Imagine a series of 2D images with $t$ as 3rd dimension.

Intensity of each pixel: $E(x,y,t)$

Our flow field: $(u(x,y,t), v(x,y,t))$ at time $t$

Most naive assumption
At each, the location where some index pixel moves to has the same intensity as its index location at time $t$

$E(x,y,t) = E(x+u(x,y,t), y+v(x,y,t), t)$

2. Moving light
   - Optical flow
   - But object is still

2 ways of tracking motion:
1. Global triangulation
2. Local relationships

For mathematics & details:
Berthold Horn  Robert Vision, Ch. 12