

Recognizing 3-D Motion

Cornelia Fermüller* and Yiannis Aloimonos

Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3411

Abstract

What is Vision? Being more precise, how should we formulate the process of vision, if our goal is to equip machines with visual capabilities? Is it suitable to follow the paradigm suggested by Marr and consider vision as the recovery or reconstruction of the shapes, material properties and motions of objects in our environment? Or, should we view the process of vision as the recognition of particular situations, objects, pattern or instances in space-time relevant to the task a machine is engaged in? We argue for the latter point of view, and we demonstrate our position by proving that the famous problem of ego-motion estimation, which in the past has been addressed as an application of applied mathematics techniques, can be formulated as a problem of recognition. We show that local image measurements form global patterns in the image plane. The position of these patterns is related to the three dimensional motion parameters. By locating through a simple search technique some of these patterns, which depend on only subsets of the motion parameters, the 3-D motion parameters are found. The proposed algorithmic procedure is very robust. Since only the sign of the image velocity is employed, the direction of translation and the axis of rotation can be estimated with up to 100 % error in the image measurements.

1 Introduction

Since computer vision is considered as one of the main subfields of Artificial Intelligence (AI)¹, it has been addressed in the first years since its appearance using the general AI-methodology. The main emphasis in AI-research has been the finding of general purpose methodologies and general purpose representations that preserve

* Department for Pattern Recognition and Image Processing, Institute for Automation, Technical University of Vienna, Treitlstraße 3, A-1040 Vienna, Austria

¹About 40 % of the human brain is devoted to visual processing.

as much information as possible. A three-step approach for solving any problem has been taken for granted. The conversion of external data (sensor data, actuator commands, decision making, etc.) into an internal representation and vice versa has been separated from the phase of developing algorithms to perform computations on internal data. Most research has been devoted on processing the internal data, and as a consequence different subfields such as planning, machine learning, knowledge-based data bases and many more appeared. Therefore it is not surprising that the first influential theory of computational vision [Marr, 1982] mainly concentrated on the computational and representational aspects. Important results have been achieved at the level of computational theories and development of algorithms dealing with internal representations. Nevertheless, it has to be admitted that vision and AI in general, are far from the ultimate goal of building a machine possessing capabilities of living organisms. Critics of the general AI-philosophy argue that the study of intelligence cannot be separated from the study of the system's interaction with its environment [Aloimonos *et al.*, 1988; Aloimonos, 1990; Bajcsy, 1988; Bajcsy and Campos, 1992; Ballard, 1991].

Taking this point of view, allows us to pose any problem in vision, or perception in general, as a problem of recognition as opposed to recovery (or reconstruction). Researchers usually taxonomize visual problems into three basic categories: problems related to object recognition, to navigation and to manipulation. The first ones are obviously problems of recognition. It can easily be demonstrated that visuomotor coordination tasks (navigation and manipulation) can be formulated as problems of recognition. For example, the problem of obstacle avoidance can be addressed through depth recovery and passive ranging techniques [Bhanu and (editors), 1992], but also through the recognition of "time-to-collision" values below a threshold, where the choice of the threshold is related to the reaction time of the agent [Nelson and Aloimonos, 1989]. Similarly, the problem of detecting independent motion by a moving observer may be addressed through techniques for segmenting optic flow fields—which work only in model-based approaches [Koller *et al.*, 1992]—or through the recognition of non-rigid flow fields for a moving observer partially knowing its motion [Aloimonos, 1990;

Nelson, 1991]. In the same way, the problem of hand/eye coordination can be dealt with using stereo and other techniques to compute the depth map and then solve the inverse kinematics problem in order to move the arm. While the arm is moving, the system is blind [Brady *et al.*, 1983]. However, the same problem can be solved through the creation of a mapping from image features to the robot's joints; then, positioning of the arm is achieved by recognizing the image features [Herve *et al.*, 1991].

In general, navigation and manipulation tasks, since they inherently involve metric relationships between the observer and the environment, have been considered as subproblems of a general visual reconstruction problem, the so called "structure from motion" problem (see next section). In this paper we consider a very general, celebrated navigational problem that has been at the core of the reconstructionist school of thought, the problem of egomotion estimation from dynamic imagery. It has usually been considered as a numerical analysis problem, where sophisticated techniques (such as singular value decomposition, simulated annealing, Kalman filtering, maximum likelihood estimation, etc.) have been employed in order to estimate 3-D motion from the geometric and photometric constraints which relate local image motion to the 3-D world. In this study we take a different approach to this problem by introducing new global constraints, which take the form of patterns in the image plane. Estimating motion amounts then to recognizing particular patterns on the image through a simple search in spaces of low dimensionality. Since 3-D rigid motion consists of a translation and a rotation, finding the axes of translation and rotation, in our approach, amounts to recognizing the points where these axes intersect the image plane. It is these points/locations that are recognized by recognizing the location of patterns in the image plane.

Previous work along these lines is due to [Horn and Weldon, 1987], who treated the case of only translational motion and [Aloimonos and Brown, 1984], who treated the pure rotational case.

2 The problem : Overview

For years 3-D motion interpretation has been approached through studying the "structure from motion" problem. The goal is to recover from dynamic imagery the 3-D motion parameters and the structure of the objects in view. The suggested strategy involves to carry out the problem in two stages. First accurate image displacements between consecutive frames have to be computed, either in form of point correspondences or as dense motion fields (optical flow fields) and then the 3D-motion and the structure is computed from the equations relating it to the 2-D image velocity. (For an overview see [Aggarwal and Nandhakumar, 1988]). The "structure from motion" approach has been uncritically accepted and as a result most studies on visual motion perception are to be found at the algorithmic level of the problem; addressing either the estimation of image motion or the recovery of 3-D motion and structure from image motion.

Here we do not consider the general problem of "struc-

ture from motion", but the problem of "passive navigation", i.e. the determination of the rigid 3-D motion of an observer moving in a static environment (its direction of translation and its rotation). The method does not require the computation of exact 2-D motion. In general, the estimation of optical flow is an ill-posed problem and additional assumptions must be made in order to estimate it. If these assumptions hold, as in the case of some model-based approaches [Burt *et al.*, 1989; Koller *et al.*, 1992], optical flow can be computed accurately and in the sequel correct 3-D motion can be derived. In the general case, however, any algorithm will produce imperfect output. Therefore, we use as representation for image motion the so-called normal flow, i.e. the prection of the flow on the gradient direction $((I_x, I_y) \cdot (u, v) = -I_t)$. Recovering 3-D motion from noisy flow fields has turned out to be a problem of extreme sensitivity with researchers reporting very large amount of errors in the motion parameter estimation under small perturbations in the input. Even optimal algorithms [Spetsakis and Aloimonos, 1988] perform quite poorly in the presence of moderate noise. Although a formal proof is still lacking, it has been argued [Adiv, 1985], that the estimation of 3-D motion from image motion is itself ill-posed, because it does not continuously depend on the input. Thus, in order to estimate 3-D motion in a robust way, we have to consider the fact, that no flow measurement (neither optical flow, nor normal flow) can be perfect. In this study new global constraints of geometric nature, which relate 3-D motion to 2-D image measurements are introduced. In our approach, we first find the rotation axis and the direction of translation. After having found the axis of rotation and the direction of translation further constraints are considered, and the complete set of motion parameters is obtained.

3 Geometric constraints

If the coordinate system is fixed to the observer with the center being the nodal point of the camera and f the focal length, then the equation relating the normal flow u_n at a point (x, y) —i.e. the projection of the image velocity (u, v) on a unit vector (n_x, n_y) along the image gradient direction at (x, y) —to the 3-D velocity $((U, V, W)$ translational and (α, β, γ) rotational) and the depth z of the corresponding scene point, is:

$$u_n = ((-x_0 + xf) \frac{W}{Z} + \alpha \frac{xy}{f} - \beta (\frac{x^2}{f} + f) + \gamma y) n_x + ((-y_0 + yf) \frac{W}{Z} + \alpha (\frac{y^2}{f} + f) - \beta \frac{xy}{f} - \gamma x) n_y \quad (1)$$

The above equation demonstrates the difficulties of motion computation using normal flow. A monocular observer not being able to measure depth is confronted with a motion field of five unknown motion parameters and one scaled depth component (W/Z) at every point.

3.1 Motion field interpretation

A motion field is composed of a translational and a rotational component. If the observer undergoes only translational motion, all points in the 3D scene move along parallel lines. Translational motion viewed under perspective results in the image plane in a motion field,

where every point moves along a line through one vanishing point. This point constitutes the intersection of the image plane and the parallel line passing through the nodal point. Its image coordinates are $x_0 = Uf/W$ and $y_0 = Vf/W$, the flow there is of value zero. In cases where the sensor is approaching the scene all vectors are emanating from the vanishing point, then called Focus of Expansion (FOE).

In cases of purely rotational motion every point in 3D moves along a circle in a plane perpendicular to the axis of rotation. The perspective image of the circular path is the result of an intersection of the image plane with the cone which the circle defines with the rotation axis. Depending on the relation of the opening angle of the cone for a specific image point and the angle the image plane forms with the rotation axis, the field lines of the rotational vector field (i.e the lines which have the property that at each point the rotational flow is tangential) form different second order curves: ellipses, hyperbolas, parabolas and even circles when the rotation axis and the optical axis coincide. The specific conic sections due to a rotational motion are defined by the axis of rotation. A rotation axis, given by the two parameters $(\frac{\alpha}{\gamma})$ and $(\frac{\beta}{\gamma})$, defines the family $M(\frac{\alpha}{\gamma}, \frac{\beta}{\gamma}; x, y)$ of conic sections as:

$$M(\frac{\alpha}{\gamma}, \frac{\beta}{\gamma}; x, y) = ((\frac{\alpha}{\gamma})^2 x^2 + 2xy\frac{\alpha}{\gamma}\frac{\beta}{\gamma} + y^2(\frac{\beta}{\gamma})^2 + 2xf\frac{\alpha}{\gamma} + 2yf\frac{\beta}{\gamma} + f^2)/(x^2 + y^2 + f^2) = C \quad (2)$$

with C in $[0 \dots (1 + (\frac{\alpha}{\gamma})^2 + (\frac{\beta}{\gamma})^2)]$

Specifically, for a rotation around the Z-axis the second order curves are circles with center 0; we call them γ -circles (Figure 1a). If the rotation axis is the X- or Y-axis the conic sections are hyperbolas whose axes coincide with the coordinate axes of the image plane. For the case of rotation around the X-axis the hyperbolas' major axis is the x-axis and they are called α -hyperbolas (1b). For rotation around the Y-axis the major axis is the y-axis and we call the conic sections β -hyperbolas. (Figure 1c).

4 Global motion constraints

In this section we study the sign of the normal flow in certain directions and the localization of normal flow vectors of equal sign. Vectors which are perpendicular to rotational vector field lines are considered. For these vectors we find that the FOE and the axis of rotation separate the normal flow values in the image according to their sign by a second order curve and a straight line.

4.1 Selection of values

The normal flow vector $\vec{u}_n$ is the projection of the optical flow vector $\vec{u}$ on the gradient direction and the value of the normal flow is therefore defined by the scalar product of the optical flow vector and the unit vector in gradient direction. The flow vector can be decomposed in its translational and single rotational components and the right hand side of equation (1) can be written as a sum of scalar products.

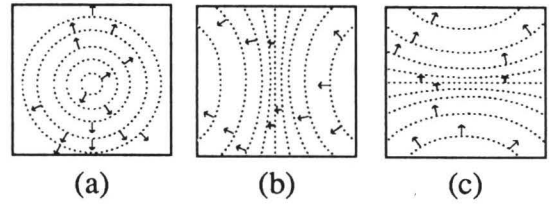


Figure 1: Positive γ -, α - and β -vectors perpendicular to the γ -circles and α -, β -hyperbolas.

$$\begin{aligned} u_n = & \frac{W}{Z}((-x_0 + xf), (-y_0 + yf))(n_x, n_y) + \\ & \alpha(\frac{xy}{f}, (\frac{y^2}{f} + f))(n_x, n_y) \\ & - \beta((\frac{x^2}{f} + f), \frac{xy}{f})(n_x, n_y) + \gamma(y, -x)(n_x, n_y) \end{aligned} \quad (3)$$

All the normal flow vectors that form right angles with the γ -circles do not contain a component due to rotation around the Z-axis. Similarly, there is no motion component due to rotation around the X-axis for the normal flow vectors perpendicular to the α -hyperbolas, and no component due to rotation around the Y-axis for vectors perpendicular to the β -hyperbolas. The motion estimation problem becomes easier when analyzing only the normal flow vectors perpendicular to one of these families of conic sections. It is reduced by one parameter for these subsets of the normal flow vectors.

We call these three subsets of the normal flow vectors the α -, β -, and γ -vectors. It is convenient to agree upon conventions for the vectors' orientations. A γ -vector at point (x, y) is said to have positive orientation if it is pointing in the direction (x, y) ; otherwise, its orientation is said to be negative. Similarly, we call an α -vector (or a β -vector) originating from a point (x, y) positive if it points in the direction $(-f^2 + y^2, xy)$ (or $(xy, -(f^2 + x^2))$) (see Figure 1).

4.2 Properties of the subsets

Let us first concentrate on the γ -vectors. These vectors do not contain a component due to rotation around the Z-axis. Along the positive direction, the two remaining rotational components contribute

$$v_{\text{rot}}(r, \phi) = -\alpha(r^2/f + f) \sin \phi + \beta(r^2/f + f) \cos \phi,$$

where r is distance from the image center and the angle ϕ is measured from the x-axis. Thus, the rotational component of the normal flow along a vector pointing away from the image center can be described by a trigonometric function with amplitude $\max(\alpha, \beta)$ and period 2π . Along the line which passes through the image center and makes angle $\phi = \arctan(\beta/\alpha)$ with the x-axis the values are zero. This line divides the plane into two halves. In one half the vectors point in the positive direction, and in the other half they point in the negative direction; in the future we simply refer to them as positive and negative vectors (Figure 2a).

The translational component of the motion field is characterized by the location of the FOE in the image plane. The γ -vectors lie on lines passing through the

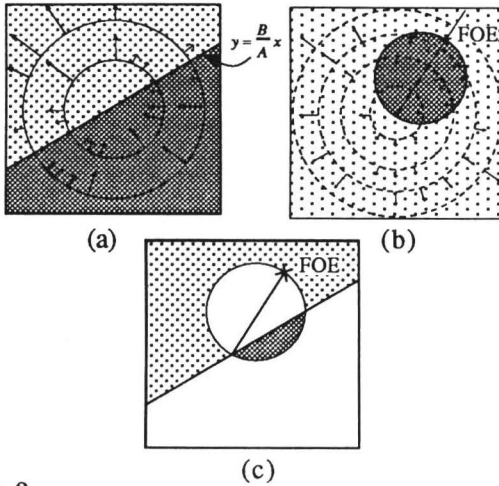


Figure 2:

- a: The γ -vectors due to rotation separate the image plane in a halfplane of positive values and a halfplane of negative values.
- b: The γ vectors due to translation are negative, if they lie within the circle defined by the FOE and the image center and are positive at all other locations.
- c: A general rigid motion defines an area of positive γ -vectors and an area of negative γ -vectors. The rest of the image plane is not considered.

image center and the optical flow values due to translation lie on lines passing through the FOE. These two lines are at right angles for all points on a circle which has the the FOE and the image center as diametrically opposite points. At these points the γ vectors' translational components vanish. Thus, the geometric locus of all points on the γ -hyperbolas where there is zero translational normal flow, is a circle. The diameter of this circle is the line segment connecting the image center and the FOE. At all points inside this circle the two lines enclose an angle greater than 90° and the normal flow along the γ -vector therefore has a negative value. The normal flow values outside the circle are positive (Figure 2b).

In order to investigate the constraints associated with a general motion, the geometrical relations derived from rotation and from translation have to be combined. A circle separating the plane into positive and negative values and a line separating the plane into two halfplanes of opposite sign always intersect (in two points or one point in case the line is tangential to the circle), because both the line and the circle pass through the origin. This splits the plane into areas of only positive, of only negative γ -vectors, and into areas in which the rotational and translational flows have opposite signs. In the latter areas, unless we make depth assumptions, no information is derivable (Figure 2c).

We thus obtain the following geometrical result for the case of general motion. Points in the image plane at which the gradient direction is perpendicular to circles around the image center can be separated into two classes. For a given FOE, and for a line through the image center which represents the quotient of two of the

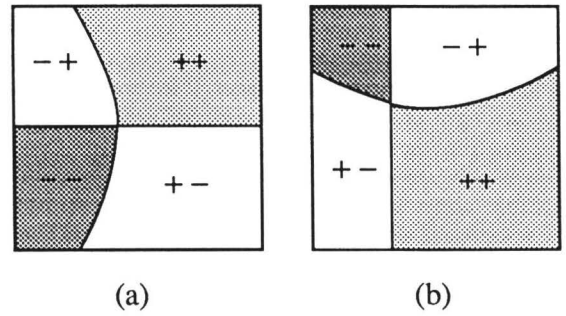


Figure 3: α - and β -patterns for general rigid motion.

three rotational parameters, there are two geometrically defined areas in the plane, one containing positive and one containing negative values. We call this structure on the γ -values the γ -pattern. It depends on the three parameters x_0 , y_0 and $\frac{B}{A}$.

Similar relations can be derived when eliminating the motion components due to rotation around the X- and Y- axes. The α - and β -vectors due to rotation are also separable into positive and negative vectors. In both cases the locus of zero normal flow which separates the two classes is a line. For the α -hyperbolas the line is parallel to the x-axis and is defined by the equation $y = \frac{B}{A}$, for the β -hyperbolas it is parallel to the y-axis and is defined by $x = \frac{A}{B}$. The translational components of the α - and β -vectors are separated by hyperbolas. The α -vectors, which are perpendicular to lines through the FOE, and which therefore have zero normal flow lie on a hyperbola of the form

$$h(x_0, y_0; x, y) = x_0 y^2 - x y y_0 - x f^2 + x_0 f^2 = 0$$

When $h(x_0, y_0) > 0$, the normal flow values are positive; in the other part of the plane they are negative. Symmetrical relations hold for the β -vectors. The curve of zero normal flow is defined by

$$g(x_0, y_0; x, y) = x^2 y_0 - x_0 x y - y f^2 + y_0 f^2 = 0$$

and in areas of positive $g(x_0, y_0)$ the β -vectors are positive.

The superposition of translational and rotational values again defines patterns in the plane each of which consists of a negative and a positive area. These patterns, called α - and β -patterns, are uniquely described by three parameters: x_0 and y_0 , the coordinates of the FOE and the quotients $\frac{B}{A}$ (and $\frac{A}{B}$) (Figure 3).

The α - β - and γ -patterns are obtained as special cases from a more general pattern concept defined for normal flow vectors in all directions [Fermüller, 1992].

5 Motion recognition: The technique

In this section we describe a search technique, which utilizes the geometrical constraints on the α -, β - and γ -vectors to recognize the direction of translation (x_0 , y_0) and the axis of rotation ($\frac{A}{B}$, $\frac{B}{A}$). By fitting a small number of patterns to the normal flow field, candidate solutions for the axis of translation and the direction of the

rotation axis are obtained. In this fitting process the input is used in a qualitative way, since only the sign of the normal flow is employed. The search for motion-patterns forms the core of our motion recognition strategy. The whole motion recognition technique requires two more processes: the ones of "detranslation" and "derotation"; additional constraints are checked, which estimate the third rotational component and narrow down the solution space to a unique 3-D motion vector.

5.1 Pattern fitting: Search in 3D parameter spaces

The geometrical constraints developed in the last section are used in a search process to estimate the directions of the translation and the rotational axis. Finding these two directions is a four-dimensional problem, but through selection of values and use of geometrical constraints the problem is reduced in dimensionality. With each subset of the normal flow vectors is associated a three-dimensional parameter space that spans the possible locations of the FOE and of a line defined by the quotient of two of the three rotational parameters. Instead of searching a four-dimensional space, three three-dimensional subspaces are searched for the solution.

The search in the three-dimensional subspaces is accomplished by checking the patterns which the subspaces' parameter triples define on selected values of the normal flow field. The α -patterns are fitted to the α -vectors; this provides possible solutions for the coordinates of the FOE: X_0 , Y_0 , and the quotient $\frac{\beta}{\alpha}$. Similarly, the fitting of the β - or γ -patterns yields solutions for X_0 , Y_0 and $\frac{\beta}{\gamma}$ or $\frac{\alpha}{\gamma}$. The objective is to find the four parameters defining the directions of the translational and rotational axes which give rise to three successfully fitted patterns. Therefore the three subspaces' patterns are combined and the parameter quadruples which define possible solutions are determined. Since only subsets of the normal flow values are considered in the fitting process, the fitting does not uniquely define the motion, but just constitutes a necessary condition. Usually a number of parameter quadruples $\{x_0, y_0, \alpha/\gamma, \beta/\gamma\}$ are selected as candidate solutions through pattern fitting.

The range of values for the coordinates of the FOE and for $\frac{\beta}{\gamma}$ and $\frac{\alpha}{\gamma}$ is $[-\infty, +\infty]$. To cope with all possible cases a coordinate transformation on the sphere is performed, in which case the coordinates are expressed by two angles.

5.2 Detranslation

Proper selection of normal flow vectors also makes it possible to eliminate the normal flow's translational components. If the location of the FOE is given the directions of the translational motion components are also known. The optical flow vectors lie on lines passing through the FOE. The normal flow vectors perpendicular to these lines do not contain translational components; they have only rotational components. For each of the possible solutions computed by pattern fitting, the normal flow vectors perpendicular to the lines passing through the FOE have to be tested to determine if they are really

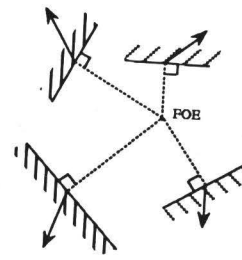


Figure 4: Normal flow vectors due to translation are constrained to lie in halfplanes.

only due to rotation. This results in solving an overdetermined system of linear equations. Since two of the rotational parameters are already computed, there is only one unknown. If the chosen normal flow vectors are due only to rotation then the solution to the overdetermined system gives the γ value. Usually "detranslation" will not result in only one solution, but will provide a set of possible parameter quintuples.

5.3 Derotation

The computational steps described so far considered only subsets of the normal flow vectors. In order to eliminate all motion parameters which are not consistent with the given normal flow field, every normal flow vector has to be checked.

This check is performed using a "derotation" technique. For every parameter quintuple derived so far, a possible FOE and a rotation is defined. The three rotational parameters are used to derotate the normal flow vectors by subtracting the rotational component. At every point the flow vector (u_{der}, v_{der}) is computed:

$$\begin{aligned} u_{der} &= u_n n_x - u_{rot} n_x \\ v_{der} &= v_n n_y - v_{rot} n_y \end{aligned} \quad (4)$$

If the parameter quintuple defines the correct solution, the remaining normal flow is purely translational. Thus the corresponding optic flow field consists of vectors that all point away from one point, the FOE. Since the direction of optical flow for a given FOE is known, the possible directions of the normal flow vectors can be determined. The normal flow vector at every point is confined to lie in a half plane (see Figure 4). The technique checks all points for this property and eliminates solutions that cannot give rise to the given normal flow field. It may happen that more than one solution is possible, since in general normal flow fields do not uniquely define 3-D motion.

6 Experiments

Several experiments have been performed on synthetic and real data. For different 3D motion parameters normal flow fields have been generated, where the depth value within an interval and the gradient direction are chosen randomly. In all experiments on noiseless data the correct solution is found as the best one. Figure 5 shows the optical flow field and the normal flow field for one of the generated data sets: The image size is

100×100 , the focal length is 150, the image coordinates of the FOE are $(-5, +30)$ and the relationship of the rotational components is $\alpha : \beta : \gamma = 10 : 11 : 150$. In figure 6 the fitting of the circle and the hyperbolas to the α - β and γ -vectors. $\frac{\beta}{\gamma}$ is displayed. Points with positive normal flow values are rendered in light color and points of negative value are dark. Since only the sign of the normal flow is utilized in the pattern fitting step for finding the axes of rotation and translation, the procedure can handle errors of up to 100 percent.

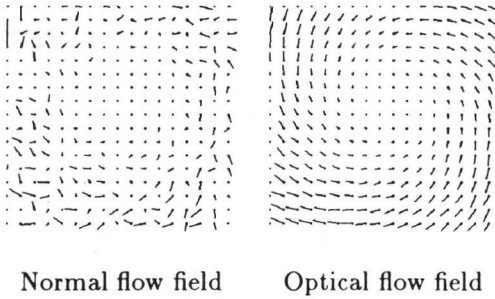


Figure 5: Flow fields for synthetic data

As a natural scene the NASA-Ames sequence ² has been chosen. The camera undergoes only translational motion, and we added different amounts of rotation: For all points at which translational motion can be found the rotational normal flow is computed, the new position of each pixel is evaluated. The “rotated” image is then generated by computing the new greyvalues through bilinear interpolation. The images were convolved with a Gaussian of kernel size 5×5 and standard deviation $\sigma = 1.4$. The normal flow was computed by using 3×3 large Sobel operators to estimate the spatial derivatives in x- and y-directions and by subtraction of the 3×3 box-filtered values of consecutive images for the temporal derivatives.

When adding rotational normal flow in the order of a third to three times the amount of translational flow the exact solution always has been found among the best fitted parameter sets. In figure 7 the computed normal flow vectors and the fitting of the α -, β - and γ -vectors for one of the “rotated” images is shown. Areas of negative normal flow vectors are marked by horizontal lines and areas of positive values with vertical lines. The ground truth for the FOE is $(-5, -8)$, the focal length is 599 pixels and the rotation between two image frames is $\alpha = 0.0006$, $\beta = 0.0006$ and $\gamma = 0.004$. The algorithm computed the solution exactly.

7 Conclusion

We argue that every vision problem can be formulated as a recognition problem with two questions to be solved. The first question, *what to recognize*, is related to the task, which the visual system is supposed to perform and the second question *how to recognize* is concerned with

²This is a calibrated motion sequence made public for the Workshop on Visual Motion, 1991

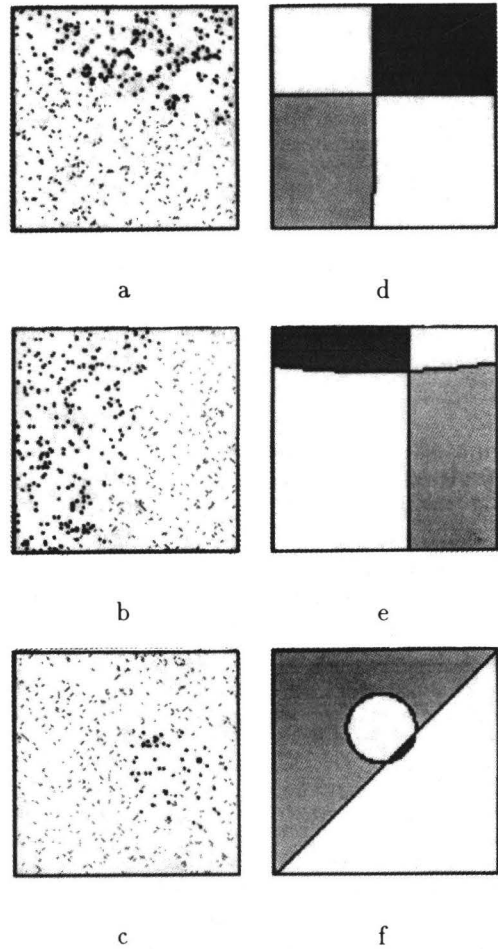


Figure 6:
a,b,c: Positive and negative α -, β - and γ -vectors for synthetic data.
d,e,f: Fitting of α -, β - and γ -patterns.

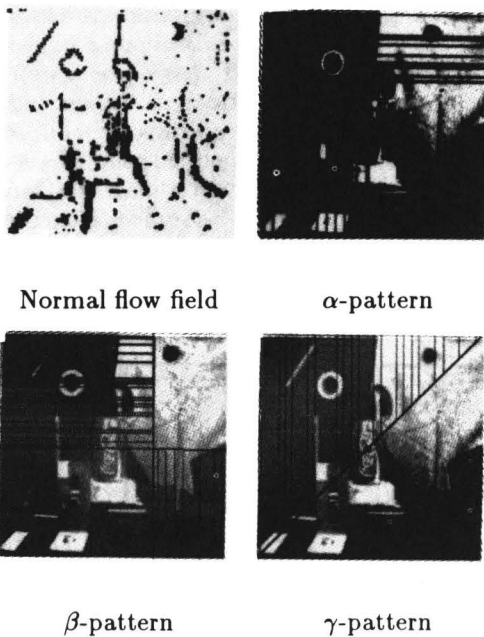


Figure 7: Real scene: Normal flow field and fitting of α - β - and γ -vectors.

the computational approach to be taken. Our position is that any visual problem should be addressed in a robust, non-local and qualitative manner. As an example we treated the celebrated problem of passive navigation as a recognition problem.

We presented geometric constraints on normal flow fields due to rigid motion, which take the form of patterns in the image plane. These constraints were exploited to solve the problem of recognizing the 3D-motion of an observer relative to the scene in a robust way.

References

- [Adiv, 1985] G. Adiv. Inherent ambiguities in recovering 3D motion and structure from a noisy flow field. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 70 – 77, 1985.
- [Aggarwal and Nandhakumar, 1988] J.K. Aggarwal and N. Nandhakumar. On the computation of motion from sequences of images. *Proceedings of the IEEE*, 76(8):917 – 935, 1988.
- [Aloimonos and Brown, 1984] J. Aloimonos and C.M. Brown, “Direct processing of curvilinear sensor motion from a sequence of perspective images”, *Proc. Workshop on Computer Vision: Representation and Control*, Annapolis, MD, 7277, 1984.
- [Aloimonos et al., 1988] J. Aloimonos, I. Weiss, and A. Bandopadhyay. Active vision. *International Journal of Computer Vision*, 2:333–356, 1988.
- [Aloimonos, 1990] J. (Y.) Aloimonos. Purposive and qualitative active vision. In *Proc. DARPA Image Understanding Workshop*, pages 816 – 828, 1990.
- [Bajcsy and Campos, 1992] R. Bajcsy and M. Campos. Active and exploratory perception. *CVGIP: Image Understanding*, Special Issue on Purposive, Qualitative and Active Vision, Y. Aloimonos, editor. 56:31 – 40, 1992.
- [Bajcsy, 1988] R. Bajcsy. Active perception. *Proceedings of the IEEE*, 76(8):996 – 1005, 1988.
- [Ballard, 1991] D.H. Ballard. Animate vision. *Artificial Intelligence*, 48:57 – 86, 1991.
- [Bhanu and (editors), 1992] B. Bhanu and B. Sridhar (editors). Special issue on passive ranging for robotic systems. *Journal of Robotic Systems*, 9(6), September 1992.
- [Brady et al., 1983] M. Brady, J.M. Hollerbach, T.L. Johnson, T. Lozano-Perez, and M.T. Mason, editors. *Robot Motion*. MIT Press, Cambridge, MA, 1983.
- [Burt et al., 1989] P.J. Burt, J.R. Bergen, R. Kolczynski, W.A. Lee, A. Leung, J. Lubin, and H. Shvaytser. Object tracking with a moving camera. In *Proc. IEEE Image Understanding Workshop*, pages 2 – 12, 1989.
- [Fermüller, 1992] Cornelia Fermüller. Motion constraint patterns. *Proc. DARPA Image Understanding Workshop*, 1993.
- [Herve et al., 1991] J-Y. Herve, R. Sharma, and P. Cucka. The geometry of visual coordination. In *National Conference on Artificial Intelligence*, pages 732 – 737, July 1991.
- [Horn and Schunck, 1981] B.K.P. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185 – 203, 1981.
- [Horn and Weldon, 1987] B.K.P. Horn and E.J. Weldon, “Computationally efficient methods of recovering translational motion”, *Proc. International Conference on Computer Vision*, 1987, 2–11.
- [Koller et al., 1992] D. Koller, K. Daniilidis, T. Thorhallson, and H.H. Nagel. Model-based object tracking in traffic scenes. In *Proc. Second European Conference on Computer Vision*, pages 437 – 452, 1992.
- [Longuet-Higgins and Prazdny, 1980] H. C. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proceedings of the Royal Society, London B*, 208:385–397, 1980.
- [Marr, 1982] D. Marr. *Vision*. W.H. Freeman, San Francisco, 1982.
- [Nelson and Aloimonos, 1989] R. C. Nelson and J.Y. Aloimonos. Obstacle avoidance using flow field divergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(10):1102–1106, 1989.
- [Nelson, 1991] R. C. Nelson. Qualitative detection of motion by a moving observer. *International Journal of Computer Vision*, 7(1):33 – 46, 1991.
- [Spetsakis and Aloimonos, 1988] M. E. Spetsakis and John Aloimonos. Optimal computing of structure from motion using point correspondence. In *Proc. International Conference on Computer Vision*, pages 449 – 453, 1988.