# How to do Research

***(Research Plan: Brainstorming Idea and Implementation)***

Vinh Dinh Nguyen
PhD in Computer Science

# This Week

**Week 1**
How to read Research Paper

**Week 2**
How to find Research Topic and Develop Research Questions

**Week 3**
**Research Plan: Study and Implement Existing Algorithms**

**This Week**
Research Plan: *Brainstorming Idea and Implementation*

**Next Week**
Performance Evaluation and Plan to Write a Research Paper
*(If Any)*

# Outline

- Research Plan
- Challenges of the Research Topics
- Brainstorming idea
- Performance Evaluation
- Assignment

# Outline

- Research Plan
- Challenges of the Research Topics
- Brainstorming idea
- Performance Evaluation
- Assignment

# Research Plan

Develop Your Research Questions

↓

Background Research

↓

Study and Implement Existing Methods

↓

Develop a Research Plan

↓

Implement Your Proposed Methods

↓

Evaluate Performance of Compared Methods

1. Research Title:
   Provide a clear and concise title that reflects the main focus of your research
2. Introduction
   Briefly introduce your research topic, explaining its significance and relevance.
3. Background and Literature Review
   Summarize the existing knowledge and research related to your topic. Highlight key studies, theories, or gaps in the literature that your research aims to address.
4. Research Objectives
   What do you aim to achieve or discover through your study?
5. Research Questions
   Formulate the research questions or hypotheses that will guide your investigation
6. Methodology
   Detail the research design, methods, and procedures you plan to employ
7. Timeline
   Provide a timeline for your research activities. Include milestones, deadlines, and key events.
8. Resources
   Outline the resources you'll need for your research, such as access to specific databases, equipment, or funding.
9. Expected Result
   Discuss the anticipated outcomes of your research.
10. Significant and Contribution
    Explain the potential significance of your research.

# Research Plan

Develop Your Research Questions

Background Research

Study and Implement Existing Methods

Develop a Research Plan

Implement Your Proposed Methods

Evaluate Performance of Compared Methods

*1. Research Title:*

Robust Stereo Matching Various Weather Conditions

*2. Introduction*

Stereo matching is a fundamental computer vision task that involves the computation of dense disparity maps from stereo image pairs. The goal is to establish correspondences between pixels in the left and right images, providing depth information crucial for 3D perception. However, the accuracy and reliability of stereo matching algorithms are often challenged by various real-world factors such as occlusions, varying lighting conditions, and textureless regions.

In response to these challenges, the concept of robust stereo matching has emerged as a critical area of research within computer vision. Robust stereo matching aims to enhance the accuracy and efficiency of disparity estimation under adverse conditions, ensuring the reliability of depth information in a wide range of scenarios.

# Research Plan

Develop Your Research Questions

Background Research

Study and Implement Existing Methods

Develop a Research Plan

Implement Your Proposed Methods

Evaluate Performance of Compared Methods

*1. Research Title: (done)*

*2. Introduction*

...

Robust stereo matching techniques incorporate advanced algorithms and methodologies to address common issues encountered in stereo vision. These may include the development of adaptive matching cost functions, the integration of machine learning approaches for robust feature extraction, and the incorporation of strategies to handle occlusions and textureless regions effectively.

This field plays a pivotal role in applications such as autonomous vehicles, robotics, and augmented reality, where precise depth perception is essential for making informed decisions. As researchers delve into the complexities of real-world stereo vision challenges, the development of robust stereo matching algorithms becomes increasingly crucial to unlocking the full potential of stereo imaging systems in diverse and dynamic environments.

# Research Plan

Develop Your Research Questions

↓

Background Research

↓

Study and Implement Existing Methods

↓

Develop a Research Plan

↓

Implement Your Proposed Methods

↓

Evaluate Performance of Compared Methods

1. Research Title (done)
2. Introduction (done)
3. Background and Literature Review

### 3.a Introduction

Stereo matching, a vital component of computer vision, aims to extract depth information from stereo image pairs. This literature review explores the historical evolution, foundational concepts, and recent advancements in stereo matching algorithms.

### 3.b Historical of Stereo Matching

Early stereo matching methods, such as block matching and dynamic programming, laid the groundwork for depth perception. Despite their simplicity, these approaches struggled with challenges like occlusions and varying lighting conditions.

### 3.Foundational Concepts in Stereo Matching

Key concepts, including epipolar geometry and the stereo correspondence problem, form the basis of stereo matching. Understanding these principles is essential for developing and evaluating stereo matching algorithms.

# Research Plan

Develop Your Research Questions

↓

Background Research

↓

Study and Implement Existing Methods

↓

Develop a Research Plan

↓

Implement Your Proposed Methods

↓

Evaluate Performance of Compared Methods

1. Research Title (done)
2. Introduction (done)
3. Background and Literature Review

**3.d Traditional Stereo Matching Techniques**

Classical methods like block-based matching and graph cuts provided initial solutions to stereo correspondence. However, they faced limitations in handling textureless regions and were sensitive to noise.

**3.e Historical of Stereo Matching**

Early stereo matching methods, such as block matching and dynamic programming, laid the groundwork for depth perception. Despite their simplicity, these approaches struggled with challenges like occlusions and varying lighting conditions.

**3.f Challenges in Traditional Stereo Matching**

Discuss challenges faced by traditional stereo matching techniques, such as sensitivity to noise, computational inefficiency, and difficulties in handling real-world scenarios with dynamic lighting conditions.

# Research Plan

Develop Your Research Questions

↓

Background Research

↓

Study and Implement Existing Methods

↓

Develop a Research Plan

↓

Implement Your Proposed Methods

↓

Evaluate Performance of Compared Methods

1. Research Title (done)
2. Introduction (done)
3. Background and Literature Review

**3.g State-of-the-Art Stereo Matching Approaches**

Highlight recent advancements in stereo matching algorithms. State-of-the-art approaches leverage machine learning, advanced matching cost functions, and efficient optimization techniques to address challenges and achieve high accuracy.

# Research Plan

Develop Your Research Questions

↓

Background Research

↓

Study and Implement Existing Methods

↓

**Develop a Research Plan**

↓

Implement Your Proposed Methods

↓

Evaluate Performance of Compared Methods

1. Research Title (done)
2. Introduction (done)
3. Background and Literature Review (done)
4. Research Objectives

Investigate and address challenges such as occlusions, varying lighting conditions, and textureless regions. Research could focus on developing algorithms that are robust in real-world scenarios, ensuring reliable depth estimation even in adverse conditions.

Explore the integration of machine learning, including deep learning techniques, to enhance feature extraction and matching cost computation. This objective aims to leverage the power of neural networks to improve the overall performance of stereo matching algorithms.

Develop real-time stereo matching algorithms for applications that require fast and efficient processing, such as robotics or autonomous vehicles. This objective involves optimizing algorithms for speed without compromising accuracy.

# Research Plan

Develop Your Research Questions

Background Research

Study and Implement Existing Methods

Develop a Research Plan

Implement Your Proposed Methods

Evaluate Performance of Compared Methods

1. Research Title
2. Introduction
3. Background and Literature Review
4. Research Objectives
5. Research Questions

   RQ1: How to improve the performance of existing stereo matching methods under various lighting conditions and complex material properties?

   RQ2: How to obtain a real-time processing for current state-of-the-art stereo matching-based deep learning methods?

   RQ3: How to design an automatically learning the network architecture, its activation functions, and its parameters from data ?

# Research Plan

Develop Your Research Questions

Background Research

Study and Implement Existing Methods

Develop a Research Plan

Implement Your Proposed Methods

Evaluate Performance of Compared Methods

1. Research Title:
2. Introduction
3. Background and Literature Review
4. Research Objectives
5. Research Questions
6. Methodology

a. Data Preprocessing:

Noise Reduction: Apply preprocessing techniques to reduce noise in the stereo images. This may include Gaussian blurring, median filtering, or other denoising methods to improve the quality of input data.

Color Correction: Normalize lighting conditions and correct color disparities between the stereo images to ensure consistency in feature extraction.

b. Feature Extraction:

Use advanced feature extraction techniques to capture meaningful information from the stereo images.

Explore local features (e.g., SIFT, SURF) and global features (e.g., color histograms) that are robust to variations in lighting and texture.

c. Matching Cost Computation:

Employ adaptive matching cost functions that dynamically adjust to local image characteristics.

Experiment with different cost aggregation methods, such as cross-based or semi-global aggregation, to improve the accuracy of disparity maps in challenging regions.

# Research Plan

Develop Your Research Questions

Background Research

Study and Implement Existing Methods

Develop a Research Plan

Implement Your Proposed Methods

Evaluate Performance of Compared Methods

1. Research Title:
2. Introduction
3. Background and Literature Review
4. Research Objectives
5. Research Questions
6. Methodology

d. Machine Learning Integration:

Explore the integration of machine learning techniques, such as Convolutional Neural Networks (CNNs), to enhance feature extraction and matching cost computation.
Train models on diverse datasets to improve the network's ability to handle variations in lighting, occlusions, and textureless regions.

e. Occlusion Handling:

Develop strategies to handle occlusions effectively. Investigate methods for detecting and managing occluded regions to prevent them from affecting disparity estimation.
Use confidence measures or segmentation techniques to identify unreliable disparities in occluded areas.

f. Adaptive Window Sizes:

Implement adaptive window sizes based on local image characteristics. This helps in focusing computational efforts where detailed disparities are needed and using larger windows in textureless or uniform areas.

# Research Plan

Develop Your Research Questions

↓

Background Research

↓

Study and Implement Existing Methods

↓

Develop a Research Plan

↓

Implement Your Proposed Methods

↓

Evaluate Performance of Compared Methods

1. Research Title:
2. Introduction
3. Background and Literature Review
4. Research Objectives
5. Research Questions
6. Methodology
7. Timeline
   - Study and implement existing methods: 1 week
   - Implement the proposed method: 1 week
   - Evaluate and analysis the performance of compared methods: 1 week

# How to do Research

Develop Your Research Questions

⬇

Background Research

⬇

Study and Implement Existing Methods

⬇

Develop a Research Plan

⬇

Implement Your Proposed Methods

⬇

Evaluate Performance of Compared Methods

1. Research Title:
2. Introduction
3. Background and Literature Review
4. Research Objectives
5. Research Questions
6. Methodology
7. Timeline
8. Resources
   - Dataset
   - Hardware
   - Software
   - Human resource

# Research Plan

Develop Your Research Questions

Background Research

Study and Implement Existing Methods

Develop a Research Plan

Implement Your Proposed Methods

Evaluate Performance of Compared Methods

1. Research Title:
2. Introduction
3. Background and Literature Review
4. Research Objectives
5. Research Questions
6. Methodology
7. Timeline
8. Resources
9. Expected Result

   Improve the performance of existing methods under difficult conditions by 1%.

   Reduce the processing time of the-state-of-the-art method to 30ms/frame

# Research Plan

Develop Your Research Questions

⬇

Background Research

⬇

Study and Implement Existing Methods

⬇

Develop a Research Plan

⬇

Implement Your Proposed Methods

⬇

Evaluate Performance of Compared Methods

1. Research Title:
2. Introduction
3. Background and Literature Review
4. Research Objectives
5. Research Questions
6. Methodology
7. Timeline
8. Resources
9. Expected Result
10. Significant and Contribution

Significance: The proposed stereo matching methodology addresses the limitations of existing techniques by specifically targeting challenging conditions, including occlusions, varying lighting conditions, and textureless regions.

Contribution: The methodology introduces novel strategies, such as adaptive matching cost functions and machine learning integration, to enhance the accuracy and robustness of disparity estimation under difficult scenarios.

# Research Questions: Stereo Matching

RQ1: How to improve the performance of existing stereo matching methods under various lighting conditions and complex material properties?

RQ2: How to obtain a real-time processing for current state-of-the-art stereo matching-based deep learning methods?

RQ3: How to design an automatically learning the network architecture, its activation functions, and its parameters from data ?

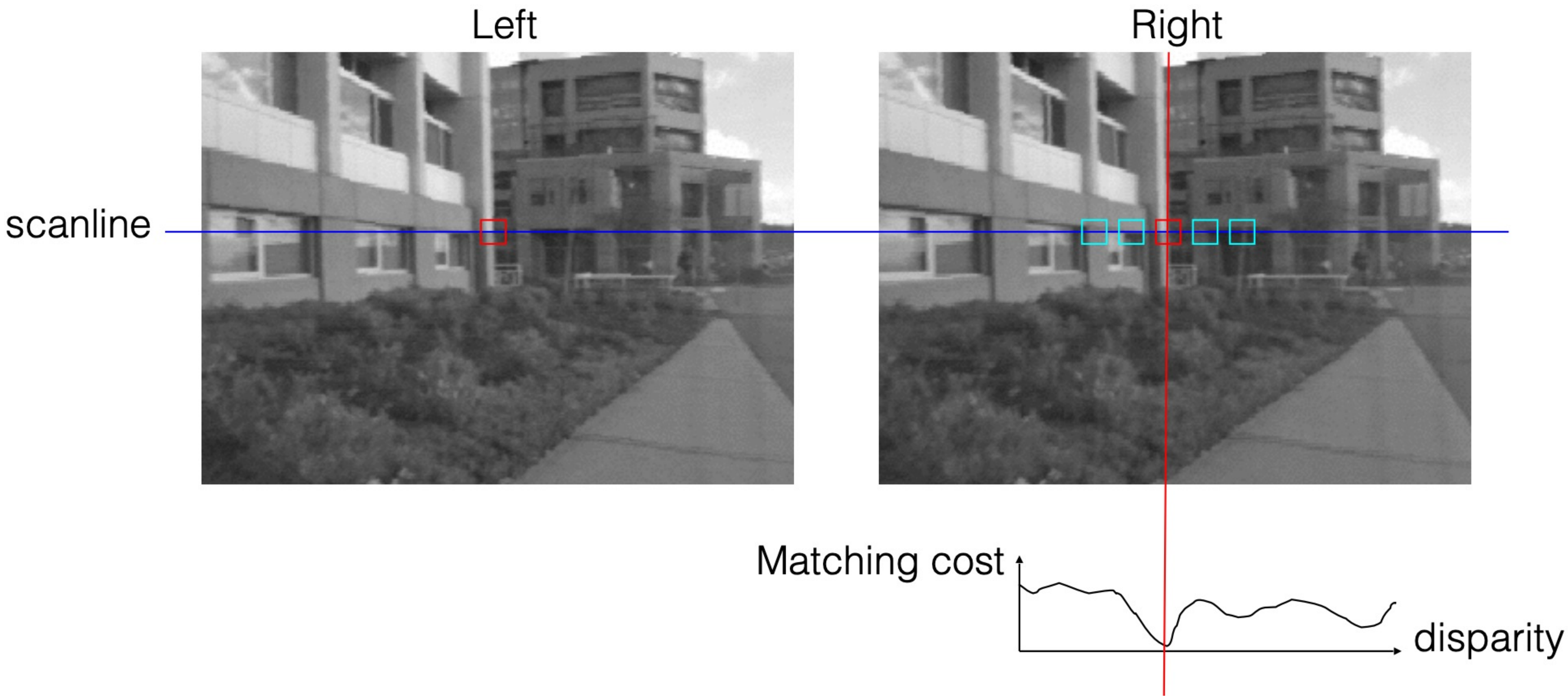
Please select the simplest question to solve


In this module

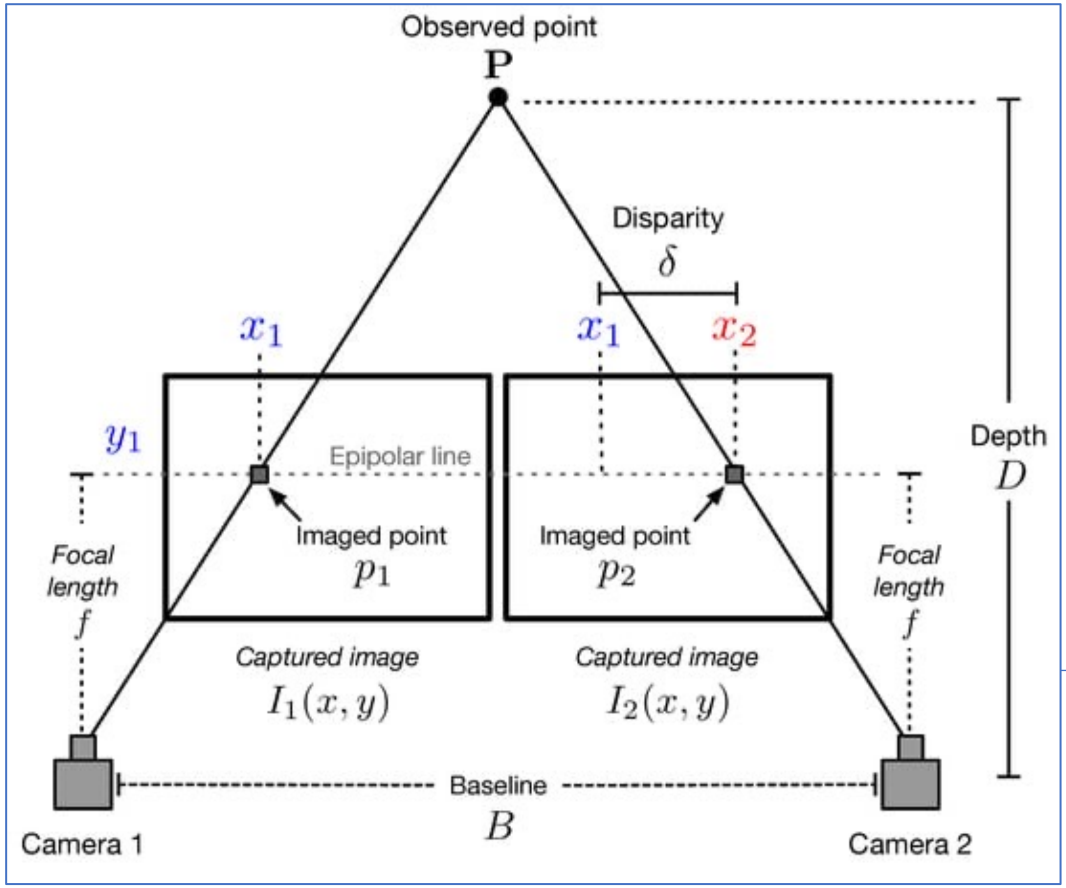
# Outline

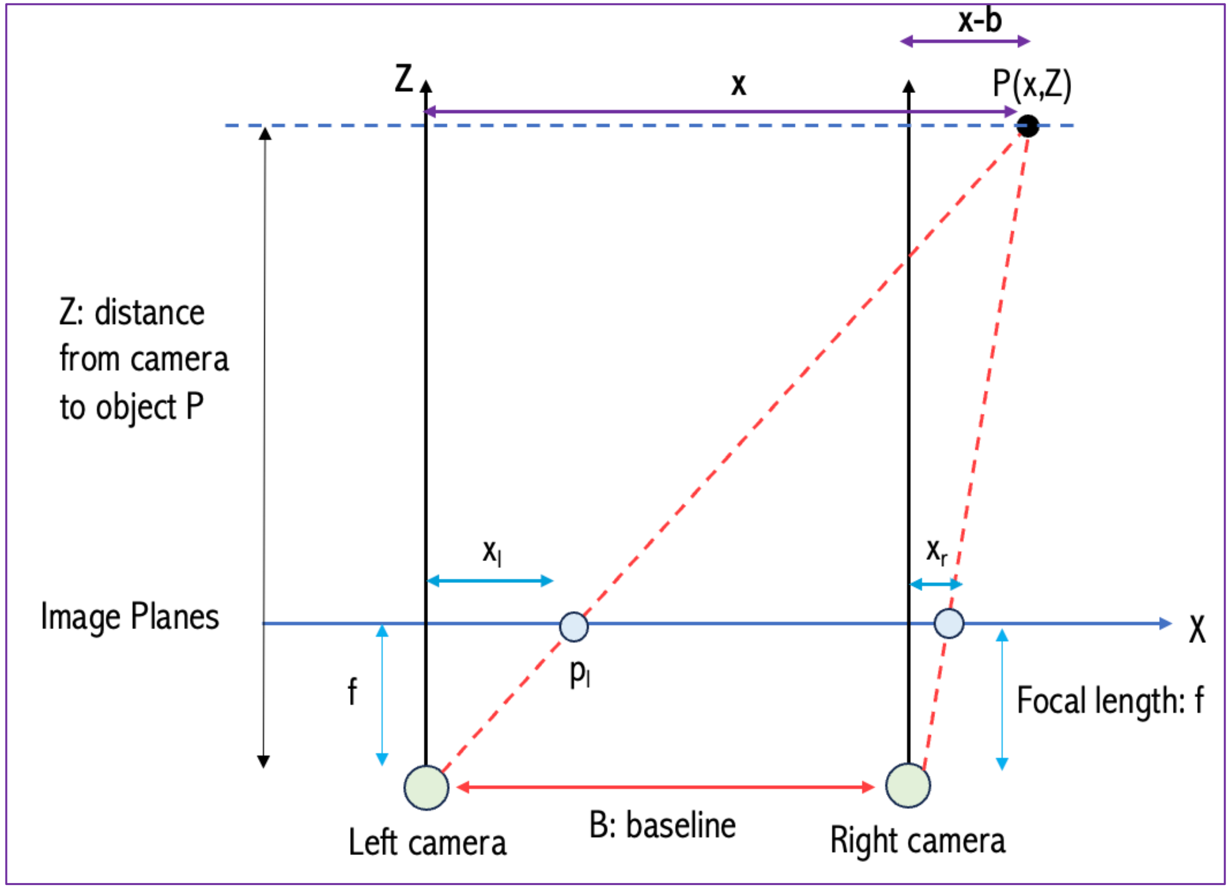
- Research Plan
- **Challenges of the Research Topics**
- Brainstorming idea
- Performance Evaluation
- Assignment

# Stereo Matching: Background



# Stereo Matching: Background



# Stereo Matching: Background

❖ **Method 1**

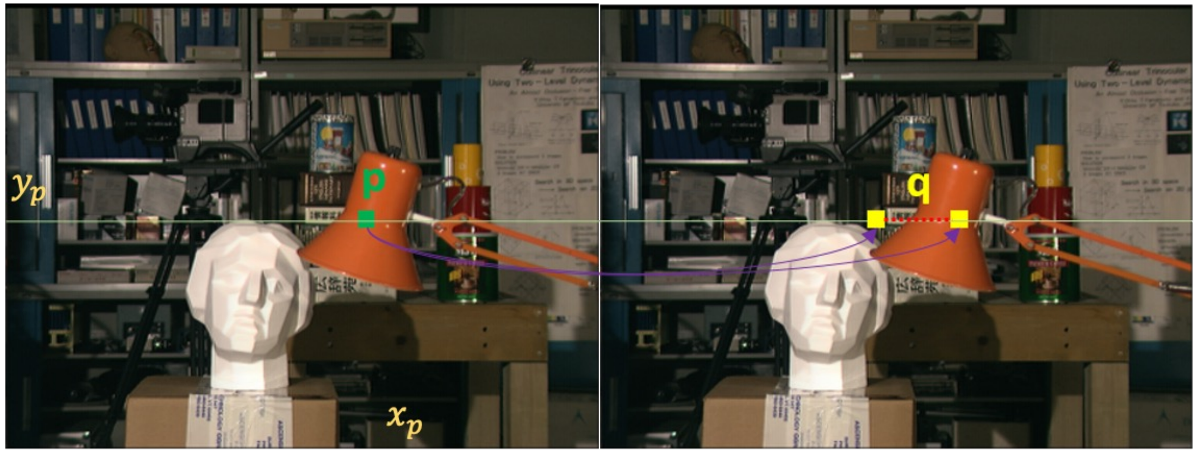
$L$ is the left image

$R$ is the right image

$L(\boldsymbol{p})$ is the (vector) value of **p**

$$\boldsymbol{p} = \begin{bmatrix} x_p \\ y_p \end{bmatrix} = \begin{bmatrix} 234 \\ 140 \end{bmatrix}$$

$$D = 16$$




Left Image Right Image

$\boldsymbol{q}$

$$\begin{bmatrix} x_p - 0 \\ y_p \end{bmatrix} = \begin{bmatrix} 234 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 8 \\ y_p \end{bmatrix} = \begin{bmatrix} 226 \\ 140 \end{bmatrix}$$

$$\begin{bmatrix} x_p - 1 \\ y_p \end{bmatrix} = \begin{bmatrix} 233 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 9 \\ y_p \end{bmatrix} = \begin{bmatrix} 225 \\ 140 \end{bmatrix}$$

$$\begin{bmatrix} x_p - 2 \\ y_p \end{bmatrix} = \begin{bmatrix} 232 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 10 \\ y_p \end{bmatrix} = \begin{bmatrix} 224 \\ 140 \end{bmatrix}$$

$$\begin{bmatrix} x_p - 3 \\ y_p \end{bmatrix} = \begin{bmatrix} 231 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 11 \\ y_p \end{bmatrix} = \begin{bmatrix} 223 \\ 140 \end{bmatrix}$$

$$\begin{bmatrix} x_p - 4 \\ y_p \end{bmatrix} = \begin{bmatrix} 230 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 12 \\ y_p \end{bmatrix} = \begin{bmatrix} 222 \\ 140 \end{bmatrix}$$

$$\begin{bmatrix} x_p - 5 \\ y_p \end{bmatrix} = \begin{bmatrix} 229 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 13 \\ y_p \end{bmatrix} = \begin{bmatrix} 221 \\ 140 \end{bmatrix}$$

$$\begin{bmatrix} x_p - 6 \\ y_p \end{bmatrix} = \begin{bmatrix} 228 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 14 \\ y_p \end{bmatrix} = \begin{bmatrix} 220 \\ 140 \end{bmatrix}$$

$$\begin{bmatrix} x_p - 7 \\ y_p \end{bmatrix} = \begin{bmatrix} 227 \\ 140 \end{bmatrix} \quad \begin{bmatrix} x_p - 15 \\ y_p \end{bmatrix} = \begin{bmatrix} 219 \\ 140 \end{bmatrix}$$

# Stereo Matching: Background

$$\boldsymbol{d} = \begin{bmatrix} d \\ 0 \end{bmatrix} \qquad d \in D$$

$$C_1(\boldsymbol{p}, \boldsymbol{d}) = |L(\boldsymbol{p}) - R(\boldsymbol{p} - \boldsymbol{d})|$$

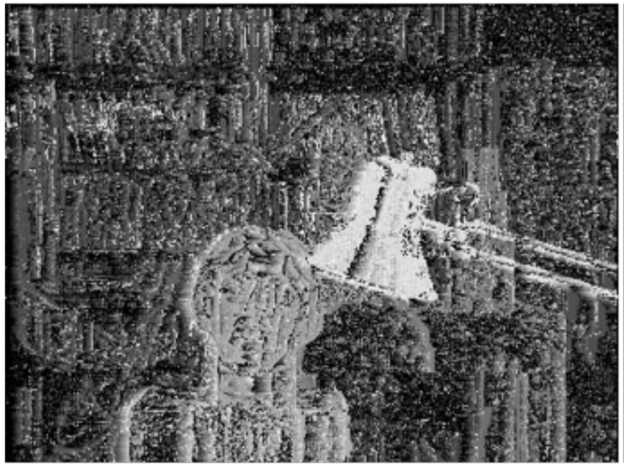
Finding d so that $C_1(\boldsymbol{p}, \boldsymbol{d})$ is minimum.

$$d = \underset{d \in D}{\operatorname{argmin}}\left(C_1(\boldsymbol{p}, \boldsymbol{d})\right)$$

Then, d is the value for the pixel **p** in disparity map

Left Image

Right Image



Disparity Map

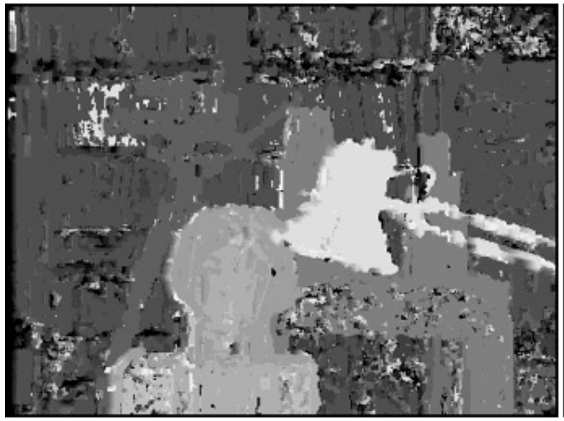
Ground Truth

# Stereo Matching: Background

## ❖Normal


Left Image

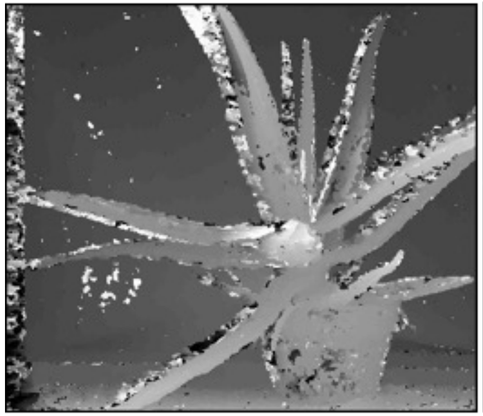

Right Image

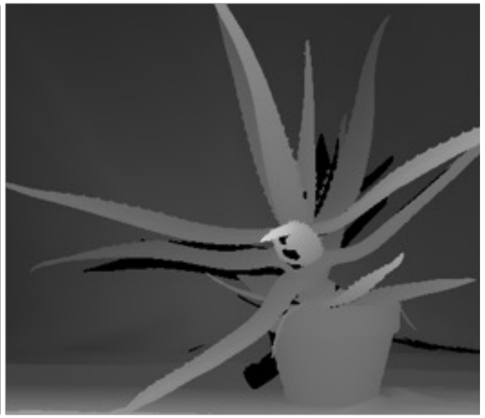

Disparity Map


Ground Truth


Left Image


Right Image

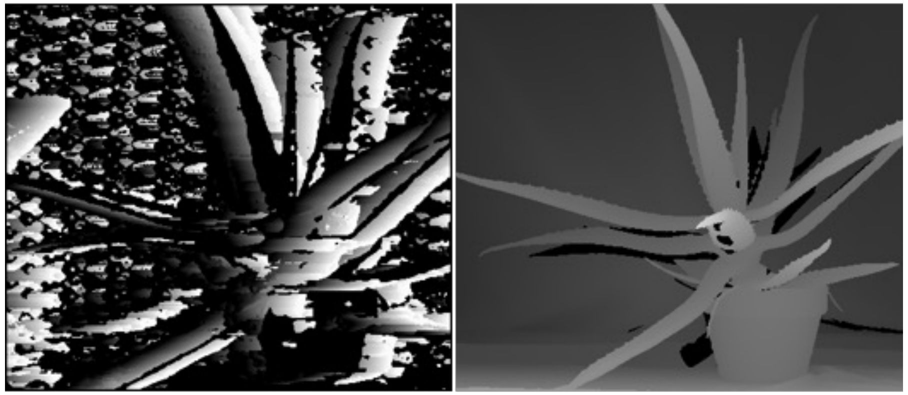

Disparity Map


Ground Truth

# Stereo Matching: Background

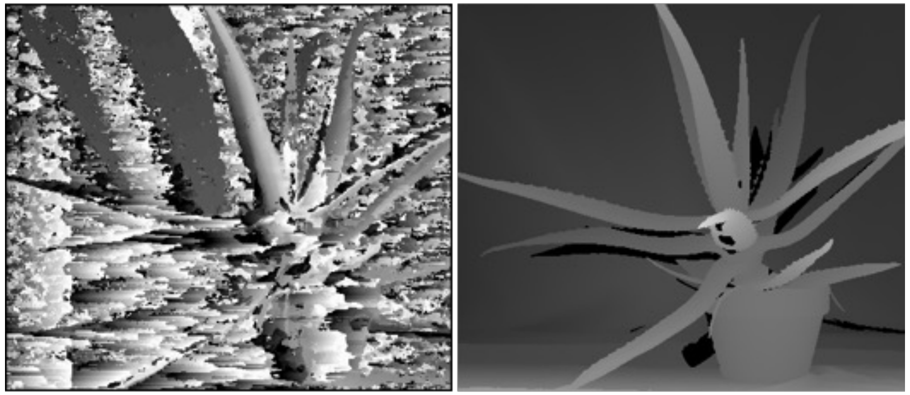
## ❖Challenges


Left Image | Right Image



Disparity Map | Ground Truth

Left Image | Right Image

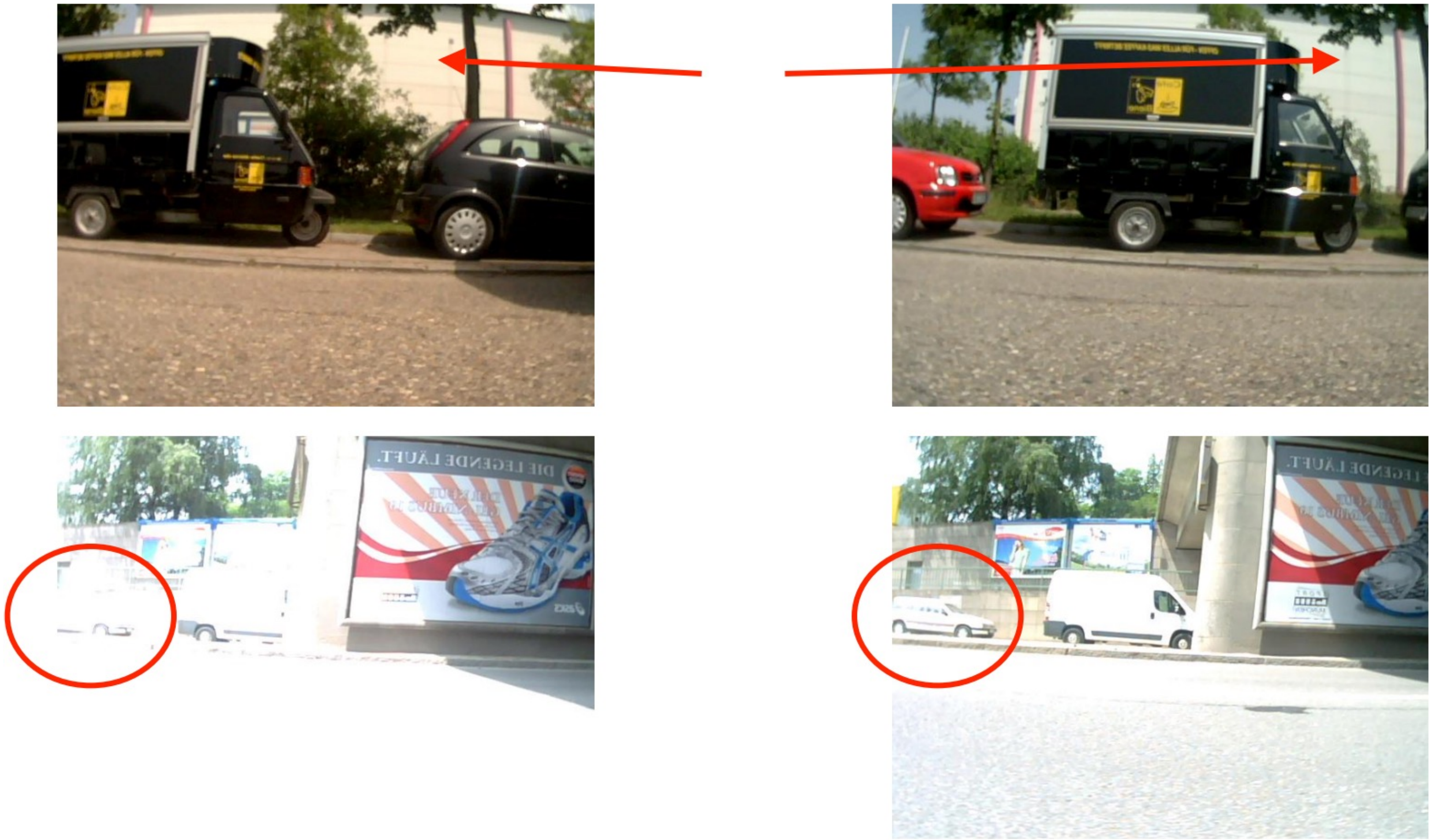


Disparity Map | Ground Truth

# Stereo Matching: Background

## ❖Challenges


Photometric
Variations.



# Stereo Matching: Background

## ❖Challenges


Image Sensor Noise

# Stereo Matching: Background

## ❖Challenges


Specularities

# Stereo Matching: Background

## ❖Challenges


Foreshortening and the Uniqueness Constraint

# Stereo Matching: Background

## ❖Challenges


Perspective Distortions.

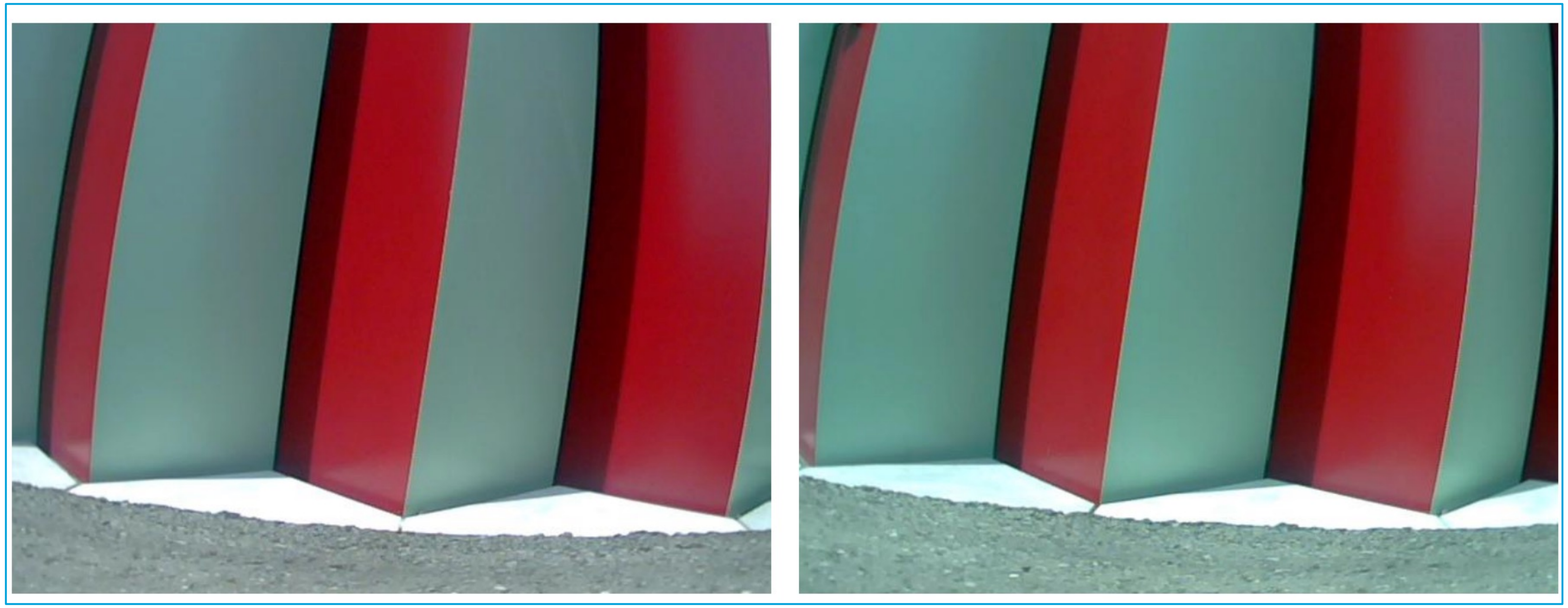
# Stereo Matching: Background

## ❖Challenges


Textureless Regions.

# Stereo Matching: Background

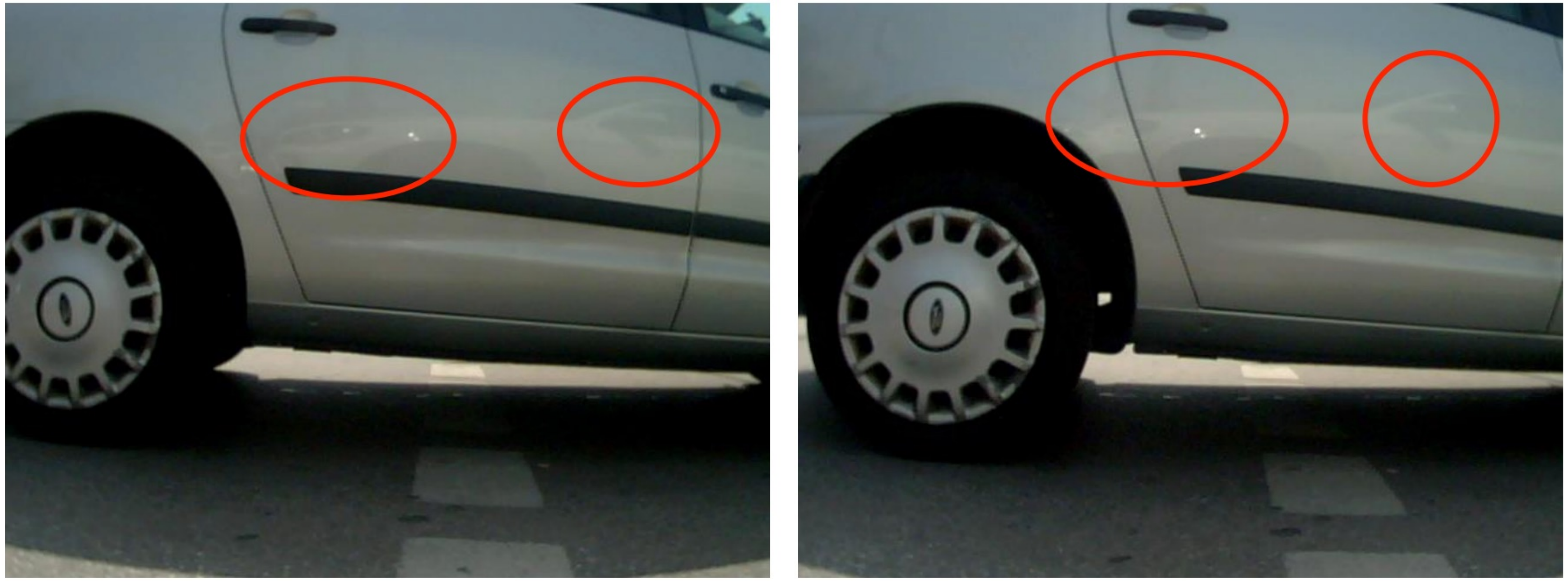
## ❖Challenges

Repetitive Structures and Textures

# Stereo Matching: Background

## ❖Challenges

Reflections

# Stereo Matching: Background

## ❖Challenges


Transparency

# Stereo Matching: Background

## ❖Challenges

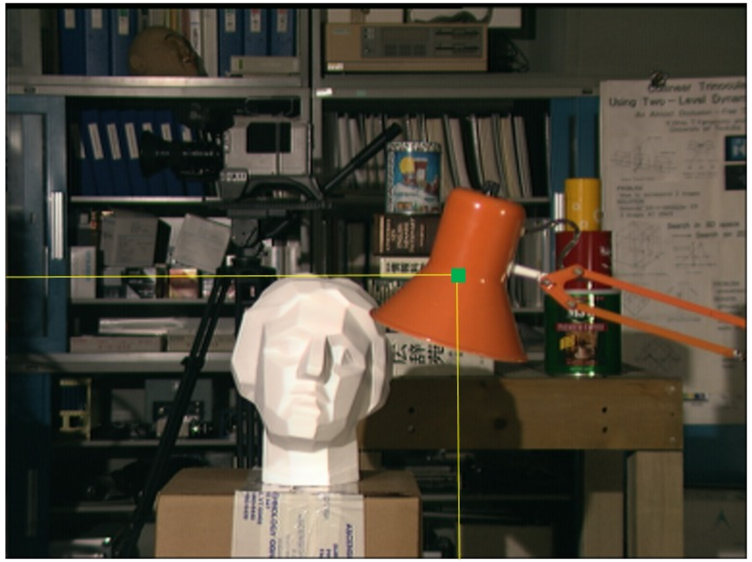

Occlusions

# Challenges in Stereo Matching

RQ1: How to improve the performance of existing stereo matching methods under various lighting conditions and complex material properties?

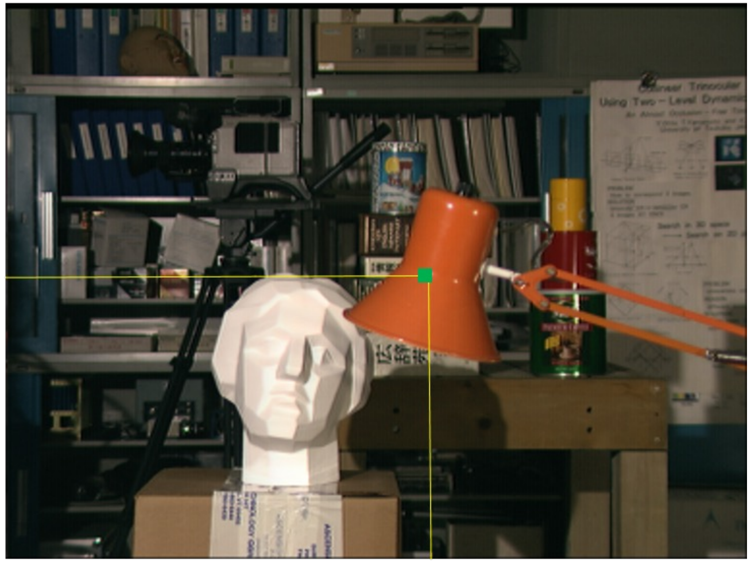
**p** = (234, 140) **q** = (220, 140)

y_p=140



x_p = 234

y_q=140



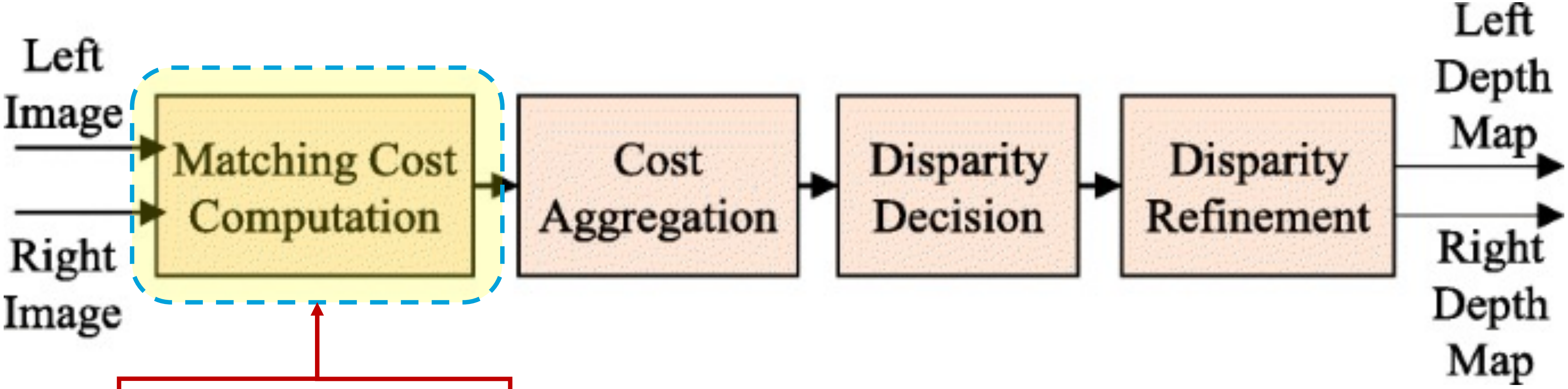
x_q = 220

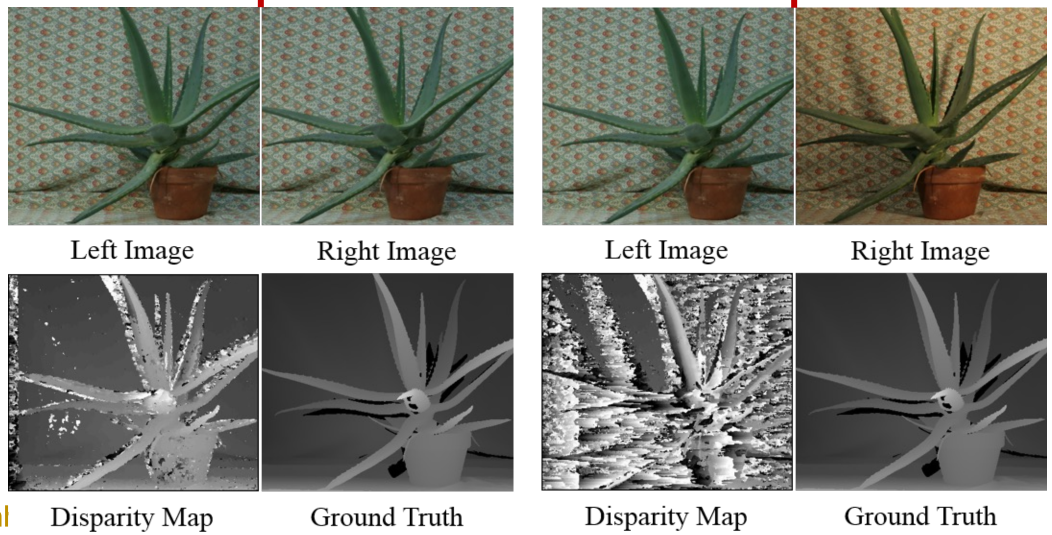
disparity_q = x_p – x_q = 14

# Outline

- Research Plan
- Challenges of the Research Topics
- Brainstorming idea
- Performance Evaluation
- Assignment

# Stereo Matching Workflow




Indentify where you can improve the performance of existing algorithms?

# Stereo Matching Cost

## ❖ AD Matching Cost

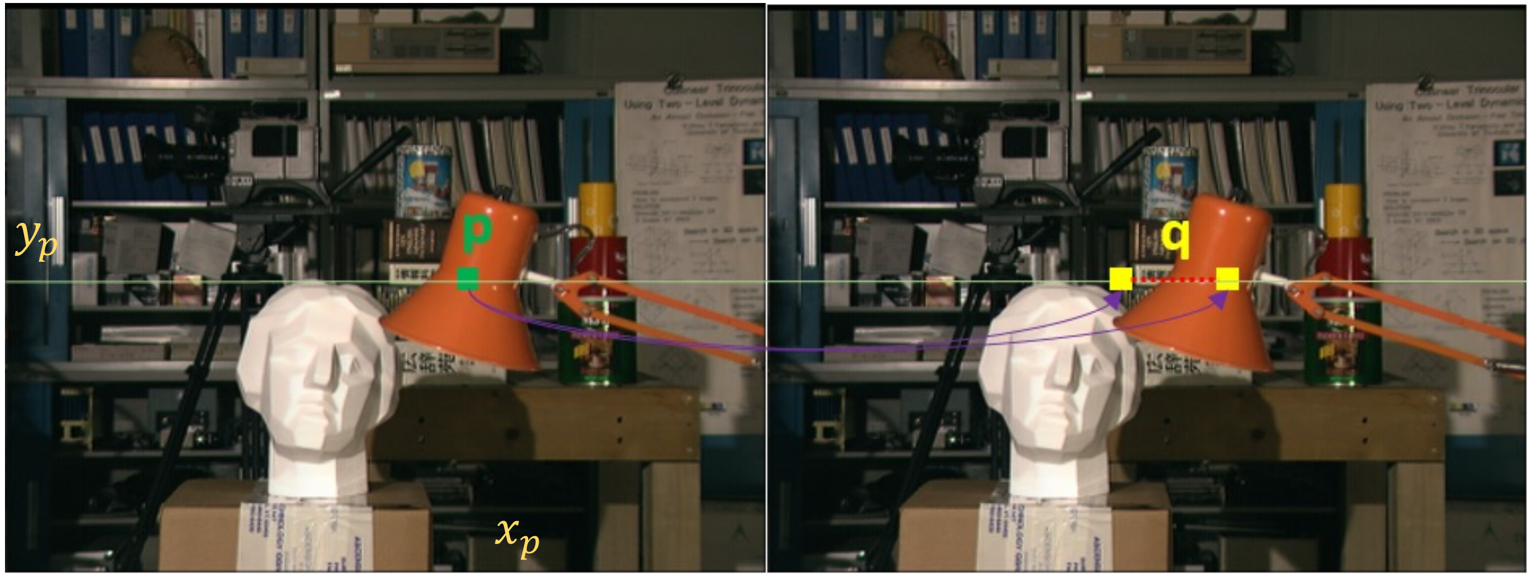
$L$ is the left image

$R$ is the right image

$L(\boldsymbol{p})$ is the (vector) value of **p**

$$\boldsymbol{p} = \begin{bmatrix} x_p \\ y_p \end{bmatrix} \qquad d \in D$$

$$\boldsymbol{d} = \begin{bmatrix} d \\ 0 \end{bmatrix} \qquad C_1(\boldsymbol{p}, \boldsymbol{d}) = |L(\boldsymbol{p}) - R(\boldsymbol{p} - \boldsymbol{d})|$$




Left Image

Right Image

Finding d so that $C_1(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{d})$ is minimum.

$$d = \underset{d \in D}{\operatorname{argmin}}\left(C_1(\boldsymbol{p}, \boldsymbol{d})\right)$$

Then, d is the value for the pixel **p** in disparity map

# Stereo Matching

## ❖SAD Matching Cost

$L$ is the left image

$R$ is the right image

$L(\boldsymbol{p})$ is the (vector) value of **p**

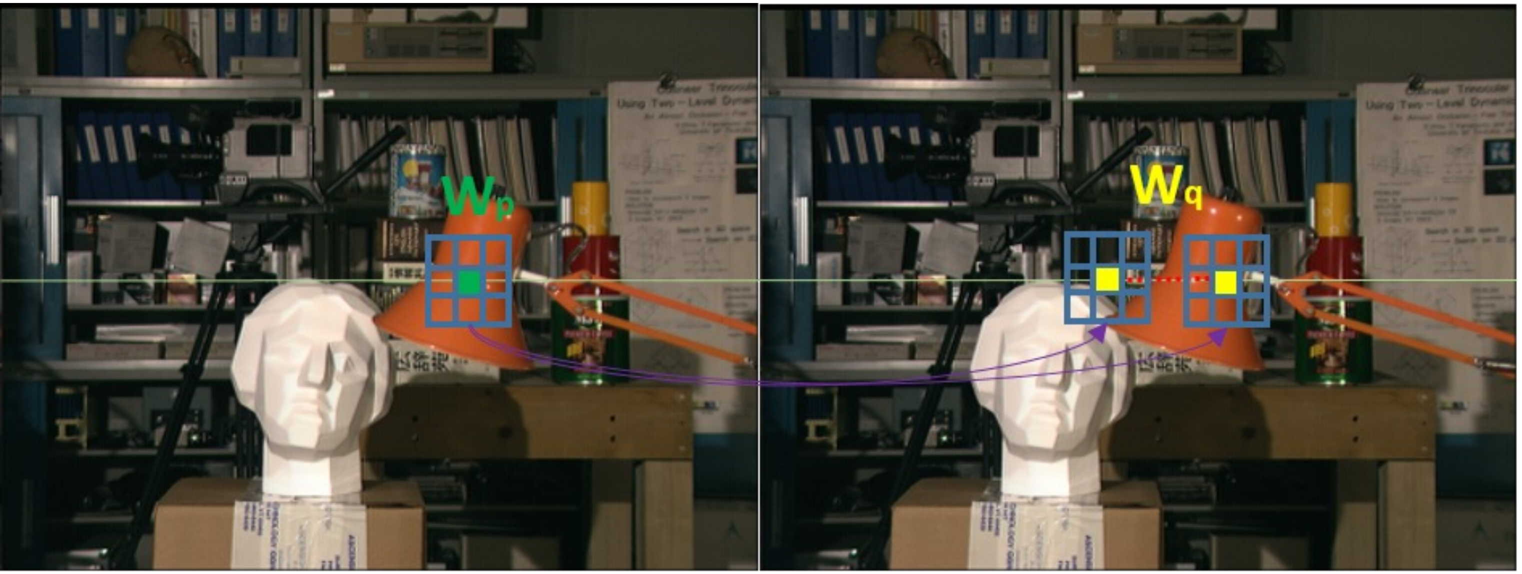
$$\boldsymbol{p} = \begin{bmatrix} x_p \\ y_p \end{bmatrix}$$

$W_{\boldsymbol{p}}$ is a window centered at p

$$\boldsymbol{d} = \begin{bmatrix} d \\ 0 \end{bmatrix}$$

$$C_2(\boldsymbol{p}, \boldsymbol{d}) = \sum_{\boldsymbol{u} \in W_{\boldsymbol{p}}} |L(\boldsymbol{u}) - R(\boldsymbol{u} - \boldsymbol{d})|$$

$d \in D$




Left Image

Right Image

Finding d so that $C_2(\boldsymbol{p}, \boldsymbol{d})$ is minimum.

$$d = \underset{d \in D}{\operatorname{argmin}}\big(C_2(\boldsymbol{p}, \boldsymbol{d})\big)$$

Then, d is the value for the pixel **p** in disparity map

# Stereo Matching Cost

## Parametric Matching Costs

$$C_{AD}(\mathbf{p},\mathbf{d}) = |I_L(\mathbf{p}) - I_R(\mathbf{p}-\mathbf{d})|, \tag{1}$$

$$C_{SAD}(\mathbf{p},\mathbf{d}) = \sum_{\mathbf{q}\in N_{\mathbf{p}}} |I_L(\mathbf{q}) - I_R(\mathbf{q}-\mathbf{d})|. \tag{2}$$

Sampling-insensitive absolute difference of Birchfield and Tomasi

$$\begin{aligned}
C_{BT}(\mathbf{p},\mathbf{d}) &= \min(A,B),\\
A &= \max(0, I_L(\mathbf{p}) - I_R^{max}(\mathbf{p}-\mathbf{d}), I_R^{min}(\mathbf{p}-\mathbf{d}) - I_L(\mathbf{p})),\\
B &= \max(0, I_R(\mathbf{p}-\mathbf{d}) - I_L^{max}(\mathbf{p}), I_L^{min}(\mathbf{p}) - I_R(\mathbf{p}-\mathbf{d})),\\
I^{min}(\mathbf{p}) &= \min(I^-(\mathbf{p}), I(\mathbf{p}), I^+(\mathbf{p}))\\
I^{max}(\mathbf{p}) &= \max(I^-(\mathbf{p}), I(\mathbf{p}), I^+(\mathbf{p}))\\
I^-(\mathbf{p}) &= (I(\mathbf{p} - [1\ \ 0]^T) + I(\mathbf{p}))/2,\\
I^+(\mathbf{p}) &= (I(\mathbf{p} + [1\ \ 0]^T) + I(\mathbf{p}))/2.
\end{aligned}$$

# Stereo Matching Cost

## Parametric Matching Costs

The zeromean sum of absolute differences (ZSAD) s

$$C_{ZSAD}(\mathbf{p},\mathbf{d}) = \sum_{\mathbf{q}\in N_{\mathbf{p}}} |I_L(\mathbf{q}) - \bar{I}_L(\mathbf{p}) - I_R(\mathbf{q}-\mathbf{d}) + \bar{I}_R(\mathbf{p}-\mathbf{d})|$$

$$\bar{I}(\mathbf{p}) = \frac{1}{|N_{\mathbf{p}}|}\sum_{\mathbf{q}\in N_{\mathbf{p}}} I(\mathbf{q}).$$

Normalized cross-correlation

$$C_{NCC}(\mathbf{p},\mathbf{d}) = \frac{\sum_{\mathbf{q}\in N_{\mathbf{p}}} I_L(\mathbf{q}) I_R(\mathbf{q}-\mathbf{d})}{\sqrt{\sum_{\mathbf{q}\in N_{\mathbf{p}}} I_L(\mathbf{q})^2 \sum_{\mathbf{q}\in N_{\mathbf{p}}} I_R(\mathbf{q}-\mathbf{d})^2}}.$$

NCC is another window-based matching technique that is commonly used. NCC compensates for gain changes and is statistically the optimal method for dealing with Gaussian noise. However, NCC tends to blur depth discontinuities more than many other matching costs because outliers lead to high errors within the NCC calculation

# Stereo Matching Cost

## Parametric Matching Costs

Zero-mean variant ZNCC

$$C_{ZNCC}(\mathbf{p},\mathbf{d}) = \frac{\sum_{\mathbf{q}\in N_{\mathbf{p}}}(I_L(\mathbf{q}) - \bar{I}_L(\mathbf{p}))(I_R(\mathbf{q}-\mathbf{d}) - \bar{I}_R(\mathbf{p}-\mathbf{d}))}{\sqrt{\sum_{\mathbf{q}\in N_{\mathbf{p}}}(I_L(\mathbf{q}) - \bar{I}_L(\mathbf{p}))^2 \sum_{\mathbf{q}\in N_{\mathbf{p}}}(I_R(\mathbf{q}-\mathbf{d}) - \bar{I}_R(\mathbf{p}-\mathbf{d}))^2}}.$$

## Non-parametric Matching Costs

Rank Filter

$$I_{Rank}(\mathbf{p}) = \sum_{\mathbf{q}\in N_{\mathbf{p}}} T[I(\mathbf{q}) < I(\mathbf{p})].$$

The function T[] is defined to return 1 if its argument is true and 0 otherwise. The transformed images are matched with the absolute difference. The Rank filter is known to be susceptible to noise in textureless areas.

# Stereo Matching Cost

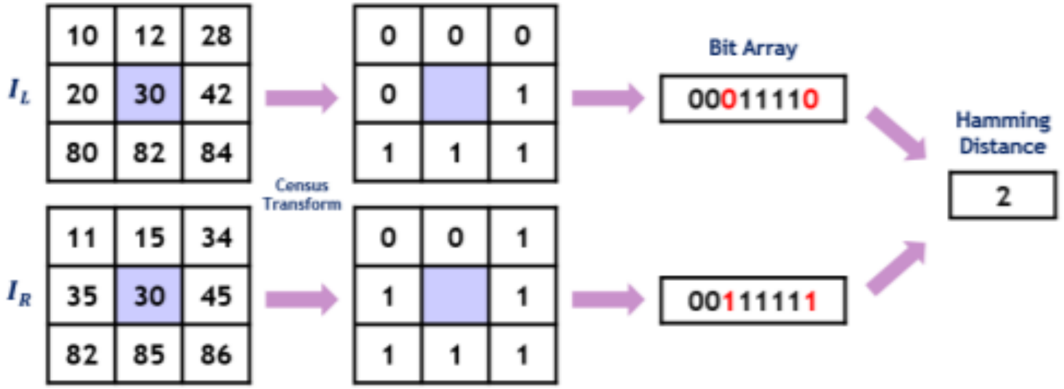
## Non-parametric Matching Costs

Soft Rank filter

$$I_{SoftRank}(\mathbf{p}) = \sum_{\mathbf{q}\in N_{\mathbf{p}}} \min\left(1, \max\left(0, \frac{I(\mathbf{p}) - I(\mathbf{q})}{2t} + \frac{1}{2}\right)\right).$$

The Soft Rank filter was proposed by Zitnick to reduce the problem of Rank filter by defining a linear, soft transition zone between 0 and 1 for values that are close together

Census Filter



# Cosine similarity

## Cosine similarity (cs) được dùng để đo mức độ giống nhau/tương đồng giữa hai vector

**Gọi $\vec{x}$ và $\vec{y}$ là hai vector, cs được tính như sau**

$$cs(\vec{x},\vec{y}) = \frac{\vec{x}.\vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \frac{\sum_1^n x_i y_i}{\sqrt{\sum_1^n x_i^2}\sqrt{\sum_1^n y_i^2}}$$

**Tính chất 1:** $cs(\vec{x},\vec{y}) = cs(a\vec{x}, b\vec{y})$

$$cs(a\vec{x}, b\vec{y}) = \frac{a\vec{x}.b\vec{y}}{\|a\vec{x}\|\|b\vec{y}\|} = \frac{\sum_1^n ax_i by_i}{\sqrt{\sum_1^n a^2x_i^2}\sqrt{\sum_1^n b^2y_i^2}}$$

$$= \frac{ab\sum_1^n x_i y_i}{\sqrt{a^2\sum_1^n x_i^2}\sqrt{b^2\sum_1^n y_i^2}}$$

$$= \frac{\sum_1^n x_i y_i}{\sqrt{\sum_1^n x_i^2}\sqrt{\sum_1^n y_i^2}} = cs(\vec{x},\vec{y})$$

**Tính chất 2:** $cs(\vec{x},\vec{y}) \neq cs(\vec{x}+c, \vec{y}+d)$

**Ví dụ:**

$\vec{x} = [4, 2, 1, 2]^T$

$\vec{y} = [1, 2, 2, 0]^T$

$\vec{u} = 2\vec{x} = [8, 4, 2, 4]^T$

$\vec{v} = 3\vec{y} = [3, 6, 6, 0]^T$

$$cs(\vec{x},\vec{y}) = \frac{4*1+2*2+1*2+2*0}{\sqrt{4^2+2^2+1^2+2^2}\sqrt{1^2+2^2+2^2+0}} = \frac{10}{\sqrt{25}\sqrt{9}} = \frac{10}{15} = 0.67$$

$$cs(\vec{u},\vec{v}) = \frac{8*3+4*6+2*6+4*0}{\sqrt{8^2+4^2+2^2+4^2}\sqrt{3^2+6^2+6^2+0}} = \frac{60}{\sqrt{100}\sqrt{81}} = \frac{60}{90} = 0.67$$

$$= cs(\vec{x},\vec{y})$$

# Cosine Similarity for Matching Cost

**Cosine similarity (cs) is used to measure the similarity between two vectors**

Let $\vec{x}$ and $\vec{y}$ be two vectors, cs is defined as

$$cs(\vec{x},\vec{y}) = \frac{\vec{x}.\vec{y}}{\|\vec{x}\|\|\vec{y}\|} = \frac{\sum_1^n x_i y_i}{\sqrt{\sum_1^n x_i^2}\sqrt{\sum_1^n y_i^2}}$$

**Property 1:** $cs(\vec{x},\vec{y}) = cs(a\vec{x}, b\vec{y})$; $ab > 0$

$$cs(a\vec{x}, b\vec{y}) = \frac{a\vec{x}.b\vec{y}}{\|a\vec{x}\|\|b\vec{y}\|} = \frac{\sum_1^n a x_i b y_i}{\sqrt{\sum_1^n a^2 x_i^2}\sqrt{\sum_1^n b^2 y_i^2}}$$

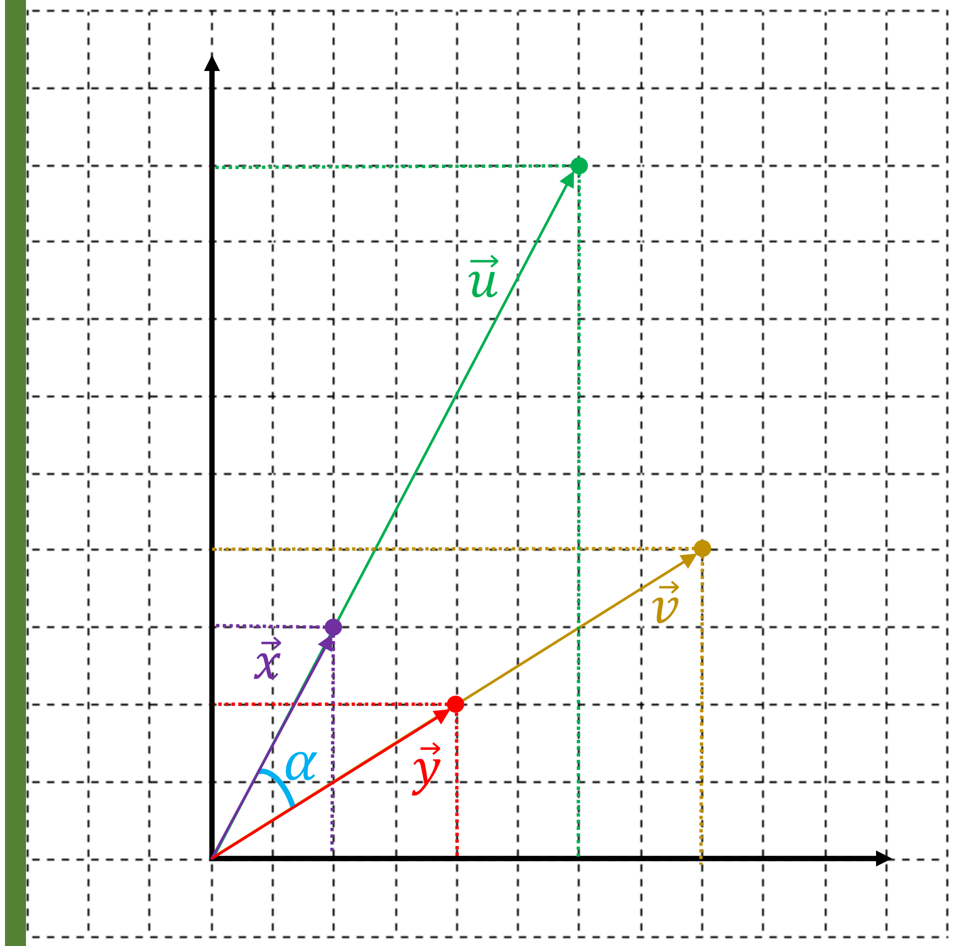
$$= \frac{ab\sum_1^n x_i y_i}{\sqrt{a^2\sum_1^n x_i^2}\sqrt{b^2\sum_1^n y_i^2}}$$

$$= \frac{\sum_1^n x_i y_i}{\sqrt{\sum_1^n x_i^2}\sqrt{\sum_1^n y_i^2}} = cs(\vec{x},\vec{y})$$

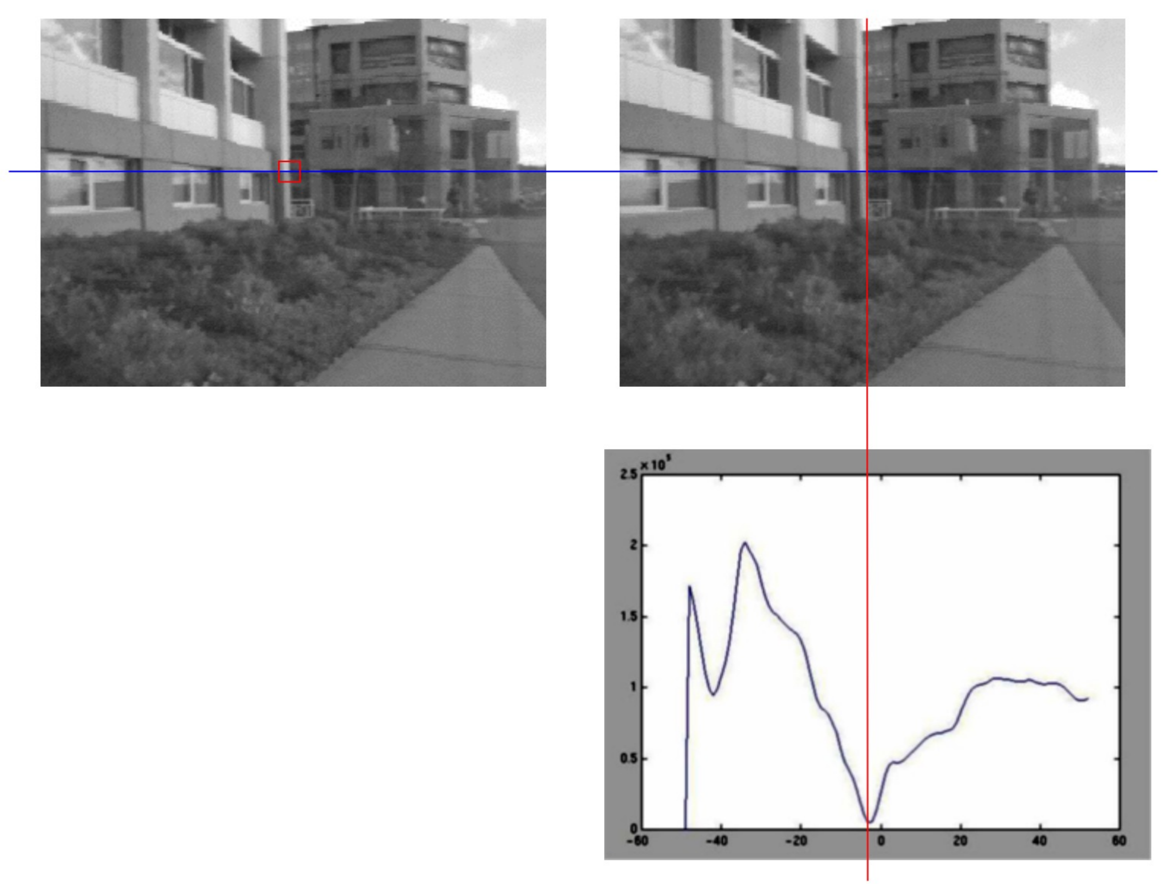
**Property 2:** $cs(\vec{x},\vec{y}) \neq cs(\vec{x}+c, \vec{y}+d)$

$\vec{x} = \binom{2}{3}$ $\quad \vec{u} = 3 * \vec{x} = \binom{6}{9}$

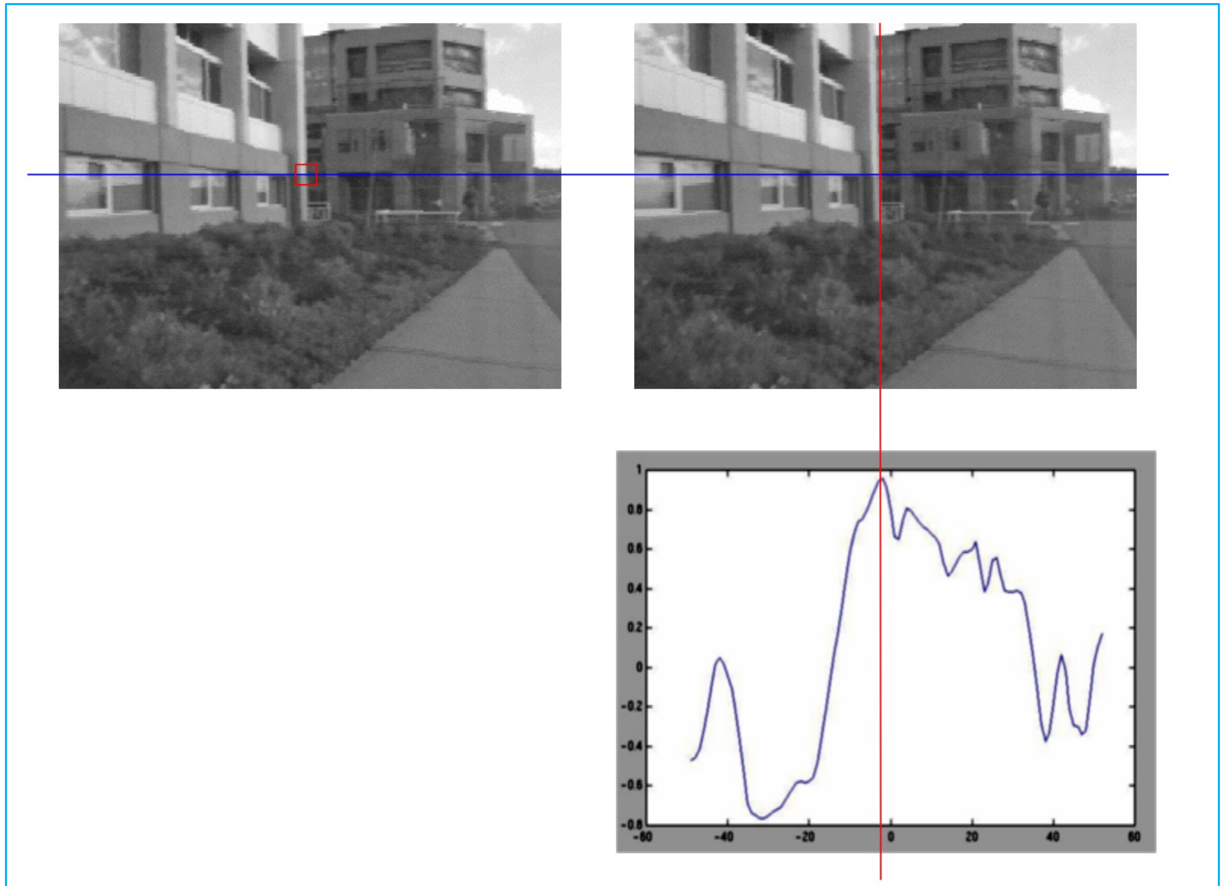
$\vec{y} = \binom{4}{2}$ $\quad \vec{v} = 2 * \vec{y} = \binom{8}{4}$



# Stereo Matching Cost




SSD



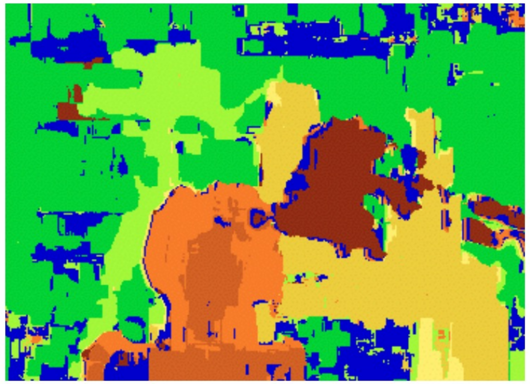

Normalized cross-correlation

# When will stereo block matching fail?

# How can we improve depth estimation?

Too many discontinuities. We expect disparity values to change slowly.

Block matching



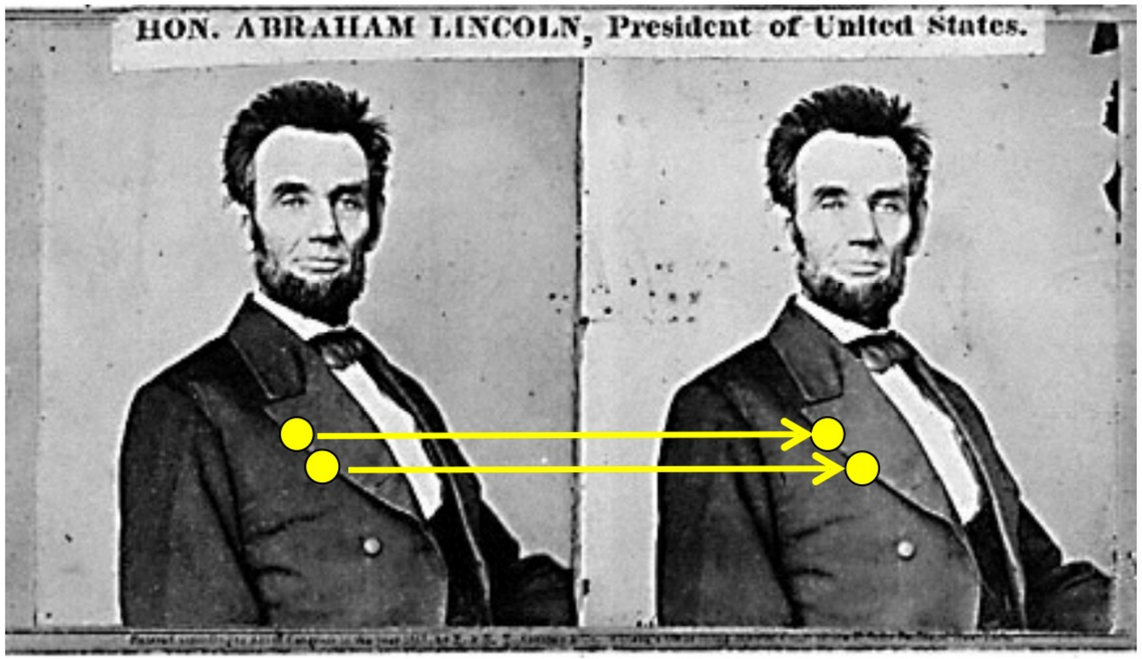
Ground truth

Let's make an assumption: depth should change smoothly

# How can we improve depth estimation?

Stereo matching as …

## Energy Minimization




**What defines a good stereo correspondence?**

1. **Match quality** – Want each pixel to find a good match in the other image
2. **Smoothness** – If two pixels are adjacent, they should (usually) move about the same amount

energy function
(for one pixel)

$$E(d) = \underbrace{E_d(d)}_{\text{data term}} + \lambda \underbrace{E_s(d)}_{\text{smoothness term}}$$

Want each pixel to find a good match in the other image

(block matching result)

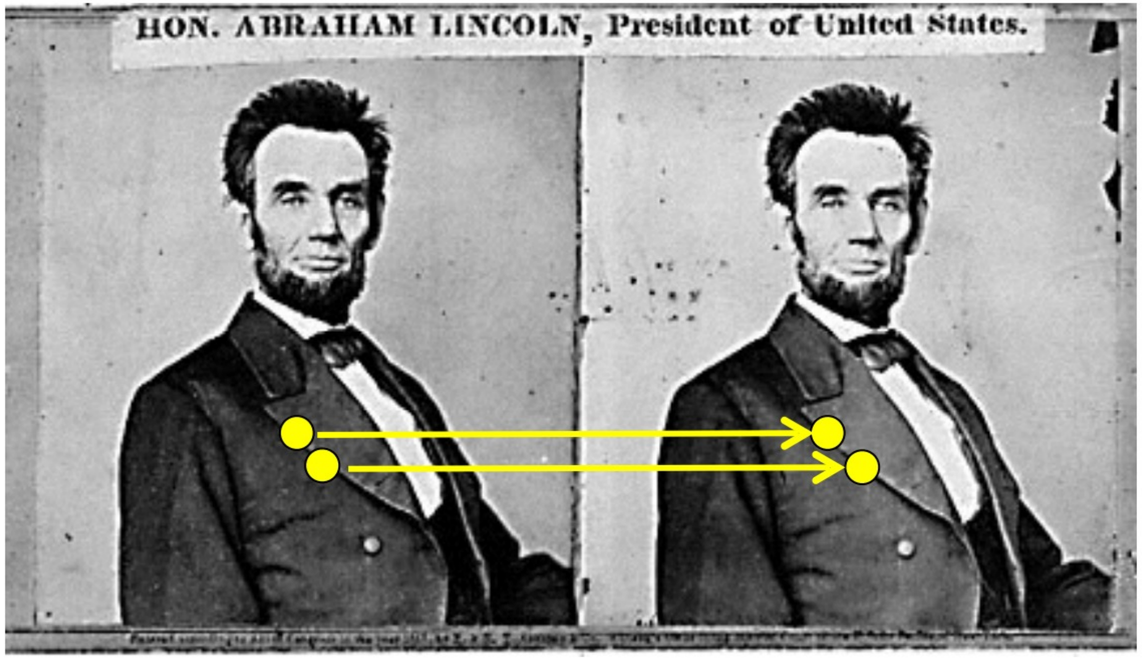
Adjacent pixels should (usually) move about the same amount

(smoothness function)

# How can we improve depth estimation?

Stereo matching as …

## Energy Minimization




$$E(d) = E_d(d) + \lambda \boxed{E_s(d)}$$

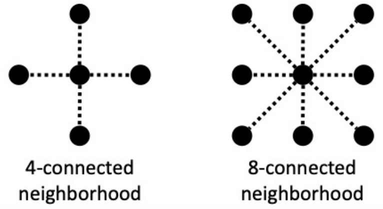
$$E_d(d) = \sum_{(x,y) \in I} C(x, y, d(x, y))$$

SSD distance between windows centered at I(x, y) and J(x+ d(x,y), y)

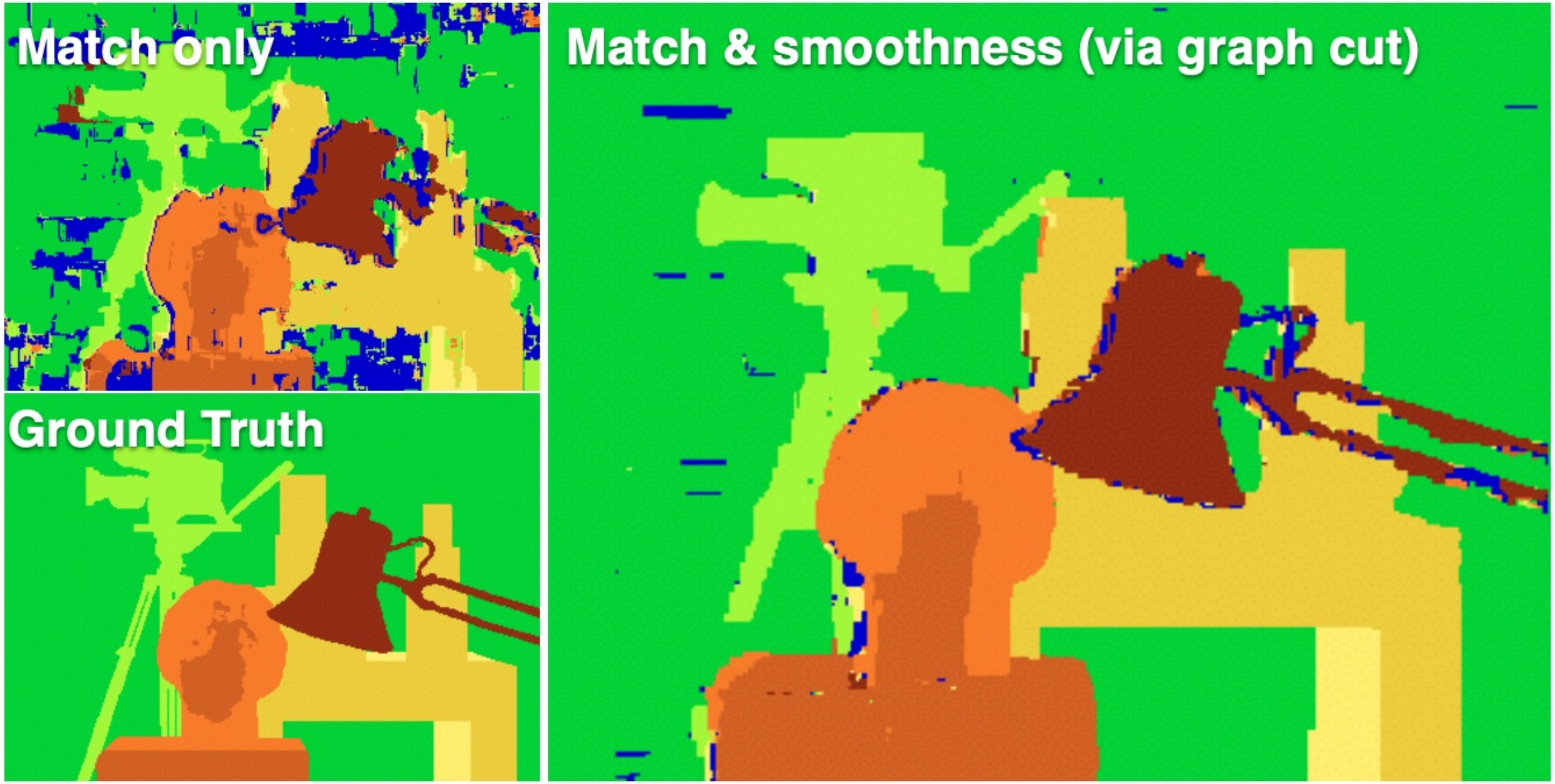
$$E_s(d) = \sum_{(p,q) \in \mathcal{E}} V(d_p, d_q)$$

smoothness term

$\mathcal{E}$ : set of neighboring pixels

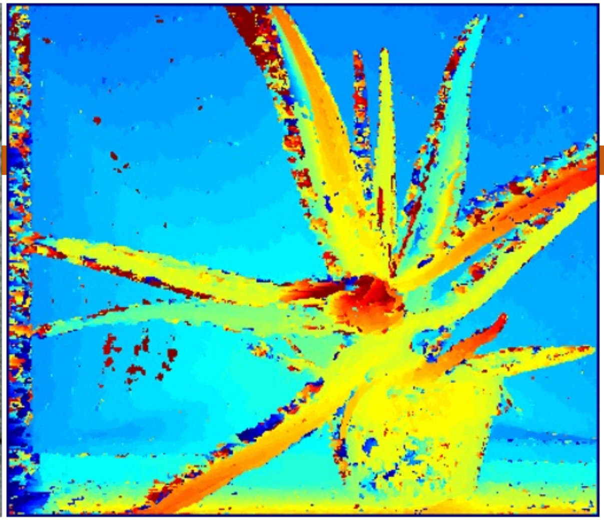


# How can we improve depth estimation?

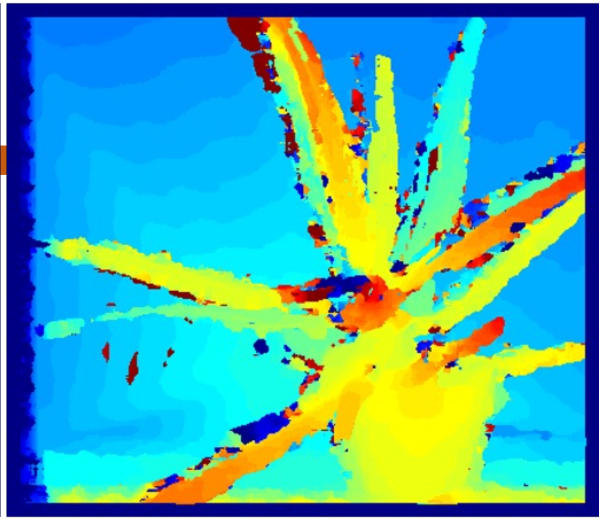



Y. Boykov, O. Veksler, and R. Zabih, Fast Approximate Energy Minimization via Graph Cuts, PAMI 2001

Ảnh stereo có cùng độ sáng

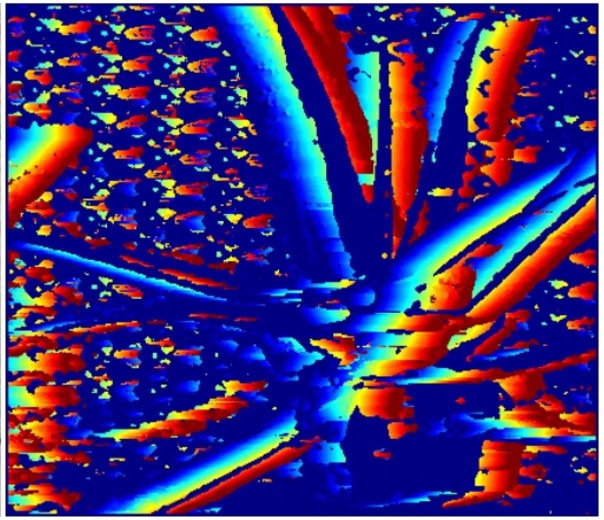


Absolute difference

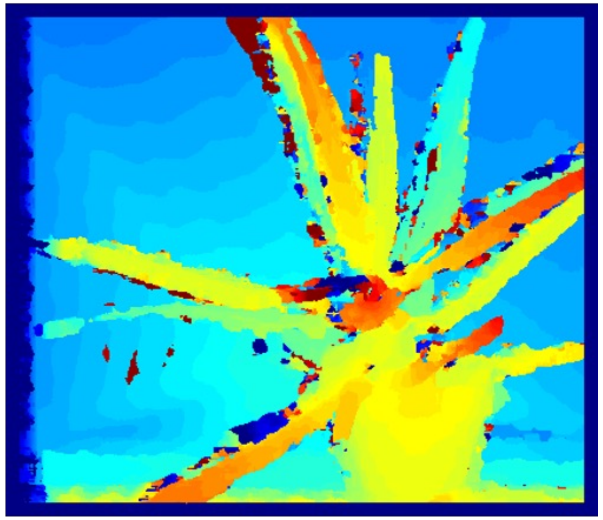


Cosine Similarity

Ảnh stereo khác độ sáng



Absolute difference



Cosine Similarity

**Cosine similarity hoạt động ổn định khi ảnh stereo thay đổi độ sáng**

## Correlation Coefficient Matching Cost



$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$$

ignore the differences between population and sample

### ❖ Definition

**Công thức: Gọi x,y là hai biến ngẫu nhiên**

$$\rho_{xy} = \frac{E[(x-\mu_x)(y-\mu_y)]}{\sqrt{var(x)}\sqrt{var(y)}}$$
$$= \frac{n(\sum_i x_i y_i) - (\sum_i x_i)(\sum_i y_i)}{\sqrt{n\sum_i x_i^2 - (\sum_i x_i)^2}\sqrt{n\sum_i y_i^2 - (\sum_i y_i)^2}}$$

**Tính chất 1**

$$-1 \leq \rho_{xy} \leq 1$$

Tương quan nghịch

Tương quan thuận

**Tính chất 2**

$$\rho_{xy} = \rho_{uv}$$

trong đó

$$u = ax + b$$
$$v = cy + d$$

**Ví dụ 1**

x = [7, 18, 29, 2, 10, 9, 9]
y = [1, 6, 12, 8, 6, 21, 10]

$$\rho_{xy} = \frac{E[(x-\mu_x)(y-\mu_y)]}{\sqrt{var(x)}\sqrt{var(y)}}$$
$$= \frac{n*818 - 84*64}{\sqrt{n*1480 - 7056}\sqrt{n*822 - 4096}} = 0.149$$

**Ví dụ 2**

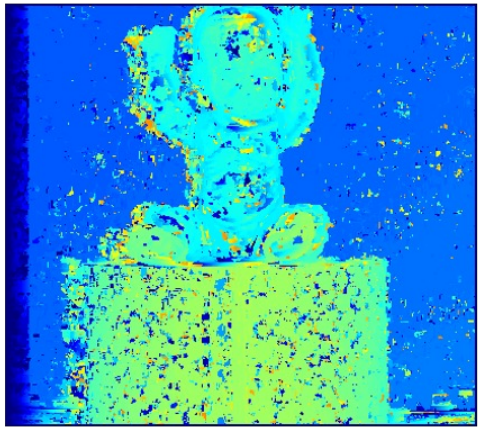
u=2*x-14 = [0, 22, 44, -10, 6, 4, 4]
v=y+2 = [3, 8, 14, 10, 8, 23, 12]

$$\rho_{uv} = \frac{E[(u-\mu_u)(v-\mu_v)]}{\sqrt{var(u)}\sqrt{var(v)}}$$
$$= \frac{n*880 - 70*78}{\sqrt{n*2588 - 4900}\sqrt{n*1106 - 6084}} = 0.149$$

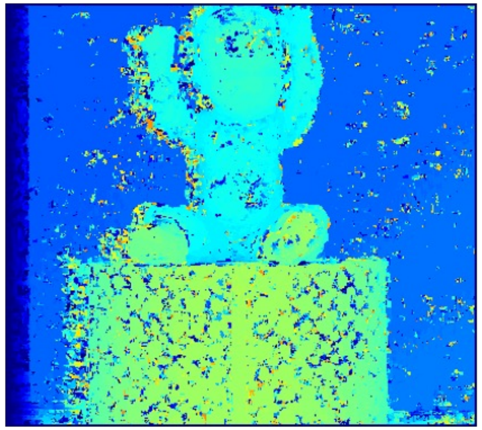
# Correlation for Stereo Matching Cost

Left image

Right image



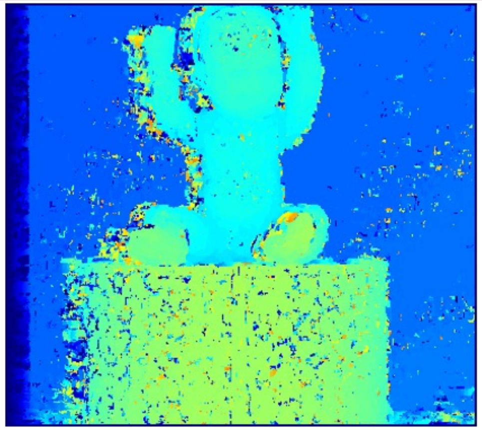
Cosine Result/ RMS = 5.6



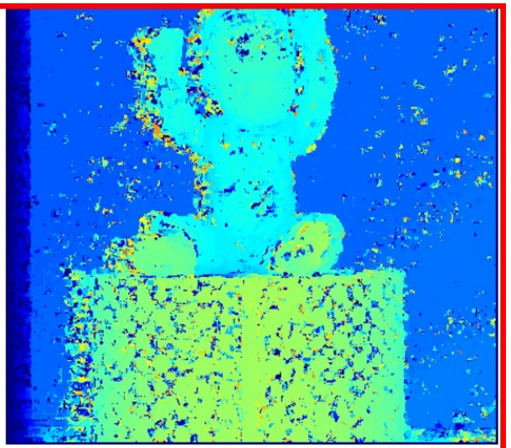
Correlation Result / RMS = 5.38

Left image

Right image



Cosine Result / RMS = 4.98



Correlation Result / RMS = 5.36

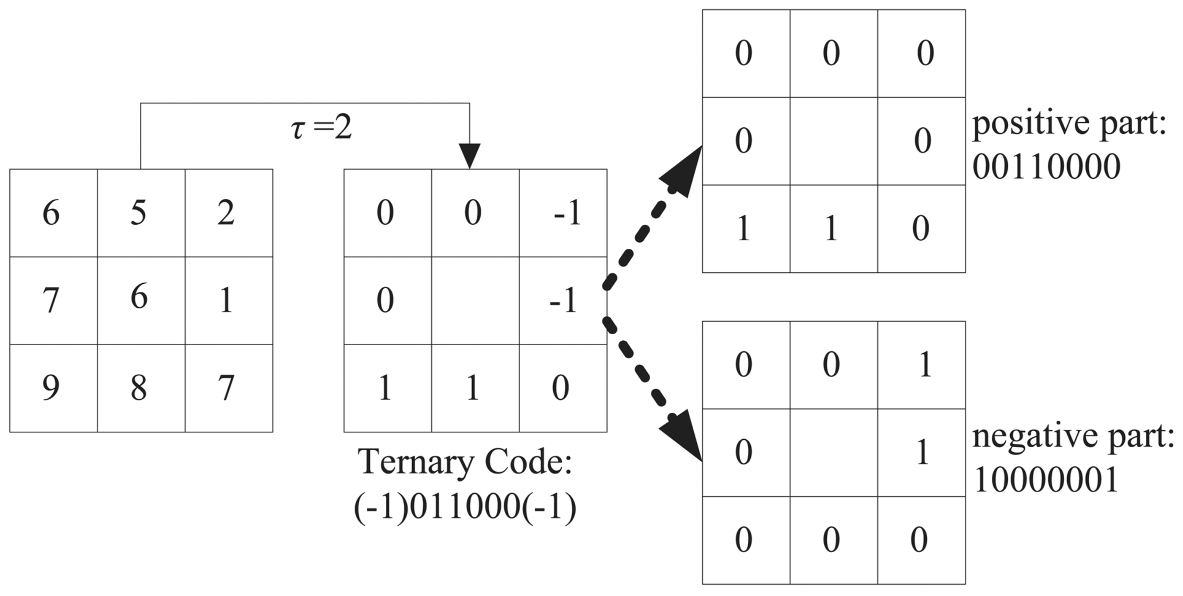
# LTP Matching Cost

$$\mathrm{LTP}_{S,P} = \sum_{y=0}^{S-1} 2^y h(i_y - i_c)$$

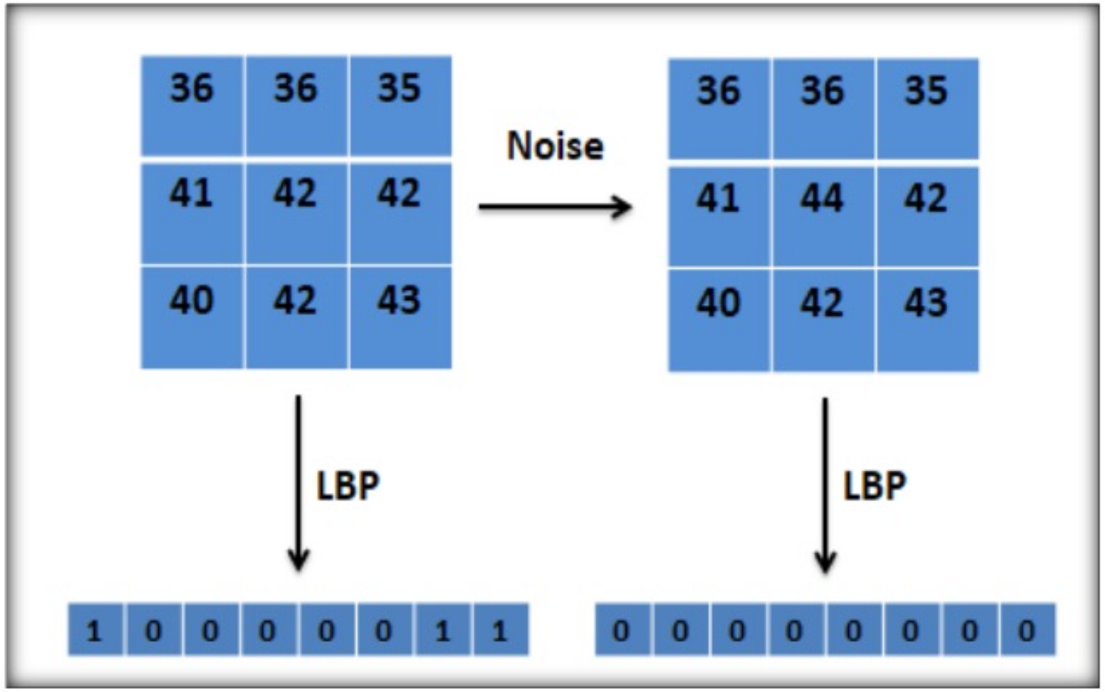
$$\text{Where} \quad h(i_y - i_c) = \begin{cases} 1 & i_y - i_c \geq t \\ 0 & -t < i_y - i_c < t \\ -1 & i_y - i_c < -t \end{cases}$$

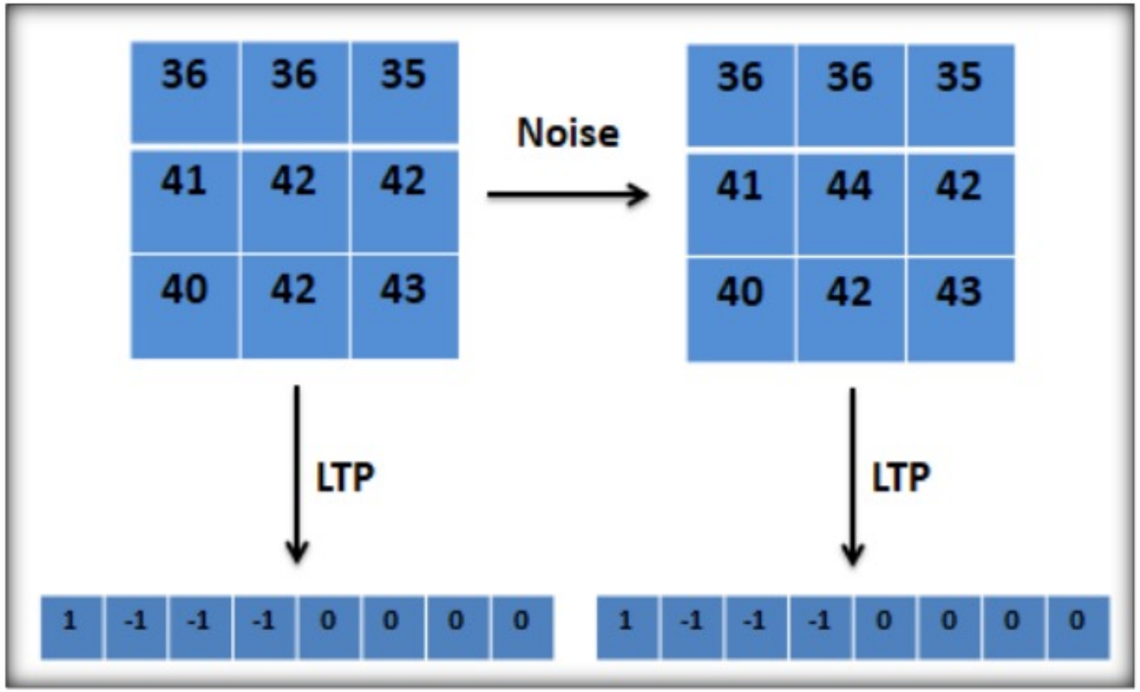
$$\mathrm{LBP}_{S,P} = \sum_{y=0}^{S-1} 2^y h(i_y - i_c)$$

$$\text{Where,} \quad h(i_y - i_c) = \begin{cases} 1 & i_y - i_c \geq 0 \\ 0 & i_y - i_c < 0 \end{cases}$$



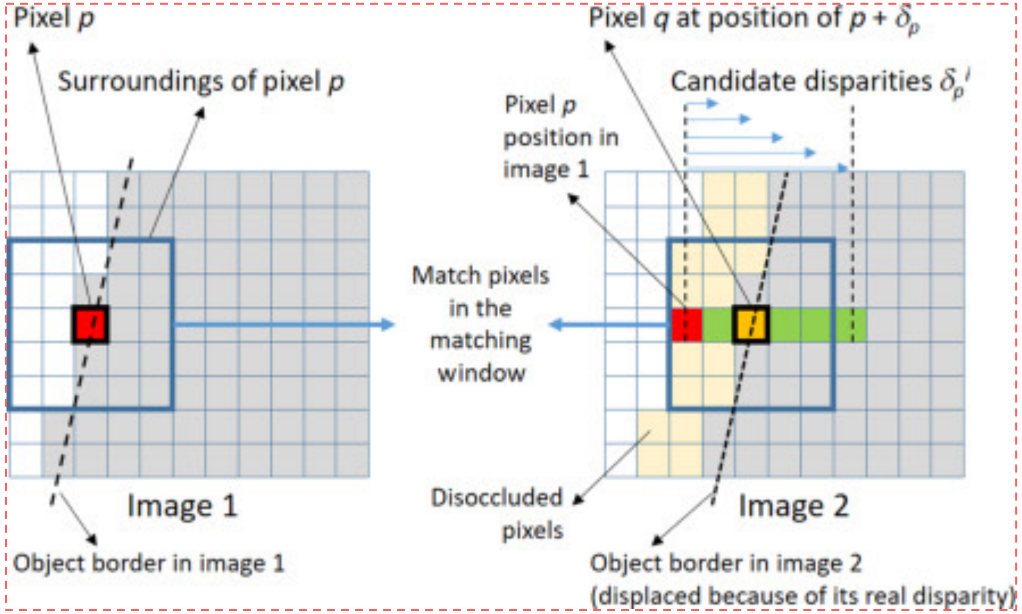
# LTP Matching Cost

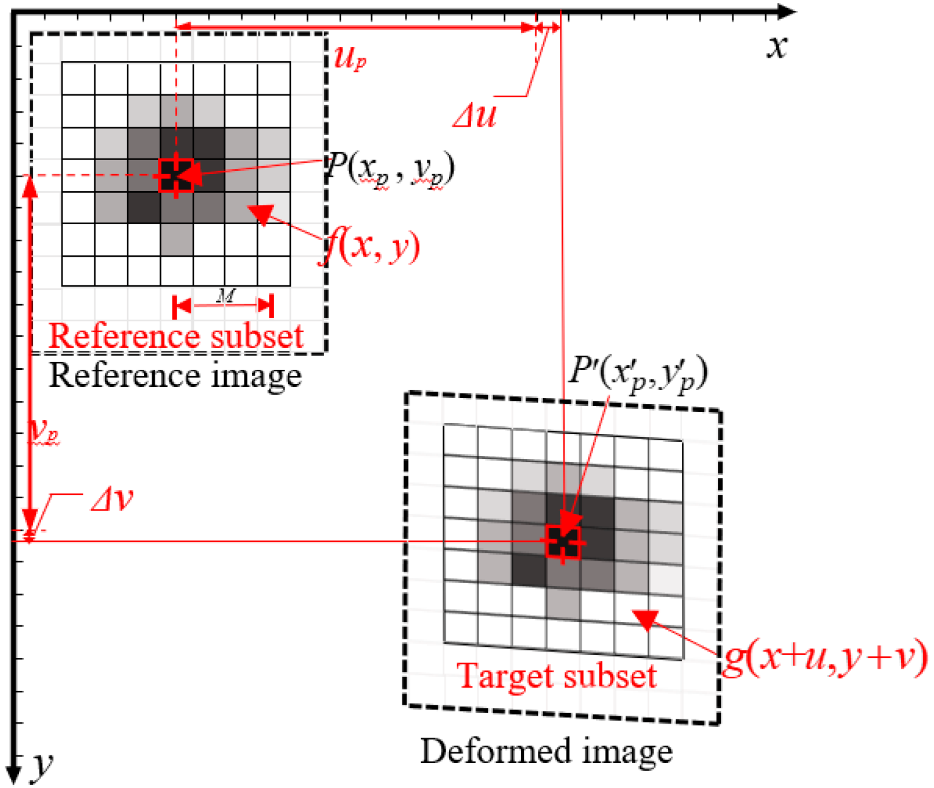


# Stereo Matching

## ❖ Local/Window-based stereo matching

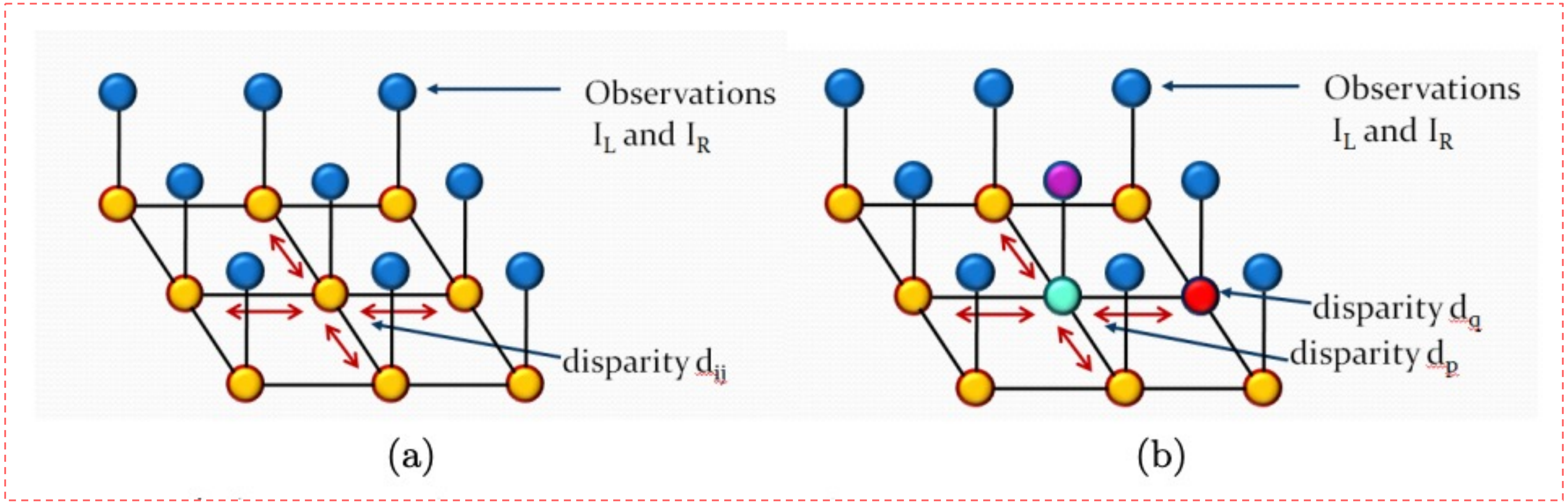



In the case of local approaches, the search area for finding a match is confined to a fixedsize window within the search image

# Stereo Matching

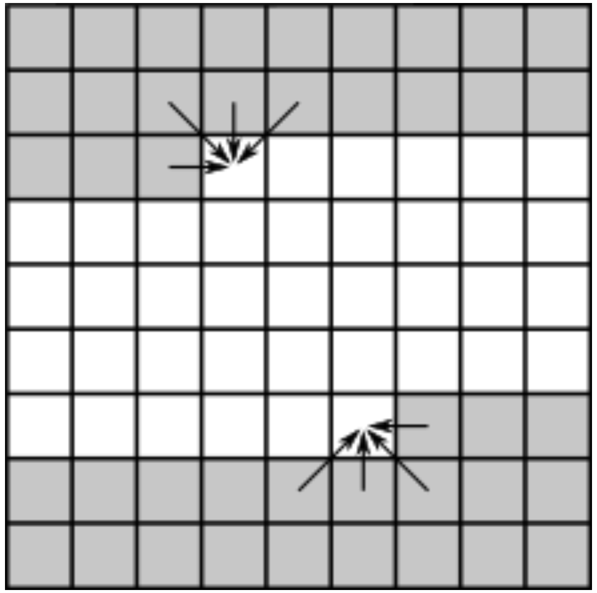
## ❖ Global-based stereo matching

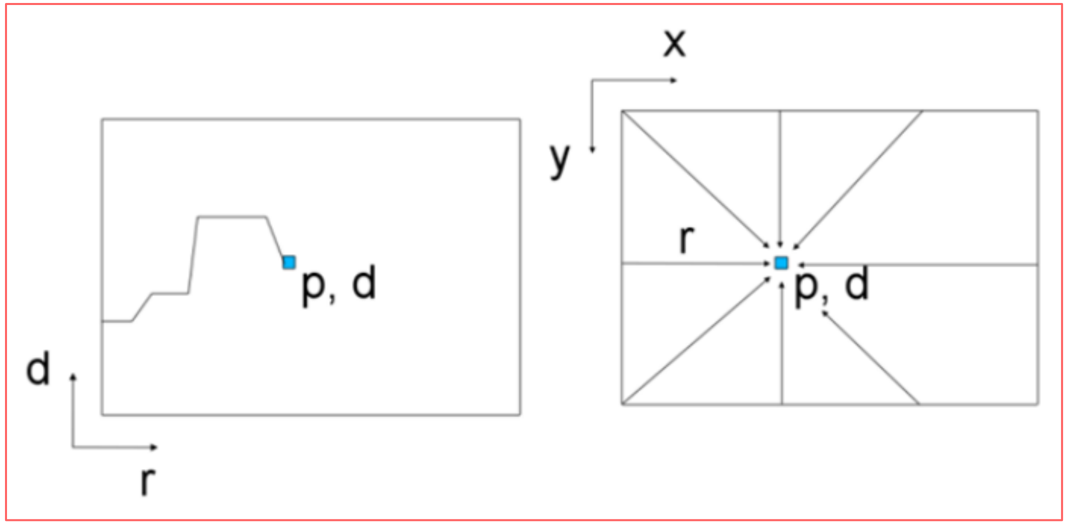



$$E(d) = \sum_{p \in I} D(d_p) + \lambda \sum_{(p,q) \in N} S(d_p, d_q),$$

# Stereo Matching

## ❖ Semi-global stereo matching

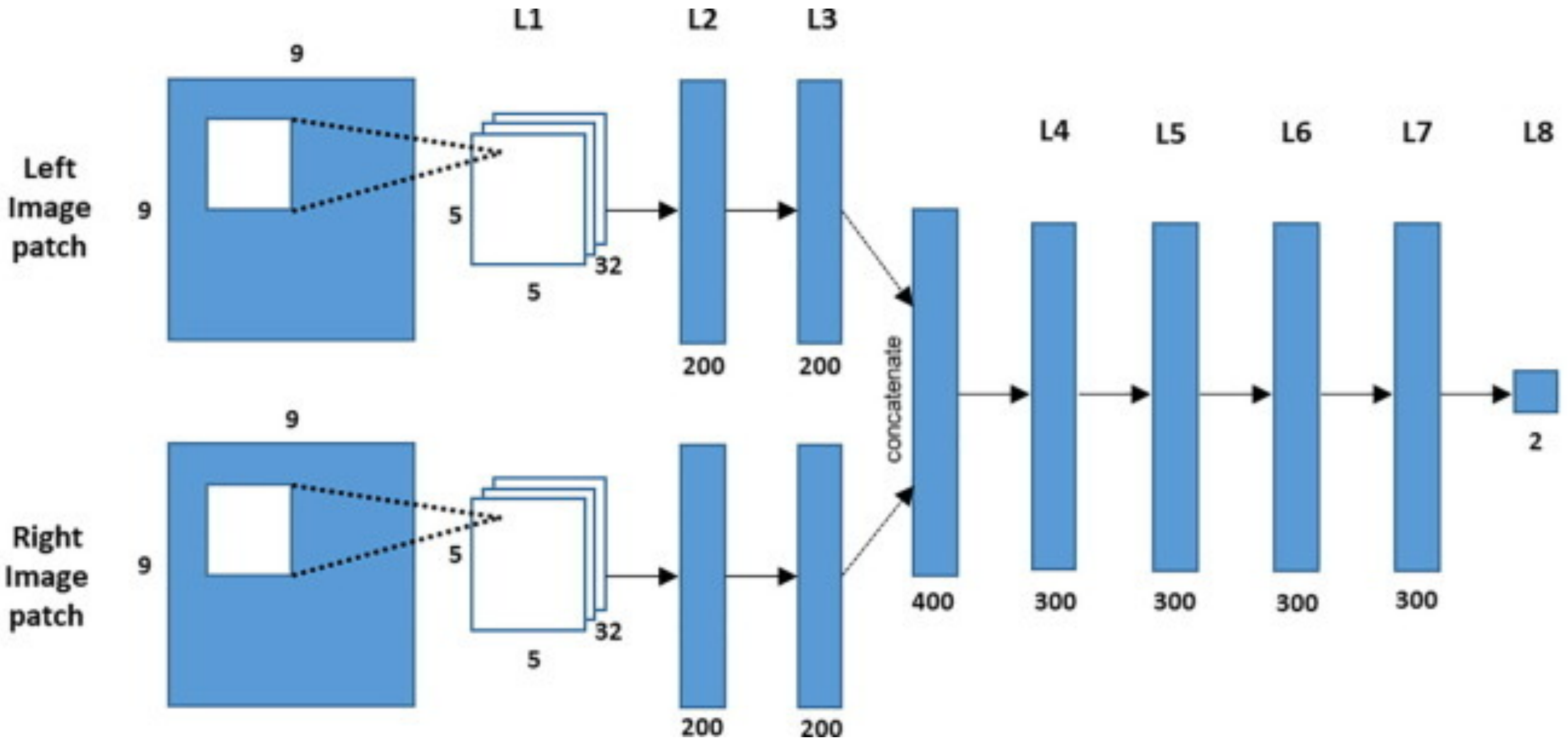





$$E(D) = \sum_{p} \Big( C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + \sum_{q \in N_p} P_2 T[|D_p - D_q| > 1] \Big). \qquad (2)$$

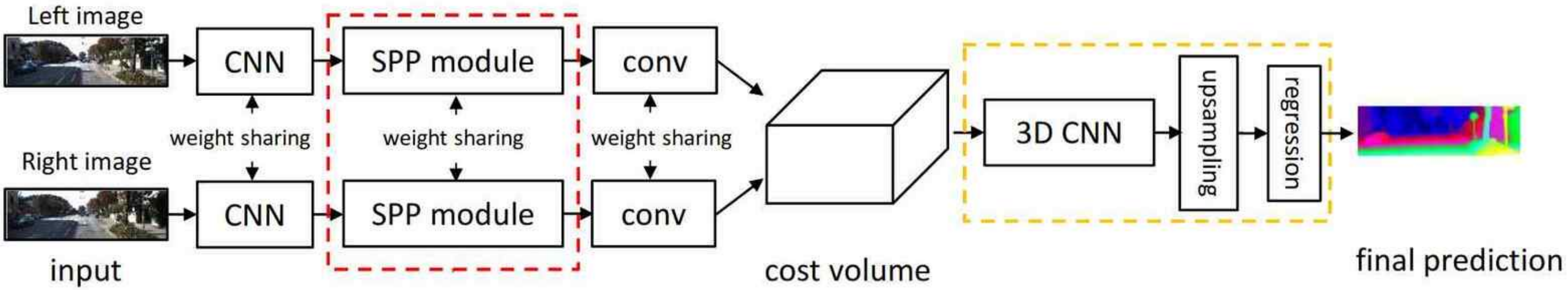
# Stereo Matching

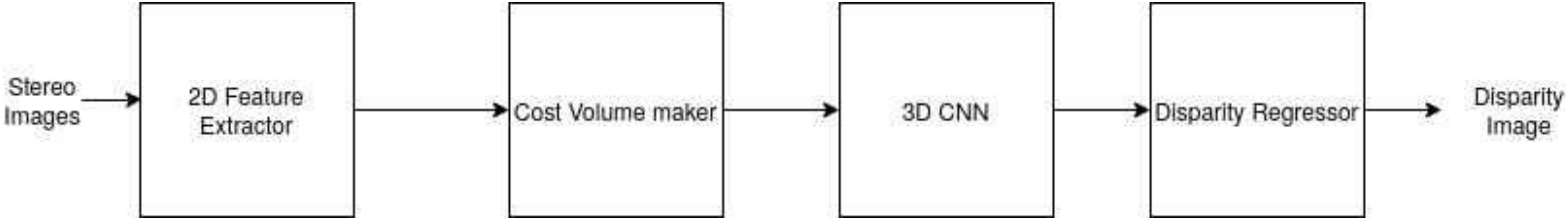
## ❖ CNN-based Stereo Matching



# Stereo Matching

## ❖ CNN-based Stereo Matching



# Outline

- Research Plan
- Challenges of the Research Topics
- Brainstorming idea
- Performance Evaluation
- Assignment

AI VIETNAM
All-in-One Course

# Datasets

The KITTI Vision Benchmark Suite
A project of Karlsruhe Institute of Technology and Toyota Technological Institute at Chicago

Karlsruhe Institute of Technology

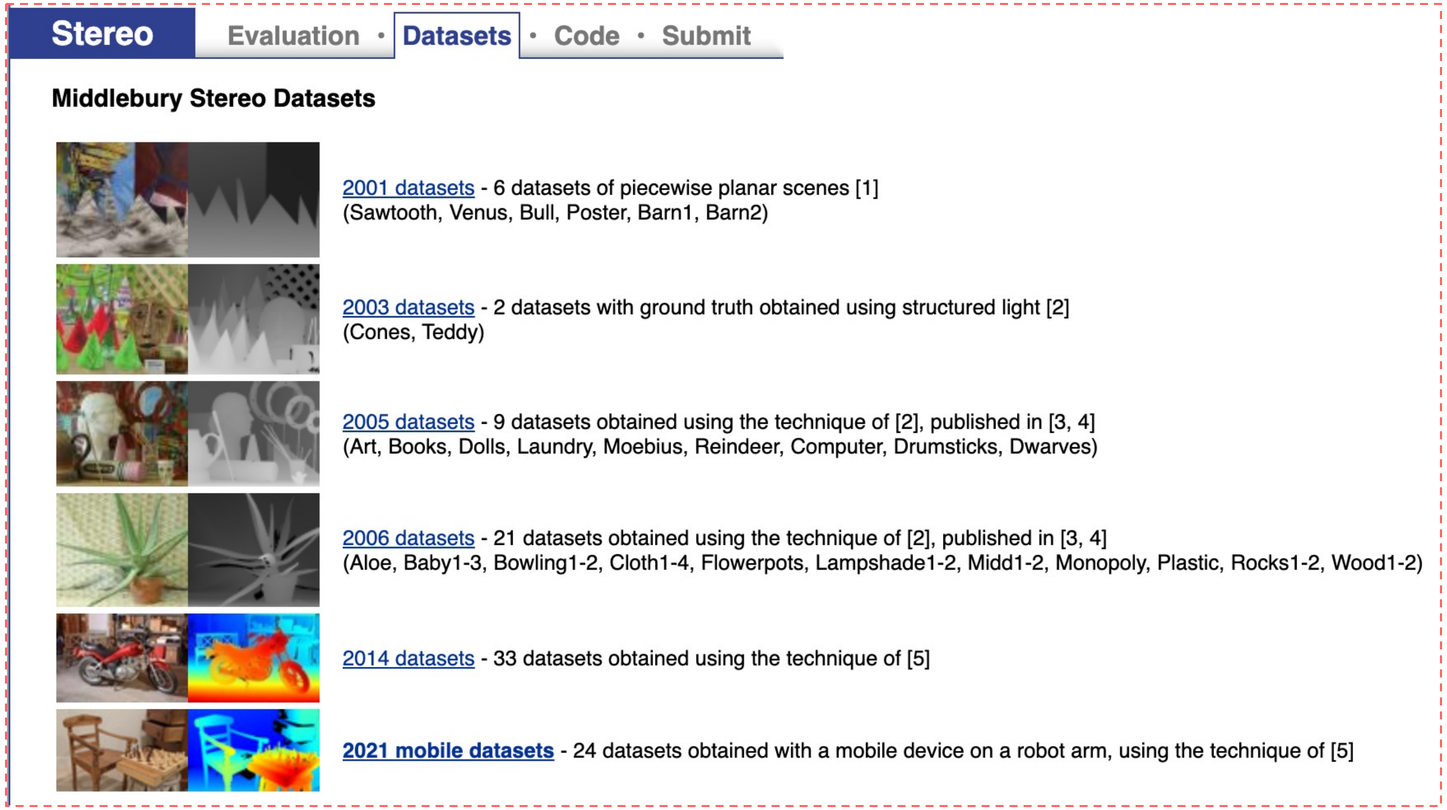
home setup stereo flow sceneflow depth odometry object tracking road semantics raw data submit results

A. Geiger | P. Lenz | C. Stiller | R. Urtasun Log in

## Stereo Evaluation 2015

The **stereo 2015** / **flow 2015** / **scene flow 2015** benchmark consists of 200 training scenes and 200 test scenes (4 color images per scene, saved in loss less png format). Compared to the stereo 2012 and flow 2012 benchmarks, it comprises dynamic scenes for which the ground truth has been established in a semi-automatic process. Our evaluation server computes the percentage of bad pixels averaged over all ground truth pixels of all 200 test images. For this benchmark, we consider a pixel to be correctly estimated if the disparity or flow end-point error is **<3px or <5%** (for scene flow this criterion needs to be fulfilled for both disparity maps and the flow map). We require that all methods use the same parameter set for all test pairs. Our development kit provides details about the data format as well as MATLAB / C++ utility functions for reading and writing disparity maps and flow fields. More details can be found in Object Scene Flow for Autonomous Vehicles (CVPR 2015).

- Download stereo 2015/flow 2015/scene flow 2015 data set (2 GB)
- Download calibration files (1 MB)
- Download multi-view extension (20 frames per scene) (14 GB)
- Download development kit (3 MB)



Stereo Evaluation • Datasets • Code • Submit

### Middlebury Stereo Datasets

2001 datasets - 6 datasets of piecewise planar scenes [1]
(Sawtooth, Venus, Bull, Poster, Barn1, Barn2)

2003 datasets - 2 datasets with ground truth obtained using structured light [2]
(Cones, Teddy)

2005 datasets - 9 datasets obtained using the technique of [2], published in [3, 4]
(Art, Books, Dolls, Laundry, Moebius, Reindeer, Computer, Drumsticks, Dwarves)

2006 datasets - 21 datasets obtained using the technique of [2], published in [3, 4]
(Aloe, Baby1-3, Bowling1-2, Cloth1-4, Flowerpots, Lampshade1-2, Midd1-2, Monopoly, Plastic, Rocks1-2, Wood1-2)

2014 datasets - 33 datasets obtained using the technique of [5]

**2021 mobile datasets** - 24 datasets obtained with a mobile device on a robot arm, using the technique of [5]

# Evaluation Metrics

| | Method | Setting | Code | D1-bg | D1-fg | D1-all | Density | Runtime | Environment |
|---|---|---|---|---|---|---|---|---|---|
| 1 | MoCha-Stereo | | | 1.36 % | 2.43 % | **1.53 %** | 100.00 % | 0.27 s | NVIDIA Tesla A6000 (PyTorch) |
| 2 | DiffuVolume | | | 1.35 % | 2.51 % | 1.54 % | 100.00 % | 0.36 s | GPU @ 2.5 Ghz (Python) |
| 3 | GANet+ADL | | | 1.38 % | 2.38 % | 1.55 % | 100.00 % | 0.67s | NVIDIA RTX 3090 (PyTorch) |
| 4 | Selective-IGEV | | | 1.33 % | 2.61 % | 1.55 % | 100.00 % | 0.24 s | 1 core @ 2.5 Ghz (Python) |
| 5 | MC-Stereo | | | 1.36 % | 2.51 % | 1.55 % | 100.00 % | 0.40 s | 1 core @ 2.5 Ghz (Python) |
| 6 | IGEV-ICGNet | | | 1.38 % | 2.55 % | 1.57 % | 100.00 % | 0.18 s | NVIDIA Tesla A5000 (Pytorch) |
| 7 | yjlig | | | 1.37 % | 2.62 % | 1.58 % | 100.00 % | 0.35 s | 1 core @ 2.5 Ghz (C/C++) |
| 8 | MDA | | | 1.37 % | 2.64 % | 1.58 % | 100.00 % | 0.32 s | 1 core @ 2.5 Ghz (Python) |
| 9 | UGNet | | | 1.34 % | 2.77 % | 1.58 % | 100.00 % | 0.2 s | GPU @ 3.0 Ghz (Python) |
| 10 | Any-IGEV | | | 1.43 % | 2.35 % | 1.58 % | 100.00 % | 0.32 s | GPU @ 2.5 Ghz (Python) |
| 11 | OpenStereo-IGEV | | code | 1.44 % | **2.31 %** | 1.59 % | 100.00 % | 0.18 s | NVIDIA-3090 |

# Evaluation Metrics

**Percentage of Bad Pixels:** Measures the percentage of pixels with disparity errors exceeding a specified threshold. Common thresholds include 1, 2, or 3 pixels.

$$BP = \frac{\text{Number of Bad Pixels}}{\text{Total Number of Pixels}} \times 100\%$$

**Root Mean Squared Error (RMSE):** Calculates the square root of the average squared disparity errors across all pixels. It provides a measure of the overall accuracy.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(D_i - \hat{D}_i)^2}$$

**Mean Absolute Error (MAE):** Computes the average absolute disparity errors. It is less sensitive to outliers than RMSE.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|D_i - \hat{D}_i|$$

**The "d1-bg"** error is specifically defined as the percentage of disparity errors (in pixels) where the absolute disparity error is less than 1 pixel and the error occurs in regions classified as foreground or background

$$\text{d1-bg} = \frac{\text{Number of pixels with disparity error} < 1 \text{ pixel and in fg/bg}}{\text{Total number of pixels in fg/bg}} \times 100\%$$

# Common Challenges

### *Lack of Clear Objectives*

*Problem:* Unclear research goals and objectives can lead to confusion and make it difficult to stay focused.
*Solution:* Clearly define your research questions and objectives before starting.

### *Literature Review Challenges*

- *Problem:* Difficulty in finding relevant literature or understanding the existing body of knowledge.
- *Solution:* Invest time in a thorough literature review, use academic databases, and seek guidance from mentors.

### *Data Collection Issues*

*Problem*: Obtaining accurate and sufficient data can be a significant challenge.
*Solution*: Plan your data collection method carefully, consider potential biases, and validate your data sources.

### *Methodological Challenges:*

*Problem*: Issues related to the research design, sampling, or experimental procedures.
*Solution*: Consult with experienced researchers, follow established methodologies, and be open to adapting your approach.

# Common Challenges

### *Lack of Clear Analytical Difficulties*

•Problem: Challenges in analyzing data or interpreting results.

•Solution: Seek statistical or methodological advice, use appropriate software, and consider collaborating with experts.

### *Time Management*

•*Problem:* Balancing research with other commitments can be challenging.

•*Solution:* Develop a realistic timeline, set milestones, and prioritize tasks effectively.

### *Ethical Dilemmas*

•*Problem*: Unforeseen ethical issues or conflicts may arise.

•*Solution*: Consult with an ethics committee, adhere to ethical guidelines, and prioritize participant well-being.

### *Communication Challenges:*

•*Problem:* Difficulty in communicating research findings or collaborating with peers.

•*Solution:* Work on effective communication skills, present at conferences, and collaborate with colleagues.

# Fail to Improve Performance of Existing Methods

### *Thorough Analysis*

•Examine your research method and results critically. Identify the specific aspects where improvement is needed. Understand the limitations and challenges you are facing.

### *Literature Review*

•Revisit the literature to ensure you are aware of the latest developments in your field. Identify if there are new methods or techniques that could be applied to your research problem

### *Thorough Analysis*

•Seek feedback from experienced researchers, colleagues, or mentors. Discuss your challenges and ask for their insights. They may offer valuable suggestions or alternative approaches.

### *Collaborate with Peers*

•Collaborating with other researchers working on similar problems can provide fresh perspectives. Sharing ideas and collaborating on solving challenges can lead to innovative solutions.

### Consider a Multidisciplinary Approach:

•Explore whether incorporating knowledge from other disciplines could benefit your research. Sometimes, a multidisciplinary approach can provide novel solutions.

### *Experiment with Variations*

•Try variations of your existing method to identify potential improvements. Experiment with different parameters, data preprocessing techniques, or algorithms to see if any adjustments lead to better results.

# Fail to Improve Performance of Existing Methods

### *Validation and Benchmarking*

- Ensure that your evaluation metrics are appropriate and that you are comparing your method against suitable benchmarks. Reassess the validity of your experimental setup.

### *Document and Report Findings*

- Be transparent about the challenges you faced and the steps you took to address them. Document your findings, even if the improvements are minimal. This transparency contributes to the scientific community's collective knowledge.

### *Consider Alternative Approaches*

- If, despite your efforts, you are unable to achieve significant improvements, consider exploring alternative research approaches. Sometimes, a change in direction can lead to valuable insights.

# Outline

- Research Plan
- Challenges of the Research Topics
- Brainstorming idea
- Performance Evaluation
- Assignment

# Assignment 4

- Requirement:
  - Suggest an idea to solve limitations of existing systems (just pickup simplest idea)
  - Implement and evaluate the performance of the idea
  - Display several exprimental results
  - Submit to: aivnresearch@gmail.com
- Deadline: 12:00 09/12/2023
- Receiving Feedback: 1 week from TA Research Team of AIVN

# This Week

**Week 1**
How to read Research Paper

**Week 2**
How to find Research Topic and Develop Research Questions

**Week 3**
**Research Plan: Study and Implement Existing Algorithms**

**This Week**
Research Plan: *Brainstorming Idea and Implementation*

**Next Week**
Performance Evaluation and Plan to Write a Research Paper
*(If Any)*


Thank you