Automatic Image Registration using SIFT-NCC

ViniVidyadharan
Mtech Student
Department of Computer Science & Engineering
SreeChitraThirunal College of Engineering

SubuSurendran
Assistant Professor
Department of Computer Science & Engineering
SreeChitraThirunal College of Engineering

ABSTRACT
Accurate, robust and automatic image registration is critical task in many typical applications that employ multi-sensor and/or multi-date imagery information. The main content of this paper is an algorithm for the registration of digital images. Some multi–sensed or temporal images contain large number of speckles and noise, or image can have some distortion by some means. For these reasons, we need to remove the noises, speckle and to recover from distortion. We register two to find the similarity between the images. This paper discusses techniques for image registration based on SIFT. In this proposed framework we use NCC metrics for optimizing the matching work. Best bin first search using kd tree is used for feature matching and RANSAC is used for outlier elimination.

Keywords
Scale Invariant Feature Transform, NCC, RANSAC, KD tree, BBF.

1. INTRODUCTION
Image registration is an importantpre-processing step in many image processing applications such as computer vision, remote sensing and medical image processing [9]. Its purpose is to overlay two or more images of the same scene taken at different times or from different viewpoints and/or by different sensors [10][11]. Image registration is achieved by aligning two or more images according to the transformation between them [9]. The image registration techniques can be generally categorized into area-based and feature based methods. In the area-based methods, a small region of points in the source image is compared with a region of the target image. The measure of matching is usually the mutual information or normalized cross-correlation. Mutual informationand normalized cross-correlation are established multimodal registration similarity metrics. In contrast, the feature-based method, utilizes extracted features (usually keypoints), where common structures from two images are matched [11].

Finding feature correspondences between two or more images of the same scene has been one of the fundamental problems in computer vision such as image registration, object tracking and object recognition [5]. From the early years itself there are many approaches to find the feature correspondences between the images. The main difficulties in this problem are the ambiguity in detecting the same points in both images and matching them based on their local features [5]. Recent studies in the literature use local invariant descriptors to describe and match the image features under various changes [2]. These local descriptors enable the detected points to be more distinctive [5].

The classical difficulty involved in building such a registration system isthe problem of aligning the frames fromdifferent camerasin real time. Even if the frames are properly aligned, there can be artifacts at the alignment boundaries asthe frames may differ in color, and in lighting conditions. To minimize these artifacts and to provide a seamless panoramic output stream, a reasonably good registration technique needs to be applied at the alignment boundaries.

In this paper we propose an algorithm where SIFT [2], RANSAC [4] and NCC together is used to determine the corresponding points in the overlapping areas of both the images. These corresponding points definethe underlying transformation matrix between the frames.The paper is organized as follows. Section 2 does a literature survey of various image registration algorithms. The proposed algorithm is described in Section 3. Experimental results are analyzed in Section 4 and Section 5 concludes the work..

2. LITERATURE SURVEY
In order to find matches between the images under various circumstances, many approaches have been developed in the literature. Each of these approaches first detects feature points and then calculates feature descriptors for similarity measurement. The feature correspondences are then used to recover the geometric transformation that registers the image [11]. Schmid and Mohr [6] showed that rotated images can be matched by using Harris' corner detector and rotationally invariant local descriptor. Montesinos et al [12] proposed the use of differential descriptor of the neighborhood of the detected feature points. Dufournaud et al. [7] proposed a multi-scale framework. If the two images of the same scene differ largely by a camera's viewpoint, focal length, orientation, or illumination, conventional corner detectors may not detect the same points in both images very well and correlation based similarity measures tend to fail to correctly distinguish the detected points [3].A high resolution image is smoothed at different scale levels and the scale ratio between two images is estimated by initial matching. Then, the higher resolution image is smoothed to the scale of a lower resolution image and feature points between two images are compared.

Lowe [2] used the local extrema of difference-of- Gaussian in scale-space as feature points. He proposed a local descriptor, which is computed by local image gradients around feature points and this is known as Scale Invariant Feature Transform (SIFT).Mikojclyk and Schmid [13] did the comparative study on the various local invariant feature descriptors, and found that the SIFT gives the best performance with respect to other features.

In the El Rube's [9] work he combined the multi-scale wavelet transform with SIFT, where the images are decomposed into multiple scales using wavelet transform, then the low frequency image at certain level is input to SIFT algorithm. He assures that the algorithm speeds up the calculation of the correspondences between images and his algorithm works perfectly more on remote sensing images.

3. METHOD DESCRIPTION

As we have discussed earlier, the performance of the local descriptor found by SIFT algorithm is much better than other descriptors due to is invariant features, but the accuracy of the feature correspondences between two images needs to be improved. Moreover, the number of keypoints obtained by the SIFT is very large which increases the complexity of the registration algorithm.

In order to solve this problem an algorithm for image registration that combines SIFT, RANSAC and NCC is proposed in this paper.

An overview of the algorithm is as follows: First the keypoints of both images are found using SIFT. Using RANSAC the outlier elimination is done and NCC is used to optimize the keypoints found. Best Bin First Search is used to find the matching keypoints which is far better than Euclidean distance, SSD or other measures.

3.1 Sift Algorithm Overview

In this feature extraction and matching, first step is to detect each point that has to be matched with the features of another image. Each detected point and its neighbourhood are described by feature descriptor which is used to match with the points in another image. It is important that the same points must be detected in both images and the detected points should be characterized as uniquely as possible [5].

The SIFT algorithm developed by Lowe is invariant to image translation, scaling rotation, and partially invariant to illumination changes and affine *3D* projection. The SIFT algorithm has four main steps: (1) Scale Space Extrema Detection, (2) Keypoint Localization, (3) Orientation Assignment and (4) Description Generation.

The first stage is to identify location and scales of keypoints using scale space extreme in the difference-of-Gaussian (DoG) functions with different values of σ; the DoG function is convolved of image in scale space separated by a constant factor k as in the following equation.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) \times I(x, y) \qquad (1)$$

Where, G is the Gaussian function and I is the image.

Then the Gaussian images are subtracted to produce a DoG, after that the Gaussian image subsampled by factor 2 and produce DoG for sampled image. A pixel compared of 3×3 neighborhood to detect the local maxima and minima of $D(x, y, \sigma)$.

In second stage, keypoints candidate are localized and refined by eliminating the keypoints where the unstable is found. In the third stage, the orientation of keypoint is obtained based on local image gradient. The last stage is to compute the local image descriptor for each keypoint based on image gradient magnitude and orientation at each image sample point in a region centered at keypoint; these samples building 3D histogram of gradient location and orientation; with 4×4 array location grid and 8 orientation bins in each sample. That is 128-element dimension of keypoint descriptor.

3.2 Normalized Cross-Correlation [1]

In the coarse matching stage, we adopt normalized cross correlation (NCC) algorithm. The normalized measure is defined as

$$C = \frac{\sum_{x=-w}^{w}\sum_{y=-h}^{h}[I_1(m+x,n+y)-\bar{I}_1(x,y)]\times[I_2(m+x,n+y)-\bar{I}_2(x,y)]}{\sqrt{\sum_{x=-w}^{w}\sum_{y=-h}^{h}[I_1(m+x,n+y)-\bar{I}_1(x,y)]^2}\times\sqrt{\sum_{x=-w}^{w}\sum_{y=-h}^{h}[I_2(m+x,n+y)-\bar{I}_2(x,y)]^2}} \qquad (2)$$

where $(2w+1)\times(2h+1)$ is the size of correlation window.And C is the correlation coefficient value, its range is in [−1, 1].If C is greater than α (a threshold, $\alpha = 0.95$ in this paper),we deem the corners are matched.

For each features in reference image must calculate thecorrelation coefficient with all features in another imageaccording to the traditional NCC algorithm. In practice, for one feature in reference image there is at most one real matched feature even zero in another image. One can imagine that themost of calculations are invalid. Hence, we can improve thespeed if we reduce the unnecessary calculations.

3.3 RANSAC

RANSAC (RANdom SAmple Consensus) algorithm is to estimate parameters of a mathematical model from a set of observed data which contains outliers [16]. We use this algorithm to eliminate mismatches after searching initial homonymy point-pairs from BBF method, the mapping relationship of two images is denoted by H, which is a 3×3 matrix with 8 degrees of freedom, and can be computed from 3 homonymy keypoint-pairs which are not collinear at least for affine transformation.

3.4 Best Bin First Search

The function of BBF (Best Bin First) algorithm is to find KNN (K-nearest neighbor) in the KD-tree [15], BBF algorithm uses the priority queue (here using the minimum heap) to achieve, when scanning the KD-tree from the root to leaf finding the road, the missed points enter the priority queue firstly; and then remove the current minimum key value from the queue (this is the smallest distance at k_i-dimension). Repeat the above process until the queue is empty, or the repetition is 200 times.

Actually, K-Dimension tree is a balanced binary tree, General K-D tree construction process is as follows:

```
functionkdtree (list of points pointList, int depth)
{
        ifpointList is empty
                return nil;
        else {
        // Select axis based on depth so that axis cycles through all valid values
                varint axis := depth mod k;
        // Sort point list and choose median as pivot element
                select median by axis from pointList;
        // Create node and construct subtrees
                vartree_node node;
                node.location := median;
                node.leftChild := kdtree(points in pointList before median, depth+1);
                node.rightChild := kdtree(points in pointList after median, depth+1);
                return node;
        }
}
```

The process of BBF arithmetic is as follows:

Step1: built KD-tree with the features of image 2.

Step2: search for k-nearest neighbor of each feature in image 1, where k = 2.

Step3: If you find a k-nearest neighbor (k = 2), then determine whether they are effective features. If $d_0/d_1 < 0.49$ (d_0 and d_1 are respectively the nearest and the second nearest distance), they are.

4. EXPERIMENTAL RESULTS

The system has been tested with different set of images and the results were promising. The approach worked successfully on various images such as real images with various scale, rotation and illumination. It also worked successfully on satellite images.

In this paper we are considering a source and target image for registration. Figure 1 shows the source and the target image which are the inputs of the system.

(a) (b)

Fig 1: (a) Source Image (b) Target Image

In the first phase, the keypoints of both the images are extracted using SIFT method. Figure 2 shows the extracted keypoints of both the images. The SIFT descriptor are stored using the data structure KD tree so that the matching pairs can be found out using BBF algorithm which is explained above.

(a) (b)

Fig 2: (a) Keypoints identified from source image, (b) Keypoints identified from target image

In the second phase, the extracted keypoints are matched by comparing the SIFT descriptors. Figure 3 shows the raw matching pairs identified between the images. These matching pairs are found using the BBF algorithm. For optimization and outlier elimination we are using NCC and RANSAC. Figure 4 shows the final matching pairs between the images.

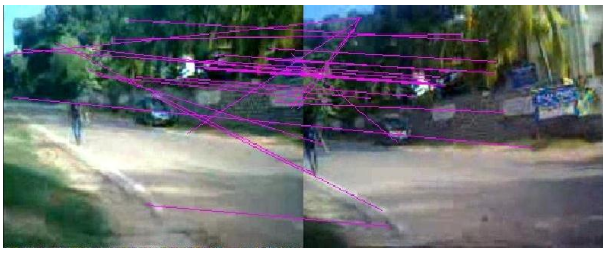


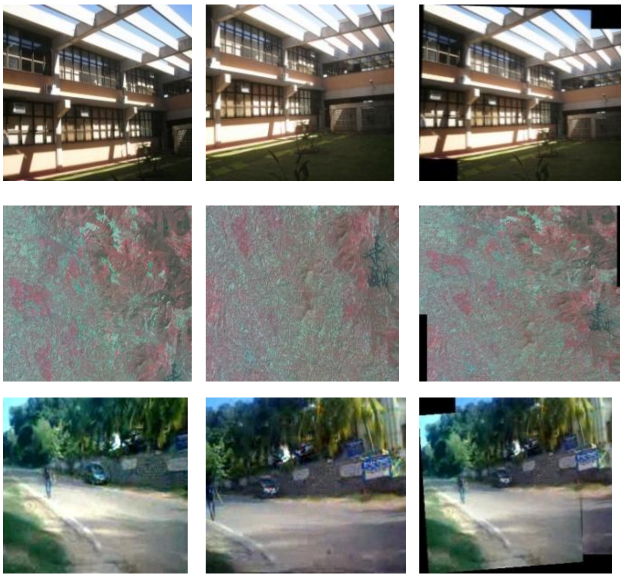
Fig 3: Matching pairs identified between the images

Fig 4: Final matching pairs identified between the images

In the final phase the target image blended to the source image using the transformation matrix found using the matching pairs found in the second phase. Figure 5 shows the final registered image.

Fig 5: Final Registered Image

The resultant registered images obtained for three different data sets are shown below.



(a) (b) (c)

Figure 6: (a) Source Images, (b) Target Images, (c) the resultant registered images obtained from the corresponding source and target images.

5. CONCLUSION AND FUTURE SCOPE

This image registration algorithm is faster and accurate than most of the algorithms. Our future scope is to make this algorithm work for the video registration. This algorithm works for most of the differently invariant images.

6. REFERENCES

[1] Haifeng Liu, Chuangbai Xiao, Mike Deng and Zhifei Tang, "A Faster Image Registration Algorithm", 2010

3rd International Congress on Image and Signal Processing.

[2] D. Lowe, "Distinctive image features from scale-invariant keypoints", International Jr. of Computer Vision, 60(2), pp.91-110, 2004.

[3] M. Brown and D. Lowe, "Invariant features from interest point groups", Proc. British Machine Vision Conference, pp.253-262, 2002.

[4] Fischler, M. A., and Bolles, R. C. "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM,* Vol. 24, no. 6,pp. 381–395, June 1981.

[5] H. Jeon, A. Basso, and P.F. Driessen, "A Global Correspondence for Scale Invariant Matching using Mutual Information and the Graph Search", in Proc. ICME, 2006, pp.1745-1748.

[6] C. Schmid and R. Mohr, "Local gray-value invariants for image retrieval", IEEE Transactions of Pattern Analysis and Machine Intelligence, 19 (5) pp.530-534, 1997.

[7] Y. Dufournaud et al, "Image matching with scale adjustment", Computer vision and image understanding 93 (2004), pp.175-194, 2004.

[8] Yuquan Wang, Guihua Xia, Qidan Zhu and Tong Wang, "Modified Scale Invariant Feature Transform in omnidirectional images", in Proceedings of ICMA, 2009, pp.2632-2636.

[9] Ibrahim A. El Rube, Maha A. Sharks, and Ashor R. Salem, "Image Registration Based on Multi-Scale SIFT for Remote Sensing Images"

[10] Lisa G, Brown, "A survey of image registration techniques", ACM Computer Survey.24, Vol. 24, Issue 4, pp. 325-376, 1992.

[11] Barbara, Zitova, J., Flusser, "Image Registration methods: a survey", Image and Vision Computing, Vol. 21, No. 11, pp. 977-1000, 2003.

[12] P. Montesinos, V. Gouet, R. Deriche, D. Pele, "Matching color uncalibrated images using differential invariants", Image and Vision Computing 18, 2000, pp. 659-671.

[13] Mikolajczyk K, Schmid C, "A performance evaluation of local descriptors", IEEE Transaction on Pattern and Machine Intelligence, Vol 27, Issue 10, 2005, pp. 1615-1630.

[14] CHUM O. "Randomized RANSAC with Td,d test". BMVC, 2002.

[15] LEE J, HUANG C, LIU L. "A modified soft-shape-context ICP registration system of 3-D point data", ICONIP, 2007.

[16] Chen, H., Varshney, P.K., and Arora, M.K., 2003 "Mutual information based image registration for remote sensing data", International Journal of Remote Sensing, 24(18):3701-3706.

[17] Torr, P. 1995, "Motion Segmentation and Outlier Detection", Ph.D. Thesis, Dept. of Engineering Science, University of Oxford, UK.

[18] Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.T. 1995, "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry", Artificial Intelligence, 78:87-119.

[19] Hui Lin, Peijun Du, Weichang Zhao, Lianpeng Zhang, Huasheng Sun, "Image Registration Based on Corner Detection and Affine Transformation", 2010 3rd International Congress on Image and Signal Processing.

[20] Hernâni Gonçalves, Luís Corte-Real, Member, IEEE, and José A. Gonçalves, "Automatic Image Registration Through Image Segmentation and SIFT", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 49, NO. 7, JULY 2011

[21] Liu Jinxia Qiu Yuehong, "Application of SIFT Feature Extraction Algorithm on the Image Registration**",** The Tenth International Conference on Electronic Measurement &Instruments (ICEMI'2011).