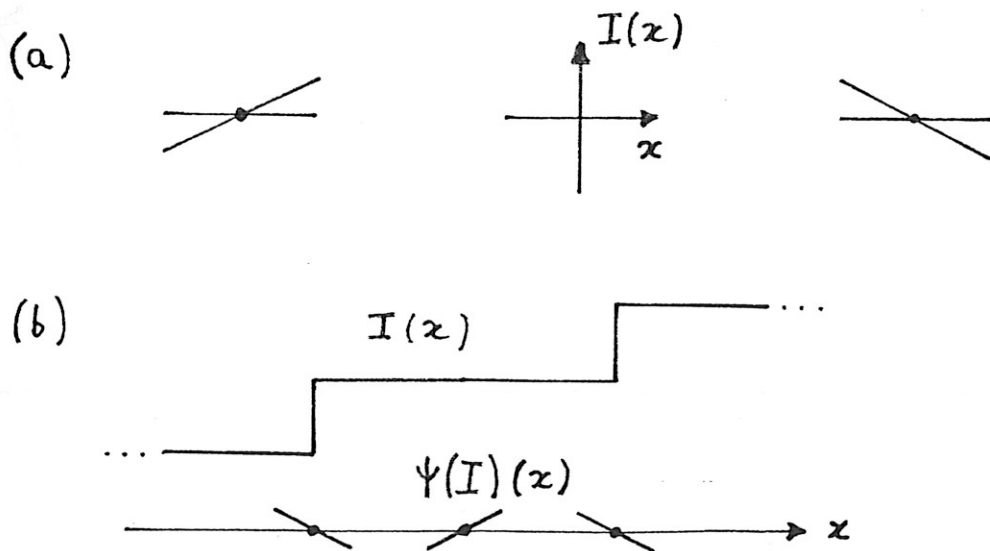


FIGURE 4.2. (a) Zero-crossings along a direction are of two types: negative-to-positive (left) and positive-to-negative (right). (b) In a staircase step edge, between the two zero-crossings of one type found at two successive steps lies a spurious zero-crossing of the other type.

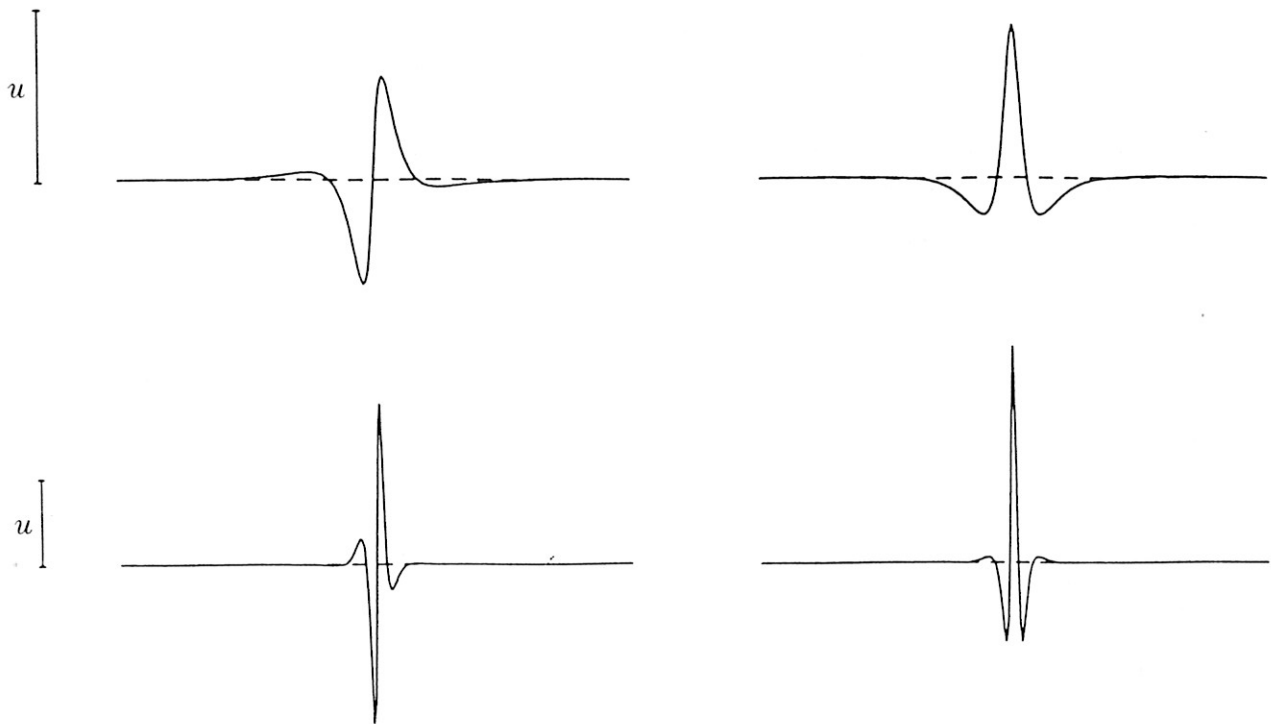


FIGURE 5.1. *Odd-symmetric and even-symmetric detection functions used by Morrone and Burr. Each line corresponds to a given size scale, namely the second and third in a series of four, with u representing one unit of amplitude. Note that the functions of the second row are not magnifications of those of the first one.*

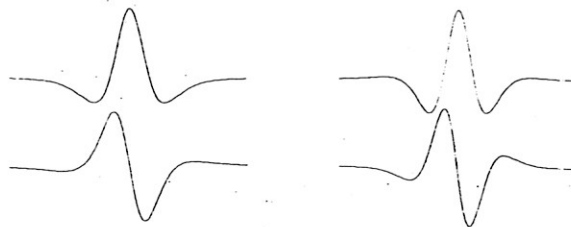


FIGURE 5.2. *Some odd-symmetric and even-symmetric detection functions proposed by Klein and Levi.*

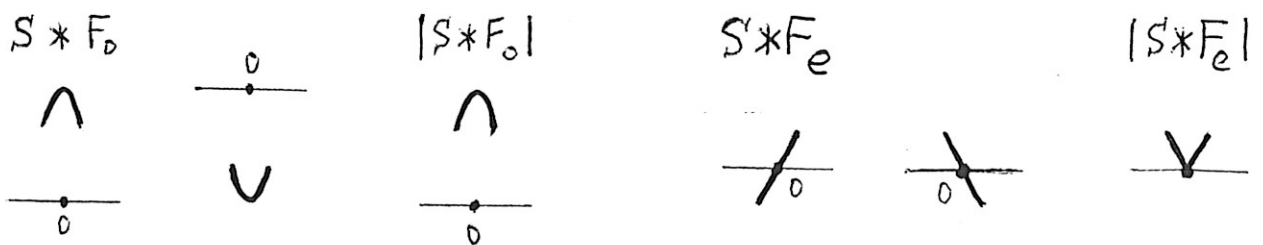


FIGURE 6.1. $|S * F_o|'(0) = \pm(S * F_o)'(0) = 0$, while the left and right derivatives of $|S * F_e|$ at 0 are $-|(S * F_e)'(0)|$ and $+(S * F_e)'(0)|$ respectively.

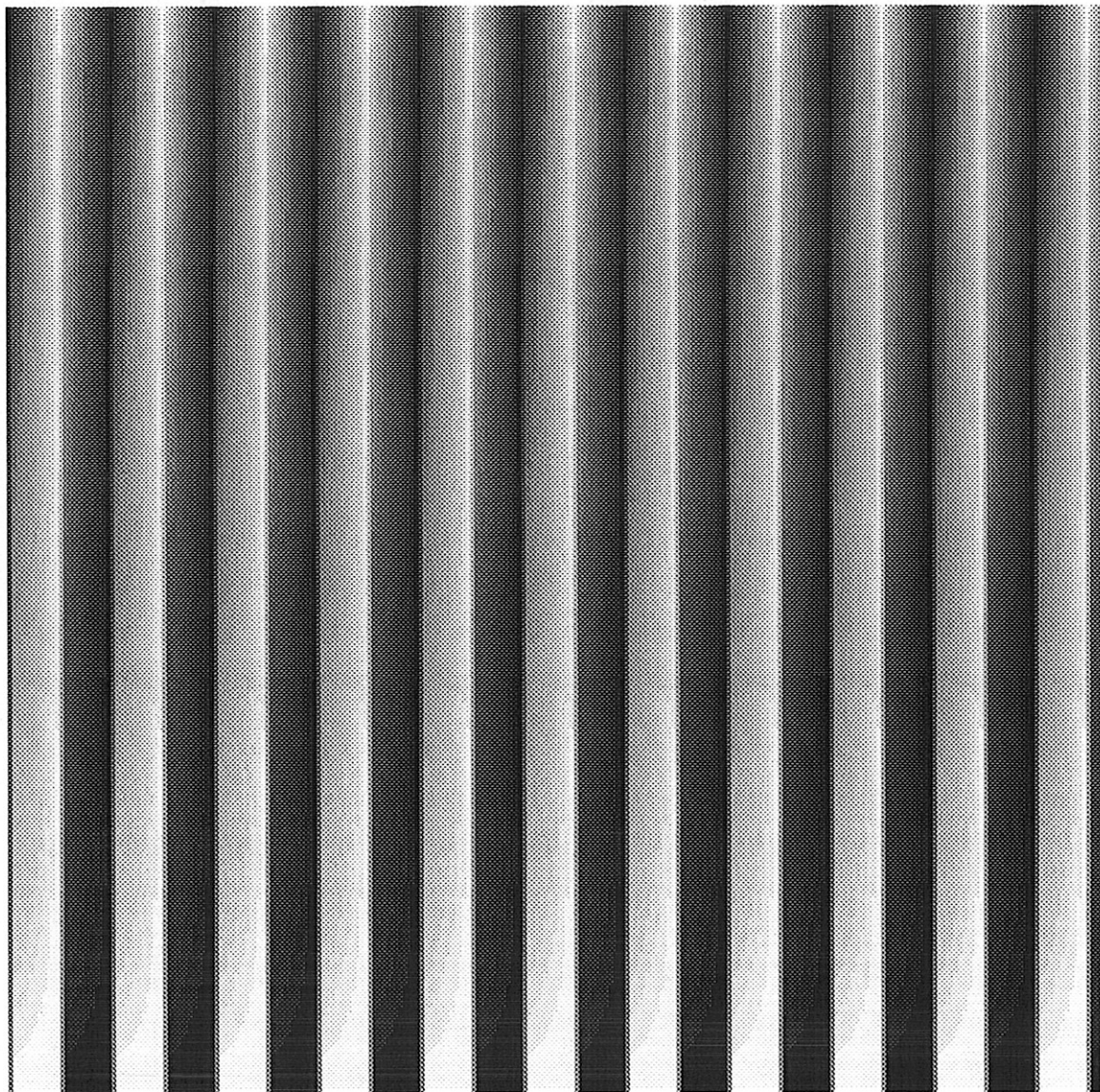


FIGURE 6.2. Vertical grating whose horizontal sections are varying from a square wave to a triangular one. In the middle, where both waves are superposed, the edges (combining a step and a roof) are seen slightly to the left of their true location. This is consistent with the Morrone-Burr model of human perception of local features at points of Fourier phase congruency.