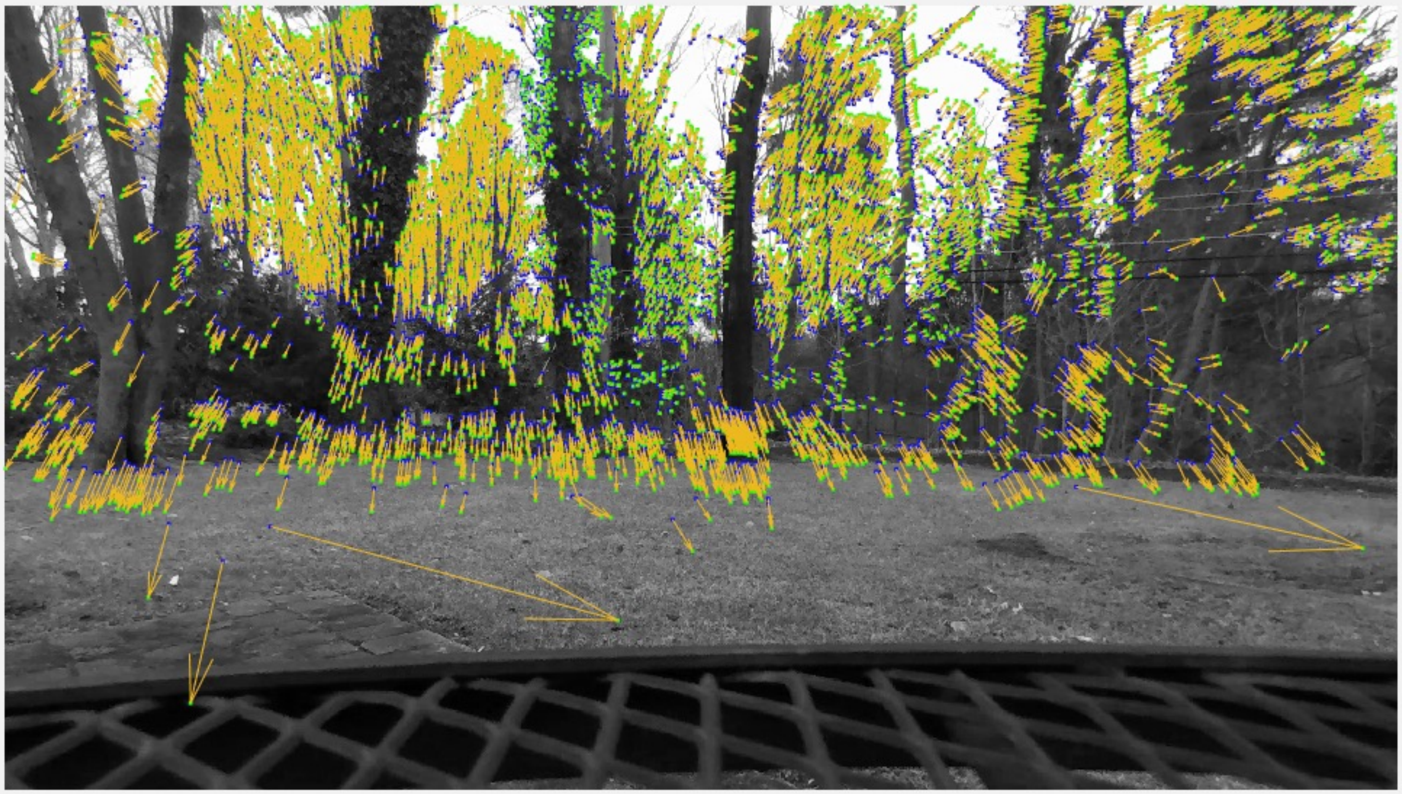
Optical Flow

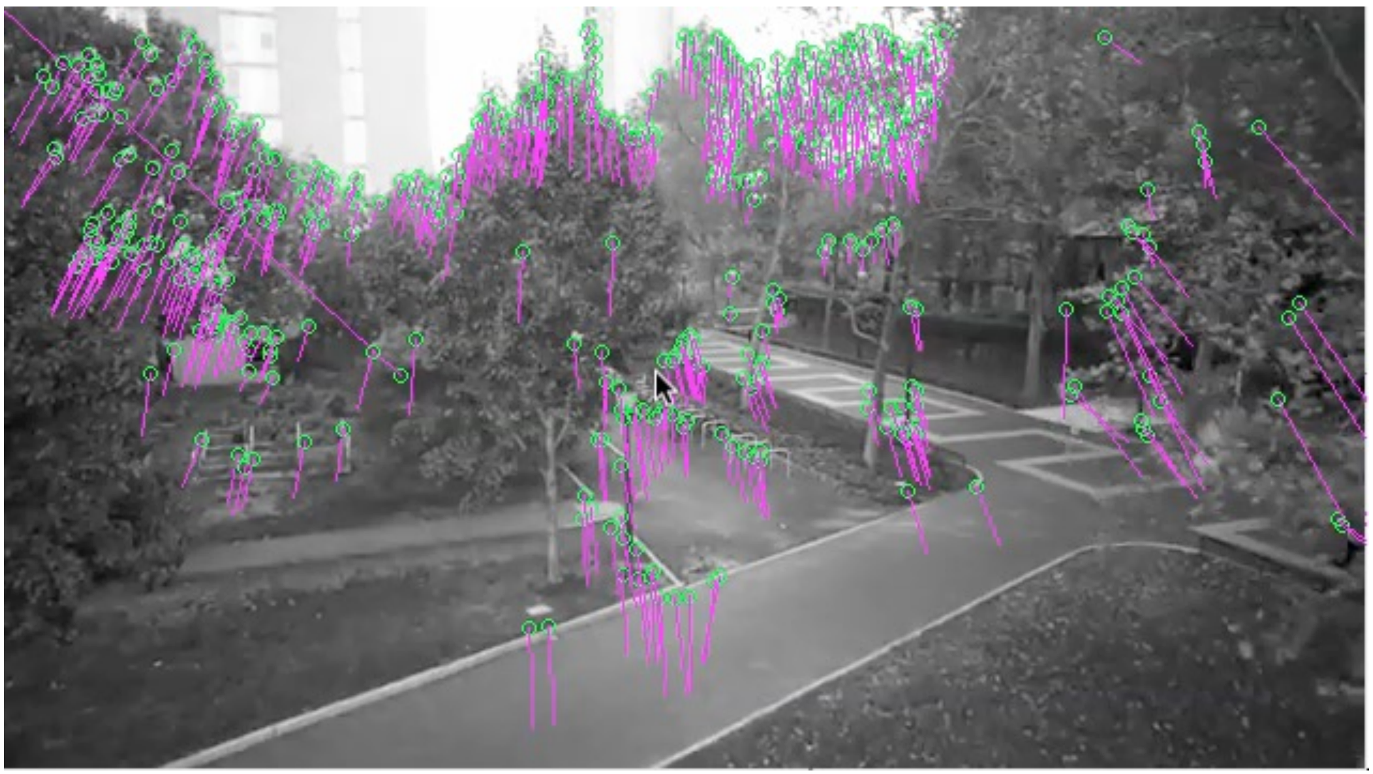
Kostas Daniilidis

A robot is always in motion

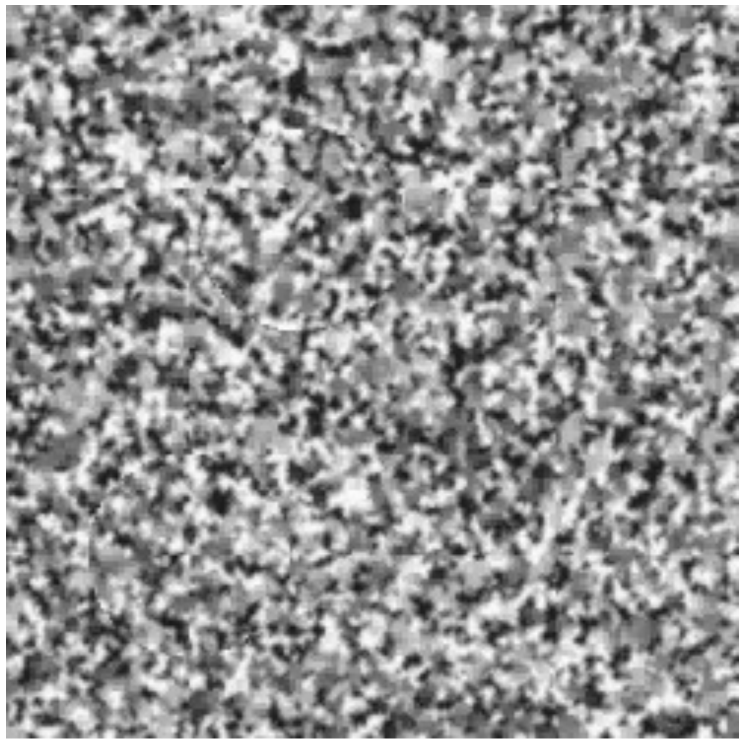
Optical Flow

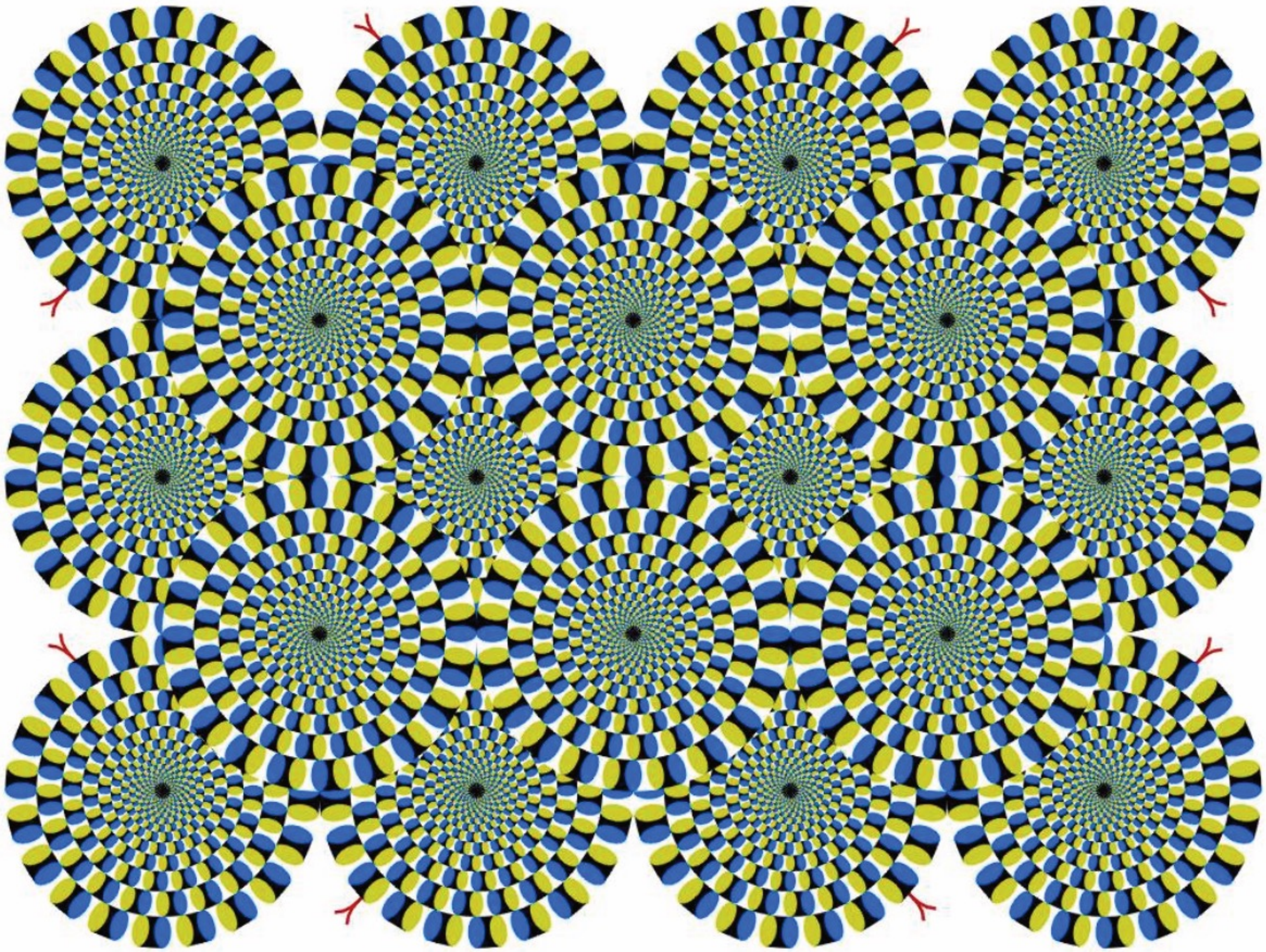


This is what we are up to !!



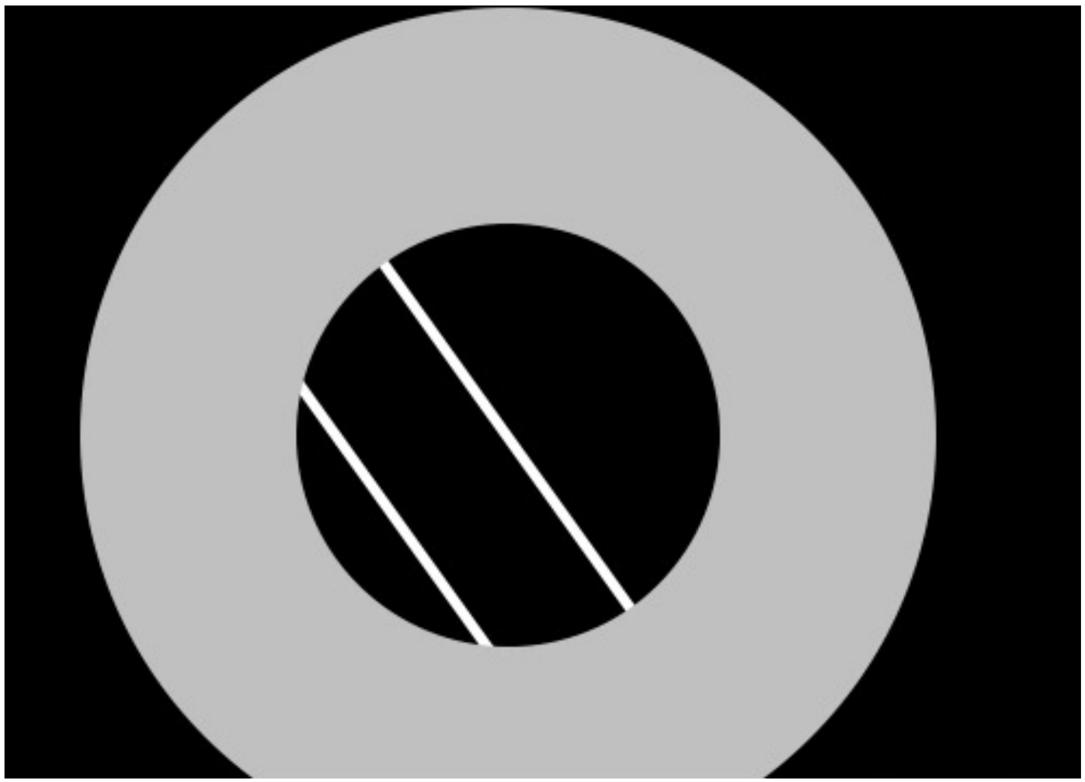
Segmentation cue





Ouchi

Aperture Problem

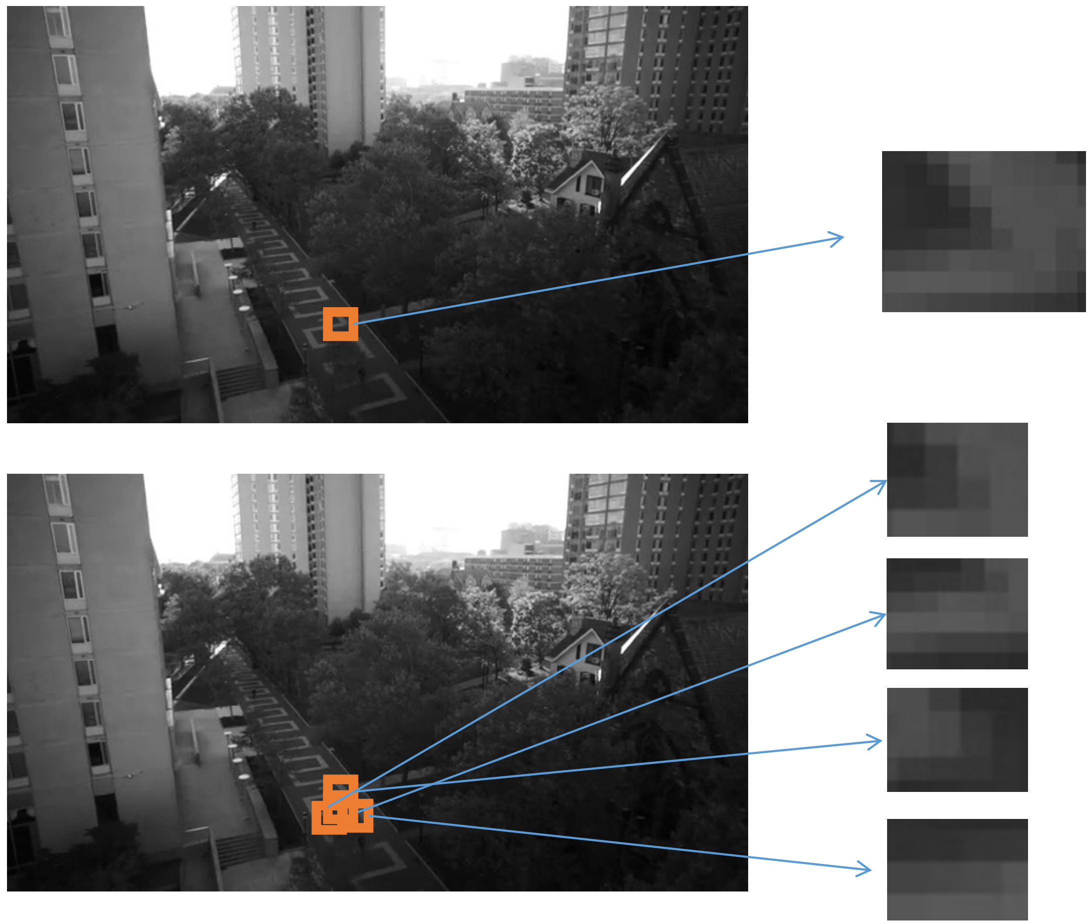


Barber Pole Illusion

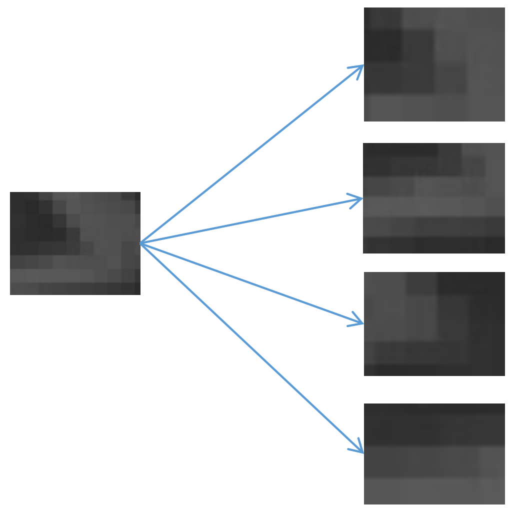
Temporal aliasing

Matching cross consecutive frames

Matching cross consecutive frames



Correspondence as local search

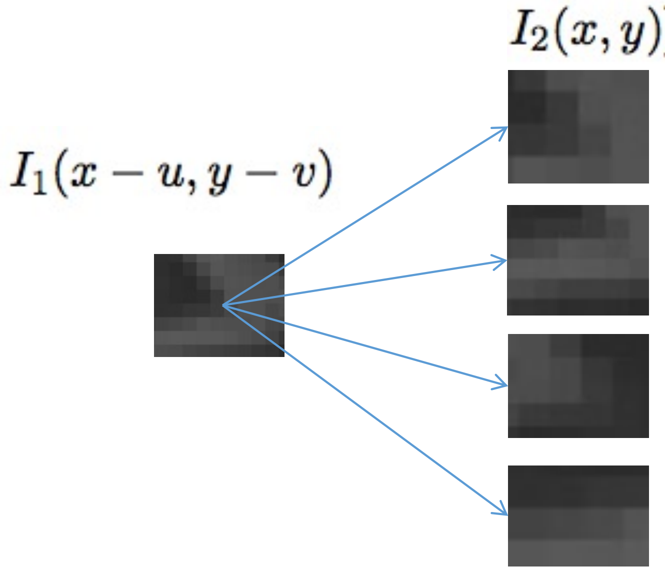


Which patch in the same neighborhood at the next frame looks more similar to the path in the first frame?

Challenges:

- Intensity remains the same (illumination, Lambertian surface)
- Geometric deformation (scaling, rotation) minimal
- Patch might get occluded
- Patch might be displaced very far

Brightness Constancy

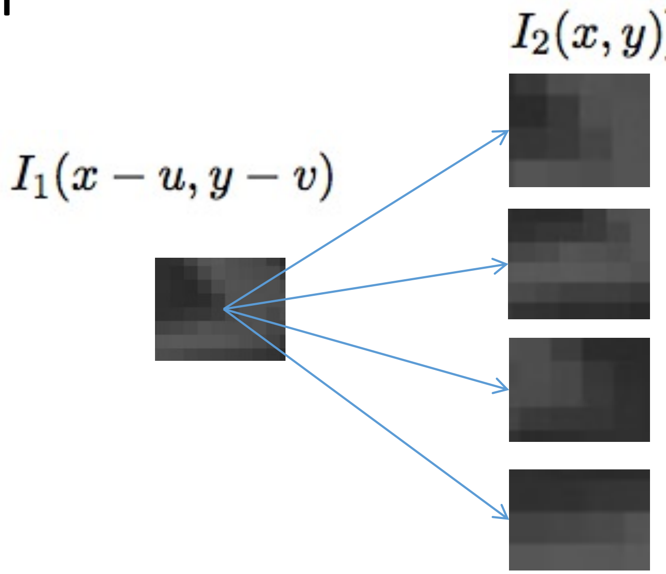


Which patch in the same neighborhood at the next frame looks more similar to the path in the first frame?

Brightness (intensity) remains the same if image is shifted by vector (u, v):
Brightness Change Constraint Equation

$$I_1(x-u, y-v) = I_2(x, y)$$

Correspondence as local search



Which patch in the same neighborhood at the next frame looks more similar to the path in the first frame?

Assuming that a patch remains approximately the same we seek to minimize

$$(u, v) = \arg\min_{(u,v)} \varepsilon(u, v)$$

where

$$\varepsilon(u, v) = \int\int_{(x,y)\in\aleph} [I_1(x-u, y-v) - I_2(x, y)]^2 \, dx \, dy$$

Possible solution: Exhaustive search!

If we regard $\varepsilon(u, v)$ as a continuous function of (u, v) then calculus of multiple variables says that a minimum exists at (u, v) iff
the gradient vanishes

$$\frac{\partial \varepsilon}{\partial u}(u, v) = 0$$

$$\frac{\partial \varepsilon}{\partial v}(u, v) = 0$$

and the Hessian is positive definite

$$\mathrm{Hessian}(\varepsilon) = \begin{pmatrix} \frac{\partial^2 \varepsilon}{\partial u^2} & \frac{\partial^2 \varepsilon}{\partial u \partial v} \\ \frac{\partial^2 \varepsilon}{\partial u \partial v} & \frac{\partial^2 \varepsilon}{\partial v^2} \end{pmatrix}$$

Let's simplify our notation using vectors for optical flow

$$\vec{d} = \begin{pmatrix} u \\ v \end{pmatrix}$$

image position

$$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

and the brightness difference:

$$E(\vec{d}) = I_1(\vec{x} - \vec{d})\text{-}I_2(\vec{x})$$

Then minimization reads as

$$\arg\min_{\vec{d}} \int\int E(\vec{d})^2 \, dx \, dy$$

How can we get rid of the exhaustive search for $(u, v) = \vec{d}$?
By taking the Taylor expansion of the error function $E(\vec{d})$:

$$\begin{aligned} E(\vec{d}) &= I_1(\vec{x} - \vec{d}) - I_2(\vec{x}) \\ &= I_1(\vec{x})\text{-}\nabla I_1(\vec{x})^T \vec{d} - I_2(\vec{x}) \\ &= -\Delta I(\vec{x})\text{-}\nabla I_1(\vec{x})^T \vec{d} \end{aligned}$$

where $\Delta I = I_2(x, y)\text{-}I_1(x, y)$ is the difference between two pixels intensities at the exact same location in both images
and ∇I_1 is the spatial gradient of I_1.

Going back to the minimization problem, we seek $\vec{d}$ such that

$$\frac{\partial \varepsilon(\vec{d})}{\partial \vec{d}} = \int\int 2E(\vec{d})\frac{\partial E}{\partial \vec{d}}\,dx\,dy = 0$$

we replace E with each Taylor expansion

$$\begin{aligned}\frac{\partial \varepsilon(\vec{d})}{\partial \vec{d}} &= 2\int\int (\Delta I(\vec{x}) + \nabla I(\vec{x})^T\vec{d})\nabla I(\vec{x})\,dx\,dy \\ &= 0\end{aligned}$$

This results

$$(\int\int \nabla I(\vec{x})\nabla I(\vec{x})^T\,dx\,dy)\vec{d} = -\int\int \Delta I(\vec{x})\nabla I(\vec{x})\,dx\,dy$$

It is easy to prove that the Hessian of $\varepsilon(x,y)$ is the matrix at the left hand side which is positive semidefinite.

Let us discretize the integral for the sake of implementation

$$(\sum_{i\in\aleph} \nabla I_i \nabla I_i^T)\vec{d} = -\sum_{i\in\aleph} \Delta I_i \nabla I_i$$

or

$$\sum_{i\in\aleph} \begin{pmatrix} I_{x,i}^2 & I_{x,i}I_{y,i} \\ I_{x,i}I_{y,i} & I_{y,i}^2 \end{pmatrix} = \sum_{i\in\aleph} \begin{pmatrix} \Delta I_i \, I_{x_i} \\ \Delta I_i \, I_{y_i} \end{pmatrix}$$

This is a 2x2 linear system that has to be solved for each pixel!

Existence and uniqueness

$$M\vec{d} = \vec{m}$$

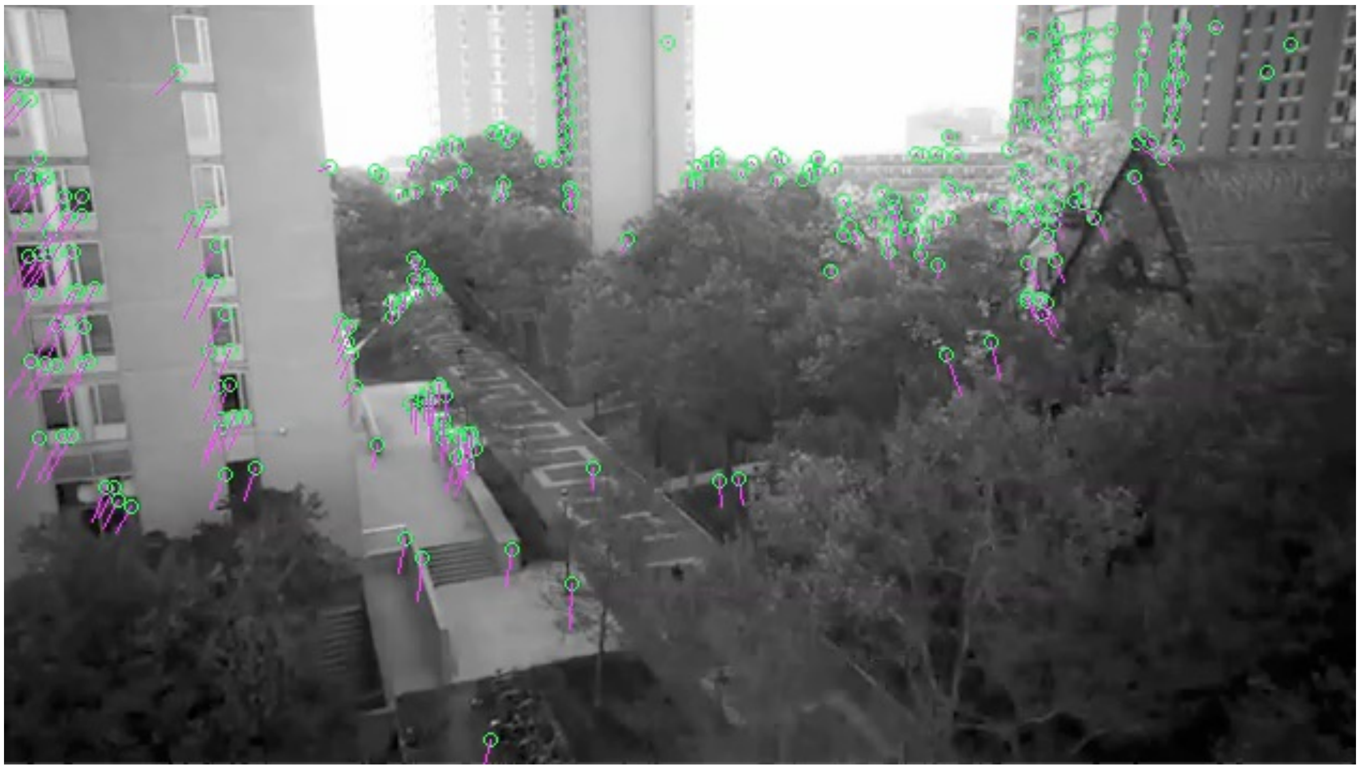
where M is a 2 by 2 matrix and $\vec{m}$ is a 2 by 1 vector.

$(M, \vec{m})$ must have a rank of 2 i.e. $\vec{m}$ belong to the range of M for a solution to *exist*.

A solution will be *unique* iff rank$(M) = 2$ or $det(M) \neq 0$ or $\lambda_{min} > 0$.
We now realize that uniqueness is the same as cornerness:
The matrix

$$\sum_{i \in \aleph} \begin{pmatrix} I_{x,i}^2 & I_{x,i}I_{y,i} \\ I_{x,i}I_{y,i} & I_{y,i}^2 \end{pmatrix}$$

to be inverted is the same as the Harris corner matrix!



Extensions

- Apply several iterations to find (u,v)
- Apply the Taylor expansion symmetrically with respect to the two images
- To capture larger motions use an image pyramid
- Assume the optical flow is an affine transformation not just a translation

Proceedings DARPA Image Understanding Workshop, April 1981, pp. 121-130

essentially the same but shorter version of this paper was presented and included in the Proc 7th Intl Joint Conf on Artificial Intelligence (IJCAI) 1981, August 24-28, Vancouver, British Columbia, pp.674-679.
when you refer to the work, please refer to the IJCAI paper.

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

Shape and Motion from Image Streams: a Factorization Method—Part 3

Detection and Tracking of Point Features

Technical Report CMU-CS-91-132

Carlo Tomasi Takeo Kanade

April 1991

Good Features to Track

Jianbo Shi
Computer Science Department
Cornell University
Ithaca, NY 14853

Carlo Tomasi
Computer Science Department
Stanford University
Stanford, CA 94305

Optical Flow in the Fourier Domain

Consider an xyt-signal (video), in which we keep only one line (fixed y) of each frame. The effect of a translation in the video can be seen as a *wave* moving along the line at fixed y at constant speed u (in px/frame):

$$f(x,t) = f_0(x - ut)$$

Example: $f_0(x) = \cos(\omega_0 x)$, and $f(x,t) = \cos(\omega_0(x - ut))$
The Fourier transform of f is the following:

$$F(\omega_x, \omega_t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_0(x - ut)e^{-j\omega_x x + \omega_t t}dxdt$$

If we note $f_0(x) \circ\!\!-\!\!\bullet\, F_0(\omega_x)$ and make the substitution $x' = x - ut$, we have

$$
\begin{aligned}
F(\omega_x, \omega_t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_0(x') e^{-j\omega_x(x'+ut)+\omega_t t} dx' dt \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_0(x') e^{-j\omega_x x'} e^{-j(\omega_x u+\omega_t)t} dx' dt \\
&= F_0(\omega_x)\delta(\omega_x u + \omega_t)
\end{aligned}
$$

Representing the spectrum $F(\omega_x, \omega_t)$ of the wave is equivalent to taking the original $F_0(\omega_x)$ spectrum, rotating it and stretching it by $\sqrt{1+u^2}$.

- In our initial example $F_0(\omega_x) = \frac{1}{2}(\delta(\omega_x + \omega_0) + \delta(\omega_x - \omega_0))$: the $F_0(\omega_x, \omega_t)$ is made of two diracs located at $(\omega_0, -u\omega_0)$ and $(-\omega_0, +u\omega_0)$, which are at distance $\omega_0\sqrt{1+u^2}$ of the origin.