Video Compression
視訊壓縮

Yu-Lun (Alex) Liu
劉育綸

with slides by Wen-Hsiao Peng,
Shao-Yi Chien,
Hsueh-Ming Hang,
and Aggelos K. Katsaggelos

Week	Date	Topic	Assignments
1	2025-09-01		
2	2025-09-08	Introduction to Image and Video Processing	
3	2025-09-16	Signals and Systems	#1 – Color Transform, due: 2025-09-29 1:19pm
4	2025-09-22	Fourier Transform and Sampling	
5	2025-09-29	教師節補假	
6	2025-10-06	中秋節	
7	2025-10-13	Fourier Transform and Sampling	#2 – 2D-DCT, due: 2025-10-27 1:59pm
8	2025-10-20	Motion Estimation	Final project assigned (group together in fours)
9	2025-10-27	Lossless Compression	#3 – MEMC, due: 2025-11-10 1:59pm
10	2025-11-03	Image Compression	
11	2025-11-10	Video Compression	#4 – Entropy coding, due: 2025-11-24 1:59pm
12	2025-11-17	Learning-based Image/Video Compression	
13	2025-11-24	Paper Presentation	
14	2025-12-01	Guest Lecturer – MEDIATEK Meta NETFLIX	
15	2025-12-08	Guest Lecturer – MEDIATEK Meta NETFLIX	
16	2025-12-15	Final Project Presentation	

Motion Estimation

with slides by Wen-Hsiao Peng, Shao-Yi Chien, Hsueh-Ming Hang, and Aggelos K. Katsaggelos

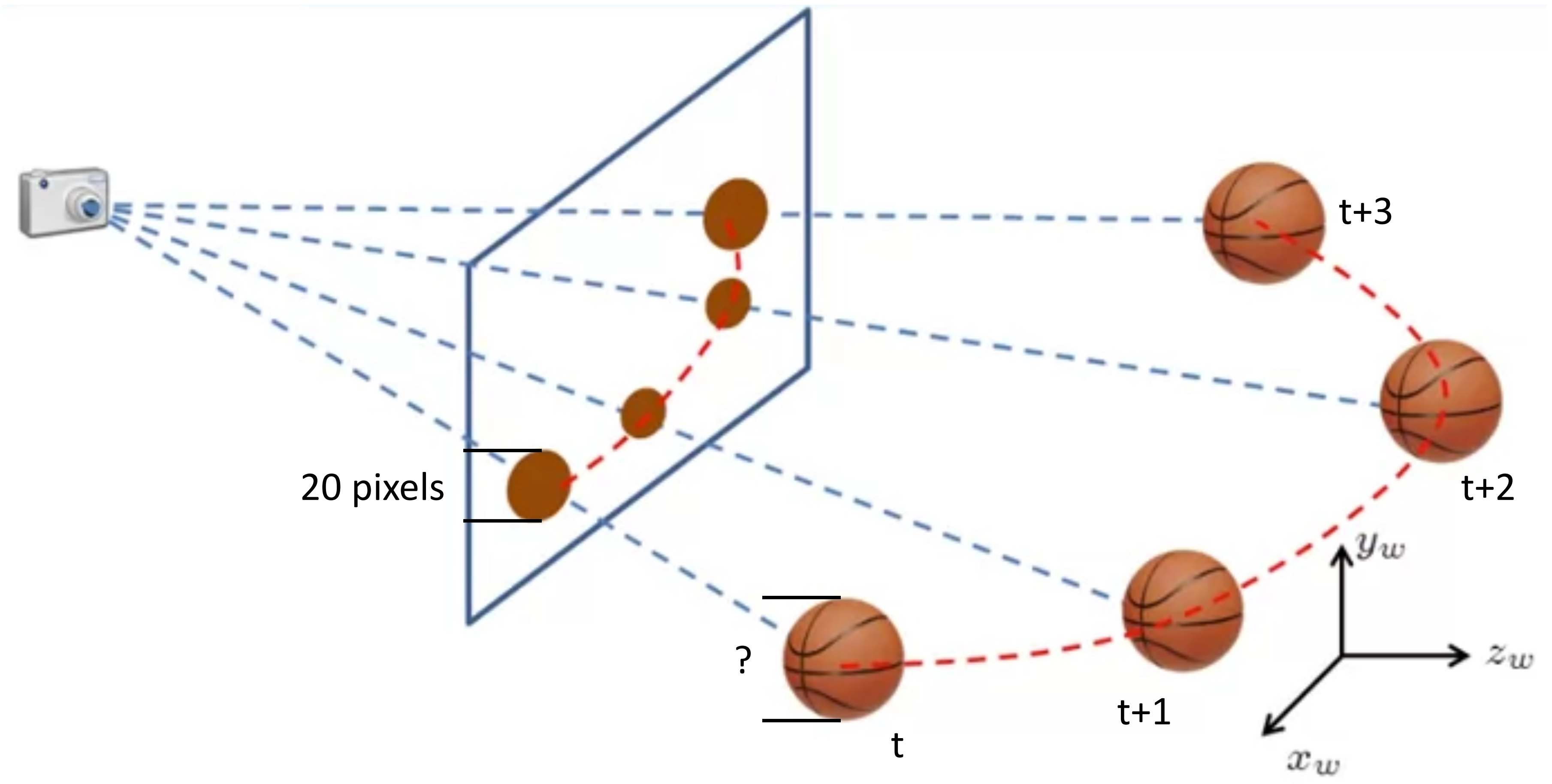
Motion Estimation

- Applications of Motion Estimation
- Phase Correlation
- Block Matching
- Spatio-Temporal Gradient Methods
- Motion Estimation and Compensation
- Learning-based Optical Flow

Motion Estimation

- **Applications of Motion Estimation**
- Phase Correlation
- Block Matching
- Spatio-Temporal Gradient Methods
- Motion Estimation and Compensation
- Learning-based Optical Flow

2D vs. 3D Motion



Basic Idea

$x_{k-1}(n_1, n_2)$

$x_k(n_1, n_2)$

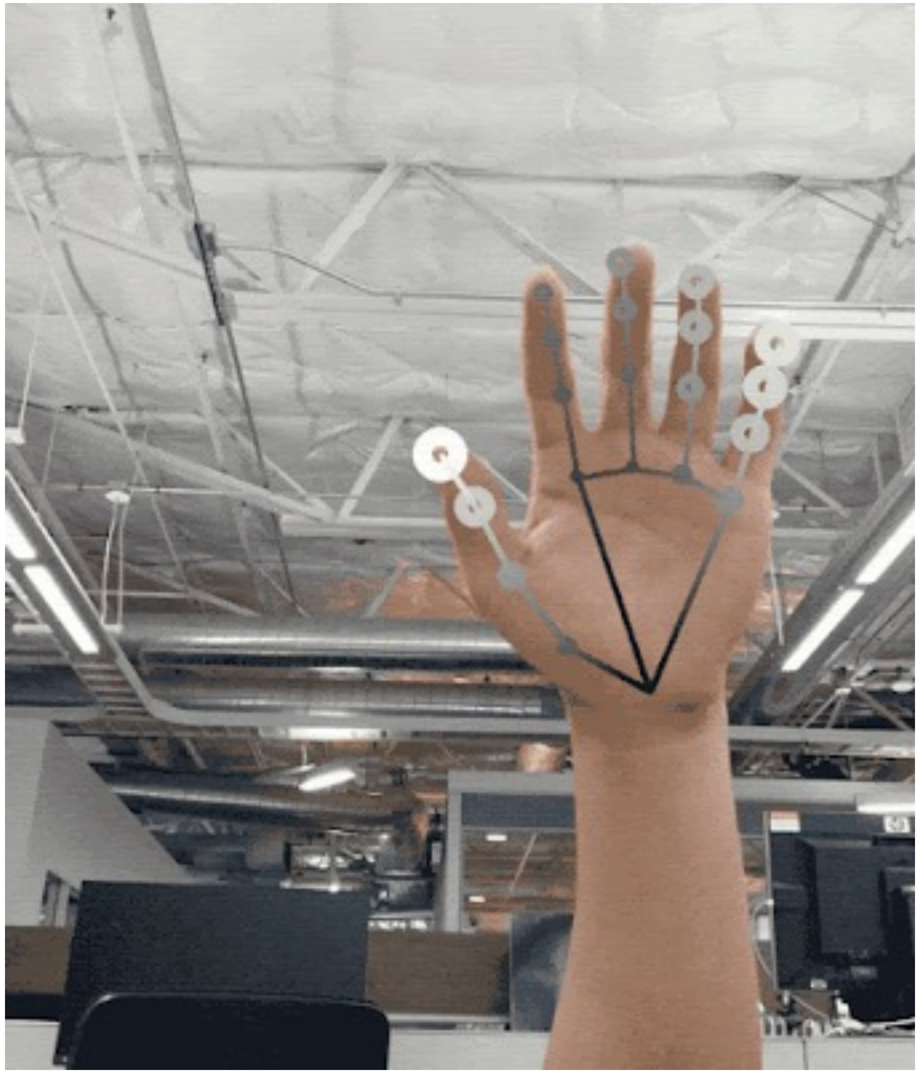
$x_{k+1}(n_1, n_2)$

Motion Estimation Applications

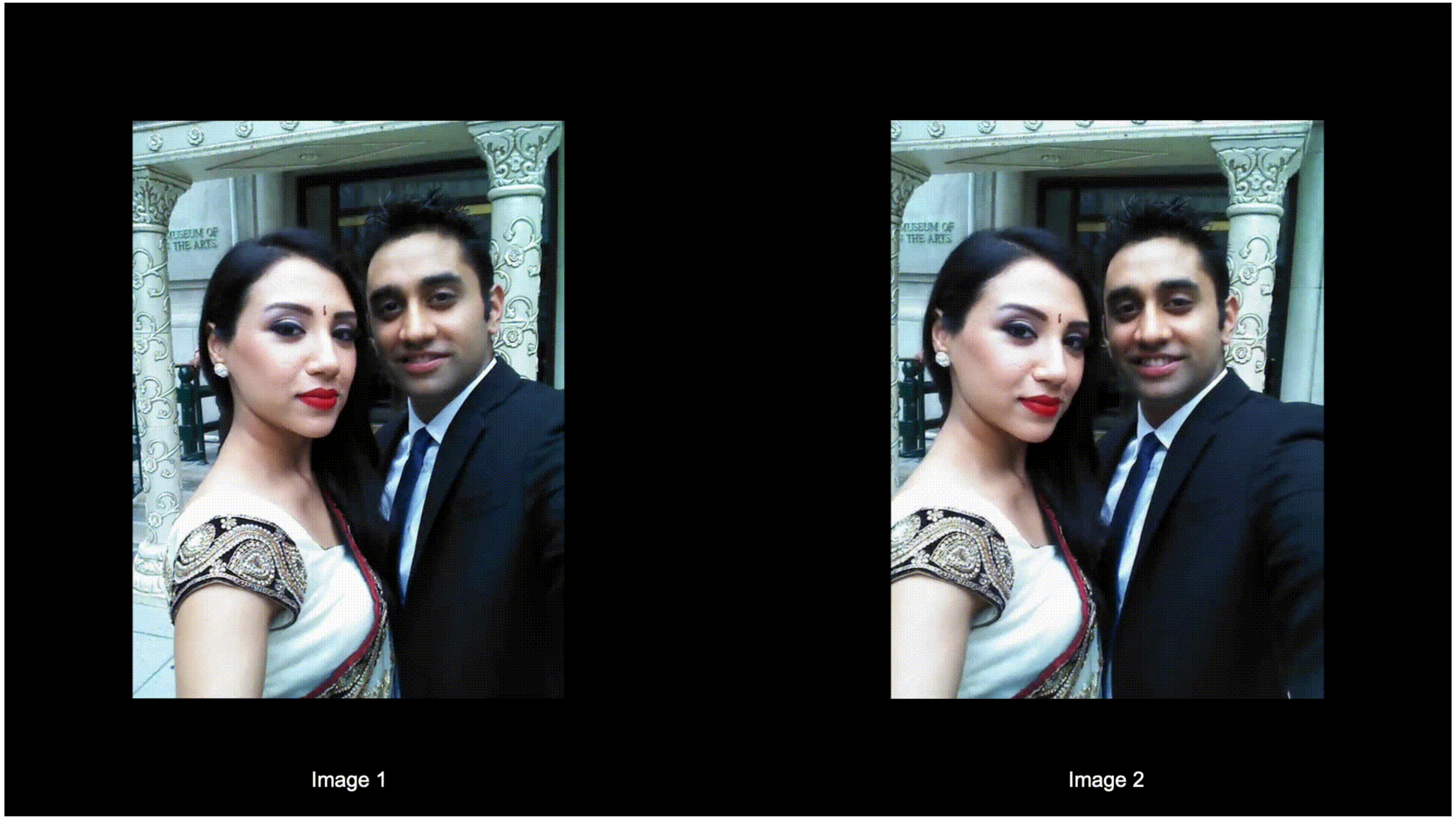
- Object Tracking
- Human Computer Interaction (HCI)
- Temporal Interpolation
- Spatio-Temporal Filtering
- Compression

Object Tracking

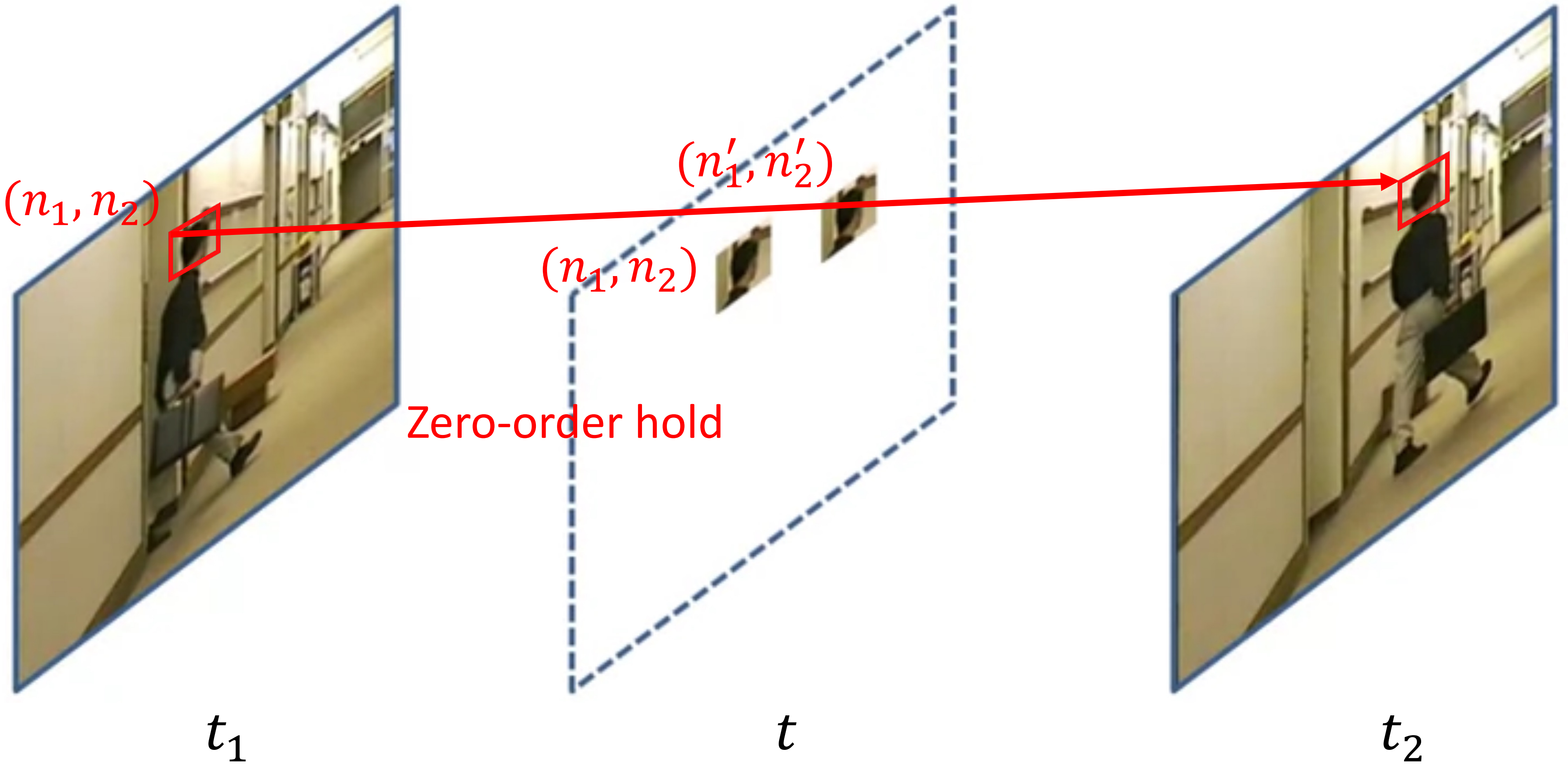
Human Computer Interaction (HCI)



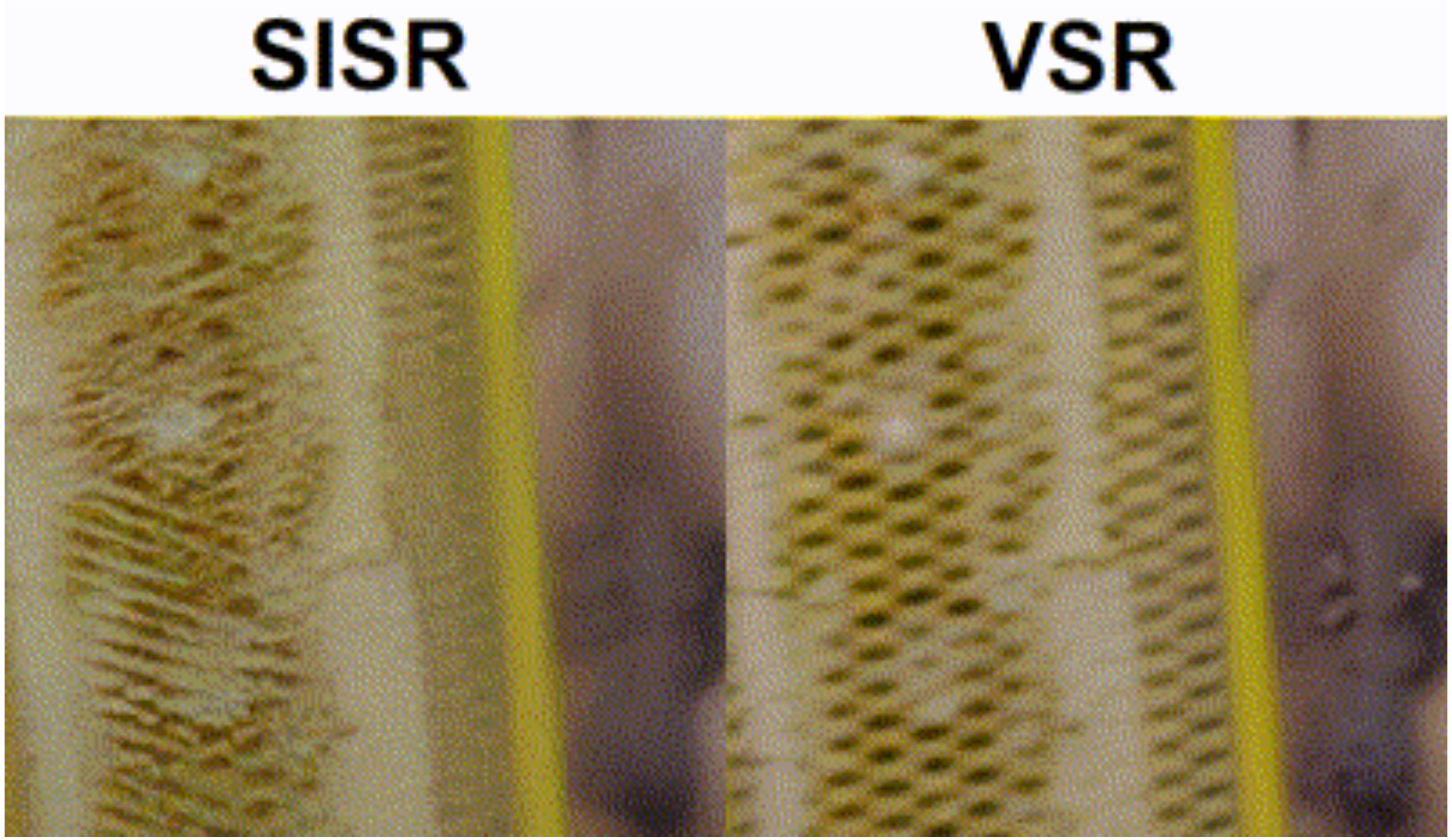
Temporal Interpolation



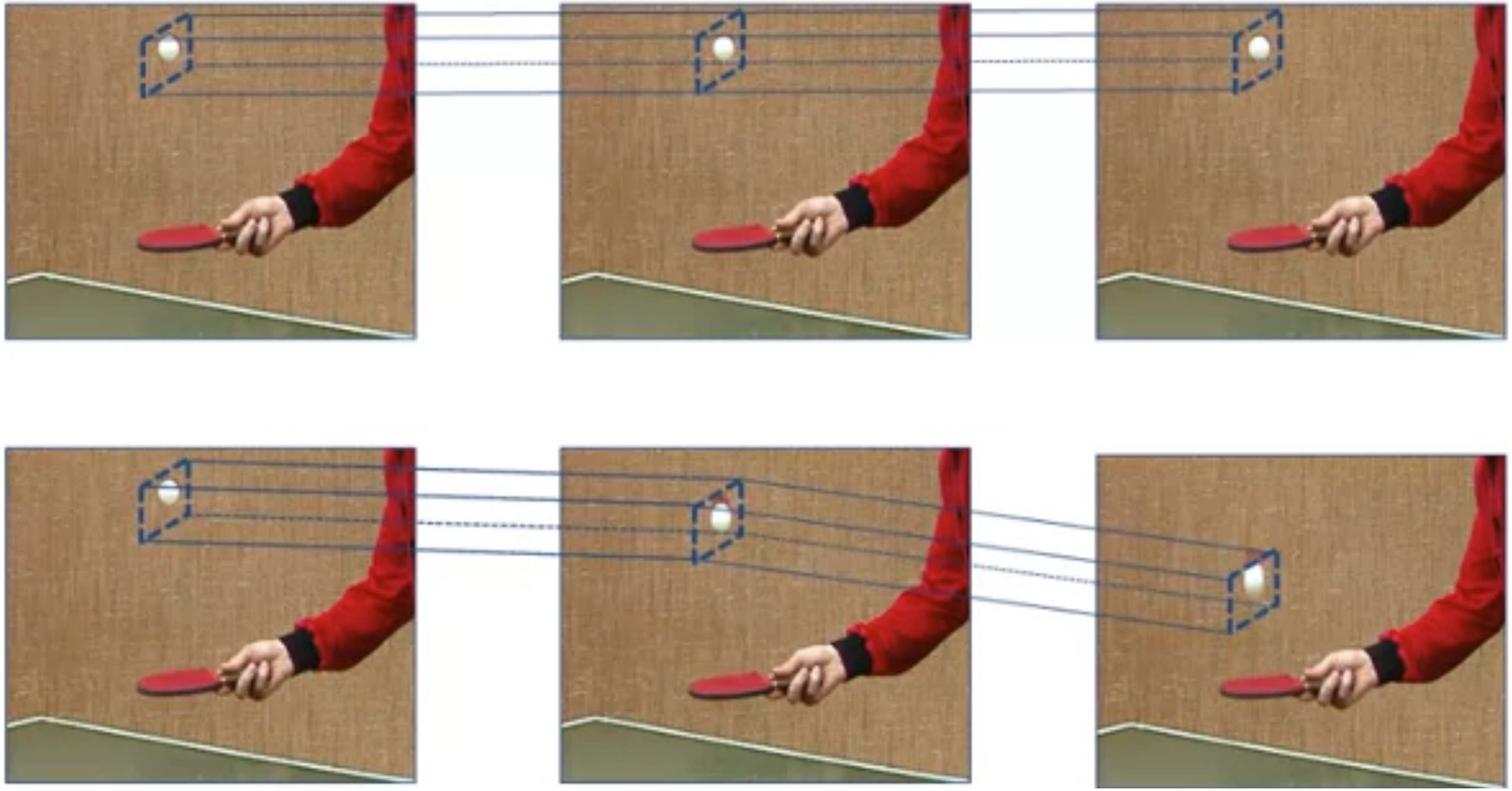
MC Temporal Interpolation



Spatio-Temporal Filtering



Motion Compensated Temporal Filtering



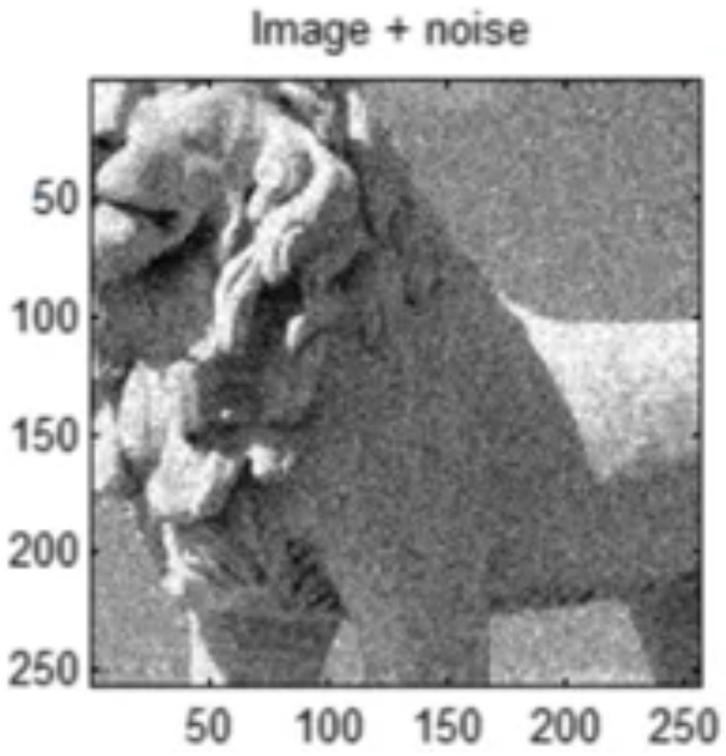
Motion Estimation

- Applications of Motion Estimation
- **Phase Correlation**
- Block Matching
- Spatio-Temporal Gradient Methods
- Motion Estimation and Compensation
- Learning-based Optical Flow

Classification of ME Methods

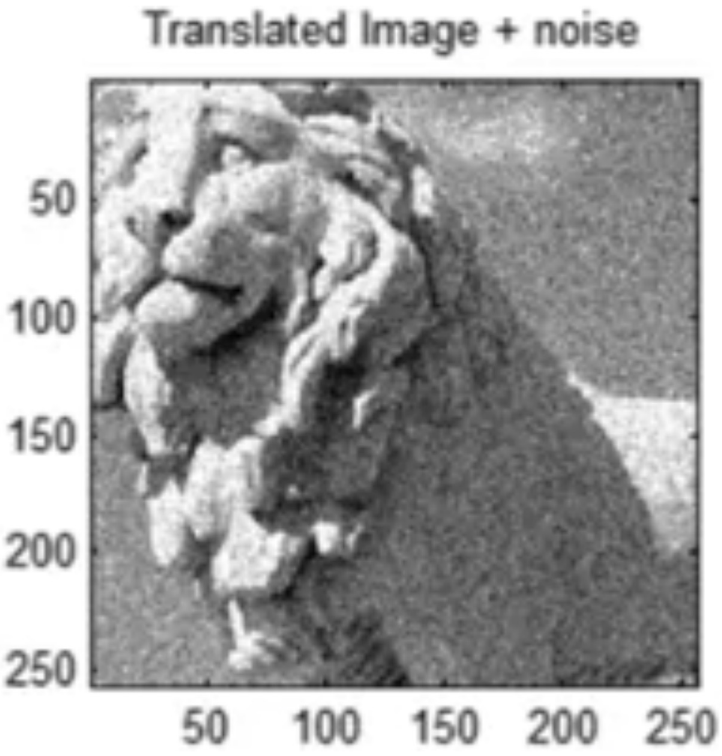
- Direct Methods
  - Phase Correlation
  - Block Matching
  - Spatio-temporal gradient
    - Optical Flow
    - Per-Recursive
- Indirect Methods
  - Feature Matching

Phase Correlation Example



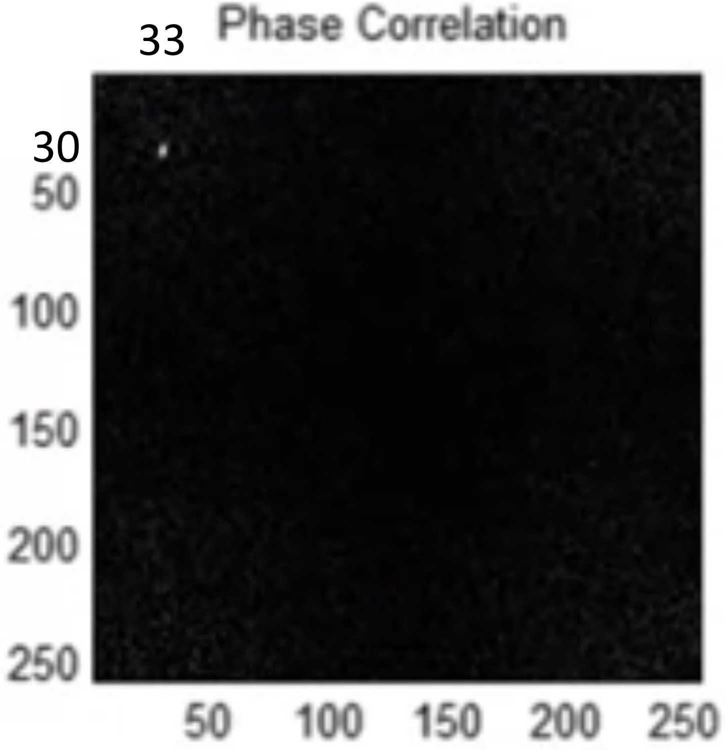
$$x(n_1, n_2)$$

$$X(\omega_1, \omega_2)$$



$$x(n_1 - m_1, n_2 - m_2)$$

$$e^{-j\omega_1 m_1} e^{-j\omega_2 m_2} X(\omega_1, \omega_2)$$



Phase Correlation

An image registration method

$x_{t-1}(n_1, n_2) \leftrightarrow X_{t-1}(k_1, k_2) \quad x_t(n_1, n_2) \leftrightarrow X_t(k_1, k_2)$

Assume $x_t(n_1, n_2) = x_{t-1}((n_1 - m_1)_{N_1}, (n_2 - m_2)_{N_2})$

Then $X_t(k_1, k_2) = X_{t-1}(k_1, k_2) e^{-j\frac{2\pi}{N_1} m_1 k_1} \ e^{-j\frac{2\pi}{N_2} m_2 k_2}$

Form: $C(k_1, k_2) = \frac{X_t(k_1,k_2) \cdot X_{t-1}^*(k_1,k_2)}{\left|X_t(k_1,k_2) X_{t-1}^*(k_1,k_2)\right|}$

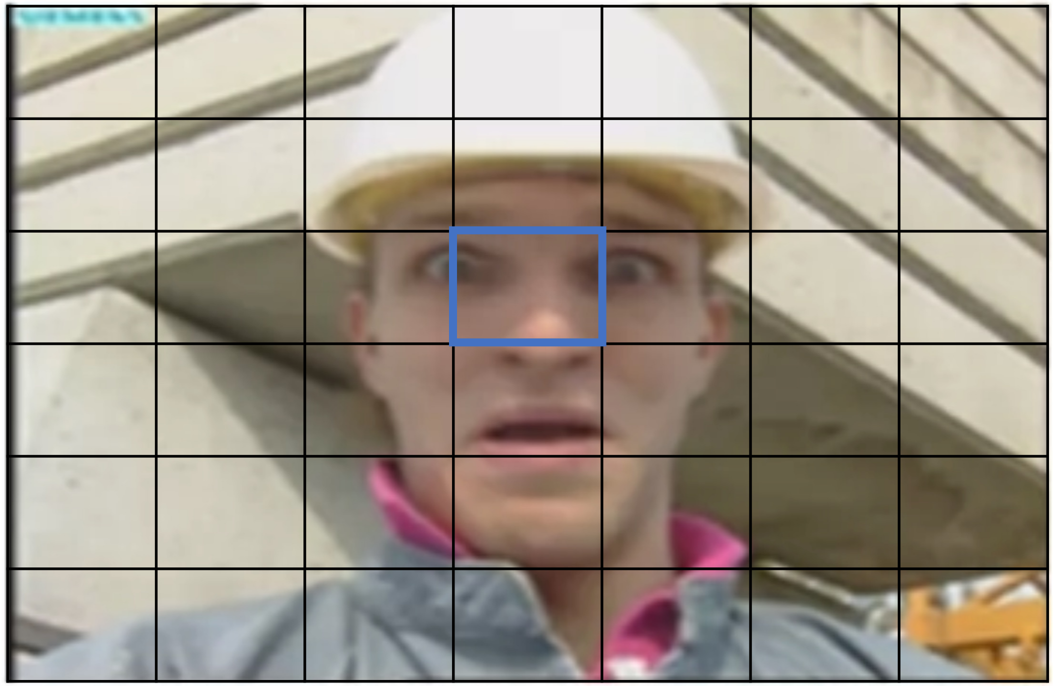
$c(n_1, n_2) = \delta(n_1 - m_1, n_2 - m_2)$

Motion Estimation

- Applications of Motion Estimation
- Phase Correlation
- **Block Matching**
- Spatio-Temporal Gradient Methods
- Motion Estimation and Compensation
- Learning-based Optical Flow

Block Matching

x_{t-1}


x_t

Basic underlying assumptions:

1. No change in the ambient lighting
2. Objects are rigid
3. Objects are translated in the 3D world on a plane parallel to the image place
4. No objects appeared or left the scene

Matching Criteria

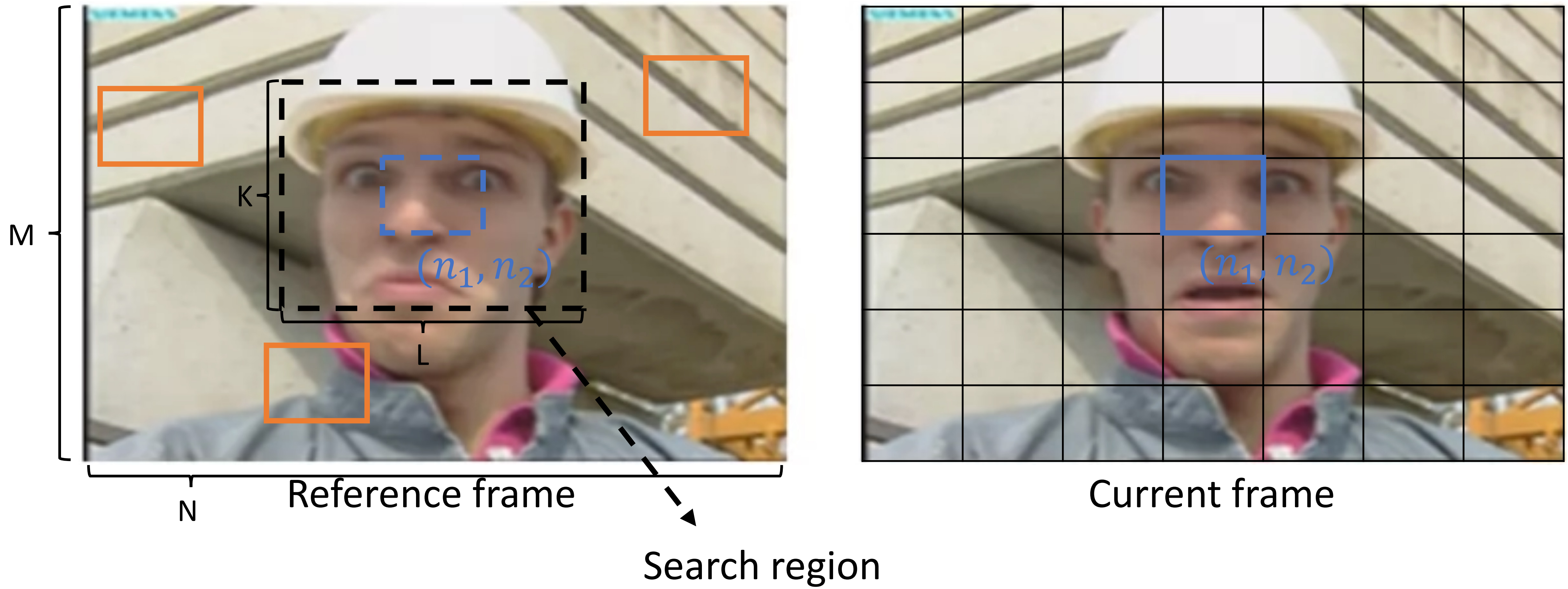
- A similarity of dissimilarity measure between regions (blocks)

$$\epsilon(d_1, d_2) = \sum_{(m_1,m_2)\in\mathcal{N}} \Phi(x_t(n_1 + m_1, n_2 + m_2), x_{t-1}(n_1 + m_1 + d_1, n_2 + m_2 + d_2))$$

- Examples of $\Phi()$
  - Correlation Function
  - Mean Squared Error (MSE)
  - Mean Absolute Error (MAE) or Mean Absolute Difference (MAD)

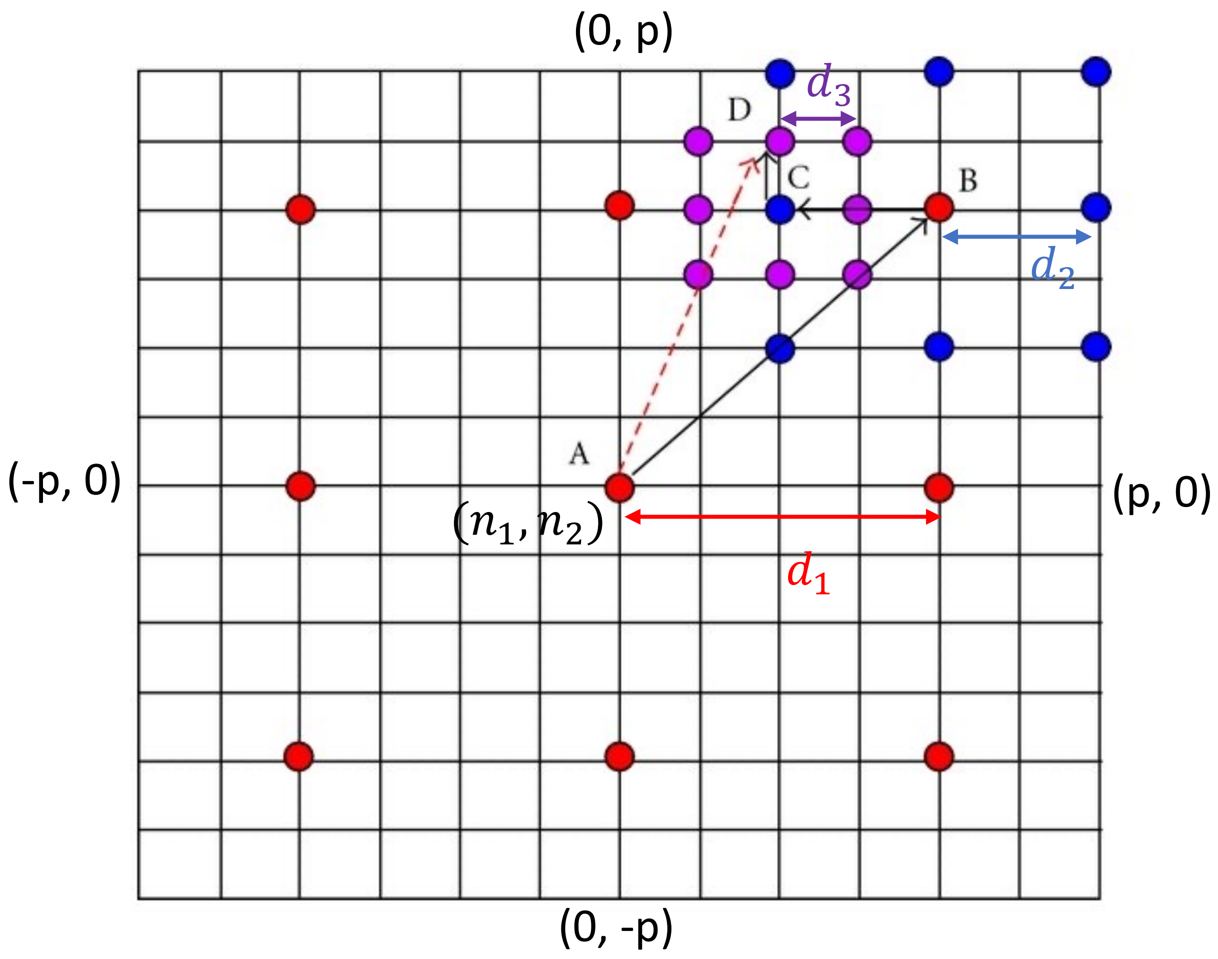
$$\epsilon(d_1, d_2) = \sum_{(m_1,m_2)\in\mathcal{N}} |x_t(n_1 + m_1, n_2 + m_2) - x_{t-1}(n_1 + m_1 + d_1, n_2 + m_2 + d_2)|$$

Search Region



$M \times N$ matching comparisons

2D Logarithmic Search



$$(2p+1)\times(2p+1)$$

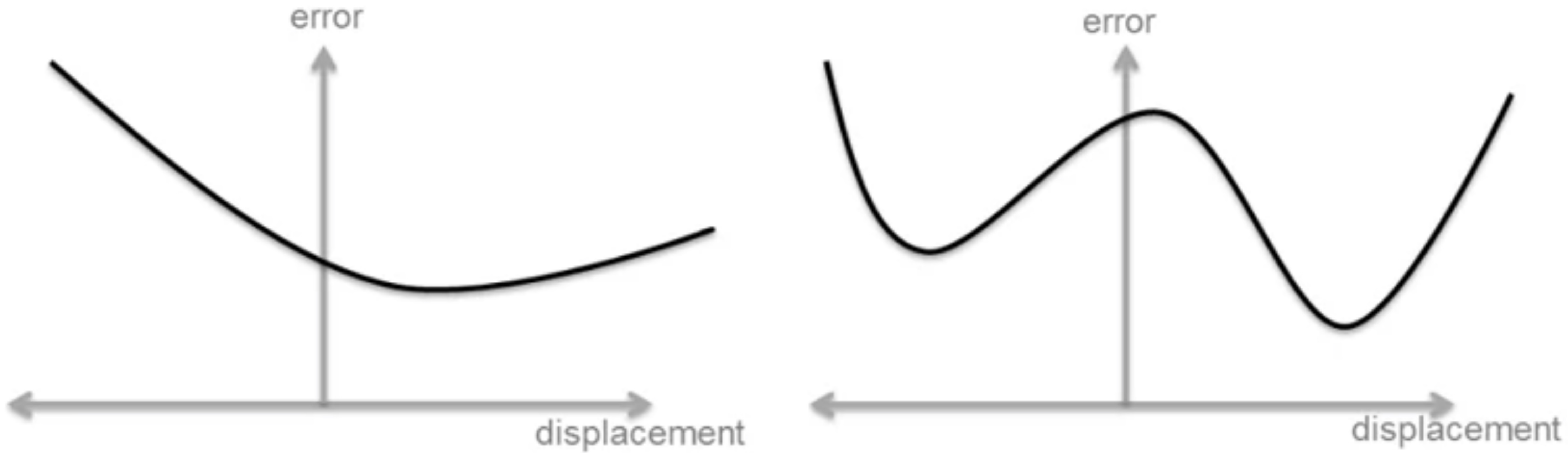
$$k = \lceil \log_2 p \rceil$$

$$p = 6, \qquad 13\times 13$$

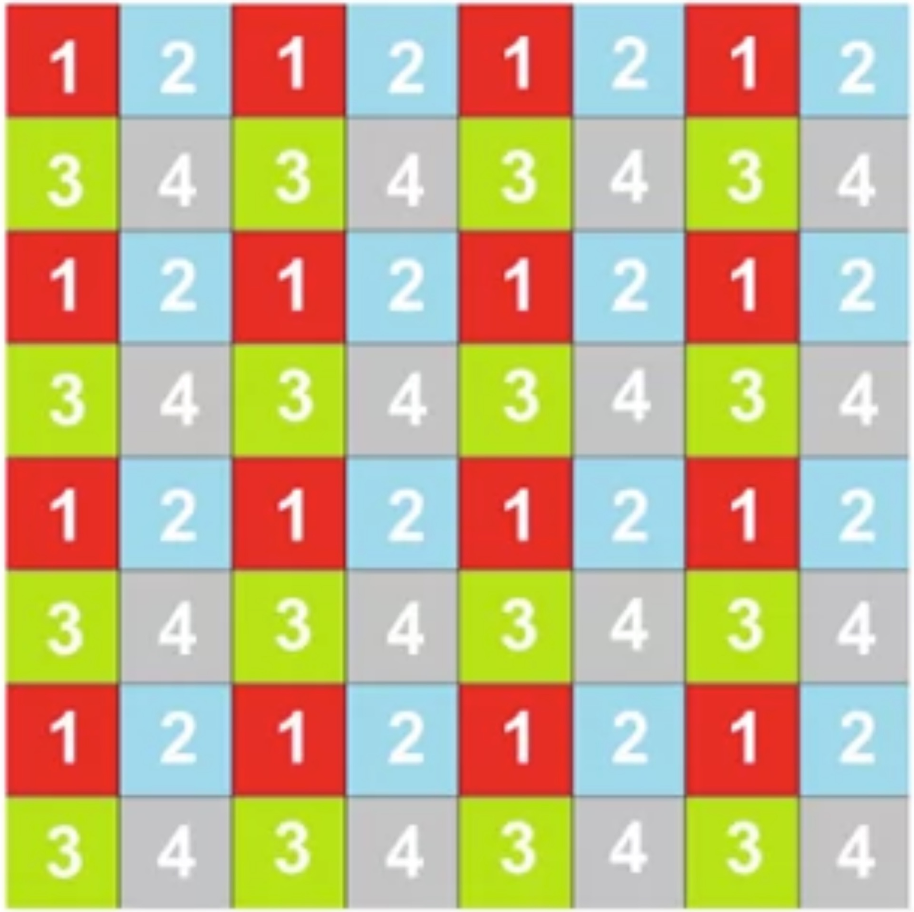
$$k = 3$$

$$k = \begin{cases} d_1 = 2^{k-1} \\ d_2 = 2^{k-2} \\ d_3 = 2^{k-3} \end{cases}$$

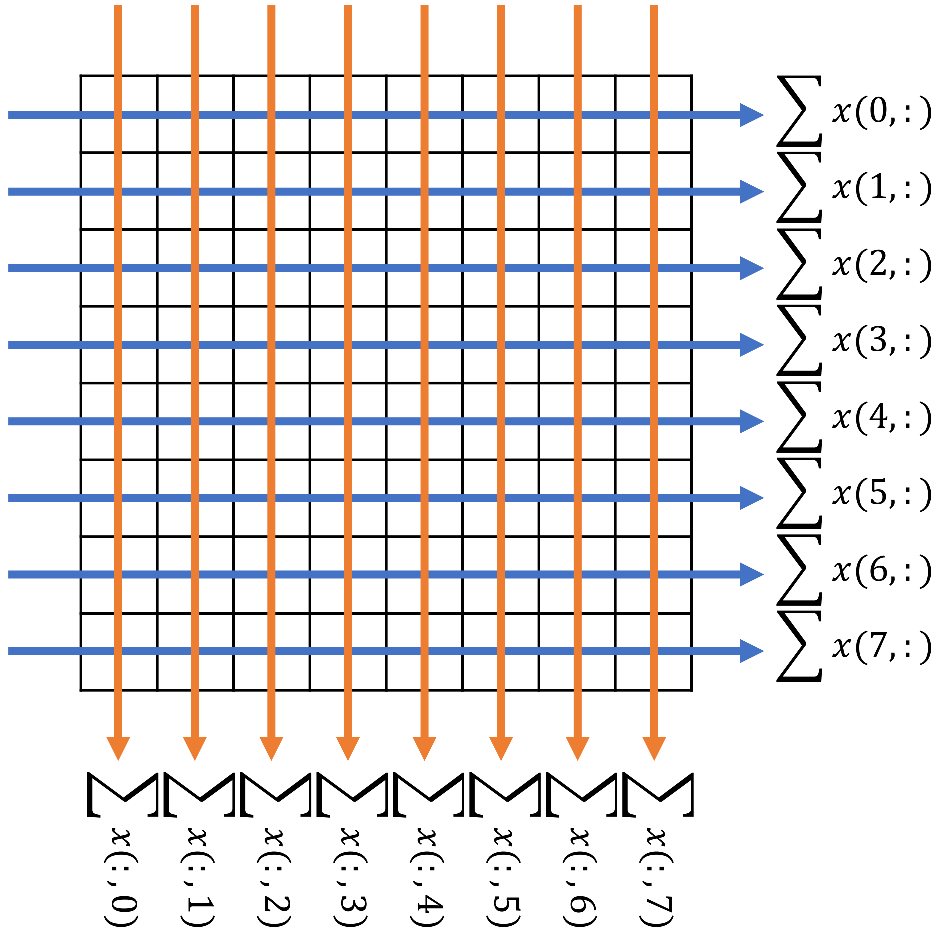
Global vs. Local Search



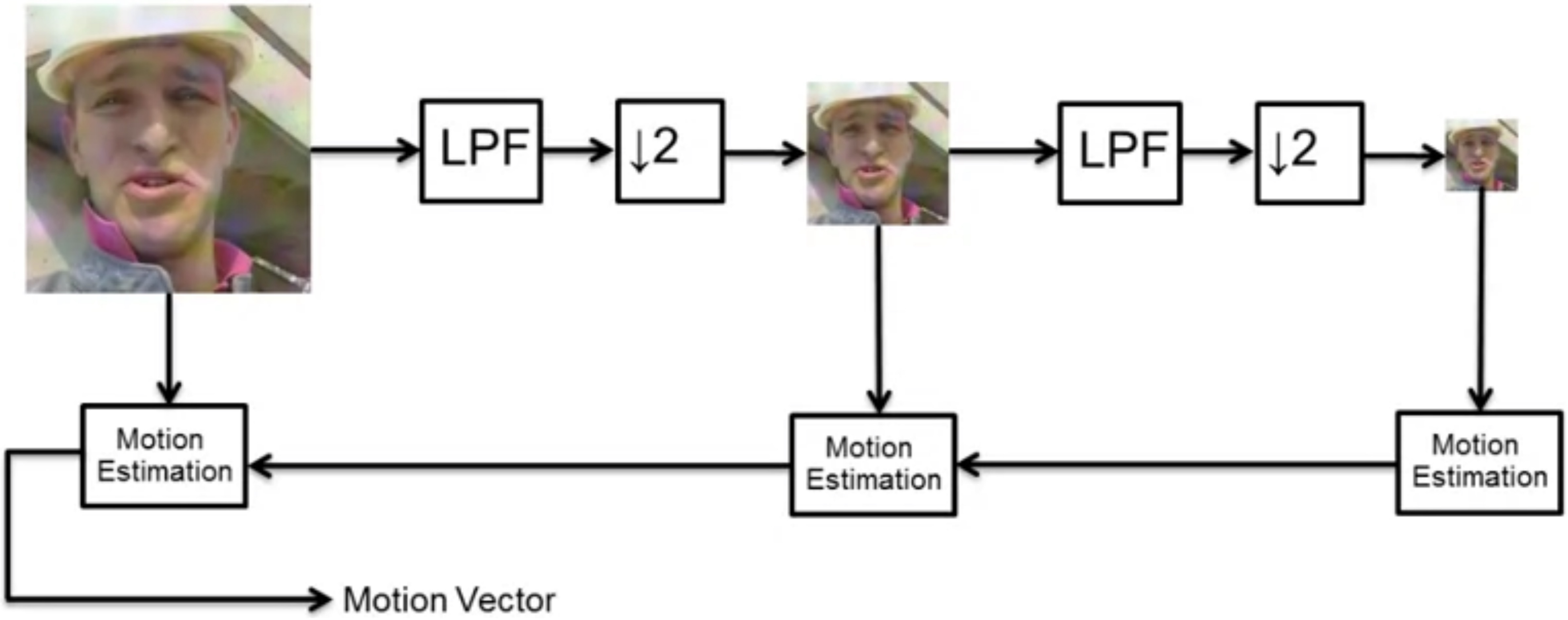
Pixel Sub-Sampling



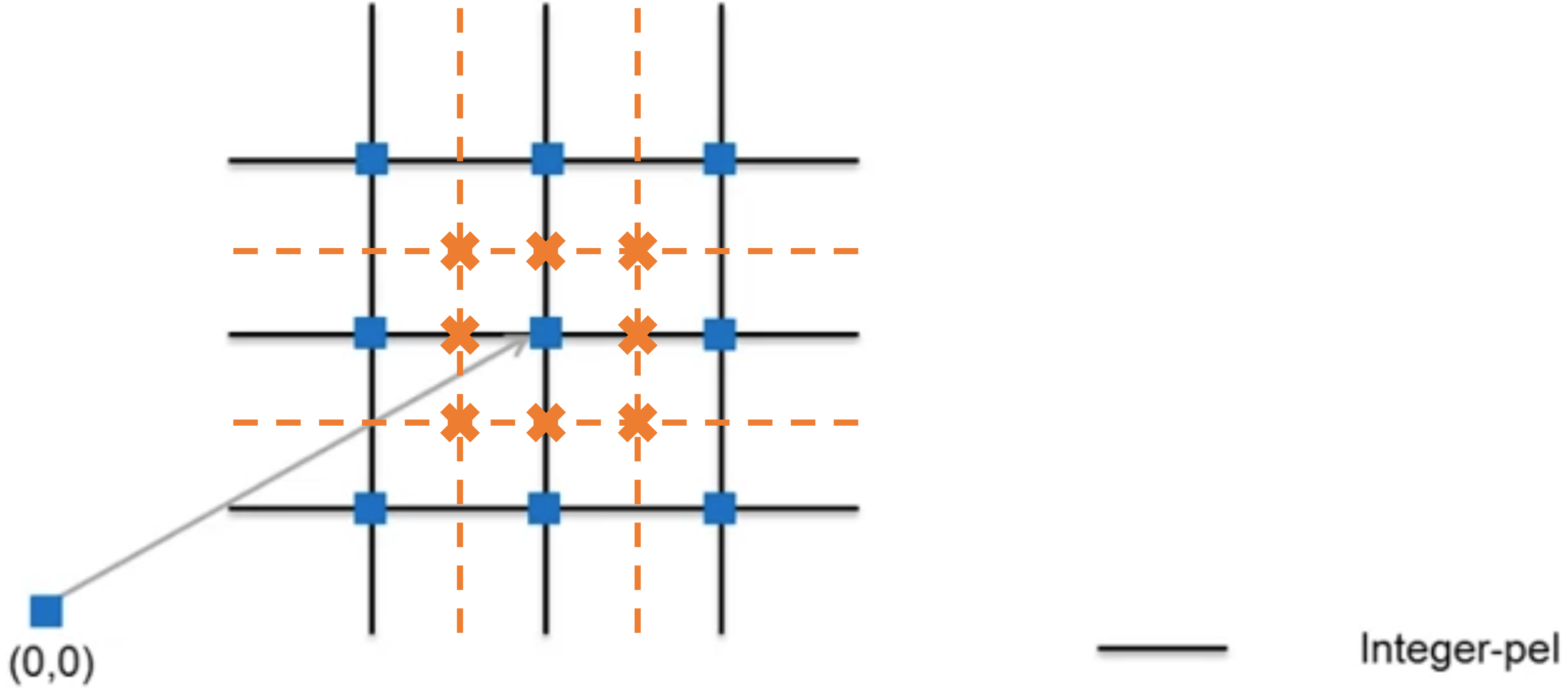
Pixel Projection



Hierarchical Motion Estimation



Sub-pixel Motion Estimation

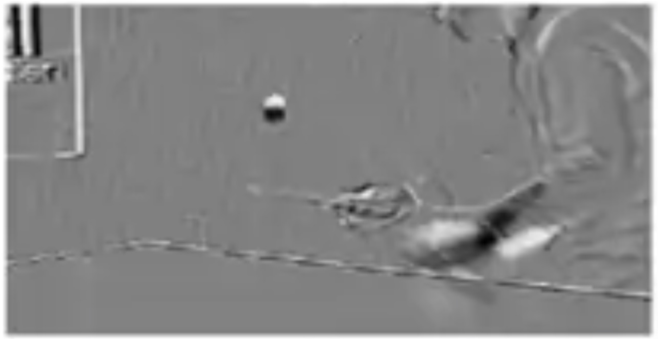


Experimental Comparison

(a)

(b)

(a) Reference picture; (b) current picture;



(c)

(c) Frame difference;

(e)

(e) Logarithmic search;

Motion Estimation

- Applications of Motion Estimation
- Phase Correlation
- Block Matching
- **Spatio-Temporal Gradient Methods**
- Motion Estimation and Compensation
- Learning-based Optical Flow

Optical Flow Approach

- Constant brightness constraint

$$I(x, y, 0) = I(x + u, y + v, \tau)$$

- Taylor series expansion

$$I(x + u, y + v, \tau) = I(x, y, 0) + \frac{\partial I(x, y, 0)}{\partial x} u + \frac{\partial I(x, y, 0)}{\partial y} v + \frac{\partial I(x, y, 0)}{\partial t} \tau + \mathrm{H.O.T.}$$

$$I_x u + I_y v + I_t \tau = 0$$

Or

$$I_x V_x + I_y V_y + I_t = 0$$

Optical Flow Approach

- Consider a neighborhood of the pixel:

$$
\begin{aligned}
I_x(q_1)V_x + I_y(q_1)V_y &= I_t(q_1)\\
I_x(q_2)V_x + I_y(q_2)V_y &= I_t(q_2)\\
\vdots \qquad\qquad \vdots \qquad &\qquad \vdots\\
I_x(q_n)V_x + I_y(q_n)V_y &= I_t(q_n)
\end{aligned}
\qquad
\begin{bmatrix} I_x(q_1) & I_y(q_1)\\ I_x(q_2) & I_y(q_2)\\ \vdots & \vdots \\ I_x(q_n) & I_y(q_n) \end{bmatrix}_{n\times 2}
\begin{bmatrix} V_x \\ V_y \end{bmatrix}_{2\times 1}
=
\begin{bmatrix} I_t(q_1)\\ I_t(q_2)\\ \vdots \\ I_t(q_n) \end{bmatrix}_{n\times 1}
$$

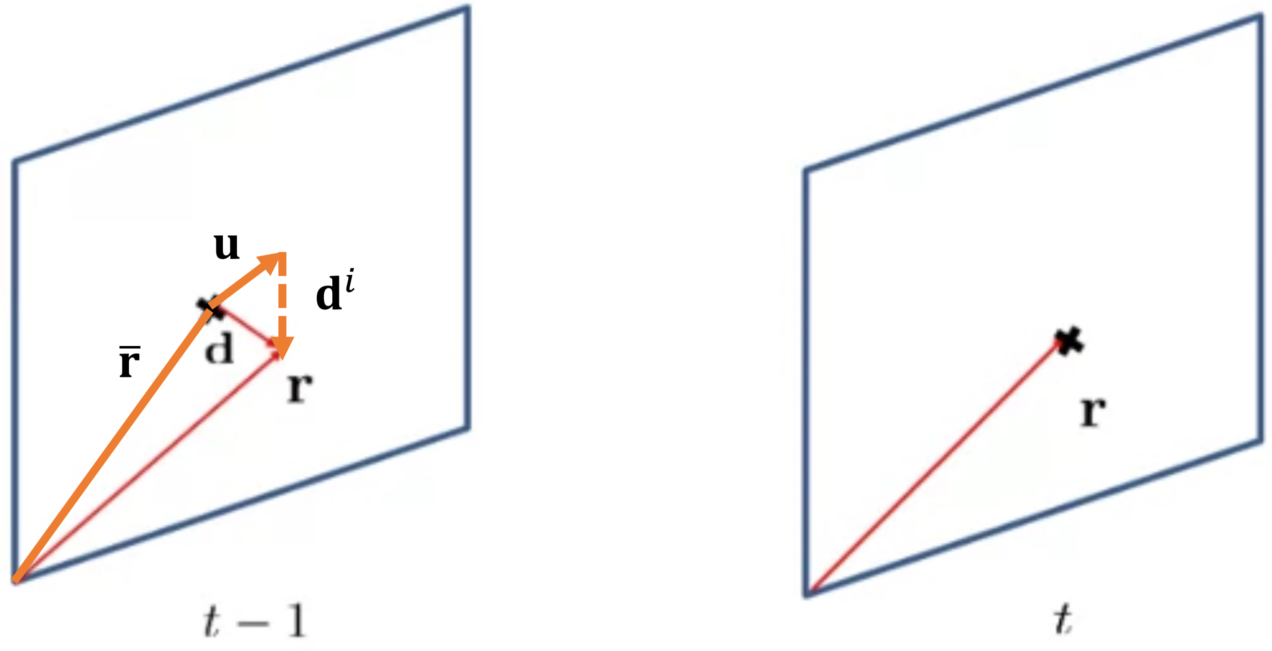
- We obtain $Ax = b$
- Min-norm Least-Square Solution $A^T Ax = A^T b \rightarrow x = (A^T A)^{-1} A^T b$
- Regularized Solution $(A^T A + \lambda C^T C)x = A^T b \rightarrow x = (A^T A + \lambda C^T C)^{-1} A^T b$

Pel-Recursive Algorithm

- Constant brightness constraint

$$I(\mathbf{r}, t) = I(\mathbf{r} - \mathbf{d}, t - 1) = I(\bar{\mathbf{r}}, t - 1)$$

- Assume an initial estimate $\mathbf{d}^i$
- and set $\mathbf{u} = \mathbf{d} - \mathbf{d}^i$



Pel-Recursive Algorithm

- Displaced Frame Difference

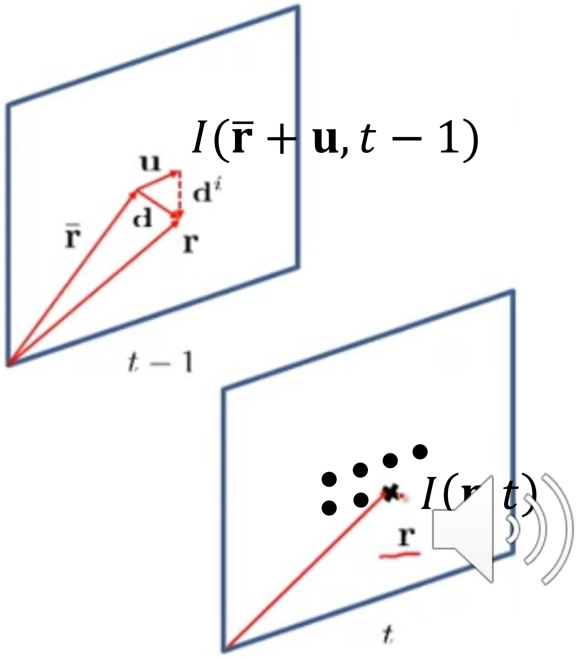
$$\Delta(\mathbf{r},\mathbf{u}) = I(\mathbf{r},t) - I\left(\mathbf{r}-\mathbf{d}^i, t-1\right) = I(\mathbf{r},t) - I(\bar{\mathbf{r}}+\mathbf{u}, t-1)$$

- Taylor series expansion

$$I(\bar{\mathbf{r}}+\mathbf{u}, t-1) = I(\bar{\mathbf{r}}, t-1) + \nabla^T I(\bar{\mathbf{r}}+\mathbf{u}, t-1)\mathbf{u} + \epsilon(\mathbf{r},\mathbf{u})$$

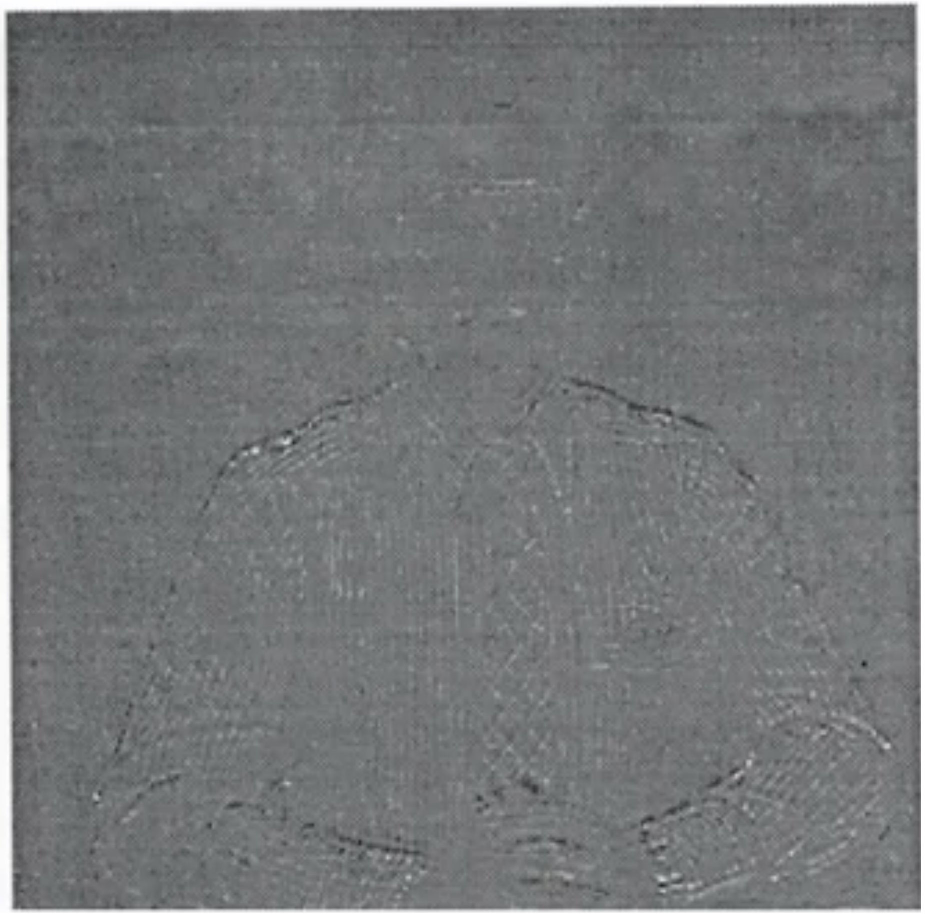
- Finally $\Delta(\mathbf{r},\mathbf{u}) = -\nabla^T I\left(\mathbf{r}-\mathbf{d}^i, t-1\right)\mathbf{u} - \epsilon(\mathbf{r},\mathbf{u})$

1. Recursive computability
2. $\mathbf{d}^i \rightarrow \mathbf{u} \quad \mathbf{u} = \mathbf{d} - \mathbf{d}^i$



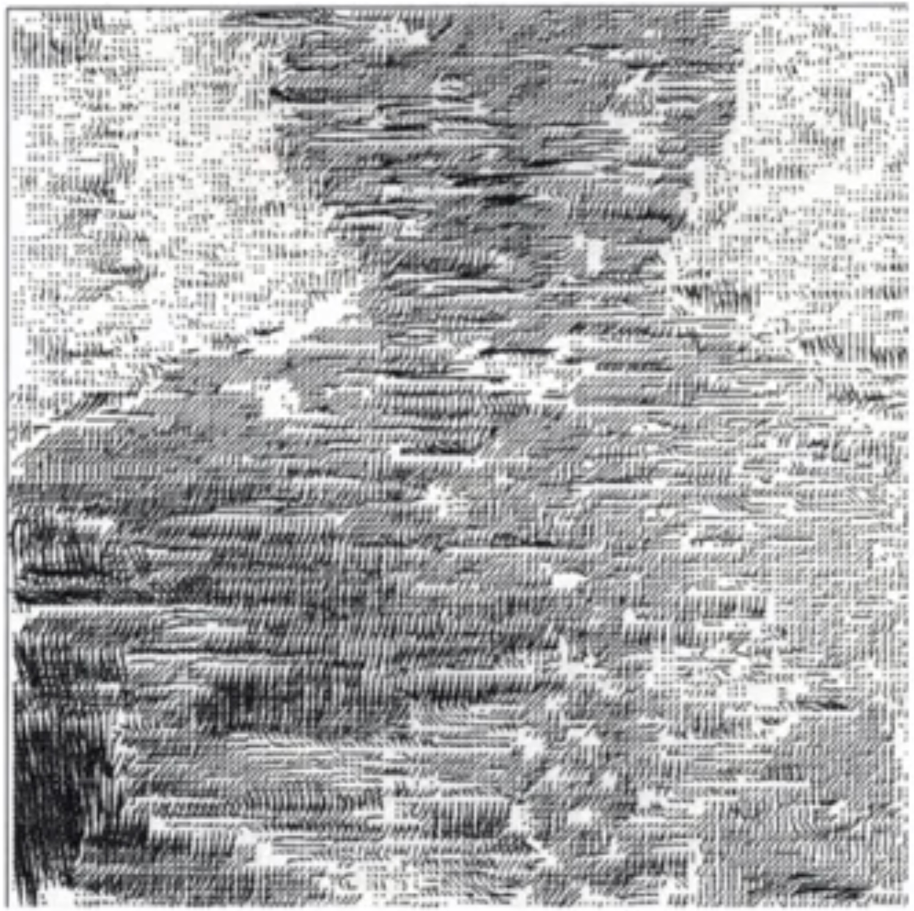
Frame Difference – DFD

FD



DFD

Estimated Motion Vectors



Feature-Based Methods

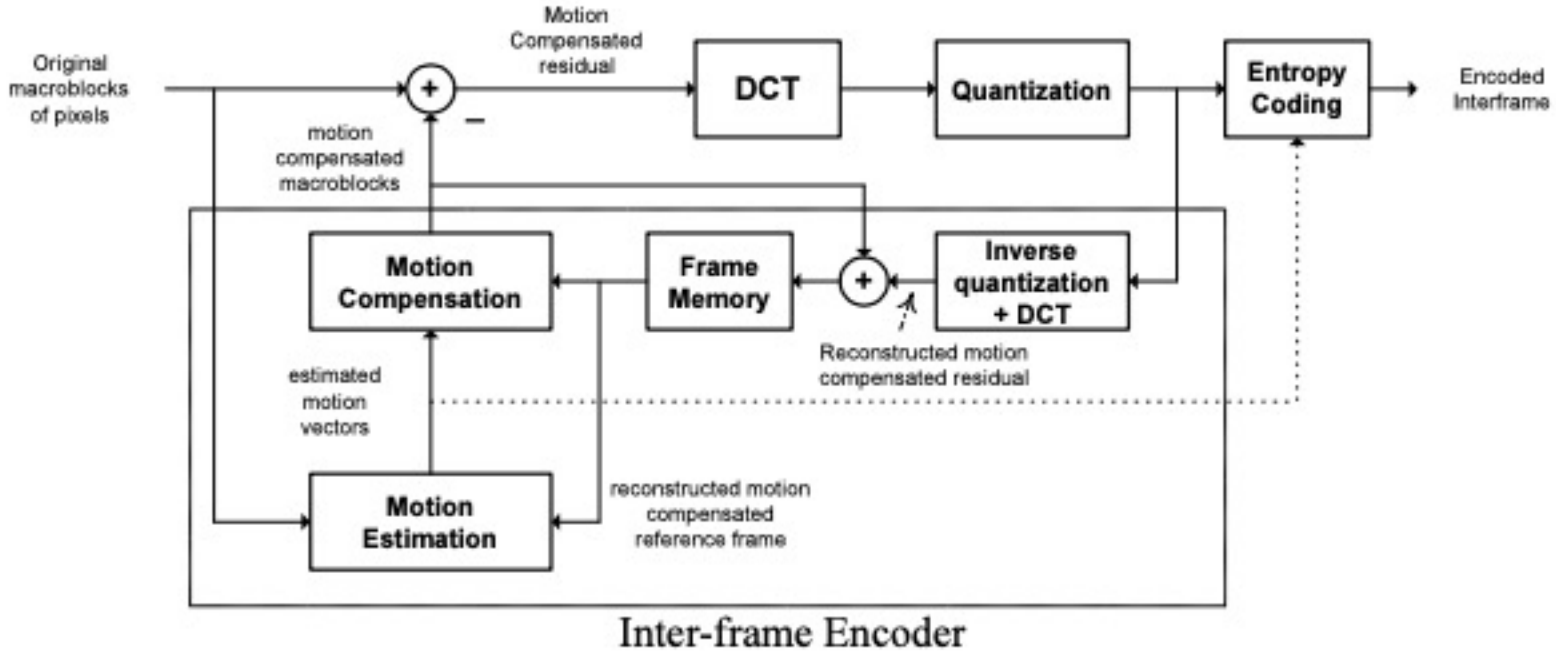
Homework Assignment #2 Extended!

- From 3/20 to 3/27

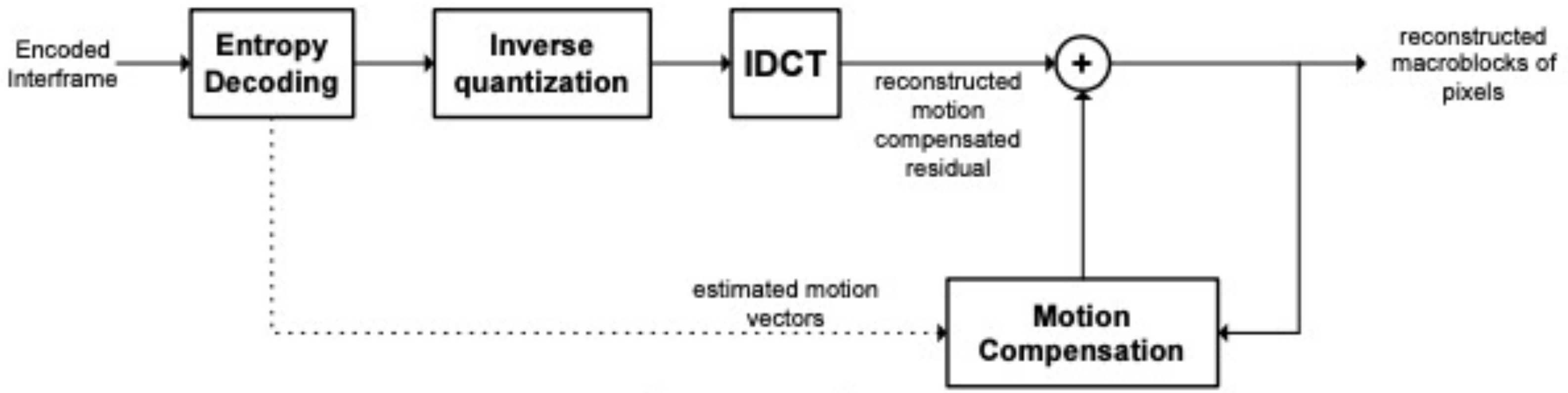
Motion Estimation

- Applications of Motion Estimation
- Phase Correlation
- Block Matching
- Spatio-Temporal Gradient Methods
- **Motion Estimation and Compensation**
- Learning-based Optical Flow

MPEG Video - Motion Estimation and Compensation

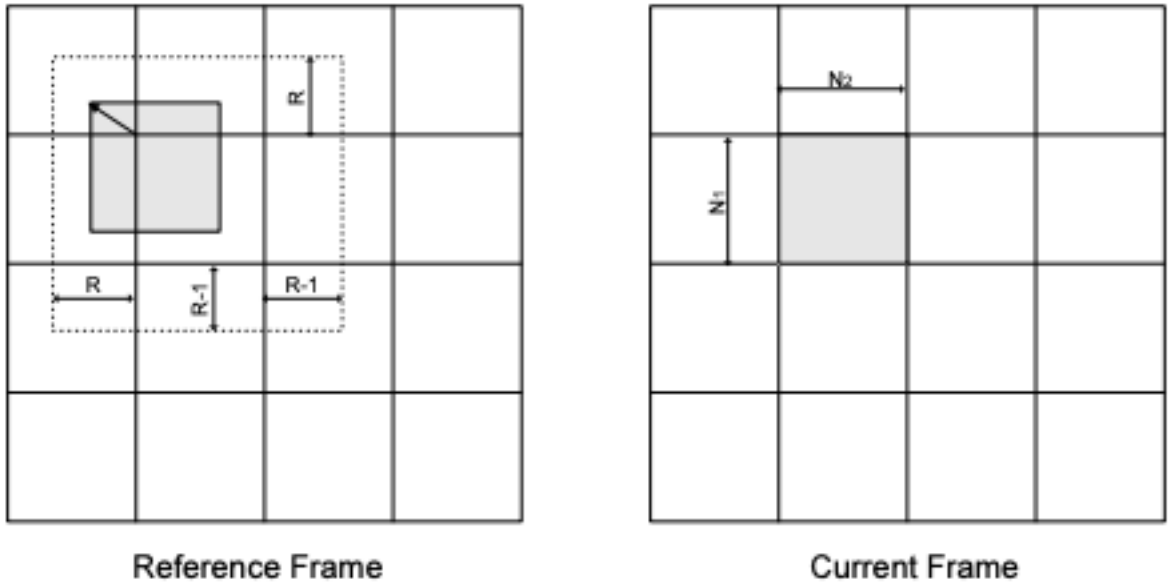


MPEG Video - Motion Estimation and Compensation



Inter-frame Decoder

Exhaustive Block Matching

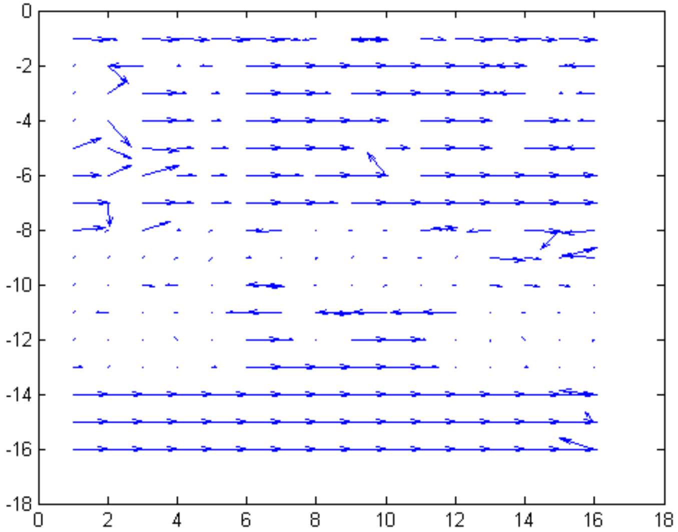


$$MAD(d_1, d_2) \equiv \frac{1}{N_1 N_2} \sum_{(n_1, n_2) \in B} \left| S_{ref}(n_1 + d_1, n_2 + d_2) - S_{cur}(n_1, n_2) \right|$$

Motion Estimation

Car 1

Car 2



Motion Vectors

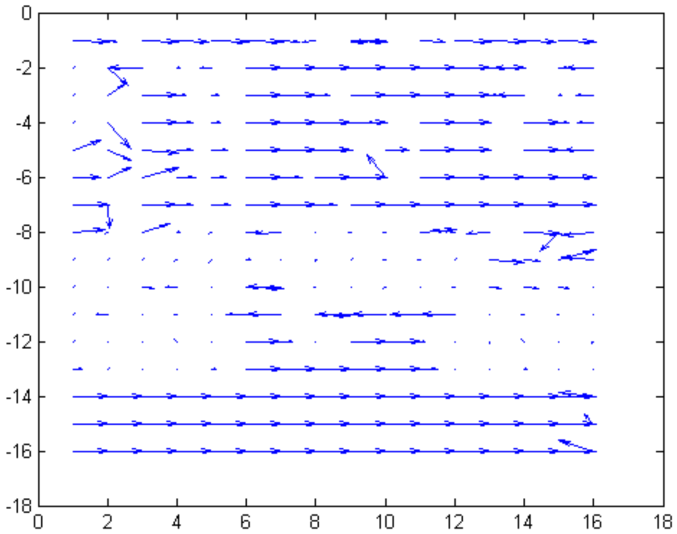
Motion Compensation

Actual Car

Estimated Car

Residual Car

Three-step Algorithm



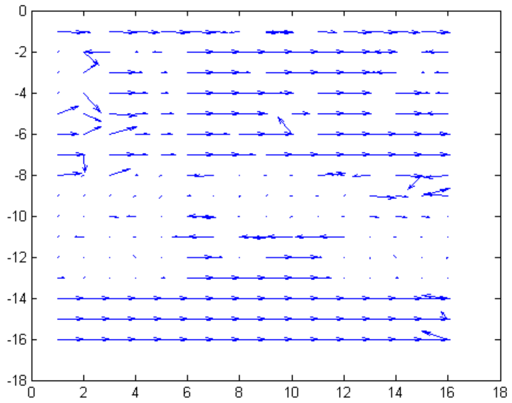
Motion Vectors

Estimated Car

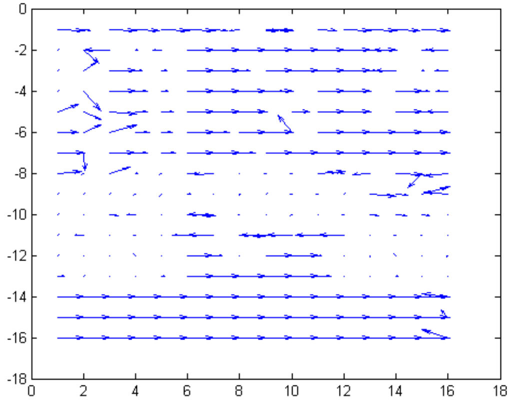
Residual Car

Advantages / Disadvantages between the Exhaustive Block Matching Algorithm and the Three-step Algorithm

Exhaustive Block Matching



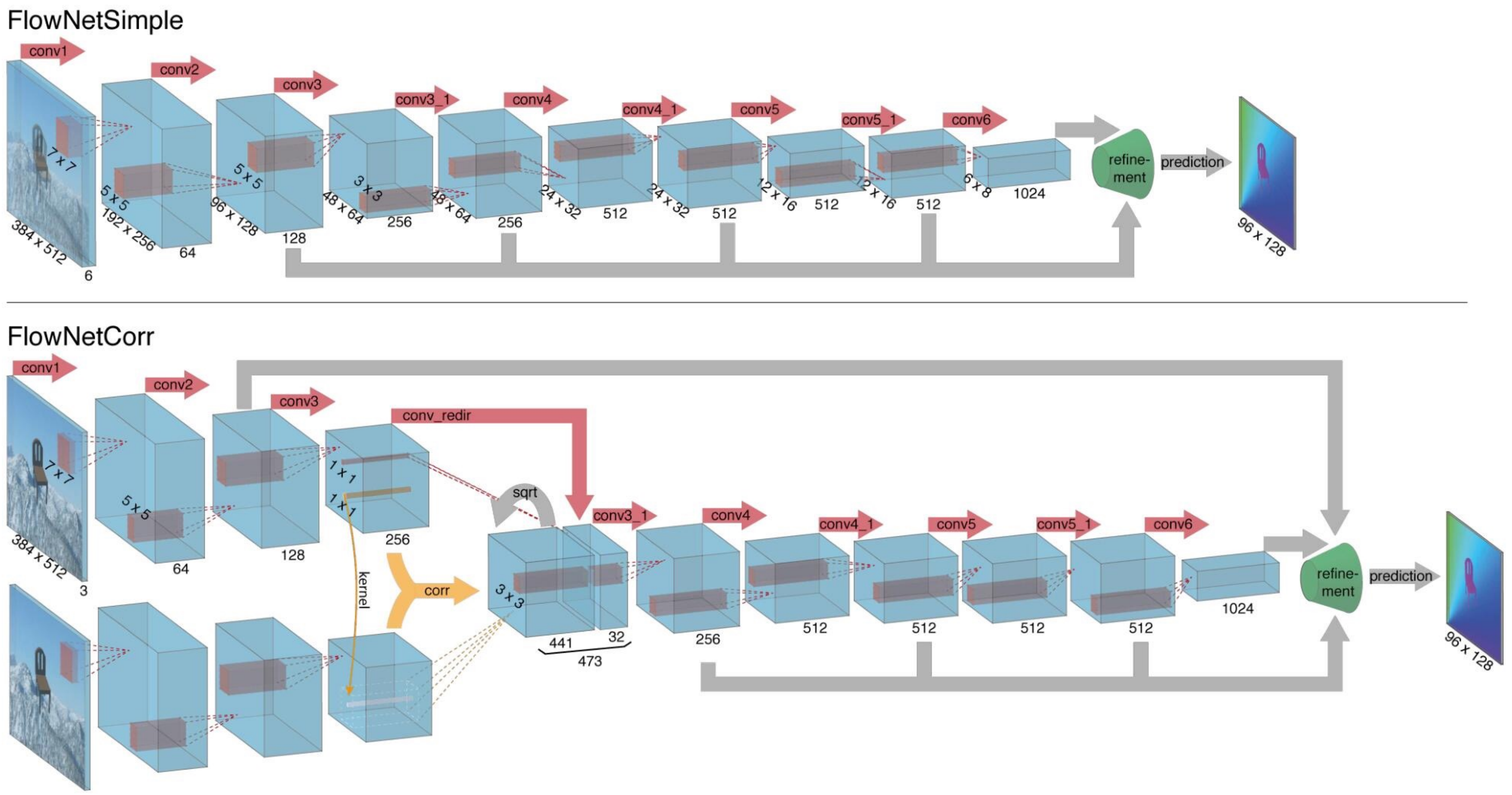
Three-step Algorithm



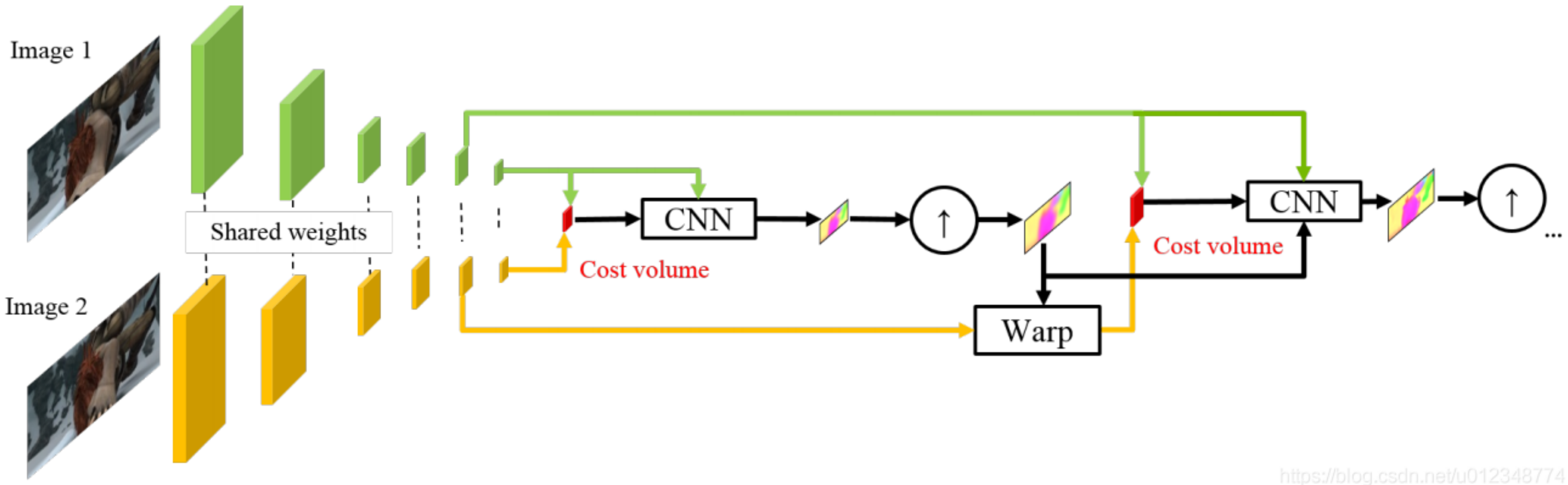
Motion Estimation

- Applications of Motion Estimation
- Phase Correlation
- Block Matching
- Spatio-Temporal Gradient Methods
- Motion Estimation and Compensation
- **Learning-based Optical Flow**

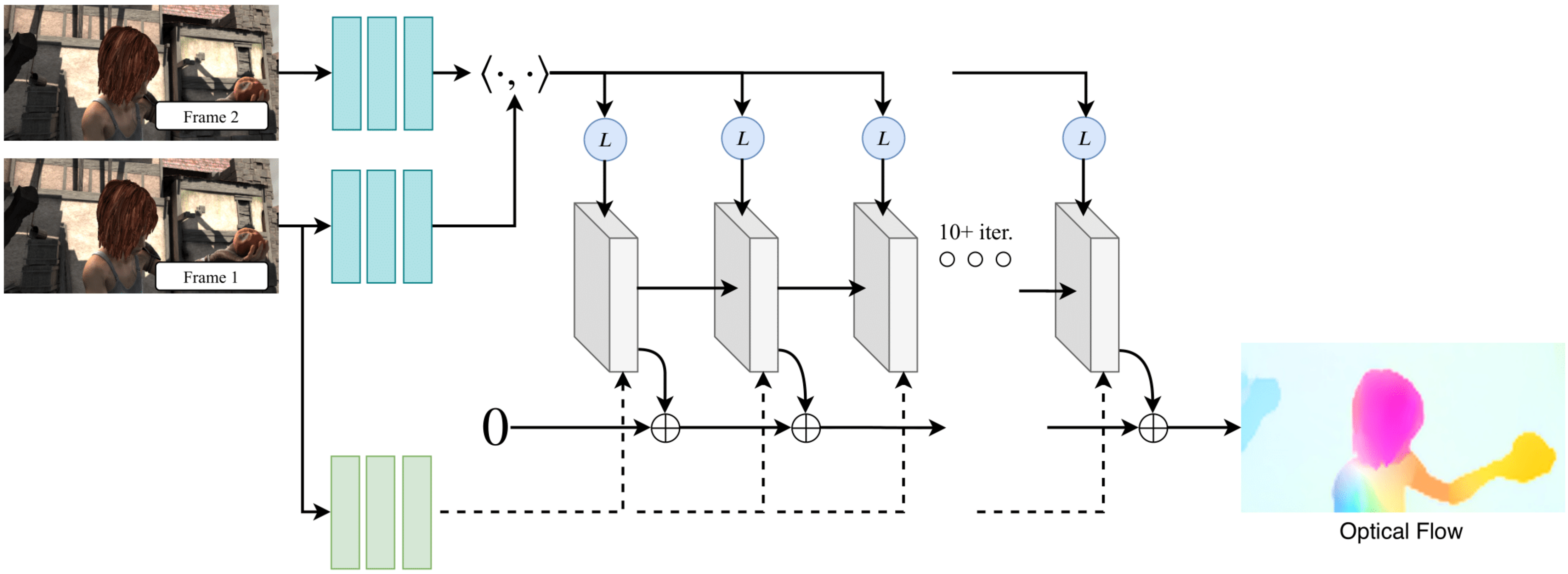
Optical Flow – FlowNet [ICCV 2015]



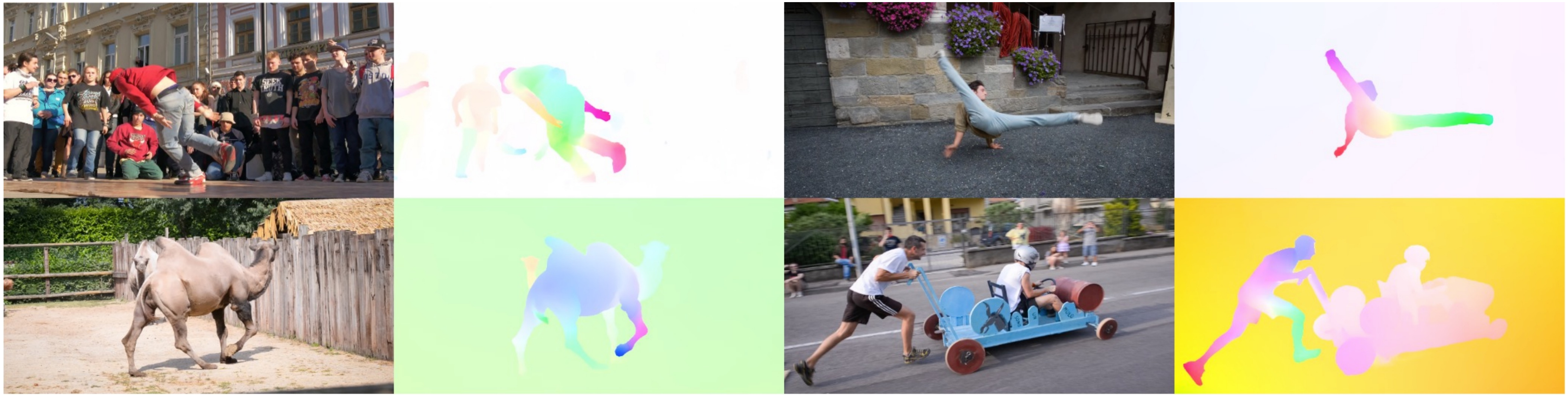
Optical Flow – PWC-Net [CVPR 2018]



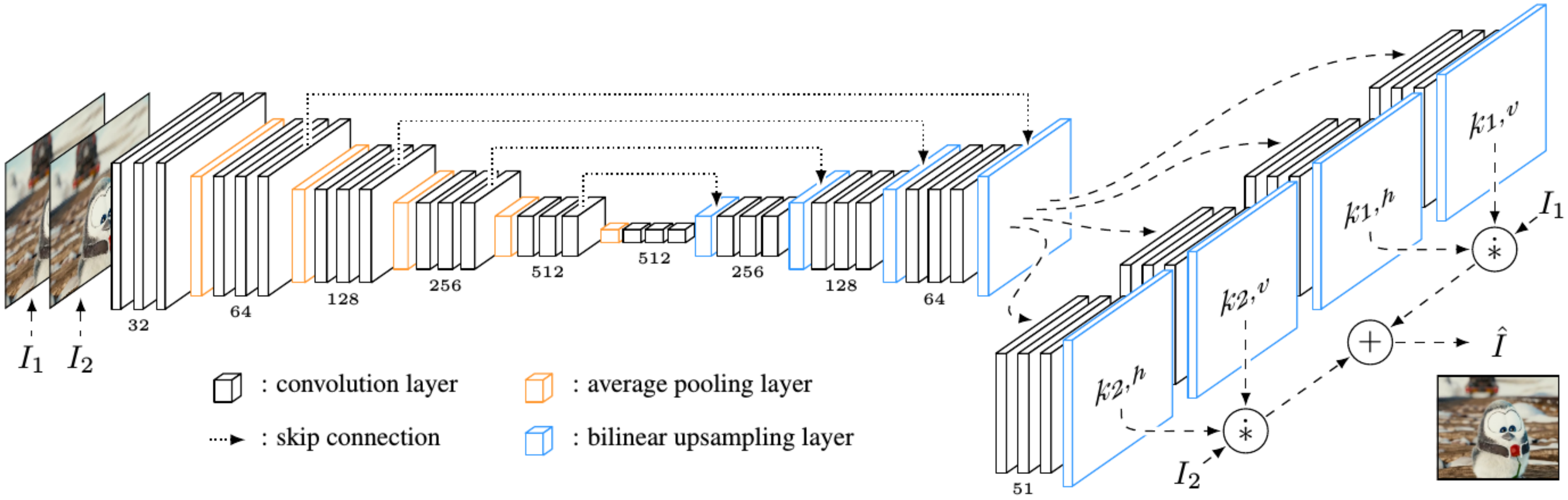
Optical Flow – RAFT [ECCV 2020]



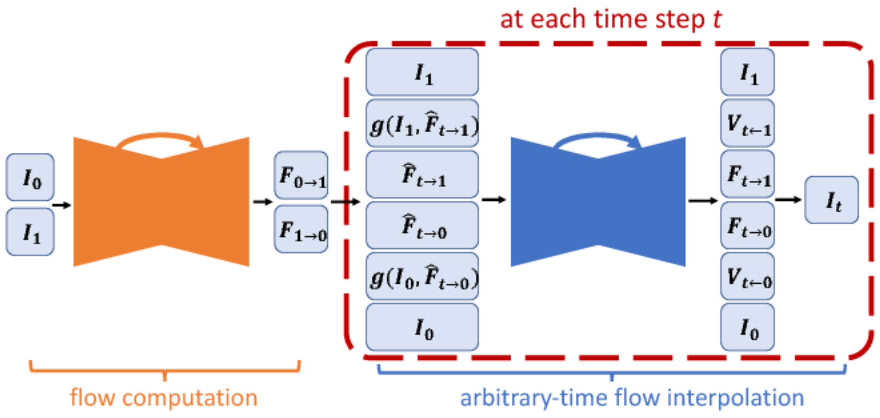
Optical Flow – RAFT [ECCV 2020]



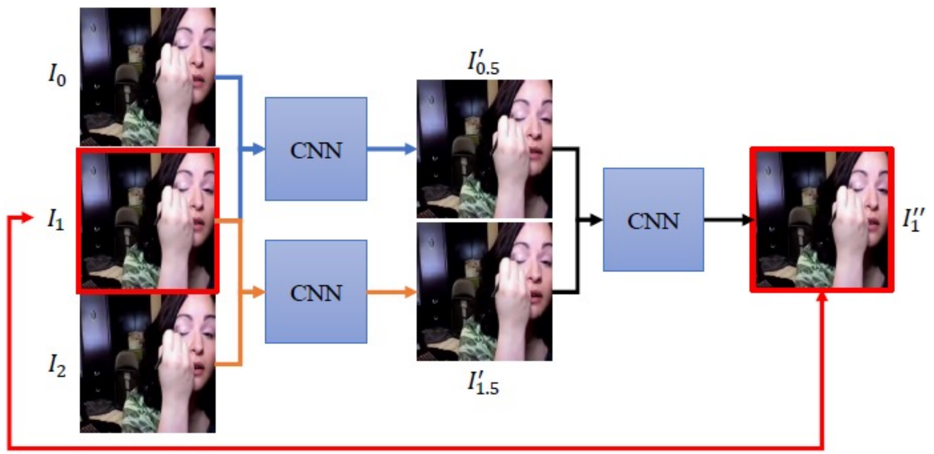
Frame Interpolation – SepConv [ICCV 2017]



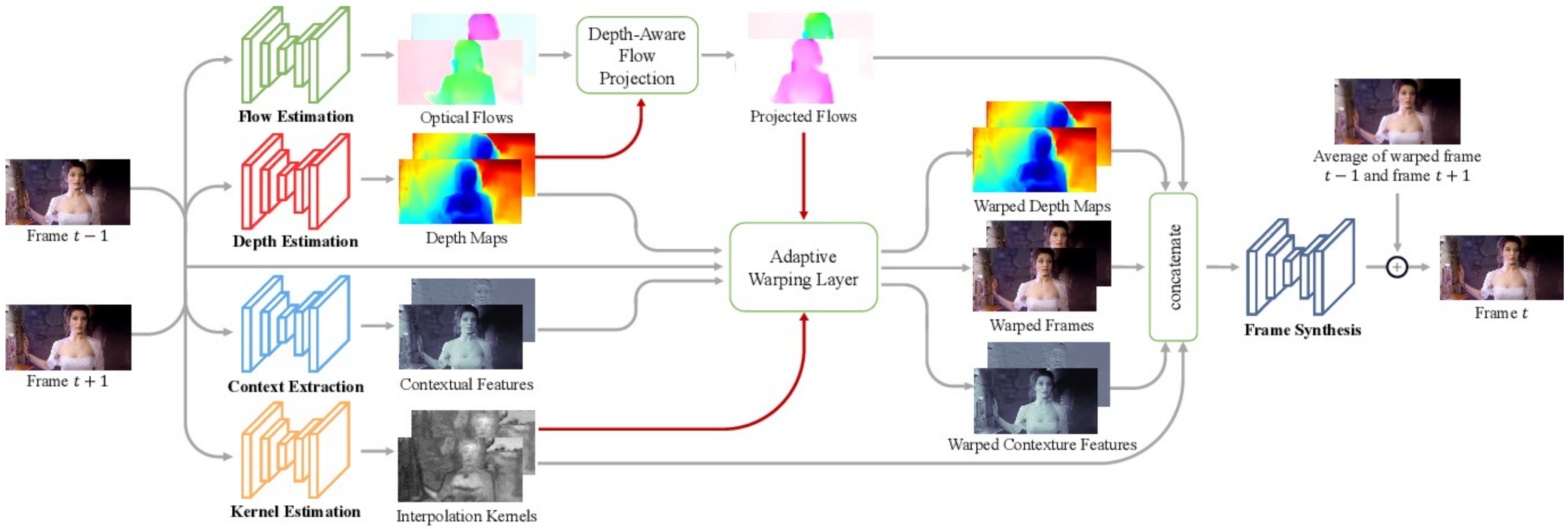
Frame Interpolation – Super SloMo [CVPR 18]



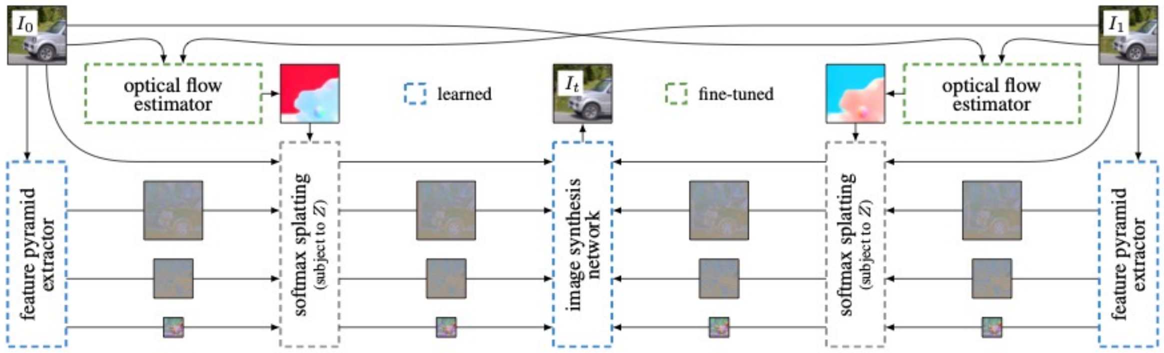
Frame Interpolation – CyclicGen [AAAI 2019]



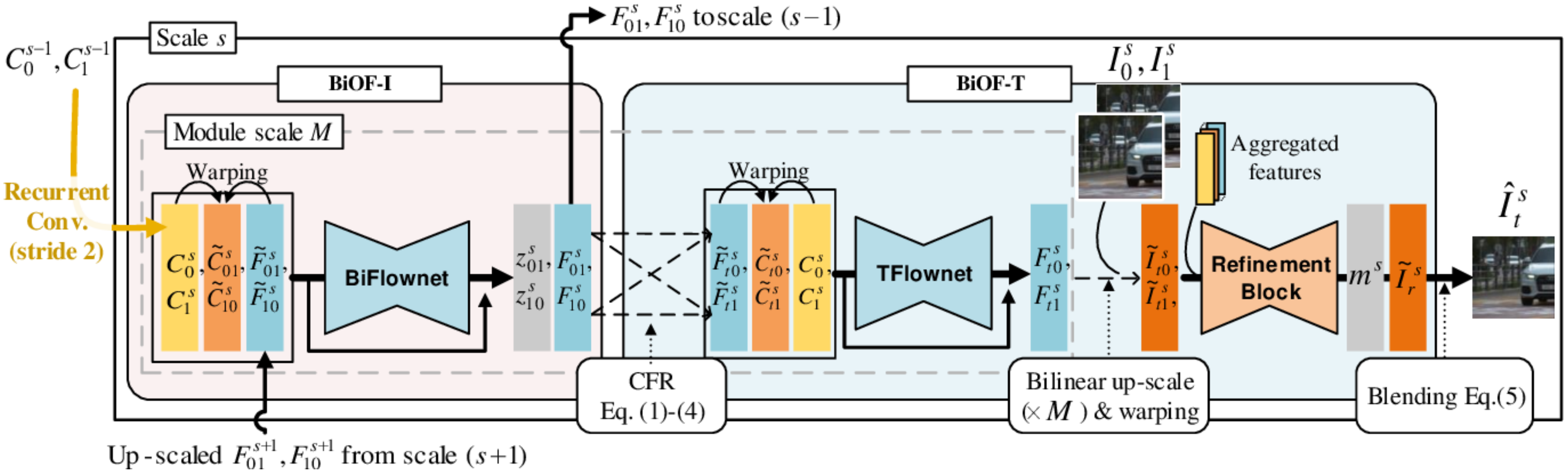
Frame Interpolation – DAIN [CVPR 2019]



Frame Interpolation – SoftSplat [CVPR 2020]



Frame Interpolation – XVFI [ICCV 2021]



Frame Interpolation – XVFI [ICCV 2021]