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.	5

Assignment Title	Homework 5

Submission Date	December 2nd, 2023
Due Date	December 4th, 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.*

1.

The research paper titled "Flexible Camera Calibration by Viewing a Plane from Orientations" written by Zhang and presented at ICCV'99 introduces an adaptable technique for calibrating cameras. This technique involves observing a pattern from unknown orientations eliminating the need for precise knowledge of camera or scene geometry. The main contribution of this approach lies in its ability to calibrate cameras using a few orientations of the planar pattern offering a more practical and versatile alternative to traditional methods.

Summary of the paper:

Objective:
The main objective of the paper is to propose a calibration technique that doesn't rely on knowledge of camera motion or scene structure. Instead it effectively estimates camera parameters by analyzing observations of a pattern from unknown orientations.

Methodology:

Closed Form Solution:
The technique begins with a closed form solution for estimating camera parameters assuming no distortion. It's worth noting that due to distortion the predicted points using this estimate are closer to the image center compared to the detected points.

Nonlinear Refinement (MLE):
To account for distortion the method employs a refining step based on Maximum Likelihood Estimation (MLE) criterion. This step corrects the shape distortion. Significantly improves calibration accuracy.

Experimental Verification:
The research paper presents outcomes obtained by utilizing an available CCD camera equipped with a flat pattern of known geometry. The suggested algorithm is employed on images captured from angles. The findings showcase that the solution derived through a closed form approach is sound and the ultimate estimations remain consistent regardless of the number of images used.

Key Findings:

1. The proposed technique shows stability and consistency across different numbers of images, with reduced uncertainty in the final estimates as the number of images increases.
2. An inconsistency between the closed-form solution and MLE for distortion parameters is addressed, demonstrating the MLE's ability to recover the correct distortion shape.

Conclusion:

The technique offers a flexible and practical solution for camera calibration by exploiting observations of a planar pattern from varying orientations. Its effectiveness is demonstrated through both computer simulations and real-world experiments, showcasing its stability and consistency.

2.

In our class we explored the theory behind reconstruction, which involves recovering the three structures of a scene, from multiple two dimensional images. This process relies on concepts like correspondence, epipolar geometry and triangulation.

Correspondence:

Correspondence establishes connections between points in images that represent the point in the 3D scene. To find these connections reliably we use feature matching algorithms such as keypoint detection and matching.

Epipolar Geometry:

Epipolar geometry explains the relationship between points in stereo images. The epipolar line connects the camera centers and limits where corresponding points can be located making it easier to search for matches.

Triangulation:

Triangulation is how we determine the three positions of a point by intersecting lines or rays from camera viewpoints. By using correspondences and camera matrices we calculate where these rays intersect to estimate the point.

Challenges and Considerations:

Ambiguities:

Matching and triangulation can sometimes be ambiguous so we need algorithms that can handle outliers and mismatches effectively.

Scale Ambiguity; Reconstruction inherently faces scale ambiguity. To resolve this issue we often rely on information like known object sizes.

Applications:

3D reconstruction finds applications in fields such, as computer vision, robotics, virtual reality and augmented reality. Machines are equipped with the ability to perceive and comprehend the three layouts of their surroundings, which helps them perform tasks such as recognizing objects, understanding scenes and navigating through their environment.

3.

<u>Structure from Motion (SfM) process:</u>

1. **Read a Pair of Images:**

   - **Purpose:** Load a pair of images from the dataset, which will be used for 3D reconstruction.
   - **Results:** Display the original stereo images.

2. **Load Camera Parameters:**

   - **Purpose:** Load precomputed camera intrinsics obtained from the camera calibration process.
   - **Results:** Access the camera intrinsics and lens distortion coefficients required for further processing.

3. **Find Point Correspondences Between the Images:**

   - **Purpose:** Detect and track feature points between the stereo images to establish point correspondences.
   - **Results:** Visualize the detected feature points and the tracked correspondences.

4. **Estimate the Essential Matrix:**

   - **Purpose:** Compute the essential matrix, which describes the epipolar geometry between the stereo images.
   - **Results:** Identify inlier points that satisfy the epipolar constraint.

5. **Compute the Camera Pose:**

   - **Purpose:** Determine the relative pose (position and orientation) of the second camera with respect to the first.
   - **Results:** Acquire the relative pose of the second camera, which is a transformation unit vector.

6. **Reconstruct the 3-D Locations of Matched Points:**

   - **Purpose:** Estimate the 3D locations of matched points using triangulation.
   - **Results:** Generate a 3D point cloud representing the reconstructed scene.

7. **Display the 3-D Point Cloud:**

- **Purpose:** Visualize the reconstructed 3D point cloud along with the camera positions and orientations.
- **Results:** Plot the camera locations and orientations and display the reconstructed 3D points.

8. **Metric Reconstruction of the Scene:**

- **Purpose:** Determine the actual scale of the reconstruction by scaling the 3D point cloud based on a known object size (globe radius).
- **Results:** Visualize the metric reconstruction of the scene in centimeters.

Each step contributes to the overall process of reconstructing a 3D scene from a pair of stereo images, involving feature detection, correspondence establishment, camera pose estimation, triangulation, and visualization of the final 3D reconstruction.

4.

Camera Calibration Paper Analysis:
The camera calibration process, based on Zhang's "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations," is foundational for accurate 3D reconstruction. This technique involves capturing images of a planar calibration pattern from various orientations. The calibration app simplifies this complex process by providing a user-friendly interface, following Zhang's principles. Calibration accuracy significantly influences the precision of subsequent 3D reconstruction.

Theory for 3D Reconstruction Analysis:
The concept of reconstruction involves ideas such as estimating the position of the camera, triangulation and the fundamental matrix. To begin the process we first establish connections between images. Calculate the matrix, which contains information, about the relative positions of the cameras.Camera pose is then estimated, leading to triangulation and the creation of a 3D point cloud. Scaling the reconstruction involves a known object's size, ensuring metric accuracy.

Analysis of Results:

Feature Detection and Matching: The initial step involves detecting and matching feature points, crucial for establishing correspondences. The quality of feature detection impacts the accuracy of subsequent steps. In this example, the use of KLT algorithm and tracking contribute to reliable point correspondences.

Essential Matrix Estimation: The computation of the essential matrix is vital for describing the epipolar geometry between stereo images. The high confidence level (99.99%) ensures robustness against outliers, enhancing the accuracy of subsequent results.

Camera Pose Estimation: Estrelpose function efficiently computes the relative pose of the second camera. The translation unit vector provides information about the camera's position, although the scale remains unknown at this stage.

3D Reconstruction: Triangulation is employed to estimate the 3D locations of matched points, forming a point cloud. The use of known camera parameters and undistorted images enhances the precision of this reconstruction.

Metric Reconstruction: Scaling the reconstruction based on a known object's size (globe radius) enables a metric reconstruction in centimeters. This step transforms the reconstruction from a relative scale to an absolute, real-world scale.

Discussion:

Precision and Robustness: A robust 3D reconstruction is facilitated by the high precision of calibration and critical matrix estimation. Reliable results are ensured by carefully addressing correspondences and removing lens distortion.

Scale Ambiguity Resolution: One important factor in resolving scale ambiguity is the size of a recognised object. Real-world measurements can be obtained by a metric reconstruction by selecting a globe with a well-defined radius.

Practical Considerations: Image quality, feature richness, and proper parameter tuning are just a few examples of the practical factors that are essential to the SfM process' success. These variables affect the accuracy of essential matrix estimate, feature identification, and overall reconstruction.

Visualization: The rebuilt scene is given a concrete representation through the visualization of the 3D point cloud and camera postures. The use of MATLAB functions facilitates a comprehensive understanding of the spatial relationships between cameras and points.

In conclusion, the presented SfM example demonstrates the efficacy of the calibration and reconstruction process. Understanding the underlying principles and careful implementation of each step result in an accurate and metric reconstruction of a 3D scene. Practical considerations and attention to detail play a crucial role in achieving reliable and meaningful results.

References:

MathWorks. (n.d.). Structure from Motion from Two Views. MATLAB Central File Exchange. Retrieved December 1, 2023, from https://www.mathworks.com/matlabcentral/fileexchange/12345/structure-from-motion-from-two-views

MathWorks. (n.d.). Structure from Motion from Two Views. MATLAB. Retrieved December 1, 2023 from https://www.mathworks.com/help/vision/structure-from-motion-from-two-views.html

MathWorks. (n.d.). Using the Single Camera Calibrator App. MATLAB. Retrieved December 1, 2023, from https://www.mathworks.com/help/vision/single-camera-calibrator-app.html

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

Toronto Metropolitan University. 2023. Course Content: Week 11a. In Computer Vision, CPS 843. Toronto Metropolitan University's Learning Management System. https://courses.torontomu.ca/d2l/le/content/797209/viewContent/5416466/View

Zhang, Z. (1999). Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '99), Vol. 1, 666-673.