



# 关于视觉运动分析中凝视与跟踪作用的研究<sup>1)</sup>

杨敬安

(合肥工业大学人工智能应用研究所 合肥 230009)

**摘要** 提出主动观察者利用摄象机旋转,凝视并跟踪环境点使视场中心保持固定,能提供求解问题所需的许多有用信息,从而缩减待解问题的维数及运动参数数目,把不适定问题转变为适定问题,大大简化计算的复杂性,有效地进行3-D运动估计、自运动复原及TTC的计算.

**关键词** 凝视约束,主动自运动,法向流模式,单目主动观察者.

## 1 引言

近年来,仅用法向流场而不经光流计算的中间阶段,能够估计3-D平移运动<sup>[1]</sup>.对于纯旋转运动,最近也导出运动与法向流相关的线性等式,产生对纯旋转及纯平移运动的估计与结构计算的几种方法<sup>[2]</sup>,其中Negahdaripour使用正深度约束估计纯平移运动的FOE(Focus of Expansion); Aloimonos. J等对绕平移方向旋转的观察者成功地实现平移及旋转运动的估计<sup>[3,4]</sup>.然而,他们都用光流或图像强度函数的时空导数作为输入,并施加某些约束实现3-D运动估计,而没有着重研究凝视与跟踪对复原与理解主动自运动的重要作用.因而3-D运动估计仍较复杂,并且有时会出现较大误差.针对以上问题,本文提出基于凝视与跟踪的视觉运动分析方法.

对于单目主动观察者在3-D世界进行自由刚体运动,坐标系 $(X, Y, Z)$ 固定在以坐标原点 $O$ 为摄象机结点的观察者上,如果观察者相对景物的平移运动用 $(U, V, W)$ 、旋转运动用 $(\alpha, \beta, \gamma)$ 表示,则图像内任意点 $p(X, Y, Z)$ 的速度分量为

$$\dot{X} = -U - \beta Z + \gamma y, \quad (1a)$$

$$\dot{Y} = -V - \gamma X + \alpha Z, \quad (1b)$$

$$\dot{Z} = -W - \alpha y + \beta X. \quad (1c)$$

如图1所示的透视投影成像模型,图像平面平行于 $XY$ 平面,观察方向沿 $Z$ 轴正方向.这样,三维点 $p(X, Y, Z)$ 的图像位置 $p(x, y)$ 用关系式

$$(x, y) = (f X/Z, f Y/Z) \quad (2)$$

1)国家自然科学基金资助项目.

收稿日期 1995-06-19

确定. 式中  $f$  为成像系统焦距, 图像点  $p$  与 3-D 速度的关系等式从(1)式求得, 并由(2)式代替得

$$u = (-Uf + xW)/Z + \alpha(xy/f) - \beta(x^2/f + f) + \gamma y, \quad (3)$$

$$v = (-Vf + yW)/Z + \alpha(y^2/f + f) - \beta(xy/f) - \gamma x. \quad (4)$$

利用透视投影, 单目观察者能计算的参数限 5 个, 即 3 个旋转参数和 2 个平移参数. 现引入平移方向的坐标  $(x_0, y_0) = (Uf/W, Vf/W)$ , 并改写(3)和(4)式右边为平移及旋转分量之和, 即

$$u = u_{\text{trans}} + u_{\text{rot}} = (-x_0 + x)W/Z + \alpha(xy/f) - \beta(x^2/f + f) + \gamma y, \quad (5)$$

$$v = v_{\text{trans}} + v_{\text{rot}} = (-y_0 + y)W/Z + \alpha(y^2/f + f) - \beta(xy/f) - \gamma x. \quad (6)$$

由于我们仅能计算法向流, 即光流在梯度方向上的投影  $(n_x, n_y)$ , 并且在任何已知点仅能导出一个约束, 这样沿梯度方向的法向流向量值  $u_n$  由下式给出

$$u_n = un_x + vn_y$$

或

$$u_n = ((-x_0 + x)W/Z + \alpha(xy/f) - \beta(x^2/f + f) + \gamma y)n_x + ((-y_0 + y)W/Z + \alpha(y^2/f + f) - \beta(xy/f) - \gamma x)n_y. \quad (7)$$

此式表明, 被动观察者利用法向流计算运动的困难在于运动场合五个未知参数, 且在每个新点又引入一个未知量  $(-W/Z)$ , 故单目被动观察者不能直接计算运动参数.

## 2 凝视约束

如前所述, 设作刚性运动的主动观察者正在观察某环境点, 该点的图像  $(x, y)$  在视场中心  $((x, y) = (0, 0))$ . 若在很小时间间隔  $[t_1, t_2]$  内, 观察者运动保持恒定, 并且为能正确跟踪, 摄像机以旋转速度  $(\omega_x(t), \omega_y(t))$  ( $t \in [t_1, t_2]$ ) 分别绕其  $x$  和  $y$  轴旋转. 跟踪旋转把旋转流场  $(u_{tr}, v_{tr})$  加到现流场  $(u, v)$ , 其中

$$u = (-Uf + xW)/Z + \alpha(xy/f) - \beta(x^2/f + f) + \gamma y, \quad (8)$$

$$v = (-Vf + yW)/Z + \alpha(y^2/f + f) - \beta(xy/f) - \gamma x, \quad (9)$$

$$u_{tr} = \omega_x(xy/f) - \omega_y(x^2/f + f),$$

$$v_{tr} = \omega_x(y^2/f + f) - \omega_y(xy/f).$$

式中  $\omega_x, \omega_y$  为观察时刻的跟踪速度,  $Z$  是被跟踪点的深度. 如果  $(x_0, y_0) = (Uf/W, Vf/W)$  确定为 FOE, 并可改写成  $y_0/x_0 = V/U = T$ , 这是一种线性约束, 它把 FOE 的位置约束到过图像坐标系原点斜率为  $T$  的一条直线. 换句话说, 跟踪不仅能提供 FOE 所在的线  $y_0/x_0 = T$ , 也定义分离正、负旋转流场的直线  $y = (B/A)x = \frac{\beta + \omega_y}{\alpha + \omega_x}x$ , 这样便可把搜索模式减少到一维.

## 3 解平移运动过程与完全解旋转

### 3.1 解平移运动过程

所谓解平移指已知 FOE 的位置, 选择仅由旋转产生的法向流向量. 根据(7)式, 如果 FOE 的位置已知, 平移运动分量的方向也就已知, 并且位于过 FOE 的直线上. 垂直这些线的法向流向量仅含旋转分量. 若在点  $(x, y)$  处选择的梯度方向为  $((y_0 - y), (-x_0 + x))$ , 那么平移运

动分量和在梯度方向上向量的标积为0, 这种消去平移分量的方法称作解平移 (detranslation), 通常用于计算第三个旋转分量并进一步减少可能解的数目.

### 3.2 完全解旋转

如果前面的处理没能提供唯一解, 而是一组解  $\mathcal{S} = \{s_1, s_2, \dots, s_n\}$ , 候选自运动参数.  $s_i = (x_{0_i}, y_{0_i}, \alpha_i, \beta_i, \gamma_i)$ . 为消除与已知法向流场相矛盾的所有运动参数, 必须检查每个法向流向量, 这可用解旋转 (derotation) 技术实现. 对于每个  $\mathcal{S}$  的一组5个参数, 定义可能的 FOE 和旋转, 并且使3个旋转参数分别减去旋转分量 ( $u_{rot}, v_{rot}$ ), 对法向流向量解旋转. 由于 FOE 的光流方向已知, 故能确定法向流向量的可能方向, 并且每个点处的法向流向量准位于半平面上. 利用这种技术能够检查所有点的这些特性, 并且消去不能产生给定法向流场的解.

## 4 算 法

设刚性运动观察者能以速度  $(\omega_x, \omega_y)$  跟踪其图像在坐标系统原点处的某环境点, 那么利用以下跟踪算法输出观察者的运动, 其步骤如下:

- 1) 跟踪加速度提供一条 FOE 所在的直线  $y = Tx$ , 也给出比例系数  $(\alpha + \omega_x) / (\beta + \omega_y)$ ;
- 2) 利用前一步结果, 实现沿直线  $y = Tx$  的一维搜索, 找出 FOE 的某些解;
- 3) 前一步可提供一组  $S = \{(x_{0_1}, y_{0_1}), (x_{0_2}, y_{0_2}), \dots, (x_{0_n}, y_{0_n})\}$ ;
- 4) 第三步可能提供一组平移和旋转的候选解

$$\mathcal{S} = \{(x_{0_1}, y_{0_1}), (x_{0_2}, y_{0_2}), \dots, (x_{0_n}, y_{0_n}, A_n, B_n, C_n)\}.$$

为消除不可能解, 必须实现完全的解旋转, 以便检查每单个法向流向量和运动参数间的一致性.

## 5 实验结果及结论

本文采用 Swivel 图像系统模拟物体及摄像机运动, 计算综合图像上的运动物体. 为分析算法性能, 计算图像中心法向流值的精度, 逐点确定已知光流值在计算的梯度方向上投影  $v_{act}$ . 法向流值的误差 (error) 定义为  $v_{act}$  和法向流值  $v_{meas}$  间标准差 ( $error = (v_{act} - v_{meas}) / v_{act} \%$ ). 用此方法, 计算得到 36.14% 的平均误差以及跟踪过程开始时运动序列 79.64% 标准偏差, 此误差与实图像中出现的噪声相近. 图2所示的物体在中心为  $u = 0.004$ ,  $v = 0.02$  焦点单位随图像运动向  $U/W = 4$  和  $V/W = 2$  的方向运动. 我们在100帧以上的图像序列对其跟踪, 并在图像中心  $9 \times 9$  像素区域的10个方向计算光流测量值. 当用其并行平移测验第一个模型时, 使用 0.0002 焦点单位的阈值, 算法收敛非常快, 仅需 2-3 次迭代. 对于较大的旋转, 算法仍收敛很快, 例如物体离摄像机 25 个单位, 每次以  $U = 0.1$ ,  $V = 0.05$ ,  $W = 0.025$  单位的平移速度运动, 此方法收敛于每单位时间绕  $x, y$  和  $z$  轴高达  $0.3^\circ$  旋转. 图3示出两个最大的法向流向量序列, 其中 (a) 无旋转; (b) 旋转为  $\omega_x = 0.1^\circ$ ,  $\omega_z = 0.1^\circ$ . 由于利用该模型的并行平移估计及超过百步的连续跟踪, 使 FOE 值误差小于 15%, 如在不旋转时, 计算得到  $U/W = 4.21$ ,  $V/W = 1.97$  的 FOE. 这些实验证明本文计算 3-D 运动算法与技术能耐强噪声, 即使输入法向流有大量噪声, 仅使用法向流跟踪即可成功实现, 并且跟踪加速度可用于参数的鲁棒估计.

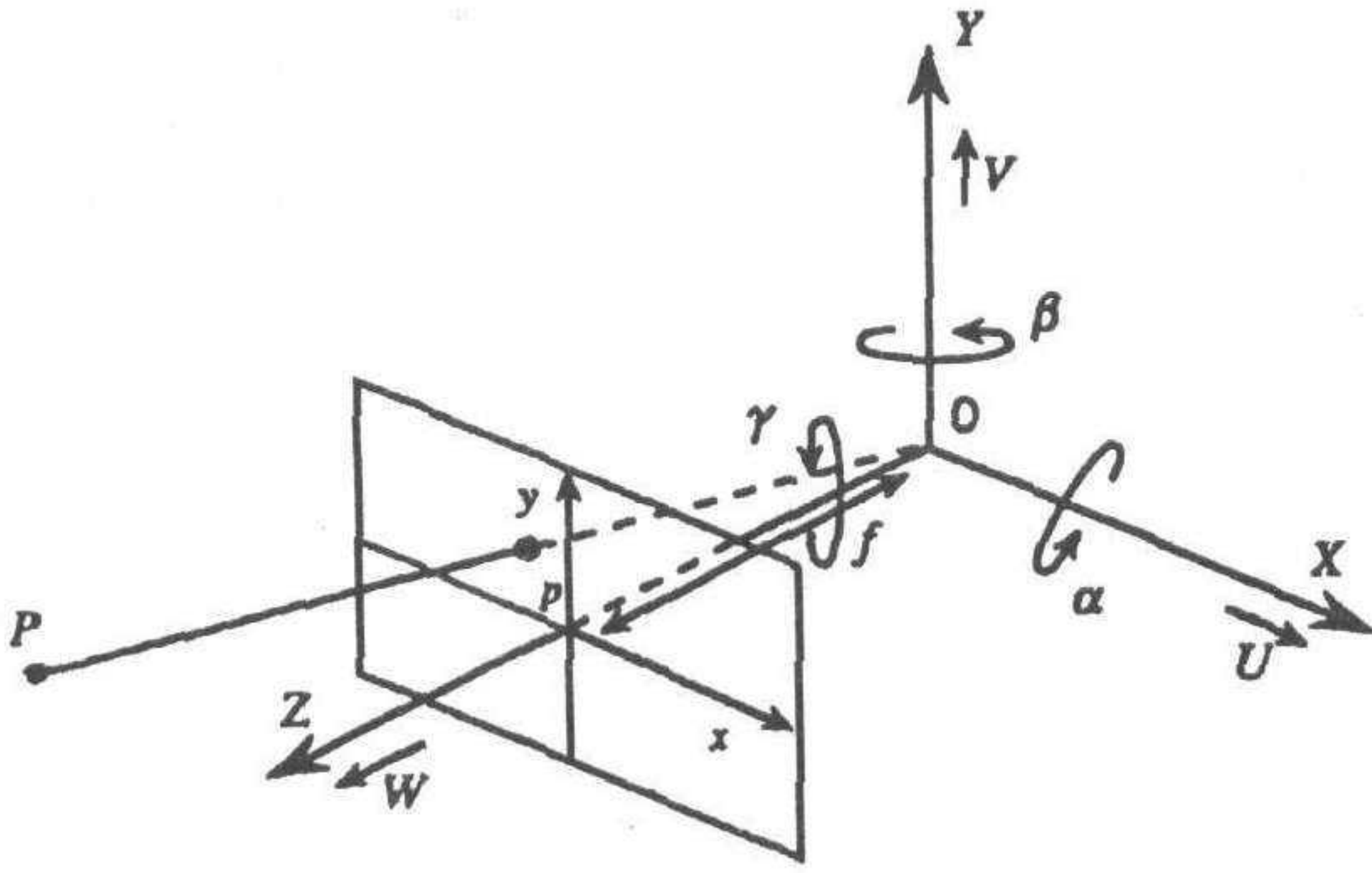


图1 以摄像机为中心的几何成像与运动表示

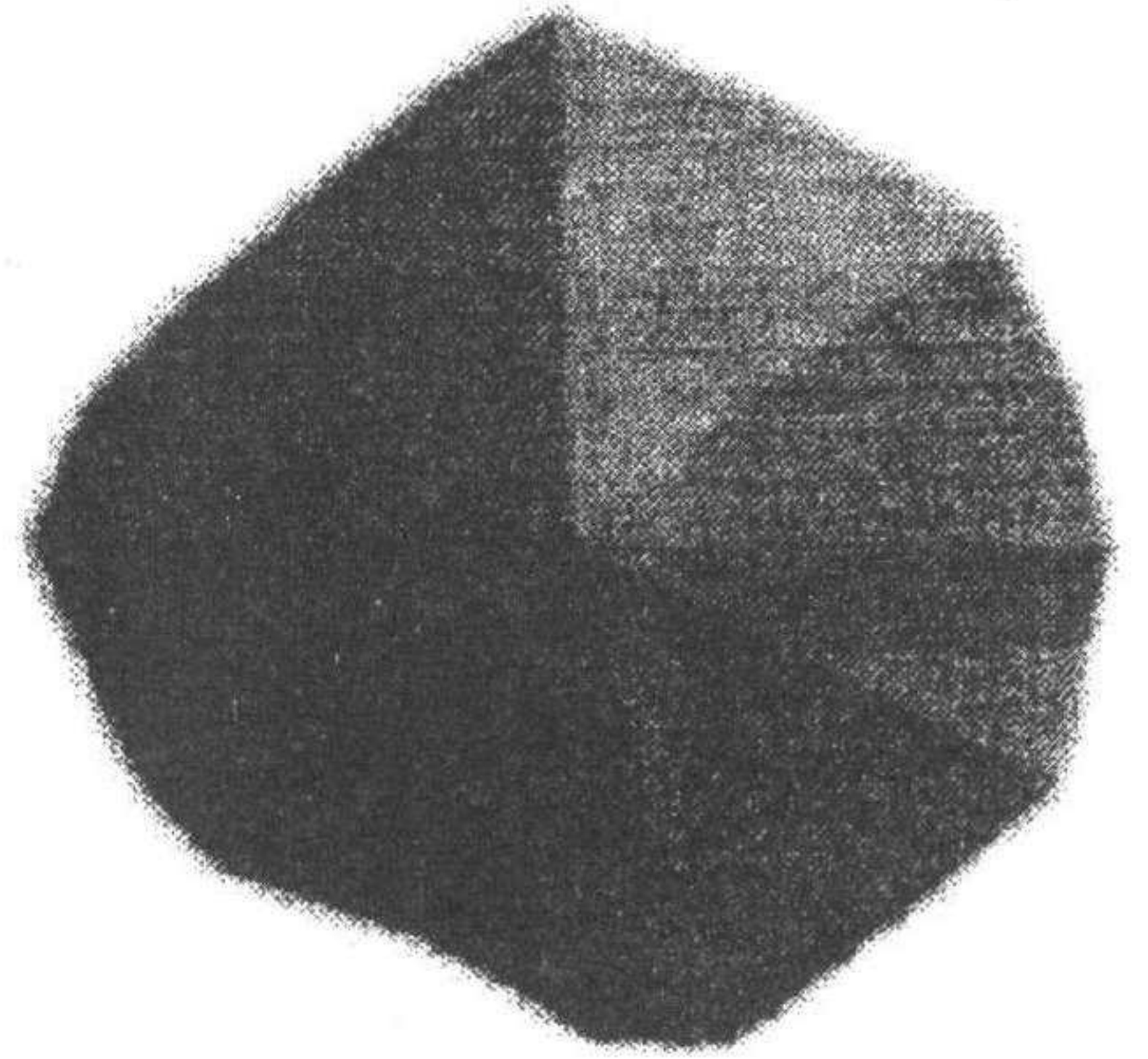


图2 跟踪使用图象序列的第一帧

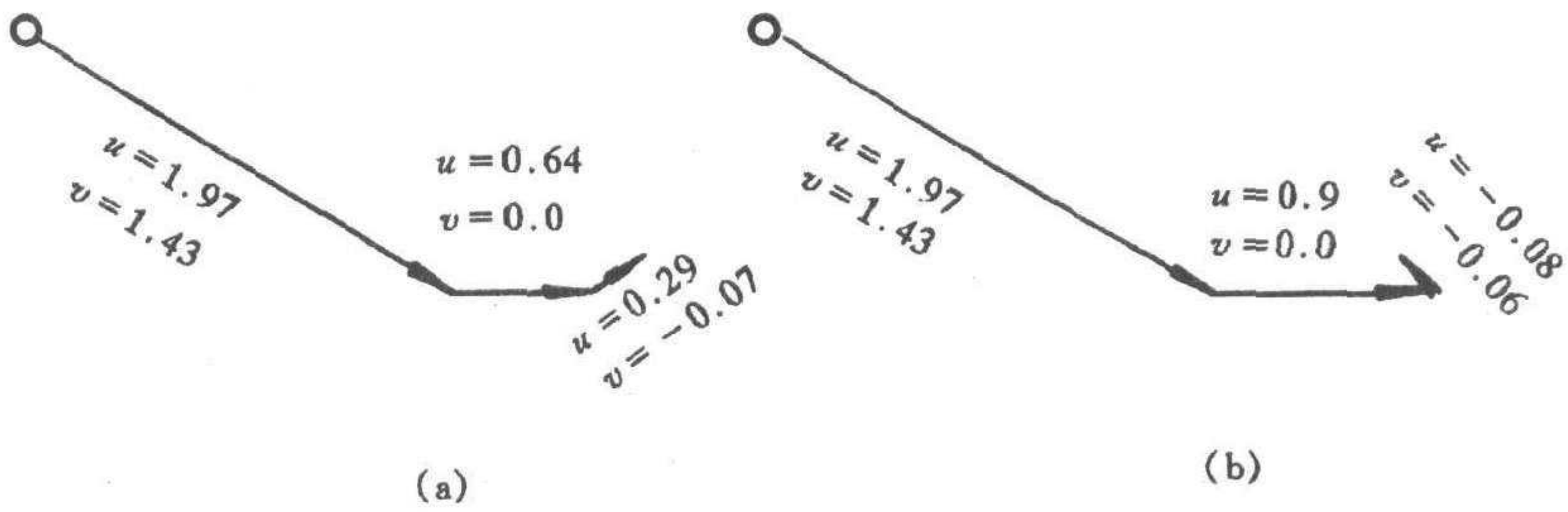


图3 最大法向流向量

这里对综合图像进行实验, 产生了不同3-D 运动参数的法向流场, 区间内深度值及梯度方向任意选择. 利用假设 FOE 所在的直线  $y = Tx$  和跟踪约束提供的直线  $y = (B/A)x$  测验模式匹配. 实验选择的图像为  $100 \times 100$ , FOE 位置  $[-40, -40]$ , 旋转分量比率为  $A : B : \gamma = 1 : -1 : 15$ . 图4示出  $\gamma$ -模式拟合到  $\gamma$ -向量, 正法向流点用淡色表示, 而负法向流点为深黑色, 法向流向量长度受干扰达50% 仍不会影响算法求出正确解.

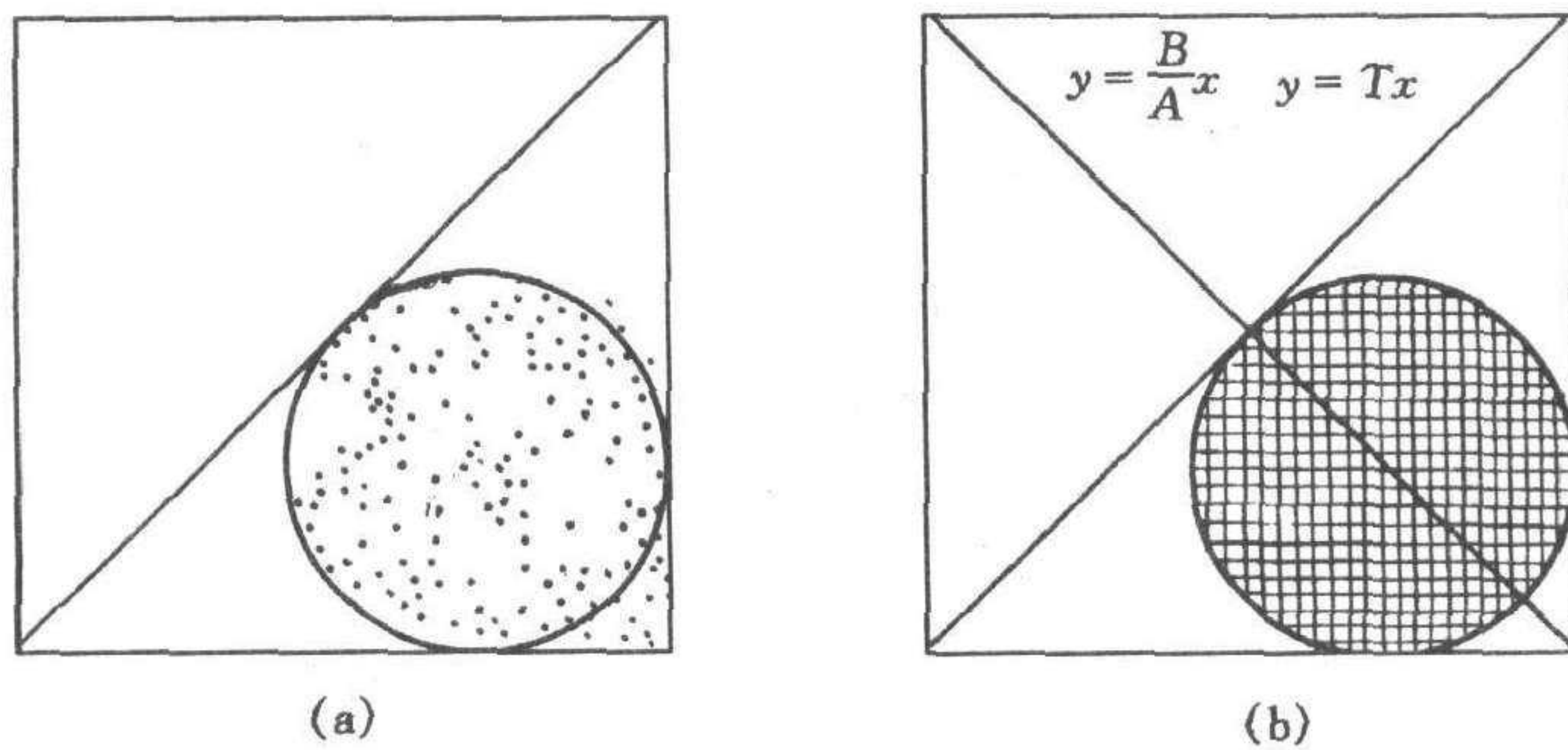


图4  $\gamma$ -模式拟合到  $\gamma$ -向量 (a)正和负  $\gamma$ -向量; (b) $\gamma$ -模式的拟合: 利用凝视约束找出 FOE 所在的线  $y = Tx$  和分离旋转分量的线  $y = (B/A)x$

本文提出的算法由三个模型组成:①用跟踪与凝视的组合,估计与图像平面平行的运动分量;②跟踪用于计算垂直的运动分量和 FOE;③利用这些模型的输出估计碰撞时间. 算法用主动视觉系统实现且具有很好的收敛性,对综合图像的测验结果证明算法能有效地进行运动估计,避开复杂的计算,求解某些对于被动观察者为不适定的一类问题.

本文部分工作完成于美国马里兰大学自动化研究中心 CV 实验室.

### 参 考 文 献

- 1 Spetsakis M E, Aloimonos J. Optimal visual motion estimation. *IEEE Trans. PAMI*, 1992, **14**(9): 959—964
- 2 Bandopadhyay A *et al.* Egomotion perception using tracking. *Computational Intelligence*, 1991, **7**:39—47
- 3 Aloimonos J *et al.* Active vision. *Inter. Journal of Computer Vision*, 1988, **2**:333—356
- 4 Brunnstrom K *et al.* Active fixation for scene exploration. *Inter. Journal of Computer Vision*, 1996, **17**(2):137—162

## A STUDY ON THE ROLE OF FIXATION AND TRACKING IN VISUAL MOTION ANALYSIS

YANG JINGAN

(*Institute of AI, Hefei University of Technology, Hefei 230009*)

**Abstract** This paper proposes an algorithm for fixation and tracking and gaze control through camera rotation. An active observer with ability to control its gaze and keep an environmental feature stationary on its image plane can thus transform some ill-posed problems into well-posed problems. Problems of motion estimation, egomotion recovery, and estimation of time to collision can be solved very efficiently due to reduced dimensions, less motion parameters, and absence of complicated computation.

**Key words** Fixation constraints, active egomotion, normal flow patterns, active monocular observer.