New Region Feature Descriptor-based Image Registration Method

Francis Bowen, *Student Member IEEE*, Eliza Du*, *Senior Member IEEE*
Dept. of Electrical and Computer Engineering,
Indiana University-Purdue University Indianapolis
Indianapolis, IN USA

Jianghai Hu, *Member IEEE*
Dept. of Electrical and Computer Engineering
Purdue University
West Lafayette, IN USA

***Abstract*—Successful content-based image registration relies on the accurate identification of corresponding features across images. Geometric and photometric transformations between images may hinder an algorithm's ability to precisely match features. In this work, we propose a novel region descriptor detection and matching algorithm for use with image registration. The detection process utilizes invariant feature points, as well as their spatial relationships and textural characteristics to create a connected graph whose structure represents an invariant region descriptor. With such a framework, feature matching can be accomplished by graph matching with a defined similarity metric. Subsequent image registration steps are outlined that employ the invariant region descriptors. The results provide strong evidence of the region descriptor's effectiveness in applications involving image registration. Several scenarios are presented including the registration of general objects, aerial photography, as well as scenes before and after a disaster.**

Keywords-feature-based; image registration; invariant feature point; region descriptors; minimum spanning tree

I. INTRODUCTION

Image registration is an important processing task that serves many purposes. In medical image analysis the registration from multiple types of images provide an effective means for diagnosis [1,2] while applications involving remote sensing can benefit from the fusion of registered information [3]. Similarly, computer vision tasks such as image stitching [4] utilize automatic registration to accomplish their goal.

The process of image registration attempts to determine the homography between two images. By determining the geometric relationship between data sets, one image can be projected into the same perspective as another using the estimated transformation parameters.

Registration algorithms can be broadly classified as feature-based or area-based approaches [5]. In the former, distinct image features such as corners [6], gradient edges [7], or shape descriptors [8] are used to define the geometrical mapping between images. Feature-based approaches rely on the detected feature's invariance to affine, rotational, and translational transformations. With area-based methods, pixel intensities are compared directly for a sub-region of an image. In such a scenario, a similarity measure is coupled with an optimization algorithm in an attempt to identify the closest mapping of pixels [9,10].

In this work a new feature-based approach is proposed that employs invariant feature points and their spatial relationships to identify the suitable control points for the estimation of transformation parameters required for accurate registration. A minimum spanning tree (MST) constructed from Speeded-up Robust Feature (SURF) keypoints is described along with the matching criteria for the region descriptor and subsequent control point identification. Previous work in [11] show the effectiveness of the proposed region descriptor for image matching while this work extrapolates the core concepts to image registration. Finally, the direct linear transform (DLT) approach is utilized for the transformation parameter estimation.

Section II provides a brief overview of previous work pertinent to the proposed method, Section III outlines the proposed algorithm, Section IV presents the results and discussion, while a conclusion is offered in Section V.

II. PREVIOUS WORK

A. Speeded-up Robust Features

In [12], Bay proposes the SURF descriptor as an alternative to the Scale Invariant Feature Transform (SIFT) descriptor which aims to compute multi-scale feature points that are invariant to scale, rotation and translational deformations. The SURF descriptors are computed efficiently by exploiting integral images which are defined in (1) as,

(1)

In doing so, the area within a bounded region (A,B,C,D) of the original image can be computed using four memory accesses and three additions, regardless of the size of the area. With integral images operations, such as convolution, are calculated in a fixed amount of time, greatly improving on the computational requirements.

(2)

Interest points are established for a pixel location (x,y) using a multi-scale Hessian detector defined in (3).

(3)

where — for a given scale, . , , and are estimated using discrete box filters. Figure 1 illustrates the discrete box filters for and . Equation 3 includes a weighting factor, , that is shown to help conserve energy between the approximate and actual Gaussian kernels [5].

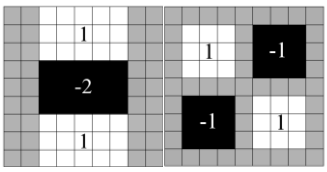


Figure 1. Box filters for and

The multi-scale image representations are estimated through the convolution of integer-based box filters that increase in size for each scale. This is in contrast to SIFT where the Gaussian pyramid and subsequent scale-space is constructed by convolving images of varying size with a Gaussian of constant dimensions. The set of images that constitute the scale-space is searched using a non-maximal suppression technique to identify candidate feature points within a 3x3x3 neighborhood.

After feature point detection, a descriptor is formed from the surrounding pixels using estimated Haar wavelet responses. The integral image is used to compute the fast convolution of an image for scale and an integer-based box filter for the Haar wavelet in the x and y directions. The size of the search neighborhood is dictated by the scale of the detected feature point. Within equally spaced 4x4 regions of the -pixel search space, a vector, , is computed from the sum of the Haar wavelets in the x and y directions. Each vector is composed of components , , , and , where and are the responses in the x and y direction, respectively. The descriptors for all sub-regions are concatenated to form the 64-dimensionsed descriptor.

(4)

Figure 2 provides an example of the two most common disadvantages of matching SURF feature points. First there may exist situation where multiple keypoints of one image map to the same keypoint of the other image. This may be a result of comparing similar textures. Second, SURF keypoints may be mismatched due to the repetitive nature of some textures or structures such as urban landscapes or nature scenes.

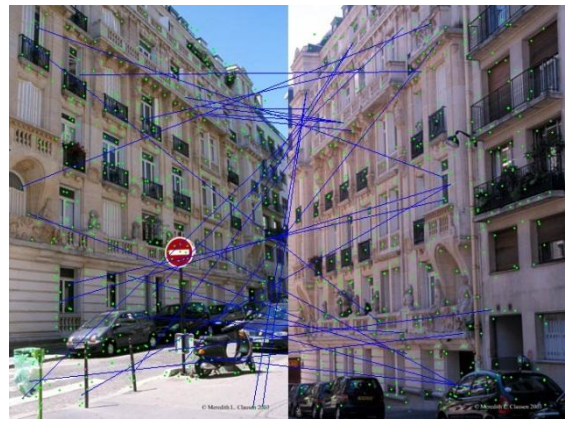


Figure 2. Matching SURF feature points using the Euclidian distance. Using point to point matching alone may create several mismatched pairs.

B. *Region Descriptors*

Region descriptors have become a popular method for image matching due to their higher discriminative power over individual feature points. Cho *et al.* [14] proposes a bag-of-words technique for grouping pertinent invariant features to form an effective region descriptor. Chen and Sun [15] describe an elliptical region descriptor based on the invariant Zernike moments. As compared to SURF points alone, the Zernike region descriptor is shown to be more effective for image matching.

III. Proposed Approach

This work presents a new approach for image registration that involves an effective graph-based region descriptor to identify suitable control points for the transformation parameter estimation. Figure 3 provides an overview of the entire system. Given a reference and query image, a novel set of region descriptors are extracted and matched across images. Using the matched regions, suitable control points are selected and utilized for estimating the transformation parameters.

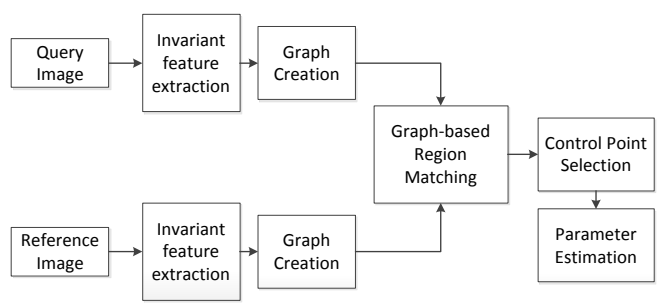


Figure 3. Image registration system overview

A. *Invariant Feature Extraction and Graph Creation*

Initially the invariant SURF keypoints are extracted from each image and clustered together such that all feature points within a cluster have similar descriptors and color characteristics. Descriptor similarity is defined as the Euclidean distance between the 64-dimensioned descriptor vectors while the color similarity is determined by the following,

(5)

where is color channel for descriptor . If features and are similar, the invariant keypoints are added to set according to the following,

(6)

For a given search window centered at each SURF keypoint, , the set of descriptors within the given neighborhood is defined in (8).

(7)

The descriptors within a search window are clustered if 75% of the descriptors are similar to all other descriptors within the window. We define the set of clustered descriptors as

(8)

Given two clusters, and , which share feature points, the clusters are merged if the number of common keypoints is at least 2.

(9)

Initial clustering results are provided in Figure 4 where SURF points are marked in green and the clusters are outlined in red.

Figure 4. Invariant feature point clustering.

For each set , a connected graph can be created if each invariant feature point is regarded as a vertex in a connected graph. Taking the origin as the upper left-hand corner of a cluster, a minimum spanning tree is created using Prim's algorithm [16]. The algorithm is greedy with complexity , where represents the number of points in the graph. The pseudo code for Prim's algorithm given as follows.

```
MST = Prim(V)
begin
```

Figure 6. Prim's algorithm pseudo code

In the above pseudo code, the output graph is created by selecting edges from with minimum weight. This is accomplished with the *minElement()* function, where the similarity metric used to calculate weights in given in (10).

(10)

An edge is added to *MST* if the addition of that edge does not create a cycle with the pre-existing edges in the set. Figure 5 provides an example minimum spanning tree composed of invariant feature points.

Figure 6. Minimum spanning tree constructed utilizing Prim's algorithm on a cluster of invariant feature points.

B. Graph-based Region Matching

With the proposed region descriptor, the process of descriptor matching can be approached as a problem of determining if one graph is a sub graph of another. To accomplish this task, a similarity measure is established to make the comparison.

(11)

Given two graphs, and , where , the similarity metric is computed by comparing the similarity of the vertices and the spatial distance between adjacent vertices. is defined as the Euclidean distance between invariant feature points while is the spatial distance.

(12)

Letting and , where , the comparison of two region descriptors requires the computation of ( ) similarity scores. The minimum score is chosen as the closest match and is stored in a scoring matrix , where when descriptors of one image is compared against descriptors of another.

The overall complexity of comparing two region descriptors is given as,

(13)

C. Control Point Selection

Each pair of matched region descriptors is used to determine the pair of control points used for the transformation parameter estimation. The centroid of each matched region is chosen as the control point. Let be the spatial coordinates of the *i-th* SURF keypoint in the region descriptor. Then the centroid for the descriptor is given in (13).

– – (13)

Figure 7 provides an example of two images whose control points are determined from the centroids of the proposed region descriptor. The selected control points are denoted in yellow while the blue vectors illustrate the matched pair of region descriptors that are represented by the blue graphs surrounded by red outlines.

Figure 7. Matched region descriptors and identified control points.

D. Parameter Estimation

Image registration is a process that aims to map one image to another through an estimation of a transformation matrix. To do so, at least four non-collinear matched control points are required in each image. Let be the coordinates in the reference image, and be the control points in the query image, where . Then visually, the goal is to find the tranformation matrix which maps one shape to the other as illustrated in Figure 8.

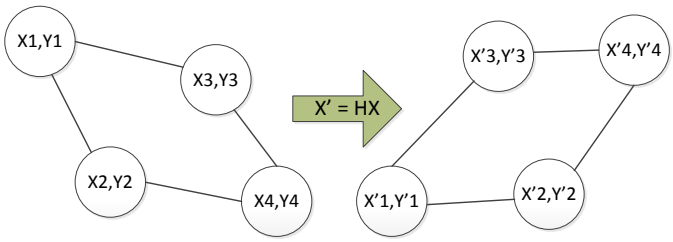


Figure 8. Control point matching.

The Direct Linear Transform (DLT) algorithm outlined in [17] provides a novel approach for estimating the homography matrix, through the Singular Value Decomposition. Let and be the normalized control point pairs such that . Using the normalized control points, a matrix is constructed such that,

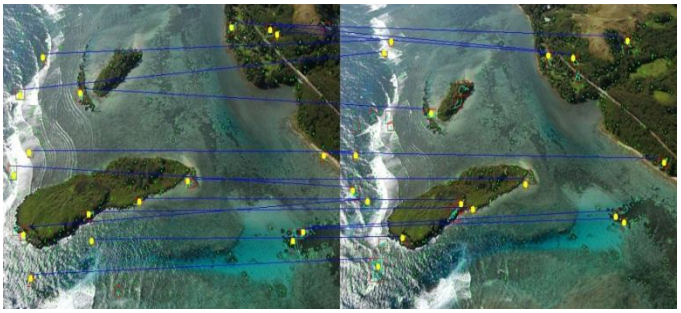
(14)

(15)

The singular value decomposition of yields the homography parameters in the last column of the right singular vector matrix, .

IV. Results and Discussion

The proposed image registration algorithm was thoroughly tested with aerial images acquired from an Unmanned Air Vehicle (UAV) as well as images taken before and after a natural disaster. Figure 8 offers two concrete examples where the proposed region descriptor provides accurate control point locations for registration. In the case of the aerial images, the registration technique must mitigate geometric transformations while the disaster-scene images exhibited both geometric and photometric variations.



a. b.

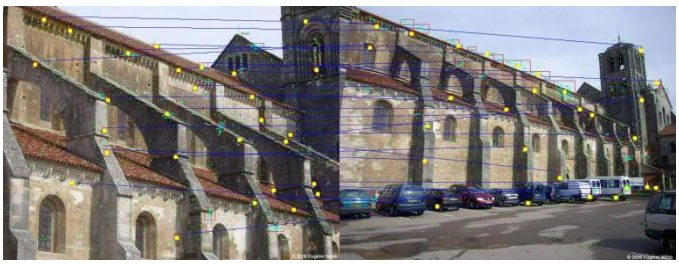
c.

d. e.

f.

g. h.

i.



j. k.

l.

Figure 9a, b, j, k. Reference and query aerial images. c, l. Registration result for aerial images. d, e, g, h. Reference and query image for scenes before and after a natural disaster. f, i. Registration result of disaster scene images.

As another baseline for comparison, the mutual information metric was used to determine the effectiveness of the proposed algorithm as demonstrated in [20-22]. The function is optimally 1.0 when two images are identical and therefore must be maximized for accurate registration. For comparison, the reference image and query image are converted to grayscale while the MI score is calculated from (16).

, (16)

where is the entropy measure for image x and is the joint entropy of images x and y.

(17)

(18)

The probability density function, , is estimated from the intensity histogram of image while is calculated from the joint histogram of images and .

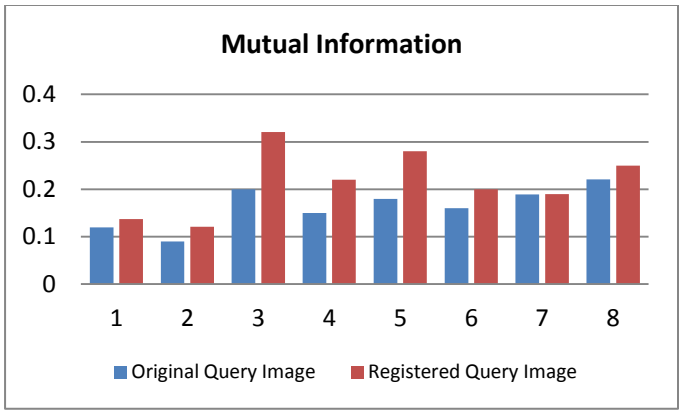


Figure 10. Mutual Information comparison before and after image registration

V. CONCLUSION

In this work, an effective invariant region descriptor and image registration technique has been proposed. Registration control points are identified by employing the graph-based approach for region matching. This method offers a higher discriminative power over traditional SURF feature points due to the coupling of descriptors and spatial relationships through a minimum spanning tree. The results for natural and urban scenes provide strong evidence that the proposed method is both robust and accurate.

REFERENCES

[1] Rueckert, D.; Aljabar, P.; , "Nonrigid Registration of Medical Images: Theory, Methods, and Applications [Applications Corner]," *Signal Processing Magazine, IEEE* , vol.27, no.4, pp.113-119, July 2010.

[2] He, Y.P.; Gu, L.X.; , "Medical Image Registration Using Normal Vector and Intensity Value," *Intelligent Computation and Bio-Medical Instrumentation (ICBMI), 2011 International Conference on* , vol., no., pp.46-49, 14-17 Dec. 2011.

[3] Zhi Li Song; Junping Zhang; , "Remote Sensing Image Registration Based on Retrofitted SURF Algorithm and Trajectories Generated From Lissajous Figures," *Geoscience and Remote Sensing Letters, IEEE* , vol.7, no.3, pp.491-495, July 2010.

[4] Senarathne, C.N.; Ransiri, S.; Arangala, P.; Balasooriya, A.; De Silva, C.; , "A faster image registration and stitching algorithm," *Industrial and Information Systems (ICIIS), 2011 6th IEEE International Conference on* , vol., no., pp.66-69, 16-19 Aug. 2011.

[5] Gang Hong; Yun Zhang; , "Combination of feature-based and area-based image registration technique for high resolution remote sensing image," *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International* , vol., no., pp.377-380, 23-28 July 2007.

[6] Yijian Pei; Hao Wu; Jiang Yu; Guanghui Cai; , "Effective Image Registration based on Improved Harris Corner Detection," *Information Networking and Automation (ICINA), 2010 International Conference on* , vol.1, no., pp.V1-93-V1-96, 18-19 Oct. 2010.

[7] Orchard, J.; , "Globally Optimal Multimodal Rigid Registration: An Analytic Solution using Edge Information," *Image Processing, 2007. ICIP 2007. IEEE International Conference on* , vol.1, no., pp.I-485-I-488, Sept. 16 2007-Oct. 19 2007.

[8] Xiaolong Dai; Khorram, S.; , "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *Geoscience and Remote Sensing, IEEE Transactions on* , vol.37, no.5, pp.2351-2362, Sep 1999.

[9] Chen Xing; Peihua Qiu; , "Intensity-Based Image Registration by Nonparametric Local Smoothing," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.33, no.10, pp.2081-2092, Oct. 2011.

[10] Omer, O.A.; Tanaka, T.; , "Robust image registration based on local standard deviation and image intensity," *Information, Communications & Signal Processing, 2007 6th International Conference on* , vol., no., pp.1-5, 10-13 Dec. 2007.

[11] F. Bowen, E. Du, J. Hu, "A Novel Graph-based Invariant Region Descriptor for Image Matching," International Conference on Electronics/Information Technology, May 2012.

[12] [5] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool, "Speeded-Up Robust Features (SURF)," in International Journal of Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359, 2008.

[13] [6] Zhi Li Song; Junping Zhang; , "Remote Sensing Image Registration Based on Retrofitted SURF Algorithm and Trajectories Generated From Lissajous Figures," Geoscience and Remote Sensing Letters, IEEE , vol.7, no.3, pp.491-495, July 2010.

[14] Ayoung Cho; Won-Keun Yang; Dong-Seok Jeong; Weon-Geun Oh; , "Bag-of-features signature using invariant region descriptor for object retrieval," Frontiers of Computer Vision (FCV), 2011 17th Korea-Japan Joint Workshop on , vol., no., pp.1-4, 9-11 Feb. 2011.

[15] Zen Chen; Shu-Kuo Sun; , "A Zernike Moment Phase-Based Descriptor for Local Image Representation and Matching," Image Processing, IEEE Transactions on , vol.19, no.1, pp.205-219, Jan. 2010.

[16] Alpert, C.J.; Hu, T.C.; Huang, J.H.; Kahng, A.B.; , "A direct combination of the Prim and Dijkstra constructions for improved performance-driven global routing," Circuits and Systems, 1993., ISCAS '93, 1993 IEEE International Symposium on , vol., no., pp.1869-1872 vol.3, 3-6 May 1993.

[17] [1] Hartley, R., Zisserman, A., "Multiple View Geometry in Computer Vision," Cambridge University Press, 2000..

[18] Jian Wu; Heng-jun Yue; Yan-yan Cao; Zhi-ming Cui; , "Video Object Tracking Method Based on Normalized Cross-correlation Matching," Distributed Computing and Applications to Business Engineering and Science (DCABES), 2010 Ninth International Symposium on , vol., no., pp.523-527, 10-12 Aug. 2010.

[19] Peng Fan; Aidong Men; Mengyang Chen; Bo Yang; , "Color-SURF: A surf descriptor with local kernel color histograms," Network Infrastructure and Digital Content, 2009. IC-NIDC 2009. IEEE International Conference on , vol., no., pp.726-730, 6-8 Nov. 2009.

[20] Ming Guo; Jian Xu; Xiaodong Zhou; , "Automatic registration method for distant ship infrared and visible image," Electrical and Control Engineering (ICECE), 2011 International Conference on , vol., no., pp.1008-1011, 16-18 Sept. 2011.

[21] Pluim, J.P.W.; Maintz, J.B.A.; Viergever, M.A.; , "f-information measures in medical image registration," Medical Imaging, IEEE Transactions on , vol.23, no.12, pp.1508-1516, Dec. 2004.

[22] Loeckx, D.; Slagmolen, P.; Maes, F.; Vandermeulen, D.; Suetens, P.; , "Nonrigid Image Registration Using Conditional Mutual Information," Medical Imaging, IEEE Transactions on , vol.29, no.1, pp.19-29, Jan. 2010.