Final Exam Review Questions - Perception(Week 5)

Q1 What is the purpose of the pinhole camera model in Perception?

A. To simulate the effects of image distortion

B. To represent the lens distortion in real-world cameras

C. To simplify the mathematical modeling of camera projection

D. To model the color reproduction capabilities of a camera

Q2.What kind of information does LiDAR capture effectively compared to other cameras?

a) Color

b) Texture

c) Depth

d) Semantic meaning

Q3.What is a key advantage of using infrared cameras in certain industrial applications?

A. Higher color accuracy

B. Better depth perception

C. Ability to see through solid objects

D. Insensitivity to ambient lighting conditions

Q4. Which application is NOT commonly associated with depth sensors?

A. Virtual reality and augmented reality

B. Autonomous vehicles

C. Image denoising

D. Human-computer interaction

Q5.How does the baseline distance between stereo cameras affect stereo matching accuracy?

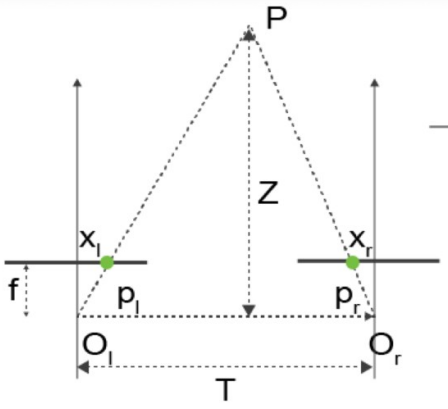
A. A shorter baseline improves accuracy

B. A longer baseline improves accuracy

C. Baseline distance has no impact on accuracy

D. Accuracy is independent of baseline distance

Basic Math



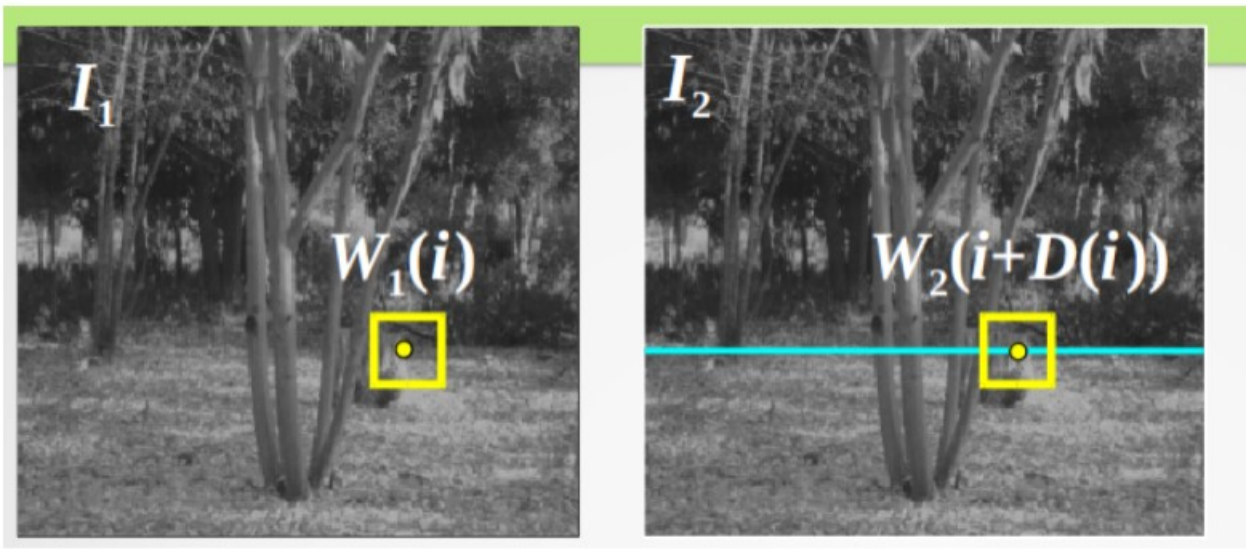
$$\frac{T + x_r - x_l}{Z - f} = \frac{T}{Z} \quad Z = f\frac{T}{x_l - x_r}$$

Disparity: $d = x_l - x_r$

$$Z = f\frac{T}{d}$$

Then given T and d, we can compute Z

- **T** is the stereo baseline
- **d** measures the difference in retinal position between corresponding points



Q6 How does stereo matching leverage a stereo pair of images?

A. By estimating camera poses

B. By finding corresponding points between the images

C. By applying edge detection algorithms

D. By analyzing color histograms.

Q7. How is optical flow typically computed in a video sequence?

A. Calculating the color histogram

B. Estimating the depth map

C. Analyzing frame-by-frame differences

D. Applying edge detection algorithms

Q8. What is the relationship between optical flow and the apparent velocity of objects in an image?

A. Directly proportional

B. Inversely proportional

C. No consistent relationship

D. Depends on the object's color

Q9.Which of the following is a common output of a Structure from Motion system?

A. Segmented images

B. 3D point cloud

C. Camera calibration matrix

D. Class labels for objects

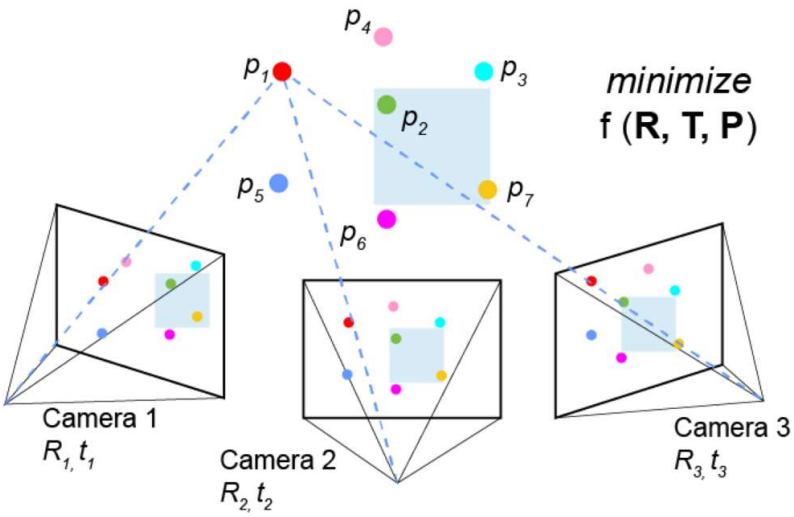
Q10. In Bundle Adjustment, what is being "adjusted" or optimized?

A. Camera poses and 3D points

B. Image resolution

C. Lighting conditions

D. Feature descriptors



Q11. What is the primary focus of feature-based recognition techniques?

A. Extracting overall visual appearance

B. Analyzing global image properties

C. Identifying and matching distinctive local features

D. Utilizing depth information for recognition

Q12. What loss function is commonly used for multi-class image classification?

a) Hinge loss

b) Cross Entropy Loss

c) Mean Absolute Error

d) Mean Squared Error

Q13.In the context of object detection, what does NMS (Non Maximum Suppression) aim to eliminate?

A. True negatives

B. False negatives

C. False positives

D. Background noise

Q14. What is the key difference between instance segmentation and semantic segmentation?

A. Instance segmentation focuses on classifying pixels, while semantic segmentation focuses on identifying individual instances.

B. Semantic segmentation assigns a class label to each pixel, while instance segmentation identifies and distinguishes individual object instances.

C. Instance segmentation is used for 3D reconstruction, while semantic segmentation is used for object detection.

D. Semantic segmentation is used for image classification, while instance segmentation is used for image segmentation.

Q15. How can image captioning models be improved to generate more descriptive and accurate captions?

a) Use larger neural network architectures

b) Train on more image-caption pairs

c) Optimize loss functions to specifically promote descriptive language

d) All of the above