1. How does finding corresponding points between image pairs contribute to 3D depth inference?

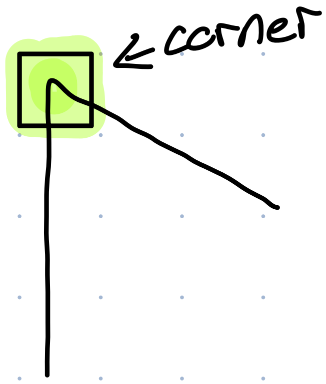
Das finden korrespondierender Punkte zwischen Bildpaaren ermöglicht die Berechnung der Disparität (Pixelverschiebung zwischen Bildern). Dies ist entscheidend für die 3D-Tiefenschätzung (3D depth inference), da die Disparität mit bekannten Kameraparametern (Brennweite f & Basislinie B) verwendet wird, um die Tiefe (Z) eines Punktes zu berechnen: $Z = \frac{f \cdot B}{d}$

Korrespondierende Punkte liefern somit die Grundlage für Tiefenkarten und die Rekonstruktion der 3D-Szene.

(2) Differentiate between sparse correspondance and dense correspondence in image matching.

- Spars Correspondence:
  - -> Korrespondenzen werden nur für wenige markante Punkte (features) gefunden.
  - -> Effizient, aber keine vollstandige Information
- Dense Correspondence:
  - -> Korrespondenzen für jedes Pixel im Bild
  - -> Rechenintensiv, liefert vollständige Karten.

③ What image characteristics define a „corner" as an easy-to-match feature?



Definition: a patch is „easy-to-match" if small shifts always produce a large difference from the original patch

„corner" has significant change in all directions

④ Describe the characteristics of an image patch that make it „easy to match" using the concept of the Sum of Squared Differences (SSD) error.

Definition: a patch is „easy-to-match" if
small shifts always produce a large
difference from the original patch -> for example,
a large SSD error

↳ Eigenvector = direction [u, v] of shifting the patch
Eigenvalue = SSD error

Corner: both eigenvalues are large

linear Edge: 1 eigenvalue is small
1 eigenvalue is large

smallest shift has to be large, then
it is a corner

$$\text{cornerness}(x_0, y_0) = \min_{u^2+v^2=1} E_{x_0,y_0}(u,v)$$

(5) In the mathematical formulation of cornerness, what does the expression $E(u,v)$ represent, and what role do u & v play?

↳ $E(u,v)$: quantifies how much intensity changes for a patch when shifted

↳ u, v : Represent the direction and magnitude (größe) of the shift

→ $E(u,v)$ small : flat region
$E(u,v)$ large in one direction : Edge
$E(u,v)$ large in all directions : Corner

(6) How do eigenvalues and eigenvectors relate to the identification of corners in an Image

- Eigenvalues:
  > Measure the strength of intensity variation.
  > Large eigenvalues in both directions (λ_1, λ_2) signify corners

- Eigenvectors:
  > Indicate the directions of intensity changes
  > Help determine the orientation of features like edges or corners

- Corners have strong variations in all directions, edges vary in one direction, and flat regions show no significant variation

7. Explain how the determinant and trace of a matrix A can be used to efficiently compute cornerness without directly calculating eigenvalues.

- $\det(A)$: Measures the product of eigenvalues $(\lambda_1 \cdot \lambda_2)$, indicating the "magnitude" of variation

- $\mathrm{trace}(A)$: Measures the sum of eigenvalues $(\lambda_1 + \lambda_2)$, summarizing total variation.

  trace = sum of diagonal entries of a Matrix

- The Harris Corner Detector uses $\det(A)$ und $\mathrm{trace}(A)$ to compute coneroness efficiently, without directly calculating eigenvalues. This makes it suitable for fast image processing tasks like realtime feature detection.

(8) Why is scale invariance a challenge in corner detection, and what stratagies can be employed to adress this issue?

Challenge: Eigenvalues are not stable under small changes in scale

Solutions:
- search over image pyramid scales
- Scale selection in 2D

9) What is the purpose of coarse-to-fine search selection for interest point detection?

It is expensive to compute (x,y, sigma) over all frectional scales and the purpose of coars-to-fine search is to optimize this.

1. Optimize cornerness (x, y, sigma) over coarse set of locations and scales
   grob

2. Fine-tune „sub-pixel" accuracy by iterating the following:
   i. Given (x,y), we can find maximal sigma with finer search
   ii. Given Sigma, we can find maximal (x,y) of cornerness

(10) Briefly outline the overall pipeline for finding correspondences between two images using intrest points.

- find N "easy-to-match" patches in image No.1 and M „easy-to-match" patches in image No.2

- Compute $O(NM)$ pairwise patch similarities and candidate correspondences
    - for each patch on image No.1, find the best match in image No.2

- Extract subset of correspondences that are consistent with a motion model

(11) Why are feature descriptors important in computer vision?

We use them to match good features and describe an image patch. Because patches with similar content should have similar descriptors.

(12) What is the main limitation of using raw pixel values as feature descriptors?

> They are very sensitive to intensity values, to be more clear sensitive to intensity changes.

- for example changes in the brightness of a picture can alter these values, making the descriptor unable to identify the matching patches

(13) How do image gradients help creating more robust feature descriptors?

They use pixel differences instead of the raw values, what makes these descriptors invariant to absolute intensity values (and it also provides directional and magnitude information)

(14) Explain the concept of binary descriptor and its advantages.

- A binary descriptor encodes image features as a compact binary string using pixel intensity comparisons

- Advantages: Compactness, low memory usage, fast matching and efficiency in real-time applications.

15 What is the primary advantage of using a color histogram as a feature descriptor?

They are invariant to changes in scale and rotation.

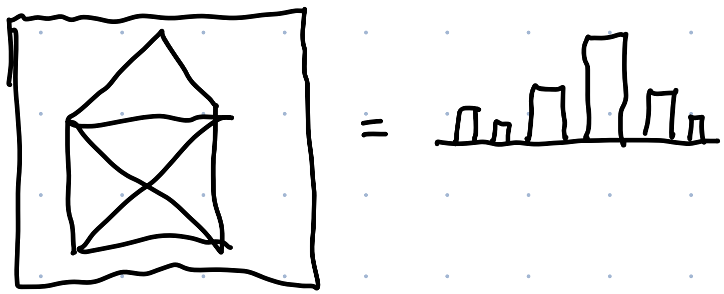
16 What is the key difference between a color histogram and a spatial histogram?

> A color histogram creates a single histogram over the entire patch

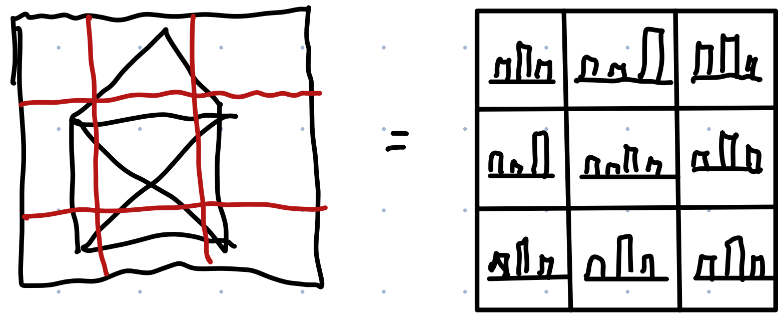
> A spatial histogram computes histograms over spatial „cells"

↳ difference: The spatial histogram preserves some spatial layout information

color histogram:



spatial histogram:



some spatial layout

(17) How does orientation normalization contribute to achieving rotation invariance?

It uses the dominant image gradient direction to normalize the orientation of the patch

> it also saves the orientation angle θ along with (x, y, s)

(18) What is the significance of designing feature descriptors that are invariant to photometric transformations?

> Designing feature descriptors that are invariant to photometric transformations ensures robustness to lighting and contrast changes

> They improve matching accuracy in real-world conditions

(19) Why is it challenging to design feature descriptors that are robust to geometric transformations?

> Designing descriptors robust to geometric transformations is challenging due to changes in pixel arrangement, scale, perspective

> objects will appear at different scales, translation and rotation often in complex ways

20 Describe a scenario where a combination of different feature descriptor types might be beneficial.

In Autonomous driving, the system must preceive and understand its surroundings using multiple types of features to ensure robustness, accuracy and adaptability in diverse scenarios.
A combination of feature descriptor types can adress the varying needs for object recognition, lane detection, and environment mapping in a unified pipeline.