The Plenoptic Function and the Elements of Early Vision

Article · August 1997

Source: CiteSeer

CITATIONS
1,198

READS
1,001

2 authors, including:

James R. Bergen
SRI International
82 PUBLICATIONS **13,989** CITATIONS

SEE PROFILE

From M. Landy and J. A. Movshon (eds), *Computational Models of Visual Processing* (pp. 3-20). Cambridge, MA: MIT Press (1991).

1

The Plenoptic Function and the Elements of Early Vision

Edward H. Adelson and
James R. Bergen

What are the elements of early vision? This question might be taken to mean, What are the fundamental atoms of vision?—and might be variously answered in terms of such candidate structures as edges, peaks, corners, and so on. In this chapter we adopt a rather different point of view and ask the question, What are the fundamental *substances* of vision? This distinction is important because we wish to focus on the first steps in extraction of visual information. At this level it is premature to talk about discrete objects, even such simple ones as edges and corners.

There is general agreement that early vision involves measurements of a number of basic image properties including orientation, color, motion, and so on. Figure l.l shows a caricature (in the style of Neisser, 1976), of the sort of architecture that has become quite popular as a model for both human and machine vision. The first stage of processing involves a set of parallel pathways, each devoted to one particular-visual property. We propose that the measurements of these basic properties be considered as the elements of early vision. We think of early vision as measuring the amounts of various kinds of visual "substances" present in the image (e.g., redness or rightward motion energy). In other words, we are interested in how early vision measures “stuff” rather than in how it labels “things.”

What, then, are these elementary visual substances? Various lists have been compiled using a mixture of intuition and experiment. Electrophysiologists have described neurons in striate cortex that are selectively sensitive to certain visual properties; for reviews, see Hubel (1988) and DeValois and DeValois (1988). Psychophysicists have inferred the existence of channels that are tuned for certain visual properties; for reviews, see Graham (1989), Olzak and Thomas (1986), Pokorny and Smith (1986), and Watson (1986). Researchers in perception have found aspects of visual stimuli that are processed pre-attentively (Beck, 1966; Bergen & Julesz, 1983; Julesz & Bergen,

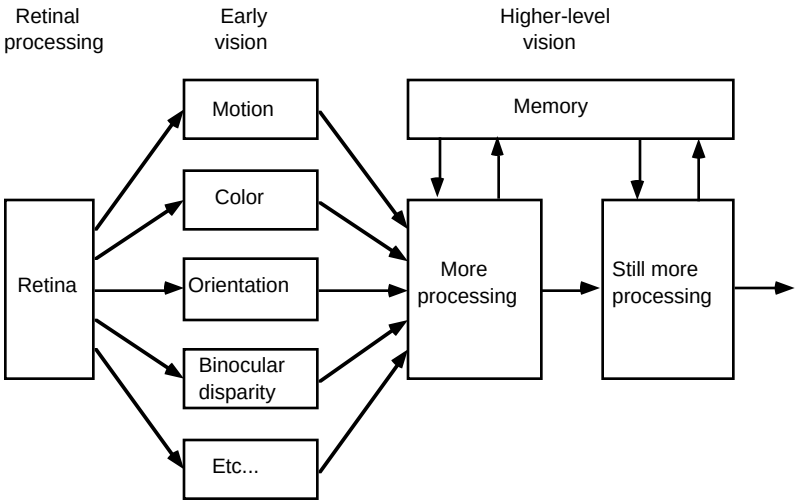


Fig. 1.1
A generic diagram for visual processing. In this approach, early vision consists of a set of parallel pathways, each analyzing some particular aspect of the visual stimulus.

1983; Treisman, 1986; Treisman & Gelade, 1980). And in computational vision, investigators have found that certain low-level measurements are useful for accomplishing vision tasks; for examples, see Horn (1986), Levine (1985), and Marr (1982).

These various approaches have converged on a set of superficially similar lists, but there is little sense of structure. Why do the lists contain certain elements and not others? Or, indeed, are there other unknown visual elements waiting to be discovered?

Our interest here is to derive the visual elements in a systematic way and to show how they are related to the structure of visual information in the world. We will show that all the basic visual measurements can be considered to characterize local change along one or more dimensions of a single function that describes the structure of the information in the light impinging on an observer. Since this function describes everything that can be seen, we call it the *plenoptic function* (from *plenus*, complete or full, and *optic*). Once we have defined this function, the measurement of various underlying visual properties such as motion, color, and orientation fall out of the analysis automatically.

Our approach generates a list of the possible visual elements, which we think of as somewhat analogous to Mendeleev's periodic table in the sense that it displays systematically all the elemental substances upon which vision can be based. This table catalogues the basic visual substances and clarifies their relationships.

This cataloging process makes no assumptions about the statistics of the world and no assumptions about the needs of the observing organism. The periodic table lists every sim-ple visual measurement that an observer could potentially make, given the structure of the ambient light expressed in the plenoptic function. A given organism will probably not measure all of the elements, and of those that it measures it will devote more resources to some than to others.

In what follows we will make reference to some of the relevant psychophysical and physiological literature on early vision. Our topic is quite broad, however, and if we were to cite all of the relevant sources we would end up listing hundreds of papers. Therefore our references will be sparse, and readers are encouraged to consult the various books and review articles cited above.

The Plenoptic Function

We begin by asking what can potentially be seen. What information about the world is contained in the light filling a region of space? Space is filled with a dense array of light rays of various intensities. The set of rays passing through any point in space is mathematically termed a *pencil*. Leonardo da Vinci refers to this set of rays as a "radiant pyramid":

The body of the air is full of an infinite number of radiant pyramids caused by the objects located in it. These pyramids intersect and interweave without interfering with each other

during the independent passage throughout the air in which they are infused. (Kemp, 1969)

If a pinhole camera happens to be positioned at a given point, it will select the pencil of rays at that point and will reveal that they form an image. An observer's eye acts in the same way (neglecting the finite aperture of the pupil for now): It reveals the structure of the pencil of light at the pupil's location. The fact that image information fills space is expressed by Leonardo in his notebooks; he invites the reader to perform a thought experiment:

I say that if the front of a building—or any open piazza or field—which is illuminated by the sun has a dwelling opposite to it, and if, in the front which does not face that sun, you make a small round hole, all the illuminated objects will project their images through that hole and be visible inside the dwelling on the opposite wall which may be made white; and there, in fact, they will be upside down, and if you make similar openings in several places in the same wall you will have the same result from each. Hence the images of the illuminated objects are all everywhere on this wall and all in each minutest part of it. (Richter, 1970)

J. J. Gibson referred to a similar notion when he spoke of the structure of ambient light (Gibson, 1966) "The complete set of all convergence points . . . constitutes the permanent possibilities of vision, that is, the set of all points where a mobile individual might be."

Let us follow this line of thought a bit further and consider the parameters necessary to describe this luminous environment. Consider, first, a black and white photograph taken by a pinhole camera. It tells us the intensity of light seen from a single viewpoint, at a single time, averaged over the wavelengths of the visible spectrum. That is to say, it records the intensity distribution P within the pencil of light rays passing through the lens. This distribution may be parameterized by the spherical coordinates, $P(\theta,\phi)$, or by the Cartesian coordinates of a picture plane, $P(x,y)$ (figure 1.2; see discussion below). A color photograph adds some infor-mation about how the intensity vanes with wavelength λ, thus: $P(\theta,\phi,\lambda)$. A color movie further extends the information to include the time dimension t: $P(\theta,\phi,\lambda,t)$. A color holographic movie, finally, indicates the observable light intensity at every viewing position, V_x, V_y, and V_z: $P(\theta,\phi,\lambda,t,V_x,V_y,V_z)$. A true holographic movie would allow reconstruction of every possible view, at every moment, from every position, at every wavelength, within the bounds of the space-time-wavelength region under consideration. The plenoptic functions is equivalent to this complete holographic representation of the visual world.

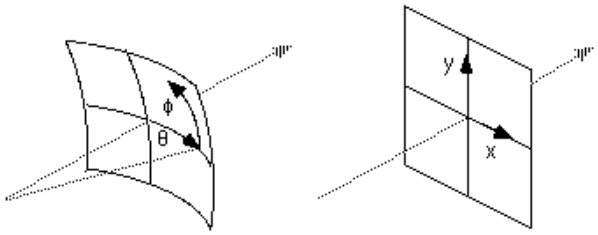


Fig. 1.2
The image information available from a single viewing position is defined by the pencil of light rays passing through the pupil. The rays may be parameterized in angular coordinates or in Cartesian coordinates. The Cartesian approach is commonly used in machine vision and computer graphics, but the angular approach can more easily represent the full sphere of optical information impinging on a point in space.

Such a complete representation would contain, implicitly, a description of every possible photograph that could be taken of a particular space-time chunk of the world (neglecting the polarization and instantaneous phase of the incoming light). Note that the plenoptic function need not contain any parameters specifying the three viewing angles describing the direction of gaze and orientation of the eye, since rotating the eye without displacing it does not affect the distribution of light in the bundle of rays impinging on the pupil, but merely changes the relative positions at which they happen to strike the retina. The fact that some rays are behind the eye and are therefore blocked is irrelevant to the present discussion, which is intended to characterize the optical information potentially available at each point in space, as if the hypothetical eye had a 360° field of view. Figure 1.3 shows a pair of samples from this function, with the eye placed at different positions in a natural scene.

To measure the plenoptic function one can imagine placing an idealized eye at every possible (V_x,V_y,V_z) location and recording the intensity of the light rays passing through the center of the pupil at every possible angle (θ,ϕ), for every wavelength, λ, at every time t. It is simplest to have the eye always look in the same direction, so that the angles (θ,ϕ) are always computed with respect to an optic axis that is parallel to the V_z axis. The resulting function takes the form:

$$P = P(\theta,\phi,\lambda,t,V_x,V_y,V_z). \quad (1)$$

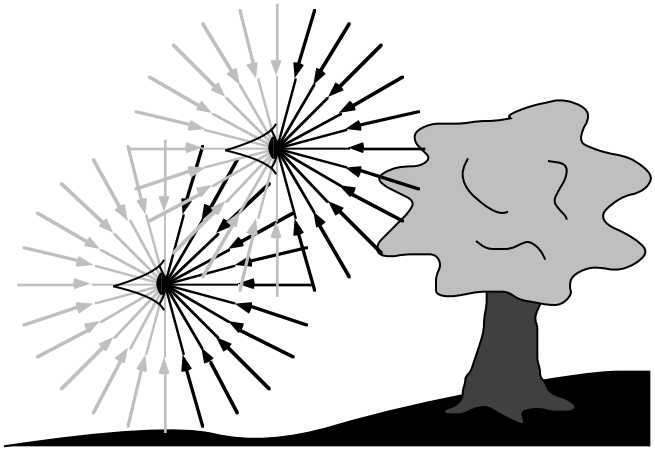


Fig. 1.3
The plenoptic function describes the information available to an observer at any point in space and time. Shown here are two schematic eyes-which one should consider to have punctate pupils-gathering pencils of light rays. A real observer cannot see the light rays coming from behind, but the plenoptic function does include these rays.

Alternatively, one may choose to parameterize the rays entering the eye in terms of (x,y) coordinates, where x and y are the spatial coordinates of an imaginary picture plane erected at a unit distance from the pupil. This is the approach commonly adopted in computer graphics and machine vision. The parameterization then becomes:

$$P = P(x,y,\lambda,t,V_x,V_y,V_z). \quad (2)$$

The spherical parameterization more easily suggests the fact that the light impinges on a given point in space from all directions and that no direction has special status. However, the Cartesian parameterization is more familiar, and we will use it in the discussion that follows.

The plenoptic function is an idealized concept, and one does not expect to completely specify it for a natural scene. Obviously one cannot simultaneously look at a scene from every possible point of view, for every wavelength, at every moment of time. But, by describing the plenoptic function, one can examine the structure of the information that is potentially available to an observer by visual means.

The significance of the plenoptic function is this: The world is made of three-dimensional objects, but these objects do not communicate their properties directly to an observer. Rather, the objects fill the space around them with the pattern of light rays that constitutes the plenoptic function, and the observer takes samples from this function. The plenoptic function serves as the sole communication link between physical objects and their corresponding retinal images. It is the intermediary between the world and the eye.

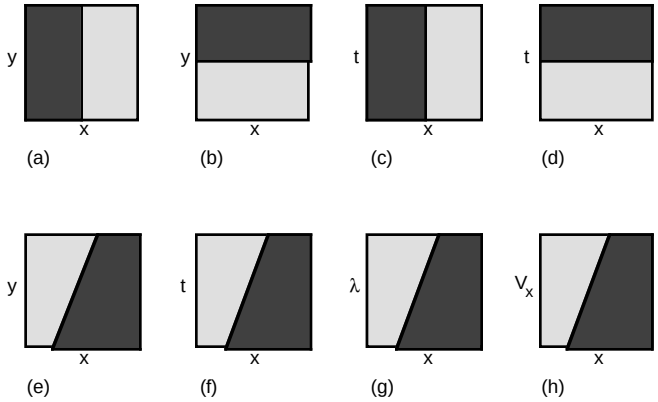


Fig. 1.4
Some edgelike structures that might be found along particular planes within the plenoptic function (note the venous axes, as labeled on each figure): (a) a vertical edge; (b) a horizontal edge, (c) a stationary edge; (d) a full-field brightening; (e) a tilting edge; (f) a moving edge; (g) a color sweep; (h) an edge with horizontal binocular parallax.

Plenoptic Structures

It may initially appear that the plenoptic function is extremely complicated. Since it has seven dimensions, it is difficult to visualize. However, much of the information that the plenoptic function contains describes familiar structures of the visual world. One can develop a sense of the structure of the plenoptic function as a whole by considering some planar slices.

Figure l.4 shows a variety of slices along various planes passing through the x-axis. The figure illustrates a number of edgelike structures that might occur within a given plenoptic function. Part A, in the (x,y) plane, shows a vertical edge. Part B, also in the (x,y) plane, shows a horizontal edge. Part C, in the (x,t) plane, shows a stationary edge. Part D, also in the (x,t) plane, shows a temporal edge a sudden increase in intensity. Part E, in the (x,y) plane, shows a tilted edge. Part F, in the (x, t) plane, shows a moving edge. Part G, in the (x,λ) plane, shows a patch that changes color across space. Part H, in the (x,V_x) plane, shows an edge that projects a changing

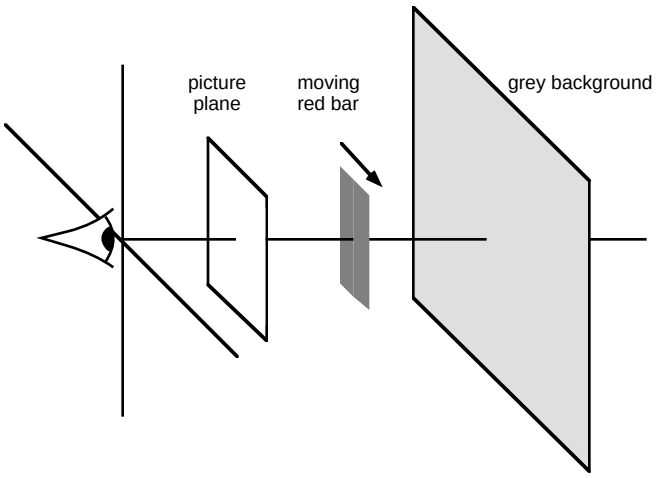


Fig. 1.5
A hypothetical scene that produces a variety of simple plenoptic structures. See figure 1.6.

retinal image as the eye is translated in the V_x direction—that is, an edge with horizontal parallax.

Figure 1.5 shows a concrete example of a particular simple scene. The eye views a red bar at a certain distance, moving against a grey background at a larger distance. The figure also shows the imaginary picture plane at a unit distance from the eye.

Figure 1.6 shows a variety of slices through the plenoptic function for this scene. All the slices shown pass through the *x*-axis; thus all of the parameters other than the two diagrammed in the slice are taken to be zero if not otherwise specified. The only exception is wavelength, λ, which we will let default to 550 nm, which is the approximate center of the human visible spectrum.

Part B of figure 1.6 shows an (x,y) slice, which is simply the view at 550 nm, at the moment $t = 0$, given the eye is in position (0, 0, 0). It consists of a darker bar against a lighter background. Part C shows an (x,t) slice, which can be thought of as the temporal record of horizontal raster lines appearing at height y = 0; wavelength is 550 nm, and the eye is at (0, 0, 0). This image consists of a spatiotemporally "tilted" bar, the tilt corresponding to the changing position over time. Part D shows an (x,λ) slice, which one might call a "spatiospectral" or "spatiochromatic" slice, taken at height $y = 0$, time $t = 0$, and viewing position (0, 0, 0). The fact that the bar is red leads to the variation in intensity along the wavelength dimension in the middle region of the slice: Long wavelengths have high intensities while short wavelengths have low intensities. Part E shows an (x, V_x) slice, representing the series of views obtained as the eye position shifts from left to right. Part F shows a similar slice for (x, V_y), as the eye position shifts up and down. Finally, part G shows an (x, V_z) slice, representing the changing image as the eye moves backward or forward.

It is clear from examining these slices that similar structures are to be found along various planes in the plenoptic function. In the case of the example shown, extended edgelike features appear in all planes, with different tilts and curvatures. Each plane offers useful information about different aspects of the stimulus. The (x,y) plane contains information about instantaneous form; the (x,λ) plane contains information about chromatic variation in the x direction; the (x, Y_x) plane contains information about horizontal parallax information that could be gathered either through head motion or through stereo vision; the (x,V_z) plane contains information about the looming that occurs when the eye advances along the *z*-axis, and so on. All of the information in the plenoptic function is potentially useful.

The Plenoptic Function and Elemental Measurements in Early Vision

The Task of Early Vision

We suggest that the first problem of early vision is to extract as much information as possible about the structure of the plenoptic function. The first task of any visual system is to measure the state of the luminous environment. As pointed out by Koenderink and van Doorn (1987), only by representing this information internally can all potential visual information be made available for subsequent analysis. By definition, the state of the luminous environ-ment is described by the plenoptic function. Clearly, only a small portion of the potential information present in this environment can be extracted.

As noted above, much of the structure of the plenoptic function describing simple stimulus configurations may take the form of oriented patterns at various angles within plenoptic space. An oriented pattern is one that changes in one direction more rapidly than in another direction; characterizing this anisotropy is a useful first step in analyzing the local structure of a signal. We therefore offer the following propositions:

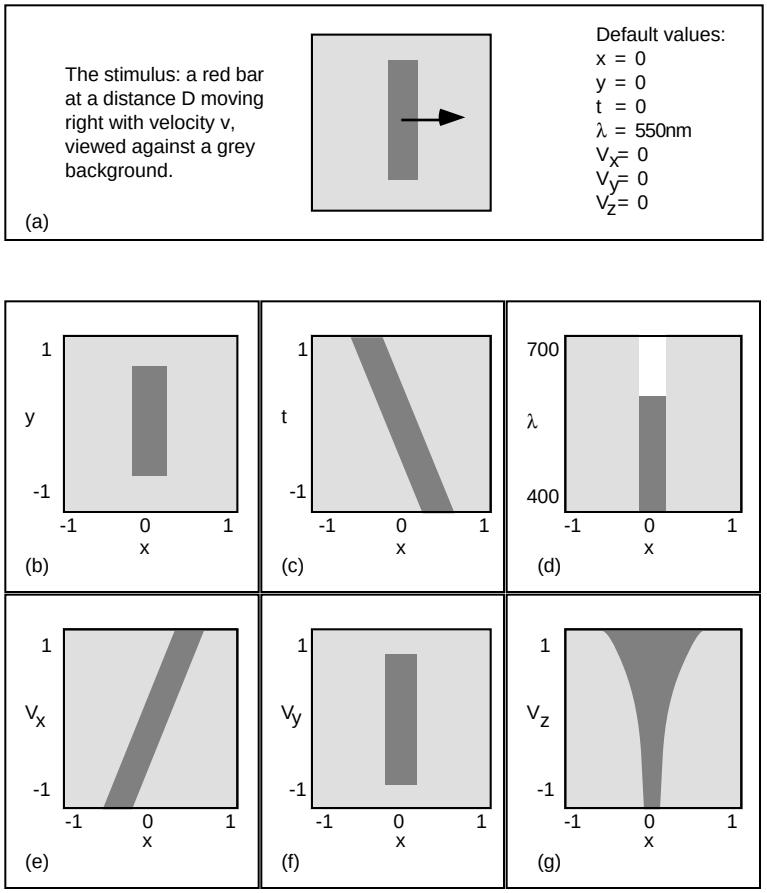


Fig. 1.6
The plenoptic structures found along various planes for the scene illustrated in figure 1.5. Each picture represents a slice through the plenoptic function, where all the unspecified parameters take on their default values.

• *Proposition 1*. The primary task of early vision is to deliver a small set of useful measurements about each observable location in the plenoptic function.

• *Proposition 2*. The elemental operations of early vision involve the measurement of local change along various directions within the plenoptic function.

These two propositions establish a useful "null hypothesis" regarding the existence of early visual mechanisms. If some local descriptor of the plenoptic function exists, then we may (in general) expect a visual system to compute this descriptor. If the system does not do this, we have reason to ask why this particular aspect of visual information is not being extracted. By comparing what is extracted by a visual system (as specified by the basic measurements that it makes) with what might be extracted (specified by the structure of the plenoptic function), we can learn something about the nature of the system's visual process. What exactly are the fundamental measurements implied by these propositions?

Extraction of Information from the Plenoptic Function

Efficient extraction of information from the plenoptic function requires that each sample contain as much information as possible about the most important aspects of the function. An additional desideratum is that the information represented by each sample be simply interpretable without reference to many other samples. Representations motivated by these requirements can take many forms. Since local change defined in arithmetic terms is equivalent to a derivative, a reasonable choice for conceptual purposes consists of the low order directional derivatives of the plenoptic function at the sample points. This set of measures fits well with the goal of capturing the simple structure of the plenoptic function since it con-

sists of oriented linear filters in the plenoptic hyperplanes. In this sense, these derivatives may be thought of as "feature detectors," corresponding to the structures shown in figure 1.6 However, since these are linear filters they are not really detecting features; rather they are measuring the amount of a particular type of local structure.

Another way to think about this type of representation is as a truncated Taylor series expansion around each sample point. Koenderink and van Doorn (1987) have developed a theory of early representation based on *local jets* which pursues such an interpretation. If we compute all of the derivatives at each point, each set of derivatives contains the information present in all of the samples (assuming that the function is smooth). By computing only the low order derivatives, we construct a synopsis of the local structure of the function at each point. This explicit representation of local structure allows analysis of salient local characteristics without repeated examination of multiple samples. The appropriateness of this representation depends on the plenoptic function having the kind of locally correlated structure described in the previous section. It would not make sense if the plenoptic function looked like uncorrelated random noise or were otherwise chaotic in structure.

Mathematically, a derivative is taken at a point, but for characterizing a function within a neighborhood it is more useful to work with the local average derivative of the function, or the derivative of the local average of the function, which is the same thing. The equivalence of these alternate processes follows trivially from the commutativity of linear systems, but it is worth emphasizing, as illustrated in figure 1.7. Let the original function $f(s)$, where s is the variable of interest, have the form of a trapezoid for purposes of illustration. The figure shows three paths to the same result. In one path, a local average is taken by convolving $f(s)$ with a smoothing function, $g(s)$; then this smoothed function is differentiated to produce the final result, $f(s) \vee g'(s)$. In a second path, the derivative is taken, and then this function is smoothed by local averaging. In the third path, shown on the diagonal, the two steps are combined into one: A single linear filter (the derivative of the smoothing function) is used to produce the final output.

Along with the three paths are three verbal descriptions: taking a derivative of an averaged signal; taking an average of a differentiated signal; and convolving the signal with a single filter. The result is the same in each case.

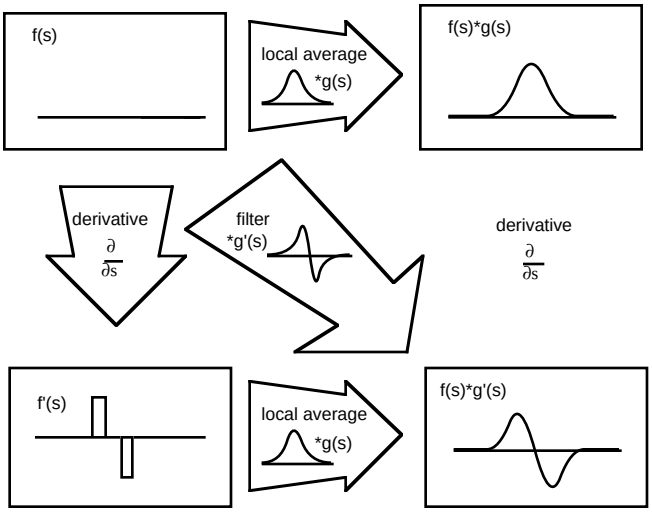


Fig. 7.7
The local average of a derivative is the same as the derivative of a local average.

We can use the standard notation of vector calculus to describe the first few directional derivatives of the plenoptic function. By convention, we assume that any dimensions not mentioned are averaged over. In other words, D_x denotes a filter that averages locally (say, with a Gaussian weighting function) in all directions, and then differentiates in the x direction. Similarly, $D_{x,}$ means differentiation in x and in y (the sequence is irrelevant). For second derivatives we write D_{xx} (for example) or $D_{\lambda.\lambda.}$. For derivatives in directions other than along coordinate axes, such as "up and to the right," we can write D_{x+y} (a diagonal derivative) or D_{x+t} (a horizontally moving one). These latter are in fact equal to $D_x + D_y$ and $D_x + D_t$, respectively, but it is useful to use the other notation sometimes to emphasize the fact that these are no more complicated than D_x or D_y, just taken in a different direction.

Examples of these operators are shown in figure 1.8. If we designate the axes shown as u and u, the operator in the upper left corner is D_v. To the right of that are D_{vv} and D_{uv}. The lower row consists of D_{u+v}, $D_{(u+v)(u+v)}$, D_{uvv} and D_{uuvv}.

The only slightly more complicated case that we need to consider is the Laplacian operator, shown in the upper right corner. This is not a directional derivative per se, but it is the sum of second directional derivatives: $L_{uv} =$

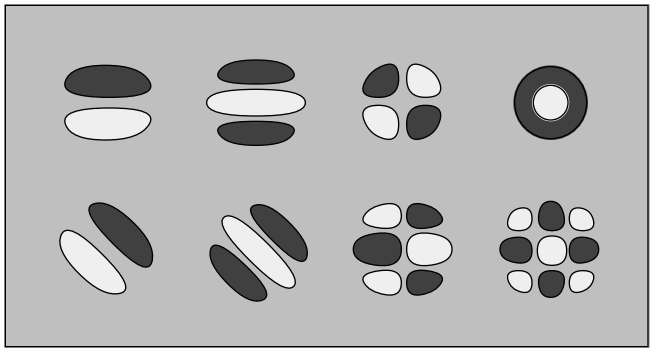


Fig. 1.8
The low order derivative operators lead to a small number to two-dimensional receptive field types.

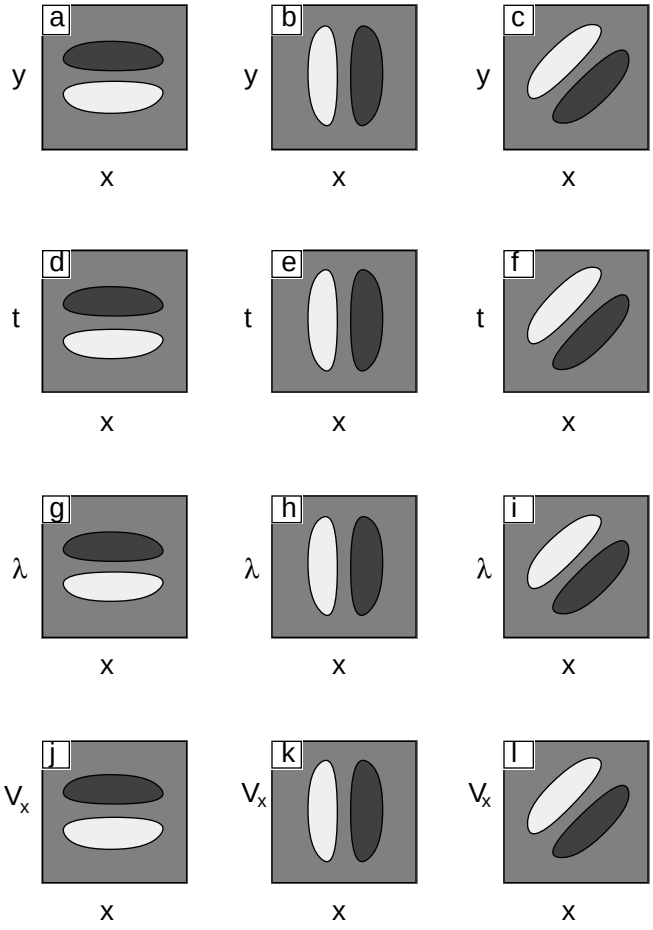


Fig. 1.9
The same receptive field structures produce different visual measurements when placed along different planes in plenoptic space.

$D_{uu} + D_{vv}$. We do not wish to suggest that center-surround structures in biological visual systems are necessarily constructed in this way; we are simply describing the formal relationships among the various types of operators.)

Visual Mechanisms for Extracting Plenoptic Structure

The visual mechanisms suggested by this approach include some familiar receptive field structures, as well as some that are more novel (cf. Young, 1989). Figure l.9 shows some examples of idealized receptive fields that one could construct to analyze change in various directions in plenoptic space—ignoring implementational constraints for the moment. These particular receptive fields represent only two dimensions of information, and one of the dimensions shown is always the spatial dimension x. All receptive fields have an implicit shape in the full set of plenoptic dimen-sions; they are assumed to be bloblike in all of the dimensions not shown.

Although these measurements do not precisely correspond to properties that are described by ordinary language, it is possible to assign approximate labels for them: (a) horizontally oriented structure (edgelike); (b) vertically oriented structure (edgelike); (c) diagonally oriented structure (edgelike); (d) full-field brightening, (e) static spatial structure; (f) moving edgelike structure; (g) full-field bluish color; (h) achromatic edgelike structure; (i) spatiochromatic variation; (j) full-field intensity change with eye position; (k) absence of horizontal parallax (edgelike structure); (1) horizontal parallax (edgelike structure).

Plenoptic Measurements in the Human Visual System

We have presented an idealized view of the basic structure available in the plenoptic function, and of the measurements that early vision could employ to characterize that structure. We now ask how these measurements, or closely related measurements, might be implemented in the human visual system. (While we refer to the "human" visual system, much of the evidence upon which our analysis is based comes from physiological studies of other mammals. We will assume without arguing the point that the early stages of processing are similar across species.)

At any given moment, a human observer has access to samples along five of the seven axes of the plenoptic function. A range of the x and y axes are captured on the surface of the retina; a range of the λ-axis is sampled by the three cone types; a range of the t-axis is captured and processed by temporal filters; and two samples from the V_x-axis are taken by the two eyes. In order to sample the V_y-axis at an instant, we would need to have a third eye, vertically displaced, and in order to sample the V_z-axis, we would need an extra eye displaced forward or backward. It is possible to accumulate information about V_y and V_z over time by moving the head—that is, using ego-motion—but the information is not available to a static observer at a given moment, and therefore we exclude any detailed consideration of this information from our discussion here.

For human observers then, the available plenoptic function involves the five parameters, x, y, t, λ, and V_x and may be parameterized:

$$P = P(x,y,t,\lambda,V_x). \tag{3}$$

Each dimension is analyzed by the human visual system with a limited resolution and a limited number of samples. Given that the visual system can only extract a finite number of samples, it is quite interesting to observe where the-sampling is dense and where it is sparse. The sampling in x and y (or visual angle, if one prefers), is by far the densest, corresponding to hundreds of thousands of distinct values. The sampling in wavelength is much cruder: The cones extract only three samples along the wavelength dimension. The sampling in horizontal viewing position is cruder still with only two samples. Time is represented continuously, but the dimensionality of the representation at any given instant is probably quite low.

The fact that different dimensions are sampled with different densities may reflect a number of factors, such as: (1) some dimensions have greater physical variance than others and thus simply contain more information (Maloney, 1986), (2) some dimensions contain information that is more important for survival than other dimensions, and (3) some dimensions are easier to sample than others, given biological constraints. It is worth noting that the eyes of certain mantis shrimps have been reported to analyze the spectrum with as many as ten distinct photoreceptor types (Cronin & Marshall, 1989); these organisms can presumably make much finer chromatic distinctions than humans can, although their spatial resolution is much poorer than that of humans. Denser sampling in one domain requires sparser sampling in another, and so the tradeoffs chosen in a given organism reflect the way in which visual information is weighted in that organism's niche. It is also significant that in the human visual system the sampling is not uniform across space, being far denser in the fovea than in the periphery. Thus the direction of gaze has a major impact on the actual information available to the observer, even though it does not affect the information potentially available in the pencil of rays entering the pupil.

We now consider the way that human (and other mammalian) visual systems analyze the various dimensions of the plenoptic function. Both the commonalities and the variations are quite interesting.

Space

The two spatial dimensions (or visual angle dimensions), which we will consider together, are analyzed in far more detail than any other dimensions. Spatial receptive fields may have more lobes than the two or three expected from the lowest order derivative operators: indeed, neurons with six or more lobes have been reported in visual cortex (De Valois, Thorell & Albrecht, 1985; Young, 1985), which would suggest derivative orders of 5 or more. In addition, the spatial analysis is performed at many positions and at many scales, and the analysis is more detailed in the fovea than in the periphery. The extensive and diverse analysis that is devoted to the spatial domain indicates that it is far more important than any other sampled dimension for primate (and presumably for human) vision.

Spatial receptive fields at the level of the retina and the lateral geniculate nucleus (LGN) tend to be circular or slightly elliptical. But the receptive fields of cells devoted to spatial analysis in primary visual cortex are almost invariably oriented, except for the cells in layer 4, which receive direct input from LGN. As far as we know, cells with "crisscross" receptive fields, such as would result from separable analysis of two spatial directions, are not found.

Time

The time dimension is unusual in two ways. First, all filtering must be causal, which is to say that the impulse

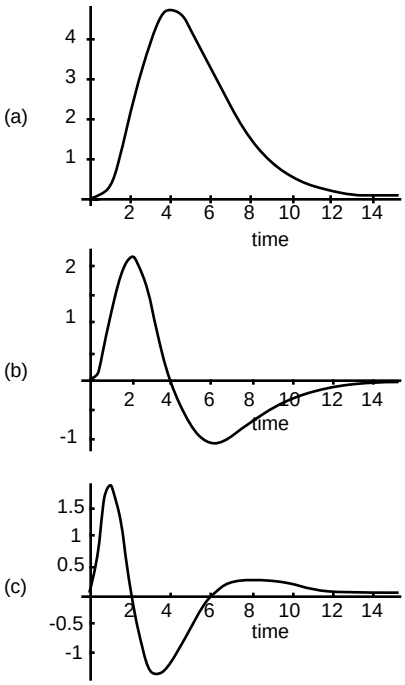


Fig.1.10
Since temporal impulse responses must be causal, they do not have the same form as the weighting functions observed along other axes. Shown here are the zeroth, first, and second derivatives of a fourth-order Butterworth filter. The amplitude spectra of these filters are similar to those of the corresponding Gaussian and Gaussian derivative filters, although the phase spectra are quite different.

response can only include events in the past and not in the future. This means that (unless unreasonable amounts of delay are introduced) the temporal weighting functions cannot have the symmetrical form found in the space domain. Rather, they tend to look more like those shown in figure 1.10, which shows impulse responses corresponding to a lowpass filter of the form t^4e^{-t}, along with its first and second derivatives.

A second peculiarity of the time dimension is that it is not discretely sampled. All other dimensions are sampled by a finite number of neurons, but since these neurons produce continuously varying outputs (neglecting the fact that they must convey this output via a spike train) there seems to be no equivalent sampling in time. The amount of information conveyed in the cell responses is, of course, limited by the stochastic character of the spike train; moreover, there is a temporal frequency bandlimit on the cell's response, and this can be used to compute an effective temporal sampling rate.

One can also ask about the dimensionality of the temporal analysis that is available at a given retinal position at a given time. If only a single cell were responding the dimensionality would be unity, but if several cells are responding simultaneously to different temporal epochs (as with lagged cells) or to different temporal frequency bands (as with flicker and motion cells), then it becomes reasonable to speak of the dimensionality of the temporal representation at an instant. This dimensionality is likely to be small, as there seem to be only two or three broadly tuned temporal channels at a given location (Graham, 1989).

It is worth noting that, from the standpoint of the plenoptic function, simple temporal change (brightening, dimming, and flicker) is at least as fundamental as is motion, which involves correlated change across space and time.

Wavelength

In humans the wavelength axis is sampled at only three points, by three cone types (neglecting the rods). The broad cone action spectra can be thought of as prefilters required by this sparse sampling that prevent high frequency aliasing in the wavelength domain (cf. Barlow, 1982). An averaging across wavelength, without any derivatives, leads to an achromatic signal, as shown in figure 1.11A. A first derivative operator in wavelength corresponds to a blue-yellow opponent signal, as shown in figure 1.11B. A second derivative operator corresponds to a red-green opponent signal, as shown in figure 1.11C. Note that this red-green signal receives positive input from both the short-wave and long-wave cones, which is consistent with many results in psychophysics and physiology. These three idealized weighting functions are qualitatively similar to the ones that are actually found experimentally, although they differ in detail.

It is also interesting to characterize common chromatic neurons with spatiochromatic receptive fields. Color opponent "blob" cells (those without spatial structure) correspond to derivatives in wavelength without any derivatives in space (D_λ); figure 1.12A illustrates an (x,λ) slice of an idealized cell. Double opponent cells correspond to derivatives in wavelength and circular (Laplacian) derivatives in space (in one dimension $D_{xx\lambda}$); a slice is shown in figure 1.12B. Single opponent cells, as in figure 1.12C, do

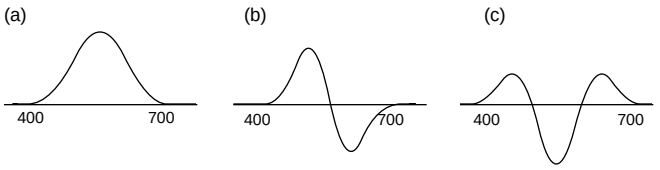


Fig.1.11
The zeroth, first, and second derivatives of a Gaussian weighting function along the wavelength axis are similar to the luminance, blue-yellow, and red-green weighting functions found in the human visual system.

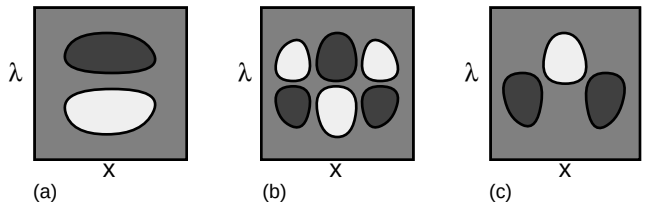


Fig.1.12
Spatiochromatic receptive fields may be used to characterize various color selective neurons: (a) an opponent color cell with no spatial structure; (b) a double opponent cell; (c) a single opponent cell. Note that this cell type does not correspond to a standard derivative type, although it may be constructed by taking sums of derivatives.

not have any simple interpretation in terms of derivatives, although one can synthesize them as sums of derivatives (in fact $D_\lambda + D_{xx} + D_{yy}$). Cells that are spatio-chromatically oriented (not illustrated here) would respond to chromatic ramps or edges.

Horizontal Parallax (Binocular)

The axis corresponding to horizontal eye position (V_x) is sampled at only two discrete positions by the two eyes. This creates potential ambiguities in the interpretation of this dimension by the visual system. The finite size of the pupil offers a very small amount of spatial prefiltering, which is entirely inadequate to prevent the aliasing introduced by such sparse sampling. Aliasing in the V_x-axis is the source of the "correspondence problem" in stereo vision (also known as the problem of ghosts). If the eyes were somehow able to prefilter properly along the V_x-axis, the correspondence prob-lem would vanish. However, this would be optically impos-sible, and even if it were possible it would greatly reduce the accuracy with which stereo disparity could be measured.

Figure 1.13 shows examples of how the (x, V_x) plane can be analyzed with simple derivative operators. We show

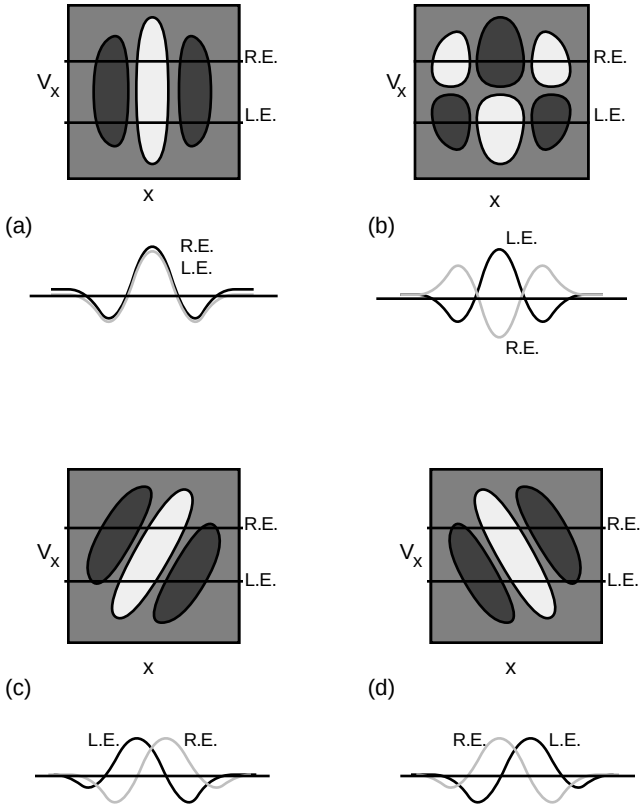


Fig.1.13
Four examples of binocular receptive fields. Humans only take two samples from the V_x axis, as shown by the two lines labeled R.E. and L.E. for right eye and left eye. The curves beneath each receptive field indicate the individual weighting functions for each eye alone. The four receptive field types shown here correspond to: (a) binocular correlation; (b) binocular anti-correlation; (c) uncrossed disparity; (d) crossed disparity. These could be used to produce cells that would be categorized as "tuned excitatory," "tuned inhibitory," "far," and "near."

the plane itself, along with the two linear samples that correspond to the images acquired by the left eye (L.E.) and right eye (R.E). For these examples we assume that the two eyes are fixating a common point at a finite distance; note that this condition differs from our earlier convention that the eye's axis is always parallel to the z-axis. When the eyes are fixating a point at an intermediate distance, then near points and far points exhibit parallax of different signs. Below each of the receptive fields is a diagram showing the individual responses of the right eye and left eye alone; these individual responses would be linearly summed to produce sampled versions of the underlying receptive fields.

Part A of figure 1.13 shows a receptive field for zero disparity, which will respond well when an object lies

in the plane of fixation. Part B shows a receptive field for binocular anticorrelation, which will respond well to an object that is giving different or opposite contrasts to the two eyes (leading to the percept known as "luster"). Part C shows a receptive field sensitive to objects beyond the plane of fixation (i.e., uncrossed disparity); part D shows a receptive field sensitive to objects before the plane of fixation (i.e., crossed disparity). Poggio and Poggio (1984) describe "tuned excitatory," "tuned inhibitory," "near," and "far", cells, which might correspond to the receptive field types shown here.

Parallax in Y_x, Y_y, and V_z

A free observer can gather information about parallax along arbitrary directions in space by actively translating the eye through a range of viewing positions. The information is known as motion parallax, and it includes horizontal and vertical parallax for the V_x and V_y direction as well as the looming effects of the V_z direction. To produce a receptive field for motion parallax, one must combine information about the observer's motion with information about the changing retinal image; such mechanisms are more complex than the ones we have chosen to discuss under the heading of early vision, but in principle they could be analyzed in similar terms.

Periodic Tables for Early Vision

We have considered the structure of information in the luminous environment and how this information appears to be sampled in mammalian visual systems. We are now in a position to list systematically the possible elements of early vision, i.e. to build our periodic table of visual elements. Figure 1.14 shows a set of measurements that can be made by taking derivatives along single axes of the plenoptic function (D_u and D_{uu}). Since there are five axes, there are five entries for each. For the first derivative we find measure-ments as follows: x-axis: vertical "edge"; y-axis: "horizontal "edge"; t-axis: flicker (brightening); λ-axis blue-yellow opponency; V_x-axis: binocular anticorrelation (luster). For the second derivative, which we have shown with on-center polarity, we find: x-axis: vertical "bar"; y-axis: horizontal "bar", t-axis: flicker (pulse); λ-axis: green-red opponency. There is no meaningful measurement for the V_x-axis given only the samples from two eyes.

The range of possible measurements becomes richer when we allow variation in two dimensions instead of just one. A given receptive field type now may occur with any pair of different axes. Thus we arrive at a 5 by 5 matrix of possible measurements for a given receptive field type. An example is shown in figure 1.15. The receptive field consists of a diagonal second derivative ($D_{(u+v)(u+v)}$). The possible axes are listed along the sides of the matrix; each pair of axes leads to a preferred stimulus, which is listed in the matrix. The diagonal entries are meaningless, and the upper triangle is identical to the lower triangle in this particular case by symmetry. Thus only the lower triangle is filled.

Each entry has four descriptors, corresponding to the spatial, temporal, chromatic, and binocular aspects of the preferred stimulus. The default values for the descriptors are: full-field (no spatial variation); static (no temporal variation); achromatic (no wavelength variation), and no binocular disparity. For each entry, the "interesting values" are shown in boldface, while the default values remain in plain text.

Some of the entries correspond to well-known cell types that respond to common experimental stimuli, such as tilted or moving bars. Both horizontal and vertical disparity also appear in the matrix. Some of the chromatic measurements are unexpected wavelength variation correlated with space, time, and eye position lead to measurements of what we call spatial and temporal hue-sweeps, and to chromatic luster. Although these chromatic properties may sound obscure, they correspond to properties that an observer can encounter in the world. For example, a spatial hue-sweep can occur when an object's color varies across space; a temporal hue-sweep can occur when an object changes color over time; and a binocular hue difference can occur with certain kinds of lustrous objects, especially iridescent ones. Interestingly, Livingstone and Hubel (1984) reported a cell in the striate cortex of a monkey that was chromatically double opponent in both eyes but with opposite sign; that is, in the right eye it was R+G- center, R-G+ surround, while in the left eye it was R-G+ center, R+G- surround.

Another example is shown in figure 1.16. In this case, the measurement consists of a pair of single derivatives taken separably along two dimensions (D_{uv}). Once again the upper and lower triangles are identical symmetry, and the diagonal entries are not meaningful. This leaves us with the ten measurements listed in the matrix. Some

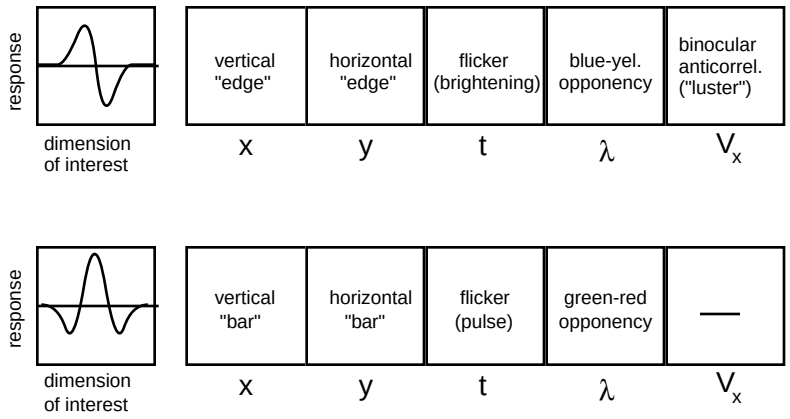
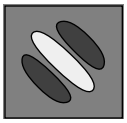
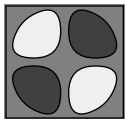


Fig.1.14
Derivatives along single dimensions lead to a number of basic visual measurements.



	x	y	t	λ	V_x
x	—				
y	**diag. "bar"** static achromatic no dispar	—			
t	**vert. "bar" leftward** achromatic no dispar	**hor. "bar" downward** achromatic no dispar	—		
λ	**vertical** static **hue-sweep** no dispar	**horiz.** static **hue-sweep** no dispar	full-field **sequential hue-sweep** no dispar	—	
V_x	**vert. "bar"** static achromatic **hor. dispar**	**hor. "bar"** static achromatic **vert. dispar**	full-field **sequential** achromatic **eye-order**	full-field static **hue-shift luster**	—

Fig.1.15
The ten entries in the section of the periodic table corresponding to a tilted second derivative. The entries on the diagonal are meaningless, and the entries in the upper triangle are the same as those in the lower triangle by symmetry.



	x	y	t	λ	V_x
x	—				
y	**criss-cross** static achromatic no dispar	—			
t	**vert."edge" reversing** achromatic no dispar.	**hor "edge" reversing** achromatic no dispar.	—		
λ	**vert."edge"** static **blue-yel.** no dispar.	**hor. "edge"** static **blue-yel.** no dispar.	full-field **reversing blue-yel.** no dispar.	—	
V_x	**vert."edge"** static achromatic **anticorr.**	**hor. "edge"** static achromatic **anticorr.**	full-field **reversing** achromatic **anticorr.**	full-field static **blue-yel. anticorr.**	—

Fig.1.16
The entries in the periodic table corresponding to separable first derivatives along both axes. There are ten elementary measurements of this type

of the measurements are familiar, such as reversing (i.e., counterphase flickering) edges. The spatial "crisscross" measurement would seem to be straightforward enough, but cells showing this receptive field are rarely if ever found in cortex. Chromatic flicker and binocular luster (anticorrelation) are among the other measurements appearing in the table.

For each receptive field type one can construct a similar table. For a given shape of two-dimensional receptive field, there correspond twenty basic measurements, or ten when there is symmetry about the diagonal. Consider the receptive fields (RF) shown in figure 1.8. The seven oriented and separable RF types will lead to a total of some 100 distinct elemental measurements. (The exact number depends on how one chooses to treat symmetries.)

These 100 measurements can be considered to be the first stage of a periodic table of visual elements. Note that these measurements only include derivatives of order one and two, and that they do not include measurements involving derivatives along three or more dimensions of the plenoptic function. Thus, for example, a moving chromatic bar would not be properly analyzed at this level.

Several points emerge from this exercise. First, the plenoptic function is quite rich in information: At every point within it there are some 100 distinct and independent local measurements that can be used to characterize is structure, even when the measurements are restricted to the simplest sort. Second, it would be an enormous burden for the human visual system to sample and represent all of these measurements at high density, and so it is necessary that only the more important ones be analyzed in detail. Third, as one allows higher derivatives and more plenoptic dimensions, the number of potential elements grows through a combinatorial explosion.

Psychophysical experiments can be performed to determine human sensitivity to change along various axes in plenoptic space. Indeed, many psychophysical tasks can be shown to have unexpected relationships through such an analysis. Some interesting examples are discussed in the appendix.

Further Computations

The linear derivative-like computations we have discussed offer a starting point for early vision. Later stages in vision can make use of these initial measurements by combining them in various ways.

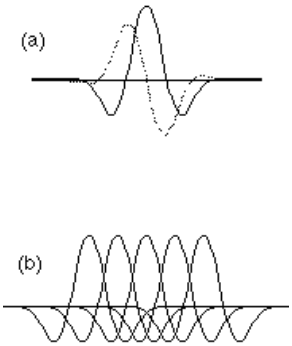


Fig. 1.17
Energy computation: (A) a pair of receptive fields that are in quadrature phase can be used to compute a local energy measure; (B) complex cells in striate cortex may compute local energy measures by combining the rectified outputs of several adjacent receptive fields.

In some cases one may wish to measure the magnitude of local change within the plenoptic function in a region of the visual field, without specifying the exact location or spatial structure within that region. One may wish to know, for example, that there exists an oriented contour without specifying whether it is an edge, a dark line, or a light line. In this case, a "local energy measure" can be computed that pools similar derivative-like signals within a spatial region. The term "local energy" is motivated by the fact that this pooling can be accomplished by summing the squared outputs of two linear receptive fields differing in phase by 90°. This arrangements is shown in figure 1.17A. This nonlinear combination gives a single-peaked positive response to edges and lines, regardless of sign (Adelson & Bergen, 1985; Granlund, 1978; Knutsson & Granlund, 1983; Ross, Morrone & Burr, 1989); the response extends smoothly over a patch centered on the feature of interest. Similar properties are obtained in general if an array of related linear subunits, as shown in figure 1.17B, are passed through a rectifying nonlinearity and then summed over some region. Complex cells in striate cortex seem to be performing a computation similar to this (Movshon, Thompson & Tolhurst, 1978; Spitzer & Hochstein, 1985). In addition, motion-selective complex cells may compute spatiotemporal energy measures by pooling the outputs of receptive fields that are oriented in space-time (Adelson & Bergen, 1985; Emerson,

Bergen & Adelson, 1987, 1991). By performing similar operations along other axes in the plenoptic function one can compute local energy measures for binocular disparity (Jacobson & Gaska, 1990; Ohzawa, DeAngelis & Freeman, 1990; cf. Sanger, 1988), flicker, chromatic saturation, and so on.

Energy measures can also be combined and cascaded to form more complex measurements (Granlund, 1978). Such an approach may be used to form cells that are end-stopped or side-stopped, or which are selective for more complex properties such as curvature. There is abundant evidence for this sort of complexity in striate cortex and beyond, but a discussion is outside our present scope (cf. Dobbins, Zucker & Cynader, 1987; Koenderink & van Doorn, 1987).

Conclusion

The world is filled with light, and the structure of this light is determined by the physical arrangement of the materials that fill the world. An observer can learn about the world by making measurements of the structure of the light passing through a given point.

We introduce the plenoptic function to specify formally the way that light is structured. For a given wavelength, a given time, and a given viewing position in space, there exists a pencil of light rays passing through the viewing point. Each ray has an intensity, and the collection of rays constitutes a panoramic image. This panoramic image will vary with time, viewing position, and wavelength. The plenoptic function thus has seven dimensions and may be parameterized as $P(x,y,t,\lambda, V_x,V_y,V_z)$.

One of the tasks of early vision is to extract a compact and useful description of the plenoptic function's local properties; the low order derivatives offer such a description under reasonable assumptions about world statistics. If one takes locally weighted first and second derivatives along various axes in plenoptic space, a full range of elemental visual measurements emerges, including all of the familiar measurements of early vision.

The actual information extracted by the human visual system is a small subset of the information that is physically available in the plenoptic function. Humans gather information from only two viewing positions at a time, obtaining two samples along the V_x axis, both taken at a single value of V_y and V_z. Humans sample the wavelength axis with only three cone types. The most densely sampled axes are those corresponding to visual angle, namely, the spatial axes of the retinal image. Time is the only axis that is represented continuously.

One can develop a taxonomy of derivative types, and for each type one can construct a table of visual measurements corresponding to different choices of axes. In this way one can construct a kind of periodic table of the visual elements. A basic table (constructed with the simple derivative types applied to one or two axes at a time) contains 100 elementary measurements.

The periodic table is constructed from first principles, using only the constraints imposed by the available optical information. It is not dependent on the statistics of the world or on the needs of any particular organism. Thus the periodic table offers a null hypothesis about the measurements that an organism could make, with no further assumptions about the environment or the organism's place in it. It then becomes interesting to compare the list of potential measurements with the list of actual measurements that a given organism makes.

The entries in the periodic table include all of the basic visual properties that are commonly considered to constitute early vision, such as orientation, color, motion, and binocular disparity. Other less popular properties also appear in the list, including flicker, binocular correlation and anticorrelation, and hue-shift in space and time. Finally, there are some unexpected properties, such as binocular chromatic anticorrelation.

Many psychophysical and physiological experiments can be considered as explorations of sensitivity to the elemental measurements listed in the periodic table. Over the years, experimenters have filled in the entries of the table in a somewhat random fashion; but it is possible to approach the problem systematically. Physiologists can look for neurons that fall into the various elemental categories, and psychophysicists can look for channels selective for the various elemental properties.

Appendix

Psychophysical Experiments in Plenoptic Space

A variety of paradigms have been developed over the years to assess the limits of human visual performance.

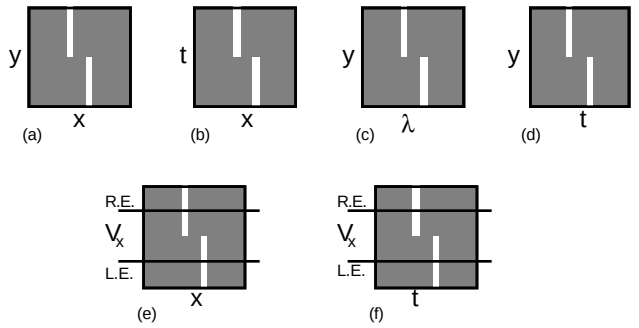


Fig. 1.18
A variety of psychophysical stimuli that seem unrelated turn out to have the same plenoptic structure. The basic spatial vernier task (a) transforms into: (b) temporal displacement discrimination; (c) spatial (top-bottom bipartite field) wavelength discrimination; (d) biparite simultaneity or temporal order judgment; (e) stereoacuity; (f) dichoptic simultaneity or eye-order judgment.

When these experiments are depicted in plenoptic space, a number of unexpected parallels emerge. As an example, the stimuli shown in figure 1.18 correspond to a variety of tasks involving spatial acuity, motion thresholds, wavelength discrimination, and so on. The tasks seem diverse but their structure in plenoptic space is quite similar, as we will now describe.

Part A of figure 1.18 shows the case in which the axes are (x,y); this is just the classic vernier offset task, in which one vertical line is displaced with respect to another. Part B shows the (x, t) case. This corresponds to a vertical line undergoing a jump in position; in other words, it measures sensitivity to small sudden displacements. Part C shows the (λ,y) case. This is a bipartite field, split into top and bottom halves, in which the top half has a different wavelength than the bottom half. In other words, this is a wavelength discrimination task. Part D shows the (t,y) case. This is a bipartite field in which the top half is briefly pulsed and then the bottom half is briefly pulsed. The task is to determine the temporal order in which the two half-fields are pulsed.

Parts E and F of figure 1.18 involve the eye position axis, V_x. Since humans have only two eyes, they take only two samples from this axis, and the locations of these samples are indicated by the horizontal lines marked R.E. and L.E. Part E depicts a stimulus of two vertical lines, one in each eye, with a horizontal displacement—in other words, a test for the detection of stereoscopic disparity. Part F depicts a stimulus in which a full-field flash is presented first to one eye and then to the other.

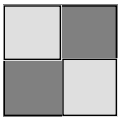
Figure 1.19 enumerates all the variants of such a stimulus, using pairwise combinations of the five dimensions of x, y, t, λ, and V_x. Every stimulus makes sense, although some of them seem a bit odd at first. For example, the (λ,V_x) stimulus represents a static field of one wavelength in the left eye and a different wavelength in the other eye; the task is to detect the wavelength difference.

Other stimulus classes are shown in figures 1.20 and 1.21. Figure 1.20 shows a stimulus that looks like a piece of a checkerboard in (x,y) but which takes on various sorts of chromaticity, temporal steps, and binocular anti-correlation when the different axes are used. The cells in the upper right triangle are left blank, as they are identical to the lower left triangle by symmetry. Figure 1.21 shows a stimulus that is a diagonal bar in the (x,y) plane, but can become a rainbow, a moving bar, or other stimuli in other planes.

The stimuli we have depicted in the tables above can be

	x	y	t	λ	V_x
x	—	**hor. vernier** static achromatic no dispar.	**vert. bipart. pulse order** achromatic no dispar.	**vert. edge** static λ **diff.** no dispar.	—
y	**vert. vernier** static achromatic no dispar.	—	**hor. bipart. pulse order** achromatic no dispar.	**hor. edge** static λ **diff.** no dispar.	—
t	**vert. line jump left** achromatic no dispar.	**horiz. line jump down** achromatic no dispar.	—	full-field **sequential** λ **change** no dispar.	—
λ	**2 vert. lines** static λ **change** no dispar.	**2 hor. lines** static λ **change** no dispar.	full-field **pulse order blue-yel.** no dispar.	—	—
V_x	**vert. line** static achromatic **h disp.**	**horiz. line** static achromatic **v disp.**	full-field **pulse order** achromatic **anticorr.**	full-field static λ **diff. anticorr.**	—

Fig. 1.19
A table of the psychophysical stimuli corresponding to the basic vernier structure.



	x	y	t	λ	V_x
x	—				
y	**checker** static achromatic no dispar	—			
t	**vert. edge reversing** achromatic no dispar.	**horiz. edge reversing** achromatic no dispar.	—		
λ	**vert. edge** static **blue-yel.** no dispar.	**horiz. edge** static **blue-yel.** no dispar.	full-field **exchange blue-yel.** no dispar.	—	
V_x	**vert. edge** static achromatic **anticorr.**	**horiz. edge** static achromatic **anticorr.**	full-field **exchange** achromatic **anticorr.**	full-field static **blue-yel. anticorr.**	—

Fig.1.20
A table of the psychophysical stimuli corresponding to the checkerboard structure.

	x	y	t	λ	V_x
x	—				
y	**diag. bar** static achromatic no dispar	—			
t	**vert. bar leftward** achromatic no dispar	**horiz. bar downward** achromatic no dispar	—		
λ	**vertical** static **rainbow** no dispar	**horiz.** static **rainbow** no dispar	full-field **sequential rainbow** no dispar	—	
V_x	**vert. bar** static achromatic **hor. dispar**	**horiz. bar** static achromatic **vert. dispar**	full-field **sequential** achromatic **eye-order**	full-field static **hue-shift luster**	—

Fig.1.21
A table of psychophysical stimuli corresponding to a tilted bar structure.

described with only two axes, but many psychophysical stimuli require more. For example, a diagonal chromatic moving bar in depth would require a full five-dimensional description.

References

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2,* 284-299.

Barlow, H. B. (1982). What causes trichromacy? a theoretical analysis using comb-filtered spectra. *Vision Research, 22,* 635-643.

Beck, J. (1966). Perceptual grouping produced by changes in orientation and shape. *Science, 154,* 538-540.

Bergen, J. R. & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man and Cybernetics, 13,* 857-863.

Cronin, T. W. & Marshall, N. J. (1989). A retina with at least ten spectral types of photoreceptors in a mantis shrimp. *Nature, 339,* 137-140.

De Valois, R. L. & De Valois, K. K. (1988). *Spatial Vision*. New York: Oxford University Press.

De Valois, R. L., Thorell, L. G., & Albrecht, D. G. (1985). Periodicity of striate-cortex-cell receptive fields. *Journal of the Optical Society of America A, 2,* l115-1123.

Dobbins, A., Zucker, S. W., & Cynader, M. (1987). Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature, 329,* 438-441.

Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1987) Movement models and directionally selective neurons in the cat's visual cortex. *Society for Neuroscience Abstracts, 13,* 1623.

Emerson, R. C., Bergen, J. R, & Adelson, E. H. (1991). Directionally selective complex cells and the calculation of motion energy in cat visual cortex. *Vision Research,* in press

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Graham, N. V. S. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Granlund, G. (1978). In search of a general picture processing operator. *Computer Graphics and Image Processing,* 8, 155-173.

Horn, B. K. P. (I986). *Robot vision*. Cambridge, MA: The MIT Press.

Hubel, D. H. (1988). *Eye, brain and vision*. San Francisco: Freeman.

Jacobson, L. D. & Gaska, J. P. (1990). A s/sf energy model for estimating binocular disparity. *Investigative Ophthalmology and Visual Science (Suppl.), 31,* 91.

Julesz, B. & Bergen, J. R. (l983). Textons, the fundamental elements in preattentive vision and perception of textures. *Bell Systems Technical Journal, 62,*1619-1645.

Kemp, M. (1989). *Leonardo on painting*. New Haven: Yale University Press.

Knutsson, H. & Granlund, G. H. (1983). Texture analysis using two dimensional quadrature filters. In *IEEE CAPAIDM*, p. 206-213, Silver Spring, MD.

Koenderink, J. J. & van Doorn, A. J. (1987). Representation of local geometry in the visual system. *Biological Cybernetics, 55,* 367-375.

Levine, M. D. (1985). *Vision in man and machine*. New York: McGraw-Hill.

Livingstone, M. S. & Hubel, D. H. (I984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience, 4,* 309-356.

Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *Journal of the Optical Society of America A, 3,* 1673-1683.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology, 283,* 79-99.

Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.

Ohzawa, I., DeAngelis, G., & Freeman, R. (1990). Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science, 249,* 1037-1041.

Olzak, L. & Thomas, J. P. (l986). Seeing spatial patterns. In K. R. Boff, L. Kaufman, & J. P. Thomas. (Eds.), *Handbook of perception and human performance:* Vol. (pp. 7-1-7-56). New York: Wiley.

Poggio, G. F. & Poggio, T. (1984). The analysis of stereopsis. *Annual Review of Neuroscience, 7,* 397-412.

Pokorny, J. & Smith, V. C. (1986). Colorimetry and color discrimination. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance*: Vol. 1 (pp. 8-1-8-51). New York: Wiley.

Richter, J. P. (1970). *The notebooks of Leonardo da Vinci*: Vol. 1. New York: Dover.

Ross, J., Morrone, M. C., & Burr, D. C. (1989). The conditions under which mach bands are visible. *Vision Research. 29,* 699-715.

Sanger, T. D. (1988). Stereo disparity computation using Gabor filters. *Biological Cybernetics, 59,* 405-418.

Spitzer, H. & Hochstein, S. (1985). A complex cell receptive field model. *Journal of Neurophysiology, 53,* 1266-1286.

Treisman, A. (1986). Properties, parts and objects. In K. R. Boff, L. Kaufman, & J. P. Thomas, (Eds.), *Handbook of perception and human performance:* Vol. 2 (pp. 35-1-35-70). New York: Wiley.

Treisman, A. and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-137.

Watson, A. B. (1986). Temporal sensitivity. In K. R. Boff, L. Kaufman, & J. P. Thomas, (Eds.), *Handbook of perception and human performance*: Vol. 1 (pp. 6-l-6-43). New York: Wiley.

Young, R. A. (1985). *The gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line-weighting profiles* (Tech. Rep. GMR-4920). General Motors Research Publication.

Young, R. A. (1989). Quantitative tests of visual cortical receptive field models. In *OSA Annual Meeting, 1989, Technical Digest,* p. 93, Washington, DC.