FEATURE-BASED IMAGE REGISTRATION IN LOG-POLAR DOMAIN

Saikiran S. Thunuguntla and *Bahadir K. Gunturk*

Electrical and Computer Engineering
Louisiana State University
Baton Rouge, LA 70803
Email: sthunu1@lsu.edu *and* bahadir@ece.lsu.edu

ABSTRACT

Image registration is a necessary step in a variety of computer vision applications. One of the recent focus areas in image registration is extracting and matching features that are invariant to affine transformation. This is critical in various applications, including 3D reconstruction and object recognition. In this paper, we present a feature-based image registration method that is robust to scaling and rotation. This is achieved by extracting and matching features in log-polar domain, where rotation and scale correspond to translation. Registration parameters are then estimated by applying the RANSAC technique to the feature correspondences. The RANSAC technique provides a robust estimation even there are moving objects within the scene. Experimental results with synthetic and real images are provided.

1. INTRODUCTION

Image registration is a necessary part of a variety of applications, including 3D reconstruction from multiple images, stereo matching, object recognition, compression, tracking, etc. Image registration problem has been studied extensively, yet there are still issues that need to be addressed. One such issue is the robust registration of images with large variations of scale and rotation. Among various image registration approaches, direct optimization methods may fail unless images are close to each other in the registration-parameters space. Hierarchical search techniques may provide robustness to some degree. In [10], scale and rotation robustness problem is addressed by a two-step algorithm. In the first step, images are transformed to log-polar domain, where scale and rotation becomes translation. To have the same characteristics, centers of log-polar transformation must match for both images. Therefore, log-polar transformation is repeated for every point near the geometric center, and cross correlation is checked in log-polar domain. The center with the maximum cross correlation gives the translation amount between the images. Scale and rotation are also determined for the correct translation. As a result a rough estimate of translation, rotation, and scale is obtained between the images. In the second step, a non-linear optimization method that is based on the Levenberg-Marquardt technique is used to fine-tune the affine parameters.

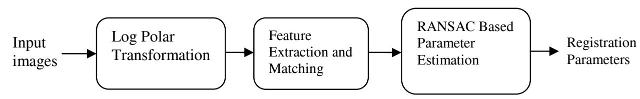

Fig. 1. The basic idea is to find and match features in log-polar domain so that scale and rotation invariance is achieved.

An alternative to direct optimization methods, feature-based methods are more robust to scale and rotation variations, especially as a result of recent developments. One of the most popular feature extractors is the Harris corner detector [4]. In [8], Schmid and Mohr used rotationally invariant descriptor of local image regions to match Harris corners. This allowed features to be matched under arbitrary rotation. However, the Harris corner detector is sensitive to changes in scale, and feature-based algorithms based on the Harris corner detector may not work under large scale variations. In [5], Lowe extended the local feature approach to achieve scale invariance. Recently, research on affine-invariant feature extraction and matching has taken off. In [1], an adaptive procedure based on isotropy of the second moment gradient matrix is proposed. Mikolajczk and Schmid proposed a similar method in [6]. There are also other algorithms [9, 7, 2].

In this paper, we present a feature-based registration algorithm. Inspired by Wolberg's work in [10], the features will be extracted and matched in log-polar domain. Since rotation and scale becomes translation, robustness to rotation and scale is achieved. Once the features are extracted and matched, a robust estimation algorithm based on the RANSAC technique is employed to find geometric transformation parameters. The RANSAC technique provides robust model fitting [3]. Figure 1 shows the block diagram of this idea. Note that log-polar transformation must have the same centers as mentioned above. Handling this issue will be explained shortly.

In Section 2, we present the proposed algorithm. Experimental results are given in Section 3. Section 4 concludes the paper.

2. REGISTRATION ALGORITHM

2.1. Features Log-Polar Domain and Robust Estimation of Registration Parameters

Since log-polar transformation converts scaling and rotation into translation, we can deal with any angle of rotation and very large scaling in the log-polar domain. Consider a point (x, y) in Cartesian coordinate system. In polar coordinate system, (x, y) is represented by (ρ, θ) where ρ is the radial distance of the point from center and θ is angle. Let (x_c, y_c) be the center of polar transformation, then the relationship between Cartesian and polar coordinates is

$$\rho = \sqrt{(x - x_c)^2 + (y - y_c)^2} \tag{1}$$

and

$$\theta = \tan^{-1}\left(\frac{y - y_c}{x - x_c}\right). \tag{2}$$

Rotation in Cartesian domain becomes translation in polar coordinate system. We further take the logarithm of the radial distance ρ to obtain the log-polar representation $(log(\rho), \theta)$ of an image. Obviously, scaling also becomes a translation in log-polar domain. (If s is a scale factor, $log(s\rho) = log(s) + log(\rho)$.)

To obtain rotation and scale invariant features, we need to solve one more problem. Log-polar transformation centers should be same for both reference and target images to have the same features. Therefore, for a fixed reference image center, different target image centers are tested iteratively to find the distance between image centers. At each iteration, Fast Fourier Transform (FFT) based cross-correlation is calculated. The distance with the largest correlation is the correct offset amount between the image centers. An alternative to the FFT-based based correlation is a feature-matching-based approach. We first find the corners in both images using the Harris corner detector. We then match these corners with a 7×7 correlation window. We count the number of corners for which the matching error is less than a threshold. The offset with the largest corner match count gives us the translation amount between images. Hierarchical search could be used to reduce the computational complexity of the offset finding. Gradient descent type of search (instead of exhaustive search) is also an option. RANSAC estimation is then applied to the feature matches to determine rotation and scale. (Note that some of the features detected by the Harris technique are due to the discontinuity along image borders. These features are removed and not used in the RANSAC estimation step.) We tested both the FFT-based and feature-matching-based offset calculation methods. Initial experiments showed similar performances for both techniques. The results in this paper are with the feature-matching-based method.

2.2. Extension to Perspective Registration

The algorithm explained so far can find translation, rotation, and scaling parameters only. The question is how to generalize it. There might be several ways. One of them is to find the (rotation and scale invariant) features as explained in the previous subsection, and then to back to Cartesian coordinates and fit any parametric model (such as perspective model) to the matched features. Application of RANSAC would again remove outliers.

2.3. Complete Algorithm

The complete algorithm is as follows:

- Apply log-polar transformation to reference image in Cartesian coordinates $I_{cartesian}$ to obtain $I_{logpolar}$. Let (x_I, y_I) be the geometric center of the log-polar transformation.
- Apply the Harris corner detector to find corners in $I_{logpolar}$.
- Let d_x and d_y be the search range for the offset between the images.
- For $m = x_I - d_x$ to $x_I + d_x$
    - For $n = y_I - d_y$ to $y_I + d_y$
        - Apply log-polar transformation to target image $J_{cartesian}$ to obtain $J_{logpolar}$. (m, n) is the geometric center of the log-polar transformation.
        - Find the correlation between $I_{logpolar}$ and $J_{logpolar}$ using either the FFT-based or the feature matching-based approaches explained in Section 2.1. (m, n) with maximum correlation or corner match count is the correct center of the target image.
    - End
- End
- Apply RANSAC to the feature matches between $I_{logpolar}$ and $J_{logpolar}$ with the correct center to find horizontal and vertical translations. Obtain scale and rotation from the translation parameters.
- Instead of the previous step, transform images to Cartesian coordinates and apply RANSAC to the matched features to find perspective parameters.

Also note the that computational complexity may be reduced by a hierarchical (coarse-to-fine multi-resolution) search scheme for testing different image centers.

Fig. 2. The first set of images used to demonstrate the algorithm are shown.

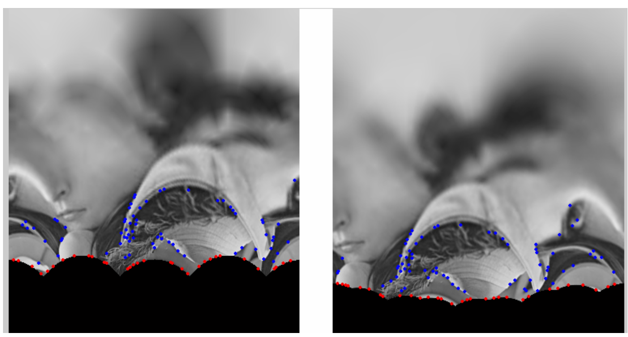

Fig. 3. Extracted features are shown. The red features, which are the border features, are removed automatically before applying RANSAC.

3. EXPERIMENTAL RESULTS

We provide experimental results with two image sets. The first image set is given in Figure 2. Extracted features are shown in Figure 3. The features which are along the border are removed before applying RANSAC. The inliers as a result of RANSAC are shown in Figure 4. The final registered image is shown in Figure 5. The results for the second image set are shown in Figures 6, 7, 8, and 9. For the first image set, we fit perspective parameters to the matched features in Cartesian coordinates. For the second image set, rotation and scale parameters are found in the log-polar domain. Notice that in the second image set, the location of the cellular phone is not fixed. However, the final registration is not affected as RANSAC removes the outliers. For both image sets, we used the feature-matching-based method (explained in Section 2.1) to determine the offset amounts.

4. CONCLUSIONS

In this paper, we present a feature-based registration algorithm. Inspired by Wolberg's work in [10], the features are be extracted and matched in log-polar domain. Therefore, robustness to rotation and scale is achieved. Once the fea-

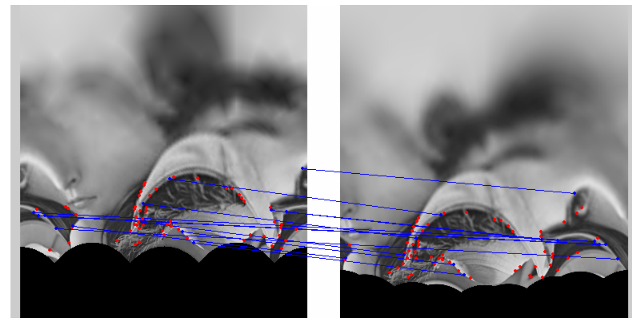

Fig. 4. Features are matched. The inliers are shown in blue. The red features are the outliers.

Fig. 5. Reference and input images are aligned and blended on top of each other.

Fig. 6. The second set of images used to demonstrate the algorithm are shown. These images are real unlike the synthetically generated first image set.

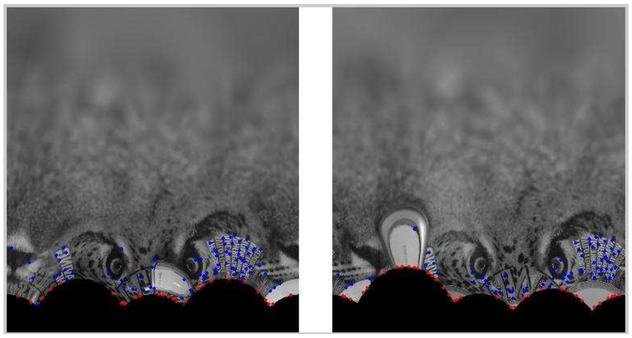

Fig. 7. Extracted features are shown. The red features, which are the border features, are removed automatically before applying RANSAC.

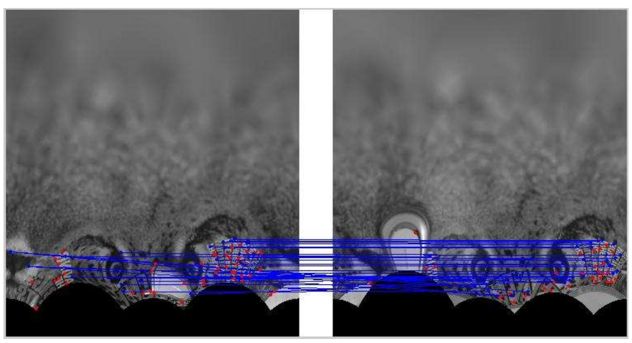

Fig. 8. Features are matched. The inliers are shown in blue. The red features are the outliers.

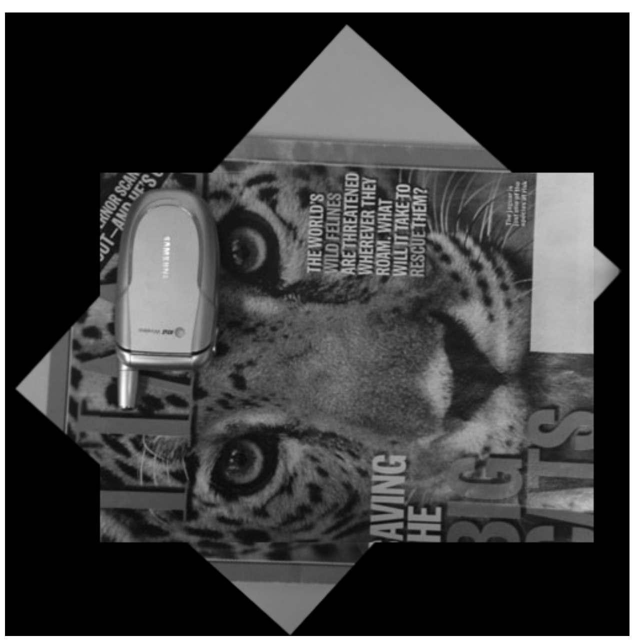

Fig. 9. Reference and input images are aligned on top of each other.

tures are extracted and matched, the RANSAC technique is employed to find registration parameters, either in the log-polar domain or in the Cartesian domain. The advantage of the feature-based approach is the robustness even there are moving objects within the scene.

5. REFERENCES

[1] A. Baumberg, “Reliable feature matching across widely separated views,” in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2000, pp. 774–781.

[2] M. Brown and D. G. Lowe, “Invariant features from interest point groups,” in *Proc. of the British Machine Vision Conference*, 2002, pp. 656–665.

[3] M. A. Fischler and R. C. Bolles, “Random sampling consensus - a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. Assoc. Comp. Mach.*, vol. 24, no. 6, pp. 381–395, 1981.

[4] C. Harris and M. Stephens, “A combined corner and edge detector,” in *Proc. of the Fourth Alvey Vision Conference*, 1988, pp. 147–151.

[5] D. G. Lowe, “Object recognition from local scale-invariant features,” in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 1999, pp. 1150–1157.

[6] K. Mikolajczyk and C. Schmid, “An affine invariant interest point detector,” in *Proc. of the European Conf. on Computer Vision*, 2002, pp. 128–142.

[7] F. Schaffalitzky and A. Zisserman, “Multi-view matching for unordered image sets, or ’how do i organize my holiday snaps?’,” in *Proc. of the European Conf. on Computer Vision*, 2002, pp. 414–431.

[8] C. Schmid and R. Mohr, “Local grayvalue invariants for image retrieval,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–534, 1997.

[9] T. Tuytelaars and L. V. Gool, “Wide baseline stereo based on local, affinely invariant regions,” in *Proc. of the British Machine Vision Conference*, 2000, pp. 412–422.

[10] G. Wolberg and S. Zokai, “Robust image registration using log-polar transform,” in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, 2000, pp. 493–496.