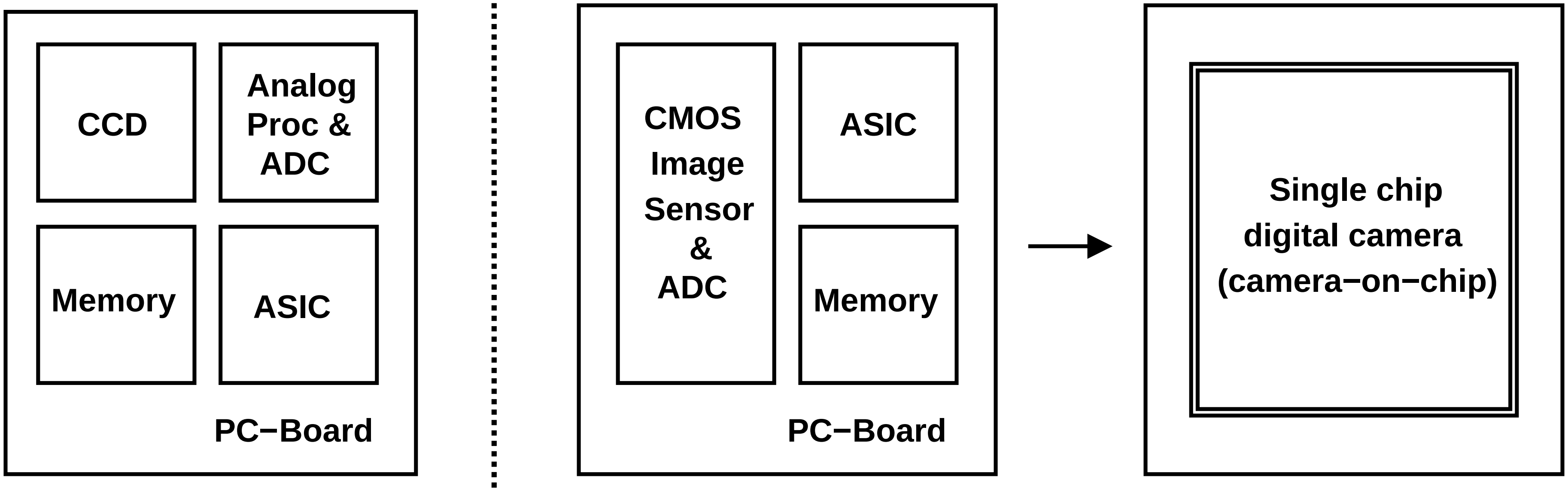
Optical Flow Estimation Using High Frame Rate Sequences

Suk Hwan Lim and Abbas El Gamal

Programmable Digital Camera Project

Department of Electrical Engineering, Stanford University, CA 94305, USA

Digital Imaging System Implementation



- CMOS image sensor:
  - Integration of camera functions with sensor on same chip
  - Low power consumption
  - High frame rate imaging

High Speed CMOS Image Sensor Examples

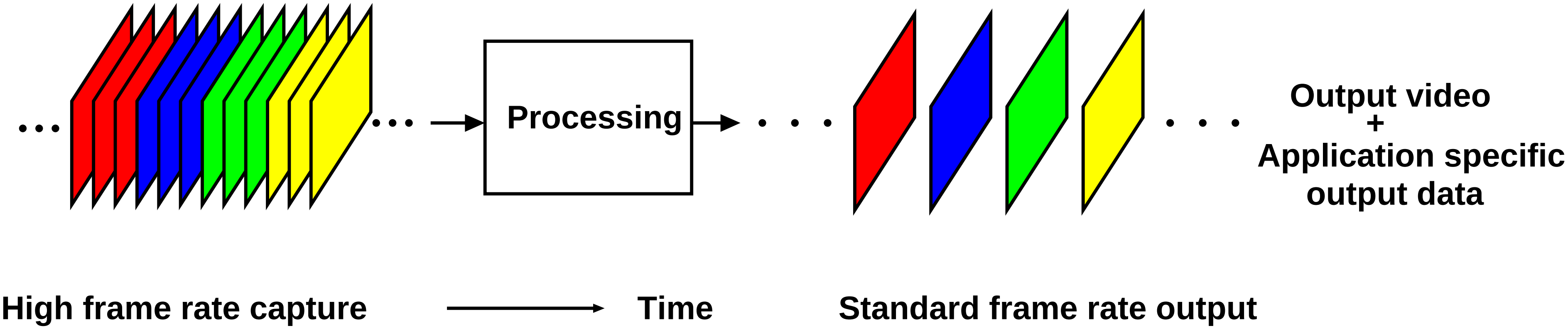
- **Fossum** *et al.* **(VLSI symposium 1999):**
  - **1024×1024 array with 500 fps**
  - **APS with 10μm×10μm pixel size**
- **Stevanovic** *et al.* **(ISSCC 2000):**
  - **256×256 array with 1024 fps**
  - **APS with 30μm×30μm pixel size**
- **Kleinfelder** *et al.* **(ISSCC 2001):**
  - **352×288 array with 10,000 fps**
  - **DPS with 9.4μm×9.4μm pixel size**
  - **ADC on each pixel**

Motivation

- Exploit high speed imaging capability to improve still and standard video rate imaging applications
  - Dynamic range enhancement
  - Motion blur-free capture
  - Optical flow estimation
  - Video stabilization
  - Super-resolution

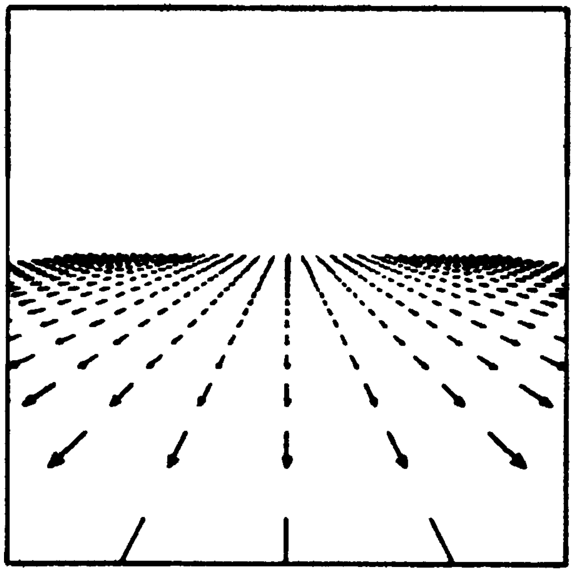
Integration of capture and processing on same chip makes system implementation feasible

Multiple Capture for Video/Data Enhancement

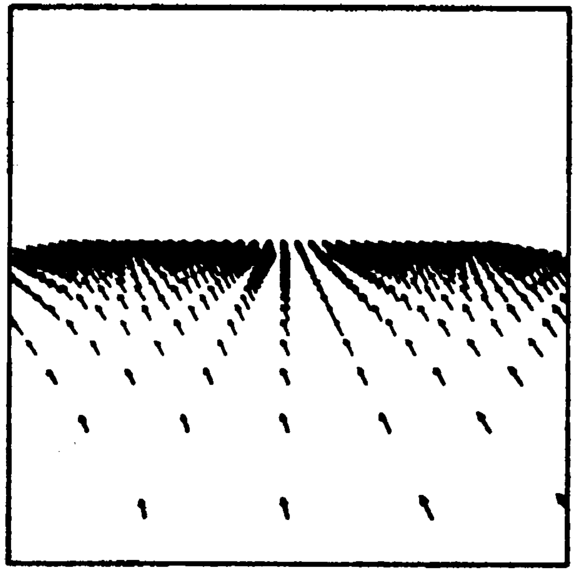


- Operate the sensor at high frame rate
- Process high frame rate data on-chip
- Output video with any application specific data at standard frame rate

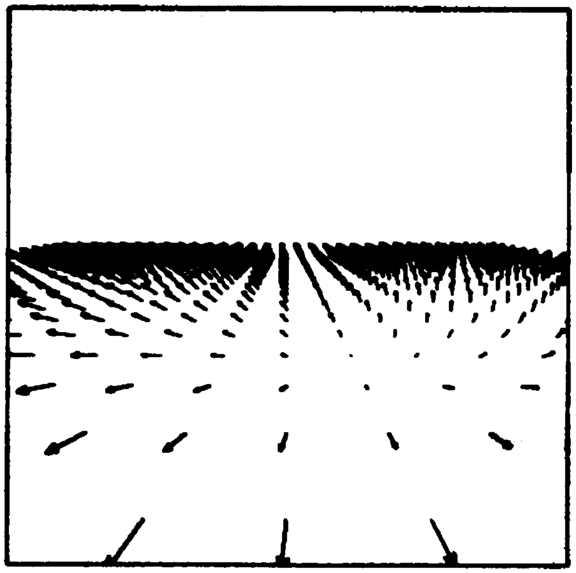
Optical Flow Estimation (OFE)



A



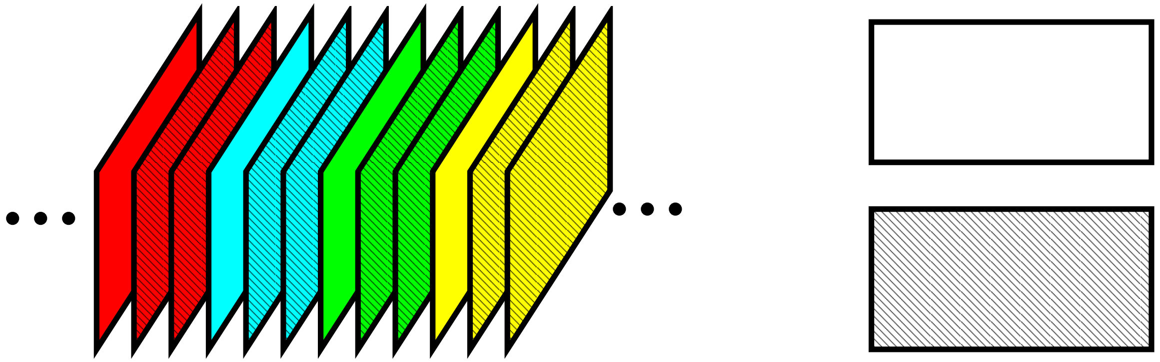
B



C

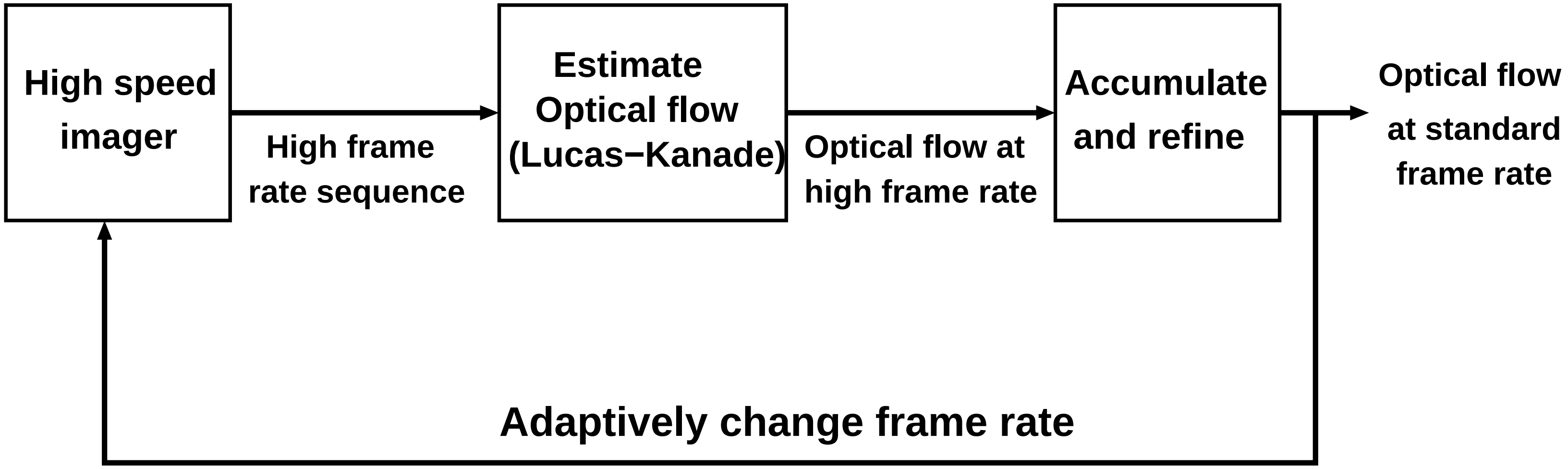
- Applications
  - 3D motion and structure estimation
  - Super-resolution
  - Image restoration
- Accuracy is of primary concern

Block Diagram of Our OFE Method



☐ Output frames

▨ Intermediate frames



High speed imager

High frame rate sequence

Estimate Optical flow (Lucas−Kanade)

Optical flow at high frame rate

Accumulate and refine

Optical flow at standard frame rate

Adaptively change frame rate

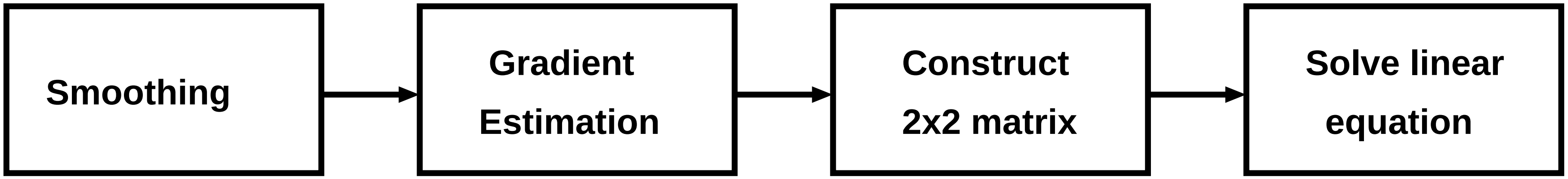
Effect of High Frame Rate on Optical Flow Estimation

- **Advantages for gradient-based methods**
  - **Brightness constancy assumption, *i.e.***

    $$\frac{dI(x,y,t)}{dt} = \frac{\partial I}{\partial x}v_x + \frac{\partial I}{\partial y}v_y + \frac{\partial I}{\partial t} = 0$$

    becomes more valid with higher frame rate
  - **Less temporal aliasing**
  - **Temporal derivatives better estimated**
  - **Smaller kernel size needed**
- **Disadvantages**
  - **Lower SNR**

Lucas-Kanade OFE Method



Smoothing

Gradient Estimation

Construct 2x2 matrix

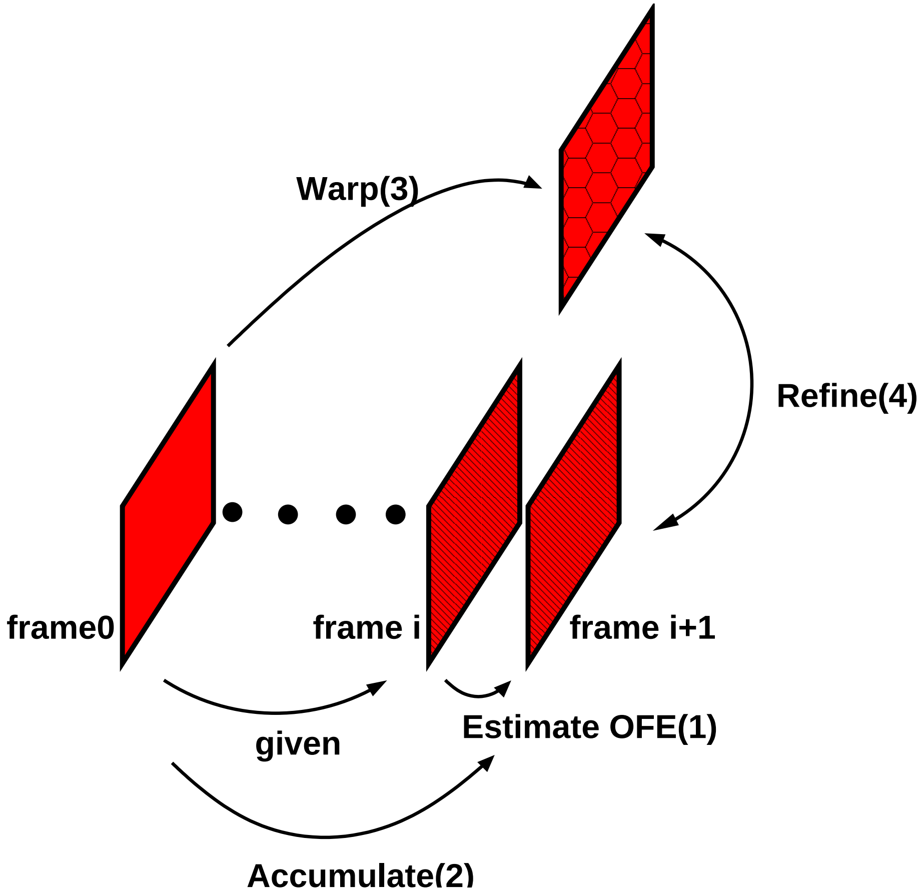
Solve linear equation

$$\begin{bmatrix} \sum wI_x^2 & \sum wI_xI_y \\ \sum wI_xI_y & \sum wI_y^2 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} = - \begin{bmatrix} \sum wI_xI_t \\ \sum wI_yI_t \end{bmatrix}$$

- I_x,I_y and I_t are partial derivatives computed using 5-tap filters
- $w(x, y)$ puts more weight to the center of neighborhood (5×5)

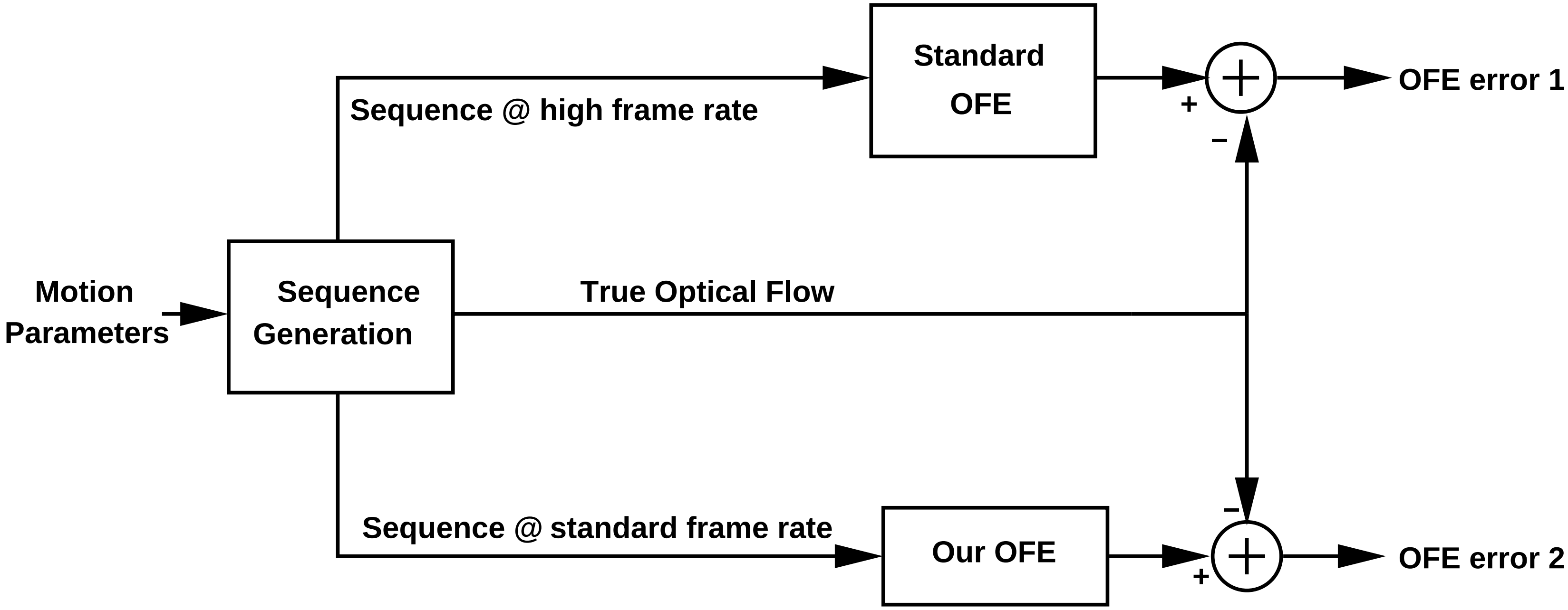
Accumulate and Refine

For $i = 0, \ldots, OV$:



OV is the temporal oversampling ratio

Experimental setup



- Displacement can be controlled and are known
- Motion blur and noise added
- Effect of frame rate on image quality included
- Standard OFE implemented by Barron *et. al*

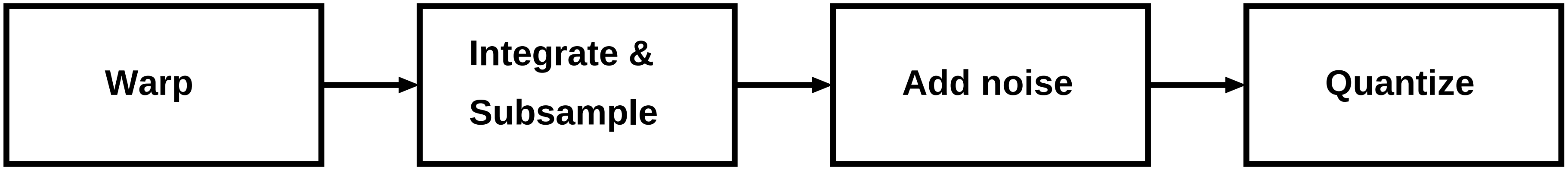
Video Sequence Model

- **The output charge from each pixel:**

$$Q(m,n) = \int_0^T \int_{ny_0}^{ny_0+Y} \int_{mx_0}^{mx_0+X} j(x,y,t)dxdydt + N(m,n)$$

  - **(m,n) is the pixel index**
  - **x_0 and y_0 are the pixel dimensions**
  - **X and Y are the photodiode dimensions**
  - **T is the exposure time**
  - **$j(x,y,t)$ A/cm^2 is the photocurrent density**
  - **$N(m,n)$ is the noise charge**
- **Pixel intensity, I, proportional to $Q(m,n)$**

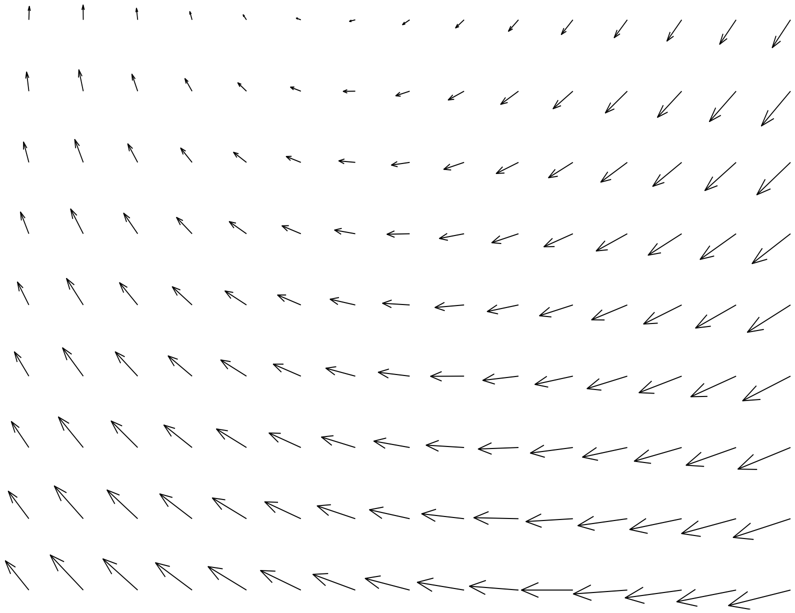
Synthetic Sequence Generation



1. Warp a high resolution (1312 × 2000) image using perspective warping
2. Integrate and subsample spatially (4 × 4) and temporally (10)
3. Add readout noise and shot noise according to the model
4. Quantize the sequence

Example Original Scene and Optical Flow

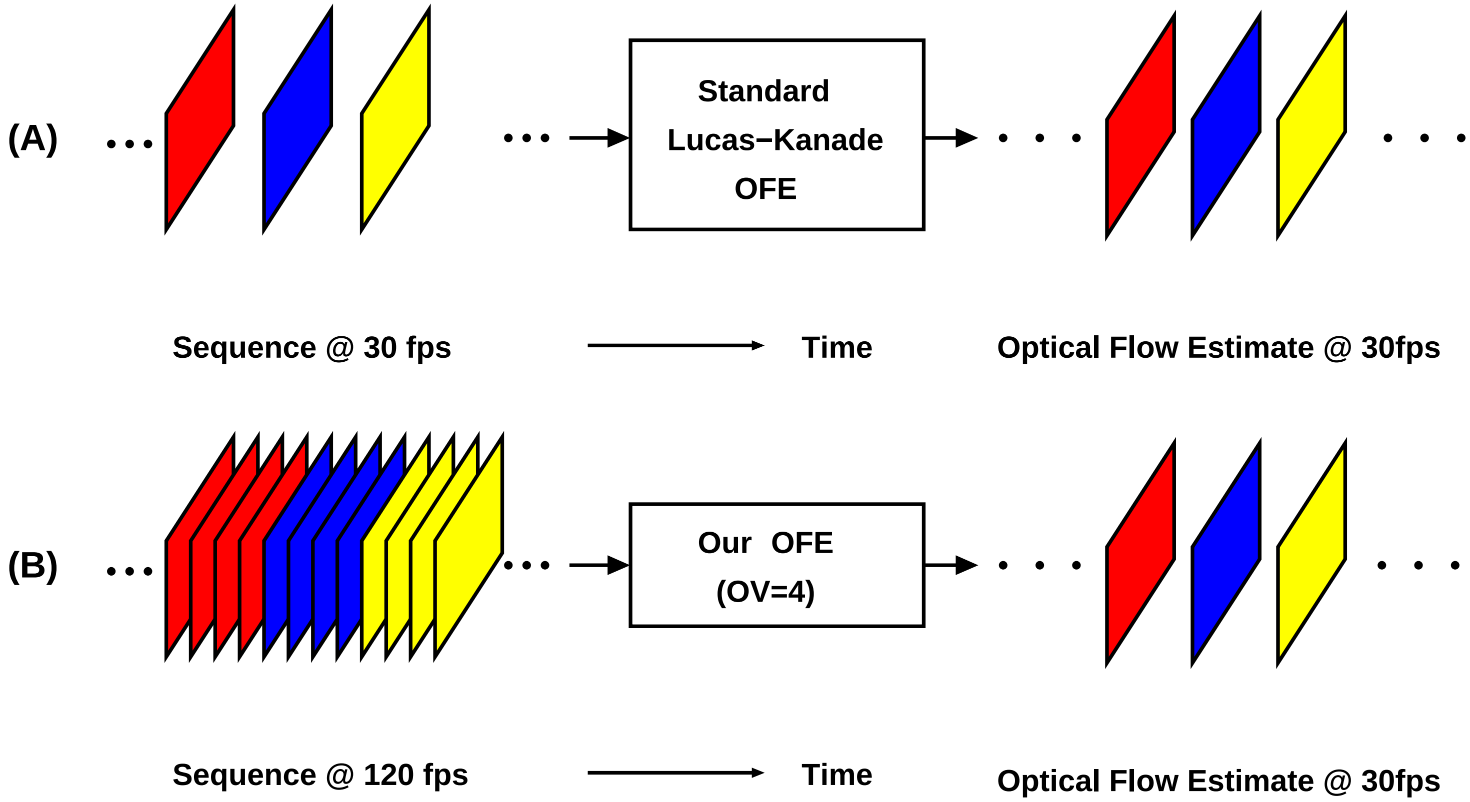
Original scene



Optical flow

Experiment I

Compare standard Lucas-Kanade OFE and our OFE



Displacement < 4 pixels/frame at standard frame rate

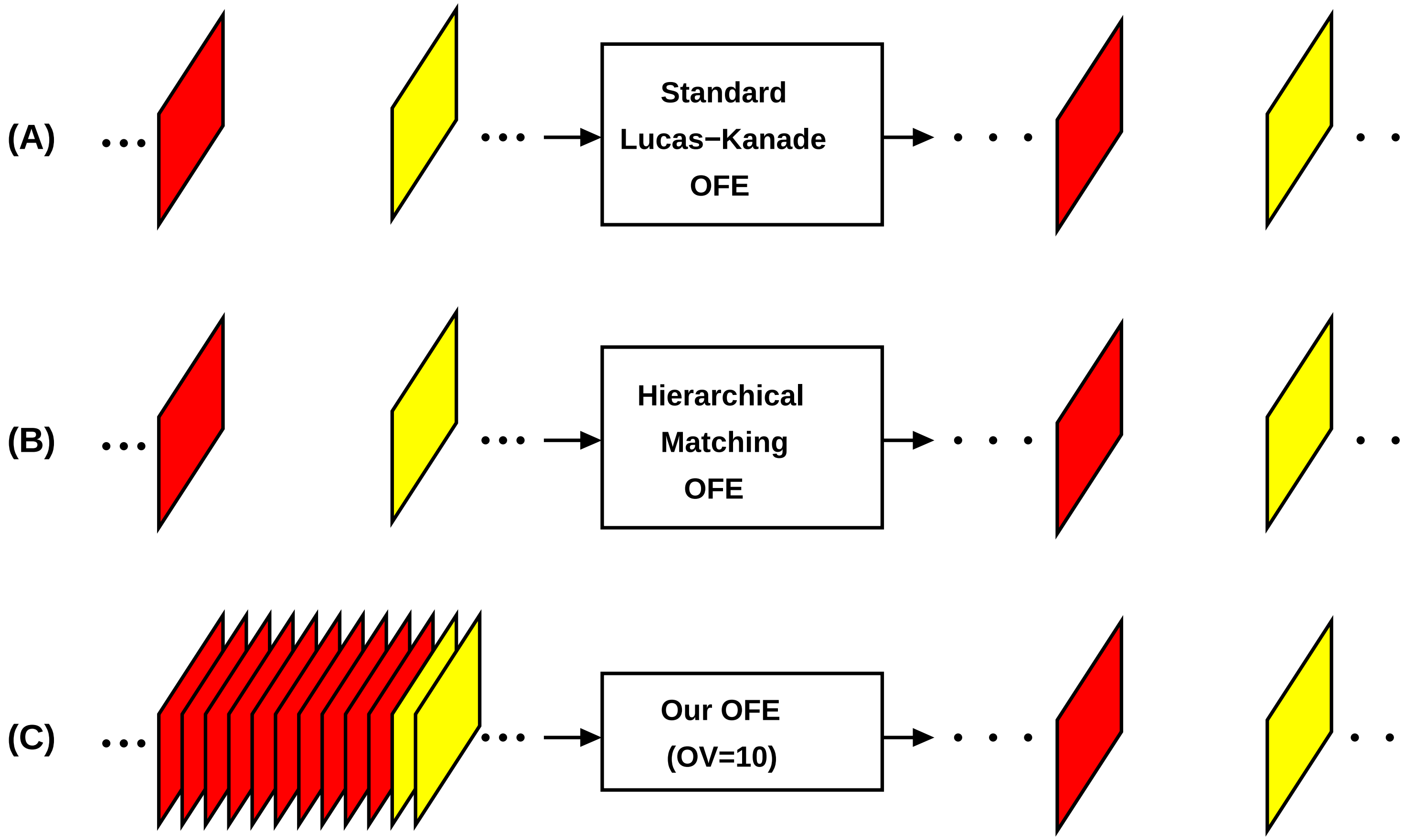
Result I

Scene	Lucas−Kanade method (A)		Our method (B)	
	Angular error	Density	Angular error	Density
1	4.43°	55.0%	3.43°	55.7%
2	3.94°	53.0%	2.91°	53.4%
3	4.56°	53.5%	2.67°	53.4%

- Higher accuracy achieved with our method
- More difference when brightness constancy does not hold
- Temporal filters
  - Our method: 2-tap
  - Lucas-Kanade method: 5-tap

Experiment II

Investigate accuracy gain for large displacements



Displacement $<$ 10 pixels/frame at standard frame rate

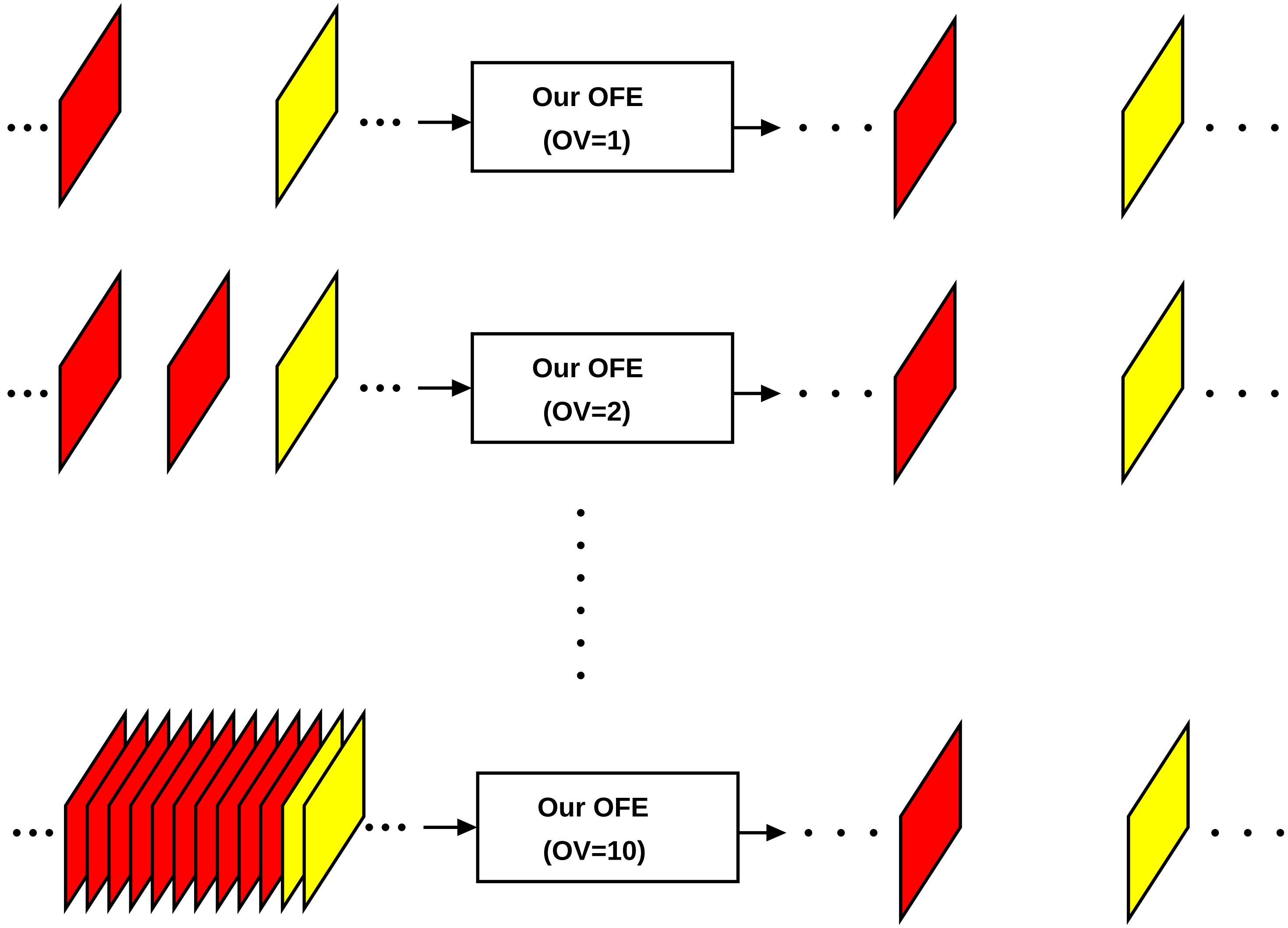
Result II

	Angular error	Density
Lucas−Kanade method	9.18°	50.81%
Hierarchical matching method	4.72°	100%
Our method (OV=10)	1.82°	50.84%

- Standard Lucas-Kanade method deteriorates for large displacements
- Hierarchical matching method has 100% density but lower accuracy
- Our method works well for both small and large displacements

Experiment III

Investigate the effect of varying OV on accuracy

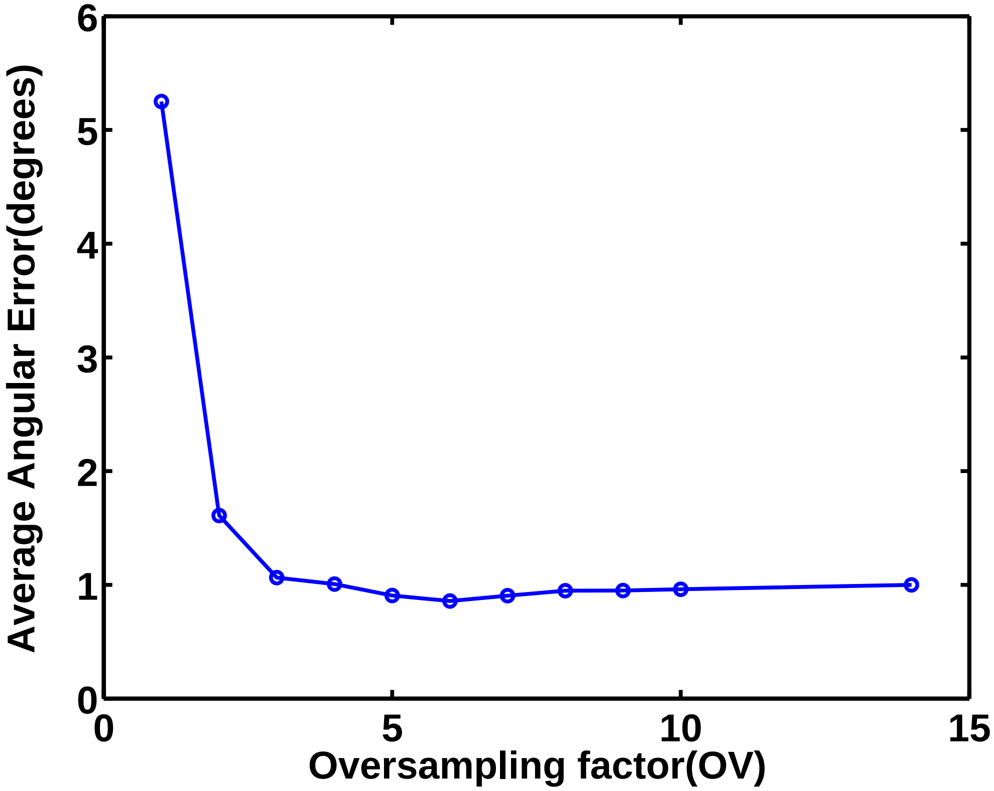


Our OFE (OV=1)

Our OFE (OV=2)

Our OFE (OV=10)

Result III



Temporal aliasing, temporal gradient estimation error, failure in brightness constancy and sensor SNR are affected by OV

Hardware Complexity

	Our method	**Lucas−Kanade method**
Memory (bytes)	$12mn$	$16mn$
Operations	$190mnOV$	$105mn$

- Assumptions:
  - $m \times n$ image with oversampling factor of OV
  - 5-tap spatial filter for gradient estimation and smoothing
  - 2-tap temporal filter for our method and 5-tap for Lucas-Kanade method
- Note memory requirement is independent of OV since our method is iterative

Conclusion

- High frame rate and integration capabilities of CMOS image sensors can be exploited to improve the performance of video processing applications
- Developed a method for accurate optical flow estimation using high frame rate sequences
- Demonstrated that our method
  - obtains higher accuracy than OFE using standard frame rate sequences
  - works well for large displacements
  - requires modest memory and computational power since our method is iterative