Started on	Tuesday, 27 February 2024, 12:08 PM
State	Finished
Completed on	Tuesday, 27 February 2024, 12:32 PM
Time taken	23 mins 21 secs
Grade	**85.83** out of 100.00

Question 1

Correct

Mark 5.00 out of 5.00

The essential matrix is applied to the normalized image plane! Correct or not?

Select one:

- True ✓
- False

The correct answer is 'True'.

Question 2

Incorrect

Mark 0.00 out of 5.00

We said that for affine structure from motion and ***n*** point correspondences / ***m*** cameras we get ***2nm*** equations in ***8m+3n*** unknowns. Due to the ambiguity in ***Q***, we can add constraints: ***2nm>8m+3n-12***. This implies, that for reconstructing ***2*** views, at least ***4*** points are needed. Note, that points and cameras are reconstructed in affine space. For a transformation to Euclidean space, an additional Euclidean update is needed. We want to ignore this here.

The same idea can be applied to projective structure from motion: A perspective projection matrix is ***3x4*** (with lower right coefficient being ***1***). Thus, ***M*** has ***11*** unknowns. The projective transform ambiguity matrix ***Q*** is ***4x4*** (with the lower right coefficient being ***1***). Thus it has ***15*** constraints. As for the affine case explained above, we can infer from this that: ***2nm > 11m+3n-15***.

How many point correspondences are sufficient (minimum) to reconstruct two cameras and all the needed (minimum number of) points in projective space? Thus, if ***4*** points is the minimum for two cameras in the affine case, how many points are needed for two cameras in the projective case?

Please provide your answer in the following format: ***NM***, where ***N*** is the number of points and ***M*** is the number of views.

Answer: 24 ✗

The correct answer is: 72

Question 3

Correct

Mark 5.00 out of 5.00

Learning-based approaches that perform camera calibration based on "in-the-wild" captured images:

- a. Learn pixel-wise mapping using U-Nets. ✓
- b. Learn from geometric representations, such as vanishing points. ✓
- c. Regress intrinsic and extrinsic parameters using CNNs. ✓

Die Antwort ist richtig.

The correct answers are: Regress intrinsic and extrinsic parameters using CNNs., Learn pixel-wise mapping using U-Nets., Learn from geometric representations, such as vanishing points.

Question 4

Correct

Mark 5.00 out of 5.00

What is the diameter of a **80mm** focal length **f/4** lens? Your answer should be in **mm**.

Answer: 20 ✓

The correct answer is: 20

Question 5

Correct

Mark 5.00 out of 5.00

The complexity of IPC's (Iterative Closest Point Method) correspondence match, if no optimal data structures are used is ***O(N^2)*** (for every point the closest neighbor has to be found).
What is its complexity roughly if k-d trees are used?

Select one:

- a. ***O(N log N)*** ✓
- b. ***O(N^3)***
- c. ***O(N^2)***
- d. ***O(N)***

Die Antwort ist richtig.

The correct answer is: ***O(N log N)***

Question 6

Correct

Mark 5.00 out of 5.00

Given are two kernels ***H1*** and ***H2***, as well as an image ***I.*** Which of the following rules are correct (multiple choices possible).

Note: ***k*** indicates a scaling factor and * convolution.

Select one or more:

- a. ***H1*(H2*I)=(H1*H2)*I*** ✓
- b. ***I+(H1*H2)=I*(H1+H2)***
- c. ***H1*I+H2*I=(H1+H2)*I*** ✓
- d. ***H1*(H2*I)=H2*(H1*I)*** ✓
- e. ***(kH1)*I=k(H1*I)*** ✓

Die Antwort ist richtig.

The correct answers are: ***H1*I+H2*I=(H1+H2)*I***, ***(kH1)*I=k(H1*I)***, ***H1*(H2*I)=H2*(H1*I)***, ***H1*(H2*I)=(H1*H2)*I***

Question 7

Correct

Mark 5.00 out of 5.00

We said, that the transport of light is geometrically invertible (Helmholz Reciprocity). This is used for dual photography, where ***p*** is the image a projector projects onto a scene, ***c*** the image that a camera records, and ***T*** the light-transport matrix.

Let's assume that ***p*** is a uniform white image. We said that with ***c=Tp*** (forward light-transport), ***c*** shows the scene from the perspective of the camera under a white projector illumination. Assuming that ***c*** is a uniform white image, the result of ***p=T''c*** (where ***T''*** is the inverse of ***T***) implements what is called the inverse light transport. It computes an image that shows the scene from the perspective of the projector, but where the colors are inverted. Projecting these inverse colors compensates the actual modulation of light within the scene and the result is a uniform white image when observed from the perspective of the camera. The principle is used for radiometric compensation (i.e., for image correction of projectors that have to display images at textured surfaces).

However, what is the result of ***p=T'c*** (where ***T'*** is the transpose of ***T***), assuming that ***c*** is a uniform white image?

Select one:

- a. It computes an image that shows the scene from the perspective of the projector under a white illumination from the perspective of the camera. ✓
- b. In a uniform white image from the perspective of the projector.
- c. In an image that shows the scene from the perspective of the camera, but where the scene depth appears reversed.

Die Antwort ist richtig.

The correct answer is: It computes an image that shows the scene from the perspective of the projector under a white illumination from the perspective of the camera.

Question 8

Correct

Mark 5.00 out of 5.00

The filter kernel:

(-1, -1, -1,
-1, 17, -1,
-1, -1, -1)/9

is a ...

Select one:

- a. edge filter.
- b. smoothening filter.
- c. sharpening filter. ✓
- d. Gaussian filter.

Die Antwort ist richtig.

The kernel is a sharpening filter. The kernel sums to 1 and accentuates differences.

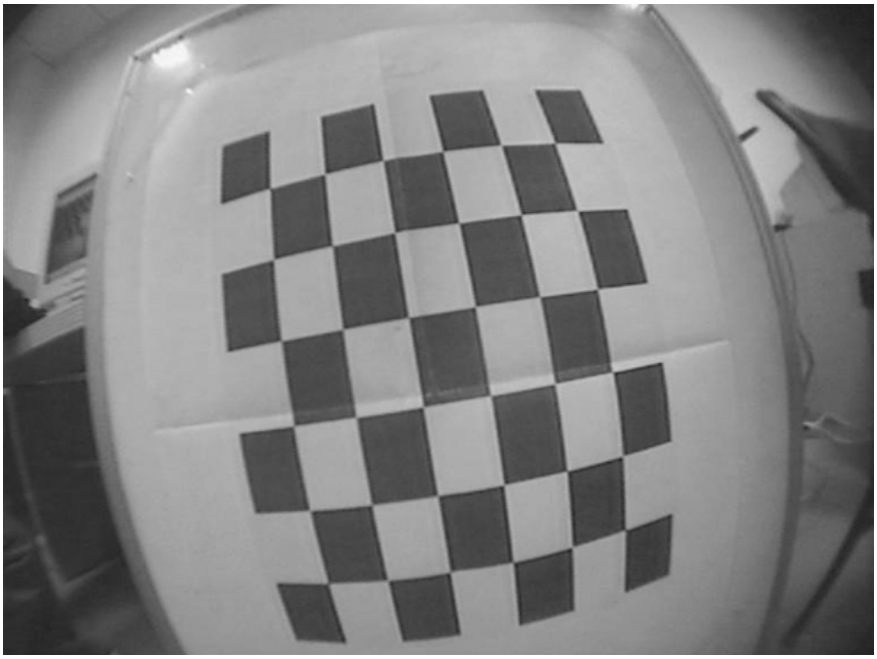
The correct answer is: sharpening filter.

Question 9

Correct

Mark 5.00 out of 5.00

The image below shows a radial distortion for which the first coefficient in the polynomial is positive. Correct or wrong?



Select one:

- True ✓
- False

The correct answer is 'True'.

Question 10

Correct

Mark 5.00 out of 5.00

For Vision Transformers....

- a. ...only the decoder part is usually used.
- b. ...linear projections of reshaped image patches + positional embeddings are usually used as input. ✓
- c. ...only the encoder part is usually used. ✓
- d. ...both, encoder and decoder are usually used.

Die Antwort ist richtig.

The correct answers are:

...only the encoder part is usually used.,

...linear projections of reshaped image patches + positional embeddings are usually used as input.

Question 11

Correct

Mark 5.00 out of 5.00

Multiplying the Intrinsic Matrix, K ...

a. ... maps coordinates from the physical image plane to the normalized image plane.

b. ... maps coordinates from the normalized image plane to the physical image plane. ✓

c. ... maps coordinates between the physical image planes of two cameras.

Die Antwort ist richtig.

The correct answer is:
... maps coordinates from the normalized image plane to the physical image plane.

Question 12

Partially correct

Mark 3.33 out of 5.00

Which of the following statements is correct (multiple choices possible)?

Select one or more:

a. The gradient field of an image has zero curl.

b. The gradient field of an image is a conservative vector field. ✓

c. After processing the gradient field, the corresponding image can always be reconstructed by integration.

d. The divergence of the gradient field is the Laplacian of the corresponding image. ✓

Die Antwort ist teilweise richtig.

You have correctly selected 2.
The correct answers are: The gradient field of an image has zero curl., The gradient field of an image is a conservative vector field., The divergence of the gradient field is the Laplacian of the corresponding image.

Question 13

Correct

Mark 5.00 out of 5.00

The color filter of a sensor array (i.e., in a camera) is arranged in such a way that 50% red, 25% green, and 25% blue sub-pixels are imaged. Correct or Wrong?

Select one:

True

False ✓

The correct answer is 'False'.

Question 14

Partially correct

Mark 2.50 out of 5.00

Cost-Volumes in PSMNet (and other models) are used for...

- a. ...extracting features from multiple views at multiple scales with spatial pyramid pooling (SPP).
- b. ...computing matching costs with a 3D CNN.
- c. ...concatenating features from multiple views at all possible disparities. ✓

Die Antwort ist teilweise richtig.

You have correctly selected 1.

The correct answers are: ...concatenating features from multiple views at all possible disparities., ...computing matching costs with a 3D CNN.

Question 15

Correct

Mark 5.00 out of 5.00

U-Nets are better suited for segmentation tasks than unconnected autoencoders because...

- a. ...of their additional pooling.
- b. ...of their additional convolutions.
- c. ...of their additional skip-connections. ✓
- d. ...of their additional transposed convolutions.

Die Antwort ist richtig.

The correct answer is:

...of their additional skip-connections.

Question 16

Incorrect

Mark 0.00 out of 5.00

The Hough space for finding lines in a 2D image is two-dimensional.
Which statement about the Hough transform is correct?

Select one:

- a. The Hough space for finding circles in a 2D image is two-dimensional.
- b. The Hough space for finding (infinitely large) planes in a 3D volume is two-dimensional. ✗
- c. The Hough transform can only be applied to find lines.
- d. The Hough space for finding ellipsoids (3D version of ellipses) in a 3D volume is five-dimensional.

Die Antwort ist falsch.

The correct answer is: The Hough space for finding ellipsoids (3D version of ellipses) in a 3D volume is five-dimensional.

Question 17

Correct

Mark 5.00 out of 5.00

Without considering the individual intrinsic and extrinsic camera parameters, you have learned that a 3D point with coordinate ***P=(x,y,z,1)*** can be mapped to its 2D image coordinate ***p=(u,v,1)*** with a 4x3 projection matrix ***M: p=1/z*M*P*** — or: ***u=(m1*P)/(m3*P) , v=(m2*P)/(m3*P)***.

Obviously, by knowing correspondences between enough ***(P,p)***-pairs (i.e., 3D points and their corresponding 2D image points on the sensor), we can solve this linear equation systems in the coefficients of ***M*** (e.g., using least-squares). How many corresponding pairs ***(P,p)*** are needed at least for a solution in ***M***?

Answer: 6 ✓

The correct answer is: 6

Question 18

Correct

Mark 5.00 out of 5.00

We cannot apply an order constraint when searching for matching features along scanlines (i.e., the feature order a,b,c,d,... along a scanline can not be expected to be the same in each image). Correct or wrong?

Select one:

- True ✓
- False

The correct answer is 'True'.

Question 19

Correct

Mark 5.00 out of 5.00

What is the Hamming distance of the Gray Code (often used for active range scanning / structured light), and why.

Select one:

- a. It is 2 for better compression.
- b. It is 2 to detect errors in transmission.
- c. It is 1 for faster transmission compared to ordinary binary codes.
- d. It is 1 to detect errors in transmission. ✓

Die Antwort ist richtig.

The correct answer is: It is 1 to detect errors in transmission.

Question 20

Correct

Mark 5.00 out of 5.00

How can neural networks be used for depth reconstruction?

- a. Fully connected layers can be used to learn the cost function. ✓
- b. Autoencoders (U-Nets) can be used for end-to-end learning. ✓
- c. U-Nets can be used to learn epipolar constraints.
- d. CNNs can be used for feature extraction. ✓

Die Antwort ist richtig.

The correct answers are:

CNNs can be used for feature extraction.,

Fully connected layers can be used to learn the cost function.,

Autoencoders (U-Nets) can be used for end-to-end learning.