计算机视觉

视觉运动

中国传媒大学
COMMUNICATION UNIVERSITY OF CHINA

本节主题：

亮度恒常约束

本节主题：

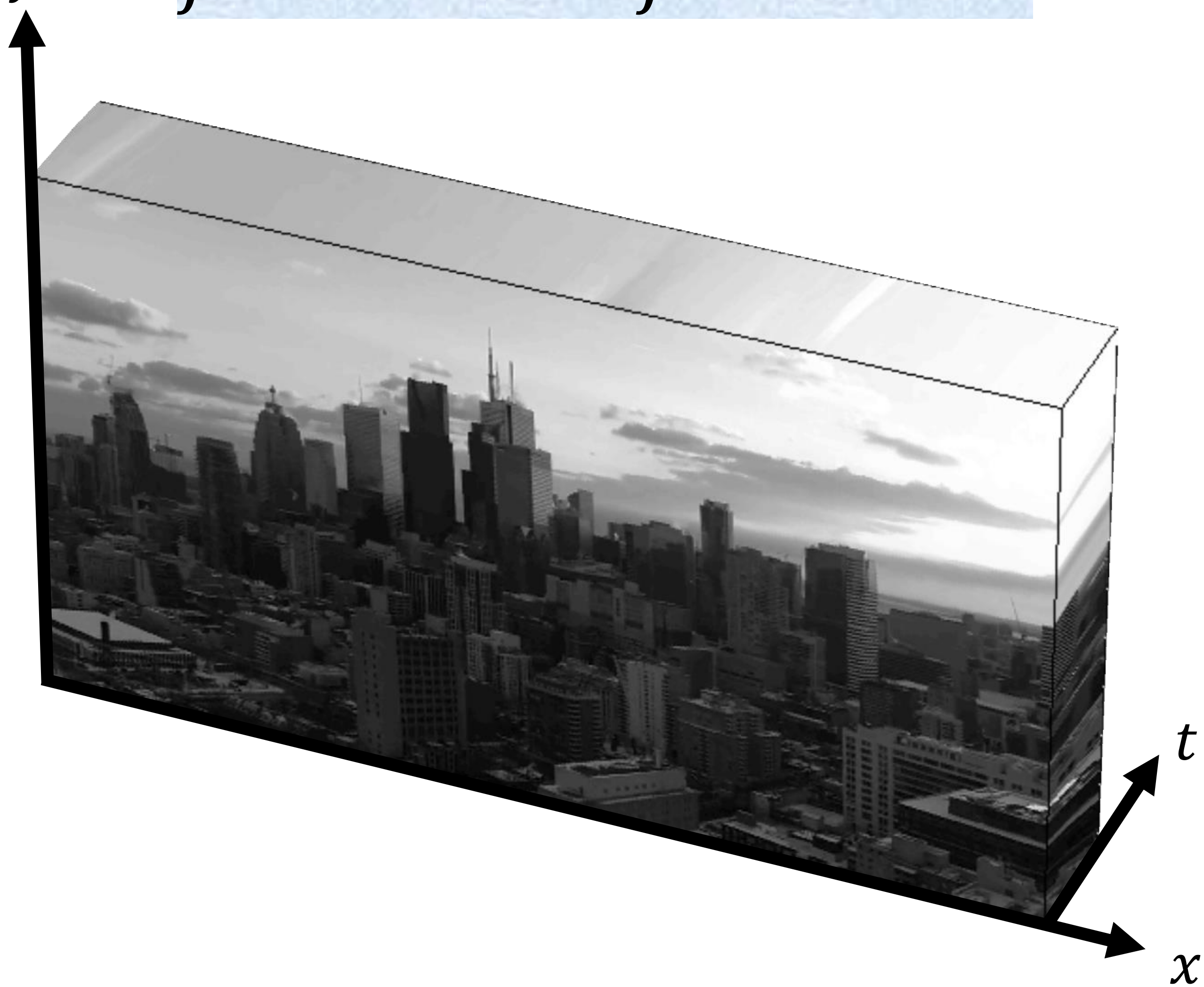
亮度恒常约束

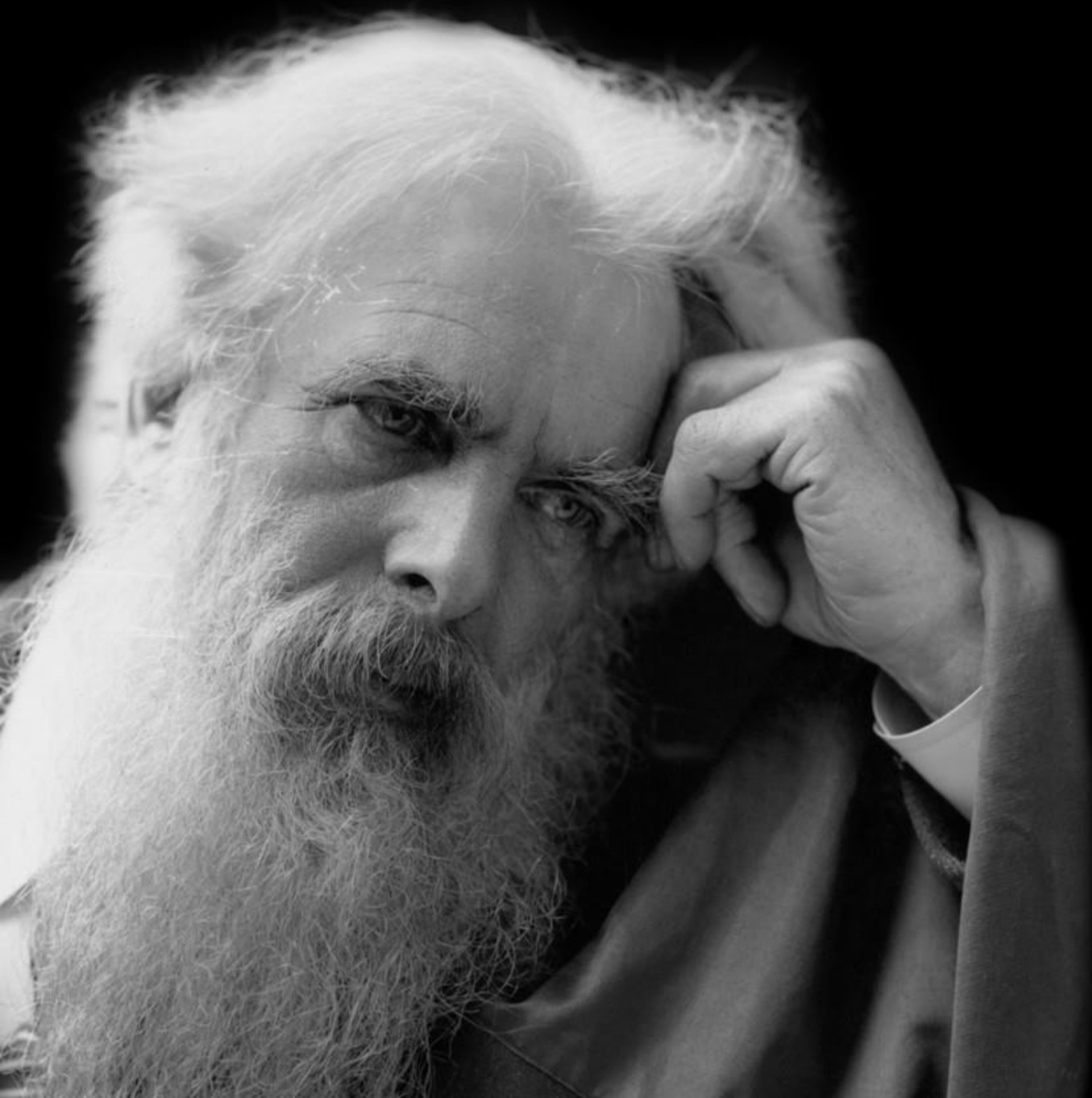
光流估计

图像即函数

$$f: \mathbb{R}^2 \to \mathbb{R} \text{ or } f: \mathbb{R}^2 \to \mathbb{R}^3$$

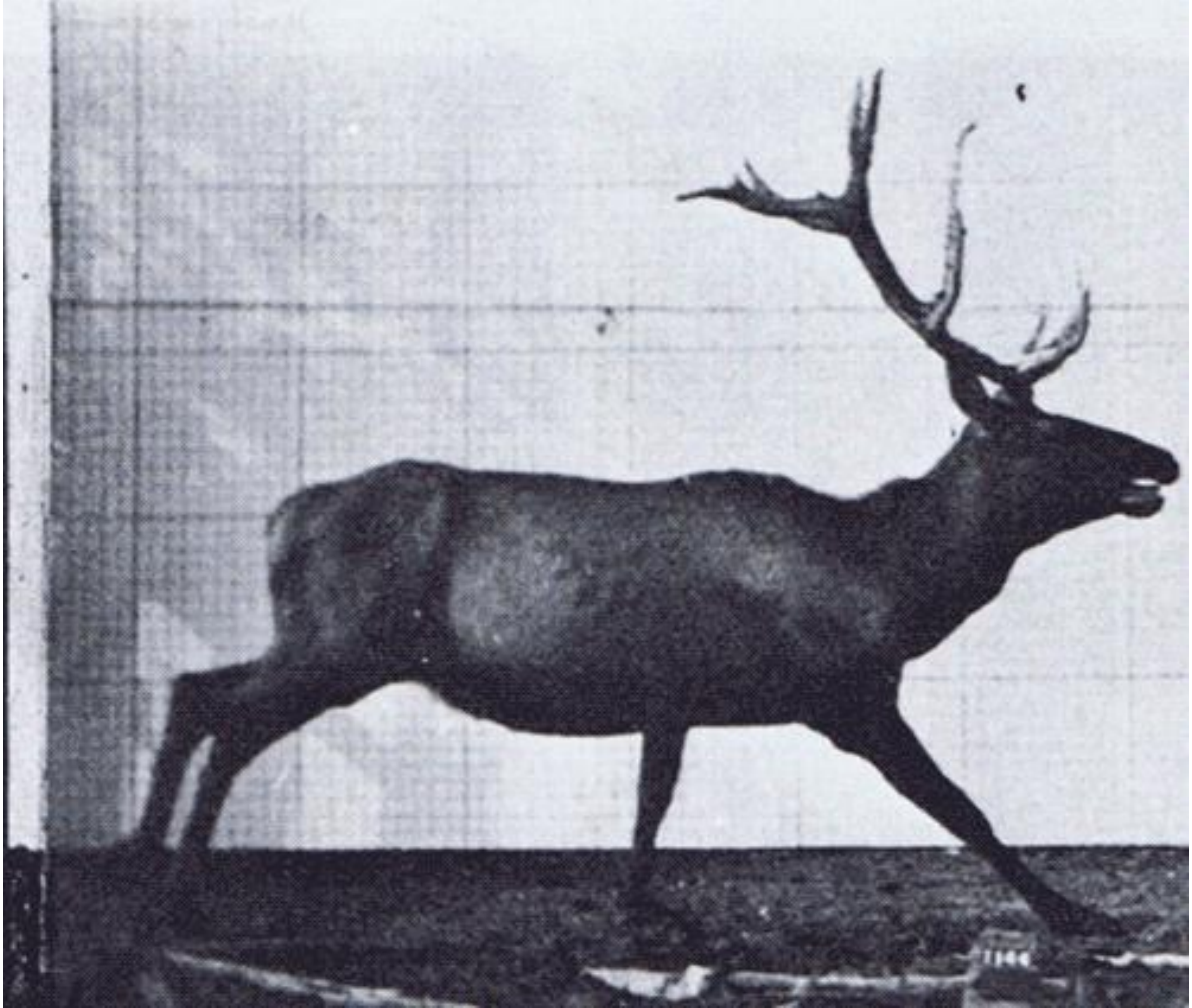
$f: \mathbb{R}^3 \rightarrow \mathbb{R}$ **or** $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$

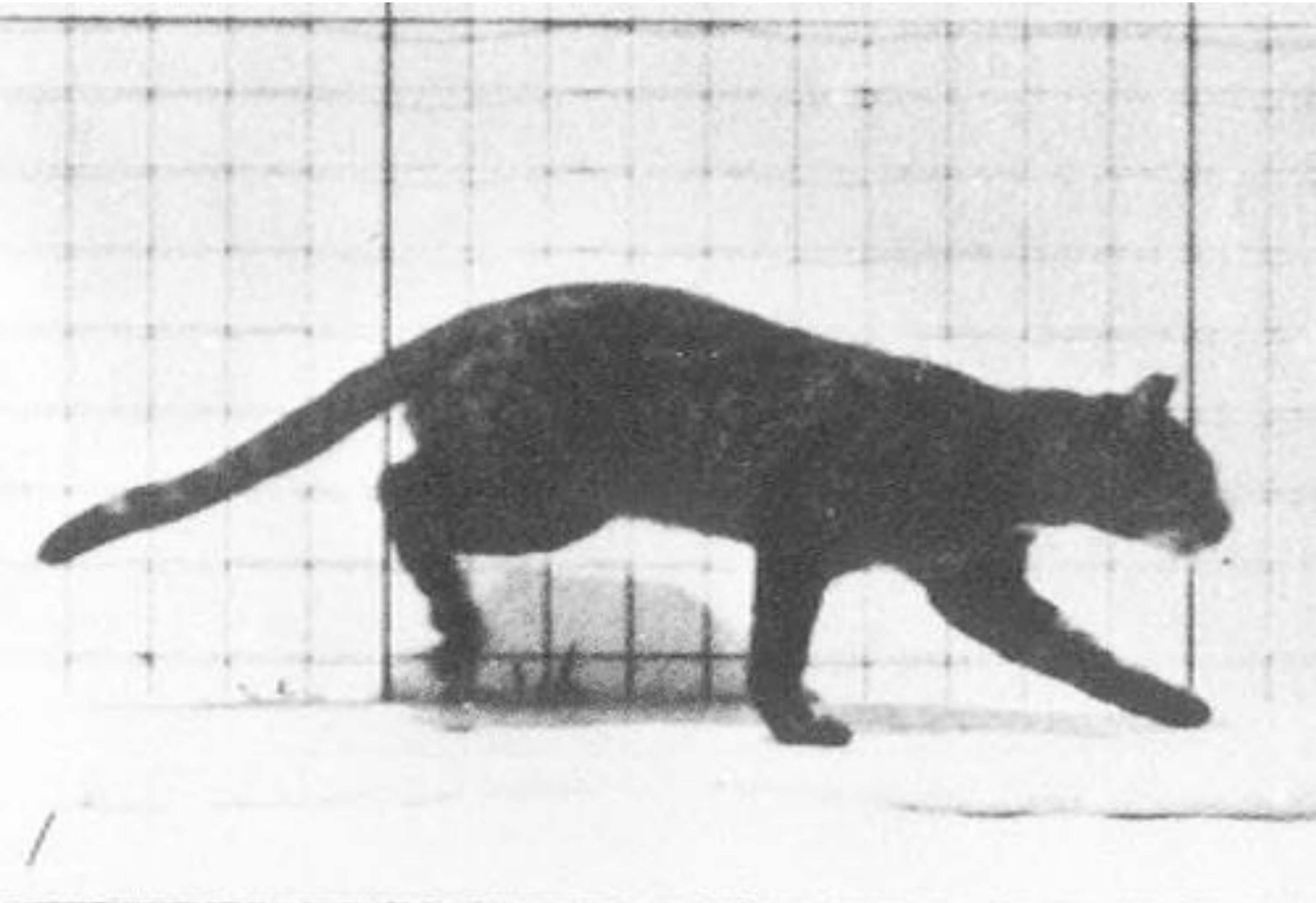


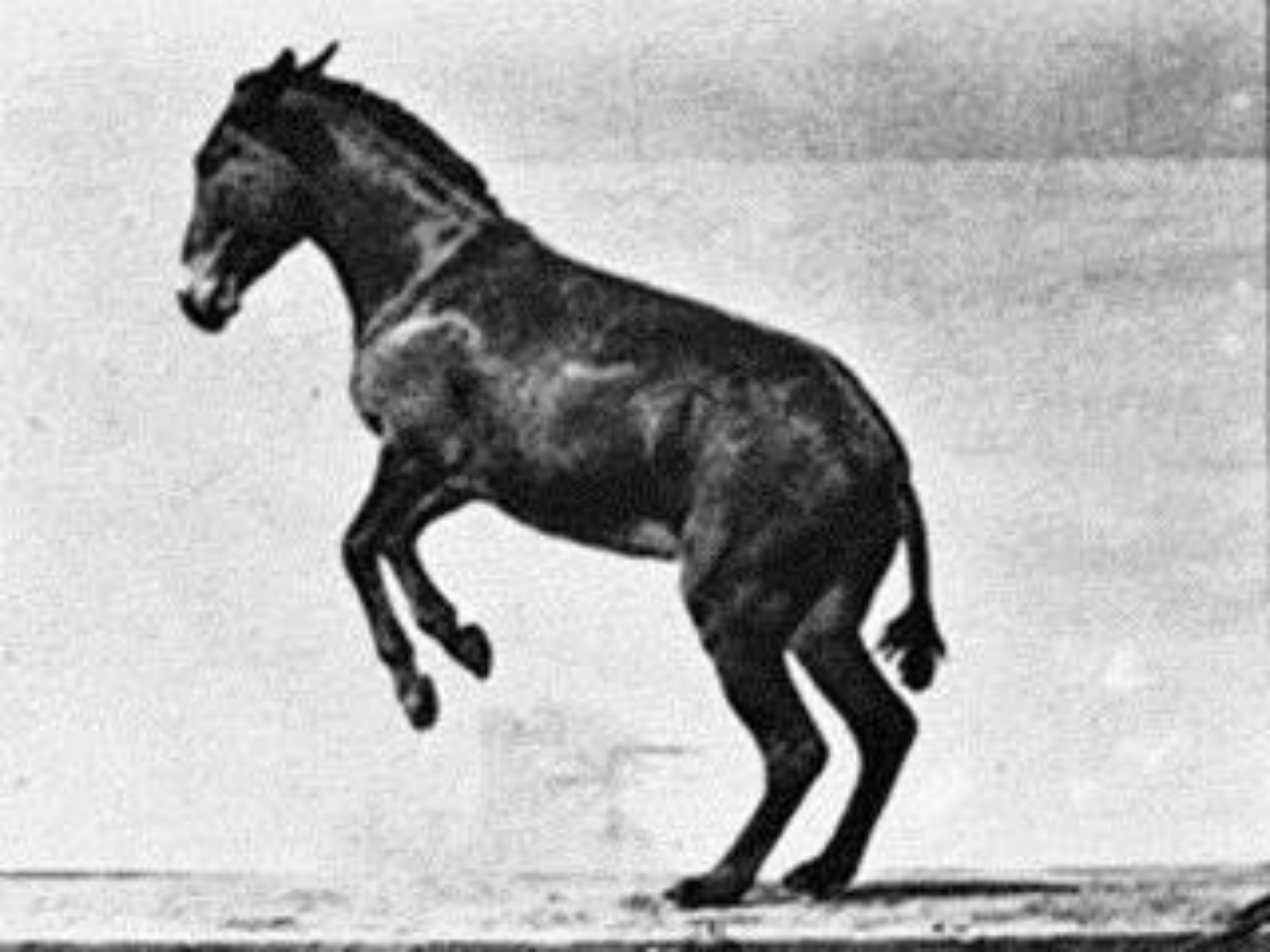


埃德沃德·迈布里奇
1830 - 1904


所有的蹄子都同时离开地面了吗?







运动分析

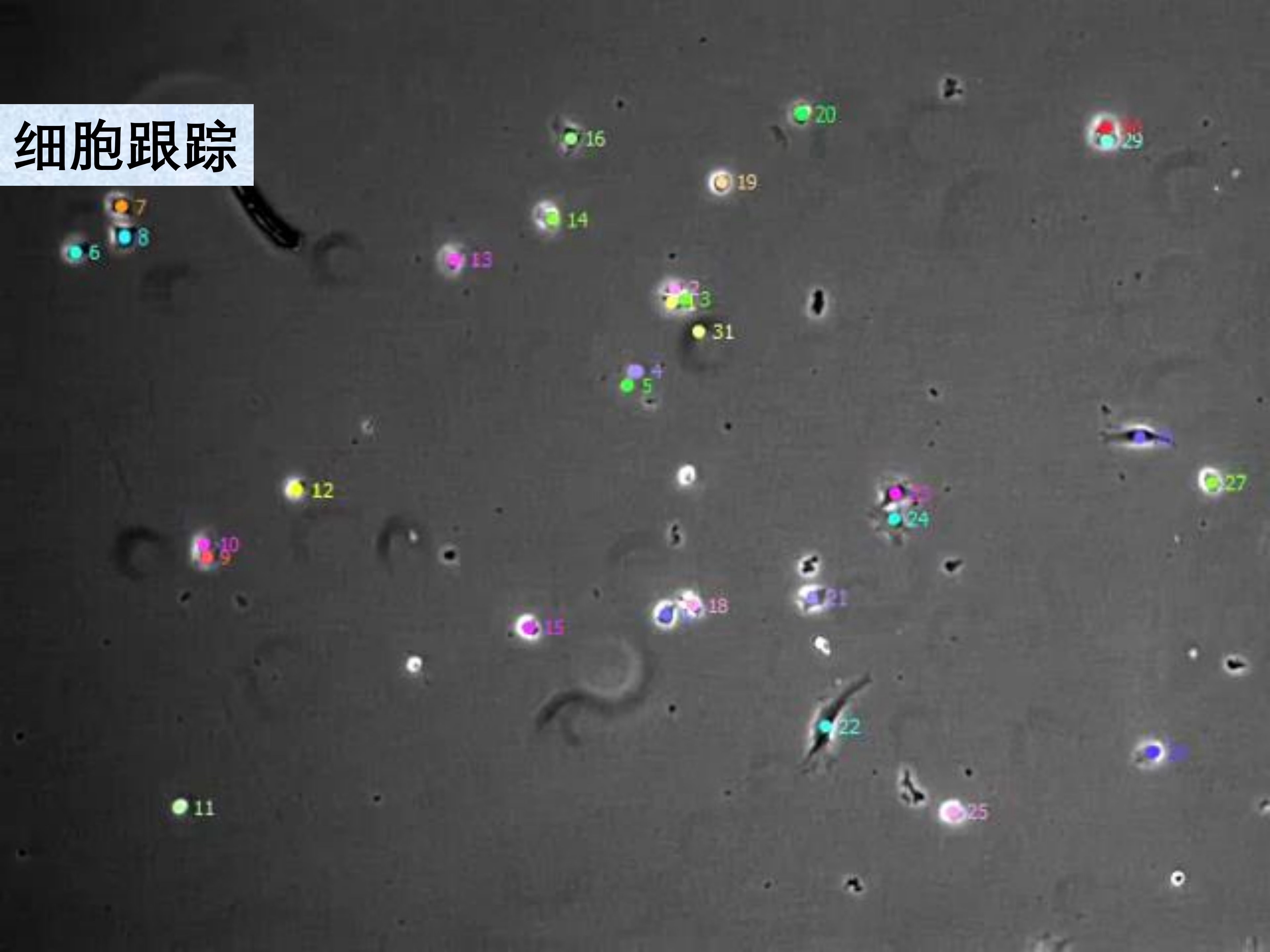

You see that an octopus would
stand out very easily there
你看章鱼在那里很容易脱颖而出

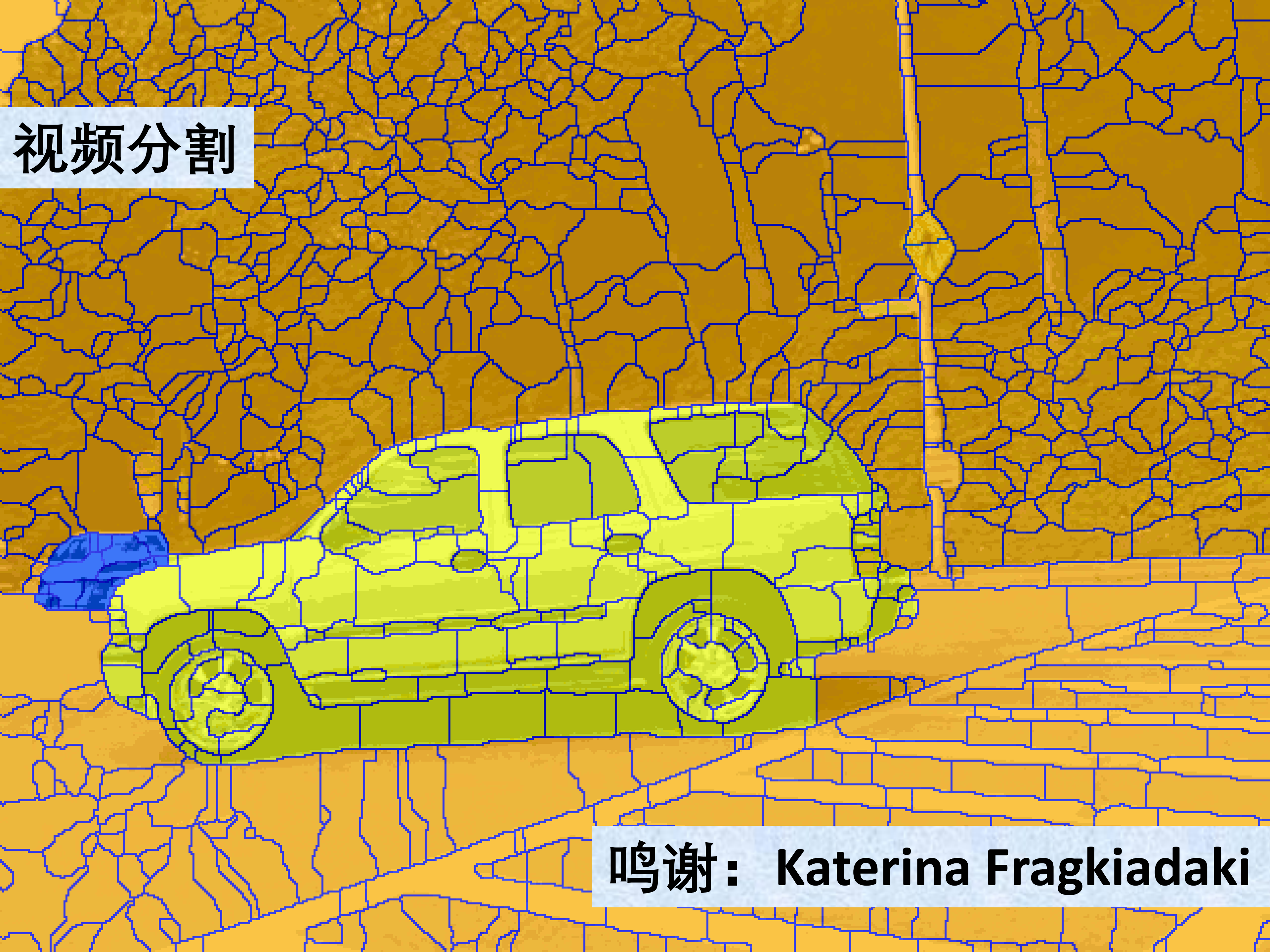
跟踪： **Babenko et al., CVPR, 2009**

最先进的跟踪器


鸣谢： Boris Babenko

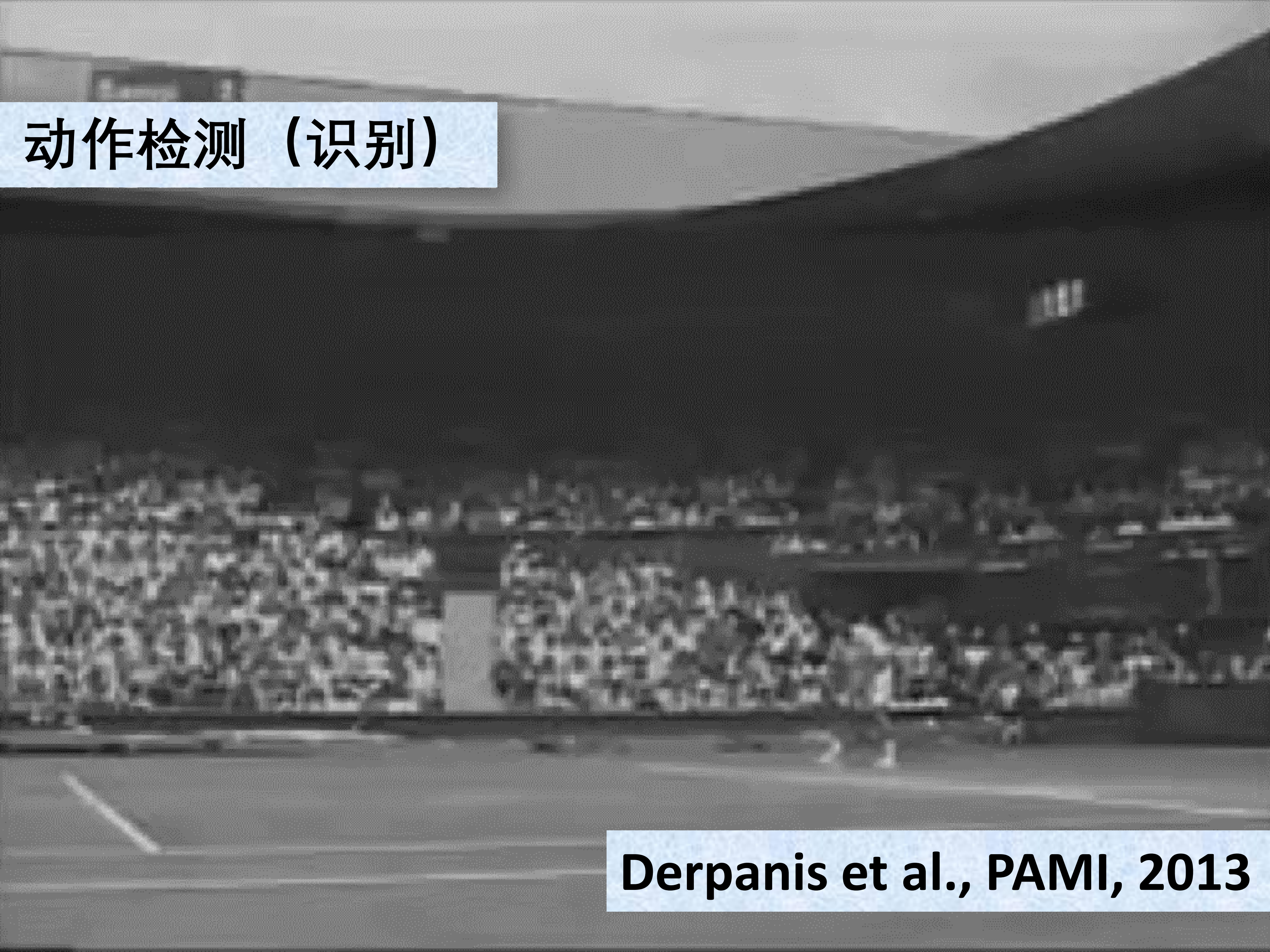
细胞跟踪




视频分割
鸣谢： Katerina Fragkiadaki

视频稳定（去抖动）

动作检测（识别）



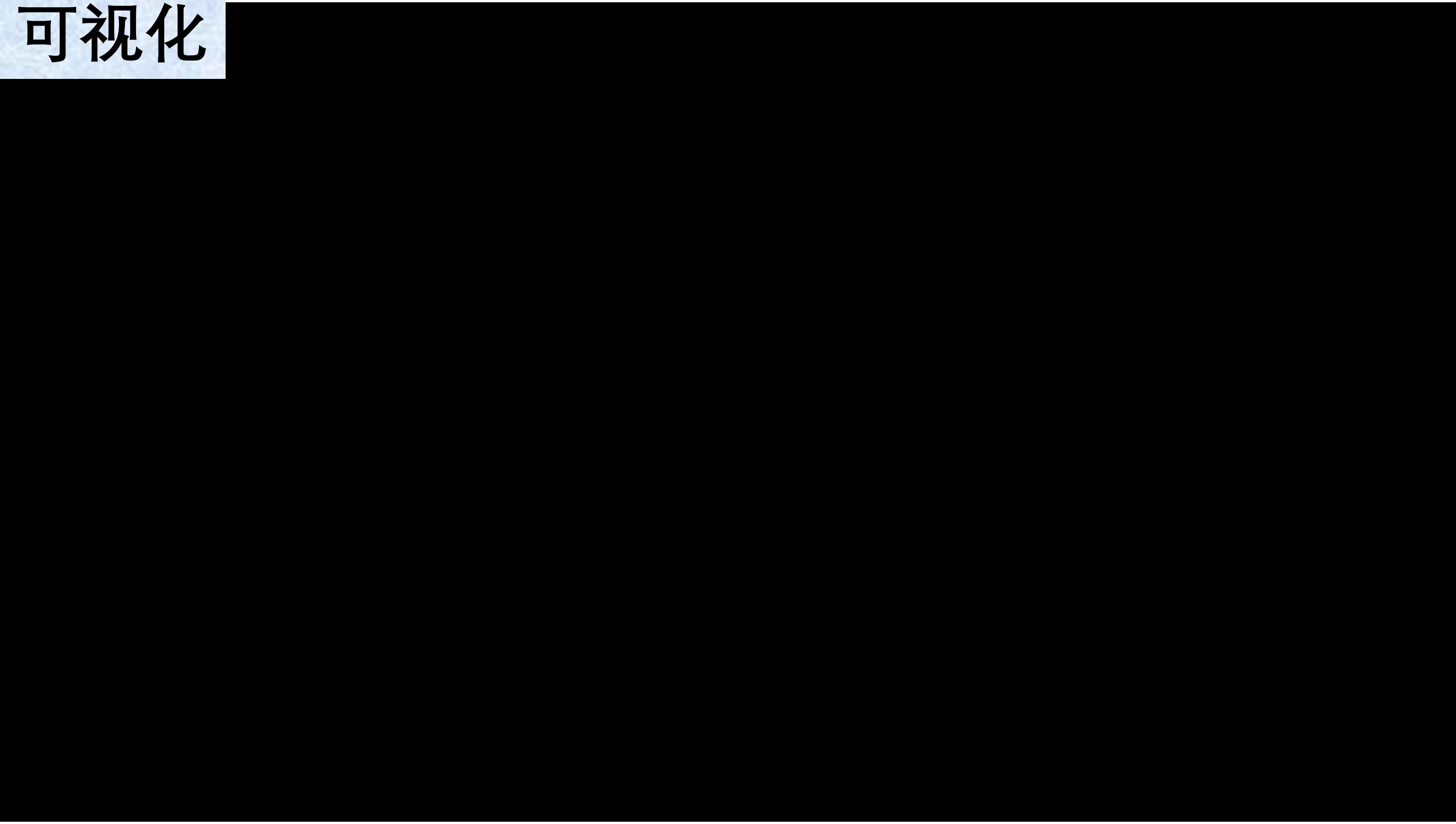
Derpanis et al., PAMI, 2013

自动导航

鸣谢：Andreas Geiger

可视化

可视化



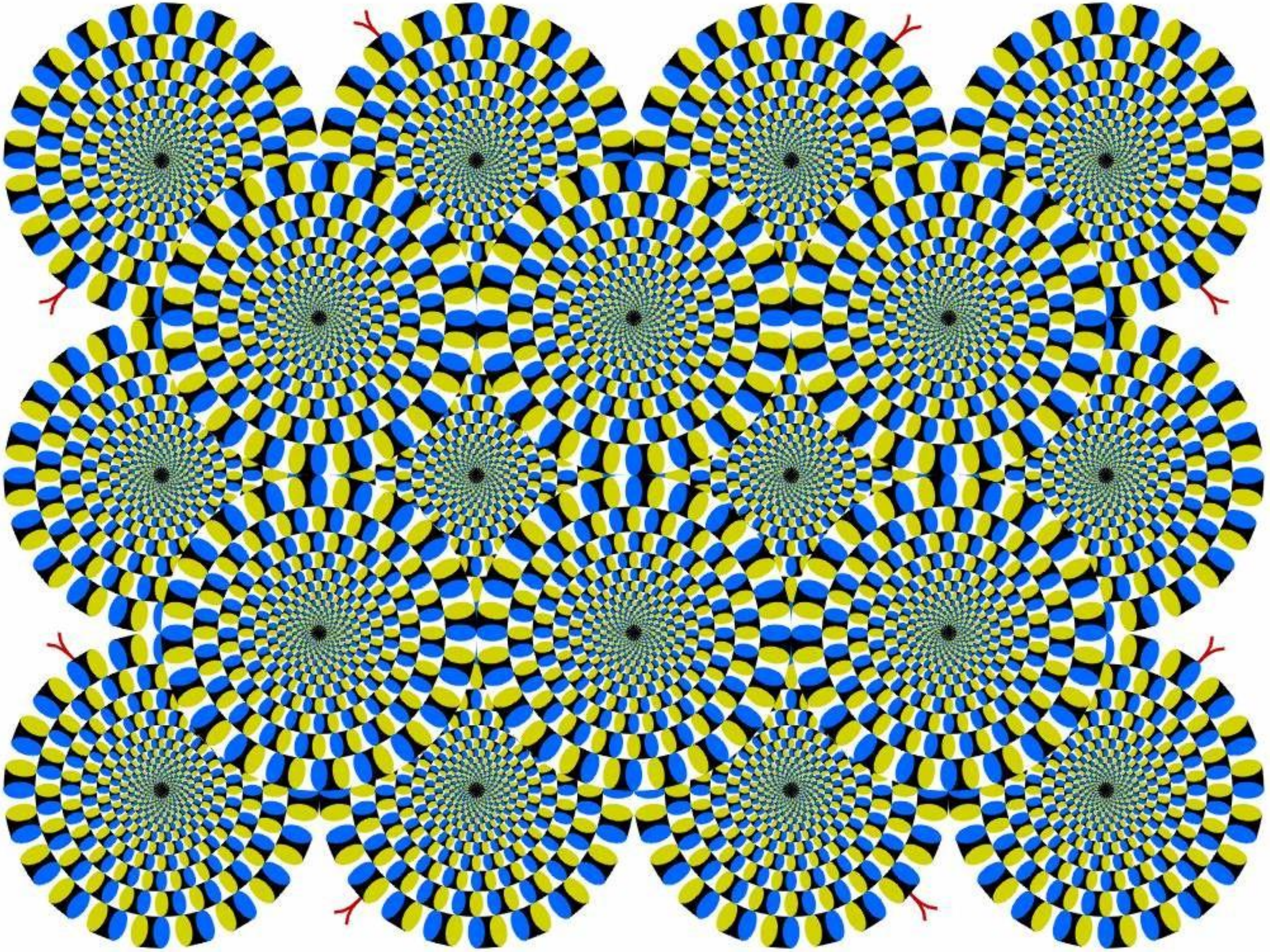
Perception & Psychophysics
1973, Vol. 14, No. 2, 201-211

Visual perception of biological motion and a model for its analysis*

GUNNAR JOHANSSON
University of Uppsala, S:t Larsgatan 2, S-752 20 Uppsala, Sweden

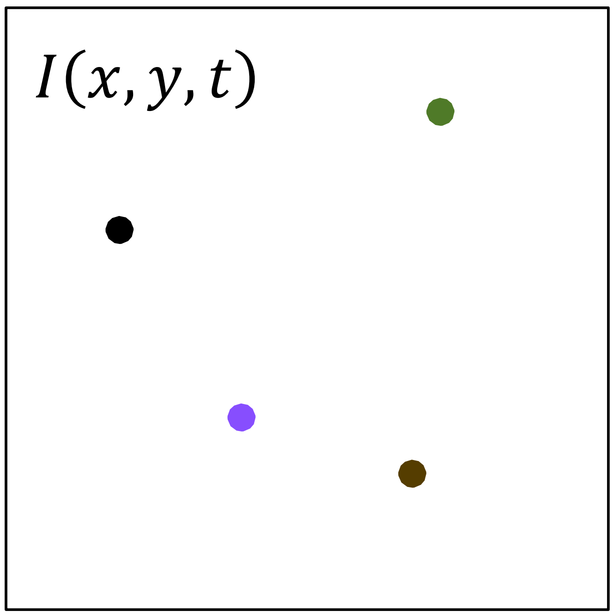
This paper reports the first phase of a research program on visual perception of motion patterns characteristic of living organisms in locomotion. Such motion patterns in animals and men are termed here as biological motion. They are characterized by a far higher degree of complexity than the patterns of simple mechanical motions usually studied in our laboratories. In everyday perceptions, the visual information from biological motion and from the corresponding figurative contour patterns (the shape of the body) are intermingled. A method for studying information from the motion pattern per se without interference with the form aspect was devised. In short, the motion of the living body was represented by a few bright spots describing the motions of the main joints. It is found that 10-12 such elements in adequate motion combinations in proximal stimulus evoke a compelling impression of human walking, running, dancing, etc. The kinetic-geometric model for visual vector analysis originally developed in the study of perception of motion combinations of the mechanical type was applied to these biological motion patterns. The validity of this model in the present context was experimentally tested and the results turned out to be highly positive.

Perception & Psychophysics, 1973

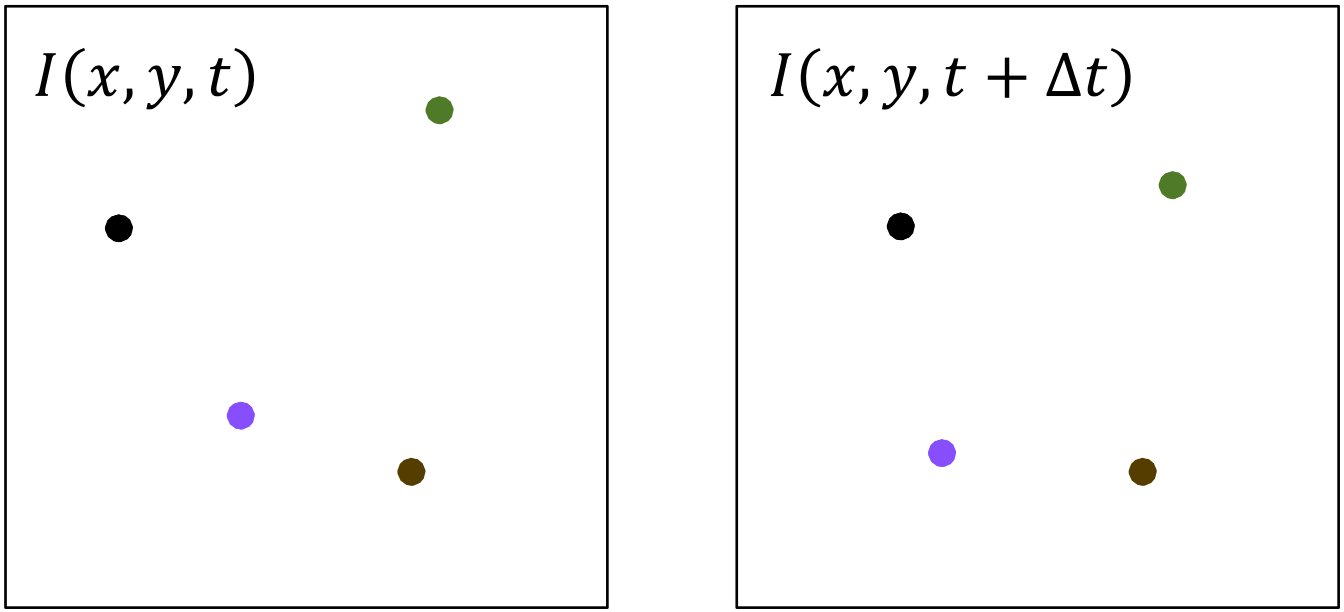


问题定义

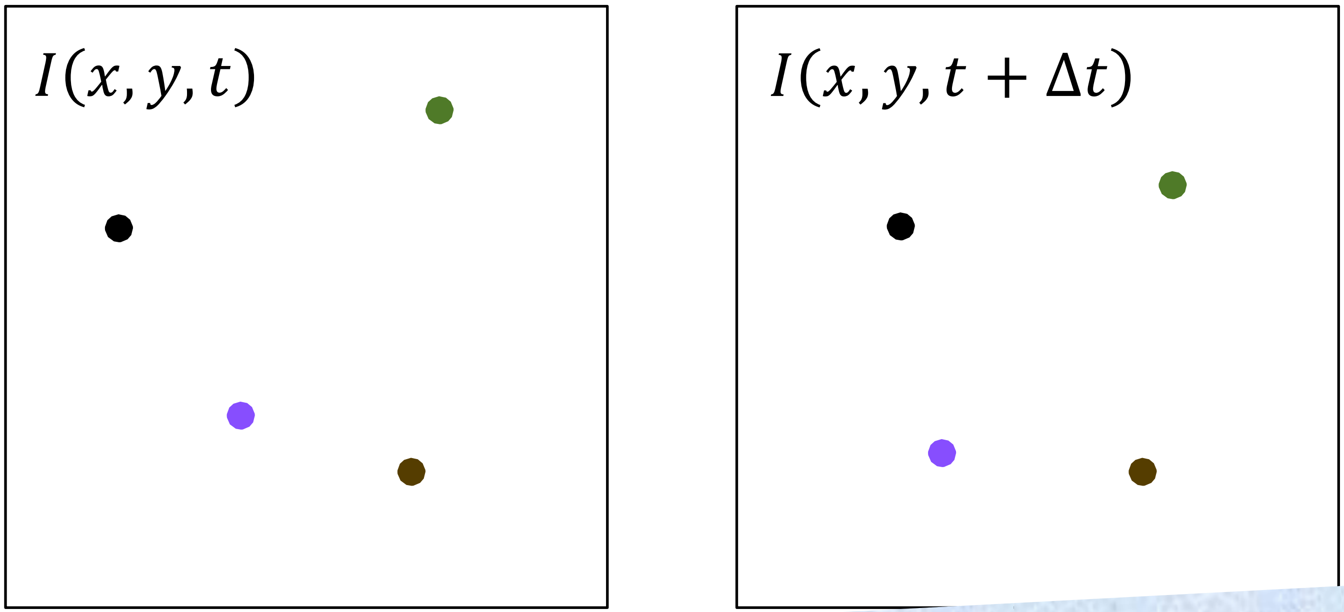
问题定义



问题定义

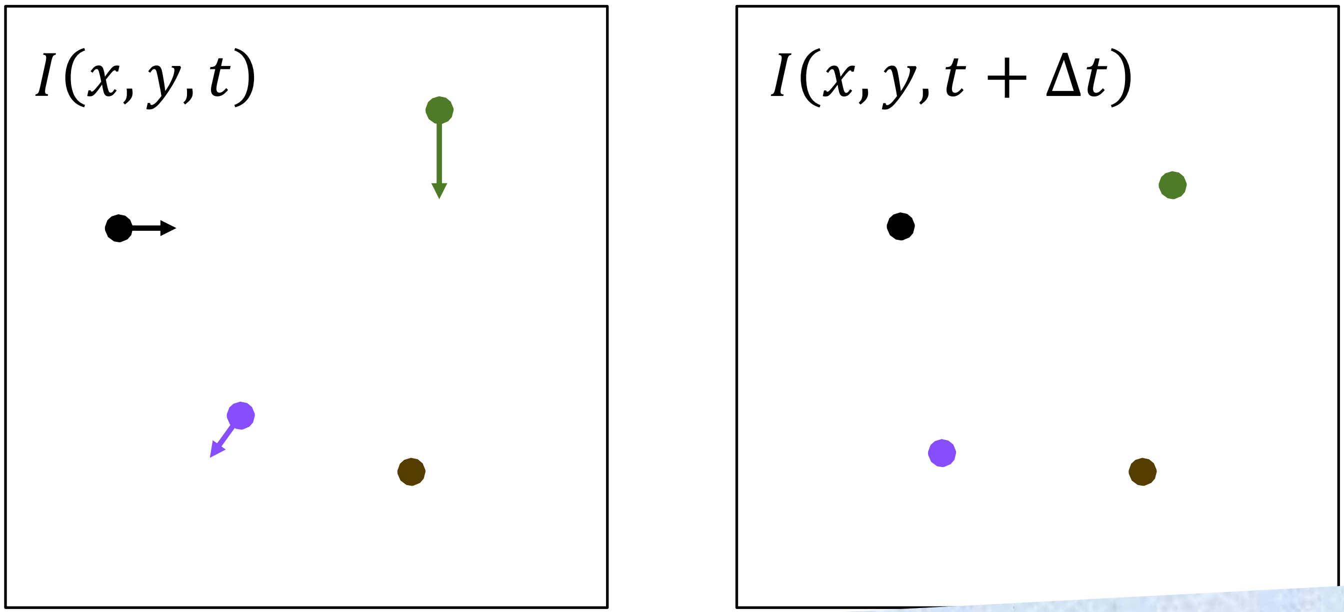


问题定义



目标：估计图像之间的像素位移

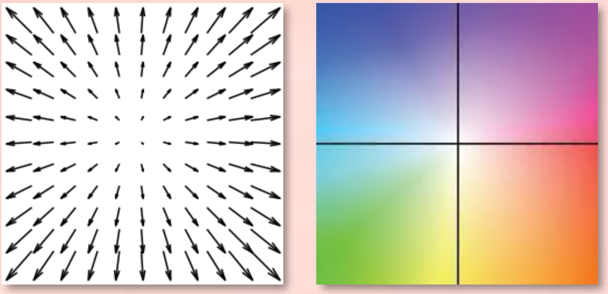
问题定义



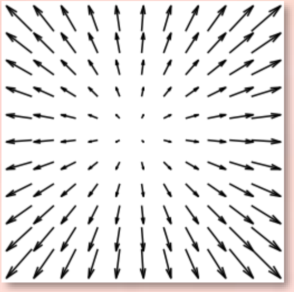
目标：估计图像之间的像素位移


输入序列
30

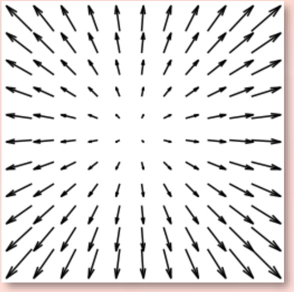
光流

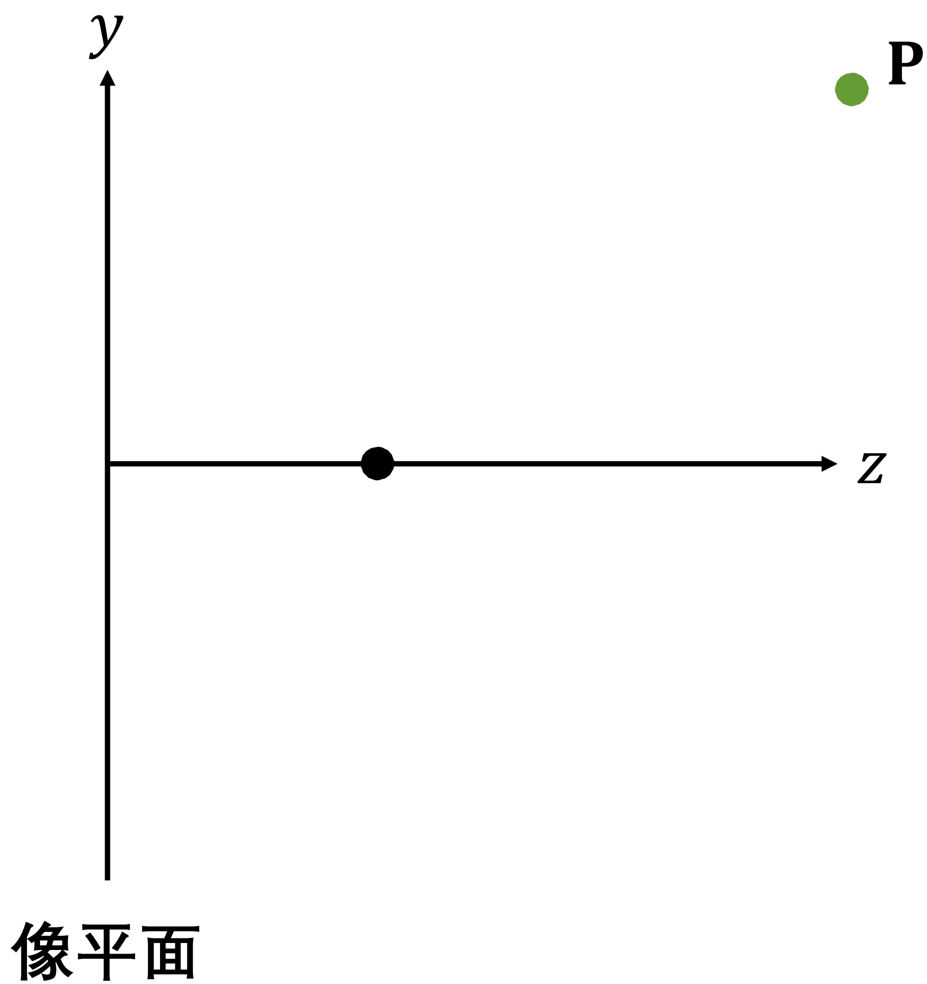


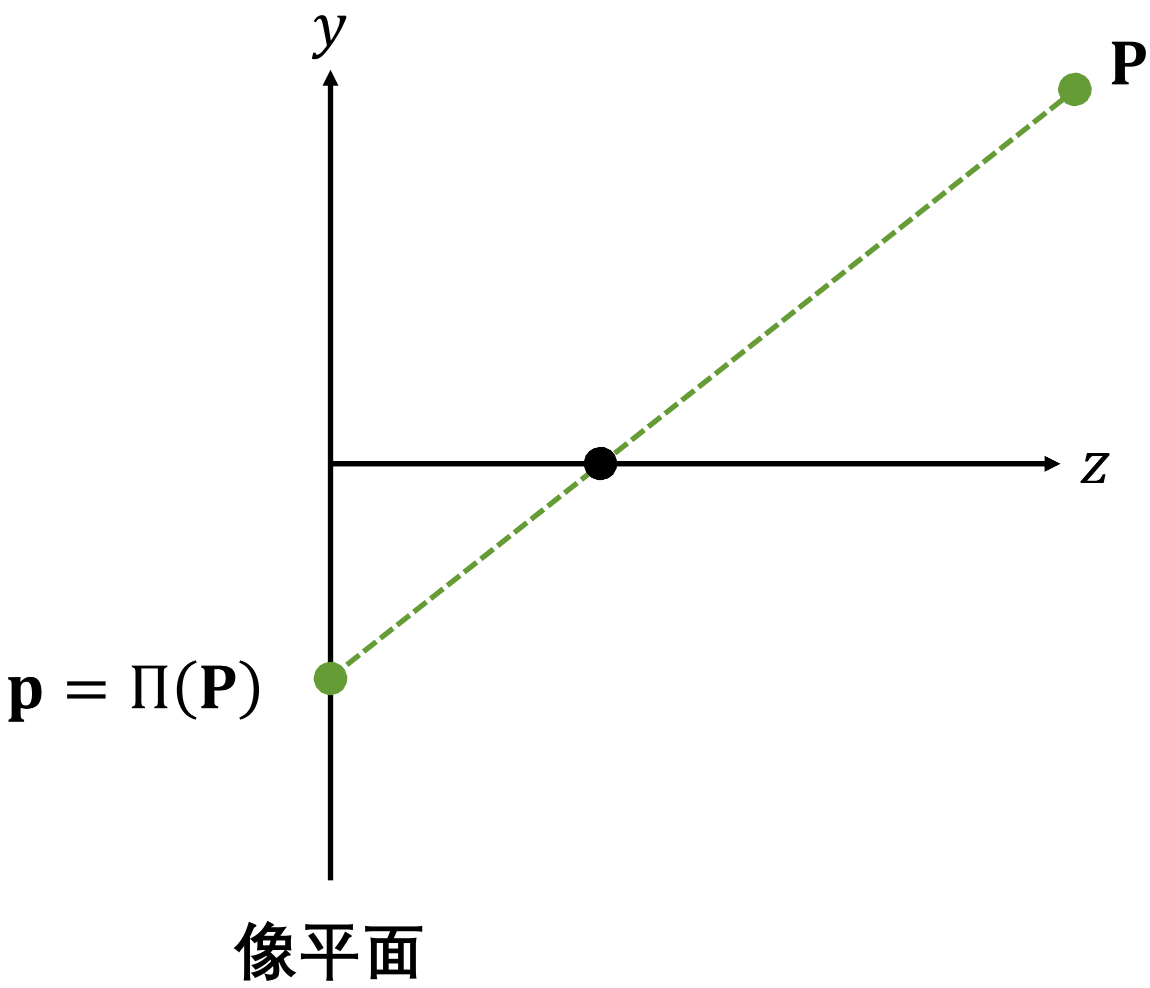
光流



光流



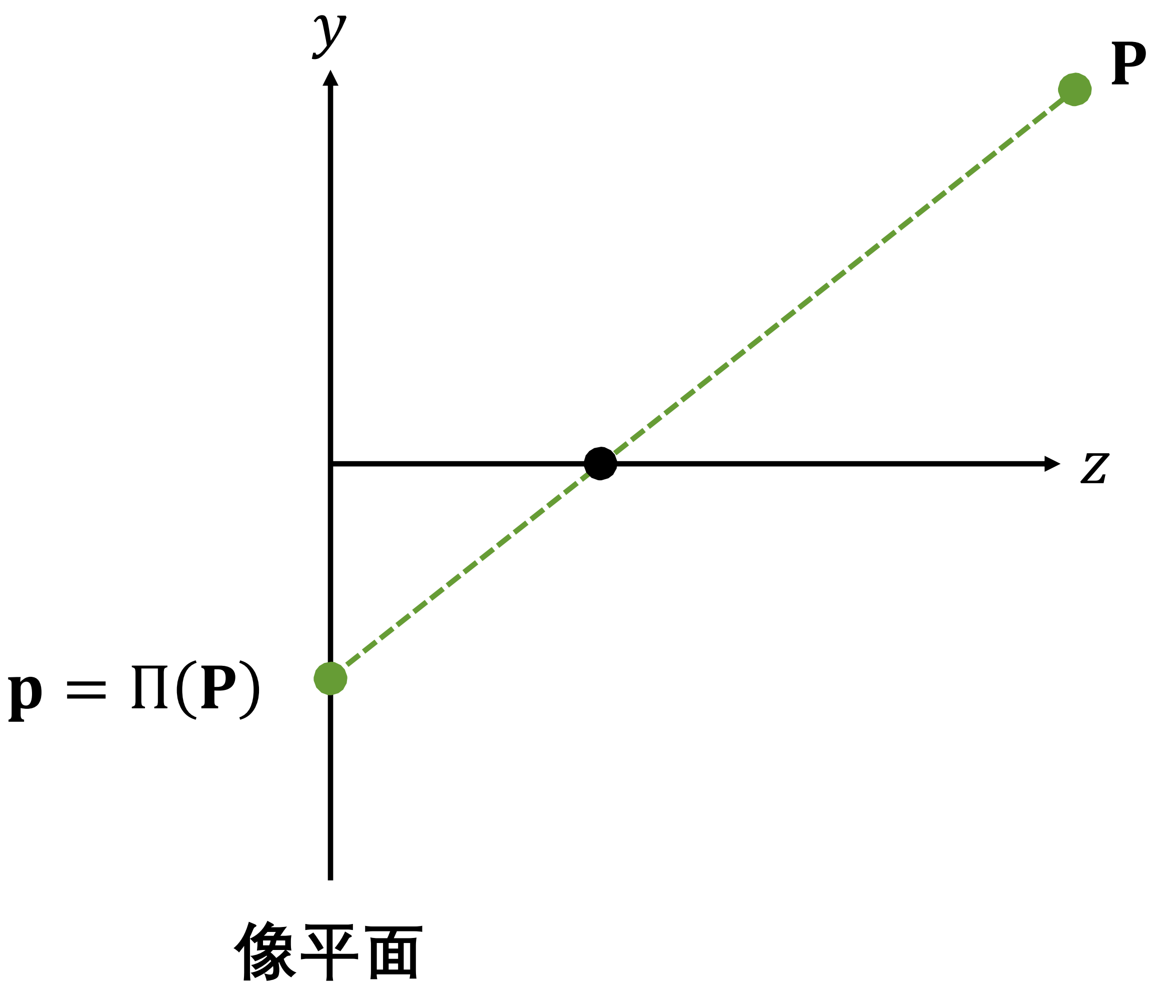

y
P
z
像平面



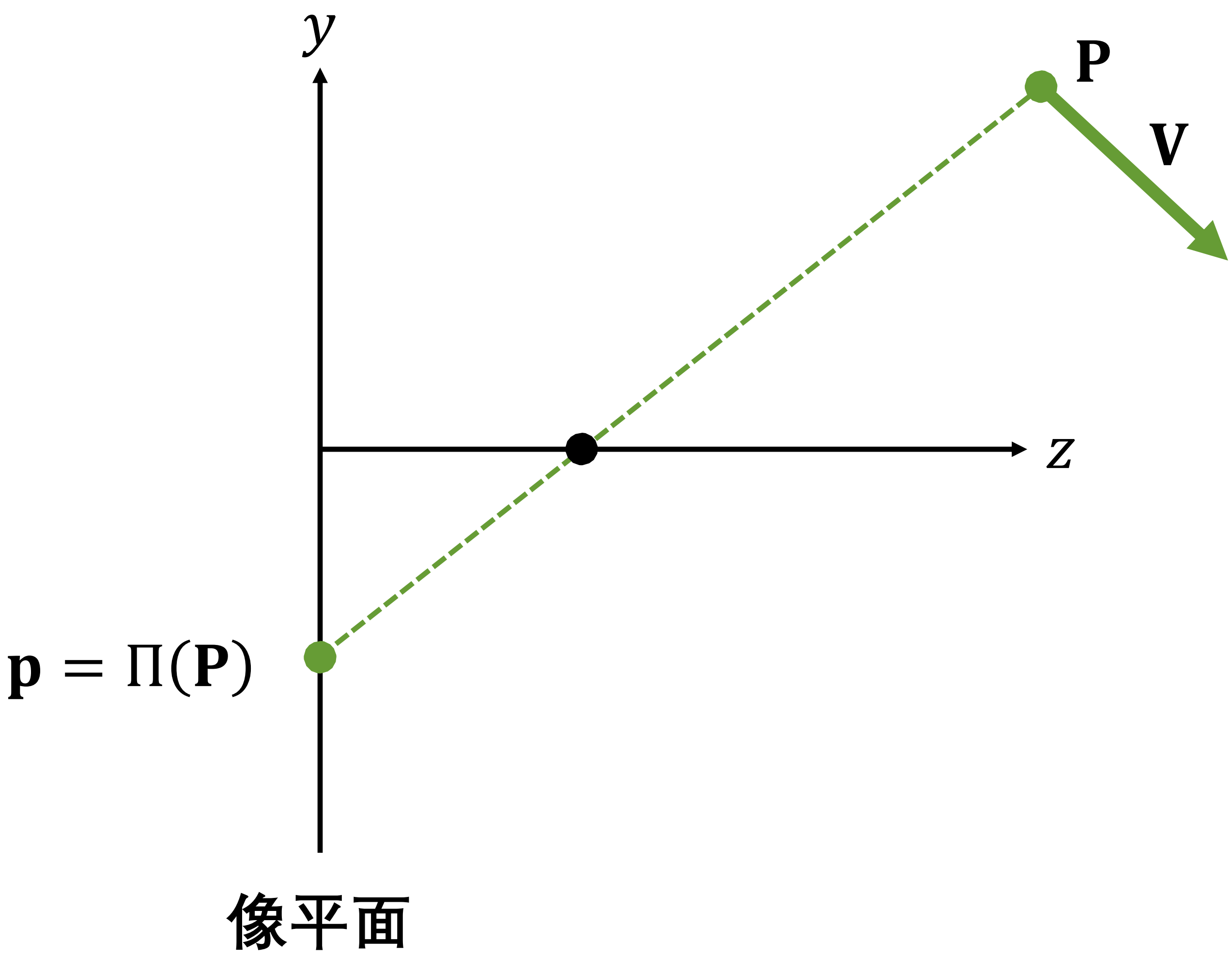
y
P
z
p = Π(P)
像平面

运动场

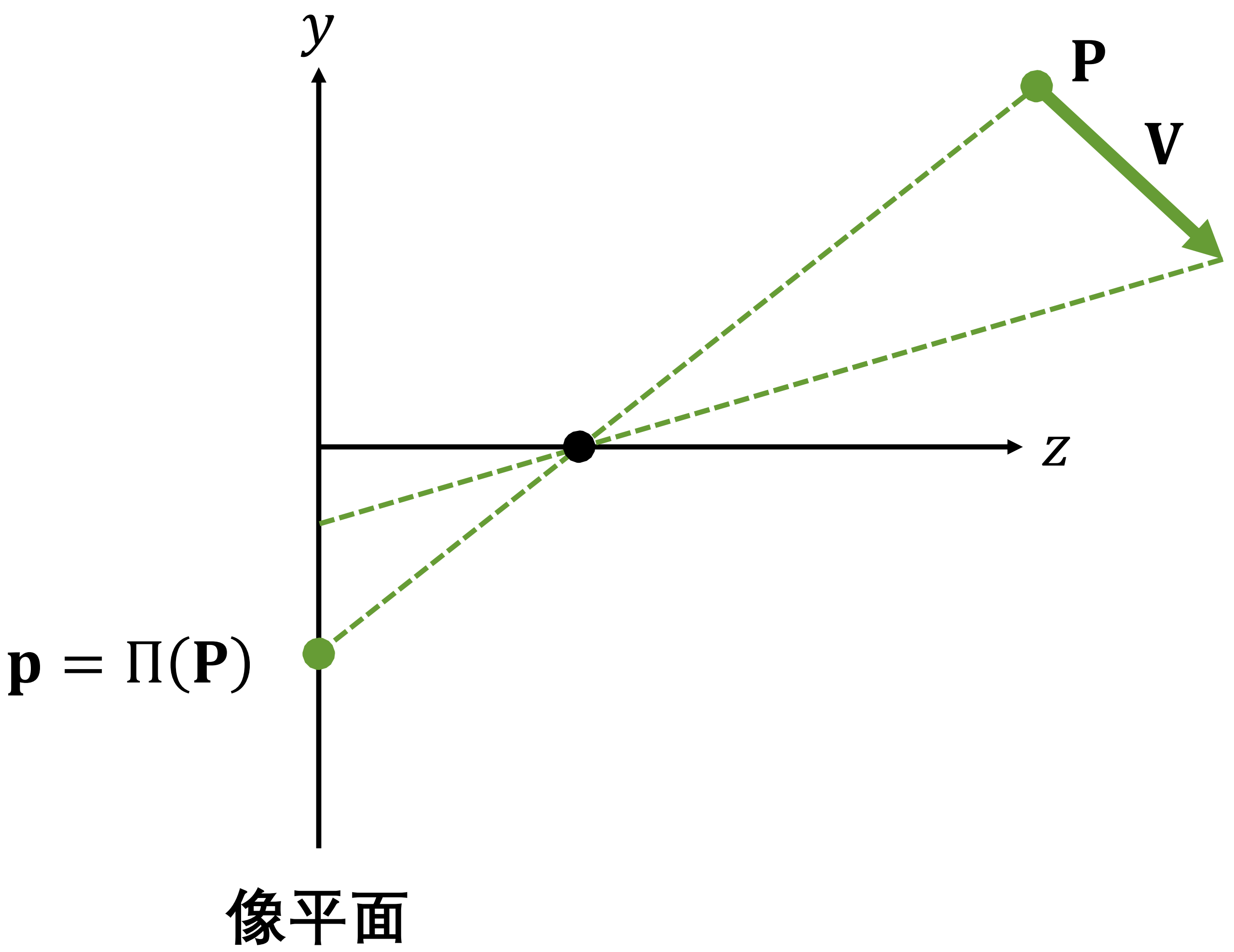
3D场景速度在像平面上的投影



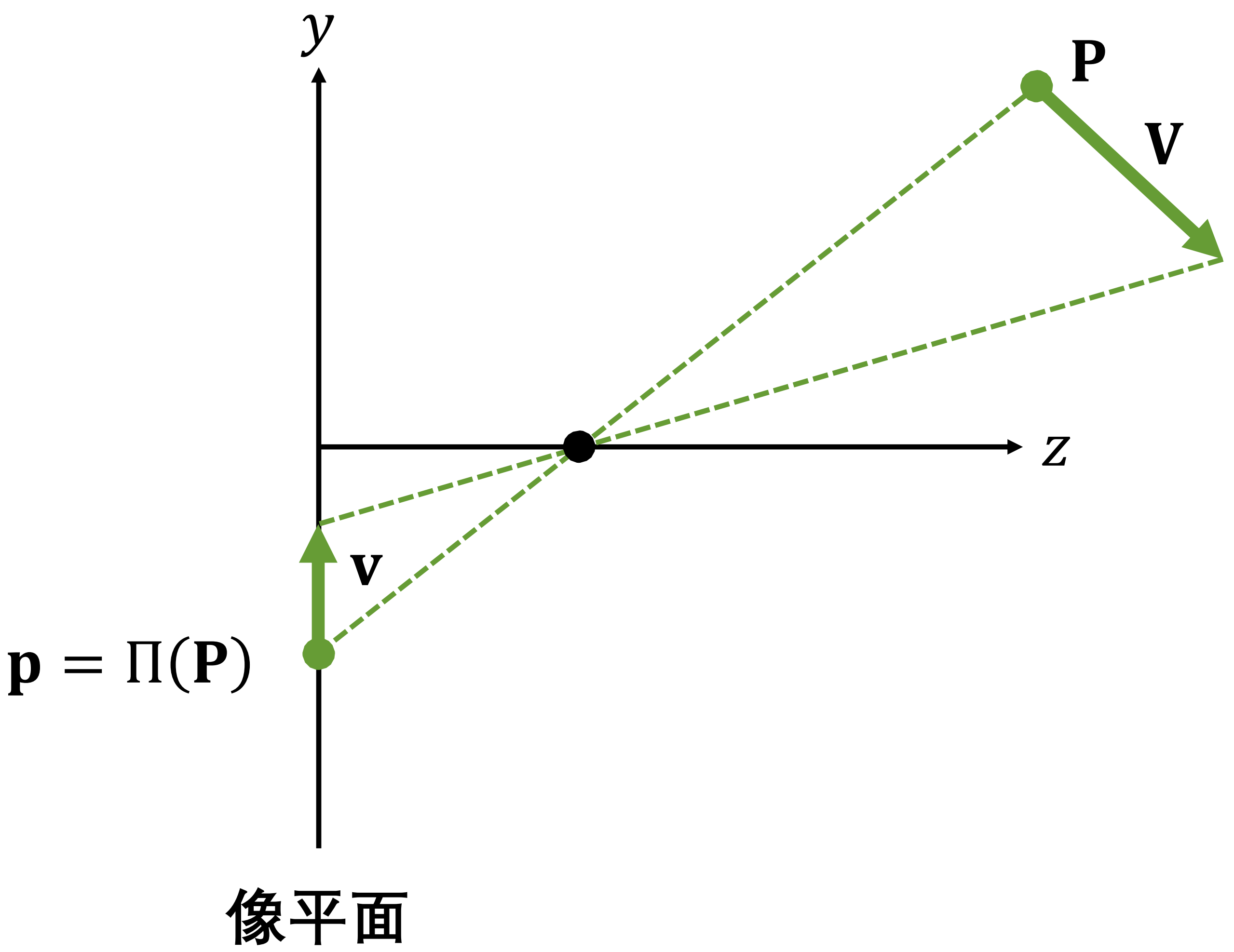
y
P
z
p = Π(P)
像平面



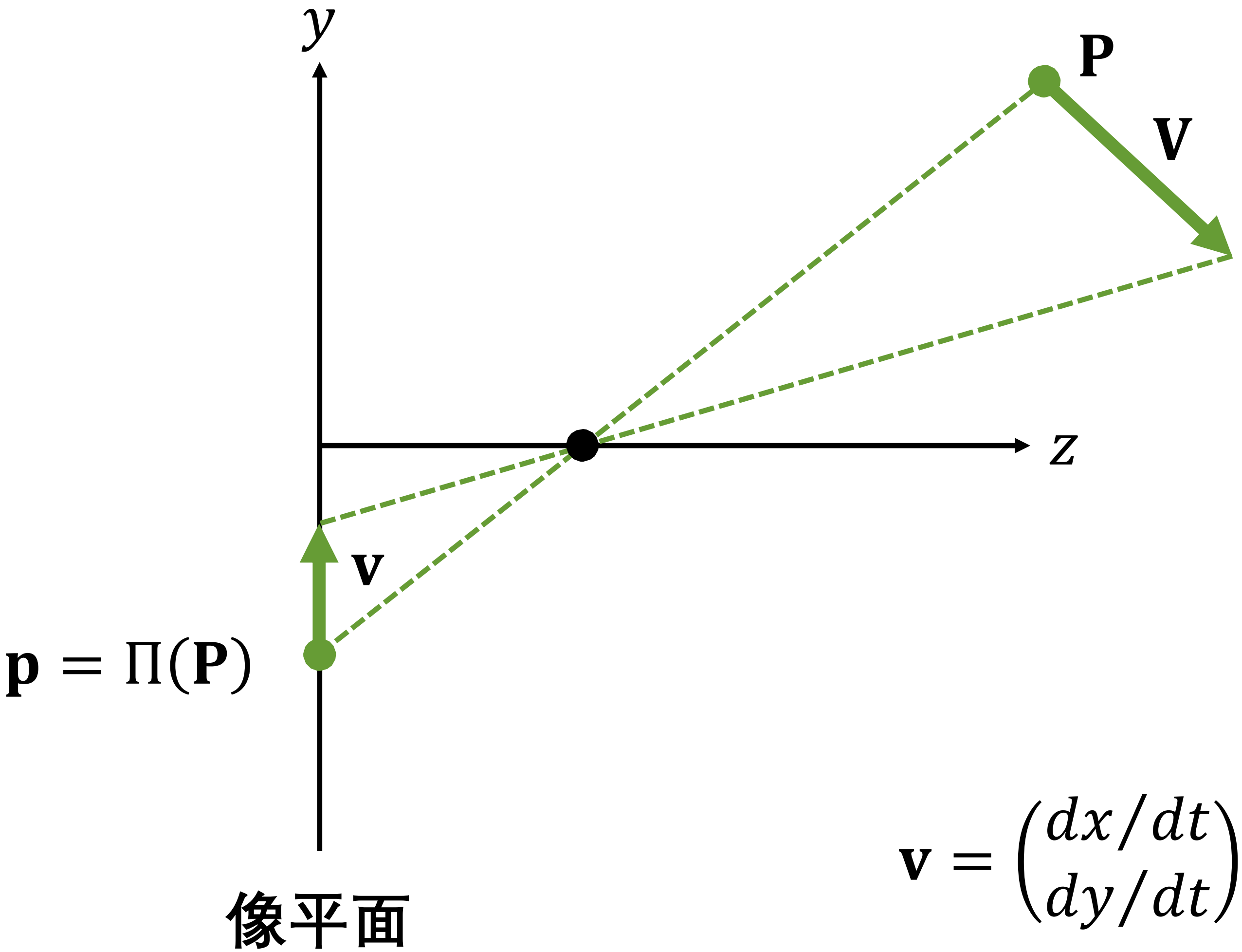
y
P
V
z
p = Π(P)
像平面



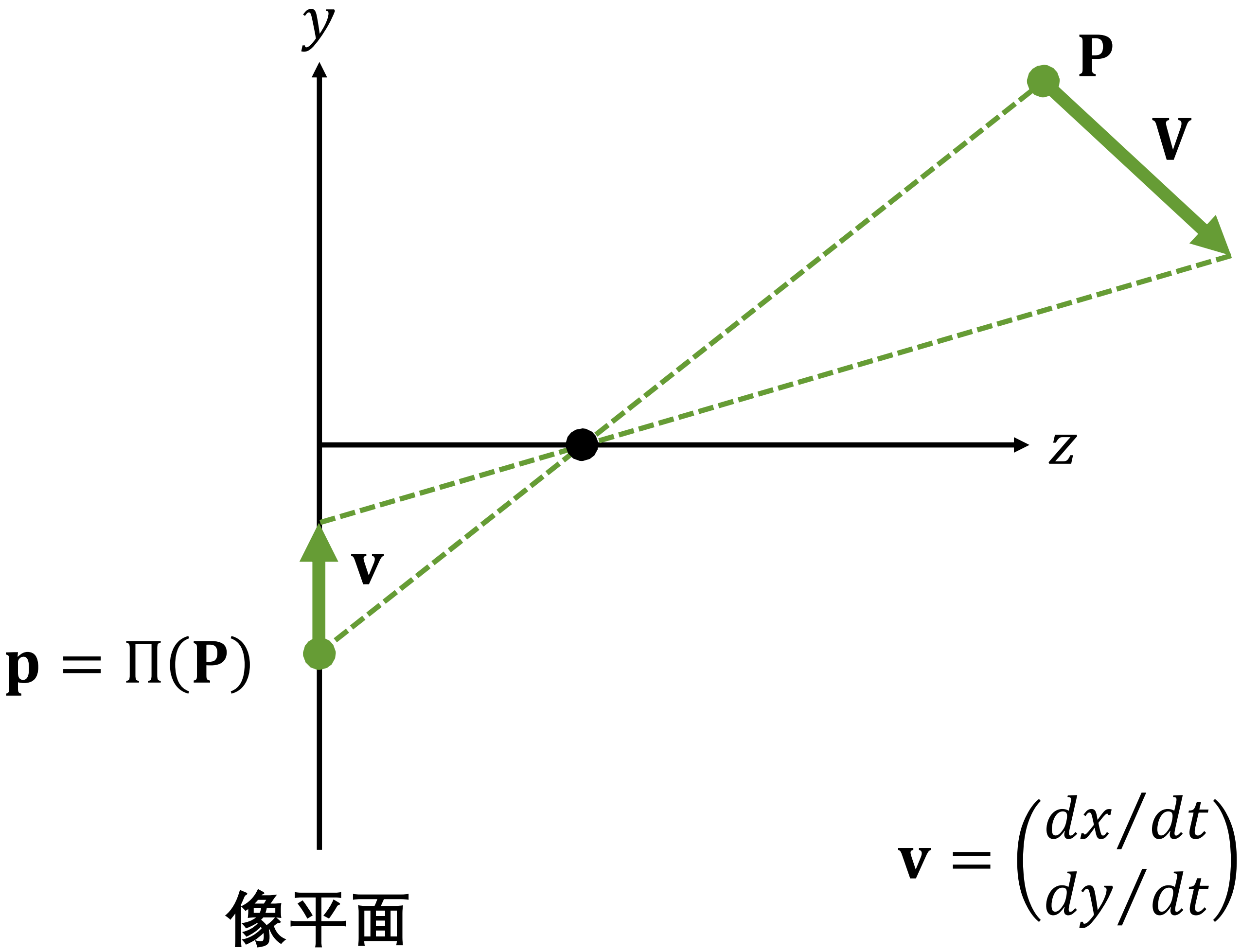
y
P
V
z
p = Π(P)
像平面



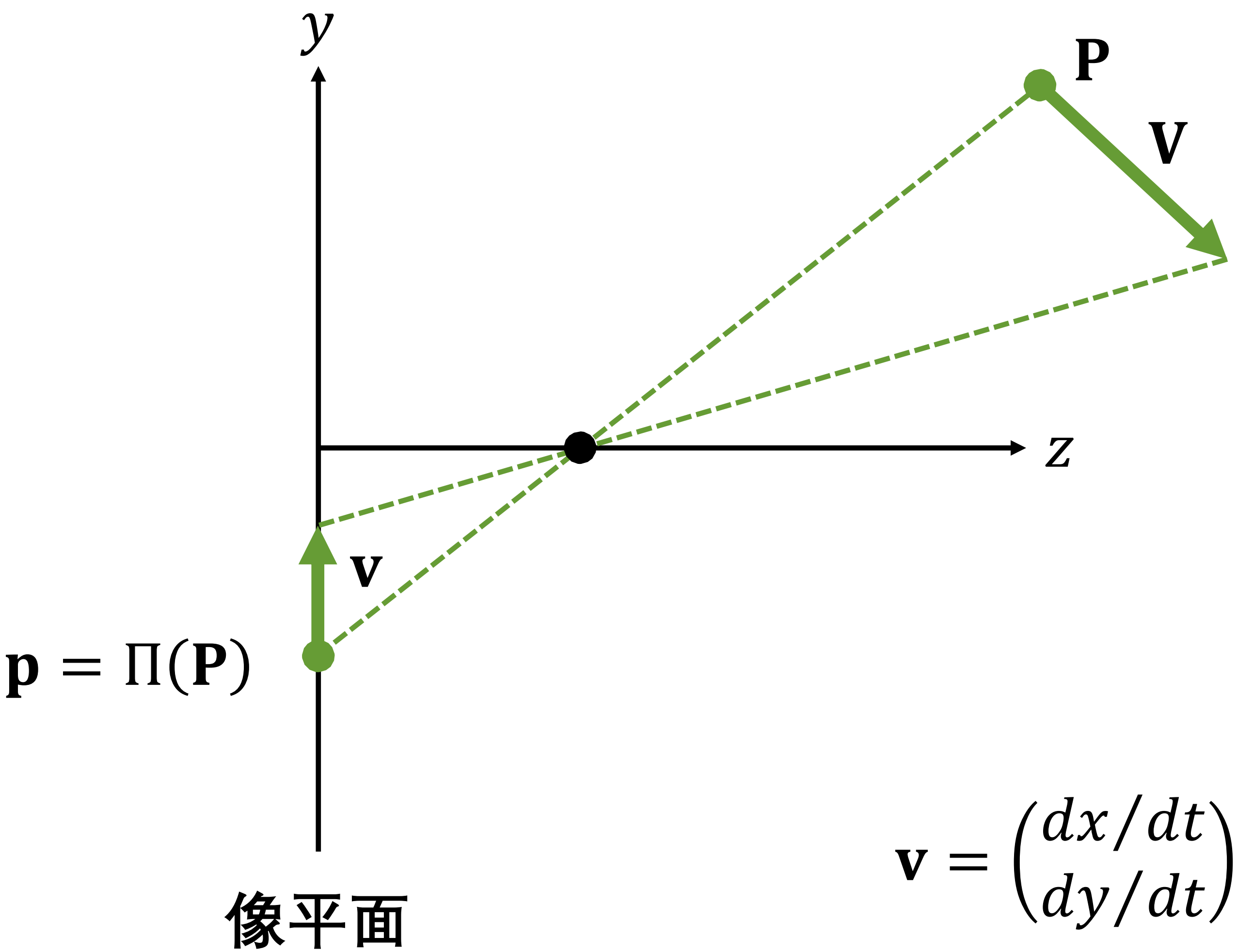
y
$\mathbf{P}$
$\mathbf{V}$
z
$\mathbf{v}$
$\mathbf{p} = \Pi(\mathbf{P})$
像平面



y
$\mathbf{P}$
$\mathbf{V}$
z
$\mathbf{v}$
$\mathbf{p} = \Pi(\mathbf{P})$
$\mathbf{v} = \begin{pmatrix} dx/dt \\ dy/dt \end{pmatrix}$
像平面



$$\mathbf{v} = \begin{pmatrix} dx/dt \\ dy/dt \end{pmatrix} = \frac{d\mathbf{p}}{dt}$$



$$\mathbf{v} = \begin{pmatrix} dx/dt \\ dy/dt \end{pmatrix} = \frac{d\mathbf{p}}{dt}$$

$$= \frac{d\Pi(\mathbf{P})}{dt}$$

光流

图像中亮度图案的表观运动速度的分布

光流

图像中亮度图案的表观运动速度的分布

由图像测量得出的运动

我们**想要**运动场

我们想要运动场

几何概念

我们计算光流

我们计算光流

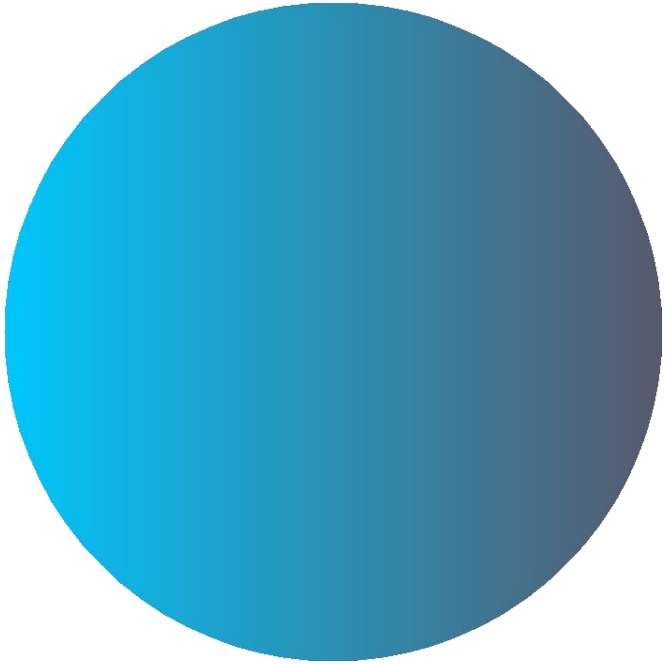
光度概念

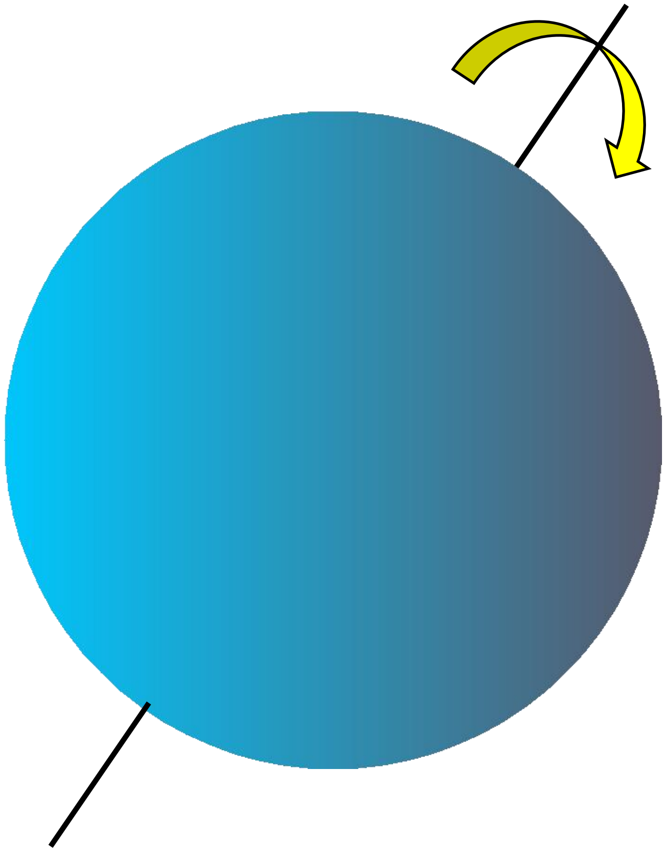
理想的光流和运动场是等价的

运动场

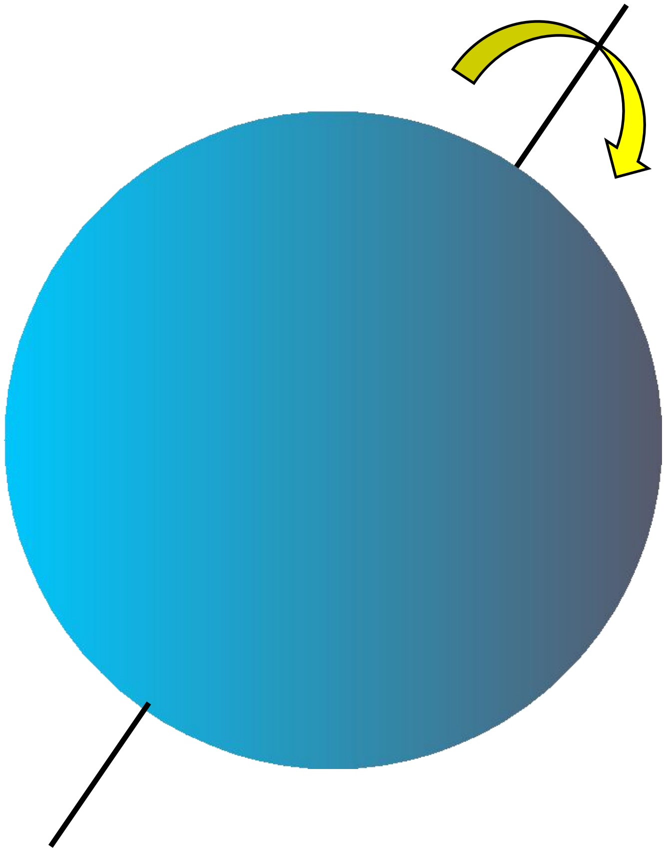
$\neq$

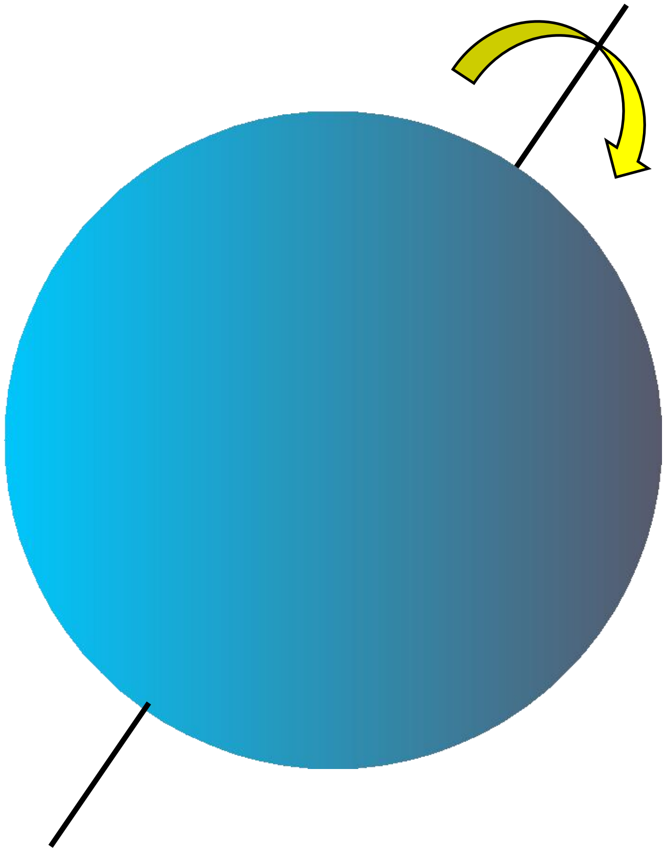
光流



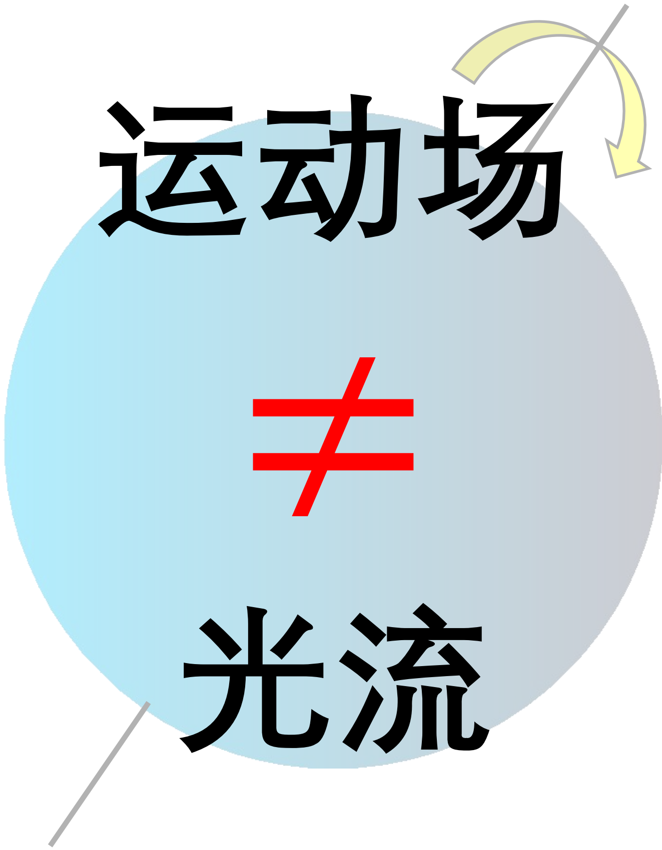


运动场是什么?

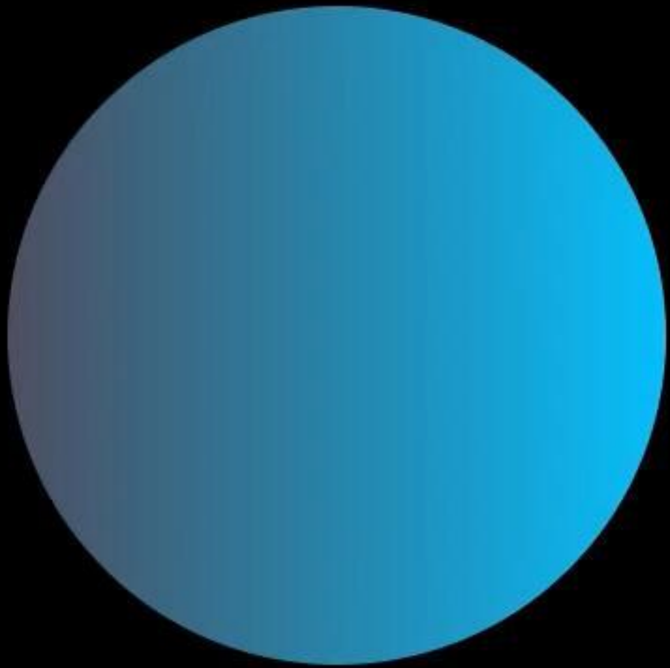


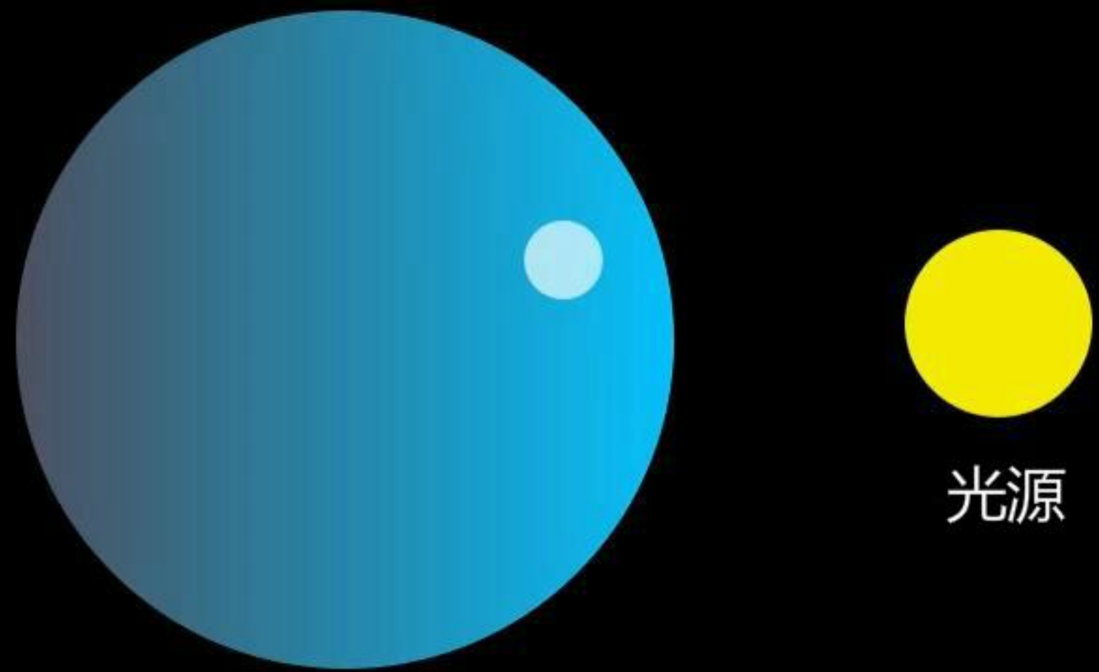


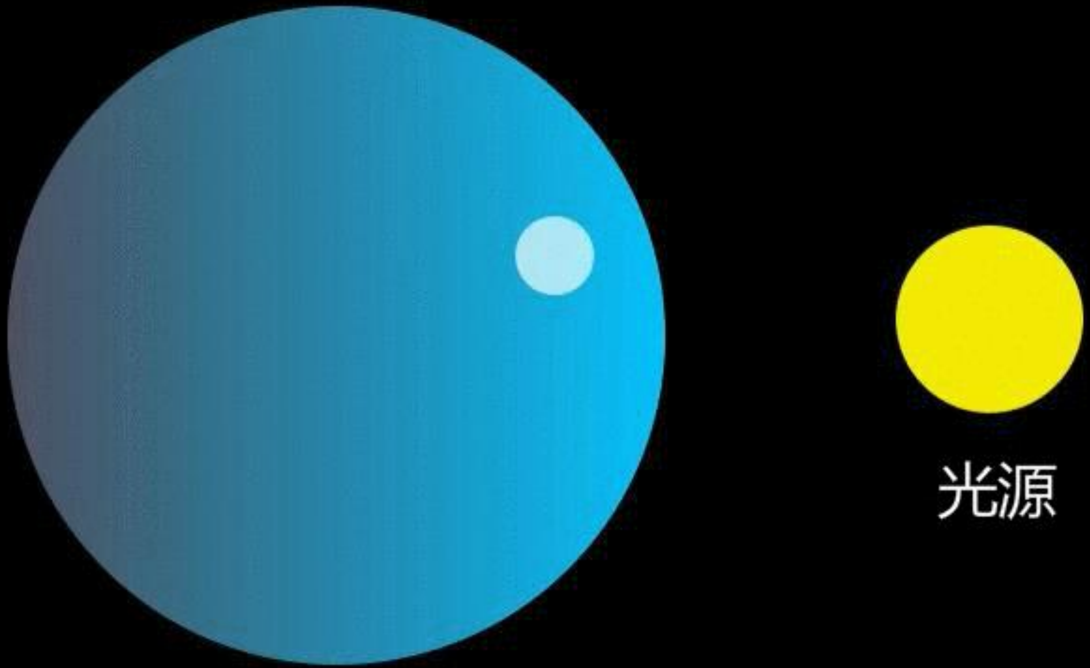
光流场是什么？


运动场
≠
光流

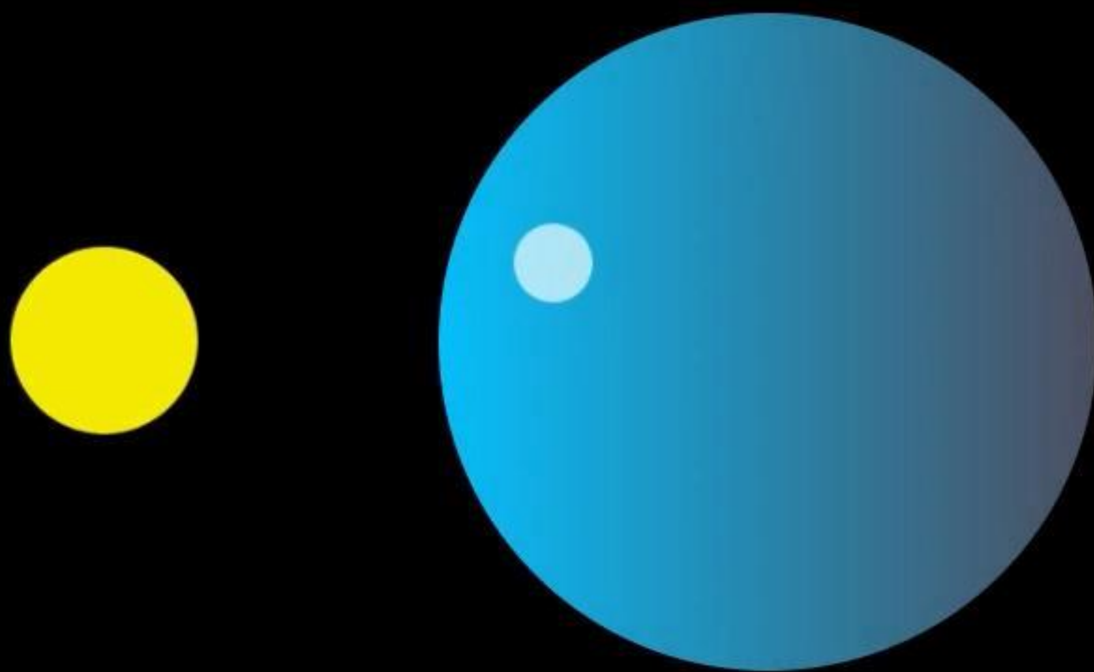
鸣谢： GoPro

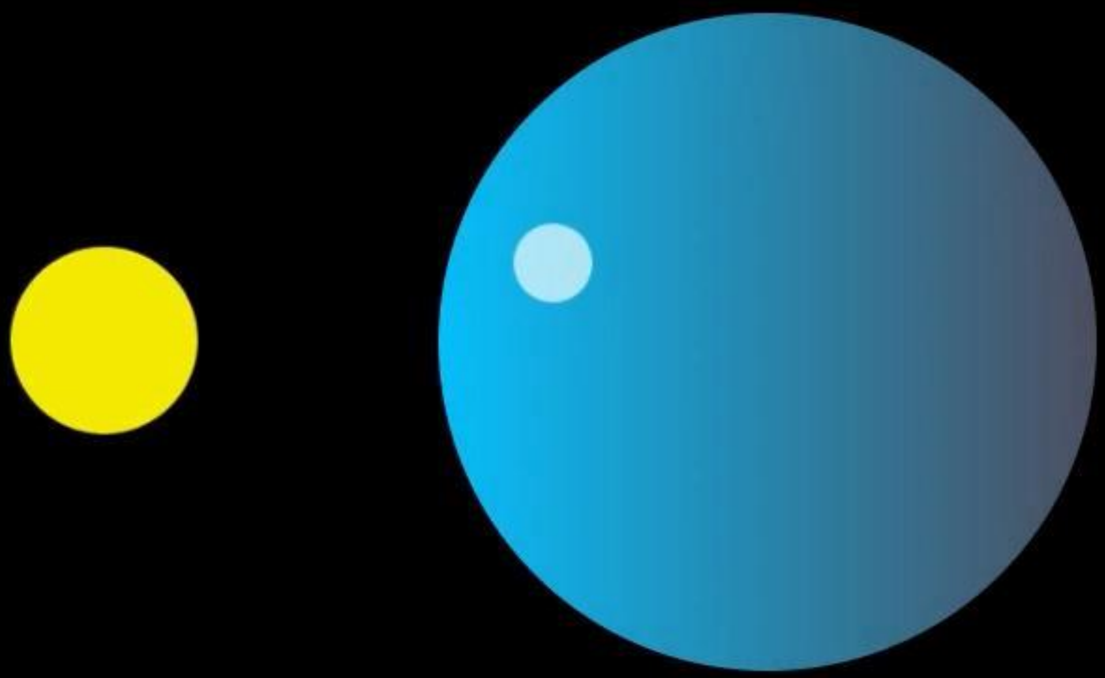



光源

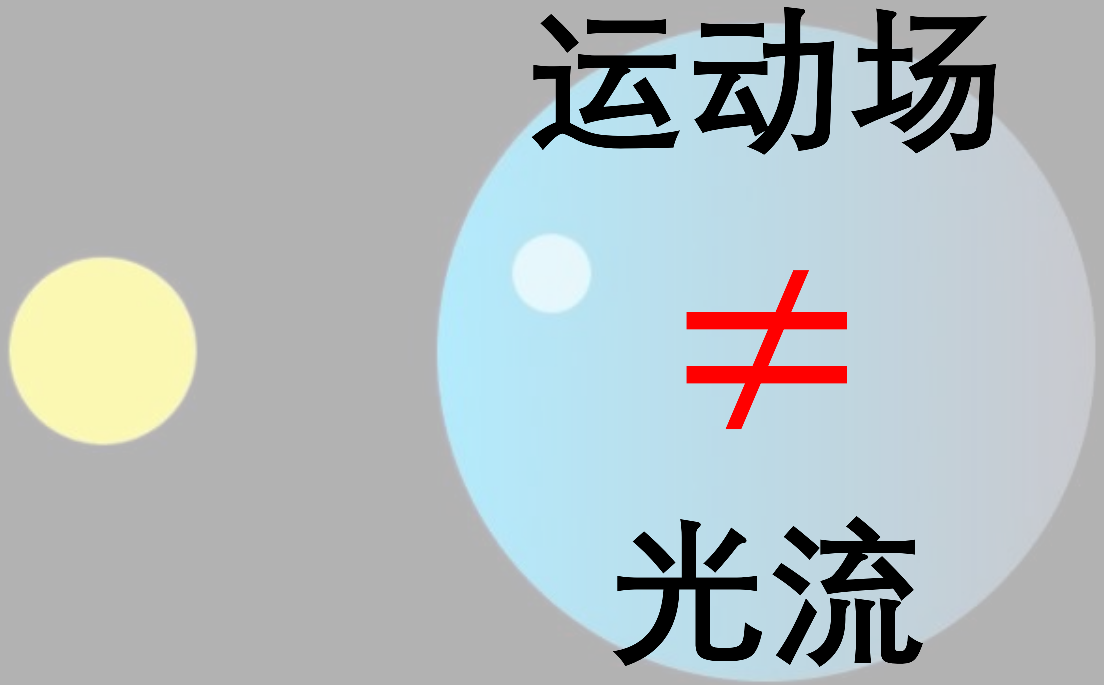

光源

运动场是什么?





光流场是什么？


运动场
≠
光流

光流为运动场提供了一个良好的近似

假设

假设

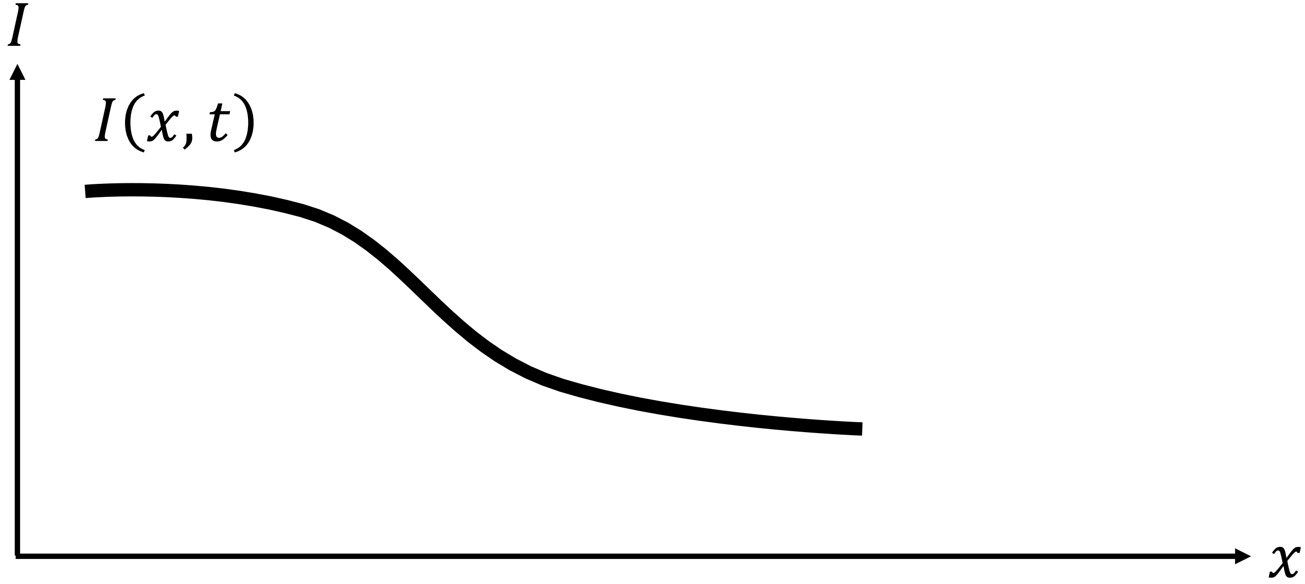
图像点的亮度随时间保持恒定

假设

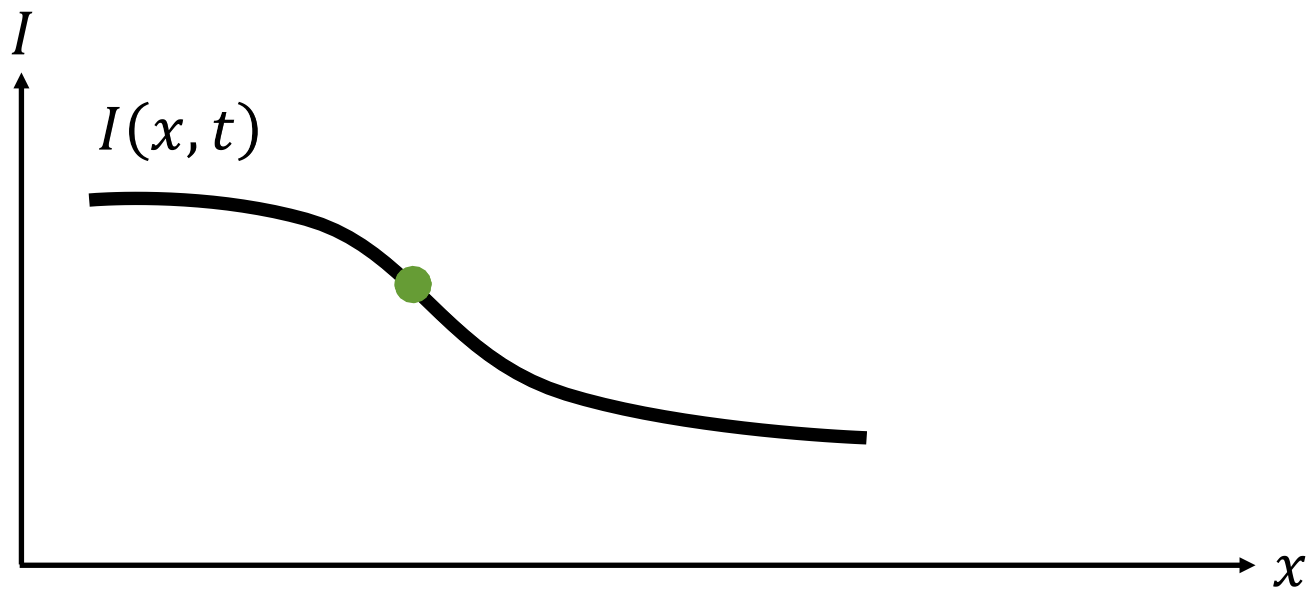
图像点的亮度随时间保持恒定

估计其他图像属性成为可能

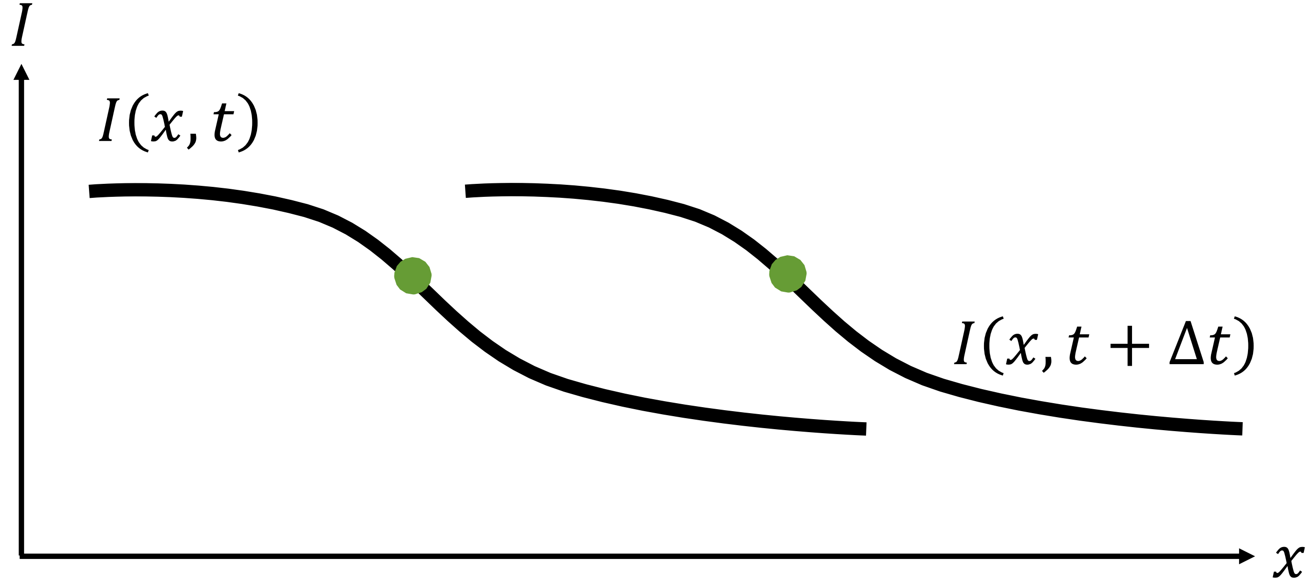
亮度
恒常性



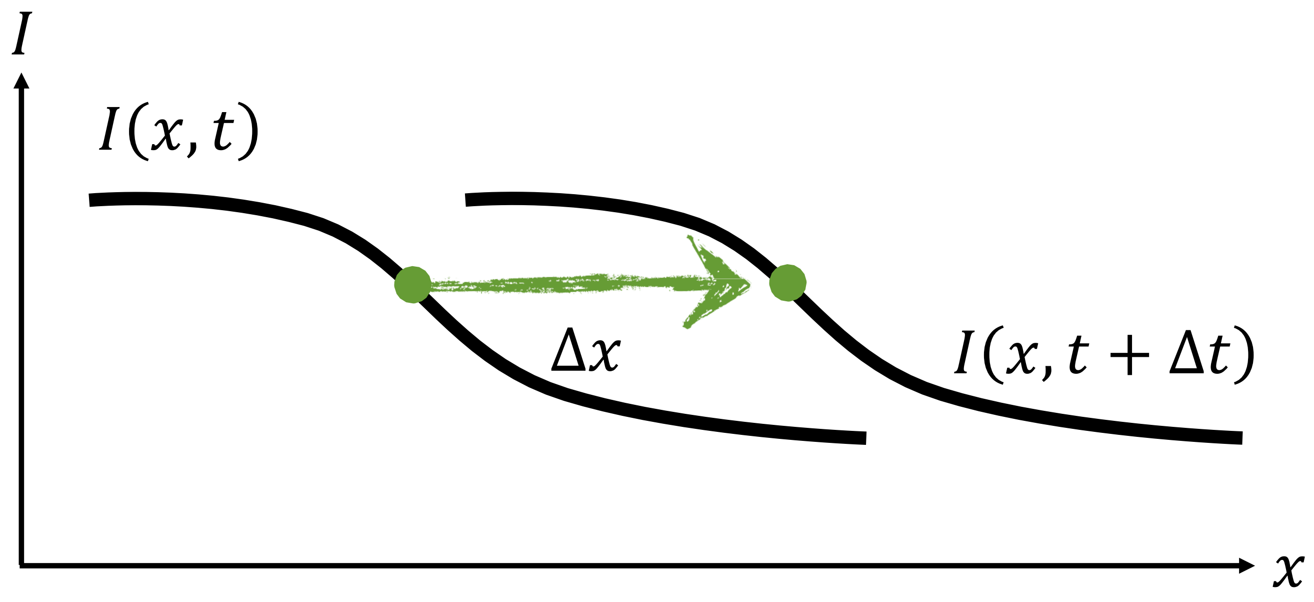
亮度恒常性



亮度
恒常性



亮度恒常性



亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

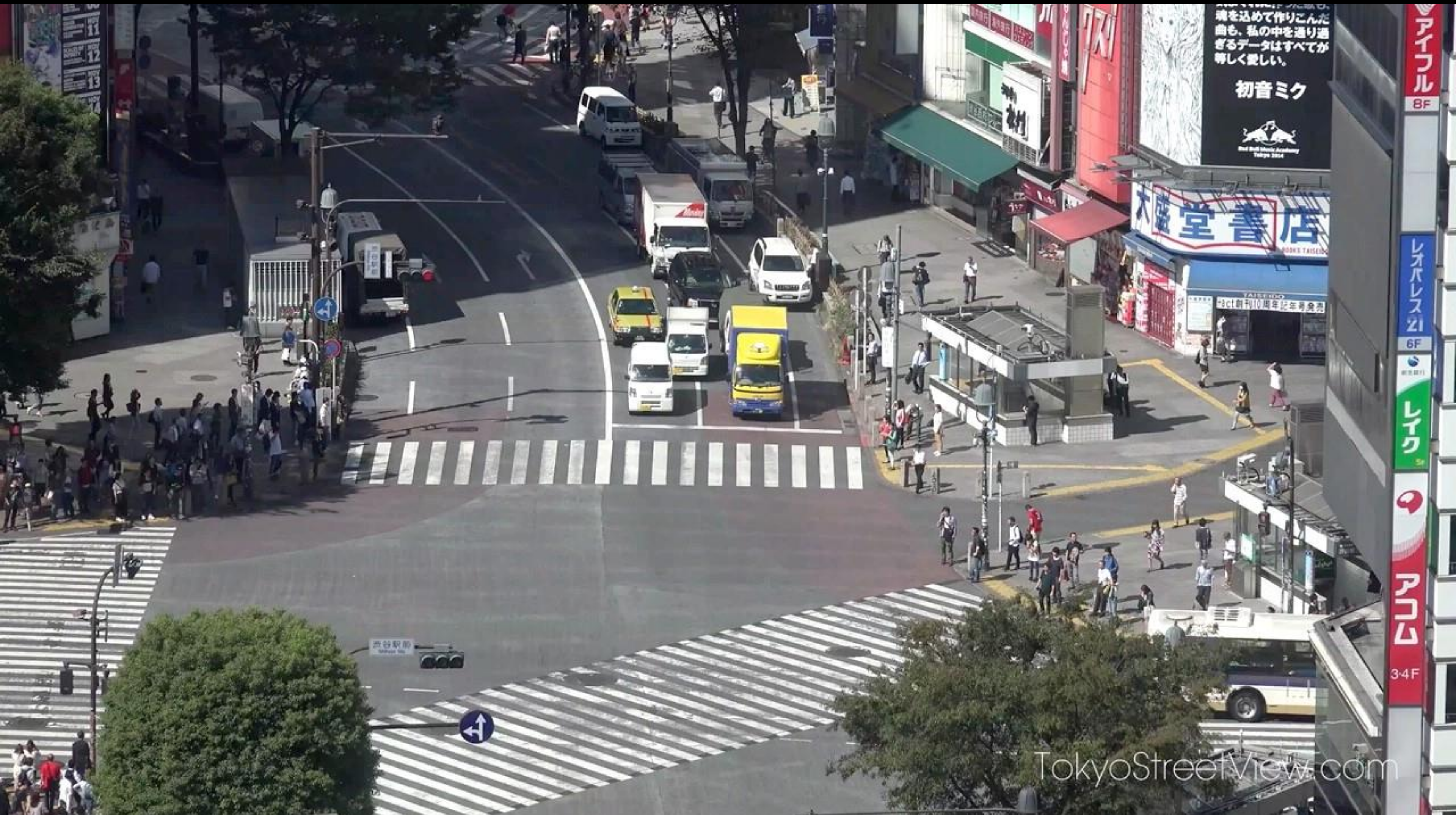
亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

假设时间间隔很短

亮度恒常假设

$$I(x,y,t) = I(x+u\Delta t, y+v\Delta t, t+\Delta t)$$


魂を込めて作りこんだ
曲も、私の中を通り過
ぎるデータはすべてが
等しく愛しい。
初音ミク
大盛堂書店
アイフル
8F
レオパレス21
6F
レイク
アコム
3·4F
TokyoStreetView.com

魂を込めて作りこんだ
曲も、私の中を通り過
ぎるデータはすべてが
等しく愛しい。
初音ミク
大盛堂書店
アイフル
8F
レオパレス21
6F
レイク
アコム
3-4F
TokyoStreetView.com

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

由于图像的流是非线性的，所以流很难恢复

回顾： 泰勒级数

$$f(x + \Delta x, y + \Delta y, z + \Delta z)$$

$$f(x+\Delta x, y+\Delta y, z+\Delta z) = f(x,y,z)$$

$$+f_x(x,y,z)\Delta x + f_y(x,y,z)\Delta y + f_z(x,y,z)\Delta z$$

$$+\mathrm{h.o.t.}$$

回顾：泰勒级数

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

由于图像的流是非线性的，所以流很难恢复

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

泰勒级数

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

泰勒级数

$$I(x, y, t) = I(x, y, t) + \frac{\partial I}{\partial x} u\Delta t + \frac{\partial I}{\partial y} v\Delta t + \frac{\partial I}{\partial t} \Delta t + \mathrm{h.o.t.}$$

亮度恒常假设

$$I(x,y,t) = I(x+u\Delta t, y+v\Delta t, t+\Delta t)$$

泰勒级数

$$I(x,y,t) = I(x,y,t) + \frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.}$$

化简

亮度恒常假设

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

泰勒级数

$$\cancel{I(x, y, t)} = \cancel{I(x, y, t)} + \frac{\partial I}{\partial x} u\Delta t + \frac{\partial I}{\partial y} v\Delta t + \frac{\partial I}{\partial t} \Delta t + \mathrm{h.o.t.}$$

化简

$$I(x,y,t) = I(x+u\Delta t, y+v\Delta t, t+\Delta t)$$

泰勒级数

$$\cancel{I(x,y,t)} = \cancel{I(x,y,t)} + \frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.}$$

化简

$$\frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.} = 0$$

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t)$$

泰勒级数

$$\cancel{I(x, y, t)} = \cancel{I(x, y, t)} + \frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.}$$

化简

$$\frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.} = 0$$

除以Δt

$$I(x,y,t) = I(x+u\Delta t, y+v\Delta t, t+\Delta t)$$

泰勒级数

$$\xcancel{I(x,y,t)} = \xcancel{I(x,y,t)} + \frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.}$$

化简

$$\frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.} = 0$$

除以Δt

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} + \mathrm{h.o.t.} = 0$$

$$\cancel{I(x,y,t)} = \cancel{I(x,y,t)} + \frac{\partial I}{\partial x} u \Delta t + \frac{\partial I}{\partial y} v \Delta t + \frac{\partial I}{\partial t} \Delta t + \mathrm{h.o.t.}$$

化简

$$\frac{\partial I}{\partial x} u \Delta t + \frac{\partial I}{\partial y} v \Delta t + \frac{\partial I}{\partial t} \Delta t + \mathrm{h.o.t.} = 0$$

除以Δt

$$\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} + \mathrm{h.o.t.} = 0$$

$$\Delta t \to 0$$

$$\cancel{I(x,y,t)} = \cancel{I(x,y,t)} + \frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.}$$

化简

$$\frac{\partial I}{\partial x}u\Delta t + \frac{\partial I}{\partial y}v\Delta t + \frac{\partial I}{\partial t}\Delta t + \mathrm{h.o.t.} = 0$$

除以Δt

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} + \mathrm{h.o.t.} = 0$$

$\Delta t \to 0$

亮度恒常约束： $I_x u + I_y v + I_t = 0$

亮度恒常约束： $I_x u + I_y v + I_t = 0$

亮度恒常约束： $I_x u + I_y v + I_t = 0$

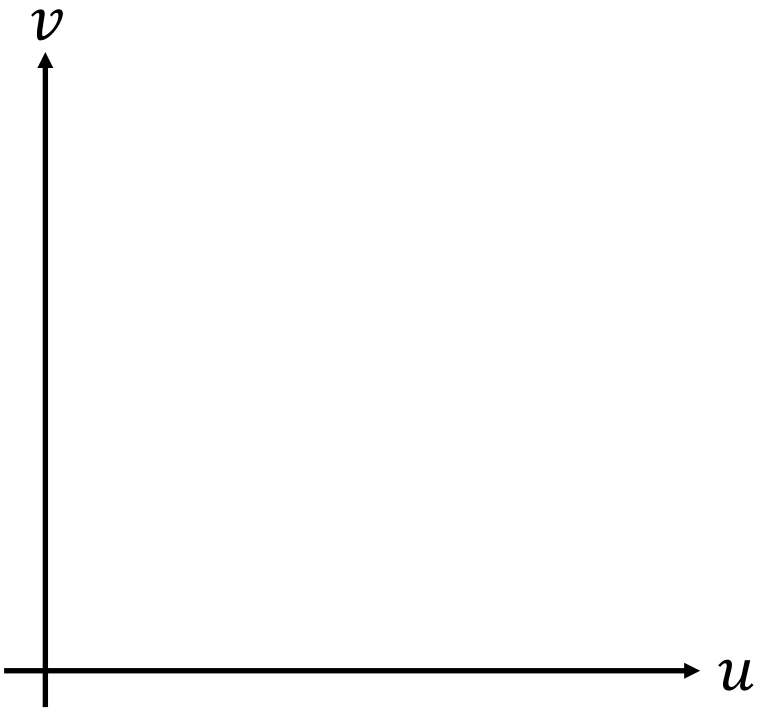
如何恢复速度？

亮度恒常约束：$I_x u + I_y v + I_t = 0$

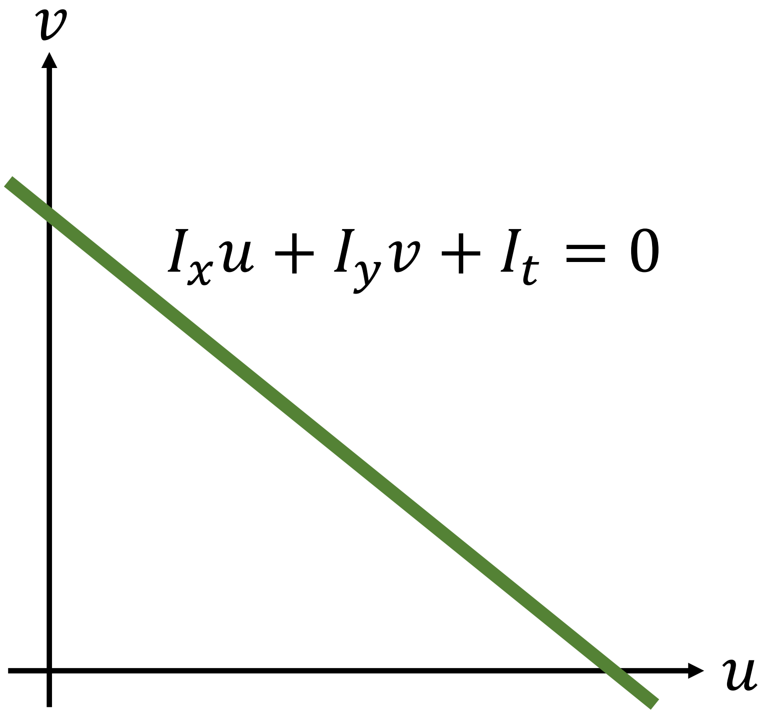
如何恢复速度?

解是欠约束的

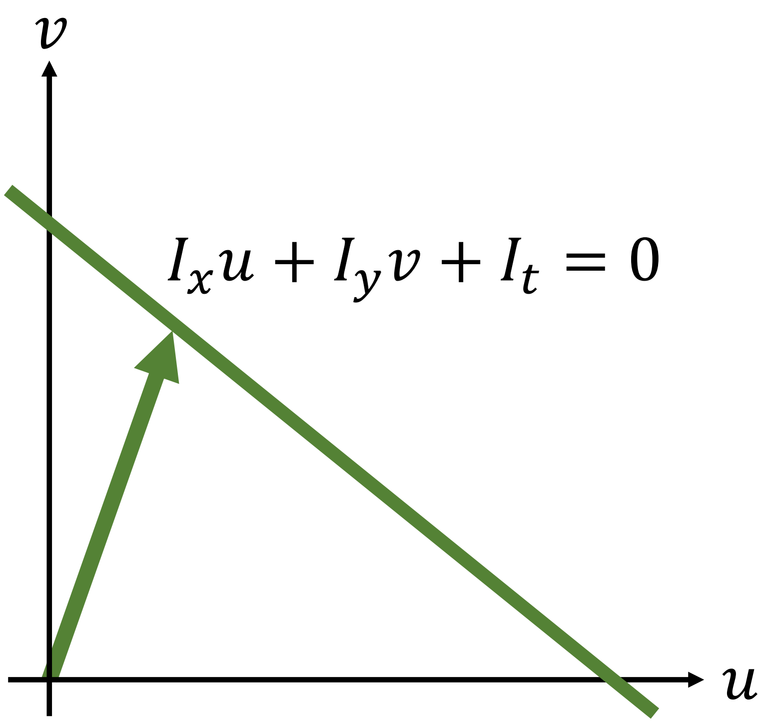
亮度恒常约束：$I_x u + I_y v + I_t = 0$



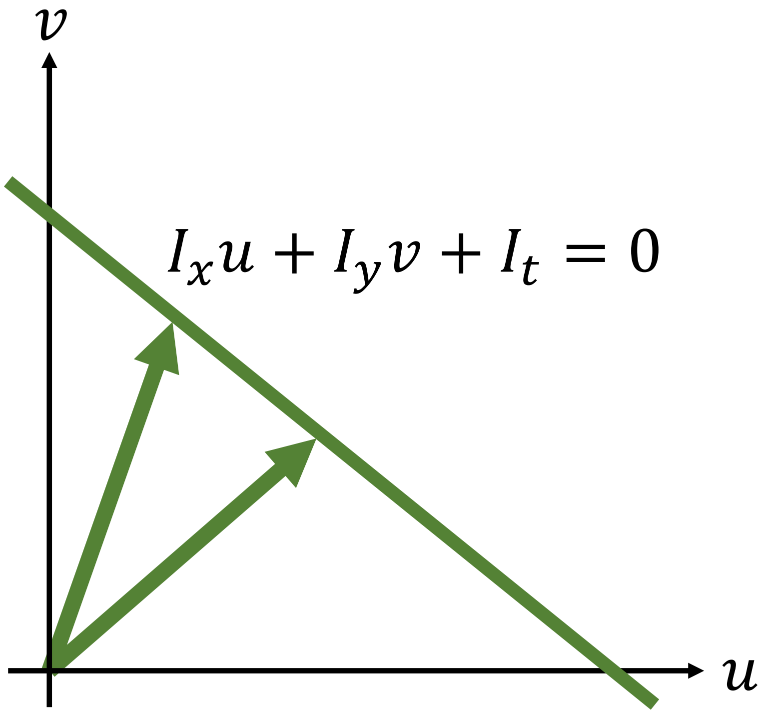
亮度恒常约束： $I_x u + I_y v + I_t = 0$



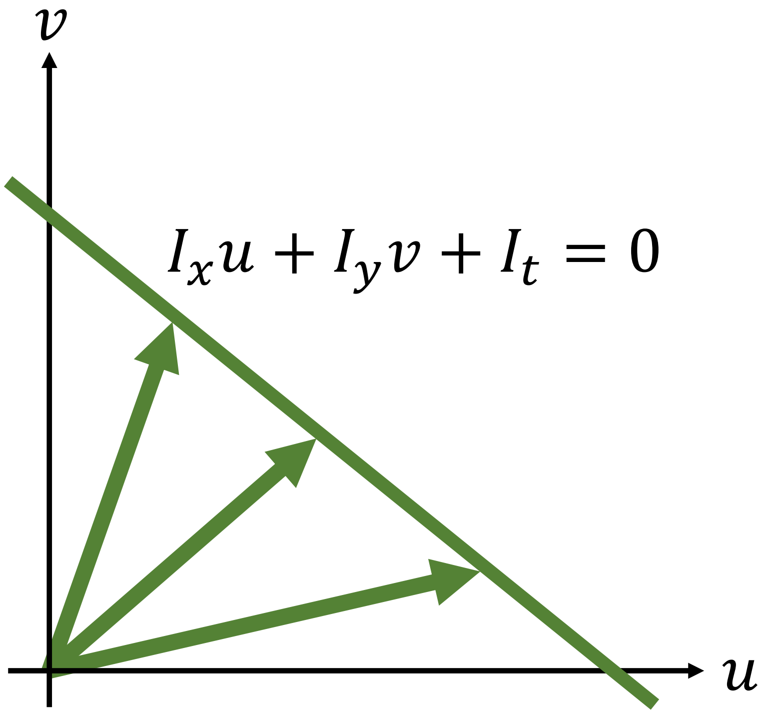
亮度恒常约束： $I_x u + I_y v + I_t = 0$



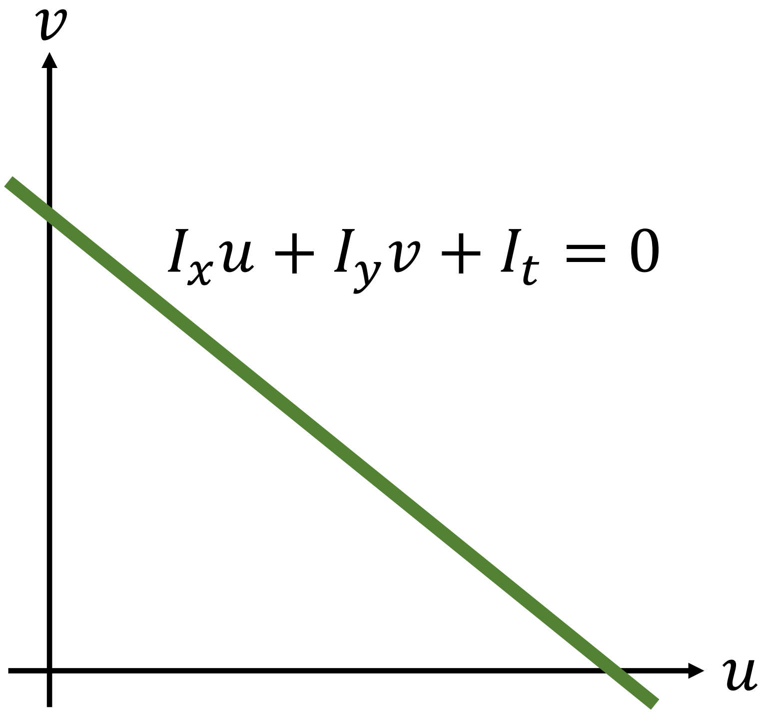
亮度恒常约束： $I_x u + I_y v + I_t = 0$



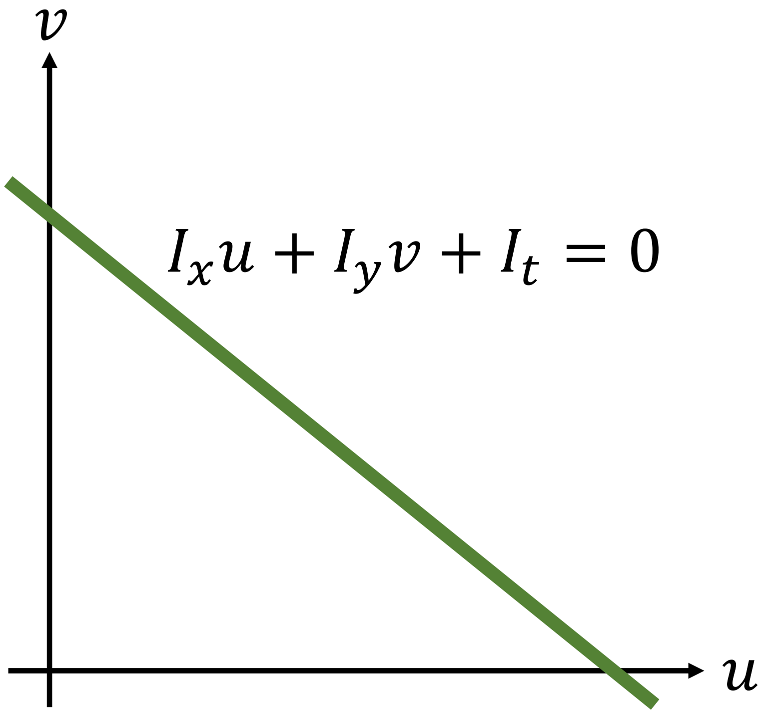
亮度恒常约束：$I_x u + I_y v + I_t = 0$



亮度恒常约束：$I_x u + I_y v + I_t = 0$

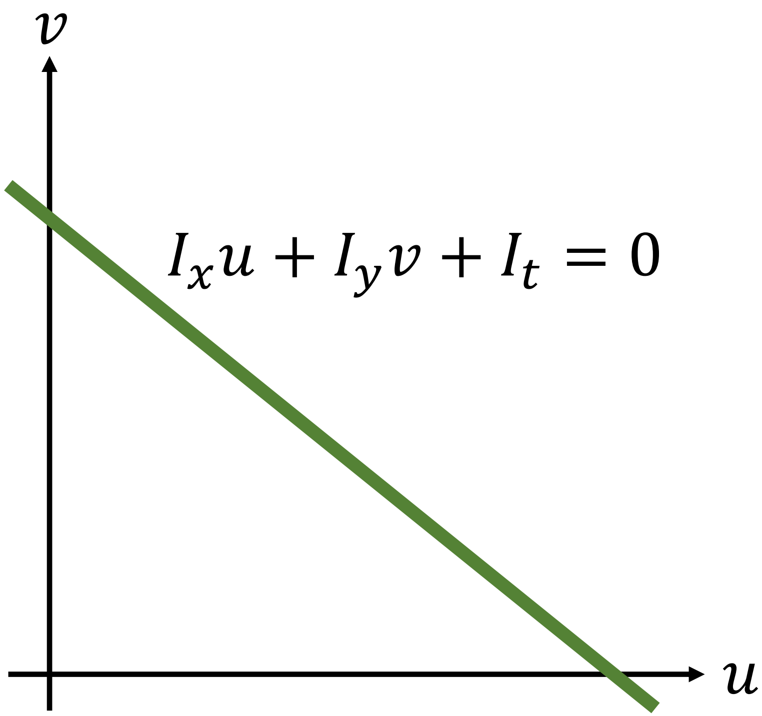


亮度恒常约束：$I_x u + I_y v + I_t = 0$



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

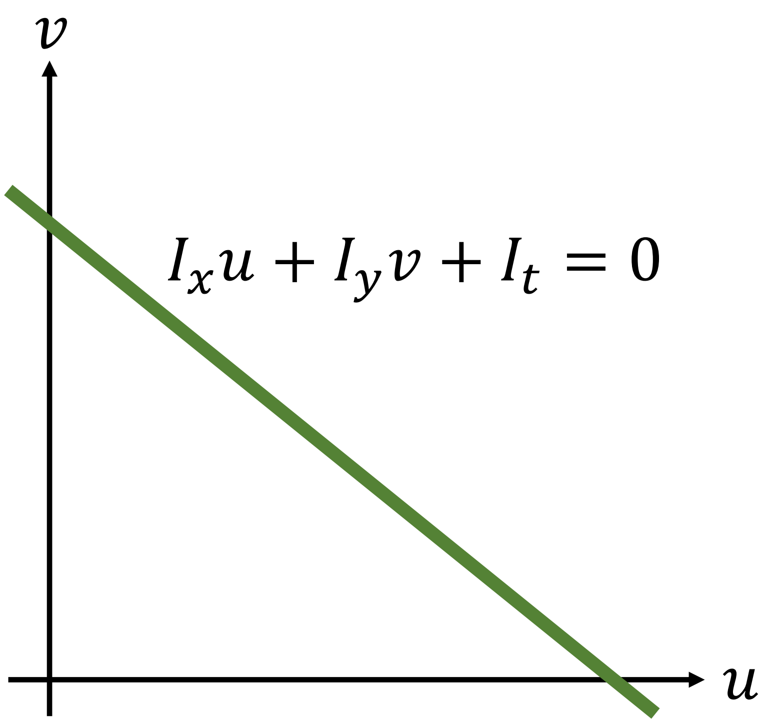
亮度恒常约束：$I_x u + I_y v + I_t = 0$



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

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

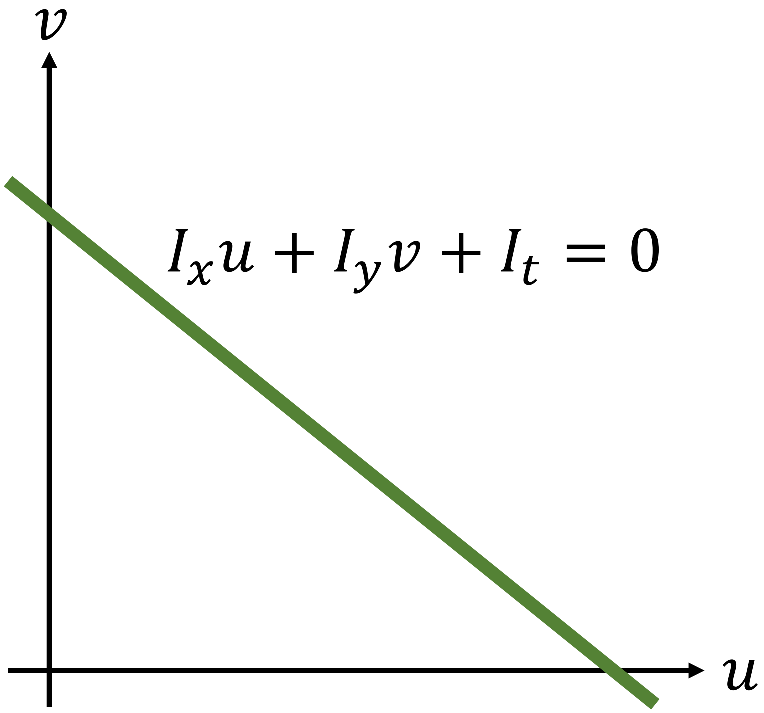
亮度恒常约束：$I_x u + I_y v + I_t = 0$



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

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

亮度恒常约束：$I_x u + I_y v + I_t = 0$

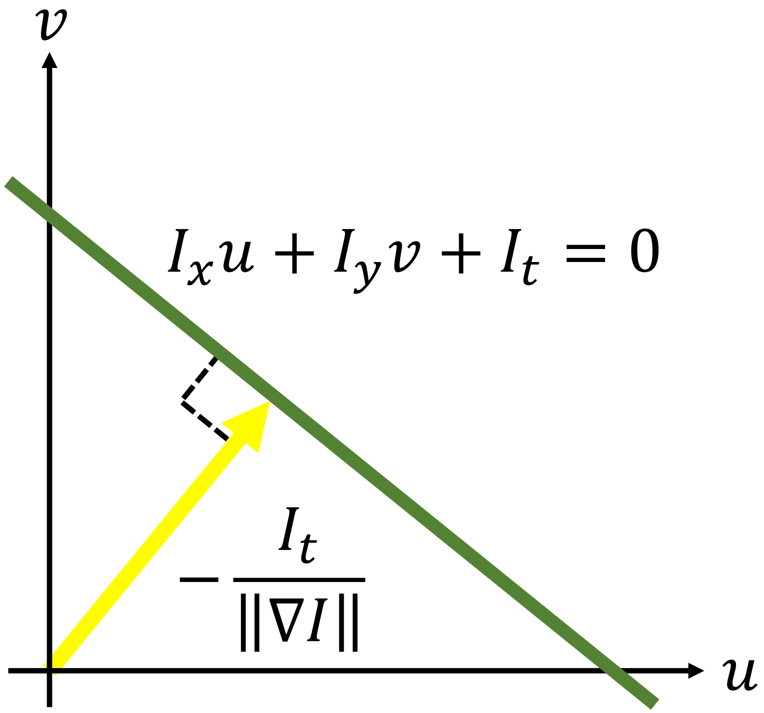


$$\mathbf{u}^{\mathrm{T}} \nabla I = -I_t$$

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

向量投影

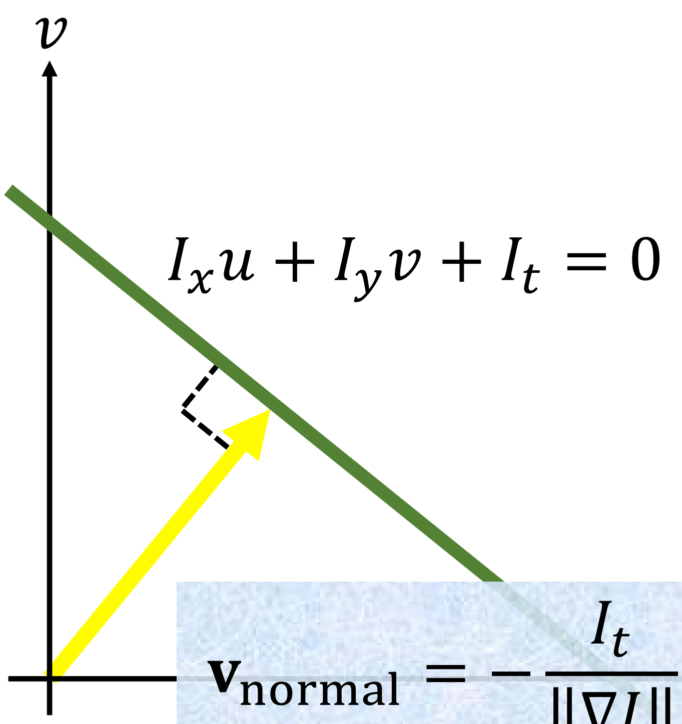
亮度恒常约束：$I_x u + I_y v + I_t = 0$



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

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

亮度恒常约束：$I_x u + I_y v + I_t = 0$

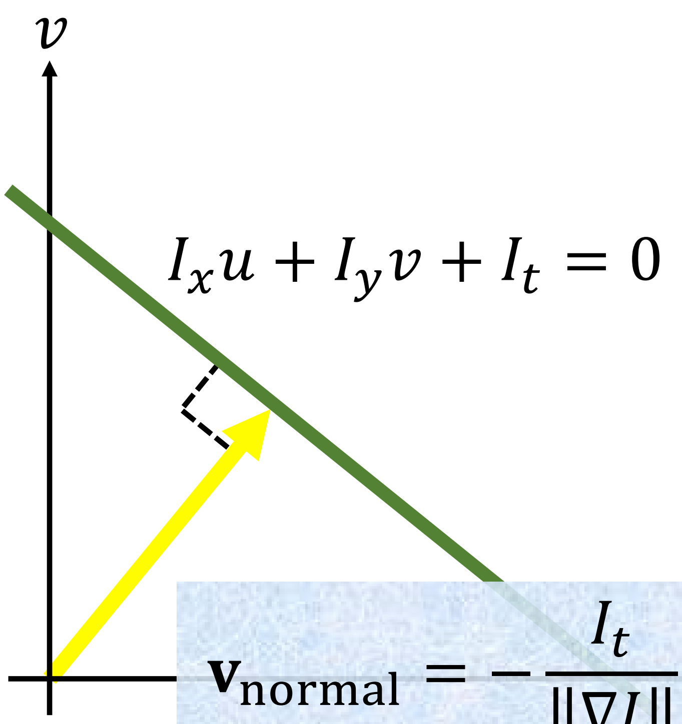


$$\mathbf{v}_{\text{normal}} = -\frac{I_t}{\|\nabla I\|}\frac{\nabla I}{\|\nabla I\|}$$

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

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

亮度恒常约束：$I_x u + I_y v + I_t = 0$



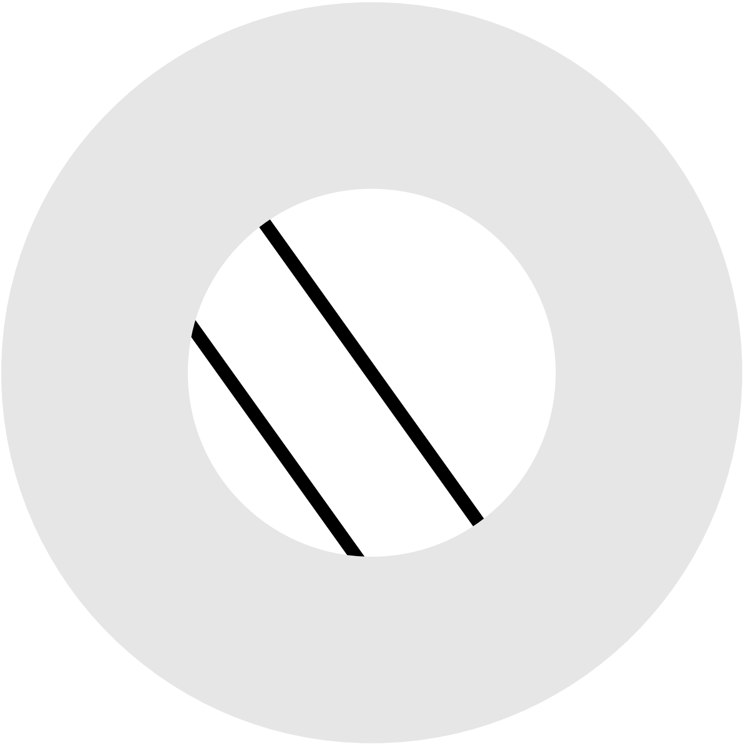
$$\mathbf{u}^{\mathrm{T}} \nabla I = -I_t$$

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

$$\mathbf{v}_{\text{normal}} = -\frac{I_t}{\|\nabla I\|} \frac{\nabla I}{\|\nabla I\|}$$

法向流

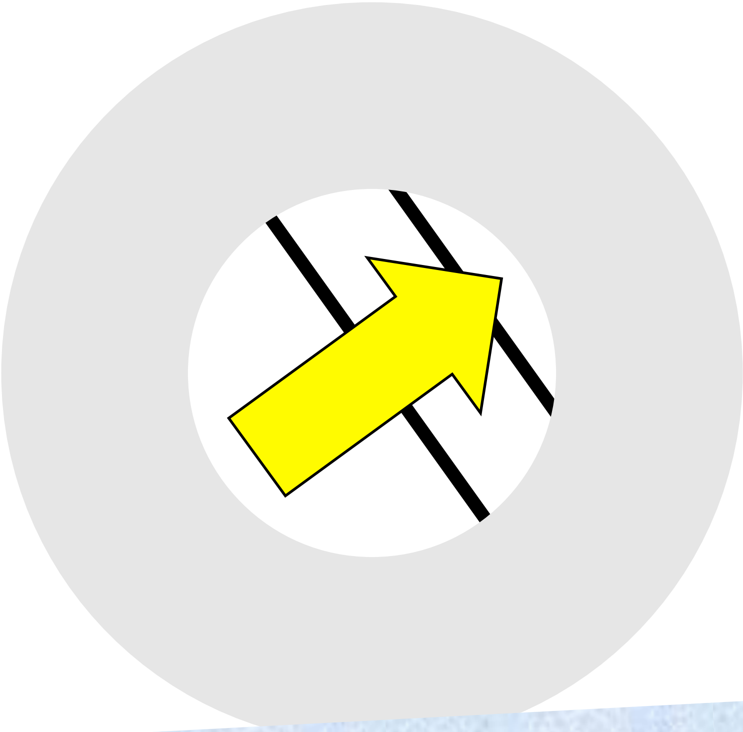
孔径问题



孔径问题

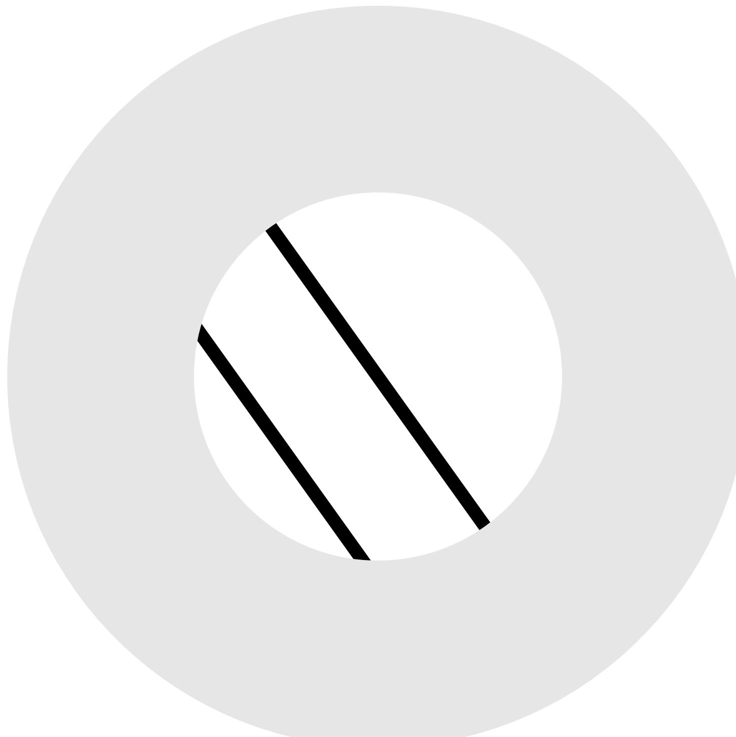
线往哪边移动？

孔径问题

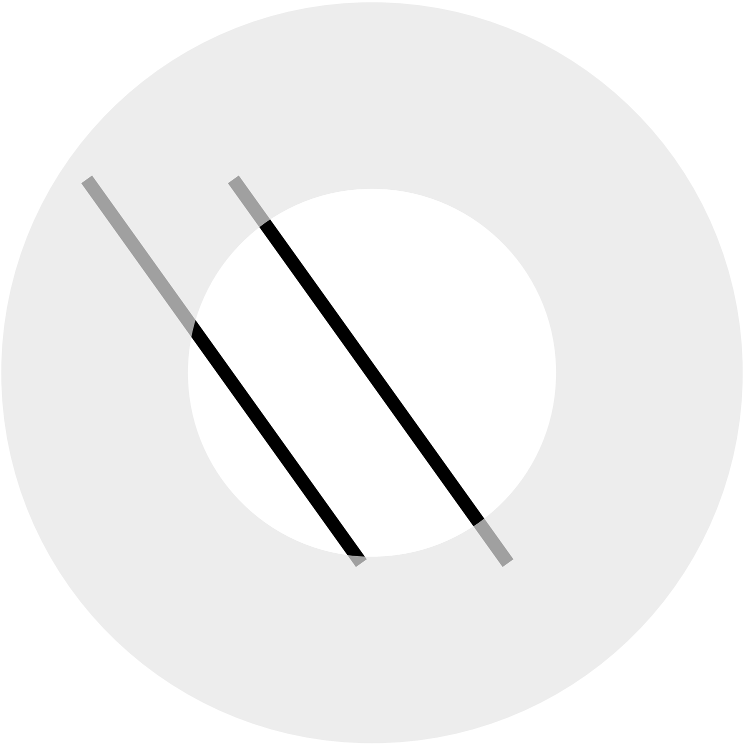


线往哪边移动?

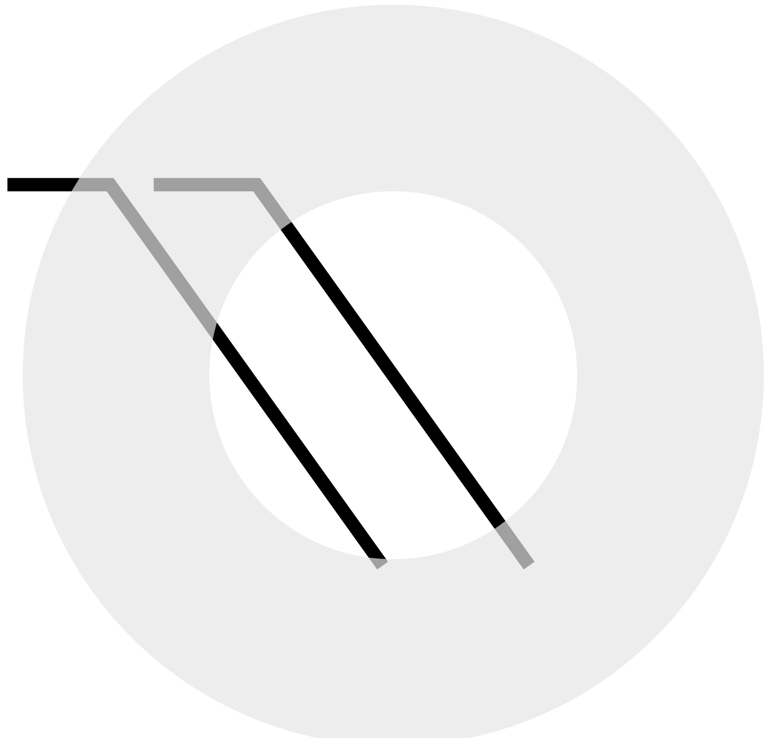
孔径问题



孔径问题



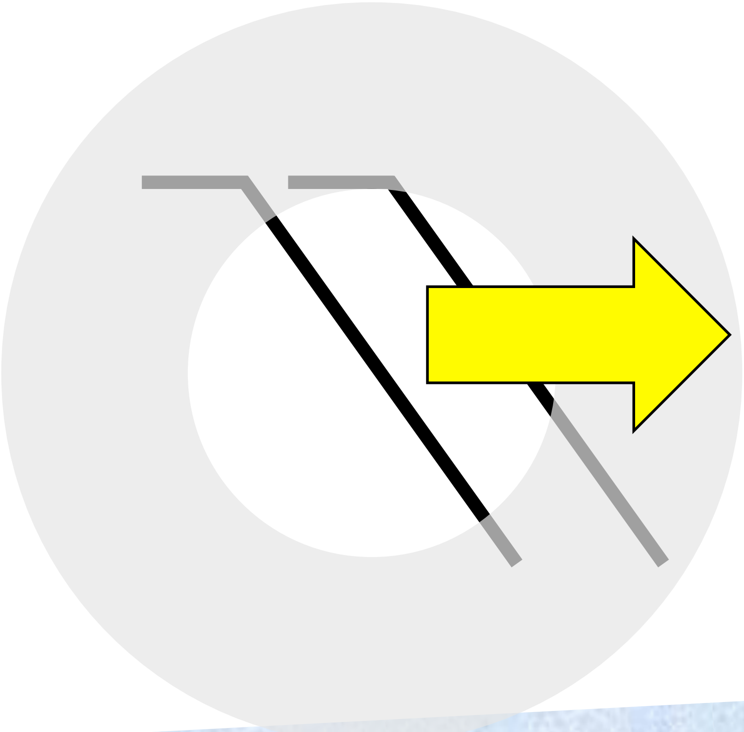
孔径问题



孔径问题

线往哪边移动?

孔径问题



线往哪边移动?

孔径问题

孔径问题

图案往哪边移动？


孔径问题
表观运动
图案往哪边移动?


孔径问题
真实运动
图案往哪边移动？

解决方案：引入其他约束

估计光流


局部


局部
全局


局部

局部方法

An Iterative Image Registration Technique with an Application to Stereo Vision

Bruce D. Lucas
Takeo Kanade

Computer Science Department
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

International Joint Conference on AI, 1981

e Image Registration Techniq

Application to Stereo Vision

Bruce D. Lucas
Takeo Kanade

Computer Science Department
Carnegie-Mellon University


30


30


假设一个像素邻域内的速度是恒定的


假设一个像素邻域内的速度是恒定的
$N \times N$

Lucas-Kanade
方法

$$\begin{pmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{N^2}) & I_y(\mathbf{p}_{N^2}) \end{pmatrix} \binom{u}{v} = \begin{pmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{N^2}) \end{pmatrix}$$

Lucas-Kanade 方法

$$\begin{pmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{N^2}) & I_y(\mathbf{p}_{N^2}) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{N^2}) \end{pmatrix}$$

$$\mathbf{A}_{N^2\times 2}$$

Lucas-Kanade方法

$$\begin{pmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{N^2}) & I_y(\mathbf{p}_{N^2}) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{N^2}) \end{pmatrix}$$

$$\mathbf{A}_{N^2\times 2} \qquad \mathbf{v}_{2\times 1}$$

Lucas-Kanade
方法

$$\begin{pmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{N^2}) & I_y(\mathbf{p}_{N^2}) \end{pmatrix} \binom{u}{v} = \begin{pmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{N^2}) \end{pmatrix}$$

$\mathbf{A}_{N^2\times 2}$ $\mathbf{v}_{2\times 1}$ $\mathbf{b}_{N^2\times 1}$

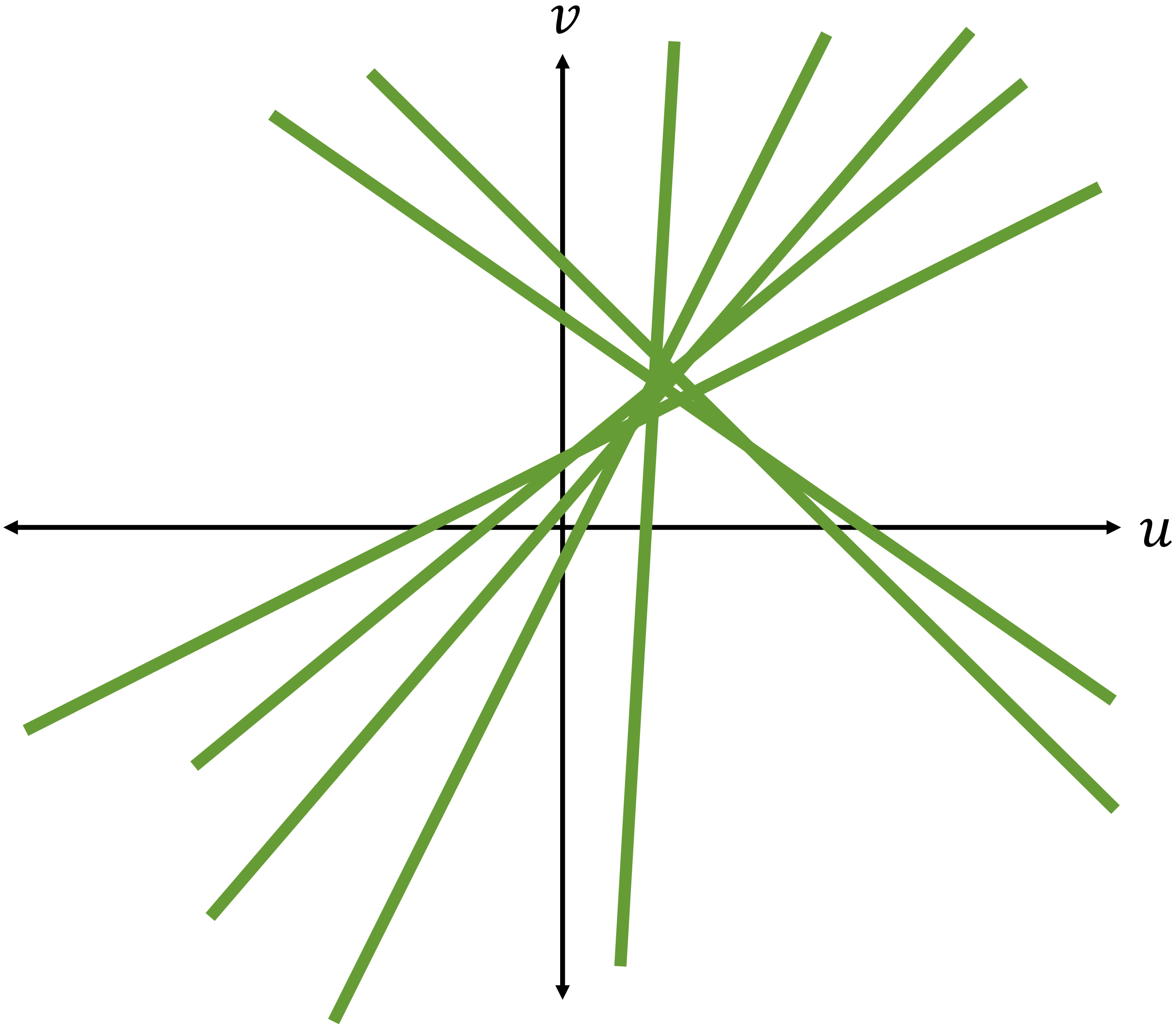
Lucas-Kanade
方法

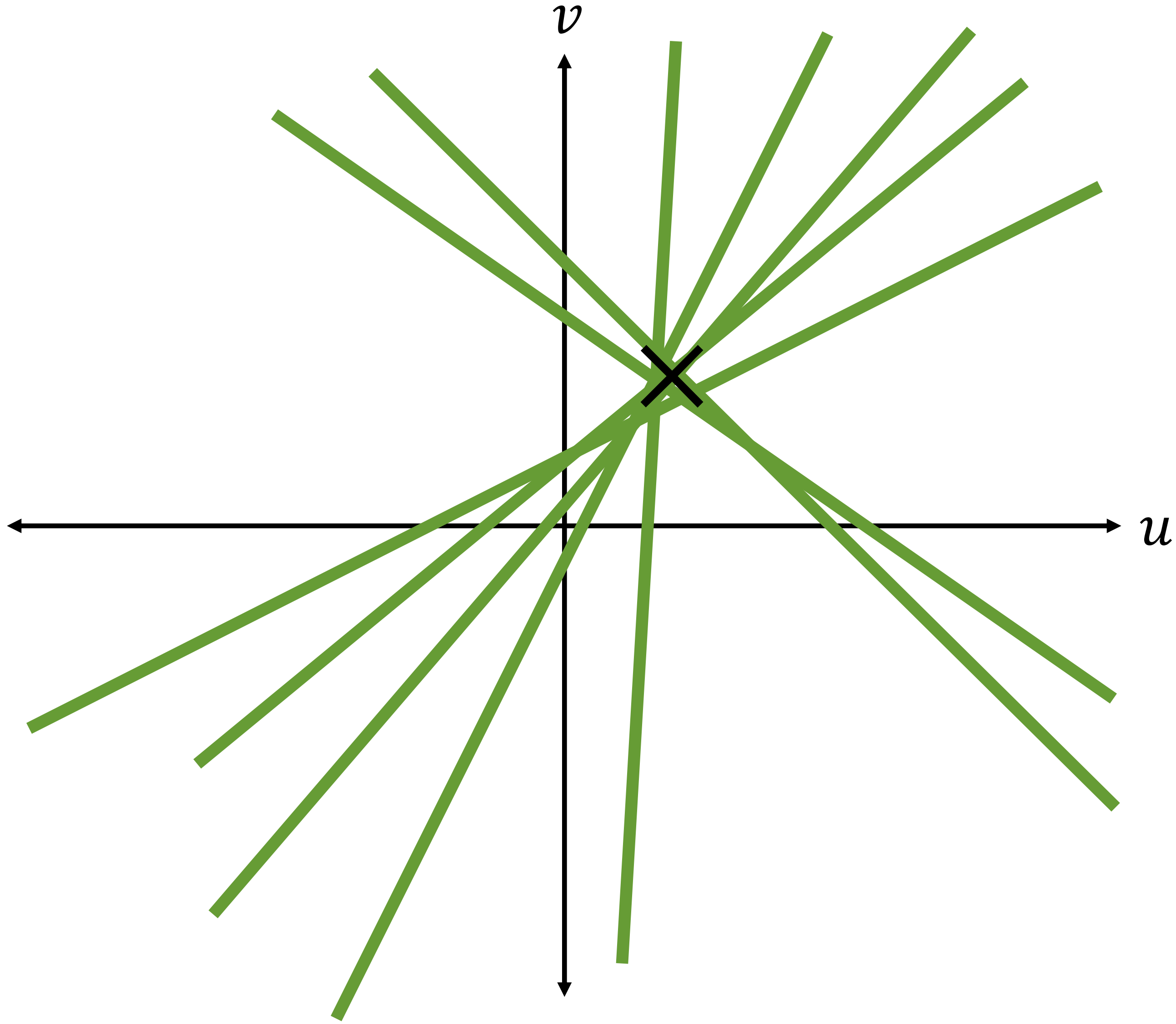
$$\begin{pmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{N^2}) & I_y(\mathbf{p}_{N^2}) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{N^2}) \end{pmatrix}$$

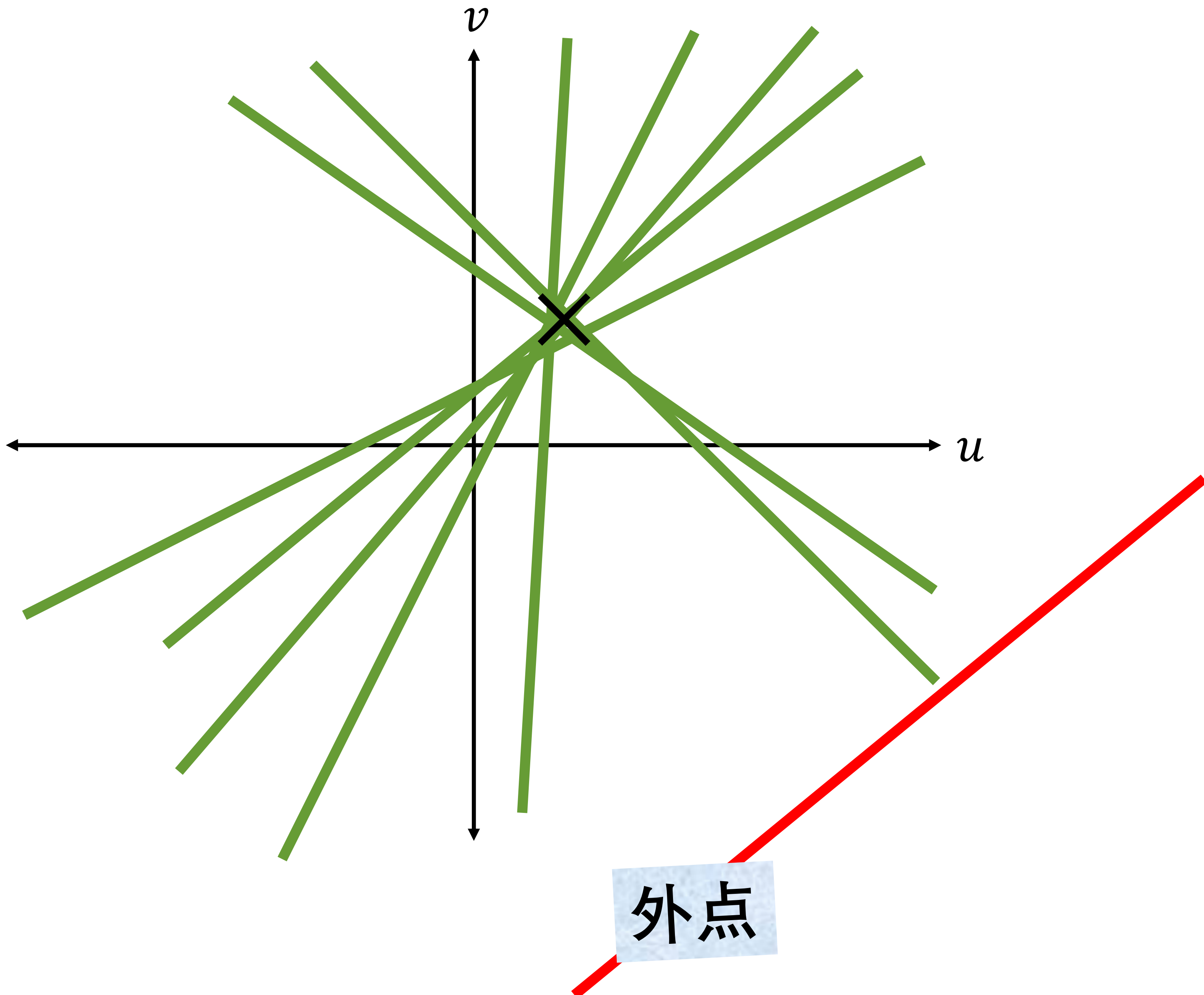
$$\mathbf{A}_{N^2\times 2} \qquad \mathbf{v}_{2\times 1} \qquad \mathbf{b}_{N^2\times 1}$$

怎样求解该系统?

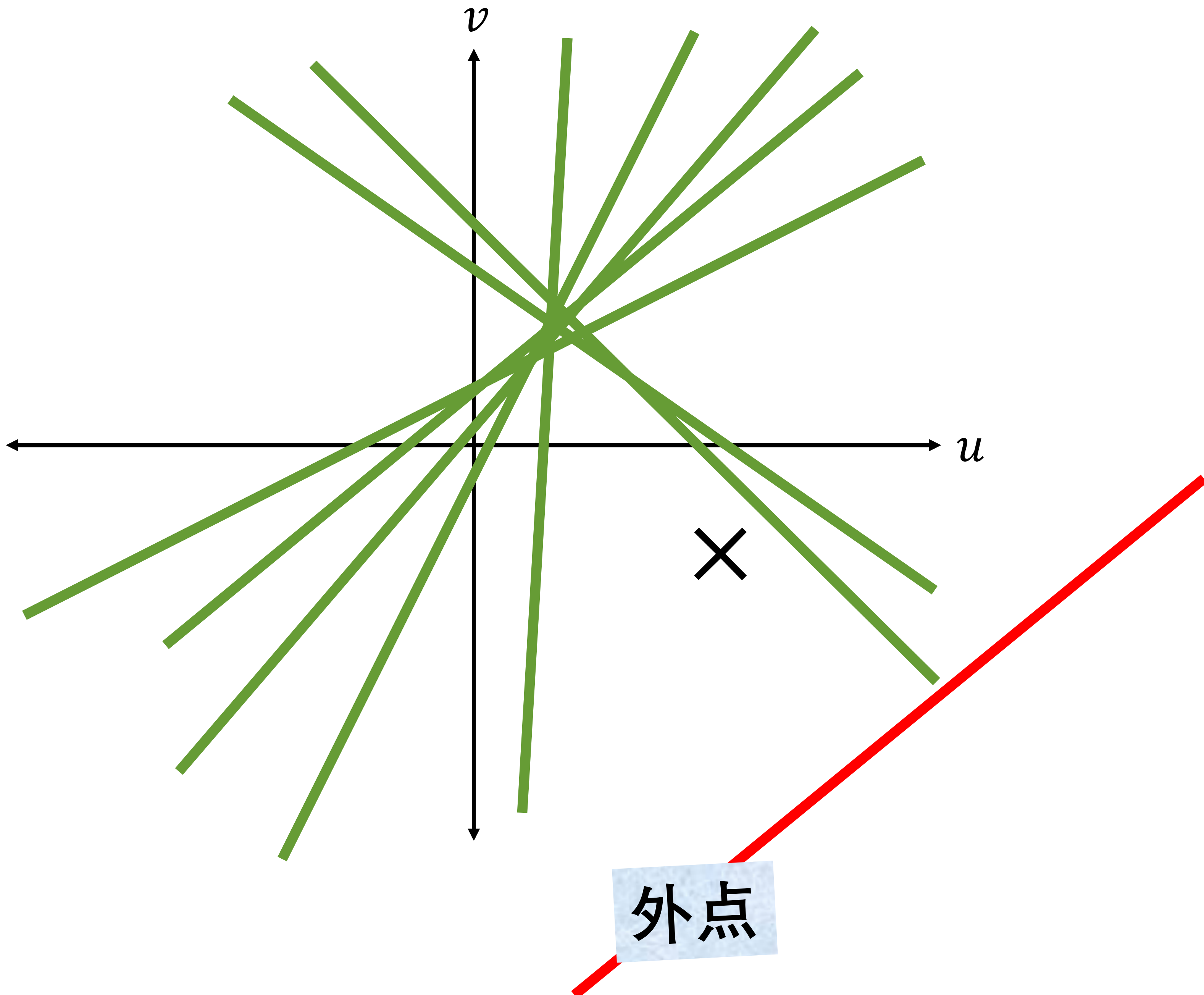
$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$


v
u


v
u



v
u
外点



v
u
外点

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{v} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

正规方程

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{v} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = - \begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}_{2\times 2} \qquad \mathbf{v}_{2\times 1} \qquad \mathbf{A}^{\mathrm{T}}\mathbf{b}_{2\times 1}$$

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{v} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\underbrace{\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{A}_{2\times2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times1}} = - \underbrace{\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{b}_{2\times1}}$$

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{v} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\underbrace{\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{A}_{2\times 2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times 1}} = -\underbrace{\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{b}_{2\times 1}}$$

看起来眼熟吗？

Harris角点

$$E(\Delta x, \Delta y) = \begin{pmatrix} \Delta x & \Delta y \end{pmatrix} \mathbf{M} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

其中

$$\mathbf{M} = \sum_{x,y} w(x,y) \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

Harris角点

$$E(\Delta x, \Delta y) = (\Delta x \quad \Delta y)\mathbf{M}\begin{pmatrix}\Delta x \\ \Delta y\end{pmatrix}$$

其中

$$\mathbf{M} = \sum_{x,y} w(x,y)\begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

M捕捉局部块中的结构

$$\arg\min_{\mathbf{v}} \|\mathbf{Av} - \mathbf{b}\|^2$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{Av} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\underbrace{\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{A}_{2\times 2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times 1}} = - \underbrace{\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{b}_{2\times 1}}$$

看起来眼熟吗？

$$\arg\min_{\mathbf{v}} \|\mathbf{A}\mathbf{v} - \mathbf{b}\|^2$$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{v} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\underbrace{\begin{pmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{A}_{2\times 2}} \underbrace{\begin{pmatrix} u \\ v \end{pmatrix}}_{\mathbf{v}_{2\times 1}} = -\underbrace{\begin{pmatrix} \sum I_x I_t \\ \sum I_y I_t \end{pmatrix}}_{\mathbf{A}^{\mathrm{T}}\mathbf{b}_{2\times 1}}$$

$$\mathbf{v} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-\mathbf{1}}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

$$\mathbf{v} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

可解性条件?

$$\mathbf{v} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

可解性条件?

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$是可逆的

$$\mathbf{v} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{\mathbf{-1}}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

可解性条件?

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$是可逆的

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$的特征值$\lambda_1, \lambda_2 \gg 0$

$$\mathbf{v} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

可解性条件?

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$是可逆的

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$的特征值$\lambda_1, \lambda_2 \gg 0$

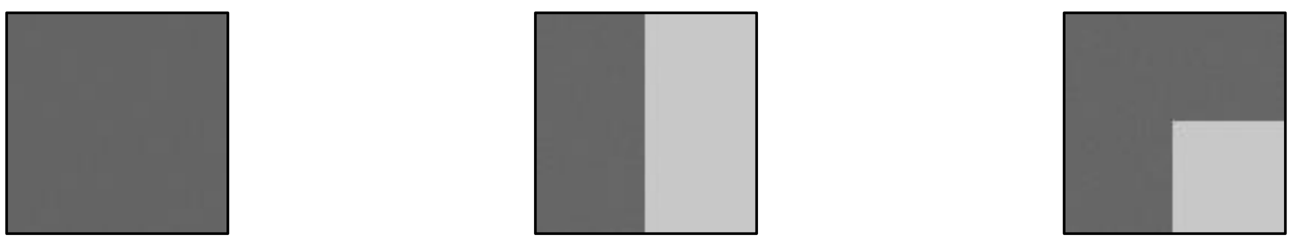
$\mathbf{A}^{\mathrm{T}}\mathbf{A}$ 应该是良置的，λ_1/λ_2 不太大

可解性条件?

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$是可逆的

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$的特征值$\lambda_1, \lambda_2 \gg 0$

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$ 应该是良置的，λ_1/λ_2 不太大



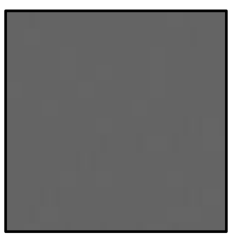
“平坦”　“边缘”　“角点”

可解性条件？

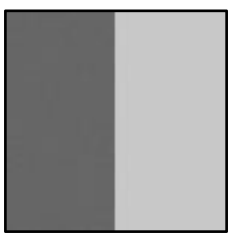
$\mathbf{A}^{\mathrm{T}}\mathbf{A}$是可逆的

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$的特征值$\lambda_1, \lambda_2 \gg 0$

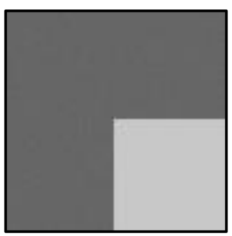
$\mathbf{A}^{\mathrm{T}}\mathbf{A}$ 应该是良置的，λ_1/λ_2 不太大



“平坦”



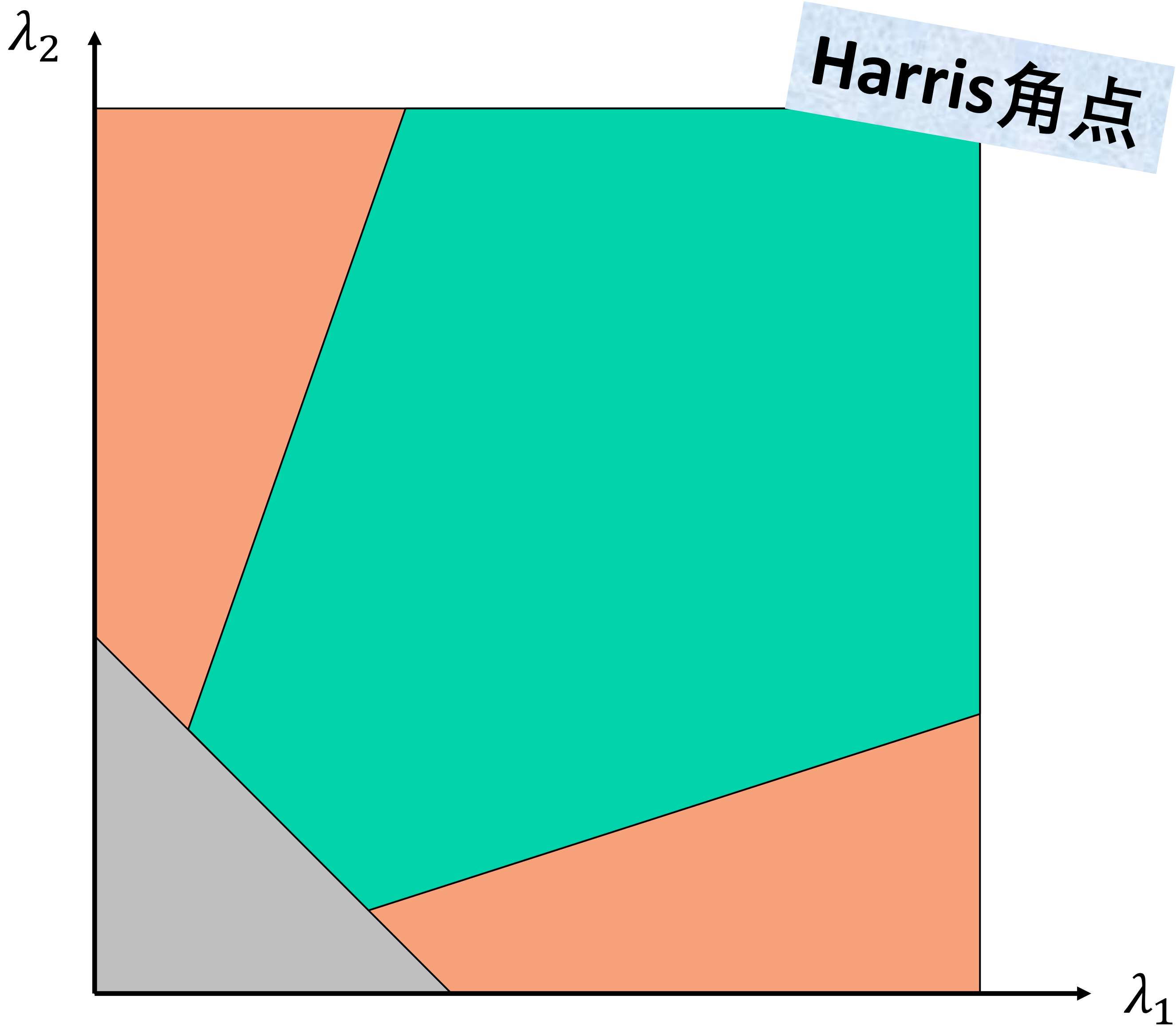
“边缘”



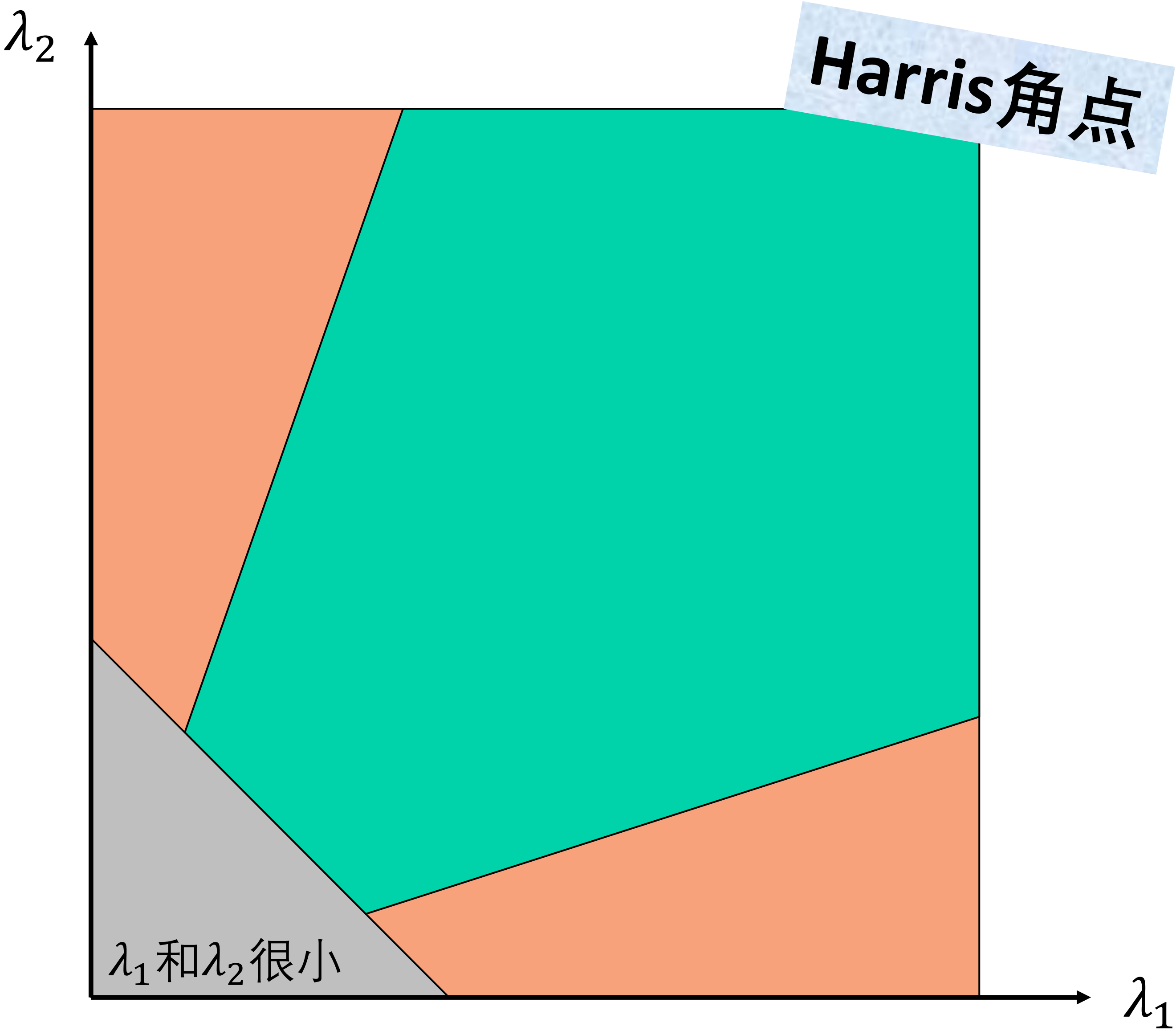
“角点”

什么图案满足这些条件？

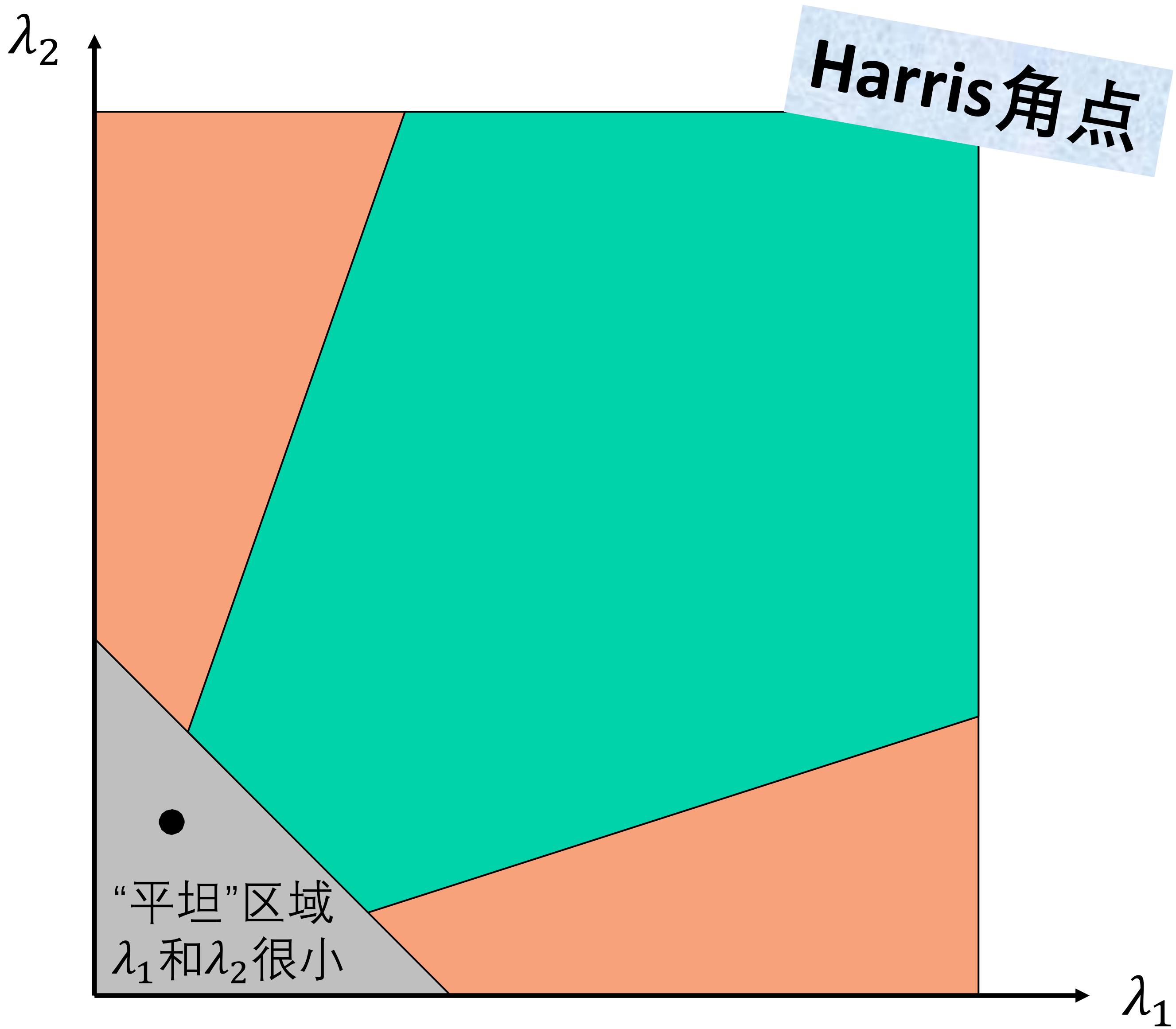
使用M的特征值对图像点分类



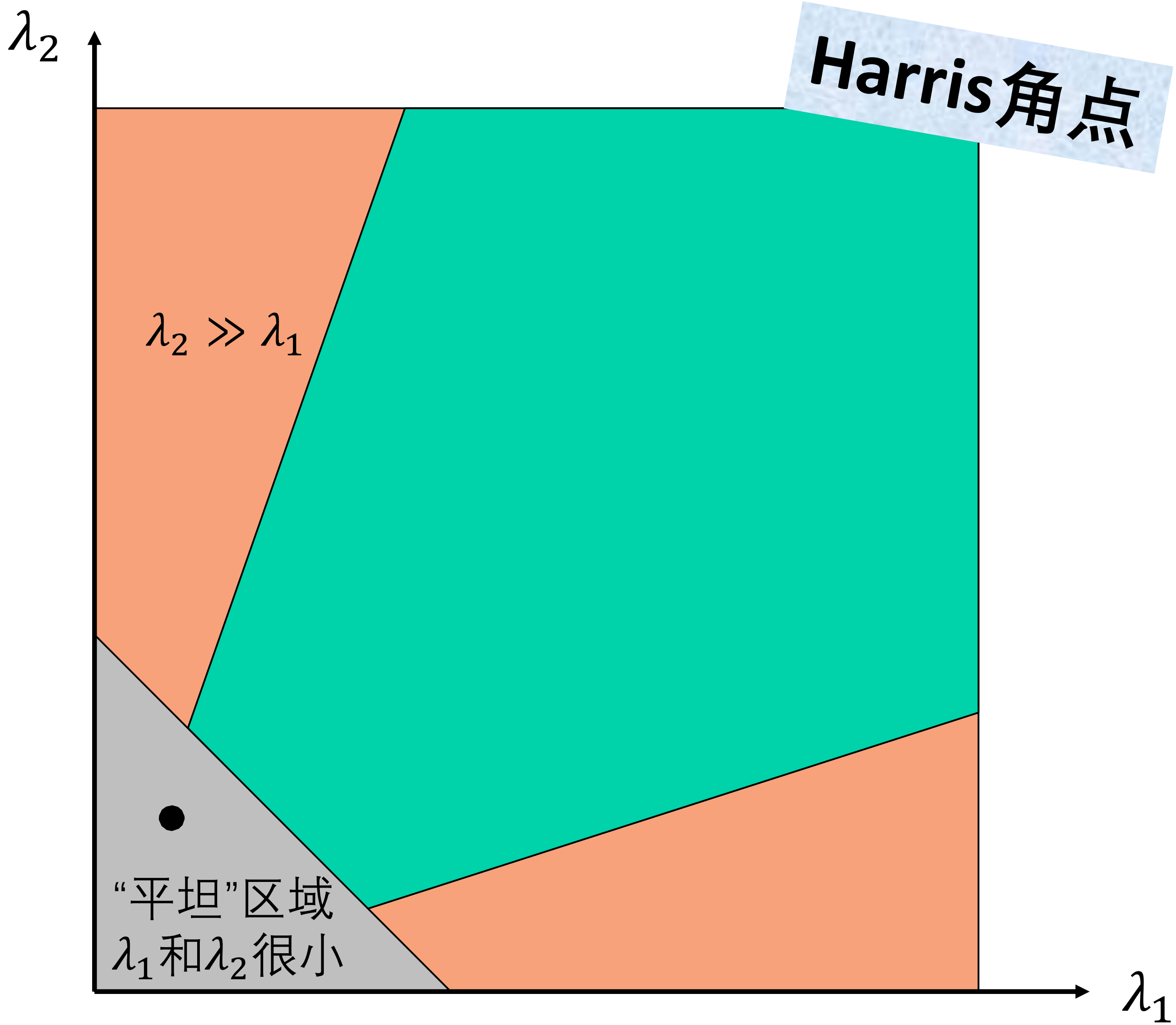
使用M的特征值对图像点分类



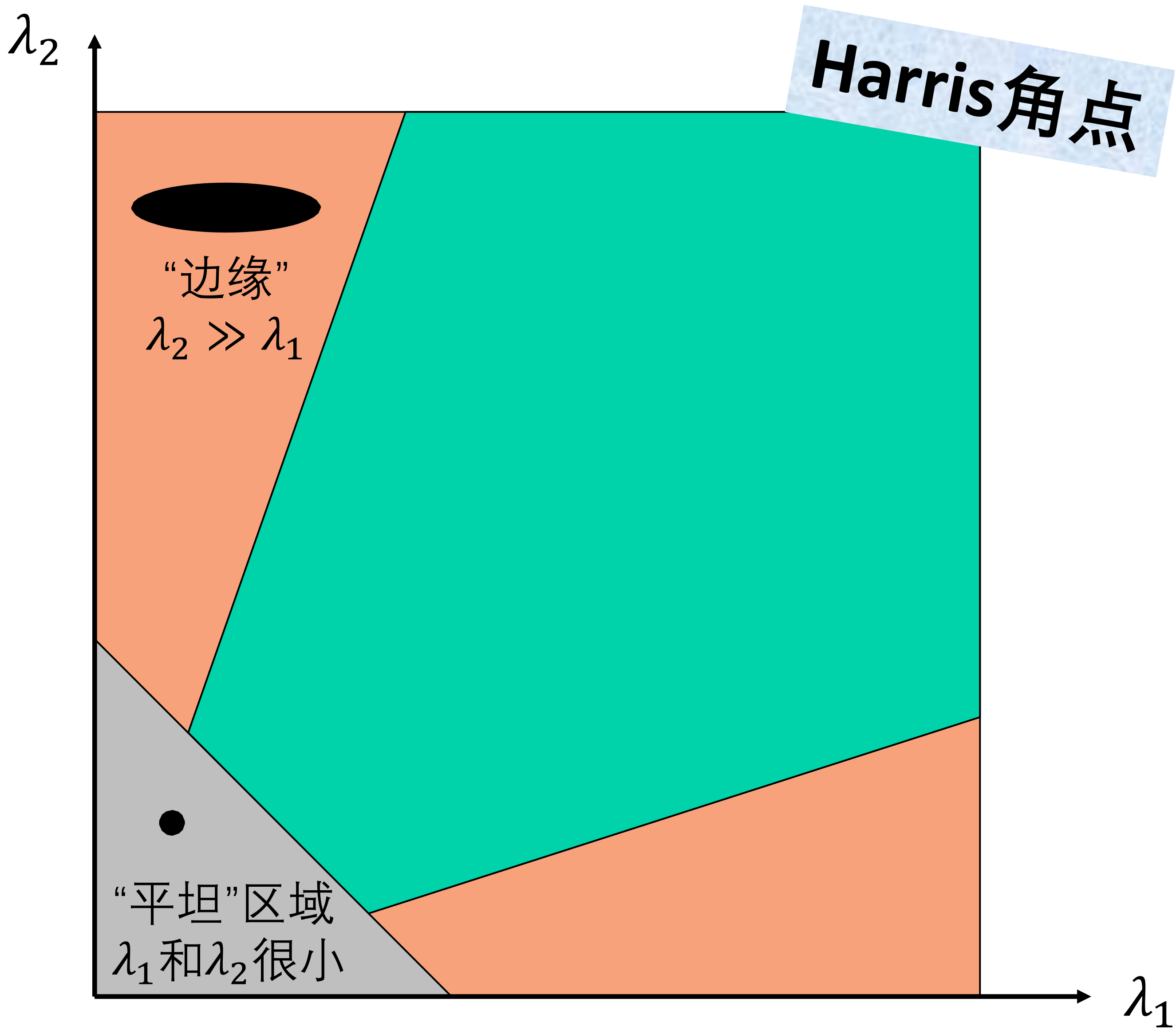
使用M的特征值对图像点分类



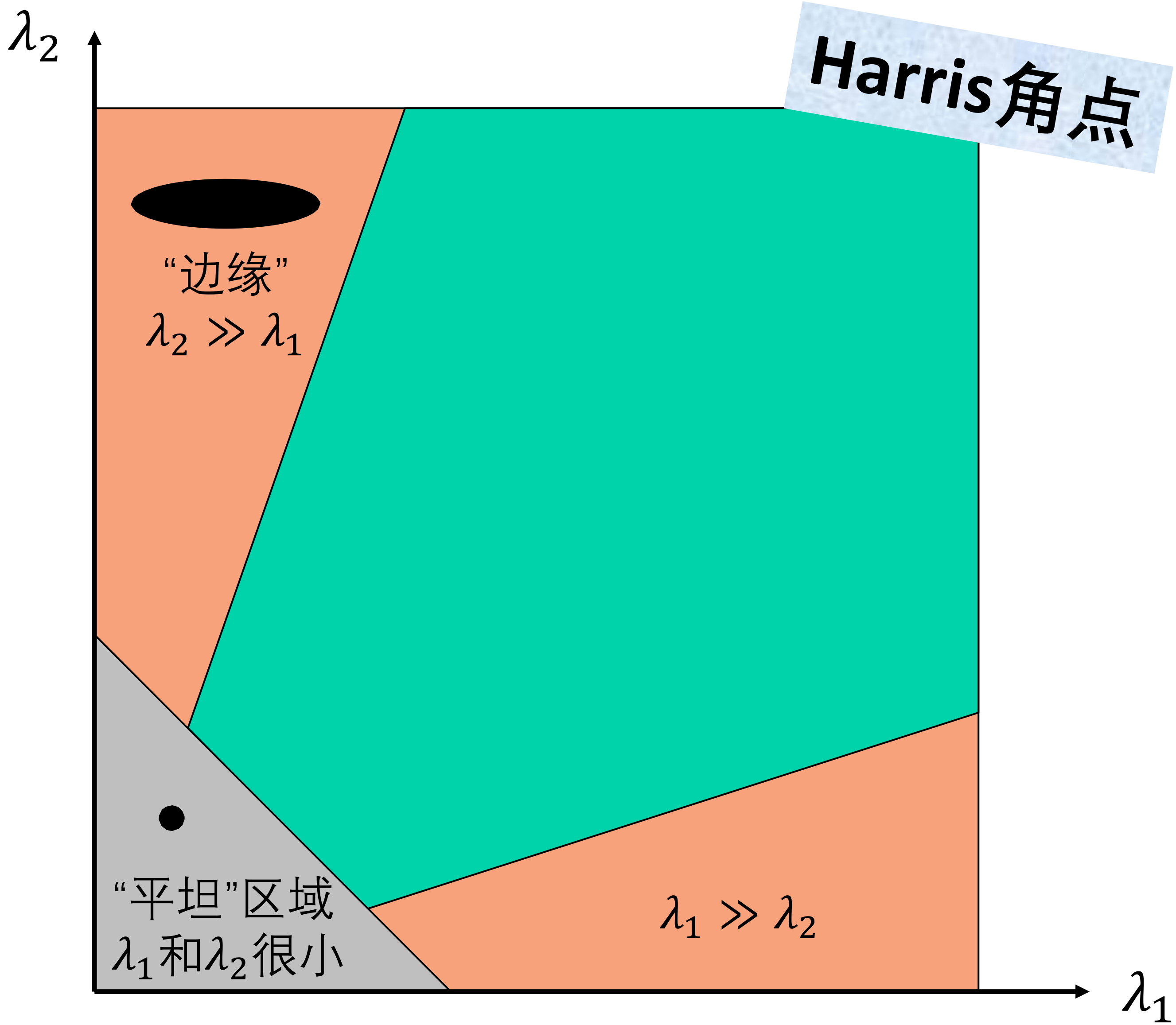
使用M的特征值对图像点分类



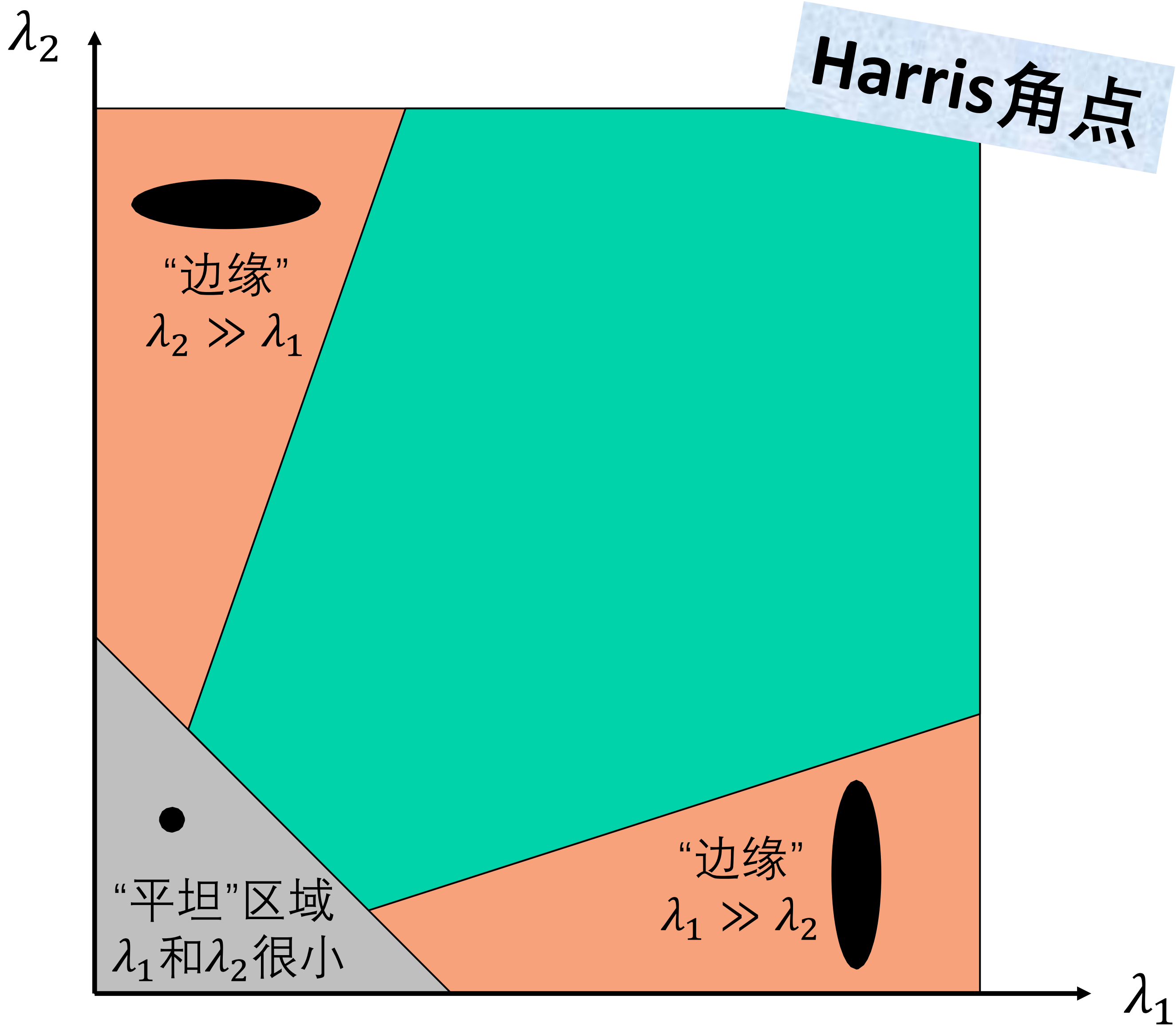
使用M的特征值对图像点分类



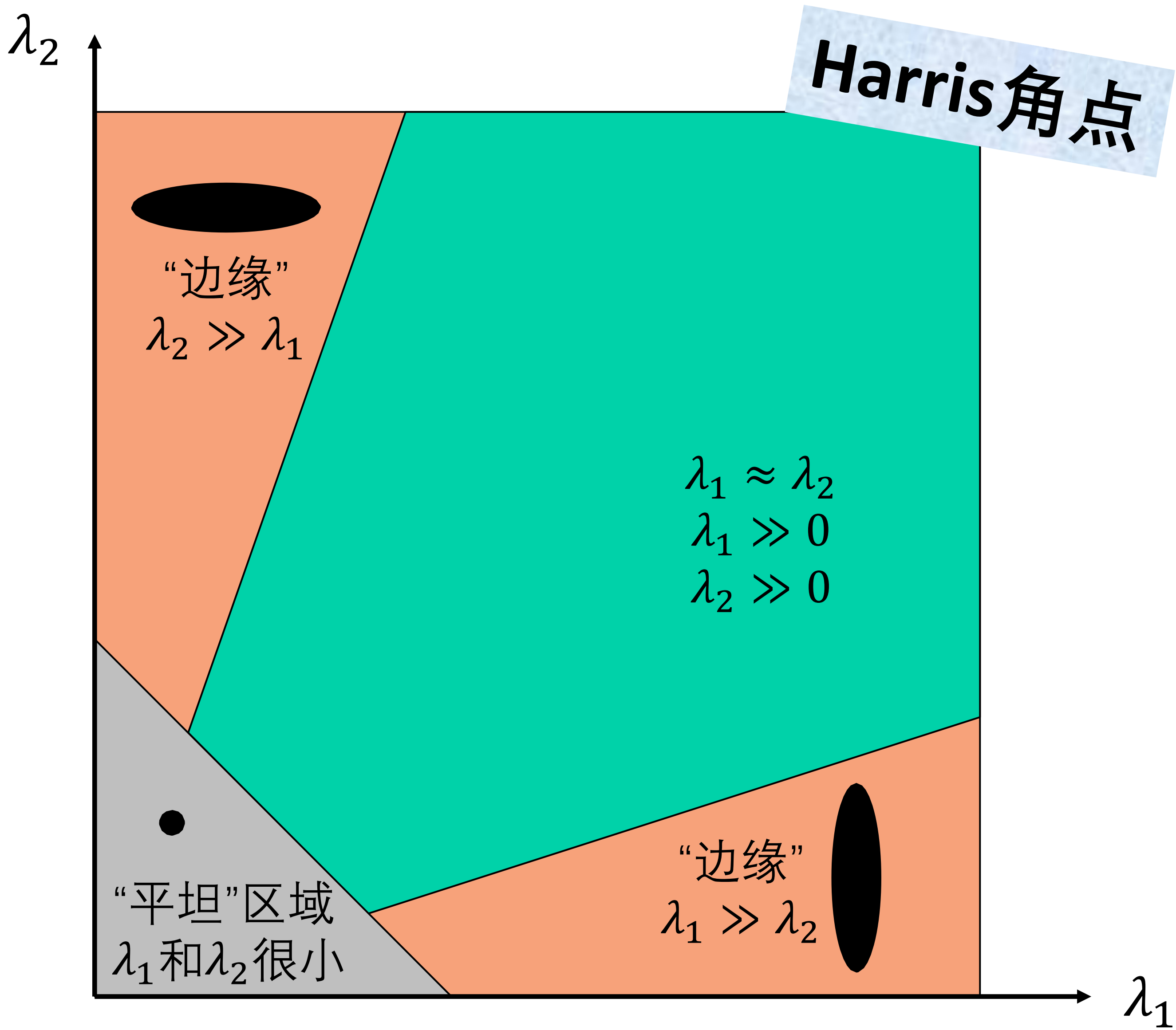
使用M的特征值对图像点分类



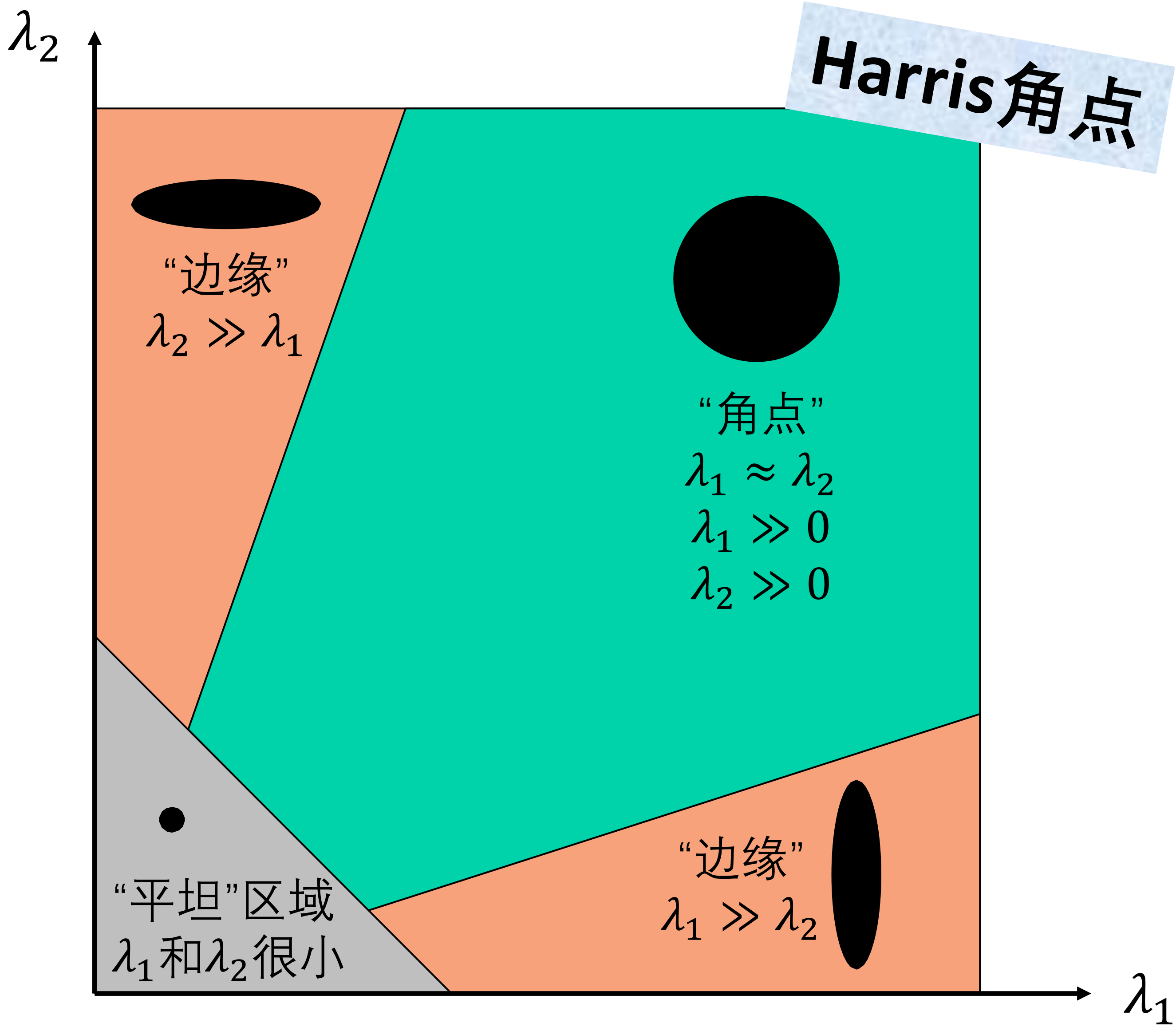
使用M的特征值对图像点分类



使用M的特征值对图像点分类



使用M的特征值对图像点分类

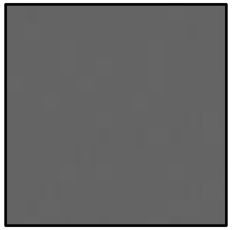


可解性条件?

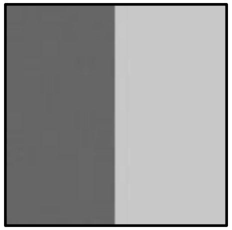
$\mathbf{A}^{\mathrm{T}}\mathbf{A}$是可逆的

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$的特征值$\lambda_1, \lambda_2 \gg 0$

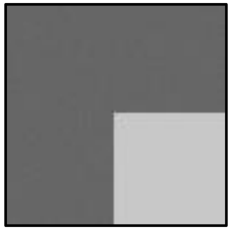
$\mathbf{A}^{\mathrm{T}}\mathbf{A}$ 应该是良置的，λ_1/λ_2 不太大



“平坦”



“边缘”



“角点”

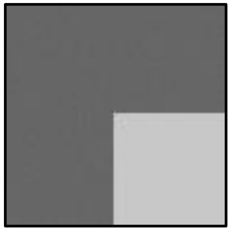
什么图案满足这些条件?

可解性条件?

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$是可逆的

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$的特征值$\lambda_1, \lambda_2 \gg 0$

$\mathbf{A}^{\mathrm{T}}\mathbf{A}$ 应该是良置的，λ_1/λ_2 不太大



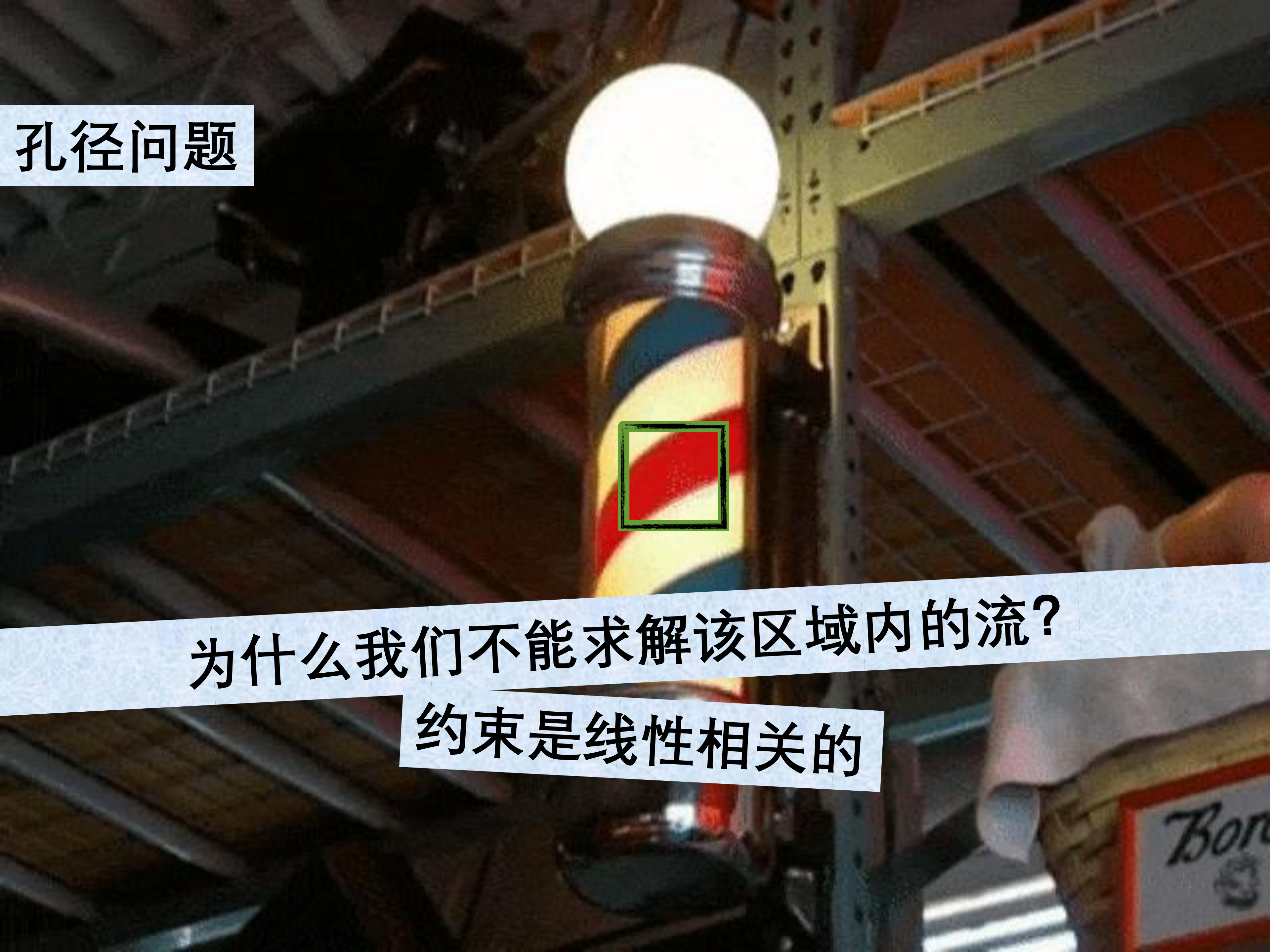
“角点”

孔径问题

孔径问题

孔径问题

为什么我们不能求解该区域内的流？


孔径问题
为什么我们不能求解该区域内的流？
约束是线性相关的

优点

优点

容易实现

优点

- 容易实现
- 快速计算

缺点

缺点

亮度不恒常时失效

缺点

亮度不恒常时失效

当窗口内的强度结构较差时失效

缺点

- 亮度不恒常时失效
- 当窗口内的强度结构较差时失效
- 沿着边界失效

缺点

亮度不恒常时失效

当窗口内的强度结构较差时失效

沿着边界失效

速度太快时失效

缺点

亮度不恒常时失效

当窗口内的强度结构较差时失效

沿着边界失效

速度太快时失效

局部的恒定平移是受限制的

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t)$$

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t)$$

平移块运动是受限的

$$I(x, y, t) = I(x + \Delta x(\mathbf{x}; \mathbf{p}), y + \Delta y(\mathbf{x}; \mathbf{p}), t + \Delta t)$$

$$I(x, y, t) = I(x + \Delta x(\mathbf{x}; \mathbf{p}), y + \Delta y(\mathbf{x}; \mathbf{p}), t + \Delta t)$$

参数模型描述的运动

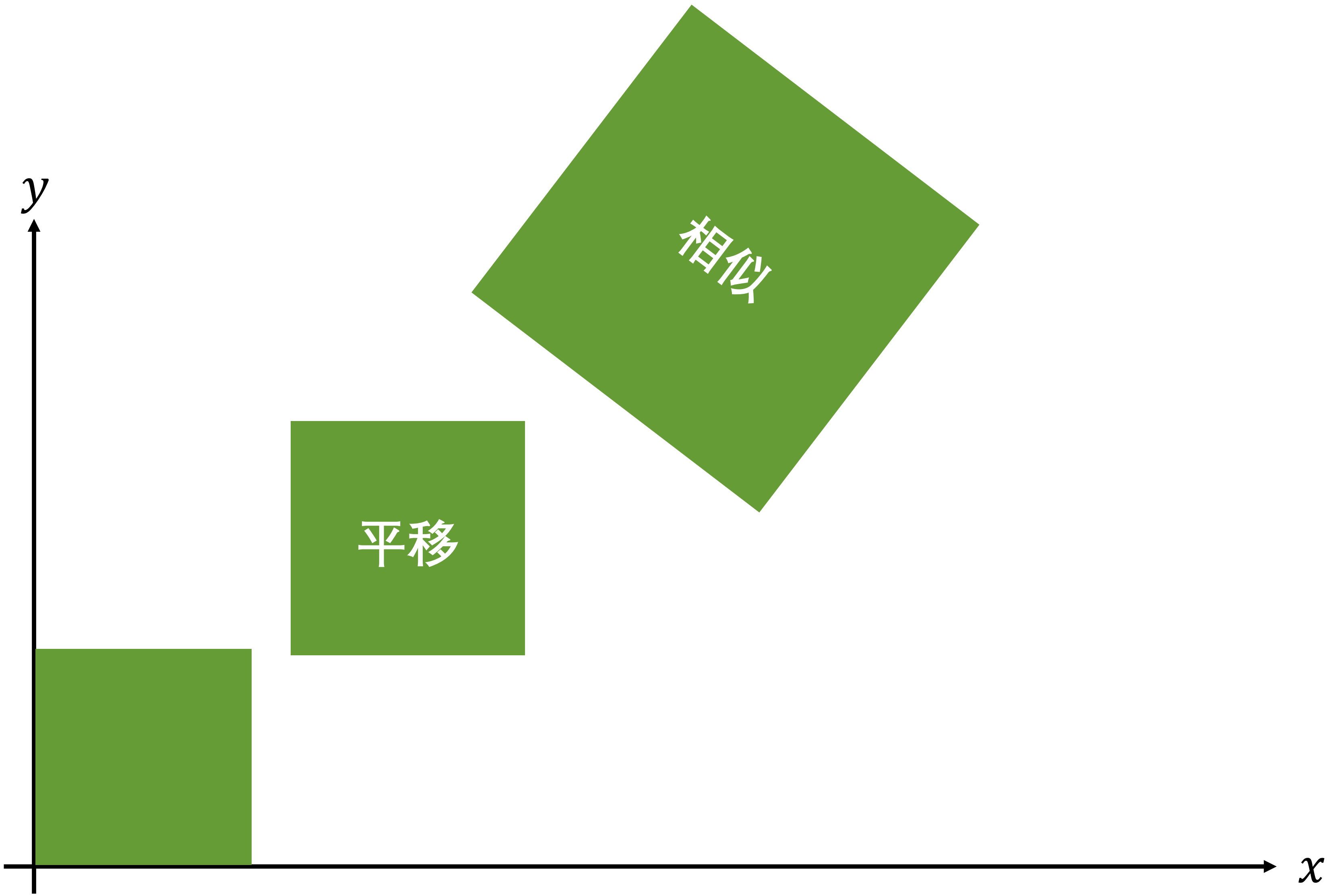
$$I(x, y, t) = I(x + \Delta x(\mathbf{x}; \mathbf{p}), y + \Delta y(\mathbf{x}; \mathbf{p}), t + \Delta t)$$

参数模型描述的运动

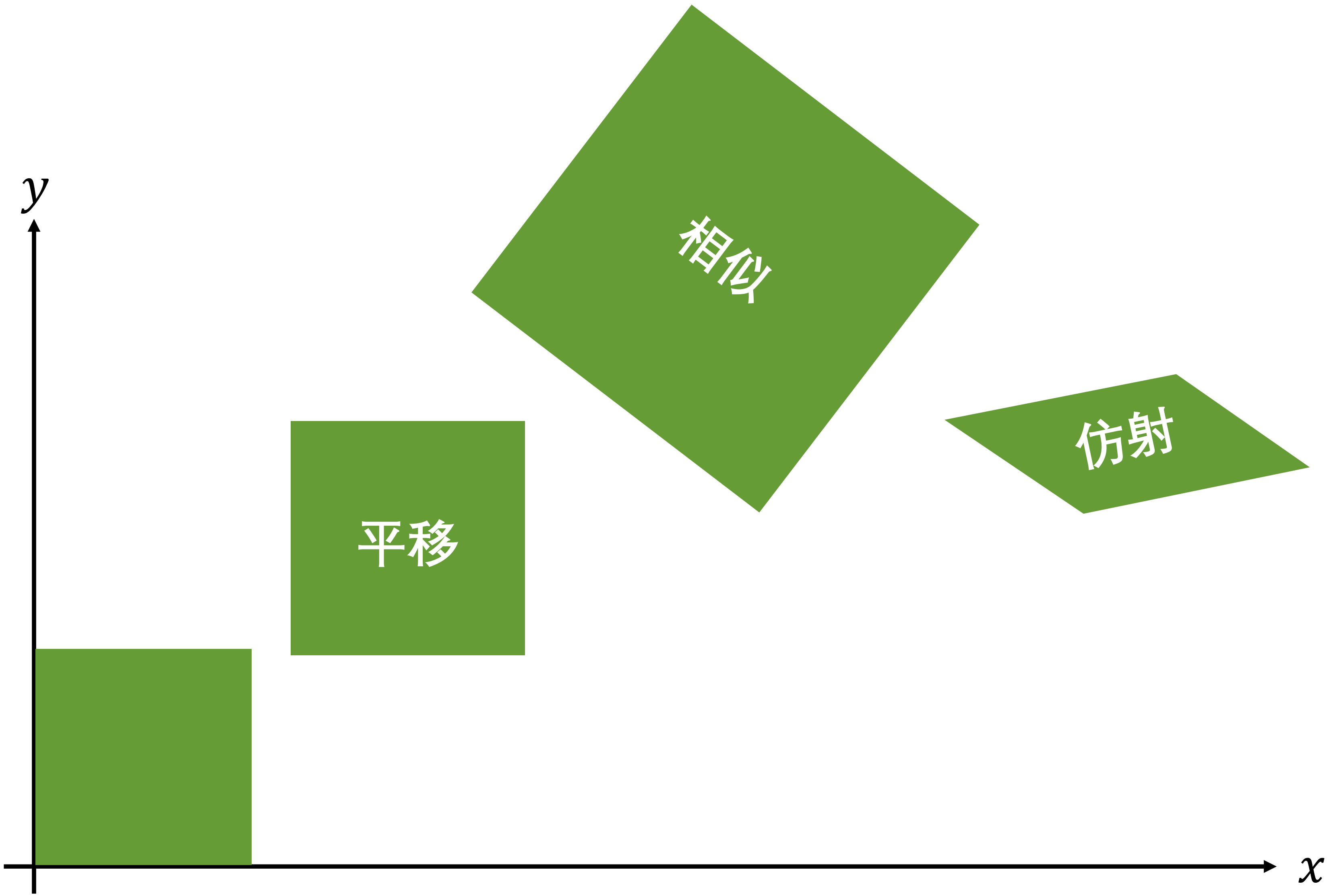
参数化运动示例

参数化运动示例

参数化运动示例



参数化运动示例



仿射运动

$$u(x,y) = a_1 + a_2 x + a_3 y$$

$$v(x,y) = a_4 + a_5 x + a_6 y$$

仿射运动

$$u(x,y) = a_1 + a_2 x + a_3 y$$
$$v(x,y) = a_4 + a_5 x + a_6 y$$

代入亮度恒常方程

仿射运动

$$u(x, y) = a_1 + a_2 x + a_3 y$$

$$v(x, y) = a_4 + a_5 x + a_6 y$$

代入亮度恒常方程

$$I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t \approx 0$$

仿射运动

$$u(x,y) = a_1 + a_2 x + a_3 y$$
$$v(x,y) = a_4 + a_5 x + a_6 y$$

代入亮度恒常方程

$$I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t \approx 0$$

仿射运动

$$u(x,y) = a_1 + a_2 x + a_3 y$$
$$v(x,y) = a_4 + a_5 x + a_6 y$$

代入亮度恒常方程

$$I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t \approx 0$$

每像素有六个未知数构成一个线性约束

仿射运动

$$u(x,y) = a_1 + a_2 x + a_3 y$$
$$v(x,y) = a_4 + a_5 x + a_6 y$$

代入亮度恒常方程

$$I_x(a_1 + a_2 x + a_3 y) + I_y(a_4 + a_5 x + a_6 y) + I_t \approx 0$$

每像素有六个未知数构成一个线性约束

在像素邻域中使用最小二乘法求解

KLT

Kanade-Lucas-Tomasi特征跟踪器

Good Features to Track

Jianbo Shi
Computer Science Department
Cornell University
Ithaca, NY 14853

Carlo Tomasi
Computer Science Department
Stanford University
Stanford, CA 94305

Abstract

No feature-based vision system can work unless good features can be identified and tracked from frame to frame. Although tracking itself is by and large a solved problem, selecting features that can be tracked well and correspond to physical points in the world is still hard. We propose a feature selection criterion that is optimal by construction because it is based on how the tracker

even good features can become occluded, and trackers often blissfully drift away from their original target when this occurs. No feature-based vision system can be claimed to really work until these issues have been settled.

In this paper we show how to monitor the quality of image features during tracking by using a measure of feature *dissimilarity* that quantifies the change of appearance of a feature between the first and the current

CVPR, 1994

KLT跟踪总结

1. 找到好的特征来跟踪

KLT跟踪总结

1. 找到好的特征来跟踪

   Harris角点或两个特征值皆大于某阈值

KLT跟踪总结

1. 找到好的特征来跟踪

   Harris角点或两个特征值皆大于某阈值

2. 计算帧间特征位移

KLT跟踪总结

1. 找到好的特征来跟踪

Harris角点或两个特征值皆大于某阈值

2. 计算帧间特征位移

使用Lucas-Kanade方法估计块的平移

KLT跟踪总结

1. 找到好的特征来跟踪

   Harris角点或两个特征值皆大于某阈值

2. 计算帧间特征位移

   使用Lucas-Kanade方法估计块的平移

3. **验证块的外观**

KLT跟踪总结

1. 找到好的特征来跟踪

Harris角点或两个特征值皆大于某阀值

2. 计算帧间特征位移

使用Lucas-Kanade方法估计块的平移

3. 验证块的外观

计算当前块和它在第一帧中的块之间的仿射配准。
如果仿射补偿图像太不相似，则丢弃该特征

KLT跟踪总结

1. **找到好的特征来跟踪**

   Harris角点或两个特征值皆大于某阈值

2. **计算帧间特征位移**

   使用Lucas-Kanade方法估计块的平移

3. **验证块的外观**

   计算当前块和它在第一帧中的块之间的仿射配准。如果仿射补偿图像太不相似，则丢弃该特征


局部


全局

全局方法

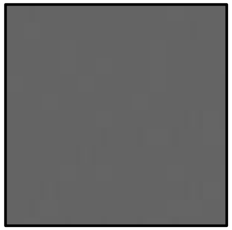
Determining Optical Flow

Berthold K.P. Horn and Brian G. Schunck
Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Artificial Intelligence, 1981

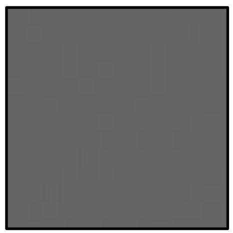
局部运动本质上是有歧义的

局部运动本质上是有歧义的

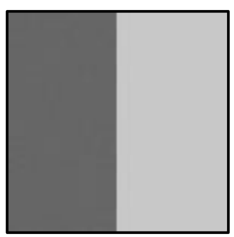


完全不明确的

局部运动本质上是有歧义的

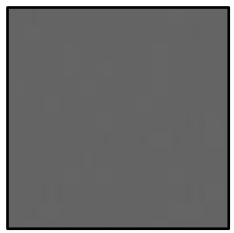


完全不明确的

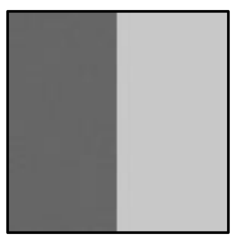


沿着法线方向是确定的，
沿着切线方向是不明确的

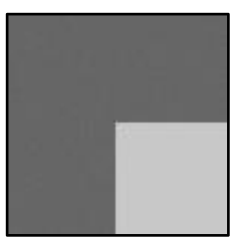
局部运动本质上是有歧义的



完全不明确的



沿着法线方向是确定的，
沿着切线方向是不明确的

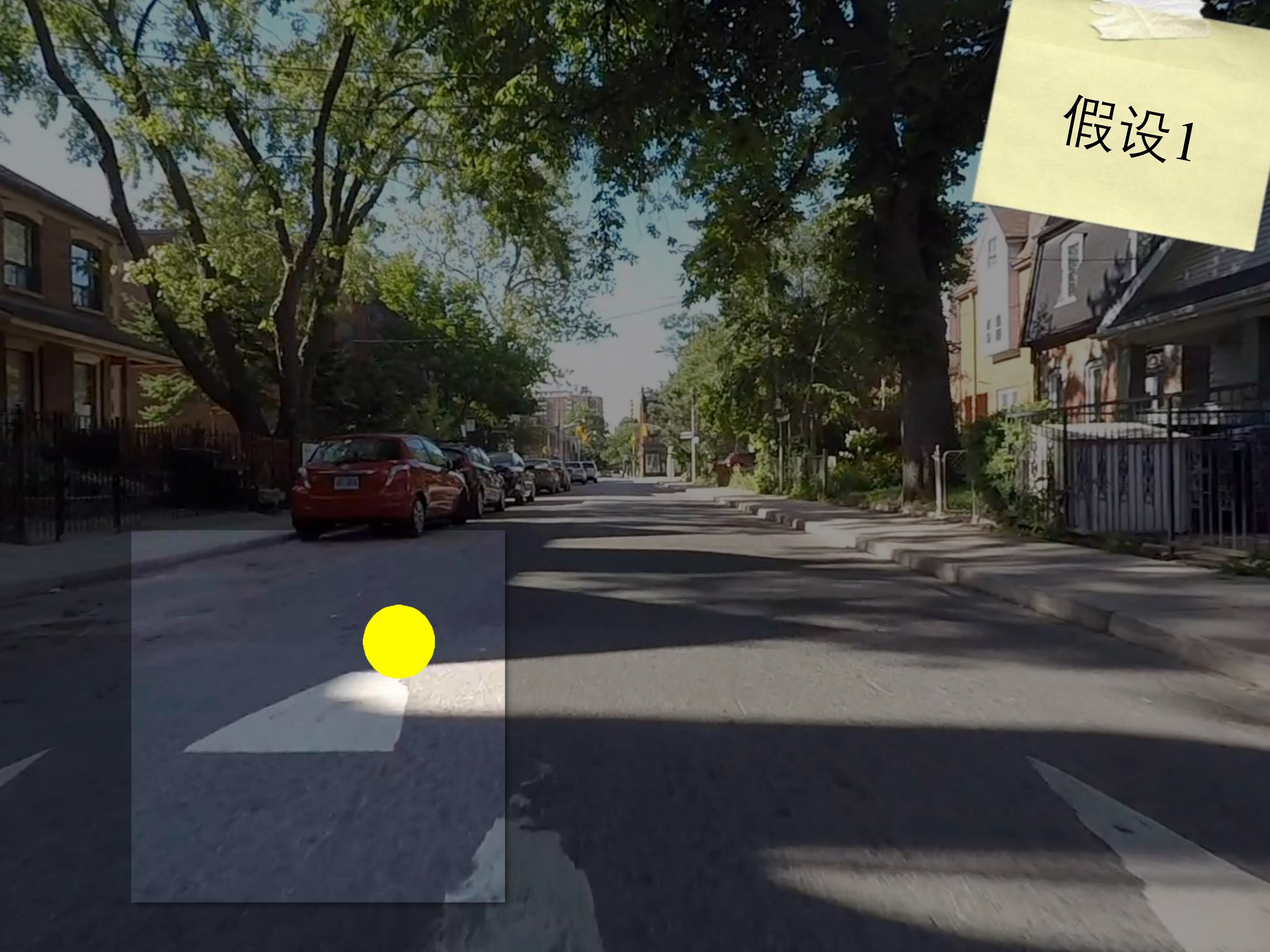


无歧义

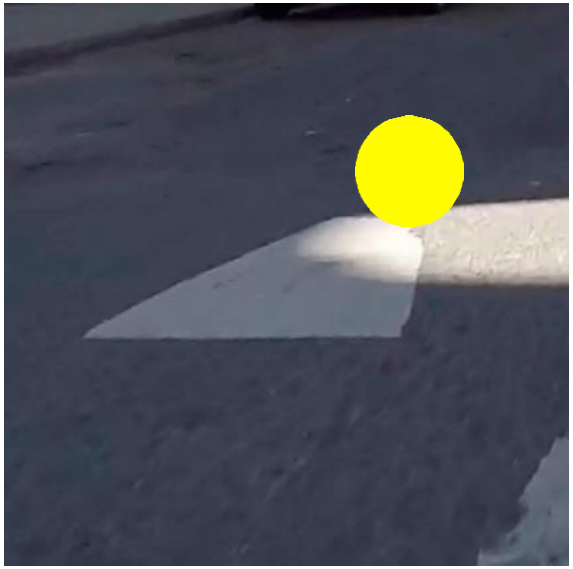
2

假设


假设1


假设1

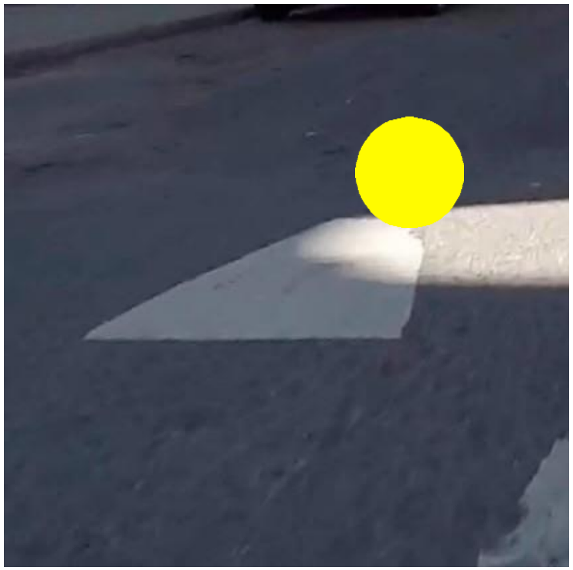
假设1



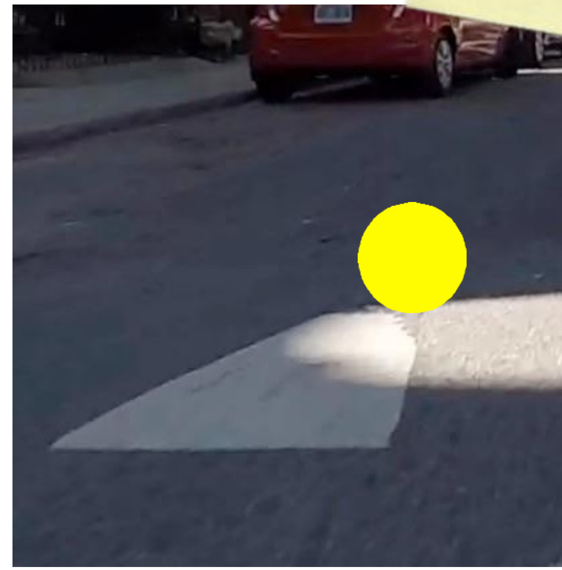
第t帧

第$t+1$帧

假设1



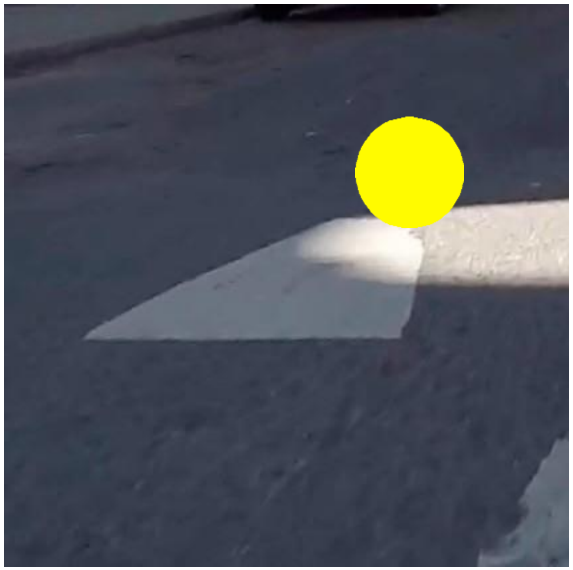
第t帧



第$t+1$帧

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

假设1



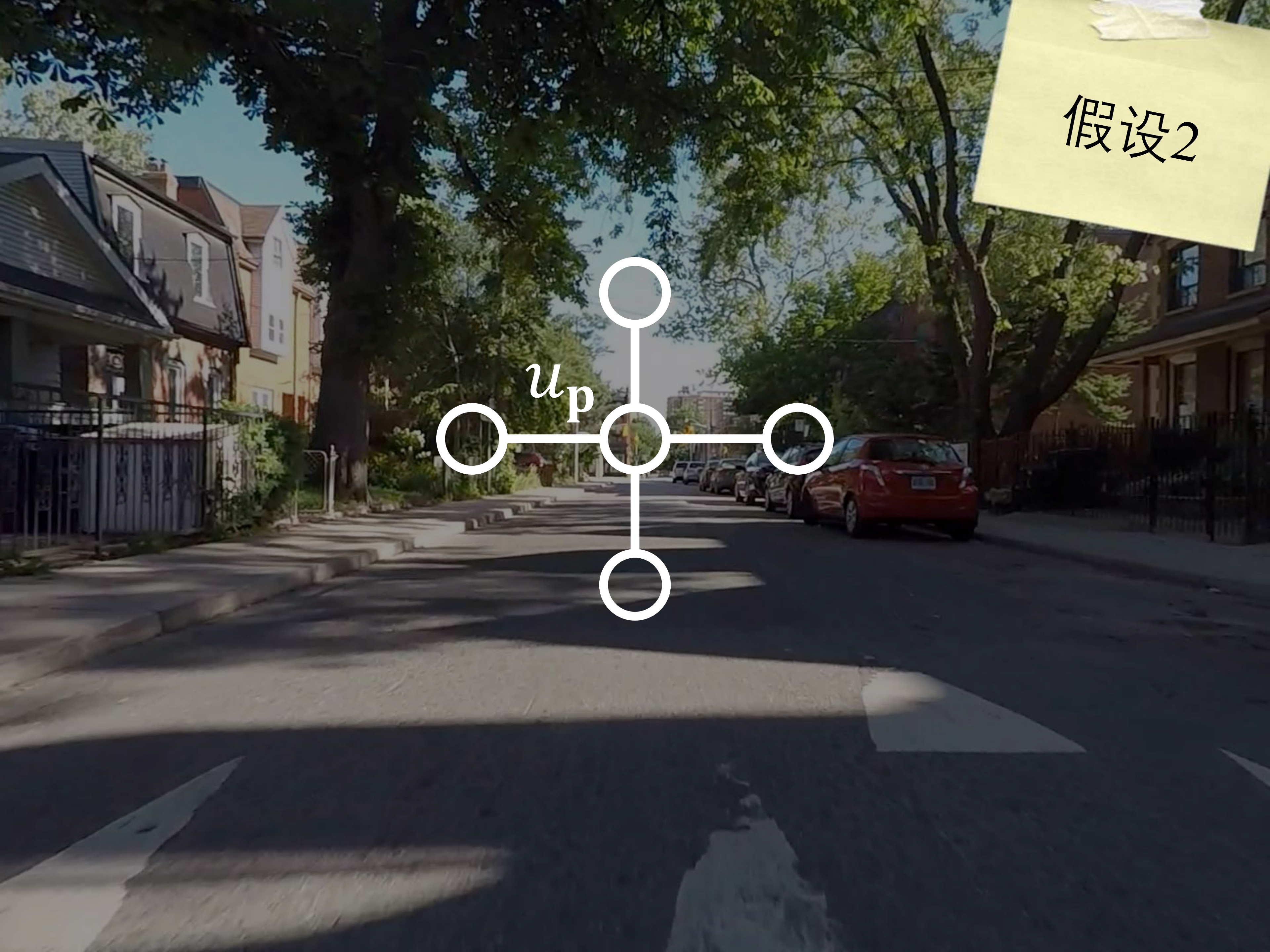
第t帧

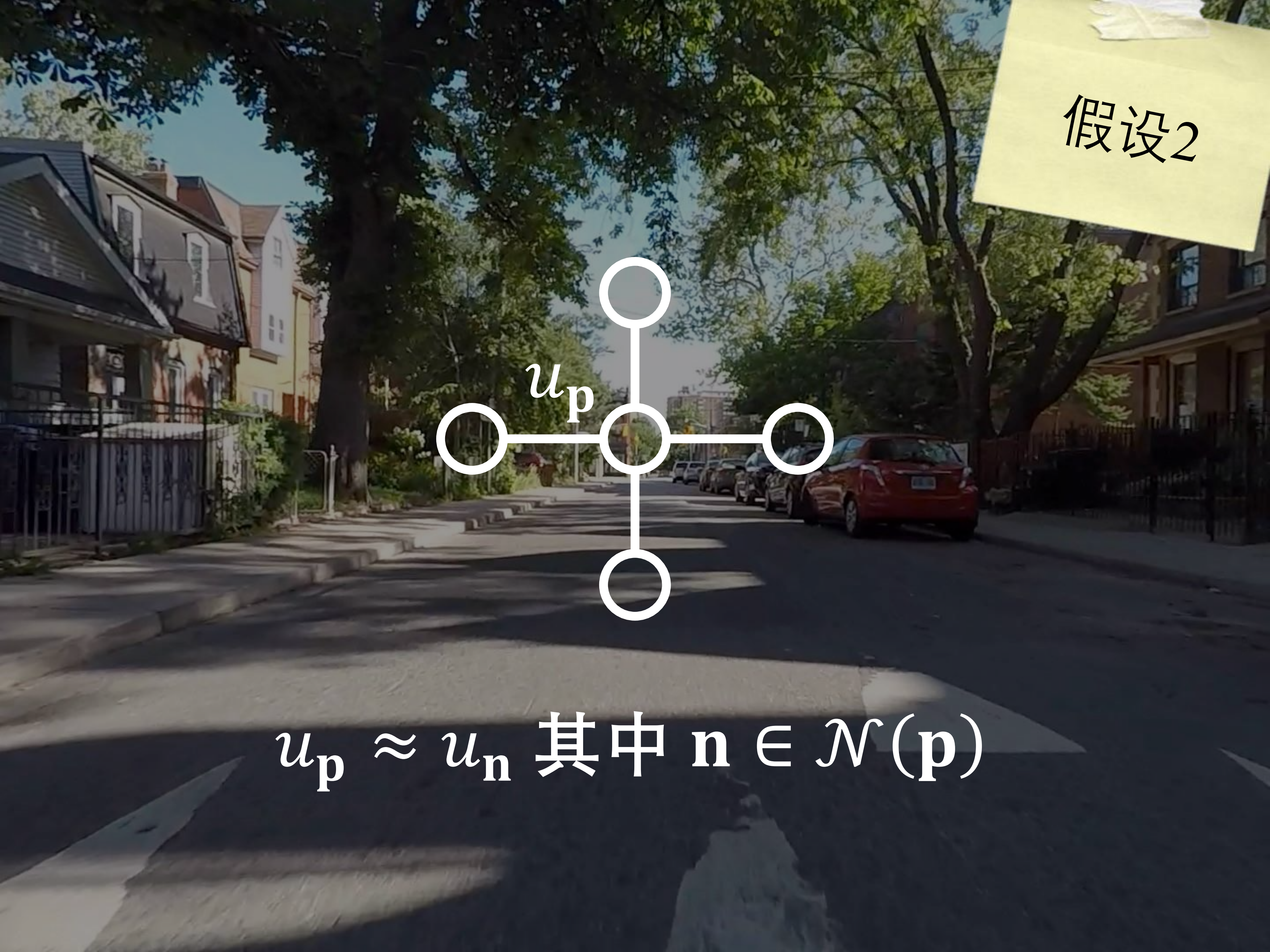
第$t+1$帧

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

外观恒常性假设


假设2


假设2
$u_{\mathbf{p}}$


假设2
$u_{\mathbf{p}}$
$u_{\mathbf{p}} \approx u_{\mathbf{n}}$ 其中 $\mathbf{n} \in \mathcal{N}(\mathbf{p})$


假设2
$u_{\mathbf{p}}$
$u_{\mathbf{p}} \approx u_{\mathbf{n}}$ 其中 $\mathbf{n} \in \mathcal{N}(\mathbf{p})$
相邻像素的速度变化缓慢

光度损失+平滑度损失

光度损失+平滑度损失

使图像对上速度场的损失最小化

Horn-Schunck
方法

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2 + \lambda \Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2 + \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \Big] \Big\}$$

Horn-Schunck 方法

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2 + \lambda \Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2 + \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \Big] \Big\}$$

每个像素都与其自身的速度变量相关联

$$\sum_{i,j}\Big\{\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2$$

$$+\lambda\Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} -\right.$$

数据项

$$\sum_{i,j}\left\{\left(I_x u_{i,j}+I_y v_{i,j}+I_t\right)^2\right.$$

$$+\lambda\left[\left(u_{i,j}-u_{i-1,j}\right)^2+\left(u_{i,j}-\right.\right.$$

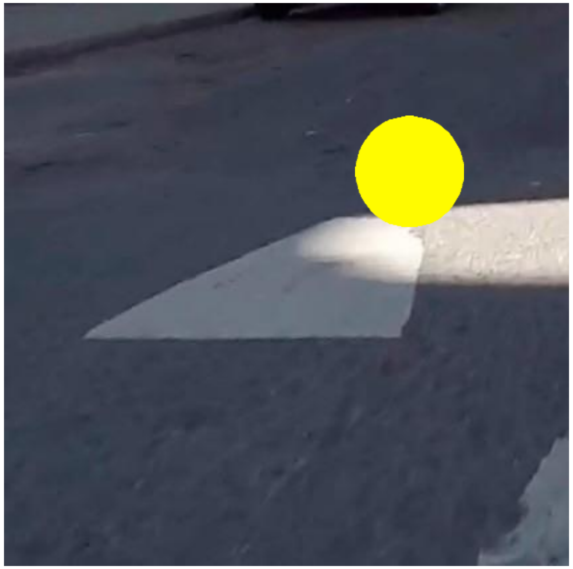
数据项

$$\sum\left\{\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2\right.$$

测量亮度恒常性的偏差

$$+\lambda\left[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} -\right.\right.$$

假设1



第t帧

第$t+1$帧

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

外观恒常性假设

数据项

$$\sum\left\{\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2\right.$$

测量亮度恒常性的偏离

$$+\lambda\left[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} -\right.\right.$$

$$
\mathrm{n} \sum_{i,j} \left\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t \right)^2 + \lambda \left[\left(u_{i,j} - u_{i-1,j} \right)^2 + \left(u_{i,j} - u_{i,j-1} \right)^2 + \left(v_{i,j} - v_{i+1,j} \right)^2 + \left(v_{i,j} - v_{i,j+1} \right)^2 \right] \right\}
$$

平滑项

$$
\text{n}\sum_{i,j}\Big\{\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2 + \lambda\Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2 + \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2\Big]\Big\}
$$


假设2
$u_{\mathbf{p}}$
$u_{\mathbf{p}} \approx u_{\mathbf{n}}$ 其中 $\mathbf{n} \in \mathcal{N}(\mathbf{p})$
相邻像素的速度变化缓慢

$$
\mathrm{n}\sum_{i,j}\Big\{\big(I_x u_{i,j} + I_y v \ldots I_t\big)^2
$$

平滑项

$$
+\lambda\Big[\big(u_{i,j} - u_{i-1,j}\big)^2 + \big(u_{i,j} - u_{i,j-1}\big)^2
$$

$$
+\big(v_{i,j} - v_{i+1,j}\big)^2 + \big(v_{i,j} - v_{i,j+1}\big)^2\Big]\Big\}
$$

$$\mathrm{n}\sum_{i,j}\left\{\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2\right.$$

平滑项

$$+\lambda\left[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2\right.$$

$$\left.\left.+\left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2\right]\right\}$$

测量相邻像素之间的流平滑度的偏差

$$
\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2
+ \lambda \Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2
+ \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \Big] \Big\}
$$

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \left\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t \right)^2 + \lambda \left[\left(u_{i,j} - u_{i-1,j} \right)^2 + \left(u_{i,j} - u_{i,j-1} \right)^2 + \left(v_{i,j} - v_{i+1,j} \right)^2 + \left(v_{i,j} - v_{i,j+1} \right)^2 \right] \right\}$$

相对权重


数据项


数据项
平滑项


数据项
平滑项

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \left\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2 + \lambda \left[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2 + \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \right] \right\}$$

相对权重

Horn-Schunck 方法

$$
\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2
+ \lambda \Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2
+ \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \Big] \Big\}
$$

Horn-Schunck 方法

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2 + \lambda \Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2 + \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \Big] \Big\}$$

怎么求解这个最优化问题?

Horn-Schunck 方法

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2 + \lambda \Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2 + \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \Big] \Big\}$$

对未知数求微分并令其等于零

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t\right)^2 + \lambda \Big[\left(u_{i,j} - u_{i-1,j}\right)^2 + \left(u_{i,j} - u_{i,j-1}\right)^2 + \left(v_{i,j} - v_{i+1,j}\right)^2 + \left(v_{i,j} - v_{i,j+1}\right)^2 \Big] \Big\}$$

对未知数求微分并令其等于零

$$\arg\min_{\mathbf{u},\mathbf{v}} \sum_{i,j} \Big\{ \left(I_x u_{i,j} + I_y v_{i,j} + I_t \right)^2 + \lambda \Big[\left(u_{i,j} - u_{i-1,j} \right)^2 + \left(u_{i,j} - u_{i,j-1} \right)^2 + \left(v_{i,j} - v_{i+1,j} \right)^2 + \left(v_{i,j} - v_{i,j+1} \right)^2 \Big] \Big\}$$

对未知数求微分并令其等于零

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t \right) I_x + \lambda \left[4u_{i,j} - \left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} \right) \right] = 0$$

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t \right) I_y + \lambda \left[4v_{i,j} - \left(v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1} \right) \right] = 0$$

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_x + \lambda\left[4u_{i,j} - \left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}\right)\right] = 0$$

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_y + \lambda\left[4v_{i,j} - \left(v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}\right)\right] = 0$$

$$(I_x u_{i,j} + I_y v_{i,j} + I_t) I_x + \lambda[4u_{i,j} - (u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1})] = 0$$

$$(I_x u_{i,j} + I_y v_{i,j} + I_t) I_y + \lambda[4v_{i,j} - (v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1})] = 0$$

每个图像点获得一对线性方程

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_x + \lambda\left[4u_{i,j} - \left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}\right)\right] = 0$$

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_y + \lambda\left[4v_{i,j} - \left(v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}\right)\right] = 0$$

方程组可以使用高斯消元法求解

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_x + \lambda\left[4u_{i,j} - \left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}\right)\right] = 0$$

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_y + \lambda\left[4v_{i,j} - \left(v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}\right)\right] = 0$$

方程组可以使用高斯消元法求解

理论上

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_x + \lambda\left[4u_{i,j} - \left(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}\right)\right] = 0$$

$$\left(I_x u_{i,j} + I_y v_{i,j} + I_t\right) I_y + \lambda\left[4v_{i,j} - \left(v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1}\right)\right] = 0$$

在实践中使用迭代法求解

数值分析

优点

优点

整合全局信息

优点

整合全局信息

密集流

缺点

缺点

亮度不恒常时失效

缺点

亮度不恒常时失效

缓慢的迭代估计

缺点

亮度不恒常时失效

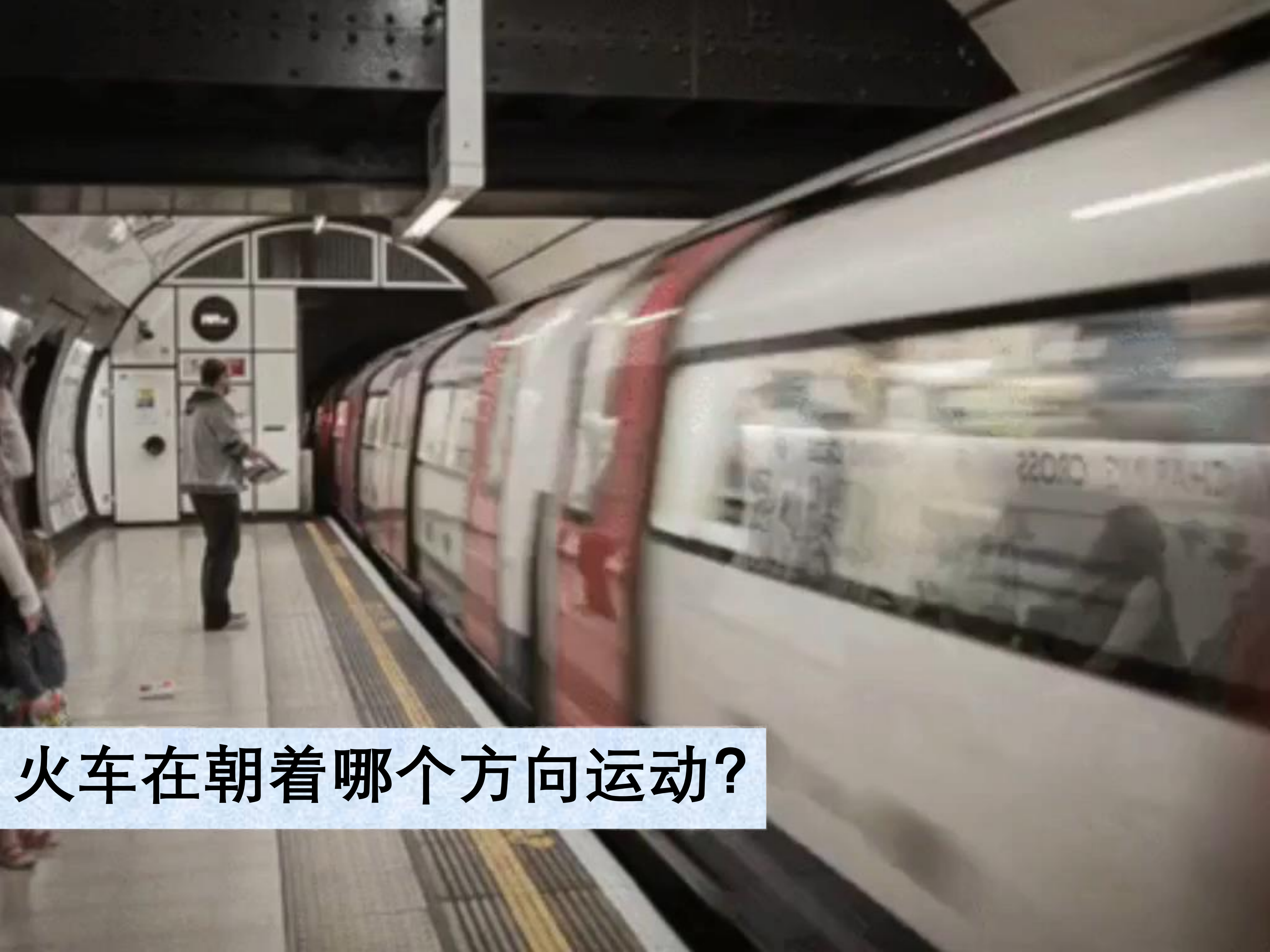
缓慢的迭代估计

平滑了边界

当速度较大时，输入在时间上混叠了

当速度较大时，输入在时间上混叠了

时间导数不可靠


火车在朝着哪个方向运动?

解决方案：由粗到细处理

$I(x, y, t)$

$I(x, y, t + 1)$

$I(x, y, t)$

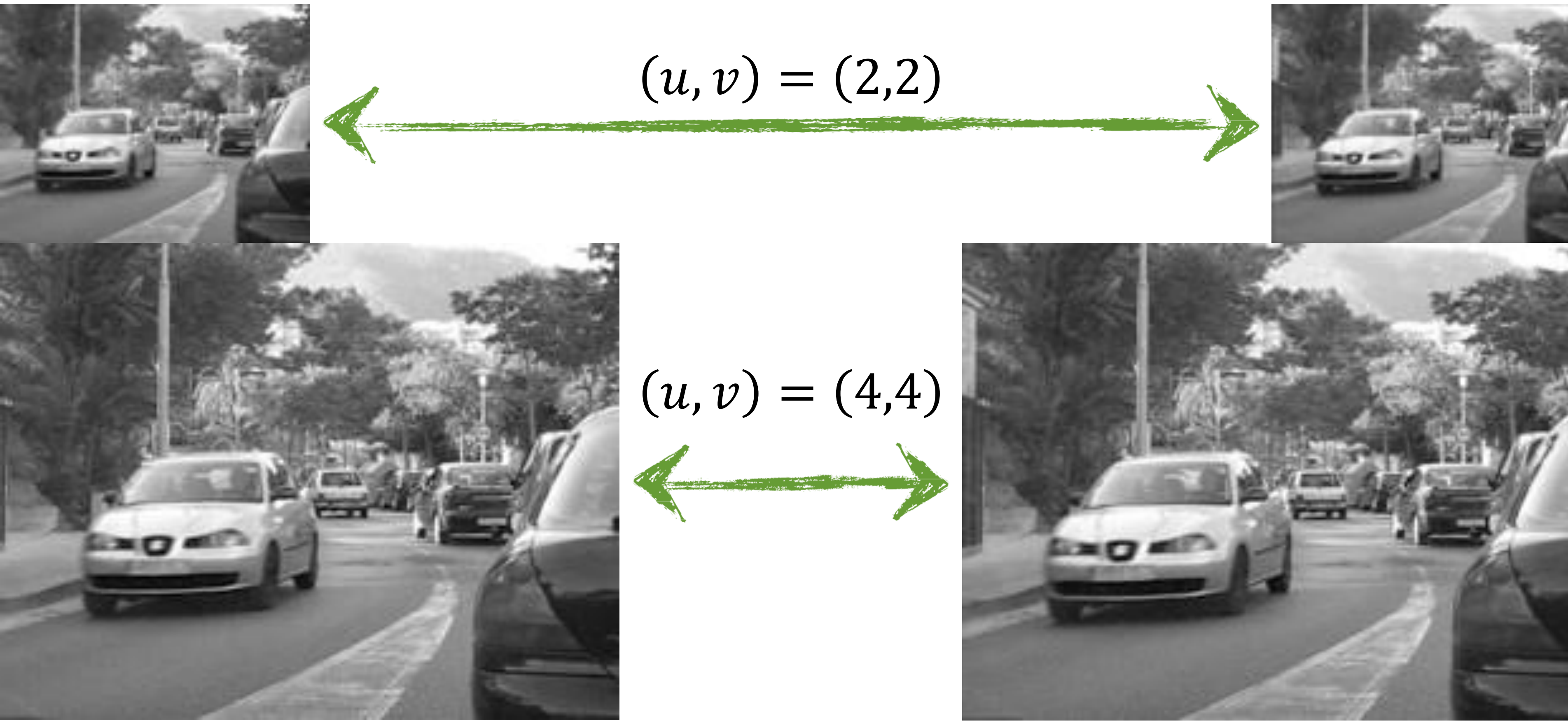
$(u, v) = (4,4)$

$I(x, y, t + 1)$

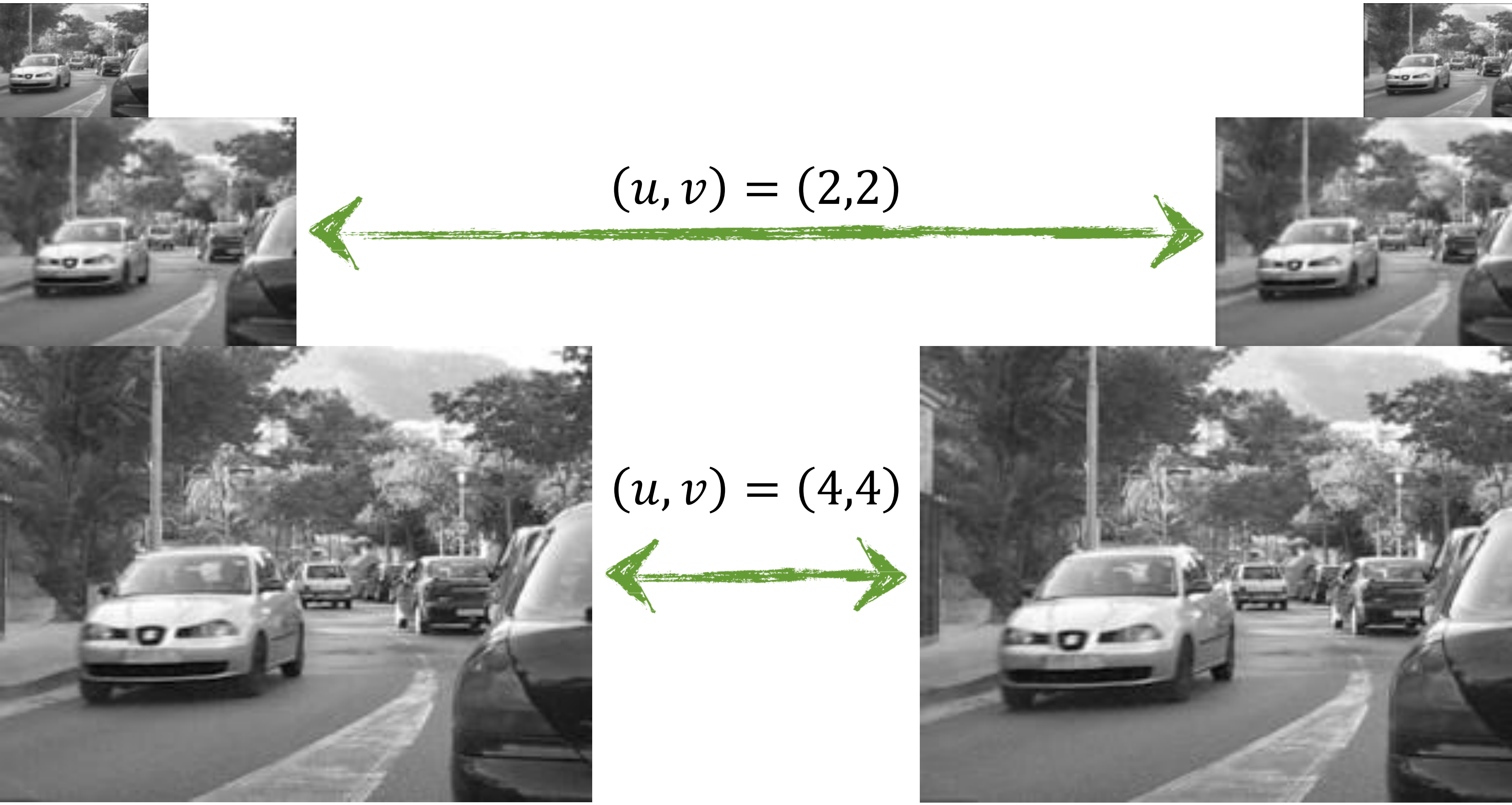
$I(x, y, t)$

$(u, v) = (4,4)$

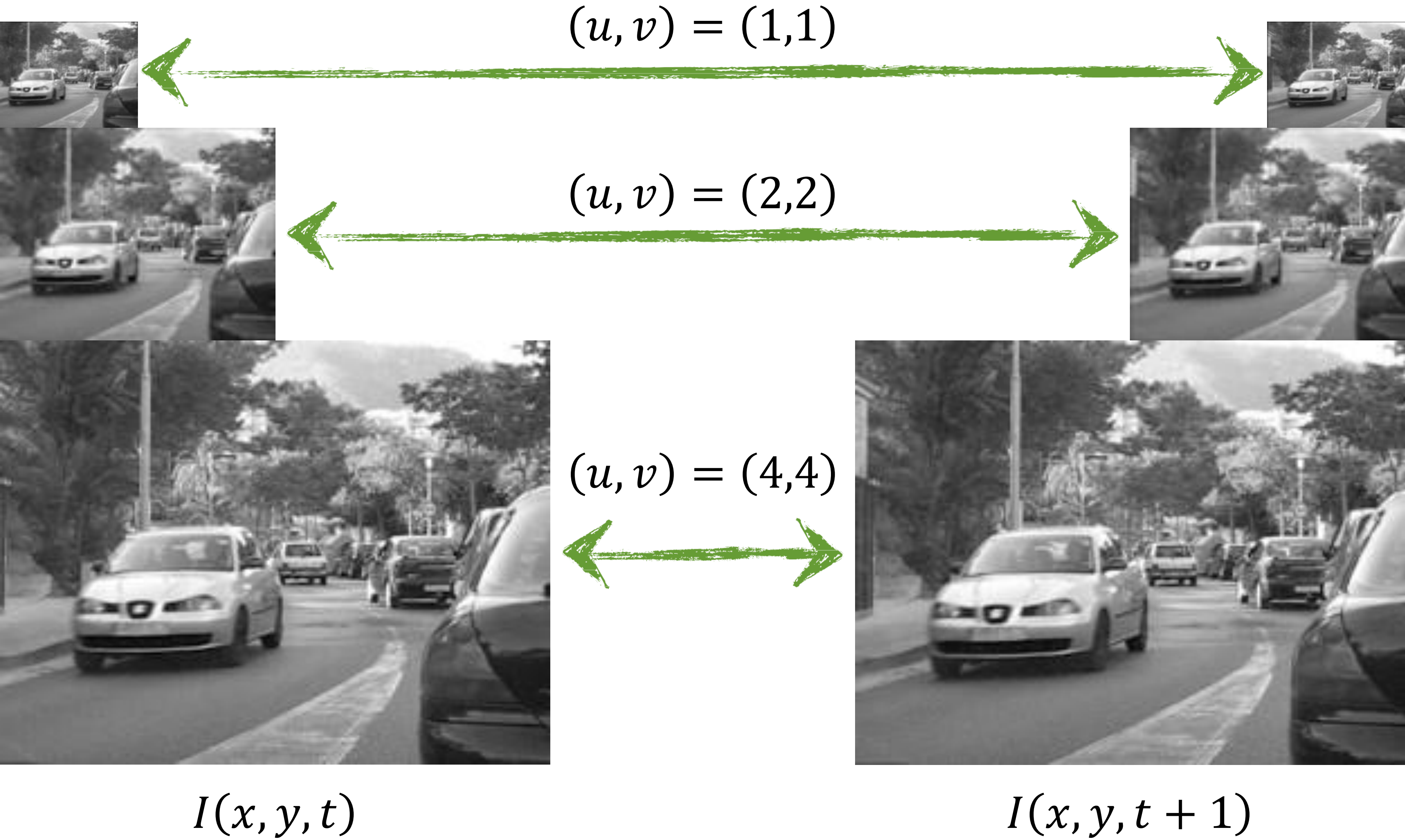
$I(x, y, t + 1)$



$I(x, y, t)$ $I(x, y, t + 1)$



$I(x, y, t)$ $I(x, y, t + 1)$



$(u, v) = (1,1)$
$(u, v) = (2,2)$
$(u, v) = (4,4)$
$I(x, y, t)$
$I(x, y, t + 1)$

$I(x, y, t)$

$I(x, y, t+1)$

估计流并传播

$I(x,y,t)$

$I(x,y,t+1)$

$I(x, y, t)$

$I(x, y, t + 1)$

转换、估计流和传播

$I(x, y, t)$

$I(x, y, t+1)$

$I(x, y, t)$

$I(x, y, t+1)$

转换、估计流和传播

A Quantitative Analysis of Current Practices in Optical Flow Estimation and the Principles behind Them

Deqing Sun · Stefan Roth · Michael J. Black

the date of receipt and acceptance should be inserted later

Abstract The accuracy of optical flow estimation algorithms has been improving steadily as evidenced by results on the Middlebury optical flow benchmark. The typical formulation, however, has changed little since the work of Horn and Schunck. We attempt to uncover what has made recent advances possible through a thorough analysis of how the objective function, the optimization method, and modern implementation practices influence accuracy. We discover that "classical" flow formulations perform surprisingly well

To take advantage of the trend towards video in wide-screen format, we further introduce an asymmetric pyramid downsampling scheme that enables the estimation of longer range horizontal motions. The methods are evaluated on Middlebury, MPI Sintel, and KITTI datasets using the same parameter settings.

Keywords Optical flow estimation · Practices · Median filtering · Non-local term · Motion boundary

International Journal of Computer Vision, 2014


THE
MATRIX

Raw Bullet Time Footage

视频重新着色

鸣谢： Michael Tao

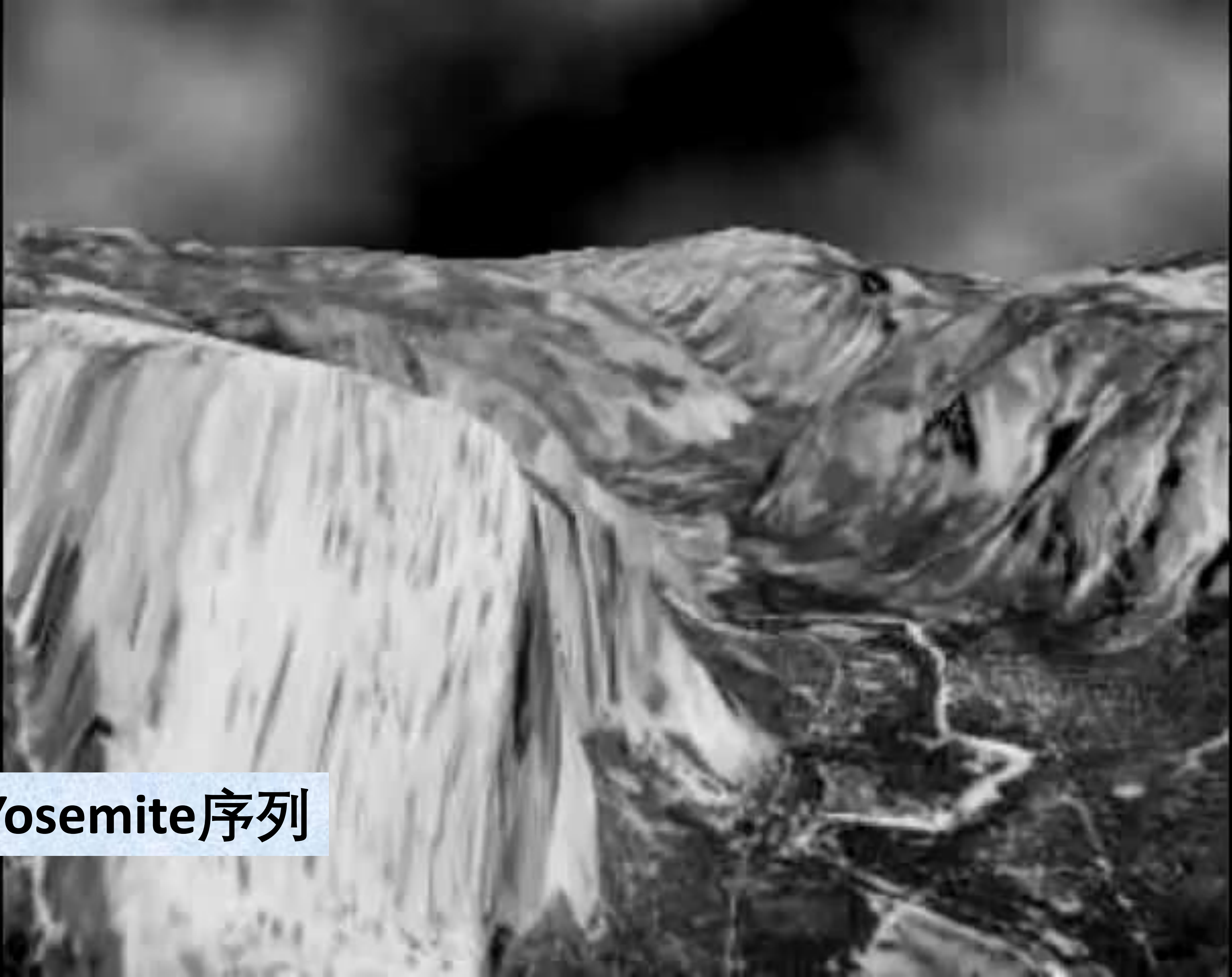
esults and Rankings

lts for methods appear here after users upload them and approve them for public display.

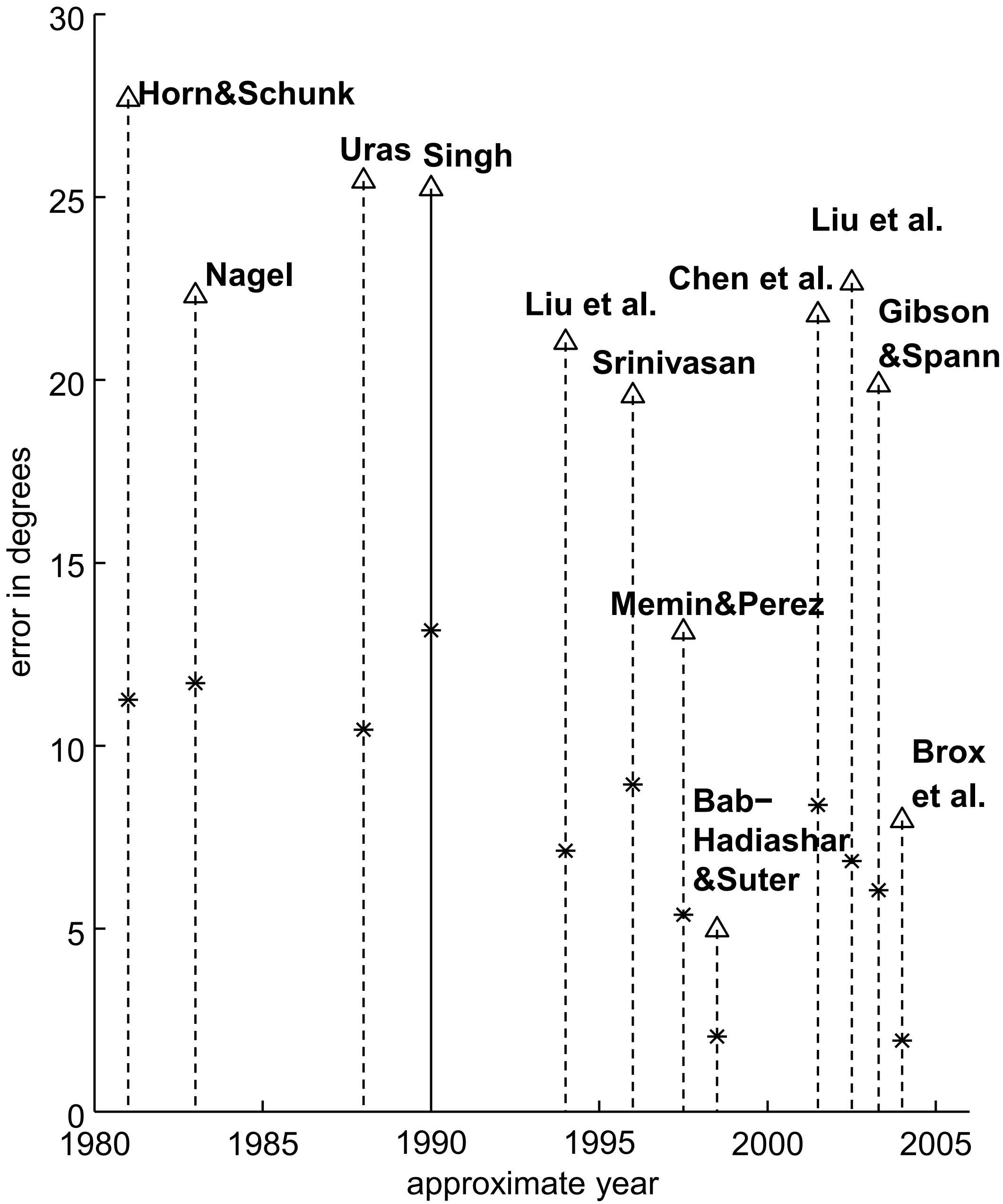
评价

运动估计

	EPE all	EPE matched	EPE unmatched	d0-10	d10-60	d60-140	s0-10	s10-40	s40+	
1]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	Visualize Results
	6.469	3.157	33.477	5.112	2.555	2.352	1.180	4.000	38.687	Visualize Results
[3]	6.986	3.590	34.631	5.520	3.486	2.821	1.499	3.911	41.622	Visualize Results
	7.212	3.336	38.781	5.650	3.144	2.208	1.284	4.107	44.118	Visualize Results
	7.249	2.973	42.088	4.896	2.817	2.218	1.159	4.183	44.866	Visualize Results
	7.617	3.690	39.613	5.509	3.583	3.010	1.791	4.067	45.128	Visualize Results
[7]	7.872	3.918	40.093	5.975	3.815	2.851	1.172	4.695	48.782	Visualize Results
F [8]	8.137	4.261	39.723	6.537	4.257	2.946	1.034	4.835	51.349	Visualize Results
sion [9]	8.231	4.274	40.460	6.221	4.252	3.193	1.702	5.701	46.696	Visualize Results
	8.287	4.165	41.905	6.345	4.127	2.996	1.312	5.122	50.540	Visualize Results
11]	8.291	4.287	40.925	6.520	4.265	2.984	1.208	5.090	51.162	Visualize Results
	8.377	4.286	41.695	6.556	4.024	3.323	1.834	4.955	49.083	Visualize Results



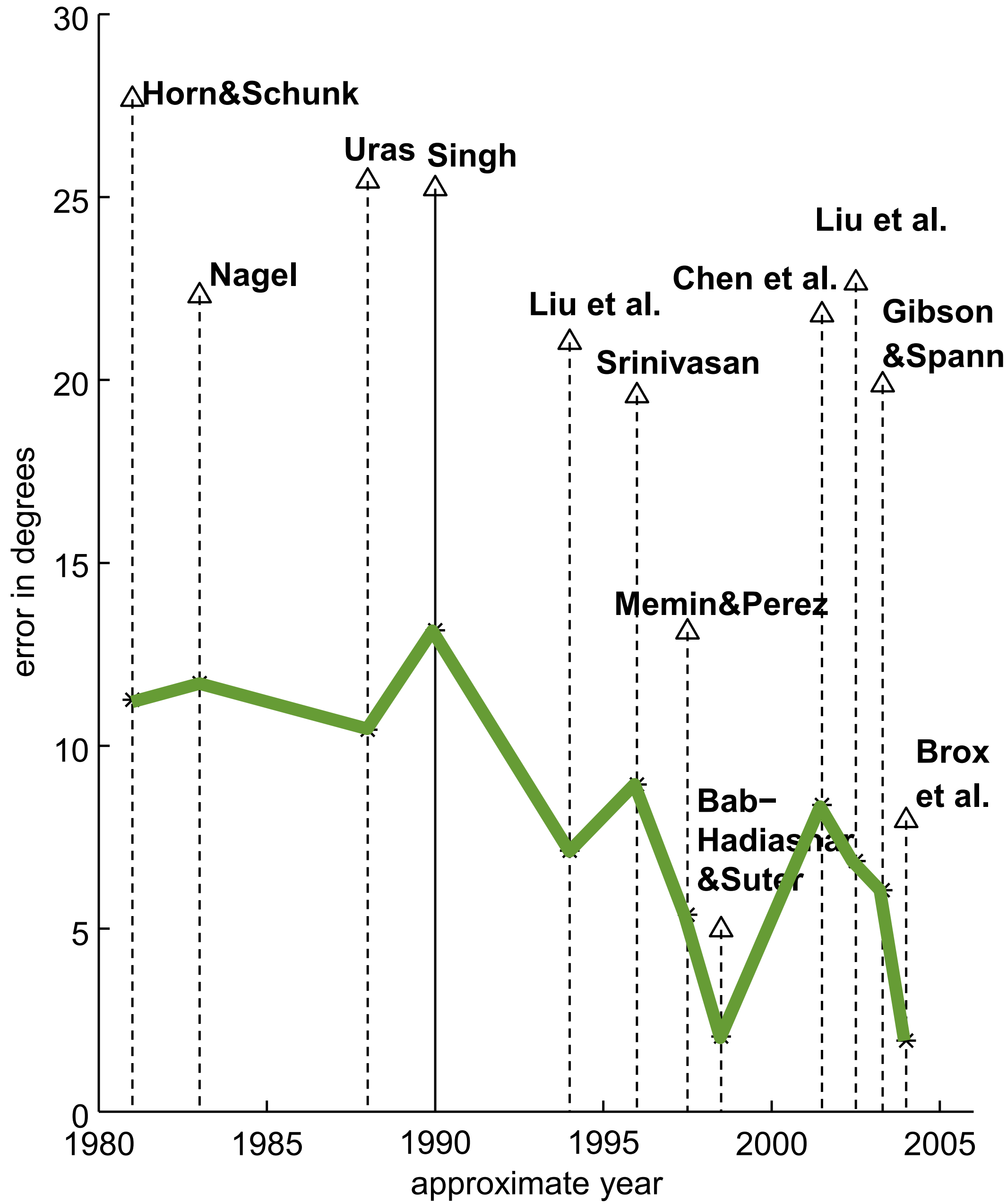
Yosemite序列



30
25
20
15
10
5
0
error in degrees
1980
1985
1990
1995
2000
2005
approximate year
Horn&Schunk
Nagel
Uras
Singh
Liu et al.
Srinivasan
Memin&Perez
Bab−
Hadiashar
&Suter
Chen et al.
Liu et al.
Gibson
&Spann
Brox
et al.

Yosemite
回顾

Yosemite 回顾



error in degrees
approximate year
30
25
20
15
10
5
0
1980
1985
1990
1995
2000
2005
Horn&Schunk
Nagel
Uras
Singh
Liu et al.
Srinivasan
Chen et al.
Liu et al.
Gibson
&Spann
Memin&Perez
Bab-
Hadiashar
&Suter
Brox
et al.

vision.middlebury.edu/flow/data

vision.middlebury.edu

stereo · mview · MRF · flow · color

Optical Flow — Evaluation · Datasets · Submit

Evaluation Datasets

With hidden ground-truth flow

	Hidden Texture				Synthetic			Stereo	High-speed camera (no GT)			
	Army	Mequon	Schefflera	Wooden	Grove	Urban	Yosemite	Teddy	Backyard	Basketball	Dumptruck	Evergreen
# frames	8	8	8	8	8	8	8	2	8	8	8	8
Flow Eval	yes	yes	yes	yes	yes	yes	yes	yes	--	--	--	--
Interp Eval	--	yes	yes	--	--	yes	--	yes	yes	yes	yes	yes

Download:	Color	Gray
All frames (usually 8)	eval-color-allframes.zip (31MB)	eval-gray-allframes.zip (12MB)
Two frames only	eval-color-twoframes.zip (8MB)	eval-gray-twoframes.zip (3MB)

Instructions for participating in the evaluation can be found on the Submit page.

Other Datasets

With public ground-truth flow. These can be used for training.

	Hidden Texture			Synthetic				Stereo	High-speed camera (no GT)			
	RubberWhale	Hydrangea	Dimetrodon	Grove2	Grove3	Urban2	Urban3	Venus	Beanbags	DogDance	MiniCooper	Walking
# frames	8	8	2	8	8	8	8	2	8	8	8	8

Download:	Color	Gray
All frames (usually 8)	other-color-allframes.zip (32MB)	other-gray-allframes.zip (12MB)
Two frames only	other-color-twoframes.zip (9MB)	other-gray-twoframes.zip (3.4MB)
Ground-truth flow	other-gt-flow.zip (13MB)	
Ground-truth interpolation	other-gt-interp.zip (4.5MB)	other-gt-interp-gray.zip (1.7MB)

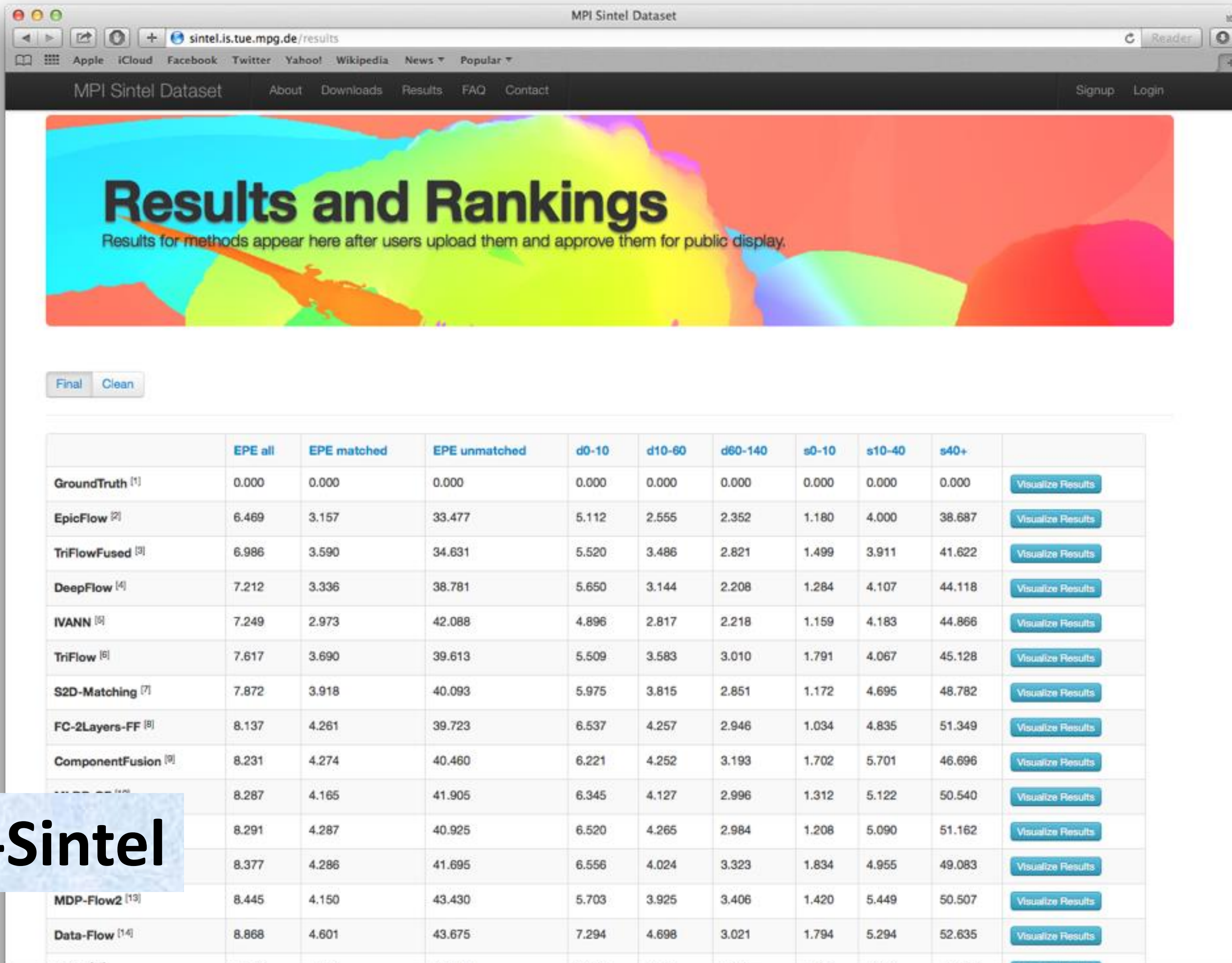
sion is available in flow-code-matlab.zip.

l Images, Pixar, Doug Creel, and Luca Fascione for their help with the synthetic data.

明德大学（米德尔伯里学院）


橡胶鲸


异齿龙



MPI Sintel Dataset

sintel.is.tue.mpg.de/results

MPI Sintel Dataset | About | Downloads | Results | FAQ | Contact | Signup | Login

Results and Rankings

Results for methods appear here after users upload them and approve them for public display.

Final | Clean

	EPE all	EPE matched	EPE unmatched	d0-10	d10-60	d60-140	s0-10	s10-40	s40+	
GroundTruth [1]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	Visualize Results
EpicFlow [2]	6.469	3.157	33.477	5.112	2.555	2.352	1.180	4.000	38.687	Visualize Results
TriFlowFused [3]	6.986	3.590	34.631	5.520	3.486	2.821	1.499	3.911	41.622	Visualize Results
DeepFlow [4]	7.212	3.336	38.781	5.650	3.144	2.208	1.284	4.107	44.118	Visualize Results
IVANN [5]	7.249	2.973	42.088	4.896	2.817	2.218	1.159	4.183	44.866	Visualize Results
TriFlow [6]	7.617	3.690	39.613	5.509	3.583	3.010	1.791	4.067	45.128	Visualize Results
S2D-Matching [7]	7.872	3.918	40.093	5.975	3.815	2.851	1.172	4.695	48.782	Visualize Results
FC-2Layers-FF [8]	8.137	4.261	39.723	6.537	4.257	2.946	1.034	4.835	51.349	Visualize Results
ComponentFusion [9]	8.231	4.274	40.460	6.221	4.252	3.193	1.702	5.701	46.696	Visualize Results
[illegible] [10]	8.287	4.165	41.905	6.345	4.127	2.996	1.312	5.122	50.540	Visualize Results
[illegible]	8.291	4.287	40.925	6.520	4.265	2.984	1.208	5.090	51.162	Visualize Results
[illegible]	8.377	4.286	41.695	6.556	4.024	3.323	1.834	4.955	49.083	Visualize Results
MDP-Flow2 [13]	8.445	4.150	43.430	5.703	3.925	3.406	1.420	5.449	50.507	Visualize Results
Data-Flow [14]	8.868	4.601	43.675	7.294	4.698	3.021	1.794	5.294	52.635	Visualize Results
LDOF [15]	9.116	5.037	42.344	6.849	4.928	4.003	1.485	4.839	57.296	Visualize Results

MPI-Sintel

MPI-Sintel

大位移很难估计


为什么分层流估计不充分呢?

Large Displacement Optical Flow*

Thomas Brox[1] Christoph Bregler[2] Jitendra Malik[1]

[1]University of California, Berkeley
Berkeley, CA, 94720, USA
{brox,malik}@eecs.berkeley.edu

[2]Courant Institute, New York University
New York, NY, 10003, USA
bregler@courant.nyu.edu

Abstract

The literature currently provides two ways to establish point correspondences between images with moving objects. On one side, there are energy minimization methods that yield very accurate, dense flow fields, but fail as displacements get too large. On the other side, there is descriptor matching that allows for large displacements, but correspondences are very sparse, have limited accuracy, and due to missing regularity constraints there are many outliers. In this paper we propose a method that can combine the advantages of both matching strategies. A region hierarchy is established for both images. Descriptor matching on these regions provides a sparse set of hypotheses for correspondences. These are integrated into a variational approach

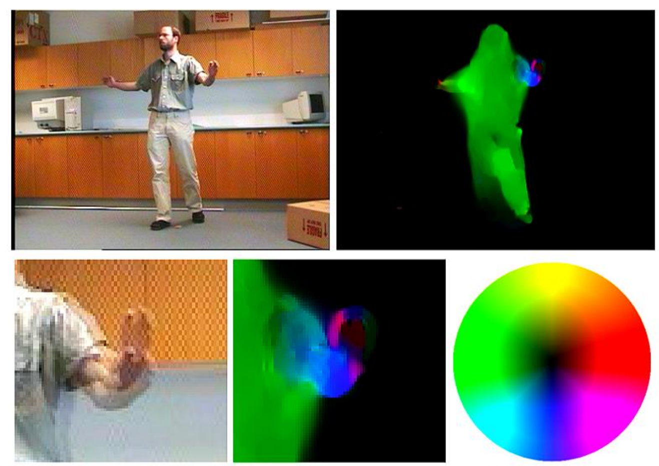
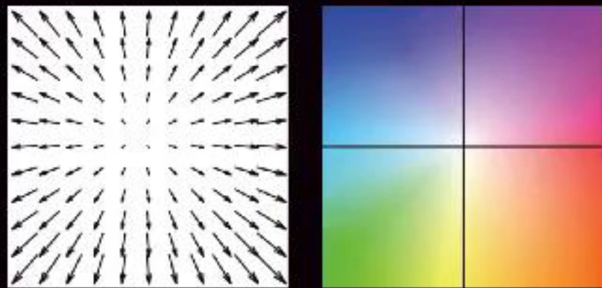


Figure 1. **Top row:** Image of a sequence where the person is stepping forward and moving his hands. The optical flow estimated

CVPR, 2009



鸣谢：Thomas Bros

深度学习

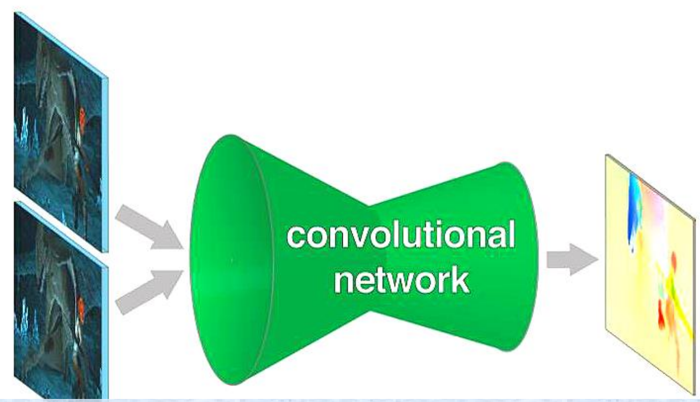
FlowNet: Learning Optical Flow with Convolutional Networks

Philipp Fischer*‡, Alexey Dosovitskiy‡, Eddy Ilg‡, Philip Häusser, Caner Hazırbaş, Vladimir Golkov*
University of Freiburg — Technical University of Munich
{fischer,dosovits,ilg}@cs.uni-freiburg.de, {haeusser,hazirbas,golkov}@cs.tum.edu

Patrick van der Smagt, Technical University of Munich, smagt@brml.org
Daniel Cremers, Technical University of Munich, cremers@tum.de
Thomas Brox, University of Freiburg, brox@cs.uni-freiburg.de

Abstract

Convolutional neural networks (CNNs) have recently been very successful in a variety of computer vision tasks, especially on those linked to recognition. Optical flow estimation has not been among the tasks where CNNs were successful. In this paper we construct appropriate CNNs which are capable of solving the optical flow estimation problem as a supervised learning task. We propose and compare two architectures: a generic architecture and another one including a layer that correlates feature vectors at different image locations.



ICCV, 2015

FlowNet 2.0 vs FlowFields

FlowNet 2.0 generates sharper boundaries,
achieves comparable error scores,
and runs ca. **200x faster**

“我们是瘾君子！注释数据是我们的**海洛因**。”

——吉滕德拉·马利克（加利福尼亚大学伯克利分校）

Back to Basics: Unsupervised Learning of Optical Flow via Brightness Constancy and Motion Smoothness

Jason J. Yu, Adam W. Harley and Konstantinos G. Derpanis
Department of Computer Science
Ryerson University, Toronto, Canada
{jjyu, aharley, kosta}@scs.ryerson.ca

Abstract

Recently, convolutional networks (convnets) have proven useful for predicting optical flow. Much of this success is predicated on the availability of large datasets that require expensive and involved data acquisition and laborious labeling. To bypass these challenges, we propose an unsupervised approach (i.e., without leveraging groundtruth flow) to train a convnet end-to-end for predicting optical flow between two images. We use a loss function that combines a data term that measures photometric constancy over time with a spatial term that models the expected variation of flow across the image. Together these losses form a proxy measure for losses based on the groundtruth flow. Empirically, we show that a strong convnet baseline trained with

flow for training, we use the images alone. In particular, we use a loss function that combines a data term that measures photometric constancy over time with a spatial term that models the expected variation of flow across the image. The photometric loss measures the difference between the first input image and the (inverse) warped subsequent image based on the predicted optical flow by the network. The smoothness loss measures the difference between spatially neighbouring flow predictions. Together, these two losses form a proxy for losses based on the groundtruth flow.

Recovering optical flow between two frames is a well studied problem, with much previous work founded on variational formulations [7, 2, 13, 12]. Our loss is similar to the objective functions proposed for two-frame motion es-

ECCV Workshops, 2016


1080-60
03
HERO4


无限的数据

光流不适用于所有时间图像

超出光流范围


PARK
OPEN
24
HOURS

Spacetime Texture Representation and Recognition Based on a Spatiotemporal Orientation Analysis

Konstantinos G. Derpanis, *Member*, *IEEE*, and Richard P. Wildes, *Member*, *IEEE*

Abstract—This paper is concerned with the representation and recognition of the observed dynamics (i.e., excluding purely spatial appearance cues) of spacetime texture based on a spatiotemporal orientation analysis. The term "spacetime texture" is taken to refer to patterns in visual spacetime, (x, y, t), that primarily are characterized by the aggregate dynamic properties of elements or local measurements accumulated over a region of spatiotemporal support, rather than in terms of the dynamics of individual constituents. Examples include image sequences of natural processes that exhibit stochastic dynamics (e.g., fire, water, and windblown vegetation) as well as images of simpler dynamics when analyzed in terms of aggregate region properties (e.g., uniform motion of elements in imagery, such as pedestrians and vehicular traffic). Spacetime texture representation and recognition is important as it provides an early means of capturing the structure of an ensuing image stream in a meaningful fashion. Toward such ends, a novel approach to spacetime texture representation and an associated recognition method are described based on distributions (histograms) of spacetime orientation structure. Empirical evaluation on both standard and original image data sets shows the promise of the approach, including significant improvement over alternative state-of-the-art approaches in recognizing the same pattern from different viewpoints.

Index Terms—Spacetime texture, image motion, dynamic texture, temporal texture, time-varying texture, textured motion, turbulent flow, stochastic dynamics, distributed representation, spatiotemporal orientation.

TPAMI, 2012