Computer Vision | Homework 5
Research Paper Review

Ruchi Manikrao Dhore - W1652116

Friday 17^{th} February, 2023

1 Introduction

As part of this homework, we would study and review the paper *"Automatic Panoramic Image Stitching using Invariant Features"*. The authors of the research paper are *Brown, M., Lowe, D.G.*. Published by *Springer Science* in the year 2007. We will provide an overview of the research area, the main research question, and the significance of the research. The key findings and contributions of the research papers will be reviewed.

2 Rationale For Research

Image stitching in two dimensions or with multiple rows is considered to be more complex. There was a need to devise automatic stitching as a multi-image matching problem and employ invariant local features to compare all of the photos and identify matches. The solution must be indifferent to the input images' scale, lighting, orientation, ordering, and noise levels. Using direct or featured-based approaches for **digital cameras**, the solution shall recognize numerous panoramas in an unordered image dataset.

3 Major Objective Of Research

To find a solution that will allow cameras to assemble panoramic images completely automatically.

4 Literature Survey

4.1 Group I

The scientific literature on panoramic image stitching is rather broad. It was majorly researched by Szeliski, 2004; Milgram, 1975; Brownand Lowe, 2003, Chen, 1995; Realviz, Hartley and Zisserman, 2004; Szeliski and Shum, 1997.

4.1.1 Strength

The main strength was that the problem's fundamental geometry was clearly grasped. It involved estimating a 3x3 camera matrix or homography for each image, which usually requires user input to align the images or a fixed image ordering.

4.1.2 Future Scope

The future scope is to improve human intervention in the registration process and image order may not be static.

4.2 Group II

In the academic literature, there are two types of methods for automatically aligning and stitching images: direct and feature-based methods. Direct approaches require a close initialization but utilise all available image data and offer very accurate registration. While traditional feature matching approaches lack the invariance properties necessary for accurate matching of arbitrary panoramic image sequences, feature-based methods do not https://www.overleaf.com/project/63f019337d06da354aa81df8need setup. It was majorly researched by Szeliski and Kang, 1995; Irani and Anandan, 1999; Sawhney and Kumar, 1999; Shum and Szeliski, 2000, Zoghlami et al., 1997; Capel and Zisserman, 1998; McLauchlan and Jaenicke, 2002.

4.2.1 Strength

The main strength was the automated registration as well as automatic image alignment and stitching.

4.2.2 Future Scope

The future scope is to the lack of invariance features to enable reliable matching of arbitrary panoramic image sequences.

5 Author's Approach

This paper presented a new approach to fully automated panoramic image stitching using invariant features. This approach offered several advantages over previous methods, including reliable matching despite changes in rotation, zoom, and illumination of the input images, automatic discovery of matching relationships, and the ability to recognize panoramas in unordered datasets. Additionally, the paper introduced gain compensation and automatic straightening steps, and an efficient bundle adjustment implementation to optimize camera parameters. The multi-band blending technique is used to produce seamless output panoramas of high quality.

5.1 Advantages of approach

- Despite input image rotation, zooming, and illumination changes, successful matching of panoramic image sequences was made possible by the application of invariant characteristics
- Considering image stitching as a multi-image matching issue, which can automatically identify panoramas in unordered datasets and determined the matching associations between the photos
- Used multi-band blending to produce high-quality results

6 Key Steps from Author's Approach

The fundamental stages of the author's methodology.

1. The development of the geometry of the problem which motivates choice of invariant features
2. A probabilistic model for image match verification and the image matching technique RANSAC
3. Algorithm for image alignment (bundle adjustment) that jointly optimizes the variables for each camera
4. Automatic straightening, gain correction, and multi-band blending are all included in the rendering pipeline

7 Mathematics Used

7.1 Feature Matching [SIFT]

The extraction and matching of Scale-Invariant Feature Transform (SIFT) features across all of the images is the initial stage in the panoramic recognition method. Despite input image rotation, zooming, and illumination changes, successful matching of panoramic image sequences is made possible by the application of invariant characteristics. Our approach can handle images with shifting orientation and zoom since SIFT features are invariant under rotation and scale changes. Furthermore noted is the partial invariance of SIFT features to affine transformation modifications. For image matching after feature extraction, they applied KNN.

Features must be matched after being extracted in linear time from all n photos. Each feature is matched to its k nearest neighbors in feature space (we assume k = 4) because numerous ages may overlap a single ray. Finding the approximate nearest neighbors using a k-d tree can be done in O(n log n) time (Beis and Lowe, 1997). An axis aligned binary space partition known as a k-d tree recursively divides the feature space into the largest variance dimensions at the mean.

7.2 Image Matching [RANSAC]

To estimate the parameters of a mathematical model from a set of observed data that contains, for every Image Matching [RANSAC - Random Sample Consensus] is an iterative procedure.

It described the process of finding matching images for panoramic stitching. It explains that only a small number of overlapping images need to be matched to get a good solution for the image geometry, and a constant number of m images, with the greatest number of feature matches to the current image, are considered as potential image matches. The RANSAC algorithm is used to select a set of inliers that are compatible with a homography between the images, followed by applying a probabilistic model to verify the match.

Two steps are involved as part of this:

1. Robust Homography Estimation using RANSAC: A set of inliers consistent are choosen with a homography between the pictures using RANSAC.

2. Probabilistic Model for Image Match Verification: A probabilistic model is used to confirm the match.

For Image Verification [Probability Model], they utilized binominal probability distribution to find inliers and outliers.

7.3 Bundle Adjustment [Levenberg-Marquardt Algorithm]

The bundle adjustment is used to solve for all camera parameters jointly. Concatenation of pairwise homographies would result in errors and ignore multiple constraints between images. The new image is added to the adjuster with the same rotation and focal length as the image to which it best matches. Levenberg-Marquardt is used to update the parameters. A robustified sum squared projection error is used as the objective function to project each feature into all the matching images and minimize the squared image distances. It's also possible to estimate the unknown ray directions jointly with the camera parameters.

Concatenation of pairwise homographies would result in cumulative errors and disregard various restrictions between pictures, therefore this step is crucial. The Levenberg-Marquardt algorithm was used to tackle this non-linear least squares issue.

7.4 Automatic Panorama Straightening [Covariance Matrix]

The rationale is that since it is uncommon for humans to turn the camera in relation to the horizon, the horizontal axis of the camera usually lies in a plane. We may get the **up-vector** by locating the null vector of the covariance matrix of the camera X vectors.

7.5 Gain Compensation

The procedure of calculating the camera's photometric parameter—the overall gain between images—is was outlined. Similar to how the geometric camera parameters are calculated, it uses an error function that is determined over all images. The total gain normalized intensity errors for all overlapping pixels make up the error function.

7.6 Multi-Band Blending

Notwithstanding each sample's compensation for gain along a ray in each intersecting image, a sound blending technique was covered. Image edges may be seen due to unmodeled phenomena like radial distortion, parallax effects, vignetting, and mis-registration mistakes. Multiband bending algorithm was applied for averaging the intensities.

8 Strengths Observed in this Approach

Several panoramas may be recognized in unordered image sets, and they can be automatically stitched together without the need for user input thanks to the use of invariant local features and a probabilistic model to check image matches.

The technology is resistant to camera zoom, image orientation changes brought on by flash and exposure/aperture settings, and variations in illumination. Despite variations in illumination, a multi-band mixing approach allows seamless transitions between images while keeping high frequency features.

9 Experimentation

For experimentation purpose, the dataset contained two different set of images. In first input, four panoramas along with four noise images made up the input image collection. The second input was a larger image sequence captured using a camera. The camera's automated mode was used to take these pictures, which gave the aperture and exposure time considerable leeway as well as letting the flash fire on some of them. The tests were conducted on a 1.6 GHz Pentium M.

10 Results

The system produced 4 blended panoramas after identifying related elements in photos that were matched and unmatched. In a larger example shown in Figure 5, the algorithm matched and registered all 57 input images automatically, and solved a 4 x 57 parameter optimization problem for the final registration. The output was a seamless 23 megapixel panorama rendered in 15 minutes, and a 2000 x 573 preview was rendered in 57 seconds.

11 Conclusion

The paper presented an autonomous panorama stitching system that recognizes and stitches numerous panoramas in unordered image sets without any user input by utilizing local characteristics and a probabilistic model. The technology is resistant to variations in lighting brought on by flash, exposure/aperture settings, camera zoom, and orientation. Despite variations in illumination, a multi-band mixing method allows seamless transitions between images while preserving high-frequency features.

Four possible future scopes have been suggested:

1. Camera motion: There is a problem with panoramic photos because of the camera movement. It indicates that slight camera movement causes motion blur, which can be reduced by using photographs that are close by and in focus. In order to eliminate parallax errors brought on by minute movements of the optical center, it is proposed to resolve camera translations and scene depths.

2. Scene motion: There is an issue with the multi-band blending strategy when dealing with large object motions in the scene. This causes visible artifacts when blending multiple images. The solution proposed is to use an alternative approach that involves finding optimal seam lines based on differences between images.

3. Advanced camera modelling: The significance of taking radial distortion into account in high-quality image stitching applications is underlined. The trials have demonstrated that the system can recognize panoramas and approximate alignment even under mild degrees of radial distortion, despite not being explicitly described by the algorithm.

4. Photometric modelling: The possibility of predicting camera photometric factors like vignetting, a common source of artifacts, especially in areas of uniform color, is discussed.