Multi-View Point Panorama Construction With Wide-Baseline Geo-Graphical Images

[1]Rony M Mariam, [2]M Sivasubramanian, [3]S Meenakshi
[1]Research Scholar, [2]Assistant Professor, [3]Head Of The Department
JP College Of Arts And Science Agarakattu Tenkasi

Abstract: **Here we present an image stitching approach, which can produce visually plausible panoramic images with input taken from different viewpoints. Unlike previous methods, our approach allows wide baselines between images and non-planar scene structures. Instead of 3D reconstruction, we design a mesh based framework to optimize alignment and regularity in 2D. By solving a global objective function consisting of alignment and a set of prior constraints, we construct panoramic images, which are locally as perspective as possible and yet nearly orthogonal in the global view. We improve composition and achieve good performance on misaligned areas. Experimental results on challenging data demonstrate the effectiveness of the proposed method.**

***Index Terms*—Image stitching, multi-view panorama, image alignment, wide-baseline images.**

I. INTRODUCTION

1.1 IMAGE PROCESSING:

An image is an array, or a matrix, of square pixels (picture elements) arranged in columns and rows. Pictures are the most common and convenient means of conveying or transmitting information. A picture is worth a thousand words. Pictures concisely convey information about positions, sizes and inter relationships between objects.

1.2 PANOROMIC IMAGE

Panoramic image is the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama of high-resolution image. It is commonly performed through the use of computer software; most approaches to image stitching require nearly exact overlaps between images and identical exposures to produce seamless results

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It is alignment of two images geometrically - the reference and sensed images.

II. MATERIALSE & METHOD

2.1EXISTING SYSTEM

A number of methods have been developed to perform image matching for registration and change detection. Closer to the work presented in this paper, researchers have also investigated local features for detection and classification. Sirmacek and Unsalan use local features to detect buildings and urban areas in 1-m resolution IKONOS imagery. Xu compare quantized color and texture features with local features for classifying 0.25-m resolution aerial image regions into four LULC classes. Chen also compares local features with standard color and texture features to classify 0.5-m Digital Globe imagery into 19 LULC classes. Skurikhin investigates attention-based saliency detection to perform local feature-based classification of 0.5-m resolution Digital Globe and Google Earth imagery into anthropogenic or natural regions.

2.2 DISADVANTAGES

- The above features lead to only an average performance.
- There is a chance to have a mismatch between the images.
- Excess for noise present at output image.

2.3. PROPOSED SYSTEM

The proposed system consists of five different modules as listed below and a brief description of all the modules is followed.

- Detect Key-points using SURF Detector
- Feature Extraction
- SURF Histogram Features
- Compute Euclidean Distance

2.4 DETECTION OF KEY-POINTS

SURF is an algorithm to detect and describe local features in images. SURF descriptors are extracted from an image in two steps. First, a detection step locates points that are identifiable from different views. This process ideally locates the same regions in an object or scene regardless of viewpoint or illumination. Second, these locations are described by a

descriptor that is distinctive yet invariant to viewpoint and illumination. Following are the major stages of computation used to generate the set of image features:

2.4.1Scale-space extrema detection

The first stage of computation searches over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation. Candidate locations are initially selected from local extrema in difference of Gaussian (DoG) filtered images in scale space. The DoG images are derived by subtracting two Gaussian blurred images with different σ

$D(x, y, \sigma) = L(x,y,k\sigma) - L(x,y,\sigma)$---------(2.1)

Where $L(x, y, \sigma)$ is the image convolved with a Gaussian kernel with standard deviation σ, and k represents the different sampling intervals in scale space. Each point in the 3-D DoG scale space is compared with its eight spatial neighbors at the same scale, and with its 18 neighbors at adjacent higher and lower scales. The local maximum or minimum are further screened for minimum contrast and poor localization along elongated edges. The last step of the detection process uses a histogram of gradient directions sampled around the interest point to estimate its orientation. This orientation is used to align the descriptor to make it rotation invariant (RI).

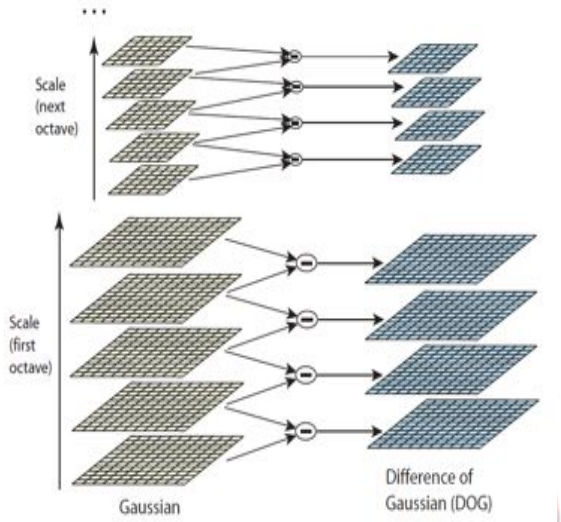

Fig. 2.2. Scale-space extrema detection

Fig 2.2 shows that for each octave of scale space, the initial image is repeatedly convolved with Gaussians to produce the set of scale space images shown on the left. Adjacent Gaussian images are subtracted to produce the difference-of-Gaussian images on the right. After each octave, the Gaussian image is down-sampled by a factor of 2, and the process repeated.

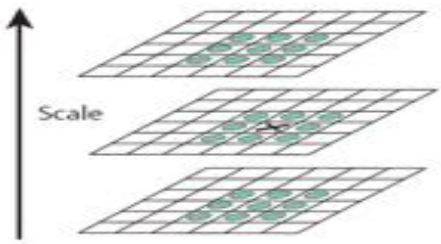

Fig. 2.3. Scale

Fig 2.3 shows that maxima and minima of the difference-of-Gaussian images that are detected by comparing a pixel (marked with X) to its 36 neighbors in 3x3 regions at the current and adjacent scales (marked with circles).

2.4.2 Keypoint localization

Determine the scale and locations. Key points are selected based on measures of their stability. It can be used to identify locations in image scale space that are invariant with respect to image translation, scaling, and rotation, and are minimally affected by noise and small distortions.

2.4.3 Key point descriptor

The local image gradients are measured at the selected scale in the region around each key point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

A SURF descriptor is extracted from the image patch centered at each interest point. The size of this patch is determined by the scale of the corresponding extreme in the DoG scale space. This makes the descriptor scale invariant. The

feature descriptor consists of histograms of gradient directions computed over a 4 × 4 spatial grid. The interest point orientation estimate is used to align the gradient directions to make the descriptor RI. The gradient directions are quantized into eight bins so the final feature vector has dimension 128 (4 × 4 × 8). Consider extracting SURF descriptors from a fixed grid instead of from the salient interest points. This approach has been named the Scale Invariant Feature Transform (SURF), as it transforms image data into scale-invariant coordinates relative to local features.

- First, the SURF detector is translation, rotation, and scale invariant which is the level of invariance needed.
- Second, an extensive comparison with other local descriptors found that the SURF descriptor performed the best in an image matching task.

2.5 FEATURE EXTRACTION

Three standard image features are considered: simple statistics, homogeneous texture, and color histogram features.

2.5.1 Simple Statistics

A 2-D feature vector is computed for each ground truth image consisting of the mean and standard deviation of the grayscale values,

fss= (μ,σ) ---------- (5.2)

This is referred to as the simple statistics feature and serves as a baseline for the experiments.

Mean is the average value and the standard deviation is the square root of variance.

2.5.2 Homogeneous Texture

Homogeneous Texture Descriptors compliant with the Multimedia Content Description Interface are extracted using banks of Gabor filters tuned to five scales and six orientations. A 60-dimensional feature vector is formed from the mean and standard deviation of the 30 filters .where μ_{RS} and σ_{RS} are the mean and standard deviation of the output of the filter tuned to orientation R and scale S. To account for differences in range, normalized versions of the features are also produced in which each of the 2RS components is scaled to have a mean of zero and a standard deviation of one over the ground truth data set.

2.5.3 Color Histogram

The color histogram serves as an effective representation of the color content of an image if the color pattern is unique compared with the rest of the data set. The color histogram is easy to compute and effective in characterizing both the global and local distribution of colors in an image. In addition, it is robust to translation and rotation about the view axis and changes only slowly with the scale, occlusion and viewing angle.

2. BOVW

The SURF detector, adopt a standard approach, termed BOVW, to summarize the descriptors without regard to where they appear in an image. The analogy to representing text documents as word count frequencies is made possible by quantizing the 128 dimension SURF descriptors. Apply standard k-means clustering to a large number of SURF descriptors to create a dictionary of visual words or codebook.

2.6 SURF HISTOGRAM FEATURES

SURF histogram features are calculated for each ground truth image by using a codebook to quantize the SURF descriptors
extracted from the image. The histogram features thus range in length from 10 to 20 000 components. Three versions of the histogram features are considered:

- 1 Un-normalized SURF histogram features which simply contain the codeword counts;
- L1 normalized SURF histogram features where the components are normalized to sum to one; and
- L2 normalized SURF histogram features where the components are normalized so the feature vectors have length one.

2.7 SURF ALGORTHM DESCRIPTION

Speeded up robust feature (or SURF) is an algorithm in computer vision to detect and describe local features in images. For any object in an image, interesting points on the object can be extracted to provide a "feature description" of the object.Another important characteristic of these features is that the relative positions between them in the original scene shouldn't change from one image to another.

Lowe's patented method can robustly identify objects even among clutter and under partial occlusion, because his SURF feature descriptor is invariant to uniform scaling, orientation, and partially invariant to affine distortion and illumination changes. This section summarizes Lowe's object recognition method and mentions a few competing techniques available for object recognition under clutter and partial occlusion.

- Key stages
- Scale-invariantfeature detection

2.7.1 Feature matching and indexing

Indexing consists of storing SURF keys and identifying matching keys from the new image. Lowe used a modification of the k-d tree algorithm called the Best-bin-first search method that can identify the nearest neighbors with high

probability using only a limited amount of computation. The BBF algorithm uses a modified search ordering for the k-d tree algorithm so that bins in feature space are searched in the order of their closest distance from the query location. This search order requires the use of a heap-based priority queue for efficient determination of the search order. The best candidate match for each keypoint is found by identifying its nearest neighbor in the database of keypoints from training images. The nearest neighbors are defined as the keypoints with minimum Euclidean distance from the given descriptor vector

2.7.2 Cluster identification

Hough Transform is used to cluster reliable model hypotheses to search for keys that agree upon a particular model pose. Hough transform identifies clusters of features with a consistent interpretation by using each feature to vote for all object poses that are consistent with the feature. When clusters of features are found to vote for the same pose of an object, the probability of the interpretation being correct is much higher than for any single feature. An entry in a hash table is created predicting the model location, orientation, and scale from the match hypothesis. The hash table is searched to identify all clusters of at least 3 entries in a bin, and the bins are sorted into decreasing order of size.

5.2 MODEL VERIFICATION

Each identified cluster is then subject to a verification procedure in which a linear least squares solution is performed for the parameters of the affine transformation relating the model to the image. The affine transformation of a model point $[x\ y]^T$ to an image point $[u\ v]^T$ can be written as below

5.3 OUTLIER DETECTION

Outliers can now be removed by checking for agreement between each image feature and the model, given the parameter solution. Given the linear least squares solution, each match is required to agree within half the error range that was used for the parameters in the Hough transform bins. As outliers are discarded, the linear least squares solution is re-solved with the remaining points, and the process iterated. If fewer than 3 points remain after discarding outliers, then the match is rejected. The final decision to accept or reject a model hypothesis is based on a detailed probabilistic model. This method first computes the expected number of false matches to the model pose, given the projected size of the model, the number of features within the region, and the accuracy of the fit.

5.6. DISCARDING LOW-CONTRAST KEYPOINTS

To discard the keypoints with low contrast, the value of the second-order Taylor expansion $D(x)$ is computed at the offset $\hat{\mathbf{x}}$. If this value is less than 0.03, the candidate keypoint is discarded. Otherwise it is kept, with final scale-space location $\mathbf{y} + \hat{\mathbf{x}}$, where $\mathbf{y}$ is the original location of the key point.

5.7 ELIMINATING EDGE RESPONSES

The DoG function will have strong responses along edges, even if the candidate keypoint is not robust to small amounts of noise. Therefore, in order to increase stability, we need to eliminate the keypoints that have poorly determined locations but have high edge responses. Finding these principal curvatures amounts to solving for the eigenvalues of the second-order Hessian matrix, **H**:

5.8 COMPARISON OF SURF FEATURES

SURF and SURF like GLOH features exhibit the highest matching accuracies (recall rates) for an affine transformation of 50 degrees. After this transformation limit, results start to become unreliable. Distinctiveness of descriptors is measured by summing the Eigen values of the descriptors, obtained by the Principal components analysis of the descriptors normalized by their variance. This corresponds to the amount of variance captured by different descriptors, therefore, to their distinctiveness. PCA-SURF (Principal Components Analysis applied to SURF descriptors), GLOH and SURF features give the highest values. SURF-based descriptors outperform other contemporary local descriptors on both textured and structured scenes, with the difference in performance larger on the textured scene. For scale changes in the range 2-2.5 and image rotations in the range 30 to 45 degrees, SURF and SURF-based descriptors again outperform other contemporary local descriptors with both textured and structured scene content. Introduction of blur affects all local descriptors, especially those based on edges, like shape context, because edges disappear in the case of a strong blur. But GLOH, PCA-SURF and SURF still performed better than the others. This is also true for evaluation in the case of illumination changes.

3.4 SYSTEM ARCHITECTURE

System investigates one analysis method wherein local image regions are characterized by features designed to be invariant to differences in appearance resulting from geometric transformations such as rotation or scaling as well as from photometric transformations such as changes in illumination. The image regions themselves are also detected in an invariant manner. The local invariant features have been successfully applied to a range of standard (no geographic) computer vision problems, and there has been increasing interest in using them for overhead image analysis. the output image is retrieved from the database by computing SURF Algorithm and also by using Feature extraction key points.

Systems perform an extensive evaluation of local invariant features for image retrieval of land-use/land-cover (LULC) classes in high-resolution aerial imagery. The effects of a number of design parameters on a bag-of-visual-words (BOVW) representation including saliency- versus grid-based local feature extraction, the size of the visual codebook, the clustering algorithm used to create the codebook, and the dissimilarity measure used to compare the BOVW representations**.**

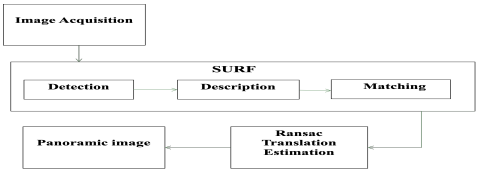

Fig. 3.1. Proposed System Architecture

IV.EXPERIMENTAL RESULT

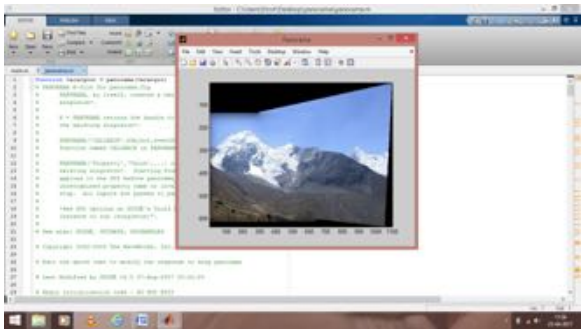

Fig. 6. Panorama Window

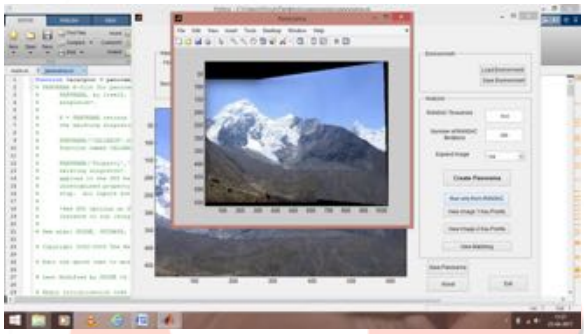

Fig. 7. Run only from RANSAC

Fig. 8. View image1 Key Points

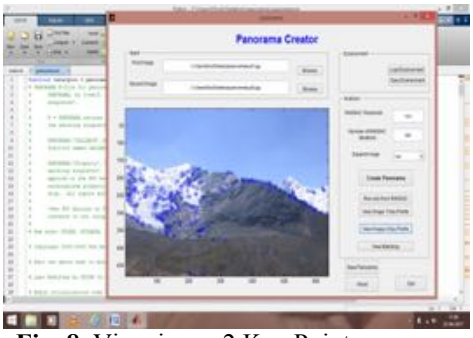

Fig. 8. View image2 Key Points

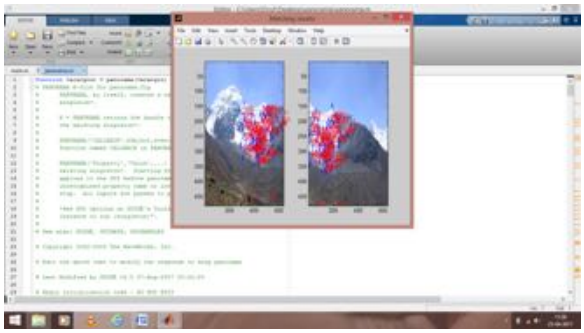

Fig. 8. Matching Results

V.CONCLUSION

Our proposed work demonstrates that the local invariant features are more effective than standard features such as color and texture for image retrieval of LULC classes in high-resolution aerial imagery. Speeded Up Robust Features is the most advanced algorithm for image mosaicking. It is a robust local feature detector. This algorithm use 2D Haar wavelet response and make an efficient use of integral images. This algorithm works effectively in presence of noise and other minor variation. Also this algorithm is scale and illumination invariant. This algorithm can be used for the real time application as because the standard version of SURF is several times faster than SIFT.For the retrieved image, restoration, deconvolution, deblurring has been done to recover the image using SURF algorithm. Finally the PSNR ratio has been calculated for the recovered image.

VI. REFERENCES

[1]. Duc-Son Pham and Svetha Venkatesh,"Efficient Algorithms for Robust Recovery of Images from Compressed Data'', IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 22, NO. 12, DECEMBER 2013

[2] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale invariant feature matching for optical remote sensing images," IEEE Trans. Geosci. Remote Sens., vol. 49, no. 11, pp. 4516–4527, Nov. 2011.

[3] X. Jianbin, H. Wen, and W. Yirong, "An efficient rotation-invariance remote image matching algorithm based on feature points matching," in Proc. IEEE Int. Geosci. Remote Sens. Symp, 2005, vol. 1, pp. 647–649.

[4] J. Dai, W. Song, L. Pei, and J. Zhang, "Remote sensing image matching via Harris detector and SURF descriptor," in Proc. Int. Congr. Image Signal Process, 2010, vol. 5, pp. 2221–2224.

[5] A. Mukherjee, M. Velez-Reyes, and B. Roysam, "Interest points for hyper spectral image data," IEEE Trans. Geosci. Remote Sens., vol. 47, no. 3, pp. 748–760, Mar. 2009.

[6] L. Dorado-Munoz, M. Velez-Reyes, A. Mukherjee, and B. Roysam, "A vector SURF operator for interest point detection in hyper spectral imagery," in Proc.Workshop Hyperspectr. Image Signal Process.—Evolution Remote Sensing, 2010, pp. 1–4.

[7] Z. Xiong and Y. Zhang, "A novel interest-point-matching algorithm for high-resolution satellite images," IEEE Trans. Geosci. Remote Sens., vol. 47, no. 12, pp. 4189–4200, Dec. 2009.

[8] C. Huo, Z. Zhou, Q. Liu, J. Cheng, H. Lu, and K. Chen, "Urban change detection based on local features and multiscale fusion," in Proc. IEEE Int. Geosci. Remote Sens. Symp, 2008, vol. 3, pp. 1236–1239.

[9] F. Tang and V. Prinet, "Computing invariants for structural change detection in urban areas," in Proc. Urban Remote Sens. Joint Event, 2007, pp. 1–6.

[10] B. Sirmacek and C. Unsalan, "Urban-area and building detection using SURF keypoints and graph theory," IEEE Trans. Geosci. Remote Sens., vol. 47, no. 4, pp. 1156–1167, Apr. 2009.