2022 Spring

计算机视觉

袁丁
北京航空航天大学
图像中心
dyuan@buaa.edu.cn

第六章

运动视觉

本章主要知识点

- 光流的概念
- 光流场与运动场
- 相机的运动与光流
- 相机的运动与运动场
- 光流约束方程及其几何含义
- 光流的求解方法

运动视觉概述

“生命和世界都在运动。没有运动就没有时间、空间和物质”

-------亚里士多德

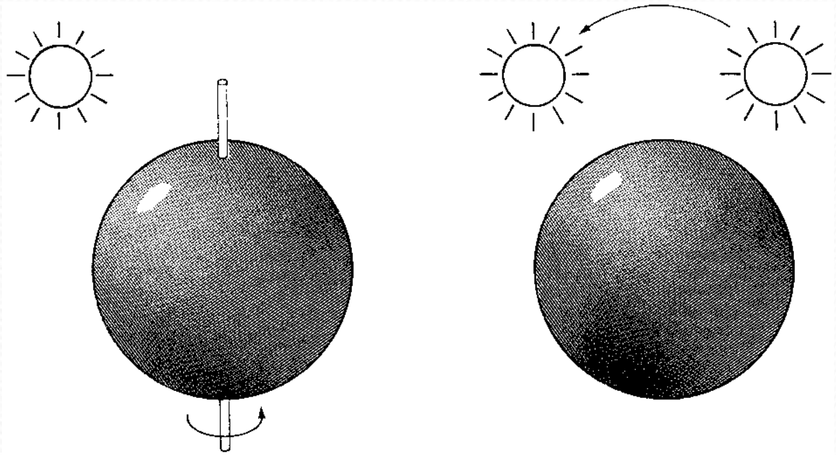
- 人类的视知觉需要在运动中获取

- 运动视觉的目的：提取决策用的视觉信息
- 实现运动视觉的目的：
  - ✓ 实时监视：对场景中的目标进行分析，实现对运动目标的检测、测量、估计、识别与跟踪
  - ✓ 自动导航：分析场景和自身运动，实现机器人、运输工具、飞行器的控制、制导、自主导航
- 处理对象：按时间顺序排列的瞬间采样的图像序列

光流（Optical Flow）的概念

- 光流：也可称图像流（image flow），是指两帧图像中亮度模式的视运动。用来描述多幅灰度图像之间的变化。
- 光流：侧重建立两帧图像灰度值变化与速度场之间的关系

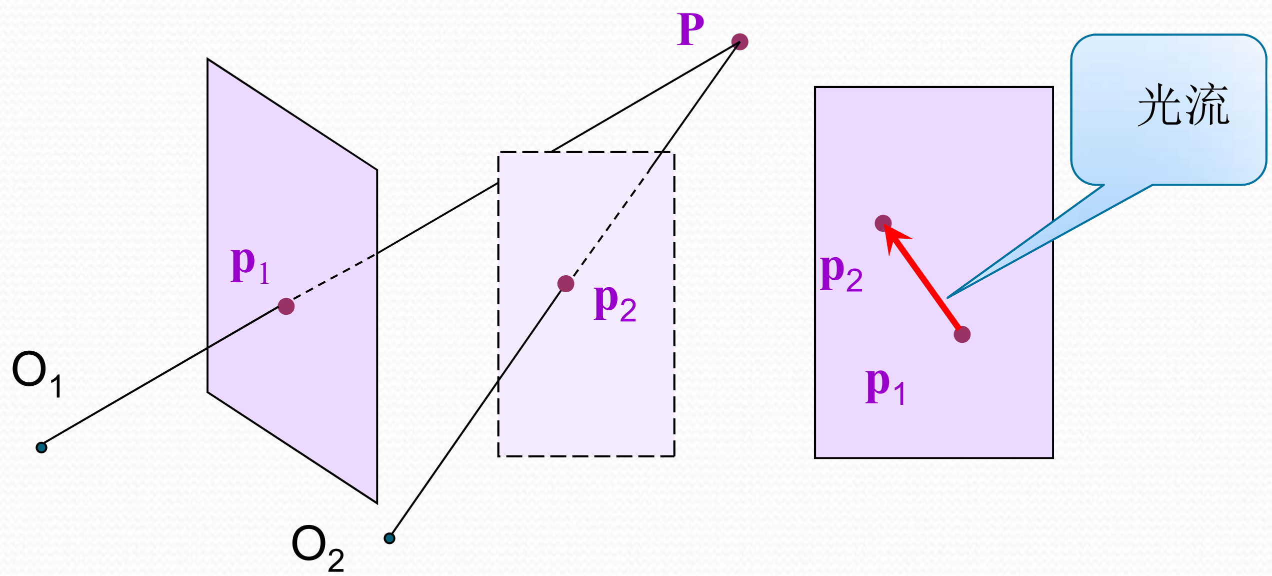
  图像流：侧重建立二维速度场与三维运动的关系



(a)固定光源　　　　(b)运动光源

光流（Optical Flow）的概念

- 如果光源固定不变，相机和景物的 相对运动 是两帧灰度图像之间的变化原因。
- 下图假设相机运动，景物静止



- 光流的定义：光流是一幅图像到下一幅图像对应像素点对间的位移量，也可以认为是空间运动的点在投影焦平面上的二维瞬时速度场

运动场（Motion Field）

相对运动：

（1）相机静止；景物运动

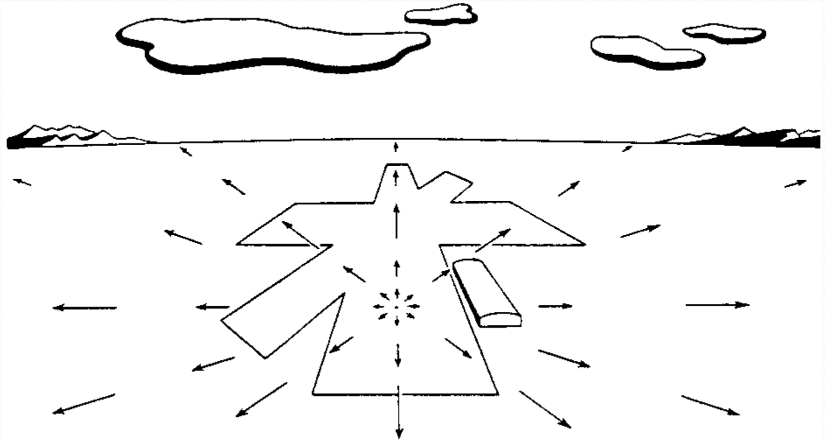
（2）相机运动；景物静止

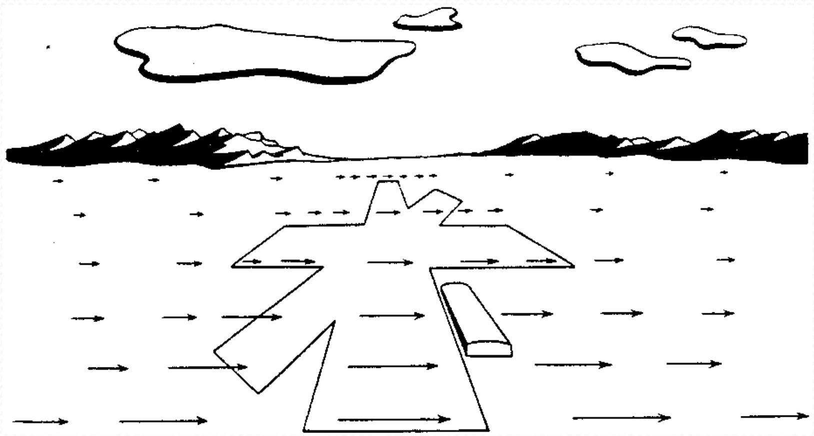
（3）相机运动；景物运动

- 运动场内，每一个像素都被赋予一个速度矢量
- 这些速度矢量源自于相机和景物之间的相对运动
- 运动场可以“*被当做*”是三维速度矢量在二维图像平面的投影
- 理想情况下，光流场等同于运动场

运动场举例

飞机将要降落在跑道上某点。降落前，飞机在跑道上方飞行



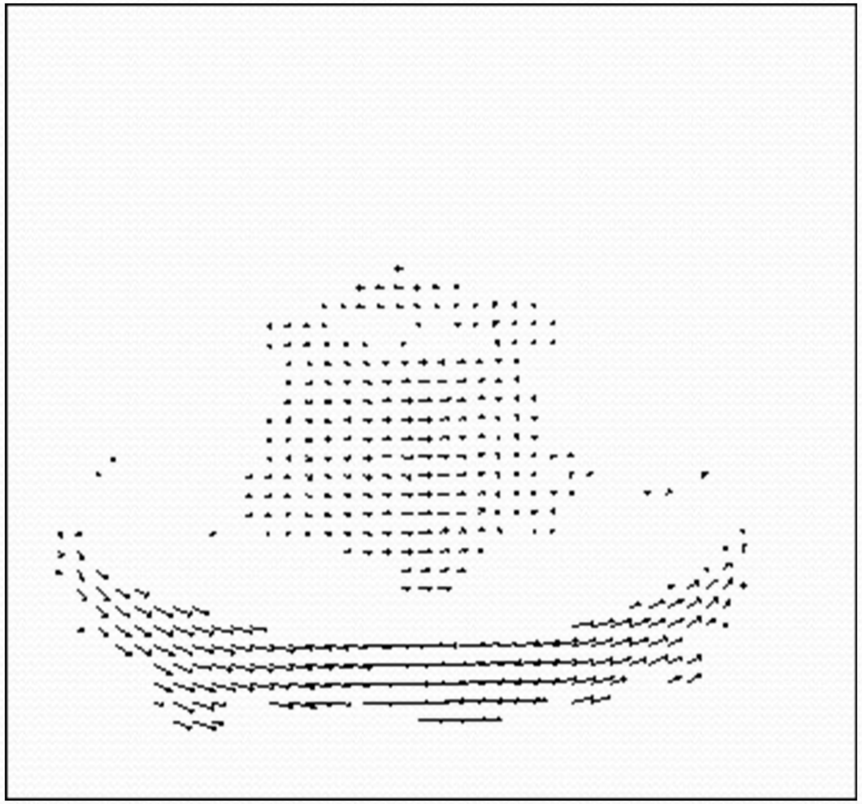


（a）飞行员向前方看时，感觉到的运动场

（a）飞行员向右看时，感觉到的运动场

运动场举例

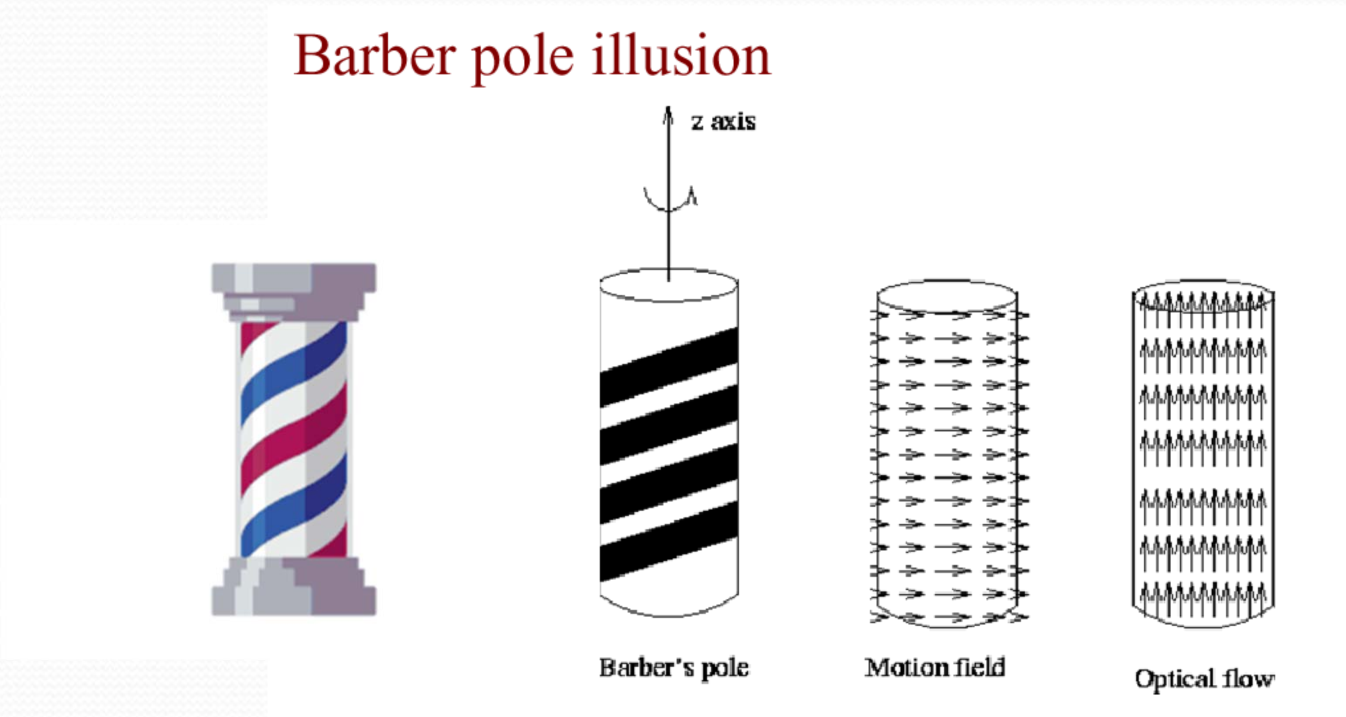
输入的两帧图像：旋转的物体



计算求得运动场（光流场）

光流场 VS. 运动场

- 光流场并不总是等同于运动场
- 光流场是运动场的一种近似



相机的运动 vs. 物体的运动

- 从观察者的角度来看：

当外界景物静止时，相机运动	≈	相机静止时，其视场内所有景物运动（方向相反）

因此：视觉信息描述了相机与景物之间的相对运动

- 当景物静止，相机运动时，对相机的运动所做的研究，叫做：

  相机自运动的研究（Ego-motion）

  应用：导航

相机的运动→二维运动场

- 相机运动，景物静止

t: 相机的平移 $\mathbf{t}=\left(t_x,t_y,t_z\right)^T$

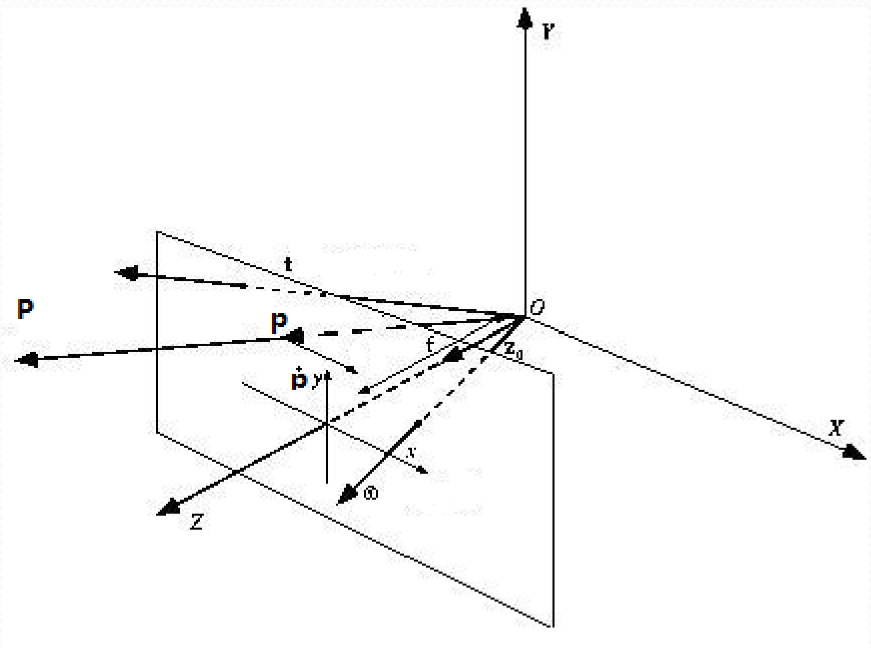
ω: 相机的转动 $\boldsymbol{\omega}=\left(\omega_x,\omega_y,\omega_z\right)^T$

（经过光心的向量）

P: 三维物点 $\mathbf{P}=\left(X,Y,Z\right)^T$

p: 像平面二维像点 $\mathbf{p}=\left(x,y,f\right)$

$\dot{\mathbf{p}}$: 像平面二维点速度 $\dot{\mathbf{p}}=\left(u,v,0\right)$



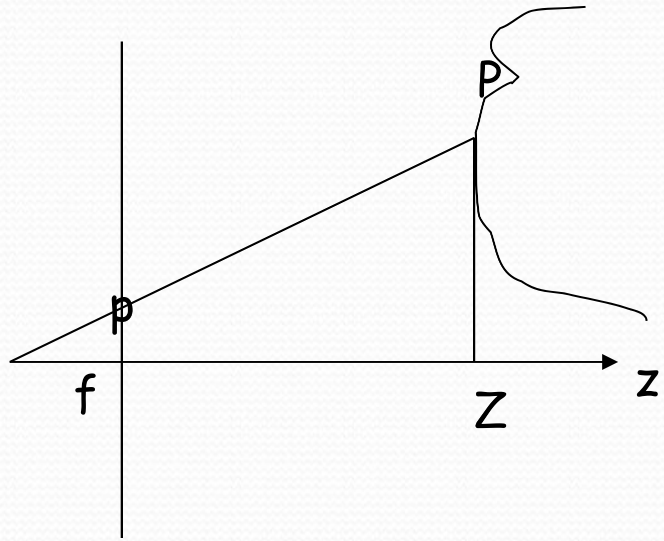
O-YXZ: 相机坐标系

o-xy: 成像平面坐标系

相机的运动→二维运动场

- 透视投影模型：

$$\mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad \mathbf{p} = \begin{bmatrix} x \\ y \\ f \end{bmatrix} \Bigg\} \quad \mathbf{p} = \frac{f\,\mathbf{P}}{Z}$$

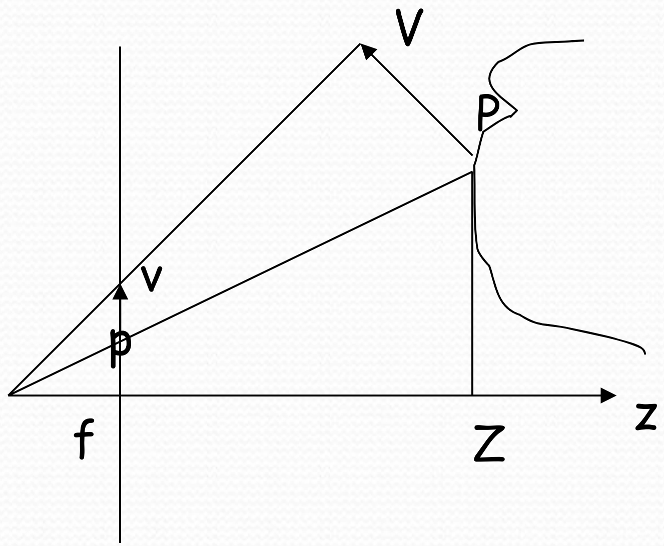


- 物点与相机间的相对运动：

$$\mathbf{V} = -\mathbf{t} - \boldsymbol{\omega} \times \mathbf{P}$$

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad \boldsymbol{\omega} = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad \mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

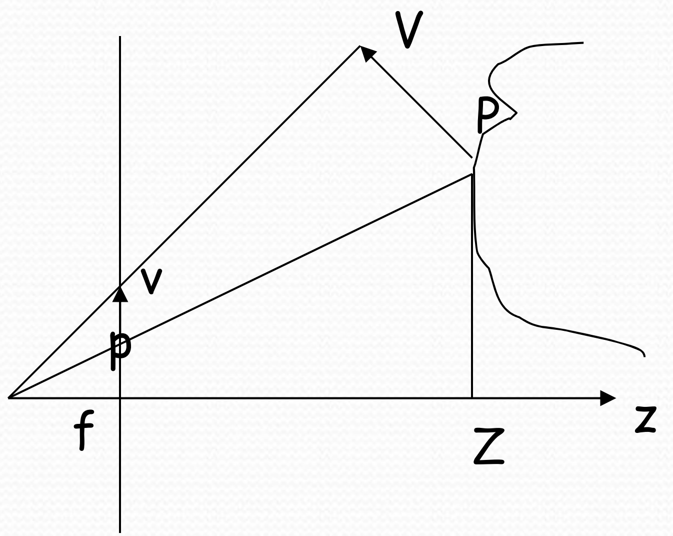
$$\begin{aligned} V_x &= -t_x - \omega_y Z + \omega_z Y \\ V_y &= -t_y - \omega_z X + \omega_x Z \\ V_z &= -t_z - \omega_x Y + \omega_y X \end{aligned}$$



像点的速度

- 对时间求导数：

由 $\mathbf{p} = \frac{f\,\mathbf{P}}{Z}$



$$\frac{d\mathbf{p}}{dt} = \mathbf{v} = \frac{d\frac{f\mathbf{P}}{Z}}{dt} = \frac{f}{Z^2}\left[\frac{d\mathbf{P}}{dt}Z - \mathbf{P}\frac{dZ}{dt}\right]$$

$$= \frac{f}{Z^2}\left[\mathbf{V}Z - \mathbf{P}V_z\right]$$

$$\mathbf{v} = f\frac{\mathbf{V}}{Z} - \mathbf{p}\frac{V_z}{Z}$$

$$\begin{cases} v_x = u = f\dfrac{V_x}{Z} - x\dfrac{V_z}{Z} \\ v_y = v = f\dfrac{V_y}{Z} - y\dfrac{V_z}{Z} \\ v_z = f\dfrac{V_z}{Z} - f\dfrac{V_z}{Z} = 0 \end{cases}$$

像点的速度

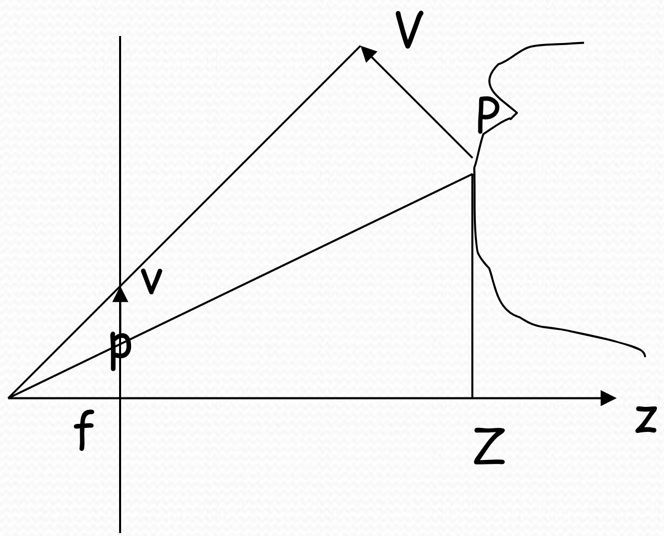
- 由：

$$
\begin{aligned}
u &= f\frac{V_x}{Z} - x\frac{V_z}{Z} \\
v &= f\frac{V_y}{Z} - y\frac{V_z}{Z}
\end{aligned}
$$

$$V_x = -t_x - \omega_y Z + \omega_z Y$$

$$V_y = -t_y - \omega_z X + \omega_x Z$$

$$V_z = -t_z - \omega_x Y + \omega_y X$$



光流：

$$
\boxed{
\begin{aligned}
u &= \frac{t_z x - t_x f}{Z} + \frac{\omega_x x y}{f} - \omega_y\left(f + \frac{x^2}{f}\right) + \omega_z y \\
v &= \frac{t_z y - t_y f}{Z} + \omega_x\left(f + \frac{x^2}{f}\right) - \frac{\omega_y x y}{f} - \omega_z x
\end{aligned}
}
$$

运动场(Motion Field): 像点的速度

平移分量　　　　旋转分量

$$
\begin{aligned}
u &= \frac{t_z x - t_x f}{Z} + \frac{\omega_x x y}{f} - \omega_y\left(f + \frac{x^2}{f}\right) + \omega_z y \\
v &= \frac{t_z y - t_y f}{Z} + \omega_x\left(f + \frac{x^2}{f}\right) - \frac{\omega_y x y}{f} - \omega_z x
\end{aligned}
$$

- 平移分量： 与相机平动，三维物点深度信息相关。具有“尺度”的不确定性（Scaling ambiguity）。

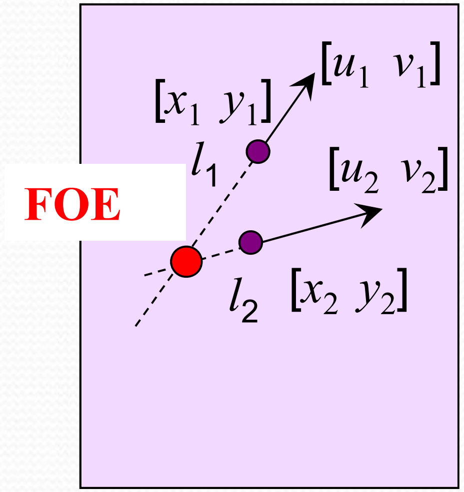
- 旋转分量：与三维物点深度信息无关

运动场举例

- 平移运动场（Translational flow field）

  相机做单纯的平动， 此时，$\boldsymbol{\omega}=0$，则光流为：

（假设 t_z不为0）　$u=\dfrac{xt_z-t_xf}{Z}$　　$v=\dfrac{yt_z-t_yf}{Z}$



FOE & 像点、光流

图像上两个像点和其所对应的光流：

像点 $[x_1\ \ y_1]$ 和其对应的光流 $[u_1\ \ v_1]$

像点 $[x_2\ \ y_2]$ 和其对应的光流 $[u_2\ \ v_2]$

直线方程　l_1:

$$\frac{x-x_1}{y-y_1}=\frac{u_1}{v_1}=\frac{-t_xf+x_1t_z}{-t_yf+y_1t_z}=\frac{f\dfrac{t_x}{t_z}-x_1}{f\dfrac{t_y}{t_z}-y_1}$$

l_2:

$$\frac{x-x_2}{y-y_2}=\frac{u_2}{v_2}=\frac{-t_xf+x_2t_z}{-t_yf+y_2t_z}=\frac{f\dfrac{t_x}{t_z}-x_2}{f\dfrac{t_y}{t_z}-y_2}$$

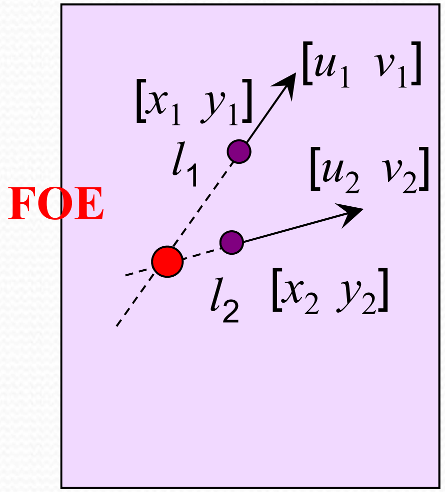
运动场举例

- 求解两条直线 l_1 和 l_2 的交点：

$$[\ f\frac{t_x}{t_z}\ ,\ f\frac{t_y}{t_z}\ ]$$

交点与像点的位置无关

当相机做单纯平动时，所有光流（或者其延长线）相交于一点：FOE（或 FOC）



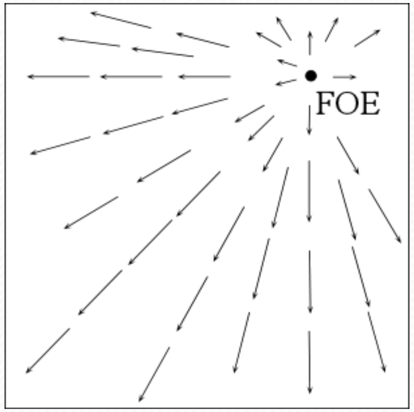
FOE & 像点、光流

$$FOE = \begin{bmatrix} f\frac{t_x}{t_z} & f\frac{t_y}{t_z} \end{bmatrix}$$

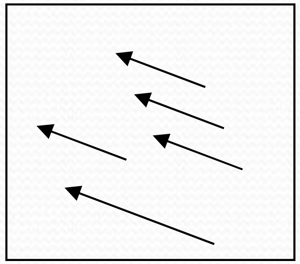
FOE（Focus of Expansion）： 相机向光轴的正方向运动（散焦点）

FOC (Focus of Contraction) ： 相机向光轴的反方向运动（聚焦点）

相机单纯平动的运动场



$t_z \neq 0$



$t_z = 0$

每一像素点的光流：

$$u = \frac{xt_z - t_x f}{Z} \qquad v = \frac{yt_z - t_y f}{Z}$$

相交于：

$$FOE = \begin{bmatrix} f\frac{t_x}{t_z} & f\frac{t_y}{t_z} \end{bmatrix}$$

每一像素点的光流：

$$u = -f\frac{t_x}{Z}$$

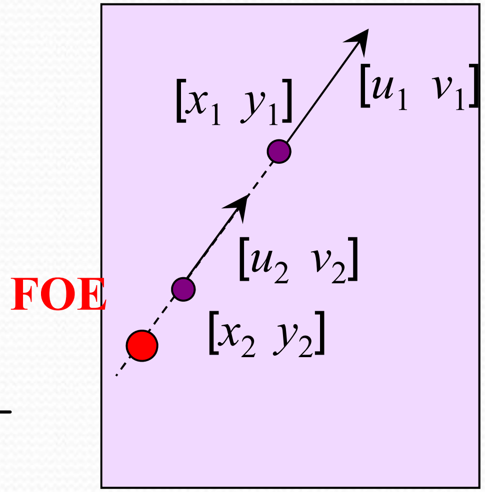
$$v = -f\frac{t_y}{Z}$$

相交于：　无穷远处

相机单纯平动的运动场

- 假设图像中两光流处深度相同：$Z_1=Z_2=Z$
- 以x方向的光流分量为例：

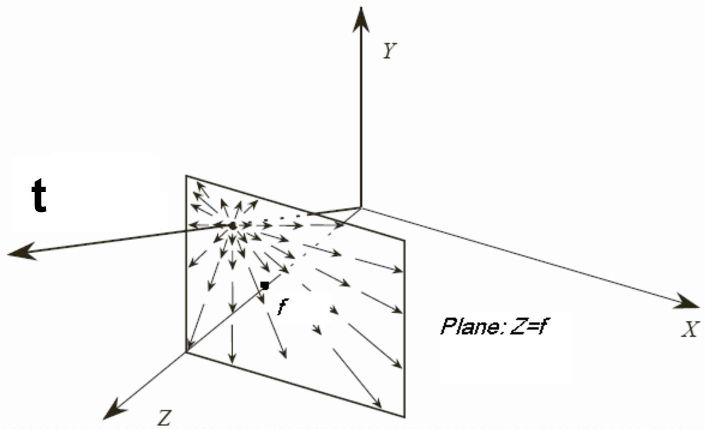
$$\frac{u_1}{u_2}=\frac{\dfrac{x_1t_z-ft_x}{Z}}{\dfrac{x_2t_z-ft_x}{Z}}=\frac{x_1-\dfrac{ft_x}{t_z}}{x_2-\dfrac{ft_x}{t_z}}=\frac{x_1-FOE_x}{x_2-FOE_x}=\frac{\Delta x_1}{\Delta x_2}$$



FOE & 像点、光流

显然：当空间点对应的各像点离开**FOE**时，像点的 x, y方向位移量与光流的分量大小成正比

单纯平动：运动场（MF）的性质

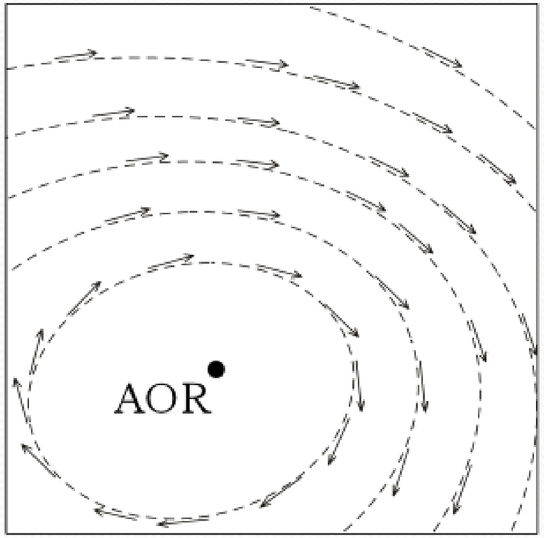


- 如果 $t_z \neq 0$，MF 呈辐射状分布，且所有光流矢量指向（或背离）同一点FOC（或 FOE）；如果 $t_z = 0$， MF 平行分布。
- MF中的光流矢量的大小和其所对应的三维物点的深度信息 Z 成反比。如果 $t_z \neq 0$，光流矢量的大小和其像点 **p** 与 FOC（或 FOE）的距离成正比。
- FOE (或 FOC) 是相机平动方向的矢量与成像平面的交点
- FOE (或 FOC) 也称为***<u>消失点</u>***（Vanishing Point），指示了相机平移运动的方向

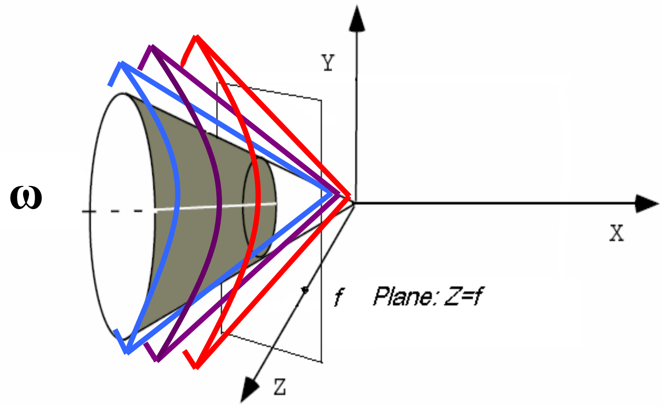
运动场举例

- 转动运动场（Rotational flow field）：

令：
$$AOR=\left(f\frac{\omega_x}{\omega_z}\quad f\frac{\omega_y}{\omega_z}\right)$$



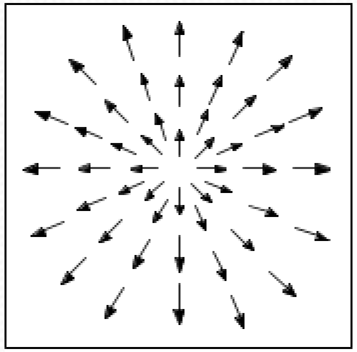
AOR (Axis of Rotation): 当相机以向量 **ω** 为轴转动， AOR 为 **ω** 与成像平面的交点。



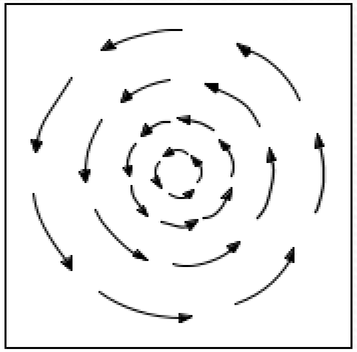
- 右图中的虚线所示为：以 **ω** 为轴的圆锥与成像平面的交线（二次曲线）
- MF中的光流矢量的方向为：二次曲线的切线方向
- MF中的光流矢量的大小：与像点所在位置相关

运动场举例

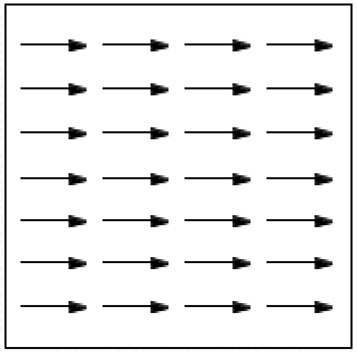
由运动场判断相机大致的运动：



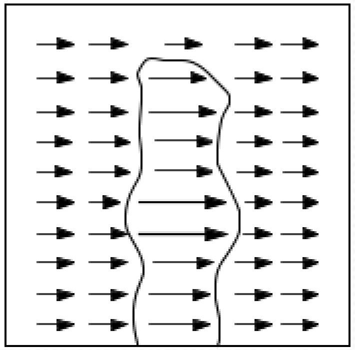
(a)



(b)



(c)



(d)

(a) 相机垂直于图像平面、沿光轴正方向的平移运动

(b) 相机绕光轴顺时针单纯旋转

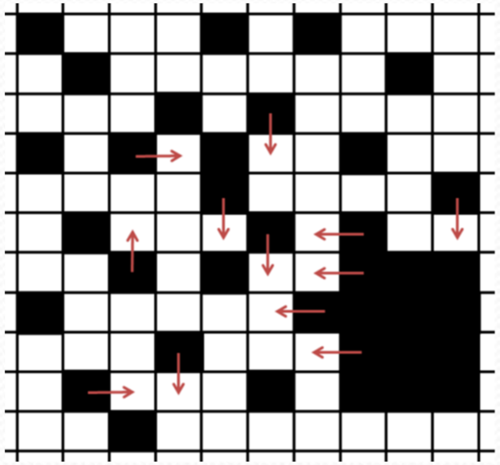
(c) 相机在某平面前，平行于该平面平移运动

(d) 相机在某目标前，平行于该目标平移运动，目标与背景有一定的距离

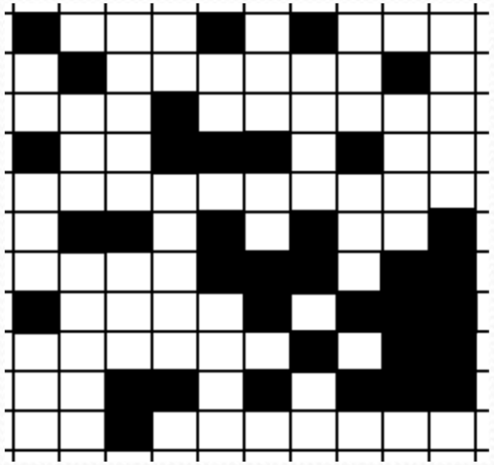
图像内光流的求解

- 光流：图像亮度模式的“显著”变化➔ 像素点的“运动”

- 问题：像素的“运动”引起了图像亮度模式的“显著”变化，从而生成了一张新的图像。



原图像

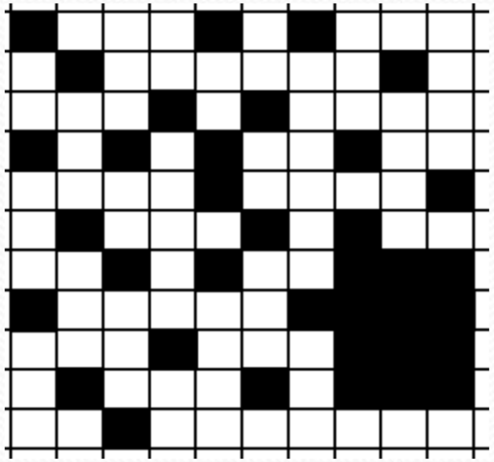


新图像

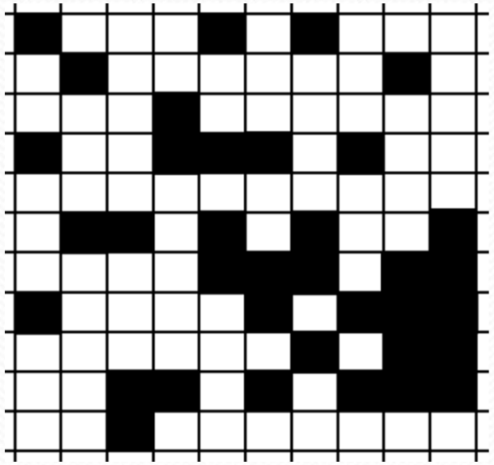
图像内光流的求解

- 根据运动前后的两张图像，能否推断出像素点的“运动”？
- 该推断结果即为“光流”
- 问题：

1. 解是否存在？
2. 解是否唯一？
3. 如何设计有效算法求解？



原图像



新图像

图像内光流的求解

- 光流的求解计算是极有研究价值和挑战性的难题
- 计算光流的先验知识: 三维物点 **P** 相对应的图像点及其附近的灰度变化规律在运动过程中保持不变。
- 光流的计算方法：

（1）基于特征法

主要运算：匹配。（类似立体视觉的匹配问题）

（2）基于梯度法

主要运算：图像梯度的计算+约束条件。

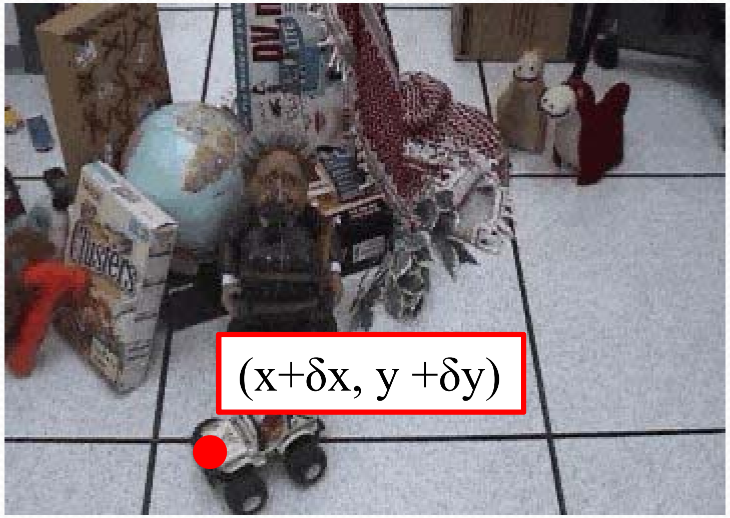
图像内光流的求解

- 假设空间物点 **P** 在 t 时刻, 在图像平面内的投影为(x, y)， 在 $t+\delta t$ 时刻在图像平面内的投影为 $(x+\delta x, y+\delta y)$；

- 假设光照条件不变，假设场景内该物点在图像内的投影的像素灰度值不随时间而改变， 令 E 表示图像像素的灰度值：

$$E(x,y,t)=E(x+\delta x,y+\delta y,t+\delta t)$$

Time = t



Time = $t+\delta t$

光流约束方程

等式右边利用一阶Taylor级数展开：

$$E(x,y,t)+\delta x\frac{\partial E}{\partial x}+\delta y\frac{\partial E}{\partial y}+\delta t\frac{\partial E}{\partial t}=E(x,y,t)$$

用 δt 除以等式两边, 并取极限 $\delta t \to 0$

$$\frac{\partial E}{\partial x}\frac{dx}{dt}+\frac{\partial E}{\partial y}\frac{dy}{dt}+\frac{\partial E}{\partial t}=0$$

令 $u(x,y)=\dfrac{dx}{dt}$, $v(x,y)=\dfrac{dy}{dt}$, $E_x=\dfrac{\partial E}{\partial x}$, $E_y=\dfrac{\partial E}{\partial y}$, $E_t=\dfrac{\partial E}{\partial t}$

得到光流约束方程（ **Optical Flow Constraint Equation**）：

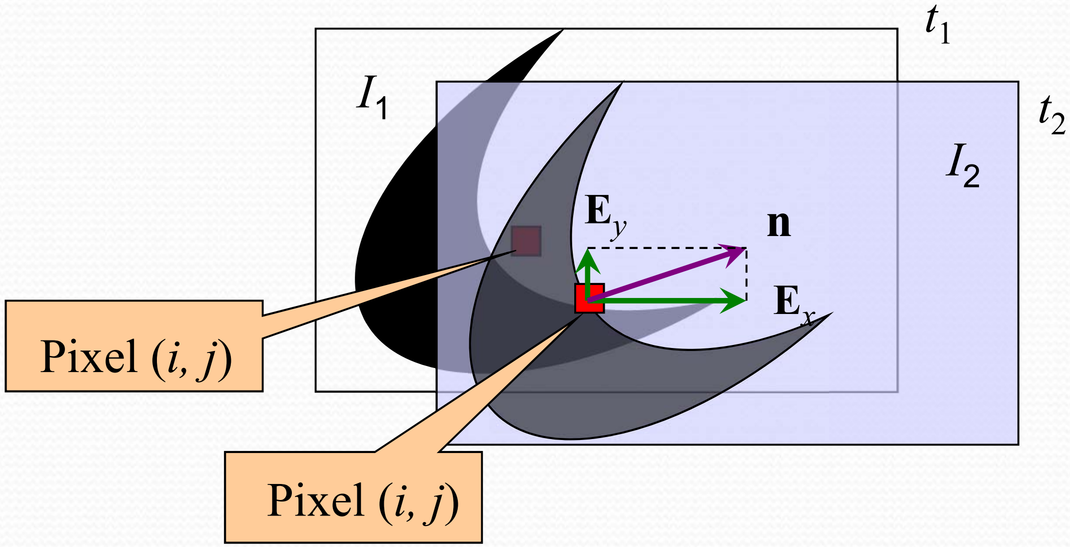
$$\boxed{E_x u+E_y v+E_t=0}$$

其中 $[u(x,y) \quad v(x,y)]^T$ 表示为该像素点的光流矢量

光流约束方程

- 图像序列中光流约束方程系数的计算：

E_x, E_y, E_t可由图像序列的灰度信息直接计算



$E_x = I_2(i, j+1) - I_2(i, j)$

$E_y = I_2(i, j) - I_2(i+1, j)$

$E_t = I_2(i, j) - I_1(i, j)$

梯度方向和E_t

光流约束方程的几何意义

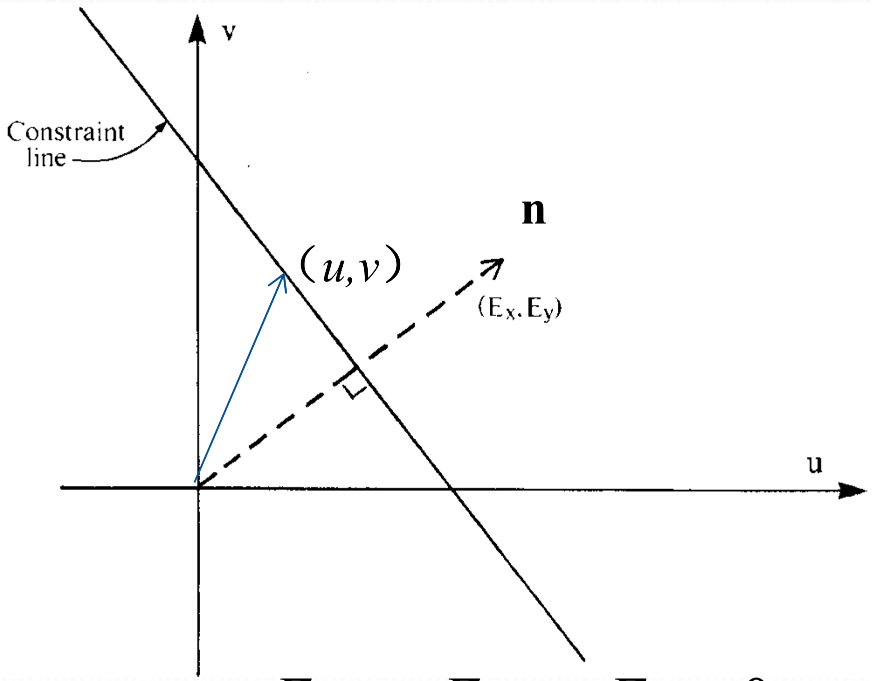
- 在解空间 u-v 平面, 光流的解为一条直线。直线上任意一点都可能是所求光流的解。

- 所求的光流（u, v） 在向量 **n** （垂直于直线）上的投影（点乘）：

$$\left(E_x, E_y\right) \cdot \left(u, v\right) = -E_t$$

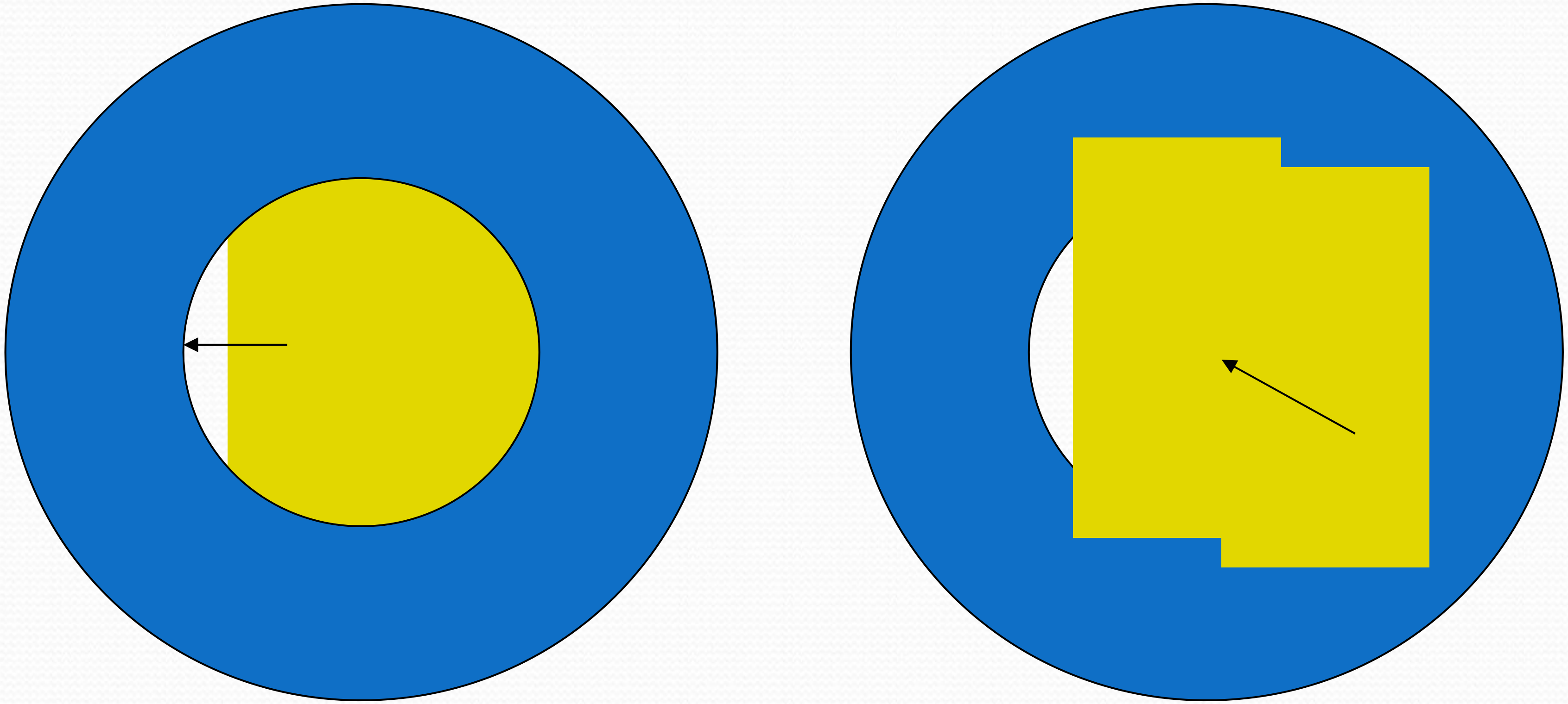
其中单位法向量：

$$\vec{\mathbf{n}} = \frac{\left(E_x \quad E_y\right)^T}{\left\|\left(E_x, E_y\right)^T\right\|} \Rightarrow (u \quad v)^T \bullet \vec{\mathbf{n}} = \frac{-E_t}{\left\|\left(E_x, E_y\right)^T\right\|}$$



$$E_x u + E_y v + E_t = 0$$

孔径问题（Aperture Problem）



通过小孔观察物体的运动，是否能反映物体运动的真实情况？

光流的求解的附加约束

（1）光流的平滑性。即图像邻域内的诸点有相近的图像速度值。

（2）光流在图像中的某一局域上为常数，例如对应一个物体的某平面的局域内为一个常数。

（3）光流是有限制的运动的结果，例如，物体只在限定的平面内运动。

作用：利用以上的约束条件，唯一的确定某点的光流

光流的求解方法: H-S算法

- Horn and Schunck 算法：利用两个速度分量的差分来测定平滑度。

邻域内的光流均值（即邻域内的平均速度值）：$\overline{u}$ ，$\overline{v}$

记： $\Delta u=u-\overline{u}$ $\Delta v=v-\overline{v}$

所谓平滑，即在给定邻域内，使得：

$$\Delta u^2+\Delta v^2=(u-\overline{u})^2+(v-\overline{v})^2$$

达到最小。

光流的求解方法: H-S算法

- 采用拉格朗日待定系数法求解，上述问题等效为求解以下问题：求解（u，v），使得：

$$\iint e(u(x,y),v(x,y))dxdy = \iint [\underbrace{(E_x u + E_y v + E_t)^2}_{\text{Constant Brightness term}} + \underbrace{\lambda^2(\Delta u^2 + \Delta v^2)}_{\text{Smooth Flow term}}]dxdy$$

达到最小。

λ 为系数，为一正数，用以控制平滑约束的权重

光流的求解方法：H-S算法

$$\iint e(u(x,y),v(x,y))dxdy = \iint [(E_x u + E_y v + E_t)^2 + \lambda^2(\Delta u^2 + \Delta v^2)]dxdy$$

利用微分法求上式最小化时的光流矢量 $(u, v)^T$

$$\Delta u = u - \bar{u}\,; \quad \Delta v = v - \bar{v}$$

误差 $e(u, v)$ 关于u，v 的偏微分为：

$$\frac{\partial e}{\partial u} = 2\left(E_x u + E_y v + E_t\right)E_x + 2\lambda^2(u - \bar{u})$$

$$\frac{\partial e}{\partial v} = 2\left(E_x u + E_y v + E_t\right)E_y + 2\lambda^2(v - \bar{v})$$

误差达到最小值时，偏微分为 0 , 可得：

$$\left[E_x^{\,2} + \lambda^2\right]u + E_x E_y v = \lambda^2 \bar{u} - E_x E_t$$

$$E_x E_y u + \left[E_y^{\,2} + \lambda^2\right]v = \lambda^2 \bar{v} - E_y E_t$$

光流的求解方法：H-S算法

解以上线性方程组，稀疏矩阵行列式为：$\lambda^4+\lambda^2 E_x{}^2+\lambda^2 E_y{}^2$，

可得：

$$u=\frac{\lambda^2\bar{u}E_y{}^2+\lambda^4\bar{u}-\lambda^2 E_xE_t-\lambda^2\bar{v}E_xE_y}{\lambda^4+\lambda^2E_x{}^2+\lambda^2E_y{}^2}$$

$$=\frac{\bar{u}E_x{}^2+\bar{u}E_y{}^2+\lambda^2\bar{u}-E_xE_t-\bar{v}E_xE_y-\bar{u}E_x{}^2}{\lambda^2+E_x{}^2+E_y{}^2}$$

$$=\bar{u}-\frac{E_x\bar{u}+E_y\bar{v}+E_t}{\lambda^2+E_x{}^2+E_y{}^2}E_x$$

$$v=\bar{v}-\frac{E_x\bar{u}+E_y\bar{v}+E_t}{\lambda^2+E_x{}^2+E_y{}^2}E_y$$

由于上式中，邻域光流平均值 $\bar{u}$ 、$\bar{v}$ 未知，采用迭代法求解。

光流的求解方法: H-S算法

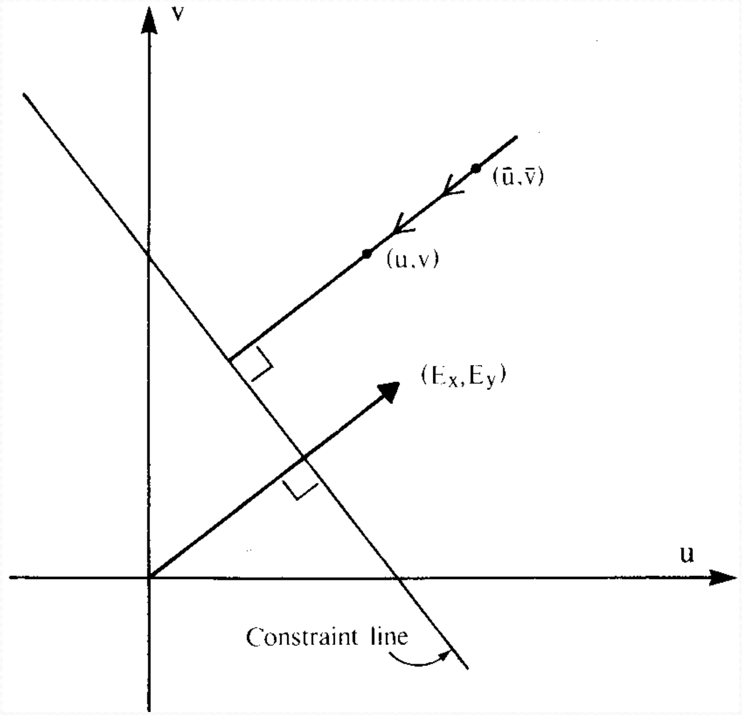
- 对两个光流变量求偏导数，并令其为零。整理可得松弛算法的迭代递推公式：

$$u^{n+1} = \bar{u}^n - \frac{\left(E_x \bar{u}^n + E_y \bar{v}^n + E_t\right)}{\lambda^2 + E_x{}^2 + E_y{}^2} E_x$$

$$v^{n+1} = \bar{v}^n - \frac{\left(E_x \bar{u}^n + E_y \bar{v}^n + E_t\right)}{\lambda^2 + E_x{}^2 + E_y{}^2} E_y$$

$\bar{u}, \bar{v}$ ： 邻域内的平均速度值

$\bar{u}^0, \bar{v}^0$ ： 初始值可以选为0



- 计算时，在一个时间间隔反复迭代，直至满足误差要求。
- 几何含义：在每一次迭代过程中，新的光流值由其邻域的光流平均值出发，在此基础上，沿着梯度方向，减去一个调整值，向光流约束方程代表的直线逼近。

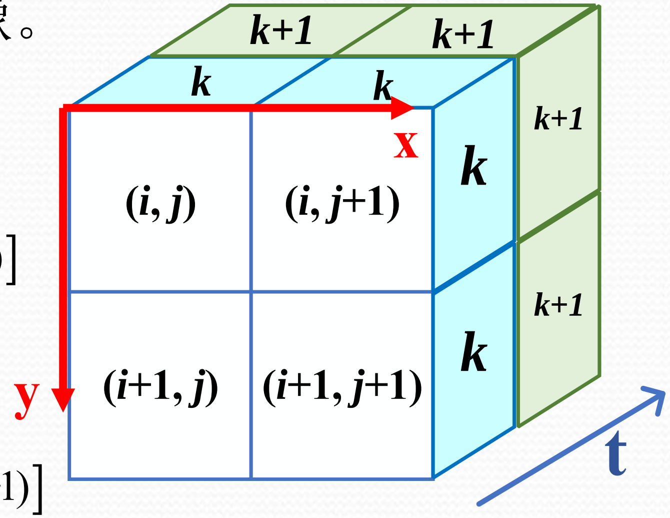
光流的求解方法：H-S算法

- 光流约束方程系数的计算：

要估计某一像素点的光流，需要得到像素附近邻域内的亮度值。

针对视频，“邻域”为本帧图像、及下一帧图像。

E_x, E_y, E_t 可由图像序列的灰度信息直接计算：



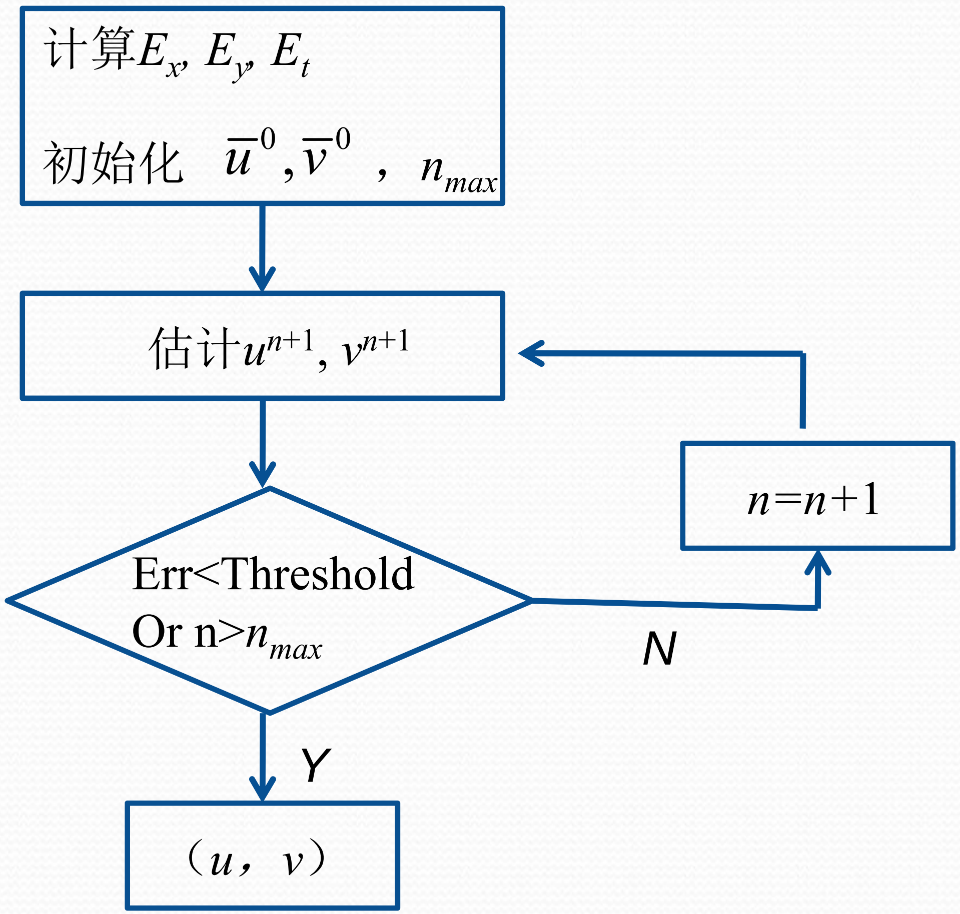
$$E_x = \frac{1}{4}\left[E(i,j+1,k)+E(i+1,j+1,k)+E(i,j+1,k+1)+E(i+1,j+1,k+1)\right]$$
$$-\frac{1}{4}\left[E(i,j,k)+E(i+1,j,k)+E(i,j,k+1)+E(i+1,j,k+1)\right]$$

$$E_y = \frac{1}{4}\left[E(i+1,j,k)+E(i+1,j+1,k)+E(i+1,j,k+1)+E(i+1,j+1,k+1)\right]$$
$$-\frac{1}{4}\left[E(i,j,k)+E(i,j+1,k)+E(i,j,k+1)+E(i,j+1,k+1)\right]$$

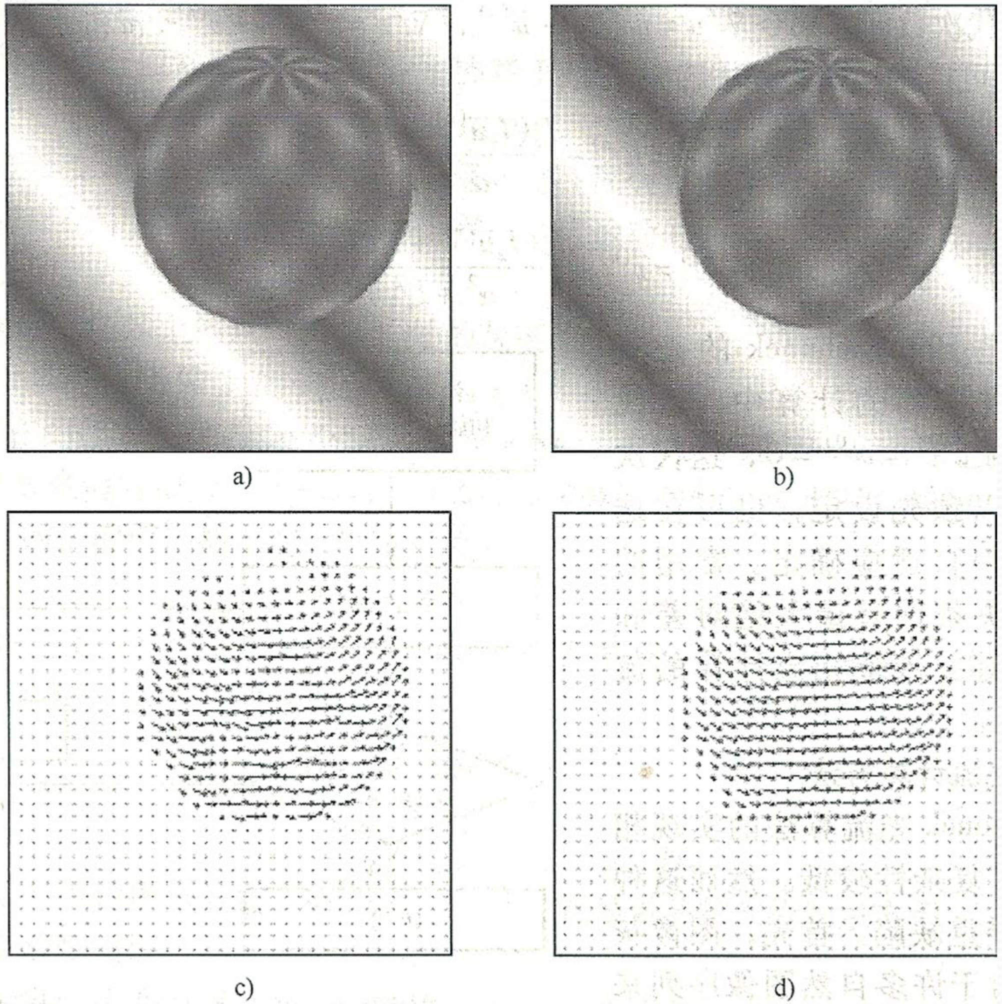
$$E_t = \frac{1}{4}\left[E(i,j,k+1)+E(i,j+1,k+1)+E(i+1,j,k+1)+E(i+1,j+1,k+1)\right]$$
$$-\frac{1}{4}\left[E(i,j,k)+E(i,j+1,k)+E(i+1,j,k)+E(i+1,j+1,k)\right]$$

光流的求解方法：H-S算法

求解流程：



光流的求解举例

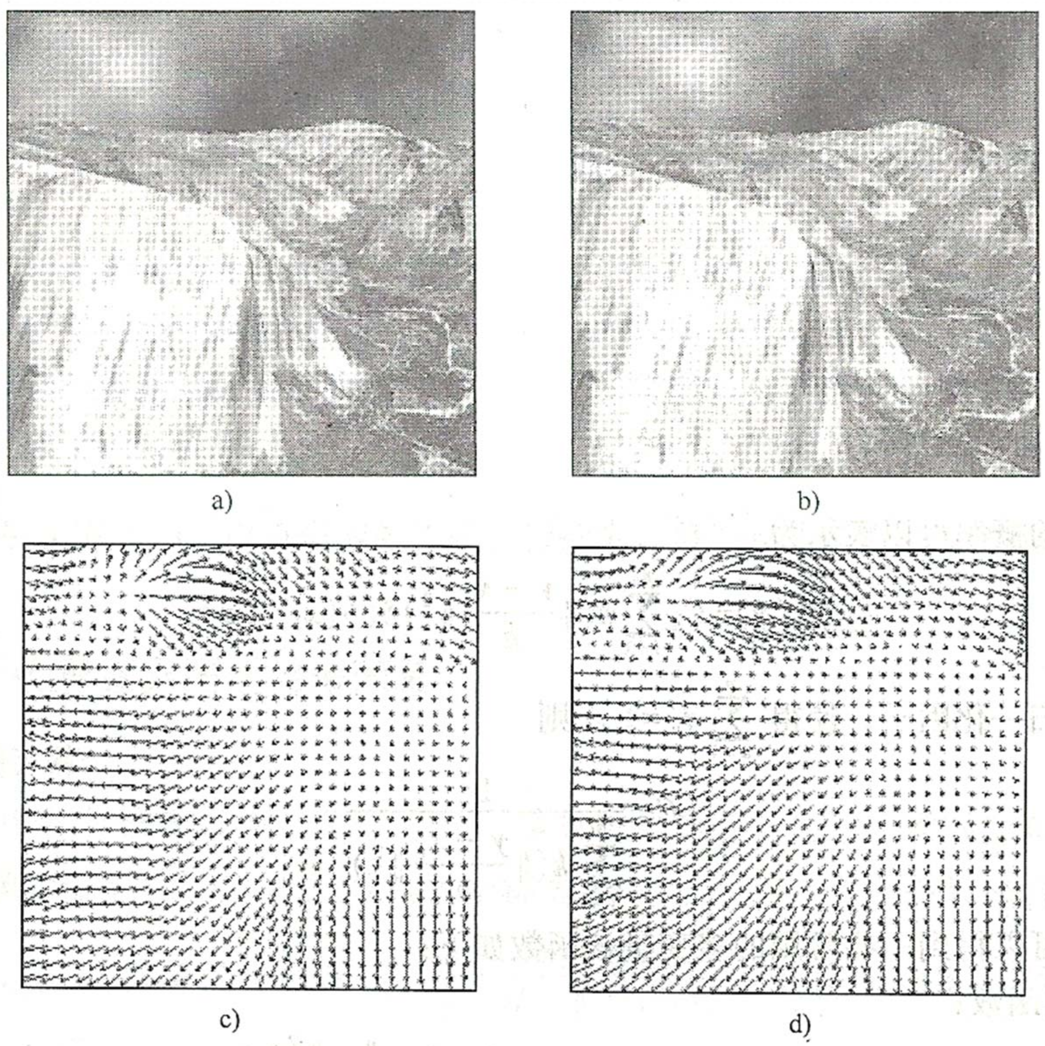


模板大小对光流估计的影响

a）序列图像第 5 帧　b）序列图像第 6 帧　c）模板大小 3×3　d）模板大小 5×5

模板越大，所求光流在图像边缘处越平滑

光流的求解举例



迭代次数对光流估计的影响：

(a)输入图像第5帧
(b)输入图像第6帧
(c)迭代次数n=100
(d)迭代次数n=200
迭代次数越多，所求光流的方向信息越清晰

光流的求解方法: H-S算法

- 影响光流估计结果的主要参数：平滑模板的尺寸和迭代次数
- H-S算法实现简单，计算复杂性低，但是存在以下技术缺陷：

（1）图像灰度不变假设对于很多自然情况不成立。当运动速度较高时，基于灰度保持假设的约束存在较大的误差。

（2）在图像的遮挡区域，速度场是突变的，总体平滑约束平滑掉了关于物体形状的重要信息。