Application of Local Beam Search with Scale Invariant Feature Transform and Random Sample Consensus to Generate Panoramic Image

Shubhankit Saxena
Department of Computer Science and Engineering
National Institute Of Technology, Agartala, India

Pranjal
Department of Computer Science and Engineering
National Institute Of Technology, Agartala, India

Vivek Patel
Department of Computer Science and Engineering
National Institute Of Technology, Agartala, India

Rakesh Kumar Jaiswal
Department of Computer Science and Engineering
National Institute Of Technology, Agartala, India

Abstract- This Paper concern with the implementation of local beam search in fully automatic image stitching. Several previous approaches have been used for panoramic image stitching, in this work we implement local beam search in addition to the several previous algorithm due to which our feature extraction and matching is much faster.

Keywords- Panorama, SIFT, RANSAC, Homograph, Local Beam Search, Image Stitching

I. INTRODUCTION

In today's digital world there has been an exponential growth in panoramic image generation. Panorama is a Greek word which means "all sight". Panorama is a physical broad view representation of any space. Panoramic image generation is being widely used in paintings, photography, seismic images and also in film industries. For example, many tourist websites provides local street views & Bing maps provides a visualization as if the user is present in a moving vehicle. For desired panoramic generation many orderly snapshots covering the whole area is required. Like all other processes panoramic image generation is also having some phases for its completion, those are image acquisition, image registration and image blending.
The previous work done on panoramic view generation [1] is based on the fully automatic panoramic image stitching. Invariant features makes reliable matching of images under consideration. High quality results are obtained using multi-band blending and automatic discovery of matching relationships between the images.

In our work what we will be examining is given below:
1. Can Local beam search be introduced in panoramic image generation?
2. If RQ 1 is true then how much effective is it?

II. LITERATURE REVIEW

Significantly we implement Local beam search (LBS) on the blob points obtained in SIFT [3] in order to find points which have high describing strength of detected feature.

Local beam search

- Initialization begins with x random states.
- At every step, successor of all x states are produced.
- If any one of them is our goal then it will halt.
- Else x best successors are selected and process is repeated.

In accordance with our base paper [1], SIFT (Scale invariant feature transform) is used to find the blob points located at scale space maxima/minima of a difference of Gaussian function. At each feature location a characteristic scale, orientation & metric is established. These blob points are in fact are all the discrete points on the surface of our image with some of the points restricted in the SIFT due to the use of standard scaling. After getting the blob points by applying SIFT, we will apply Local beam search (LBS) on the blob points to find those points which have high describing strength of detected feature. After applying LBS we will extract and match features and finally use random sample consensus (RANSAC) [4] to estimate image transformation parameters and find a solution that has a best consensus with the data.

III. MOTIVATION

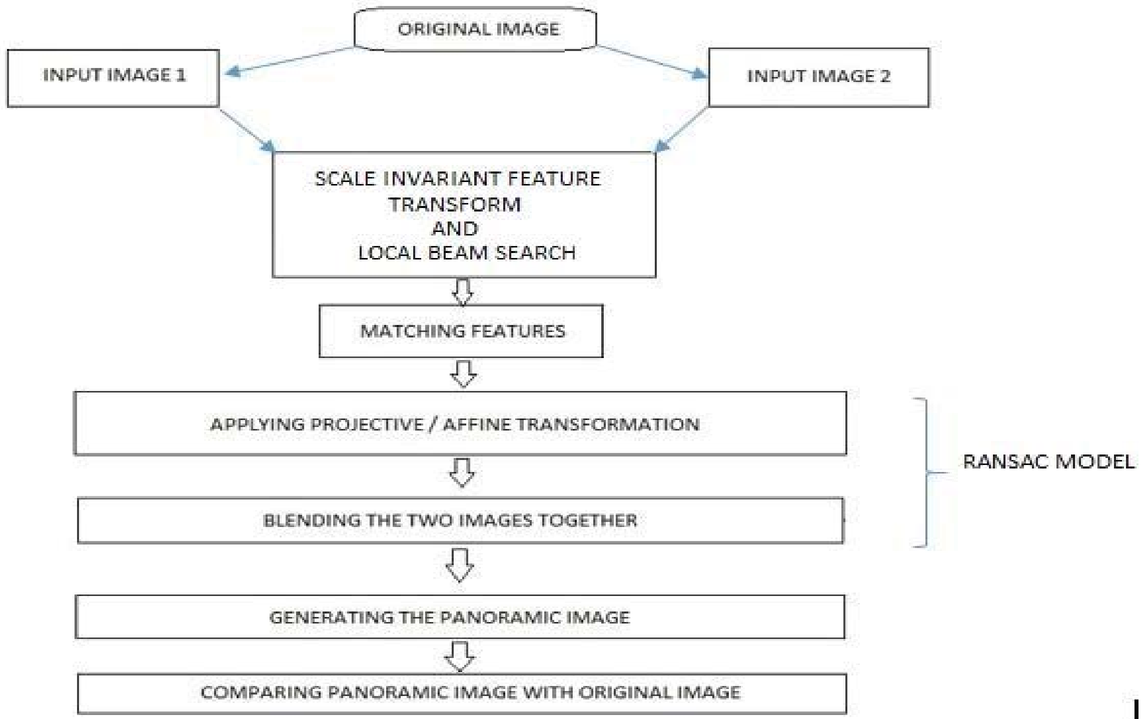
The base paper [1], on which this project's foundation is laid on, its ground work is fundamentally concentrating on the Scale Invariant Feature Transform (SIFT) and Random sample consensus algorithm (RANSAC). SIFT and RANSAC algorithm are already well established in the implementation of panorama generation in the real life.

SIFT: The first step in the panoramic recognition algorithm is to extract and match the features between all of the input images. SIFT features are located at scale space maxima/ minima of a difference of Gaussian function. At each feature location, a characteristic scale and orientation is established. This gives a similarity invariant frame in which to make measurements. Although simply sampling intensity values in this frame would be similarity invariant, the invariant descriptor is actually computed by accumulating local gradients in orientation histograms. This allows edges to shift slightly without altering the descriptor vector, giving some robustness to affine change. This spatial accumulation is also important for shift invariance, since the interest point locations are typically only accurate in the 0-3 pixel range. Illumination invariance is achieved by using gradients (which eliminates bias) and normalizing the descriptor vector.

RANSAC (random sample consensus) is a robust estimation procedure that uses a minimal set of randomly sampled correspondences to estimate image transformation parameters, and finds a solution that has the best consensus with the data. In the case of panoramas we select sets of $r = 4$ feature correspondences and compute the homography H between them using the direct linear transformation (DLT) method. We repeat this with $n = 500$ trials and select the solution that has the maximum number of inliers. The above research paper does not deal with the comparison of the final generated panorama image with the original non panoramic image in order of quality and effectiveness. The second limitation we encountered was that the final generated panoramic image differs in the size with the original non panoramic image. That's why we were motivated to come up with an experimental setup to study the quantitative outcome of the conventional SIFT and RANSAC algorithm, to check the difference between the final generated panoramic image with the original non panoramic image. We will also implement the Local Beam Search with the RANSAC algorithm in place of conventional SIFT and RANSAC model.

IV. ARCHITECTURE

System diagram



Description

The first step for panoramic view generation using LBS is to divide an image firstly in 60:40 ratio and secondly in 40:60 ratio and taking the 60:60 portion of both the images as two input images. Now SIFT features are found using SURF (Speeded up robust feature) algorithm. SIFT features are located at scale space maxima/minima of a difference of Gaussian function. Each feature has its own scale orientation and metric value. Now LBS is being applied on the SIFT features. The whole image is divided in dynamic grids (since the total number of grids can be adjusted in accordance with total number of SIFT features). Grids are formed in order to select subordinate features from SIFT features,

On the basis describing strength of detected SIFT feature. The SURF algorithm uses a determinant of an approximated hessian. Now the features are extracted from the found points (blob points), which are obtained on application of LBS on those points.

RANSAC [4] (Random sample consensus) is robust estimation procedure which is used to estimate image transformation parameters which is applied on the extracted subordinate features. RANSAC is essentially a sampling approach to find H, which is the homography between the set of features using Direct Linear Transformation (DLT) method [5].

Finally we have a set matched features that are geometrically consistent (RANSAC inliers) and a set of features that are inside the area of overlap but not consistent (RANSAC inliers), we use bundle adjustments [6] to solve for all of the camera parameters jointly and after automatic panorama straightening and multiband blending the final panorama is generated.

Algorithm

Automatic Panorama Stitching Using Local Beam Search
Input: n unordered images

I. Extract SIFT features from all n images.

II. Dynamically divide the whole grid into several small grids.

III. For each small grid

(i) Find the highest value of the describing strength of the detected feature, using the local beam search algorithm.

IV. Find k nearest-neighbors for each feature using a k-d tree

V. For each image:

(i) Select m candidate matching images that have the most feature matches to this image

(ii) Find geometrically consistent feature matches using RANSAC to solve for the homography between pairs of images

(iii) Verify image matches using a probabilistic model

VI. Find connected components of image matches

VII. For each connected component:

(i) Perform bundle adjustment to solve for the rotation theta 1, theta 2, theta 3 and focal length f of all cameras

(ii) Render panorama using multi-band blending

VIII. Compare the generated panorama with the original non panoramic image using the Euclidean Distance vector method.

Output: Euclidean distance between the two images.

V. EXPERIMENTAL SETUP RESULT

Our experimental setup is made to analyze the panorama generated by the 60:60 part of original input images with the original non panoramic image.

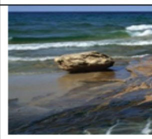
WITHOUT LBS								
SL No.	Input		Output	Euclidean distance	extractFeature (Total CPU time(in sec.))	matchFeature(Total CPU time (in sec.))	Total number of points	Total number of filtered points
	Right Crop	Left Crop	Output					
1				5.5732e-05	0.026	0.014	L=57 R=90	L=57 R=90
2				1.0683e-04	0.035	0.013	L=48 R=161	L=48 R=161
3				0.09904	0.081	0.12	L=348 R=222	L=348 R=222

Table 1

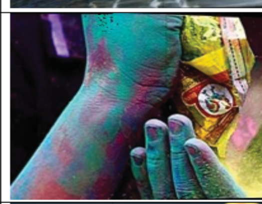
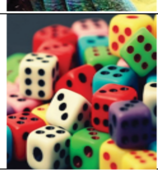
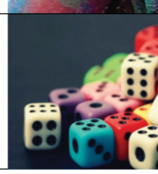
WITH LBS								
SL No.	Input		Output	Euclidean distance	extractFeature (Total CPU time)	matchFeature(Total CPU time)	Total number of points	Total number of filtered points
	Right Crop	Left Crop	Output					
1				4.8021e-05	0.021	0.012	R=90 L=57	L=36 R=36
2				2.4924e-04	0.025	0.011	R=161 L=48	L=40 R=60
3				0.10042	0.044	0.013	L= 222 R =348	L=80 R=80

Table 2

The right and left crop are the two input images we have taken as shown the system diagram.
The output of (1) and (2) in the tabular result shown below are the panoramic view generated by the algorithm as presented in our base paper [1]. Their Euclidean distance with the original input image is shown. Total number of SIFT features received and total number of SIFT features passed are shown different columns. In the second half of the experimental setup we have used our algorithm on the same set of input images to find the Euclidean distance between generated panoramic image and the original non panoramic image. The setup also represents the final grid size which we incorporated and total number of subordinate features received after applying LBS.
In the implemented methodology of applying LBS in the base paper [1], the final CPU runtime of extracting features and matching features is comparatively lower than the original methodology of base paper.

VI. CONCLUSION

This paper has introduced the application of local beam search (LBS) in accordance with SIFT and RANSAC in order to obtain effective extracting and matching of SIFT features. The experimental setup evaluates the generated panoramic image using LBS and without using LBS and compares with the original non panoramic image using Euclidean Distance vector method.

VII. FUTURE WORK

Possible areas for future work include:
Artificial Neural Network (ANN): ANN can be introduced in the proposed work in order to automate the grid positioning according to the describing strength of detected features value (Metric value), which currently is being done manually. We can also apply ANN for restoring the digital movement of images which could be considered for panorama generation [2].The operating principal of artificial neural networks is to learn and record the information or knowledge in the memory.

VIII. ACKNOWLEDGEMENT

The authors would like to thank Sharmistha Majumder, Assistant Professor, Computer Science and Engineering Department, National Institute of Technology, Agartala, India and National Institute of Technology, Agartala computer laboratory for supporting and guiding us for the completion of this paper.

REFERENCES

[1] Matthew Brown and David G. Lowe, Automatic Image stitching using in variant feature {mbrown | lowe} @Cs.ubc.ca . Department of Computer Science,University of British Columbia,Vancouver, Canada.
[2] Hajji Tarik* , Mohcine Kodad, Jaara El Miloud, Digital movement image restoring by Artificial neural networks, Laboratoire de Recherche Informatiques (LaRI), Faculté des Sciences d'Oujda (FSO), Université Mohammed Première (UMP) Oujda, 60000, Maroc (LaRI-FSO-UMP).
[3] D. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, 2004
[4] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography.Communications of the ACM, 24:381–395, 1981.
[5] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, ISBN:0521540518, second edition,2004.
[6] W. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle adjustment: A modern synthesis. In Vision Algorithms: Theory and Practice, number 1883 in LNCS, pages 298–373. Springer-Verlag, Corfu, Greece, September1999.