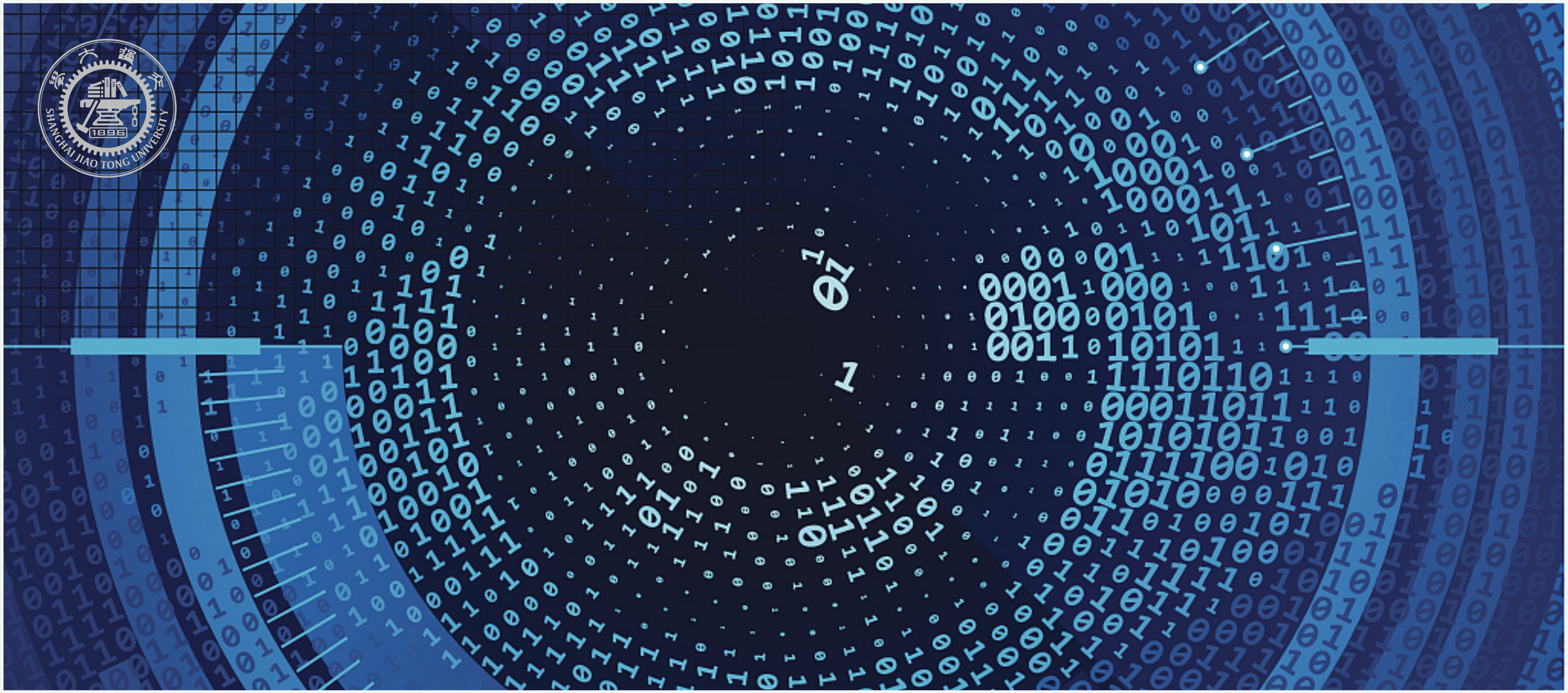


上海交通大学

计算机视觉

教师：赵旭

班级：AI4701

2024 春

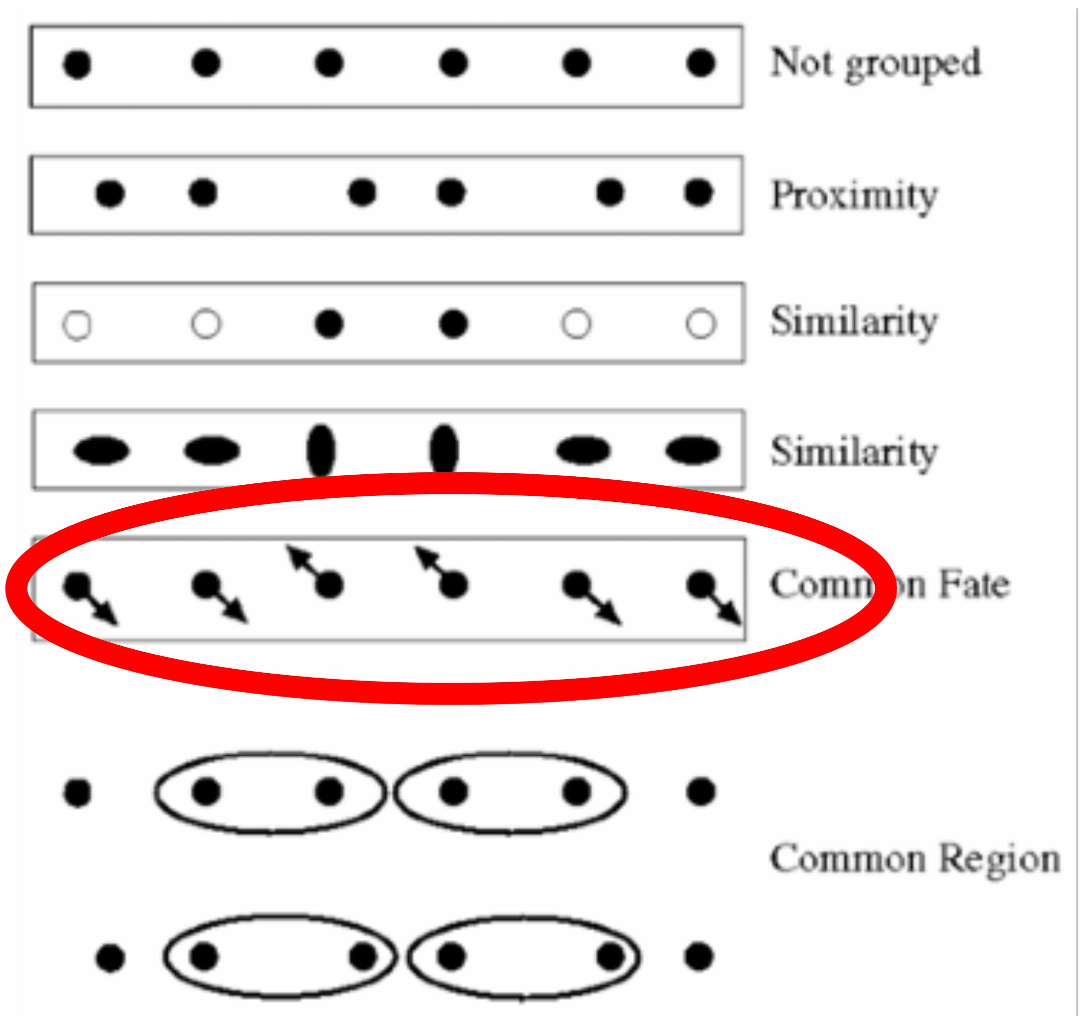
11. 运动特征-光流

主要内容

- 运动场
- 光流
- Lucas-Kanade 算法

运动和感知组织

- 有些情况下，运动是唯一线索



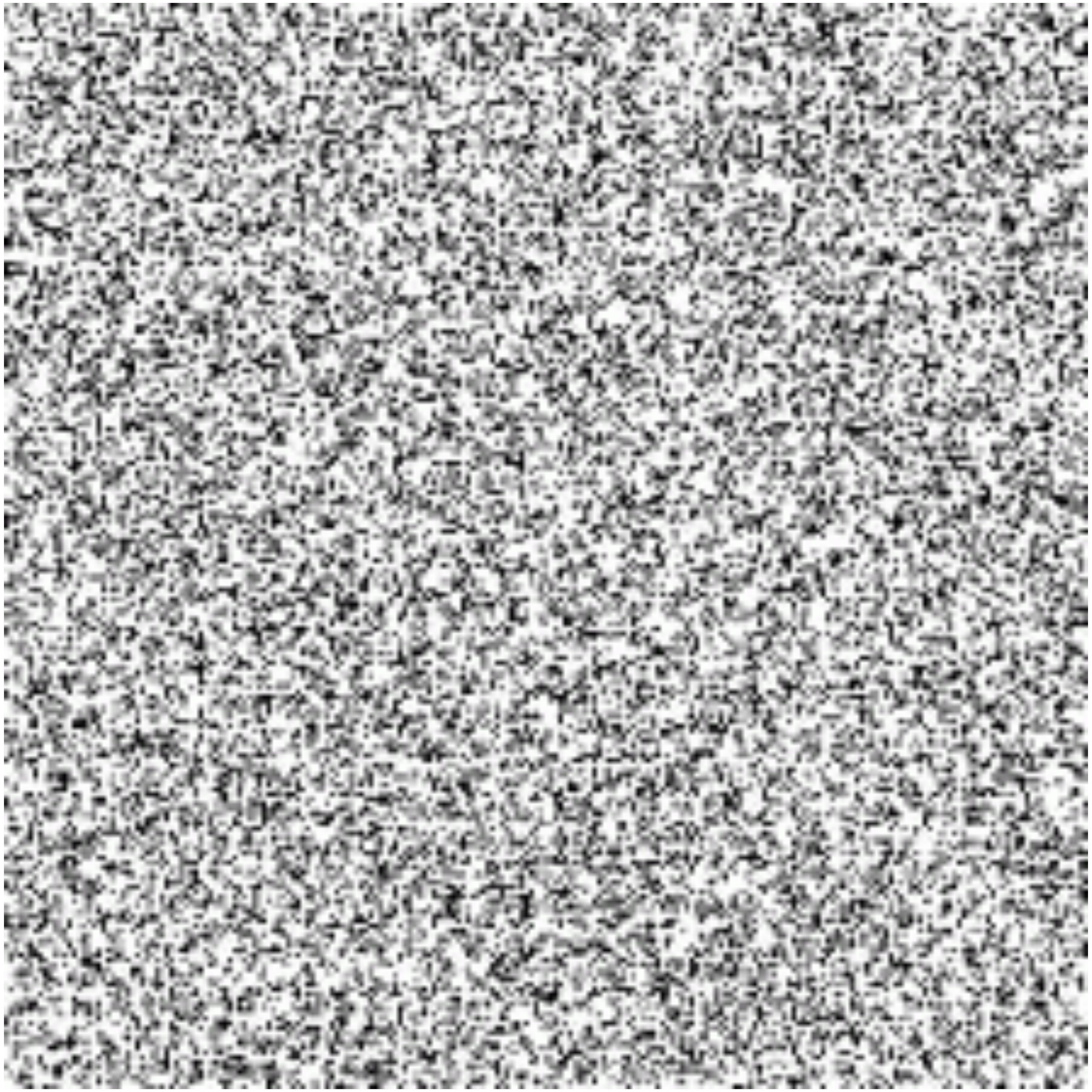
运动与格式塔聚类法则

slides from Grauman

运动和感知组织

- 甚至很少的运动数据就可以唤起强烈的感知

运动和感知组织

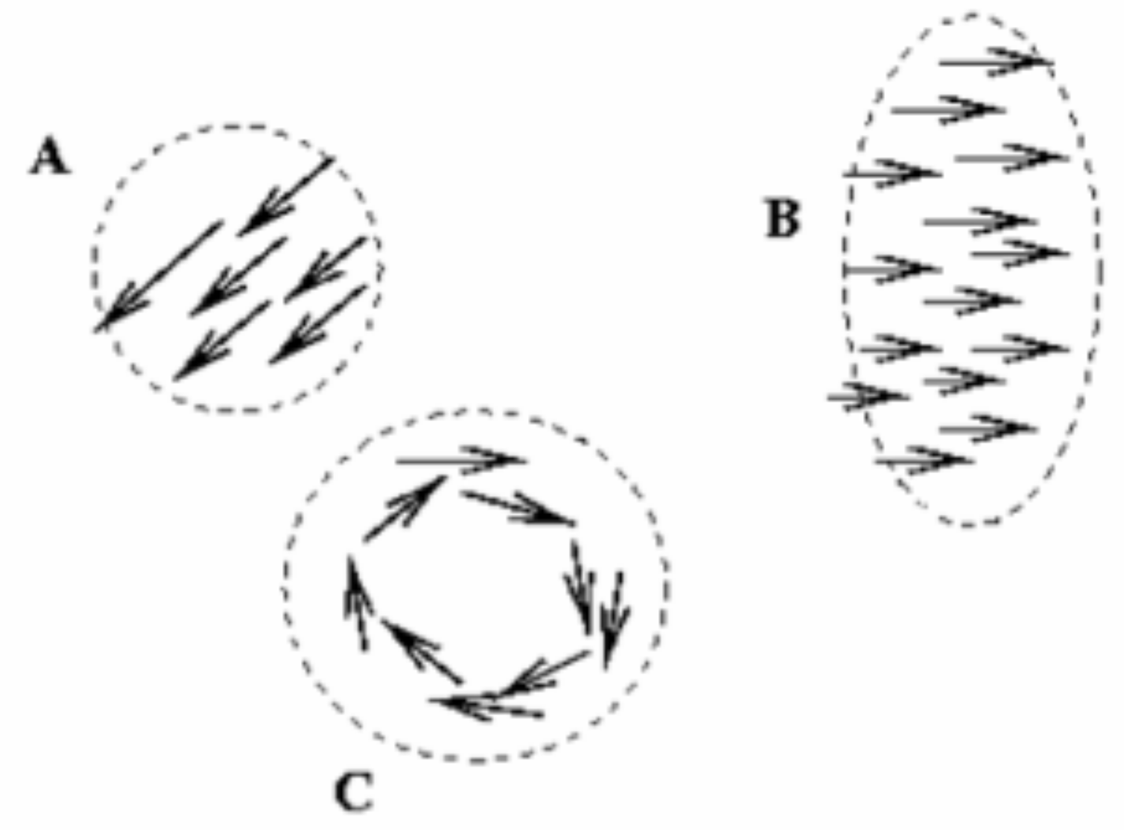


运动信息可以做什么?

- 估计三维结构
- 物体分割
- 学习动态模型
- 识别动作和事件
- 提升视频质量 (运动补偿)

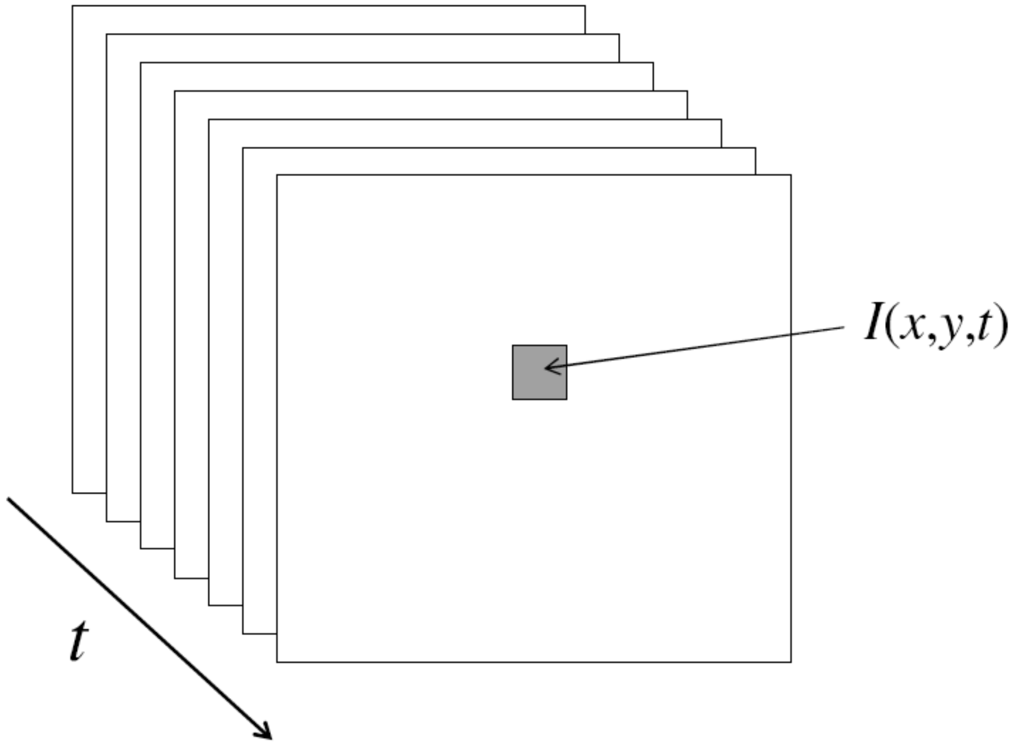
举例: 运动分割

- 运动分割: 将视频分割为多个具有运动一致性的运动物体



运动: 视频角度

- 视频：时序图像序列

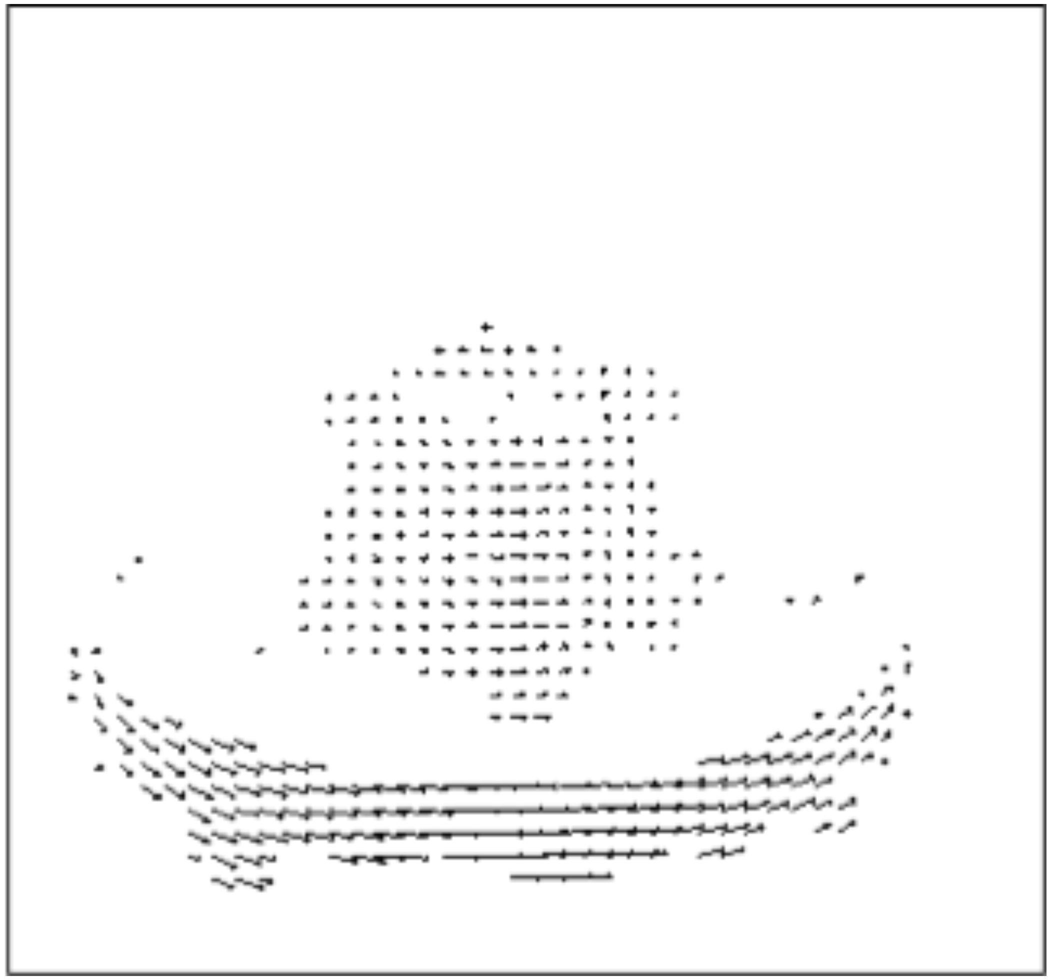


运动估计的方法

- **基于稀疏特征的方法**: 点对应
  - 提取视觉特征 (角点、纹理区域)，然后跟踪这些特征
  - 虽然是稀疏的运动场，但是更易达成鲁棒的跟踪
  - 适合场景：图像运动较大
- **密集的像素级运动估计**
  - 从图像像素的时空亮度变化中恢复图像运动
  - 不需要计算特征
  - 密集运动场，对表观的变化更敏感
  - 适合场景：图像运动较小 (< 10 像素)

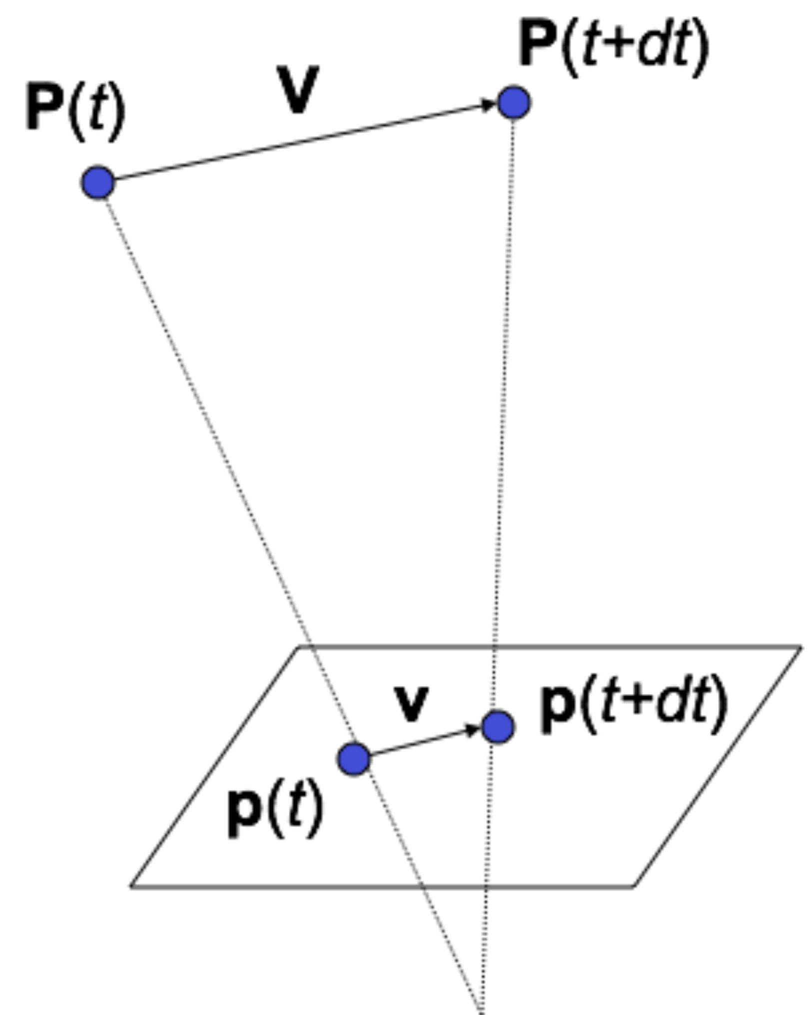
运动场

- 二维图像运动场是三维场景运动的二维图像投影



运动场和视差

- **P**(*t*)： 三维运动点
  - 速度: **V** = d**P**/d*t*
- **p**(*t*) = (*x*(*t*),*y*(*t*))： **P** 的图像投影
  - 图像点的运动速度 **v** = （v_x = d*x*/d*t*,　v_y = d*y*/d*t*)
  - 这些构成了图像运动场



运动场和视差

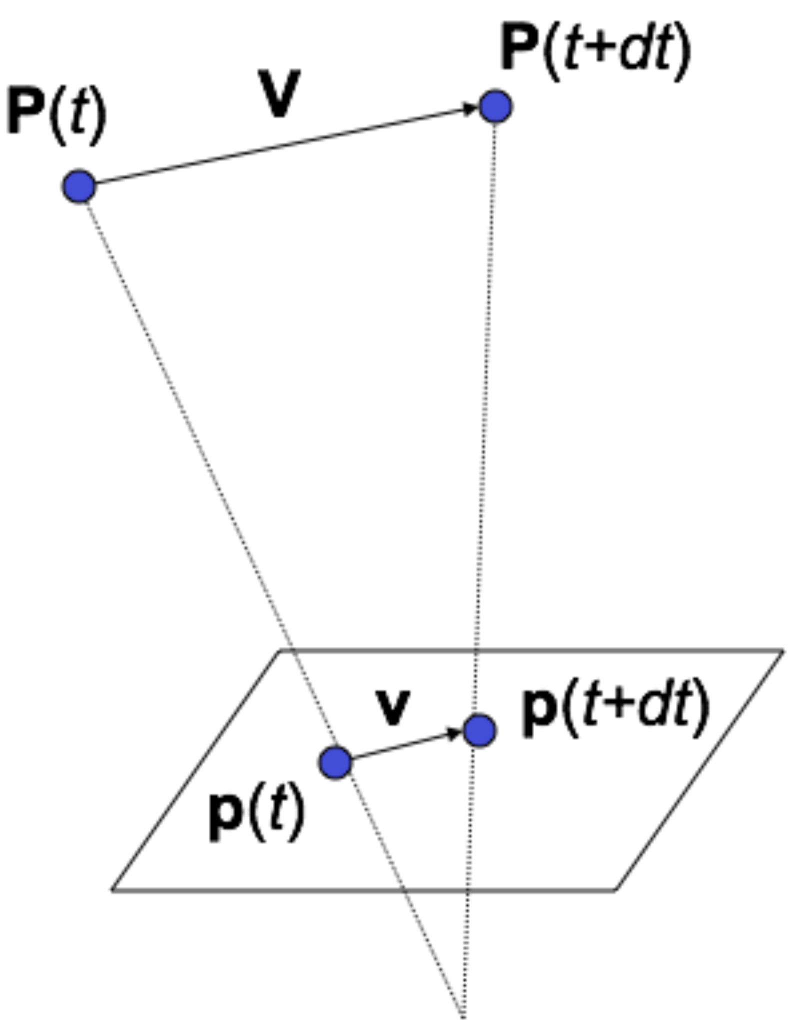
$$\mathbf{V} = (V_x, V_y, V_Z) \quad \mathbf{p} = f\frac{\mathbf{P}}{Z}$$

为了求解 **v**, 对p关于t求导:

$$\mathbf{v} = f\frac{Z\mathbf{V} - V_z\mathbf{P}}{Z^2}$$

$$v_x = \frac{fV_x - V_z x}{Z} \qquad v_y = \frac{fV_y - V_z y}{Z}$$

图像运动是三维运动和三维点深度的函数



Quotient rule:

D(f/g) = (g f’ – g’ f)/g^2

运动场和视差

纯平移: V 处处为常数

$$v_x = \frac{fV_x - V_z x}{Z}$$
$$v_y = \frac{fV_y - V_z y}{Z}$$

$$\mathbf{v} = \frac{1}{Z}(\mathbf{v}_0 - V_z \mathbf{p}),$$
$$\mathbf{v}_0 = (fV_x, fV_y)$$

Vz 非零: 所有的运动向量指向（或远离）v_0,
平移方向的消影点

Vz 为零: 运动与图像平面平行，所有的运动向量平行

运动视差

- 运动向量的长度与深度成反比关系，距离摄像头越近，运动越快（在图像平面内）

$$v_x = \frac{fV_x - V_z x}{Z}$$

$$v_y = \frac{fV_y - V_z y}{Z}$$

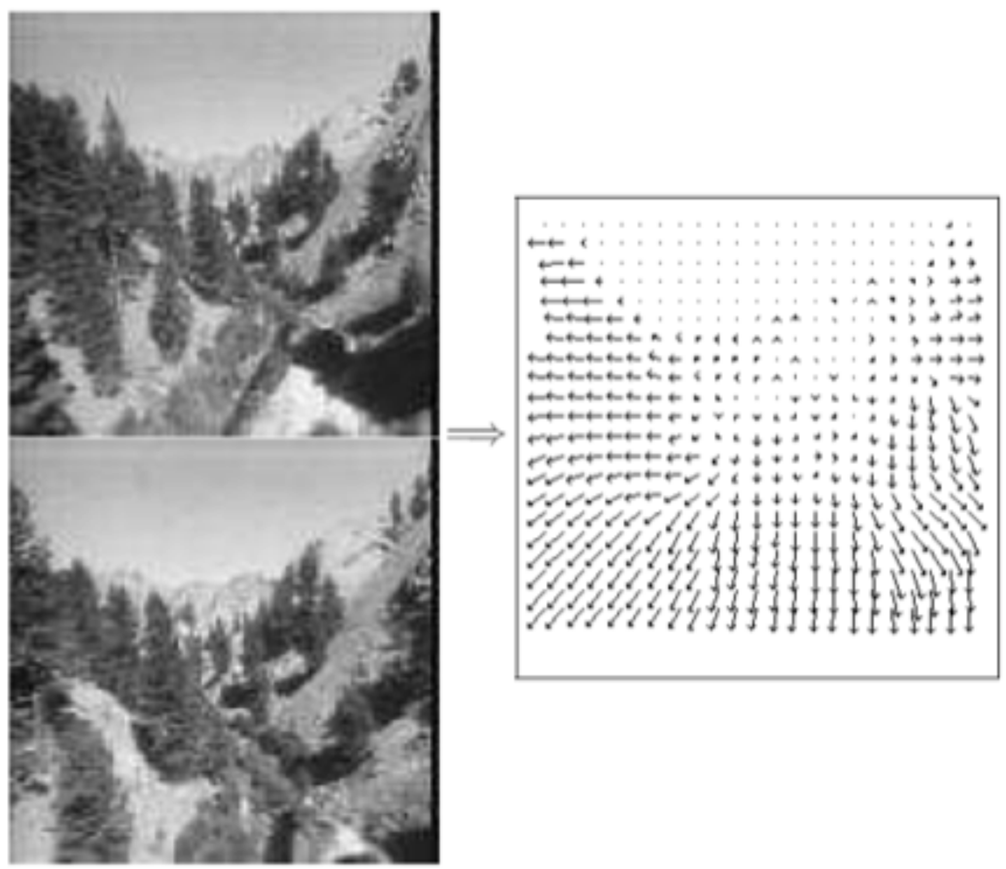
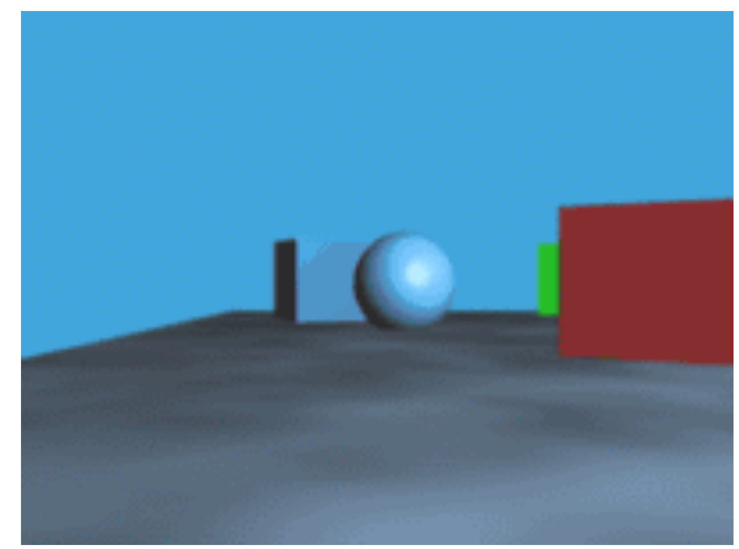
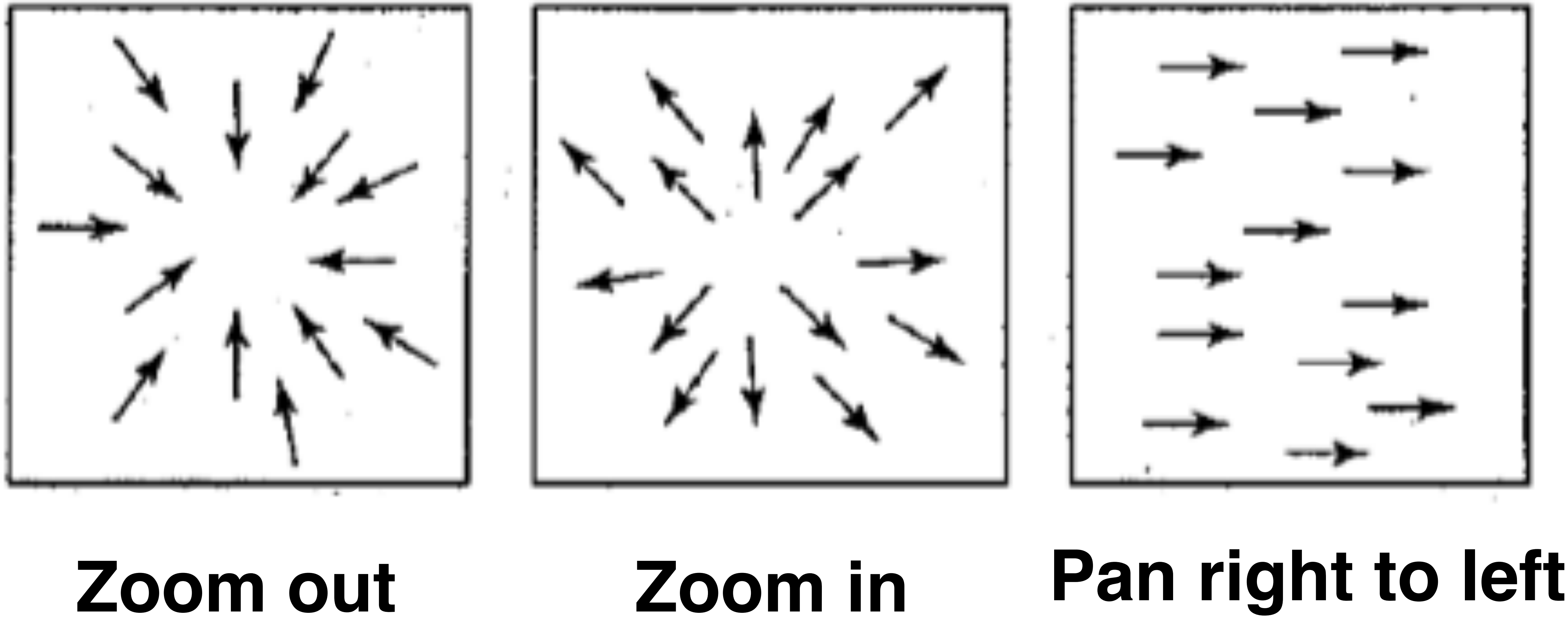


Figure 1.2: Two images taken from a helicopter flying through a canyon and the computed optical flow field.



http://psych.hanover.edu/KRANTZ/MotionParallax/MotionParallax.html

摄像机运动引起的运动场



光流

- **定义**: 图像中亮度模式的显著运动
- 理想情况下，光流和运动场相同
- 但是，即使没有实际运动，光照的变化也会引起亮度模式的显著运动

显著运动 ~= 运动场

Apparent motion ~= motion field

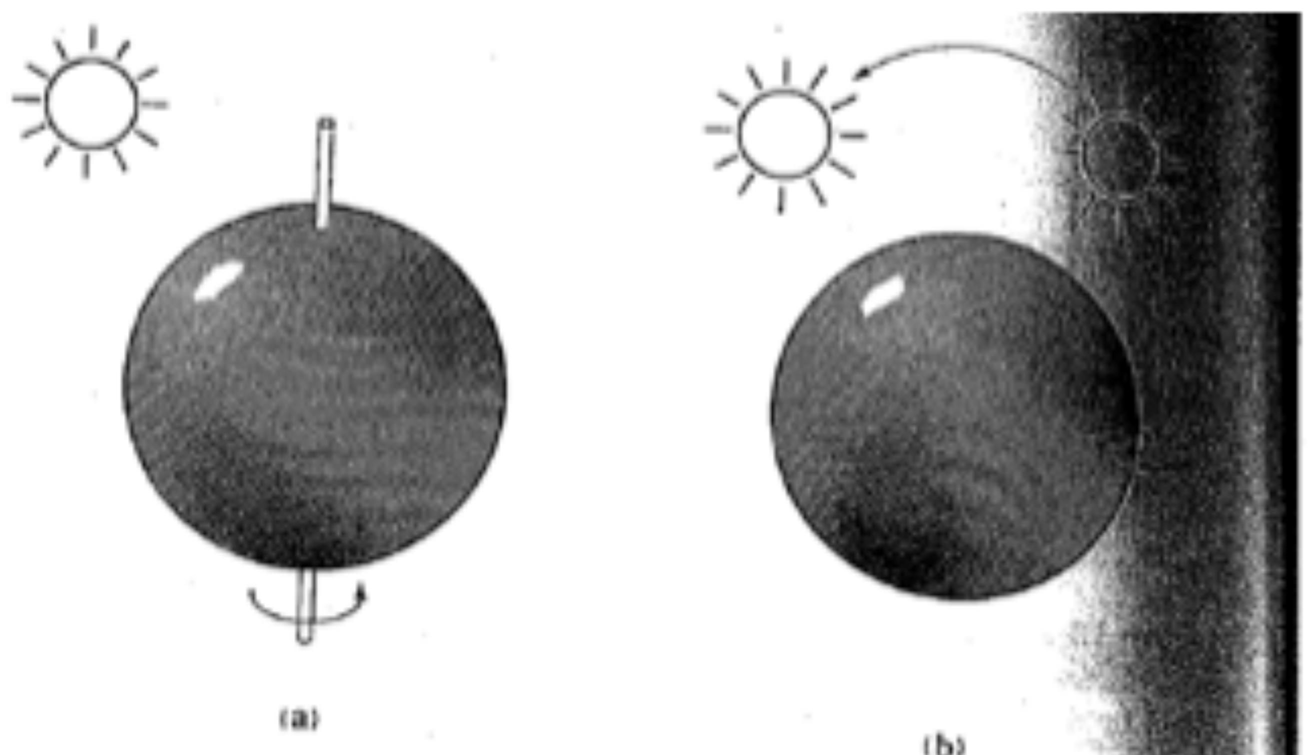
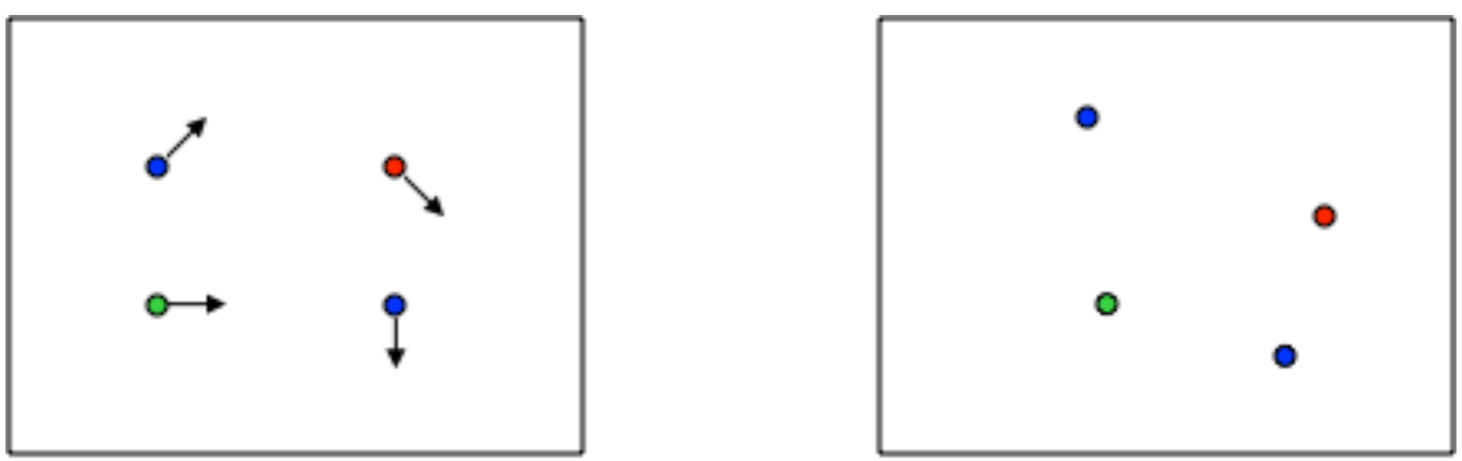


Figure 12-2. The optical flow is not always equal to the motion field. In (a) a smooth sphere is rotating under constant illumination—the image does not change, yet the motion field is nonzero. In (b) a fixed sphere is illuminated by a moving source—the shading in the image changes, yet the motion field is zero.

Figure from Horn book

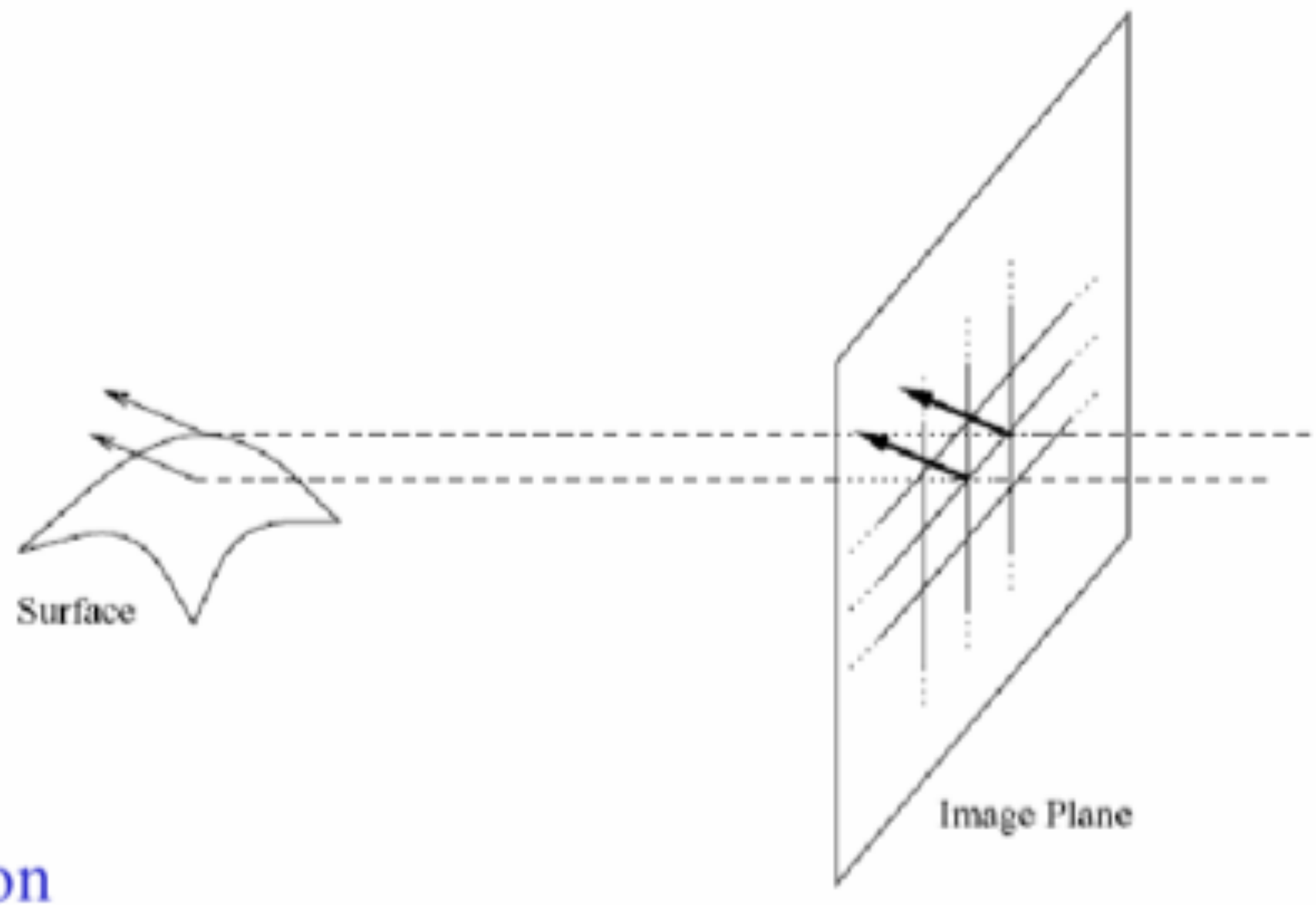
光流估计

- 给定时序相邻的两帧图像，估计显著运动场
- 光流三假设
  - **亮度恒常假设：**相同的点的投影在每帧图像中看起来相同
  - **微小运动假设：** 每个点的运动幅度较小
  - **空间一致性:** 相同局部邻域内的点具有相近的运动



$I(x,y,t-1)$ $\quad$ $I(x,y,t)$

空间一致性



Assumption

* Neighboring points in the scene typically belong to the same surface and hence typically have similar motions.
* Since they also project to nearby points in the image, we expect spatial coherence in image flow.

亮度恒常

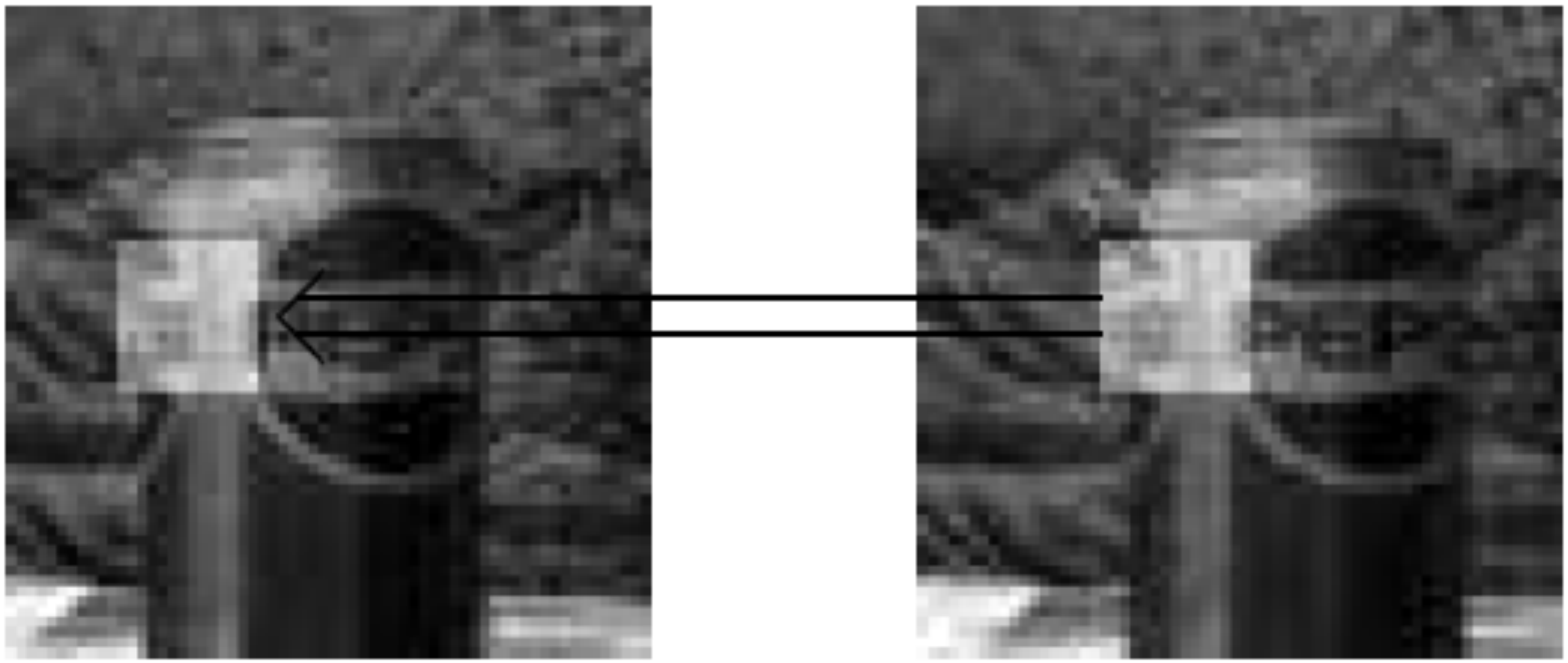
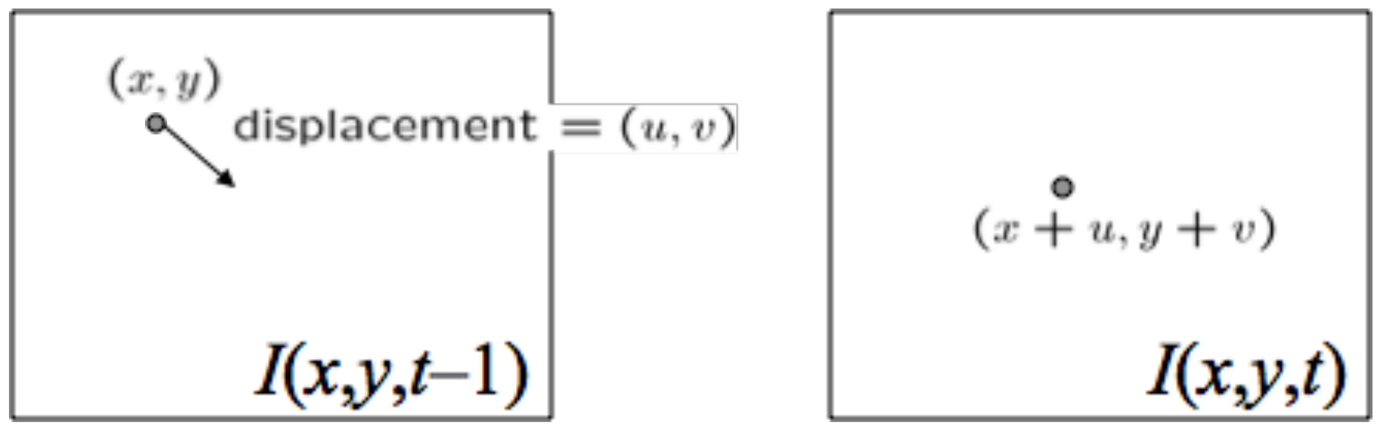


Figure 1.5: Data conservation assumption. The highlighted region in the right image looks roughly the same as the region in the left image, despite the fact that it has moved.

亮度恒常约束



Brightness Constancy Equation:

$$I(x, y, t-1) = I(x + u(x, y), y + v(x, y), t)$$

Linearizing (assuming small *(u,v)*) using Taylor series expansion:

$$I(x, y, t-1) \approx I(x, y, t) + I_x \cdot u(x, y) + I_y \cdot v(x, y)$$

So

shorthand: $I_x = \frac{\partial I}{\partial x}$

$$I_x \cdot u + I_y \cdot v + I_t \approx 0$$

- How many equations and unknowns per pixel?
  One equation, two unknowns

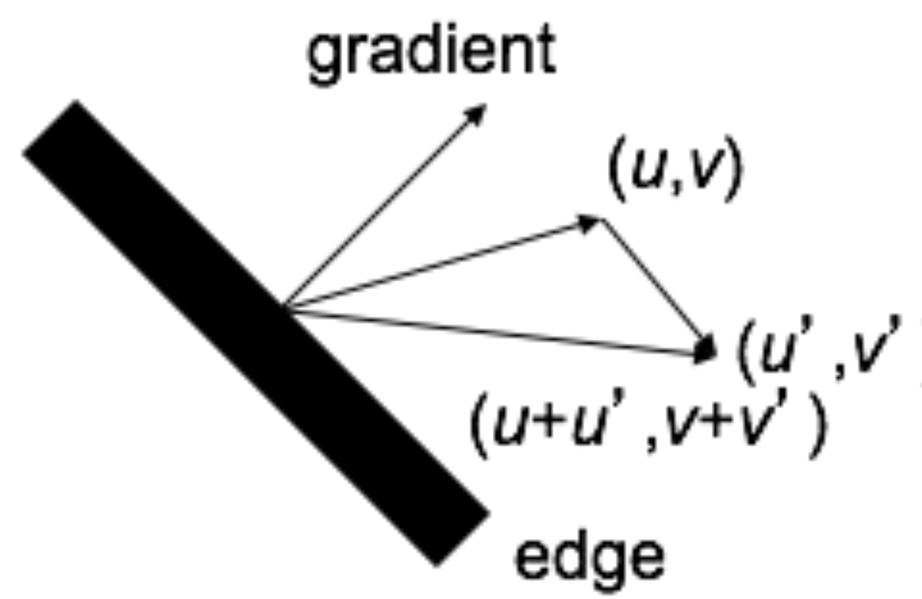
$$I_x \cdot u + I_y \cdot v + I_t = 0$$

- Intuitively, what does this constraint mean?

$$\nabla I \cdot (u, v) + I_t = 0$$

- The component of the flow perpendicular to the gradient (i.e., parallel to the edge) is unknown

If (u, v) satisfies the equation,
so does ($u+u'$, $v+v'$) if $\nabla I \cdot (u', v') = 0$

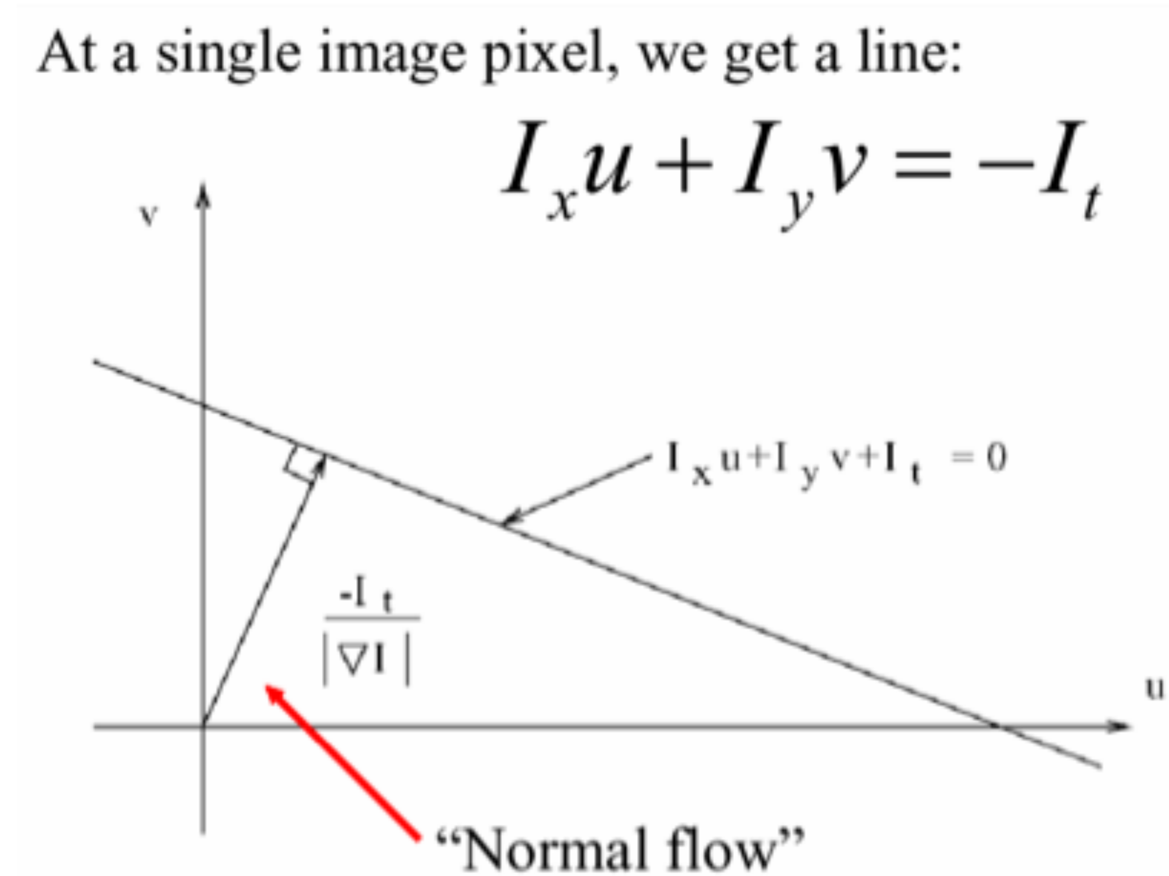


$$I_x u + I_y v + I_t = 0$$

$$\downarrow$$

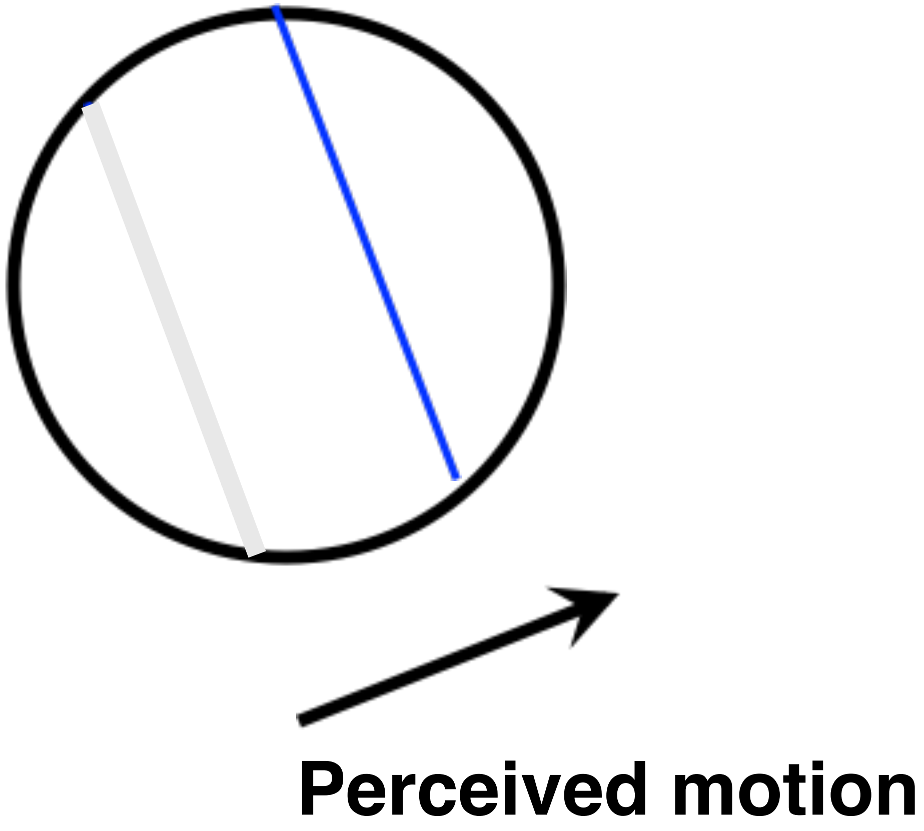
$$\nabla I^T \mathbf{u} = -I_t$$

$$\mathbf{u} = \begin{bmatrix} u \\ v \end{bmatrix} \quad \nabla I = \begin{bmatrix} I_x \\ I_y \end{bmatrix}$$

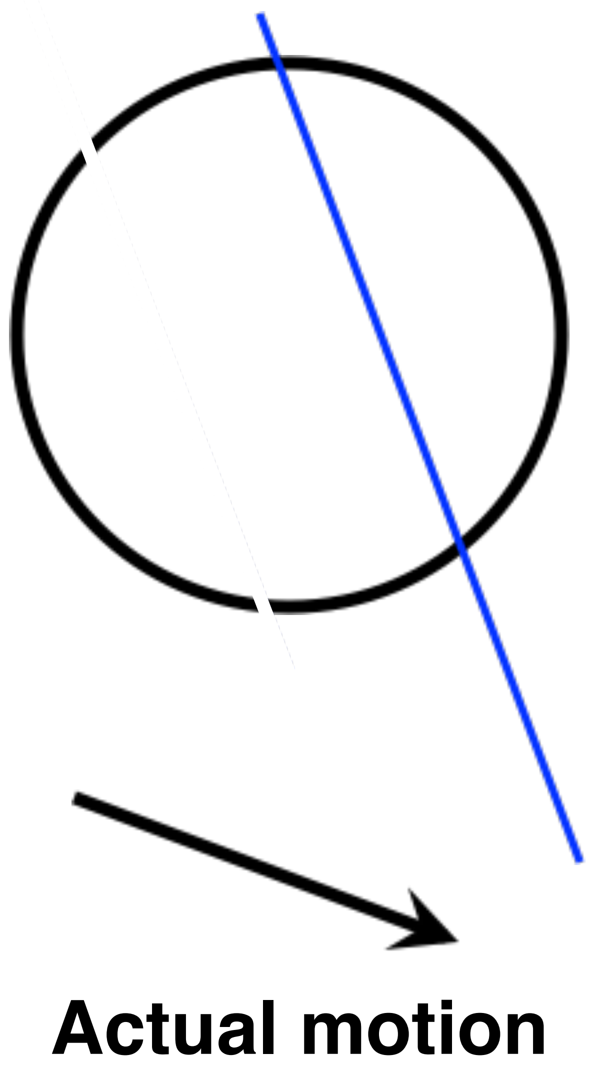


- 因此，我们最多可以得到“法向量流” – 在一点上，仅可检测到垂直于亮度梯度的运动
- 解决方案：在中心像素的周围取更多的像素点

光圈问题



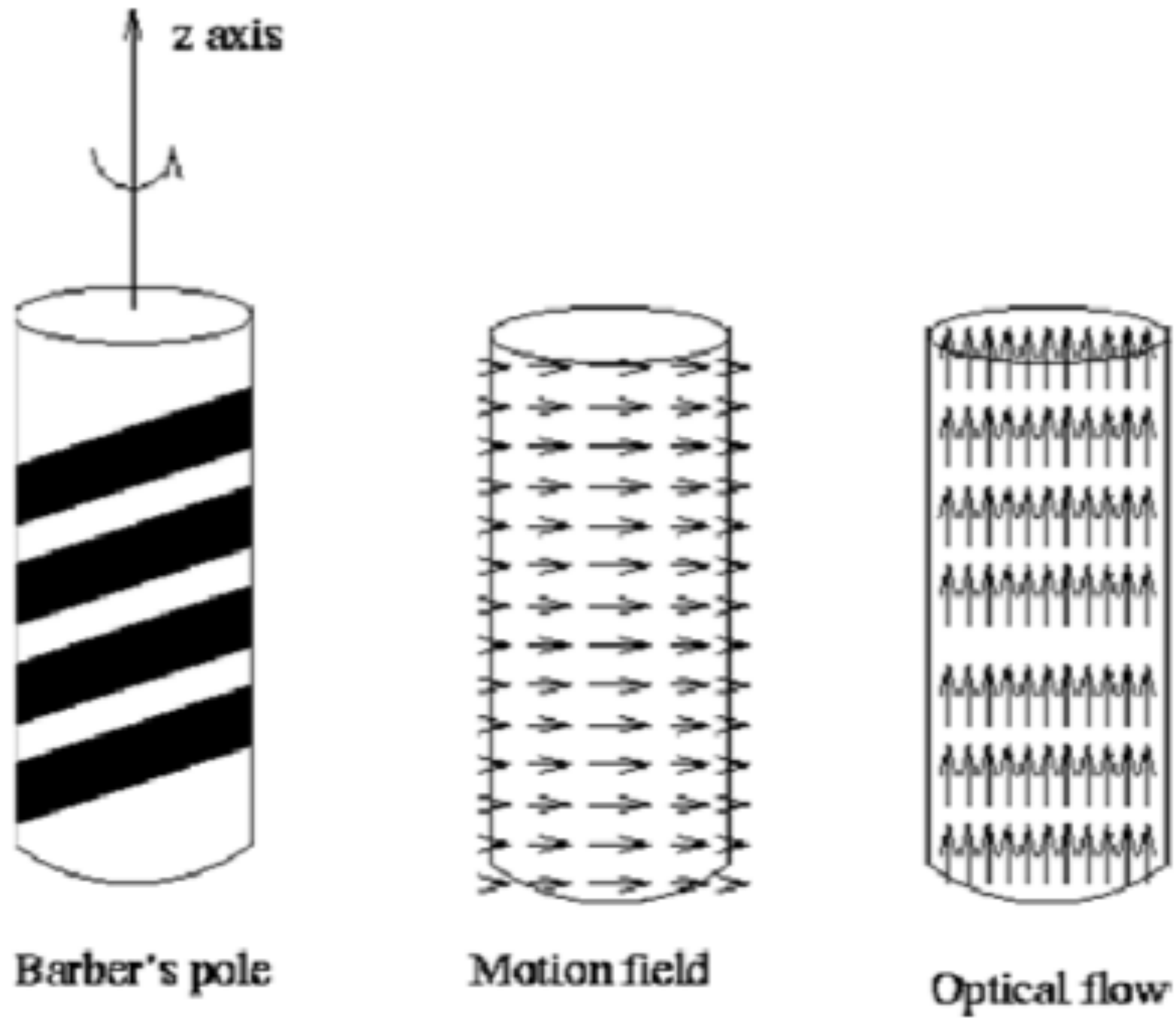
光圈问题



“理发店杆”错觉

http://en.wikipedia.org/wiki/Barberpole_illusion

Barber pole illusion



光圈问题求解

- 如何得到更多的方程?
- **空间一致性假设:** 假设像素周边邻域的其他像素具有相同的 (u,v)
  - 因此，5x5 窗口可以给出 25个方程

$$0 = I_t(\mathbf{p_i}) + \nabla I(\mathbf{p_i}) \cdot [u \;\; v]$$

$$\begin{bmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{25}) & I_y(\mathbf{p}_{25}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{25}) \end{bmatrix}$$

$$\underset{25\times2}{A} \quad \underset{2\times1}{d} = \underset{25\times1}{b}$$

RGB 版本：25*3=75

$$0 = I_t(\mathbf{p_i})[0,1,2] + \nabla I(\mathbf{p_i})[0,1,2] \cdot [u \;\; v]$$

$$\begin{bmatrix} I_x(\mathbf{p}_1)[0] & I_y(\mathbf{p}_1)[0] \\ I_x(\mathbf{p}_1)[1] & I_y(\mathbf{p}_1)[1] \\ I_x(\mathbf{p}_1)[2] & I_y(\mathbf{p}_1)[2] \\ \vdots & \vdots \\ I_x(\mathbf{p}_{25})[0] & I_y(\mathbf{p}_{25})[0] \\ I_x(\mathbf{p}_{25})[1] & I_y(\mathbf{p}_{25})[1] \\ I_x(\mathbf{p}_{25})[2] & I_y(\mathbf{p}_{25})[2] \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(\mathbf{p}_1)[0] \\ I_t(\mathbf{p}_1)[1] \\ I_t(\mathbf{p}_1)[2] \\ \vdots \\ I_t(\mathbf{p}_{25})[0] \\ I_t(\mathbf{p}_{25})[1] \\ I_t(\mathbf{p}_{25})[2] \end{bmatrix}$$

75x2 75x1

Note that RGB is not enough to disambiguate because R, G & B are correlated just provides better gradient

光圈问题求解

- **问题**：方程数多于未知数

$$\underset{25\text{x}2}{A}\ \underset{2\text{x}1}{d} = \underset{25\text{x}1}{b} \longrightarrow \text{minimize } \|Ad - b\|^2$$

Solution: solve least squares problem

- 最小二乘解（关于d）：

$$\underset{2\text{x}2}{(A^T A)}\ \underset{2\text{x}1}{d} = \underset{2\text{x}1}{A^T b}$$

$$\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}$$

  - 求和：K x K 窗口内的所有像素
  - 算法： Lucas & Kanade (1981)

有解的条件

- 什么情况下可求解?
  - **A^TA** s可逆
  - **A^TA** 不是太小
    - **A^TA**的特征值 λ_1 和 λ_2 不是太小
  - **A^TA** 为非病态矩阵
    - $\frac{\lambda_1}{\lambda_2}$ 不太大 （λ_1= 大特征值)

边缘

- 大 λ_1, 小 λ_2

$$\sum \nabla I (\nabla I)^T$$

低纹理区域

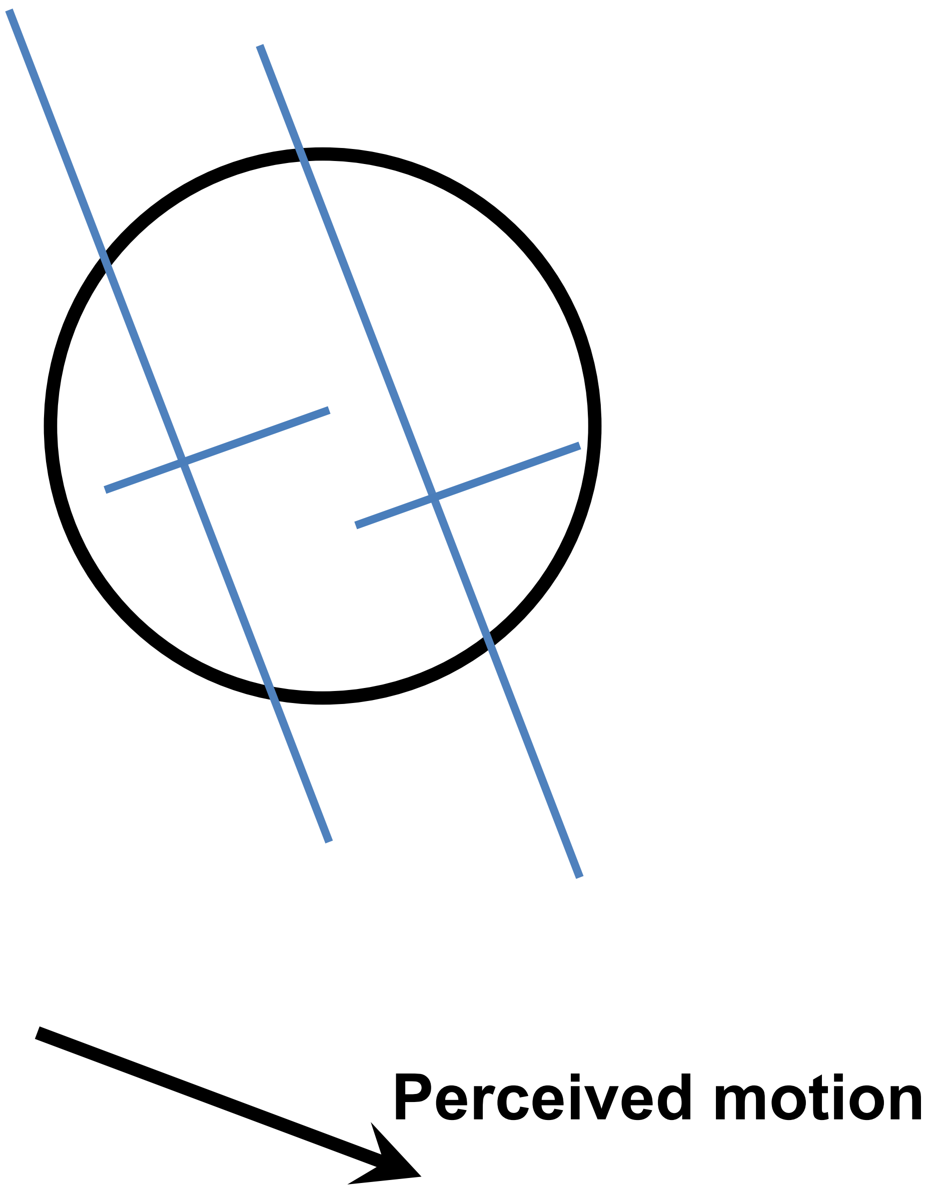
- 小 λ_1, 小 λ_2 $\quad \sum \nabla I (\nabla I)^T$

高纹理区域

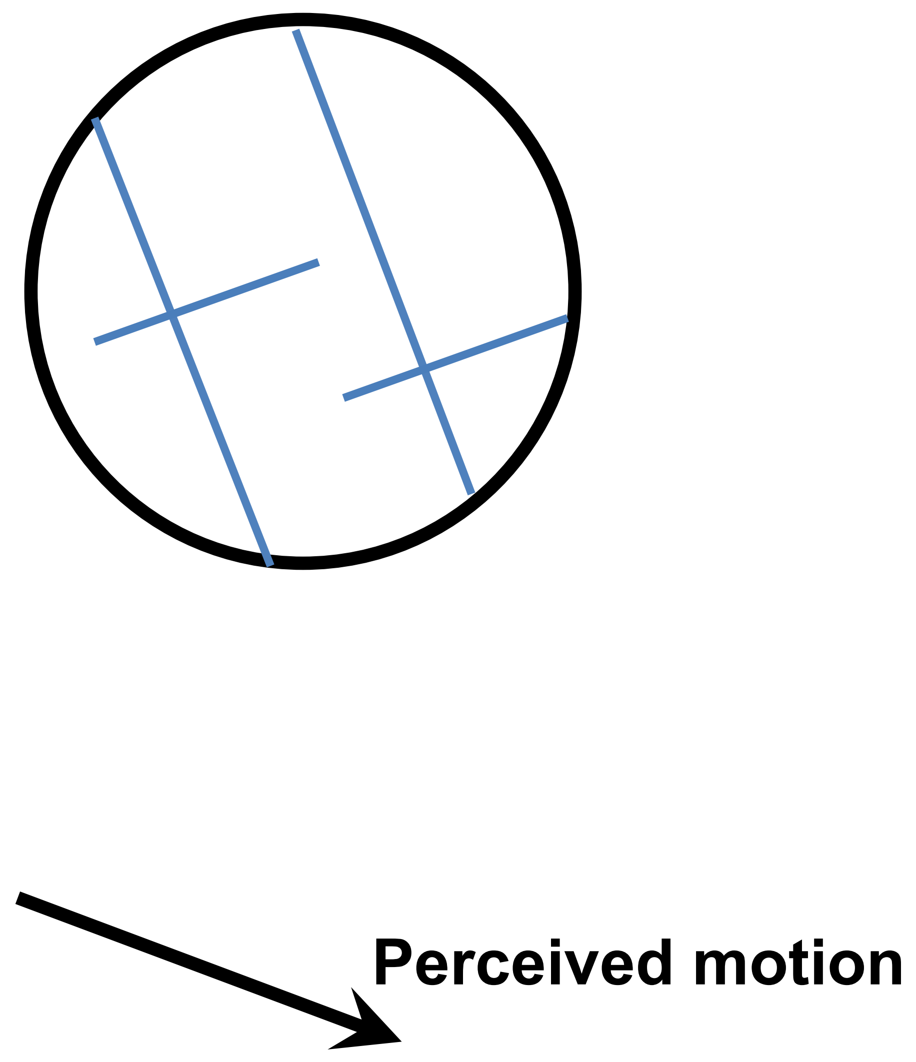
- 大 λ_1, 大 λ_2

$$\sum \nabla I (\nabla I)^T$$

光圈问题得到解决！



光圈问题得到解决！



思考

- 本质上，光流问题需用到两幅图像，但是
  - 可通过观察一副图像评估敏感性!
  - 可判断哪些像素更容易跟踪，哪些更难
  - 在特征跟踪时很有用

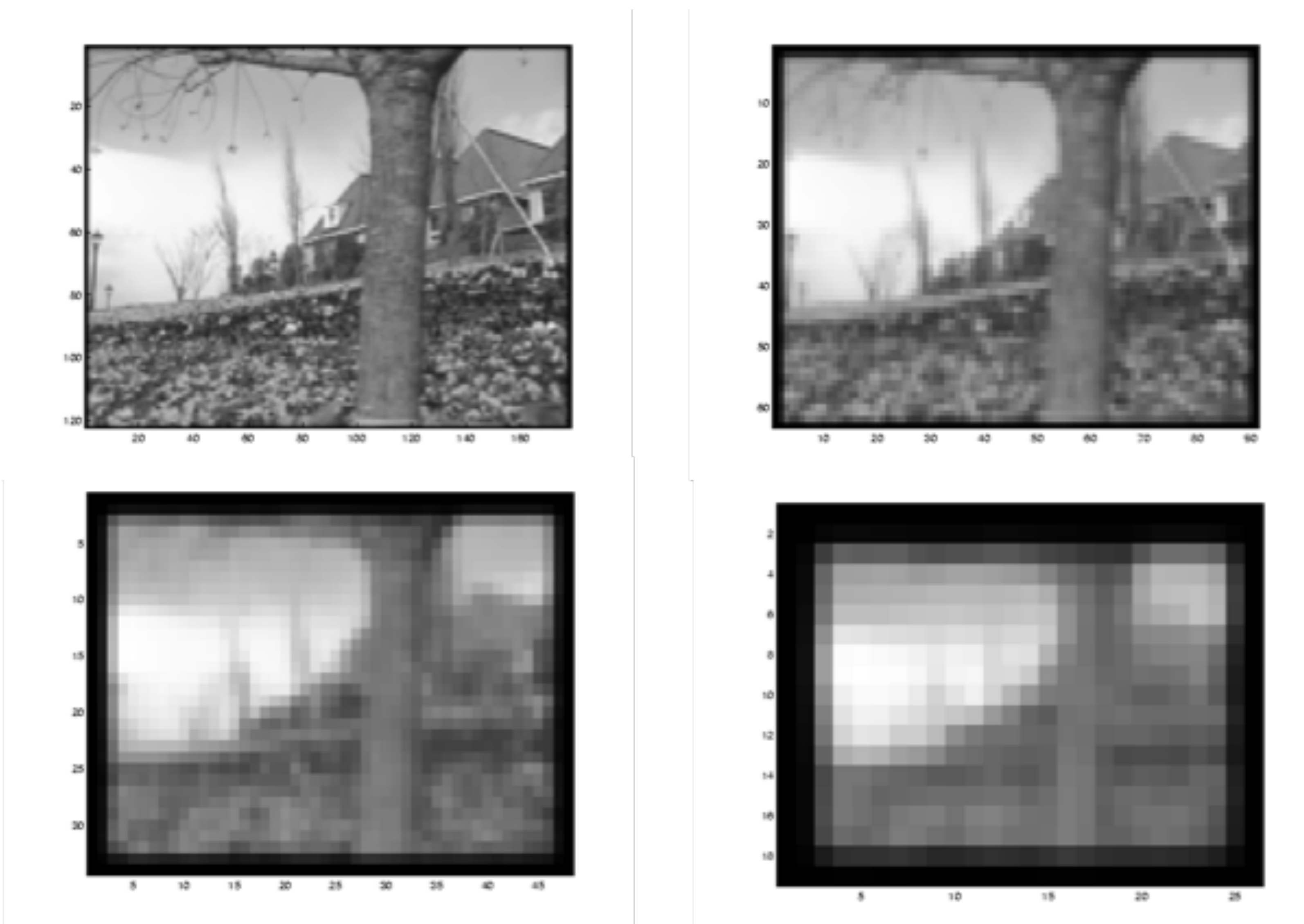
Lukas-Kanade算法的问题

- 出现问题的潜在诱因?
  - 假设 A^TA 是性质良好的可逆阵
  - 假设图像中的噪声不多
- 下列假设不满足时
  - 亮度恒常性不满足
  - 不是小的运动
  - 一个点的运动和邻域点不一致
  - 窗口太大
    - 怎么确定理想的窗口大小

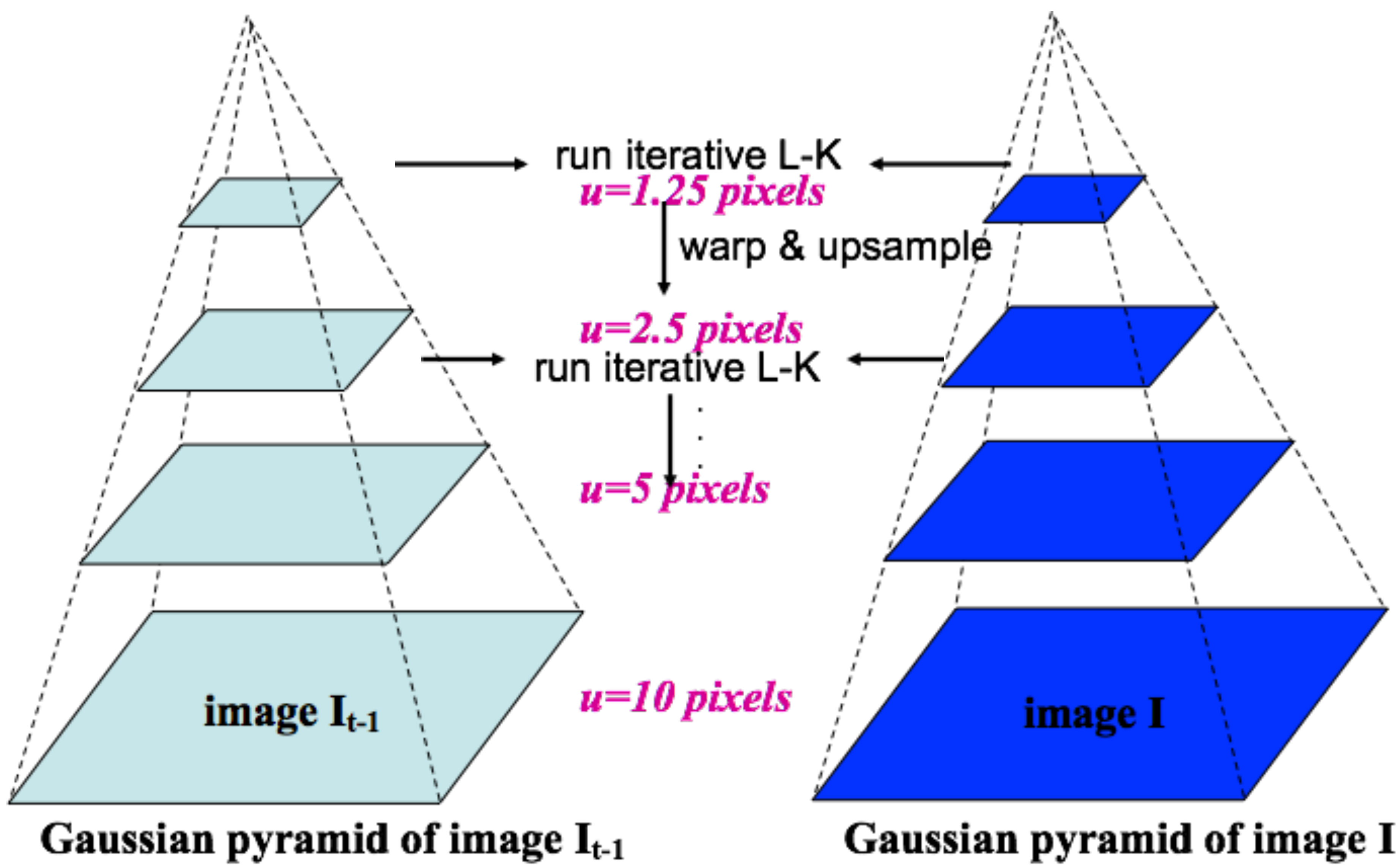
- 这样的运动是不是足够小?
  - 大于一个像素
  - 如何解决?

小运动假设

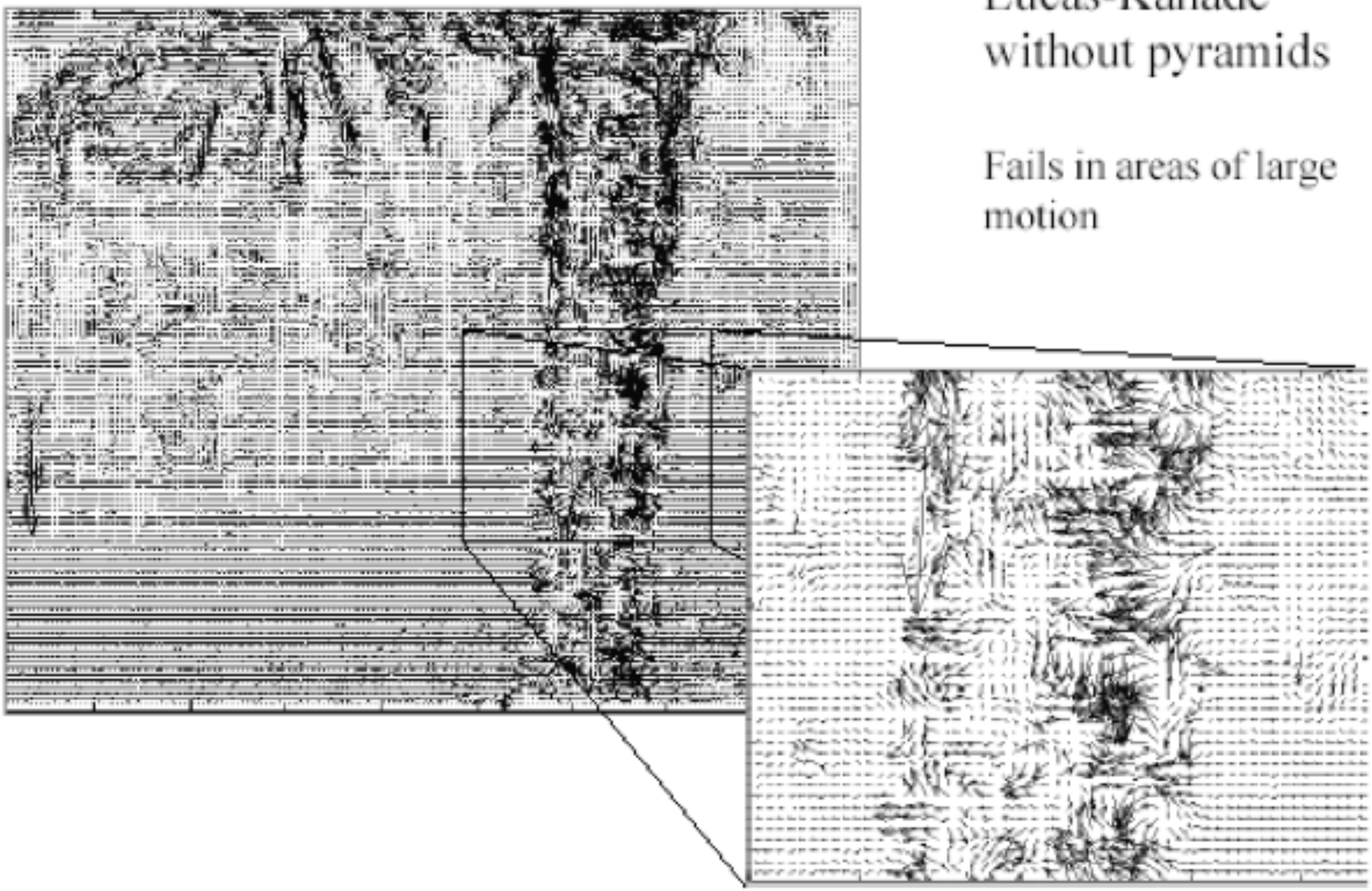
- 当小运动假设不满足时，如何解决?
  - 降低分辨率!



Coarse-to-fine 光流估计



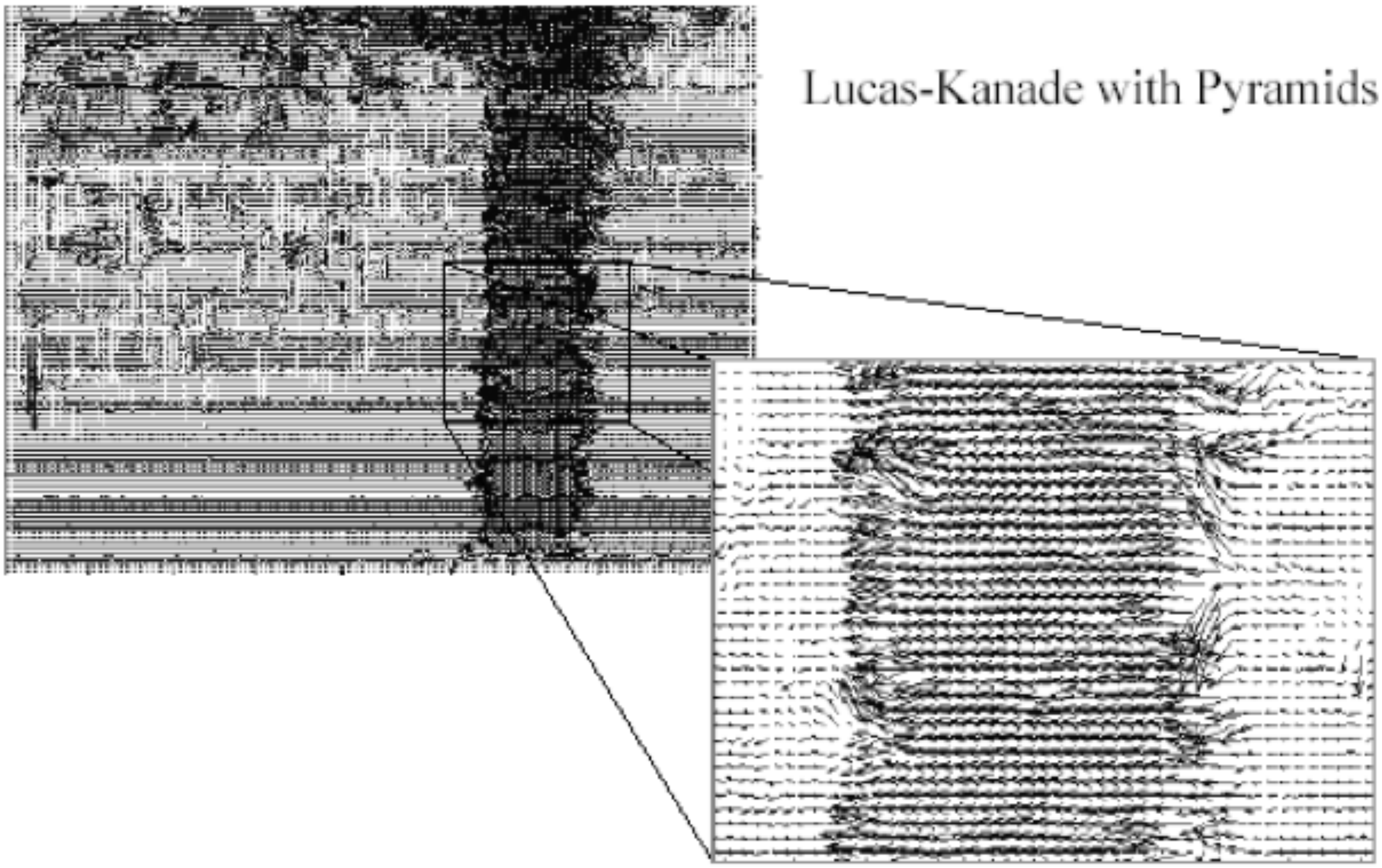
结果



Lucas-Kanade
without pyramids

Fails in areas of large
motion

结果



From Khurram Hassan-Shafique CAP5415 Computer Vision 2003

光流

- 运动还是表观?
  - 小运动假设
  - 亮度恒常性