MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 761 March, 1984

COMPUTATIONS UNDERLYING THE MEASUREMENT OF VISUAL MOTION

Ellen C. Hildreth

ABSTRACT: The organization of movement in a changing image provides a valuable source of information for analyzing the environment in terms of objects, their motion in space, and their three-dimensional structure. This movement may be represented by a two-dimensional velocity field that assigns a direction and magnitude of velocity to elements in the image. This paper presents a method for computing the velocity field, with three main components. First, initial measurements of motion in the image take place at the location of significant intensity changes, which give rise to zero-crossings in the output of the convolution of the image with a $\nabla^2 G$ operator. The initial motion measurements provide the component of velocity in the direction perpendicular to the local orientation of the zero-crossing contours. Second, these initial measurements are integrated along contours to compute the two-dimensional velocity field. Third, an additional constraint of smoothness of the velocity field, based on the physical constraint that surfaces are generally smooth, allows the computation of a unique velocity field. The details of an algorithm are presented, with results of the algorithm applied to artificial and natural image sequences.

© Massachusetts Institute of Technology, 1984

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for this research was provided by the Air Force Office of Scientific Research under contract F49620-83-C-0135.

1. INTRODUCTION

The organization of movement in a changing two-dimensional image provides a valuable source of information for analyzing the environment in terms of objects, their motion in space, and their three-dimensional structure. It is not surprizing, therefore, that the analysis of motion plays a central role in biological vision systems, and in recent years, has played an increasingly important role in computer vision systems as well. For biological systems, the analysis of movement is crucial for such basic tasks as detecting and tracking prey, responding quickly to an approaching predator, and guiding locomotion through a complex environment. Perhaps the most remarkable use of visual motion is the recovery of three-dimensional structure from the changing two-dimensional projection that a moving surface casts onto the eye. The ability of the human visual system to perform this recovery was first demonstrated in the studies of Wallach and O'Connell (1953) and Johansson (1973, 1975). For animals without binocular vision, motion is a primary cue to the three-dimensional structure of the world around them.

The extensive use of motion by biological systems, and in particular the human visual system, demonstrates the feasibility of carrying out certain information processing tasks and helps to establish specific goals for the analysis of movement in time-varying imagery. This analysis divides naturally into two parts. The first is the measurement of motion; for example, the assignment of direction and magnitude of velocity to elements in the image, on the basis of the changing intensity pattern. The second is the use of motion measurements; for example, to separate the scene into distinct objects, and infer their three-dimensional structure.

This paper presents a computational study of the measurement of visual motion. It is a problem that was found to be surprizingly difficult, both in computer vision, and in modelling biological vision systems. The main theoretical problems posed by this measurement are introduced in Section 2. Section 3 presents a particular formulation of the input and output representations for the underlying computations. A theoretical analysis of the computation is then presented in Section 4, followed by the discussion of a specific algorithm in Section 5. Examples of the results of the algorithm suggest that the proposed computation is feasible for computer vision systems, and also yields solutions that are consistent with human motion perception.

2. WHY IS THE MEASUREMENT OF MOTION DIFFICULT?

The changing image may be represented by a two-dimensional array of time-dependent light intensities, $I(x, y, t)$. Motion in the image may be described by a two-dimensional vector field $\mathbf{V}(x, y, t)$ that specifies the direction and magnitude of velocity at points with coordinates (x, y) at

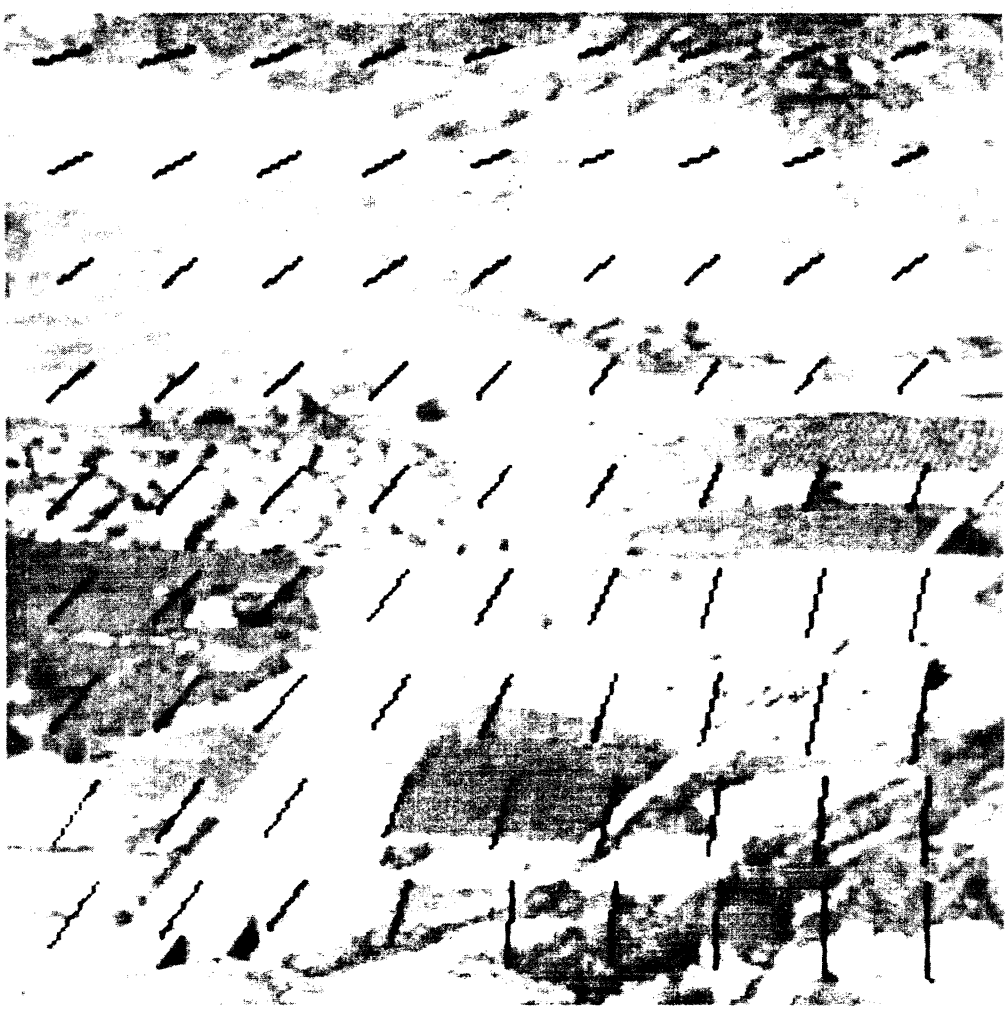


Figure 1. A velocity field. The black lines represent local velocity vectors, displayed at evenly spaced points over the original image.

time t. The measurement of visual motion may then be formulated as the computation of $\mathbf{V}(x, y, t)$ from $I(x, y, t)$. Figure 1 illustrates a velocity field that was computed from a sequence of aerial photographs, using the algorithm that is developed in this paper. In the background, a single image is shown with reduced contrast. The superimposed black lines represent local velocity vectors that describe the movement of features in the image, as the plane flies over the scene. For some visual tasks, it may be sufficient to compute only certain properties of the velocity field; for example, to respond quickly to a moving object, motion must be detected, but not necessarily measured. Other tasks, such as the recovery of three-dimensional structure, require a more complete and accurate

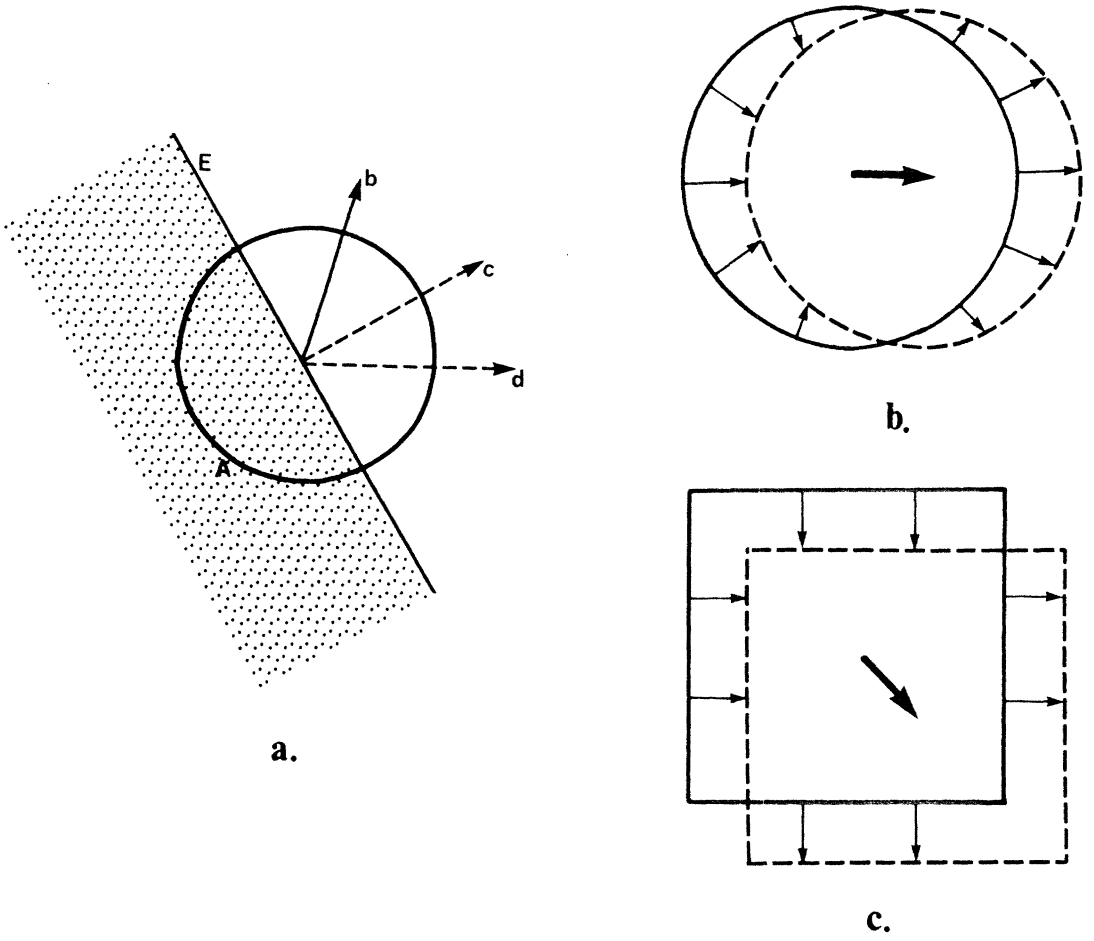


Figure 2. The aperture problem. (a) An operation that views the moving edge E through the local aperture A can compute only the component of motion **c** in the direction perpendicular to the edge. (b) and (c) The perpendicular components of velocity for a translating circle and square.

computation of the velocity field, as illustrated in Figure 1 (Ullman 1979a, 1983a, b; Clocksin 1980; Prazdny 1980; Longuet-Higgins 1981; Longuet-Higgins and Prazdny 1981).

The measurement of motion poses significant theoretical problems for a computational study. First, local motion measurements, obtained directly from the changing image, in general only provide one component of local velocity. This is a consequence of the *aperture problem*, illustrated in Figure 2(a) (Wallach 1976; Fennema and Thompson 1979; Horn and Schunck 1981; Marr and Ullman 1981; Adelson and Movshon 1982). If the motion of the edge E is to be measured by a local motion detector that examines only an area A that is small compared to the overall extent of the edge, the only motion that can be extracted is the component **c** in the direction perpendicular to the local orientation of the edge. A local detector cannot distinguish between motions in the directions **b**, **c** and **d** in Figure 2(a). In Figures 2(b) and 2(c), a circle and square undergo pure

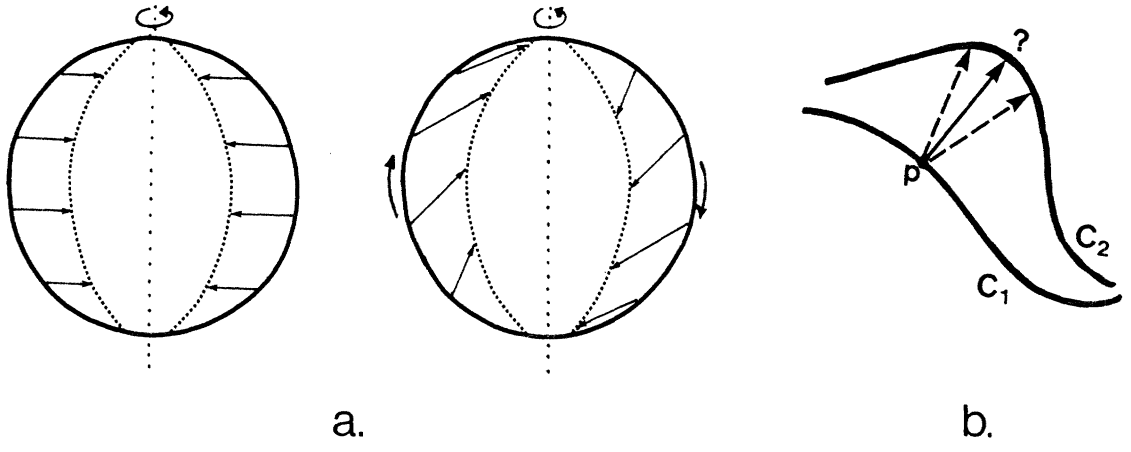


Figure 3. Ambiguity of the velocity field. (a) The arrows represent two possible velocity fields that are consistent with the changing image. (b) The curve C_1 rotates, translates and deforms over time to yield the curve C_2. The velocity of the point p is ambiguous.

translation in the directions given by the vectors at the center of the figures. The vectors along the contours represent the local perpendicular components of velocity that can be obtained directly from the changing image. To compute the true motion of the figure, a second stage of analysis is required, that combines these local measurements.

This combination stage faces a deeper theoretical problem, however; the movement of elements in the image is not determined uniquely by the pattern of changing intensities. Thus, the true velocity field is not determined uniquely from the initial local motion measurements. Two factors contribute to this ambiguity of motion. The first is the loss of information due to the projection of the three-dimensional world onto a two-dimensional image; multiple surfaces, undergoing different motions in space, may project to the same two-dimensional image. The second factor is the loss of information due to the projection into a pattern of changing intensity. The image that a surface projects onto the eye may not be sufficient to determine its movement in space. As an extreme example, a matte white sphere, rotating about a central axis, cannot be determined as such, on the basis of its projected image.

Figure 3 presents two simple examples that illustrate the ambiguity of the velocity field. In Figure 3(a), the solid and dotted lines represent the image of a moving circle, at different instants of time. In the first frame (solid line), the circle lies parallel to the image plane, while in the second frame (dotted line), the circle is slanted in depth. One velocity field that is consistent with the two frames is derived from pure rotation of the circle about the central vertical axis, as shown to the left in Figure 3(a). (The arrows represent a sample of the local velocity vectors along the

circle.) There could also be a component of rotation in the plane of the circle, about its center, as shown to the right in Figure 3(a). In addition, this changing image might represent the projection of a different three-dimensional curve that is deforming over time, giving rise to yet another projected velocity field. This ambiguity is not peculiar to symmetric figures such as circles; it is a fundamental problem that is always present. In Figure 3(b), the curve C_1 rotates, translates and deforms over time, to yield the curve C_2. The motion of points from C_1 to C_2 is again ambiguous (consider, for example, different possible velocities for the point p). In general, there are infinitely many two-dimensional velocity fields that are consistent with the changing image.

To compute motion uniquely, additional constraint is therefore required, in the form of basic assumptions about the physical world that generally hold true. The main focus of this paper is the derivation of a particular constraint, the *smoothness constraint*, which relies on the physical assumption that surfaces are generally smooth, compared with their distance from the viewer. A smooth surface usually generates a smoothly varying velocity field when it moves. Thus, intuitively, we seek a velocity field that is consistent with the changing image, and varies smoothly. Unfortunately, there are still infinitely many possibilities. A single solution may be obtained, however, by finding the velocity field that exhibits the least amount of variation. In Section 4.3, this constraint is formulated more precisely, in a way that yields a velocity field solution that is mathematically unique and physically plausible.

3. THE INPUT AND OUTPUT REPRESENTATIONS

The measurement of motion has been formulated as the computation of an instantaneous two-dimensional velocity field; such a formulation requires that motion in the image be roughly continuous. There are alternate representations of visual motion that are not so restricted. For example, motion can be described by an explicit *correspondence* over time, between elements in the image that represent the same physical feature under motion (Ullman 1979a). Motion measurement in this case requires locating identifiable elements in the changing image, and matching them over time. The input for a correspondence scheme may consist of a set of discrete frames, with large spatial separations between corresponding elements. The perception of motion by the human visual system also does not require that a stimulus move continuously across the visual field. With appropriate spatial and temporal presentation parameters, a stimulus presented sequentially can produce the impression of smooth, uninterrupted motion, as in motion pictures (Wertheimer 1912). Why, then, have we chosen a formulation of the motion measurement problem that relies on roughly continuous motion?

Recent psychophysical investigation has suggested that in the human system, motion may

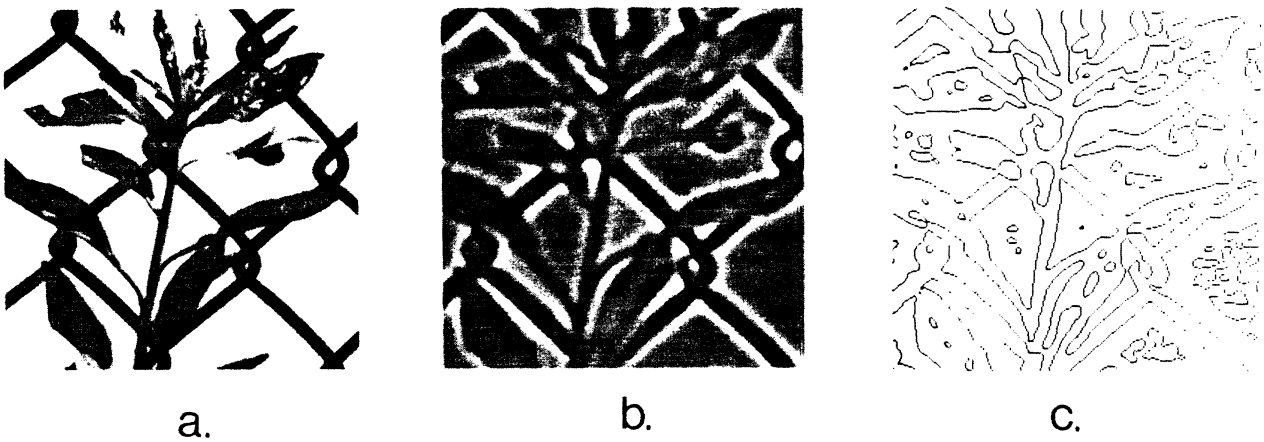


Figure 4. Initial processing of an image. (a) The original image, containing 320×320 picture elements. (b) The convolution of the image with a $\nabla^2 G$ operator. (c) The resulting zero-crossing contours.

be analyzed by two different systems, termed short range and long range processes by Braddick (1974, 1980). It has been proposed that the short range process analyzes continuous motion, or motion that is presented discretely, but with spatial displacements at most $10' - 15'$ of visual arc, and temporal intervals up to $60 - 100$ milliseconds. The long range process would then analyze motion over larger spatial and temporal intervals. The ability of the human visual system to infer motion from the discrete displacement of image elements, over considerable distances and temporal intervals, suggests an underlying correspondence computation. Under these conditions, there is no continuous motion of elements across the image to be measured directly. Short range motion is roughly continuous, however; we propose that the measurement of motion by the short range process may be appropriately formulated as the computation of an instantaneous velocity field.

With regard to the input representation for the velocity field computation, in order to detect movement in a changing image, there must be a variation of intensity over space and time; the combination of the two variations can be used to measure the direction and magnitude of velocity. The explicit comparison of spatial and temporal derivatives of intensity forms the basis of a class of motion measurement schemes referred to as *gradient* schemes (Limb and Murphy 1975; Cafforio and Rocca 1976, Fukinuki *et al.* 1976; Fennema and Thompson 1979; Horn and Schunck 1981; Marr and Ullman 1981). In principle, motion measurements may be obtained wherever intensity varies. Marr and Ullman (1981) proposed, however, that initial motion measurements in the human system are made only at the locations of significant intensity changes. To detect these intensity changes, Marr and Hildreth (1980) proposed that a powerful operator for the initial

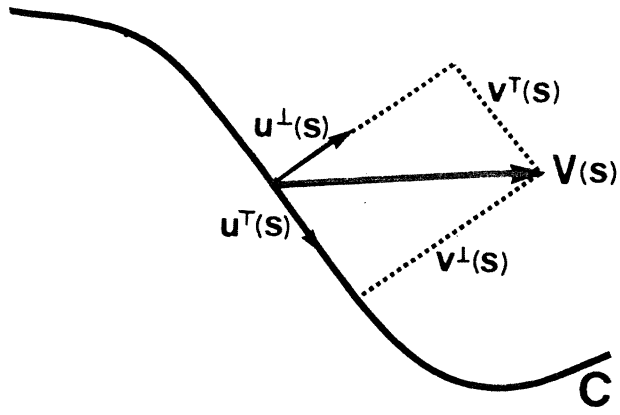


Figure 5. Decomposition of velocity. $\mathbf{u}^{\perp}(s)$ and $\mathbf{u}^{\top}(s)$ are unit direction vectors, perpendicular and tangent to the curve, and $v^{\perp}(s)$ and $v^{\top}(s)$ are the two velocity components.

filtering of an image is the Laplacian of a Gaussian, $\nabla^2 G$ (approximated in shape by the difference of two Gaussian functions). The elements in the output of an image convolved with $\nabla^2 G$, which correspond to the locations of intensity changes, are the zero-crossings (Marr and Poggio 1979). Figure 4 shows an image that has been convolved with a $\nabla^2 G$ operator, and the resulting zero-crossing contours. Marr and Ullman (1981) proposed that initial motion measurements take place at the locations of these zero-crossings, using a mechanism that combines spatial and temporal gradients of the filtered image. Further motivation for this input representation can be found in Hildreth (1984a).

In two dimensions, the initial measurements face the aperture problem. For the case of contours, local motion measurements provide only the component of motion in the direction perpendicular to the orientation of the contour. The component of velocity along the contour remains undetected. More formally, the two-dimensional velocity field along a contour may be described by the vector function $\mathbf{V}(s)$, where s denotes arclength. $\mathbf{V}(s)$ can be decomposed into components tangent and perpendicular to the contour, as illustrated in Figure 5. $\mathbf{u}^{\top}(s)$ and $\mathbf{u}^{\perp}(s)$ are unit vectors in the directions tangent and perpendicular to the curve, and $v^{\top}(s)$ and $v^{\perp}(s)$ denote the two components:

$$\mathbf{V}(s) = v^{\top}(s)\mathbf{u}^{\top}(s) + v^{\perp}(s)\mathbf{u}^{\perp}(s). \tag{1}$$

The component $v^{\perp}(s)$, and direction vectors $\mathbf{u}^{\top}(s)$ and $\mathbf{u}^{\perp}(s)$, are given directly by the initial measurements from the changing image. The component $v^{\top}(s)$ is not, and must be recovered,

to compute $V(s)$. Intuitively, the set of measurements given by $v^{\perp}(s)$ over an extended contour should provide considerable constraint on the motion of the contour. Additional constraint is still required, however, to determine this motion uniquely.

4. ADDITIONAL CONSTRAINT FOR MOTION MEASUREMENT

In this section, the computation of the velocity field is discussed from a theoretical viewpoint. The section is organized by the type of additional constraint that is used for this computation. Three types of constraint are considered: (1) velocity is constant over an area of the image; (2) the velocity field is consistent with rigid rotation and translation of objects in the image plane; and (3) the velocity field is smooth, and exhibits the least variation among the set of velocity fields consistent with the changing image.

The discussion of the constant velocity assumption, in Section 4.1, is primarily a review of previous work on motion measurement, as this assumption underlies most methods that have been proposed, both for computer and biological vision systems. The class of rigid motions in the image plane has been addressed in correspondence schemes for motion measurement. In Section 4.2, we present a scheme for analyzing the instantaneous velocity field for this class of motions. The constraint on smoothness of the velocity field, discussed in Section 4.3, was motivated by the work of Horn and Schunck (1981) on the optical flow computation, and allows the analysis of more general classes of motion.

4.1 The Constant Velocity Constraint

Much of the previous work in motion analysis has assumed that velocity is constant, at least over small areas of the image. If the projection of the scene onto the image plane is orthographic, the constant velocity constraint is strictly valid only when objects undergo pure translation. Examples of methods that use this constraint include those based on the cross-correlation of intensity, used both in computer vision (Leese *et al.* 1970; Lillestrand 1972; Smith and Phillips 1972; Wolferts 1974) and in modelling biological vision systems (Reichardt 1961; Lappin and Bell 1976; Petersik, Hicks and Pantle 1978; Anstis 1980), and correspondence schemes (Potter 1975, 1977; Mutch and Thompson 1982; Lawton 1983). In addition, most gradient schemes assume that velocity is constant over an area of the image (Limb and Murphy 1975; Fennema and Thompson 1979; Thompson and Barnard 1981; Marr and Ullman 1981). For gradient schemes, the constraint on velocity imposed by a single measurement of $v^{\perp}(s)$ can be illustrated graphically in *velocity space*, a space in which the x and y axes represent the x and y components of velocity, which we denote by V_x and V_y. This is shown in Figure 6. When mapped to velocity space, the velocity

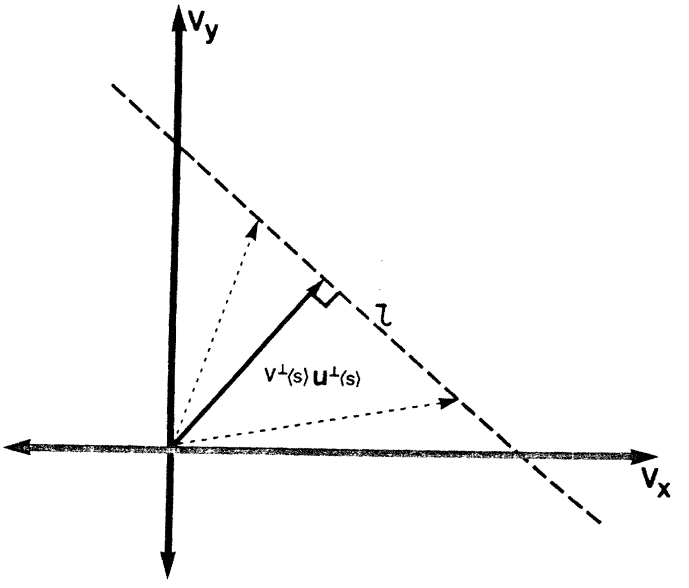


Figure 6. Velocity constraints in velocity space. $v^{\perp}(s)$ is the perpendicular component of velocity, and $\mathbf{u}^{\perp}(s)$ is the unit direction vector, at a point p on the image curve. The velocity vector at p must project to the line l; examples are shown with dotted lines.

vector at a point on the contour must terminate along the line l perpendicular to the vector $v^{\perp}(s)\mathbf{u}^{\perp}(s)$; examples are shown by the dotted arrows. For the case of uniform translation, the lines of constraint formed by multiple measurements of $v^{\perp}(s)$ along a contour intersect at a single point in velocity space, corresponding to the endpoint of the true velocity vector for the contour.

Some gradient schemes for motion measurement make explicit use of this intersection point (Fennema and Thomspon 1979; Thompson and Barnard 1981; Adelson and Movshon 1982). Marr and Ullman (1981) proposed a zero-crossing based scheme, in which each local motion measurement restricts the true direction of velocity of a patch to lie within a 180° range of directions to one side of the zero-crossing contour. A set of measurements taken at different orientations along the contour further restrict the allowable velocity directions, until a single direction is obtained, which is consistent with all the local measurements.

In many natural dynamic scenes, particularly those arising from motion of the observer, the velocity field can be approximated locally by pure translation. Most of the above schemes involve local operations for computing displacements or velocities, and can be effective for analyzing this type of motion. In more general situations, for example, when a nearby surface rotates in depth and deforms, the assumption of local translation is no longer sufficient.

4.2 Rigid Motion in the Image Plane

Some motion measurement schemes allow objects to undergo rigid rotation and translation in the image plane. For example, Davis *et al.* (1983) present an iterative gradient scheme that combines motion constraints along contours, starting from points of known velocity, and utilizing the perpendicular components of velocity. Several correspondence schemes also allow the analysis of this class of motions (Ullman 1979a; Aggarwal and Duda 1975; Chow and Aggarwal 1977; Barnard and Thompson 1980; Nagel 1982).

In the remainder of this section, we analyze the instantaneous velocity field for this class of motions, and present a simple geometric construction for computing the velocity field. Suppose that a rigid curve undergoes general motion in space. Its instantaneous motion may be described as the combination of: (1) a rotation with angular velocity ω about a single axis in space, which we denote by the unit vector $\mathbf{n} = [n_x, n_y, n_z]^T$ (T denotes the transpose of a vector), and (2) a translation, which we denote by the vector $\mathbf{t} = [t_x, t_y, t_z]^T$. Let the curve be given by $C = (x(s), y(s), z(s))$, where s denotes arclength. The location of a point on the curve may be given by the position vector $\mathbf{r} = [x(s), y(s), z(s)]^T$. If it is assumed that the axis of rotation passes through the origin of the coordinate system, the velocity of a point given by the position vector $\mathbf{r}$ is equal to the cross product of $\mathbf{r}$ with the vector $\omega\mathbf{n}$. If we then let the optical axis lie along the z-axis (with the positive z-axis directed toward the viewer), and let the projection of the curve onto the image plane (the (x, y) plane) be orthographic, then the two-dimensional velocity field $\mathbf{V}(s)$ along the contour is given by:

$$\mathbf{V}(s) = \mathbf{M}(\mathbf{r} \times \omega\mathbf{n} + \mathbf{t}) = -\omega z(s)\begin{bmatrix} n_y \\ -n_x \end{bmatrix} - \omega n_z \begin{bmatrix} -y(s) \\ x(s) \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \tag{2}$$

$\mathbf{M}$ denotes the matrix that performs the orthographic projection. The first term in the resulting expression describes the component of the velocity field due to rotation in depth about an axis parallel to the image plane (the axis $\mathbf{n} = [n_x, n_y, 0]^T$); the second term is the component due to motion in the image plane (rotation about the axis $\mathbf{n} = [0, 0, n_z]^T$), and the third term is the translation component.

Consider the restricted case of rigid motion in the image plane; the velocity field now corresponds to the combination of a translation, and rotation about the axis $\mathbf{n} = [0, 0, 1]^T$. Thus, $\mathbf{V}(s)$ is given by:

$$\mathbf{V}(s) = -\omega \begin{bmatrix} -y(s) \\ x(s) \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \tag{3}$$

$\mathbf{V}(s)$ is simply a translation, rotation and scaling of the image curve $(x(s), y(s))$, as illustrated in Figure 7. In Figure 7(a), the curve C_1 undergoes a small rotation and translation in the image

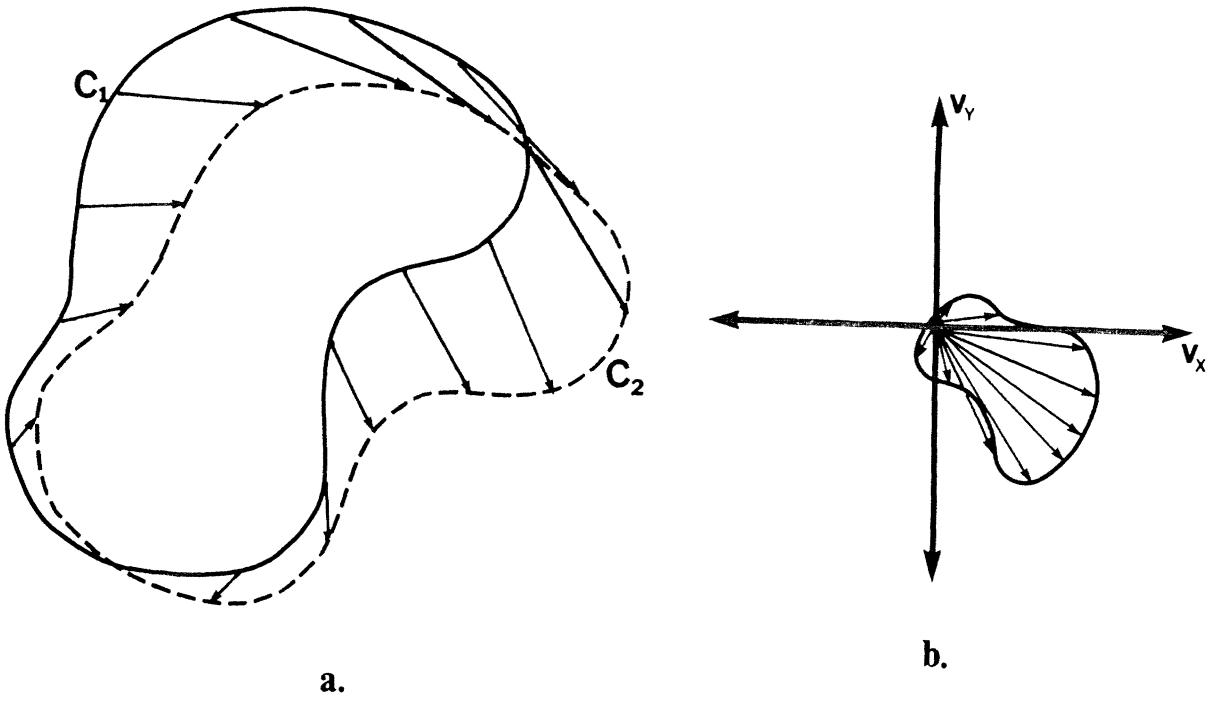


Figure 7. Velocity field for rigid motion in the image plane. (a) The image curve C_1 undergoes a small rotation and translation in the image to yield the curve C_2. (b) The velocity vectors drawn in velocity space.

plane to yield the curve C_2. The arrows indicate a sampling of the local velocity vectors along the curve. In Figure 7(b), these velocity vectors have been translated to a common origin in velocity space. The curve in velocity space has the same shape as the image curve C_1; its size is proportional to angular velocity ω, and it is rotated 90° with respect to C_1 (this relationship is also used in kinematics (Hartenberg and Denavit 1964)). The additional translation of the curve in the image yields the same translation of the curve in velocity space. In the case of pure translation, the image of the velocity field in velocity space degenerates to a single vector. For the case of discrete motion of a curve, a similar relationship exists between the shape of the image curve, and its *displacement field*, the set of vectors that describe the discrete displacement of points on the curve (Hildreth 1984a).

A simple geometric construction can be used to compute the velocity field, for the class of rigid motions in the image plane. If the true direction of velocity is known at two points on the contour, the direction of velocity can be computed everywhere as follows: (1) construct the lines perpendicular to the direction of velocity at the two known points, (2) compute the intersection of these two lines, (3) from every point p_i along the contour, construct the line to the intersection

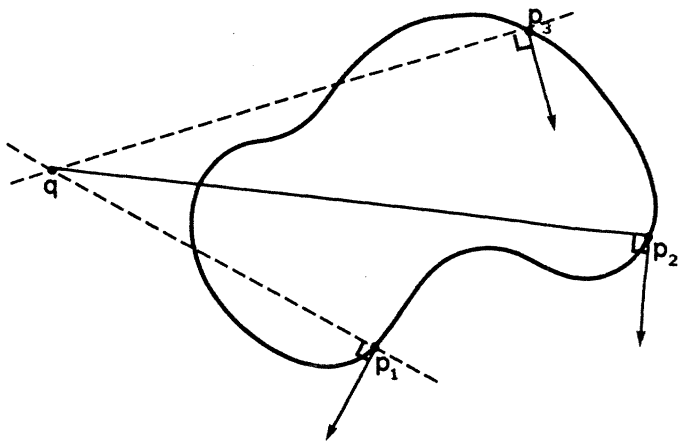


Figure 8. Constructing the velocity field. The direction of velocity is known at points p_1 and p_3, and is computed for the point p_2, using the construction scheme described in the text.

point; the true direction of velocity is perpendicular to this line. In Figure 8, the direction at p_2 is derived, given known directions at p_1 and p_3. This construction is simply locating the point about which the motion can be described as pure rotation. For pure translation, the two lines, from points of known direction of velocity, are parallel, so the direction of motion everywhere is equal to the direction of motion of the known points. Certainly, if both the direction and magnitude of velocity were known at two points on the contour, the global motion parameters, and hence direction and magnitude of velocity everywhere, could be computed. From the direction of velocity alone at two points on the contour, however, the direction of velocity can be obtained everywhere. If, in addition, the perpendicular components of velocity are known, then both direction and magnitude of velocity can be computed along the contour.

There are at least two sources for points of known velocity direction in the image. First, identifiable features, such as points of high curvature, may be tracked in two dimensions (Lawton 1983). Second, points at which the perpendicular component of velocity is zero are constrained to move along the direction of the tangent to the curve. For the case of a smooth, closed curve derived from a single object moving rigidly in the image plane, there exists at least two points on the curve for which the perpendicular component of velocity is zero (Hildreth 1984a). Suppose the velocity field computation is focused at the locations of zero-crossings derived from the image. Zero-crossing contours are generally closed (except at image boundaries), so there is usually sufficient constraint from the image to compute the velocity field, for this simple class of motions.

The methods described in this section, for computing the velocity field for the restricted

class of rigid motion in the image plane, are not sufficient for general motion analysis; however, they may be useful for the initial detection and rough measurement of motion in the periphery, the analysis of motion during smooth pursuit eye movements, or the recovery of observer motion from optical flow (Prazdny 1980; Longuet-Higgins and Prazdny 1981). In computer vision, there are restricted applications for these techniques, such as the tracking of objects, or computation of camera motion (Brüss and Horn 1983; Lawton 1983).

4.3 The Smoothness Constraint

In this section we derive a more general constraint on the velocity field, that allows the computation of the projected motion of three-dimensional surfaces that move freely in space, and deform over time. We rely on the physical assumption that the real world consists predominantly of solid objects, whose surfaces are generally smooth compared with their distance from the viewer. A smooth surface in motion usually generates a smoothly varying velocity field. Thus, intuitively, we seek a velocity field that is consistent with the motion measurements derived from the changing image, and which varies smoothly. Unfortunately, there is an infinity of velocity fields that satisfy these two properties. Horn and Schunck (1981), in their work on the optical flow computation, suggest that a single solution may be obtained by finding the velocity field that varies as little as possible. In the remainder of this section, we show how this constraint may be formulated more precisely, in a way that guarantees a velocity field solution that is mathematically unique and physically plausible.

4.3.1 Measuring Variation in Velocity

To find the velocity field that varies the least, some means of measuring the variation in velocity along a contour is required. This can be accomplished in many ways. For example, we could measure the change in direction of velocity as we trace along the contour. Total variation in velocity could then be defined as the total change in direction over the entire contour. A second possibility is to measure the change in magnitude of velocity along the contour. Third, the change in the full velocity vector could be measured, incorporating both the direction and magnitude of velocity. Other measures are also possible. The goal of the computation is to find a velocity field that is consistent with the changing image, and minimizes one of these measures of variation in velocity.

A measure of variation may be described more formally by defining a mathematical functional, Θ, which maps the space of all possible vector fields (along the contour), V, into the real numbers: $\Theta: V \mapsto \Re$. This functional should be such that the smaller the variation in the velocity field, the

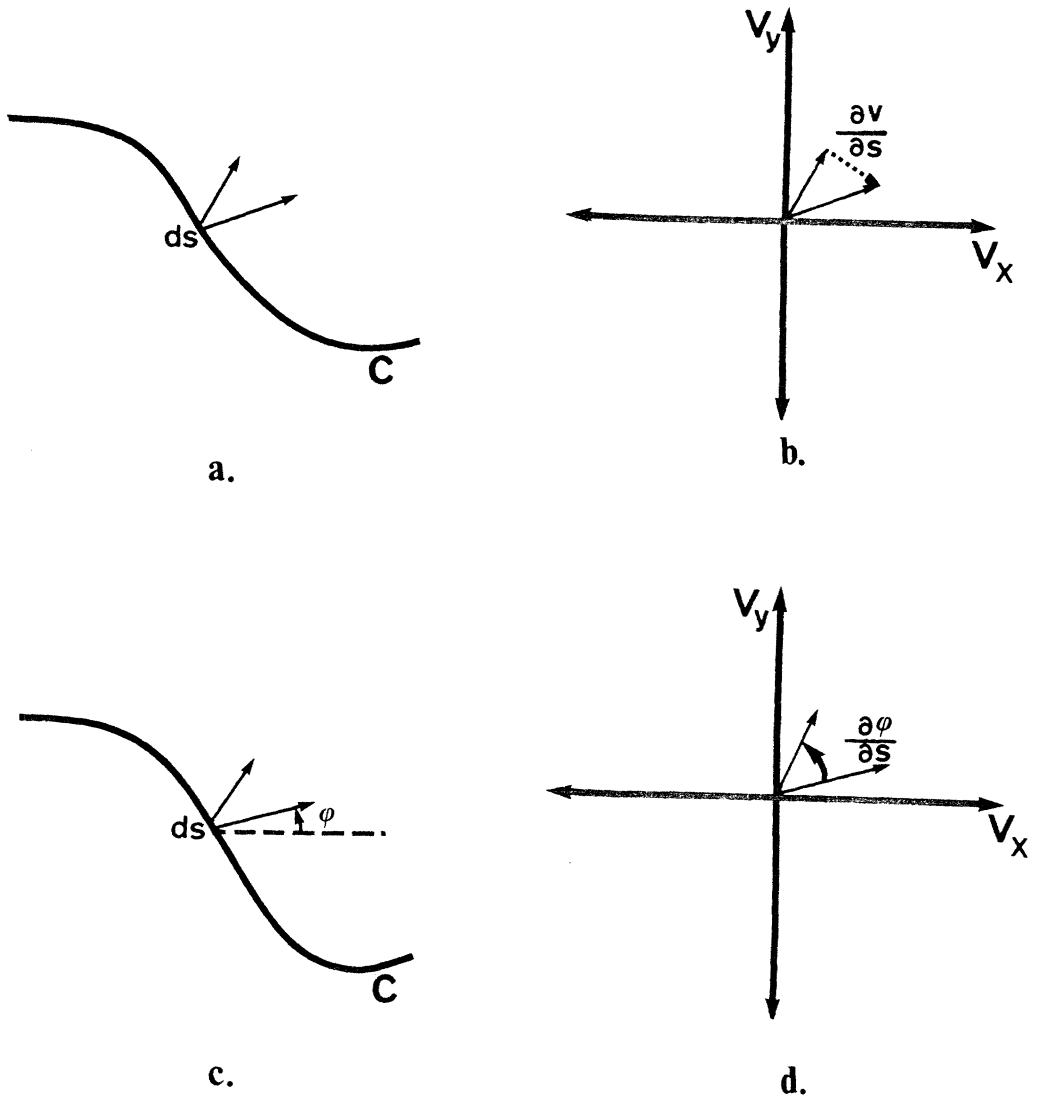


Figure 9. Measuring variation in velocity. (a) The vectors $\mathbf{V}(s)$ are displayed at two nearby points on the image curve C. (b) The velocity vectors drawn in velocity space, where $\frac{\partial \mathbf{V}}{\partial s}$ is indicated by the dotted arrow. (c) The direction of velocity for points on the contour is represented by the angle φ. (d) The velocity vectors of (c) are drawn in velocity space, where $\frac{\partial \varphi}{\partial s}$ is shown.

smaller the real number assigned to it. Two candidate velocity fields may then be compared, by comparing their corresponding real numbers. This formulation allows the development of an explicit method for computing the velocity field of least variation. A set of functionals can now be derived, based on the measures of variation that were previously mentioned informally: (1) variation in the full velocity vector, $\mathbf{V}(s)$, (2) variation in the direction of velocity, and (3) variation in the magnitude of velocity, all with respect to the curve.

1. Variation in $\mathbf{V}(s)$

The local change in $\mathbf{V}(s)$ with respect to the contour is given by the vector $\frac{\partial \mathbf{V}}{\partial s}$. A scalar measure may be obtained by taking its magnitude: $\left|\frac{\partial \mathbf{V}}{\partial s}\right|$. In Figure 9(a), two nearby velocity vectors along the image contour C are shown. The vectors are translated to a common origin in velocity space in Figure 9(b), where the vector $\frac{\partial \mathbf{V}}{\partial s}$ is shown as a dotted arrow. A measure of the total variation in the velocity field over an entire contour may be derived by integrating this local measure, suggesting a functional such as:

$$\Theta(\mathbf{V}) = \int \left|\frac{\partial \mathbf{V}}{\partial s}\right| ds.$$

Variations on this functional may also be considered, involving higher order derivatives, or higher powers, such as:

$$\Theta(\mathbf{V}) = \int \left|\frac{\partial^2 \mathbf{V}}{\partial s^2}\right| ds \qquad \text{or} \qquad \Theta(\mathbf{V}) = \int \left|\frac{\partial \mathbf{V}}{\partial s}\right|^2 ds.$$

2. Variation in Direction

Let the direction of velocity be described by the angle φ, measured in the counterclockwise direction from the horizontal, as shown in Figure 9(c). In Figure 9(d), the local change in direction for two nearby velocity vectors along the image contour, given by $\frac{\partial \varphi}{\partial s}$, is shown in velocity space. Total variation in direction along the contour may again be obtained by integrating this local measure, leading to functionals such as the following:

$$\Theta(\mathbf{V}) = \int \left|\frac{\partial \varphi}{\partial s}\right| ds$$

or variations involving higher order derivatives, or higher powers.

3. Variation in Magnitude

Finally, the total change in magnitude of velocity alone could be measured, using functionals such as:

$$\Theta(\mathbf{V}) = \int \frac{\partial |\mathbf{V}|}{\partial s} ds.$$

Again, we could also consider variations on this measure.

The functional that is used to measure variation may also incorporate a measure of the velocity field itself, rather than strictly utilizing changes in the velocity field along the contour, by incorporating a term that is a function of $|\mathbf{V}|$. This might be useful if we sought a velocity field that also exhibits the least total motion. In addition, the functional could become arbitrarily complex in its combination of $|\frac{\partial \mathbf{V}}{\partial s}|, |\frac{\partial \varphi}{\partial s}|, \frac{\partial |\mathbf{V}|}{\partial s}$, or higher order derivatives.

Given that there are many possible measures of variation, what criteria can be used to choose a single measure? First, from a mathematical point of view, there should exist a unique velocity field that minimizes the particular measure of variation; this requirement imposes a set of mathematical constraints on the functional. Second, the velocity field computation should yield solutions that are physically plausible. These two criteria are important for the evaluation of any assumption that is proposed for the motion measurement computation. If, in addition, we suggest that such a constraint underlies the motion computation in the human visual system, the minimization of this measure of variation should yield a velocity field that is consistent with human motion perception.

4.3.2 Mathematical Uniqueness of the Velocity Field

An examination of these measures of variation from a mathematical viewpoint suggests that a measure incorporating the change in the full velocity vector is necessary for the velocity field computation. The use of functionals that incorporate only a measure of direction or magnitude of velocity, for example, does not in general lead to a unique velocity field solution. It can be shown, however, that given a simple condition on the constraints that are derived from the image, there exists a unique velocity field that satisfies these constraints, and minimizes the particular measure of variation given by: $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$. The basic mathematical question is, what conditions on the form of the functional, and the structure of the space of velocity fields, are needed to guarantee the existence of a unique solution? These conditions are captured by the following theorem from functional analysis (Rudin 1973):

> Theorem: *Suppose there exists a complete semi-norm Θ on a space of functions H, and that Θ satisfies the parallelogram law. Then, every nonempty closed convex set $E \subset H$ contains a unique element v of minimal norm, up to an element of the null space. Thus, the family of minimal functions is*
>
> $$\{v + s \mid s \in S\}$$
>
> *where*
>
> $$S = \{w \mid v + w \in E\} \cap \mathcal{N}$$

and $\mathcal{N}$ is the null space of the functional

$$\mathcal{N} = \{u \mid \Theta(u) = 0\}.$$

The functional $\Theta = \{\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds\}^{\frac{1}{2}}$ is a complete semi-norm that satisfies the parallelogram law, and the space of possible velocity fields that satisfy the constraints derived from the changing image, is convex (Hildreth 1984a, b) It then follows from the above theorem that this space contains a unique element of minimal norm, up to possibly an element of the null space. The null space in this case is the set of constant velocity fields, because the condition that $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds = 0$ implies that $\left|\frac{\partial \mathbf{V}}{\partial s}\right| = 0$ everywhere, which implies that $\mathbf{V}(s)$ is constant. In addition, the smoothness measure is non-negative, so that minimizing $\{\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds\}^{\frac{1}{2}}$ is equivalent to minimizing $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$. These further observations lead to the following result (Hildreth 1984a, b):

> *If $v^{\perp}(s)$ is known everywhere along the contour, and there exists at least two points at which the local orientation of the contour is different, then there exists a unique velocity field that satisfies the known velocity constraints and minimizes $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$.*

An extended straight line does not yield measurements at two different orientations, but in all other cases, there is sufficient information along a contour to guarantee a unique velocity field solution. The smoothness constraint can be used to compute a projected two-dimensional velocity field for any three-dimensional surface, whether rigid or nonrigid, undergoing general motion in space. While it is not yet clear whether this general formulation of the smoothness constraint, or the particular measure $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$, is the most appropriate for the motion computation, it is important that this measure satisfies certain essential mathematical requirements, that the other measures do not. The use of a functional incorporating only a measure of velocity direction, for example, which attempts to make the local velocity vectors as parallel as possible, does not lead to a unique velocity field solution. In Appendix A.1, it is shown that for the simple case of a line segment moving rigidly in space, with known velocity vectors at its endpoints, there does not exist a unique velocity field that minimizes either variation in direction, or variation in magnitude of velocity. This result is not surprising; velocity has two degrees of freedom, and can be decomposed in various ways (for example, into direction and magnitude, x and y components, normal and tangential components, or radial and rotational components). If the measure of variation in velocity constrains only one of these degrees of freedom, such as direction, the other is allowed to vary freely, and cannot be specified uniquely. The simpler functional, $\int |\frac{\partial \mathbf{V}}{\partial s}| ds$, also does not yield a unique velocity field in general. In Appendix A.2, an example is presented for which a velocity field that minimizes this measure does not exist. Measures involving higher order derivatives impose a higher degree of smoothness on the velocity field, which may not be necessary for this computation. Our approach

is to consider first the simplest measure, $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$. If this is not sufficient for the velocity field computation, we can then consider a more complex measure.

4.3.3 Physical Plausibility of the Velocity Field Solution

The second criterion for evaluating a particular measure of variation in velocity is the physical plausibility of the resulting solution. One question that can be asked is, under what conditions will the velocity field that minimizes $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$ be the correct physical velocity field? If we assume orthographic projection of the scene onto the image, there are at least two classes of motion for which this is true. The first consists of arbitrary rigid objects undergoing pure translation. In this case, $\frac{\partial \mathbf{V}}{\partial s} = 0$ everywhere along contours in the image, and hence $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds = 0$. Since zero is the smallest value that the measure can obtain, it follows that if there exists a valid solution that is consistent with pure translation, then this solution minimizes $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$. Consequently, motion measurement schemes that rely on pure translation address a special case of this more general method.

The second class of motions includes rigid polyhedra, undergoing general motion in space. In particular, the following can be shown (Hildreth 1984a, b):

> *Suppose that a rigid three-dimensional object, consisting of straight lines intersecting in space, projects onto the image plane, using orthographic projection, in such a way that line intersections are preserved (that is, two lines intersect in the image if and only if their generators intersect in space). Further, suppose that this object undergoes a general displacement in space. Then the two-dimensional velocity field that satisfies* $v_\perp(s)$ *measured along lines in the image, and minimizes* $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$, *is the correct projected two-dimensional velocity field.*

The proof of this result takes advantage of the fact that the velocity field varies linearly along a straight line that is moving rigidly in space. The classes of *blocks world* objects (Roach and Aggarwal 1979), and simple polygons in the image plane (Aggarwal and Duda 1975), are special cases of this class of motions, for which an algorithm that minimizes $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$ is guaranteed to compute the correct velocity field.

Recently, Yuille (1983) derived a general condition under which the velocity field that minimizes $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$ is the correct velocity field. Let $\mathbf{V}'(s)$ denote the true projected two-dimensional velocity field for a curve in motion, and let $\mathbf{T}(s)$ denote the tangent vector along the curve. If the following relationship holds at every point on the curve:

$$\mathbf{T} \cdot \frac{\partial^2 \mathbf{V}'}{\partial s^2} = \mathbf{0} \tag{4}$$

then the velocity field $\mathbf{V}(s)$ that satisfies the constraints imposed by $v^{\perp}(s)$ and minimizes $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$ is the true velocity field $\mathbf{V}'(s)$. The two classes of motion mentioned above correspond to cases for which $\frac{\partial^2 \mathbf{V}'}{\partial s^2} = \mathbf{0}$ along the curve, so that this general condition holds trivially. Yuille (1983) notes that this condition is also satisfied when a curve undergoes pure expansion in the image plane.

For the class of smooth curves, undergoing general motion in space, we rely on an empirical investigation for demonstrating the physical plausibility of the velocity field solution. In general, the velocity field of least variation, in this case, is not the physically correct one. It is often qualitatively similar to the true velocity field, however. As we will show in Section 5.4.1, when the two velocity fields differ significantly, it appears that the smoothest velocity field may be more consistent with human motion perception.

A possible constraint on the velocity field, that is not considered explicitly, is the rigidity of the underlying surface. The computation of a *smoothest* velocity field does not necessarily seek a solution that corresponds to rigid motion, in either two or three dimensions. This may at first seem physically implausible. When a three-dimensional curve rotates in space, however, its two-dimensional projection may undergo significant distortion in the image. Without knowing the three-dimensional structure of the curve, it is very difficult, if not impossible, to find a two-dimensional velocity field that corresponds to a single rigid motion in three dimensions. It is also the case that some of the motion that we encounter arises from surfaces that are nonrigid. If the analysis of motion is a two-stage process, with the measurement of two-dimensional motion preceding the derivation of three-dimensional structure from motion, a constraint such as smoothness may be the most restrictive type of constraint that may be used, which yields a unique solution, and still allows the analysis of general motion.

5. AN ALGORITHM AND EXAMPLES

In the previous section, the use of the smoothness constraint led to the formulation of the velocity field computation as an optimization problem. We seek a velocity field solution that satisfies the constraints derived from the changing image, and minimizes the measure of variation along contours given by: $\int |\frac{\partial \mathbf{V}}{\partial s}|^2 ds$. The computation may also be described as seeking a solution for which neighboring velocity vectors are as similar as possible. To further test the adequacy of this approach, it is necessary to specify an algorithm that embodies the smoothness constraint, and examine the results of the algorithm for a number of motion sequences. An algorithmic study first allows us to demonstrate the biological feasibility of this approach, by showing that

this formulation of the velocity field computation lends itself naturally to algorithms that involve simple, local, parallel operations (Ullman 1979b; Grimson 1981; Marr 1982). Second, an empirical study provides a further means for testing the physical plausibility of the resulting velocity field solution.

In general, there are many algorithms that implement a given computational theory. For the velocity field computation, there are many choices for the design of an algorithm that embodies the additional smoothness constraint. In this section and in Appendix B, we use a standard, iterative algorithm from mathematical programming, known as the *conjugate gradient algorithm* (Luenberger 1973), to implement this computation. This particular algorithm is certainly not appropriate as a model for human vision. Our aim is to test the basic idea of computing the velocity field of least variation. If the results of the algorithm support the feasibility of this idea, we can then explore alternative algorithms to implement the theory, that may be more appropriate for the human system.

5.1 The Discrete Formulation

In Section 3.3, a continuous formulation of the smoothness constraint was presented; the image, however, is discrete, and therefore image contours consist of a set of discrete points. The first step in the design of an algorithm is to convert the continuous formulation into a discrete one. To do this, the functional can first be expressed in terms of the x and y components of velocity, which we denote by V_x and V_y. The continuous functional becomes:

$$\Theta = \int \left[\left(\frac{\partial \mathrm{V}_x}{\partial s} \right)^2 + \left(\frac{\partial \mathrm{V}_y}{\partial s} \right)^2 \right] ds. \tag{5}$$

Suppose there are n points along a contour, whose coordinates are given by:

$$\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}.$$

Let $(\mathrm{V}_{x_i}, \mathrm{V}_{y_i})$ denote the velocity at the point (x_i, y_i). Assuming for now that the contour is closed, and the points are evenly spaced along the contour, the following function defines a discrete correlate to the continuous functional:

$$\Phi_1 = \sum_{i=2}^{n} [(\mathrm{V}_{x_i} - \mathrm{V}_{x_{i-1}})^2 + (\mathrm{V}_{y_i} - \mathrm{V}_{y_{i-1}})^2] + [(\mathrm{V}_{x_1} - \mathrm{V}_{x_n})^2 + (\mathrm{V}_{y_1} - \mathrm{V}_{y_n})^2]. \tag{6}$$

For the case of an open contour, the discrete differences at the endpoints change. If points on the contour are not evenly spaced, the discrete differences must be divided by the distances

between adjacent points on the contour. The goal of the computation is to compute the set of x and y components of velocity, $\{(\mathbf{V}_{x_1}, \mathbf{V}_{y_1}), (\mathbf{V}_{x_2}, \mathbf{V}_{y_2}), ..., (\mathbf{V}_{x_n}, \mathbf{V}_{y_n})\}$, which minimizes the discrete function Φ_1.

5.2 Satisfying the Image Constraints

The second step toward defining an algorithm is to incorporate the constraints on velocity that are derived from the changing image. Here there is a choice; the velocity field can be forced to satisfy the constraints exactly, or satisfy them approximately. In the first case, the constraints can be expressed as follows:

$$\mathbf{V} \cdot \mathbf{u}^{\perp} - v^{\perp} = 0. \tag{7}$$

The equation states explicitly that the perpendicular component of velocity for the computed velocity field, $\mathbf{V} \cdot \mathbf{u}^{\perp}$, is equal to the measured perpendicular component, $v^{\perp}$. (For simplicity, we omit the argument, s, to these functions, throughout the remainder of the paper.) Letting $\mathbf{V} = (\mathbf{V}_x, \mathbf{V}_y)$ and $\mathbf{u}^{\perp} = (\mathbf{u}_x^{\perp}, \mathbf{u}_y^{\perp})$, the constraints are simple linear constraints:

$$\mathbf{V}_{x_i}\mathbf{u}_{x_i}^{\perp} + \mathbf{V}_{y_i}\mathbf{u}_{y_i}^{\perp} - v_i^{\perp} = 0. \tag{8}$$

This equation can be used to express the y components of velocity in terms of the x components:

$$\mathbf{V}_{y_i} = \frac{v_i^{\perp} - \mathbf{V}_{x_i}\mathbf{u}_{x_i}^{\perp}}{\mathbf{u}_{y_i}^{\perp}}. \tag{9}$$

This expression is valid for $\mathbf{u}_{y_i}^{\perp} \neq 0$. If $\mathbf{u}_{y_i}^{\perp} = 0$, then $\mathbf{u}_{x_i}^{\perp} = 1$, and it follows that $\mathbf{V}_{x_i} = v_i^{\perp}$, and $\mathbf{V}_{y_i}$ is unknown. This situation arises at points on the contour for which the vector $\mathbf{u}^{\perp}$ is oriented horizontally (the contour is oriented vertically). At such a point, $v^{\perp}$ specifies the x component of velocity directly, and the y component is unconstrained. For points at which $\mathbf{u}_{y_i}^{\perp} \neq 0$, the above expressions for $\mathbf{V}_{y_i}$ can be substituted directly into the discrete function Φ_1. At points for which $\mathbf{u}_{y_i}^{\perp} = 0$, the known values of $\mathbf{V}_{x_i}$ and unknown parameters $\mathbf{V}_{y_i}$ can be substituted into Φ_1. Given n points on the contour, there remains n unknown parameters to compute, with no further constraint on these parameters.

In general, there will be error in the measurements of $v^{\perp}$. From a practical standpoint, it may be advantageous to require that the velocity field only approximately satisfy these image constraints. This can be accomplished by requiring that the difference between $\mathbf{V} \cdot \mathbf{u}^{\perp}$ and the measured $v^{\perp}$ be small, rather than requiring this difference to be zero, as in Equation (7). The continuous functional can be extended as follows:

$$\Theta = \int \left[\left(\frac{\partial \mathbf{V}_x}{\partial s} \right)^2 + \left(\frac{\partial \mathbf{V}_y}{\partial s} \right)^2 \right] ds + \beta \int \left[\mathbf{V} \cdot \mathbf{u}^{\perp} - v^{\perp} \right]^2 ds. \tag{10}$$

β is a weighting factor that expresses the confidence in the measured velocity constraints. In this particular formulation, β is constant over the entire contour, but in principle, β could vary, if confidence in the precision of individual measurements of $v^{\perp}$ varies along the contour. The second term describes the least squares difference between the computed and measured perpendicular components of velocity. There will also be error in the measurement of the direction vector $\mathbf{u}^{\perp}$; some of this error is captured implicitly in the second term of the above functional. The continuous functional leads to the following discrete function:

$$\Phi_2 = \Phi_1 + \beta \sum_{i=1}^{n} \left[\mathrm{V}_{x_i} \mathbf{u}_{x_i}^{\perp} + \mathrm{V}_{y_i} \mathbf{u}_{y_i}^{\perp} - v_i^{\perp} \right]^2. \tag{11}$$

The goal of the computation in this case is to find the set of x and y components of velocity, $\{(\mathrm{V}_{x_1}, \mathrm{V}_{y_1}), (\mathrm{V}_{x_2}, \mathrm{V}_{y_2}), ..., (\mathrm{V}_{x_n}, \mathrm{V}_{y_n})\}$, which minimizes the discrete function Φ_2. These components satisfy the system of $2n$ linear equations given by:

$$\frac{\partial \Phi_2}{\partial \mathrm{V}_{x_i}} = 0, \qquad \frac{\partial \Phi_2}{\partial \mathrm{V}_{y_i}} = 0, \qquad 1 \leq i \leq n. \tag{12}$$

In general, there may be several hundred points on a given contour, giving rise to a system of several hundred linear equations. To solve this system of equations, we use techniques from mathematical programming.

5.3 An Algorithm from Mathematical Programming

The general mathematical programming problem can be stated as:

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & h_i(\mathbf{x}) = 0 & i = 1, ..., m \\ & g_j(\mathbf{x}) \leq 0 & j = 1, ..., r \\ & \mathbf{x} \in S & \end{array}$$

where $\mathbf{x}$ is an n-dimensional vector of unknowns, $\mathbf{x} = (x_1, x_2, ..., x_n)$. The objective function f, and constraints h_i, $i = 1, ..., m$ and g_j, $j = 1, ..., r$ are real-valued functions of the variables $x_1, x_2, ..., x_n$. The set S is a subset of the n-dimensional space. A particular optimization problem may or may not give rise to constraints of the form $h_i(\mathbf{x}) = 0$ or $g_j(\mathbf{x}) \leq 0$; that is, the problem may be a *constrained* or *unconstrained* optimization problem.

For the case in which the image constraints are satisfied approximately, the velocity field computation can be formulated as an unconstrained optimization problem. In this case, we let:

$$\mathbf{x} = \{\mathrm{V}_{x_1}, \mathrm{V}_{x_2}, ..., \mathrm{V}_{x_n}, \mathrm{V}_{y_1}, \mathrm{V}_{y_2}, ..., \mathrm{V}_{y_n}\}.$$

The objective function $f(\mathbf{x}) = \Phi_2$. We seek a vector $\mathbf{x}$ that minimizes $f(\mathbf{x})$, with no further constraints on $\mathbf{x}$.

For the particular examples presented in this paper, the only constraints obtained from the image are the measurements of $v^\perp$ along contours. In general, there may be points at which $\mathbf{V}(s)$, or the direction of velocity alone, is known. This information may be the result of a process that tracks localizable features in the image, such as corners, isolated points, or line endpoints. These additional constraints may be incorporated easily into the velocity field algorithm (Hildreth 1984a).

There are several standard algorithms from mathematical programming that can be used to solve this particular optimization problem (Luenberger 1973). For the examples in the next section, the *conjugate gradient algorithm* was used to obtain the solution. This is an iterative algorithm that utilizes the gradient of the objective function $f(\mathbf{x})$ to choose an optimal path to follow toward the final solution. The initial velocity field is given by the perpendicular velocity vectors, $v^\perp \mathbf{u}^\perp$, along a contour. If there are n parameters to be computed, the conjugate gradient algorithm is guaranteed to converge to the exact solution in at most n iterations. Details of the application of the conjugate gradient algorithm to the velocity field computation are given in Appendix B.

5.4 Examples of the Smoothest Velocity Field

In this section, we present examples of the velocity field of least variation, for a variety of motion sequences. The conjugate gradient algorithm was chosen for its fast convergence and because it is guaranteed to compute the velocity field of least variation. The algorithm provides us with a means of testing whether the use of the smoothness constraint is a reasonable approach to the velocity field computation.

5.4.1 Ideal Smooth Curves

We begin with some simple curves, undergoing rigid motion. For this first set of examples, the curves and their perpendicular components of velocity were generated analytically from a known velocity field, and therefore represent ideal input for the algorithm. Many of the examples were chosen because perceptual studies indicate that human observers see motions that differ from the true motion of the curves.

1. Rigid translation in the image plane

In Figure 10(a), a sampling of the true velocity field is shown for a simple polygon, translating rigidly in the image plane. Figure 10(b) illustrates the perpendicular velocity vectors $v^\perp \mathbf{u}^\perp$, which

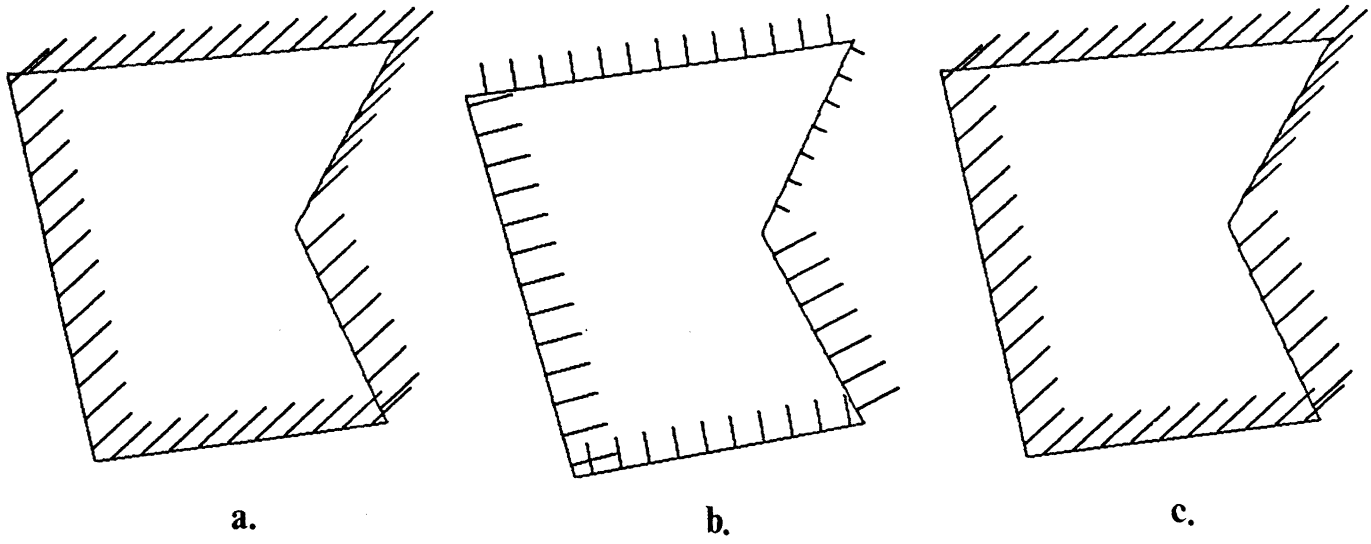


Figure 10. Pure translation. (a) The arrows represent a sampling of the true velocity field for a translating polygon. (b) The initial perpendicular velocity vectors. (c) The computed velocity field.

form the input to the algorithm. In this case, the velocity field of least variation, shown in Figure 10(c), is identical to the true velocity field, as expected.

2. Polygon Rotating in the Image Plane

Figure 11 illustrates the true velocity field, initial perpendicular velocity vectors, and smoothest velocity field for a polygon rotating rigidly in the image plane, about the point O. From the theoretical results of Section 4.3.3, we expect to obtain the correct velocity field in this case. Figure 11(c) shows that the computed smoothest velocity field is in fact the correct one.

3. Rotating Ellipses

Figure 12 illustrates the true velocity field, initial perpendicular velocity vectors, and smoothest velocity field for an ellipse rotating rigidly in the image plane, about the point O. In this case, the smoothest velocity field is quite different from the true velocity field. There is a reduced rotational component of velocity, and added radial component in the computed velocity field. The difference in total variation is significant.

At first glance, one might not consider the smoothest velocity field in this case to be a plausible solution. In some earlier perceptual experiments by Wallach *et al.* (1956), however, they noted that a rigid ellipse does not appear rigid under rotation; it appears to deform continuously. In their experiments, simple geometric figures were placed on a rotating turntable, and observers described the perceived motion of the figures while fixating the center of the turntable (conditions

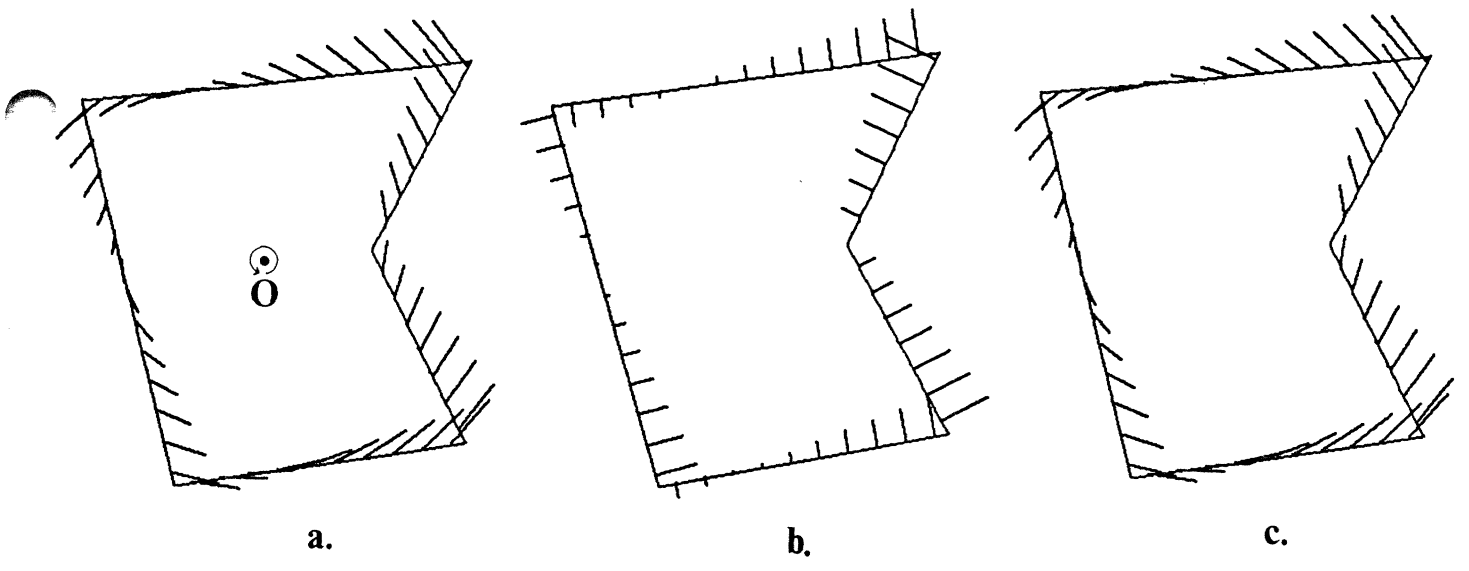


Figure 11. Rotating polygon. (a) The true velocity field for a polygon rotating rigidly in the image about the point *O*. (b) The initial perpendicular velocity vectors. (c) The computed velocity field.

of free and tracking eye movements were also used). Ellipses of various aspect ratios were observed. It was found that when an ellipse whose axes measured 25 and 23.5 cm was rotated about its center, it appeared to stand still while its contour pulsated. The largest effects were observed for an ellipse whose aspect ratio was 3:2; the entire figure appeared fluid, undergoing a strong deformation, as well as a rotation. For some observers, the deformation was more restricted, and did not occur in the immediate vicinity of the poles of the major axis of the ellipse. As the aspect ratio of the ellipse was increased, it appeared more rigid. The perceived deformation of the ellipse was the same, regardless of whether it rotated about its center, or was placed eccentrically on the turntable.

For the ellipse of Figure 12, the aspect ratio is 2:1. The computed velocity field of least variation clearly implies a significant distortion of the contour, in addition to a rotation. In the immediate vicinity of the poles of the major axes, the smoothest velocity field is very similar to the true velocity field. In Figure 13, the true and smoothest velocity fields are shown for rotating ellipses whose aspect ratios are 25:23.5 and 5:1. When the ellipse is nearly circular, the smoothest velocity field indicates a strong inward and outward pulsation of the contour (Figure 13(b)). The smoothest velocity field for the narrower ellipse, shown in Figure 13(d), is closer to the true velocity field than in the case of the ellipse with aspect ratio 2:1, implying less distortion of the contour. Finally, the smoothest velocity field for an eccentrically rotated ellipse differs from that

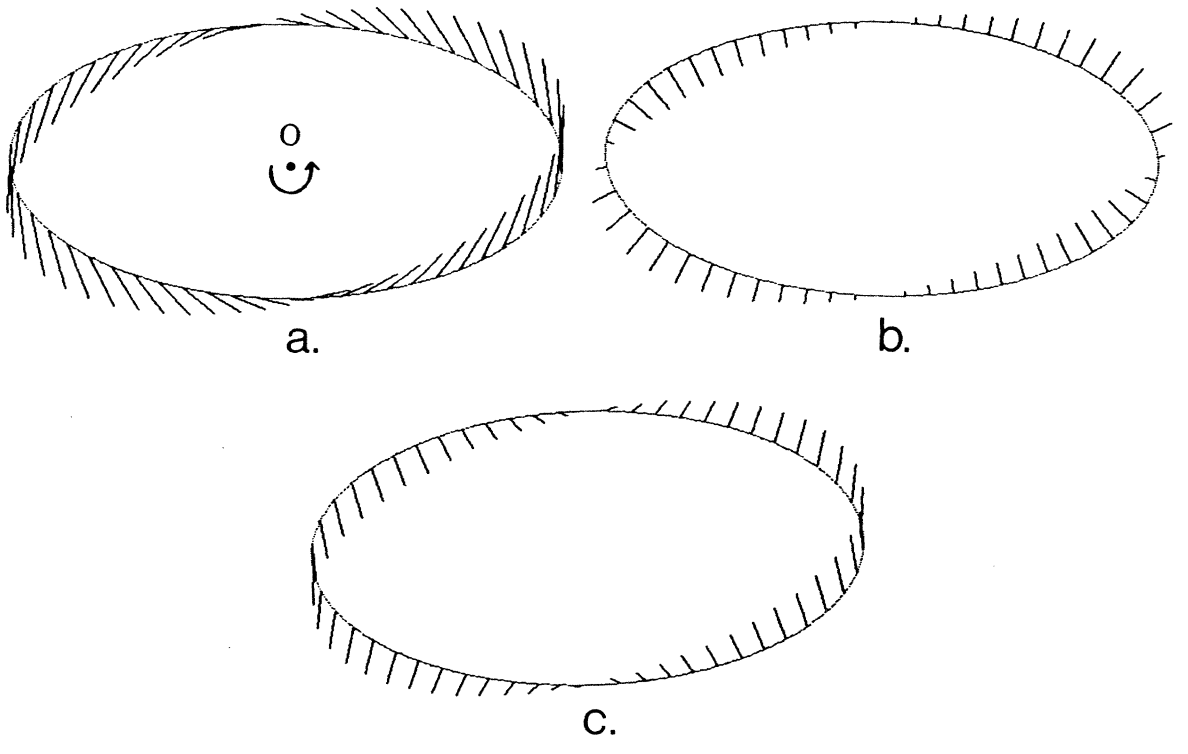


Figure 12. Rotating ellipse. (a) The true velocity field for an ellipse rotating rigidly in the image about the point *O*. (b) The initial perpendicular velocity vectors. (c) The computed velocity field.

of the centrally rotated ellipse only by the addition of a uniform translation along the contour. Thus, the same deformations of the ellipses is implied by the smoothest velocity field obtained for eccentric rotations. We conclude that the perception of the movement of rotating ellipses is at least qualitatively consistent with the computation of the velocity field of least variation.

4. Rotating Spiral

It is well known that a spiral appears to expand or contract, when it undergoes pure rotation about its center (Holland 1965). The perceived velocity field thus contains a large radial component, while the true velocity field contains only a rotational component of velocity. Figure 14 illustrates the true velocity field, initial perpendicular velocity vectors, and computed smoothest velocity field, for a single arm of a rotating logarithmic spiral. The smoothest velocity field exhibits a large radial component of motion at the center of the spiral, which decreases toward the periphery. The perception of the movement of the spiral is qualitatively more consistent with the smoothest velocity field, than with either the true or initial velocity fields.

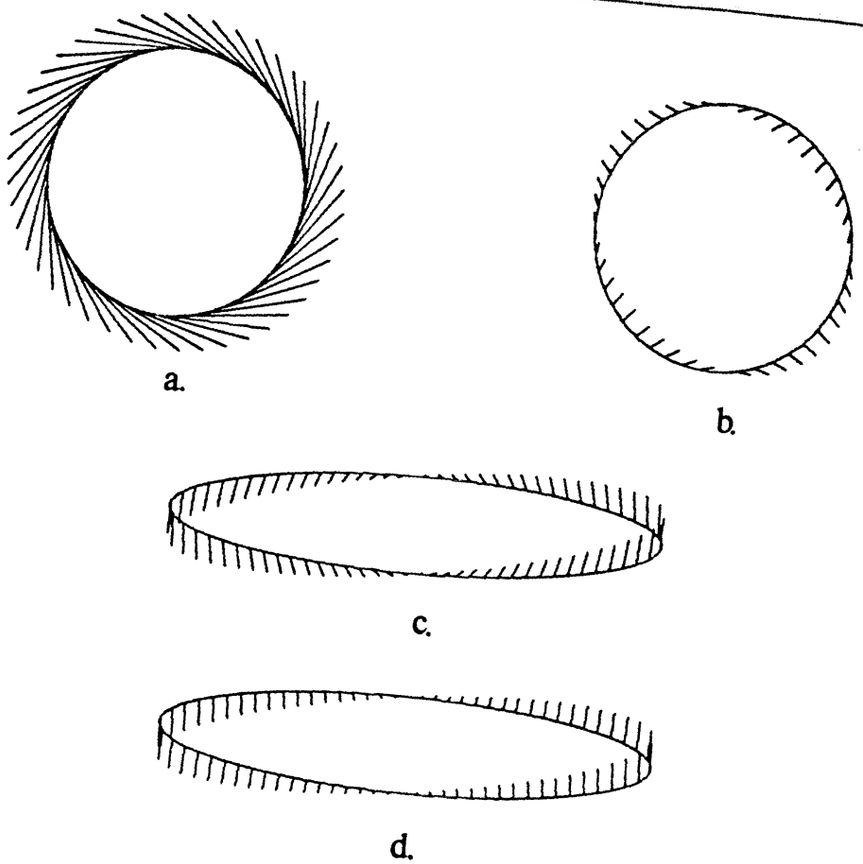


Figure 13. Rotating ellipses of different aspect ratios. (a) and (c) The true velocity fields for rotating ellipses with aspect ratios of 23:23.5 and 5:1. (b) and (d) The computed velocity fields for the ellipses in (a) and (c).

6. The Moving Dot Demonstration

Another motion demonstration of interest was presented by Anstis and Ramachandran (Anstis 1983). The first frame of a motion pair consisted of a sparse array of random dots. In the second frame, each of the dots was replaced by two dots, with one displaced horizontally, and the other vertically. The local configuration of dots is shown in Figure 15(a). The filled circle represents a dot from the first frame, and open circles represent dots in the second frame. The horizontal displacement was half the vertical displacement. The inserted scale indicates the size of the displacements used in the actual presentation, in minutes of visual arc. Anstis and Ramachandran found that when the two frames were alternated, the whole field of dots appeared to move obliquely as a unit. The perceived direction of velocity varied with the absolute displacement of the dots. Viewed closely (with large displacements), perceived direction of velocity was roughly

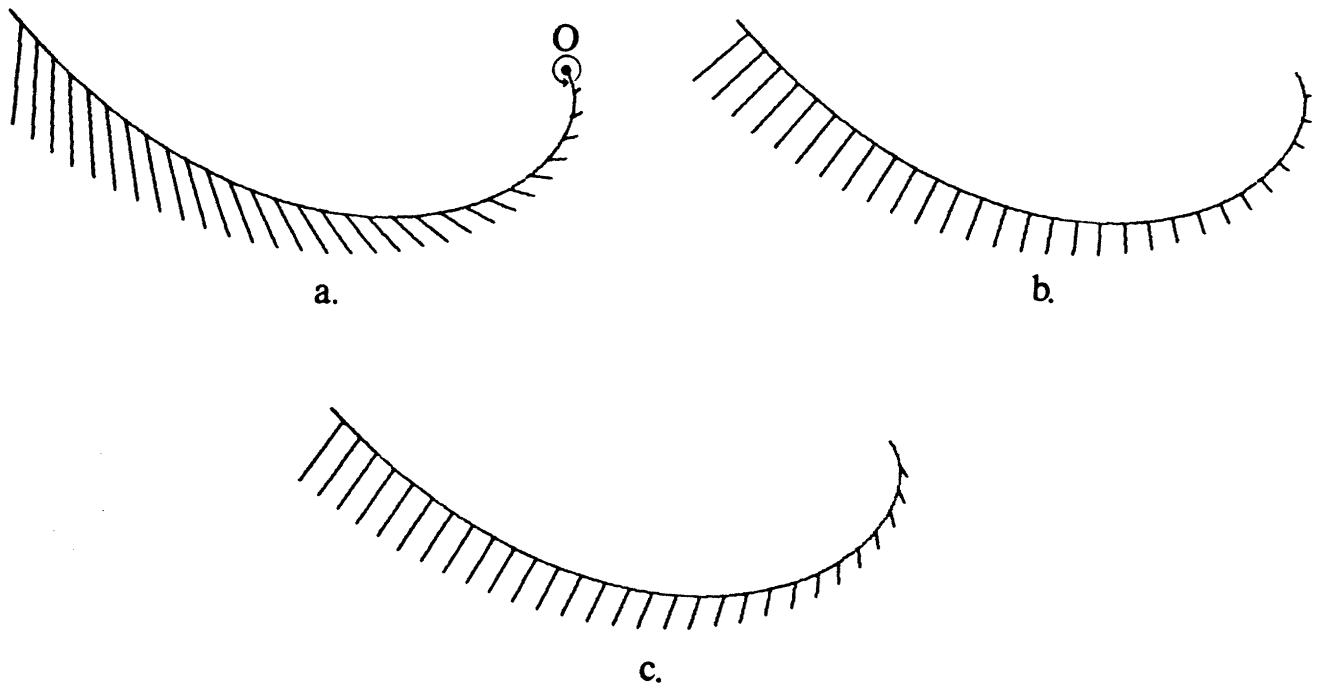


Figure 14. Rotating spiral. (a) The true velocity field for a logarithmic spiral rotating in the image about the point O. (b) The initial perpendicular velocity vectors. (c) The computed velocity field.

horizontal. As viewing distance increased (displacements decreased), the apparent direction of motion became increasingly more vertical.

The predictions of the velocity field algorithm are shown in Figures 15(b) through 15(e). In this case, the two images were convolved with a $\nabla^2 G$ operator, and the zero-crossings (which encircle individual dots) were obtained. The time derivative, $\frac{\partial}{\partial t}(\nabla^2 G * I)$ was computed by subtracting the two filtered images. At the location of a zero-crossing, the local vector $\mathbf{u}^\perp$ was computed from the gradient of the filtered image, and the perpendicular components of velocity, $v^\perp$, were computed as follows:

$$v^\perp = \frac{-\frac{\partial}{\partial t}(\mathbf{I}')}{|\nabla \mathbf{I}'|} \tag{13}$$

where ∇ denotes the gradient operator, and $\mathbf{I}'$ is the filtered image. In the leftmost column, the actual displacements between dots are given as a function of w, the diameter of the central positive region of the initial $\nabla^2 G$ operator. The resulting smoothest velocity fields are shown enlarged in the rightmost column of Figures 15(b) through 15(e). As the absolute displacement of the dots decreases, the predicted direction of motion becomes increasingly more vertical, consistent with the perception.

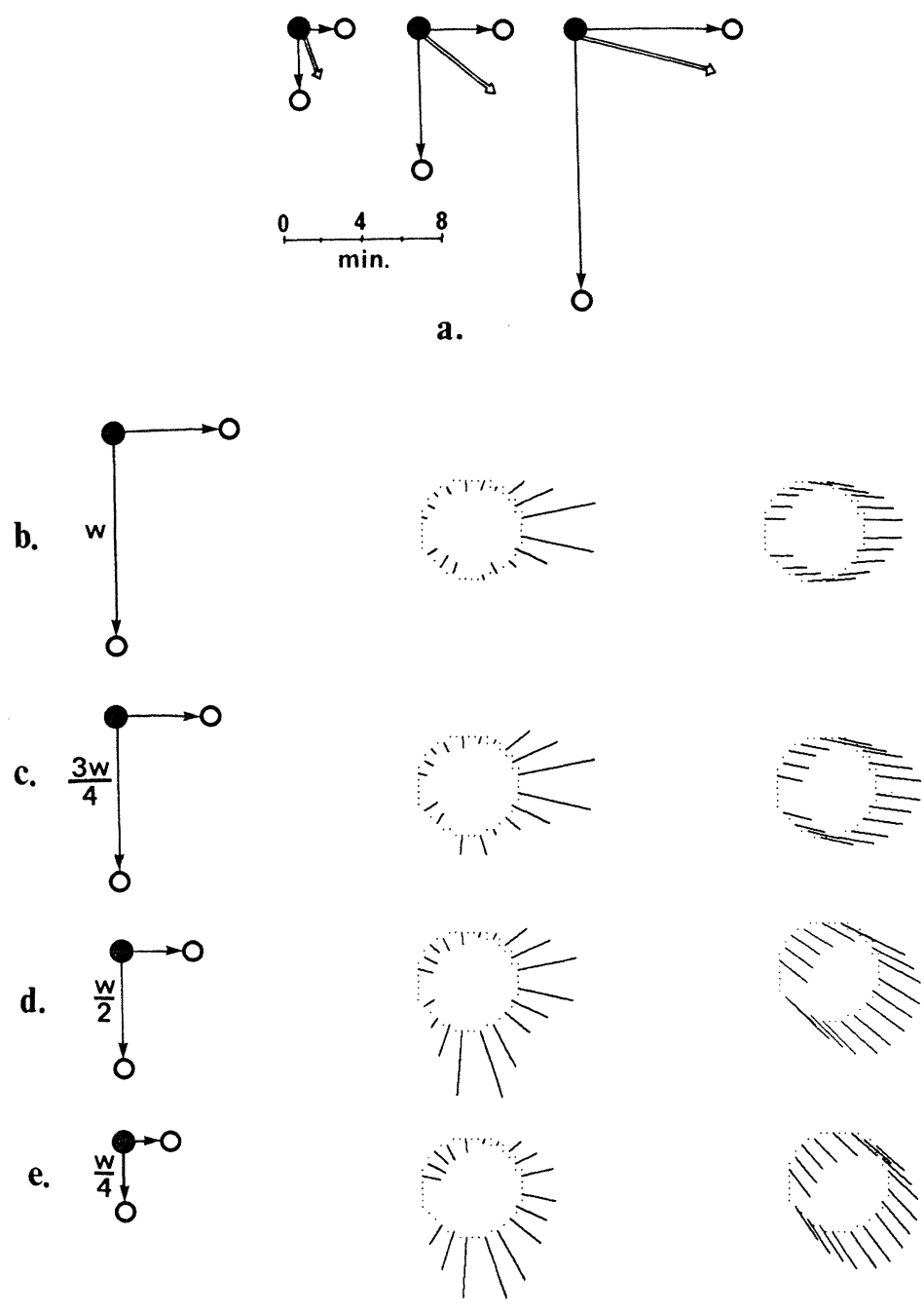


Figure 15. The moving dot demonstration. (a) Closed circles represent a dot from the first frame, and open dots represent dots from the second. The double arrows show the perceived direction of motion of the full dot pattern. The inserted scale indicates the size of the displacements. (b) through (e) The leftmost column indicates displacements as a function of w, the middle column illustrates the initial perpendicular components of velocity, and the rightmost column illustrates the computed velocity fields.

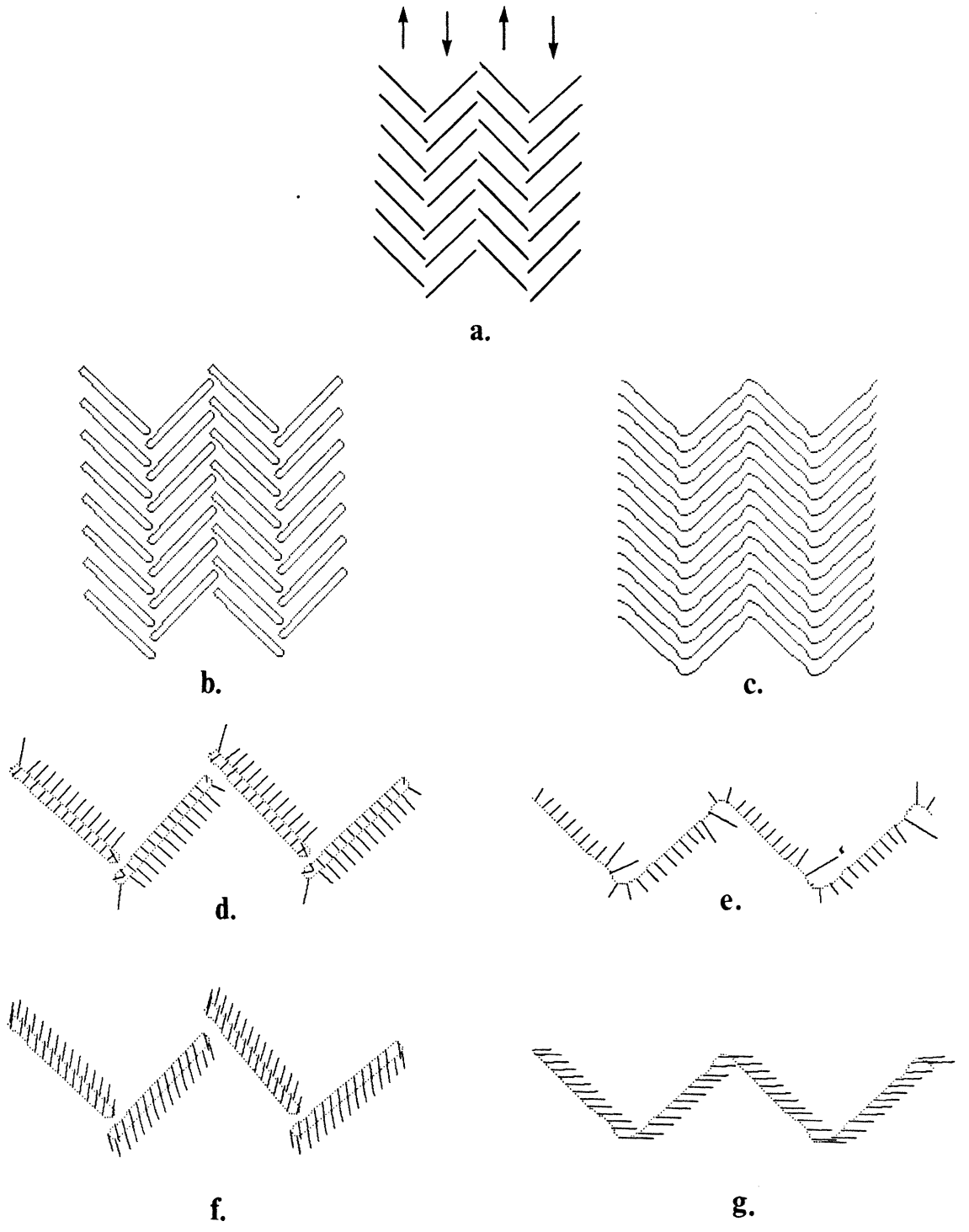


Figure 16. The split herringbone demonstration. (a) Alternating columns of a herringbone pattern of lines move in opposite directions. (b) and (c) The zero-crossing contours derived for two different size $\nabla^2 G$ operators. (d) and (e) The initial perpendicular components of velocity. (f) and (g) The computed velocity fields.

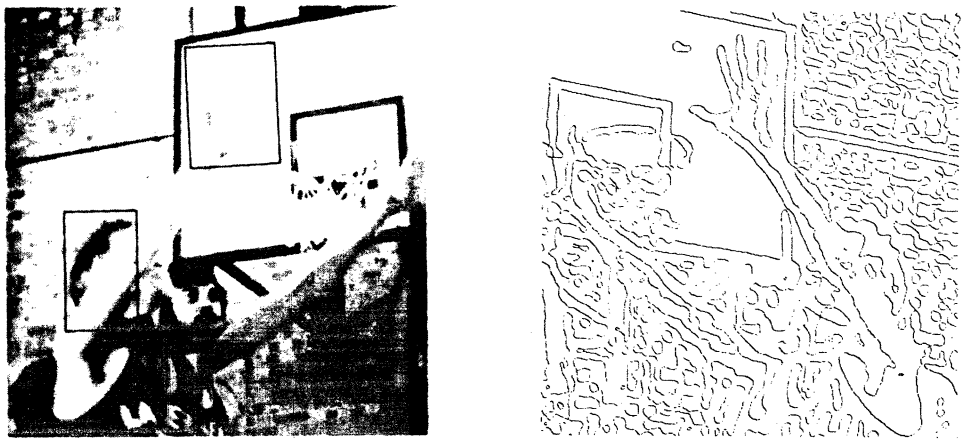
7. The "split herringbone" demonstration

Adelson and Movshon (1983) present a motion demonstration, referred to as the "split herringbone." The display, shown in Figure 16(a), consists of alternating columns of line segments that tilt left or right in the odd and even columns. The odd (left-tilting) columns move upward, while the even (right-tilting) columns move downward. There are two perceptions of motion that result, depending on the viewing conditions. When the display is of high contrast, sharp, and centrally fixated, the columns are seen moving vertically, in opposite directions. When the display is of low contrast, blurred with a diffusion screen, or viewed peripherally, an illusion of rightward motion is seen. To explain this phenomenon, Adelson and Movshon (1983) distinguish two kinds of velocity field computations. The first is a local computation, that involves the tracking of localizable features, such as the line endpoints in the herringbone pattern. The second is a more "global" process that combines velocity constraints derived from different parts of the image. They argue that the perception of vertically moving columns represents the results of the first type of process, while the perception of rightward movement represents the results of the second.

It is possible that both perceptions arise from a single underlying velocity field computation that combines velocity constraints along contours in the image, which is applied to image descriptions derived from initial $\nabla^2 G$ operators of different size. (Evidence for the existence of different size filters in early human vision can be found, for example, in Wilson and Bergen (1979).) Figures 16(b) and 16(c) show the zero-crossing contours derived from the convolution of the image of Figure 16(a) with $\nabla^2 G$ operators whose central positive diameter, w, is 6 and 16 picture elements, respectively. With $w = 6$, the zero-crossing contours surround individual bars, while in the case where $w = 16$, the zero-crossings merge from one column to the next, forming continuous contours across the image. Figures 16(d) and 16(e) illustrate the initial perpendicular components of velocity, obtained from the temporal derivative and gradient of $\nabla^2 G * I$. Figures 16(f) and 16(g) illustrate the resulting smoothest velocity field in each case. For the smaller operator, the velocity directions are roughly vertical, while for the larger operator, the smoothest velocity field indicates horizontal movement of the contours. Horizontal motion was perceived under the conditions of peripheral viewing or blurring, for which a coarser initial filtering of the image is expected (Wilson and Bergen 1979).

5. Other Examples of Ideal Smooth Curves

Other examples of ideal smooth curves in motion, which illustrate the consistency between the velocity field of least variation, and human motion perception, shown in (Hildreth 1984a, b), include the *kinetic depth effect* stimulus, used by Wallach and O'Connell (1953) to demonstrate the



a. b.

Figure 17. A natural image sequence. (a) The original image, containing 576 × 576 picture elements. (b) The resulting zero-crossing contours.(The images in (a) and (b) are mirror reversed.)

ability of the human visual system to derive three-dimensional structure from motion, the rotating deformed circle studied in (Wallach *et al.* 1956), and a rotating helix, which yields a perception of motion that is similar to the *barberpole illusion* (the perceived downward motion of the stripes of a barberpole).

5.4.2 Natural Image Sequences

The final set of examples includes contours that have been extracted from natural images. For the first two examples, the motion was constructed artificially. That is, a single natural image was translated or rotated to produce a second image. In this case, the known velocity field allows a rigorous evaluation of the correctness of the resulting solution. The final example is a sequence of aerial photographs. Here, we rely on a qualitative assessment of the results.

1. Pure translation of a curve from a natural image

The image of Figure 17(a) was translated to the right by three picture elements, to yield a second image. The two images were convolved with a $\nabla^2 G$ operator, whose central positive diameter was 12 picture elements, and the zero-crossings, shown in Figure 17(b), were obtained. The time derivative, $\frac{\partial}{\partial t}(\nabla^2 G * I)$ was computed by subtracting the two filtered images. At the location of a zero-crossing, $\mathbf{u}^{\perp}$ was computed from the gradient of the filtered image, and $v^{\perp}$ was computed using Equation (13).

The analysis of the velocity field is presented for the zero-crossing contours extracted from

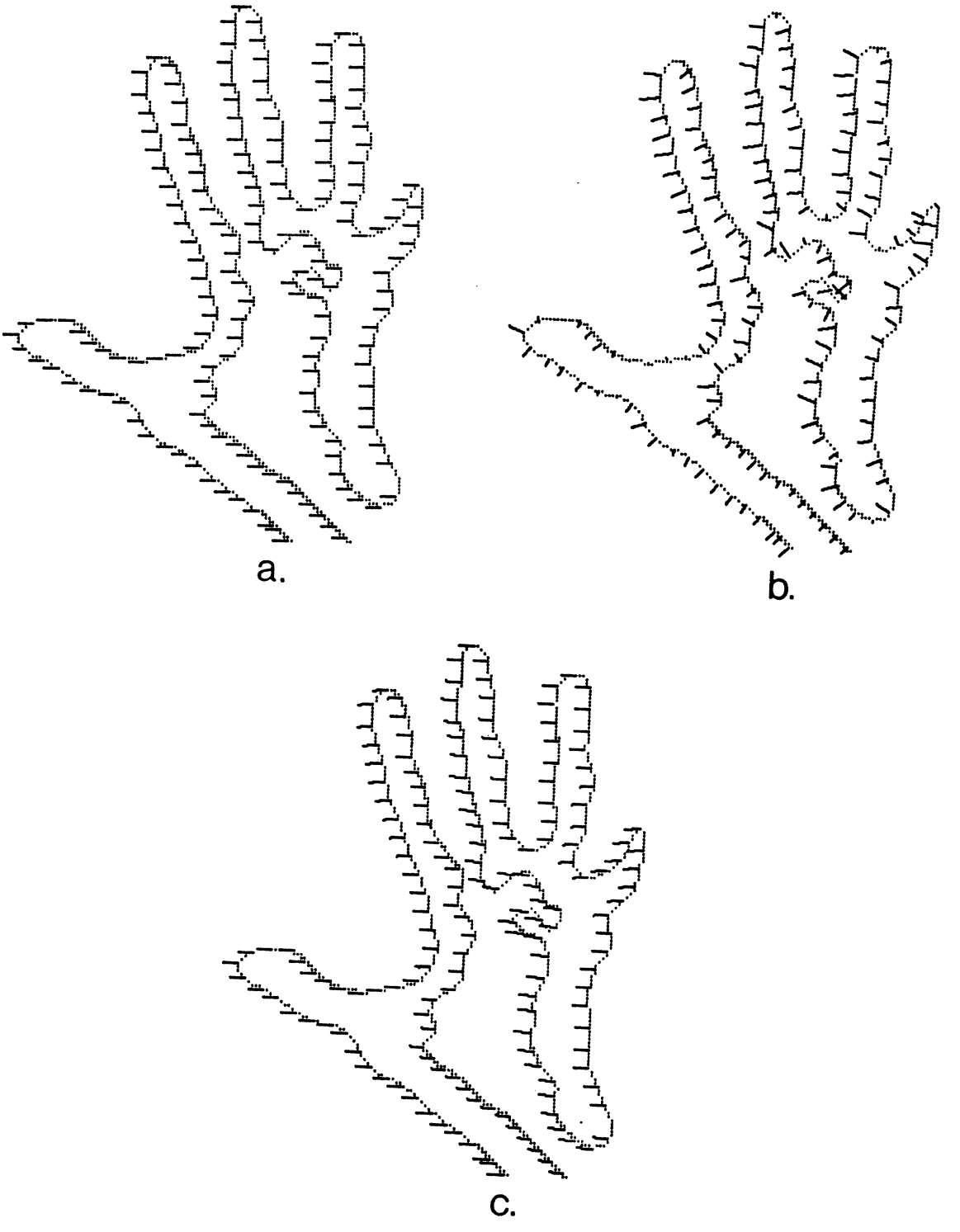


Figure 18. Translating contours. (a) Zero-crossing contours, shown with the true velocity field. (b) The initial perpendicular components of velocity. (c) The computed velocity field.

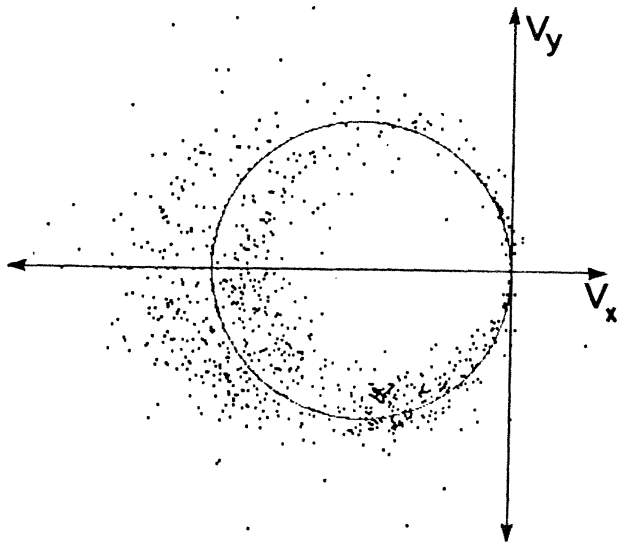


Figure 19. Constraints in velocity space. The vectors $v^{\perp}\mathbf{v}^{\perp}$, measured along the contours in Figure 18, are indicated as points in velocity space.

the area outlined by the rectangle in Figure 17(a). The contours, together with the true velocity field, are shown in Figure 18(a). The initial perpendicular velocities are shown in Figure 18(b). In Figure 19, the perpendicular vectors $v^{\perp}\mathbf{u}^{\perp}$ along the contours are displayed as points in velocity space. If the initial measurements were perfect, the points in velocity space would lie along the solid circle shown. The deviation of the position of the points from this circle provides a visual demonstration of the large error in the initial motion measurements. The resulting velocity field solution, with $\beta = 0.001$, is shown in Figure 18(c). The average error in the direction of velocity is 2.2°, and average error in magnitude is 2.8% of the true magnitude of velocity. The directions of all resulting velocity vectors are within 5° of the true direction of velocity. It is significant that a velocity field solution can be obtained, for which the error in velocity is so small, given the large error in the perpendicular components of velocity that formed the input to the algorithm. This error in the input underscores the need to design an algorithm that only requires the velocity field to satisfy the image constraints approximately.

2. Rotation of a curve from a natural image

For the next example, the original image was rotated rigidly to obtain a second image. Figure 20(a) shows the zero-crossing contour that is derived from the rectangle of Figure 17(a), and its true velocity field. The original $\nabla^2 G$ operator had a central positive diameter of 16 picture

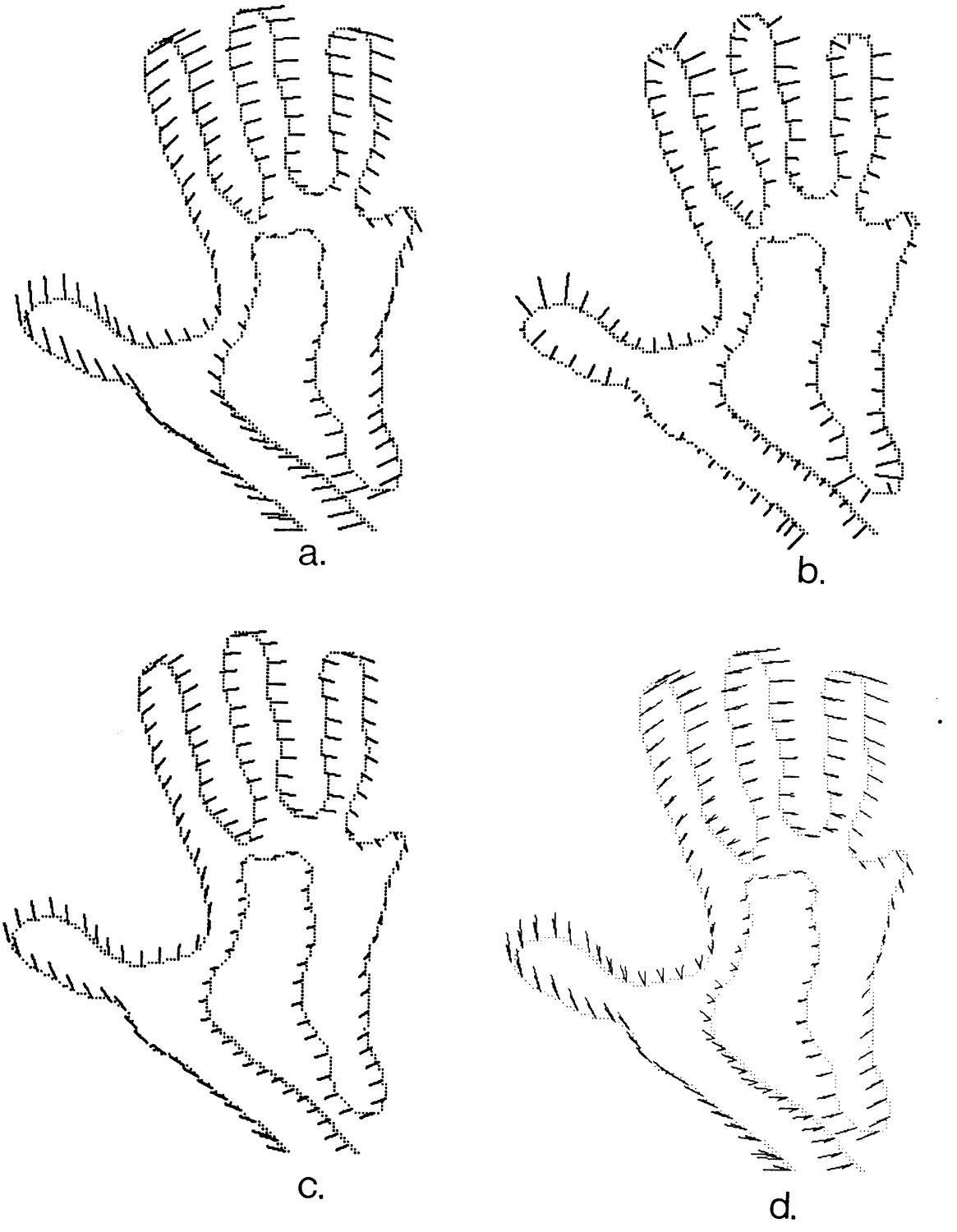


Figure 20. Rotating contour. (a) A zero-crossing contour, shown with its true velocity field. (b) The initial perpendicular components of velocity. (c) The computed velocity field. (d) The true and computed velocity fields superimposed.

elements. The image was rotated 3° in the clockwise direction, about a point in the center of the palm of the hand. (The length of the displacement vectors in Figure 20 is slightly exaggerated.) Again, the initial perpendicular components of velocity were computed from the time derivative between the two filtered images, and the gradient along the zero-crossing contours of the first filtered image; these components are illustrated in Figure 20(b). The computed smoothest velocity field, with $\beta = 0.001$, is shown in Figure 20(c). The true and computed velocity fields are shown superimposed in Figure 20(d). The directions of velocity for the true and computed motions agree quite well around the fingers, thumb, and rightmost boundary of the hand. Around the palm of the hand, and the contour running from the thumb to the forefinger, there is considerable error in the direction of velocity. Error in the area of the palm of the hand may be due to a reduced contrast along the contour. The average error in the direction of velocity over the entire contour is 6.1°. The magnitudes of velocity are generally smaller for the smoothest velocity field; average error in magnitude is 13.1% of the true magnitude of velocity. The solution varies with the choice of β; a larger value of β leads to better agreement in the magnitudes of velocity between the true and smoothest velocity fields, but the error in direction of velocity increases.

In general, the velocity field of least variation, for smooth curves in rotation, is expected to differ from the true velocity field. This is true for rotation in the image, as well as rotation in space, and was demonstrated for a number of ideal smooth curves in the previous section. The error in the velocity field of Figure 20 is due in part to error in the initial motion measurements, and in part to the fact that even with ideal input, we do not expect to obtain the correct velocity field in this case.

3. A natural motion sequence

Figures 21(a) and 21(b) show two aerial photographs, taken in sequence from an airplane. The images contain 256 × 256 picture elements. The two images were convolved with a $\nabla^2 G$ operator whose central diameter was 12 picture elements. The resulting zero-crossings, for the two images, are shown in Figures 22(a) and 22(b). The contours cover a large extent of the image in this case; this is a consequence of the small size of the image, and large displacements, which required a large initial $\nabla^2 G$ operator. If the image were analyzed with higher spatial and temporal resolution, the two-dimensional velocity field could also be obtained with higher resolution. In Figure 23, the two zero-crossing descriptions are shown superimposed, in order to illustrate the relative displacements between the two images. The zero-crossings from the first image are shown in white, those from the second are black, and the background is grey. Points at which zero-crossings occurred in both images are shown in white. Qualitatively, it can be seen

Figure 21. A natural image sequence. (a) and (b) Two natural images, containing 256 × 256 picture elements, taken in sequence from an airplane.

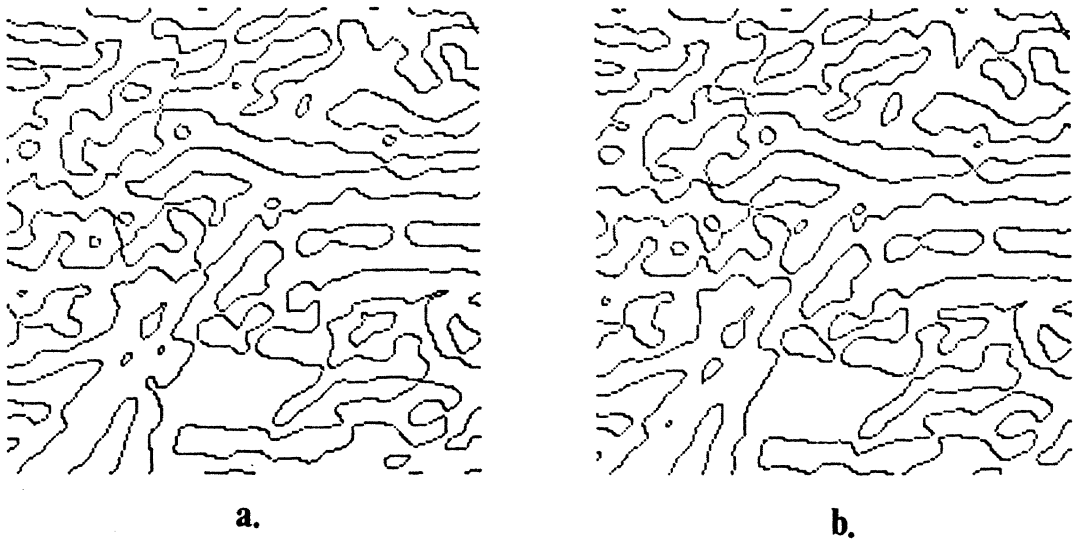


Figure 22. Initial zero-crossing descriptions, derived from the images in Figure 21.

that displacement increases in magnitude as we move from the top to the bottom of the image. In addition, the displacements vary in direction over the image, having a larger horizontal component

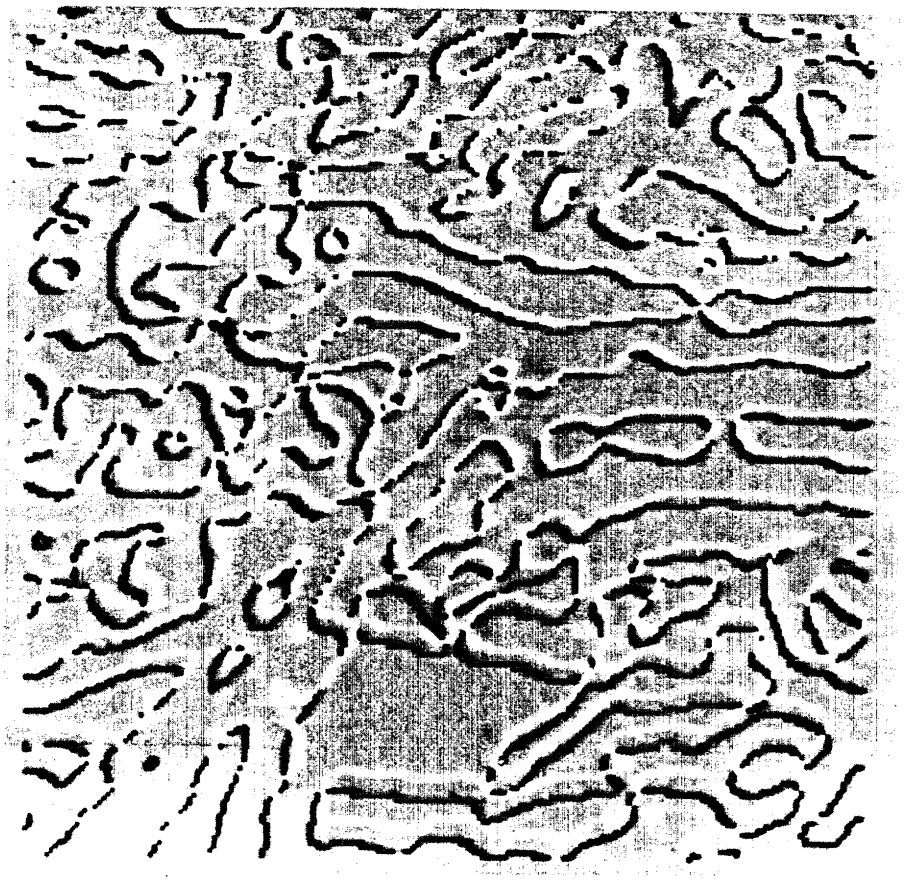


Figure 23. Superimposed zero-crossings. The zero-crossings of Fig. 22(a) are shown in white, and those of Fig. 22(b) are shown superimposed in black.

toward the top of the image, and a larger vertical component toward the bottom.

As before, the initial perpendicular components of velocity were computed along the zero-crossing contours in the first filtered image by comparing the local gradient at the zero-crossings with the time derivative, obtained by subtracting the two filtered images. The algorithm was then run over each contour separately, to obtain the velocity field along the zero-crossing contours of the first image. In Figure 24, the zero-crossings are shown in black, with the resulting velocity vectors in white. The vectors shown represent a sampling of the velocity field along the contours. Very small contours, and some of the contours that occurred at the image boundary, were not included in this analysis. The length of the vectors in this display is equal to their displacement in the image. Qualitatively, the magnitude of the displacement vectors increases from the top to the bottom of the image, and their directions also vary in a way that reflects the displacement of the zero-crossing contours in Figure 23.

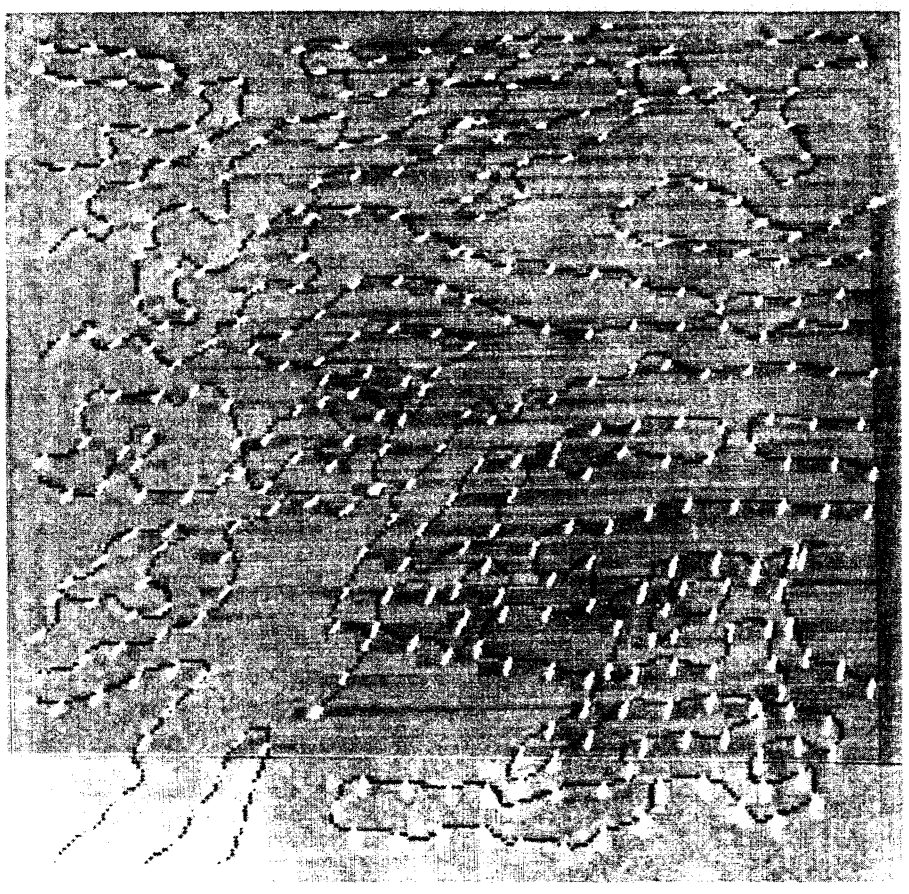


Figure 24. The computed velocity field. The zero-crossing contours are shown in black, and a sampling of the computed velocity field is shown in white.

In Figure 1, shown in Section 2, the first image of this sequence, reduced in contrast, was displayed with the velocity field superimposed with black lines. The velocity vectors are evenly spaced on the image, and were computed by taking the average of the velocity vectors within a neighborhood of size 48×48 picture elements, centered at each point. The length of the velocity vectors was doubled in this display, in order to emphasize the variation in direction and magnitide of velocity over the image. The resulting velocity field agrees qualitatively with the displacements of the contours shown in Figure 23. It also agrees well with the perceived movement between the images, when viewed in rapid succession.

6. CONCLUSIONS

In this paper, we have examined the computation of the two-dimensional velocity field that results from the changing projection of three-dimensional surfaces in motion. A theory for

this computation was developed, with three main components. First, initial measurements of motion in the image take place at the location of significant intensity changes, which give rise to zero-crossings in the output of the convolution of the image with a $\nabla^2 G$ operator. The initial motion measurements provide the component of velocity in the direction perpendicular to the local orientation of the zero-crossing contours. Second, these initial measurements are integrated along contours to compute the two-dimensional velocity field. Third, an additional constraint of smoothness of the velocity field, based on the physical constraint of smoothness of surfaces, allows the computation of a unique velocity field.

There are three aspects of this work that are of fundamental importance. First, the additional constraint for the velocity field computation that was formulated here allows the analysis of general motion, while providing a velocity field solution that is unique, and physically plausible. Second, this formulation of the computation leads naturally to algorithms that are biologically feasible, in that they require only simple, local, parallel operations. Finally, the computation appears to yield solutions that are consistent with human motion perception, and therefore may contribute toward our understanding of the measurement of motion in the human visual system.

ACKNOWLEDGEMENT

I thank Shimon Ullman for valuable supervision of this work.

APPENDIX A. EVALUATING OTHER SMOOTHNESS MEASURES

In this appendix, we first present an example for which the measures of variation given by $\int |\frac{\partial \varphi}{\partial s}| ds$ (variation in direction) and $\int \frac{\partial |\mathbf{V}|}{\partial s} ds$ (variation in magnitude) do not yield unique solutions. We then present an example for which a velocity field that minimizes $\int |\frac{\partial \mathbf{V}}{\partial s}| ds$ does not exist. In both cases, the formulation of velocity constraints in velocity space is used to analyze the behavior of the smoothness measures.

1. A Line Segment in Motion

Consider a simple line segment that moves rigidly in the image plane, as shown in Figure 25(a). Let us assume that the perpendicular component of velocity along the line, $v^{\perp}(s)$, and the velocities of the two endpoints, $\mathbf{V}(s_0)$ and $\mathbf{V}(s_n)$, are known. The velocity of a particular point, $\mathbf{V}(s_i)$ on the line is constrained to project to a line L_i in velocity space, whose orientation is equal to the orientation of the line in the image, and whose perpendicular distance from the origin in velocity space is given by $v^{\perp}(s_i)$. This constraint is illustrated in Figure 25(b). A set of evenly spaced points along the line yields a set of constraint lines in velocity space, which are evenly spaced from the origin, as shown in Figure 25(c). In addition, the velocity vectors associated with the endpoints are shown in velocity space.

The true velocity field for the moving line segment corresponds to the straight line l in Figure 26(a), connecting the two known velocity vectors. The velocity of a point on the line segment, $\mathbf{V}(s_i)$, is given by the intersection of l and L_i. Any valid velocity field must correspond to a path in velocity space that begins at the endpoint of $\mathbf{V}(s_0)$, intersects each constraint line in the order that they arise along the image curve, and ends at the endpoint of $\mathbf{V}(s_n)$. Three examples of valid velocity fields correspond to the curves shown in Figure 26(b). Computation of the velocity field in this case involves finding the valid velocity field that exhibits the least variation, given one of the measures of variation.

Suppose that we want to find the velocity field that is consistent with the constraints shown in Figure 25(c), and minimizes the total change in direction of velocity, given by: $\int |\frac{\partial \varphi}{\partial s}| ds$. For the case of the true velocity field, φ increases monotonically; its total change is given by the angle α, shown in Figure 26(a); hence $\int |\frac{\partial \varphi}{\partial s}| ds = \alpha$. α is the smallest total variation in direction that a velocity field consistent with these constraints may obtain. There are other velocity fields, however, that distribute the total change α over the curve monotonically and smoothly, in a way that is consistent with $v^{\perp}(s)$. Other examples correspond to the curved paths in Figure 26(b); for these velocity fields, it is also the case that $\int |\frac{\partial \varphi}{\partial s}| ds = \alpha$. Hence, in the simple case of a line segment

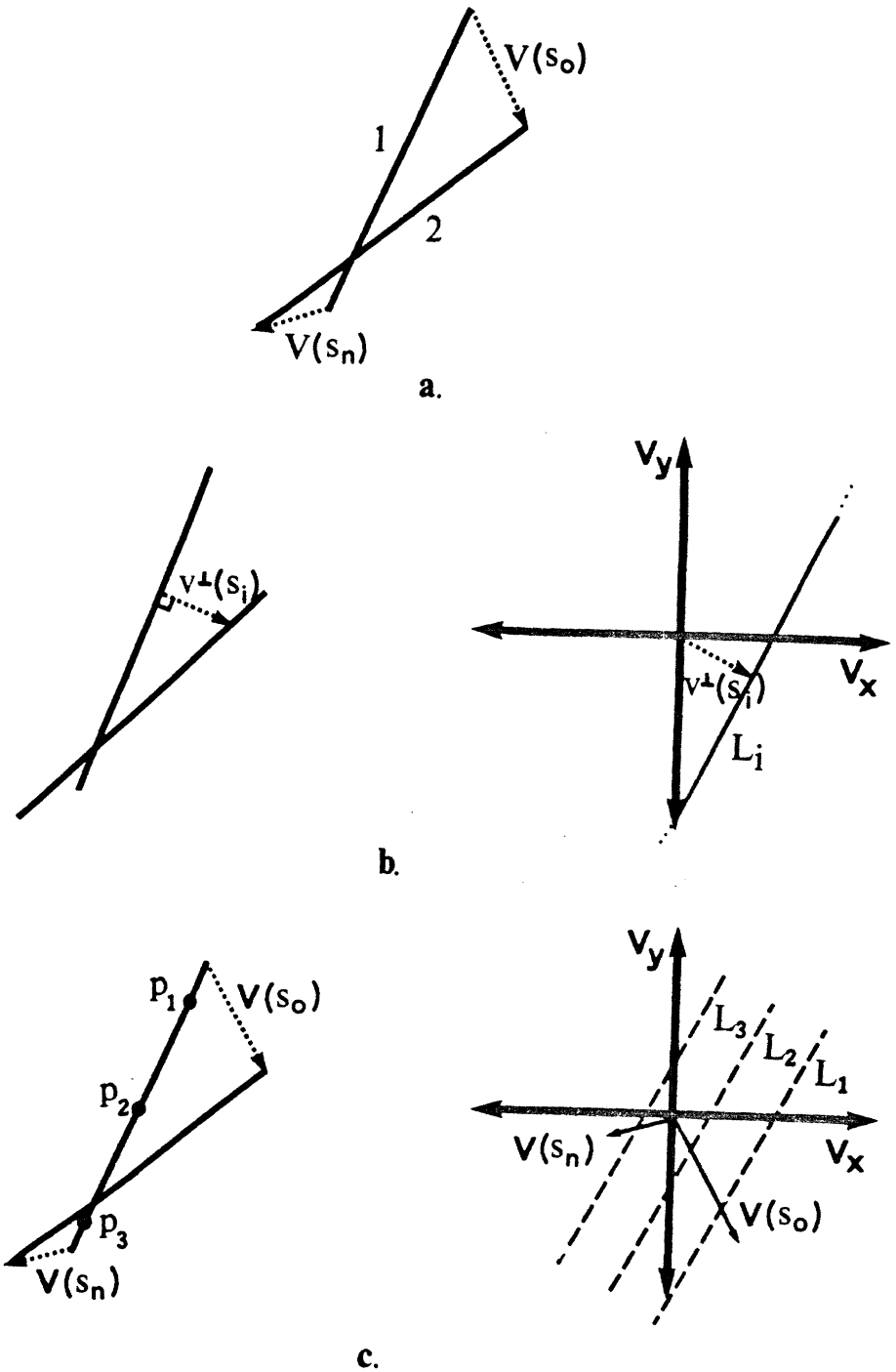


Figure 25. A line segment in motion. (a) A line segment moving rigidly in the image, with known velocity vectors at its endpoints. (b) The constraint imposed by a single measurement of $v^{\perp}(s)$ (shown in the image and in velocity space). (c) The constraints imposed by muluple measurements of $v^{\perp}(s)$ along the line segment.

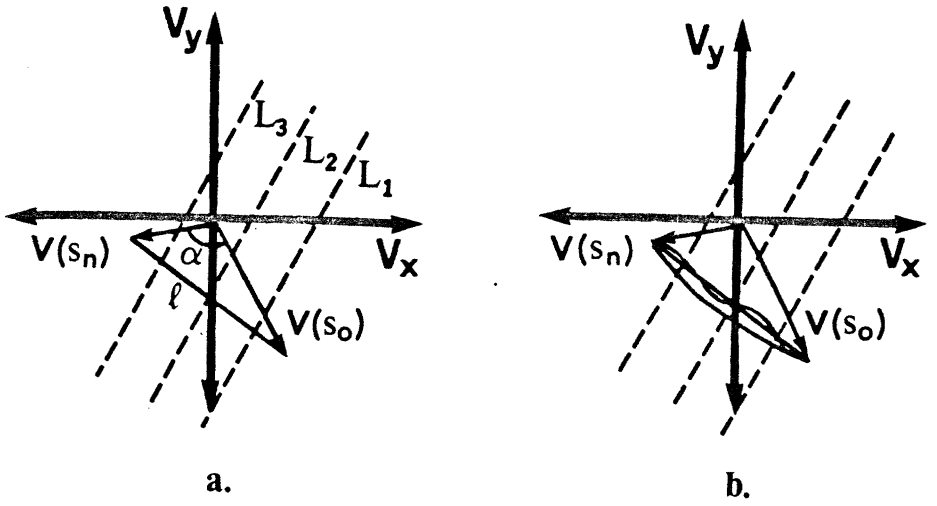


Figure 26. Velocity fields for a line segment in motion. (a) The true velocity field in the image projects to the line l in velocity space. (b) The curved paths correspond to other valid velocity field solutions.

moving rigidly in the image plane, minimizing the total change in direction of velocity alone does not yield a unique velocity field solution.

Now consider minimizing the variation in magnitude of velocity, given by the measure: $\int \frac{\partial |\mathbf{V}|}{\partial s} ds$. Let m_1 and m_2 denote the magnitudes of the velocity vectors associated with the endpoints. It then follows that $\int \frac{\partial |\mathbf{V}|}{\partial s} ds \geq |m_1 - m_2|$. Any velocity field that is consistent with the image constraints, and changes monotonically in magnitude along the line will yield $\int \frac{\partial |\mathbf{V}|}{\partial s} ds = |m_1 - m_2|$. This solution is again not unique.

2. A Rotating Semi-circle

As a second example, consider a semi-circle rotating about its center, as shown in Figure 27(a). Assume that the velocity vectors $\mathbf{V}(s_0)$ and $\mathbf{V}(s_n)$ at the endpoints are known. Along the contour, $v^{\perp}(s) = 0$, so these measurements give rise to lines of constraint in velocity space that intersect at the origin, as shown in Figure 27(b). The known velocity vectors at the endpoints are also shown in Figure 27(b). Consider the measure of variation given by: $\int |\frac{\partial \mathbf{V}}{\partial s}| ds$. As before, a valid velocity field must correspond to a path in velocity space that begins at the endpoint of $\mathbf{V}(s_0)$, intersects the lines of constraint in the order that they arise from the image curve, and ends at the endpoint of $\mathbf{V}(s_n)$. The value of the measure $\int |\frac{\partial \mathbf{V}}{\partial s}| ds$ is equal to the length of the path in velocity space. Thus, a velocity field that minimizes this measure must give rise to a path of minimum length in velocity space. For this particular example, however, a minimum length path satisfying the constraints does not exist. In Figure 27(c), we show three paths in velocity space, labeled

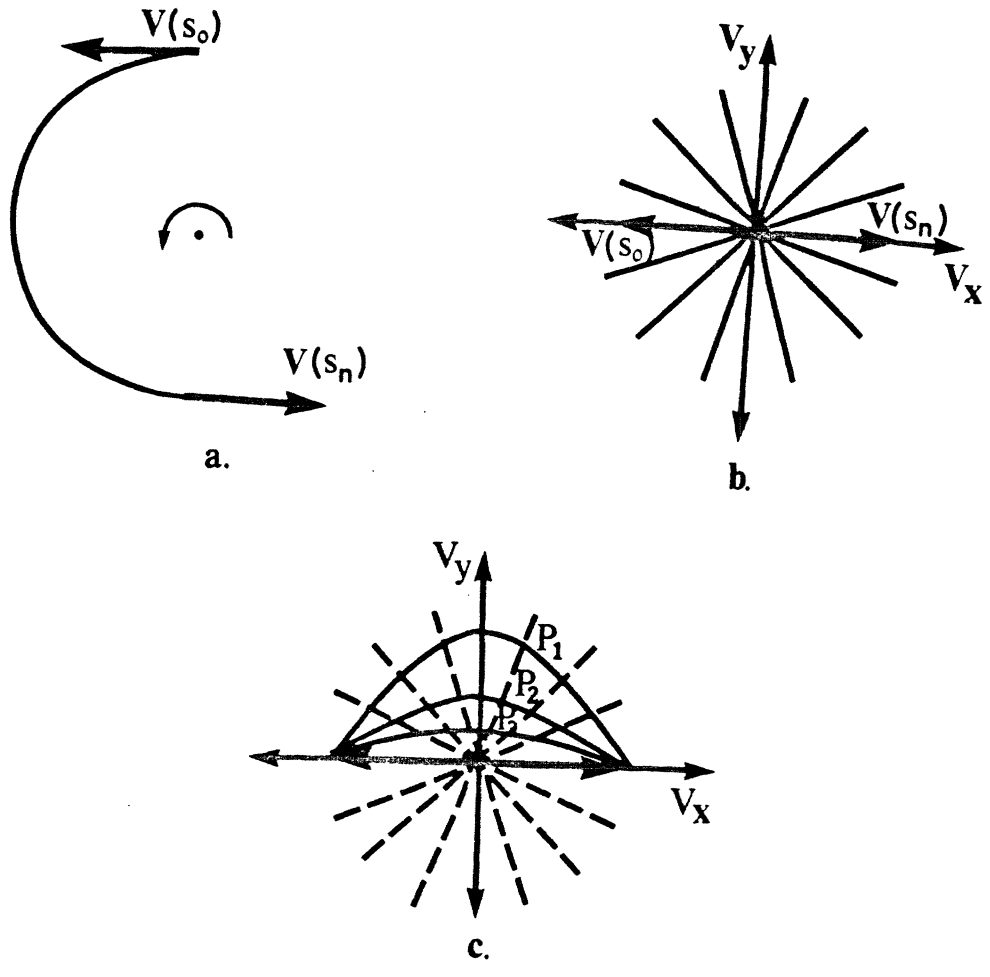


Figure 27. A rotating semi-circle.(a) A semi-circle rotating in the image about its center, with known velocity vectors at its endpoints. (b) The constraint imposed by the measurements of $v^{\perp}(s)$, shown in velocity space. (c) Possible paths corresponding to valid velocity fields.

P_1, P_2 and P_3, that correspond to valid velocity fields. As we move from P_1 to P_3, the length of the path decreases. Thus, the value of the measure $\int |\frac{\partial \mathbf{V}}{\partial s}| ds$ decreases, for the corresponding velocity fields. We can continue indefinitely, however, to find paths of shorter length that satisfy the constraints. The paths will approach the horizontal line, but can never reach it exactly, because the horizontal line does not correspond to a valid velocity field. Thus, a velocity field of least variation, given this particular measure of variation, does not exist.

APPENDIX B. THE CONJUGATE GRADIENT ALGORITHM

In Section 5, it was shown that the velocity field computation, in which the constraints imposed by the measurements of $v^{\perp}(s)$ are only approximately satisfied, can be formulated as an unconstrained optimization problem, and the conjugate gradient algorithm can be used to obtain a solution. The computation finds the x and y components of velocity for each of the n points on a contour. Each possible solution can be considered as a point in a $2n$-dimensional vector space, with each dimension representing the x or y component of velocity at one point on the contour. The $2n$-dimensional vector space can be embedded in a $(2n+1)$-dimensional vector space, where the final dimension corresponds to the total variation of velocity along the contour. By associating a measure of variation with each possible velocity field, we can construct a *hypersurface*, called the *objective surface*, in the $(2n+1)$-dimensional space. The goal of the velocity field computation is to find the "lowest" point on this surface; this point corresponds to the velocity field of least variation.

The conjugate gradient algorithm is an iterative descent algorithm; a sequence of approximations $\mathbf{V}^{(1)}, \mathbf{V}^{(2)}, \ldots$ to the exact solution $\mathbf{V}$ is computed, given an initial approximation $\mathbf{V}^{(0)}$, and this sequence has the property that each new approximation decreases the value of the objective function. The basic steps for a descent method are as follows:

(1) Start at an initial point $\mathbf{V}^{(0)}$, which is usually $\mathbf{V}^{(0)} = \mathbf{0}$.

(2) According to a fixed rule, determine a direction of movement along the objective surface that will reduce the value of the objective function.

(3) Move in this direction to a (relative) minimum of the objective function.

(4) If the final solution has not yet been reached, return to step 2.

One of the most common descent methods is that of *steepest descent*, in which the direction of movement in step 2 above is given by the negative gradient of the objective function, evaluated at the present point. The steepest descent method has the disadvantage of relatively slow convergence. A much improved rate of convergence can be obtained with the method of conjugate gradients.

The velocity field of least variation is given by the solution of a system of linear equations. Let this system be denoted in matrix form by:

$$\mathbf{Qx} = \mathbf{b}$$

where $\mathbf{Q}$ is an $n \times n$ symmetric positive definite matrix. The unique solution to this equation is equivalent to the unique solution of the quadratic problem:

$$\text{minimize} \qquad \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{b}^T\mathbf{x}$$

$$\text{subject to} \qquad \mathbf{x} \in \Omega \subset E^n.$$

E^n denotes the n-dimensional Euclidean space. Conjugate direction methods were invented largely for the analysis of quadratic problems. We present a brief development of these methods, and the conjugate gradient algorithm. This development is taken from Luenberger (1973), where further details may be found.

The basic idea behind conjugate direction methods is to optimize the direction in which each new step is taken in the algorithm. To show how this optimization is achieved, we begin by defining the notion of Q-orthogonality:

> **Definition.** *Given a symmetric matrix* $\mathbf{Q}$, *two vectors* $\mathbf{d}_1$ *and* $\mathbf{d}_2$ *are said to be* $\mathbf{Q}$*-orthogonal, or conjugate, with respect to* $\mathbf{Q}$ *if* $\mathbf{d}_1^\top \mathbf{Q}\mathbf{d}_2 = 0$.

The following can then be shown:

> **Proposition.** *If* $\mathbf{Q}$ *is positive definite and the set of nonzero vectors* $\mathbf{d}_0$, $\mathbf{d}_1$,...,$\mathbf{d}_k$ *are* $\mathbf{Q}$*-orthogonal, then these vectors are linearly independent.*

Given the positive definite matrix $\mathbf{Q}$, and a set of non-zero $\mathbf{Q}$-orthogonal vectors, $\mathbf{d}_0, \mathbf{d}_1, \ldots \mathbf{d}_{n-1}$, their linear independence allows us to express the solution $\mathbf{x}'$ in terms of these vectors:

$$\mathbf{x}' = \alpha_0\mathbf{d}_0 + \cdots + \alpha_{n-1}\mathbf{d}_{n-1}$$

for some set of scalars α_i. The orthogonality of the vectors implies that multiplying by $\mathbf{Q}$ and taking the scalar product with $\mathbf{d}_i$ causes all the terms, except the i^{th} to vanish, yielding:

$$\alpha_i = \frac{\mathbf{d}_i^\top \mathbf{Q}\mathbf{x}'}{\mathbf{d}_i^\top \mathbf{Q}\mathbf{d}_i} = \frac{\mathbf{d}_i^\top \mathbf{b}}{\mathbf{d}_i^\top \mathbf{Q}\mathbf{d}_i}.$$

The solution is then given by the following:

$$\mathbf{x}' = \sum_{i=0}^{n-1} \frac{\mathbf{d}_i^\top \mathbf{b}}{\mathbf{d}_i^\top \mathbf{Q}\mathbf{d}_i}\mathbf{d}_i.$$

The expansion for the solution $\mathbf{x}'$ can be considered as the result of an iterative process of n steps, where at the i^{th} step, the term $\alpha_i\mathbf{d}_i$ is added. In this way, the following result can be proven (Luenberger 1973):

Conjugate Direction Theorem: *Let* $\{\mathbf{d}_i\}_{i=0}^{n-1}$ *be a set of non-zero* $\mathbf{Q}$*-orthogonal vectors. For any* $\mathbf{x}_0 \in E^n$ *the sequence* $\{\mathbf{x}_k\}$ *generated according to*

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \qquad k \leq 0$$

with

$$\alpha_k = -\frac{\mathbf{g}_k^\top \mathbf{d}_k}{\mathbf{d}_k^\top \mathbf{Q}\mathbf{d}_k}$$

and

$$\mathbf{g}_k = \mathbf{Q}\mathbf{x}_k - \mathbf{b}$$

converges to the unique solution, $\mathbf{x}'$*, of* $\mathbf{Q}\mathbf{x} = \mathbf{b}$ *after* n *steps.*

The conjugate gradient algorithm provides one method for generating the sequence of $\mathbf{Q}$-orthogonal direction vectors; the successive direction vectors are selected as a conjugate version of the successive gradients obtained as the method proceeds. Hence, at each step, the next direction vector is determined, based on the current state of the objective function and its gradient vector. The algorithm, which is guaranteed to converge in at most n steps, is outlined below:

The Conjugate Gradient Algorithm

Starting at any point $\mathbf{x}_0 \in E^n$ define $\mathbf{d}_0 = -\mathbf{g}_0 = \mathbf{b} - \mathbf{Q}\mathbf{x}_0$ and

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$$

$$\alpha_k = -\frac{\mathbf{g}_k^T \mathbf{d}_k}{\mathbf{d}_k^T \mathbf{Q}\mathbf{d}_k}$$

$$\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_k \mathbf{d}_k$$

$$\beta_k = \frac{\mathbf{g}_{k+1}^T \mathbf{Q}\mathbf{d}_k}{\mathbf{d}_k^T \mathbf{Q}\mathbf{d}_k}$$

where $\mathbf{g}_k = \mathbf{Q}\mathbf{x}_k - \mathbf{b}$.

For the velocity field computation, we showed in Section 5.2 that the system of $2n$ linear equations to be solved is given by:

$$\frac{\partial \Phi_2}{\partial \mathrm{V}_{x_i}} = 0, \qquad \frac{\partial \Phi_2}{\partial \mathrm{V}_{y_i}} = 0, \qquad 1 \leq i \leq n$$

where Φ_2 is given by:

$$\Phi_2 = \sum_{i=2}^{n} [(\mathrm{V}_{x_i} - \mathrm{V}_{x_{i-1}})^2 + (\mathrm{V}_{y_i} - \mathrm{V}_{y_{i-1}})^2] + [(\mathrm{V}_{x_1} - \mathrm{V}_{x_n})^2 + (\mathrm{V}_{y_1} - \mathrm{V}_{y_n})^2]$$

$$+\beta\sum_{i=1}^{n}\left[\mathbf{V}_{x_i}\mathbf{u}_{x_i}^{\perp}+\mathbf{V}_{y_i}\mathbf{u}_{y_i}^{\perp}-v_i^{\perp}\right]^2.$$

From the partial derivatives with respect to the x components of velocity, we have the equations:

$$(4+2\beta(\mathbf{u}_{x_i}^{\perp})^2)\mathbf{V}_{x_i}-2\mathbf{V}_{x_{i+1}}-2\mathbf{V}_{x_{i-1}}+2\beta\mathbf{u}_{x_i}^{\perp}\mathbf{u}_{y_i}^{\perp}\mathbf{V}_{y_i}=2\beta v_i^{\perp}\mathbf{u}_{x_i}^{\perp},\qquad 1\leq i\leq n.$$

From the partial derivatives with respect to the y components of velocity, we have:

$$(4+2\beta(\mathbf{u}_{y_i}^{\perp})^2)\mathbf{V}_{y_i}-2\mathbf{V}_{y_{i+1}}-2\mathbf{V}_{y_{i-1}}+2\beta\mathbf{u}_{x_i}^{\perp}\mathbf{u}_{y_i}^{\perp}\mathbf{V}_{x_i}=2\beta v_i^{\perp}\mathbf{u}_{y_i}^{\perp},\qquad 1\leq i\leq n.$$

The coefficients on the left hand side of the equations form the symmetric positive definite matrix **Q**, and the values on the right hand side form the column vector **b**. It should be noted that for the case of an extended straight line, **Q** is singular, so that a unique solution does not exist. The initial velocity field consists of the normal velocity vectors; hence:

$$\mathbf{x}_0=(v_1^{\perp}\mathbf{u}_{x_1}^{\perp},v_2^{\perp}\mathbf{u}_{x_2}^{\perp},\ldots,v_n^{\perp}\mathbf{u}_{x_n}^{\perp},v_1^{\perp}\mathbf{u}_{y_1}^{\perp},v_2^{\perp}\mathbf{u}_{y_2}^{\perp},\ldots,v_n^{\perp}\mathbf{u}_{y_n}^{\perp}).$$

The conjugate gradient algorithm is then applied straightforwardly.

REFERENCES

Adelson, E. H. and Movshon, J. A., Phenomenal coherence of moving visual patterns, *Nature* **300** (1982) 523-525.

Adelson, E. H. and Movshon, J. A., The perception of coherent motion in two-dimensional patterns, *Proc. ACM Interdisciplinary Workshop on Motion*, Toronto (1983) 11-14.

Aggarwal, J. K. and Duda, R. O., Computer analysis of moving polygonal images, *IEEE Trans. Comp.* **c-24** (1975) 966-976.

Anstis, S. M., The perception of apparent motion, *Phil. Trans. Roy. Soc. Lond. B.* **290** (1980) 153-168.

Anstis, S. M., Visual coding of position and motion, in *Physical and Biological Processing of Images*, O.J. Braddick and A.C. Sleigh (Eds.) (Springer-Verlag, Berlin, 1983).

Barnard, S. T. and Thompson, W. B., Disparity analysis of images, *IEEE Trans. Patt. Anal. Mach. Intell.* **2** (1980) 333-340.

Braddick, O. J., A short-range process in apparent motion, *Vis. Res.* **14** (1974) 519-527.

Braddick, O. J., Low-level and high-level processes in apparent motion, *Phil. Trans. Roy. Soc. Lond. B.* **290** (1980) 137-151.

Bruss, A. and Horn, B. K. P., Passive navigation, *Comp. Vis. Graph. Image Proc.* **21** (1983) 3-20.

Cafforio, C. and Rocca, F., Methods for measuring small displacements of television images, *IEEE Trans. Inform. Theory* **IT-22** (1976) 573-579.

Chow, W. K. and Aggarwal, J. K., Computer analysis of planar curvilinear moving images, *IEEE Trans. Comp.* **c-26** (1977) 179-185.

Clocksin, W. F., Perception of surface slant and edge labels from optical flow: a computational approach, *Percept.* **9** (1980) 253-269.

Davis, L., Wu, Z. and Sun, H., Contour-based motion estimation, *Comp. Vis. Graph. Image Proc.* **23** (1983) 313-326.

Fennema, C. I. and Thompson, W. B., Velocity determination in scenes containing several moving objects, *Comp. Graph. Image Proc.* **9** (1979) 301-315.

Fukinuki, T., Yoshigi, H. and Fukushima, K., Improvement of inter-frame predictive coding of TV-signals by utilizing visual properties for moving objects, *Trans. IECE Japan* **59-A** (1976) 764-771. [Japanese]

Grimson, W. E. L., *From Images to Surfaces. A Computational Study of the Human Early Visual System* (MIT Press, Cambridge, 1981).

Hartenberg, R. S. and Denavit, J., *Kinematic Synthesis of Linkages* (McGraw-Hill, New York, 1964).

Hildreth, E. C., *The Measurement of Visual Motion* (MIT Press, Cambridge, 1984a).

Hildreth, E. C., The computation of the velocity field, *Proc. Roy. Soc. Lond. B.* (1984b) in press.

Holland, H. C., *The Spiral After-Effect* (Pergamon Press, Oxford, 1965).

Horn, B. K. P. and Schunck, B. G., Determining optical flow, *Artificial Intelligence* **17** (1981) 185-203.

Johansson, G., Visual perception of biological motion and a model for its analysis, *Percept. and Psych.* **14** (1973) 201-211.

Johansson, G., Visual motion perception, *Sci. Am.* **232** (1975) 76-88.

Lappin, J. S. and Bell, H. H., The detection of coherence in moving random dot patterns, *Vis. Res.* **16** (1976) 161-168.

Lawton, D. T., Processing translational motion sequences, *Comp. Vis. Graph. Image Proc.* **22** (1983) 116-144.

Leese, J. A., Novak, C. S. and Taylor, V. R., The determination of cloud pattern motion from geosynchronous satellite image data, *Patt. Recog.* **2** (1970) 279-292.

Lillestrand, R. L., Techniques for change detection, *IEEE Trans. Comp.* **c-21** (1972) 654-659.

Limb, J. O. and Murphy, J. A., Estimating the velocity of moving objects from television signals, *Comp. Graph. Image Proc.* **4** (1975) 311-327.

Longuet-Higgins, H. C., A computer algorithm for reconstructing a scene from two projections, *Nature* **293** (1981) 133-135.

Longuet-Higgins, H. C. and Prazdny, K., The interpretation of moving retinal images, *Proc. Roy. Soc. Lond. B.* **208** (1981) 385-397.

Luenberger, D. G., *Introduction to Linear and Nonlinear Programming* (Addison-Wesley, Reading, MA, 1973).

Marr, D., *Vision* (W. H. Freeman Co., San Francisco, 1982).

Marr, D. and Hildreth, E. C., Theory of edge detection, *Proc. Roy. Soc. Lond. B.* **207** (1980) 187-217.

Marr, D. and Poggio, T., A computational theory of human stereo vision, *Proc. Roy. Soc. Lond. B.* **204** (1979) 301-328.

Marr, D. and Ullman, S., Directional selectivity and its use in early visual processing, *Proc. Roy. Soc. Lond. B.* **211** (1981) 151-180.

Mutch, K. M. and Thompson, W. B., Hierarchical estimation of spatial properties from motion, Univ. Minnesota Computer Science Department Tech. Rep. 82-22, 1982.

Nagel, H.-H., On change detection and displacement vector estimation in image sequences, *Patt. Recog. Letters 1* (1982) 55-59.

Petersik, J. T., Hicks, K. I. and Pantle, A. J., Apparent movement of successively generated subjective patterns, *Percept.* 7 (1978) 371-383.

Potter, J. L., Velocity as a cue to segmentation, *IEEE Trans. Sys., Man, Cyb.* **SMC-5** (1975) 390-394.

Potter, J. L., Scene segmentation using motion information, *Comp. Graph. Image Proc.* **6** (1977) 558-581.

Prazdny, K., Egomotion and relative depth map from optical flow, *Biol. Cyb.* **36** (1980) 87-102.

Reichardt, W., Autocorrelation, a principle for the evaluation of sensory information by the central nervous system, in: *Sensory Communication*, W. A. Rosenblith (ed.) (MIT Press, Cambridge, 1961).

Roach, J. W. and Aggarwal, J. K., Computer tracking of objects moving in space, *IEEE Trans. Patt. Anal. Mach. Intell.* **PAMI-1** (1979) 127-135.

Rudin, W., *Functional Analysis* (McGraw-Hill Co., New York, 1973).

Smith, E. A. and Phillips, D. R., Automated cloud tracking using precisely aligned digital ATS pictures, *IEEE Trans. Comp.* **c-21** (1972) 715-729.

Thompson, W. B. and Barnard, S. T., Lower-level estimation and interpretation of visual motion, *IEEE Comp.*, August (1981) 20-28.

Ullman, S., *The Interpretation of Visual Motion* (MIT Press, Cambridge, 1979a).

Ullman, S., Relaxation and constrained optimization by local processes, *Comp. Graph. Image Proc.* **9** (1979b) 115-125.

Ullman, S., Computational studies in the interpretation of structure and motion: summary and extension, MIT A.I. Lab. Memo 706, 1983a.

Ullman, S., Maximizing rigidity: the incremental recovery of 3-D structure from rigid and rubbery motion, MIT A.I. Lab. Memo 721, 1983b.

Wallach, H., On perceived identity: 1. The direction of motion of straight lines", in: *On Perception*, H. Wallach (Ed.) (Quadrangle, New York, 1976).

Wallach, H. and O'Connell, D. N., The kinetic depth effect, *J. Exp. Psych.* **45** (1953) 205-217.

Wallach, H., Weisz, A. and Adams, P. A., Circles and derived figures in rotation, *Am. J. Psych.* **69** (1956) 48-59.

Wertheimer, M., Experimentelle studien uber das sehen von bewegung, *Zeitchrift fur psychologie* **61** (1912) 161-265.

Wilson, H. R. and Bergen, J. R., A four mechanism model for threshold spatial vision, *Vis. Res.* **19** (1979) 19-32.

Wolferts, K., Special problems in interactive image processing for traffic analysis, *Proc. Second Int. Joint Conf. Patt. Recog.* (1974) 1-2.

Yuille, A. L., The smoothest velocity field and token matching schemes, MIT A.I. Lab. Memo 724, 1983.