


Multiple View Image Denoising using 3D Focus Image Stacks

Shiwei Zhou, Yu Hen Hu, Hongrui Jiang

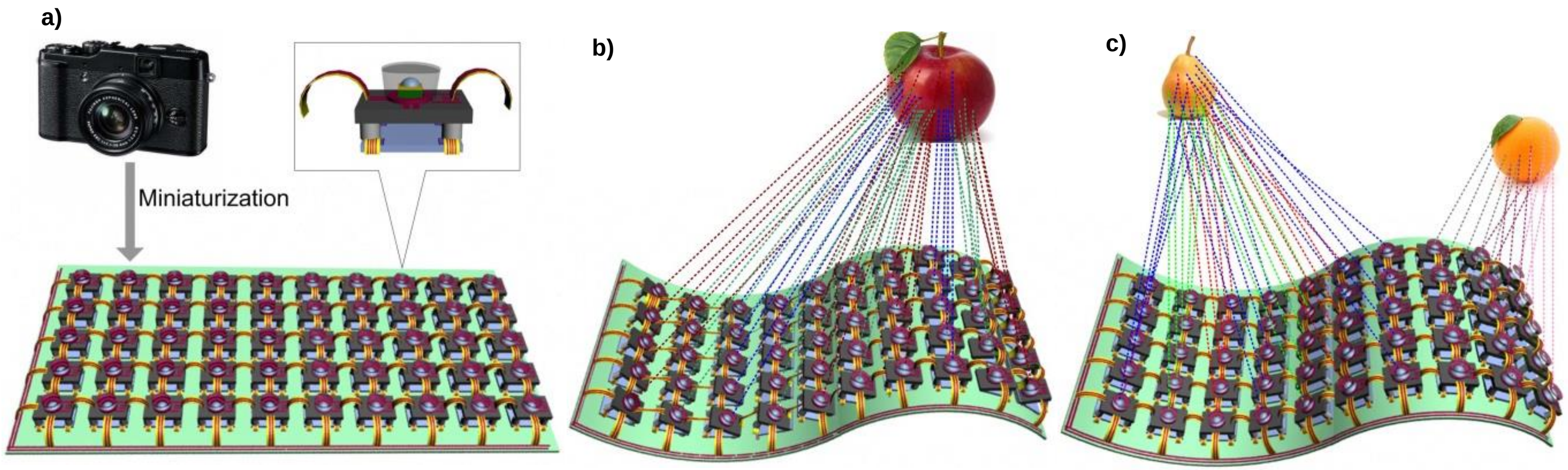
University of Wisconsin-Madison

Outline

- Introduction
- Related Work
- Denoising Scheme
  - 3D Focus Image Stack
  - Disparity Map Estimation
  - Preliminary Denoising
  - Reliability Map
  - Handling Unreliable Pixels
- Experiments

Motivation

- CPS project: smart flexible camera sheet [6]
  - Ultra-thin semantic-guided cooperative micro-camera array



Introduction

- Multiple view imaging

Introduction

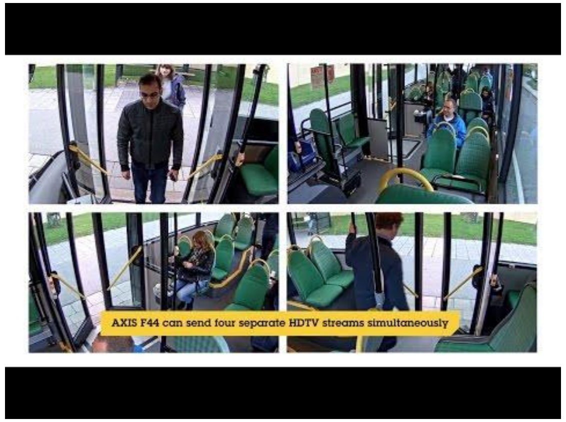
- Multiple view imaging
  - 3D scene reconstruction

Introduction

- Multiple view imaging
  - 3D scene reconstruction
  - Object tracking and recognition

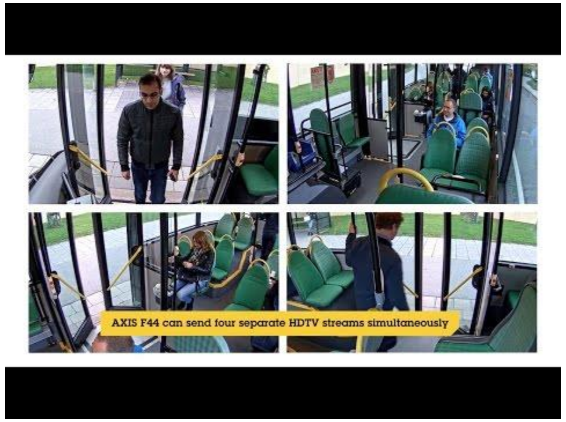
Introduction

- Multiple view imaging
  - 3D scene reconstruction
  - Object tracking and recognition
  - Environmental surveillance



Introduction

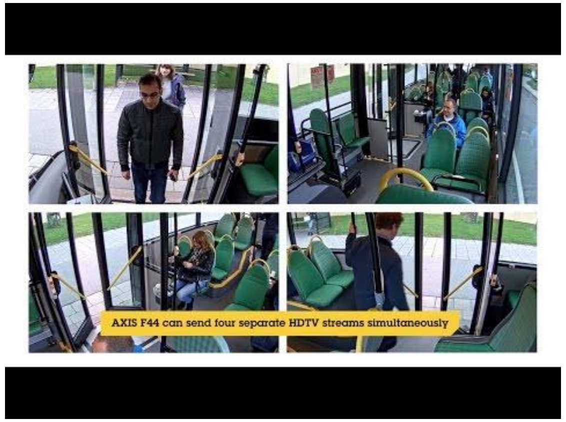
- Multiple view imaging
  - 3D scene reconstruction
  - Object tracking and recognition
  - Environmental surveillance
  - 3DTV
  - Etc.



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Introduction

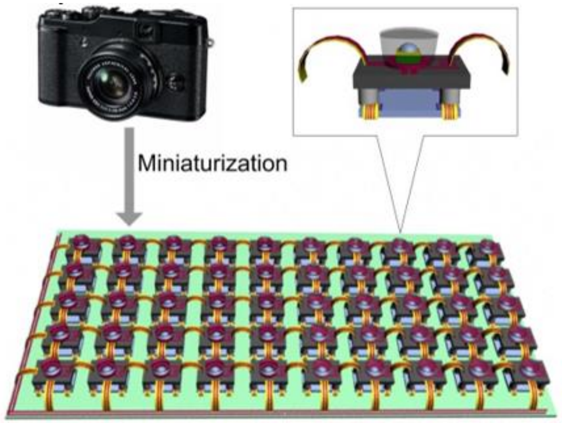
- Multiple view imaging
  - 3D scene reconstruction
  - Object tracking and recognition
  - Environmental surveillance
  - 3DTV
  - Etc.



- Multi-views: helps exploit redundancy and 3D information

Introduction

- Capture: array of cameras
  - Small miniature cameras
  - Limited exposure – avoid motion blur
  - Small aperture – large depth of field

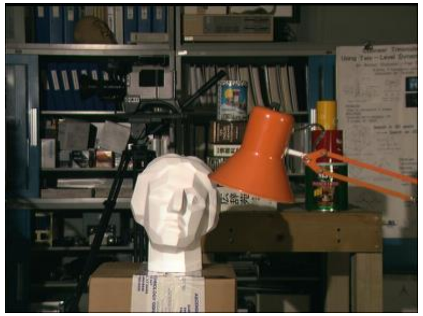


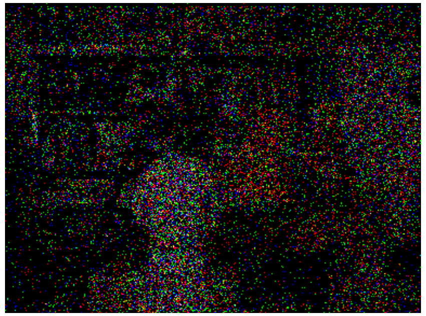
CPS smart flexible camera sheet

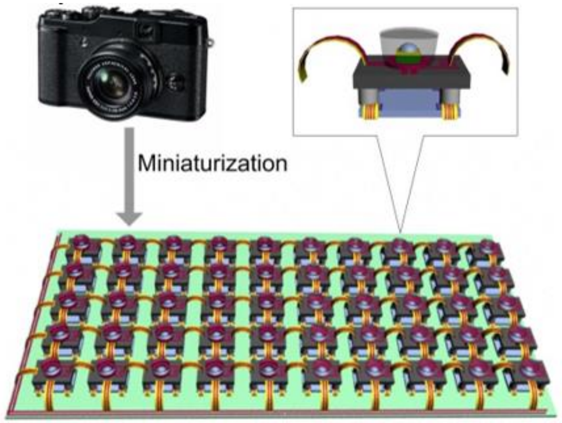
The Stanford Multi-camera Array

Introduction

- Capture: array of cameras
  - Small miniature cameras
  - Limited exposure – avoid motion blur
  - Small aperture – large depth of field
- Drawback: noisy image







CPS smart flexible camera sheet

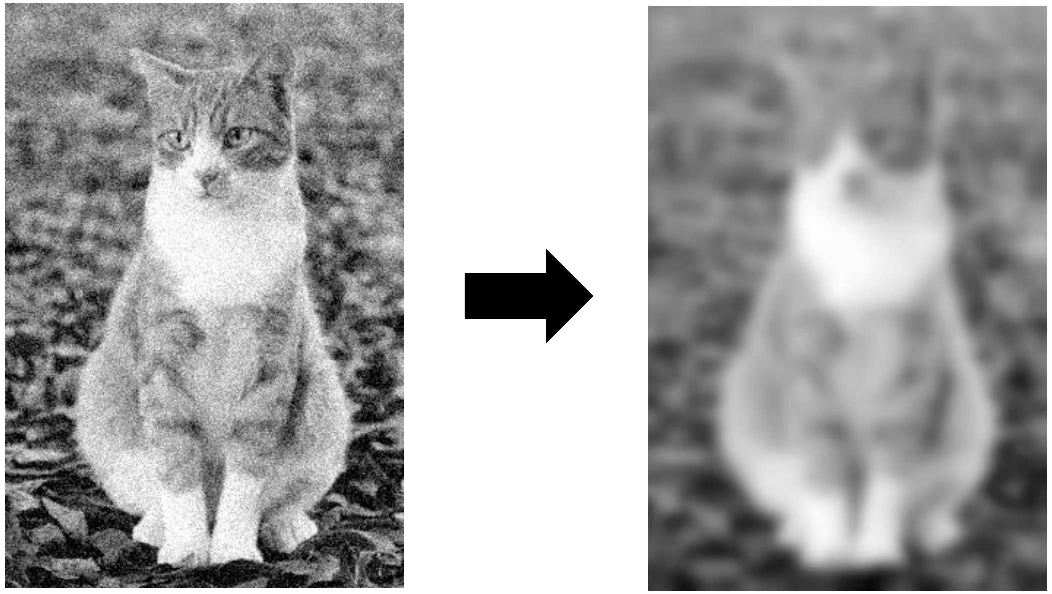
The Stanford Multi-camera Array

Related Work

- Single view image denoising

Related Work

- Single view image denoising

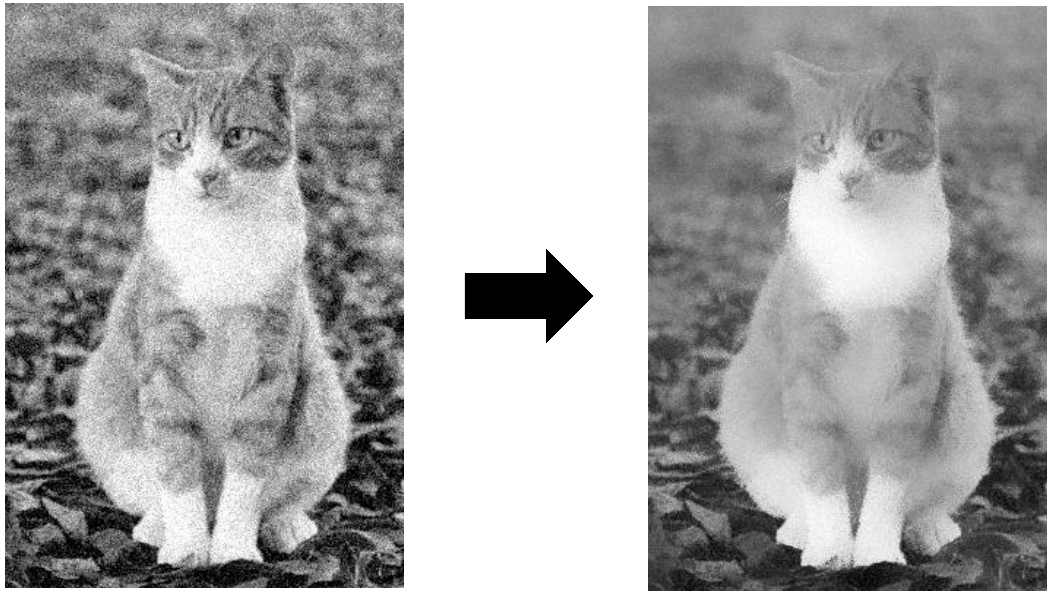


Gaussian Filtering

Source: http://graphics.stanford.edu/courses/cs478/lectures/02132012_denoising.pdf

Related Work

- Single view image denoising

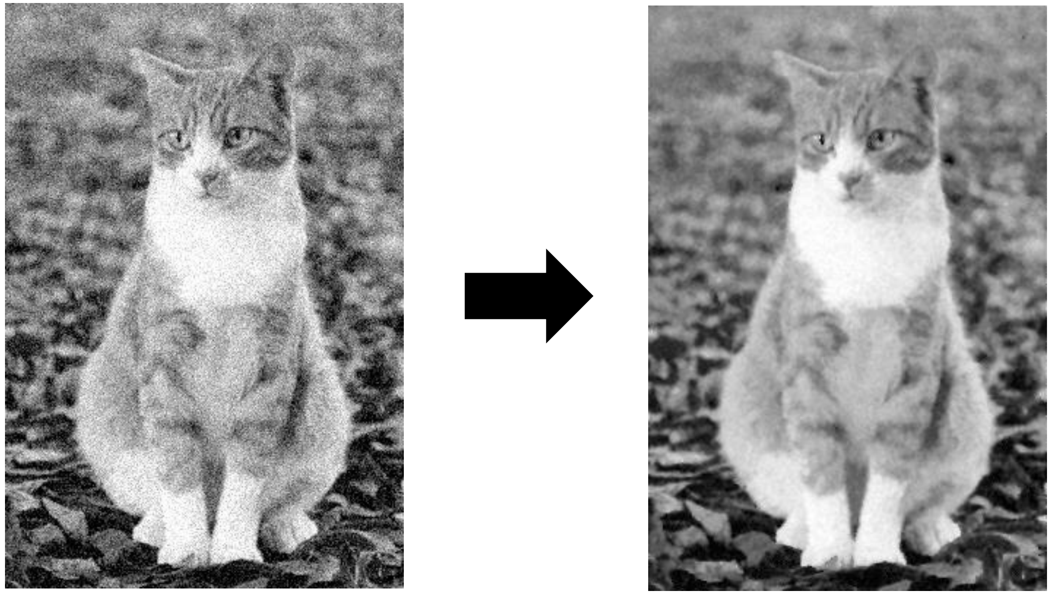


Bilateral Filtering

Source: http://graphics.stanford.edu/courses/cs478/lectures/02132012_denoising.pdf

Related Work

- Single view image denoising

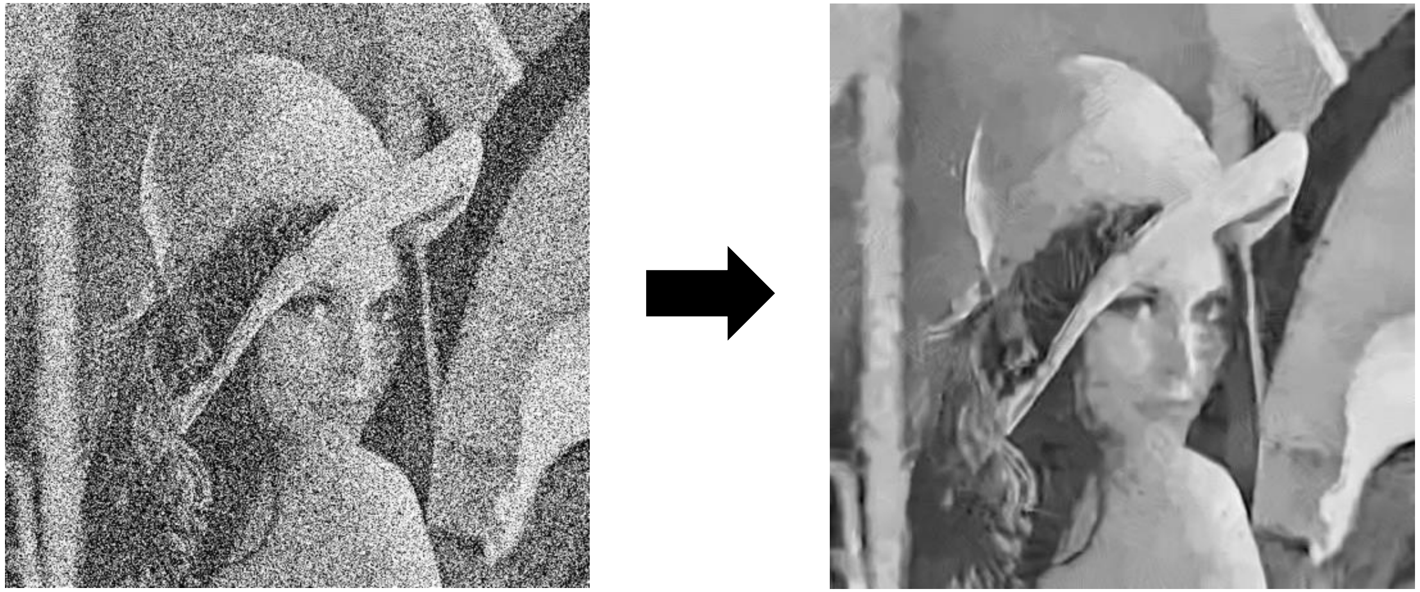


Non-local Means

Source: http://graphics.stanford.edu/courses/cs478/lectures/02132012_denoising.pdf

Related Work

- Single view image denoising



Block Matching 3D

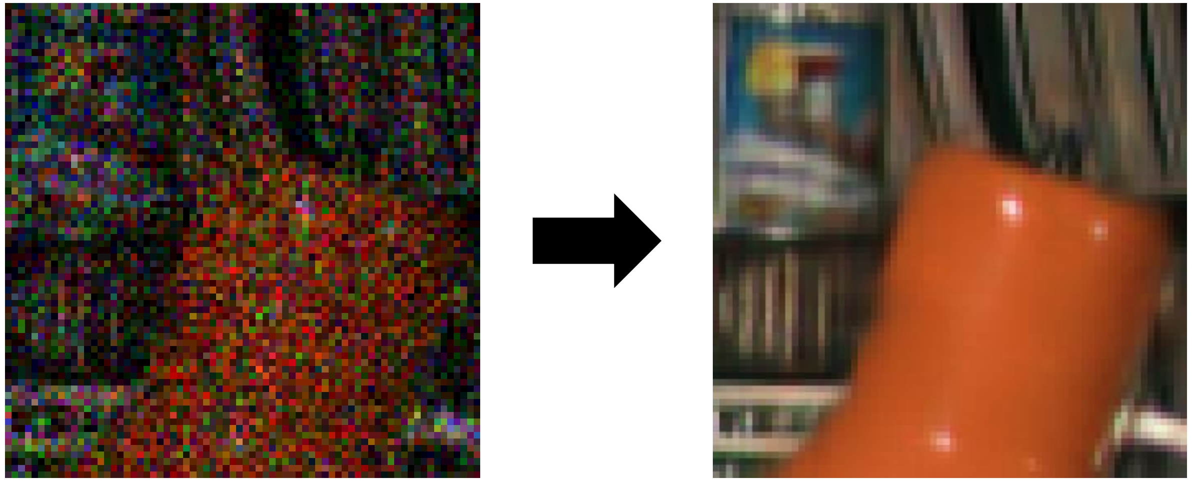
Source: http://graphics.stanford.edu/courses/cs478/lectures/02132012_denoising.pdf

Related Work

- Multi-view image denoising

Related Work

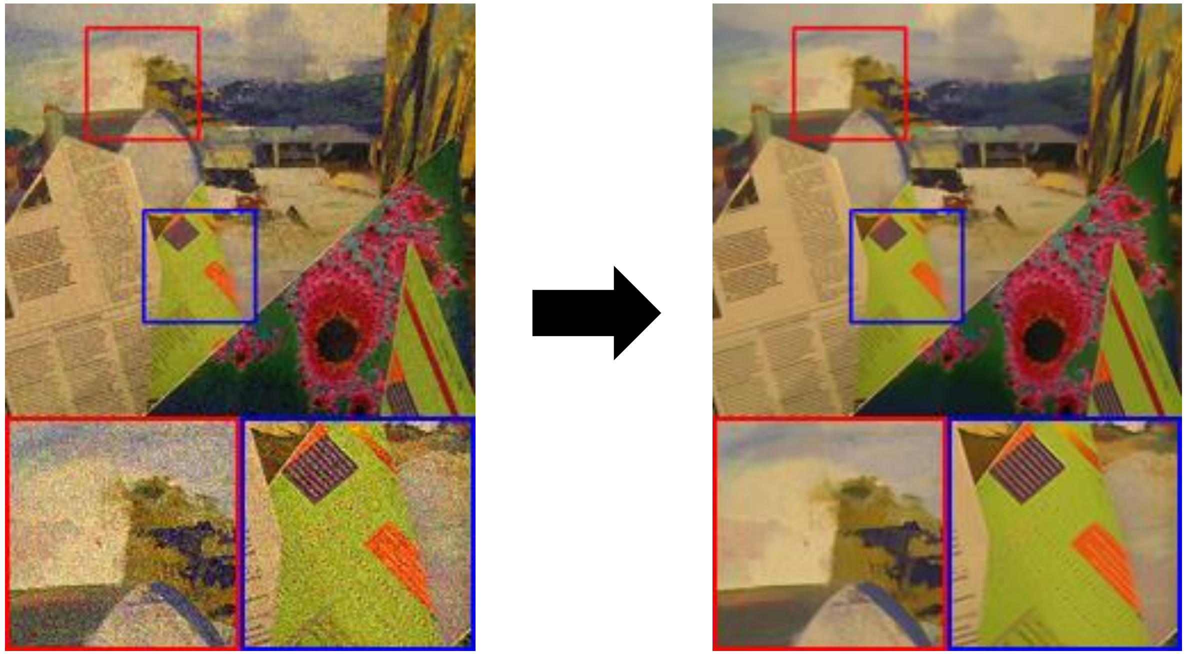
- Multi-view image denoising



Zhang *et al*., "Multiple View Image Denoising"

Related Work

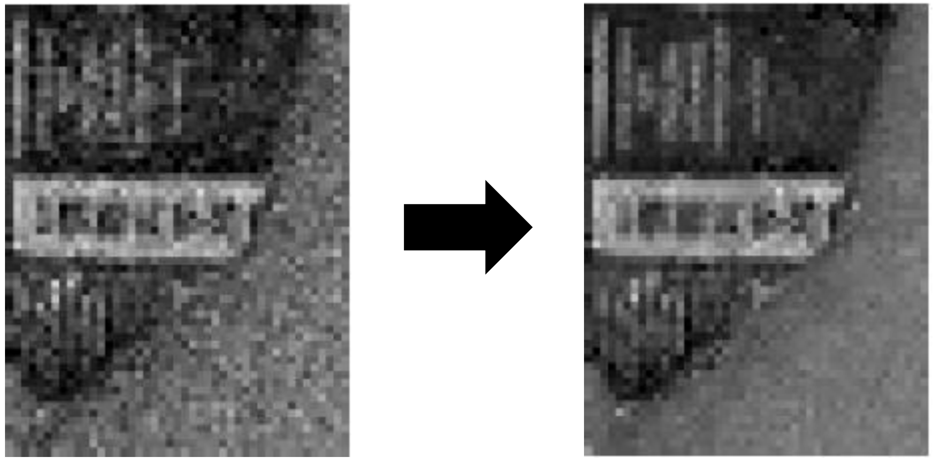
- Multi-view image denoising



Luo *et al*., “Adaptive Non-local Means for Multiview Image Denoising: Search for the Right Patches via a Statistical Approach”

Related Work

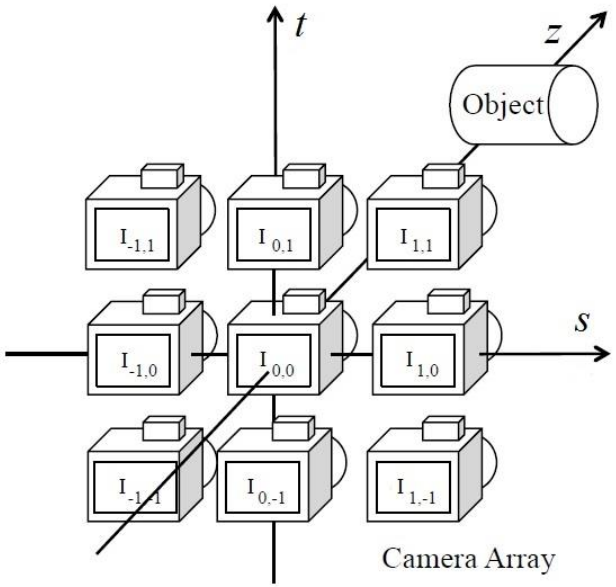
- Multi-view image denoising



Miyata *et al*., “Fast Multi-view Image Denoising Based on Image Reconstruction by Plane Sweeping”

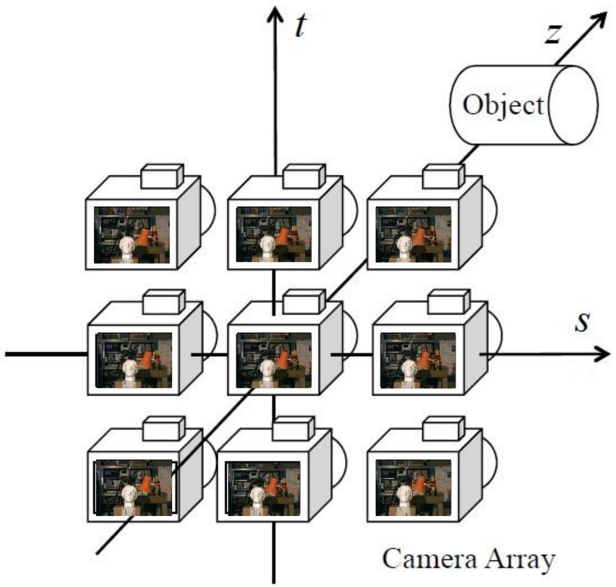
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



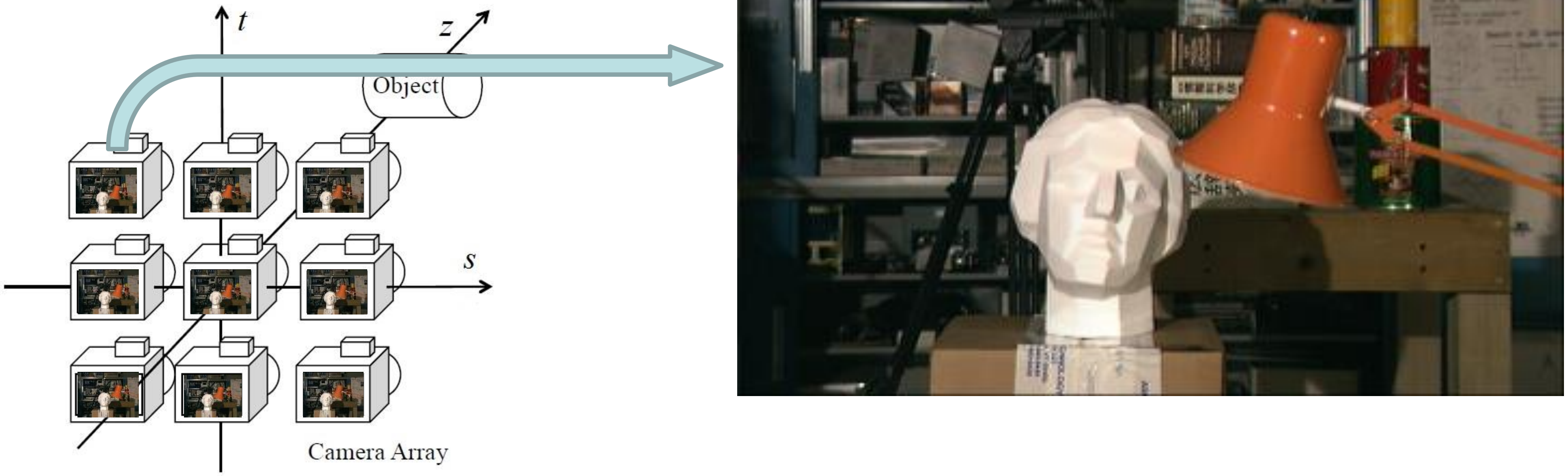
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



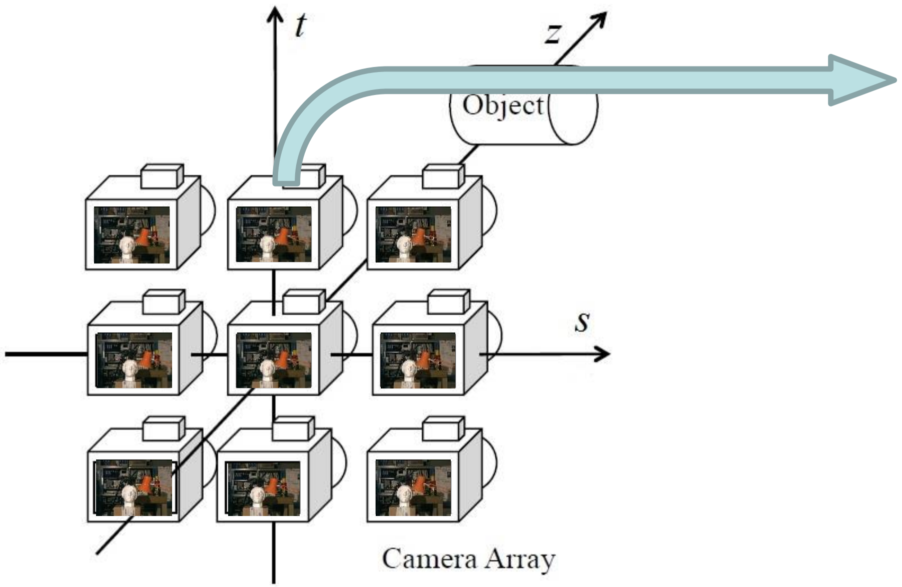
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



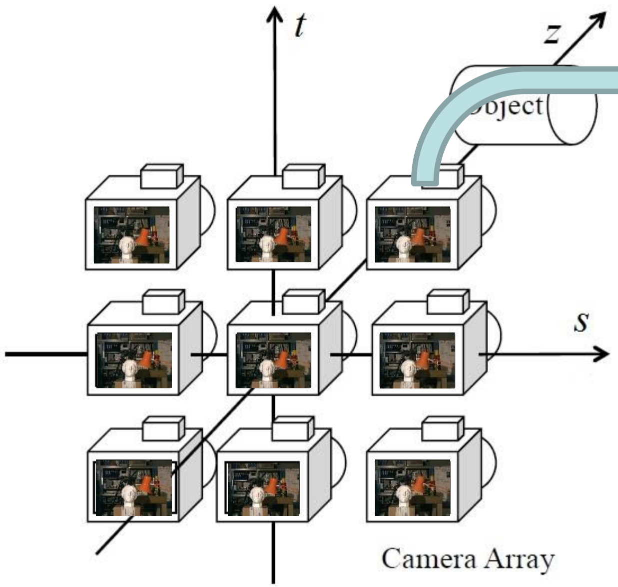
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



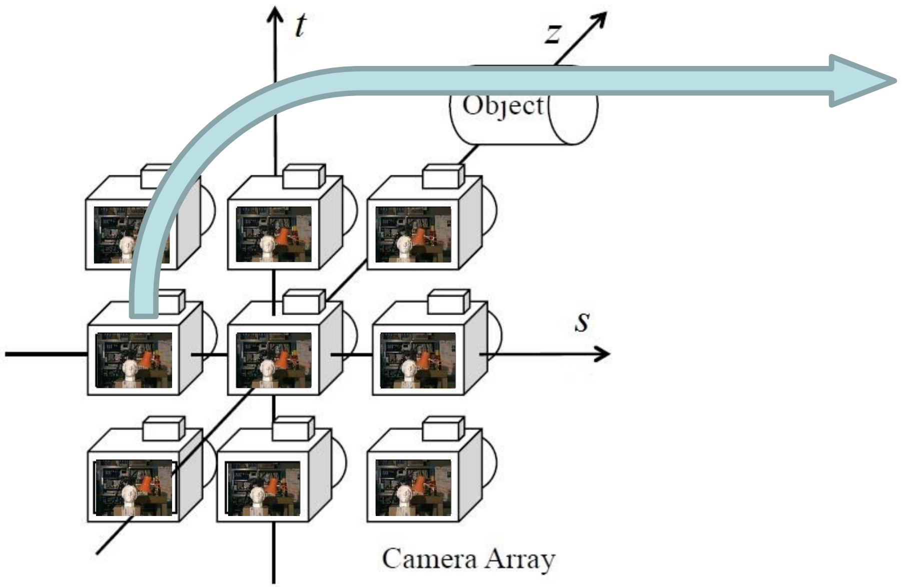
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



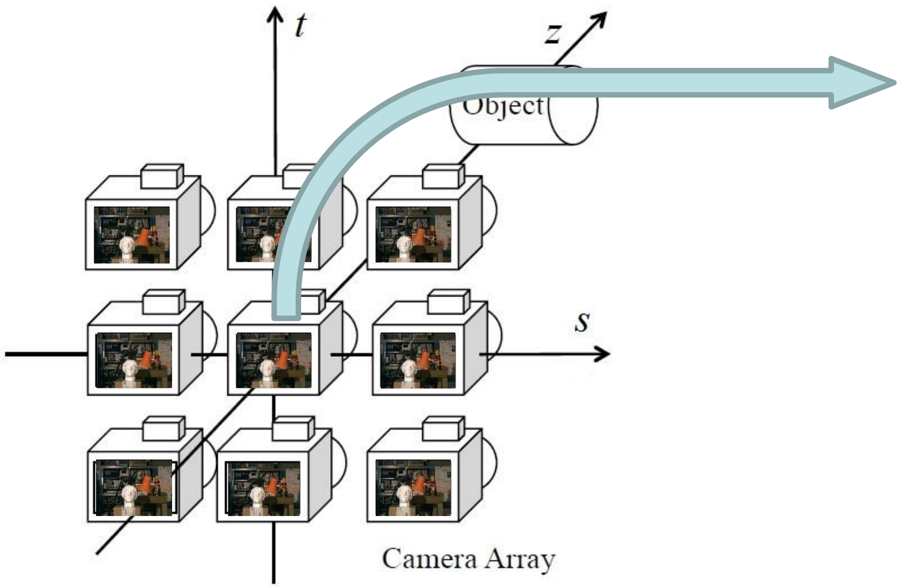
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



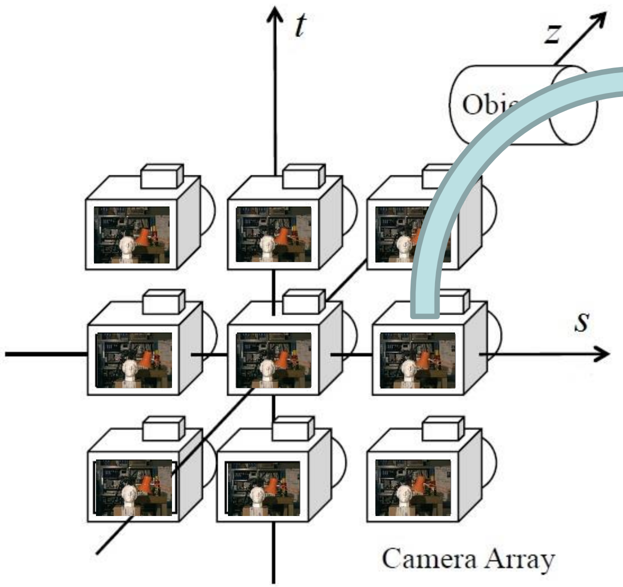
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



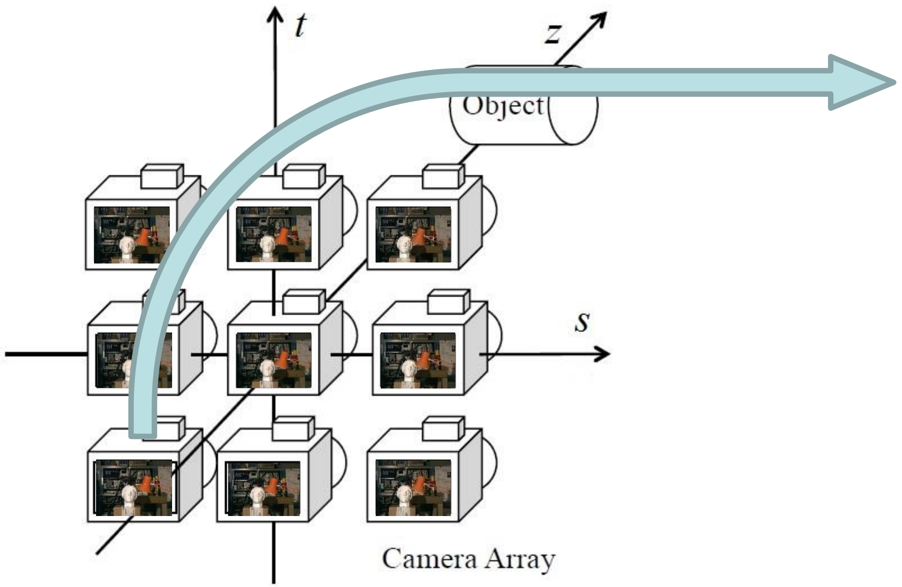
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



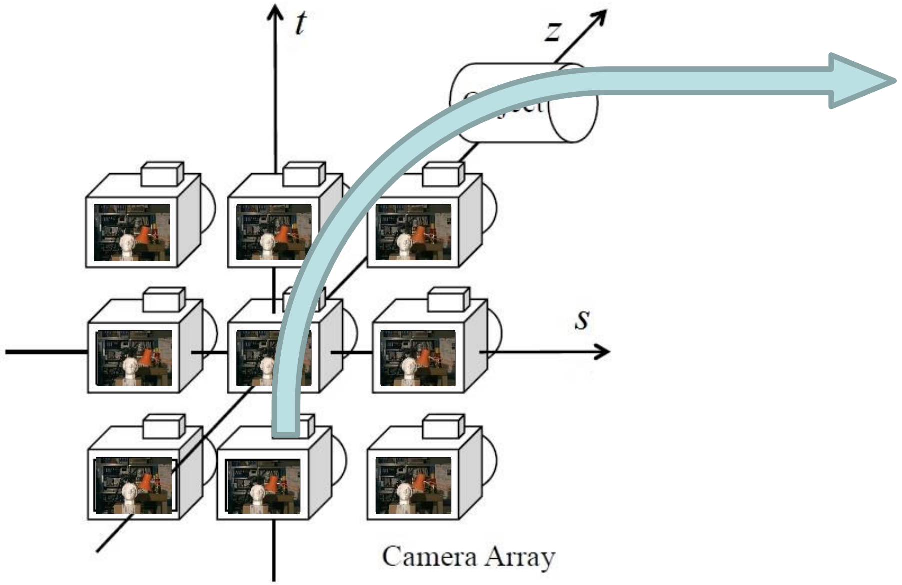
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



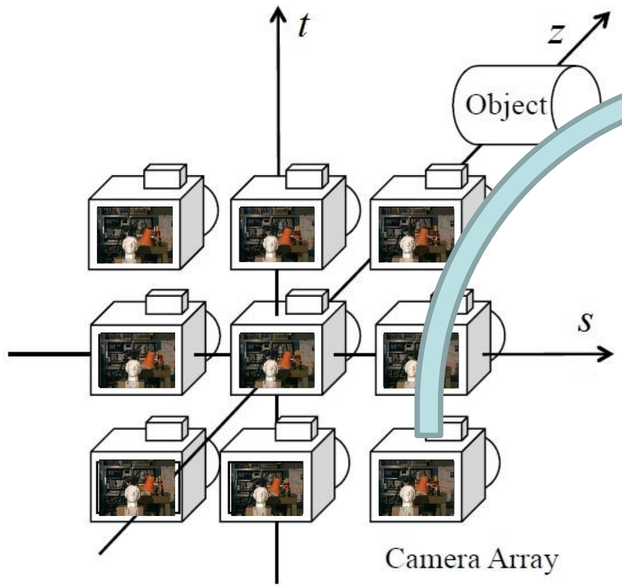
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates



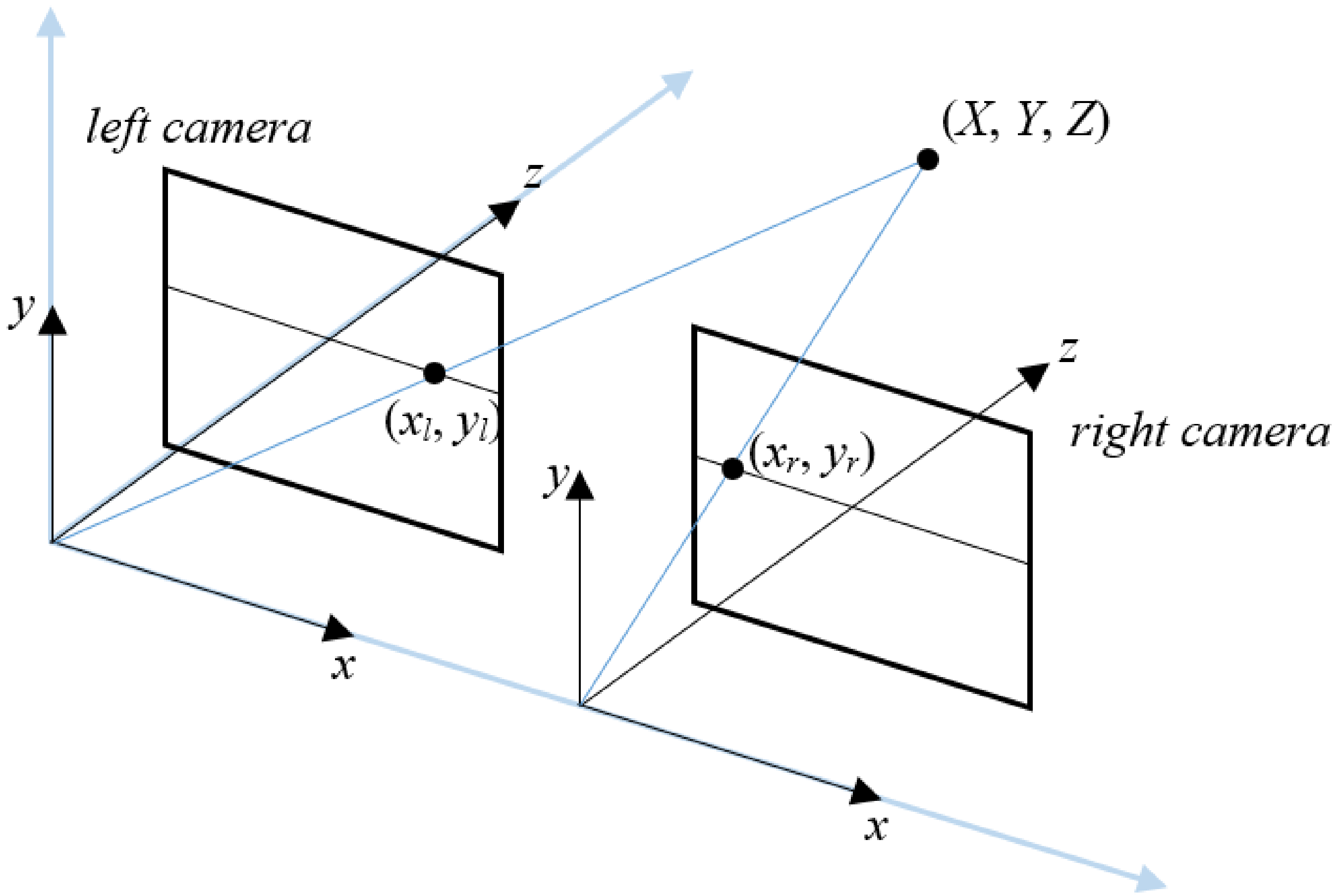
Multi-view Images

- Multi-view images $I_{s,t}(x, y)$
  - x, y – image coordinates
  - s, t – camera coordinates

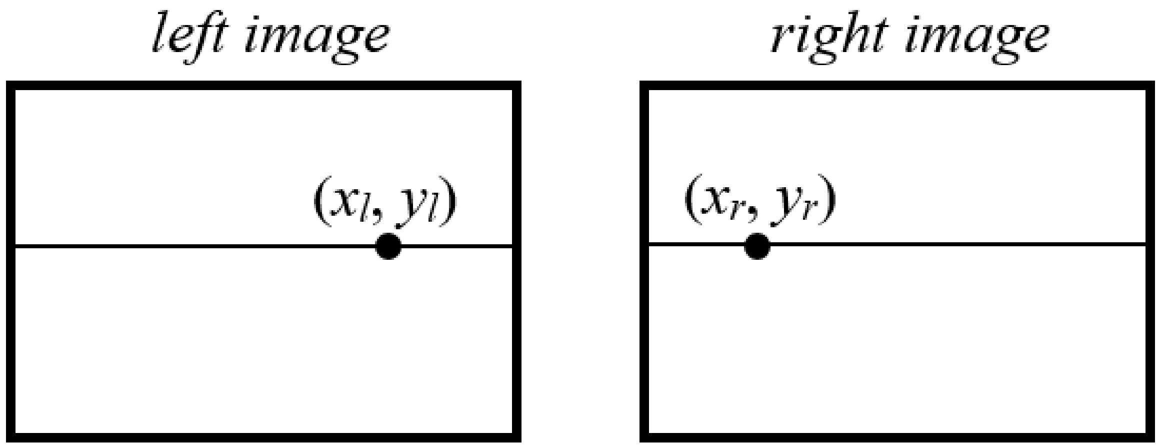


Disparity

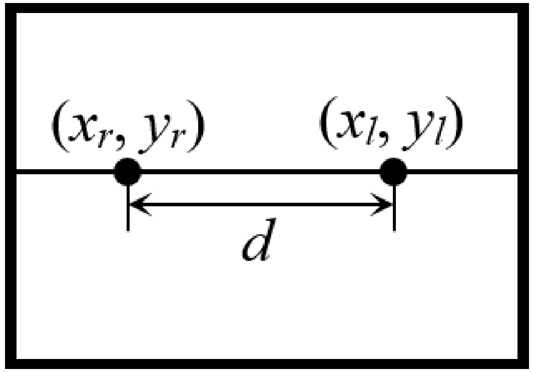
Disparity



Disparity



Disparity

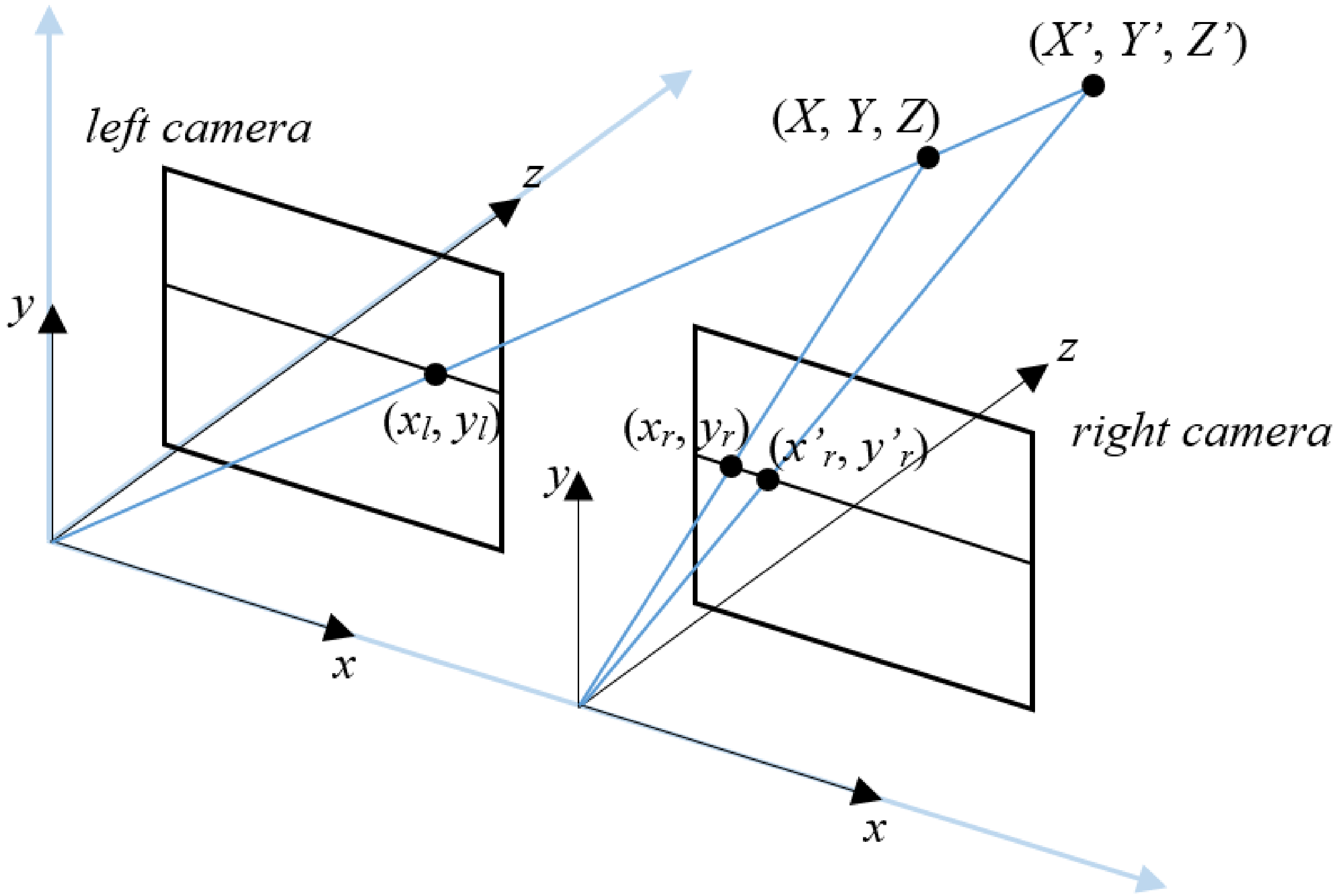


Disparity

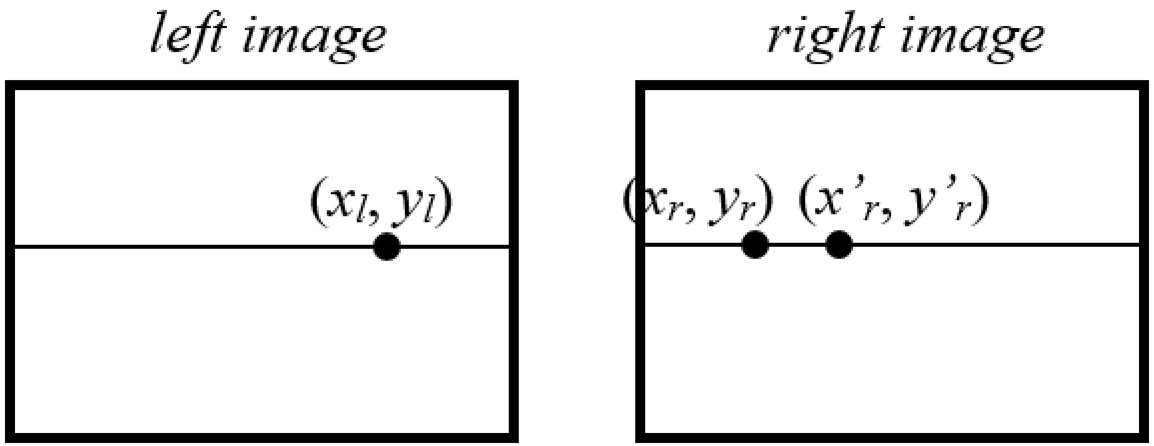
- Disparity: the distance between corresponding points in the left and right image

Disparity

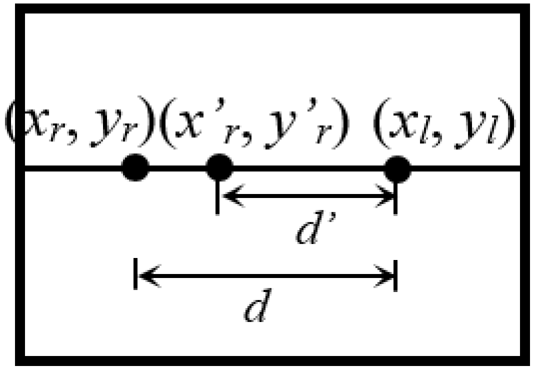
Disparity



Disparity



Disparity



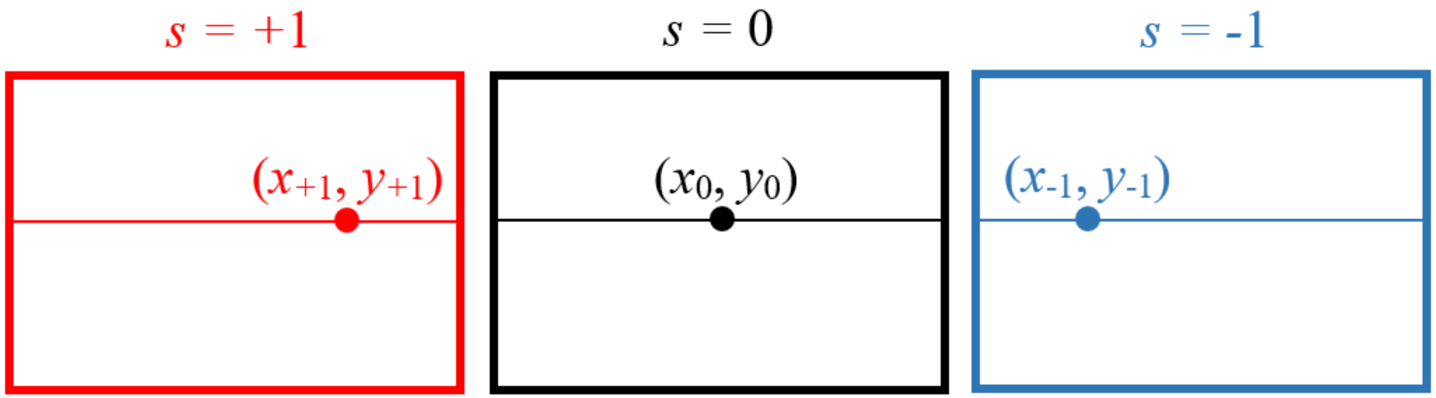
Disparity

- Disparity is inversely proportional to depth

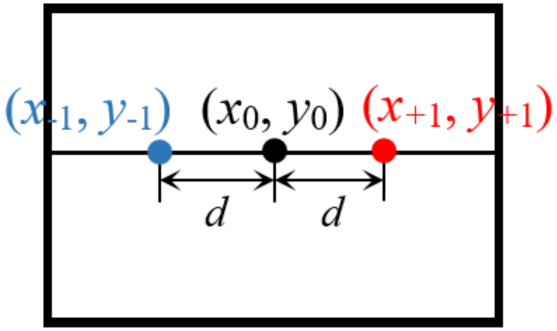
Denoising Scheme

Acquire multi-view images of the same scene → Construct 3D focus image stacks → Disparity map estimation → Preliminary denoising using disparity map → Compute reliability map → Handle unreliable pixels

3D Focus Image Stacks



3D Focus Image Stacks

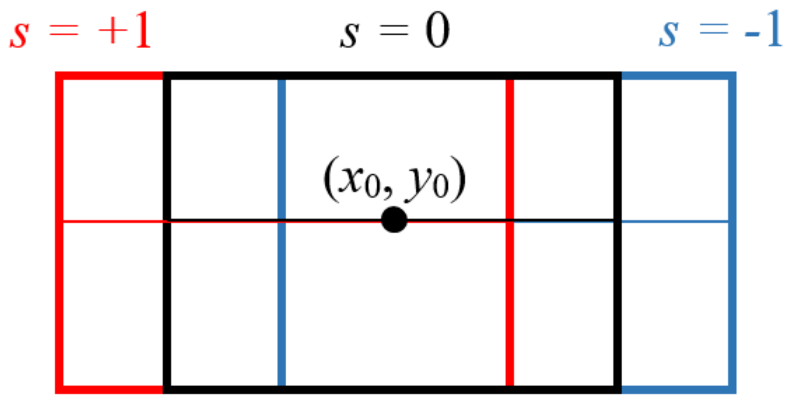


3D Focus Image Stacks

- For each disparity *d*,
  - Translate each pixel using

$$I_{s,t}^{d}(x,y) = I_{s,t}(x + (s - s_0)d, y + (t - t_0)d)$$

- Pixels with correct disparity,
  - Corresponding pixels are stacked
  - Clear image, a.k.a. in-focus



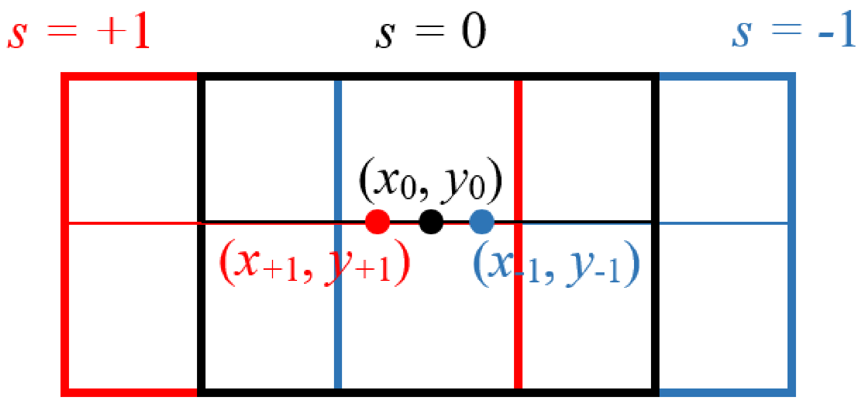
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

3D Focus Image Stacks

- For each disparity *d*,
  - Translate each pixel using

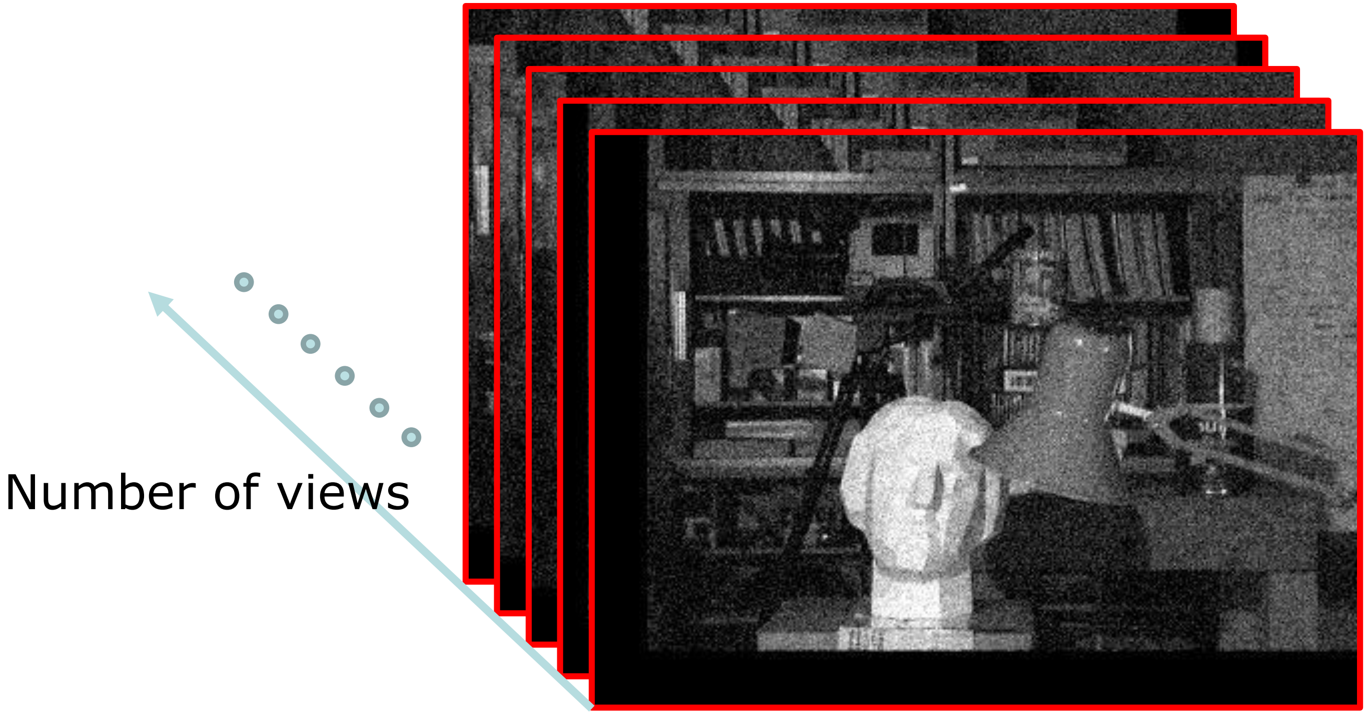
$$I_{s,t}^{d}(x,y) = I_{s,t}\left(x+(s-s_0)d,\, y+(t-t_0)d\right)$$

- Pixels with incorrect disparity,
  - Corresponding pixels have displacement
  - Blurred image, a.k.a. out-of-focus



3D Focus Image Stacks

- Stack $I^{d}_{s,t}(x, y)$ into $F_d(x, y)$
  - F_d is called 3D focus image stacks



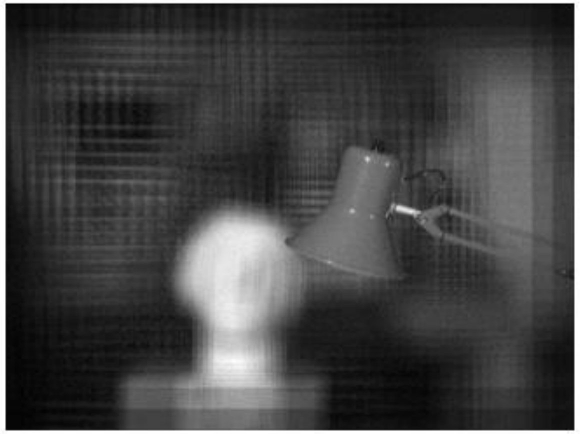
3D Focus Image Stack

- Visualization
  by simple averaging

Disparity = 5

Disparity = 6

Disparity = 10



Disparity = 14

Depth Estimation

- Similarity measure

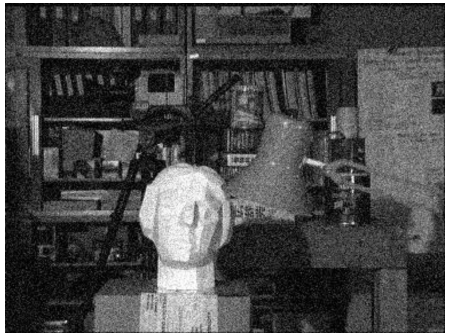
$$S_d(x,y) = \frac{1}{N}\sum_{k=1}^{N}\sum_{(i,j)\in W_{x,y}}\left|F_d(i,j,k) - I_{s_0,t_0}(i,j)\right|$$

  - *N* – Number of views in focus image stack
  - *k* – *k*th layer (view) of the focus image stack
  - $W_{i,j}$ – Window centered at pixel (*i*, *j*)

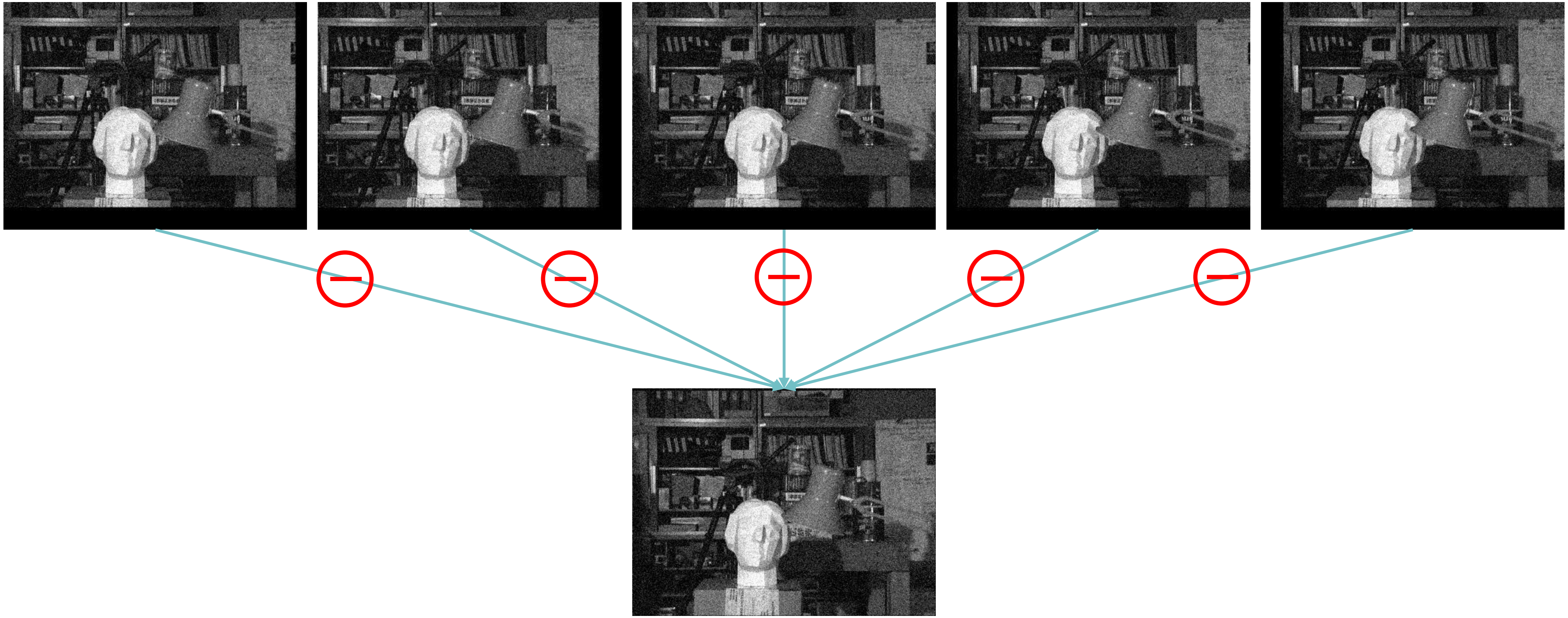
Depth Estimation

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

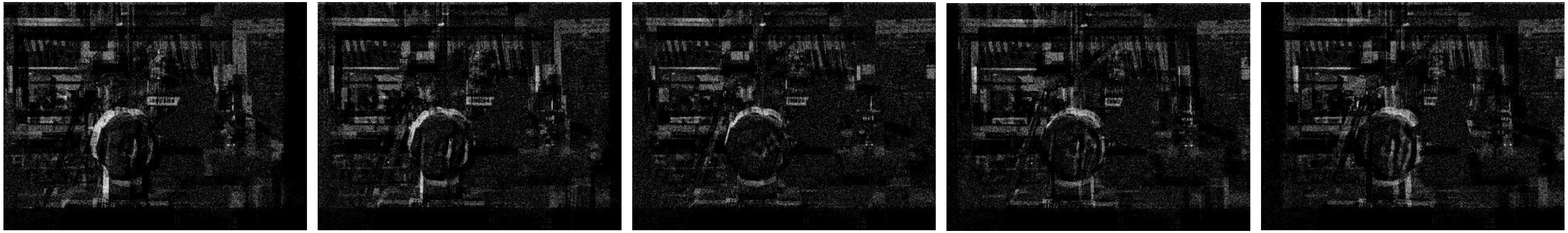
Depth Estimation



Depth Estimation

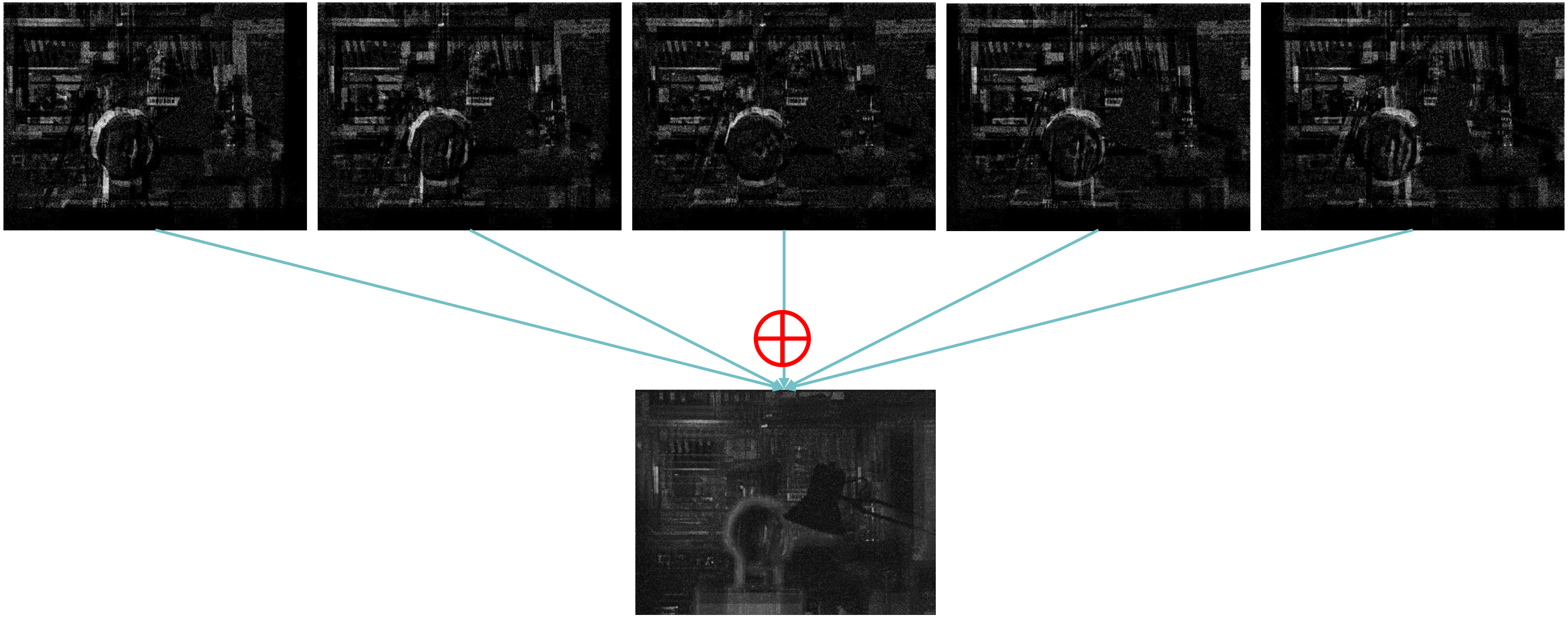


Depth Estimation

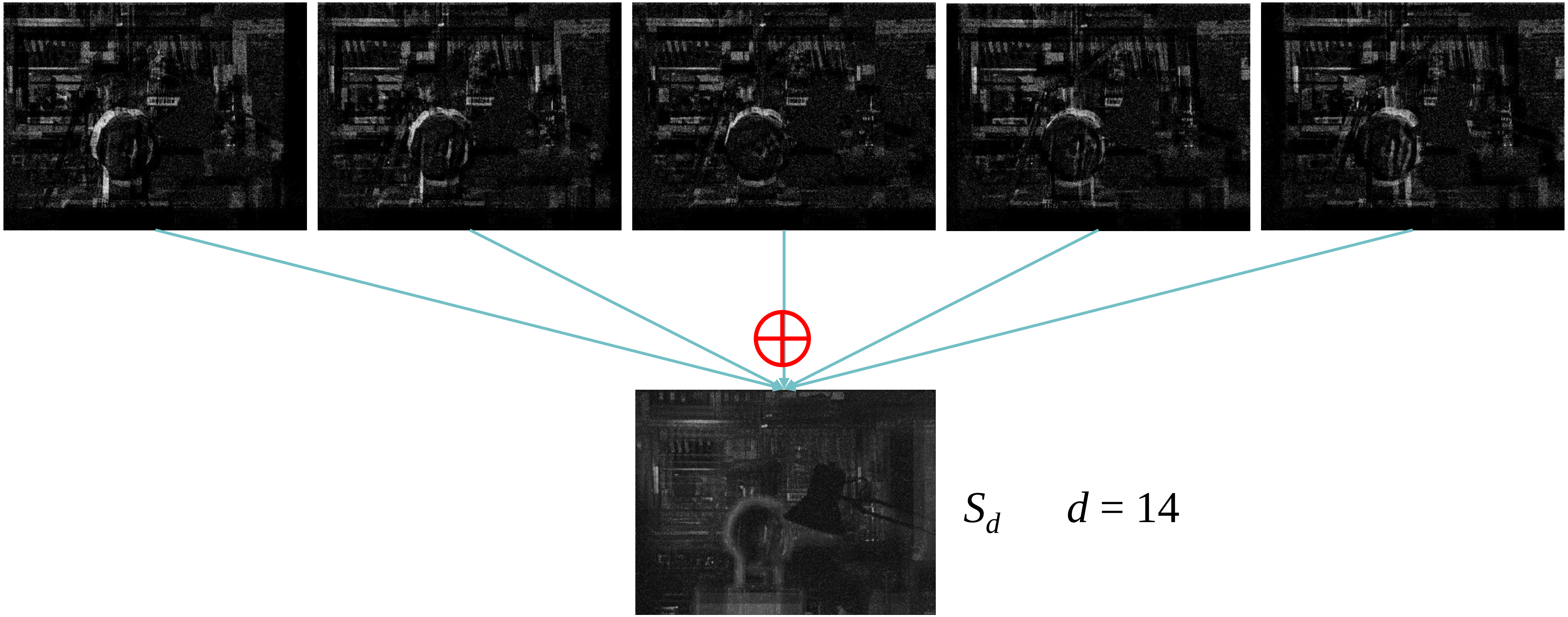


WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Depth Estimation



Depth Estimation

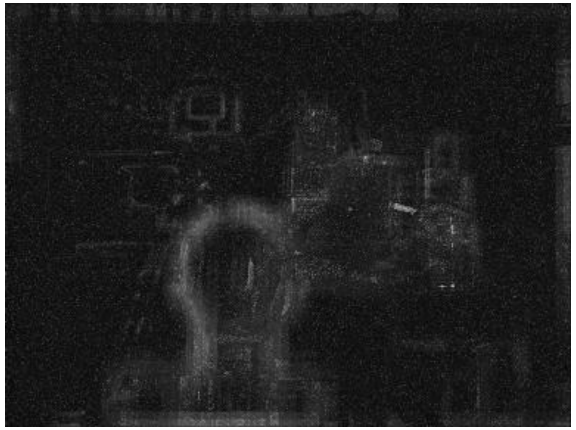


$S_d \qquad d = 14$

Depth Estimation

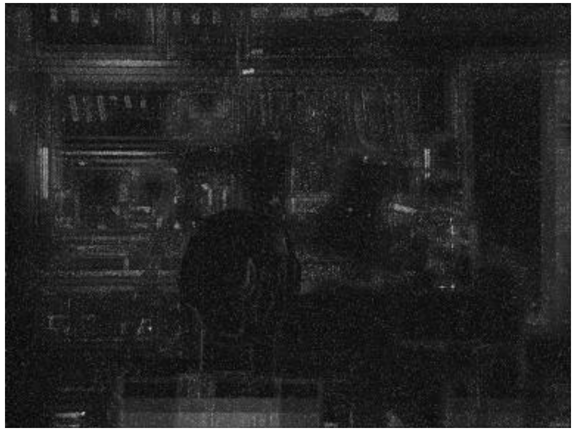
- Similarity measure

$$S_d(x,y) = \frac{1}{N}\sum_{k=1}^{N}\sum_{(i,j)\in W_{x,y}} \left|F_d(i,j,k) - I_{s_0,t_0}(i,j)\right|$$



$d = 5$

$d = 6$

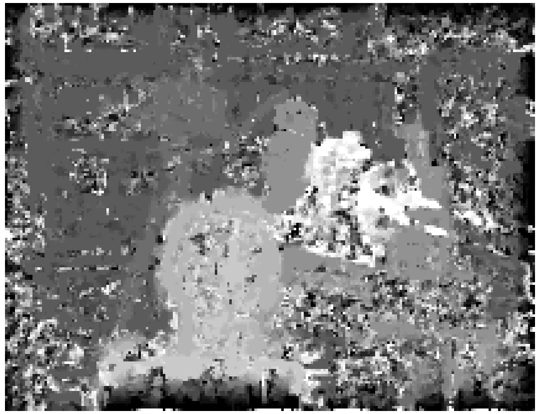


$d = 10$

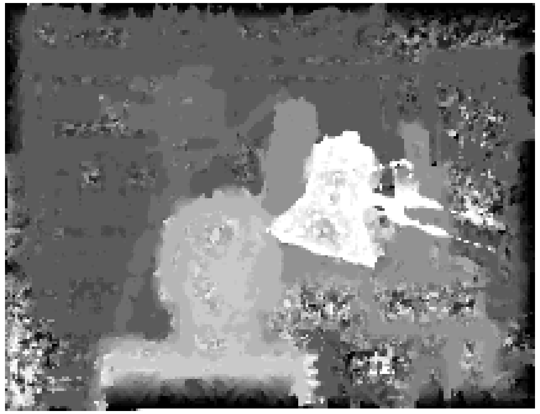
$d = 14$

Depth Estimation

- Disparity map $d\left(x,y\right) = \arg\min_{d} S_{d}\left(x,y\right)$



Disparity map in [1]



Disparity map using focus image stacks

Ground truth

Preliminary Denoising

$$I_{est}(x,y) = \frac{1}{w_{sum}} \sum_{k=1}^{N} w_k F_{d(x,y)}(x,y,k)$$

Preliminary Denoising

$$I_{est}(x,y) = \frac{1}{w_{sum}} \sum_{k=1}^{N} w_k F_{\boxed{d(x,y)}}(x,y,k)$$

- For each pixel (x, y):
  - Find its disparity d

Preliminary Denoising

$$I_{est}(x,y) = \frac{1}{w_{sum}} \sum_{k=1}^{N} w_k \boxed{F_{d(x,y)}(x,y,k)}$$

- For each pixel (x, y):
  - Find its disparity d
  - Extract corresponding focus image stack F_d
    - This pixel will be in-focus at this disparity

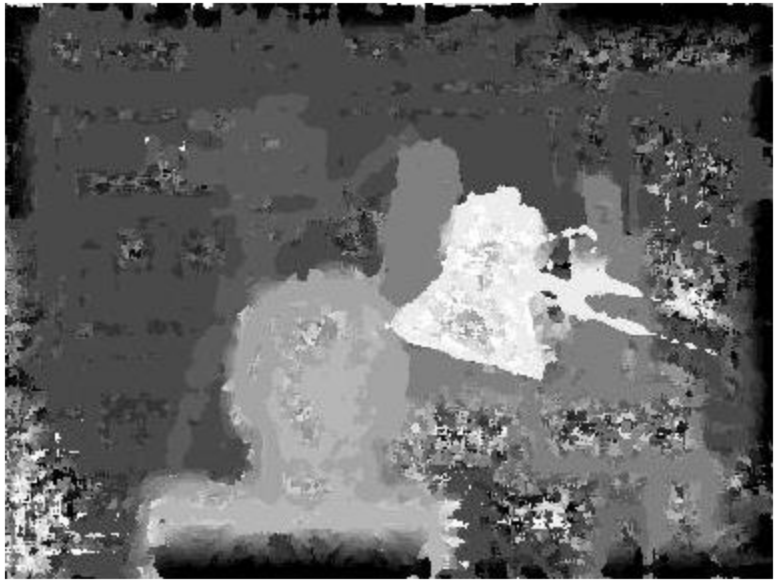
Preliminary Denoising

$$I_{est}(x,y) = \boxed{\frac{1}{w_{sum}} \sum_{k=1}^{N} w_k F_{d(x,y)}(x,y,k)}$$

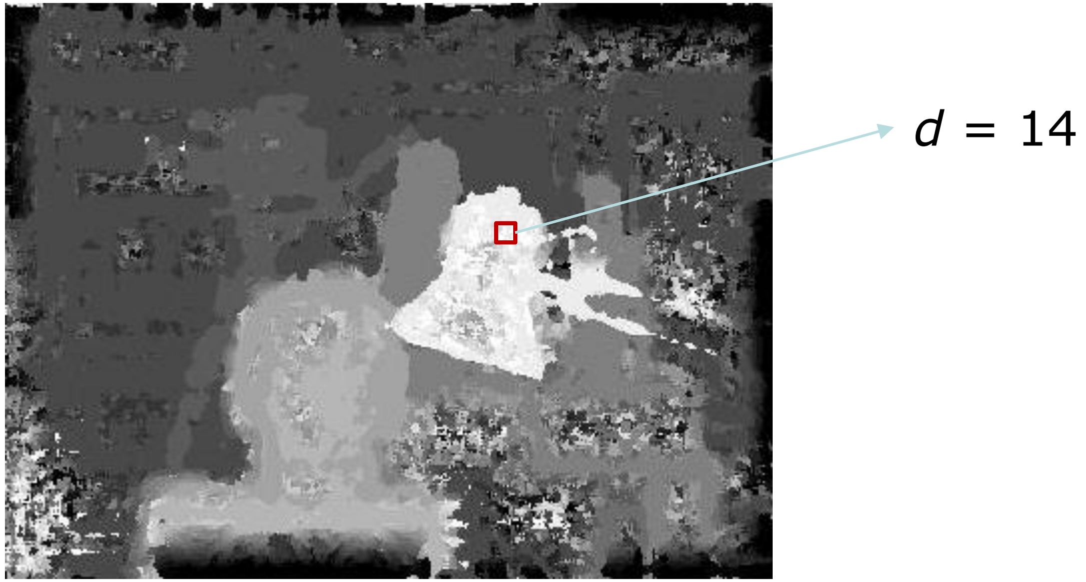
- For each pixel (x, y):
  - Find its disparity d
  - Extract corresponding focus image stack F_d
    - This pixel will be in-focus at this disparity
  - Take weighted average of each pixel at (x, y) in F_d
    - N – Number of views in F_d
    - k – k^{th} layer (view) of F_d
    - w_k – Weight of each layer (view) in F_d depending on distance to the center view

Preliminary Denoising

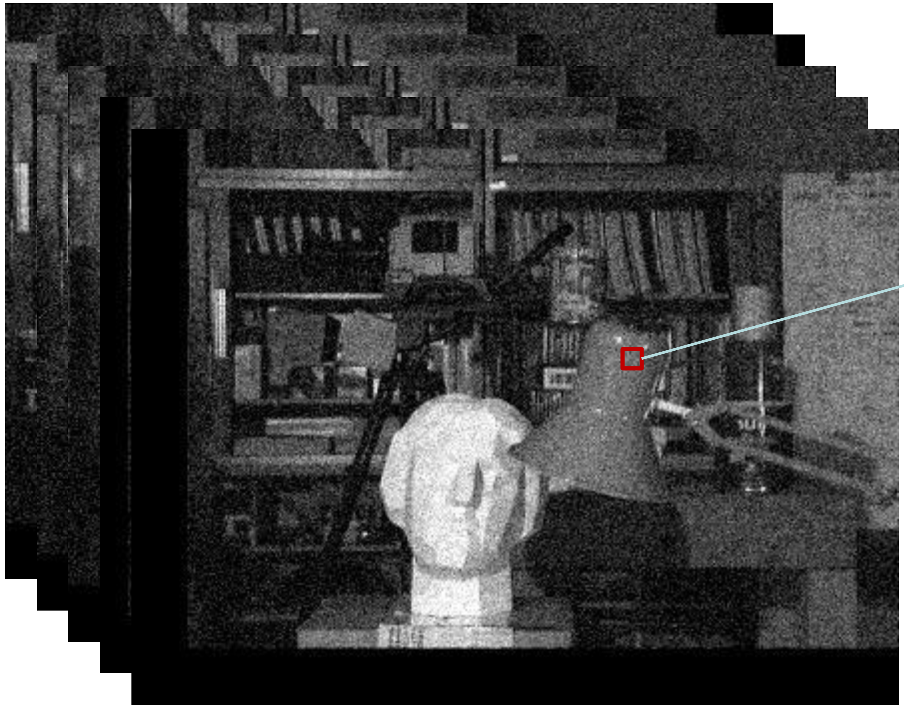
Preliminary Denoising



Preliminary Denoising

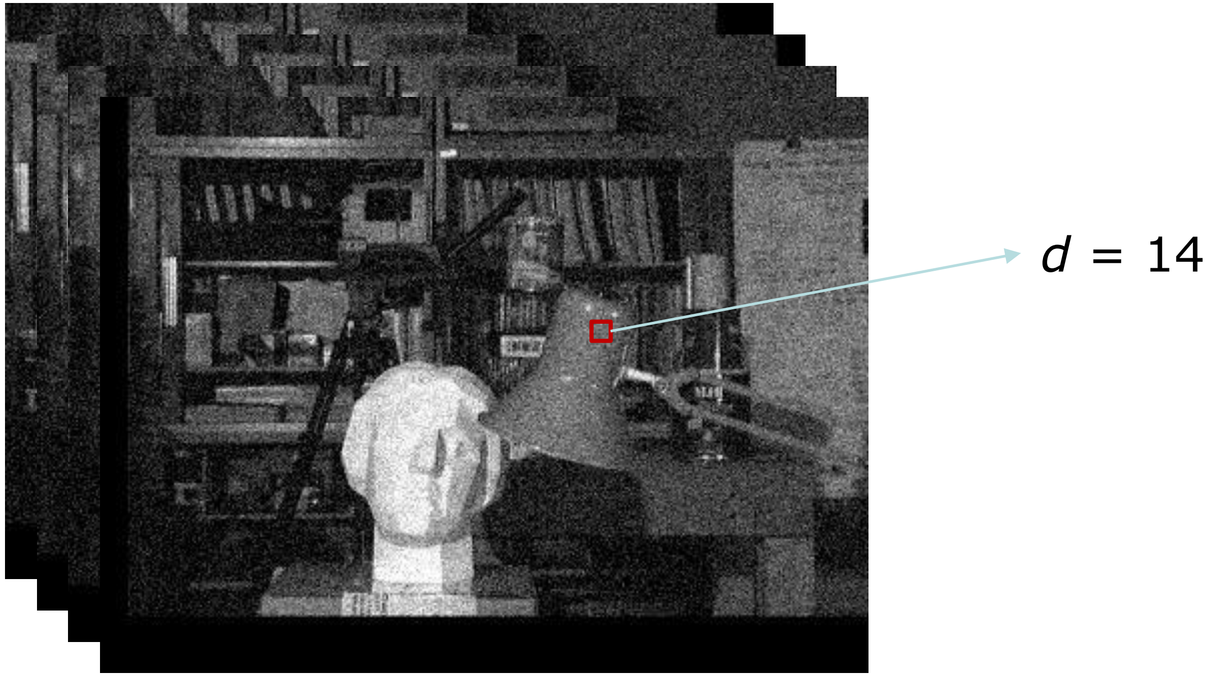


Preliminary Denoising



d = 14

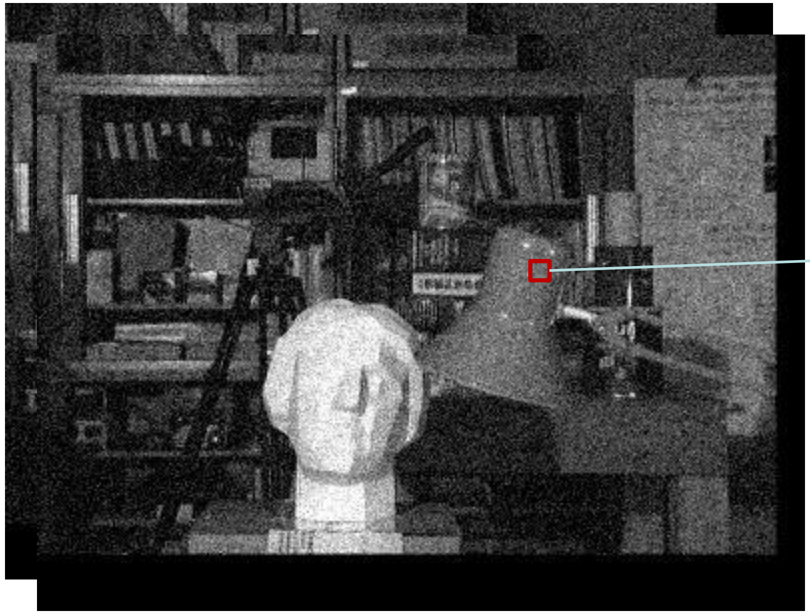
Preliminary Denoising



Preliminary Denoising

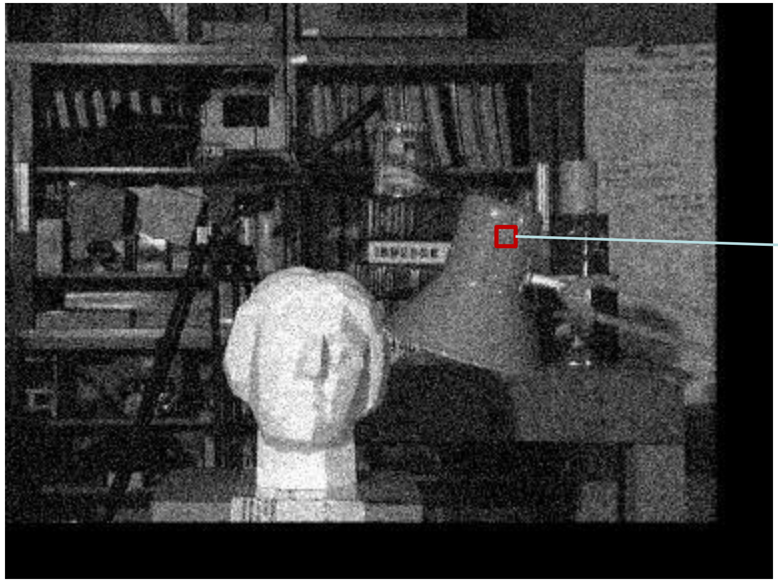
$d = 14$

Preliminary Denoising



$d = 14$

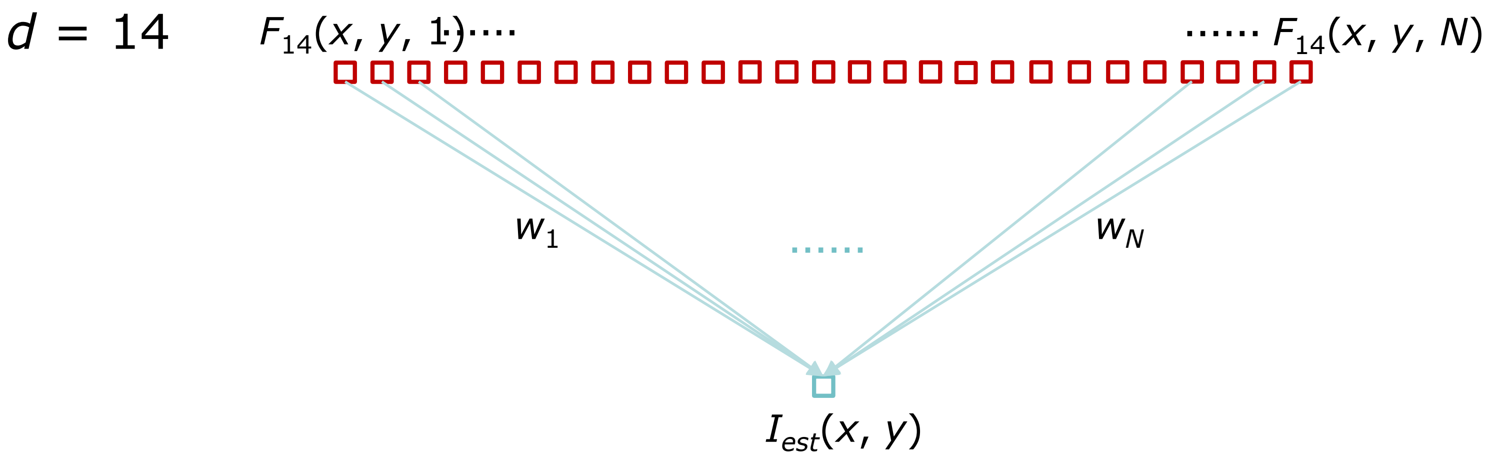
Preliminary Denoising



$d = 14$

Preliminary Denoising

Preliminary Denoising



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Preliminary Denoising

Noiseless true image

Noisy image

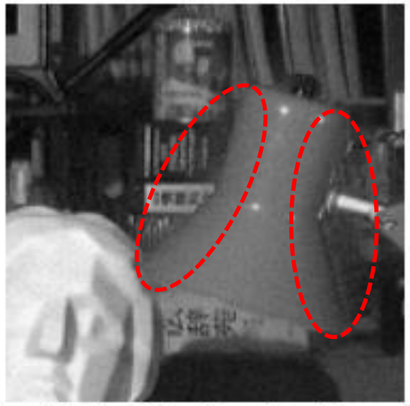
Preliminarily denoised image

Reliability map (binarized)

Preliminary Denoising

Noiseless true image

Noisy image



Preliminarily denoised image

Reliability map (binarized)

Reliability Evaluation

- Reliability map

$$R(x,y) = S_{d(x,y)}(x,y) / L^2$$

  - S_d - Similarity measure
  - L - Window size for computing S_d
  - Binarize $R(x, y)$ using a threshold
  - Morphological transformations
    - Dilation and erosion
    - Remove isolated outliers

Noiseless true image

Noisy image

Preliminarily denoised image

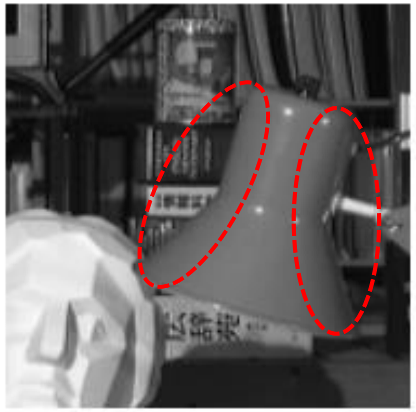
Reliability map (binarized)

Reliability Evaluation

- Reliability map

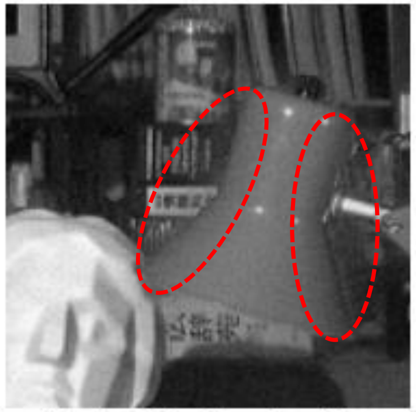
$$R(x,y) = S_{d(x,y)}(x,y) / L^2$$

  - S_d - Similarity measure
  - L - Window size for computing S_d
  - Binarize $R(x, y)$ using a threshold
  - Morphological transformations
    - Dilation and erosion
    - Remove isolated outliers

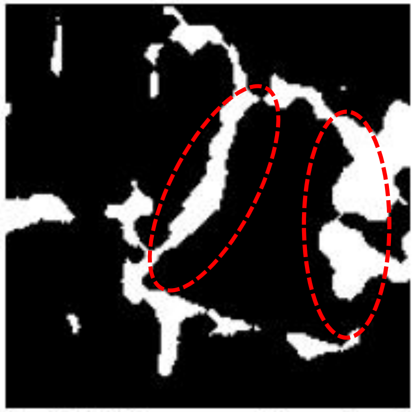


Noiseless true image

Noisy image



Preliminarily denoised image



Reliability map (binarized)

Handling Unreliable Pixels

- Non-local means

$$NL(i) = \sum_{j \in W_i} w(i, j) I(j)$$

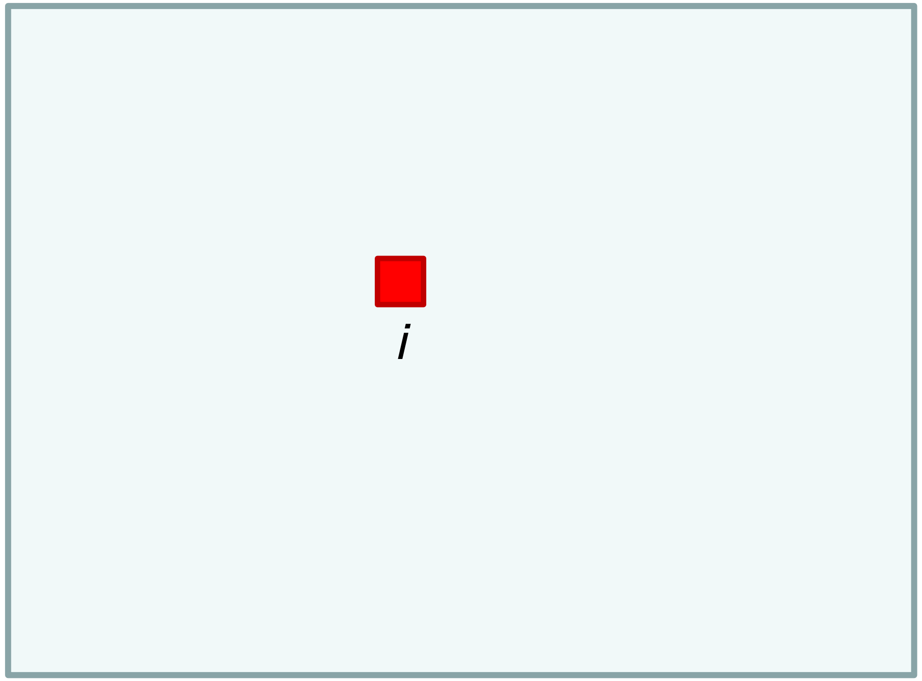
$$w(i, j) = \frac{1}{Z(i)} \exp\left(-\frac{\left\| I(N_i) - I(N_j) \right\|^2}{h^2} \right)$$

Handling Unreliable Pixels

- Non-local means

$$NL(i) = \sum_{j \in W_i} w(i, j) I(j)$$

$$w(i, j) = \frac{1}{Z(i)} \exp\left(- \frac{\left\| I(N_i) - I(N_j) \right\|^2}{h^2} \right)$$



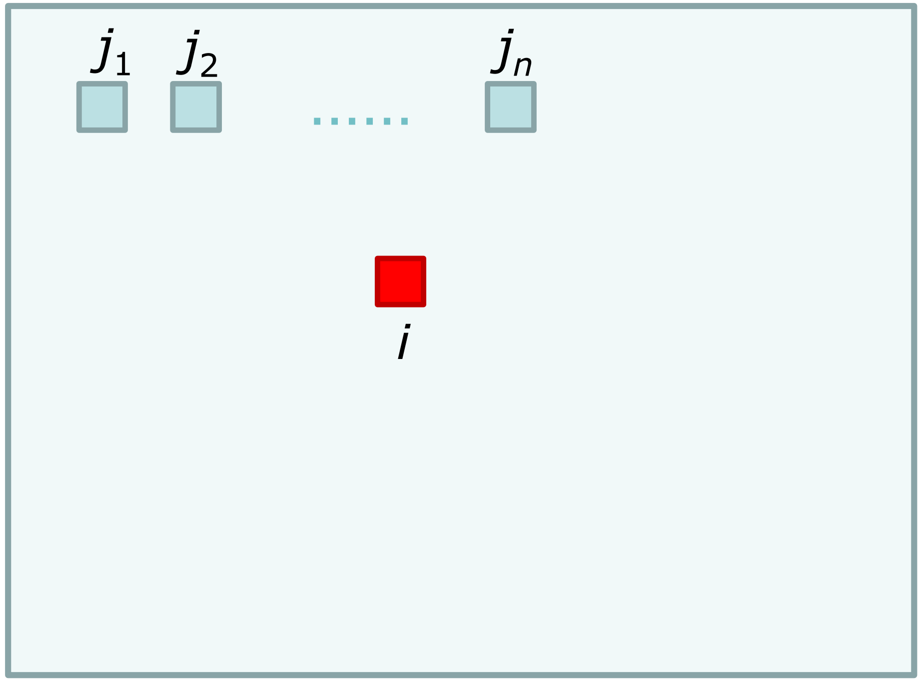
WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

Handling Unreliable Pixels

- Non-local means

$$NL(i) = \sum_{j \in W_i} w(i, j) I(j)$$

$$w(i, j) = \frac{1}{Z(i)} \exp\left(-\frac{\left\| I(N_i) - I(N_j) \right\|^2}{h^2} \right)$$

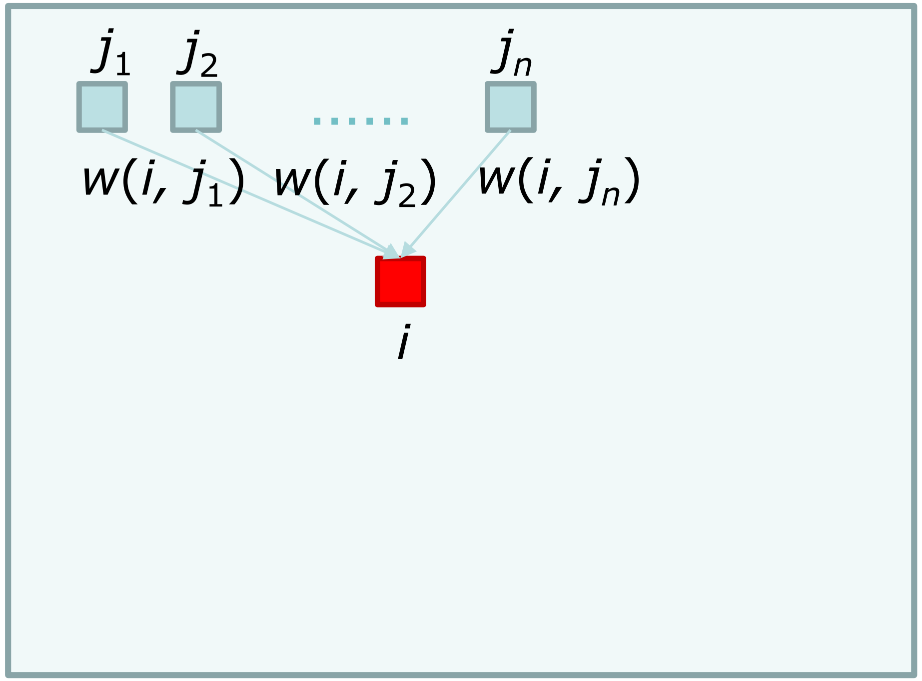


Handling Unreliable Pixels

- Non-local means

$$NL(i) = \sum_{j \in W_i} w(i, j) I(j)$$

$$w(i, j) = \frac{1}{Z(i)} \exp\left(-\frac{\left\| I(N_i) - I(N_j) \right\|^2}{h^2} \right)$$



Experiments

- PSNR (dB)

$$PSNR = 10\log_{10}\left(\frac{MAX^2}{MSE}\right)$$

  - *MSE* – mean squared error between denoised image and original image
  - *MAX* – maximum possible pixel value, i.e. 255

- Visual comparison

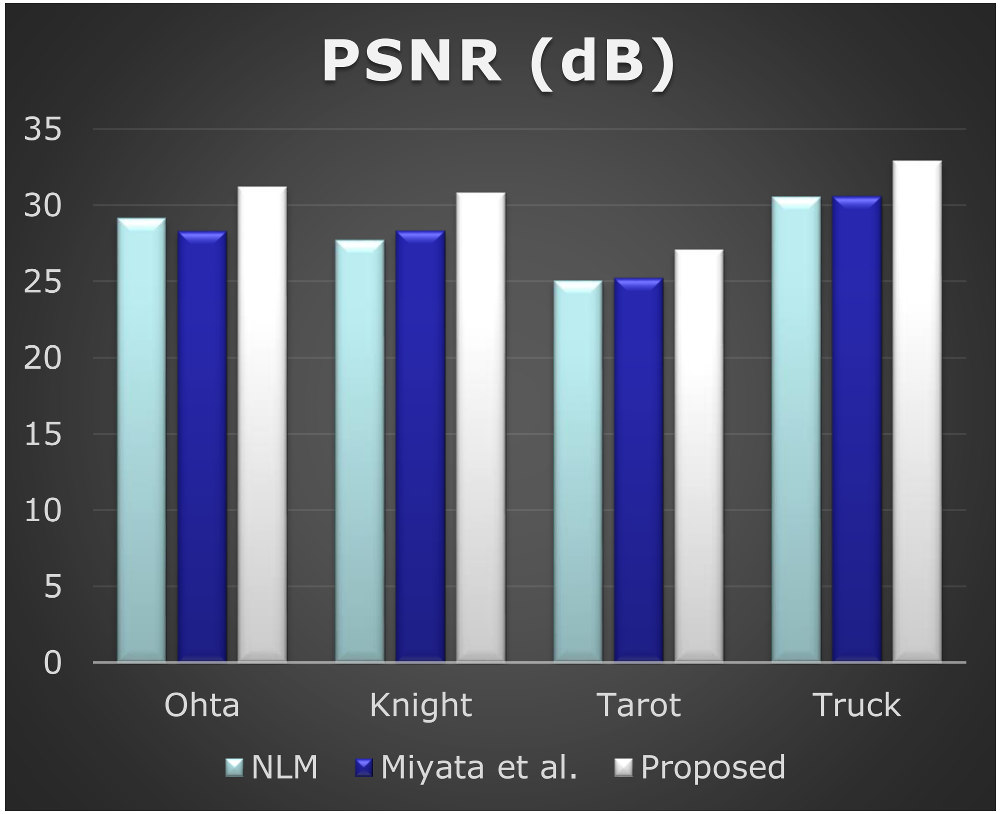
Ohta

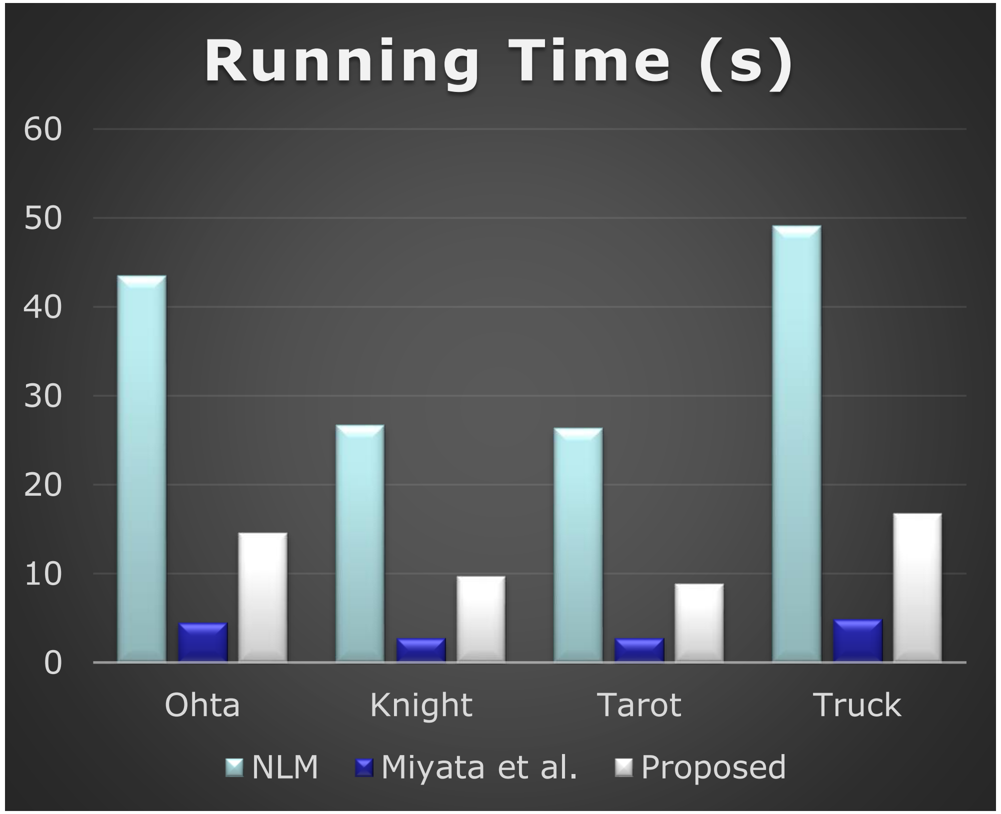
Knight

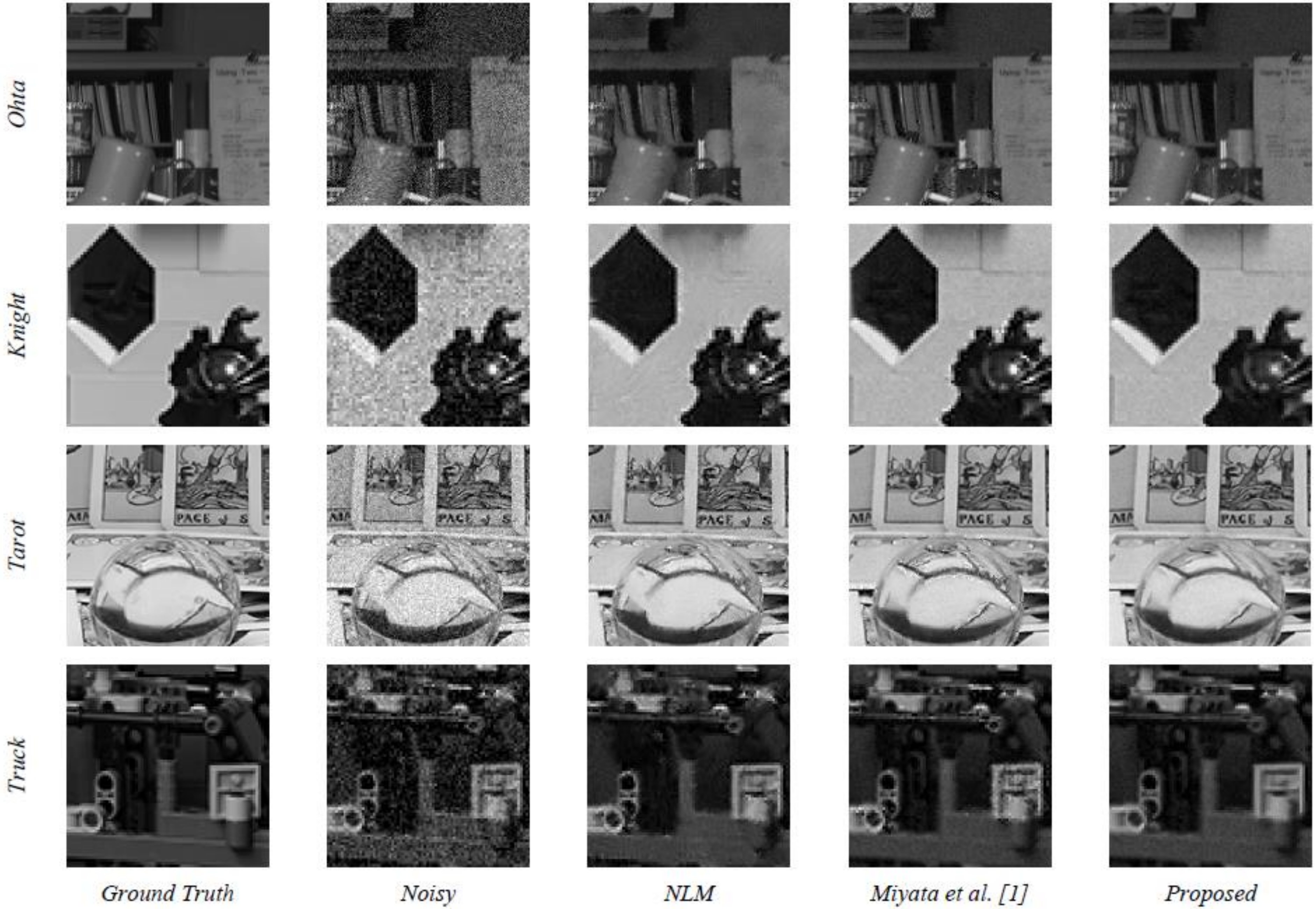
Tarot

Truck

Experiments






Ohta
Knight
Tarot
Truck
Ground Truth
Noisy
NLM
Miyata et al. [1]
Proposed


Ohta
Ground Truth
Noisy
NLM
Miyata et al. [1]
Proposed

Ohta
Ground Truth
Noisy
NLM
Miyata et al. [1]
Proposed

Ohta
Ground Truth
Noisy
NLM
Miyata et al. [1]
Proposed

Thank you!