RYERSON UNIVERSITY

Faculty of Engineering, Architecture and Science

Department of Electrical and Computer Engineering

Course Number	CPS 843
Course Title	Introduction to Computer Vision
Semester/Year	F2023

Instructor	Dr. Guanghui Richard Wang

ASSIGNMENT No.	4

Assignment Title	Homework 4

Submission Date	November 19th, 2023
Due Date	November 20th, 2023

Student Name	Abdulrehman Khan
Student ID	500968727
Signature*	A.K.

**By signing above you attest that you have contributed to this written lab report and confirm that all work you have swung the lab contributed to this lab report is your own work.*

Part 1:

Problem 1:

(1), **(2)**, and **(3)**:

Abdulrehman Khan
S009687737

Problem 1. (1): The canonical form of the plane at infinity is $\pi_\infty = (0, 0, 0, 1)^T$

Problem 1. (2): Proof by constuction: if the plane at infinity is a fixed plane (P), then it is under a 3D affine transformation (Q)

i) Assuming that the plane at infinity is a fixed plane under the projective transformation H, if and only if, H is an affinity.

ii) For a given H, H is proved to be an affinity when it satisfies certain properties. Therefore, 3D affine transformations can be considered as H, since they have the same properties in terms of preserving points and straight lines in space (translation, rotation, scaling, and shearing).

Abdulrehman Khan
500968727

Problem 1. (3): If $X' = HX$, then a 3D plane can be transformed as $\pi' = H^{-T}\pi$

Three planes define a point: $\begin{bmatrix} \pi_1^T \\ \pi_2^T \\ \pi_3^T \end{bmatrix} X = 0$

where, $Ax = 0$. This point is obtained as the 1D right null-space of the 3×4 matrix. The projective transformation begins with $X' = HX$ where it signifies the points that are under transformation. Then the plane proceeds to transform as $\pi' = H^{-T}\pi$.

Problem 2:

(1), **(2)**, and **(3)**:

[redacted]

Abdulrehman Khan
500968727

Problem 2. (1): The finite projective camera with matrix is:
$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$
with the matrix $P = KR[I \mid -\tilde{C}]$,

a finite projective camera has a camera centre with coordinates $(\tilde{C}^T, 1)^T$ that has 11 degrees of freedom. The set of camera matrices of finite projective cameras is the same as the set of homogeneous 3×4 matrices. Where, the left hand 3×3 submatrix is non-singular.

We can decompose this matrix into the general form using the PQ matrix decomposition

i) $P = K[R \mid t]$ ii) $P = [M \mid P_4]$

iii) $P = M[I \mid M^{-1} P_4] = KR[I \mid -\tilde{C}]$

iv) $P = [M \mid P_4]$ v.) Decomposing to the

final: $M = KR$

Where, R is an orthogonal matrix, C (tilde) is a camera centre with coordinates in a world coordinate system, K is a camera with a calibration matrix (finite projective camera) for this matrix:

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$ These are the focal length in pixels for two directions $\alpha_x = f m_x$ and $\alpha_y = f m_y$ and s represents a skew parameter.

Problem 2. (2): Given a projection matrix P, the value K represents the internal camera parameters R and t represent the external camera parameters: $P = K[R|t]$ $t = -R\tilde{C}$, where this camera projection matrix is 9 degree of freedom (3 for K, 3 for R, and 3 for t (or C)).

In order to recover the internal parameters, we must use RQ decomposition: $M = KR$ and then reverse the QR decomposition

For example, for the following given matrix P:

$$\mathtt{P} = \begin{bmatrix} 3.53553\,e{+}2 & 3.39645\,e{+}2 & 2.77744\,e{+}2 & -1.44946\,e{+}6 \\ -1.03528\,e{+}2 & 2.33212\,e{+}1 & 4.59607\,e{+}2 & -6.32525\,e{+}5 \\ 7.07107\,e{-}1 & -3.53553\,e{-}1 & 6.12372\,e{-}1 & -9.18559\,e{+}2 \end{bmatrix}$$

We will calculate the value of the center (C tilde) which comes out to be:

$$\widetilde{\mathbf{C}} = -\mathtt{M}^{-1}\mathbf{p}_4.$$

$$\widetilde{\mathbf{C}} = (1000.0, 2000.0, 1500.0)^{\mathsf{T}}$$

And using the value to conduct the RQ decomposition:

$$\mathtt{M} = \mathtt{KR} = \begin{bmatrix} 468.2 & 91.2 & 300.0 \\ & 427.2 & 200.0 \\ & & 1.0 \end{bmatrix} \begin{bmatrix} 0.41380 & 0.90915 & 0.04708 \\ -0.57338 & 0.22011 & 0.78917 \\ 0.70711 & -0.35355 & 0.61237 \end{bmatrix}$$

Abdulrehman Khan
500968727

Problem 2 (3): In forward projection, the space point is mapped to an image point, or otherwise known as the Principal point offset: $x = PX$. Where, x is the homogeneous image point, P is the camera projection matrix, and X is the homogeneous coordinate of the world point.

Let us assume that the image of a point at infinity is only affected by the sub matrix M. The mapping point, $D = (d^T, 0)^T$ can be substituted in the formula defined above, transforming to $x = PD$, now, P is substituted by the properties of a projective camera, $[M \mid P_4]$. Where, P_4 is the image of the coordinate origin. This concludes to: $X = PD = [M \mid P_4] D = Md$

Problem 3:

(1), **(2)**, and **(3)**:

Abdulrehman
Khan
500968727

Problem 3. (1): To locate the optical centre using projection matrix P; first use singular value decomposition: $PC = 0$

Then using direct computation:

$$x = \det([P_2, P_3, P_4])$$
$$y = -\det([P_1, P_3, P_4])$$
$$z = -\det([P_1, P_2, P_4])$$
$$t = -\det([P_1, P_2, P_3])$$

Problem 3. (2): The first three columns of the projection matrix correspond to the vanishing points of the X, Y, and Z axes of the world system. This is so as the values multiplied by the respective plane at infinity determine the X, Y, and/or Z values.

$$S_x \tilde{V}_x = P\tilde{x}_w = [P_1, P_2, P_3, P_4][1,0,0,0]^T = P_1,$$
$$S_y \tilde{V}_y = P\tilde{y}_w = [P_1, P_2, P_3, P_4][0,1,0,0]^T = P_2,$$
$$S_z \tilde{V}_z = P\tilde{z}_w = [P_1, P_2, P_3, P_4][0,0,1,0]^T = P_3,$$
$$S_o \tilde{V}_o = P\tilde{O}_w = [P_1, P_2, P_3, P_4][0,0,0,1]^T = P_4$$

Problem 3. (3): The last row of the projection matrix corresponds to the principal plane of the camera as this row represents a perpendicular vector to the axis of the lens.

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} = \begin{bmatrix} P^{1T} \\ P^{2T} \\ P^{3T} \end{bmatrix}$$

The row vectors may be interpreted geometrically as particular world planes principal plane.

<u>Problem 4:</u>

(1), **(2)**, and **(3)**:

Abdul rehman Khan
500968727

Problem 4 (1): We can verify that coplanar 3D points and their images are related by a 2D homography given by $H = K[r_1, r_2, t]$ because $x = PX = [p_1\ p_2\ p_3\ p_4](X\ Y\ 0\ 1)^T = [p_1\ p_2\ p_3\ p_4](X\ Y\ 1)^T$ and $P = K[R|T]$,

Problem 4 (2): We can verify that the back-projection of an image line is a 3D plane given by $\pi = P^T l$ because it is the plane that the camera matrix uses to map a set of points in space to a line through $\pi^T X = l^T PX$.

Problem 4(3): We can verify that the images captured by a zooming camera are related by a 2D homography given by $H = K'K^{-1}$ because when you are zooming you multiply $x' = K'[I|0]X = K'K^{-1}x = Hx$

Problem 5:

(1), **(2)**, **(3)**, and **(4)**:

Abdul rehman khan
500968727

Problem 5 (1): We can verify the image of the absolute conic is: $\omega = (KK^T)^{-1}$. Since the IAC depends only on the internal parameters, we calculate the angle between two rays which is given by

$$\cos \theta = \frac{x_1^T \omega x_2}{\sqrt{x_1^T \omega x_1} \sqrt{x_2^T \omega x_2}}$$

Problem 5(2): The two image points correspond to the orthogonal directions: $x_1^T \omega x_2 = 0$. This results in the dual image of the absolute conic $\omega^* = \omega^{-1} = KK^T$. The image of the absolute conic is finally obtained by performing the Cholesky decomposition from the DIAC $\omega^* = KK^T$

Problem 5(3): We can verify that two constraints on the image of the absolute conic can be obtained from homography 3D plan and image $H = [h_1, h_2, h_3]$. This metric plane is imaged with the known homography H as mentioned. where, $x = Hx$ $x' = H'x'$ which transforms to $I = Fx$ then to $F = H^{-T} F H^{-1}$

Problem 5 (4): We can show the constraints on the image of the absolute conic obtained as the square pixels accumulate to $\omega 12 = \omega 21 = 0$, $\omega 11 = \omega 22$
$\omega = [(w_1 \; 0 \; w_2)], (0 \; w1 \; w3), [(w_2 \; w_3 \; w_4)]$ where $\omega = (KK^T)^{-1}$

Part 2:

Technical Overview of SIFT Feature Matching and Image Stitching

SIFT Feature Matching:
SIFT (Scale-Invariant Feature Transform) is a computer vision algorithm used for extracting distinctive features from images. It identifies keypoints, which are areas in an image that are invariant to scale, rotation, and illumination changes. These keypoints are characterized by their scale, orientation, and descriptors. SIFT utilizes a series of operations, including Gaussian blurring, gradient computation, and histogram generation, to create a robust feature representation.

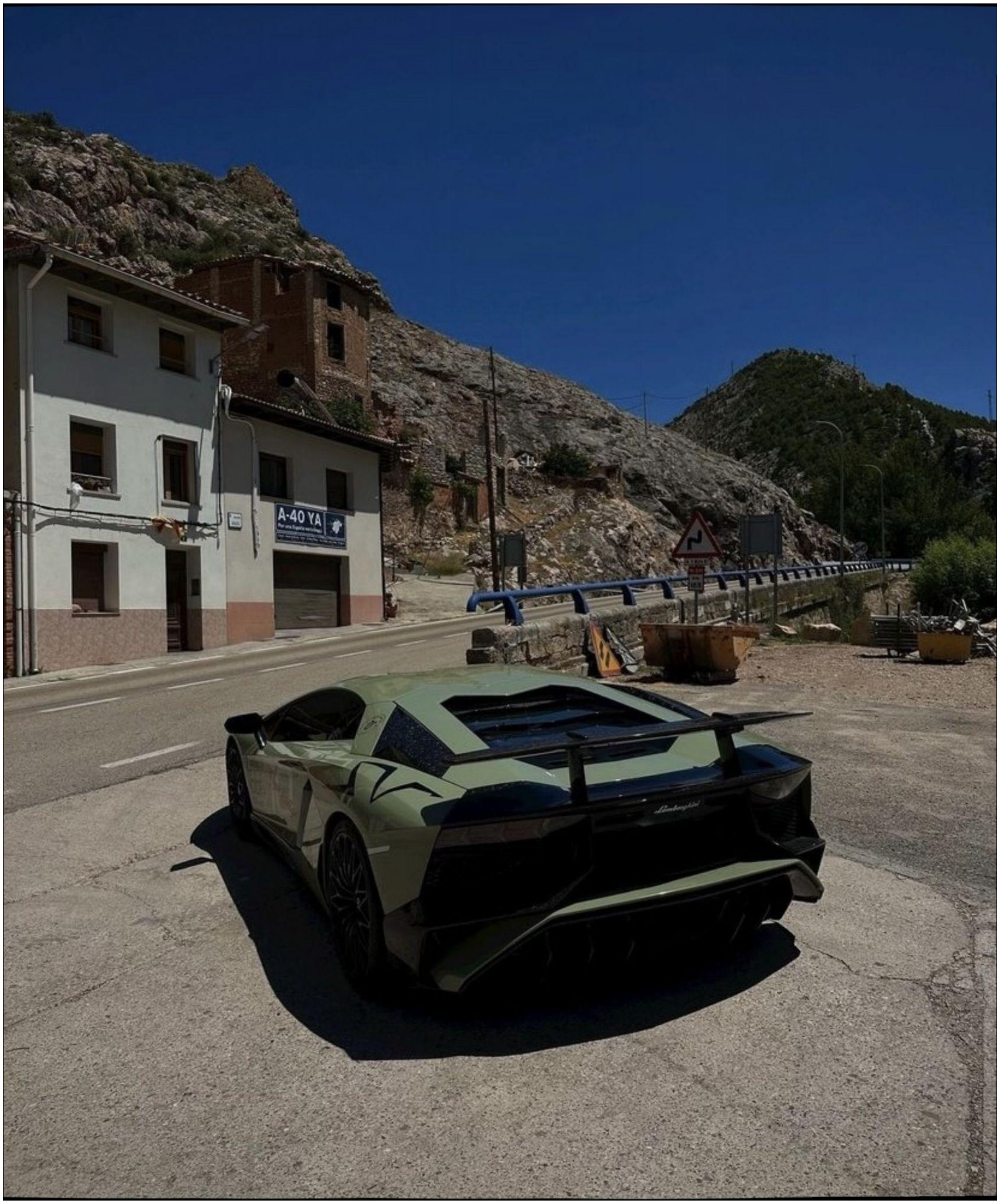
In the context of AutoStitch, the SIFT algorithm is employed to automatically find matching keypoints between unordered collections of images. By identifying common features across images, AutoStitch establishes correspondences, allowing it to align the images accurately.

Image Stitching:
The process of image stitching involves combining multiple images to create a panoramic view. AutoStitch accomplishes this by first identifying matching keypoints using SIFT. Once correspondences are established, the software robustly aligns the images. Advanced blending algorithms are then applied to seamlessly merge the images, resulting in a composite panorama.

Sample Original Images and Stitching Result:

Below are six sample original images and the stitching result:

Stitching Result: [Panorama Image]:



Analysis and Discussion:

The image stitching process aims to seamlessly merge images with overlapping content, creating a panoramic view. The success of this process depends on the accuracy of feature matching and alignment.

Observations:
1. The SIFT algorithm plays a crucial role in identifying key features across images, enabling AutoStitch to find correspondences accurately.
2. The stitching result demonstrates the effectiveness of AutoStitch in aligning and blending images, producing a visually coherent panorama.
3. The software's ability to handle images with significant overlap contributes to the overall quality of the panorama.

Challenges:
1. The success of AutoStitch may be influenced by the quality and characteristics of the input images. Poorly captured images or those with insufficient overlap might lead to stitching artifacts.
2. The demo version's limitation to spherical projection may affect the stitching of certain scenes, particularly those requiring cylindrical or planar projections.

Conclusion:
AutoStitch offers a user-friendly approach to panoramic image stitching, automating the process through advanced algorithms like SIFT. The stitching result demonstrates the capability of the software to handle image collections with sufficient overlap, providing a valuable tool for creating panoramic views effortlessly.

References:

AutoStitch. (n.d.). AutoStitch: a new dimension in automatic image stitching. Matthew Alun Brown. Retrieved from http://matthewalunbrown.com/autostitch/autostitch.html

Lowe, D. G. (n.d.). Distinctive image features from scale-invariant keypoints. University of British Columbia. Retrieved from https://www.cs.ubc.ca/~lowe/keypoints/

Toronto Metropolitan University. 2023. Course Content: Week 8. In Computer Vision, CPS 843. Toronto Metropolitan University's Learning Management System. https://courses.torontomu.ca/d2l/le/content/797209/viewContent/5397373/View

Toronto Metropolitan University. 2023. Course Content: Week 9. In Computer Vision, CPS 843. Toronto Metropolitan University's Learning Management System. https://courses.torontomu.ca/d2l/le/content/797209/viewContent/5403297/View

Toronto Metropolitan University. 2023. Course Content: Week 10. In Computer Vision, CPS 843. Toronto Metropolitan University's Learning Management System. https://courses.torontomu.ca/d2l/le/content/797209/viewContent/5408223/View

Wu, P. (2016, Month Day). How to Write a Panorama Stitcher. ppwwyyxx Blog. Retrieved from https://ppwwyyxx.com/blog/2016/How-to-Write-a-Panorama-Stitcher/