From: AAAI-92 Proceedings. Copyright ©1992, AAAI (www.aaai.org). All rights reserved.

Grouping Iso-Velocity Points for Ego-Motion Recovery *

Yibing Yang and Alan Yuille
Division of Applied Sciences
Harvard University
Cambridge, MA 02138
yibing@hrl.harvard.edu

Abstract

The instantaneous image motion field due to a camera moving through a static environment encodes information about ego-motion and environmental layout. For pure translational motion, the motion field has a unique point termed focus of expansion/contraction where the image velocity vanishes. We reveal the fact that for an arbitrary 3D motion the zero-velocity points, whose number can be large, have the regularity of being approximately cocircular. More generally, all the image points with the same velocity **u** are located approximately on a *circle* (termed the *iso-velocity circle* (IVC)) determined solely by **u** and the ego-motion, except for the pathological cases in which the circle degenerates into a straight line. While IVCs can be recovered from 3 or more *pairs* of iso-velocity points in the motion field using a linear method, estimating ego-motion reduces to solving systems of linear equations constraining iso-velocity point pairs (Yang 1992).

Introduction

When a camera moves through a fixed environment, the image points are endowed velocities, resulting in a time-varying motion field. A monocular observer can make use of the motion field information to recover the shapes of the objects and the motion relative to it (Gibson 1950). Mathematical studies further show that in general motion and relative depth are uniquely determined by the instantaneous motion field (Horn 1987). It has be found difficult, however, to bestow a computer the ability of ego-motion recovery, despite many efforts (e.g., (Longuet-Higgins & Prazdny 1980)(Waxman & Ullman 1985) (Jepson & Heeger 1991)). The main difficulty comes from the highly nonlinear relationship between the observables (i.e., the motion field values) and the unknowns (i.e., motion and depth).

*This research was supported in part by the Brown/Harvard/MIT Center for Intelligent Control Systems with U.S. Army Research Office grant number DAAL03-86-K-0171.

The Gestalt psychologists have formulated a number of principles of perceptual organization among which are grouping by proximity and grouping by common fate (Koffka 1935). We would rather refer grouping to grouping by proximity but applied to two conjugate spaces, the *image space* and the *velocity space*. It is generally believed that the Gestalt laws may work because they reflect sensible assumptions that can be made about the world of physical and biological objects (Marr 1982). For example, because matter is cohesive, adjacent regions are likely to belong together and move together.

In this paper, we will demonstrate the usage of grouping iso-velocity points (IVPs) for ego-motion perception, thanks to the constraints resulting from the rigid environment assumption. As a theoretical contribution of this work, the *Iso-Velocity Point Theorem* reveals the fact that any image point of velocity **u** must lie approximately on a semi-circle termed the *iso-velocity circle* (IVC) which is determined by the motion and the velocity **u**. With the geometry of the circle not dependent on depth information, structure and motion are segregated. (For the pathological circumstances of ego-motion, the circle may degenerate into a straight line.) Applicable to any 3D ego-motion and rigid surfaces, this theorem provides new insight into the image motion perceived by a moving observer.

Our other contribution is a computer algorithm for motion/structure from optical flow. The proposed algorithm computes the IVCs from three or more *pairs* of IVPs, thus eliminating the need of identifying at least three IVPs in order to determine a single IVC. The computation of ego-motion, separated from a subsequent depth recovery procedure, involves no differentiating/searching/iterating but grouping IVP pairs and solving systems of largely over-constrained *linear* equations each of which is specified by one pair of IVPs.

Motion Field

Motion Field Equation

Like many researchers, we adopt a camera-based coordinate system with the origin being the projection center and the optical axis running along the Z-axis.

Under perspective projection, a world point $\mathbf{P} = (X, Y, Z)^T$ is imaged at $\mathbf{p} = (x, y, f)^T = \frac{f}{Z}\mathbf{P}$, where f is the focal length. Suppose the camera moves with instantaneous translational velocity $\mathbf{t} = (t_x, t_y, t_z)^T$ and rotational velocity $\mathbf{r} = (r_x, r_y, r_z)^T$ relative to a static environment. Then the induced image velocity at $\mathbf{p}$ is

$$f^{-1}\dot{\mathbf{p}} = \frac{1}{Z}(f^{-1}(\mathbf{t}\cdot\hat{\mathbf{z}})\mathbf{p} - \mathbf{t}) + f^{-2}(\mathbf{p}\times\mathbf{r}\cdot\hat{\mathbf{z}})\mathbf{p} - f^{-1}\mathbf{p}\times\mathbf{r}$$

where $\hat{\mathbf{z}}$ is the unit vector in the Z-direction. This is the well-known *motion field equation* which specifies the image velocity $\dot{\mathbf{p}}$ as a function of motion $\{\mathbf{t}, \mathbf{r}\}$ and depth Z, as well as the image position $\mathbf{p}$.

To gain more insight into the relation between $\dot{\mathbf{p}}$ and $\mathbf{p}$, we rewrite the motion field equation in the following form:

$$\mathrm{MF}_{pers} : \mathbf{u} = \begin{pmatrix} -t_x/Z - r_y \\ -t_y/Z + r_x \end{pmatrix} + \begin{pmatrix} t_z/Z & r_z \\ -r_z & t_z/Z \end{pmatrix} f^{-1}\mathbf{x} + f^{-2}\mathbf{x}\mathbf{x}^T \begin{pmatrix} -r_y \\ r_x \end{pmatrix} \quad (1)$$

where $\mathbf{x} = (x, y)^T$ and $\mathbf{u} = (u, v)^T = f^{-1}\dot{\mathbf{x}}$. If both the translational velocity $\mathbf{t}$ and depth Z are magnified or reduced by the same factor, the motion field remains unchanged. This *scale factor ambiguity* shows that the absolute magnitude of the translational motion is not encoded in the motion field. Only the relative values $t_x/t_z, t_y/t_z$ and Z/t_z take effects, assuming $t_z \neq 0$.

Suppose the image size is K pixels. The objective of motion recovery is to estimate the 5 motion parameters $\{t_x/t_z, t_y/t_z, r_x, r_y, r_z\}$, and the K relative depth values, given the $2K$ image velocity values $\{u_k, v_k\}$.

Motion Fields in A Small Field of View

The motion field value in eq. (1) is quadratic in $\mathbf{x}$ for constant depth. If the *field of view* (FOV) is sufficiently small,[1] or $\max\|\mathbf{x}\|/f \ll 1$ in the image, then the motion field equation can be expressed in the following simplified form:

$$\mathrm{MF}_{para} : \mathbf{u} = \underbrace{\begin{pmatrix} t_z/Z & r_z & -t_x/Z - r_y \\ -r_z & t_z/Z & -t_y/Z + r_x \end{pmatrix}}_{\Phi} \begin{pmatrix} f^{-1}\mathbf{x} \\ 1 \end{pmatrix} \quad (2)$$

The above equation holds if world point $\mathbf{P}$ is mapped to the image point $\mathbf{p}$ under the so-called *paraperspective projection* (Ohta et al. 1981)(Aloimonos 1988) which is an approximation of perspective projection under the condition of small FOV.

Note that if we further assume lateral translation, i.e., $t_z = 0$, then we get the motion field equation under *orthographic projection*:

$$\mathrm{MF}_{orth} : \mathbf{u} = \begin{pmatrix} r_z y - t_x/Z - r_y \\ -r_z x - t_y/Z + r_x \end{pmatrix} \quad (3)$$

This motion field is ambiguous (up to scale); different motion/structures can give rise to the same motion field. Indeed, if r_x, r_y, Z are replaced by $r'_x = r_x + ct_y, r'_y = r_y - ct_x, Z' = (c + 1/Z)^{-1}$, ($c$ = const.) respectively, then MF_{orth} remains unchanged. In this case, it is well understood that motion/structure can be recovered from the position information of features at three or more time instants (Ullman 1979).

[1]The visual information within a small FOV is of particular importance. The human eye has much better resolution near the optical axis. It has a high-resolution fovea where over a 1° range the resolution is better by an order of magnitude than that in the periphery.

Fixed Points on Image Plane

A well-known fact concerning the pure translational motion is the existence of a unique image point called the *focus of expansion/contraction* (FOE) where the image velocity vanishes. Viewing MF_{pers} as an autonomous differential equation system, we expect more than one fixed points in the general case that $\mathbf{r} \neq \mathbf{0}$. Local analysis has shown the fixed points contain useful information about the ego-motion (Verri et al. 1989). What we are going to explore in this paper is the global regularity of the fixed points.

Let us consider the fixed points of MF_{para}, categorizing them (Jordan 1987) in the extreme cases. When $Z \to \infty$, we have

$$\mathbf{u} \to (r_z y/f - r_y, -r_z x/f + r_x)^T$$

The motion field looks as if it were induced by pure rotation. It has a unique fixed point $\mathbf{x} = \frac{f}{r_z}(r_x, r_y)^T$, which is obviously a *center*.

On the other hand, as $Z \to 0$, we have

$$\mathbf{u} \to \frac{1}{Z}(t_z x/f - t_x, t_z y/f - t_y)^T$$

In this case, the rotational effects are negligible and the unique fixed point is a *node* $\mathbf{x} = \frac{f}{t_z}(t_x, t_y)^T$, corresponding to the direction of translation. At any image position, the image velocity points to this node.

Definition. The point $\mathbf{N} \triangleq (t_x, t_y)^T f/t_z$ on the image plane is called the *node of translation* and the point $\mathbf{C} \triangleq (r_x, r_y)^T f/r_z$, the *center of rotation.*

As a matter of fact, the node of translation has the same definition as the FOE. We deem it necessary to give a new name in order to avoid confusion, because conventionally the latter is used only in the context of pure translation. The center of rotation is the rotational counterpart of the FOE which has been largely overlooked in the computer vision community. Although $\mathbf{C}$ by itself is linked to the trivial case of pure rotational motion, the alliance of $\mathbf{C}$ with $\mathbf{N}$ will prove very useful for characterizing general motion, as illustrated below.

Theorem 1. All fixed points of MF_{para} lie exclusively on one of the two semi-circles defined by $\mathbf{N}$ and $\mathbf{C}$.

Proof. The key to the proof is the observation that in eq. (2), matrix Φ specifies a *similitude.* Let $\xi = x + iy, \mu = u + iv$ be the complex representations of $\mathbf{x}$ and $\mathbf{u}$, respectively. Then eq. (2) can be rewritten as

$$\mu = (t_z/Z - ir_z)f^{-1}\xi - t_x/Z - r_y + i(-t_y/Z + r_x) \quad (4)$$

Let $\mu = 0$ and then the fixed point is

$$\breve{\xi}(Z) = \frac{(t_x + it_y) + (r_y - ir_x)Z}{t_z - ir_z Z} f$$

Notice that the depth Z takes values on the real axis (of the complex plane) only and Z is mapped to $\breve{\xi}$ by a *bilinear transformation* which maps a straight line to a circle biuniformly. When Z changes from 0 to ∞, $\breve{\xi}$ changes accordingly from $\xi_{\mathbf{N}} = (t_x + it_y)f/t_z$ to $\xi_{\mathbf{C}} = (r_x + ir_y)f/r_z$ on a circle which has a diameter specified by $\xi_{\mathbf{N}}$ and $\xi_{\mathbf{C}}$. Clearly, $\xi_{\mathbf{N}}$ and $\xi_{\mathbf{C}}$ are the complex representations of $\mathbf{N}$ and $\mathbf{C}$, respectively. ∎

Iso-Velocity Points Theorem

Inspired by the regularity obeyed by the fixed points, we proceed to investigate the iso-velocity points (IVPs) of an arbitrary velocity. Roughly speaking, the relation between a fixed point and an IVP is like the relation between a zero-crossing and a level-crossing.

Iso-Velocity Circles

We shall show that all the IVPs of MF_{pers} are approximately co-circular and all the IVPs of MF_{para} are exactly co-circular. Here an important question that has to be answered is how good the approximation is.

Theorem 2. (Iso-Velocity Points Theorem: Non-degenerate case). All the IVPs $\mathscr{X}_{\mathbf{u}} = \{\mathbf{x} : \mathbf{u}(\mathbf{x}) = \mathbf{u}, ||\mathbf{x}||/f \leq \tau\}$ of MF_{pers} are nearly co-circular in the sense that for any $\mathbf{x} \in \mathscr{X}_{\mathbf{u}}$, there is a point $\mathbf{C}_{\mathbf{u},\mathbf{x}}$ within the circle centered at $\mathbf{C}_{\mathbf{u}} = \mathbf{C} + (-v, u)^T f/r_z$ with radius $\varepsilon = \tau^2||\mathbf{C}||$, such that $\mathbf{x}$ spans the right angle with respect to $\mathbf{N}$ and $\mathbf{C}_{\mathbf{u},\mathbf{x}}$ (see Fig. 1).

Proof. Eq. (1) can be rewritten as

$$u = -t_x/Z - r_y + t_z x/(Zf) + r_z y/f + \tilde{u} \quad (5)$$

$$v = -t_y/Z + r_x - r_z x/f + t_z y/(Zf) + \tilde{v} \quad (6)$$

where $(\tilde{u}, \tilde{v})^T = f^{-2}\mathbf{x}\mathbf{x}^T(-r_y, r_x)^T$. Algebraic manipulations yield

$$\frac{t_z}{r_z Z}\begin{pmatrix} x - t_x f/t_z \\ y - t_y f/t_z \end{pmatrix} = \begin{pmatrix} -y + (r_y + u - \tilde{u})f/r_z \\ x - (r_x - v + \tilde{v})f/r_z \end{pmatrix}$$

Notice that the LHS can be written as $(\mathbf{X}-\mathbf{N})t_z/(r_z Z)$ and the RHS is perpendicular to $(\mathbf{x} - \mathbf{C}_{\mathbf{u},\mathbf{x}})$, where

$$\mathbf{C}_{\mathbf{u},\mathbf{x}} = \begin{pmatrix} (r_x - v + \tilde{v})f/r_z \\ (r_y + u - \tilde{u}f/r_z \end{pmatrix} = \mathbf{C}_{\mathbf{u}} + \frac{f}{r_z}\begin{pmatrix} \tilde{v} \\ -\tilde{u} \end{pmatrix} \quad (7)$$

Hence we have

$$(\mathbf{x} - \mathbf{N}) \perp (\mathbf{x} - \mathbf{C}_{\mathbf{u},\mathbf{x}})$$

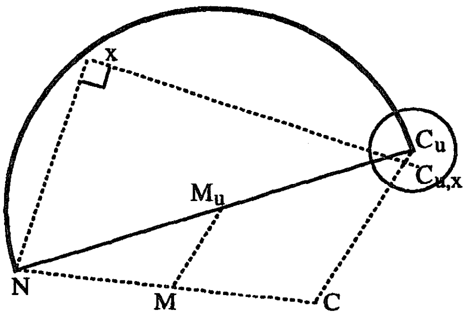


Figure 1: The iso-velocity circle.

From eq. (7) we get

$$||\mathbf{C}_{\mathbf{u},\mathbf{x}} - \mathbf{C}_{\mathbf{u}}|| = ||\frac{f}{r_z}(\tilde{v}, -\tilde{u})^T|| = ||\frac{f}{r_z}\frac{1}{f^2}\mathbf{x}\mathbf{x}^T(-r_y, r_x)^T||$$

$$\leq ||\frac{1}{f^2}\mathbf{x}\mathbf{x}^T||\ ||\mathbf{C}|| \leq \tau^2||\mathbf{C}||$$

This completes the proof. ∎

As illustrated in Fig. 1, the IVPs are located approximately on a circle defined by $\mathbf{N}$ and $\mathbf{C}_{\mathbf{u}}$. The former is independent of $\mathbf{u}$ and the latter is not. The circle is centered at $\mathbf{M}_{\mathbf{u}} = (\mathbf{N} + \mathbf{C}_{\mathbf{u}})/2$ and has a radius $R_{\mathbf{u}} = ||\mathbf{N} - \mathbf{C}_{\mathbf{u}}||/2$. We denote the IVC as $\{\mathbf{N}, \mathbf{C}_{\mathbf{u}}\}$ or $\{\mathbf{M}_{\mathbf{u}}, R_{\mathbf{u}}\}$.

Theorem 2 shows that the relative uncertainty of $\mathbf{C}_{\mathbf{u}}$ has an upper bound proportional to the square of the FOV value (i.e., $\max||\mathbf{x}||/f$) and the magnitude of the center of rotation. Therefore, the uncertainty of both the center and the radius of the IVC is $\tau^2||\mathbf{C}||/2$.

Points on the Iso-Velocity Circle

Since eq. (2) corresponds to $\tilde{u} = \tilde{v} = 0$ in eqs. (5) and (6), the IVPs of MF_{para} lie exactly on the IVC . For ease of discussion, we now concentrate on the study of the behavior of the IVPs of MF_{para} rather than those of MF_{pers}.

Exactly speaking, the IVPs can be located on only half of the circle, specified by $\mathbf{N}$ and $\mathbf{C}_{\mathbf{u}}$, because of the positive-definiteness of depth Z. Let $\mathbf{x}_{\mathbf{u}}$ denote an image point with velocity $\mathbf{u}$. The position of $\mathbf{x}_{\mathbf{u}}$ on the IVC $\{\mathbf{N}, \mathbf{C}_{\mathbf{u}}\}$ encodes the depth information $Z(\mathbf{x}_{\mathbf{u}})$.

Theorem 3. If at image position $\mathbf{x}_{\mathbf{u}}$ the image velocity of MF_{para} is $\mathbf{u}$ and the depth value is Z, then on the IVC $\{\mathbf{M}_{\mathbf{u}}, R_{\mathbf{u}}\}$ the angular position of $\mathbf{x}_{\mathbf{u}}$ is $\angle \mathbf{x}_{\mathbf{u}}\mathbf{M}_{\mathbf{u}}\mathbf{N} = 2\arctan2(t_z/Z, r_z)$.

Proof. We know that eq. (2) can be written as eq. (4) or

$$\xi = \frac{(t_x + it_y) + (r_y + u - i(r_x - v))Z}{t_z - ir_z Z} f$$

Without loss of generality, we assume vector $\mathbf{C_u N}$ points to the positive direction of the x-axis. (We can always achieve this configuration by rotating the viewer-based coordinate system around the optical axis.) Let $\mathbf{N} = B + C + iA$ and $\mathbf{C_u} = B - C + iA$, where $C > 0$. Then the above equation reduces to

$$\xi = iA + \frac{(B+C)t_z - i(B-C)r_z Z}{t_z - ir_z Z}$$

Since the IVC is centered at $\mathbf{M_u} = B + iA$ we have

$$\angle \xi \mathbf{M_u N} = \angle(\xi - \mathbf{M_u}) = \angle \frac{C(t_z + ir_z Z)}{t_z - ir_z Z}$$

Because C is assumed to be positive, the angular position of $\mathbf{x_u}$ is

$$\angle \mathbf{x_u M_u N} = 2\mathrm{arctan2}(t_z/Z, r_z)$$

This completes the proof. ∎

It is worth emphasizing that the angular position of $\mathbf{x_u}$ does not change with $\mathbf{u}$. As Z increases from 0 to ∞, $\mathbf{x_u}(Z)$ moves in the counterclockwise (clockwise) direction from $\mathbf{N}$ to $\mathbf{C_u}$ if t_z and r_z have the same (opposite) sign. When Z is small, θ changes linearly with Z; as Z becomes bigger, θ gradually approaches the saturation value π. As a result, the angular variation of $\mathbf{x_u}$ does not depend solely on the depth variation.

Obviously, the co-circular property is a necessary condition and does not mean that points on the IVC have the same velocity. Suppose $\mathbf{x}$ is a point on the IVC $\{\mathbf{N}, \mathbf{C}_\alpha\}$ with angular position θ, then $\mathbf{u}(\mathbf{x}) = \alpha$ only if $Z(\mathbf{x}) = t_z \tan\theta/r_z$. For this reason, it is possible that no point on $\{\mathbf{N}, \mathbf{C}_\alpha\}$ has image velocity $\mathbf{u}(\mathbf{x}) = \alpha$. Especially, the motion field can have no fixed points.

Degenerate Cases

Up to now we have assumed that $t_z r_z \neq 0$; otherwise the IVC does not exist. Here we discuss the degenerate cases where either t_z or r_z vanishes.

Theorem 4. (Iso-Velocity Points Theorem: Degenerate cases). **(A)** If $r_z = 0, t_z \neq 0$, then all the IVPs $\mathscr{X}_\mathbf{u} = \{\mathbf{x} : \mathbf{u}(\mathbf{x}) = \mathbf{u}, \mathbf{x}/f \leq \tau\}$ of MF_{pers} lie approximately on the straight line passing through $\mathbf{N}$ with direction $\mathbf{d_u} = \mathbf{u} - (r_y, -r_x)^T$ in the sense that for any $\mathbf{x} \in \mathscr{X}_\mathbf{u}$, there is a point $\mathbf{d_{u,x}}, \|\mathbf{d_{u,x}} - \mathbf{d_u}\| \leq \tau^2\|\mathbf{r}\|$, such that the triple points $\mathbf{N}, \mathbf{N}+\mathbf{d_{u,x}}, \mathbf{x}$ are collinear. **(B)** If $t_z = 0, r_z \neq 0$, then all the IVPs $\mathscr{X}_\mathbf{u} = \{\mathbf{x} : \mathbf{u}(\mathbf{x}) = \mathbf{u}, \|\mathbf{x}\|/f \leq \tau\}$ of MF_{pers} lie approximately on the straight line passing through $\mathbf{C_u}$ with direction $\mathbf{d} = (-t_y, t_x)^T$ in the sense that for any $\mathbf{x} \in \mathscr{X}_\mathbf{u}$, there is a point $\mathbf{C_{u,x}}, \|\mathbf{C_{u,x}} - \mathbf{C_u}\| \leq \tau^2\|\mathbf{C}\|$, such that the triple points $\mathbf{C_{u,x}}$, $\mathbf{C_{u,x}} + \mathbf{d}, \mathbf{x}$ are collinear.

Proof. Similar to that of Theorem 2. ∎

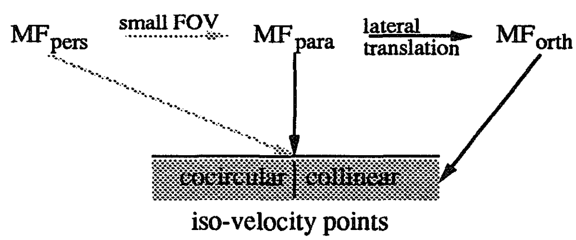
For a 3D motion with either lateral translation or lateral rotation, the IVC degenerates to the *iso-velocity line* (IVL). When $r_z = 0$ the IVLs intersect at $\mathbf{N}$; when $t_z = 0$ the IVLs are perpendicular to $\mathbf{t}$. As indicated before, orthographic projection corresponds to para-perspective projection plus lateral translation. Hence the iso-velocity points of MF_{orth} are (exactly) collinear and the IVLs are perpendicular to the translational direction.

Numerical Evaluations

To gain some empirical understanding, we solve the motion field equation for a sequence of depth values. Without loss of generality we assume $f = 1$. For motion $\mathbf{t} = (2, 0, 10)^T, \mathbf{r} = (-3, -5, 10)^T$, the IVCs intersect at $\mathbf{N} = (.2, 0)^T$. The three IVCs shown in Fig. 2 (left) correspond to different velocities $\mathbf{u}_1 = (0,0)^T, \mathbf{u}_2 = (4,1)^T, \mathbf{u}_3 = (8,-1)^T$. For each and every velocity $\mathbf{u}_j$, we compute a set of image positions $\mathscr{X}_j = (\mathbf{x}_{j,k})$ corresponding to $Z_k = t_z/r_z \tan\theta_k/2$, where $\theta_k = k\pi/10, k = -9, -8, \cdots, 8, 9, 10$, by solving eq. (1). (Somewhat remarkably, the solution is unique for each Z.) Points in $\mathscr{X}_j, j = 1, 2, 3$ are denoted by $\circ, +, \times$ respectively. We can see that the quadratic term in eq. (1) makes the IVPs deviate more or less from the IVCs. The deviations can hardly be perceived in the vicinity of the origin. In Fig. 2 (right) we show the angle $\theta_{j,k} = \angle \mathbf{x}_{j,k}\mathbf{M}_j\mathbf{N}$ vs. depth Z_k. Without the quadratic term in eq. (1), θ should be an arctangent function of Z.

Summary

For a given image velocity, the contributions from motion and depth can be segregated; motion determines the circle which is the set of feasible image positions having that velocity and depth affects the relative position on that circle. Our results obtained so far are illustrated in the following diagram where the gray arrow implies approximation.



Ego-Motion Recovery

The recovery of ego-motion/structure from noisy measurements of motion field is an important problems in machine vision which has been intensively studied in the past two decades. Thanks to the Iso-Velocity Point Theorem, the task can be accomplished by first estimating the five motion parameters $\mathbf{N}, \mathbf{r}$. We will show two important results: (A) the IVCs can be estimated from *pairs* of IVPs and (B) the problem of ego-motion recovery reduces to solving two systems of linear equations.

Computing IVCs

An IVC can be estimated from a set of three or more IVPs. Problem arises when the motion field may have only pairs of IVPs or in other words, there may only be two-fold overlapping in the velocity space (see Fig. 3 (upper-right) for an appreciation of the uv-space). As a matter of fact, the IVCs can be computed from (three or more) pairs of IVPs. Let $\mathbf{u}^{\perp} = (-v, u)^T, \lambda = f/(2r_z)$. The center of an IVC can be written as (see Fig. 1) $\mathbf{M_u} = \mathbf{M} + \lambda \mathbf{u}^{\perp}$, where λ and $\mathbf{M} = (M_x, M_y)^T$ are to be solved. Suppose $\mathbf{x_u} = (x, y)^T$ and $\mathbf{x'_u}$ are a pair of IVPs and $\bar{\mathbf{x}}_\mathbf{u} = (\mathbf{x_u} + \mathbf{x_u}')/2$. Then we have

$$\mathbf{M_u} - \bar{\mathbf{x}}_\mathbf{u} \perp \mathbf{x_u} - \mathbf{x'_u}$$

Therefore the inner product of LHS and RHS is zero:

$$\begin{pmatrix} M_x - \lambda v - (x + x')/2 \\ M_y + \lambda u - (y + y')/2 \end{pmatrix} \cdot \begin{pmatrix} x - x' \\ y - y' \end{pmatrix} = 0$$

Let $\mathbf{w} = (x - x', y - y', (y - y')u - (x - x')v)$ and $h = \frac{1}{2}(x^2 + y^2 + x'^2 + y'^2)$. Then we get

$$\mathbf{w}(M_x, M_y, \lambda)^T = h \qquad (8)$$

With each pair of IVPs providing such a constraint, $\mathbf{M}$ and λ can be solved from at least 3 pairs of IVPs. In practice, we often have a large number of IVP pairs and thus we can resort to a Least Squares (LS) estimator or a robust estimator. In the following we use the LS method. In order to treat all sample pairs equally, eq. (8) needs to be normalized before computing pseudo-inverse.

Computing Ego-Motion

After been solved for, $\mathbf{M}$ and λ may be used to compute $\mathbf{N}$ and $\mathbf{C}$. The idea is *to remove the motion field component induced by camera roll* which is now known. By virtues of Theorem 4, the resulting motion field have IVPs collinear with the node of translation. By estimating the intersection of (two or more) straight lines passing through pairs of IVPs we can locate $\mathbf{N}$. Then we can solve for $\mathbf{r}$ easily, noting that $\mathbf{C} = 2\mathbf{M} - \mathbf{N}$.

Experiments

With the 256 × 256 depth map shown as an intensity image in Fig. 3 (upper-left), a motion field $\mathbf{u} = \mathbf{u}(\mathbf{x})$ is generated for the camera to move at $\mathbf{t} = (-1, 0, 10)^T$ and $\mathbf{r} = (0, 1, 10)^T$. The focal length is 5 times the image size or FOV=11.42°. The interior orientation of the camera is a such that the optical axis passes through the image center. Each image grid point $\mathbf{x}$ is mapped to $\mathbf{u}(\mathbf{x})$ in the velocity space shown in Fig. 3 (upper-right) (note the folding). The IVPs are computed from the manifolds of constant u and constant v. In Fig. 3 (lower-left), conjugate IVPs are linked by line segments. We solve for $\mathbf{M}$, λ using the LS method and get $\mathbf{M} = (-53.54, 67.10)^T, \lambda = 60.95$ (in cells).

After removing the flow component induced by camera roll, the resulting flowfield has the IVPs shown in Fig. 3 (lower-right). Clearly the IVLs seem to intersect at a common point. Apply the LS technique again and we get $\mathbf{N} = (-126.85, 19.44)^T$ (in cells). More experiments can be found in (Yang 1992).

Concluding Remarks

We have seen that the image points with the same velocity provide useful information for motion perception. The IVPs corresponding to ego-motion can be distinguished from those corresponding to independently moving objects by taking into account contingency in the image space. While the latter indicate exterior motion(s), the former can be used for recovering ego-motion in an efficient, stable way.

References

Aloimonos, J. 1988. Shape from texture. *Biological Cyberbetics.* 58:345-360.

Gibson, J.J. 1950. The Perception of the Visual World. Boston, MA: Houghton Mifflin.

Horn, B.K.P. 1987. Motion fields are hardly ever ambiguous. *Int. Journal of Computer Vision.* 1:259-274.

Jepson, A.D. & Heeger, D.J. 1991. A fast subspace algorithm for recovering rigid motion. In Proc. IEEE Workshop on Visual Motion. 124-131. Princeton, NJ.

Jordan, D.W. 1987. Nonlinear Differential Equations. New York: Oxford University Press.

Koffka, K. 1935. Principles of Gestalt Psychology. New York: Harcourt Brace.

Longuet-Higgins, H.C. & Prazdny, K. 1980. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B.* 208:385-397.

Maybank, S.J. 1985. The angular velocity associated with the optical flow field arising from motion though a rigid environment. *Proceedings of the Royal Society of London A.* 401:317-326.

Marr, D. 1982. Vision. San Francisco: W.H. Freeman and Co.

Ohta, Y.K., Maenobu, K. & Sakai, T. 1981. Obtaining surface orientation from texels under perspective projection. In Proc. Int. Joint Conf. Artif. Intell. 746-751. Vancouver, Canada.

Ullman, S. 1979. The Interpretation of Visual Motion. Cambridge, MA: MIT Press.

Verri, A., Girosi, F. & Torre, V. 1989. Mathematical properties of the 2-D motion field: from singular points to motion parameters. *J. Opt. Soc. Amer.*

Waxman, A.M. & Ullman, S. 1985. Surface structure and 3-D motion from image flow: A kinematic analysis. *Int. Journal of Robotics Research.* 4(3):72-94.

Yang, Y. 1992. Ph.D. diss., Division of Applied Sciences, Harvard Univ. Forthcoming.

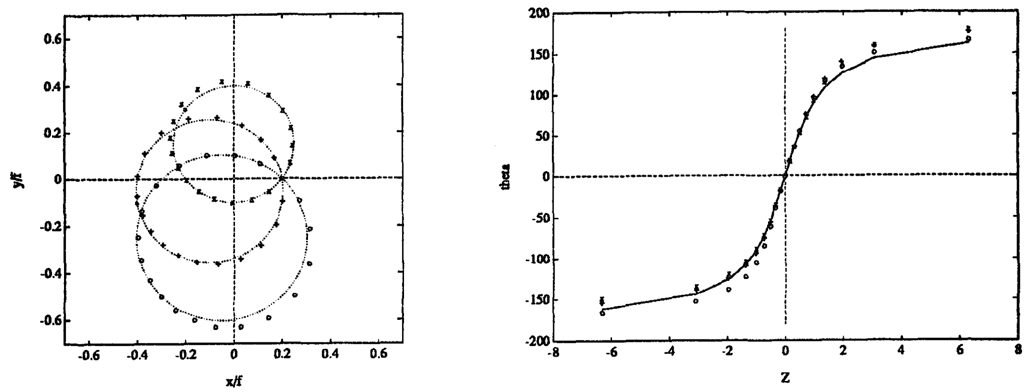


Figure 2: IVPs and IVCs (left) and angular positions (right).

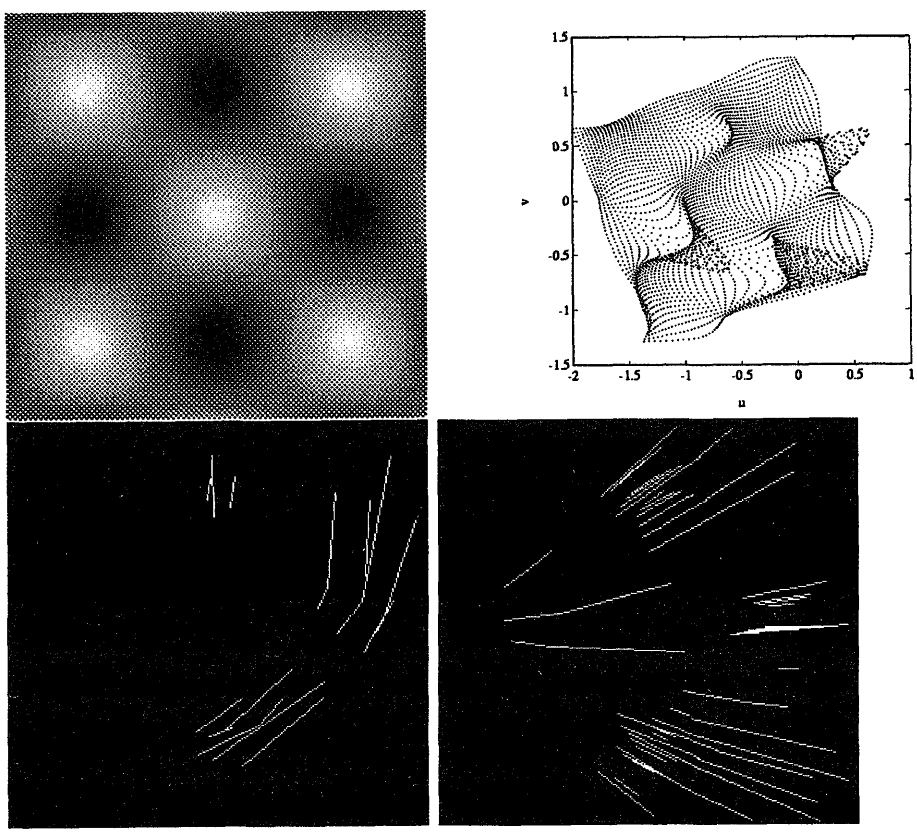


Figure 3: Computing the node of translation.