Image Motion

COMPSCI 527 — Computer Vision
Outline

1. Image Motion
2. Occlusion, Correspondence, Motion Boundaries
3. Constancy of Appearance
4. Motion Field and Optical Flow
5. The Aperture Problem
6. Estimating the Motion Field
Sensor Irradiance → Pixel Values

- **Irradiance** is the patterns of colors on the image sensor, compressed as an \((r, g, b)\) triple \(\mathbf{e}(\mathbf{x}, t)\)
- A pixel value is a noisy and quantized version of the integral of irradiance over a volume of size \(P_s \times P_s \times T_e\):

\[
f(i, n) = Q \int_{nT - T_e/2}^{nT + T_e/2} \left[ \int_{iP - P_s/2}^{iP + P_s/2} \mathbf{e}(\mathbf{x}, t) \, d\mathbf{x} \right] dt + \nu(i, n)
\]
Motion Field and Displacement

- Image trajectory of a world point that projects to $x$ at time $s$: $y(x, s, t)$
- So in particular $y(x, s, s) = x$
- Image velocity of $y(x, s, t)$ at time $t$: $w(x, s, t) \overset{\text{def}}{=} \frac{\partial y(x, s, t)}{\partial t}$
- Motion field at $x$ and at time $s$: $v(x, s) \overset{\text{def}}{=} w(x, s, s)$
- The true image velocity of a world point
- The vector difference $d(x, s, t) \overset{\text{def}}{=} y(x, s, t) - y(x, s, s)$ is the displacement at $x$ between times $s$ and $t$
Euler and Lagrange Viewpoints

Lagrange: $w(x, s, t)$
- Follow the point that is at $x$ at time $s$, as $t$ varies

Euler: $v(x, s) \overset{\text{def}}{=} w(x, s, s)$
- Stay at $x$ and observe velocities of points going by, as $s$ varies
The Displacement Field is not a 1 – 1 Map

- Point visible at $\mathbf{x}$ at time $s$ that becomes hidden at time $t$ (with $s < t$) forms an occlusion
- When $s > t$, this is called a disocclusion
- If points $\mathbf{x}$ at time $s$ and $\mathbf{y}$ at time $t$ do not form an occlusion and are projections of the same point in the world, they correspond to each other
- The displacement field is generally not integer-valued, so we cannot compute a 1 – 1 map between image pixels even if no occlusions or disocclusions exist
- A displacement field is typically given as a map $\mathbb{Z}^2 \rightarrow \mathbb{R}^2$, undefined at occlusions
- Sometimes two maps, in the two temporal directions
Constancy of Appearance

- What is assumed to remain constant across images?
- Motion estimation is impossible without such an assumption
- Most generic assumption: The appearance of a point does not change with time or viewpoint
- If two image points in two images correspond, they look the same
- If \( \mathbf{x} \) at time \( s \) and \( \mathbf{y} \) at time \( t \) correspond, then \( e(\mathbf{x}, s) = e(\mathbf{y}, t) \) (finite-displacement formulation)
- Equivalently, \( \frac{de(\mathbf{x}(t), t)}{dt} = 0 \) (differential formulation)
- This is the key constraint for motion estimation
Motion Field and Optical Flow

- Extreme violations of constancy of appearance:

- Ill-defined distinction:
  - Motion field $\approx$ true motion
  - Optical flow $\approx$ locally observed motion
The Optical Flow Constraint Equation

- The appearance of a point does not change with time or viewpoint: \[ \frac{de(x(t), t)}{dt} = 0 \]
- Total derivative, not partial (Lagrange viewpoint):
  \[ \frac{de(x(t), t)}{dt} \overset{\text{def}}{=} \lim_{\Delta t \to 0} \frac{e(x(t+\Delta t), t+\Delta t) - e(x(t), t)}{\Delta t} \]
- Use chain rule on \( \frac{de(x(t), t)}{dt} = 0 \) to obtain the Optical Flow Constraint Equation (OFCE)
  \[ \frac{\partial e}{\partial x} \frac{dx}{dt} + \frac{\partial e}{\partial t} = 0 \]
- \( \mathbf{v} \overset{\text{def}}{=} \frac{dx}{dt} \) is the unknown motion field (Euler viewpoint)
- This is the key constraint for motion estimation
The Aperture Problem

• Issues arise even when the appearance is constant:

\[ \frac{\partial e}{\partial x} \mathbf{v} + \frac{\partial e}{\partial t} = 0 \]

• Three equations in two unknowns:

• However, changes in irradiance are often caused by shading or shadows, which affects \( r, g, b \) similarly:

\[ \frac{\partial e_{\text{def}}}{\partial x} \begin{bmatrix} \frac{\partial e_1}{\partial x_1} & \frac{\partial e_1}{\partial x_2} \\ \frac{\partial e_2}{\partial x_1} & \frac{\partial e_2}{\partial x_2} \\ \frac{\partial e_3}{\partial x_1} & \frac{\partial e_3}{\partial x_2} \end{bmatrix} \]

• This degeneracy is called the \textit{aperture problem}.

\[ \frac{\partial e}{\partial t} = 0 \]

\[ \frac{\partial n}{\partial t} = 0 \]

\[ \frac{\partial e}{\partial t} = 0 \]

\[ \frac{\partial b}{\partial t} = 0 \]
The Aperture Problem for Black-and-White Video

- The aperture problem is extreme for black-and-white images, for which \( e \in \mathbb{R}^2 \):

\[
\frac{\partial e}{\partial x} + \frac{\partial e}{\partial t} = 0
\]

(OFCE is one scalar equation in the two unknowns in \( \mathbf{v} \))

- We cannot recover motion based on local measurements alone

- Only recover the *normal component* along the gradient

\[
\nabla e(x) = \frac{\partial e}{\partial x}^T
\]

(if the gradient is nonzero):

\[
\nu(x) \overset{\text{def}}{=} \| \nabla e(x) \|^{-1} [\nabla e(x)]^T \mathbf{v}(x)
\]

- In practice, this is very often the case also with color video
Smoothness and Motion Boundaries

- The assumption of constancy of appearance yields about one equation in two unknowns at every point in the image.
- To solve for $v$, we need further assumptions.
- The motion field $v : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is usually modeled as piecewise smooth.
- OFCE is solved in the LSE sense, and an additional regularization term is added to penalize deviations from smoothness.
- Smoothness holds almost everywhere, but not everywhere.
- Motion discontinuities are smooth image curves called motion boundaries.
Estimating the Motion Field

- Because of the aperture problem, we can only estimate several displacement vectors \( \mathbf{d} \) or motion field vectors \( \mathbf{v} \) simultaneously.
- Local methods
  - The image displacement \( \mathbf{d} \) in a small window around a pixel \( \mathbf{x} \) is assumed to be constant (extreme local smoothness).
  - Write one constancy of appearance equation for every pixel in the window.
  - Solve for the one displacement that satisfies all these equations as much as possible (in the LSE sense).
- Global methods
  - A data term measures deviations from constancy of appearance at every pixel in the image.
  - A smoothness term measures deviations of the motion field \( \mathbf{v}(\mathbf{x}) \) from smoothness.
  - Minimize a linear combination of the two types of terms.